

CURRICULUM VITAE

Dokyoon Kim, Ph.D.

Ritchie Lab
Center for Systems Genomics
The Pennsylvania State University
PA, USA

Office: +1 814 863 4467
Fax: +1 814 863 6699
duk27@psu.edu

EDUCATION

- Feb, 2013 **Ph.D.**, Biomedical Informatics
College of Medicine, Seoul National University, Korea
Advisor: Ju Han Kim, M.D., Ph.D.
- Dissertation: Graph- and kernel-based integrative analyses of multi layers of heterogeneous genomic data
- Feb, 2006 **B.S.**, Computer Science
Korea University, Korea

RESEARCH AND PROFESSIONAL EXPERIENCES

- Mar, 2013 – Present **Postdoctoral Fellow** (PI: Marylyn D. Ritchie, Ph.D.)
Center for Systems Genomics
Department of Biochemistry & Molecular Biology
The Pennsylvania State University, PA, USA
- Mar, 2006 – Feb, 2013 **Ph.D. Candidate & Researcher** (PI: Ju Han Kim, M.D., Ph.D.)
Division of Biomedical Informatics, College of Medicine
Seoul National University, Seoul, Korea
- Dec, 2009 - Aug, 2011 **Visiting Researcher** (PI: Hyunjung Shin, Ph.D.)
Department of Industrial Engineering
Ajou University, Suwon, Korea
- Jun, 2005 - Aug, 2005 **Intern** (PI: Ju Han Kim, M.D., Ph.D.)
Division of Biomedical Informatics, College of Medicine
Seoul National University, Seoul, Korea
- Jan, 2004 - Feb, 2004 **Intern**
Sun Microsystems, Colorado, USA

PEER-REVIEWED RESEARCH PAPERS

1. **Dokyo Kim**, Heather Volk, Santhosh Girirajan, Sarah Pendergrass, Irva Hertz-Picciotto, Molly Hall, Shefali Setia, Rebecca J. Schmidt, Robin L. Hansen, Debashis Ghosh, Marylyn Ritchie*, Scott Selleck*. Interactions between air pollution exposure and copy number variation confer significant risk for autism. *JAMA Psychiatry* (submitted)
2. **Dokyo Kim**, Ruowang Li, Scott M. Dudek, Marylyn D. Ritchie. Predicting censored survival data based on the interactions between meta-dimensional omics data in breast cancer. *J Biomed Inform* (In press)
3. Marylyn D. Ritchie*, Emily R. Holzinger, Ruowang Li, Sarah A. Pendergrass, **Dokyo Kim**. Systems genomics analyses and data integration for exploring the genetic architecture of complex traits. *Nature Reviews Genetics*, 16, 85-97, 2015
4. **Dokyo Kim**, Ruowang Li, Scott M. Dudek, John R. Wallace, Marylyn D. Ritchie. Binning somatic mutations based on biological knowledge for predicting survival: an application in renal cell carcinoma. *Pacific Symposium on Biocomputing (PSB)*, 96-107, 2015
5. **Dokyo Kim**, Ruowang Li, Scott M. Dudek, Alex T. Frase, Sarah A. Pendergrass, Marylyn D. Ritchie. Knowledge-driven genomic interactions: an application in ovarian cancer. *BioData Mining*, 7, 20, 2014
6. **Dokyo Kim**, Je-Gun Joung, Kyung-Ah Sohn, Hyunjung Shin, Marylyn D. Ritchie, Ju Han Kim. Knowledge Boosting: A graph-based integration approach with multi-omics data and genomic knowledge for cancer clinical outcome prediction. *J Am Med Inform Assoc*, doi:10.1136/amiajnl-2013-002481, 2014
7. Ruowang Li, **Dokyo Kim**, Scott M. Dudek, Marylyn D. Ritchie. An integrated analysis of genome-wide DNA methylation and genetic variants underlying etoposide-induced cytotoxicity in European and African populations. *Lect Notes Comput Sci*, 8602:928-938, 2014
8. Je-Gun Joung, **Dokyo Kim**, Su-Yeon Lee, Hwa Jung Kang, Ju Han Kim. Integrated analysis of microRNA-target interactions with clinical outcomes for cancers. *BMC Med Genomics*, 7(Suppl 1):S10, 2014
9. **Dokyo Kim**, Marylyn D. Ritchie. Data integration for cancer clinical outcome prediction. *J Health Med Informat*, doi:10.4172/2157-7420.1000e122, 2014
10. **Dokyo Kim**, Hyunjung Shin, Kyung-Ah Sohn, Anurag Verma, Marylyn D. Ritchie, Ju Han Kim. Incorporating inter-relationships between different levels of genomic data into cancer clinical outcome prediction. *Methods*, S1046-2023(14)00033-4, 2014

