SILAC-based Quantitative Proteome Analysis with PEAKS

Wen Zhang, Ziaur Rahman, Yifang Chen, Lei Xin, and Baozhen Shan
Bioinformatics Solutions Inc., Waterloo, Canada

Aims

To overview SILAC (Stable Isotope Labeling with Amino Acids in Cell Culture)-based data analysis in PEAKS Q.

Summary:

PEAKS supports complex SILAC data analysis (e.g. time series, super-SILAC, etc.). A published dataset of SILAC-based comparative quantification of global proteomes was used to demonstrate SILAC data analysis in PEAKS.

Benefits

- Accurate and sensitive detection and association of SILAC feature pairs
- Transfer IDs between associated SILAC pairs and different MS runs
- Flexible experimental design and statistical tools to facilitate data analysis

2017. Aug. Revision

Introduction

➢ SILAC and Super-SILAC

SILAC (stable isotope labeling by amino acids in cell culture) has become one of the most popular labeling techniques for mass spectrometry (MS)-based quantitative proteomics [1]. In this metabolic labeling strategy, differential isotope labeled samples (proteins/peptides) are combined early in the experimental procedure and analyzed together by LC -MS/MS. Therefore, the variation introduced from sample processing is minimized. Since stable-isotope labeled peptides have almost the same physicochemical properties as their natural counterparts, the same peptides with differential labeling co-elute from the liquid chromatography (LC) column and their amounts can be accurately quantified relative to each other.

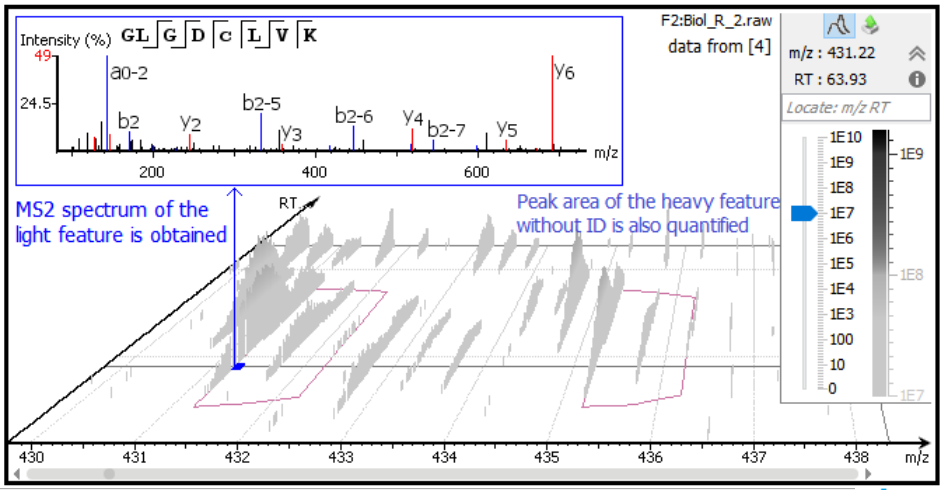
An extension of this labeling strategy is named super-SILAC [2, 3], where labeled samples can be produced separately and spiked into each of the experimental samples that are not amenable to metabolic labeling, e.g. human tissues. This pool of spiked-in heavy proteins is used as an internal standard for quantification.

PEAKS Q supports analyses of SILAC and super-SILAC types of data.

SILAC Quantification Algorithms in PEAKS Q

➢ ID transfer between associated SILAC pairs

PEAKS Q detects and associates 2 or 3-plex SILAC feature pairs that have the same charge, similar MS1 peak area correlation over retention time, expected mass shifts caused by labeling and fall within certain mass errors. If an identification is obtained from one of the labeled states, then the whole SILAC pair feature can be quantified and used for peptide and protein ratio calculations. For example, the light form of peptide GLGDCLVK was fragmented and identified in sample R_2 [4]. Although no MS2 spectrum was obtained from the K8 labeled heavy counterpart, the SILAC pair could still be quantified in PEAKS Q as highlighted in the feature vectors table. Id Count indicates the number of MS2 identified for each SILAC pair.



