



# **Quality Control (QC) Analysis with PEAKS 11**

Yiming Xiao, PhD, Application Scientist Qing Zhang, MSc, Software Systems Manager Kyle Hoffman, PhD, Applications Manager

Bioinformatics Solutions Inc., Waterloo, Canada

#### Abstract:

PEAKS 11 supports QC analysis to monitor attributes of LC-MS/MS data that fail to meet a specified standard. In this study, we utilized published label-free quantification (LFQ) datasets to effectively showcase the application and functionality of the QC module within PEAKS 11.

in BIOINFORMATICS-SOLUTIONS-INC-/

## Introduction:

LC-MS-based protein/peptide quantification has garnered considerable attention in the fields of physiopathology and pharmaceutical research. In particular, clinical studies rely on a substantial number of samples to ensure adequate statistical power, often necessitating over 50 samples per group due to the significant interindividual variation [1]. For cohorts with such large sample sizes, the implementation of a rigorous Quality Control (QC) step becomes indispensable for unobstructed and reliable statistical analyses. Only through this rigorous approach can experimental conclusions be confidently validated.

Addressing this need, PEAKS 11 introduces a specialized QC module that enhances protein/peptide identification and quantification results. The module offers sophisticated QC analysis, ensuring the presentation of interactive and user-friendly outputs tailored to the specific requirements of the user's research.

## Case study: benchmarking data set analysis with PEAK 11:

A published data set [2] was used as an example to demonstrate the capability of label free quantification (LFQ) and QC analysis in PEAKS 11.

## Study aims and background:

LC-MS-based protein/peptide quantification has been a subject of continuous interest in the fields of physiopathology and pharmaceutical studies. Label-free quantification (LFQ) allows for different MS acquisition strategies, namely, data-dependent acquisition (DDA) and data-independent acquisition (DIA). While DDA represents the more traditional approach, it is susceptible to high missing values and under-sampling. On the other hand, DIA offers increased robustness and can mitigate biases associated with DDA. However, DIA also presents challenges, such as a high false positive rate in complex MS2 spectra.

In this investigation, identical samples were analyzed using high-resolution DDA and DIA (SWATH) methods, and subsequently compared. The quantification of high-resolution DDA (HS-DDA) demonstrated comparable accuracy and precision to DIA, and in certain cases, even outperformed DIA for proteins characterized by low abundance and small fold-changes.

For the purpose of this application note, only DDA data was utilized to demonstrate the capabilities of the QC module in PEAKS 11. This choice was made to emphasize the software's comprehensive QC analysis and its adeptness in presenting results effectively.

# **Experimental design:**

Five groups of samples were prepared with three different proteomes (human, E.coli and yeast) and run with five technical replicates (n=25). The human protein amount proportion was 60% across all samples. The portion of E. coli and yeast protein amounts were as follows:

A: 5%/35%, B: 7.5%/32.5%, C: 10%/30%, D: 15%/25%, E: 20%/20%.

DDA data was acquired with MS1 240000 resolution and MS2 15000 resolution. The gradient was 180 min long. A detail of LC-MS method can be found in [2].

# LFQ Data Analysis:

MS data (5 runs \*5 samples) was analyzed in PEAKS 11 as LFQ with PEAKS Q module. LFQ was applied with the implementation of match-between-run and total ion current (TIC) normalization techniques. Further specifics on the search parameters and the configuration of the quality control (QC) analysis are presented in Figure 1a (LFQ) and Figure 1b (QC).

While using at least two peptides per protein for quantification, 5157 protein groups are quantifiable. Statistics of filtered results are shown in Figure 2a; the volcano plot is shown in Figure 2b.

