BIOGRAPHICAL SKETCH

Provide the following information for the key personnel and other significant contributors in the order listed on Form Page 2.
Follow this format for each person. **DO NOT EXCEED FOUR PAGES.**

NAME Aaron Jon Mackey	POSITION TITLE Assistant Professor of Public Health Sciences
eRA COMMONS USER NAME (credential, e.g., agency login) AJM6QNIH	

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	YEAR(s)	FIELD OF STUDY
Reed College, Portland OR	B.A.	1991-1995	Biochemistry and Molecular Biology
Washington University, St. Louis, MO	M.S.	1995-1998	Immunology
University of Virginia, Charlottesville, VA	Ph.D.	1998-2004	Microbiology/ Computational Biology
University of Pennsylvania, Philadelphia, PA		2004-2006	postdoctoral research

A. Personal Statement

My somewhat eclectic career path has provided me a diverse skillset in experimental and computational biology. I began my graduate research studying autoimmune T cell receptor biology in mouse models of type 1 diabetes, where I used combinatorial peptide libraries to define novel peptide antigens for T cell activation. Later, as a computational graduate student, I developed novel, statistically motivated bioinformatics methods for improved ORF finding and evolutionary sequence analyses in partially finished prokaryotic genome sequence, which today are used for aspects of metagenomic sequence analysis. I moved to a eukaryotic parasitology laboratory for my postdoctoral fellowship, where I developed a now-popular method to combine disparate sources of gene exon/intron structure information to automatically generate a first-pass, consensus structural genome annotation; the GLEAN algorithm has been used for automated annotation in many high-profile eukaryotic genome projects to date, including the honeybee, sea urchin, and flour beetle genomes. I also mentored the development of a similar statistical integration technique to better predict orthologous relationships between proteins. At this time I began work integrating functional genomic datasets (SAGE tags at the time, then spotted cDNA arrays, then various microarray chips) to explore transcriptomic changes related to host-pathogen interactions. I moved to GlaxoSmithKline to undertake an ambitious computational genomics project in their Oncology group, where I was responsible for integrated analyses of accumulated genomics data (transcriptome, SNP-chip, CNVs, driver mutations, phosphoproteomics) on hundreds of cell lines and phase 2 clinical trial patients, with the aim of identifying and validating biomarkers predictive of response to particular therapies in development for phase 1 clinical studies. It was at GSK that I developed my interest in using biological pathways and networks as the means to computationally integrate information across multiple experimental domains, and to identify not just lists of interesting genes, but to directly identify specific mechanistic hypotheses revealed and supported by the available data. As tenure-track faculty at UVA, my research continues to exploit various probabilistic graphical models for the statistical integration of disparate data; we exercise these concepts in many diverse research domains, including genetic variant discovery associated with complex human disease, molecular causal network reconstruction for understanding basic biological mechanisms of action, and consensus eukaryotic gene structure prediction. My recent work with HemoShear has focused on implementing robust, quality-controlled bioinformatics pipelines for large-scale RNAseq-based expression profiling and predictive classification of drug responses in the HemoShear vascular system, including the development of integrative statistical expert systems.

B. Positions and Honors

Positions and Employment

9/2012 – present Director of Quantitative and Computational Sciences HemoShear, LLC, Charlottesville VA

7/2009 – present Assistant Professor of Public Health Sciences Center for Public Health Genomics, University of Virginia, Charlottesville VA
2/2009 – 6/2009 Research Scientist (Contractor)
    Human Genome Sequencing Center, Baylor College of Medicine, Houston TX
3/2006 – 2/2009   Investigator, Computational Biology, GlaxoSmithKline, Collegeville PA
2/2009 – 7/2009   Research Scientist (Contractor), Human Genome Sequencing Center, Baylor College of Medicine, Houston TX
7/2009 – *present* Assistant Professor of Public Health Sciences, Division of Clinical Informatics and Center for Public Health Genomics, University of Virginia, Charlottesville, VA
5/2012 – *present* Director of Quantitative and Computational Sciences, HemoShear LLC, Charlottesville, VA

