Analyst

An accurate and high throughput isobaric MS2 quantification strategy based on metabolic labeling and trypsin digestion.
39x19mm (600 x 600 DPI)

Global *in vivo* terminal amino acid labeling for exploring differential expressed proteins induced by dialyzed serum cultivation

*Li-Qi Xie[§, 1,2], Ai-Ying Nie[§,1], Shu-Jun Yang,[1] Chao Zhao[2], Lei Zhang[1], Peng-Yuan Yang[1] and Hao-Jie Lu[*1,2]*

1. Shanghai Cancer Center and Department of Chemistry, Fudan University, Shanghai 200032, P. R. China

2. Key Laboratory of Medical Molecular Virology and Institutes of Biomedical Sciences, Shanghai Medical College, Fudan University, Shanghai 200032, P. R. China

§ These authors contributed equally to this work

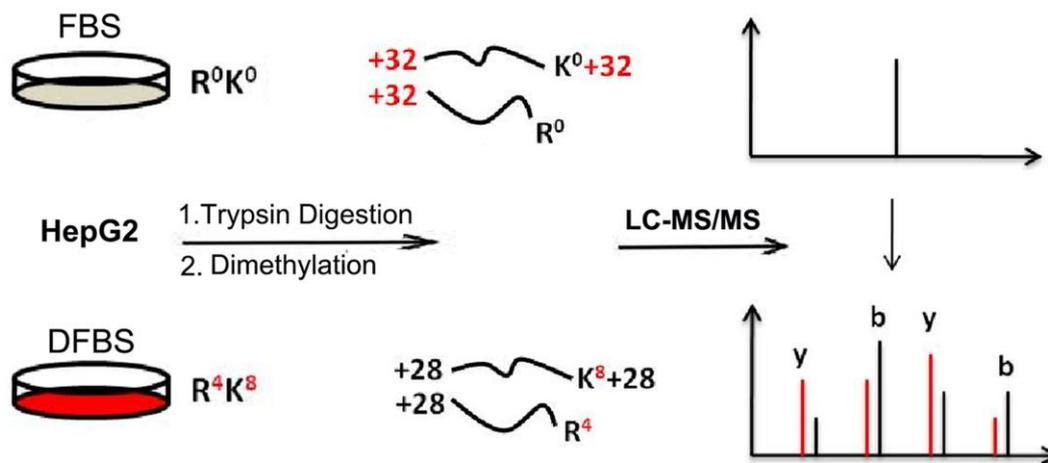**Corresponding Author*** Prof. Lu Haojie

Tel: (086) 021-54237618.

Fax: (086) 021-54237961.

E-mail: luhaojie@fudan.edu.cn.

Table of contents entry

An accurate and high throughput isobaric MS2 quantification strategy based on metabolic labeling and trypsin digestion.

.

Abstract

Taking advantage of reliable metabolic labeling and accurate isobaric MS2 quantification, we developed a global *in vivo* terminal amino acid labeling (G-IVTAL) strategy by combining metabolic labeling and isotopic dimethyl labeling for quantifying tryptic peptides. With G-IVTAL, the scale of qualitative and quantitative data can be increased by two fold when compared with *in vivo* termini amino acid labeling (IVTAL) in which Lys-N and Arg-C are used for digestion. As a result, up to 81.78% of identified proteins have been confidently quantified in G-IVTAL labeled HepG2 cells. Dialyzed serum has been used in most SILAC studies to ensure complete labeling. However, dialysis requests to remove the low molecular weight hormone, cytokines and cellular growth factors which are essential for cell growth of certain cell lines. To address the influence of dialyzed serum in HepG2 growth, the G-IVTAL strategy was applied to quantify the expression differences between dialyzed serum cultured and normal serum cultured HepG2 cells. Finally, we discovered 111 differentially expressed proteins, which could be used as references to improve reliability of SILAC quantification. Among them, the differential expressions of MTDH, BCAP31 and GPC3 were confirmed to be influenced by dialyzed serum using western blotting. The experimental results demonstrate that G-IVTAL strategy is a powerful tool to achieve accurate and reliable protein quantification.

Introduction

Quantitative proteomics is a rapid-developing area [1]. With the emergence of new quantitative

methods, a lot of candidate biomarkers have been identified [2, 3]. In terms of spectra, current

quantification methods can be classified into precursor ions (MS1) based and fragmental ions

(MS2/MS3) based methods [1]. MS1 based quantification methods, such as SILAC/AACT,

have been used for protein quantification by comparing intensities or areas of paired

heavy/light isotopic MS1 peaks[4]. SILAC can achieve accurate quantification because of

reduced uncompensated losses during parallel processing by mixing specimens at early stage[4-6].

However, due to the multiplied MS1 peaks, the complexity of MS spectra is increased, the

sensitivity is reduced and the under sampling problem is aggravated[7]. To avoid these

disadvantages, isobaric tags based MS2 quantification methods, such as iTRAQ and TMT,

were developed [8, 9]. This kind of methods generate isobaric precursor ions and uses reporter

ion pairs in the low mass range of MS2 spectra for relative quantification [8]. Unfortunately,

repression effects on mass tags in the low mass range dramatically affect the accuracy and

dynamic range of the method [10].

Isobaric MS2 quantification using b, y fragment ion pairs in the whole mass range provides

an ideal solution for quantitative proteomics. As isobaric tags based MS2 quantification, the

method allows sensitive and reproducible quantification without increasing MS1 complexity

[11]. Meanwhile, several fragmental ion pairs can be used for quantification simultaneously,

which can increase reliability as well as avoid signal repression effects in the low mass range

[12, 13]. For example, by labeling the N-term and C-term of Lys-C digested peptides with

different isotope reagents, isobaric peptide termini labeling (IPTL) produces identical MS

4

peaks in MS1 spectra and fragmental ion pairs in MS2 scan [11, 14]. Zou et.al optimized the

multi-processes procedure by using single site-selective N-terminus dimethylation reaction to

produce isobaric peptides [15]. Despite these improvements, such approaches can only be

applied in Lys-C digested peptides, which are unfavorable for CID based MS identification [16].

Accordingly, we developed isobaric MS2 quantification strategy for tryptic peptides by

combining ^{18}O labeling, dimethylation and guanidination in a former study [17]. However, the

method included three labeling reactions with multiple steps, which may lead to insufficient

labeling and byproducts [18].

