Stephen D. Turner, Ph.D.

Assistant Professor of Public Health Sciences
Bioinformatics Core Director
University of Virginia School of Medicine
P.O. Box 800741, Charlottesville, VA 22908

Phone: 434-982-4208
E-mail: bioinformatics@virginia.edu
Web: http://stephenturner.us
Blog: http://gettinggeneticsdone.com
Twitter: @genetics_blog

SUMMARY

I am faculty in the Department of Public Health Sciences at the University of Virginia, and director of the Bioinformatics Core in the School of Medicine. I provide expert bioinformatics, biostatistics, and computational biology consulting services, working closely with investigators in the experimental design and analysis of genome-scale research, including studies of gene expression, protein-DNA binding, DNA methylation, DNA variation, and metagenomics, using both array and next-generation sequencing technologies.

EDUCATION

Ph.D.	Human Genetics	Vanderbilt University, Nashville TN
M.S.	Applied Statistics	Vanderbilt University, Nashville TN
B.S.	Biology	James Madison University, Harrisonburg VA

PROFESSIONAL EXPERIENCE

2011-present	Bioinformatics Core Director, Assistant Professor of Public Health Sciences, University of Virginia School of Medicine
2011-2011	Postdoctoral Research Fellow, University of Hawaii Cancer Center, Epidemiology
2006-2010	Graduate Research Fellow, Vanderbilt University, Center for Human Genetics Research
2002-2006	Research Assistant, James Madison University, Dept. of Biology

OTHER QUALIFICATIONS & PROFESSIONAL ACTIVITIES

2013-present	U.S. Government Security Clearance
2010-present	Registered U.S. Patent Agent
2010-present	Faculty of 1000
2010-present	Advisor, Wiley-Blackwell Scientific Advisory Group
2010-present	American Society of Human Genetics Communications Committee
2009-present	Founder & contributing editor, GettingGeneticsDone.com, a genetics & bioinformatics blog with over 7,000 subscribers and over 20,000 pageviews monthly from 169 countries worldwide.

RESEARCH INTERESTS

Biostatistics & Bioinformatics
Complex Disease Genetics
Next-Gen Sequencing
Metagenomics
RNA-Seq
Systems Biology
Personal Genomics
Microbial Forensics
High-Throughput Biology
Technology Management
Patent Law
RNA-seq

FUNDING

ACTIVE FUNDING

HHM402-11-D-0031-DO-0007 (Turner) 2012-2014 DoD subcontract $93,080
Title: Microbial Forensics Technical & Scientific Process Subject Matter Expert Support
Role: Principal investigator

W911NF-12-1-0599 (Hewlett) 2012-2013 DoD / ARO $116,832
Title: Microbial Forensics Technical & Scientific Process Subject Matter Expert Support
Role: Co-investigator

(Grant number pending) 2013-2014 Gates Foundation
Title: Novel Intestinal Biomarkers for Environmental Enteropathy in Pakistani Children Using a Pathophysiologic Approach
Role: Statistical consultant / advisor

COMPLETED FUNDING

F31 NS066638-01/02 (Turner) 2009-2011 NIH/NINDS $112,194
Title: Genomewide Analysis of Interactions in Alzheimer Disease Using Neural Networks
Role: Principal investigator

T32 GM080178 (Haines) 2007-2009 NIH/NIGMS $160,000
Title: Training Program on Genetic Variation and Human Phenotypes
Role: Trainee

PEER REVIEW

2013	Referee for *G3: Genes, Genomes, Genetics*
2012-2013	Study Section / International College of Reviewers and Review Committee member for Genome Canada and Canadian Institutes of Health Research *Genomics and Personalized Health* RFA
2012	Program committee for EvoBio 2013: 11th European Conference on Evolutionary Computation, Machine Learning and Data Mining in Computational Biology
2011	Program committee for EvoBio 2012: 10th European Conference on Evolutionary Computation, Machine Learning and Data Mining in Computational Biology
2011	Referee for Elsevier for book: *R and Data Mining: Examples and Case Studies*.
2011	Referee for *Bioinformatics*
2011	Referee for *American Journal of Epidemiology*
2011	Referee for *BioData Mining*
2010	Referee for *Respiratory Research*

PUBLICATIONS

Scian MJ, Maluf DG, Archer KJ, **Turner SD**, Suh JL, David KG, King AL, Posner M, Brayman KL, Mas VR. Identification of Biomarkers to Assess Organ Quality and Predict Post-Transplant Outcomes. *Transplantation*. 94(8): 1-8 (2012). PMID: 22992769.

