Journal of proteome research

Article

pubs.acs.org/jpr

Analysis of Reproducibility of Proteome Coverage and Quantitation Using Isobaric Mass Tags (iTRAQ and TMT)

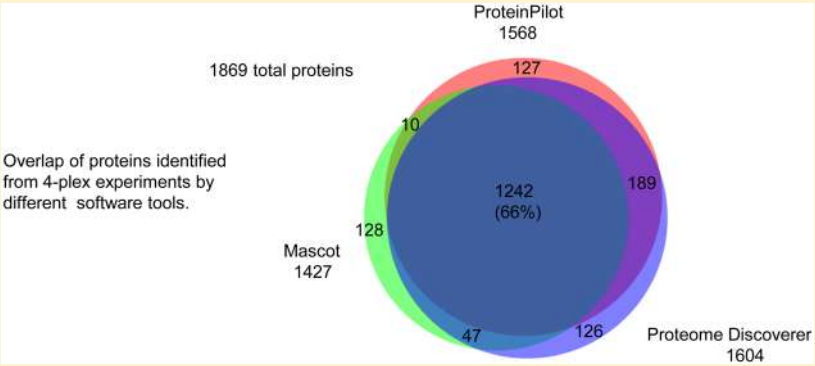
Tammy M. Casey,[†,‡] Javed M. Khan,[†,‡] Scott D. Bringans,[*,†] Tomas Koudelka,[†] Pari S. Takle,[†] Rachael A. Downs,[†] Andreja Livk,[†] Robert A. Syme,[§] Kar-Chun Tan,[§] and Richard J. Lipscombe[†]

[†]Proteomics International, Perth, Western Australia 6009, Australia

[‡]Harry Perkins Institute of Medical Research, QEII Medical Centre, The University of Western Australia, Nedlands, Western Australia 6009, Australia

[§]Centre for Crop and Disease Management, Department of Environment and Agriculture, Curtin University, Bentley, Western Australia 6102, Australia

Supporting Information



ABSTRACT: This study aimed to compare the depth and reproducibility of total proteome and differentially expressed protein coverage in technical duplicates and triplicates using iTRAQ 4-plex, iTRAQ 8-plex, and TMT 6-plex reagents. The analysis was undertaken because comprehensive comparisons of isobaric mass tag reproducibility have not been widely reported in the literature. The highest number of proteins was identified with 4-plex, followed by 8-plex and then 6-plex reagents. Quantitative analyses revealed that more differentially expressed proteins were identified with 4-plex reagents than 8-plex reagents and 6-plex reagents. Replicate reproducibility was determined to be ≥69% for technical duplicates and ≥57% for technical triplicates. The results indicate that running an 8-plex or 6-plex experiment instead of a 4-plex experiment resulted in 26 or 39% fewer protein identifications, respectively. When 4-plex spectra were searched with three software tools—ProteinPilot, Mascot, and Proteome Discoverer—the highest number of protein identifications were obtained with Mascot. The analysis of negative controls demonstrated the importance of running experiments as replicates. Overall, this study demonstrates the advantages of using iTRAQ 4-plex reagents over iTRAQ 8-plex and TMT 6-plex reagents, provides estimates of technical duplicate and triplicate reproducibility, and emphasizes the value of running replicate samples.

KEYWORDS: *iTRAQ, TMT, reproducibility*

INTRODUCTION

Quantitative mass spectrometry-based proteomic methods enable the measurement of protein abundance changes across multiple proteomes. Isobaric tags for relative and absolute quantification (iTRAQ) has become the technique of choice in the biomarker discovery field.[1] The quantitative isobaric labeling method was first described by Ross et al.[2] and allows up to eight samples to be measured simultaneously.[3] Tandem mass tagging is another isobaric labeling technique, which follows the same principle as iTRAQ quantitation; however, the labels are slightly different with respect to structure and mass.[4] Despite the widespread use of iTRAQ and tandem mass tagging (the terms "iTRAQ" and "tandem mass tag" appeared in 723 PUBMED title/abstract searches in 2014 and 2015), values for protein identification reproducibility and differential expression reproducibility between replicates have not been widely reported in the literature.

In this study, the intracellular proteomes of *P. nodorum* SN15 and *gna1* were used as a base to compare iTRAQ 4-plex, iTRAQ 8-plex, and TMT 6plex reagents with respect to the depth and reproducibility of proteome coverage and differential protein expression. The fungus *Parastagonospora* (syn. ana, *Stagonospora*; teleo, *Phaeosphaeria*) *nodorum* (Berk.)[5] is the

Received: December 23, 2015
Published: December 7, 2016

DOI: 10.1021/acs.jproteome.5b01154
J. Proteome Res. 2017, 16, 384−392

causal agent of septoria nodorum blotch (SNB) on wheat.[5,6] Heterotrime G-protein signaling is required by *P. nodorum* to successfully carry out the infection cycle on wheat. Deletion of genes associated with this pathway results in mutants carrying severe developmental impairments that account for a reduction in pathogenicity fitness.[7,8] The genome of *P. nodorum* is well-characterized.[9−11] This has enabled the use of proteomics to dissect the role of heterotrimeric G-protein signaling in *P. nodorum* by comparing subproteomes of the SN15 wildtype with a mutant lacking the putative Gα subunit gene (*Gna1*).[12−14] Several *Gna1*-regulated genes were identified, and their involvement in sporulation,[12,13] pycnidial development,[12] cell wall/protein degradation,[7,14] and mycotoxin production[15] was characterized. Therefore, *P. nodorum* can be considered an ideal model fungal necrotroph for the development of proteomic tools.

Spectra were searched with three software tools—ProteinPilot, Mascot, and Proteome Discoverer—to compare the overlap between the algorithms in terms of proteome coverage and differentially expressed proteins.

MATERIALS AND METHODS

Growth and Maintenance of *P. nodorum*

P. nodorum SN15 wild-type (Department of Agriculture, Western Australia) and the *gna1-35* strain carrying a disrupted *Gna1* (Genbank accession number: EAT82421) were grown and maintained as previously described.[13]

Experimental Design

Technical replicates for iTRAQ 4-plex experiments were generated according to the scheme depicted in Figure 1. In brief, following labeling the combined desalted sample was split into two aliquots (aliquots 1 and 2), which were subsequently individually separated by SCX and then desalted. Aliquot 1 was split into a further two aliquots before all three samples were analyzed by mass spectrometry. The process was repeated for iTRAQ 8-plex and TMT 6-plex experiments.