11. Jong Jin Oh, Seok Soo Byun, Sang Eun Lee, Sung Kyu Hong, Chang Wook Jeong, Won Suk Choi, **Dokyoony Kim**, Hae Jong Kim, Soon Chul Myung. Genetic variants in the CYP24A1 gene associated with prostate cancer risk and aggressiveness in a Korean study population. *Prostate Cancer P D*, doi:10.1038/pcan.2014.1, 2014
12. Jong Jin Oh, Seok Soo Byun, Sang Eun Lee, Sung Kyu Hong, Chang Wook Jeong, **Dokyoony Kim**, Hae Jong Kim, Soon Chul Myung. Genetic variations in VDR associated with prostate cancer risk and progression in a Korean population. *Gene*, 533(1):86-93, 2014
13. **Dokyoony Kim**, Ruowang Li, Scott M. Dudek, Marylyn D. Ritchie. ATHENA: Identifying interactions between different levels of genomic data associated with cancer clinical outcomes using grammatical evolution neural network. *BioData Mining*, 6(1):23, 2013
14. Kyung-Ah Sohn[§], **Dokyoony Kim**[§], Jaehyun Lim, Ju Han Kim*. Relative impact of multi-layered genomic data on gene expression phenotypes in serous ovarian tumors. *BMC Systems Biology*, 7(Suppl 6):S9, 2013
§: Joint-first author
15. **Dokyoony Kim**, Sungeun Kim, Shannon L. Risacher, Li Shen, Marylyn D. Ritchie, Michael W. Weiner, Andrew J. Saykin, Kwangsik Nho. A graph-based integration of multimodal brain imaging data for the detection of early mild cognitive impairment (E-MCI). *Lect Notes Comput Sci*, 8159:159-169, 2013
16. **Dokyoony Kim**[§], Hyunjung Shin^{§,*}, Su-Yeon Lee, Je-Gun Jeong, Ju Han Kim*. Intra-relation reconstruction from inter-relation: miRNA to gene expression. *BMC Systems Biology*, 7(Suppl 3):S8, 2013
§: Joint-first author
17. Kanghee Park, Amna Ali, **Dokyoony Kim**, Yeolwoo An, Minkoo Kim, Hyunjung Shin. Robust predictive model for evaluating breast cancer survivability. *Eng. Appl. Artif. Intel*, 26(9):2194-2205, 2013
18. Je-Gun Jeong, **Dokyoony Kim**, Kyung Hwa Kim, Ju Han Kim. Extracting coordinated patterns of DNA methylation and gene expression in ovarian cancer. *J Am Med Inform Assoc*, 20(4):637-642, 2013
19. Heewon Seo, **Dokyoony Kim**, Jong-Hee Chae, Hee Gyung Kang, Byung Chan Lim, Hae Il Cheong, Ju Han Kim. Development of Korean rare disease knowledge base. *Health Inform Res*, 18(4):272-278, 2012
20. **Dokyoony Kim**[§], Hyunjung Shin^{§,*}, Young Soo Song, Ju Han Kim*. Synergistic effect of different levels of genomic data for cancer clinical outcome prediction. *J Biomed Inform*, 45(6):1191-1198, 2012
§: Joint-first author
21. **Dokyoony Kim**, Young Soo Song, Ju Han Kim, Hyunjung Shin. Genomic data comparison: Which data is more informative?. *Proc. of KSH International*

Conference on Internet (ICONI 2010), 79-83, 2010

22. Young Soo Song, Hye Won Lee, Yu Rang Park, **Dokyoon Kim**, Jaehyun Sim, Hyunseok Peter Kang, Ju Han Kim. TMA-TAB: A spreadsheet-based document for exchange of tissue microarray data based on the Tissue Microarray-Object Model. *J Biomed Inform*, 43:435-441, 2010
-

