JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v1i1.50

Large-scale 2-D DIGE studies - guidelines to overcome pitfalls and challenges along the experimental procedure

Franziska Dautel[1], Stefan Kalkhof[1], Saskia Trump[2], Irina Lehmann[2], Andreas Beyer[3], Martin von Bergen*[1,4].

[1]Department of Proteomics, UFZ, Helmholtz-Centre for Environmental Research, Permoserstr. 15, 04318 Leipzig, Germany; [2]Department of Environmental Immunology, UFZ, Helmholtz-Centre for Environmental Research, Permoserstr. 15, 04318 Leipzig, Germany; [3]Biotechnology Center, Technische Universität Dresden, Tatzberg 47/49, 01307 Dresden, Germany; [4]Department of Metabolomics, UFZ, Helmholtz-Centre for Environmental Research, Permoserstr. 15, 04318 Leipzig, Germany.

Received: 29 October 2010 **Accepted:** 14 December 2010 **Available Online**: 15 December 2010

ABSTRACT

In large 2-D DIGE proteomic studies with a large number of samples, it is essential to design the experimental setup to detect statistically significant protein changes under consideration of experimental variances. Herein are presented guidelines and general remarks on the extraction of protein expression data by following protein spots on their way from first spot synchronization, detection, quantification and statistical analysis until excision and identification. Further discussion addresses common difficulties, potential pitfalls and strategies for dealing with gel-to-gel discrepancies, labeling inefficiencies, and dye- and batch effects which might not be obvious to novices and even more experienced users of DIGE technology.

Keywords: Protein expression analysis, Proteomics; Differential gel electrophoresis, 2-D-DIGE.

Abbreviations.

2-D-DIGE, 2-D-difference gel electrophoresis; **B[a]P,** benzo[a]pyrene.

1. Introduction

A typical task for researches with a variety of biological queries is to detect the up- or downregulation of proteins belonging to two or more biological groups such as treatment versus control. Therefore, it is widely accepted and recommended to conduct more than one biological and at least three technical replicates of each group. A popular technique for protein quantification is the two-dimensional polyacrylamide gel electrophoresis (2D-GE) which allows both the separation and visualization of thousands of protein species and the protein identification and quantification of their expression patterns. The greatest achievement of 2-D differential in gel electrophoresis (DIGE) lies in the simultaneous separation of more than one sample per gel as well as having an internal standard for the relative quantification of spot intensities [1]. Samples are labeled prior to electrophoresis with spectrally resolvable fluorescent cyanine dyes Cy2, Cy3 and Cy5, mixed prior to isoelectric focusing (IEF) and resolved on the same 2-D gel [2]. Although the proteome can be assessed by different means, DIGE has shown to be a sensitive, accurate, reproducible, financially affordable and easy to handle approach. For many laboratories it is the method of choice for a quantitative proteome analysis - especially for the reliable detection of minor changes in protein abundances that are not detectable by other staining methods [3]. The technology allows the processing of large numbers of samples simultaneously, e.g. for measuring dense time courses. However, the data processing is still not fully supported by the available software. In particular, there are no established standard procedures to process the quantitative information extracted from 2-D DIGE experiments involving a large number of gels.

The main focus of this study is to present strategies for

***Corresponding author:** PD Dr. Martin von Bergen. UFZ, Helmholtz-Centre for Environmental Research, Department of Proteomics, Permoserstr. 15, 04318 Leipzig, Germany. Fax: +49-341-2351786. Email Address: Martin.vonbergen@ufz.de.

dealing with gel-to-gel discrepancies, labeling inefficiencies, and dye- and batch effects. Gel-to-gel discrepancies arise from run-time differences, variances in the loaded protein amounts or dye-front deformations [4]. Accounting for these differences is important for both 2-DE and 2-D DIGE. However, the dye-effect is specific for DIGE-projects, as the application of three different fluorophores can cause preferential dye-protein binding, variances in the fluorescent signal and background and differences in gel migration of the labeled proteins. As a result, protein abundances are not directly comparable when the proteins are labeled differently in various samples [5-7]. In addition, the experimental execution for a large number of samples is often divided into several batches of 6 or 12 gels. As a consequence, results of protein expression often cluster with the performed batches rather than with the individual samples and replicates.

The goal is to identify spots that are truly differentially expressed, while accounting for statistical issues such as the multiple testing problem. This multiple testing problem states the accumulation of false positives as a general property of confidence-based statistical tests. These tests are applied across multiple features such as individual spots in DIGE to detect significantly altered changes in protein abundance [8].

This study reports an experimental design for a 2-factor analysis (time and concentration): murine hepatoma cells (Hepa1c1c7) were treated with the procarcinogen benzo-*a*-pyrene (B[a]P) and protein concentrations were quantified using 2-D DIGE (Fig. 1). Differential protein expression induced by B[a]P (or active B[a]P-metabolites) has previously been studied in different cellular models using one incubation time point and several B[a]P- or B[a]P-metabolite concentrations [9-13]. In contrast, this B[a]P-protein expression analysis sampled four incubation time points at one toxic (5 μM) and one sub-acute B[a]P-concentration (50 nM), which required the processing of 36 samples in total [14]. In order to process the data originating from these experiments, a statistical analysis pipeline was developed to account for dye- and batch effects and to extract concentration- and time-dependent protein profiles.

2. Material and Methods

2.1 *Cell culture and BaP exposure*

Murine hepatoma cells (Hepa1c1c7, ATCC No. CRL-2026; LGC Promochem, Wesel, Germany) were cultured as described elsewhere [15]. The cells were exposed to 50 nM B[a]P (Sigma-Aldrich, Steinheim, Germany), 5 μM B[a]P or DMSO for 2, 4, 12 and 24 h. Three independent biological replicates of all treatments were prepared.