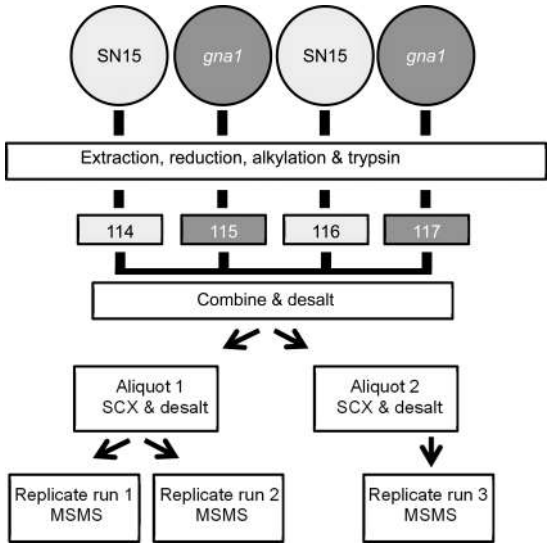


Figure 1. Schematic diagram showing the workflow used in the iTRAQ 4-plex experiment, whereby three technical replicates were measured by mass spectrometry. The same workflow was used in the iTRAQ 8-plex and TMT 6-plex experiments.

Sample Preparation and iTRAQ Labeling

Soluble intracellular proteins were extracted from freeze-dried mycelia, as previously described,[13] and protein concentration was determined using the Pierce BCA protein assay kit. Proteins (2 mg) from SN15 and *gna1* strains were individually precipitated by adding five volumes of acetone, incubating for 1 h at −20 °C, and pulse centrifuging for 10 s. The protein pellets were resuspended in 0.5 M triethylammonium bicarbonate (pH 8.5) and 0.1% SDS before digestion with 13 μg trypsin at 37 °C for 3 h. Samples were reduced with TCEP and alkylated with methyl methanethiosulfonate (MMTS) according to the manufacturer's instructions (AB Sciex, Framingham, MA) before undergoing a second digestion with 25 μg trypsin at 37 °C for 24 h. Tryptic digestion occurred at the expected termini over 93% of the time, as calculated by the Proteomics System Performance Evaluation Pipeline (PSPEP)[16] algorithm that is embedded in the software. Samples were centrifuged at 13 000*g* for 10 min at room temperature before the supernatant was removed and assayed for peptide concentration (Direct Detect Spectrometer, Merck Millipore, Billerica, MA). The following aliquots were removed for labeling: for iTRAQ 4plex experiments −2 × 100 μg SN15 and 2 × 100 μg *gna1* aliquots; for iTRAQ 8plex experiments −4 × 50 μg SN15 and 4 × 50 μg *gna1* aliquots; and for TMT 6plex experiments −3 × 70 μg SN15 and 3 × 70 μg *gna1* (Table 1). Peptides were labeled according to the manufacturer's instructions (iTRAQ: AB Sciex; TMT: Thermo Fisher Scientific, Rockford, IL) before samples were combined and desalted on Strata-X 33 mm polymeric reverse phase columns (Phenomenex, Torrance, CA). The efficiencies of iTRAQ and TMT labeling of N-terminus and lysine side chains of peptides were calculated by the PSPEP algorithm as being over 91%, 98%, and 93% for 4-plex, 8-plex, and 6-plex experiments, respectively.

Strong Cation-Exchange Chromatography (SCX)

Peptides were separated by SCX on an Agilent 1100 HPLC system (Agilent Technologies, Palo Alto, CA) using a PolySulfethyl column (4.6 × 100 mm, 5 μm, 300 Å, Nest Group, Southborough, MA). Peptides were eluted with a linear gradient of Buffer B (1 M KCl, 10% ACN and 10 mM KH_2PO_4, pH 3; flow rate 0.5 mL/min) and were collected as one fraction per minute. A total of 40 subfractions were collected and pooled into eight fractions, desalted on Strata-X 33 mm polymeric reverse-phase columns, and dried.

5600 TripleTOF

Fractions were desalted on Strata-X columns and loaded onto a Agilent Zorbax 300SB-C18, 3.5 μm apparatus (Agilent Technologies) running on an Shimadzu Prominence nano HPLC system (Shimadzu, Kyoto, Japan). Peptides were resolved with a gradient of 10−40% acetonitrile (0.1% formic acid) at 300 nL/min over 160 min and eluted through a nanospray interface into a 5600 TripleTOF mass spectrometer (AB Sciex).

The data were acquired in an information-dependent acquisition (IDA) mode with Analyst TF 1.6 software (AB Sciex). The MS settings were as follows: Ionspray Voltage Floating (ISVF) = 2300 V, curtain gas (CUR) = 20, ion source gas 1 (GS1) = 20, interface heater temperature (IHT) = 150, and declustering potential (DP) = 70 V. The first TOF MS scan (experiment 1) was performed in the mass range of 400−1250 Da with a 0.25 s TOF MS accumulation time, whereas the MS/MS product ion scan was performed in the mass range of 100−1800 Da with a 0.1 s accumulation time. The criteria for product ion fragmentation was set as follows: ions (>400 m/z and <1250 m/z)

DOI: 10.1021/acs.jproteome.5b01154
J. Proteome Res. 2017, 16, 384−392

Table 1. Labelling Scheme for *P. nodorum* SN15 and *gna1*

sample	iTRAQ 4-plex reagent (100 μg/sample)	iTRAQ 8-plex reagent (50 μg/sample)	TMT 6-plex reagent (70 μg/sample)
gna1		113	
SN15	114	114	126
gna1	115	115	127
SN15	116	116	128
gna1	117	117	129
SN15		118	130
gna1		119	131
SN15		121	

with charge states of 2 to 5 and an abundance threshold of >250 cps. Former target ions were excluded for 10 s after one occurrence. The maximum number of candidate ions per cycle was 20 spectra. IDA advanced "rolling collision energy (CE)" and "Adjust CE When Using iTRAQ Reagent" were applied for subsequent MS and MS/MS scans.

Data Analysis

Because one of the aims of this study was to compare the performance of different software tools, spectra were analyzed with three different algorithms: ProteinPilot, Mascot, and Proteome Discoverer.