Lim U, **Turner SD**, Franke AA, Cooney RV, Wilkens LR, Ernst T, Albright CL, Novotny R, Chang L, Kolonel LN, Murphy SP, Le Marchand L. Predicting total, abdominal, visceral and hepatic adiposity with circulating biomarkers in caucasian and Japanese american women. *PLoS ONE*. 7(8):e43502 (2012). PMID: 22912885.

Zuvich RL, Armstrong LL, Bielinski S, Bradford Y, Carlson CS, Clayton E, Crawford DC, Crenshaw AT, de Andrede M, Doheny K, Haines JL, Hayes G, Jarvik G, Jiang L, Kullo I, Li R, Ling H, Matsumoto M, McCarty CA, McDavid AN, Mirel DB, Olson L, Paschall J, Pugh E, Rasmussen LV, Rasmussen-Torvik L, **Turner SD**, Wilke RA, Ritchie MD. Lessons learned in merged GWAS data in the eMERGE Network: Pitfalls of merging GWAS data: lessons learned in the eMERGE network and quality control procedures to maintain high data quality. *Genetic Epidemiology*. 35(8):887-898 (2011). PMID:22125226.

Turner SD, Ritchie MD, Berg D, Peissig P, Linneman J, McCarty CA, Wilke RA. Knowledge-Driven Multi-Locus Analysis Reveals Gene-Gene Interactions Influencing HDL Cholesterol Level in Two Independent EMR-Linked Biobanks. *PLoS ONE*. 6(5):e19586 (2011). PMID: 21589926.

Turner SD, Armstrong L, Bradford Y, Carlson C, Crawford DC, Crenshaw AT, de Andrede M, Doheny K, Haines JL, Hayes G, Jarvik G, Jiang L, Ling H, Kullo I, Li R, Manolio TA, Matsumoto M, McCarty CA, McDavid A, Mirel D, Paschall J, Pugh E, Rasmussen LV, Wilke RA, Zuvich RL, Ritchie MD. Quality Control procedures for Genome-Wide Association Studies. *Current Protocols in Human Genetics*. Chapter 1, Unit 1.19 (2011). PMID:21234875.

Turner SD, Bush WS. A Multivariate Analysis of Regulatory SNPs: Empowering Personal Genomics by Considering Cis-Epistasis and Heterogeneity. *Pacific Symposium in Biocomputing* (2011). *Pac Symp Biocomput*. 2011:276-287 (2011). PMID:21121055.

Turner SD, Dudek SM, Ritchie MD. ATHENA: A Knowledge-Based Hybrid Backpropagation-Grammatical Evolution Neural Network Algorithm for Discovering Epistasis among Quantitative Trait Loci. *BMC BioData Mining*. 3:5 (2010). PMID:20875103.

Zabaleta J, Camargo MC, Ritchie MD, Piazuelo MB, Sierra RA, **Turner SD**, Delgado A, Fontham ETH, Schneider BG, Correa P, Ochoa AC. Association of haplotypes of inflammation-related genes with gastric preneoplastic lesions in African Americans and Caucasians. *International Journal of Cancer.* 128(3):668-675 (2011). PMID:20473875.

Turner SD, Dudek SM, Ritchie MD. Incorporating Domain Knowledge into Evolutionary Computing for Discovering Gene-Gene Interaction. *11th Int'l Conference on Parallel Problem Solving From Nature (PPSN), Lecture Notes in Computer Science*. 6238(I):394-403 (2010).

Holzinger ER, Buchanan C, **Turner SD**, Dudek SM, Torstenson ES, Ritchie MD. Initialization Parameter Sweep in ATHENA: Optimizing Neural Networks for Detecting Gene-Gene Interactions in the Presence of Small Main Effects. *Genetic and Evolutionary Computation Conference – GECCO 2010*: 203-210. ACM Press (2010). PMID:21151364.

Turner SD, Dudek SK, Ritchie MD. Grammatical Evolution of Neural Networks for Discovering Epistasis among Quantitative Trait Loci. *Lecture Notes in Computer Science.* 6023:86-97 (2010).

Edwards TS, **Turner SD**, Torstenson E, Dudek SM, Martin ER, Ritchie MD. A General Framework for Formal Tests of Interaction after Exhaustive Search Methods with Applications to MDR and MDR-PDT. *PLoS One.* 5(2):e9363 (2010). PMID: 20186329.

