



THE UNIVERSITY of TEXAS  
HEALTH SCIENCE CENTER AT HOUSTON  
SCHOOL OF PUBLIC HEALTH

HUMAN GENETICS CENTER

Kyoung-Jae Won  
Professor  
Dept. of Genetics, School of Medicine  
University of Pennsylvania  
June 4, 2014

Dear Dr. Won,

I am writing to strongly recommend my beloved student Mr. Shicheng Guo, who applied for a postdoctoral fellow position in your lab. Having been his supervisor for two years, I am proud to say that Mr. Guo is the best student I have ever taught and a truly exceptional young scientist. I believe that his creative research in biostatistics, statistical genetics and high dimensional data analysis is of extraordinary significance in biostatistics and biomedical research and he would achieve great success in his academic career.

Mr. Guo has received advanced education and training in the areas of genetics, molecular biology, bioinformatics, statistic genetics and high dimensional data analysis. He received Bachelor degree in the Department of Biotechnology from the School of Life Science at the Northeast Agriculture University, Harbin, China in 2009, and then he was recommended into five year's human medical genetics Ph.D program in Fudan University, Shanghai. Mr. Guo was an outstanding student. He made exceptional independent achievements. I am very impressed that although he took 30 courses in clinical research, bioinformatics and genetics, his GPA is still 3.37, which is very high quantile in Fudan University. I got to know Mr. Guo in a joint project for disease risk prediction models for thyroid cancer based on significant SNPs identified from GWAS study with his advisor Dr. Li Jin, vice president at the Fudan University, Shanghai, China, in 2011. His highly intelligence, diligence and perceptiveness really impressed me in the joint research. Recommended by his advisor, he came to Houston and joined the Ph. D program in biostatistics in the Division of Biostatistics, University of Texas School of Public Health in 2011, and then as my research assistant, he has focused on developing novel statistical methods for RNA-seq and image-genetics data analysis which leverage high dimensional data reduction, causal inference and functional data analysis techniques to identify risk factors and predict diseases. He was supported by my grant: 5R01HL106034-02, "Statistical Methods for Finding Missing Heritability".

As a Ph.D candidate student, Mr. Guo is very promising. During his Ph.D studies he published 2 patents and 19 papers in Nature biotechnology, PLoS biology, Clinical Cancer Research, Neoplasia, Journal of medical genetics, European Journal of Human Genetics, Genes and immunity, The FASEB Journal and Clinical epigenetics. To my surprising, he published 4 Abstracts (Posters) in the International Meeting on Human Genome Variation and Complex





THE UNIVERSITY of TEXAS  
HEALTH SCIENCE CENTER AT HOUSTON  
SCHOOL OF PUBLIC HEALTH

HUMAN GENETICS CENTER

Genome Analysis, Annual Meeting of American Society of Human Genetics, Epigenetics in Development and Diseases 9th Asian Epigenomics Meeting. He won the first place of poster award in the UT GSBS human and molecular genetics symposium in 2014 for his poster "Cluster analysis to RNA-seq data based on Function PCA reveals cancer subtype".

Although great progress in genome-wide association studies (GWAS) has been made, the performance of the most current SNP- based disease risk prediction methods was not assessed accurately; He re-analyzed his previous published association data and evaluated the prediction ability with several significant SNPs which is identified by GWAS studies. Since thyroid is one of complex diseases with highest heritability among all the cancers, his result has an important implication that the role of the prediction ability based on few significant SNPs from GWAS study is limited. Also, this research re-emphasized that associations identified by GWAS account for only a few percent of the genetic variance, more attentions should be paid to the missing heritability caused by epigenetic variations such DNA methylation and histone modifications and temporal and spatial specific gene expression.

DNA methylation was suggested as the promising biomarker for early diagnosis of lung cancer. However, it is a great challenge to search for the optimized combination of the methylation biomarkers to obtain the maximum diagnosis performance. He developed a panel of DNA methylation biomarkers and validated their diagnostic efficiency for non-small cell lung cancer (NSCLC) in a large Chinese Han NSCLC retrospective cohort. Three high-throughput DNA methylation microarray datasets were collected in the discover stage. After normalization, batch effect elimination and integration, significant differential methylated genes and best combination of the biomarkers were determined with leave-one-out support vector machine (SVM) feature selection operation. Then candidate promoters were examined by methylation status determined single nucleotide primer extension technology (MSD-SNuPET) in an independent set of NSCLC/normal tissues. He proposed an effective DNA methylation-based biomarker discover pipeline and identified a promising panel for NSCLC diagnosis. High throughput DNA methylation microarray dataset followed by batch effect elimination can be a good method to discover optimized DNA methylation diagnostic panels. Methylation profiles of AGTR1, GALR1, SLC5A8, ZMYND10 and NTSR1, could be an effective methylation-based assay for the NSCLC diagnosis. The works are prepared to be submitted to Genetics in Medicine. I fully believe that his works in these areas are of great importance and success.

RNA-seq technology provides huge biological information of gene expression and alternative splicing for biologist and medical scientists to discover diagnostic or prognostic biomarkers. Cumulative sum method was generally adopted in current RNA-seq analysis. However this analysis would ignore alternative splicing information which would play important role in the pathogenesis of the complex disease. He provided an effective novel pipeline to analysis of next-generation RNA-seq data based on Functional PCA which can identify aberrant alternative





**THE UNIVERSITY of TEXAS**  
**HEALTH SCIENCE CENTER AT HOUSTON**  
**SCHOOL OF PUBLIC HEALTH**

**HUMAN GENETICS CENTER**

splicing in specific disease or conditions and can discover specific biological variation/subtype, such as cancer or normal, drug response status. This methodology takes the spatial information in the RNA expression characteristic into the consideration, which would be a great innovation in RNA-seq analysis and biological theoretical. The work now is being prepared to the manuscript and will be submitted to "PNAS".

As one of his supervisors, I have the highly priority to give Mr. Guo overall appraisals on his academic potentials and personalities. Mr. Guo has strong motivations in scientific research. He often presents me with creative ideas and has strong abilities to implement his ideas independently. Mr. Guo always can communicate effectively the concepts, approaches and significance of his research. Mr. Guo is polite and easy to work with collaborators. We have initiated extensive national and international collaboration, including scientists at the M.D. Anderson Cancer Center, Baylor College of Medicine, University of Texas Medical School, Yale University, University of Washington, and Beijing University, Fudan University in China. He was actively involved in these collaborative research projects. Particularly, he played an extremely important role in the project: "Point-of-Care Screening and diagnosis of liver cancer in Chinese Population" with Fudan University, Medical School of Nanjing University, and Fujian Medical University Union Hospital. He is an outstanding team Player. He is an intelligent, motivated and exceptional hard working researcher. He is consistent, dedicated and passionate, enthusiastic, cheerful, trustworthy and a pleasure to work with.

In general, Mr. Guo is a very promising young scientist with multidisciplinary knowledge and skills in biostatistics, statistical genetics, clinical genomics and biomedical sciences. He has great motivation and outstanding ability to conduct both independent and collaborative research projects in clinical, epidemiological, genetic and epigenetic research, and cancer prevention. I am fully confident that he will make great success in his academic career and will continually make exceptional achievements in the future. I recommend him to be a postdoctoral fellow without any hesitation. If you have any questions or wish to speak further about Mr. Guo, please feel free to call me or write me.

Sincerely Yours,

Momiao Xiong, Ph. D  
Professor  
Division of Biostatistics  
Human Genetics Center  
University of Texas School of Public Health  
Phone: (713) 500-9894