CONFERENCE PRESENTATIONS (ORAL PRESENTATION)

1. **Dokyoon Kim**, Ruowang Li, Scott M. Dudek, John R. Wallace, Marylyn D. Ritchie. Binning somatic mutations based on biological knowledge for predicting survival: an application in renal cell carcinoma. *Pacific Symposium on Biocomputing (PSB)*, Hawaii, USA, Jan 4-8, 2015
2. **Dokyoon Kim**, Ruowang Li, Scott M. Dudek, Marylyn D. Ritchie. Predicting censored survival data based on the interactions between meta-dimensional omics data in breast cancer. *Translational Bioinformatics Conference (TBC/ISB 2014)*, Qingdao, China, Oct 24-27, 2014
3. **Dokyoon Kim**, Ruowang Li, Scott M. Dudek, Marylyn D. Ritchie. Data integration for identifying interactions between meta-dimensional omics data associated with cancer clinical outcomes using grammatical evolution neural network. *International Indian Statistical Association (IISA) conference*, Riverside, USA, July 11-13, 2014
4. **Dokyoon Kim**, Ruowang Li, Scott M. Dudek, Marylyn D. Ritchie. ATHENA: Identifying interactions between different levels of genomic data associated with cancer clinical outcomes using grammatical evolution neural network. *Translational Bioinformatics Conference (TBC/ISCB 2013)*, Seoul, Korea, Oct 2-4, 2013
5. **Dokyoon Kim**, Je-Gun Joung, Kyung-Ah Sohn, Hyunjung Shin, Marylyn D. Ritchie, Ju Han Kim. Knowledge Boosting: A graph-based integration with multi-omics data and genomic knowledge for cancer clinical outcome prediction. *Translational Bioinformatics Conference (TBC/ISCB 2013)*, Seoul, Korea, Oct 2-4, 2013
6. **Dokyoon Kim**, Hyunjung Shin, Su-Yeon Lee, Je-Gun Jeong, Ju Han Kim. Intra-relation reconstruction from inter-relation: miRNA to gene expression. *International Conference on Bioinformatics 2013 (InCoB 2013)*, Taicang, China, Sep 20-22, 2013
7. **Dokyoon Kim**, Ju Han Kim, and Hyunjung Shin. Cancer clinical outcome prediction based on inter-relationship between miRNA and gene expression. *Business Intelligence and Data Mining Conference (BIDM 2011 Fall)*, Seoul, Korea, Dec 2, 2011
8. Hyunjung Shin, **Dokyoon Kim**, Kanghee Park, and Amna Ali. Breast cancer survivability prediction with surveillance, epidemiology, and end results database.

Translational Bioinformatics Conference 2011 (TBC2011), Seoul, Korea, Nov 10-11, 2011

9. **Dokyo Kim**, Young Soo Song, Ju Han Kim, Hyunjung Shin. Integrative molecular-based classification of clinical outcomes in cancer. *Business Intelligence and Data Mining Conference (BIDM 2010 Fall)*, Seoul, Korea, Dec 3, 2010
10. **Dokyo Kim**, Young Soo Song, Ju Han Kim, and Hyunjung Shin. Cancer recurrence prediction using machine learning. *Korean Institute of Industrial Engineers Conference*, Jeju, Korea, Jun 3–4, 2010

CONFERENCE PRESENTATIONS (POSTER PRESENTATION)