2.2 *DIGE and Data Analysis*

2.2.1 Difference gel electrophoresis

Cells were washed and lysed according to the procedure previously described [16]. Protein extracts were prepared and labeled according to manufacturer's recommendations (GE Healthcare, Uppsala, Sweden). A Cy2-labeled common internal standard for all gels was prepared from a mixture of all samples IPG strips (24 cm, pH range 3-10 NL; GE Healthcare, Freiburg, Germany), which were rehydrated overnight and focused for 100,000 Vhrs using an Ettan IPGphor 3 isoelectric focusing unit (GE Healthcare, Freiburg, Germany) as described earlier [17]. Second dimension separation was performed using an Ettan DALTtwelve electrophoresis system (GE Healthcare, Uppsala, Sweden) on 12 % SDS-PAGE gels. The gels were scanned using the Ettan DIGE Imager Scanner (GE Healthcare, Uppsala, Sweden).

2.2.2 *DIGE analysis*

The gel image analysis was performed using Delta 2-D version 3.6 (Decodon GmbH, Greifswald, Germany; [18]). The gels were warped and a fusion gel was created including all gels of the experiment. Subsequent to the spot detection, the spots were manually edited and transferred to all individual gel pictures. Relative spot volumes (integrated staining intensities) were determined by normalizing the spot volumes to the total protein amount on each gel (excluding the largest spots representing ~ 5 % of the total intensity). The relative spot volumes were extracted and transformed to a log2-scale. Afterwards the log2-ratios of the Cy3 and Cy5 intensities were adjusted to the intensities of the internal standard on the Cy2 channel. A dye-specific bias was observed. In order to remove the bias, the effect of the dye type (Cy3 or Cy5) was regressed out and the residuals were used for subsequent analysis. The distribution of residual spot intensities on each gel was centered by subtracting the mean of each gel. A random effects model was fitted to each spot to account for the spot-specific intra-gel correlation between the Cy3 and Cy5 signals. To eliminate the batch effect, the mean of the DMSO samples at the respective time point were subtracted from the residuals of this model. Thus, each final measurement should only reflect perturbation due to B[a]P exposure.

A two-way ANOVA model with B[a]P-exposure time and concentration as the factors was fitted to each spot. P-values for the time main effect, the concentration main effect, and their interaction were corrected for multiple hypotheses testing using the false discovery rate (FDR). The time main effect was significant for 120 spots at FDR < 0.05. Only these spots were considered in the subsequent analysis.

2.2.3 *Preparation of 2-D-reference gels for protein identification*

Since DIGE-gels only contain 300 μg protein/gel, only very large protein spots are visible on DIGE-gels after applying the blue silver staining method - a modified Neuhoff's colloidal Coomassie Blue G-250 staining with sensitivity close to silver staining [19] (data not shown). In order to detect all identified protein spots, reference gels with 2.0 and 2.5 mg protein (equal mix of all samples) were created for protein identification (procedure as described elsewhere [16]). To avoid mistakes in protein spot identification, the images of the

reference gels were loaded in the Delta 2-D DIGE-project and warped to the DIGE-images.

2.3 *Protein Identification by MALDI-MS or nano-HPLC/ESI-MS*

Following spot excision from the reference gel and tryptic digestion, the measurement was performed with MALDI-TOF/TOF-MS (Ultraflex III, Bruker Daltonik, Bremen Daltonik, Bremen, Germany) using the HCCA matrix (0.6 mg/ml) according to Georgieva *et al.* [20]. Alternatively, if no significant identification was obtained with MALDI-TOF-MS analysis (Mascot-Score cut-off 100) the samples were measured using a nano-HPLC system (2-D-nano-HPLC, Eksigent, Dublin, CA, USA) coupled to an LTQ-Orbitrap XL ETD hybrid mass spectrometer (Thermo Fisher Scientific, USA) [21].

3. Results and Discussion

3.1 General workflow

The analysis of murine hepatoma cells (Hepa1c1c7) treated with the procarcinogen B[a]P is presented as an example showing how to conduct and analyze large 2-D DIGE studies. In this project both different B[a]P-concentrations and incubation time points are used (Table 1). In the optimized workflow (Fig. 1) Cy2 normalization (section 4.3.1) as well as corrections of the labeling (4.3.2) and gel batch effect (4.3.3) were incorporated before the final expression patterns were extracted and tested for significant regulation by ANOVA and FDR estimation. Spots of interest with significant expression changes were identified using post-stained reference gels which were warped to the equivalent DIGE-gels in the Delta2D-project to eliminate false identifications. Each step is discussed in detail in the following sections.

3.2 *Spot matching and spot detection using Delta2D*

The first and crucial step in the 2-D DIGE workflow before data processing involves the recognition of common spots across different gels (Fig. 1). In the proposed setup, this part is performed using the Delta2D software. However, free, open-source image-processing algorithms for image registration and fusion are also available [22]. The time consuming image analysis is one of the bottlenecks in 2-DE studies and the assistance of a manual operator is needed [23]. Thus, possible drawbacks of this workflow-step have an important impact on the outcome of the further analysis.

Twenty 2-D DIGE gels (24 cm, p*I* 3-10) were created to analyze two B[a]P concentrations, DMSO as a control and four different time points in triplicates (36 samples) (Table 1). All gel images were loaded into the Delta2D software. However, uploading and processing such a large number of images can lead to severe software instabilities. In this case, only the 64 bit version of Delta2D (not the 32 bit version) was able to administer 60 gel images in one project at once.

Another problem is the visualization of large gel series: it is impossible to show all gels side-by-side, which can lead to confusion and consequently mistakes in the analysis by the manual operator [18].

