Feature-based Peptide Identifications with PEAKS Studio X

Wen Zhang, Weiping Sun
Bioinformatics Solutions Inc., Waterloo, Canada

Abstract

PEAKS Studio X introduces a feature-based identification method into its unique *de novo* assisted identification workflow, which significantly maximizes the efficiency of peptide identification for shotgun analysis of complex proteomes. The deconvolution of chimeric MS2 generated from cofragmented peptides enables 1.11 identified peptides per tandem spectrum, resulting in an increase of 40% more peptides and 24% more proteins identified than PEAKS Studio 8.5.

Introduction

During the liquid chromatography-MS/MS data-dependent analysis of complex proteomes, many peptides are co-eluted. With a ± 1~2 m/z MS/MS isolation window, such co-eluted peptides can be co-fragmented, resulting in chimeric MS2 spectra. However, most conventional analysis workflows ignore the cofragmentation information. In PEAKS Studio X, a new method is added to associate nearby unfragmented features with the chimera tandem spectrum, thus deriving peptide feature-spectrum matches. This workflow gives higher identifications than PEAKS Studio 8.5 and some published methods.

Methods

A DeMix workflow was previously proposed to identify chimera spectra [1]. The dataset used in the publication was re-analyzed with PEAKS Studio 8.5 and PEAKS Studio X and the results were compared. Briefly, HeLa cells were extracted, reduced with DTT, alkylated with iodoacetamide, digested with trypsin and analyzed by LC-MS/MS on an Orbitrap Q Exactive mass spectrometer in a data-dependent manner. Three replicates were performed with a MS2 isolation window of ± 2 Da. Thermo .RAW files were analyzed with PEAKS Studio 8.5 and PEAKS Studio X. Associating feature with chimera scan was enabled on PEAKS Studio X. 1% FDR was applied at peptide-spectrum match (PSM) level.

Analysis parameters	Settings
Precursor mass tolerance	20 ppm
Fragment mass tolerance	0.03 Da
Fixed PTM	Carbamidomethylation (C)
Variable PTMs	Oxidation (M), acetylation (Protein N-terminal)
Database	Uniprot Human Complete Proteome database (v201808)

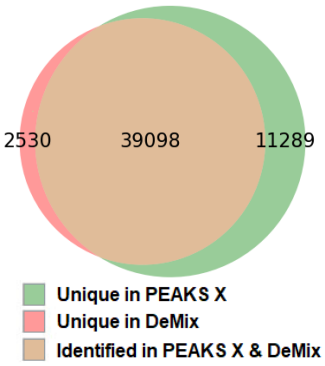
Results

➢ Increased peptide identifications in PEAKS Studio X results compared to PEAKS Studio 8.5

By enabling chimera scan analysis, on average PEAKS Studio X identifies 40% more peptides and 24% more proteins for each HeLa replicate than PEAKS Studio 8.5. In addition, more than 97% peptides and 87% proteins identified with PEAKS Studio 8.5 were also present in PEAKS Studio X results (see tables on next page).

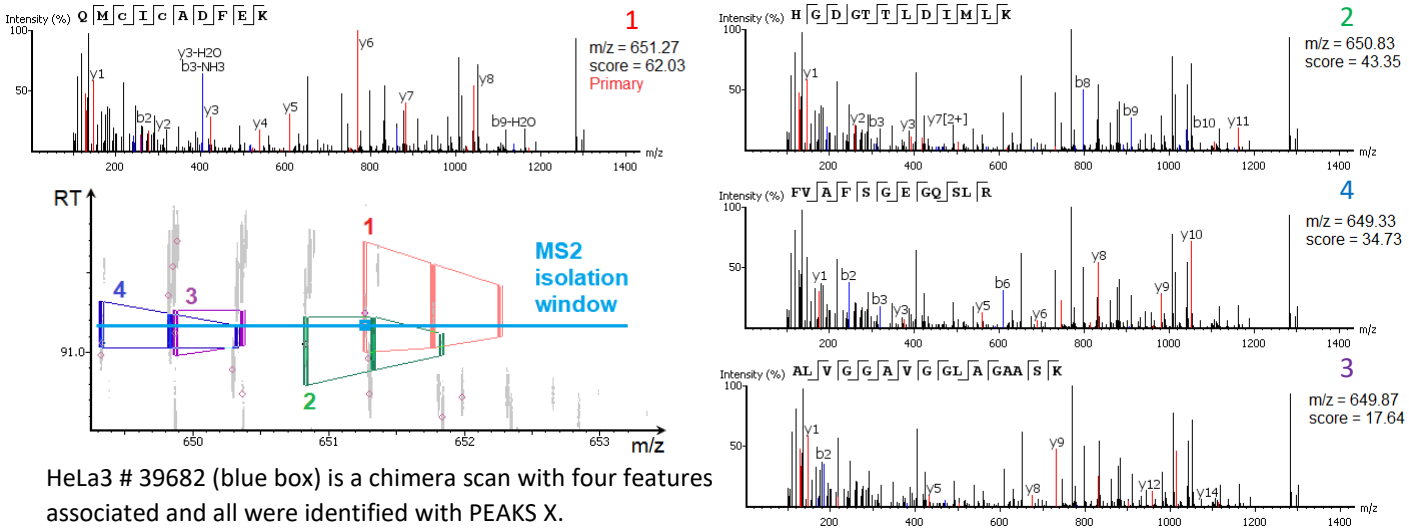
Peptides	8.5 unique	Common (%)	X unique (%)
HeLa 1	635	26810 (97.69)	11053 (40.27%)
HeLa 2	676	28011 (97.64)	11472 (39.99%)
HeLa 3	618	27042 (97.77)	10883 (39.35%)

Proteins	8.5 unique	Common (%)	X unique (%)
HeLa 1	862	5970 (87.38)	1715 (25.10)
HeLa 2	799	6150 (88.50)	1714 (24.67)
HeLa 3	734	5981 (89.07)	1795 (26.73)



In addition, we compared our results with the publication where the three replicate data were analyzed together. PEAKS Studio X identified 20% more peptides and 7% more proteins than the DeMix method and 32% more peptides and 20% more proteins than MaxQuant (table below). Furthermore, 94% peptides identified by DeMix workflow were also identified by PEAKS Studio X (left figure).

Search Engine	MaxQuant		DeMix		PEAKS Studio	
Chimera spectra analysis	-	+	-	+	8.5	X
Peptide-spectrum matches at 1% FDR	115135	131597	122208	200292	128266	179622
PSM per MS/MS	0.714	0.816	0.758	1.242	0.795	1.114
Unique peptide sequences	32646	38112	33712	41628	36608	50387
Protein groups	4409	4642	4707	5167	4638	5540



HeLa3 # 39682 (blue box) is a chimera scan with four features associated and all were identified with PEAKS X.

Conclusions

Feature-based identification workflow in PEAKS Studio X allows chimera scan analysis and enables higher numbers of identifications of peptides and proteins. This method achieves superior performance than other similar workflows such as DeMix and MaxQuant.

References

[1] Zhang B, et al., DeMix workflow for efficient identification of cofragmented peptides in high resolution data-dependent tandem mass spectrometry. Mol Cell Proteomics. 2014 Nov;13(11):3211-23.