1. **Dokyo Kim**, Ruowang Li, Scott M. Dudek, John R. Wallace, Marylyn D. Ritchie. Cancer-BioBin: Binning somatic mutations based on biological knowledge for predicting survival outcome. *TCGA' 4th Annual Scientific Symposium*, Bethesda, USA, May 11-12, 2015 (Poster)
2. **Dokyo Kim**, Anastasia Lucas, Ruowang Li, Alex T. Frase, Santhosh Girirajan, Scott B. Selleck, Marylyn D. Ritchie. Testing the genomic enrichment of common and rare copy number burden associated with autism. *Biology of Genome*, Cold Spring Harbor Laboratory, USA, May 5-9, 2015 (Poster)
3. **Dokyo Kim**, Heather Volk, Santhosh Girirajan, Sarah A. Pendergrass, Molly Hall, Shefali Verma, Rebecca Schmidt, Robin Hansen, Debashis Ghosh, Irva Hertz-Picciotto, Marylyn D. Ritchie, Scott Selleck. A gene-environment interaction between copy number burden and ozone exposure provides a high risk of autism. *International Genetic Epidemiology Society (IGES 2014)*, Vienna, Austria, Aug 28-30, 2014 (Poster)
4. **Dokyo Kim**, Ruowang Li, Scott M. Dudek, Marylyn D. Ritchie. Identifying interactions between meta-dimensional genomic data associated with cancer clinical outcomes using grammatical evolution neural network. *TCGA' 3rd Annual Scientific Symposium*, Bethesda, USA, May 12-13, 2014 (Poster)
5. **Dokyo Kim**, Ruowang Li, Scott M. Dudek, Alex T. Frase, Sarah A. Pendergrass, Marylyn D. Ritchie. Identifying knowledge-driven genomic interactions associated with cancer clinical outcome. *Biology of Genome*, 2014, Cold Spring Harbor Laboratory, USA, May 6-10 (Poster)
6. **Dokyo Kim**, Je-Gun Joung, Kyung-Ah Sohn, Hyunjung Shin, Marylyn D. Ritchie, Ju Han Kim. A graph-based integration with multi-omics data and genomic knowledge for cancer clinical outcome prediction. *American Society of Human Genetics (ASHG 2013)*, Boston, USA, Oct 22-26, 2013 (Poster)

7. **Dokyo Kim**, Chan Hee Park, Jung Hoon Woo, Ju Han Kim. Xperanto-SNP: A web-based integrated data management system for genetical genomics. *Pharmacogenomics & Personalized Medicine*, Hinxton, UK, 2009 (Poster)
 8. **Dokyo Kim**, Ji Hun Kim, Sung Bum Cho, Hee Joon Chung, Ju Han Kim. Evaluating chromosome instability in cancer using gene expression data and location information. *European Conference on Computational Biology (ECCB 2008)*, Cagliari, Italy, 2008 (Poster)
 9. **Dokyo Kim**, Ju Han Kim. ChromoViz-Web: mapping and multimodal visualizing gene expression data with integrated major bio resources onto chromosomes using scalable vector graphics. *International Conference on Intelligent Systems for Molecular Biology (ISMB 2006)*, Fortaleza, Brazil, 2006 (Poster)
-

INVITED PRESENTATIONS

1. Multi-omics data integration for predicting cancer clinical outcomes. *Capita Selecta in Complex Disease Analysis Conference and EU COST Pancreas Annual Meeting*, Liege, Belgium, Nov 25th, 2014 (Keynote speaker)
 2. Gene-environment interactions in autism: Copy number burden and air pollution exposure contribution to risk. *The Center for Children's Environmental Health (CCEH) meeting*, UC Davis, USA, July 18th, 2014
 3. Graph-based integrative analyses of multi-omics data for cancer clinical outcome prediction. *Genomic Seminar*, The Pennsylvania State University, USA, August 20th, 2013
 4. Integration of multi-layers of genomic data for cancer clinical outcome prediction. *Summit on Translational Bioinformatics (TBI2013)*, San Francisco, USA, Mar 19th, 2013 (Late Breaking Session)
 5. Empirical comparison on heterogeneous genomic data: Copy number variation, methylation, miRNA, and gene expression. *Proc. of Annual Meeting of Institute for Operations Research and the Management Sciences (INFORMS 2010)*, Austin, Texas, USA, Nov 8th, 2010
-

HONORS AND AWARDS

1. Pacific Symposium on Biocomputing 2015 Travel Award (2015)
2. Best Paper Award, Translational Bioinformatics Conference (TBC) (2014)

3. Associate Faculty Member (AFM), Pharmacogenomics, Faculty of 1000 (www.F1000.com) (2014-present)
4. Best Paper Award, Translational Bioinformatics Conference (TBC) (2013)
5. Brain Korea 21 Research Fellowship, Seoul National University (2006-2010)
6. Korean government scholarship for exchange students in USA (2004)