Turner SD, Crawford DC, Ritchie MD. Methods for optimizing statistical analyses in pharmacogenomics research. *Expert Reviews in Clinical Pharmacology.* 2(5):559-570 (2009). PMID: 20221410.

Turner SD, Ritchie MD, Bush WS. Conquering the Needle-in-a-Haystack: How Correlated Input Variables Beneficially Alter the Fitness Landscape for Neural Networks. *Lecture Notes in Computer Science*. 5483:80-91 (2009).

Edwards TL, Bush WS, **Turner SD**, Dudek SM, Torstenson E, Schmidt M, Martin ER, Ritchie MD. Generating linkage disequilibrium patterns in data simulations using GenomeSIMLA. *Lecture Notes in Computer Science.* 4973:24-35 (2008).

PENDING PUBLICATIONS

Prince J, Lundgren A, Stadnisky MD, Nash WT, Beeber A, Turner SD, Brown MG. Multi-parametric analysis of host response to murine CMV in MHC class I disparate mice reveals primacy of D^k-licensed $Ly49G_2$+ NK cells in viral control. *Journal of Immunology*. In revision.

PUBLICATIONS IN PREPARATION

Song W, Mondal P, Alonso LC, Ong BWT, Lim OC, Wolfe A, Radovick S, Wondisford FE, Farber EA, Farber CR, **Turner SD**, Hussain MA. [*Title redacted until publication*]. Preparing to submit to *Nature*.

Belyea BC, Castellanos-Riveria R, Xu F, Hu Y, **Turner SD**, Legallo R, Metrano S, Pentz ES, Sequeira-Lopez MLS, Gomez RA. Deletion of RBP-J in the Renin Progenitor: A Novel Origin for B Cell Leukemia.

Turner SD and Bush WS. Data Management. Book Chapter in *Genetic Analysis of Complex Diseases, 3rd ed.* Eds. Ritchie MD and Scott WK.

Bush WS and **Turner SD**. Complex Genetic Interactions. Book Chapter in *Genetic Analysis of Complex Diseases, 3rd ed.* Eds. Ritchie MD and Scott WK.

PRESENTATIONS

Invited Talk & Session Organizer, Association of Biomolecular Resource Facilities (ABRF) first annual Mid-Atlantic Directors and Staff of Cores (MADSSCi) meeting, Frederick MD, 6/22/2013.
Bioinformatics Cores & New Technologies: The evolving role of 21st Century Bioinformatics Core Facilities.

Invited Speaker and Panelist at UVA Center for Public Health Genomics Retreat, 4/27/2013.
Computational / Analytical Genomics: Spanning the Gap from Data to Knowledge

Invited Talk, School of Engineering, Science and Technology, Virginia State University, 2/22/2013.
Bioinformatics-as-a-Service.

Invited Talk, Department of Statistics Colloquium, University of Virginia, 10/19/2012.
Bioinformatics-as-a-Service.

Invited Talk, Center for Human Genetics Research, Vanderbilt University, 9/21/2012.
Bioinformatics-as-a-Service.

Invited Talk, 20th Annual Conference on Intelligent Systems for Molecular Biology, Long Beach, CA, 7/16/2012.
From Data to Knowledge: Extracting Biological Insight from Diverse Data Sources.

Invited Talk, Genome Sciences Seminar Series, University of Virginia, 4/23/12.
Bioinformatics at Virginia: Opportunities and Challenges with Large-Scale –Omics Data.

Invited Talk, Pacific Symposium in Biocomputing, Big Island, Hawaii, 1/5/2011.
A Multivariate Analysis of Regulatory SNPs: Empowering Personal Genomics by Considering Cis-Epistasis and Heterogeneity.

Platform Presentation, 60th Annual American Society of Human Genetics Meeting, Washington, DC 10/5/2010.
Knowledge-driven multi-locus analysis reveals gene-gene interactions influencing HDL cholesterol level in two independent biobanks.

Invited Talk, 8th European Conference on Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics, Istanbul, Turkey, 4/7/2010.
Grammatical Evolution of Neural Networks for Discovering Epistasis among Quantitative Trait Loci.

Graduate Student Research Symposium, Vanderbilt University, 3/26/2010.
Using a biobank linked to an electronic medical record to study genetic underpinnings of complex disease.

