*Review*

# Revival of 2DE-LC/MS in Proteomics and Its Potential for Large-Scale Study of Human Proteoforms

Xianquan Zhan [1, 2, 3, 4]*, Na Li [1, 2, 3], Xiaohan Zhan [1, 2, 3], Shehua Qian [1, 2, 3]

[1] Key Laboratory of Cancer Proteomics of Chinese Ministry of Health, Xiangya Hospital, Central South University, 87 Xiangya Road, Changsha, Hunan 410008, China;

[2] Hunan Engineering Laboratory for Structural Biology and Drug Design, Xiangya Hospital, Central South University, 87 Xiangya Road, Changsha, Hunan 410008, China;

[3] State Local Joint Engineering Laboratory for Anticancer Drugs, Xiangya Hospital, Central South University, 87 Xiangya Road, Changsha, Hunan 410008, China;

[4] The Laboratory of Medical Genetics, Central South University, 88 Xiangya Road, Changsha, Hunan 410008, China.

**\*Corresponding Author:** Xianquan Zhan, Key Laboratory of Cancer Proteomics of Chinese Ministry of Health, Xiangya Hospital, Central South University, 87 Xiangya Road, Changsha, Hunan 410008, China. E-mail: yjzhan2011@gmail.com.

## ABSTRACT

Two-dimensional gel electrophoresis coupled with liquid chromatography/ mass spectrometry (2DE-LC/MS) is a classic and conventional approach to separate and identify proteins in a proteome. However, this approach is conventionally looked upon as a seemingly "low-throughput" technique because only one to two proteins per 2D spot are often achieved; thus, it has gradually dimmed in the field of proteomics compared to seemingly "high-throughput" bottom-up non-gel approaches such as isobaric tags for relative and absolute quantification (iTRAQ), peptide tandem mass tag (TMT), stable isotope labeling of amino acids in cell culture (SILAC), and label-free. With the rapid development and application of high-sensitivity mass spectrometry in proteomics, an average of over 50 or even hundreds of proteins can be identified in every 2D gel spot in an analysis of a complex human proteome, mostly low-abundance proteins, and 2DE-LC/MS can detect the protein species, to break through the conventional concept of 2DE-LC/MS, and assist its revival in the field of proteomics. Splicing and post-translational modifications are fundamental to the proteome, and are the main factors to clarify proteoforms, which enrich the concept

of the proteome. The top-down and high-throughput nature of stable isotope-labeled 2DE-LC/MS for the detection, identification, and quantification of proteoforms provides a solid methodological support for the large-scale study of human proteoforms and disease-related proteoforms to clarify mechanisms of a disease and to discover reliable biomarkers for the prediction, diagnosis, and prognostic assessment of a disease.

# 1 INTRODUCTION

Two-dimensional gel electrophoresis (2DGE) coupled with mass spectrometry (2DE-MS) is a classic proteoform separation and identification approach in the field of proteomics [1–3]. Here, MS commonly includes matrix-assisted laser/desorption ionization-MS peptide mass fingerprinting (MALDI-MS PMF) [4] and -tandem mass spectrometry (MALDI-MS/MS) [5], and liquid chromatography (LC)-electrospray ionization (ESI)-MS/MS (LC-ESI-MS/MS; LC-MS/MS; LC/MS) [6]. There were thousands of 2DE-based publications before and after the concepts of proteome and proteomics appeared [7]. Almost all these publications described 2DE as a very high-resolution separation technique with only one or two proteins detected in each 2D spot in a proteome analysis. 2DE-separated proteins commonly identified with MALDI-MS PMF, MALDI-MS/MS, or LC-MS/MS with determination of the first- or second-ranked proteins in the list of searched proteins as the result proteins in a 2D spot [8–13]. These types of 2DE-based data are collected in many 2DGE resources such as the World-2DPAGE Constellation (http://world-2dpage.expasy.org) [14], which include the World-2DPAGE list (http://world-2dpage.expasy.org/list)—the most complete 2DPAGE list that contains nearly 400 gel images and 60 databases, World-2DPAGE portal (http://world-2dpage.expasy.org/portal)—the biggest gel-based proteomics dataset that contains over 250 maps of 23 species and nearly 40,000 identified 2D gel spots, World-2DPAGE Repository (http://world-2dpage.expasy.org/repository), SWISS-2DPAGE (http://world-2dpage.expasy.org/swiss-2dpage), and 2D-PAGE at MPIIB (http://www.mpiib-berlin.mpg.de/2D-PAGE/). Thus, if maximizing the total number of identified proteins in a proteome is the objective, then 2DE-MS is commonly seen as a seemingly "low-throughput" approach, compared to seemingly "high-throughput" bottom-up non-gel methods [15–17] such as isobaric tags for relative and absolute quantification (iTRAQ) [18,19], peptide tandem mass tags (TMT) [20,21], stable isotope labeling of amino acids in cell culture (SILAC) [22,23], or label free [24,25] based quantitative proteomics. When one studies very precious, hard-to-obtain human brain samples such as, for example, pituitary tissue, then 2D gels provide a crucially important method to collect, store, and archive that human proteome [3,6,10–12]. 2DLC-MS methods can never provide that necessary archival storage advantage. Of course, one can perform 2DGE in parallel, up to 12 gels (usually 3 replicas of 4 samples) at one time in less than 2 days; whereas bottom-up non-gel approach 2DLC-MS/MS cannot be run in parallel, for the same 3 replicas of 4 samples, it would take 12 × 3 hours (36 hours). Thus, two methods don't have difference in throughput if the objective is to quantify the changes of protein abundance. A recent study discovered peptides that belong to an average of over 50 or even hundreds of open-reading frames (ORFs) in every 2D gel spot when 2DE was coupled with high-sensitivity MS [26] to reach at least 500,000 ORF species identified from 10,000 spots on a 30 cm × 40 cm 2D gel [26–28], with most of ORFs identified in a spot of low- or extremely low-abundance. This discovery challenges the conventional concept of 2DE as a very high-resolution approach with only one or two proteins in each spot since it was first established in 1975 [1]. Moreover, this new 2DE-LC/MS method can identify the same gene-derived proteoforms [29] across different 2DE spots based on the isoelectric point ($pI$) and relative molecular mass ($M_r$) of each proteoform. It holds huge potential for a large-scale study of human proteoforms. Here, one must note that "spot" means a visualized spot in a 2DE map, which is actually a single very tiny spot that is clearly resolved from other detectable spots, or a spot that is part of a tight and overlapping group of spots, and that 2DE means two-dimensional gel electrophoresis.
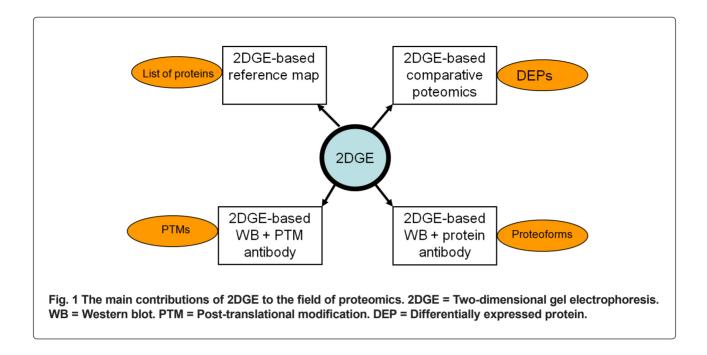
# 2 THE TRADITIONAL HIGH-RESOLUTION AND PRESUMABLY "LOW-THROUGHPUT" 2DE-MALDI/LC-MS IN THE FIELD OF PROTEOMICS

