BIOINFORMATICS APPLICATIONS NOTE Vol. 30 no. 4 2014, pages 586–587
doi:10.1093/bioinformatics/btt726

Systems biology

Advance Access publication December 15, 2013

SILVER: an efficient tool for stable isotope labeling LC-MS data quantitative analysis with quality control methods

Cheng Chang[1,2,†], Jiyang Zhang[3,†], Mingfei Han[1,2], Jie Ma[1,2], Wei Zhang[3], Songfeng Wu[1,2], Kehui Liu[1,2,‡], Hongwei Xie[3,*], Fuchu He[1,2,4,*] and Yunping Zhu[1,2,*]

[1]Department of Bioinformatics, State Key Laboratory of Proteomics, Beijing Proteome Research Center, National Center for Protein Sciences Beijing, Institute of Radiation Medicine and [2]Department of Bioinformatics, National Engineering Research Center for Protein Drugs, Beijing 102206, [3]Department of Automatic Control, College of Mechatronics and Automation, National University of Defense Technology, Changsha, Hunan 410073 and [4]Department of Chemistry, Institutes of Biomedical Sciences, 130 DongAn Road, Fudan University, Shanghai 200032, P.R. China

Associate Editor: Martin Bishop

ABSTRACT

Summary: With the advance of experimental technologies, different stable isotope labeling methods have been widely applied to quantitative proteomics. Here, we present an efficient tool named SILVER for processing the stable isotope labeling mass spectrometry data. SILVER implements novel methods for quality control of quantification at spectrum, peptide and protein levels, respectively. Several new quantification confidence filters and indices are used to improve the accuracy of quantification results. The performance of SILVER was verified and compared with MaxQuant and Proteome Discoverer using a large-scale dataset and two standard datasets. The results suggest that SILVER shows high accuracy and robustness while consuming much less processing time. Additionally, SILVER provides user-friendly interfaces for parameter setting, result visualization, manual validation and some useful statistics analyses.

Availability and implementation: SILVER and its source codes are freely available under the GNU General Public License v3.0 at http://bioinfo.hupo.org.cn/silver.

Contact: zhuyunping@gmail.com, hefc@nic.bmi.ac.cn and xhwei65@163.com

Supplementary information: Supplementary data are available at *Bioinformatics* online

Received on June 3, 2013; revised on November 22, 2013; accepted on December 11, 2013

1 INTRODUCTION

Mass spectrometry (MS) has been applied in proteomics research for >20 years. The great advantages in resolution, sensitivity and accuracy make MS the dominating technology in proteomics (Nilsson *et al.*, 2010). As the technology improves, MS-based quantitative proteomics has become an important research field in proteomics. Among the strategies developed for quantitative proteomics, stable isotope labeling strategies have been widely used thanks to their high accuracy and precision.

Recently, several software tools have been developed for stable isotope labeling quantification, such as MSQuant (Mortensen *et al.*, 2010) and Census (Park *et al.*, 2008), the well-known software package MaxQuant (Cox and Mann, 2008) embedded with its own search engine, named Andromeda (Cox *et al.*, 2011), as well as a newly published quantification pipeline IsoQuant (Liao *et al.*, 2012), which was designed especially for stable isotope labeling by amino acids in cell culture (SILAC) (Ong *et al.*, 2002). However, there has been little attention to quantitative quality control among these tools and there is hardly any proper algorithm focused on this issue (Yates *et al.*, 2012). On the other hand, the rapidly growing amount of large-scale quantitative proteomics data requires faster and more efficient tools.

Here, we present a new efficient tool, named SILVER for stable isotope labeling data analysis with special considerations for quality control of quantification. SILVER introduces novel quantification confidence filters and indices, which finally result in accurate and robust quantification for stable isotope labeling experiments. Meanwhile, SILVER is fast and flexible by supporting different types of MS data such as the original raw files and the mzXML format files. SILVER takes the quality control results of PeptideProphet (Keller *et al.*, 2002) or PepDistiller (Li *et al.*, 2012) as the input and provides user-friendly interfaces for parameter setting, result visualization, manual validation and some useful statistical analyses.