Genetics Interest Group, Vanderbilt University, 3/11/2010.
Using GWAS in an EMR-linked biobank to explore genetic and environmental determinants of HDL cholesterol.
Invited Talk, 7th European Conference on Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics, University of Tübingen, Germany, 4/15/2009.
Conquering the Needle-in-a-Haystack: How Correlated Input Variables Beneficially Alter the Fitness Landscape for Neural Networks.

Genetics Interest Group, Vanderbilt University, 2/26/2009.
Conquering the Needle-in-a-Haystack: How Correlated Input Variables Beneficially Alter the Fitness Landscape for Neural Networks.

Genetics Interest Group, Vanderbilt University, 4/1/2008.
Neural Networks for Disease Gene Identification.

TEACHING & OUTREACH

Invited *Thought Leader* at UVA Center for Public Health Genomics Retreat, 4/27/2013.
Computational / Analytical Genomics: Spanning the Gap from Data to Knowledge

Instructor at University of Miami's Genetic Analysis of Complex Human Diseases Course, 5/22/2013.
Examining Gene Expression and Methylation with Next-Gen Sequencing.

Lecture & Lab in Essentials of Translational Science (UVA Cell Biology 8401), 4/2013.
Introduction to Bioinformatics (Lecture), *Gene Expression Analysis with Bioconductor* (Lab)

Organizer, Software Carpentry Two-Day Software Skills Bootcamp. 3/2013.
Software Carpentry Two-Day Software Skills Bootcamp

Lecturer in UVA PHS 7070: Introcution to Health Informatics Graduate Course, 11/2012.
Introduction to Bioinformatics

Panel Discussion, 20th Annual Conference on Intelligent Systems for Molecular Biology (ISMB), Long Beach, CA, 7/16/2012.
Bioinformatics Core Facilities Workshop on Data Integration

Instructor at University of Miami's Genetic Analysis of Complex Human Diseases Course, 5/24/2012
Examining Gene Expression and Methylation with Next-Gen Sequencing.

Panelist at UVA Library Scientific Data Consulting Group's Data Management Day, 4/12/2012
Panel Discussion on Data Management Training and Support at UVA.

UVA Health Sciences Library Special Topics Seminar, 4/10/2012
Introduction to Bioinformatics and the UVA Bioinformatics Core.

Guest Lecturer in UVA Pharmacology 9020 Graduate Course, 4/2/2012
Bioinformatics at Virginia: Opportunities and Challenges with Large-Scale –Omics Data.

Guest Lecturer in UVA Biomedical Engineering Graduate Course, 1/25/2012
Bioinformatics at Virginia: Opportunities and Challenges with Large-Scale –Omics Data.

POSTERS & ABSTRACTS

Mid-Atlantic Chapter of the Medical Library Association Regional Meeting 2012. Paper and talk presented by BR.
The BioConnector - Research Informatics Support Through Collaboration
Ragon B, Arnold G, Horne AS, Harrison J, **Turner SD**.

Fourth Congress of the International BioIron Society, 5/25/2011.
Transferrin Gene Polymorphism rs3811647 and Iron Status in Patients with Iron Deficiency and Control Subjects
McLaren GD, McCarty CA, Ritchie MD, **Turner SD**, Gordeuk VR, Garner CP, Vulpe CD, McLachlan S, Adams PC, McLaren CE.

American Society of Hematology Meeting, 12/4/2010 (*Blood* Vol 116, No 21, p 172).
Lack of Association between Single Nucleotide Polymorphism rs3811647 in the TF Gene and Iron Status in Patients with Iron Deficiency and Control Subjects.
McLaren GD, McCarty CA, Ritchie MD, **Turner SD**, Gordeuk VR, Garner CP, Vulpe CD, McLachlan S, Adams PC, McLaren CE.

American Society of Human Genetics Meeting, 11/5/2010 (Platform presentation).
Gene-gene interactions associated with Alzheimer disease replicate in independent Alzheimer Disease Genetics Consortium datasets.
Thornton-Wells TA, Torstenson ES, **Turner SD**, Dudek SM, Bush WS, Ritchie MD, Martin E, Pericak-Vance MA, Haines, JL.