2DGE-MS has extensively used in the field of proteomics as described above. 2DGE was used as the main proteomic separation technique due to three main factors: (i) The commercial immobilized

pH gradient gel strip (IPG Strip) significantly improved the reproducibility in the first dimension—isoelectric focusing (IEF) [2]. Also, a series of wide-, medium- and narrow-ranges of IPG strips are commercially available. For example, the pH ranges 3–10, 3–7, 4–7, 6–9, 6–11, 3.5–4.5, 4–5, 4.5–5.5, 5–6, and 5.5–6.7 IPG strips with linear or nonlinear pH gradients [2,30] significantly improve the separation capability in a limited range of proteins to maximize the coverage (the total number of identified proteins) of a proteome. (ii) The vertical wet multiple homogenous-concentration-gel system (up to 12 gels at one time) has been continuously used for the second dimension SDS-PAGE separation. Although the horizontal semi-dry single precast gradient-concentration-gel system is occasionally used as the second dimension SDS-PAGE separation [31–37], the reproducibility is much better for the vertical multiple homogenous-concentration-gel system compared to the horizontal single precast gradient-concentration-gel system [31,32]. Thus, the horizontal single gradient-concentration-gel system was not commonly used or is currently seldom used for the second dimension electrophoresis. However, the HPE-FlatTop Tower horizontal electrophoresis system (Serva GmbH) can use multiple precast homogenous-concentration-gels (up to four gels at one time) to improve the reproducibility of horizontal electrophoresis for second dimension SDS-PAGE electrophoresis. (iii) The tryptic peptides produced from in-gel trypsin digestion of 2DGE-resolved proteins are compatible with soft ionization MS, including MALDI-MS and ESI-LC-MS, to characterize proteins in a gel spot [4–6]. Here, in fact, MS is most often used to identify the products of ORFs, but not proteins or proteoforms. However, the same ORF can occur at different places in the gel because there is a distinct proteoform with a specific $pI$ and $M_r$ derived from the same ORF. 2DE in combination with MS offers the capability to detect and identify proteoforms. An alternative type of 2DGE method, 2D difference in-gel electrophoresis (2D DIGE), can multiplex samples but does not result in more spots or improved resolution. For 2D DIGE, proteins are labeled with fluorescence dyes and then the labeled protein samples are mixed equally for 2DGE [38–40]. Although protein overlap in 2D DIGE is a serious practical problem, it uses 100× less amount of protein sample compared to Coomassie blue staining [41], makes spot-matching much easier and protein quantification more accurate, and significantly enhances the reproducibility and sensitivity of protein detection relative to the classic 2DGE [42]. Here, one must realize that, although 2DGE historically only detected one or two proteins

per spot, it is to determine which proteoforms change in abundance between treatments, and only those interested in 'stamp collecting' spent the effort to identify the ORF products in every detectable spot on the gel. This is the important difference between quantifying *before* identifying for the approach of 2DGE-MS, and quantifying *after* identifying for 2DLC-MS/MS methods.

In review of all those 2DE publications, four main contributions of 2DGE in the field of proteomics in the past years (Fig. 1) are: (i) 2DGE-based comparative proteomics that identified differentially expressed proteins (DEPs) in a given condition relative to its control [43,44], (ii) 2DGE-based reference map that established the database of a proteome [45,46], (iii) 2DGE-based Western blot in combination with a specific protein antibody that visually detected variants/proteoforms of that given protein in a proteome [47,48], and (iv) 2DGE-based Western blot in combination with a specific post-translational modification (PTM) antibody that visually identified a kind of PTM in a proteome [49,50], and relatively quantified a differential level of a protein in a given condition relative to its control [51,52]. Furthermore, in the past 10–15 years, 2DE protocols have been highly refined in the analysis of a proteome with the use of postfractionation strategy [53], the change of the detergent composition of the solubilization buffer [54], or the pre-extraction of sample handling with automated frozen disruption [55] to resolve those traditional 2DE problems such as proteins stacked at pH extremes, the 2D-gel-area obscured by high-abundance proteins, unresolved peptides that migrated at the front of separations, and hydrophobic membrane proteomes [53–55]. In addition, the 3rd electrophoretic separation-based post-fractionation strategy [53] clearly showed multiple proteins per 'spot' for some 2DE spots. The routine 'Deep Imaging' approach also confirmed that 2DE can resolve and detect even low-abundance species with 2DE [56]. The newer formulation and stain/wash protocol clearly showed that Coomassie was the 'gold-standard' for in-gel protein detection, with a genuine sensitivity in the femto-to-subfemtomole range to provide in-gel detection of intact proteoforms at the same level of sensitivity that routine MS detects peptides [57]. Therefore, these refined 2DE protocols not only improved the resolution and detection of proteins/proteoforms with 2DE, but substantially also improved the quality of the subsequent MS analyses. It emphasizes the power of the top-down, 2DE-LC/MS approach relative to the popular bottom-up approaches to proteomics.

**Fig. 1 The main contributions of 2DGE to the field of proteomics. 2DGE = Two-dimensional gel electrophoresis. WB = Western blot. PTM = Post-translational modification. DEP = Differentially expressed protein.**

In order to promote non-gel methods such as 2DLC-MS/MS protocols to the field of proteomics, 2DGE and 2D DIGE were commonly claimed as being very time-consuming, labor-intense, and low-throughput. They were also claimed to be inadequate to separate well the extremely high- or low-mass proteins and the extremely acidic- or basic-proteins [3], identify low-abundance proteins, and distinguish co-migrated or overlapped proteins with very similar $pI$ and $M_r$ values in one single spot [26,50,58]. However, those 2DE problems were overcome with the refined 2DE protocols described above [53–57] to maximize the coverage (the total number of proteins) of a proteome. In contrast, the 2DLC procedure basically includes enzymatic peptides from a complex proteome, the first-dimensional LC separation, second-dimensional RP-LC separation, online inputting the separated enzymatic peptides into a mass spectrometer for MS/MS analysis, followed by protein identification with database search [16,17]. 2DLC-MS/MS proteomics was developed rapidly, including (i) stable isotope-labeled 2DLC-MS/MS such as SILAC [22,23], TMT [20,21], iTRAQ [18,19], ICAT (Isotope-Coded Affinity Tags) [59–61], IPTL (Isobaric Peptide Termini Labeling) [62,63], ICPL (Isotope-Coded Protein Labeling) [64,65], and 18O [66,67], and (ii) non-labeled 2DLC-MS/MS such as label-free [24,25], SWATH (Sequential Window Acquisition Of All Theoretical Spectra) [68,69], AQUA (Absolute Quantification) [70,71], and SRM/MRM (Selected or Multiple Reaction Monitoring) [72,73]. These 2DLC-MS/MS methods are currently being extensively used in the field of proteomics because of their high-sensitivity, high-accuracy, and seemingly "high-throughput" in the analysis of a proteome. However, considering the real depth of the proteome, in particular in terms of assessing proteoforms rather

than assuming the presence of certain proteins based on one or two peptides, then 2DE-LC/MS is actually superior to 2DLC-MS/MS in the detection and identification of proteoforms.

Moreover, one must note that 2DLC is extensively used to separate the enzymatic peptide mixture, whereas 2DGE is extensively used to separate and visualize the intact protein mixture with its own advantages in visualization of the protein components of a proteome, and detection of protein variants or proteoforms that are mainly derived from splicing and PTMs [47–52]. Protein variants and proteoforms are very important issues because different variants or proteoforms of a given protein are associated with a given condition, such as a corresponding pathophysiological status, which plays important roles in multiple biological processes [74–76]. 2DLC is obviously limited in the detection and identification of protein variants and proteoforms of a given protein. Therefore, 2DE-LC/MS and 2DLC-MS/MS are two types of important and complementary approaches in the field of proteomics that cannot be replaced by each other, which has been discussed in-depth [41,77–80], and the following three approaches should flourish: 2DLC-MS/MS for rapid proteome scanning and cataloging, 2DE-MS for in-depth detailed analysis of a proteome, and 2DE-MS in combination with 2DLC-MS/MS to achieve the most definitive analysis of a proteome [41,78,80]. Thus, 2DE is still a powerful technique in the field of proteomics.