Table 1. Experimental setup for the large 2-D DIGE study of B[a]P-induced alterations in protein expression in murine Hepa1c1c7 cells. Gel batches were performed according to incubation time points and dyes were swapped among biological replicates and among the DMSO-samples of the 3rd replicate in each time point.

Sample	Time	Replicate	Label	Gel No.	Batch
5 µM B[a]P	2 h	1	Cy3	2	1
50 nM B[a]P	2 h	1	Cy5	1	1
DMSO	2 h	1	Cy3	1	1
5 µM B[a]P	2 h	2	Cy5	3	1
50 nM B[a]P	2 h	2	Cy3	3	1
DMSO	2 h	2	Cy5	2	1
5 µM B[a]P	2 h	3	Cy3	5	1
50 nM B[a]P	2 h	3	Cy5	4	1
DMSO	2 h	3	Cy3/Cy5	4+5	1
5 µM B[a]P	4 h	1	Cy5	7	2
50 nM B[a]P	4 h	1	Cy3	6	2
DMSO	4 h	1	Cy5	6	2
5 µM B[a]P	4 h	2	Cy5	8	2
50 nM B[a]P	4 h	2	Cy3	8	2
DMSO	4 h	2	Cy3	7	2
5 µM B[a]P	4 h	3	Cy3	9	2
50 nM B[a]P	4 h	3	Cy5	10	2
DMSO	4 h	3	Cy3	9+10	2
5 µM B[a]P	12 h	1	Cy3	12	3
50 nM B[a]P	12 h	1	Cy5	11	3
DMSO	12 h	1	Cy3	11	3
5 µM B[a]P	12 h	2	Cy5	13	3
50 nM B[a]P	12 h	2	Cy3	13	3
DMSO	12 h	2	Cy5	12	3
5 µM B[a]P	12 h	3	Cy3	15	3
50 nM B[a]P	12 h	3	Cy5	14	3
DMSO	12 h	3	Cy3/Cy5	14+15	3
5 µM B[a]P	24 h	1	Cy3	17	4
50 nM B[a]P	24 h	1	Cy5	16	4
DMSO	24 h	1	Cy3	16	4
5 µM B[a]P	24 h	2	Cy5	18	4
50 nM B[a]P	24 h	2	Cy3	18	4
DMSO	24 h	2	Cy5	17	4
5 µM B[a]P	24 h	3	Cy3	20	4
50 nM B[a]P	24 h	3	Cy5	19	4
DMSO	24 h	3	Cy3/Cy5	19+20	4

2-DE-gels are often affected by spatial distortions due to run-time differences and dye-front deformations. Spot matching is the first and essential step to receive a proteome map of good quality. This is achieved via 'warping' the images in the image analysis software to remove distortions from the gel images and to bring the spot patterns into congruency [18]. Delta2D has been shown to be a fast and more reliable image analysis software than comparable commercial products like Proteomweaver (Definiens), especially due to its precise warping procedure [24].