Other Experience and Professional Memberships

1998 – *present*   member, International Society of Computational Biology

C.   Selected peer-reviewed publications (selected from 27 peer-reviewed publications)

1. Elsik CG. **Mackey AJ**. Reese JT. Milshina NV. Roos DS. Weinstock GM. Creating a honey bee consensus gene list. *Genome Biology* 8(1):R13, 2007. PMCID: 1839126
2. Liu Q. **Mackey AJ**. Pereira FCN. Roos DS. Evigan: a hidden variable model for integrating gene evidence in eukaryotic gene prediction. *Bioinformatics* 24(5):597-605, 2008. PMID: 18187439
3. Mackay TFC. Richards S. Stone EA. Barbadilla A. Ayroles JF. Zhu D. Casillas S. Han Y. Magwire MM. Cridland JM. Richardson MF. Anholt RRH. Barron M. Bess C. Blankenburg KP. Carbone MA. Castellano D. Chaboub L. Duncan L. Harris Z. Javaid M. Jayaseelan JC. Jhangiani SN. Jordan KW. Lara F. Lawrence F. Lee SL. Librado P. Linheiro RS. Lyman RF. **Mackey AJ**. Mundisa M. Muzny DM. Nazareth L. Newsham I. Perales L. Pu L-L. Qu C. Ramia M. Reid JG. Rollmann SM. Rozas J. Turlapati L. Worley KC. Wu Y-Q. Yamamoto A. Zhu Y. Bergman CM. Thornton KR. Mittelman D. Gibbs RA. The Drosophila melanogaster Genetic Reference Panel. *Nature* 482(7384):173-178, 2012 PMID: 22318601
4. Roller DG. Axelrod M. Capaldo BJ. Jensen K. **Mackey A**. Weber MJ. Gioeli D. Synthetic lethal screening with small-molecule inhibitors provides a pathway to rational combination therapies for melanoma. *Molecular Cancer Therapeutics* 11(11), 2012 PMCID: 3496043
5. Marie CS. Verkerke P. Paul SN. **Mackey AJ**. Petri WA Jr. Leptin protects host cells from Entamoeba histolytica cytotoxicity by a STAT3-dependent mechanism. *Infection and Immunity* 80(5), 2012 PMID: 22331430

Additional recent publications of importance to the author's field (in chronological order)

1. Xu P. Widmer G. Wang Y. Ozaki LS. Alves JM. Serrano MG. Puiu D. Manque P. Akiyoshi D. **Mackey AJ**. Pearson WR. Dear PH. Bankier AT. Peterson DL. Tzipori S. Buck GA. The genome of *Cryptosporidium hominis. Nature* 431:1107-1112, 2004. PMID: 15510150
2. Chen F. **Mackey AJ**. Stoeckert CJ. Roos DR. OrthoMCL-DB: Querying A Comprehensive Multi-Species Collection of Ortholog Groups. *Nucleic Acids Research* 34(Database Issue):D363-D368, 2006. PMCID: 1347485
3. Wang H. Su Y. **Mackey AJ**. Kraemer ET. Kissinger JC. SynView: a Gbrowse-compatible approach to visualizing comparative genome data. *Bioinformatics* 22(18):2308-2309, 2006. PMID: 16844709
4. The Honeybee Genome Sequencing Consortium. Insights into social insects from the genome of the honey bee *Apis mellifera. Nature* 443: 931-949, 2006. PMCID: 2048586
5. The Sea Urchin Genome Sequencing Consortium. The genome of the sea urchin *Strongylocentrus purputatus. Science* 314(5801):941-962, 2006. PMID: 17095691
6. Chen F. **Mackey AJ**. Vermunt JK. Roos DR. Assessing performance of orthology detection strategies applied to eukaryotic genomes. *PLoS ONE* 2(4):e383, 2007. PMCID: 1849888
7. Lapp H. Bala S. Balhoff JP. Bouck A. Goto N. Holder M. Holland R. Holloway A. Katayama T. Lewis PO. **Mackey AJ**. Osborne BI. Piel WH. Kosakovsky Pond SL. Poon A. Qui W-G. Stajich JE. Stoltzfus A. Thierer T.