Simultaneously, although many scientists still claim that 2DE-MS and 2DLC-MS are both important in proteomics, both techniques are not mature and need further improvement in many aspects [41,78]. The throughput, namely the number of identified proteins

in a 2D spot or in an entire 2D gel, might be an important aspect that needs to be improved. Although a few publications showed that several proteins were contained in some 2D gel spots [8,9,28,53,81], most 2D spots in a 2D gel map are still identified to contain only one or two proteins in a spot. Thus, even though 2DE is recognized as an important technique in proteomics, the number of identified proteins is still small according to the traditional concept of one to two proteins in most 2D gel spots in a 2DE map. However, our recent study [26] discovered an average of fifty or even hundreds of proteins in every 2D gel spot in a human 2D map, significantly breaking through the traditional concept and improving the throughput of identified proteins in the 2DE analysis of a proteome.

Furthermore, multiple 2D gels can be carried out in parallel, whereas 2DLC methods rely on single sequential analysis on any given mass spectrometer. Therefore, although 2DLC-MS/MS actually has rather low throughput for even the simplest of studies, it is often referred to as 'high throughput' by the authors because they simply do not routinely do technical replications as it would then take about a week (or more) to carry out a full analysis of even the simplest of experiments (i.e. 3 test conditions and 3 controls). If this concept was further refined to include independent analysis of the soluble and membrane proteomes, with full technical replications—as is easily done in parallel with 2DE—then the 2DLC approach becomes even more low throughput and generally would not handle hydrophobic proteins very well [54,55].

# 3 THE CURRENT HIGH-THROUGHPUT AND REVIVED 2DE-LC-MS IN THE FIELD OF PROTEOMICS

Separation and identification are two key techniques used to maximize the coverage (the total number of proteins) of a complex proteome. Separation techniques such as 2DGE and 2DLC are mainly performed to simplify the components of a complex proteome before identification. MS is the core technology used to identify a protein. However, each MS instrument has its own sensitivity level that significantly affects the number of identified proteins from complex proteomes [26]. For example, in our laboratory, the sensitivity is at the level of 10–100 fmol for the old ESI-QTOF and MALDI-TOF-TOF, whereas it is at the level of 1–10 amol for the new OrbiTrap Velos, to identify endogenous proteins from a human proteome [26]. The higher sensitivity OrbiTrap Velos dramatically increases the capability to identify absolutely low-abundance

proteins in a complex human proteome to benefit the identification of low-abundance proteins in a 2DE gel spot in the separation of complex proteomes. High-throughput capability is an important hallmark of a proteomic approach. There are three strategies to increase throughput: the first is to significantly simplify the complexity of a protein sample; the second is to improve the sensitivity of a MS system; and the third is to reduce the LC/MS cycle time - which is why some researchers are moving back to microflow rather than nanoflow so that the cycle time can be drastically reduced at the expense of some sensitivity. Until now, 2DGE is the only one method to maximally simplify the complexity of a complex proteome sample (over one thousand spots; or over several thousands of gel pieces if cutting the entire gel by 3 × 3 mm grid) relative to 2DLC (commonly 20–30 fractions). The new mass spectrometer such as OrbiTrap Velos has its sensitivity at 1–10 amol, which can analyze the very low-abundance proteins in a complex proteome.

Recently, we integrated 2DGE and high-sensitivity MS to detect and identify the protein components in a complex human proteome [26]. As a result, over fifty or even up to several hundreds of proteins were identified in every 2D gel spot in an analysis of proteomes from complex human pituitary adenoma tissues (Fig. 2) and glioblastoma tissues (Fig. 3). Also, Fig. 2 shows that the number of identified proteins increased with the pooled two spots and three spots compared to the single spot under the identification criteria of two unique peptides. It also showed the reproducibility of the number of identified proteins in the same spots from those replicate gels (spots 1, 2, 3, 5 and 16 versus spots 1*, 2*, 3*, 5* and 16*). Fig. 3 shows the number of identified proteins under the different identification criteria of 1, 2, or 3 unique peptides for each spot. These results broke through the conventional opinion that only one or two proteins per spot are identified for most of the 2D gel spots in a 2DE map, although a few previous publications found that some 2D gel spots contained several proteins [8,9,28,53,81]. These findings challenge the use of 2DGE-based comparative proteomics to determine a differentially expressed protein (DEP) with the difference in spot volume between treatments, because if over fifty or hundreds of proteins are identified, then which one is the real DEP on earth. More importantly, most proteins identified in a 2D gel spot were extremely low-abundance or low-abundance proteins (Table 1), and many proteins were present in several of the analyzed spots to demonstrate the capability of 2DE-MS to separate proteoforms (Table 2). Furthermore, multiple proteins in a 2D gel spot confirm the necessity of isotopic labeling in a large-scale quantification of different proteoforms in a proteome.

**Table 1. The number of proteins in each emPAI range, which is the estimation of the ratio of each protein with at least 2 unique peptides identified in the analyzed glioblastoma 2-DE spot with OrbiTrap Velos MS/MS.**

| Spot No. | The number of proteins by emPAI range | | | | | Total (*n*) |
|---|---|---|---|---|---|---|
| | 0.01–0.1 (*n*) | 0.1–1 (*n*) | 1–10 (*n*) | 10–100 (*n*) | >100 (*n*) | |
| 3 | 146 | 105 | 28 | 3 | 7 | 289 |
| 7 | 122 | 105 | 19 | 6 | 2 | 254 |
| 8 | 124 | 130 | 22 | 2 | 2 | 280 |
| 16 | 97 | 89 | 32 | 8 | - | 226 |
| 22 | 81 | 122 | 27 | 3 | 1 | 234 |
| 31 | 129 | 125 | 36 | 4 | - | 294 |
| 46 | 62 | 106 | 19 | 4 | 1 | 192 |
| 48 | 68 | 132 | 29 | 1 | - | 230 |
| 49 | 94 | 139 | 29 | 2 | 3 | 267 |
| 50 | 85 | 126 | 40 | 1 | - | 252 |
| 52 | 51 | 99 | 33 | - | 1 | 184 |
| 53 | 47 | 61 | 12 | - | - | 120 |
| 58 | 71 | 129 | 31 | 4 | 2 | 237 |
| 60 | 62 | 60 | 6 | 4 | 2 | 134 |
| 61 | 136 | 116 | 28 | 6 | 1 | 287 |
| 63 | 97 | 53 | 24 | 2 | 1 | 177 |
| 64 | 130 | 93 | 18 | 1 | - | 242 |
| 65 | 150 | 85 | 19 | 1 | - | 255 |

Note: Each spot was analyzed with combined 3 matched gel spots. Each identified protein had at least 2 unique peptides identified. The emPAI is exponentially modified protein abundance index (emPAI), which was calculated for each protein in an analyzed spot to estimate the amount of each protein in a 2-DE spot. emPAI is equal to 10PAI minus one. The protein abundance index (PAI) is the number of identified tryptic peptides divided by the number of theoretically observable tryptic peptides. Modified from Zhan X, *et al.* (2018)[26], with permission from Wiley-VCH, © 2018.

**Table 2. Protein speciation recognized among 44 randomly picked spots in the glioblastoma 2-DE map and 10 randomly picked spots in the pituitary adenoma 2-DE map.**

| Pituitary adenoma (Total 10 spots analyzed) | | Glioblastoma (Total 44 spots analyzed) | |
|---|---|---|---|
| The number of proteins | Speciation (how many spots occurred) | The number of proteins | Speciation (how many spots occurred) |
| 2 | 10 | 1 | 21 |
| 1 | 9 | 1 | 17 |
| 2 | 8 | 2 | 16 |
| 6 | 7 | 3 | 13 |
| 6 | 6 | 5 | 12 |
| 9 | 5 | 3 | 11 |
| 13 | 4 | 8 | 10 |
| 30 | 3 | 10 | 9 |
| 88 | 2 | 13 | 8 |
| 604 | 1 | 18 | 7 |
| | | 24 | 6 |
| | | 41 | 5 |
| | | 87 | 4 |
| | | 150 | 3 |
| | | 436 | 2 |
| | | 1690 | 1 |
| Total: 761 | | Total: 2492 | |

Note: Modified from Zhan X, *et al.* (2018)[26], with permission from Wiley-VCH, © 2018.

| | MALDI-MS/MS | | | LC-ESI-MS/MS | | |
|---|---|---|---|---|---|---|
| | 1 spot | 2 pooled spots | 3 pooled spots | 1 spot | 2 pooled spots | 3 pooled spots |
| | 2 (1) | 1 (1*) | 2 (12) | 76 (1) | 50 (1*) | 102 (12) |
| | 1 (2) | 1 (2*) | 0 (13) | 32 (2) | 48 (2*) | 91 (13) |
| | 0 (3) | 2 (3*) | 1 (14) | 60 (3) | 93 (3*) | 141 (14) |
| | 1 (5) | 1 (5*) | 2 (15) | 86 (5) | 78 (5*) | 156 (15) |
| | 1 (16) | 1 (16*) | 1 (18) | 61 (16) | 84 (16*) | 98 (18) |
| | Averaged 1 /spot | Averaged 1 /spot | Averaged 1 /spot | Averaged 63 /spot | Averaged 71 /spot | Averaged 118 /spot |

Note: The number in the bracket is the spot number labeled in Figure 2. Spots 1*, 2*, 3*, 5*, and 16* matched to the corresponding spots 1, 2, 3, 5, and 16.