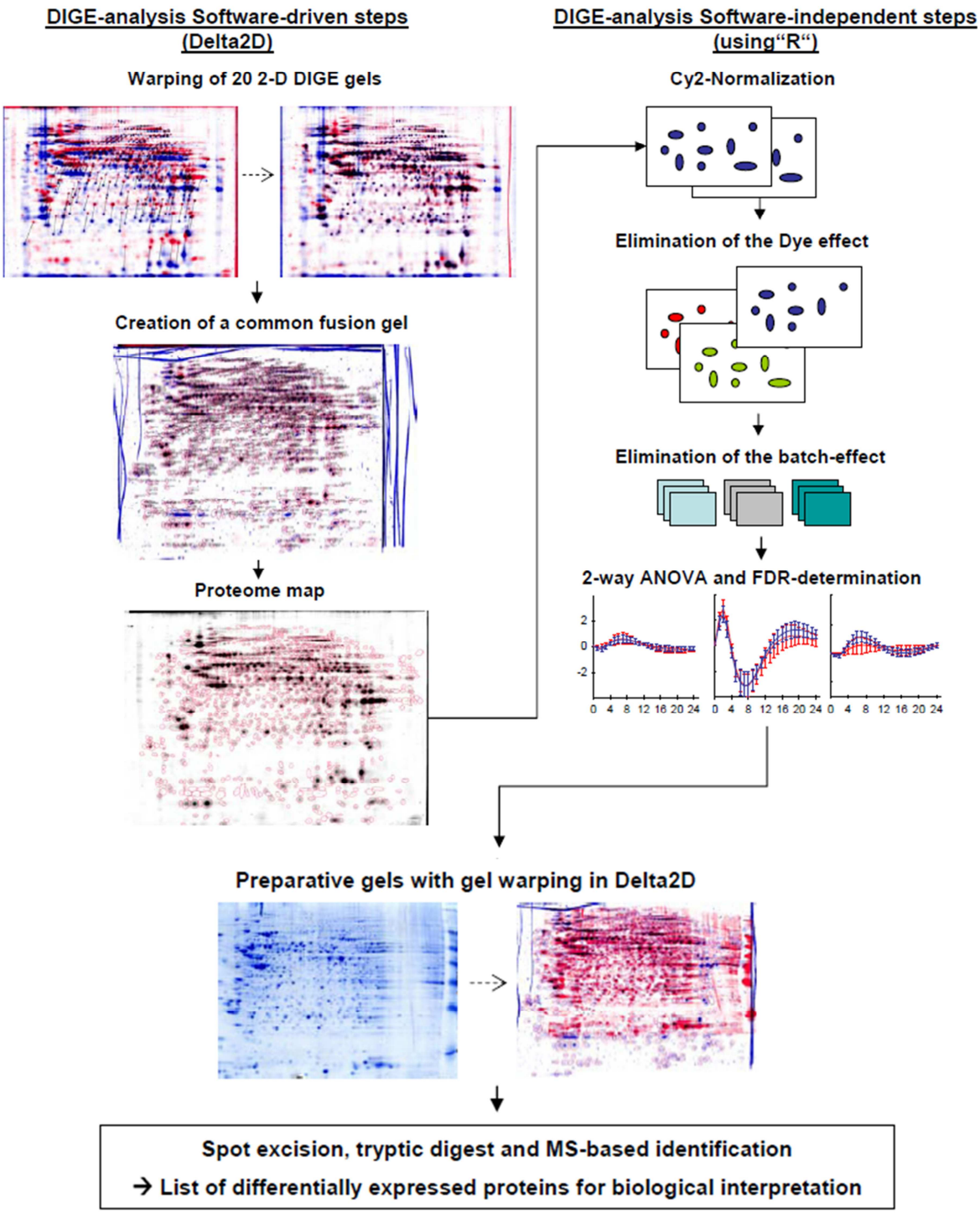

Figure 1. *Flowchart to demonstrate the individual steps in the proposed 2D-DIGE analysis.* Gel images are loaded in appropriate analysis software (e.g. Delta2D) for gel warping, creation of a fusion gel and the subsequent spot detection to create a common proteome map, which is then transferred on all gels. To correct for gel- to gel-variation, dye- and batch effects and to extract time- and concentration dependent protein expression curves, it is recommended to transfer the data to independent statistical software such as "R".

A key advantage of using a pooled internal standard is that the same sample is used for the gel-to-gel matching, as opposed to matching gels that contain different samples and that may consequently have different spot patterns [2]. The alignment of gel images establishes a spot consensus pattern by creating a composite image summarizing the whole experiment's gel information (Fig. 1). A robust reference image is particularly important for large studies since the large number of gels processed in different batches increases the likelihood of detecting artifactual differences simply by

chance. Manual inspection of this virtual fusion gel verified that it did not contain visible imperfections such as dust, air bubbles or precipitated dye. In order to be declared as a spot, a dot on the fusion gel was required to match the three-dimensional profile characteristics of a spot. With the consensus spot patterns transferred to all gels from the fusion gel, 1227 spots were detected. In comparison, Corzett *et al.* applied a more traditional approach (DeCyder Differential Analysis Software v5.01, GE Healthcare) using spot detection on each individual gel resulting in different spot patterns for each gel. Thus, only 165 (6.8 %) of all detected spots could be matched on all 12 gels loaded with protein lysates of human plasma [5]. In conclusion, the detection of differentially expressed proteins can be substantially improved by using consensus spot patterns [24]. Without a unified proteome map and thus separate spot detections on every gel, the missing values must be engrafted by statistical means such as missing value imputation [18].

However, one pitfall in the 2-D DIGE setup remains: the need for a manual operator in spite of the associated great dependency on the performance of the individual. The decision whether a spot is a spot and where it begins or ends can be a source for unintentional data manipulation. Although automated warping procedures exist, they do not yet provide sufficient quality, especially for a large number of gels in which artifacts accumulate and gel-to-gel variations increase due to different gel batches. It has been observed that vectors were set incorrectly, requiring time-consuming manual corrections [23, 24]. In addition, automated spot matching is prone to spot amalgamation in regions of variable spot resolution and can lead to erroneous measurements [25]. Thus, a future goal should be the reduction of user controlled settings, as described in [26].

3.3 *Data analysis and normalization*

3.3.1 *Data processing and Cy2-Normalization*

Existing software analysis packages specialized for 2-D DIGE evaluation lack important features required for analyzing larger sets of gels. Large 2-D DIGE datasets with various time points and concentrations are in need of a more flexible analysis than datasets with only two samples (e.g. treated versus control) in three replicates.

The following steps have been taken to deal with problems specific to large 2-D DIGE projects (Fig. 1). First, a pooled internal standard labeled with Cy2 was included in the 2-D DIGE experimental design. Second, using 3 different fluorophores for pre-protein staining can result in a dye-specific effect. In Delta2D it is possible to detect a dye-effect if a dye-swapping was performed (refer to section 4.3.2), but the data has to be transferred to external software to correct for this effect (Fig. 1). Third, a gel-specific batch effect was observed for the data which was confounded with the B[a]P-incubation time points (section 4.3.3). Again, Delta2D only allows for detecting but not correcting batch effects. The ultimate goal was to extract concentration- and time-dependent protein expression curves. Delta2D can extract either concentration- or time-dependent expression profiles; however, a combined graph can not be exported.

In conclusion, the protein spot volume data were exported and all subsequent analysis was done using external statistics software, R in this case (www.r-project.org). Delta2D reports spot volumes as percent of the total volume on the respective gel. These percentages were transformed into log2-fractions, which is more convenient since this scale is symmetrical and centered at zero. For example, a two fold increase and decrease on the log2 scale is 1 and -1, respectively, whereas on the percentage scale the same changes would be 200 % and 50 %, respectively. This lack of symmetry can pose problems for downstream-analysis. In addition, spots with missing data were removed. Next, the ratio of the Cy3 and Cy5 channels to the internal standard (Cy2) for each spot was calculated and transformed to the log2 scale, eliminating false effects arising from slightly different protein amounts loaded on the gels. In addition to having the same spot patterns for spot matching to receive a common proteome map, this is the second advantage of the internal standard approach. However, one possible problem of using the Cy5/Cy2 and Cy3/Cy2-ratios (as also suggested by the Delta2D software) is the violation of the statistical assumption of independent sampling [4]. In addition, spatial trends in the intensity measurements of the internal standard were observed and one-third of the available sample-space is lost [25]. Engelen *et al.* suggested removing the spatial bias with a strategy called spatial intensity bias removal (SIBR) using a two-dimensional nonlinear regression algorithm. Since SIBR does not need an internal standard, more biological samples could be run on a single gel, resulting in decreasing numbers of gels, lower costs, shorter sample preparation and image analysis steps [25]. Another approach suggests one sample (test or control) with a pooled standard on one gel as performed in saturation labeling also in minimal labeling approaches [27]. But with limited resources, restricted numbers of gels and many samples to be analyzed, the three-dye minimal approach is more readily applicable [28]. Furthermore, the use of an internal standard remains the most applicable approach to correct for gel-to gel variations.

