Opinion

Increasing your peptide identification sensitivity without changing your mass spectrometer

Opinion

Proteomists has always been hunger for higher sensitivity in peptide identification, no matter how fast mass spectrometers improve. For low abundant peptides, we could never have enough fragments to make a perfect match to the theoretical spectra of the sequences, whatever fragmentation methods we use. There will always be peptides which fragments produced were just below the threshold we accepted as a positive identification.

I have been wondering why the retention time was not used yet to help the identification. Retention time is a relatively stable parameter of a peptide for a particular liquid chromatography system. It is independent from m/z. We used LC-MS/MS, but we only extract information from MS/MS. If the information from LC is used, we can describe the peptide with one more dimension of parameter. I guess scientists might want to find an system that could predict retention time before using it.[1] Considering all the LC conditions especially enormous modifications, it may be hard to make a satisfactory prediction system. Why don't we just record and use the empirical retention time for each peptide before we can make a perfect prediction, if we assume we can? The number of peptide is limited anyway. In 2009, when the data was not big, the empirical retention time was demonstrated to be able to increase the sensitivity of peptide identification without changing the mass spectrometer at all.[2] For example, the empirical retention times could be taken from the identification results of mixture A of 18 known proteins (including Rabbit GAPDH and Bovine catalase both at 20nM). The empirical retention time can then be used to help to identify these two proteins at lower concentration (both at 6nM) in mixture B of the same 18 known proteins. Without empirical time, these two proteins could not be identified in mixture B. This data was downloaded instead of intentionally generated with special caution of LC just for this retention time analysis. The empirical retention time database came from only a few technical repeats of an experiment. Now with big data, the accuracy of the retention time and the efficiency of identification should be much higher.

The other reason why proteomists and bioinformatists don't use empirical retention time is probably because they believe empirical retention time changes with the system they use. The question is: how often do we change our system? Do we buy different LC system every year? Do we change our column length or resin day by day? Actually we tend to do the similar type of sample with the same system for quite a while. For people who do urine proteome analysis like myself, the

Youhe Gao

Department of Biochemistry and Molecular Biology, Beijing Normal University, China

Correspondence: Youhe Gao, Department of Biochemistry and Molecular Biology, Beijing Normal University, Gene Engineering and Biotechnology Beijing Key Laboratory, Beijing, 100875, P. R. of China, Tel 861058804382, Email gaoyouhe@bnu.edu.cn

LC system may stay the same for years even the mass spectrometer was changed. The empirical retention time information was all wasted if the software developers do not use it for peptide identification. Even we changed the system, I suspect that the sequence of coming out of the column for each peptide may be more robust than retention time, as long as the resin remained the same.

I bet for all the software for peptide identification, the one that uses empirical retention time will have a much higher sensitivity than the one without using it. In the future we may have a few standardized LC settings and a huge database of peptide retention time, including lots of peptides with post-translational modifications. At least we burn money slower with empirical retention time.

Acknowledgements

Conflict of interest

The author declares no conflict of interest.

References

1. Strittmatter EF, Kangas LJ, Petritis K, et al. Application of peptide LC retention time information in a discriminant function for peptide identification by tandem mass spectrometry. *J Proteome Res*. 2004;3(4):760–769.

2. Sun W, Yang R, Zhang L, et al. Improve peptide identification by empirical peptide retention time database. *Rapid Commun Mass Spectrom*. 2009;23(1):109–118.