**Fig. 2 Coomassie blue-stained 2DE pattern of a pituitary adenoma proteome analyzed with IPGstrip pH 3-10 NL and 12% gel concentration of SDS-PAGE. The labeled spots were analyzed with MALDI-TOF-TOF MS/MS and LC-ESI-MS/MS. Spots 1*, 2*, 3*, 5*, and 16* were matched to the corresponding spots 1, 2, 3, 5, and 16, and pooled from 2 matched gels. Modified from Zhan X, *et al.* (2018) [26], with permission from Wiley-VCH.**



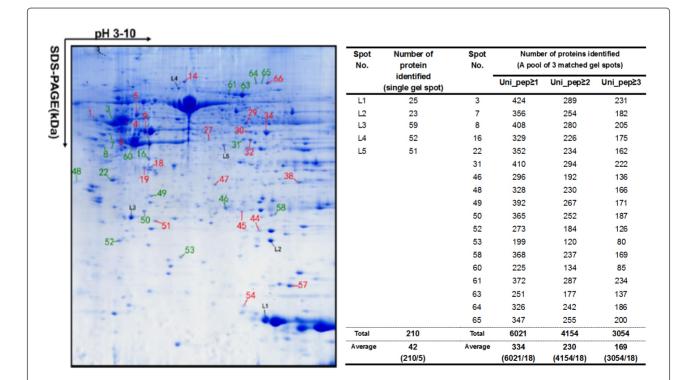| Spot No. | Number of protein identified (single gel spot) | Spot No. | Number of proteins identified (A pool of 3 matched gel spots) | | |
|---|---|---|---|---|---|
| | | | Uni_pep≥1 | Uni_pep≥2 | Uni_pep≥3 |
| L1 | 25 | 3 | 424 | 289 | 231 |
| L2 | 23 | 7 | 356 | 254 | 182 |
| L3 | 59 | 8 | 408 | 280 | 205 |
| L4 | 52 | 16 | 329 | 226 | 175 |
| L5 | 51 | 22 | 352 | 234 | 162 |
| | | 31 | 410 | 294 | 222 |
| | | 46 | 296 | 192 | 136 |
| | | 48 | 328 | 230 | 166 |
| | | 49 | 392 | 267 | 171 |
| | | 50 | 365 | 252 | 187 |
| | | 52 | 273 | 184 | 126 |
| | | 53 | 199 | 120 | 80 |
| | | 58 | 368 | 237 | 169 |
| | | 60 | 225 | 134 | 85 |
| | | 61 | 372 | 287 | 234 |
| | | 63 | 251 | 177 | 137 |
| | | 64 | 326 | 242 | 186 |
| | | 65 | 347 | 255 | 200 |
| Total | 210 | Total | 6021 | 4154 | 3054 |
| Average | 42 (210/5) | Average | 334 (6021/18) | 230 (4154/18) | 169 (3054/18) |

**Fig. 3 Coomassie blue-stained 2DE pattern of a glioblastoma proteome analyzed with IPGstrip pH 3-10 NL and 12% gel concentration of SDS-PAGE. Spots L1-L5 came from one gel for MS/MS analysis. Each red- or green-number labeled spot was combined from three matched gel spots for MS/MS analysis. Modified from Zhan X, *et al.* (2018) [26], with permission from Wiley-VCH.**

According to our result with an average of over fifty proteins per spot for an analysis of glioblastoma and pituitary adenoma tissue proteomes, then one could speculate that *ca.* 500,000 different proteoforms can be quantified with SILAC-2DE-LC/MS with a resolution power of 10,000 spots within a 30 × 40 cm large gel [26–29]. Whereas bottom-up 2DLC-MS/MS can only identify a maximum of 20,000 proteins, which is limited by the human genome, then it is impossible to reach the level of proteoforms and to identify extremely low-abundance proteins. Therefore, stable isotope-labeled 2DE-LC/MS shows its strong power to detect, identify, and quantify proteoforms in a human proteome with super high throughput and resolution capability, to assist its revival in the field of proteomics. It means that stable isotope-labeled 2DE-LC/MS will not only seek maximum proteome coverage (the total number of proteins in a proteome), but will also resolve multiple proteoforms derived from the same gene in a gel or multiple proteoforms derived from different genes at a single spot (pixel) in the gel after analysis of the entire gel by the grid including the parts where there is no visible spot.

