Ning Li et al., J Proteomics Bioinform 2013, 6:7
http://dx.doi.org/10.4172/0974-276X.S1.069

OMICS Group Conferences
Accelerating Scientific Discovery

3rd International Conference on
Proteomics & Bioinformatics
July 15-17, 2013 Courtyard by Marriott Philadelphia Downtown, USA

# Comparative proteogenomic analysis of *streptococcus suis* by high resolution mass spectrometry

**Ning Li[1,2], Chengpu Zhang[1,2], Yunping Zhu[1], Fuchu He[2] and Ping Xu[1]**
[1]Beijing Proteome Research Center, China
[2]Beijing Institute of Radiation Medicine, China

*Streptococcus suis (S.suis)* serotype 2 is an important human pathogen, causing more than 200 cases of severe human infection across the world. One of its virulent isolates in China, 05ZYH33 strain, caused the large-scale human streptococcal toxic shock syndrome outbreak in 2005. The whole genome sequencing of 05ZYH33 strain was finished in 2007. However, up to now, comprehensive proteome sequencing of the strain was not reported and accurate annotation of its genome remains challenging. In this study, in-depth proteome sequencing of *S.suis* 05ZYH33 was first implemented using high resolution mass spectrometry. Then comparative proteogenomic analysis of 05ZYH33 strain was carried out by database search of an *S.suis* complex database generated by integration of 05ZYH33 predicted protein sequences, six-frame translated sequences, and predicted protein sequences of other 16 *S.suis* strains. In all, 1,644 proteins from *S.suis* 05ZYH33 were identified representing 75.2% of its predicted genes. Besides, 157 novel peptides that only match to 05ZYH33 genome and other strains' protein sequences were identified and validated by spectra similarity comparation of the corresponding synthetic peptides. Based on these novel peptides, we discovered 27 novel protein-encoding genes in 05ZYH33 genome (5 of which are strain-specific novel genes), 41 new translation start sites, 11 frameshift mutations and 78 potential genome sequencing errors or SAPs. Finally, by the combination of bioinformatics analysis and experimental validation, a considerable proportion of new genes, new translation start sites and frameshift mutations were validated, which largely improved the genome annotation of *S.suis*.

## Biography

Ning Li has received his Bachelor Degree from Peking University in 2006 and completed his Ph.D at the age of 27 years from Beijing Institute of Radiation Medicine. He joined Beijing Proteome Research Center in 2011 and is currently working as an Assistant Research Fellow in the Department of Genomics and Proteomics. His research is primarily focused in the fields of computational proteomics and proteogenomics. He is the author of the MASCOT quality control software PepDistiller (www.bprc.ac.cn/pepdistiller), and published several papers in reputed journals such as *Proteomics and Journal of Proteome Research*.

francy.camacho@correo.uis.edu.co, hefc@nic.bmi.ac.cn, xupingghy@gmail.com