American Society of Human Genetics Meeting, 11/5/2010.
Quality Control Pipeline for Genome-Wide Association Studies in the eMERGE Network: Comparing Single Site QC to a Merged QC Approach.
Ritchie MD, Armstrong L, Bradford Y, Carlson C, Crawford DC, Crenshaw AT, de Andrede M, Doheny K, Haines JL, Hayes G, Jarvik G, Jiang L, Ling H, Kullo I, Li R, Manolio TA, Matsumoto M, McCarty CA, McDavid A, Mirel D, Paschall J, Pugh E, Rasmussen LV, Wilke RA, Zuvich RL, **Turner, SD**.

American Society of Human Genetics Meeting, 11/4/2010.
Initialization Parameter Sweep in ATHENA: Optimizing Neural Networks for Detecting Gene-Gene Interactions in Simple and Complex Disease Models.
Holzinger ER, Buchanan C, **Turner SD**, Torstenson ES, Dudek SM, Ritchie MD.

11th Annual Vanderbilt Human Genetics Symposium 10/08/2009.
Knowledge-driven multi-locus analysis reveals gene-gene interactions influencing HDL cholesterol level in two independent biobanks.
Turner SD, Bert RL, Crawford DC, Denny JC, Linneman JG, McCarty CA, Peissig PL, Rasmussen LV, Roden DM, Wilke RA, Ritchie MD.

International Genetic Epidemiology Society Meeting, 10/11/2010.
Quality Control Pipeline for Genome-Wide Association Studies in the eMERGE Network: Comparing Single Site QC to a Merged QC Approach.
Ritchie MD, Armstrong L, Bradford Y, Carlson C, Crawford DC, Crenshaw AT, de Andrede M, Doheny K, Haines JL, Hayes G, Jarvik G, Jiang L, Ling H, Kullo I, Li R, Manolio TA, Matsumoto M, McCarty CA, McDavid A, Mirel D, Paschall J, Pugh E, Rasmussen LV, Wilke RA, Zuvich RL, **Turner, SD**.

International Genetic Epidemiology Society Meeting, 10/11/2010.
Knowledge-driven multi-locus analysis reveals gene-gene interactions influencing HDL cholesterol level in two independent biobanks.
Turner SD, Bert RL, Crawford DC, Denny JC, Linneman JG, McCarty CA, Peissig PL, Rasmussen LV, Roden DM, Wilke RA, Ritchie MD.

11th International Conference on Parallel Problem Solving From Nature, Krakow, Poland, 9/11/2010.
Incorporating Domain Knowledge into Evolutionary Computing for Discovering Gene-Gene Interaction.
Turner SD, Dudek SM, Ritchie MD.

American Medical Informatics Association Summit on Translational Bioinformatics, 3/11/2010.
Using Self-Reported Data to Determine Relatedness in Biobank Subjects.
Rasmussen L, **Turner SD**, Waudby C, Ritchie MD, McCarty CA.

Pacific Symposium in Biocomputing, 1/7/2010.
Evaluating and Improving Grammatical Evolution Neural Networks for Discovering Epistasis Influencing Quantitative Traits.
Turner SD, Dudek SM, Ritchie MD.
American Society of Human Genetics Meeting, 10/22/2009.
Genome-wide association study of cataract in the Marshfield Personalized Medicine Research Project as part of the eMERGE network.
Torstenson ES, McCarty CA, **Turner SD**, Bradford Y, Berg R. Peissig P, Linneman J, Starren J, Waudby C, Chen L, Ritchie MD.

American Society of Human Genetics Meeting, 10/22/2009.
Genome-wide association study on HDL cholesterol level in the Marshfield Personalized Medicine Research Project as part of the eMERGE network.
Turner SD, McCarty CA, Bradford Y, Berg R. Peissig P, Linneman J, Starren J, Wilke RA, Ritchie MD.

American Society of Human Genetics Meeting, 10/22/2009.
Getting Genetics Done: an educational productivity blog by graduate students for getting things done in human genetics research.
Holzinger ER, Bush WS, **Turner SD**.

International Genetic Epidemiology Society Meeting, 10/19/2009.
Genome-wide association study on HDL cholesterol level in the Marshfield Personalized Medicine Research Project as part of the eMERGE network.
Turner SD, McCarty CA, Bradford Y, Berg R. Peissig P, Linneman J, Starren J, Wilke RA, Ritchie MD.

10th Annual Vanderbilt Human Genetics Symposium 10/08/2009.
Genome-wide association study on HDL cholesterol level in the Marshfield Personalized Medicine Research Project as part of the eMERGE network.
Turner SD, McCarty CA, Bradford Y, Berg R. Peissig P, Linneman J, Starren J, Wilke RA, Ritchie MD.