# 4 THE IMPORTANCE OF PROTEOFORMS IN A PROTEOME

The human genome sequence has been completed [82] and has driven researchers to move to the era of functional genomics, namely transcriptomics and proteomics, to study transcriptomes and proteomes, respectively. About 20,300 genes have been found in the human genome. However, the transcriptome is more complex than the genome because RNA splicing occurs when a gene is transcribed to RNA; thus, one gene often corresponds to multiple transcripts [83–85]. The human transcriptome is estimated to have 100,000 transcripts [29,86]. Each transcript guides the ribosome to synthesize a protein amino acid sequence. The synthesized protein in the ribosome has no functional role. It must translocate and re-distribute to the corresponding locations such as the plasma membrane, mitochondrion, endoplasmic reticulum, Golgi apparatus, *etc.*, to form a special conformation, and to interact with surrounding molecules to exert its biological roles. There are numerous protein PTMs [86,87] and even unknown factors that modify proteins in the process of translocation and re-distribution. It is speculated that about 400–600 PTMs, which are the main factors that cause the complexity and diversity of proteins (namely protein species or proteoforms) [29,86,88,89], exist in the human body. Thus, one transcript corresponds to multiple proteoforms. It is estimated
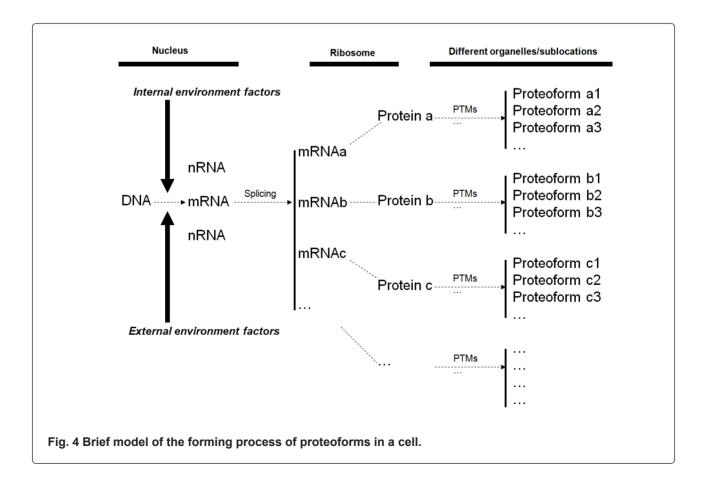
that there are about 1,000,000 proteoforms in a given condition [29]. Each proteoform has its own copy number. A proteoform is the final functional performer of a gene. Fig. 4 briefly shows the forming process of proteoforms in a cell. For example, the human growth hormone (hGH) protein is derived from the hGH gene. However, four hGH splicing variants were found, and together with different PTMs such as phosphorylation, glycosylation, and deamination, 24 hGH proteoforms with different $pI$ and $M_r$ were identified within the 2DE map of the pituitary proteome [74]. Different hGH proteoforms such as 20 KDa and 22 KDa hGH exhibit different intracellular signaling profiles and properties [90]. Another example is the human prolactin (hPRL) that is derived from the hPRL gene. hPRL does not have existing splicing variants; however, six hPRL proteoforms with different $pI$ and $M_r$ were found in the 2DE map of the pituitary proteome due to different PTMs such as glycosylation, phosphorylation, and deamination. In addition, the patterns of the six hPRL proteoforms were changed among five types of nonfunctional pituitary adenomas, and the different hPRL proteoforms acted on different short or long PRL receptor signaling pathways [91]. Also, 59 proteoforms from the HSP27 gene in human myocardium [92], 52 proteoforms from the HSP70 gene, 24 proteoforms from the gamma-enolase-2 gene, and 17 proteoforms from the lactate dehydrogenase 2B gene in the mouse brain proteome [93] were identified with 2DE-MS analysis, respectively. Therefore, proteoforms enrich the concept and content of a proteome. Studies on proteoforms provide much more in-depth insight into a proteome, and directly result in the determination of reliable biomarkers for the understanding of accurate molecular mechanisms, the discovery of effective therapeutic targets, and for effective prediction, diagnosis, and prognostic assessment. It emphasized the important scientific merits of proteoform study.
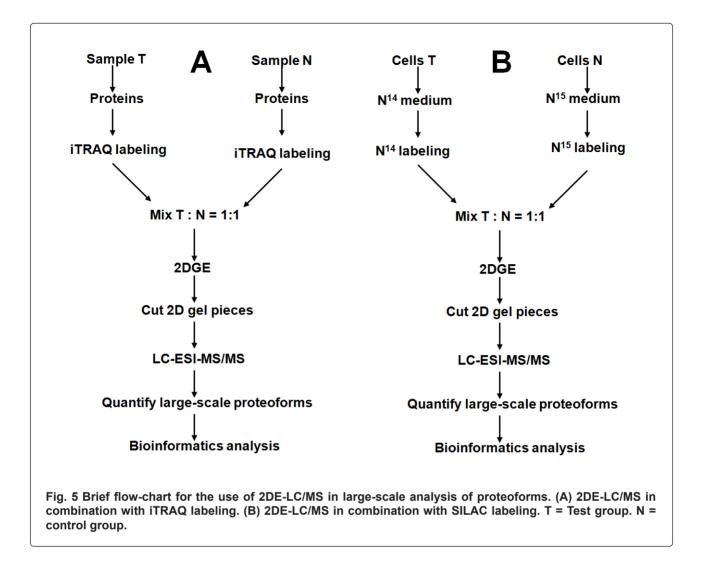
# 5 THE POTENTIAL OF CURRENT 2DE-LC/MS IN A LARGE-SCALE ANALYSIS OF PROTEOFORMS IN PROTEOMES

Detection, identification, and quantification of proteoforms in a proteome are essential to clarify their biological significance in an organism. Detection includes gel and gel-free methods [74,75]. Gel-based methods primarily include 1DGE, 2DGE, and 2D DIGE [48,74]. These gel-based methods, commonly coupled with a specific antibody, detect a given PTM [15], or a kind of proteoform of a given protein [48]. Gel-free methods include multiplexed gel-eluted liquid fraction entrapment electrophoresis (mGELFrEE; size-based

separation) with 8 parallel glass gel columns [94], C4 or C5 reverse-phase LC (RPLC) with 300 Å pore-size particles [95], capillary electrophoresis (CE)-ESI-MS [96], hydrophobic interaction chromatography (HIC) to separate large biomolecules such as proteins [97,98], and weak-cation exchange chromatography (WCX) in combination with HIC in a single column with a single phase (2D-LC; from WCX to HIC mode) [99]. MS is the vital technique used to identify PTMs and proteoforms because MS/MS can identify the modified sites and the amino acid sequence of a proteoform [48,74]. Also, MS coupled with isotopic labeling such as iTRAQ, TMT, and SILAC, or with non-isotopic labeling such as label-free, PRM, and SWATH can quantify a proteoform between two given conditions [28]. However, until now, all those methods are limited in any large-scale study of human proteoforms because their low-sensitivity and low-throughput. A high-throughput method is needed for human proteoform studies.

Our recently used 2DGE coupled with high-sensitivity MS and isotopic labeling can exactly meet these urgent needs [26]. First, a large format 2D gel can effectively array the extremely complex proteoforms in a proteome. For example, 10,000 spots can be achieved within a 30 cm × 40 cm 2D gel [27]. Second, a high-sensitivity MS can identify at least fifty proteins or even several hundreds of proteins in a 2D spot [26]; thus, at least 5,000,000

proteoforms can be achieved, and most of them are low-abundance. Third, when isotopic labeling such as iTRAQ, SILAC, or TMT is applied in the protein sample before 2DGE [28], the proteoform in a 2D gel spot can be quantified and compared between two different conditions (Fig. 5). Fig. 5 shows the brief experimental flow-chart for the use of 2DE-LC/MS in combination with isotopic labeling for a large-scale quantitative analysis of proteoforms. Protocol A—2DE-LC/MS in combination with iTRAQ labeling (Fig. 5A)—is suitable for the analysis of the extracted proteins from different treatments such as cancer and control tissues. Protocol B—2DE-LC/MS in combination with SILAC labeling (Fig. 5B)—is suitable for the cultured cells from different treatments such as the cultured cancer and control cells, or before and after a specific treatment of a cancer cell. After the labeled proteomic sample is arrayed with 2DE, the entire gel will be MS-analyzed by cutting the gel into gel pieces by grid. The stable isotope-labeled reporter ion intensities can be used to determine the difference of a proteoform between two given conditions. Furthermore, if those refined 2DE protocols [53–57] are effectively used in the protocols A and B in the Fig. 5, then it might further improve the resolution and detection of human proteoforms. Therefore, 2DE-LC/MS [26] in combination with isotopic labeling [28] has huge potential for large-scale analyses of proteoforms in proteomes.



**Fig. 4 Brief model of the forming process of proteoforms in a cell.**

**Fig. 5 Brief flow-chart for the use of 2DE-LC/MS in large-scale analysis of proteoforms. (A) 2DE-LC/MS in combination with iTRAQ labeling. (B) 2DE-LC/MS in combination with SILAC labeling. T = Test group. N = control group.**

Here, one must also be aware of the possible bias and false-positive results introduced with isotope-labeled 2DE-LC/MS as discussed in detail[26], including high protein-loading amount on the gels and pooling the matched gel spots for MS analysis, the carryover effects/memory effects of LC between two samples, and the readout effects of the very sensitive MS identifications. It would be necessary to refine the isotope-labeled 2DE-LC/MS protocol in the future studies, and corroborate the results of isotope-labeled 2DE-LC/MS with other methods. However, these seemingly negative aspects of stable isotope-labeled 2DE-LC/MS cannot shade/bury its huge potential for the large-scale detection, identification, and quantification of proteoforms.

# 6 CONCLUSIONS

2DE-LC/MS as a classic and conventional approach has been extensively used in the field of proteomics. Because it is historically looked upon as a seemingly "low-throughput" technique (only one to two proteins per 2D spot are usually achieved for most of the spots in a 2DE map), 2DE-LC/MS has gradually dimmed in the field of proteomics compared to seemingly "high-throughput" bottom-up approaches such as iTRAQ, TMT, SILAC, and label-free. The rapid development and application of high-sensitivity MS allows one to achieve an average of over fifty or even hundreds of proteins in every 2D gel spot in an analysis of a complex human proteome to break through the conventional concept on 2DE-LC/MS, and let it revive in the field of proteomics. Proteoforms that are mainly derived from splicing and PTMs are the real final functional performers of a gene, and have important scientific merits for life science and medical science. The top-down and high-throughput 2DE-LC/MS in combination with isotopic labeling provides a solid technique support for the large-scale study of human proteoforms and disease-related proteoforms, and provides insights into the mechanisms of a disease and the discovery of reliable biomarkers for prediction, diagnosis, and prognostic assessment of a disease.