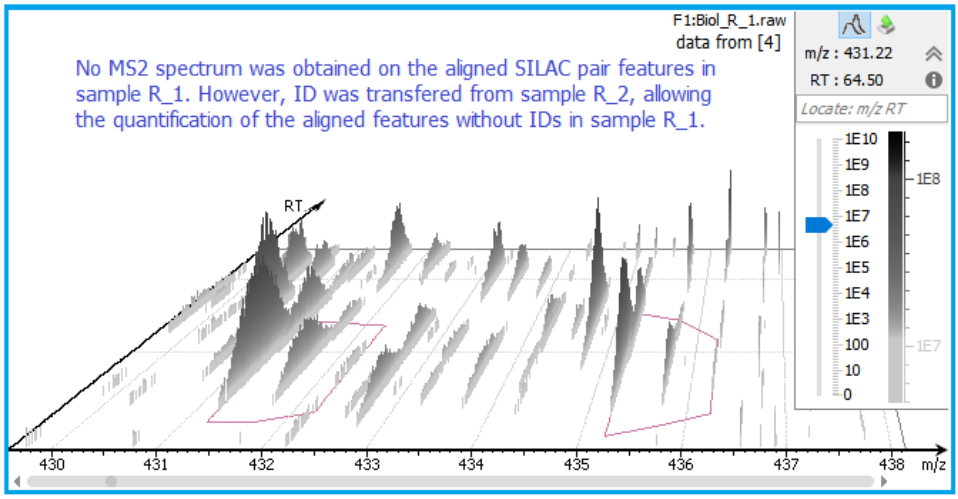
3 feature vectors

	Sample	Fraction	-10lgP	Quality Score	Light_Area	Heavy_Area	Light_Ratio	Heavy_Ratio	Id Count	m/z	z	ppm	RT	AScore
1	R_1	F 1	20.18	60.60	2.47E8	2.42E8	1.00	0.98	0	431.2292	2	1.5	63.93	
2	R_2	F 2	20.18	59.12	2.75E8	2.62E8	1.00	0.95	1	431.2292	2	1.5	63.93	
3	R_2_1	F 3	25.06	63.84	1.89E8	1.92E8	1.00	1.02	2	431.2288	2	0.5	64.30	

Peptide	-10lgP	Quality Score	Heavy : Ratio Profile	Group 1 : Light_Ratio	Group 1 : Heavy_Ratio	#Vector	Accession
GLGDC(+57.02)LVK(*)	25.06	61.19		1.00	0.98	3	P12236|ADT3_HUMAN:P05141|ADT2_HUMAN

➢ ID transfer between different MS runs after alignment

Retention times of different LC-MS runs are first aligned, then the MS/MS and ID can be matched from another run by aligning features within tight mass ranges and retention times, allowing quantification of SILAC pairs without any ID.



3 feature vectors

	Sample	Fraction	-10lgP	Quality Score	Control_Area	Treated_Area	Control_Ratio	Treated_Ratio	Id Count	m/z	z	ppm	RT	AScore
1	R_1	F 1	20.18	60.60	2.47E8	2.42E8	1.00	0.98	0	431.2292	2	1.5	63.93	
2	R_2	F 2	20.18	59.12	2.75E8	2.62E8	1.00	0.95	1	431.2292	2	1.5	63.93	
3	R_2_1	F 3	25.06	63.84	1.89E8	1.92E8	1.00	1.02	2	431.2288	2	0.5	64.30	

Case Study

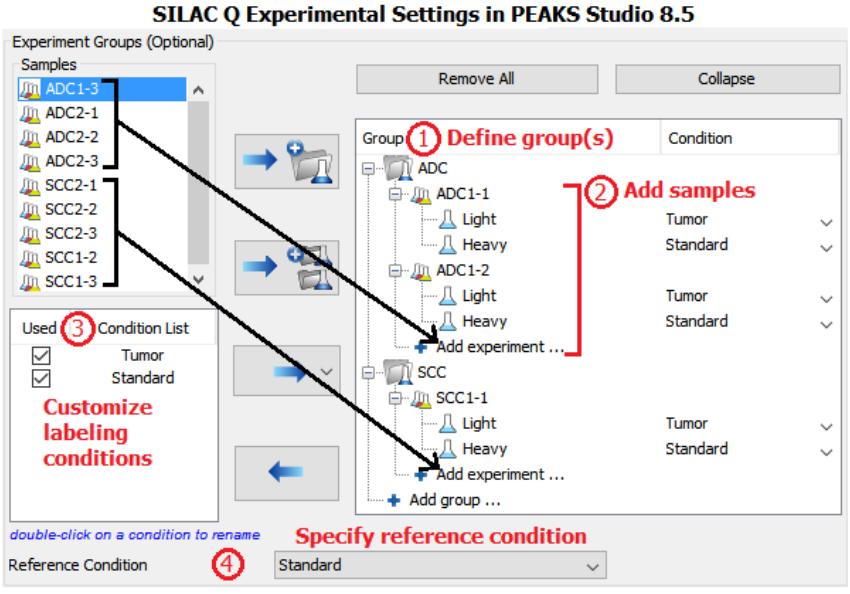
A published dataset [5] was used as an example to demonstrate the capability of SILAC-based data analysis in PEAKS Studio.

➢ Study Aims and Background

Adenocarcinoma (ADC) and squamous cell carcinoma (SCC) are the two major histological subtypes of non-small cell lung cancer (NSCLC), which are associated with patient prognosis. One of the main goals of the study was to identify signature proteins that distinguished the ADC and SCC subtypes.