Sample Group  Select All Underd All  Group  Group  Group  Group  Group  Color		Reporter ion Quantification (eg. )		Trecarbor for quantin	cation (eg. SIL/
Select All Unselect All       Ceer       Reference         Select All Unselect All       Ceer       Select All Comp         Select All Unselect All       Select All Selection       Select All Selection         Match Between Run       Select All Selection       Select All Selection       Select All Selection         Match Between Run       Match Between Run       Select All Selection       Select All Selection       Select All Selection         Match Between Run       Mass Error Tolerance       20       ppm       Retention Time Shift Tolerance(min):       Auto Detect If Feature Intensity 2 100000       If Range         RT Range       S RT S       Max       Base Sample: A01       Feature Intensity 2 100000       If Select All Selection         Project Creaston       Viorkflow Selection       Data Refines       D Search       Quantification       2 Aluto Detect If Sender All Selection         Project Creaston       Viorkflow Selection       Data Refines       D Search       Quantification       QC         Control Sample:       A01       Acceptance Tolerance (% 10.0       QC       If Selection Comps       If Se	Sample Group				
Group       Group         Image: Server S	Select All Unselect All	clear			Ren
Image: State State       Image: State State         Poptide Feature Filters: Auto Normalization.       Auto Detect Image: State State         Project Creation       Workflow Selection       Data Refine         Poptide Feature Filters: Image: State       Poptide Selection       Data Refine         Poptide Feature Filters: Image: State       Poptide Feature Filters: Image: State       Poptide Feature Filters: Image: State         Project Creation       Workflow Selection       Data Refine       DB Search       Quantification       QC         Poptide features: Image: State       Im				Group	Color
Project Creation       Workflow Selection       Data Refine       DB Search       Cuantific Projects         Project Creation       Workflow Selection       Data Refine       DB Search       Cuantific Projects         Project Creation       Workflow Selection       Data Refine       DB Search       Cuantific Projects         Project Creation       Workflow Selection       Data Refine       DB Search       Cuantific Projects         Project Creation       Workflow Selection       Data Refine       DB Search       Cuantific Projects         Project Creation       Workflow Selection       Data Refine       DB Search       Cuantific Projects         Project Creation       Workflow Selection       Data Refine       DB Search       Cuantific Projects         V CC       Antibutes       VC CA throbates:       VC QC Attributes:       VC QC Attributes:         V MAX2 Sons       V Morkflow Selection       Data Refine       DB Search       Cuantific Projects         V MAX2 Sons       VC Sons       VC Sons       VC CA throbates:       VC QC Attributes:         V MAX2 Sons       VC Sons       VC Sons       VC Sons       VC Sons         V MAX2 Sons       VC Sons       VC Sons       VC Sons       VC Sons         V Morkflow (min)       VC Sons				Gloup	color
Image:					
Image:       O       S ST 4       Max       Base Sample:       Auto Detect       Feature Intensity 2       10000       Image:         RT Range:       S ST 4       Max       Base Sample:       Auto Detect       Feature Intensity 2       10000       Image:         RT Range:       S ST 4       Max       Base Sample:       Auto Detect       Feature Intensity 2       10000       Image:         RT Range:       Auto Detect       Sentimetric       Auto Detect       Feature Intensity 2       10000       Image:       <		>	В		
Match Between Run         Mass Error Tolerance         Mass Error Tolerance         20       ppm<         Retention Time Shift Tolerance(min):         Auto Detect       Feature Intensity 2         Mass Error Tolerance       20         ppm       Retention Time Shift Tolerance(min):         Auto Detect       Feature Intensity 2         Mass Error Tolerance       20         Peptide Feature Filters       Avg. Area 2 200000, Quality 2 10, 2 5 Charge 5 5, Peptide ID Count 2 0 per group, and detected in at least 1 samples per group.         Protect       ✓ Significance Method: ANOVA, Modified Form Exclusion, Remove Outlier, Use Top 3 peptide, Significance 2 0,0,1 5 Fold change 5 64, has at least 1 used pep         Normalization       ✓ Method: Auto Normalization.         @ Project Creation       Workflow Solection       Data Refine       DB Search       Quantified Protein Groups         @ All Antibute <ul> <li># Mo2</li> <li>#</li></ul>			·) c		-
Auto Detect       Feature Intensity 2       00000 \$       I         Match Between Run       Mass Error Tolerance       20 \$ ppm * Retention Time Shift Tolerance(min): Auto Detect       Feature Intensity 2       00000 \$ I         RT Range       RT Range       RT Range       RT Range       RT Range       RT Range         RT Range       Auto Detect       Feature Intensity 2       00000 \$ I       I         Project Greature Filters       Avg. Area 2 200000, Quality 2 10, 2 5 Charge 5 5. Peptide ID Count 2 0 per group, and detected in at least 1 samples per group.         Project Greation       Mork Hood Auto Normalization.       C         Project Greation       Workflow Selection       Data Refins       DB Search       Quantification       OC         Project Greation       Workflow Selection       Data Refins       DB Search       Quantification       OC         Project Greation       Workflow Selection       Data Refins       DB Search       Quantification       OC         Project Greation       Workflow Selection       Data Refins       DB Search       Quantification       OC         Project Greation       Workflow Selection       Data Refine       DB Search       Quantified Peptides       Quantified Peptides         All Attribute       If Missing Value (%)       If Control Sample       A		· · · · · · · · · · · · · · · · · · ·	D		-