# ACKNOWLEDGEMENTS

# CONFLICT OF INTEREST

The authors have declared no conflicts of interest.

# AUTHORS' CONTRIBUTIONS

X.Z. conceived the concept, collected and analyzed references, designed, wrote, and critically revised manuscript, and was responsible for its financial supports and the corresponding works. N.L., X.H.Z, and S.Q. participated in collection and analysis of references, discussions, and partial revision. X.H.Z. also contributed to the critical revision of English language. All authors approved the final manuscript.

# ABBREVIATIONS

2DE: Two-dimensional electrophoresis

2DGE: Two-dimensional gel electrophoresis

CE: Capillary electrophoresis

DEP: Differentially expressed protein

ESI: Electrospray ionization

hGH: Human growth hormone

HIC: Hydrophobic interaction chromatography

hPRL: Human prolactin

ICAT: Isotope-Coded Affinity Tags

ICPL: Isotope-Coded Protein Labeling

IEF: Isoelectric focusing

IPG: Immobilized pH gradient gel

IPTL: Isobaric Peptide Termini Labeling

iTRAQ: Isobaric tags for relative and absolute quantification

LC: Liquid chromatography

MALDI: Matrix-assisted laser/desorption ionization

mGELFrEE: Multiplexed gel-eluted liquid fraction entrapment electrophoresis

$M_r$: Relative molecular mass

MS: Mass spectrometry

MS/MS: Tandem mass spectrometry

ORF: Open-reading frame

PMF: Ppeptide mass fingerprinting

PTM: Post-translational modification

RPLC: Reverse-phase liquid chromatography

SILAC: Stable isotope labeling of amino acids in cell culture

SRM/MRM: Selected or multiple reaction monitoring

SWATH: Sequential window acquisition of all theoretical spectra

TMT: Peptide tandem mass tags

WCX: Weak-cation exchange chromatography

# REFERENCES

1. O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. J Biol Chem. 1975; 250: 4007-4021.

2. Görg A, Obermaier C, Boguth G, Harder A, Scheibe B. The current state of two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis. 2000; 21: 1037-1053.

3. Zhan X, Desiderio DM. Comparative proteomics analysis of human pituitary adenomas: current status and future perspectives. Mass Spectrom Rev. 2005; 24: 783-813.

4. Malmström L, Malmström J, Marko-Varga G, Westergren-Thorsson G. Proteomic 2DE database for spot selection, automated annotation, and data analysis. J Proteome Res. 2002; 1(2): 135-138.

5. Lefler DM, Pafford RG, Black NA, Raymond JR, Arthur JM. Identification of proteins in slow continuous ultrafiltrate by reversed-phase

chromatography and proteomics. J Proteome Res. 2004; 3: 1254-1260.

6. Zhan X, Desiderio DM. A reference map of a human pituitary adenoma proteome. Proteomics. 2003; 3: 699-713.

7. Wasinger VC, Cordwell SJ, Cerpa-Poljak A, Yan JX, Gooley AA, Wilkins MR, Duncan MW, Harris R, Williams KL, Humphery-Smith I. Progress with gene-product mapping of the Mollicutes: Mycoplasma genitalium. Electrophoresis. 1995; 16: 1090-1094.

8. Rabilloud T. When 2D is not enough, go for an extra dimension. Proteomics. 2013; 13: 2065-2068.

9. Campostrini N, Areces LB, Rappsilber J, Pietrogrande MC, Dondi F, Pastorino F, Ponzoni M, Righetti PG. Spot overlapping in two-dimensional maps: a serious problem ignored for much too long. Proteomics. 2005; 5: 2385-2395.

10. Zhan X, Desiderio DM. Heterogeneity analysis of the human pituitary proteome. Clin Chem. 2003; 49: 1740-1751.

11. Zhan X, Desiderio DM, Wang X, Zhan X, Guo T, Li M, Peng F, Chen X, Yang H, Zhang P, Li X, Chen Z. Identification of the proteomic variations of invasive relative to non-invasive non-functional pituitary adenomas. Electrophoresis. 2014; 35: 2184-2194.

12. Wang X, Guo T, Peng F, Long Y, Mu Y, Yang H, Ye N, Li X, Zhan X. Proteomic and functional profiles of a follicle-stimulating hormone positive human nonfunctional pituitary adenoma. Electrophoresis. 2015; 36: 1289-1304.

13. Sanchez JC, Chiappe D, Converset V, Hoogland C, Binz PA, Paesano S, Appel RD, Wang S, Sennitt M, Nolan A, Cawthorne MA, Hochstrasser DF. The mouse SWISS-2D PAGE database: a tool for proteomics study of diabetes and obesity. Proteomics. 2001; 1: 136-163.

14. Hoogland C, Mostaguir K, Appel RD, Lisacek F. The World-2DPAGE Constellation to promote and publish gel-based proteomics data through the ExPASy server. J Proteomics. 2008; 71: 245-248.

15. Chahrour O, Cobice D, Malone J. Stable isotope labelling methods in mass spectrometry-based quantitative proteomics. J Pharm Biomed Anal. 2015; 113: 2-20.

16. Wolters DA, Washburn MP, Yates JR. An automated multidimensional protein identification technology for shotgun proteomics. Anal Chem. 2001; 73: 5683-5690.

17. Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature. 2003; 422: 198-207.

18. Aggarwal K, Choe LH, Lee KH. Shotgun proteomics using the iTRAQ isobaric tags. Brief Funct Genomic Proteomic. 2006; 5: 112-120.

19. Evans C, Noirel J, Ow SY, Salim M, Pereira-Medrano AG. An insight into iTRAQ: where do we stand now? Anal Bioanal Chem. 2012; 404: 1011-1027.

20. O'Brien DP, Timms JF. Employing TMT quantification in a shotgun-MS platform. Methods Mol Biol. 2014; 1156: 187-199.

21. Rauniyar N, Yates JR. Isobaric labeling-based relative quantification in shotgun proteomics. J Proteome Res. 2014; 13: 5293-5309.

22. Hoedt E, Zhang G, Neubert TA. Stable isotope labeling by amino acids in cell culture (SILAC) for quantitative proteomics. Adv Exp Med Biol. 2014; 806: 93-106.

23. Bell-Temin H, Culver-Cochran AE, Chaput D, Carlson CM, Kuehl M, Burkhardt BR, Bickford PC, Liu B, Stevens SM Jr. Novel molecular insights into classical and alternative activation states of microglia as revealed by SILAC-based proteomics. Mol Cell Proteomics. 2015; 14: 3173-3184.

24. Matzke MM, Brown JN, Gritsenko MA, Metz TO, Pounds JG, Rodland KD, Shukla AK, Smith RD, Waters KM, McDermott JE, Webb-Robertson BJ. A comparative analysis of computational approaches to relative protein quantification using peptide peak intensities in label-free LC-MS proteomics experiments. Proteomics. 2013; 13: 493-503.

25. Guo H, Isserlin R, Lugowski A, Kuzmanov U, Emili A. Large-scale label-free phosphoproteomics: from technology to data interpretation. Bioanalysis. 2014; 6: 2403-2420.

26. Zhan X, Yang H, Peng F, Li J, Mu Y, Long Y, Cheng T, Huang Y, Li Z, Lu M, Li N, Li M, Liu J, Jungblut PR. How many proteins can be identified in a 2-DE gel spot within an analysis of a complex human cancer tissue proteome? Electrophoresis. 2018; 39: 965-980.

27. Klose J, Kobalz U. Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. Electrophoresis. 1995; 16: 1034-1059.

28. Thiede B, Koehler CJ, Strozynski M, Treumann A, Stein R, Zimny-Arndt U, Schmid M, Jungblut PR. High resolution quantitative proteomics of HeLa cells protein species using stable isotope labeling with amino acids in cell culture(SILAC),

two-dimensional gel electrophoresis(2DE) and nano-liquid chromatograpohy coupled to an LTQ-OrbitrapMass spectrometer. Mol Cell Proteomics. 2013; 12: 529-538.