Different from the methods mentioned above, IVTAL generates isobaric peptides by using

a set of heavy amino acid ($^{13}C_6$-arginine or $^{13}C_6$-lysine) for cell culture and Lys-N and Arg-C

for digestion [19]. Regardless of the digestion efficiency, only one third of specific peptides

digested by Lys-N and Arg-C can be used for MS2 quantification [19]. To make better use of

metabolic labeling for its high reliability and accuracy as well as high sensitivity in MS2

quantification, we established a G-IVTAL strategy for quantifying tryptic peptides. In this

novel strategy, cells were *in vivo* labeled with heavy or normal arginine and lysine. After that,

the mostly used enzyme in proteomics study, trypsin, was selected to digest proteins so that

qualitative and quantitative information can be obtained as much as possible. Furthermore, in

order to obtain the isobaric tryptic peptides for protein quantification, isotope dimethylation

labeling, with excellent labeling efficiency, simple procedure, short reaction time and low cost

[20], was adopted.

Dialyzed serum is used in cell culture medium to ensure completed labeling in most

SILAC based quantification studies [21, 22]. Without low molecular weight hormone, cytokines

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

and cellular growth factors, dialyzed serum did not significantly influence cell growth in most

cases [4]. However, with the exception of certain cell lines such as MCF-7, certain amount of

normal serum is essential for their growth [23]. To overcome such problem, Pandey used normal

serum in SILAC labeling, and found that the influence of amino acid in normal serum is

ignorable [24, 25]. To address the influence of dialyzed serum in cell growth, we applied G-

IVTAL to explore the protein expression differences between dialyzed fetal bovine serum

(DFBS) and FBS cultivated HepG2 cell lines.

Results and discussion

Framework of G-IVTAL strategy. Trypsin is the most efficient and widely used enzyme in proteome research [26]. To achieve universally applicable isobaric MS2 quantification, G-IVTAL was developed for tryptic peptides in present study. For tryptic peptides, the C-terminus is either lysine or arginine. The two amino acids have several commercially available heavy isotope forms i.e. 4, 6, 10 or 17 Da mass increase for arginine and 4, 6, 8 or 17 Da mass increase for lysine. To cross-sequentially label the peptides, dimethylation, a simple, fast and complete reaction specifically reacts with the peptide N-terminus or ε-amino groups on lysine residues [20], was introduced. By using different isotopomers of formaldehyde and cyanoborohydride, peptides with 2, 4 or 8 Da mass differences can be produced. The side chain amino group of arginine shows no reactions during borch reductive animation [27]. Therefore, peptides with C-terminus lysine will have two fold mass increase than peptides with C-terminus arginine after dimethylation. As a result, tags with two times heavier mass were required on C-terminus lysine peptides in counterpart groups. Accordingly, isotopic arginine have 4 Da mass difference (R^0, R^4), isotopic lysine have 8 Da mass difference (K^0, K^8) and isotopic formaldehyde reagents (CD$_2$O, CH$_2$O) were selected to construct the isobaric peptides.

The workflow of G-IVTAL is shown in Figure. 1. First, cell state A and cell state B were cultured in media with lysine (K^0) and arginine (R^0) or $^{13}C_6{}^{15}N_2$-lysine (K^8) and $^{15}N_4$-arginine (R^4), respectively, for sufficient time to obtain nearly complete ($\geqslant 98\%$) protein labeling. Next, the cell lysates were digested with trypsin and dimethylated by CD$_2$O or CH$_2$O respectively. Subsequently, differentially labeled peptides were equally mixed for LC-MS/MS

analysis. The b, y fragments with 4 Da mass difference in MS2 spectra were utilized for relative peptide and protein quantification. The G-IVTAL strategy increased the number of identified and quantified proteins by trypsin digestion meanwhile retained the advantages of *in vivo* MS2 quantification (Figure 1). Compared with Lys C based isobaric MS2 quantification methods, G-IVTAL can be more competent for CID based MS identification [16]. In addition, G-IVTAL only requires one chemical labeling step, which reduces the possibility of insufficient labeling and byproducts when compared with QITL [15]. A possible limitation is that G-IVTAL can only be applied in quantification of cellular samples due to *in vivo* labeling.

Investigation of labeling efficiency and isotopic effect in LC separation. To ensure the metabolic labeling efficiency, HepG2 cells grown for 6 passages in R^4K^8 medium were extracted for SDS-PAGE separation and identified by MALDI-MS. Compared with labeled peptides, the intensity of unlabeled peptide (m/z 1132.68) is ignorable, which suggests completed and unbiased incorporation of K^8 and R^4 (supplemental Figure. S1). To evaluate the labeling efficiency of isotope dimethylation in G-IVTAL strategy, tryptic peptides of hemoglobin, BSA, myoglobin and cytochrome c were reacted with CH_2O and CD_2O, respectively. After dimethylation, peptides with C-terminal arginine such as VGAHAGEYGAEALER exhibits 28 Da or 32 Da mass increase, while peptides with C-terminal lysine such as TYFPHFDLSHGSAQVK shows 56 Da or 64 Da mass shift (supplemental Figure. S2). Similar to previous study [28], enhanced MS signals and increased peptide sequence coverage of dimethylated peptides were observed. As shown in supplemental table S1, the labeling efficiency of tryptic hemoglobin peptides are ≥99%. And the deuterated peptides have the same retention time as light counterparts in LC-MS/MS

analysis (supplemental Figure. S3). It has been reported that the deuterium atoms scattering around the polar groups in CD_2O minimized chromatography effect of deuterated peptides [20, 29]. Accordingly, the negligible isotope effect on LC separation would not affect the quantitative accuracy of G-IVTAL.

Reproducible and accurate quantification of G-IVTAL labeled HepG2 cells. RP−RP fractionation operated at different pH was used to separate equally mixed heavy and light labeled hepG2 cell following G-IVTAL strategy. Paired b, y fragment ions with mass difference of 4 Da are observed in MS2 spectra of both peptides with C-terminus lysine or arginine (Figure. 2). To ensure the reliability of quantification, at least six b, y ion pairs were used for quantification. B, y ion pairs in Figure. 2A are assigned to the peptide of [+28]ETSGNLEQLLLAVVK[+28+8] and [+32]ETSGNLEQLLLAVVK[+32]. Their average heavy/light ratio is 1.0 with a standard deviation of 0.19, which is exactly the expected ratio. Figure. 2B displays the MS2 spectra of [+28]EAEAMALLAEAER[+4] and [+32]EAEAMALLAEAER. The quantification value agreed expect ratio with a standard deviation of 0.15. All these results proved the reliability and accuracy of G-IVTAL based protein quantification.

