

ROBERT B. SCHARPF ¹

Johns Hopkins Bloomberg School of Public Health
615 N. Wolfe St, Rm E3537 · Baltimore, MD 21205-2179

Phone: 410-614-7837 · Email: rscharpf@jhsph.edu

Website: <http://www.biostat.jhsph.edu/people/postdocs/scharpf.shtml>

EDUCATION

- 2007- JOHNS HOPKINS BLOOMBERG SCHOOL OF PUBLIC HEALTH Baltimore, MD
Postdoctoral Fellow, Biostatistics
Mentors: Ingo Ruczinski, PhD (primary) · Josef Coresh, MD, PhD (co-mentor)
- 2002-2007 JOHNS HOPKINS BLOOMBERG SCHOOL OF PUBLIC HEALTH Baltimore, MD
PhD, Biostatistics
Advisors: Giovanni Parmigiani, PhD (primary) · Ingo Ruczinski, PhD (co-advisor)
Dissertation: *Combining high-throughput genomic data: methods and applications*
- 1995-1998 JOHNS HOPKINS UNIVERSITY KRIEGER SCHOOL OF ARTS AND Baltimore, MD
SCIENCES
MS, Biotechnology
- 1990-1994 JAMES MADISON UNIVERSITY Harrisonburg, VA
BS *Magna Cum Laude*, Biology

HONORS AND AWARDS

- 2008 Young Investigator Travel Award, Markov Chain Monte Carlo in Theory and Practice (“MCM-Ski”) Conference, the 2nd International Joint Meeting of the Institute of Mathematical Statistics and the International Society for Bayesian Analysis, Bormio, Italy
- 2007 Margaret Merrell Award for Excellence in Research, Johns Hopkins Bloomberg School of Public Health
- 2007 Helen Abbey Award for Excellence in Teaching, Johns Hopkins Bloomberg School of Public Health
- 2006 Travel Scholarship sponsored by the American Statistical Association, Bioconductor Conference, Seattle, WA
- 2006 Nominated for Teaching Assistant Recognition Award, Johns Hopkins Bloomberg School of Public Health
- 2001 Research Citation Finalist, Anesthesiology and Critical Care Medicine Conference, San Francisco, CA
- 1994 Finalist for Excellence in Biology Award, James Madison University

PROFESSIONAL EXPERIENCE

- 2007-2008 *Visiting Assistant Professor*
Department of Health Sciences, American University of Armenia Yerevan, Armenia
- 2006-2007 *Consultant* for Department of Health Policy and Management
Department of Biostatistics, Johns Hopkins School of Public Health Baltimore, MD
- 2002-2007 *Predoctoral Research*
Sidney Kimmel Cancer Center, Johns Hopkins School of Medicine Baltimore, MD
- 1998-2002 *Laboratory Research Assistant*
Department of Anesthesiology, Johns Hopkins School of Medicine Baltimore, MD
- 1995-1998 *Informatics Research Assistant*
Division of Biomedical Sciences, Johns Hopkins School of Medicine Baltimore, MD

¹Updated July 30, 2008

TEACHING EXPERIENCE

PRIMARY INSTRUCTOR

Department of Health Sciences, American University of Armenia

2007 (Fall) Inferential Biostatistics I

2008 (Fall) Modeling and Sampling

CO-INSTRUCTOR

Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health

2006 (Summer) The Analysis of Gene Expression Data

TEACHING ASSISTANT

Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health

2005-2007 Lead Teaching Assistant for Statistical Methods in Public Health I and II