29. Aebersold R, Agar JN, Amster IJ, Baker MS , Bertozzi CR, Boja ES, Costello CE, Cravatt BF, Fenselau C, Garcia BA, Ge Y, Gunawardena J, Hendrickson RC, Hergenrother PJ, Huber CG, Ivanov AR, Jensen ON, Jewett MC, Kelleher NL, Kiessling LL, Krogan NJ, Larsen MR, Loo JA, Ogorzalek Loo RR, Lundberg E, MacCoss MJ, Mallick P, Mootha VK, Mrksich M, Muir TW, Patrie SM, Pesavento JJ, Pitteri SJ, Rodriguez H, Saghatelian A, Sandoval W, Schlüter H, Sechi S, Slavoff SA, Smith LM, Snyder MP, Thomas PM, Uhlén M, Van Eyk JE, Vidal M, Walt DR, White FM, Williams ER, Wohlschlager T, Wysocki VH, Yates NA, Young NL, Zhang B. How many human proteoforms are there? Nat Chem Biol. 2018; 14: 206-214.

30. Görg A, Postel W, Günther S. The current state of two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis. 1988; 9: 531-546.

31. Zhan X, Desiderio DM. Differences in the spatial and quantitative reproducibility between two second-dimensional gel electrophoresis systems. Electrophoresis. 2003; 24: 1834-1846.

32. Zhan X, Desiderio DM. Spot volume vs. amount of protein loaded onto a gel: a detailed, statistical comparison of two gel electrophoresis systems. Electrophoresis. 2003; 24: 1818-1833.

33. Grøn B, Filion-Myklebust C, Bjørnsen S, Haidaris P, Brosstad F. Cross-linked alpha s gamma t-chain hybrids in plasma clots studied by 1D- and 2D electrophoresis and western blotting. Thromb Haemost. 1993; 70: 438-442.

34. Neuhuber F, Radacher M, Sorgo G. Analysis of STR-PCR products with high-resolution denaturing PAGE. Int J Legal Med. 1996; 108: 225-226.

35. Lustig S, De Kimpe J, Cornelis R, Schramel P, Michalke B. Platinum speciation in clinical and environmental samples: scrutiny of data obtained by using electrophoresis techniques (flatbed and capillary). Electrophoresis. 1999; 20: 1627-1633.

36. Beranova-Giorgianni S, Giorgianni F, Desiderio DM. Analysis of the proteome in the human pituitary. Proteomics. 2002; 2: 534–542.

37. Beranova-Giorgianni S, Pabst MJ, Russell TM, Giorgianni F, Goldowitz D, Desiderio DM. Preliminary analysis of the mouse cerebellum proteome. Brain Res Mol Brain Res. 2002; 98: 135-140.

38. Liu J, Zhan X, Li M, Li G, Zhang P, Xiao Z, Shao M, Peng F, Hu R, Chen Z. Mitochondrial proteomics of nasopharyngeal carcinoma metastasis. BMC Med Genomics. 2012; 5: 62.

39. Arentz G, Weiland F, Oehler MK, Hoffmann P. State of the art of 2D DIGE. Proteomics Clin Appl. 2015; 9: 277-288.

40. Meleady P. Two-Dimensional Gel Electrophoresis and 2D-DIGE. Methods Mol Biol. 2018; 1664: 3-14.

41. Oliveira BM, Coorssen JR, Martins-de-Souza D. 2DE: the phoenix of proteomics. J Proteomics. 2014; 104: 140-150.

42. Hunsucker SW, Duncan MW. Is protein overlap in two-dimensional gels a serious practical problem? Proteomics. 2006; 6: 1374-1375.

43. Evans CO, Moreno CS, Zhan X, McCabe MT, Vertino PM. Molecular pathogenesis of human prolactinomas identified by gene expression profiling, RT-qPCR, and proteomic analyses. Pituitary. 2008; 11: 231-245.

44. Moreno CS, Evans CO, Zhan X, Okor M, Desiderio DM. Novel molecular signaling and classification of human clinically nonfunctional pituitary adenomas identified by gene expression profiling and proteomic analyses. Cancer Res. 2005; 65: 10214-10222.

45. Wang X, Guo T, Peng F, Long Y, Mu Y, Yang H, Ye N, Li X, Zhan X. Proteomic and functional profiles of a follicle-stimulating hormone-positive human nonfunctional pituitary adenoma. Electrophoresis. 2015; 36: 1289-1304.

46. Zhan X, Desiderio DM. A reference map of a human pituitary adenoma proteome. Proteomics. 2003; 3: 699-713.

47. Prasannan L, Misek DE, Hinderer R, Michon J, Geiger JD. Identification of beta-tubulin isoforms as tumor antigens in neuroblastoma. Clin Cancer Res. 2000; 6: 3949-3956.

48. Kohler M, Thomas A, Püschel K, Schänzer W, Thevis M. Identification of human pituitary growth hormone variants by mass spectrometry.J Proteome Res. 2009; 8: 1071-1076.

49. Zhan X, Desiderio DM. The human pituitary nitroproteome: detection of nitrotyrosyl-proteins with two-dimensional Western blotting, and amino acid sequence determination with mass spectrometry. Biochem Biophys Res Commun. 2004; 325: 1180-1186.

50. Peng F, Li J, Guo T, Yang H, Li M, Sang S, Li X, Desiderio M, Zhan X. Nitroproteins in human astrocytomas discovered by gel electrophoresis and tandem mass spectrometry. J Am Soc Mass Spectrum. 2015; 26: 2062-2076.

51. Turko IV, Li L, Aulak KS, Stuehr DJ, Chang JY, Murad F. Protein tyrosine nitration in the mitochondria from diabetic mouse heart. Implications to dysfunctional mitochondria in diabetes. J Biol Chem. 2003; 278: 33972-33977.

52. Miyagi M, Sakaguchi H, Darrow RM, Yan L, West KA. Evidence that light modulates protein nitration in rat retina. Mol Cell Proteomics. 2002; 1: 293-303.

53. Butt RH, Coorssen JR. Postfractionation for enhanced proteomic analyses: routine electrophoretic methods increase the resolution of standard 2D-PAGE. J Proteome Res. 2005; 4: 982-991.

54. Churchward MA, Butt RH, Lang JC, Hsu KK, Coorssen JR. Enhanced detergent extraction for analysis of membrane proteomes by two-dimensional gel electrophoresis. Proteome Sci. 2005; 3: 5.

55. Butt RH, Coorssen JR. Pre-extraction sample handling by automated frozen disruption significantly improves subsequent proteomic analyses. J Proteome Res. 2006; 5: 437-448.

56. Wright EP, Prasad KA, Padula MP, Coorssen JR. Deep imaging: how much of the proteome does current top-down technology already resolve? PLoS One. 2014; 9: e86058.

57. Noaman N, Abbineni PS, Withers M, Coorssen JR. Coomassie staining provides routine (sub) femtomole in-gel detection of intact proteoforms: Expanding opportunities for genuine Top-down Proteomics. Electrophoresis. 2017; 38: 3086–3099.

58. Zhan X, Huang Y, Long Y. Two-dimensional gel electrophoresis coupled with mass spectrometry methods for an analysis of human pituitary adenoma tissue proteome. J Vis Exp. 2018; (134). doi: 10.3791/56739.

59. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol. 1999; 17: 994-999.

60. Shiio Y, Aebersold R. Quantitative proteome analysis using isotope-coded affinity tags and mass spectrometry. Nat Protoc. 2006; 1: 139-145.

61. Fan X, Zhou S, Wang B, Hom G, Guo M, Li B, Yang J, Vaysburg D, Monnier VM. Evidence of highly conserved ß-crystallin disulfidome that can be mimicked by in vitro oxidation in age-related human cataract and glutathione depleted LEGSKO mouse lens. Mol Cell Proteomics. 2015; 14: 3211-3223.

62. Arntzen MØ, Koehler CJ, Barsnes H, Berven FS, Treumann A. IsobariQ: software for isobaric quantitative proteomics using IPTL, iTRAQ, and TMT. J Proteome Res. 2011; 10: 913-920.