3.3.2 *Correcting the dye effect*

One systematic source of variation within the dataset could come from a dye effect (Fig. 1). Often it is assumed that a protein-specific dye effect does not occur. As a result, protein abundances should be directly comparable across dyes after normalizing the Cy3 and Cy5-values to the internal standard Cy2 as suggested by Delta2D [18]. However, several studies have demonstrated an existing dye effect in DIGE-studies due to a combination of preferential dye binding to proteins, differences in gel migration of the labeled proteins, and differences in the fluorescent signal and background [5-7]. One method of eliminating dye effects is to perform a dye-swap. In the B[a]P-DIGE dataset, the Cy3/Cy5-fluorophores were switched among the samples (Table 1) ensuring that

within the three biological replicates for each sample, dyes were swapped at least once. The dye swap would generally double the number of gels needed as each sample would have to be measured twice, or only half of the biological replicates would be left in the experiment. A sensible solution that avoids such a suboptimal experimental design is to swap the two dyes among the three biological replicates, one dye being the same in the corresponding sample of two replicates. Additionally, one sample per batch (usually the DMSO-control of the 3rd replicate) was labeled with both dyes on separate gels (Tab. 1). Thus, a distinction can be made between the dye effect and differences coming from the biological replicates. With this approach, a protein-specific dye effect could clearly be detected. Such protein-specific biases require computational correction, especially if a large number of samples are analyzed and not all samples can be swapped individually [5]. Within Delta2D, a mathematical dye correction cannot be performed. Corzett *et al.* described a method to account for spot variances, including the dye effect, based on a mixed-effects model [5]. In this case, the dye correction was done by regressing out the effect of the dyes and taking the residuals using a fixed effects model. Following this, distribution of log2 ratios for each gel were centered. In the second step, a random effects model was fitted to account for the correlation between samples which were run on the same gel.

3.3.3 *Correcting the batch effect*

In customized GE systems, it is often only possible to run a small number of gels in one batch. Eravci *et al.* extended the common available GE-apparatus to the load of 24 gels at once [24]. However, the handling of up to twelve gels at once in the commercially available instruments provides a wide variability of possible errors in the sample and gel processing, resulting in false results or artifacts and avoidable repetitions of experiments. In addition, technical failures or mistakes along the experimental procedure can lead to gel artifacts rendering entire batches useless for any further analysis. The costs for losing complete batches of 12 2-D DIGE-gels or more must be considered.

For this study, the twenty gels were divided in batches according to B[a]P-incubation time points (Tab. 1, Suppl. Fig. 1). Even after the correction of all spot intensities of Cy3- and Cy5-labelled samples with the common internal standard Cy2, the batch effect was still visible (Suppl. Fig. 1). To differentiate between the time- and the batch effect, the data of the control samples DMSO were used. Based on the results from cytotoxicity measurements and considering that the cells are not synchronized, it was concluded that any temporal effects observed in the control samples are not due to different incubation times, but artifacts caused by the batch effect [14]. Thus, the DMSO mean from each incubation time point was subtracted from the corresponding B[a]P-incubated samples. The batch effect correction leads to an increased confidence that the observed change in protein abundance is due to biological happening in the cell and not due to experimental variability (Fig. 2B). The batch effect increases with the number of samples as more runs are needed to process all samples. Furthermore, it has been demonstrated that commercially available IPG strips show large differences when they originate from different mastergels. Often the strip numbers in one package are not consecutive (GE Healthcare) or not numbered at all (other suppliers). Consequently the strips originate from different mastergels. This contributes to the batch effect as well [24].

3.3.4 *Identification of significantly changing protein spots*

After completion of the corrections outlined above, the next step is determining proteins exhibiting significant changes as functions of time or concentration. The statistical analysis of DIGE software packages like DeCyder (GE Healthcare), Progenesis SameSpots V3.0 (Nonlinear Dynamics) and Dymension 3 (Syngene) (comparable to Delta2D) have been shown to be inconsistent; results obtained for protein fold changes for one dataset were substantially different in each package. Thus, DIGE quantification is still software dependent despite the use of an internal standard [26]. This observation calls for a reevaluation and validation of the results by DIGE analysis software independent statistical methods such as the external statistics software "R", and biological methods, such as Western Blot.

Many studies exclusively focus on the calculation of protein abundance ratios, often expressed as fold changes (e.g. control versus treatment). Additionally, the student's t-test is often used to test for statistical significance of concentration changes. Such univariate methods determine whether the differences between two samples are significant. However, such methods are inappropriate for studies involving more than two conditions (i.e. more than one treatment condition). In this case two factors exist, time and concentration, time having four levels (4 time points). The t-test in such a situation would not discern whether differences between samples are caused by time or concentration effects, because most samples differ with respect to both time and concentration. Shen *et al.* reported protein changes in human amniotic epithelial FL cells incubated with 0.05 µM anti-benzo[*a*]pyrene-7,8-dihydrodiol-9,10-epoxide in which no single protein was significantly altered in all three incubation time points [12].