In total, three replicated analysis identified 31004 non-redundant peptides and 5412 proteins. The identified sequences and observed heavy/light ratios for each analysis are listed in supplemental materials-2 (https://sourceforge.net/projects/givtal/supplemental materials-2). The numbers of identified and quantified peptides as well as proteins are listed in supplemental table S2, including the mean quantification ratio and corresponding standard deviation of each replicate. The venn diagram shows over 66.94% coverage of quantified and identified proteins among replicates, which indicates considerable reproducibility of the G-

9

IVTAL strategy (Figure. S4). As expected, benefitting from trypsin digestion and good separation ability of differential pH 2D-RPLC, the ratio of quantified proteins versus identified proteins was two times higher than IVTAL strategy. 62.26% of peptides and 81.78% of proteins were confidently quantified by G-IVTAL strategy while the number are 32.03% and 47.49% respectively in IVTAL strategy.

Dialyzed serum induced protein differential expressions. G-IVTAL quantification strategy was used to evaluate the proteomic differences between HepG2 cells cultivated in DFBS medium containing K^8R^4 and in FBS medium containing light amino acids. After *in vivo* metabolic isotope labeling and isotope dimethyl labeling, the two samples were mixed equally for LC-MS analysis. Three replicate analyses were introduced to ensure the quantification reliability. The identified sequences and observed heavy/light ratio for each analysis are listed in supplemental materials-3 (https://sourceforge.net/projects/givtal/supplemental materials-3). Altogether, 17486 peptides were quantified, and the number of identified and quantified protein was 5324 and 4288, respectively (Supplemental table S3). Fig. 3 indicates reproducible and reliable quantification. After normalization and statistical analysis following previous study [19], proteins with over two-fold changes between the two specimens were defined as differentially expressed. Totally, 111 differentially expressed proteins were discovered, and the corresponding quantitative ratios and RSDs are listed in supplemental table S4. Among them, 23 proteins are up-regulated and 88 proteins are down-regulated in DFBS cultivate HepG2 cells.

Before chosen as candidate biomarkers, the influence of dialyzed serum and isotope labeling should be deduced in SILAC studies. The reference data set of 111 proteins provided

here could assist researchers to acquire more accurate and reliable quantitative information.

Based on our investigation of previous HepG2 studies using dialyzed serum for SILAC

quantification, some of the candidate protein biomarkers were identified to be differentially

expressed in our study [22, 30, 31]. For example, in SILAC quantification of HBV infected

proteome changes of HepG2 lipid rafts, Annexin III and Annexin IV were down regulated [30].

And Annexin III and Annexin IV were up-regulated in HepG2 in a SILAC study aiming for

identifying novel diagnosis biomarkers for early hepatocellular carcinoma [22]. While in our

current study, Annexin III was up-regulated and Annexin IV was down-regulated in dialyzed

serum. Similarly, PGRC1, PGRC2, VAPA, NB5R3, TMCO1, MOGS, ACSL5, RET1, CLD1

and TGM2 were identified as differentially expressed proteins in various SILAC studies [22, 30,

31]. However, these studies didn't considered the expression difference induced by dialyzed

serum cultivation. We demonstrate that their quantification ratios should be re-calculated by

wiping out the influence of dialyzed serum and isotope labeling.

Functional annotation and immunity validation of dialyzed serum induced protein

differential expression. Ingenuity pathway analysis of the 111 differentially expressed

proteins revealed that the most related functions were protein post-translational modification,

fatty acid metabolism, synthesis of sterol and lipid and cell viability. By using the quantitative

ratio of hepG2 cell lines cultured in dialyzed serum with versus without heavy amino acid to

deduce the influence of isotope labeling, the differential expression of 90 proteins were found

deeply correlated with dialyzed serum (supplemental table S4). Through downstream effect

analysis of IPA, we found that fatty acid metabolism and cell viability were significantly

inhibited after dialyzed serum cultivation. And we can conclude from network analysis that

11

the differentially expressed proteins interacted with each-other mainly through PI3K, NFkB

and JNK pathways, which plays important roles in cell cycle, apoptosis and various kinds of

diseases (Figure. 4). Three typical regulators of theses pathways, GPC3, BACP31 and MTDH

[32-34], were selected for western blotting validation.

To ensure the differential expression were resulted from dialyzed serum, HepG2 cells

were divided into 3 groups, FBS, DFBS and 10DFBS. HepG2 cells were equally passaged

into three flasks and cultivated in serum free medium for 24 hr, and then cultivated in medium

containing FBS, DFBS or 10DFBS for 48 hr, respectively. The serum mixture of 10DFBS

was prepared by mixing comparable amount of DFBS and the low molecular weight fraction

of FBS flowing through 10 kDa ultrafiltration membrane. Compared with DFBS groups, FBS

and 10DFBS groups exhibited consistent expression trends according to western blotting

images of GPC3, BACP31 and MTDH (Figure. 5). These results confirmed that the

expression differences of the three proteins were induced by dialyzed serum cultivation.

GPC3, the cell surface heparan sulfate proteoglycan involved in the suppression/modulation

of growth in the predominantly mesodermal tissues and organs [32], increased in DFBS

cultivated HePG2 cells and fell to normal expression level after adding low molecular serum

fraction below 10 kD. BCAP31, which involved in CASP8-mediated apoptosis [35], and MTDH,

the activator of NFkB [34], were both verified to be down-regulated in DFBS groups. The

consistent expression trends of western-blotting results and G-IVTAL data confirmed the

reliability of the 111 proteins as the reference data set.

Experimental

Cell culture with normal or isotopic labeled amino acids

In the establishment of G-IVTAL strategy, HepG2 cells were cultured in arginine and lysine

free Dulbecco's modified Eagle's medium (DMEM, US Biolocial) with 10% DFBS, 100

units/ml penicillin and streptomycin. Normal arginine (Sigma, A6969) and lysine (Sigma,

L8662) or heavy lysine (K^8, Cambridge Isotope Laboratories, CNLM-291) and arginine (R^4,

Cambridge Isotope Laboratories, NLM-396) were added to the medium respectively. And

normal L-proline (Sigma, P5607) was added to the media with a final concentration of 200

μg/ml to prevent arginine-proline conversion [21]. Cells were tested for fully incorporation of

isotope labels after six passages.