63. Koehler CJ, Strozynski M, Kozielski F, Treumann A, Thiede B. Isobaric peptide termini labeling for MS/MS-based quantitative proteomics. J Proteome Res. 2009; 8: 4333-4341.

64. Biniossek ML, Lechel A, Rudolph KL, Martens UM, Zimmermann S. Quantitative proteomic profiling of tumor cell response to telomere dysfunction using isotope-coded protein labeling (ICPL) reveals interaction network of candidate senescence markers. J Proteomics. 2013; 91: 515-535.

65. Brunner A, Kellermann J, Lottspeich F. A practical guide to the ICPL_ESIQuant software for ICPL-based quantitative proteomics. Biochim Biophys Acta. 2014; 1844: 21-28.

66. Miyagi M, Rao KC. Proteolytic 18O-labeling strategies for quantitative proteomics. Mass Spectrum Rev. 2007; 26: 121-136.

67. Capelo JL, Carreira RJ, Fernandes L, Lodeiro C, Santos HM, Simal-Gandara J. Latest developments in sample treatment for 18O-isotopic labeling for proteomics mass spectrometry-based approaches: a critical review. Talanta. 2010; 80: 1476-1486.

68. Lambert JP, Ivosev G, Couzens AL, Larsen B, Taipale M, Lin ZY, Zhong Q, Lindquist S, Vidal M, Aebersold R, Pawson T, Bonner R, Tate S, Gingras AC. Mapping differential interactomes by affinity purification coupled with dataindependent mass spectrometry acquisition. Nat Methods. 2013; 10: 1239-1245.

69. Gillet LC, Navarro P, Tate S, Röst H, Selevsek N. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. Mol Cell Proteomics. 2012; 11: O111.

70 Muntel J, Fromion V, Goelzer A, Maaß S, Mäder U, Büttner K, Hecker M, Becher D. Comprehensive absolute quantification of the cytosolic proteome of Bacillus subtilis by data independent, parallel fragmentation in liquid chromatography/mass spectrometry (LC/MS (E)). Mol Cell Proteomics. 2014; 13: 1008-1019.

71. Chahrour O, Cobice D, Malone J. Stable isotope labelling methods in mass spectrometry-based quantitative proteomics. J Pharm Biomed Anal. 2015; 113: 2-20.

72. Colangelo CM, Chung L, Bruce C, Cheung KH.

Review of software tools for design and analysis of large scale MRM proteomic datasets. Methods. 2013; 61: 287-298.

73. Lemoine J, Fortin T, Salvador A, Jaffuel A, Charrier JP. The current status of clinical proteomics and the use of MRM and MRM (3) for biomarker validation. Expert Rev Mol Diagn. 2012; 12: 333-342.

74. Zhan X, Giorgianni F, Desiderio DM. Proteomics analysis of growth hormone isoforms in the human pituitary. Proteomics. 2005; 5: 1228-1241.

75. Stastna M, Van Eyk JE. Analysis of protein isoforms: can we do it better? Proteomics. 2012; 12: 2937-2948.

76. Rogowska-Wrzesinska A, Le Bihan MC, Thaysen-Andersen M, Roepstorff P. 2D gels still has a niche in proteomics. J Proteomics. 2013; 88: 4-13.

77. Coorssen JR, Yergey AL. Proteomics is analytical chemistry: fitness-for-purpose in the application of top-down and bottom-up analyses. Proteomes. 2015; 3: 440-453.

78. Rogowska-Wrzesinska A, Le Bihan MC, Thaysen-Andersen M, Roepstorff P. 2D gels still have a niche in proteomics. J Proteomics. 2013; 88: 4-13.

79. Wright EP, Partridge MA, Padula MP, Gauci VJ, Malladi CS, Coorssen JR. Top-down proteomics: enhancing 2D gel electrophoresis from tissue processing to high-sensitivity protein detection. Proteomics. 2014; 14: 872-89.

80. Rabilloud T. The whereabouts of 2D gels in quantitative proteomics. Methods Mo Biol. 2012; 893: 25-35.

81. Lim H, Eng J, Yates JR 3rd, Tollaksen SL, Giometti CS, Holden JF, Adams MW, Reich CI, Olsen GJ, Hays LG. Identification of 2D-gel proteins: a comparison of MALDI/TOF peptide mass mapping to mu LC-ESI tandem mass spectrometry. J Am Soc Mass Spectrom. 2003; 14: 957-970.

82. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann Y, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blöcker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kaspryzk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowki J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ, Szustakowki J; International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature. 2001; 409: 860-921.

83. Anderson L, Seilhamer J. A comparison of selected mRNA and protein abundances in human

liver. Electrophoresis. 1997; 18: 533-537.

84. Gygi SP, Rochon Y, Franza BR, Aebersold R. Correlation between protein and mRNA abundance in yeast. Mol Cell Biol. 1999; 19: 1720-1730.

85. Zhan X, Long Y, Zhan X, Mu Y. Consideration of statistical vs. biological significances for omics data-based pathway network analysis. Med One. 2017; 1: e170002.

86. Zhan X, Long Y, Lu M. Exploration of variations in proteome and metabolome for predictive diagnostics and personalized treatment algorithms: Innovative approach and examples for potential clinical application. J Proteomics. 2017; doi: 10.1016/j.jprot.2017.08.020.

87. Zhan X, Qian S, Huang Y. The untapped potential of nitroproteomics for medicine. Med One. 2017; 2: e170027.

88. Jungblut PR, Holzhütter HG, Apweiler R, Schlüter H. The speciation of the proteome. Chem Cent J. 2008; 2: 16.

89. Schlüter H, Apweiler R, Holzhütter HG, Jungblut PR. Finding one's way in proteomics: a protein species nomenclature. Chem Cent J. 2009; 3: 11.

90. Yaoxia L, Jingyan C, Xialian T, Ping C, Min Z. The 20kDa and 22kDa forms of human growth hormone (hGH) exhibit different intracellular signalling profiles and properties. Gen Comp Endocrinol. 2017; 248: 49-54.

91. Qian S, Yang Y, Li N, Cheng T, Wang X, Liu J, Li X, Desiderio DM, Zhan X. Prolactin variants in human pituitaries and pituitary adenomas identified with two-dimensional gel electrophoresis and mass spectrometry. Front Endocrinol. 2018; doi: 10.3389/fendo.2018.00468.

92. Scheler C, Müller EC, Stahl J, Müller-Werdan U, Salnikow J, Jungblut P. Identification and characterization of heat shock protein 27 protein species in human myocardial two-dimensional electrophoresis patterns. Electrophoresis. 1997; 18: 2823-2831.

93. Klose J, Nock C, Herrmann M, Stühler K, Marcus K, Blüggel M, Krause E, Schalkwyk LC, Rastan S, Brown SD, Büssow K, Himmelbauer H, Lehrach H. Genetic analysis of the mouse brain proteome. Nat Genet. 2002; 30: 385-393.

94. Tran JC, Doucette AA. Multiplexed size separation of intact proteins in solution phase for mass spectrometry. Anal Chem. 2009; 81: 6201-6209.

95. Staub A, Zurlino D, Rudaz S, Veuthey JL, Guillarme D. Analysis of peptides and proteins using sub-2 Î¼m fully porous and sub 3-Î¼m shell particles. J Chromatogr A. 2011; 1218: 8903-8914.

96. Sikanen T, Aura S, Franssila S, Kotiaho T, Kostiainen R. Microchip capillary electrophoresis-electrospray ionization-mass spectrometry of intact proteins using uncoated Ormocomp microchips. Anal Chim Acta. 2012; 711: 69-76.

97. Cummins PM, O'Connor BF. Hydrophobic interaction chromatography. Methods Mol Biol. 2011; 681: 431-437.

98. Goheen SC, Engelhorn SC. Hydrophobic interaction high-performance liquid chromatography of proteins. J Chromatogr. 1984; 317: 55-65.

99. Geng X, Ke C, Chen G, Liu P, Wang F, Zhang H, Sun X. On-line separation of native proteins by two-dimensional liquid chromatography using a single column. J Chromatogr A. 2009; 1216: 3553-3562.