PepPre: Promote Peptide Identification & using accurate and comprehensive precursors

bridge between DIA data and DDA engines

pFindStudio

ICT

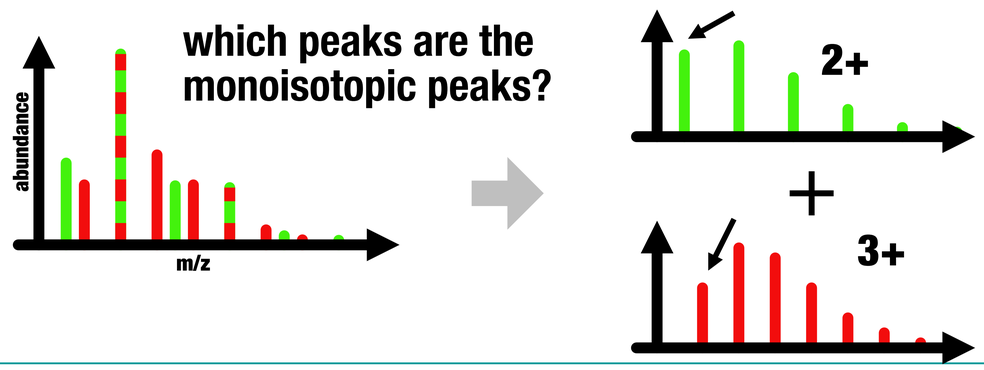
Introduction

Accurate and comprehensive peptide precursor ions are crucial to tandem mass spectrometry-based peptide identification.

- An identification engine can derive great advantages from the search space reduction hinted by credible and detailed precursors.
- By considering multiple precursors per mixed spectrum, both the number of identifications and the spectrum explainability can be substantially improved.

Methods

Question: accurate mass & charge state of a peptide



Model: a spectrum is a vector

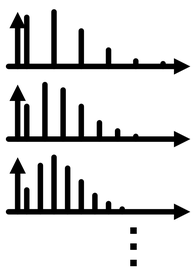
$$\mathbf{s} \approx x_1 \mathbf{p}_1 + x_2 \mathbf{p}_2 + \cdots$$

enumerate each peak & charge state

Candidate Peaks:
m/z = 1010.00 Th
m/z = 1010.25 Th
m/z = 1010.50 Th
m/z = 1010.75 Th
m/z = 1011.25 Th
m/z = 1011.50 Th
⋮

Candidate Ions:
m = 2020.00 Da, z = 2+
m = 3030.00 Da, z = 3+
m = 4040.00 Da, z = 4+
⋮
m = 2020.50 Da, z = 2+
⋮

Candidate Clusters:



⋮

Solution: linear programming

$$\mathbf{s} \approx x_1\mathbf{p}_1 + x_2\mathbf{p}_2 + \cdots + x_n\mathbf{p}_n \quad (1)$$

$$\Delta = |\mathbf{s} - \sum_{i=1}^{n} x_i \mathbf{p}_i| \quad (2)$$

$$\begin{aligned} &\min \quad \Delta \\ &\text{s.t.} \quad |\mathbf{s} - \sum_{i=1}^{n} x_i \mathbf{p}_i| = \Delta \\ &\qquad x_i \geq 0 \qquad i = 1, 2, \ldots, n \end{aligned} \quad (3)$$

Ching Tarn / 谭永卿, Email: i@ctarn.io, WeChat: TarnYeongChing

pFind Team, Institute of Computing Technology, Chinese Academy of Sciences / 中国科学院计算技术研究所pFind课题组

Made for 12th CNHUPO, 2023.9.24-27, Chengdu / 成都, China / 中国

Tarn, C. et al. (2023). PepPre: Promote Peptide Identification Using Accurate and Comprehensive Precursors. bioRxiv, 2023.2005.2013.540645. https://doi.org/10.1101/2023.05.13.540645

Applications

Promote DDA Peptide Identification

HeLa, 2 hours, no fraction

Normal Isolation Window (±1 Th):