Vilella AJ. Vos RA. Zmasek CM. Zwickl D. Vision TJ. The 2006 NESCent Phyloinformatics Hackathon: A Field Report. *Evolutionary Bioinformatics* 3: 357-366, 2007. PMCID: 2684128

8. The *Tribolium* Genome Sequencing Consortium. The genome of the model beetle and pest *Tribolium castaneum*. *Nature* 452: 949-955, 2008. PMID: 18362917

9. Gajria B. Bahl A. Brestelli J. Dommer J. Fischer S. Gao X. Heiges M. Iodice J. Kissinger JC. **Mackey AJ**. Pinney DF. Roos DS. Stoeckert CJ Jr. Wang H. Brunk BP. ToxoDB: an integrated *Toxoplasma gondii* database resource. *Nucleic Acids Research* 36: D553-D556, 2008. PMCID: 2238934

10. Coghlan A. Fiedler TJ. McKay SJ. Flicek P. Harris TW. Blasair D. The nGASP Consortium. Stein LD. nGASP – the nematode genome annotation assessment project. *BMC Bioinformatics* 9:549, 2008 PMCID: 2651883

11. Stajich JE, Wilke SK, Ahrén D, Au CH, Birren BW, Borodovsky M, Burns C, Canbäck B, Casselton LA, Cheng CK, Deng J, Dietrich FS, Fargo DC, Faraman ML, Gathman AC, Goldberg J, Guigó R, Hoegger PJ, Hooker JB, Huggins A, James TY, Kamada T, Kilaru S, Kodira C, Kües U, Kupfer D, Kwan HS, Lomsadze A, Li W, Lilly WW, Ma LJ, **Mackey AJ**, Manning G, Martin F, Muraguchi H, Natvig DO, Palmerini H, Ramesh MA, Rehmeyer CJ, Roe BA, Shenoy N, Stanke M, Ter-Hovhannisyan V, Tunlid A, Velagapudi R, Vision TJ, Zeng Q, Zolan ME, Pukkila PJ. Insights into evolution of multicellular fungi from the assembled chromosomes of the mushroom *Coprinopsis cinerea* (*Coprinus cinereus*). *Proceedings of the National Academy of Sciences of the United States of America*.107(26), 2010 PMID: 20547848

12. Huang K. Mellor KE. Paul SN. Lawson MJ. **Mackey AJ**. Timko MP. Global changes in gene expression during compatible and incompatible interactions of cowpea (*Vigna unguiculate L.*) with the root parasitic angiosperm Striga gesnerioides. *BMC Genomics* 13(1):402, 2012 PMID: 22900582

D. Invited Talks and Seminars

"Genetic mouse models of Type I Diabetes", FASEB/AAI (San Francisco, California), July 1995

"Cotransfer of DO11.10 T cells with BDC2.5 T cells prevents diabetes", FASEB/AAI (San Francisco, California), May 1998

"Managing Database Searches in BioPerl", BioPerl/ISMB (Heidelberg, Germany), July 1999

"Getting more from less: algorithms for rapid protein identification with multiple short peptide sequences", Sanger Center (Hinxton, UK); Cambridge University Dept. of Biochemistry (Cambridge, UK); Imperial College Dept. of Computer Science (London, UK), March 2001

"Integrated Data Collection, Preparation and Analysis using Perl and R", UVA Workshop on Reproducible Statistical Reporting (Charlottesville, Virginia), October 2002

"Making the most of what you've got: the utility of unfinished genomes", Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biology Laboratory (Woods Hole, MA), March 2004