American Society of Human Genetics Meeting, 11/12/2008.
Genome-wide Analysis of Gene-Gene Interaction in Alzheimer Disease.
Turner SD, Martin ER, Beecham GW, Gilbert JR, Haines JL, Pericak-Vance MA, Ritchie MD.

American Heart Association Scientific Sessions, 11/10/2008.
SCN5A Promoter and Enhancer Variants associated with Atrial Fibrillation.
Watanabe H, Darbar D, Atack TC, Yang P, Ritchie MD, **Turner SD**, Ingram CR, Roden DM.
International Genetic Epidemiology Society Meeting 9/14/2008.
Genome-wide Analysis of Gene-Gene Interaction in Alzheimer Disease.
Turner SD, Martin ER, Beecham GW, Gilbert JR, Haines JL, Pericak-Vance MA, Ritchie MD.

9th Annual Vanderbilt Human Genetics Symposium 09/11/2008.
Genome-wide Analysis of Gene-Gene Interaction in Alzheimer Disease.
Turner SD, Martin ER, Beecham GW, Gilbert JR, Haines JL, Pericak-Vance MA, Ritchie MD.

International Conference on Alzheimer's Disease, 7/26/2008.
Genome-wide Analysis of Gene-Gene Interaction in Alzheimer Disease.
Martin ER, **Turner SD**, Beecham GW, Gilbert JR, Haines JL, Pericak-Vance MA, Ritchie MD.

Program in Human Genetics Annual Retreat, 4/9/2008.
Neural Networks for Disease Gene Discovery.
Turner SD.

American Society of Human Genetics Meeting, 10/23/2007.
GenomeSIMLA: A Data Simulation Package to Explore the Human Genome.
Edwards TL, Bush WS, **Turner SD**, Torstenson ES, Dudek SM, Ritchie MD.

American Society of Human Genetics Meeting, 10/23/2007.
A Platform for the Analysis, Translation, and Organization of Whole-Genome Association Data.
Ritchie MD, **Turner SD**, Bush WS, Dudek SM.

American Society of Human Genetics Meeting, 10/23/2007.
Examination of Sortilin-related receptor SORL1 in late onset Alzheimer Disease.
Turner SD, Liang X, Martin ER, Schnetz-Boutaud N, Bartlett J, Anderson BM, Zuchner S, Gwirtsman H, Schmechel D, Carney R, Gilbert J, Pericak-Vance MA, Haines JL.

8th Annual Vanderbilt Human Genetics Symposium 09/06/2007.
Development of a sequential replication filter to minimize type I error in genetic association studies.
Turner SD, Dudek SM, Ritchie MD.

ADVANCED COURSEWORK AND EDUCATION

*Software Carpentry Bootcamp**, University of Virginia, 2013
**I invited the instructors and organized this course to be given at University of Virginia*
Description: Software Carpentry helps researchers be more productive by teaching them basic computing skills. We run boot camps at dozens of sites around the world, and also provide open access material online for self-paced instruction. Our boot camps cover the core skills every scientist needs to get more done in less time: program construction, version control, testing, the command line, and data management. Short lessons alternate with hands-on practical sessions for two full days.

*Browsing Genes and Genomes with Ensembl.**, University of Virginia, 2012
**I invited the instructor and organized this course to be given at University of Virginia*
Description: The Ensembl project (http://www.ensembl.org) provides a comprehensive and integrated source of annotation of, mainly vertebrate, genome sequences. This 1-day workshop offers participants the possibility of gaining lots of hands-on experience in the use of the Ensembl genome browser but also provides them with the necessary background information. The workshop consists of several modules with presentations, demonstrations, and ample opportunity to do exercises. Modules include: introduction to Ensembl, data retrieval with BioMart, comparative genomics, variation, regulation, and custom annotation.

*Looking at Data by Hadley Wickham**, Vanderbilt University, 2010
**I invited the instructor and organized this course to be given at Vanderbilt University*
Description: This short course in data visualization will teach you how to explore your data using statistical graphics. You will learn how to use exploratory data analysis to find unexpected features in data and build better models by iterating between fitting and plotting. The course will focus on ggplot2, a new data visualization package for R that uses a graphical grammar to create a powerful and flexible system for creating beautiful data

graphics. In addition to an introduction to ggplot2, the course will also cover topics related to graphics for large datasets, data manipulation & transformation, and how to polish graphics for maximum presentation impact.