ProteinPilot

Protein identification and quantification were performed using ProteinPilot 4.5 Beta Software (July 2012; AB Sciex). MS/MS spectra were searched against an updated and manually curated version of the *P. nodorum* SN15 protein data set described by Bringans et al.[10] The data set consisted of 15 812 predicted proteins. Search parameters were: Sample type: iTRAQ 4plex (peptide labeled) or 8-plex (peptide labeled) or TMT 6-plex ID-only (peptide labeled); Cys alkylation: MMTS; Digestion: Trypsin; Instrument: TripleTOF 5600; Special factors: None; Species: None; Quantitate tab checked; Bias correction and Background correction tabs checked (for 4-plex and 8-plex analyses); ID focus: Biological modifications; Search effort: Thorough; Detected protein threshold [Unused ProtScore (Conf)]: 0.05 (10.0%); FDR Analysis tab checked.

All identified proteins had an Unused Protscore of >1.3 (which corresponds to proteins identified with >95% confidence), as calculated by the software and a global false discovery rate (FDR) of ≤1% determined at the protein level by the PSPEP algorithm. To be considered as being differentially expressed, proteins were required to have a p value of ≤0.05, as calculated by the software.

Mascot

Spectral data were also searched using the Mascot search engine (version 2.2.4 Matrix Science, Boston, MA) against the *P. nodorum* SN15 protein data set. Search parameters were: Enzyme: Trypsin; Max missed cleavages: 1; Fixed modifications: iTRAQ4plex (K) or iTRAQ8plex (K) or TMT6plex (K), iTRAQ4plex (N-term) or iTRAQ8plex (N-term) or TMT6plex (N-term), Methylthio (C); Variable modifications: iTRAQ4-plex (Y) or iTRAQ8plex (Y), Oxidation (M); Peptide charge: 2^+, 3^+, and 4^+; Peptide tolerance: ± 0.4 Da; MS/MS tolerance: ±0.4 Da; Quantitation: iTRAQ 4plex or iTRAQ 8plex or TMT 6plex with weighted protein ratios and median normalization of ratios; Instrument: ESI-QUAD-TOF. The MOWSE algorithm (MudPIT scoring) of Mascot was used to score the significance of peptide/protein matches with $p < 0.05$ for each protein identification.

All identified proteins had a peptide level FDR of ≤1%. Identified proteins were required to have a p value of ≤0.05, as calculated by the software, to be considered as being differentially expressed.

Proteome Discoverer

Data analyses were also performed using Proteome Discoverer Version 1.4.0.288 (Thermo Fisher Scientific, Waltham, MA). The spectral data were again searched against the *P. nodorum* SN15 protein data set. The analysis workflow used included five nodes, namely, Spectrum Files (data input), Spectrum Selector (spectrum and feature retrieval), Mascot (sequence database search), Percolator (peptide spectral match or PSM Validation and FDR analysis), and Reporter Ions Quantifier (quantification). ESI-QUAD-TOF (5600 TripleTOF) relevant settings were used for all nodes. iTRAQ 4-plex/iTRAQ 8-plex/TMT 6-plex options were selected for respective analyses in all nodes. Because Proteome Discoverer Version 1.4 relies on an in-house sequence database search program such as SEQUEST or Mascot (in this case Mascot was used), the settings for the Mascot node were edited to mimic those described above (in the Mascot data analysis section) for independent Mascot searches to ensure consistency. Other settings included default parameters for Spectrum Selector (except for Max Precursor Mass value which was set at 5000), Percolator, and Reporter Ions Quantifier nodes.

All identified proteins had a FDR of ≤1%, which was calculated at the peptide level. To be considered as being differentially expressed, proteins were required to have a p value of ≤0.05. Quantitative analyses were performed as described below. Results from Proteome Discoverer were exported. Unique peptides and their quantification values (115/114, 117/116, etc.) were appended to their corresponding protein accession number. A t test was performed on each quantification ratio using a two-tailed one sample t test using the SciPy (www.scipy.org) statistics module in Python. Data were considered statistically different when $p < 0.05$.

Venn Diagrams

The Venny online Venn diagram generator (http://bioinfogp.cnb.csic.es/tools/venny/index.html) was used to create four-way Venn diagrams.[17] Area proportional Venn diagrams were generated using the BioVenn web tool (http://www.cmbi.ru.nl/cdd/biovenn/).[18]

RESULTS AND DISCUSSION

The purpose of this study was to compare the depth and reproducibility of total proteome coverage and differentially expressed protein coverage in technical duplicates and triplicates using isobaric mass tags. There are many aspects of the data that have been examined, and each has been described below. These include the depth of coverage of the total proteome and

DOI: 10.1021/acs.jproteome.5b01154
J. Proteome Res. 2017, 16, 384−392

Table 2. Unique Peptide (A) and Protein (B) Identification Results from iTRAQ 4-plex, iTRAQ 8-plex, and TMT 6-plex Experiments[a]

A

	iTRAQ 4-plex				iTRAQ 8-plex				TMT 6-plex			
	run 1	run 2	run 3	mean ± SEM	run 1	run 2	run 3	mean ± SEM	run 1	run 2	run 3	mean ± SEM
ProteinPilot	9885	8107	6565	8186 ± 959	4183	4603	5513	4766 ± 393	4726	6606	4451	5261 ± 677
Mascot	12707	8167	6842	9239 ± 1776	6020	5697	7314	6344 ± 494	6093	5532	4602	5409 ± 435
Proteome Discoverer	7511	6067	4861	6146 ± 766	3648	4234	5204	4362 ± 454	3152	4115	3181	3483 ± 316

B

	iTRAQ 4-plex				iTRAQ 8-plex				TMT 6-plex			
	run 1	run 2	run 3	mean ± SEM	run 1	run 2	run 3	mean ± SEM	run 1	run 2	run 3	mean ± SEM
ProteinPilot	1388	1223	1084	1232 ± 88	842	992	1085	973 ± 71	741	928	759	809 ± 60
Mascot	1217	1031	923	1057 ± 86	790	859	963	870 ± 50	709	802	686	732 ± 35
Proteome Discoverer	1416	1229	1097	1247 ± 92	824	955	1043	941 ± 64	710	866	724	767 ± 50

[a]Spectra were searched with either ProteinPilot, Mascot, or Proteome Discoverer.

differentially expressed proteins, overall reproducibility, and the effect of replicates on reproducibility and the result of searching data with different search programs. The percentage proteome coverage for 4-plex experiments analyzed by ProteinPilot was calculated as being 9.9% (1568 protein identified of a possible 15 812).

Depth of Total Proteome Coverage and Differentially Expressed Protein Coverage

Total proteome depth was first examined by analyzing mass spectra from iTRAQ 4-plex, iTRAQ 8-plex, and TMT 6-plex technical triplicates using ProteinPilot, Mascot, and Proteome Discoverer. All of the proteins identified are listed in Supporting Information Tables S1−S9.

