BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FIVE PAGES.**

NAME: Yun Ju Sung

eRA COMMONS USER NAME (credential, e.g., agency login): Sungyj

POSITION TITLE: Associate Professor of Biostatistics

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	Completion Date MM/YYYY	FIELD OF STUDY
Pohang University of Sci. & Tech., South Korea	BS	1995	Mathematics
University of Minnesota, Minneapolis, MN	MS	2001	Statistics
University of Minnesota, Minneapolis, MN	PhD	2003	Statistics
University of Washington, Seattle, WA	Post Doc	2006	Medical Genetics

A. Personal Statement

My research interest is in the genomic underpinning of human complex diseases. Since joining Washington University, I have been working on various methodological developments, with particular attention to their applicability to real data. I have worked on several projects dealing with GWAS consortia, imputation of genotypes in various multi-ethnic family studies for which the Division serves as the data Coordinating Center, analysis of sequence data, and analysis of rare variants using various statistical methods. I currently have a K25 award to gain fundamental knowledge in cardiovascular disease (CVD) research and in genomics so that I can more effectively participate in multi-disciplinary research programs with clinical and translational CVD researchers and be better equipped to develop and apply statistical methods to contribute more meaningfully to the field of CVD genetics. The research objective of this K25 application is to decipher the genetic and environmental architecture of cardiometabolic traits by incorporating GxE interactions and regulatory annotation information. We hypothesize that joint analysis of the environment, regulatory variants and coding variants will enhance the discovery of putative genetic variants and the functional mechanisms underlying cardiometabolic traits. To evaluate this hypothesis, we aim to identify genetic variants involving GxE interactions and identify putative functional variants by incorporating ENCODE regulation information.

B. Positions and Honors

Positions and Employment:

2001-2002 Instructor, University of Minnesota, Minneapolis, MN
2003-2006 Senior Fellow in Medical Genetics, University of Washington, Seattle, WA
2006-2015 Assistant Professor of Biostatistics, Washington University in St. Louis, St Louis, MO
2015- Associate Professor of Biostatistics, Washington University in St. Louis, St Louis, MO

Other Experience and Professional Memberships:

2005- Member of American Society of Human Genetics
2006- Member of International Genetic Epidemiology Society

Honors:

1999-2000 School of Statistics Alumni Fellowship, University of Minnesota
2001-2002 Graduate School Doctoral Dissertation Fellowship, University of Minnesota

C. Contribution to Science

1. **Development of Statistical Methods (and Software):** My early publications addressed development of novel statistical methods. I developed a Monte Carlo method to approximate the maximum likelihood estimate when the likelihood is not available in closed form. An R package *bernor* was written to implement this method for the Logit-Normal generalized linear mixed model. In collaboration with my post-doctoral mentors, Drs. Ellen Wijsman and Elizabeth Thompson, I developed a new program for parametric linkage analysis that allows for two linked genes and a polygenic component. The program was later extended to handle more complex models, such as gene-gene interaction and imprinting effects. I also developed multivariate polygenic models using conditional maximum likelihood and applied them to autism studies to properly adjust for pedigree ascertainment.

    a. **Sung YJ**, Dawson G, Munson J, Estes A, Schellenberg GD, Wijsman EM: Genetic investigation of quantitative traits related to autism: use of multivariate polygenic models with ascertainment adjustment. *American Journal of Human Genetics* 76:68-81, 2005.
    b. **Sung YJ**, Geyer CJ: Monte Carlo likelihood inference for missing data models. *Annals of Statistics* 35:990-1011, 2007.
    c. **Sung YJ**, Thompson EA, Wijsman EM: MCMC-based linkage analysis for complex traits on general pedigrees: multipoint analysis with a two-locus model and a polygenic component. *Genetic Epidemiology* 31:103-114, 2007.
    d. **Sung YJ**, Wijsman EM: Accounting for epistasis in linkage analysis of general pedigrees. *Human Heredity* 63:144-152, 2007.
    e. **Sung YJ**, Rao DC. Model-based linkage analysis with imprinting for quantitative traits: Ignoring imprinting effects can severely jeopardize detection of linkage. *Genetic Epidemiology* 32:487-496, 2008.

2. **Statistical Methods in GWAS era:** In genome-wide association studies (GWAS), statistical methods played important roles utilizing the publically available data. My research investigating such statistical methods, such as genotype imputations and analysis of rare variants, provided insights and associated challenges.

    a. Neuman RJ, **Sung YJ**. Multistage analysis strategies for genome-wide association studies: Summary of Group 3 contributions to Genetic Analysis Workshop 16. *Genetic Epidemiology*. 2009;33 Suppl 1:S19-23
    b. **Sung YJ**, Wang L, Rankinen T, Bouchard C, Rao DC: Performance of genotype imputations using data from the 1000 Genomes Project. *Human Heredity* 73:18-25, 2012. PMCID: PMC3322630
    c. **Sung YJ**, Gu CC, Tiwari HK, Arnett DK, Broeckel U, Rao DC: Genotype imputation for African Americans using data from HapMap phase II versus 1000 Genomes Projects. *Genetic Epidemiology* 36:508-516, 2012. NIHMS ID: NIHMS477089
    d. **Sung YJ**, Korthauer KD, Swartz MD, Engelman CD: Methods for collapsing multiple rare variants in whole genome sequence data. *Genetic Epidemiology* 38 Suppl 1:S13-S28, 2014.