To evaluate the effect of the dialyzed serum on the HepG2 cells, another set of

HepG2 cells culturing in the DMEM medium with10% FBS and normal arginine & lysine

were compared with those culturing in 10% DFBS and heavy arginine & lysine containing

DMEM medium.

Protein extraction and trypsin digestion

Cells were collected and lysed with 7M urea, 2M thiourea as previous described [19]. The

concentration of protein was measured using the Bradford assay. Aliquots of proteins were

reduced by 10 mM dithiothreitol at 37 °C for 1 hour, alkylated by 55 mM iodoacetamide at

room temperature for 45 min in dark and finally precipitated overnight by ice acetone. After

centrifugation, the deposited proteins were digested overnight with trypsin at a

substrate/enzyme ratio of 50:1 (w/w) in 25 mM ammonium bicarbonate, pH 7.8 at 37 °C.

Standard proteins of hemoglobin, myoglobin and cyctochrome C were digested overnight

with trypsin as described above. Peptides were lyophilized in vacuum.

Stable Isotope dimethyl labeling

Lyophilized peptides were re-dissolved in sodium acetate buffer (100 mM, pH 5–6), and then

mixed with freshly prepared 4% CD_2O or CH_2O solution, followed by the addition of freshly

prepared sodium cyanoborohydride (600 mM) [9]. Afterwards, the mixtures were incubated in a

fume hood for 1 h at room temperature. Then, 4% ammonium hydroxide was added and

incubated to quench the reaction [36]. Finally, aliquots of differently labeled standard proteins

were mixed equally, purified with μ-C18 ZipTips (Millipore, Billerica, MA) and analyzed

using MALDI-TOF/TOF-MS 5800 System (AB SCIEX, Foster City, CA).

Differential pH 2D-RPLC-ESI-MS/MS.

The labeled peptides from the HepG2 samples were reconstituted in buffer A [10 mM

ammonium formate (NH_4FA), pH 10] and injected onto Sepax PolyRP-300 column (5 μm,

300 Å, 2.1 × 150 mm, Sepax Technologies, Newark, DE) using an LC-20AD high

performance LC system (Shimadzu, Tokyo, Japan) for the first dimensional high pH-RPLC

separation. Peptides were eluted from the column with a 40 min gradient from 0 to 50%

buffer B (90% ACN/10% 10 mM NH_4FA, pH 10), followed by a 4 min gradient from 50% to

80% buffer B. A total of 24 fractions were collected based on UV absorbance at 214 nm. All

fractions were lyophilized and reconstituted in 20 μL of 0.1% formic acid (FA) before the

second dimensional low pH RPLC-ESI-MS/MS analysis.

The collected fractions were injected in turn to a nano ACQUITY UPLC system (Waters

Corporation, Milford) connected with an LTQ-Orbitrap XL mass spectrometer (Thermo

Electron Corp., Bremen, Germany). The peptide mixtures were separated with a three-step

linear gradient from 5% to 45% phase B in 95 min (phase A, water with 0.1% FA; phase B,

ACN with 0.1% FA), then increased to 80% phase B in 5 min, and finally maintained at 80%

phase B for 5 min in a 100 μm i.d. × 15 cm reverse phase column (Michrom Bioresources,

Auburn) at 500 nL/min, 35 °C. An electrospray voltage of 2.0 kV versus the mass

spectrometer inlet was used. The LTQ-Orbitrap XL mass spectrometer was operated in data-

dependent mode to switch automatically between MS1 and MS2 acquisition. MS1 scan with

one microscan (m/z 400 to 1800) was acquired in the Orbitrap with a mass resolution of

60000 at m/z 400, followed by eight sequential LTQ-MS/MS scans. Dynamic exclusion was

used with two repeat counts: a 10 s repeat duration, and 90 s exclusion duration. For MS2,

precursor ions were activated using 35% normalized collision energy at the default activation

q of 0.25. Three technical replicates were performed.

Protein identification and quantification

All MS2 spectra were searched twice against the human UniProt database (release 2009–02

with 20,331 entries) and corresponding reversed sequences using SEQUEST [v.28 (revision

1)]. Two sets of search parameters were used. For the first set of parameters, N-termini

(+32.0564 Da) and lysine (+32.0564 Da) for CD_2O-derived peptides were set up as fixed

modification, whereas for the second set of parameters, N-termini (+28.0313 Da), lysine

(+36.0550 Da) and arginine (+3.98820 Da) for CH_2O-derived peptides were set up as fixed

modification. And both sets used fixed modification on cysteine (+57.0215 Da) and variable

modification on methione (+15.9949 Da), partial tryptic cleavage with two missed cleavage

15

sites, mass tolerance of 50 ppm for precursor ions and 1.0 Da for fragment ions. Trans

Proteomic Pipeline software (revision 4.2, Institute of Systems Biology, Seattle, WA) was

utilized to score the search results, unique peptides with PeptideProphet probability over 0.90

and proteins with ProteinProphet probability over 0.95 were retained. False discovery rate

was limited to less than 1%. Results of the two sets of database searches were combined to

determine the identified peptides and proteins for quantification.

Using in-house built scripts edited by Perl (version 5.10) and MatLab (version 7.10),

the intensity values of b, y fragment ions which ranked in the top 200 in the DTA files of the

identified peptides were extracted [21]. The outlier data of b, y fragment ion pairs were

removed using box plot. The peptide quantification ratio was calculated by taking the

average ratios of assigned b, y fragment ion pairs in MS2 spectra. And only the peptides that

include more than six pairs of b, y ions were quantified. According to the database search

results, the quantified peptides were assigned to corresponding proteins, and the peptide

ratios were averaged to get protein ratios. Standard deviation and coefficient of variance of

the quantified proteins were calculated simultaneously. For the G-IVTAL labeled HepG2

samples, protein ratios larger than 2 or less than 0.5 in three replicated analyses were

determined as the up- or down-regulated proteins. Ingenuity Pathway Analysis (version 2012

Winter) was used for the functional analysis of these up- or down-regulated proteins.

Additionally, in the test of isotope effect about dimethyl labeled peptides, the CH_2O and

CD_2O labeled standard proteins were quantified by Xpress.

16

Western blotting analysis

Three differentially expressed protein, GPC3, BCAP31 and MTDH were selected for immune

validation. Equal amount of cell lysates were run on SDS-PAGE gel and transferred to PVDF

membrane before incubation with antibodies. After antibody staining, the chemiluminescent

samples were detected with a cooled CCD imaging system LAS-3000. β-actin was used as

endogenous control. All antibodies used here were purchased from Abcam.