Match Between Run         Mass Error Tolerance:         20 • ppm • Retorion Time Shift Tolerance(min):         Auto Detect         RT Range         RT Range         Peptide Feature Filters         Avg. Area 2 2000000, Quality 2 10, 2 ≤ Charge 5 5, Peptide ID Count 2 0 per group, and detected in at least 1 samples per group.         Protect Feature Filters         Avg. Area 2 2000000, Quality 2 10, 2 ≤ Charge 5 5, Peptide ID Count 2 0 per group, and detected in at least 1 samples per group.         Protect Feature Filters         Avg. Area 2 2000000, Quality 2 10, 2 ≤ Charge 5 5, Peptide ID Count 2 0 per group, and detected in at least 1 samples per group.         Protect Creation         Workflow Selection       Data Refine         DB Search       Quantification         QC         Control Sample:       Acceptance Tolerance (%):         Q C         Control Sample:       Acceptance Tolerance (%):         Q All Attributes:       Vertering Cattributes:         Vertering Creation       Workflow Selection       Data Refine         Mass 2 sample:       Q C         Project Creation       Workflow Selection       Data Refine         DB Search       Quantified Poties:       Vertibute Search         Q and thethilder Potien Groups       Quantified Potien Groups			<u>]</u> .		
Match Between Run   Mass Error Tolerance:   Ø      pm      Retention Time Shift Tolerance(min): <ul> <li>Auto Detect</li> <li>Feature Intensity 2 10000</li> <li>I</li> </ul> RT Range:   0.0      ST      Max      Base Sample: A01     Peptide Feature Filters:   Aug. Area 2 200000, Quality 2 10, 2 4 Charge 4 5, Peptide ID Count 2 0 per group, and detected in at least 1 samples per group.   Protein      Significance Method: ANOVA, Modified Form Exclusion, Remove Outlier, Use Top 3 peptide, Significance 2 0.0, 1 4 fold change 4 64, has at least 1 used per Normalization: <ul> <li>Reserve Oct Finish</li> </ul> Project Creastion Workflow Selection   Data Refine: Data Refine:   Data Refine: DB Search   Quantification: OC   OC QC   Control Sample: Auto Acceptance Tolerance (%): 10.0   All Attribute: Identification QC Attributes:   If Mode: Identified Features:   If Mode: Identified Features:   If Mode: Identified Features:   If How Mode: Identified Features:   If How Mode: Identified Features:   If How Mode: If Popteins Groups   If How Mode: If Popteins:					
Match Between Run         Masc Error Tolerance:       20          ppm          Retention Time Shift Tolerance(min):         Auto Detect         Feature Intensity 2         100000         i         RT Range         RT Range         RT Range         00					
Match Between Run         Mass Error Tolerance:       20 • ppm • Retention Time Shift Tolerance(min):  Auto Detect • Feature Intensity 2 10000 • :         RT Range       RT Kange         RT Range:       0.0 • • • • • • • • • • • • • • • • • •		<b>~</b> 🔺			
Match Between Run         Mass Error Tolerance:       20          ppm          Retention Time Shift Tolerance(min):					
Match Between Run         Mass Error Tolerance:         20 • ppm • Retention Time Shift Tolerance(min):         RT Range         RT Range:         0.0 • S RT S         Max • Base Sample:         Auto Detect:         Peptide Feature Filters:         Avg. Area 2 200000, Quality 2 10, 2 S Charge 5 5, Peptide ID Count 2 0 per group, and detected in at least 1 samples per group.         Protein       Significance Method: ANOVA, Modified Form Exclusion, Remove Outlier, Use Top 3 peptide, Significance 2 0.0, 1 S Fold change S 64, has at least 1 used per         Normalization       Method: Auto Normalization.					
Match Between Run         Mass Error Tolerance       20					
Match Between Run         Mass Error Tolerance:       20					
Match Between Run Mass Error Tolerance: 20					
Mass Error Tolerance 20 ppm Retention Time Shift Tolerance(min): ↓ Auto Detect ♥ Feature Intensity ↓ 10000 ↓ i   RT Range 0.0 ↓ SRT ≤ Max Base Sample: A01 ↓   Peptide Feature Filters Avg. Area ₺ 200000, Quality ₺ 10, 2 ≤ Charge ≤ 5, Peptide ID Count ₺ 0 per group, and detected in at least 1 samples per group. Protein Significance Method: ANOVA, Modified Form Exclusion, Remove Outlier, Use Top 3 peptide, Significance ₺ 0.0, 1 ≤ Fold change ≤ 64, has at least 1 used pep Normalization Method: Auto Normalization. Project Creation Workflow Selection Data Refine DB Search Quantification QC Control Sample: A01 • Acceptance Tolerance (%): 10.0 All Attribute Ø 401 • Acceptance Tolerance (%): 10.0 Validetified Features: Very Missing Value (%): # Used Quantified Peptides Ø 401 • Acceptance Tolerance (%): 10.0 Validetified Features: Ø 4 Identification QC Attributes: Very Missing Value (%): Full Width (min) Full Protein S Protein Groups Full Width (min) Full Width (min) Full Width (min) Full Width (min) Full Proteins Full Protein S Full Proteins Full Proteins Full Protein S Full Proteins Full Proteins Full Proteins Full Protein S Full Proteins Full Proteins Full Proteins Full Proteins Full Proteins Full Protein S Full Proteins Full Pr	Match Between Run				
RT Range   RT Range:   0.0   Stristing:   Avg. Area ≥ 200000, Quality ≥ 10, 2 ≤ Charge ≤ 5, Peptide ID Count ≥ 0 per group, and detected in at least 1 samples per group.   Protein   Significance Method: ANOVA, Modified Form Exclusion, Remove Outlier, Use Top 3 peptide, Significance ≥ 0.0, 1 ≤ Fold change ≤ 64, has at least 1 used per   Normalization   Method: Auto Normalization.   Project Creation   Workflow Selection   Data Refine   DB Search   Quantification   QC   Control Sample:   A01   Acceptance Tolerance (%):   10.0   Altributes   Data QC Attributes:   Undiffication QC Attributes:   Use The Significance of the Signification   Attribute   # MS1   # MS2   # MS2   # MS2   # Sequences   # Identified Features   # Portein Groups   # Protein Groups   # Mixing Value (%)   # Features   # Furthin Groups   # All Proteins	Mass Error Tolerance: 20 🌲 pp	m 🔹 Retention Time Shift Tolerance(min):	🗍 Auto Detect 🗸	Feature Intensity ≥ 100000	\$ [ <b>:</b> ]
