Curriculum Vitae
The Johns Hopkins University School of Medicine

Robert B. Scharpf

February 22, 2014

Demographic Information

Current Appointments

University	Assistant Professor, Department of Oncology
	Johns Hopkins University
	Baltimore, MD

Joint Appointments

	Assistant Professor, Department of Biostatistics
	Johns Hopkins Bloomberg School of Public Health
	Baltimore, MD

Personal Data

Office	Johns Hopkins University
	550 N. Broadway, Suite 1103
	Baltimore, MD 21205-2179
	Phone: 443-287-9408
	FAX: 410-955-0859
	Email: rscharpf@jhsph.edu

Education and Training

Year	Degree	Institution	Discipline
1990-1994	BS	James Madison University	Biology
1998-2002	MS	Johns Hopkins University Krieger School of Arts & Sciences	Biotechnology
2002-2007	PhD	Johns Hopkins Bloomberg School of Public Health	Biostatistics
2007-2010	n/a	Johns Hopkins Bloomberg School of Public Health	Postdoctoral Fellow

Professional Experience

1995 - 1998	Informatics Research Assistant	Johns Hopkins School of Medicine, Baltimore, MD
1998 - 2002	Laboratory Research Assistant	Johns Hopkins School of Medicine, Baltimore, MD
2006 - 2007	Statistical Consultant for PhD Students	Johns Hopkins School of Public Health, Baltimore, MD
2007 - 2008	Visiting Assistant Professor	American University of Armenia, Yerevan, Armenia
2007 - 2010	Postdoctoral fellow	Johns Hopkins University School of Public Health, Baltimore, MD
2010 -	Assistant Professor	Johns Hopkins School of Medicine, Baltimore, MD

Research Activities

Peer-Reviewed Articles

[1] Faraday N, Scharpf RB, Dodd-o JM, Martinez EA, Rosenfeld BA, and Dorman T. Leukocytes can enhance platelet-mediated aggregation and thromboxane release via interaction of P-selectin glycoprotein ligand 1 with P-selectin. Anesthesiology, 2001;94(1):145–151.

[2] Faraday N, Guallar E, Sera VA, Bolton ED, Scharpf RB, Cartarius AM, Emery K, Concord J, and Kickler TS. Utility of whole blood hemostatometry using the clot signature analyzer for assessment of hemostasis in cardiac surgery. Anesthesiology, 2002;96(5):1115–1122.

[3] Scharpf RB, Garrett ES, Hu J, and Parmigiani G. Statistical modeling and visualization of molecular profiles in cancer. Biotechniques, 2003;Suppl:22–29.

[4] Faraday N, Martinez EA, Scharpf RB, Kasch-Semenza L, Dorman T, Pronovost PJ, Perler B, Gerstenblith G, Bray PF, and Fleisher LA. Platelet gene polymorphisms and cardiac risk assessment in vascular surgical patients. Anesthesiology, 2004;101(6):1291–1297.

[5] Morrell CN, Matsushita K, Chiles K, Scharpf RB, Yamakuchi M, Mason RJA, Bergmeier W, Mankowski JL, Baldwin WM, Faraday N, and Lowenstein CJ. Regulation of platelet granule exocytosis by S-nitrosylation. Proc Natl Acad Sci U S A, 2005;102(10):3782–3787.

[6] Scharpf RB, Iacobuzio-Donahue CA, Sneddon JB, and Parmigiani G. When should one subtract background fluorescence in 2-color microarrays? Biostatistics, 2007;8(4):695–707.

[7] Scharpf RB, Ting JC, Pevsner J, and Ruczinski I. SNPchip : R classes and methods for SNP array data. Bioinformatics, 2007;23(5):627–628.

[8] Zahurak M, Parmigiani G, Yu W, Scharpf RB, Berman D, Schaeffer E, Shabbeer S, and Cope L. Pre-processing agilent microarray data. BMC Bioinformatics, 2007;8:142. PMCID: PMC1876252.

[9] Scharpf RB, Parmigiani G, Pevsner J, and Ruczinski I. Hidden Markov models for the assessment of chromosomal alterations using high-throughput SNP arrays. Annals of Applied Statistics, 2008;2(2):687–713. PMCID: PMC2710854.