When the numbers of unique peptides were compared across all experiments, the results show that a higher number of unique peptides were identified with 4-plex reagents than 8-plex and 6-plex reagents (Table 2A). Next, the numbers of identified proteins were compared across experiments, and the highest number of proteins was found to be identified with 4-plex, followed by 8-plex and then 6-plex reagents (Table 2B). This result was consistent for all software.

The higher number of protein identifications with 4-plex reagents is in agreement with a study by Pichler et al.,[19] who showed that more proteins were identified with 4-plex than 6-plex and 8-plex reagents. In contrast, Li et al.[20] demonstrated that 4-plex and TMT-duplex reagents yielded similar numbers of protein identifications, and Pottiez et al.[21] showed similar numbers of proteins were identified with 4-plex and 8-plex reagents. Pottiez et al.[21] suggested that the difference between their results and those of Pichler et al.[19] could be due to the use of different instruments (LTQ Orbitrap versus 4800 MALDI-TOF/TOF) or different search algorithms (Mascot and Proteome Discoverer versus ProteinPilot 4.0). The results presented here show that for data obtained on a 5600 TripleTOF the search algorithm had no bearing on whether more proteins were identified with 4-plex or 8-plex reagents because a higher number of proteins was identified with 4-plex reagents for all software. Both Pichler et al.[19] and Sandberg et al.[22] have shown that 6-plex reagents yielded more protein identifications than 8-plex reagents, and these findings are in contrast with the results presented here.

For quantitative analyses, ProteinPilot was used to analyze mass spectra from only 4-plex and 8-plex replicates. TMT 6-plex analyses were not possible with ProteinPilot because the software did not recognize TMT reagents. Mascot and Proteome Discoverer were used to analyze spectra from 4-plex, 8-plex, and 6-plex replicates. All of the identified differentially expressed proteins are listed in Supporting Information Tables S1−S9.

The results demonstrate that when spectra were analyzed with ProteinPilot, more differentially expressed proteins were identified with 4-plex reagents than 8-plex reagents (Table 3). For spectra analyzed with Mascot and Proteome Discoverer, a similar outcome was obtained with a higher number of differentially expressed proteins identified with 4-plex reagents than 8-plex and 6-plex reagents.

Reproducibility of Identified Proteins and Differentially Expressed Proteins

The reproducibility of protein identification between technical duplicates was determined, and it was found that 70% (4-plex), 72% (8-plex) and 69% (6-plex) of proteins were identified in both of the duplicate runs (Supporting Information, Figure S1), while for technical triplicates it was found that 58% (4-plex), 61% (8-plex), and 57% (6-plex) of proteins were identified in all three of the triplicate runs (Supporting Information, Figure S2). These results show that the reproducibility of duplicate runs was ∼12% higher than that of triplicate runs.

In accordance with the results presented here, Li et al.[20] and Besson et al.[23] previously demonstrated that 73 and 87% of proteins, respectively, were identified reproducibly between duplicate 4-plex runs. Protein identification reproducibility in *P. nodorum* biological replicates has been previously measured, and it was found that 44% of proteins were identified in all three biological replicates.[13] Others have demonstrated similar findings, with Song et al.[24] using plasma samples, Ji et al.[25] using C3H10T1/2 cells, and Wang et al.[26] using *Chlamydomonas reinhardtii* showing that 43, 46, and 61% of proteins, respectively, were identified in all three biological replicates.

Next, the reproducibility of differential protein expression between technical duplicates was assessed, and it was found that 38% (4-plex), 35% (8-plex), and 28% (6-plex) of differentially expressed proteins were identified in both of the duplicate runs (Supporting Information, Figure S3). For technical triplicates, it was found that 24% (4-plex), 21% (8-plex), and 15% (6-plex) of differentially expressed proteins were identified in all three triplicate runs (Supporting Information, Figure S4). These results demonstrate that the reproducibility of duplicate runs was ∼14% higher than that of triplicate runs. Importantly, the

DOI: 10.1021/acs.jproteome.5b01154
J. Proteome Res. 2017, 16, 384−392

Table 3. Differential Protein Expression Results from iTRAQ 4-plex, iTRAQ 8-plex, and TMT 6-plex Experiments[a,b]

		iTRAQ 4-plex					iTRAQ 8-plex					TMT 6-plex				
	ratios	run 1	run 2	run 3	mean ± sem	ratios	run 1	run 2	run 3	mean+ SEM	ratios	run 1	run 2	run 3	mean ± SEM
Protein-Pilot	115:114	242 (↑99, ↓143)	202 (↑74, ↓1128)	172 (↑60, ↓112)	205 ± 20	115:114	101 (↑48, ↓53)	158 (↑67, ↓91)	176 (↑82, ↓94)	145 ± 23	127:126	*	*	*	
	117:116	250 (↑ 126/↓124)	199 (↑92, ↓107)	166 (↑72, ↓94)	205 ± 24	117:116	105 (↑50, ↓55)	136 (↑58, ↓78)	168 (↑71, ↓97)	136 ± 18	129:128	*	*	*	
Mascot	115:114	325 (↑129/↓196)	239 (↑76/↓163)	203 (↑64/↓139)	255 ± 36	115:114	141 (↑52, ↓89)	174 (↑71, ↓103)	209 (↑711, ↓138)	174 ± 19	127:126	199 (↑14, ↓185)	261 (↑21, ↓240)	220 (↑13, ↓207)	229 ± 21
	117:116	280 (↑127/↓153)	243 (↑111/↓132)	200 (↑79/↓121)	241 ± 23	117:116	128 (↑47, ↓81)	160 (↑53, ↓107)	192 (↑65, ↓127)	160 ± 18	129:128	205 (↑14, ↓191)	274 (↑19, ↓255)	229 (↑10, ↓219)	242 ± 24
PD	115:114	226 (↑116/↓110)	165 (↑75/↓90)	153 (↑82/↓71)	181 ± 123	115:114	94 (↑52/↓42)	123 (↑57/↓66)	142 (↑61/↓81)	120 ± 14	127:126	109 (↑63, ↓46)	118 (↑66, ↓52)	95 (↑56, ↓39)	107 ± 7
	117:116	215 (↑120/↓95)	185 (↑95/↓90)	177 (↑89/↓88)	192 ± 12	117:116	82 (↑43/↓39)	122 (↑48/↓74)	140 (↑71/↓69)	115 ± 17	129:128	101 (↑163, ↓38)	138 (↑60, ↓76)	116 (↑58, ↓58)	118 ± 10

[a]Spectra were searched with either ProteinPilot, Mascot or Proteome Discoverer (PD). [b]↑ indicates the number of proteins up-regulated. ↓ indicates the number of protein down-regulated. * indicates that it was not possible to analyze TMT 6-plex data with ProteinPilot.

results show that protein identification reproducibility was higher than differentially expressed protein reproducibility for all sets of replicates; 4-plex: 70% vs 38% for duplicates and 58% vs 24% for triplicates, 8-plex; 72% vs 35% for duplicates and 61% vs 21% for triplicates; 6-plex: 69% vs 28% for duplicates and 57% vs 15% for triplicates.