RT Range   RT Range:   00 • ≤ RT ≤ Max: • Base Sample: A01 •   Peptide Feature Filters  Avg. Area ≥ 200000, Quality ≥ 10, 2 ≤ Charge ≤ 5, Peptide ID Count ≥ 0 per group, and detected in at least 1 samples per group. Protein  Significance Method: ANOVA, Modified Form Exclusion, Remove Outlier, Use Top 3 peptide, Significance ≥ 0.0, 1 ≤ Fold change ≤ 64, has at least 1 used peptide. Normalization Method: Auto Normalization.    Project Creation Workflow Selection Data Refine DB Search Quantification QC    Project Creation Workflow Selection Data Refine DB Search Quantification QC    Project Creation Workflow Selection Data Refine DB Search Quantification QC    Project Creation Workflow Selection Data Refine DB Search Quantification QC    Project Creation Workflow Selection Data Refine DB Search Quantification QC    Project Creation Workflow Selection Data Refine DB Search Quantification QC    Project Creation Workflow Selection Data Refine DB Search Quantification QC					
RT Range: 0.0 → SRT S Max → Base Sample: A01 → Peptide Feature Filters Arg. Area ≥ 200000, Quality ≥ 10, 2 ≤ Charge ≤ 5, Peptide ID Count ≥ 0 per group, and detected in at least 1 samples per group. Protein Significance Method: ANOVA, Modified Form Exclusion, Remove Outlier, Use Top 3 peptide, Significance ≥ 0.0, 1 ≤ Fold change ≤ 64, has at least 1 used peptide. Normalization Method: Auto Normalization.  Project Creation Workflow Selection Data Refine DB Search Quantification QC Control Sample: A01 → Acceptance Tolerance (%): 10.0  Attributes Data QC Attributes:  Max → Max → Base Sample: A01  Attribute  Att	RT Range				
Peptide Feature Filters       Avg. Area ≥ 200000, Quality ≥ 10, 2 ≤ Charge ≤ 5, Peptide ID Count ≥ 0 per group, and detected in at least 1 samples per group.         Protein       Significance Method: ANOVA, Modified Form Exclusion, Remove Outlier, Use Top 3 peptide, Significance ≥ 0.0, 1 ≤ Fold change ≤ 64, has at least 1 used per         Normalization       Method: Auto Normalization. <ul> <li>Reack</li> <li>QC</li> </ul> Project Creation       Workfilow Selection       Data Refine       DB Search       Quantification       QC         Qc       Control Sample:       A01       Acceptance Tolerance (%):       10.0       EQ QC Attributes:       EQ QC Attributes:         Image: MS2       Identification QC Attributes:       Image: Peptide S       Quantified Peptides       Quantified Peptides         Image: MS2       Image: MS2 Scans       Image: Peptides       Quantified Protein Groups       Quantified Features         Image: Prestures       Image: Peptides       Image: Peptides       Quantified Peptides       Quantified Features         Image: Protein Groups       Image: Protein Groups       Image: Protein Groups       Image: Protein Groups       Image: Peptides       Image: Peptides         Image: Protein Groups       Image: Protein G	RT Range: 0.0 - ≤ RT ≤	Max 👻 Base Sample: A01 👻			
Protein Creation       Workflow Selection       Data Refine       DB Search       Quantification       QC         Project Creation       Workflow Selection       Data Refine       DB Search       Quantification       QC         QC       QC       Image: Second Control Sample:       Attribute       Image: Second Control Sample:       Attribute         Image: Model Auto Normalization       Image: Second Control Sample:       Attribute       Image: Second Control Sample:       Attribute         Image: Model Auto Normalization       Image: Second Control Sample:       Attribute       Image: Second Control Sample:       Attribute         Image: Model Auto Normalization       Image: Second Control Sample:       Attribute       Image: Second Control Sample:       Attribute         Image: Model Auto: Acceptance Tolerance (%):       10.0       Image: Second Control Sample:       Attribute       Image: Second Control Sample:       Image:	Pentide Feature Filters	> 200000 Quality > 10 2 < Charge < 5 Pentide ID Cour	$nt \ge 0$ per group and determined	cted in at least <b>1</b> samples per orr	oup
Protein        Significance Method: ANOVA, Modified Form Exclusion, Remove Outlier, Use Top 3 peptide, Significance ≥ 0.0, 1 ≤ Fold change ≤ 64, has at least 1 used peptide.         Normalization		2 200000, Quarty 2 10, 2 3 charge 3 5, 1 cplace is cour	in 2 6 per group, and dete	ceco in delease i sumples per gre	oup.
Normalization       Method: Auto Normalization.	Protein 🛛 🖍 🛛 Significance Method: ANO	VA, Modified Form Exclusion, Remove Outlier, Use T	op 3 peptide, Significance	$\geq$ 0.0, 1 $\leq$ Fold change $\leq$ 64, has	s at least <b>1</b> used pept
Normalization       Control Sample:       Norkflow Selection       Data Refine       DB Search       Quantification       QC         QC       QC         Control Sample:       A01 •       Acceptance Tolerance (%):       10.0         Image: A01 •       Acceptance Tolerance (%):       10.0         All Attributes       Image: Autor	Normalization Auto Norr	nalization			
Project Creation        Workflow Selection       Data Refine       DB Search       Quantification       QC         QC       QC         Control Sample:       A01       Acceptance Tolerance (%):       10.0         ✓ All Attributes       Identification QC Attributes:       LFQ QC Attributes:         ✓ # MS1       ////////////////////////////////////		nanzauon.			
Project Creation       Workflow Selection       Data Refine       DB Search       Quantification       QC         QC       QC       Control Sample:       A01       Acceptance Tolerance (%):       10.0         Image: A01       Acceptance Tolerance (%):       10.0       Image: A01       Image: A01       Image: A01         Image: A01       Acceptance Tolerance (%):       10.0       Image: A01       Image: A01       Image: A01         Image: A01       Acceptance Tolerance (%):       10.0       Image: A01       Image: A01       Image: A01         Image: A01       Acceptance Tolerance (%):       10.0       Image: A01       Image: A01       Image: A01         Image: A01       Acceptance Tolerance (%):       10.0       Image: A01       Image: A01       Image: A01         Image: A01       Acceptance Tolerance (%):       Image: A01       Image: A01       Image: A01       Image: A01         Image: A11 Attributes:       Image: A11				< Back	QC Finish