[10] Caffo BS, Liu D, Scharpf RB, and Parmigiani G. Likelihood estimation of conjugacy relationships in linear models with applications to high-throughput genomics. The International Journal of Biostatistics, 2009;5(1).

[11] Carvalho B, Irizarry RA, Scharpf RB, and Carey VJ. Processing and analyzing Affymetrix SNP chips with Bioconductor. Statistics in Biosciences, 2009;1(2):160–180.

[12] Köttgen A, Glazer NL, Dehghan A, Hwang SJ, Katz R, Li M, Yang Q, Gudnason V, Launer LJ, Harris TB, Smith AV, Arking DE, Astor BC, Boerwinkle E, Ehret GB, Ruczinski I, Scharpf RB, Chen YDI, de Boer IH, Haritunians T, Lumley T, Sarnak M, Siscovick D, Benjamin EJ, Levy D, Upadhyay A, Aulchenko YS, Hofman A, Rivadeneira F, Uitterlinden AG, van Duijn CM, Chasman DI, Par G, Ridker PM, Kao WHL, Witteman JC, Coresh J, Shlipak MG, and Fox CS. Multiple loci associated with indices of renal function and chronic kidney disease. Nat Genet, 2009;41(6):712–717.

[13] Levy D, Ehret GB, Rice K, Verwoert GC, Launer LJ, Dehghan A, Glazer NL, Morrison AC, Johnson AD, Aspelund T, Aulchenko Y, Lumley T, Köttgen A, Vasan RS, Rivadeneira F, Eiriksdottir G, Guo X, Arking DE, Mitchell GF, Mattace-Raso FUS, Smith AV, Taylor K, Scharpf RB, Hwang SJ, Sijbrands EJG, Bis J, Harris TB, Ganesh SK, O'Donnell CJ, Hofman A, Rotter JI, Coresh J, Benjamin EJ, Uitterlinden AG, Heiss G, Fox CS, Witteman JCM, Boerwinkle E, Wang TJ, Gudnason V, Larson MG, Chakravarti A, Psaty BM, and van Duijn CM. Genome-wide association study of blood pressure and hypertension. Nat Genet, 2009;41(6):677–687. PMCID: PMC2998712.

[14] Scharpf RB, Tjelmeland H, Parmigiani G, and Nobel A. A Bayesian model for cross-study differential gene expression. JASA, 2009;104(488):1295–1310. PMCID: PMC2994029.

[15] Scharpf RB, Tjelmeland H, Parmigiani G, and Nobel A. Rejoinder. JASA, 2009;104(488):1318–1323.

[16] Leek JT, Scharpf RB, Bravo HC, Simcha D, Langmead B, Johnson WE, Geman D, Baggerly K, and Irizarry RA. Tackling the widespread and critical impact of batch effects in high-throughput data. Nat Rev Genet, 2010; 11(10):733–739. PMCID: NIHMS:322071.

[17] Scharpf RB, Iacobuzio-Donahue CA, Cope L, Ruczinski I, Garrett-Mayer E, Lakkur S, Campagna D, and Parmigiani G. Cross-platform comparison of two pancreatic cancer phenotypes. Cancer Inform, 2010;9:257–264. PMCID: PMC2978933.

[18] Scharpf RB and Ruczinski I. R classes and methods for SNP array data. Methods Mol Biol, 2010;593:67–79. PMCID: PMC2853754.

[19] Halper-Stromberg E, Frelin L, Ruczinski I, Scharpf R, Jie C, Carvalho B, Hao H, Hetrick K, Jedlicka A, Dziedzic A, Doheny K, Scott AF, Baylin S, Pevsner J, Spencer F, and Irizarry RA. Performance assessment of copy number microarray platforms using a spike-in experiment. Bioinformatics, 2011;27(8):1052–1060. PMCID: PMC3072561.

[20] Khachatryan L, Scharpf RB, and Kagan S. Influence of diabetes mellitus type 2 and prolonged estrogen exposure on risk of breast cancer among women in Armenia. Health Care for Women International, 2011;To appear.