17

Conclusion

We developed G-IVTAL strategy by combining SILAC and isotopic dimethyl labeling for quantifying tryptic peptides. Compared with IVTAL, G-IVTAL strategy maintained the advantages of high accuracy and reliability, as well as greatly improved the quantitative throughput by using trypsin for digestion. When applied in quantification of the differential expression of DFBS and FBS cultivated HepG2 cell, G-IVTAL identified 5324 proteins and 4288 of them were confidently quantified. Among them, 111 proteins were differentially expressed, including 23 up-regulated and 88 down-regulated proteins in DFBS cultivated cells. Western blotting analysis verified the expression differences of MTDH, BCAP31 and GPC3, which further confirmed reliability of the quantification information. Functional annotation revealed that fatty acid, lipid metabolism, vitamin and mineral metabolism were significantly changed by dialyzed serum cultivation. The reference data set of 111 proteins provided here could assist acquiring more accurate and reliable SILAC quantitative information. These results suggested that G-IVTAL provides a high through-put, reproducible and reliable isobaric MS2 quantification strategy.

Acknowledgements

All authors have declared no conflict of interest.

Bibliographic References

1. B. Domon and R. Aebersold, *Nat. Biotechnol.*, 2010, **28**, 710-721.

2. Y. Zhao, W. N. P. Lee and G. G. Xiao, *Expert Rev. Proteomics*, 2009, **6**, 115.

3. F. Abdi, J. F. Quinn, J. Jankovic, M. McIntosh, J. B. Leverenz, E. Peskind, R. Nixon, J. Nutt, K. Chung and C. Zabetian, *J. Alzheimer's Dis.*, 2006, **9**, 293-348.

4. S. E. Ong, B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey and M. Mann, *Mol. Cell. Proteomics*, 2002, **1**, 376-386.

5. X. Chen, L. M. Smith and E. M. Bradbury, *Anal. Chem.*, 2000, **72**, 1134-1143.

6. H. Zhu, S. Pan, S. Gu, E. M. Bradbury and X. Chen, *Rapid Commun. Mass Spectrom.*, 2002, **16**, 2115-2123.

7. M. Colzani, F. Schutz, A. Potts, P. Waridel and M. Quadroni, *Mol. Cell. Proteomics*, 2008, **7**, 927-937.

8. A. Thompson, J. Schäfer, K. Kuhn, S. Kienle, J. Schwarz, G. Schmidt, T. Neumann and C. Hamon, *Anal. Chem.*, 2003, **75**, 1895-1904.

9. P. L. Ross, Y. N. Huang, J. N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey and S. Daniels, *Mol. Cell. Proteomics*, 2004, **3**, 1154-1169.

10. A. S. Hebert, A. E. Merrill, D. J. Bailey, A. J. Still, M. S. Westphall, E. R. Strieter, D. J. Pagliarini and J. J. Coon, *Nat. Meth.*, 2013, **10**, 332-334.

11. C. J. Koehler, M. Strozynski, F. Kozielski, A. Treumann and B. Thiede, *J. Proteome Res.*, 2009, **8**, 4333-4341.

12. A. Treumann and B. Thiede, *Expert Rev. Proteomics*, 2010, **7**, 647-653.

13. G. Hopfgartner, E. Varesio, V. Tschäppät, C. Grivet, E. Bourgogne and L. A. Leuthold, *J. Mass Spectrom.*, 2004, **39**, 845-855.

14. C. J. Koehler, M. O. Arntzen, M. Strozynski, A. Treumann and B. Thiede, *Anal. Chem.*, 2011, **83**, 4775-4781.

15. H. Qin, F. Wang, Y. Zhang, Z. Hu, C. Song, R. Wu, M. L. Ye and H. Zou, *Chem. Commun.*, 2012, **48**, 6265-6267.

16. L. Q. Xie, C. P. Shen, M. B. Liu, Z. D. Chen, R. Y. Du, G. Q. Yan, H. J. Lu and P. Y. Yang, *Mol. BioSyst.*, 2012, **8**, 2692-2698.

17. S. J. Yang, A. Y. Nie, L. Zhang, G. Q. Yan, J. Yao, L. Q. Xie, H. J. Lu and P. Y. Yang, *J Proteomics*, 2012, **75**, 5797-5806.

18. X. Zhang, Q. K. Jin, S. A. Carr and R. S. Annan, *Rapid Commun. Mass Spectrom.*, 2002, **16**, 2325-2332.

19. A. Y. Nie, L. Zhang, G. Q. Yan, J. Yao, Y. Zhang, H. J. Lu, P. Y. Yang and F. C. He, *Anal. Chem.*, 2011, **83**, 6026-6033.

20. J. L. Hsu, S. Y. Huang, N. H. Chow and S. H. Chen, *Anal. Chem.*, 2003, **75**, 6843-6852.

21. S. C. Bendall, C. Hughes, M. H. Stewart, B. Doble, M. Bhatia and G. A. Lajoie, *Mol. Cell. Proteomics*, 2008, **7**, 1587-1597.

22. F. Ren, H. Wu, Y. Lei, H. Zhang, R. Liu, Y. Zhao, X. Chen, D. Zeng, A. Tong, L. Chen, Y. Wei and C. Huang, *Mol. Cancer*, 2010, **9**, 81.

23. M. L. Gehrmann, Y. Hathout and C. Fenselau, *J. Proteome Res.*, 2004, **3**, 1063-1068.

24. H. Harsha, H. Molina and A. Pandey, *Nat. Protoc.*, 2008, **3**, 505-516.

25. S. Rayavarapu, W. Coley, E. Cakir, V. Jahnke, S. Takeda, Y. Aoki, H. Grodish-Dressman, J. K.

19

Jaiswal, E. P. Hoffman, K. J. Brown, Y. Hathout and K. Nagaraju, *Mol. Cell. Proteomics*, 2013, **12**, 1061-1073.