2 METHODS

SILVER is designed for comprehensive labeling MS data analysis. The workflow of SILVER is shown in Figure 1. The algorithms of SILVER are classified into three levels: the spectrum level, peptide level and protein level. For isotope cluster detection, SILVER uses a novel dynamic isotope matching tolerance algorithm for matching the experimental and theoretical isotope clusters. During the construction of extracted ion current (XIC) (Supplementary Fig. S1), the signal-to-noise ratio of each isotope peak, the goodness of least-square fitting between the theoretical and observed isotope distribution as well as the autocorrelation coefficient of the labeled and unlabeled isotope clusters are determined as quantification confidence filters at the spectrum level (Supplementary Methods). For peptide quantification, SILVER provides two novel quantification confidence indices: the Pearson and the Spearman correlation

*To whom correspondence should be addressed.
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.
‡Present address: State Key Laboratory of Molecular and Developmental Biology, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, P.R. China.

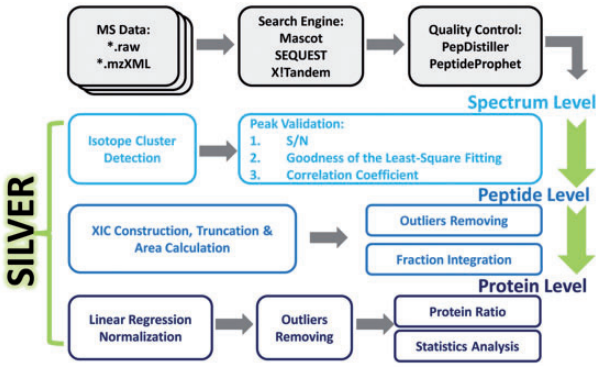

Fig. 1. Overview of the workflow and algorithms in SILVER

coefficients between the labeled and unlabeled XICs, which are calculated as measurements of XIC shape and the intensity variation, respectively. For stable isotope labeling data with fractions such as reverse-phase liquid chromatography separation or sodium dodecyl sulfate–polyacrylamide gel electrophoresis, SILVER starts by integrating the XIC areas as the peptide quantification result. Then after a global linear normalization (Callister *et al*., 2006) and outliers rejection, a weighted average of peptide ratios (Wang *et al*., 2006) is adopted as the corresponding protein quantification result. Protein abundance is determined as the sum of its unique peptide intensities. Finally, the protein and peptide quantification results are shown in flexible tables (Supplementary Fig. S2), and some statistical analyses can be further performed in user-friendly interfaces (Supplementary Figs S3 and S4). More importantly, the labeled and unlabeled XIC plots of each peptide are provided for quantification validation manually (Supplementary Fig. S5).

Meanwhile, to improve the efficiency of SILVER, a new data structure named IdxRaw is defined for fast accession of every spectrum (Supplementary Fig. S6).

3 RESULTS

SILVER was developed for fast and accurate quantification of stable isotope labeling data. The quantitative accuracy of SILVER was evaluated and compared with MaxQuant (v1.3.0.5) and Proteome Discoverer (v1.3, Thermo Scientific) using a large-scale dataset named MaxQuant dataset as well as two standard datasets (^{15}N and ^{18}O labeling) (Arsova *et al*., 2012; Liu *et al*., 2013) with different labeling ratios.