[21] Scharpf RB, Irizarry RA, Ritchie ME, Carvalho B, and Ruczinski I. Using the R package crlmm for genotyping and copy number estimation. Journal of Statistical Software, 2011;40(12):1–32. ISSN 1548-7660. PMCID: PMC3329223.

[22] Scharpf RB, Ruczinski I, Carvalho B, Doan B, Chakravarti A, and Irizarry RA. A multilevel model to address batch effects in copy number estimation using SNP arrays. Biostatistics, 2011;12(1):33–50. PMCID: PMC3006124.

[23] Laurie CC, Laurie CA, Rice K, Doheny KF, Zelnick LR, McHugh CP, Ling H, Hetrick KN, Pugh EW, Amos C, Wei Q, Wang LE, Lee JE, Barnes KC, Hansel NN, Mathias R, Daley D, Beaty TH, Scott AF, Ruczinski I, Scharpf RB, Bierut LJ, Hartz SM, Landi MT, Freedman ND, Goldin LR, Ginsburg D, Li J, Desch KC, Strom SS, Blot WJ, Signorello LB, Ingles SA, Chanock SJ, Berndt SI, Le Marchand L, Henderson BE, Monroe KR, Heit JA, de Andrade M, Armasu SM, Regnier C, Lowe WL, Hayes MG, Marazita ML, Feingold E, Murray JC, Melbye M, Feenstra B, Kang JH, Wiggs JL, Jarvik GP, McDavid AN, Seshan VE, Mirel DB, Crenshaw A, Sharopova N, Wise A, Shen J, Crosslin DR, Levine DM, Zheng X, Udren JI, Bennett S, Nelson SC, Gogarten SM, Conomos MP, Heagerty P, Manolio T, Pasquale LR, Haiman CA, Caporaso N, and Weir BS. Detectable clonal mosaicism from birth to old age and its relationship to cancer. Nat Genet, 2012;44(6):642–650.

[24] Scharpf RB, Beaty TH, Schwender H, Younkin SG, Scott AF, and Ruczinski I. Fast detection of de novo copy number variants from SNP arrays for case-parent trios. BMC Bioinformatics, 2012;13(1):330. PMCID: PMC3576329.

Ph.D. Thesis

[25] Scharpf RB. *Combining High-Throughput Data: Methods and Applications*. Ph.D. thesis, Johns Hopkins University, 2007.

Extramural Sponsorship (current, pending, previous)

Current

04/01/2013 - 10/01/2013	R00 Pathways to Independence Award R00HG005015 NHGRI $23,000 PI: Scharpf/ 14%
05/07/1997 - 04/30/2012	Regional Oncology Research Center: Bioinformatics Core P30CA006973 NCI $4,613,214 PI: Nelson/Ochs Role: Statistician/ 1%
04/01/2011 - 03/31/2015	Validation and Fine-Scale Mapping of Pancreatic Cancer Susceptibility Loci R01CA154823 NCI $2,246,569 Role: Statistician/ 4%

09/01/2009 - 07/31/2014   TGF-beta Signaling in Pancreatic Cancer Progression
R01CA140599
NCI
$201,275
PI: Iacoubzio-Donahue
Role: Statistician 5%

12/01/2011 – 11/30/2014   Software for the statistical analysis of microarray probe level data
R01RR021967
NIH
$231,294
PI: Irizarry
Role: Co-investigator 5%

09/01/2007-05/31/2012   Institutional Clinical and Translational Science Award
U54RR023561
$15,378,996
PI: Ford
Role: Statistician/ 5%

PREVIOUS

07/01/2010 - 04/01/2013   R00 Pathways to Independence Award
R00HG005015
NHGRI
$249,000
PI: Scharpf/ 75%

09/30/2009 - 08/31/2011   Next-generation medical resequencing of gout disease genes in the ARIC cohort
RC2HG005697
UTHSCH/NHGRI
$370,710
PI: Hixon/Kao
Role: Statistician/ 2%

04/01/2009-04/01/2010   K99 Pathways to Independence Award
K99HG005015
NHGRI
$83,000
PI: Scharpf / 83%