➢ Experimental Design

Four patient tumor-derived xenograft proteomes (two ADC and two SCC, each with 3 replicates) were quantified and compared by the super-SILAC approach. Three NSCLC cell lines were labeled with $^{13}C_6{}^{15}N_2$-lysine (K8) and $^{13}C_6{}^{15}N_4$-arginine (R10) and heavy proteins were combined and used as the super-SILAC standard. The standard was spiked into non-labeled xenograft tissues at 1:1 protein ratio. Therefore, protein expression levels in ADC and SCC xenografts were compared with reference to the super-SILAC standard.



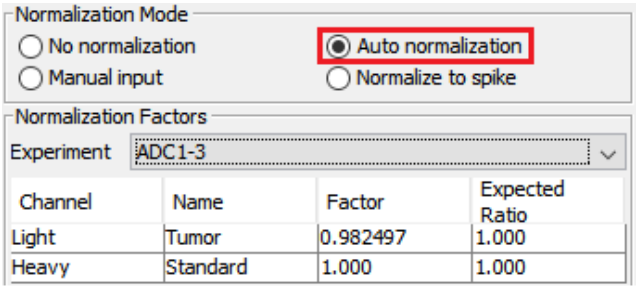
➢ Data analysis

MS data was analyzed in PEAKS Studio 8.5 using the built-in SILAC-2plex (R10, K8) method in PEAKS Q for quantification.

Results

➢ Normalization of SILAC ratios in each sample

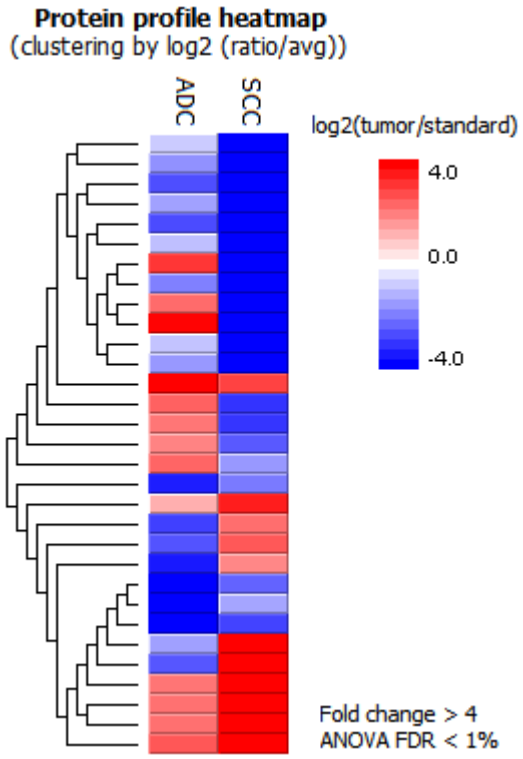
Auto normalization was first performed in PEAKS Q so that the total light and heavy intensities in each sample were equivalent since the same amount of light and heavy proteins were mixed. The adjusted normalization factor is displayed for each sample in the normalization setting window.



Channel	Name	Factor	Expected Ratio
Light	Tumor	0.982497	1.000
Heavy	Standard	1.000	1.000

➢ Identification of differentially expressed proteins between groups

SILAC ratios of proteins were calculated using the median of peptide ratios (MS1 peak area in the labeling channel relative to the reference channel). Proteins that had significantly differential expressions between ADC and SCC groups could be identified by applying a fold change filter of, e.g. at least 4, and an ANOVA test FDR 1% threshold. Welch's ANOVA is the statistical tool integrated in PEAKS Q for multi-group SILAC data analysis.



Conclusions

PEAKS Studio supports complex SILAC data analysis and offers built-in statistical tools to identify proteins with significant changes of expression levels.

Reference:

[1] Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., *et al.*, Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & cellular proteomics : MCP* 2002, *1*, 376-386.
[2] Neubert, T. A., Tempst, P., Super-SILAC for tumors and tissues. *Nature methods* 2010, *7*, 361-362.
[3] Geiger, T., Cox, J., Ostasiewicz, P., Wisniewski, J. R., Mann, M., Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nature methods* 2010, *7*, 383-385.
[4] Liberski, A. R., Al-Noubi, M. N., Rahman, Z. H., Halabi, N. M., *et al.*, Adaptation of a commonly used, chemically defined medium for human embryonic stem cells to stable isotope labeling with amino acids in cell culture. *Journal of proteome research* 2013, *12*, 3233-3245.
[5] Zhang, W., Wei, Y., Ignatchenko, V., Li, L., *et al.*, Proteomic profiles of human lung adeno and squamous cell carcinoma using super-SILAC and label-free quantification approaches. *Proteomics* 2014, *14*, 795-803.