"The genome of Toxoplasma gondii and the new ToxoDB 3.0", Gordon Research Conference: Biology of Host-Parasite Interactions (Newport, RI), July 2004

 "Phylogenetic character state data management with CDATlib", CSHL Generic Model Organism Database Workshop, Cold Spring Harbor Labs (Cold Spring Harbor, NY), October 2004

 "Improved eukaryotic gene identification via statistical consensus of gene evidence", Greater Philadelphia Bioinformatics Alliance Retreat (Great Valley, PA), October 2004

 "Improving the structural annotation of Toxoplasma gondii by statistical integration of newly available evidence", Molecular Parasitology Meeting (Woods Hole, MA), September 2005

"*GLEAN – improved eukaryotic gene prediction by statistical consensus of gene evidence*", Genome Informatics (Cold Spring Harbor, NY), October 2005
"*GLEAN – improved eukaryotic gene prediction by statistical consensus of gene evidence*", Human Genome Sequencing Center, Baylor College of Medicine (Houston, TX), May 2006

"*Progress Report: Targeted 454-Resequencing of 1000 T1D Cases and Controls in Four GWAS Regions*", Wellcome Trust (London, UK), June 2010

"*In Search of New Biology: from Functional Genomics to Mechanistic Hypotheses*", Sweet Briar College Biology department (Sweet Briar, VA), September 2010

"*Improved Variant Discovery and Allele Frequency Estimation from Pooled DNA Resequencing with Bayesian Latent Class Analysis and Compositional Bias Models*", RECOMB 2011 Satellite Workshop on RECOMB-seq (Vancouver, BC), March 2011

"*Top-down genomics: generating testable mechanistic hypotheses from genome-wide studies*", UMASS Amherst (Amherst, MA), Feb 2011
"*ENIGMA: tools for comparative, consensus gene structure annotation*", Biocuration 2012 (Washington, D.C.), April 2012

"*Mechanisms and markers of melanoma resistance to combination therapies*", Washington University in St. Louis, Laboratory and Genomic Medicine (St. Louis, MO), March 2013

E. Teaching

Teaching assistant for undergraduate *"General Biology"* at Washington University (St. Louis, Missouri), Spring 1997

Teaching assistant for the *"Computational Genomics"* course at Cold Spring Harbor Laboratory (Cold Spring Harbor, New York), November 2001, 2002

Coinstructor of *"Learning Perl for Bioinformatics"* at University of Virginia (Charlottesville, Virginia), Spring 2002

Coinstructor of *"Relational Databases for Biologists"* tutorial at ISMB 2002 (Edmonton, Alberta), August 2002

Instructor of *"Relational Databases for Bioinformaticists"* tutorial at the O'Reilly Bioinformatics Technology Conference (San Diego, California), February 2003

Lecturer, *"Genome Informatics"* course at Cold Spring Harbor Laboratories (Cold Spring Harbor, New York), October 2003, 2004.

F. Research Support

Ongoing Research Support

Young Investigator (Mackey)   05/01/11-04/30/14
Melanoma Research Alliance
This project aims to employ whole exome sequencing and DNA methylation studies to identify genetic and/or epigenetic variants that alter sensitivity to combinatorial drug therapy in a panel of human melanoma cell lines.
Role: Principal Investigator

NSF Large Proposal (Evans) 08/01/2011-07/31/2016
NSF
Trustworthy Computing

The primary goal of this project is to develop and employ secure two-way computations, using strongly encrypted message passing protocols that ensure no private information leakage beyond that revealed by the computational result.
Role: Co-Investigator

Completed Research Support

NIH; 5U54 HG003273-06 (subaward)
Genomes and Genetics at the BCM-HGSC
PI: Richard Gibbs
Role: Co-investigator

R21 HL104596-01A1 (Hall)   04/01/11-03/31/13
University of Minnesota
Genetic risk factors of myocyte contractility in human heart failure
Role: Consortium PI