3. **Genetic Architecture for Cardiometabolic Traits:** I have performed many GWAS analyses to decipher the genetic basis for multiple cardiometabolic traits. I also performed genotype imputation using HapMap and 1000 Genomes Projects for various multi-ethinic family studies; these imputed data were used by many collaborators and made significant contributions to the cardiometabolic research.

    a. Kato N, Takeuchi F, Tabara Y, Kelly TN, Go MJ, …, **Sung YJ**, …,Tai ES, Cho YS, He J: Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in east Asians. *Nature Genetics* 43:531-538, 2011. PMCID: PMC3158568
    b. Franceschini N, Fox E, Zhang Z, Edwards TL, Nalls Michael A, **Sung YJ**, …,Chakravarti A, Reiner AP, Levy D, Keating BJ, Zhu X. Genome-wide association analysis of blood-pressure traits in african-ancestry individuals reveals common associated genes in african and non-african populations. *American Journal of Human Genetics* 93:545-554, 2013
    c. He J, Kelly T, Zhao Q, Li H, Huang J, Wang L, Jaquish C, **Sung YJ**, …, Hixson J, Gu D.Genome-Wide Association Study Identifies Eight Novel Loci Associated with Blood Pressure Responses to Interventions in Han Chinese. *Circulation: Cardiovascular Genetics* 6:598-607, 2013.

d. **Sung YJ**, Basson J, Cheng N, Nguyen KD, Nandakumar P, Hunt SC, Arnett D, Davila-Roman V, Rao DC, Chakravarti A. Role of rare variants in systolic blood pressure: Analysis of exome chip data in HyperGEN African Americans. *Human Heredity* 79: 20-27, 2015 DOI: 10.1159/000375373
e. **Sung YJ**, Pérusse L, Sarzynski MA, Fornage M, Sidney S, Sternfeld B, Rice T, Terry G, Jacobs DR, Katzmarzyk P, Carr JJ, Ghosh S, Rankinen T, Rao DC, Bouchard C. Genome-wide association studies suggest sex-specific loci associated with abdominal and visceral fat. *International Journal of Obesity* 40:662–674, 2016. doi:10.1038/ijo.2015.217

4. **Genetic and Environmental Architecture of Cardiometabolic Traits:** My recent work has focused on deciphering the genetic and environmental architecture of cardiometabolic traits by incorporating GxE interactions. Understanding this architecture will contribute to our knowledge of the pathogenesis of cardiovascular disease which may have important implications for personalized medicine.

a. **Sung YJ**, Schwander K, Arnett D, Kardia S, Rankinen T, Bourchard C, Boerwinkle E, Hunt SC, Rao DC. An empirical comparison of meta-analysis and mega-analysis of individual participant data for identifying gene-environment interactions. *Genetic Epidemiology* 38: 369-378, 2014
b. **Sung YJ**, de las Fuentes L, Schwander K, Simino J, Rao DC. Gene-smoking interactions identify several novel blood pressure loci in the Framingham Heart Study. *American Journal of Hypertension*. 2015; 28: 343–354. doi:10.1093/ajh/hpu149
c. Basson J*, **Sung YJ***, de Las Fuentes L, Schwander K, Vazquez A, Rao DC. Three Approaches to Modeling Gene-Environment Interactions in Longitudinal Family Data: Gene-Smoking Interactions in Blood Pressure. *Genetic Epidemiology.* 2016; 40:73-80. doi: 10.1002/gepi.21941. *Co-first authors
d. **Sung YJ**, Winkler TW, Manning AK, … , Kardia S, Zhu X, Rice K, Borecki IB, Rao DC, Gauderman WJ, Cupples LA. An empirical comparison of joint and stratified frameworks to studying GxE interactions: Systolic blood pressure and smoking in the CHARGE Gene-Lifestyle Interactions Working Group. *Genetic Epidemiology* 40:404-415, 2016

D. Research Support
Sung, Yun Ju

R01 HL086694 Rao, D.C. 12/1/2006 -5/31/2014
JHU/NIH
Genome-Wide Association Analysis in Essential Hypertension (FEHGAS2)
This study will perform exome sequencing in select samples in the ARIC cohort, carry out metabo chip and exome chip analyses in large FBPP samples, and pursue genetic dissection of blood pressure and hypertension using collapsing and other state-of-the-art methods including interactions.
Role: Co-Investigator

R01 HL107552 Rao, D.C. 8/1/2011 -3/31/2014
NIH/NHLBI
Gene-Environment Interactions in the Longitudinal Framingham Heart Study
In this study we will use gene by environment interactions to discover novel disease loci which could lead to new diagnostic and therapeutic interventions for treatment of cardiovascular and metabolic diseases.
Role: Co-Investigator

R01 HL090668 Rao, DC 7/1/2009 -6/30/2014
NIH
Initiating Factors for Hypertension
As the Data Coordinating Center for this project, we are responsible for developing multiple analytical approaches, including network analyses, such as the Artificial Neural Networks and Bayesian Networks.
Role: Co-Investigator

R01HL111249 Rao, D.C. 7/1/2012 -6/30/2017

NIH

Rare Variants for Hypertension in Taiwan Chinese

The primary goal of the proposed research is to identify rare and low frequency variants that have large effects on blood pressure and hypertension by carrying out exome sequencing in 150 highly enriched Taiwan Chinese sib-pairs (300 subjects) and 300 unrelated controls, then to validate the top 6,000 SNPs in larger samples, and finally replicate the top 50 SNPs in nearly 45,000 multi-ethnic subjects.

Role: Co-Investigator

K25HL121091 Sung, YJ 7/1/2014 - 6/30/2019

NIH

Statistical Methods for Genomic Dissection of Cardiovascular Diseases

This mentored career development grant application proposes a training program to integrate my previous research in statistical genetics into cardiovascular disease (CVD). The research objective is to decipher the genetic and environmental architecture of cardiometabolic traits by incorporating GxE interactions and regulatory annotation information.

Role: Principal Investigator