The reproducibility of differential protein expression in iTRAQ and TMT replicates has been indirectly measured before. Using linear regression analyses, Besson et al.[23] and Ji et al.[25] reported good analytical reproducibility in iTRAQ experiments and Wang et al.[26] suggested that quantitative TMT data from replicates were quite reproducible. The results in the current study provide a more detailed evaluation of reproducibility by reporting the percentage of differentially expressed proteins in replicates.

Effect of the Workflow on Reproducibility

The workflow shown in Figure 1 demonstrates that two of the replicates used in this study correspond to repeated LC−MS/MS analyses from the same aliquot (run 1 and run 2), while the third replicate (run 3) was prepared from a separate aliquot. This strategy allowed comparison of variability from MS runs and SCX fractionation. The results show that there was no difference in either protein identification reproducibility or differential protein reproducibility between the first two replicates (run 1 vs run 2) and either run 1 vs run 3 or run 2 vs run 3, which indicates that the workflow did not introduce variation (Supporting Information, Figures S1 and S3).

Cumulative Effect of 4-plex and 8-plex Experiments and 4-plex and 6-plex Experiments

A total of 1663 proteins were identified when results from all 4-plex triplicates were combined and compared with all 8-plex

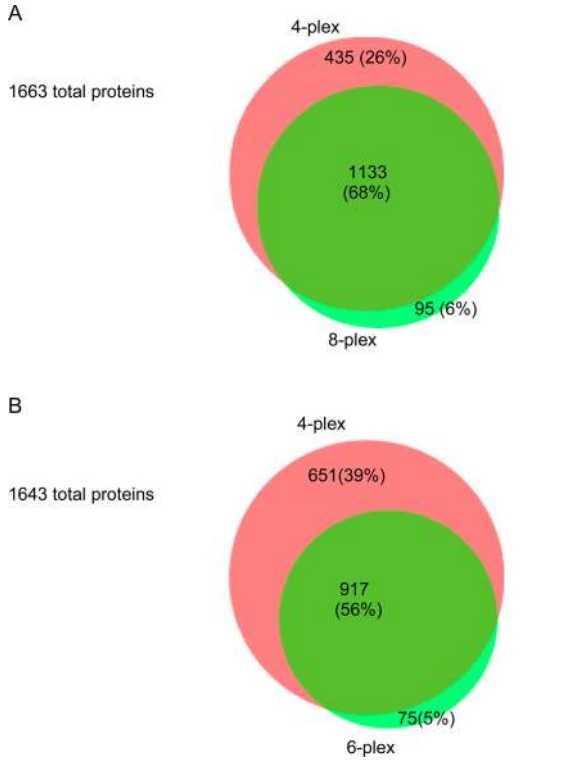


Figure 2. Overlap of proteins identified from (A) 4-plex experiment and 8-plex experiments and (B) 4-plex and 6-plex experiments. 4-plex and 8-plex spectra were searched with ProteinPilot, and 6-plex spectra were searched with Proteome Discoverer. Results were generated from all three replicates.

DOI: 10.1021/acs.jproteome.5b01154
J. Proteome Res. 2017, 16, 384−392

triplicates (Figure 2). Overall, 68% of proteins were common to both experiments, while 26% were only identified in 4-plex triplicates and 6% of proteins were only identified in 8-plex triplicates. Importantly, these results show that 94% of all proteins were identified in 4-plex triplicates and that running an 8-plex experiment instead of a 4-plex experiment would have resulted in 26% fewer protein identifications. This finding is in agreement with those of Besson et al.,[23] who when comparing independent replicates, demonstrated that the use of 8-plex reagents instead of 4-plex reagents could result in 39% fewer proteins being identified. Similarly, the results presented here show that when proteins identified in all 4-plex triplicates were combined and compared with all 6-plex triplicates a high proportion of proteins (95%) were identified in 4-plex triplicates and that running a 6-plex experiment instead of a 4-plex experiment would have resulted in 39% fewer protein identifications.

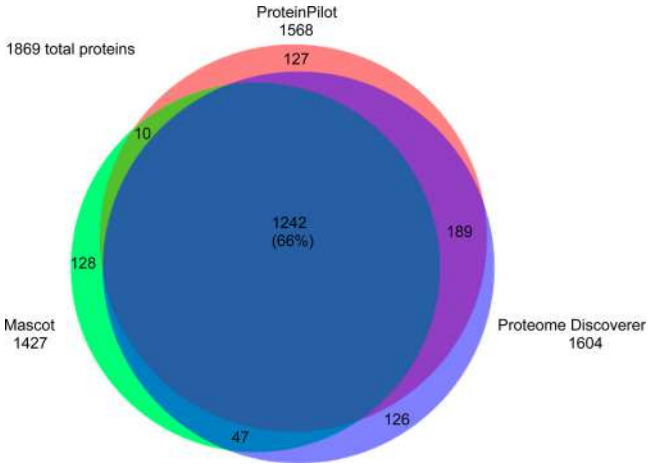


Figure 3. Overlap of proteins identified from 4-plex experiments by different search engines (ProteinPilot, Mascot, and Proteome Discoverer). Results were generated from all three replicates.

Protein Overlap between Analysis Software

A total of 1869 proteins were identified with >95% confidence when spectra from 4-plex experiments were analyzed by all software, with Proteome Discoverer identifying the most proteins (Figure 3). Each of the three replicates were analyzed individually by each software tool before results from all replicates were combined to give a list of total proteins identified. Overall, 66% of proteins were identified by all three software tools, 13% overlapped in two of the three software tools, and 20% of the proteins were only identified by one software tool.

Protein identification results for Mascot and Proteome Discoverer were different despite Mascot being used by Proteome Discoverer as the search engine. This difference is due first to the inbuilt parameter settings for the Mascot node in Proteome Discoverer and second to the "Percolator", node which is an independent FDR calculation algorithm used in Proteome Discoverer.