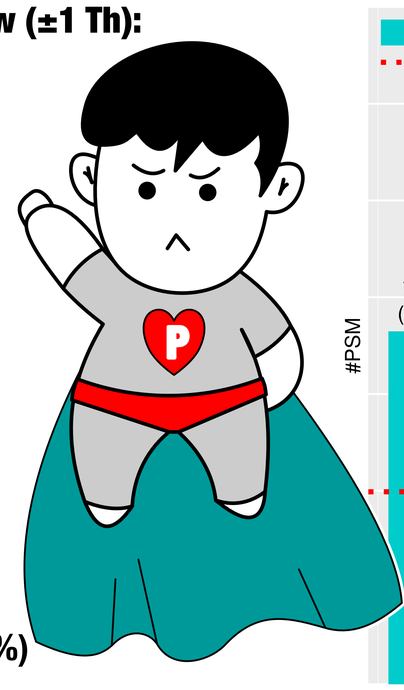
Baseline:
#PSM = 39794
#Pep. = 28635

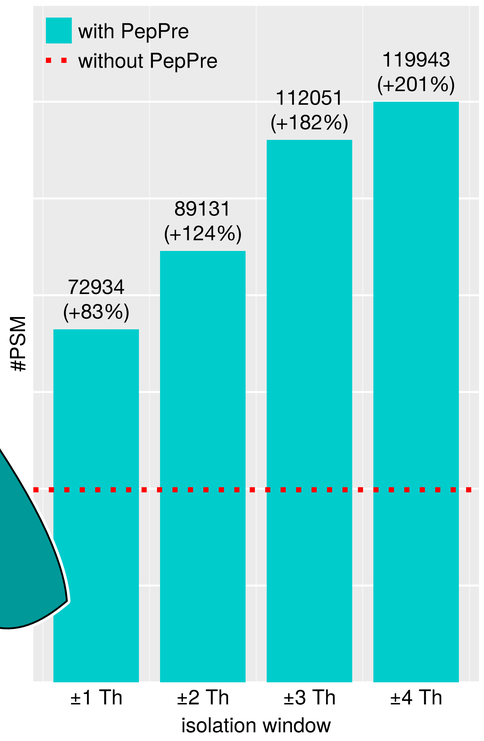
PepPre:
#PSM = 72934 (+83%)
#Pep. = 45341 (+58%)

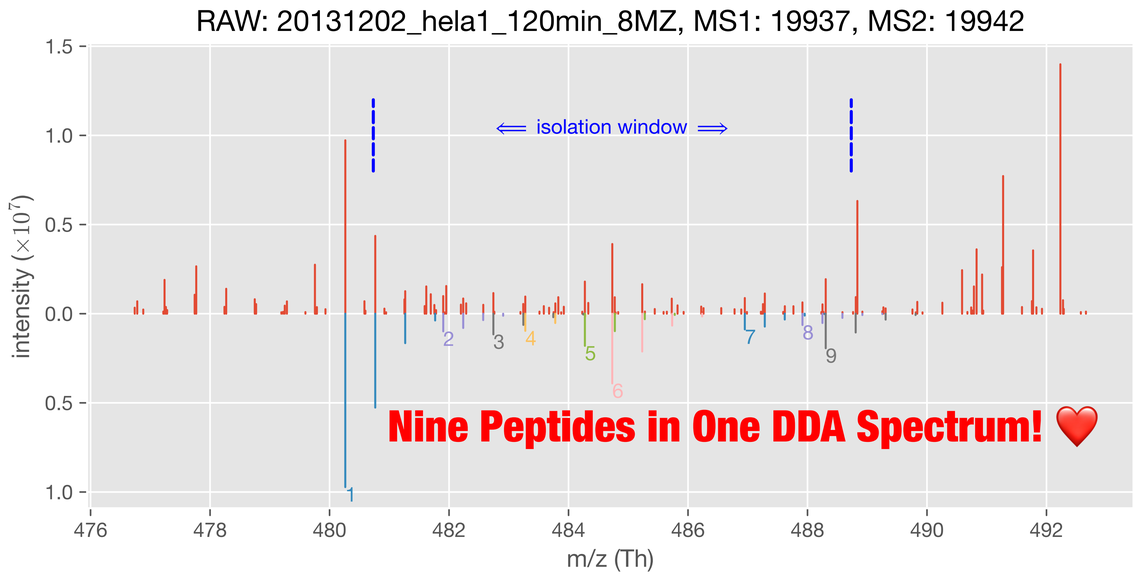
Wide Window (±4 Th):

Baseline:
#PSM = 39590
#Pep. = 28255

PepPre:
#PSM = 119943 (+203%)
#Pep. = 47370 (+65%)







	m/z (Th)	charge state	sequence	modification	q-value
#1 ion	480.263305	2	SLEVLENR	-	0.08%
#2 ion	481.903038	3	SEPHSLSSEALMR	-	0.11%
#3 ion	482.742083	2	FEEELAAR	-	0.36%
#4 ion	483.276243	2	LLSQVSYR	-	0.20%
#5 ion	484.273943	2	KEFFLQR	-	0.46%
#6 ion	484.731832	2	NGDYLCVK	Carbamidomethyl[C]@6	0.02%
#7 ion	486.950699	3	KPLPDHVSLVEPK	-	0.14%
#8 ion	487.915765	3	LSNRPAFMPSEGR	-	0.45%
#9 ion	488.305021	2	KGVALNFVK	-	0.11%

Bridge between DIA data and DDA engines

http://peppre.ctarn.io