2005 Generalized Linear Models III, IV

2004 Introduction to Probability I, II

2003-2004 Statistical Methods in Public Health I, II, III, IV

PUBLICATIONS

PEER-REVIEWED ARTICLES

1. **Scharpf RB**, Tjelmeland H, Parmigiani G, and Nobel AB (2008). A Bayesian model for cross-study differential gene expression, *Journal of the American Statistical Association*, (in press) [pdf].
2. **Scharpf RB**, Parmigiani G, Pevsner J, and Ruczinski I (2008). Hidden Markov models for the assessment of chromosomal alterations using high-throughput SNP arrays. *Annals of Applied Statistics*, 2(2) 687-713 [pdf].
3. Zahurak ML, Parmigiani G, Yu W, **Scharpf RB**, Berman D, Schaeffer E, Shabbeer S, and Cope L (2007). Pre-processing Agilent microarray data. *BMC Bioinformatics*, 8:142 [pdf].
4. **Scharpf RB**, Ting JC, Pevson J, and Ruczinski I (2007). SNPchip: R classes and methods for SNP array data. *Bioinformatics*, 23(5):627-628 [pdf].
5. **Scharpf RB**, Iacobuzio-Donahue CA, Sneddon JB, and Parmigiani G (2007). When should one subtract background fluorescence in two color microarrays? *Biostatistics*; 8(4):695-707 [pdf].
6. Morrell CN, Matsushita K, Chiles K, **Scharpf RB**, Yamakuchi M, Mason RJ, Bergmeier W, Mankowski JL, Baldwin WM 3rd, Faraday N, Lowenstein CJ (2004). Regulation of platelet granule exocytosis by S-nitrosylation. *Proceedings of the National Academy of Sciences*, 102:3782-7 [pdf].
7. Faraday N, Martinez EA, **Scharpf RB**, Kasch-Semenza L, Dorman T, Pronovost PJ, Perler B, Thompson RE, Gerstenblith G, Bray PF, Fleisher LA (2004). Platelet gene polymorphisms and perioperative cardiac risk assessment in vascular surgical patients. *Anesthesiology*, 101(6):1291-7 [PubMed].
8. **Scharpf R**, Garrett ES, Hu J, Parmigiani G. Statistical modeling and visualization of molecular profiles in cancer. *Biotechniques*, 34:S22-S29, 2003 [pdf].
9. Faraday N, Gullar E, Sera VA, Bolton ED, **Scharpf RB**, Cartarius AM, Emery K, Concord J, Kickler TS. Utility of whole blood hemostatology using the clot signature analyzer for assessment of hemostasis in cardiac surgery. *Anesthesiology*, 96(5):1115-22, 2002 [PubMed].
10. Faraday N, **Scharpf RB**, Dodd-o JM, Martinez EA, Rosenfeld BA, Dorman T. Leukocytes can enhance platelet-mediated aggregation and thromboxane release via interaction of P-selectin glycoprotein ligand 1 with P-selectin. *Anesthesiology*, 94(1):145-51, 2001 [PubMed].
11. Anagnostopoulos AV, **Scharpf RB**. It's a knockout! *Trends in Genetics*. 14(2):82, Feb 1998 [PubMed].
12. Anagnostopoulos AV, **Scharpf RB**. It's a knockout! *Trends in Genetics*. 13(12):499-500, Dec 1997 [PubMed].

BOOK CHAPTERS

1. **Scharpf RB** and Garrett-Mayer E. Models for Probability of Under and Overexpression: The POE Scale. In *Bayesian Inference for Gene Expression and Proteomics* (Kim-Anh Do, Peter Mueller, and Marina Vannucci), *Cambridge University Press*, 2006
2. **Scharpf RB** and Ruczinski I. R classes and methods for SNP array data. In *Bioinformatics in Clinical OMICs Research* (John Walker and Rune Matthiesen), *Humana Press* (in press).

PAPERS UNDER REVIEW

1. **Scharpf RB**, Parmigiani G, Nobel AB, and Tjelmeland H (2008). XDE: R software for cross-study analysis of differential gene expression using a Bayesian hierarchical model, Johns Hopkins University, Department of Biostatistics.
2. Caffo B, Liu D, **Scharpf RB**, and Parmigiani G (2008). Likelihood estimation of conjugacy relationships in linear models with applications to high-throughput genomics, Technical Report 62, Johns Hopkins University, Department of Biostatistics [[pdf](#)].

PRESENTATIONS

INVITED TALKS

- *Estimating copy number and identifying regions of copy number variants using high density SNP platforms*, Second Annual Young Investigator Symposium on Genomics and Bioinformatics, Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD (Scheduled for September 5, 2008).
- *Estimation problems in high throughput SNP platforms*, Genome-Wide Association Studies Symposium, Graduate Institute in Epidemiology and Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD (June 16, 2008).
- *Estimating copy number from high-throughput Affymetrix SNP platforms*. The Expressionists Seminar Series, Department of Biostatistics, Johns Hopkins University, Baltimore, MD (April 30, 2008).
- *Assessing copy number alterations in high throughput SNP platforms*, Welch Center for Prevention Epidemiology and Research, Johns Hopkins University, Baltimore, MD (March 21, 2008).
- *A Bayesian model for cross-study differential gene expression*. MCMSki II, Bormio, Italy (January 11, 2008).
- *A Bayesian model for cross-study differential gene expression*. 39th Conference on Interface of Statistics and Computing Science, Philadelphia, PA (May 24, 2007).
- *Combining estimates of copy number and genotype in high-throughput SNP chips*. Eastern North American Region (ENAR) International Biometric Society Conference, Atlanta, GA (March 12, 2007).
- *Visualizing and analyzing high density SNP Data with SNPscan*. Biostatistics Departmental Retreat, Baltimore, MD (May, 2006).
- *Visualizing and analyzing high density SNP data with SNPscan*. ENAR Conference, Tampa, FL (March 28, 2006).