For MaxQuant dataset, only commonly quantified proteins by the three tools were evaluated for an objective comparison. As shown in Supplementary Figure S7, the median ratio of SILVER's result is the closest to the theoretical value compared with the other two software packages. For ^{18}O labeling dataset, SILVER performs as well as MaxQuant for the 1:1 labeling data, and its results are closer to the theoretical ratio for the 1:2 labeling data (Supplementary Fig. S8). What's more, SILVER can also handle ^{15}N labeling data that is another advantage over MaxQuant and Proteome Discoverer. There are three different ratios (H/L = 1:2, 1:5, 1:10) in the ^{15}N labeling dataset. The protein ratio medians are all close to the theoretical values, suggesting the high accuracy and sensitivity of SILVER (Supplementary Fig. S9). These results demonstrate that SILVER is able to handle large-scale complex stable isotope labeling data and displays a high quantitative accuracy and sensitivity despite data complexity (Supplementary Methods).

In addition, the quantification processing time of the three software tools is compared on MaxQuant dataset. As shown in Supplementary Table S1, the time consumed by SILVER is nearly one third of that consumed by MaxQuant and even nearly one tenth of that of Proteome Discoverer under the same computing conditions. More details are provided in Supplementary Methods.

4 CONCLUSION

In summary, we have developed a new efficient tool named SILVER for stable isotope labeling LC-MS data quantification. By using the quality control methods, SILVER can provide accurate and sensitive quantification results for SILAC, ^{15}N and ^{18}O labeling data with effective validation, visualization as well as statistical analysis functions. Other labeling strategies such as isobaric labeling and multiplexed labeling methods will be included in future versions.

Funding: Chinese National Basic Research Program (2010CB912700, 2011CB910601 and 2013CB910800), National High-Tech Research and Development Program (2012AA020201 and 2012AA020409), National Natural Science Foundation of China (Grant Nos. 31000587, 311171266, 21105121 and 21275160), State Key Laboratory of Proteomics [SKLP-O201004] and Beijing Municipal Natural Science Foundation 5122013.

Conflict of Interest: none declared.

REFERENCES

Arsova,B. *et al.* (2012) Precision, proteome coverage, and dynamic range of Arabidopsis proteome profiling using (15)N metabolic labeling and label-free approaches. *Mol. Cell. Proteomics*, **11**, 619–628.

Callister,S.J. *et al.* (2006) Normalization approaches for removing systematic biases associated with mass spectrometry and label-free proteomics. *J. Proteome Res.*, **5**, 277–286.

Cox,J. and Mann,M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.*, **26**, 1367–1372.

Cox,J. *et al.* (2011) Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.*, **10**, 1794–1805.

Keller,A. *et al.* (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.*, **74**, 5383–5392.

Li,N. *et al.* (2012) PepDistiller: a quality control tool to improve the sensitivity and accuracy of peptide identifications in shotgun proteomics. *Proteomics*, **12**, 1720–1725.

Liao,Z. *et al.* (2012) IsoQuant: a software tool for stable isotope labeling by amino acids in cell culture-based mass spectrometry quantitation. *Anal. Chem.*, **84**, 4535–4543.

Liu,K. *et al.* (2013) Evaluation of empirical rule of linearly-correlated peptide selection (ERLPS) for proteotypic peptide-based quantitative proteomics. *Proteomics*, in press.

Mortensen,P. *et al.* (2010) MSQuant, an open source platform for mass spectrometry-based quantitative proteomics. *J. Proteome Res.*, **9**, 393–403.

Nilsson,T. *et al.* (2010) Mass spectrometry in high-throughput proteomics: ready for the big time. *Nat. Methods*, **7**, 681–685.

Ong,S.E. *et al.* (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics*, **1**, 376–386.

Park,S.K. *et al.* (2008) A quantitative analysis software tool for mass spectrometry-based proteomics. *Nat. Methods*, **5**, 319–322.

Wang,G. *et al.* (2006) Label-free protein quantification using LC-coupled ion trap or FT mass spectrometry: reproducibility, linearity, and application with complex proteomes. *J. Proteome Res.*, **5**, 1214–1223.

Yates,J.R. 3rd *et al.* (2012) Toward objective evaluation of proteomic algorithms. *Nat. Methods*, **9**, 455–456.