26.   R. Aebersold and M. Mann, *Nature*, 2003, **422**, 198-207.

27.   P. J. Boersema, T. T. Aye, T. A. van Veen, A. J. Heck and S. Mohammed, *Proteomics*, 2008, **8**, 4624-4632.

28.   J. L. Hsu, S. Y. Huang, J. T. Shiea, W. Y. Huang and S. H. Chen, *J. Proteome Res.*, 2005, **4**, 101-108.

29.   R. Zhang, C. S. Sioma, R. A. Thompson, L. Xiong and F. E. Regnier, *Anal. Chem.*, 2002, **74**, 3662-3669.

30.   N. Xie, K. Huang, T. Zhang, Y. Lei, R. Liu, K. Wang, S. Zhou, J. Li, J. Wu, H. Wu, C. Deng, X. Zhao, E. C. Nice and C. Huang, *J Proteomics*, 2012, **75**, 725-739.

31.   Y. Sun, W. Mi, J. Cai, W. Ying, F. Liu, H. Lu, Y. Qiao, W. Jia, X. Bi and N. Lu, *J. Proteome Res.*, 2008, **7**, 3847-3859.

32.   C. Buchanan, I. Stigliano, H. M. Garay-Malpartida, L. Rodrigues Gomes, L. Puricelli, M. C. Sogayar, E. Bal de Kier Joffe and M. G. Peters, *Breast Cancer Res. Treat.*, 2010, **119**, 559-574.

33.   A. C. Moor, *J. Photochem. Photobiol. B*, 2000, **57**, 1-13.

34.   G. Hu, Y. Wei and Y. Kang, *Clin. Cancer Res.*, 2009, **15**, 5615-5620.

35.   D. G. Breckenridge, M. Nguyen, S. Kuppig, M. Reth and G. C. Shore, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 4331-4336.

36.   L. Choe, M. D'Ascenzo, N. R. Relkin, D. Pappin, P. Ross, B. Williamson, S. Guertin, P. Pribil and K. H. Lee, *Proteomics*, 2007, **7**, 3651-3660.
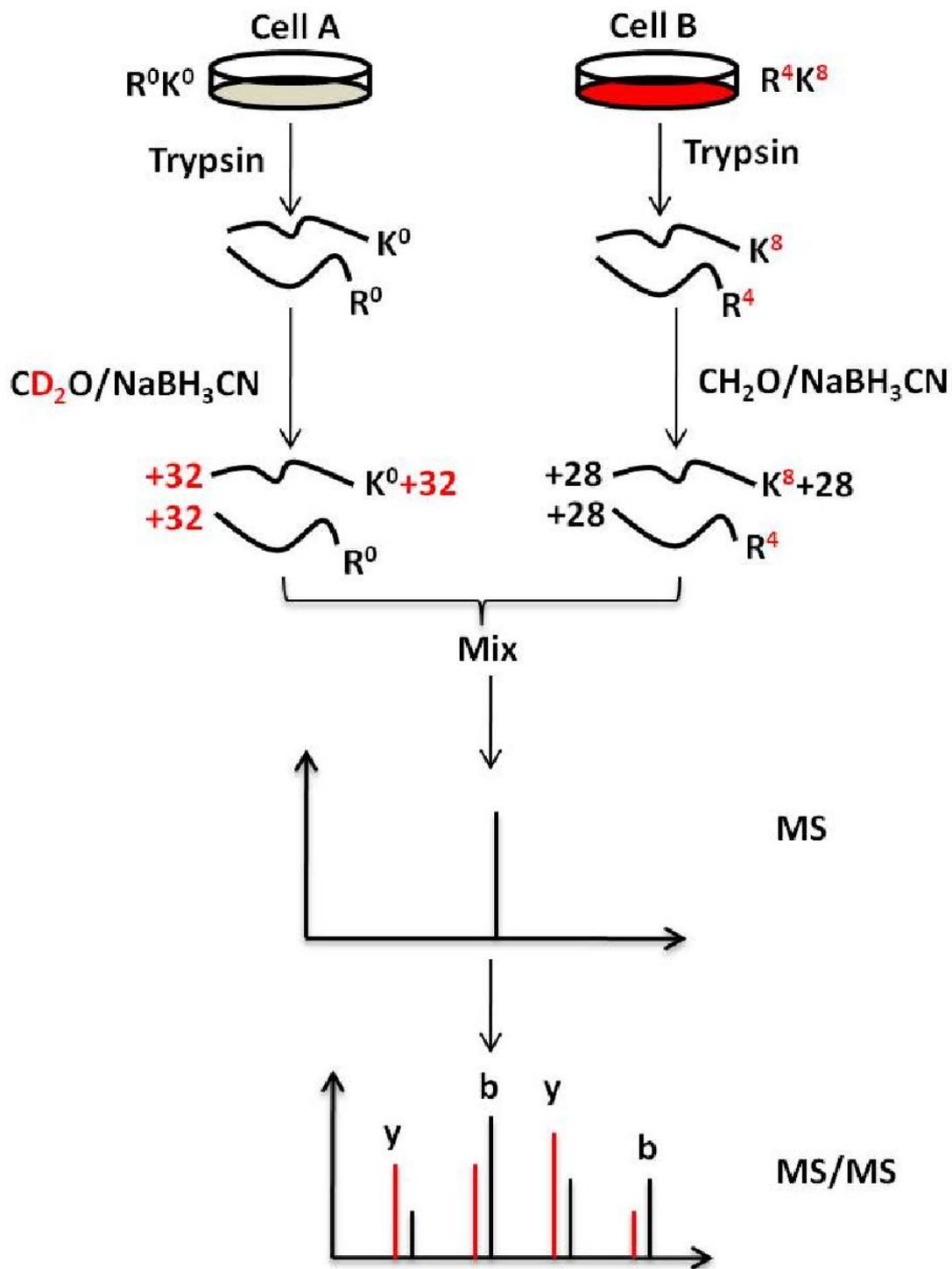
Figures
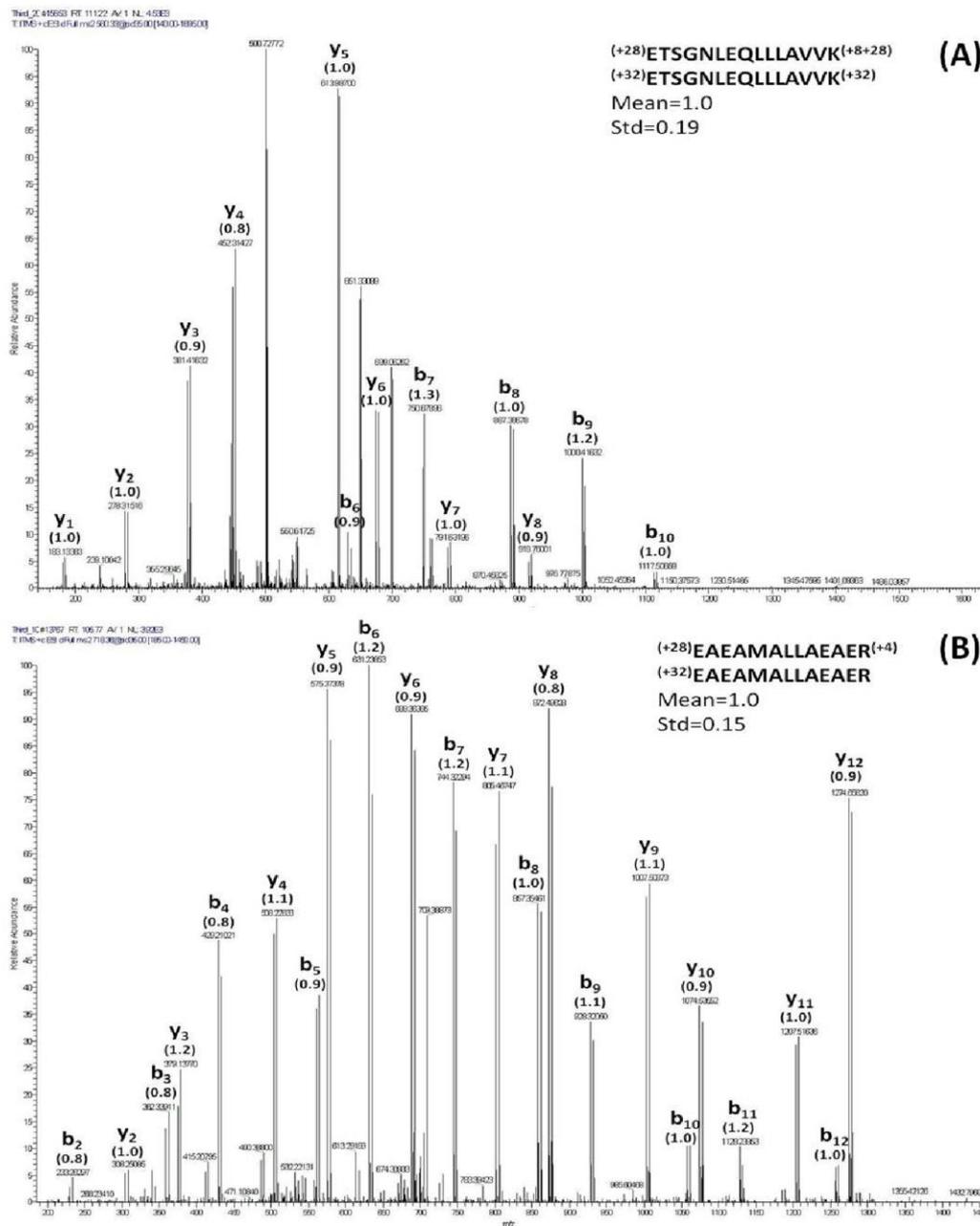