03/01/2007-03/01/2009   Cardiovascular Disease and epidemiology Fellowship
NHLBI
$687,297
PI: Coresh
Role: Trainee/100%

07/12/2004-06/30/2009   Environmental Biostatistics and Epidemiology Training Grant
NIEHS
$181,891
PI: Louis
Role: Trainee/100%

2004, 2006-2008   Bioinformatics training grant
T32GM074906
NSF
$248,856
PI: Parmigiani
Role: trainee/100%

Educational Activities

Primary Instructor
Department of Health Sciences, University of Armenia

2007 (Fall)	Inferential Biostatistics I
2008 (Fall)	Modeling and Sampling
2013 (Spring)	Bayesian Methods I

Co-instructor
Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health

2006 (Summer)	The Analysis of Gene Expression Data
2009 (Fall)	Lab for Statistical Methods in Public Health I and II
2010 (1st term)	Lab for Statistical Methods in Public Health I
2011 (Spring)	Programming in R
2011 (Fall)	Programming in R
2012 (Fall)	Programming in R
2012 (Spring)	Lab for Statistical Methods in Public Health III
2013 (Spring)	Programming in R

Teaching Assistant
Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health

2005-2007	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-IV

Guest Lectures
Department of Biostatistics, JHSPH

09/17/2009	Learning LaTeX	140.776 Statistical Computing
11/26/2009	SNP arrays, genotyping, copy number variations	140.668 Special Topics In Genomics
03/08/2010	Measuring copy number variation	140.668 Special Topics in Genomics
03/10/2010	Variation in the Human Genome and Disease	140.668 Special Topics in Genomics
06/30/2010	Statistical Approaches for CNV Analysis	Summer Epi & Biostat Institute
11/02/2010	ANOVA	140.621 Statistical Methods in Public Health II
11/04/2010	Simple Linear Regression	140.621 Statistical Methods in Public Health II
02/04/2011	Multiple Random Variables	140.615 Statistics for Laboratory Scientists
03/14/2011	ANOVA	140.615 Statistics for Laboratory Scientists
09/2011	Learning LaTeX	140.776 Statistical Computing
11/2011	Linear Regression	140.622 Statistical Methods in Public Health II
02/2012	Survival Analysis	140.623 Statistical Methods in Public Health III

Department of Epidemiology, JHSPH

10/2011	Copy number estimation	Genetic Epidemiology

Summer Institute, JHSPH

06/2011	Copy number estimation	Molecular Biology for Genetic Epidemiology

FAES Workshop, National Institutes of Health

05/25/2011	Measuring DNA copy number from SNP arrays	Advanced Microarray Analysis and Pathway Integration
03/01/2013	Confidence Intervals for Proportions	140.615 Statistics for Laboratory Scientists
03/25/2013	Goodness of Fit	140.615 Statistics for Laboratory Scientists

Mentoring

Dissertation Reader for MPH Candidates
2009 Kushagra Mahansoria   American University of Armenia
2009 Lilit Khachatryan   American University of Armenia
2009 Mikhayil Melikou   American University of Armenia
2013 Saki Takahashi   Johns Hopkins School of Public Health

Summer research students
2010 Kinjal Basu   Topic: copy number estimation   co-advised with Aravinda Chakravarti
2010 Jack Yu   Topic: copy number estimation   co-advised with Sarah Wheelan and Jonathan Pevsner

Research assistant
2010 - Lynn Mireless   Research area: copy number analysis for ARIC cohort study   co-advised with Linda Kao

Postdoctoral fellows
2012 - Rumen Kostadinov   Clonality & evolution of cancers   co-advised with Rafael Irizarry and Patrick Brown

Former students

		Current Position	Job Title
2010 - 2011	Moiz Bootwalla	University of Southern California	Programmer Analyst III

THESIS COMMITTEES

Date	PhD Candidate	Activity	Advisor
December, 2010	Eric Stevens	Committe Meeting	Jonathan Pevsner
May, 2010	Joseph Baugher	Committe Meeting	Jonathan Pevsner
December, 2011	Eric Stevens	Committe Meeting	Jonathan Pevsner
September, 2012	Eric Stevens	Committe Meeting	Jonathan Pevsner
June, 2012	Joseph Baugher	Committe Meeting	Jonathan Pevsner
December, 2012	Josepf Baugher	Dissertation reader	Jonathan Pevsner