Project Creation       Workflow Selection       Data Refine       DB Search       Quantification       QC         QC       QC         Control Sample:       A01       Acceptance Tolerance (%):       10.0         Image: All Attributes       Image: Acceptance Tolerance (%):       10.0         Image: All Attributes       Image: All Attributes       Image: All Attributes         Image: All Attribute       Image: All Attribute       Image: All Attribute         Image: All Attribute       Image: All Attribute       Image: All Attribute         Image: All Attribute       Image: All Attribute       Image: All Attribute         Image: All MS2       Image: All Attribute       Image: All Attribute         Image: All MS2       Image: All Attribute       Image: All Attribute         Image: All MS2       Image: All Attribute       Image: All Attribute         Image: All MS2       Image: All Protein Groups       Image: Quantified Protein Groups         Image: All Mode: All Mode: All Mode: All Proteins       Image: Missing Value (%)       Image: Feature Correlation         Image: Protein Groups       Image: All Proteins       Image: All Proteins       Image: All Proteins				C-Service Service Serv	
Project Creation       Workflow Selection       Data Refine       DB Search       Quantification       QC         QC       Control Sample:       A01       Acceptance Tolerance (%):       10.0         Image: All Attributes       Image: Acceptance Tolerance (%):       10.0         Image: All Attributes       Image: All Attributes       Image: All Attributes         Image: Attribute 2       Image: All Attribute 2       Image: All Attribute 2         Image: All Attribute 3       Image: All Attribute 3       Image: All Attribute 3         Image: All Attribute 3       Image: All Attribute 3       Image: All Attribute 3         Image: All Attribute 3       Image: All Attribute 3       Image: All Attribute 3         Image: All MS2       Image: All Attribute 3       Image: All Attribute 3         Image: All MS2       Image: All Attribute 3       Image: All Attribute 3         Image: All MS2       Image: All Attribute 3       Image: All Attribute 3         Image: All MS2       Image: All Protein Groups 3       Image: All Attribute 3         Image: All MS2       Image: All Protein Groups 3       Image: All Proteins 3         Image: All MS2       Image: All MS2       Image: All MS2       Image: All MS2         Image: All MS2       Image: All MS2       Image: All MS2       Image: All MS2       Image:					
QC         Control Sample:       A01 <ul> <li>Acceptance Tolerance (%):</li> <li>10.0</li> </ul> V All Attributes         Data QC Attributes:       Identification QC Attributes:         V # MS1         V # MS2         V # Peptides         V # Peptides         V # Protein Groups         V # Top Proteins         V # All Proteins		ection 🔪 Data Refine 💙 DB Search 💙 Qua	ntification QC		
Control Sample:       A01       Acceptance Tolerance (%):       10.0         Image: Control Sample:       A01       Acceptance Tolerance (%):       10.0         Image: Control Sample:       Attributes       LFQ QC Attributes:       LFQ QC Attributes:         Image: Control Sample:       Attribute       Image: Control Sample:       Control Sample:       Control Sample:         Image: Control Sample:       Image: Control Sample:       Image: Control Sample:       Image: Control Sample:       Control Sample:         Image: Control Sample:       Attribute       Image: Control Sample:       Image: Control Sample:       Image: Control Sample:         Image: Control Sample:       Attribute       Image: Control Sample:       Image: Control Sample:       Image: Control Sample:         Image: Control Sample:       Image: Control Sample:       Image: Control Sample:       Image: Control Sample:       Image: Control Sample:         Image: Control Sample:       Image: Control Sample:       Image: Control Sample:       Image: Control Sample:       Image: Control Sample:       Image: Control Sample:         Image: Control Sample:       Image: Control Sample:       Image: Control Sample:       Image: Control Sample:       Image: Control Sample:       Image: Control Sample:       Image: Control Sample:       Image: Control Sample:       Image: Control Sample:       Image: Control Sample:<	Project Creation Workflow Sele				
Control Sample:       A01       Acceptance Tolerance (%):       10.0         Identification QC Attributes:       LFQ QC Attributes:         Attribute       Identification QC Attributes:       LFQ QC Attributes:         Image: Attribute       Image: Attribute       Image: Attribute         Image: Attribute       Image: Attribute       Image: Attribute       Image: Attribute         Image: Attribute	Project Creation Workflow Sele				
V All Attributes       Identification QC Attributes:       LFQ QC Attributes:         Attribute       Attribute       Attribute         V # MS1       V # MS2 Scans       Quantified Peptides         V # MS2       V # Identified Features       Quantified Protein Groups         V # Features       V # Peptides       Quantified Peatures         V # Features       V # Protein Groups       V # Sequences         V # Fourier Groups       V # Top Proteins       V # Feature Correlation         V # All Proteins       V # All Proteins       V # Feature Correlation	Project Creation Workflow Sele				
Attributes     Identification QC Attributes:     LFQ QC Attributes:       Attribute     Attribute     Attribute       # MS1     # MS2     Attribute       # MS2     # Identified Features     Quantified Peptides       # MS2     # Peptides     Quantified Protein Groups       # Features     # Peptides     Missing Value (%)       # Full Width (min)     # Top Proteins       # WHM (sec)     # All Proteins	Project Creation Workflow Sele	stance Tolerance (%): 10.0			