Therefore study designs involving multiple factors require the application of Analysis of Variance (ANOVA). ANOVA determines the extent to which the observed variances between samples can be explained by the experimental parameters as opposed to biological or technical variation within the experiment [6]. Delta2D provides a large variety of statistical tools including a two-way ANOVA model. However, to analyze the variances among this data and account for dye and batch effect, the raw data must be extracted from Delta2D for further processing, precluding any possible return to the gel analysis software for that project. Thus, within the two-way ANOVA model, the concentration factor (5 µM and 50 nM B[a]P) and the time factor (2 h, 4 h, 12 h, 24 h) as well as their interaction were considered. This model was fit for

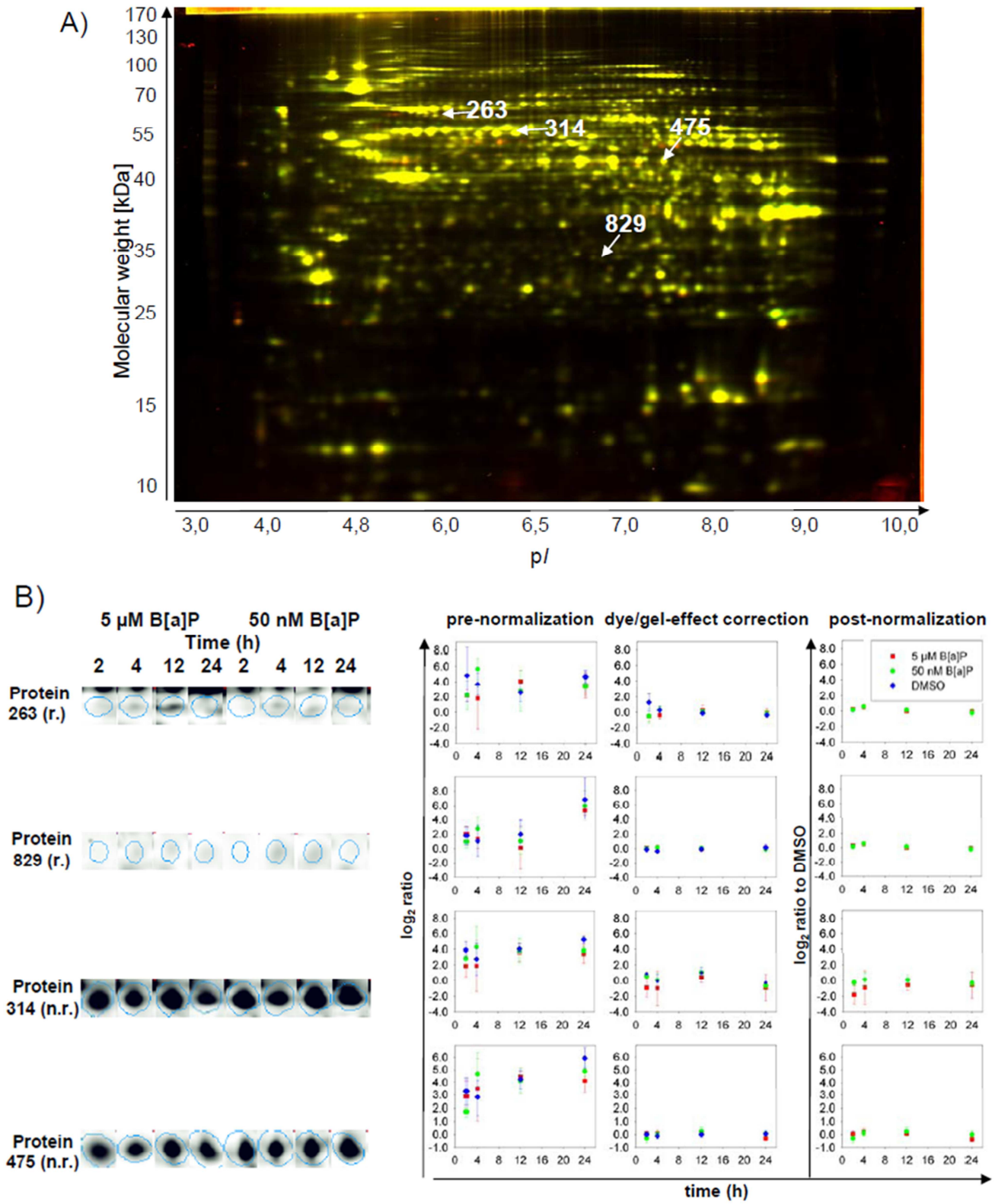

Figure 2. *Visualization of protein expression results. A)* 2-D DIGE gel of murine Hepa1c1c7cell-protein extract. The cells were incubated for 4 h with 50 nM B[a]P(Cy3) and DMSO (Cy5). *B)* Extracted protein spots (left) in comparison with the time- and concentration-dependent protein expression profile (right) of 4 selected proteins (263, 829, 314, 475). Whereas a spot album gives rather visual information about the protein abundance without any normalization, the expression curves show the entire behavior of a protein over B[a]P-concentration (5 µM: red; 50 nM: green; DMSO: blue) and exposure time based on normalized data. Displayed are the log2-ratios of the spot intensities after taking the Cy3/Cy2 and Cy5/Cy2 ratios (pre-normalization), following the dye/gel-correction and after correcting for the batch effect (post-normalization). Proteins 263 and 829 were differently regulated over B[a]P-exposure time (r.) while the expression of proteins 314 and 475 were not found to be changed (n.r.). The location of these spots on a 2-D DIGE gel is also marked in A).

each spot and the p-values for each effect were recorded. Finally, an adjustment for multiple testing was made. Each test has a certain probability of giving false positive results. In the matter of 2-D DIGE, a protein spot could be declared significantly altered in expression, despite the difference being due only to chance. This problem is exacerbated for experiments involving a large number of tests, as is commonly done when using DNA microarrays. Due to the large number of gels processed in this study, similar problems were faced, and inflated false positive rates were accounted for by computing the False Discovery Rate (FDR). The FDR is the expected rate of false positives among all results that were declared positive (i.e. 'significant') [4, 29]. After the FDR adjustment, 120 out of 1227 protein spots had a significant time effect at FDR < 0.05 and no spots exhibited a significant concentration effect or an interaction between time and concentration (Fig. 2A).