Effect of Replicates on Protein Coverage

Three technical replicates were analyzed in this study, and the results demonstrate that replicate analysis leads to an increased number of proteins being identified (Supporting Information, Figure S5). Analysis of a second technical replicate in 4-plex experiments resulted in 7, 10, and 7% more protein identifications, respectively, while a third replicate lead to further increases of 6, 14, and 5%, respectively, when spectra were searched with ProteinPilot, Mascot, and Proteome Discoverer. These results are in agreement with those of Paulo,[27] who demonstrated that a second replicate increased protein identifications by 10−15% and that a third replicate resulted in 4 to 5% more proteins being identified. In addition, Paulo[27] found that after the third replicate, gains in the number of protein identifications stabilized and were minimal for four or more replicates.

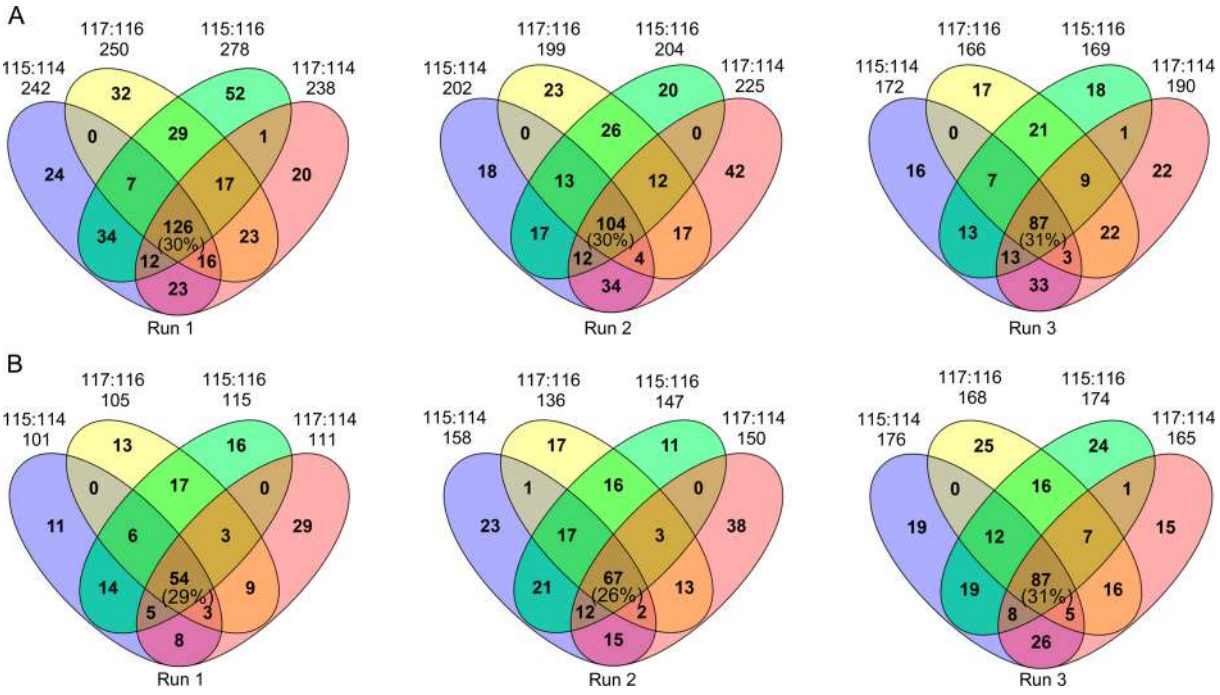


Figure 4. Overlap of differentially expressed proteins showing reproducibility within runs (intrarun reproducibility) for (A) 4-plex experiments and (B) 8-plex experiments. Spectra were searched with ProteinPilot.

DOI: 10.1021/acs.jproteome.5b01154
J. Proteome Res. 2017, 16, 384−392

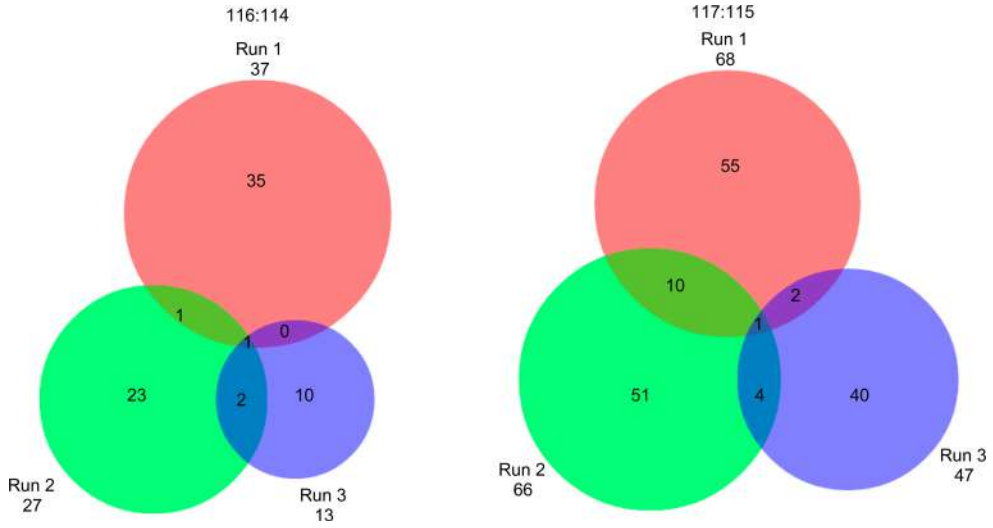


Figure 5. Overlap of differentially expressed proteins identified from 4-plex experiments in negative controls. Spectra were searched with ProteinPilot.

Intrarun Reproducibility

To assess the intrarun reproducibility of differentially expressed proteins for 4-plex experiments, all labeling combinations of SN15:*gna1* were examined; these comprised 115:114, 117:116, 117:114, and 115:116. The same labeling combinations were used for 8-plex experiments, while for 6-plex experiments the following combinations were used: 127:126, 129:128, 127:128, and 129:126. When spectra were analyzed with ProteinPilot, the intra-assay reproducibility for 4-plex and 8-plex was 30 and 29%, respectively (Figure 4). Again, TMT 6-plex analyses were not possible with ProteinPilot. Intrarun reproducibility was lower when spectral analyses were performed with Mascot (16% (4-plex), 11% (8-plex), and 18% (6-plex)) (Supporting Information, Figure S6) and Proteome Discoverer (16% (4-plex), 15% (8-plex), and 21% (6-plex)) (Supporting Information, Figure S7).