Attribute       Attribute         # MS1       # MS2 Scans         # MS2       # Identified Features         # MS2       # Identified Features         # Features       # Peptides         # Features       # Protein Groups         # Full Width (min)       # Top Proteins         # WHM (sec)       # All Proteins	Project Creation     Workflow Sele       QC     QC       Control Sample:     A01 •       All Attributor	otance Tolerance (%): 10.0			
Attribute       V # MS1       V # MS2       V # MS2       V # MS2       V # MS2       V # Sequences       V # Features       V # Forein Groups       V # Protein Groups       V # Attribute       V # Protein Groups       V # Top Proteins       V # Attribute       V # Attribute	Project Creation     Workflow Sele       QC     QC       Control Sample:     A01 • Acception       All Attributes     Data OC Attributes:	otance Tolerance (%): 10.0	IEO OC AN	tributes:	
V       # MS1         V       # Identified Features         V       # S2         V       MS2/MS1 Rate         V       # Peptides         V       # Sequences         V       # Protein Groups         V       # Top Proteins         V       # All Proteins	Project Creation     Workflow Sele       QC     QC       Control Sample:     A01 • Acception       All Attributes     Data QC Attributes:	otance Tolerance (%): 10.0 Identification QC Attributes:	LFQ QC At	tributes:	
V     MS2/MS1 Rate     V     # Peptides     V     Quantified Frotein Groups       V     # Features     V     # Sequences     V     Missing Value (%)       V     Full Width (min)     V     # Top Proteins     V     Feature Correlation       V     FWHM (sec)     V     # All Proteins     V     Feature Correlation	Project Creation       Workflow Sele         QC       QC         Control Sample:       A01 • Acception         All Attributes       Acception         Data QC Attributes:       Attribute         • Attributes       Attribute	otance Tolerance (%): 10.0 Identification QC Attributes: Attribute # MS2 Scans	LFQ QC At	ttributes: Attribute	
Image: Market with a sequences     Image: Market with a sequences       Image: With Microscope     Image: With Microscope       Image: With	Project Creation     Workflow Selection       QC     QC       Control Sample:     A01 • Acception       All Attributes     Acception       Data QC Attributes:     Attribute       # MS1     # MS2	otance Tolerance (%): 10.0 Identification QC Attributes:	LFQ QC A	ttributes: Attribute Intified Peptides	
Image: Protein Groups     Image: Protein Groups       Image: Protein Groups     Image: Protein Groups <t< td=""><td>Project Creation     Workflow Selection       QC     QC       Control Sample:     A01 • Acception       All Attributes     Acception       Data QC Attributes:     Attribute       # MS1     # MS2       # MS2     # MS2</td><td>Identification QC Attributes: W # MS2 Scans W # Identified Features W # Peptides</td><td>LFQ QC AI</td><td>ttributes: Attribute Intified Peptides Intified Protein Groups</td><td></td></t<>	Project Creation     Workflow Selection       QC     QC       Control Sample:     A01 • Acception       All Attributes     Acception       Data QC Attributes:     Attribute       # MS1     # MS2       # MS2     # MS2	Identification QC Attributes: W # MS2 Scans W # Identified Features W # Peptides	LFQ QC AI	ttributes: Attribute Intified Peptides Intified Protein Groups	
V     Full width (min)       V     # Top Proteins       V     # All Proteins	Project Creation       Workflow Sele         QC       QC         Control Sample:       A01 • Acception         All Attributes       Acception         Data QC Attributes:       Attribute         Image: I	Detance Tolerance (%): 10.0 Identification QC Attributes:	LFQ QC Ai	tributes: Attribute Intified Peptides Intified Protein Groups Intified Features	
V PVVPIW (sec) V # All Proteins	Project Creation       Workflow Sele         QC       QC         Control Sample:       A01 • Acception         All Attributes       Acception         Data QC Attributes:       Attribute         # MS1       # MS2         # MS2/MS1 Rate       # Features         # Features       Features	ptance Tolerance (%):       10.0         Identification QC Attributes: <b>Attribute #</b> MS2 Scans <b>#</b> Identified Features <b>#</b> Peptides <b>#</b> Sequences <b>#</b> Protein Groups	LFQ QC A Qua Qua Qua Qua Qua Qua Qua Qua	ttributes: Attribute Intified Peptides Intified Protein Groups Intified Features Sing Value (%)	
	Project Creation       Workflow Sele         QC       QC         Control Sample:       A01 • Acception         All Attributes       Acception         Data QC Attributes:       Attribute         # MS1       # MS2         # MS2/MS1 Rate       # Features         Full Width (min)       Tubut (min)	ptance Tolerance (%): 10.0 Identification QC Attributes:	LFQ QC A Qua Qua Qua Qua Qua Qua C Qua C Qua C Qua C Qua C Qua C Qua C Qua C A C A C A C A C A C A C A C A C A C	ttributes: Attribute Intified Peptides Intified Protein Groups Intified Features Sing Value (%) Iture Correlation	
	Project Creation       Workflow Sele         QC       QC         Control Sample:       A01 • Acception         All Attributes       Attributes         Data QC Attributes:       Attribute         * MS1       * MS1         * # MS2       MS2/MS1 Rate         * # Features       Full Width (min)         Full Width (min)       FWHM (sec)         BPC       Features	ptance Tolerance (%):       10.0         Identification QC Attributes: <b>Attribute #</b> MS2 Scans <b>#</b> Identified Features <b>#</b> Identified Features <b>#</b> Peptides <b>#</b> Sequences <b>#</b> Top Proteins <b>#</b> All Proteins <b>#</b> All Proteins <b>#</b> ID Rate	LFQ QC A Qua Qua Qua Qua Qua Qua Qua C Qua	ttributes: Attribute Intified Peptides Intified Protein Groups Intified Features sing Value (%) ture Correlation	