Figure. 1 The flow chart of G-IVTAL strategy.

21

Figure. 2 MS2 spectra of 1:1 mixed G-IVTAL labeled HepG2 peptides

ETSGNLEQLLLAVVK (A) and EAEAMALLAEAER (B). Their corresponding sequence,

mean quantification ratio and STD are shown in upper right corner.
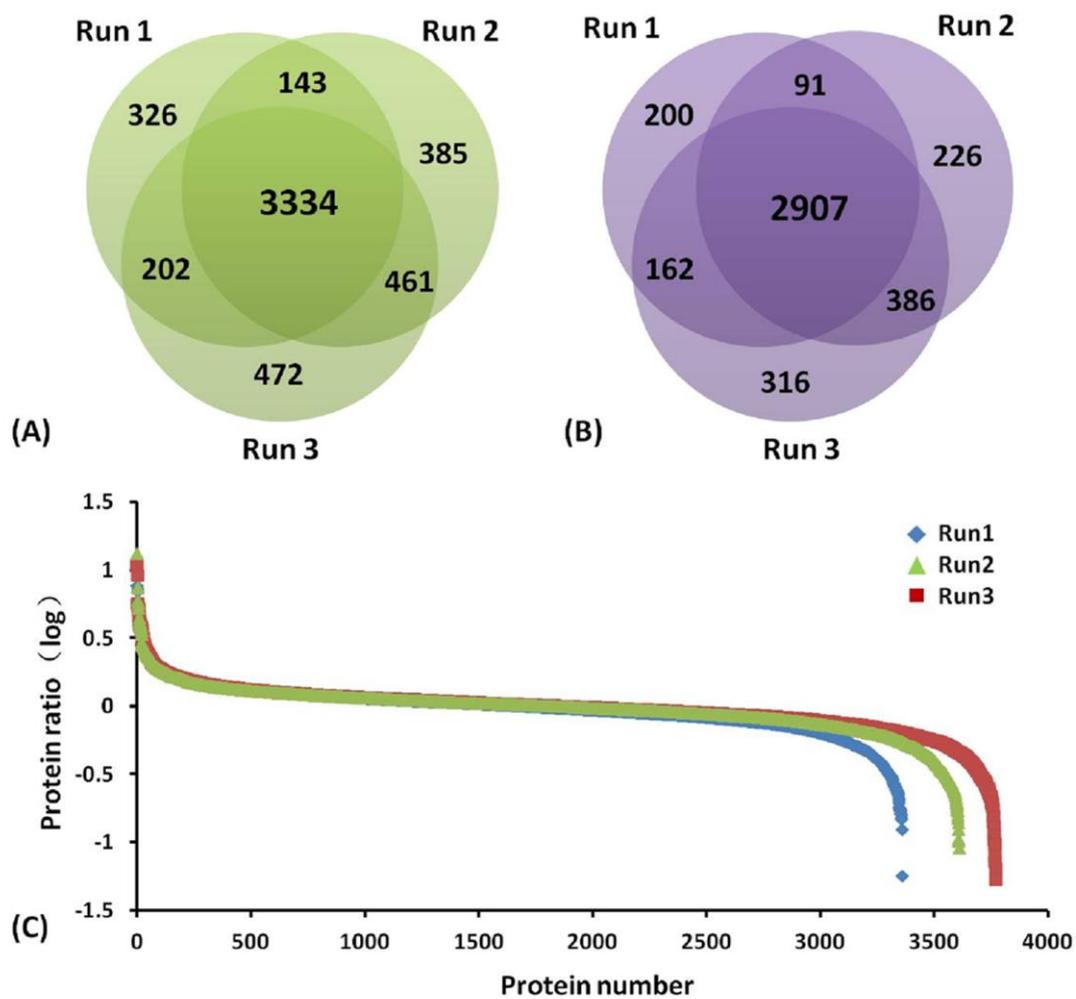
Figure. 3 The coverage of identified proteins (A), quantified proteins (B) and the distribution of quantification log ratios (C) of three replicate G-IVTAL analysis of dialyzed serum induced protein differential expression in HepG2.
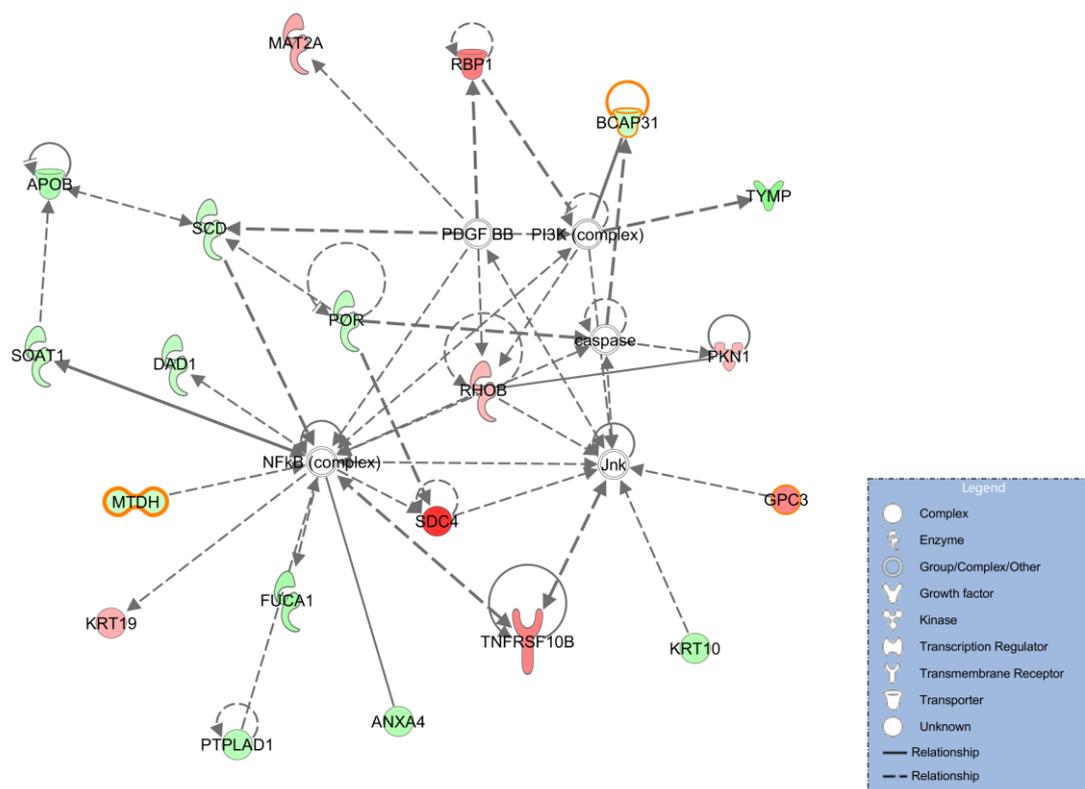
23

Figure. 4 Protein interaction network of differentially expressed proteins in DFBS cultivated