3.4 *Protein spot excision*

A crucial step for obtaining reliable results in 2-D DIGE experiments is the excision of significantly altered protein spots from gels (Fig. 1). Spot excision can be performed manually or by using spot pickers. To increase the detection rate of differentially expressed proteins of 2-D DIGE analysis without the availability of a spot picker and still eliminating false spot identifications, a two-step approach is proposed: after staining the DIGE-gels with Coomassie Brilliant Blue (CBB) or equivalent stains, the images should be loaded in the analysis software and warped within the 2-D DIGE project to the internal standard of the gel. This ensures that the same spot, labeled once with Cy3, Cy2 or Cy5 and once with a post-stain, is being detected. In most cases, approximately 100 µg of protein are labeled for each sample and thus 300 µg protein lysate in total are loaded on one 2-D DIGE gel. If the separation is increased by using large IEF-strips and long gels for the second dimension, low abundant spots are often not detectable with post-staining methods (Suppl. Fig. 2). Therefore, in the second step, 2-D reference gels are prepared carrying a larger protein amount (in this case 2.0 and 2.5 mg) from a mixture of all samples incubated with the same post-stain, and are additionally warped to the fusion gel of the 2-D DIGE project to ensure that exactly the same spots are excised. Ideally, one would use one reference gel for each type of sample (e.g. treated versus control), as the spot pattern might then be similar to the respective DIGE-gel. Each reference gel would then be warped to the respective samples e.g. using Delta2D. However, such an approach is often impossible due to limited resources.

Post-staining with other dyes prior to spot picking has been recommended even when using spot pickers [30-32]. Slight mobility differences between labeled and unlabeled species of the same protein have been reported. The unlabeled protein moves slightly faster than the labeled equivalent (about 0.5 kDa) [2] which can cause problems in the subsequent MS-based identification of the protein. However, proteins may be missed, since different proteins vary in their individual staining properties and some post-stains show relatively low sensitivity such as CBB [30]. About 40% of the differentially regulated spots discovered with DIGE could not be reliably detected after post-staining with colloidal CBB [32]. Better results were reported for SYPRO Ruby, but its excitation is achieved either with UV light or with laser scanners. Thus,

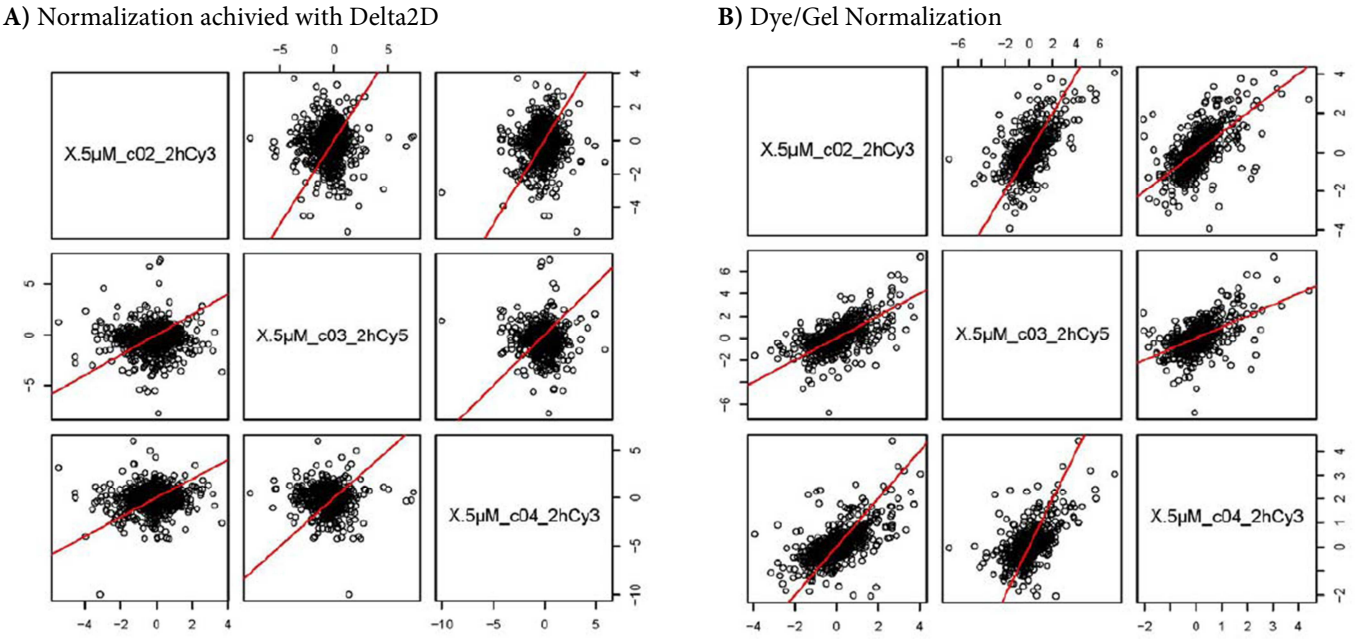

Figure 3. *Comparison between the different normalization approaches (Delta2D-based and -independent). A)* With no correction for the different variances in the gels (dye- and batch effect), the sample replicates do not show any linear trend. *B)* Note the improvement in the linear correlation between the sample replicates after correcting for gel-, dye- and batch effects in this approach.

spots would have to be picked under excitation and not under daylight condition [31]. In this study, the "blue silver" staining was used, which is faster and more sensitive in comparison to colloidal CBB, and spots can be cut in daylight under a hood to prevent ceratin contamination [19]. In the end, the excised spots were digested and identified using MALDI- or ESI-MS/(MS), as it is often useful to combine different types of MS [14, 28].