Fig 1. **Establishing protein/peptide LFQ and QC in PEAKS Studio 11** (a) The LFQ setup and parameters in PEAKS Q workflow. Top 3 peptides were selected to calculated protein quantification and auto (TIC) normalization was applied. (b) QC Setup: The QC process was configured by examining all QC attributes. The control sample, A01, was designated for reference, and a tolerance threshold of 10% was set to monitor variations within the QC samples.

MS1 Error Mean (S.D)
 MS2 Error Mean (S.D)
 Missed Cleavage Ratio
 Enzyme Specificity Ratio

# QC result of LC-MS data:

In the QC Summary tab, the QC statistics of the LFQ results were categorized into three groups: control sample, passed samples, and failed samples. Samples were classified as failed if any of their attributes fell outside the tolerance level, while samples with all attributes within +/- 10% of the control sample (A01) were considered as passed samples. A concise overview of the QC findings is presented in Figure 3(a). Notably, among the 29 samples in the dataset, only 10 samples exhibited all QC attributes falling within the 10% tolerance range.

Figure 3(b) displays the distribution of the failed attributes in this dataset, with "TIC correlation" and "Feature correlation" being the two most prevalent. These two attributes collectively imply potential concerns regarding the dataset's reproducibility. Further elaboration and discussion on these specific QC attributes will be provided in subsequent sections.



Fig 2. **LFQ result.** (a) statistics of quantification result (b) Volcano plot illustrating differentially expressed proteins. Proteins with downregulation are indicated in green, upregulated proteins in red, and proteins that do not pass the filter criteria (i.e., containing only one quantifiable peptide) are depicted in grey.





#### Mass error in QC:

The mass error emerges as a common factor contributing to imperfections in MS datasets, especially when dealing with large cohorts of samples. Environmental temperature fluctuations, MS calibrations, and power surges are potential factors responsible for the occurrence of MS errors. Thus, it is imperative to initially investigate whether the QC attribute failures can be attributed to MS errors. The PEAKS QC module offers multiple avenues for researchers to assess MS errors. The distribution of MS errors was visualized across the categories of control, passed, and failed samples (Figure 4). The violin chart analysis revealed no significant differences between passed and failed samples in both MS1 and MS2 levels. This observation suggests that mass error may not be the primary cause for the samples failing to meet the QC criteria.



Fig 4. Sample error distribution. (a) MS1 Mass Error distribution (b) MS2 Mass Error distribution.

In addition to the collective comparison between passed and failed samples, individual samples' mass errors can also be independently compared with the control sample. Figure 5 demonstrates this comparison, where the sample with the highest MS1 error (D03) was selected for evaluation. It is evident from the analysis that sample D03 remains well within the acceptable mass error range of +5/-5 ppm when compared to the control sample. This further reinforces the notion that the mass error is not a major contributing factor to the failure of samples to meet the QC criteria in the dataset.



Fig 5. Individual Sample error distribution. (a) Bar chart of MS1 error mean of all samples. D03, marked by red rectangles, has the highest MS1 error mean (b) MS1 error distribution comparison between control sample A01 and D03.

# **TIC Correlation:**

Figure 3(b) highlights the most frequently failed QC attribute as TIC correlation. This attribute serves as a measure of similarity between the selected sample and the control sample. In a set of technical replicates, the TIC chromatograms should exhibit high similarity, leading to relatively high TIC correlation values. However, in this study, the TIC correlation is not ideal, even within the same technical replicate datasets (Figure 6).

Figure 7 presents the TIC chromatograms of two samples from technical replicate group A, overlaid with the control sample. In Figure 7(a), the TIC chromatograms of the two samples are nearly identical, indicating high similarity. However, in Figure 7(b), there is a small but discernible difference between the two chromatograms, suggesting reduced similarity.

The evaluation of chromatogram similarity provides a quick and effective method to assess sample reproducibility. In PEAKS 11 QC module, multiple tools are offered to check the identification and quantification reproducibility across various sample aspects. Figure 8 illustrates Venn diagrams depicting the peptide and protein identifications between samples A03 & A01 and A05 & A01. Even between technical replicates (A01 and A03) with highly similar TIC chromatograms, the exact reproducibility is slightly below 80% (common identified proteins/peptides divided by the total): 78% for proteins and 61% for peptides. This observation highlights the stochastic nature of DDA experiments where the precursor selection is semi-random.





Fig 6. TIC correlation of each sample comparing to control sample A01. The technical replication of group A was marked with red rectangle.

Fig 7. TIC chromatogram overlay between control sample and selected sample (a) A03 vs A01, TIC correlation 0.95 (b) A05 vs A01, TIC correlation 0.86.

The identification reproducibility is somewhat lower between A01 and A05, with 76% for proteins and 54% for peptides. This decrease in reproducibility is understandable since the chromatograms are less identical between sample A01 and A05.



Fig 8. **Venn diagram of sample identification.** (a)(b), number of unique or common identified proteins or peptides in A05 and A01. (c)(d), number of unique or common identified proteins or peptides in A05 and A01.

The evaluation of chromatogram similarity provides a quick and effective method to assess sample reproducibility. In PEAKS 11 QC module, multiple tools are offered to check the identification and quantification reproducibility across various sample aspects. Figure 8 illustrates Venn diagrams depicting the peptide and protein identifications between samples A03 & A01 and A05 & A01. Even between technical replicates (A01 and A03) with highly similar TIC chromatograms, the exact reproducibility is slightly below 80% (common identified proteins/peptides divided by the total): 78% for proteins and 61% for peptides. This observation highlights the stochastic nature of DDA experiments where the precursor selection is semi-random.