NATIONAL/ INTERNATIONAL RESPONSIBILITIES

2015	Scientific Program Committee, the 17 th Portuguese Conference on Artificial Intelligence (EPIA 2015)
2014	Scientific Program Committee, the 4 th Annual Translational Bioinformatics Conference (ISB/TBC 2014)
2014	Scientific Program Committee, the 5 th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM-BCB 2014)
2014-present	Editorial Board Member, Journal of Health & Medical Informatics
2014	Grant Reviewer, “Cracking unsolved cases of neuromuscular disease”, OR 14-02, Parinses Beatrix Spierfonds
2013	Scientific Program Committee, the 3 rd Annual Translational Bioinformatics Conference (TBC/ISCB-Asia 2013)

REVIEWER FOR JOURNALS

Biodata Mining
Computers in Biology and Medicine
Genome Medicine
Journal of Biomedical Informatics
PloS ONE

TEACHING EXPERIENCES

Mar, 2013 - Present	Mentoring undergraduate/graduate students in the lab of Dr. Marylyn Ritchie
Mar, 2014	Invited lecture titled “Complex Modeling of Genetics Effects: Meta-dimensional Analysis”, Course: Human Genomics & Biomedical Informatics, The Pennsylvania State University, USA
Sep, 2012	“Advanced R graphics.” 11 th workshop of biomedical informatics: R for Bioinformatics and Biomedicine, Seoul National University, Seoul, Korea
Aug, 2012	“TCGA data analysis.” Genome Data Analysis Workshop 2012 Summer, Seoul National University, Seoul, Korea
Aug, 2012	“CNV data analysis.” Genome Data Analysis Workshop 2012 Summer, Seoul National University, Seoul, Korea
Aug, 2012	“SNP data analysis.” Genome Data Analysis Workshop 2012 Summer, Seoul National University, Seoul, Korea
Aug, 2012	“Gene-set approaches & prognostic subgroup prediction.” Genome Data Analysis Workshop 2012 Summer, Seoul National University, Seoul, Korea
Feb, 2012	“CNV data analysis.” Genome Data Analysis Workshop 2012 Winter, Seoul National University, Seoul, Korea
Feb, 2012	“SNP data analysis.” Genome Data Analysis Workshop 2012 Winter, Seoul National University, Seoul, Korea
Feb, 2012	“Gene-set approaches & prognostic subgroup prediction.” Genome Data Analysis Workshop 2012 Winter, Seoul National University, Seoul, Korea
Aug, 2011	“CNV data analysis.” Genome Data Analysis Workshop 2011 Summer, Seoul National University, Seoul, Korea
Aug, 2011	“SNP data analysis.” Genome Data Analysis Workshop 2011 Summer, Seoul National University, Seoul, Korea
Aug, 2011	“Gene-set approaches & prognostic subgroup prediction.” Genome Data Analysis Workshop 2011 Summer, Seoul National University, Seoul, Korea

Updated on May, 2015

Oct, 2008	“Introduction to bioinformatics.” Department of Health Administration and Management, Soon Chun Hyang University, Seoul, Korea
Jul, 2008	“Application of pharmacogenomics.” 2008 Asian Institute in Statistical Genetics and Genomics, Kyunghee University, Seoul, Korea
Mar-June, 2006	Teaching Assistant, Course: Bioinformatics for Genomic Medicine, Seoul National University College of Medicine, Seoul, Korea

COMPUTER SKILLS

Languages	Python, JAVA, C/C++, PHP, JSP, Perl
Statistical packages	R, Matlab, SAS, SPSS
Operating systems	Linux, MS-Windows, Mac-OS
DBMS	MySQL

REFERENCES

Marylyn D. Ritchie, Ph.D.

Director, Center for Systems Genomics

Professor, Department of Biochemistry and Molecular Biology, The Pennsylvania State University, PA, USA

Office: +1-814-863-4467

Fax: +1-814-863-6699

E-mail: marylyn.ritchie@psu.edu

Ju Han Kim, M.D., Ph.D.

Director, Systems Biomedical Informatics Research Center

Chair and Professor, Division of Biomedical Informatics, College of Medicine, Seoul National University, Seoul, Korea

Office: +82-2-740-8320

Fax: +82-2-747-8928

E-mail: juhan@snu.ac.kr

Scott B. Selleck, M.D., Ph.D.

Professor and Head, Department of Biochemistry and Molecular Biology, The Pennsylvania State University, PA, USA

Office: +1-814-867-4373

Fax: +1-814-867-4374

E-mail: sbs24@psu.edu