EDITORIAL ACTIVITIES

PEER-REVIEW FOR JOURNALS
2013: Statistics in Medicine (1)
2012: Bioinformatics (1) · Computational Statistics and Data Analysis (1) · IEEE Transactions in Computational Biology and Bioinformatics (1) · PLOS One (2)
2011: Bioinformatics (2) · BMC Bioinformatics (1) · Journal of Statistical Software (1) · Nucleic Acids Research (1) · Statistical Applications in Genetics and Molecular Biology (1)
2010: BMC Bioinformatics (3) · Bioinformatics (3) · Biostatistics (1) · SAGMB (1) · Pharmacogenomics (1)
2006-2009: American Journal of Epidemiology (2) · Bioinformatics (3) · Biometrics (1) · Biostatistics (2) · BMC Bioinformatics (1) · BMC Genomics (1) · Computational Statistics and Data Analysis (2) · Genomics (1) · JASA (2) · Statistics in Medicine (1) · Statistical Methods and Applications (1)

PEER-REVIEW FOR SOFTWARE: Bioconductor Project (1)

PEER-REVIEW FOR CONFERENCES/SYMPOSIA: Pacific Symposium on Biocomputing (1) PEER-REVIEW FOR CONSORTIA: Atherosclerosis Risk in Compmunities Publications Committee (2)

ORGANIZATIONAL ACTIVITIES

ORGANIZER OF INVITED SESSION
 March, 2010   ENAR   Statistical models and practice that improve reproducibility in genomics research

SESSION CHAIR

March, 2010	ENAR	Invited Session: Statistical models and practice that improve reproducibility in genomics research
August, 2010	Joint Statistical Meetings	Contributed Session: Genomics and Mutation Classification

PROFESSIONAL SOCIETIES

2006 - ˙	American Statistical Association
2007	Interface
2008	American Heart Association

RECOGNITION

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 Bloomerg 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

INVITED TALKS

- *Association of denovo copy number variants with oral clefts*, International Chinese Statistical Association, Boston, MA. June 24, 2012.

- *Association of denovo copy number variants with oral clefts*, FaceBase Workshop, Los Angelos, CA. June 26, 2012.

- *Measuring copy number from high-throughput SNP arrays*, ENAR International Biometric Society Conference, March 24, 2011.

- *Copy number estimation and batch effects in SNP arrays*, International Biometric Conference, Florianopolis, Brazil. December 6, 2010.

- *Copy number estimation and batch effects in SNP arrays*, BIRS Statistical Genomics In Biomedical Research, Banff, Alberta. July 20, 2010.

- *Batch effects in SNP arrays: solutions for genotyping and copy number estimation*, Vassar University, November 17, 2009.

- *Batch effects in SNP arrays: solutions for genotyping and copy number estimation*, Computational Biology and Bioinformatics Seminar, M.D. Anderson Cancer Center, Houston, TX, Sept 30, 2009.

- *Discussant*, Biostatistics Grand Rounds, Johns Hopkins Bloomberg School of Public Heath, Baltimore, MD, February 4, 2009.

- *Genome-scale estimation of copy number and inference*, Second Annual Young Investigator Symposium on Genomics and Bioinformatics, Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, 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).

- *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*. ENAR Conference, Tampa, FL (March 28, 2006).

Contributed Talks

- *Finding de novo copy number variants in case-parent trios for oral cleft*. European Bioconductor Developer's Meeting, Zurich Switzerland (December 13, 2012).

Other Professional Accomplishments

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**, Tjelmeland H, Nobel AB, and Parmigiani G. A Bayesian multilevel model for cross-study differential gene expression.
crlmm	Carvalho B and **Scharpf RB**, Genotyping and copy number estimation for high-throughput SNP arrays.
oligoClasses	Carvalho B and **Scharpf RB**, Infrastructure for analysis of high-throughput genotyping platforms.
MinimumDistance	**Scharpf RB** and others. Fast Detection of De Novo Copy Number Variants from SNP Arrays.