HepG2. Deregulating information of proteins is represented by their colors, red for up-
regulated, green for down-regulated, and the extent of differential deregulation is represented
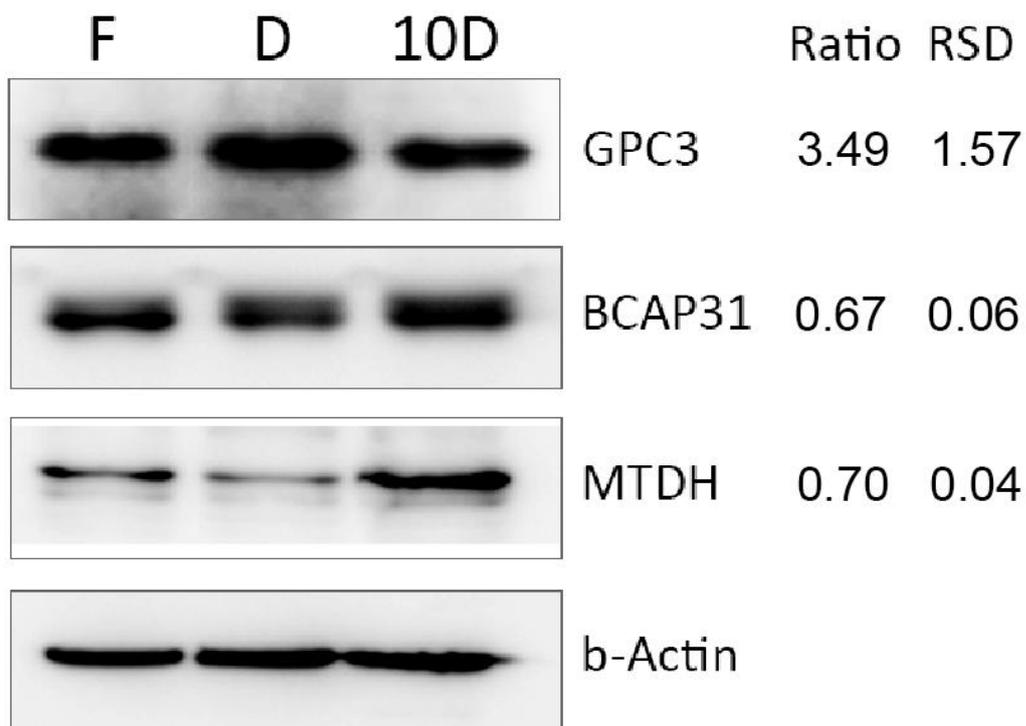by the darkness of the color.

Figure. 5. Western blotting validates the expression differences of GPC3, BCAP31 and

MTDH. HepG2 cell were divided into 3 groups, F indicates cells cultivated in FBS containing

medium, D means cells cultivated in DFBS containing medium, 10D indicates cells cultivated

in serum mixtures of DFBS and small molecules (< 10 kDa) in FBS. The right panel shows

the mean F/D ratio and RSD of MS data.