Regression Modeling Strategies Short Course by Frank Harrell, Vanderbilt University, 2010
Description: This short course is aimed at statistical practitioners, educators, and students of statistics throughout science and engineering who wish to have an overview of all the steps involved in the proper conduct of regression modeling, using modern methods and procedures. The speaker is nationally known, and is Chair of the Department of Biostatistics at the Vanderbilt University School of Medicine. He has also published a textbook by the same title as this course.

Grant & Fellowship Workshop for Graduate Students, Vanderbilt University, 2010
Description: These concurrent sessions will address grant proposal writing from inception to completion. In this workshop, Vanderbilt University's active grant recipients will discuss conceptualizing the project, composing convincing budgets, and securing necessary approval. This workshop will help graduate students better understand the timeline of the grant writing process.

An Introduction to the Fundamentals & Functionality of the R Language, Vanderbilt University, 2009
Description: In this 1-day course, we will introduce R, the free interactive programming language and environment for statistical computing and graphics. We will emphasize the fundamentals and functionality of the language. During this course, we will cover some of the fundamentals, but will also get right into the functionality - we will cover loading data, data manipulation, summarizing data with descriptive statistics, and graphical representation of data with R's plotting features. The goal is that you will be very comfortable interacting with R and will have enough knowledge to start using and learning R on your own. Basic programming knowledge is helpful.

Presenting Data and Information by Edward Tufte, Durham NC, 2009
Description: This one-day course taught by Edward Tufte covers fundamental strategies of analytical design, evaluating evidence used in presentations, credibility of presentations, and strategic graphic design of statistical data, and effective delivery of business, scientific, research, and financial presentations.

*Genetic Analysis of Complex Human Diseases**, University of Miami, 2008
**I took this course as a student in 2008, and taught as an instructor in 2012-2013.*
Description: A comprehensive four-day course directed toward physician-scientists and other medical researchers. The course will introduce state-of-the-art approaches for the mapping and characterization of human inherited disorders with an emphasis on the mapping of genes involved in common and genetically complex disease phenotypes. The overall focus is a broad-based understanding of the problems and solutions to the design and execution of disease gene mapping projects using Human Genome Project resources.

Pharmacogenomics Research Network Analysis Workshop VI, St. Jude Children's Research Hospital, 2008
Description: This is a small group, highly-interactive meeting with short oral presentations and extensive discussion. The purpose is to share state-of-the-art techniques and ideas for analysis of pharmacogenetic/genomic data.

Summer Institute in Statistical Genetics: Computing for Statistical Genetics, University of Washington, 2007
Description: This course introduces software for analysis of genetic data in the R statistical environment. Data management in R, programming concepts for R, and standard regression analyses will be discussed. These topics will be followed by analysis more specific to genetic data, including association analysis, and haplotype inference. Use of the extensive collection of genomics packages from the Bioconductor project will be introduced. Finally, the use of R as an interface to other more specialized "legacy" software will be demonstrated. Reference will be made to current analyses of whole-genome association study data.

Responsible Conduct in Research Course, Vanderbilt University, 2006
Description: This program is designed to help scientists identify and deal with ethical issues and dilemmas, and to promote open discussion of such topics as: institutional and NIH policies regarding scientific misconduct and conflicts of interest, ethical considerations of research involving human and animal subjects, responsible authorship and review of scientific publications and grants, data management, record keeping, and intellectual

property. During RCR, formal lectures on these subjects are complemented by small group discussions and case study analyses, some of which are derived from Scientific Integrity: An Introductory Text with Cases by Francis L. Macrina, American Society for Microbiology; 2nd edition.

COMPUTING AND SOFTWARE EXPERTISE

Bioinformatics:	R, BEDTools, IGV, Annovar, BLAST, Galaxy, BWA/Bowtie, Tophat, Cufflinks, edgeR, DESeq, DEXSeq, limma, Samtools, Picard, Krona, VCFTools, MAKER, GATK, …
Programming:	R, Bash, Perl, Python, MATLAB
Statistical computing:	R, R/Bioconductor, Stata, SAS, SPSS, JMP
Genetic analysis:	R, PLINK, VCFtools, GenABEL, Eigensoft, Haploview, GATK
Other:	SQL, Linux/Unix