Negative Controls

The following combinations were compared for 4-plex experiments: SN15:SN15 (116:114) and *gna1*:*gna1* (117:115) (Figure 5). These combinations serve as negative controls because no statistically significant proteins were expected to be detected for identical samples. In the worst case, 68 proteins were found to be differentially expressed (run 1 for *gna1*:*gna1*). When duplicate runs were compared, 11 proteins were differentially expressed (run 1 and run 2 for *gna1*:*gna1*), and when triplicate runs were compared, one protein was differentially expressed. These findings highlight the importance of running iTRAQ experiments as replicates to minimize false-positive results.

CONCLUSIONS

Despite the continuing popularity of isobaric tagging as a method for quantitative proteomics, data regarding the reproducibility and comparability of the technique are lacking in the literature. For this reason, the current study sought to determine the reproducibility of iTRAQ 4-plex, iTRAQ 8-plex, and TMT 6-plex reagents with respect to total proteome and differentially expressed proteome coverage. Here replicate reproducibility of the total proteome was determined to be ≥69% for technical duplicates and ≥57% for technical triplicates, while the replicate reproducibility of differentially expressed proteins was ≥28% for duplicates and ≥15% for triplicates. This study demonstrates that iTRAQ 4-plex reagents provide a greater depth of total proteome coverage and differentially expressed proteome coverage than iTRAQ 8-plex and TMT 6-plex reagents and that the use of 8-plex and 6-plex reagents instead of 4-plex reagents resulted in up to 39% fewer protein identifications.

Isobaric mass tag reproducibility values may vary in other experimental systems and are likely to be dependent upon factors such as sample complexity, workflow, and instrument. Conclusions as to the effect of changing these parameters cannot be inferred from this study. Nevertheless, the results presented here highlight the need to ensure that any differentially expressed proteins identified from iTRAQ data are validated with a secondary technique such as immunoblotting, ELISA, or a targeted mass spectrometry assay such as MRM/SRM.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.5b01154.

Supporting Information Table S-1: Proteins identified in 4-plex experiments with ProteinPilot Supporting Information. Table S-2: Proteins identified in 8-plex experiments with ProteinPilot. Supporting Information Table S-3: Proteins identified in 6-plex experiments with ProteinPilot. Supporting Information Table S-4: Proteins identified in 4-plex experiments with Mascot. Supporting Information Table S-5: Proteins identified in 8-plex experiments with Mascot. Supporting Information Table S-6: Proteins identified in 6-plex experiments with Mascot. Supporting Information Table S-7: Proteins identified in 4-plex experiments with Proteome Discoverer. Supporting Information Table S-8: Proteins identified in 8-plex experiments with Proteome Discoverer. Supporting Information Table S-9: Proteins identified in 6-plex experiments with Proteome Discoverer. (XLSX)

Supporting Information Figure S1: Reproducibility of identified proteins between duplicate runs. Supporting

DOI: 10.1021/acs.jproteome.5b01154
J. Proteome Res. 2017, 16, 384−392

Information Figure S2: Reproducibility of identified proteins between triplicate runs. Supporting Information Figure S3: Reproducibility of differentially expressed proteins between duplicate runs. Supporting Information Figure S4: Reproducibility of differentially expressed proteins between triplicate runs. Supporting Information Figure S5: Additive effect of replicate analysis on protein coverage in 4-plex experiments. Supporting Information Figure S6: Overlap of differentially expressed proteins showing reproducibility within runs. Supporting Information Figure S7L Overlap of differentially expressed proteins showing reproducibility within runs. (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: scott@proteomics.com.au. Tel: +61 8 9389 1992.

ORCID

Kar-Chun Tan: 0000-0001-6094-823X

Notes

The authors declare no competing financial interest.

All MS data are available via ProteomeXchange with identifier PXD003807 (http://massive.ucsd.edu/ProteoSAFe/status.jsp?task=0e5527b95453496ea9769ef3b2b43cb0).

ACKNOWLEDGMENTS

The analyses were performed in the WA Proteomics Facility, supported by Lotterywest and Bioplatforms Australia, at the Harry Perkins Institute for Medical Research. K.-C.T. received funding through Grains Research and Development Corporation (GRDC) research grants CUR00012 and CUR00023 (Programme 3) and Curtin University. RAS received funding through scholarships from GRDC (GRS10061) and Curtin University.

REFERENCES

(1) Paul, D.; Kumar, A.; Gajbhiye, A.; Santra, M. K.; Srikanth, R. Mass Spectrometry-Based Proteomics in Molecular Diagnostics: Discovery of Cancer Biomarkers Using Tissue Culture. *BioMed Res. Int.* **2013**, *2013*, 783131.

(2) Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **2004**, *3*, 1154−1169.

(3) Pierce, A.; Unwin, R. D.; Evans, C. A.; Griffiths, S.; Carney, L.; Zhang, L.; Jaworska, E.; Lee, C. F.; Blinco, D.; Okoniewski, M. J.; Miller, C. J.; Bitton, D. A.; Spooncer, E.; Whetton, A. D. Eight-channel iTRAQ enables comparison of the activity of six leukemogenic tyrosine kinases. *Mol. Cell. Proteomics* **2008**, *7*, 853−863.

(4) Thompson, A.; Schafer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Hamon, C. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* **2003**, *75*, 1895−904.

(5) Quaedvlieg, W.; Verkley, G. J.; Shin, H. D.; Barreto, R. W.; Alfenas, A. C.; Swart, W. J.; Groenewald, J. Z.; Crous, P. W. Sizing up Septoria. *Stud. Mycol.* **2013**, *75*, 307−390.

(6) Solomon, P. S.; Lowe, R. G. T.; Tan, K. C.; Waters, O. D. C.; Oliver, R. P. *Stagonospora nodorum*: cause of stagonospora nodorum blotch of wheat. *Mol. Plant Pathol.* **2006**, *7*, 147−156.

(7) Solomon, P. S.; Tan, K. C.; Sanchez, P.; Cooper, R. M.; Oliver, R. P. The disruption of a Ga subunit sheds new light on the pathogenicity of *Stagonospora nodorum* on wheat. *Mol. Plant-Microbe Interact.* **2004**, *17*, 456−466.

(8) Gummer, J. P.; Trengove, R. D.; Oliver, R. P.; Solomon, P. S. A comparative analysis of the heterotrimeric G-protein Ga, Gb and Gg subunits in the wheat pathogen *Stagonospora nodorum*. *BMC Microbiol.* **2012**, *12*, 131.