In DDA experiments, it is common to observe suboptimal identification reproducibility, even in technical replicates, due to semi-random precursor selection. However, in LFQ experiments, where MS1 ions are used for quantification and ID-transfer is applied, the missing value issues are significantly reduced, and quantification reproducibility should be higher between technical replicates.

Figure 9 illustrates the quantification correlation across three different levels: proteins, peptides, and features. Both A03 and A05 demonstrate excellent linear correlation, exceeding 0.95, when compared to the control sample A01. The quantification correlation in A05 is only marginally lower, approximately 0.01, compared to A03. This finding implies that despite sample A05 not passing the QC attribute check in TIC correlation, its LFQ reproducibility remains largely unaffected.

Notably, the quantification correlations within technical replicates A exhibit high values above 0.95 (data not shown). This observation suggests that our current TIC correlation tolerance level might be overly stringent.



Fig 9. Pearson correlation of quantification results. a-c, proteins, peptides and features correlation between A03 and A01; d-f, proteins, peptides and features correlation between A05 and A01

# **LC Conditions:**

While the overall sample reproducibility remains largely unaffected, it remains of interest to investigate the factors contributing to the observed decrease in TIC chromatogram correlation. One potential reason for the changes in TIC chromatogram correlation could be variations in the LC conditions.

Fortunately, PEAKS 11 QC module offers multiple tools to assess the column performance, gradient suitability, and other relevant factors. In Figure 10(a), it is demonstrated that the full width at half maximum (FWHM) across the entire sample set remains consistent, indicating that the column and LC conditions throughout the run exhibit no significant shifts. Furthermore, the comparison of MS2 count over time and LFQ retention time (RT) alignment between A01 and A05 also suggests that no major LC condition changes occur even between two samples exhibiting TIC chromatogram differences.

These findings indicate that factors other than major LC condition changes may be contributing to the observed variations in TIC chromatogram correlation. Further investigation and analysis are required to identify and address these factors effectively.

# **MS Sensitvity:**

The observed variations in TIC chromatogram correlation may also be influenced by changes in MS sensitivity during sample acquisition. Accumulation of neutral particles on the front end of the MS can lead to decreased instrument sensitivity, affecting the TIC chromatogram and resulting in low TIC correlation.

Figure 11 provides intriguing insights, revealing a trend with the base peak chromatogram (BPC) and the number of identified peptides across different samples within a technical replicate group. The decreasing trend in both BPC and peptide identification numbers suggests a potential decline in MS sensitivity.



Fig 10. Quality control of LC condition changes. (a) FWHM of all samples (b) MS2 over retention time comparison between A01 and A05 (c) LFQ RT alignment between A01 and A05.





To explore the underlying reasons for this trend, we regrouped the samples based on their acquisition time (Figure 12(b)). By plotting the BPC of each sample and the average BPC of different time groups (Figure 12(a)), a clear trend emerges. The average BPC of time groups indicates a gradual decrease in MS sensitivity over time, explaining the observed differences in TIC chromatograms within the technical replicate groups.

#### **Feature Correlation:**

The feature correlation emerges as another crucial QC attribute that falls short of meeting the threshold, and this phenomenon is exclusively observed in samples belonging to group E (Figure 13(a)). This observation indicates the existence of inherent differences between the technical groups, particularly related to the variation in sample amounts between yeast and E. coli (as indicated in the experimental design).

To gain further insights into these population differences, Figure 13(b) reveals three distinct populations representing higher E. coli protein levels and lower yeast protein levels in sample E05. This observation is consistent with the protein density ratio depicted in Figure 13(c). The observed discrepancy in protein levels between yeast and E. coli in sample E05 aligns with the variations observed in the feature correlation, potentially contributing to the failure of this QC attribute in the group E samples.



Fig 12. (a) BPC of samples grouped by acquisition time. The red rectangle highlights the average BPC of different groups. (b) the list of data acquisition time of all samples.



Fig 13. (a) All samples' feature correlation to control sample (b) The feature correlation distribution between E05 and A01 (c) Protein density ratio between group A and group E

# **Conclusion:**

This study employed PEAKS 11 to conduct Label-Free Quantification (LFQ) and Quality Control (QC) analyses on a diverse dataset. The investigation focused on changes in TIC chromatogram and revealed that it was primarily caused by a drop in MS sensitivity. While the sample reproducibility was not significantly affected in this study, it highlights the importance of consistently monitoring machine status, including factors like MS error, MS sensitivity, and LC conditions. The QC analysis in PEAKS 11 proved invaluable in identifying and understanding potential issues influencing sample reproducibility and data quality. By addressing these factors, we can draw robust and meaningful conclusions from our dataset, furthering our understanding of the underlying biological processes under investigation.

#### References:

 Overmyer, K. A. et al. Large-scale multi-omic analysis of COVID-19 severity. Cell Syst. 12, 23–40 (2021).
 Wang, X. et al. Ultra-High-Resolution IonStar Strategy Enhancing Accuracy and Precision of MS1-Based Proteomics and an Extensive Comparison with State-of-the-Art SWATH-MS in Large-Cohort Quantification. Analytical Chemistry. 11, 4884-4893 (2021).

#### **Bioinformatics Solutions, Inc.**

140 Columbia St, Suite 202 Waterloo, Ontario N2L 3K8 Canada

Tel: (519) 885-8288 Fax: (519) 885-9075

sales@bioinfor.com www.bioinfor.com



Information, descriptions, and specifications in this publication are subject to change without notice. Bioinformatics Solutions, Inc. 2022