OTHER TALKS

- *Assessing copy number alterations using high throughput SNP platforms*. American Heart Association, 48th Cardiovascular Disease Epidemiology and Prevention Conference, Colorado Springs, CO (March 12, 2008).
- *Genotype calling: uncertainty and copy number*. Atherosclerosis Risk in Communities (ARIC) Study Workshop, Houston, TX (June, 2007).
- *Combining high-throughput genomic data: methods and applications*. Thesis defense, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD (February 07, 2007).

- *Leukocytes can enhance platelet-mediated aggregation and thromboxane release via interaction of P-selectin glycoprotein ligand 1 with P-selectin.* Anesthesiology and Critical Care Medicine Conference, San Francisco, CA (February, 2001)

POSTERS

- *Hidden Markov models for characterizing chromosomal alterations with high-throughput SNP arrays.* 3rd Annual Postdoctoral Fellows' Research Poster Competition, Johns Hopkins Bloomberg School of Public Health (June, 2007).
- *Cross-linking of Fcg RIIa and Fcg RIIIb on human polymorphonuclear neutrophils inhibits platelet aggregation.* Anesthesiology/Critical Care Medicine Poster Presentation Day (October, 2001).

SOFTWARE

PEER-REVIEWED R PACKAGES AVAILABLE AT [BIOCONDUCTOR](#):

- SNPchip** **Scharpf RB** and Ruczinski I. Classes and methods for visualization of high-throughput SNP chip data.
- VanillaICE** **Scharpf RB** and Ruczinski I. Hidden Markov Models for identifying chromosomal aberrations using high-throughput SNP chip data.
- XDE** **Scharpf RB**, Parmigiani G, Nobel AB, and Tjelmeland H. A Bayesian multi-level model for identifying differentially expressed genes across multiple studies.

RESEARCH SUPPORT

ONGOING RESEARCH SUPPORT

- 2008 - present *Institute for Clinical and Translational Research* (NIH/NCRR)
Principal Investigator: Daniel Ford, MD.
- 2007 - present *Cardiovascular Disease and Epidemiology Fellowship* (NHLBI)
Director: Josef Coresh, MD, PhD.

PENDING REVIEW

- Submitted June, 2008 *K99-R00 Pathways to Independence Award* (NHGRI)
Principal Investigator: Robert B. Scharpf, PhD
Mentors: Ingo Ruczinski, PhD (primary) · Josef Coresh, MD, PhD (co-mentor)

PREVIOUS SUPPORT

- 2004, 2006-2008 *Bioinformatics training grant* (NSF)
Principal Investigator: Giovanni Parmigiani, PhD
- 2004-2006 *Environmental Biostatistics and Epidemiology Training Grant* (NIH)
Principal Investigator: Thomas Louis, PhD

SERVICE

PEER REVIEW ACTIVITIES:

- Bioinformatics (2) · Biostatistics (1) · BMC Genomics (1) · Computational Statistics and Data Analysis (2)
- Genomics (1) · Journal of the American Statistical Association (1) · Journal of Statistical Software (1)
- Statistical Methods and Applications (1)

ADDITIONAL INFORMATION

PERSONAL STATEMENT OF RESEARCH OBJECTIVES

I am interested in developing statistical methods and software for characterizing diseases and complex traits from DNA and RNA measured on a genomic scale. Two technologies that provide high throughput measurements of genomic information are the focus of my current research: single nucleotide polymorphism (SNP)

and gene expression arrays. For SNP arrays, I am developing statistical methods to estimate copy number at a single locus, and to infer regions of copy number variants (CNV) spanning multiple loci. In addition, I am developing methods to assess the association of CNV with phenotypes in genome-wide association studies (GWAS). For gene expression data, I am developing methods for the cross-study analysis of expression data to identify concordant and discordant patterns of differential gene expression. The software developed for each of these methods facilitate reproducible research, unify the steps in complex analyses of genomic data, and are freely available under the GNU General Public Licence of the Free Software Foundation.