(9) Hane, J. K.; Lowe, R. G.; Solomon, P. S.; Tan, K. C.; Schoch, C. L.; Spatafora, J. W.; Crous, P. W.; Kodira, C.; Birren, B. W.; Galagan, J. E.; Torriani, S. F. F.; McDonald, B. A.; Oliver, R. P. Dothideomycete plant interactions illuminated by genome sequencing and EST analysis of the wheat pathogen *Stagonospora nodorum*. *Plant Cell* **2007**, *19*, 3347−3368.

(10) Bringans, S.; Hane, J. K.; Casey, T.; Tan, K. C.; Lipscombe, R.; Solomon, P. S.; Oliver, R. P. Deep proteogenomics; high throughput gene validation by multidimensional liquid chromatography and mass spectrometry of proteins from the fungal wheat pathogen *Stagonospora nodorum*. *BMC Bioinf.* **2009**, *10*, 301.

(11) Syme, R. A.; Hane, J. K.; Friesen, T. L.; Oliver, R. P. Resequencing and comparative genomics of Stagonospora nodorum: sectional gene absence and effector discovery. *G3: Genes, Genomes, Genet.* **2013**, *3*, 959−969.

(12) Tan, K. C.; Heazlewood, J. L.; Millar, A. H.; Thomson, G.; Oliver, R. P.; Solomon, P. S. A signaling-regulated, short-chain dehydrogenase of *Stagonospora nodorum* regulates asexual development. *Eukaryotic Cell* **2008**, *7*, 1916−1929.

(13) Casey, T.; Solomon, P. S.; Bringans, S.; Tan, K. C.; Oliver, R. P.; Lipscombe, R. Quantitative proteomic analysis of G-protein signalling in *Stagonospora nodorum* using isobaric tags for relative and absolute quantification. *Proteomics* **2010**, *10*, 38−47.

(14) Tan, K. C.; Heazlewood, J. L.; Millar, A. H.; Oliver, R. P.; Solomon, P. S. Proteomic identification of extracellular proteins regulated by the Gna1 Ga subunit in *Stagonospora nodorum*. *Mycol. Res.* **2009**, *113*, 523−531.

(15) Tan, K. C.; Trengove, R. D.; Maker, G. L.; Oliver, R. P.; Solomon, P. S. Metabolite profiling identifies the mycotoxin alternariol in the pathogen *Stagonospora nodorum*. *Metabolomics* **2009**, *5*, 330−335.

(16) Tang, W. H.; Shilov, I. V.; Seymour, S. L. Nonlinear fitting method for determining local false discovery rates from decoy database searches. *J. Proteome Res.* **2008**, *7*, 3661−3667.

(17) Oliveros, J. C. *Venny. An interactive tool for comparing lists with Venn's diagrams*, 2007−2015. http://bioinfogp.cnb.csic.es/tools/venny/index.html.

(18) Hulsen, T.; de Vlieg, J.; Alkema, W. BioVenn − a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* **2008**, *9*, 488.

(19) Pichler, P.; Kocher, T.; Holzmann, J.; Mazanek, M.; Taus, T.; Ammerer, G.; Mechtler, K. Peptide Labeling with Isobaric Tags Yields Higher Identification Rates Using iTRAQ 4-Plex Compared to TMT 6-Plex and iTRAQ 8-Plex on LTQ Orbitrap. *Anal. Chem.* **2010**, *82*, 6549−6558.

(20) Li, Z.; Adams, R. M.; Chourey, K.; Hurst, G. B.; Hettich, R. L.; Pan, C. Systematic Comparison of Label-Free, Metabolic Labeling, and Isobaric Chemical Labeling for Quantitative Proteomics on LTQ Orbitrap Velos. *J. Proteome Res.* **2012**, *11*, 1582−1590.

(21) Pottiez, G.; Wiederin, J.; Fox, H. S.; Ciborowski, P. Comparison of 4-plex to 8-plex iTRAQ Quantitative Measurements of Proteins in Human Plasma Samples. *J. Proteome Res.* **2012**, *11*, 3774−3781.

(22) Sandberg, A.; Branca, R. M.; Lehtiö, J.; Forshed, J. Quantitative accuracy in mass spectrometry based proteomics of complex samples: The impact of labeling and precursor interference. *J. Proteomics* **2014**, *96*, 133−144.

(23) Besson, D.; Pavageau, A. H.; Valo, I.; Bourreau, A.; Bélanger, A.; Eymerit-Morin, C.; Moulière, A.; Chassevent, A.; Boisdron-Celle, M.; Morel, A.; Solassol, J.; Campone, M.; Gamelin, E.; Barré, B.; Coqueret, O.; Guette, C. A Quantitative Proteomic Approach of the Different Stages of Colorectal Cancer Establishes OLFM4 as a New Non-metastatic Tumor Marker. *Mol. Cell. Proteomics* **2011**, *10*, M111.009712.

DOI: 10.1021/acs.jproteome.5b01154
J. Proteome Res. 2017, 16, 384−392

(24) Song, F.; Poljak, A.; Kochan, N. A.; Raftery, M.; Brodaty, H.; Smythe, G. A.; Sachdev, P. S. Plasma protein profiling of Mild Cognitive Impairment and Alzheimer's disease using iTRAQ quantitative proteomics. *Proteome Sci.* **2014**, *12*, 5.
(25) Ji, Y. H.; Ji, J. L.; Sun, F. Y.; Zeng, Y. Y.; He, X. H.; Zhao, J. X.; Yu, Y.; Yu, S. H.; Wu, W. Quantitative Proteomics Analysis of Chondrogenic Differentiation of C3H10T1/2 Mesenchymal Stem Cells by iTRAQ Labeling Coupled with On-line Two-dimensional LC/MS/MS. *Mol. Cell. Proteomics* **2010**, *9*, 550−564.
(26) Wang, H.; Alvarez, S.; Hicks, L. M. Comprehensive comparison of iTRAQ and label-free LC-based quantitative proteomics approaches using two Chlamydomonas reinhardtii strains of interest for biofuels engineering. *J. Proteome Res.* **2012**, *11*, 487−501.
(27) Paulo, J. A. Practical and Efficient Searching in Proteomics: A Cross Engine Comparison. *WebmedCentral* **2013**, *4*, WMCPLS0052.

DOI: 10.1021/acs.jproteome.5b01154
J. Proteome Res. 2017, 16, 384−392