4. Concluding remarks

Evaluation of proteomic data with 2-D DIGE technology remains a multistep process. In this study, the route of the protein spots is followed along the experimental procedure. From the appropriate gel warping and spot detection to the final spot excision, the researcher faces multiple challenges for obtaining high-quality protein expression data. It has been shown that protein quantification can be improved by taking into account dye and batch effects and proper statistical analysis (Fig. 3). Some difficulties concerning 2-D DIGE experimentation and evaluation must still be overcome, particularly the large individual influence of the manual operator, robust correction for different variances during the spot analysis, and correct spot excision call for improved methods. Reliable protein expression time courses for B[a]P treated murine Hepa cells were successfully extracted (Fig. 2B). These expression curves will provide further hints on the molecular processes triggered by B[a]P-exposure.

5. Supplementary material

Supplementary material regarding this manuscript is online available in the web page of JIOMICS.

http://www.jiomics.com/index.php/jio/rt/suppFiles/50/0

Acknowledgements

The authors thank Jacob Michaelson, Dr. Janina Tomm, Stephan Müller, David Lee Harris, Yvonne Kullnick, Antje Thonig, Anna Lohse and Kerstin Krist for cooperation and technical assistance. The project was supported by the Helmholtz-Alliance on Systems Biology and by Helmholtz Impulse and Networking Fund through Helmholtz Interdisciplinary Graduate School for Environmental Research (HIGRADE). AB acknowledges funding from the Klaus Tschira Foundation.

References

1. Unlu, M., M.E. Morgan, and J.S. Minden, Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. Electrophoresis, 1997. **18**(11): p. 2071-7.
2. Alban, A., et al., A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. Proteomics, 2003. **3**(1): p. 36-44.
3. Karp, N.A., et al., Comparison of DIGE and post-stained gel electrophoresis with both traditional and SameSpots analysis for quantitative proteomics. Proteomics, 2008. **8**(5): p. 948-60.
4. Karp, N.A., et al., Experimental and statistical considerations to avoid false conclusions in proteomics studies using differential in-gel electrophoresis. Mol Cell Proteomics, 2007. **6**(8): p. 1354-64.
5. Corzett, T.H., et al., Statistical analysis of the experimental variation in the proteomic characterization of human plasma by two-dimensional difference gel electrophoresis. J Proteome Res, 2006. **5**(10): p. 2611-9.
6. Karp, N.A. and K.S. Lilley, Maximising sensitivity for detecting changes in protein expression: experimental design using minimal CyDyes. Proteomics, 2005. **5**(12): p. 3105-15.
7. Krogh, M., et al., Analysis of DIGE data using a linear mixed model allowing for protein-specific dye effects. Proteomics, 2007. **7**(23): p. 4235-44.
8. Listgarten, J. and A. Emili, Statistical and computational methods for comparative proteomic profiling using liquid chromatography-tandem mass spectrometry. Mol Cell Proteomics, 2005. **4**(4): p. 419-34.
9. Gao, Z., et al., Zinc finger proteins and other transcription regulators as response proteins in benzo[a]pyrene exposed cells. Mutat Res, 2004. **550**(1-2): p. 11-24.
10. Hooven, L.A. and W.M. Baird, Proteomic analysis of MCF-7 cells treated with benzo[a]pyrene, dibenzo[a,l]pyrene, coal tar extract, and diesel exhaust extract. Toxicology, 2008. **249**(1): p. 1-10.
11. Lau, A.T. and J.F. Chiu, Proteomic and biochemical analyses of in vitro carcinogen-induced lung cell transformation: synergism between arsenic and benzo[a]pyrene. Proteomics, 2006. **6**(5): p. 1619-30.
12. Shen, W., H. Liu, and Y. Yu, Translation initiation proteins, ubiquitin-proteasome system related proteins, and 14-3-3 proteins as response proteins in FL cells exposed to anti-benzo[a]pyrene-7,8-dihydrodiol-9,10-epoxide. Proteomics, 2008. **8**(17): p. 3450-68.
13. Shen, W., H. Liu, and Y. Yu, Proteomic analysis of cellular responses to different concentrations of anti-benzo(a)pyrene-7,8-dihydrodiol-9,10-epoxide in human amniotic epithelial FL cells. J Proteome Res, 2007. **6**(12): p. 4737-48.
14. Dautel , F., Kalkhof , S., Trump , S., Michaelson , J., Beyer , A., Lehmann , I., von Bergen, M., DIGE based protein expression analysis of BaP-exposed hepatoma cells reveals a complex stress response including alterations in oxidative stress, cell cycle control, and cytoskeleton motility at toxic and subacute concentrations. J. Proteome Res., 2010.
15. Solhaug, A., M. Refsnes, and J.A. Holme, Role of cell signalling involved in induction of apoptosis by benzo[a]pyrene and cyclopenta[c,d]pyrene in Hepa1c1c7 cells. J Cell Biochem, 2004. **93**(6): p. 1143-54.
16. Morbt, N., et al., Proteome changes in human bronchoalveolar cells following styrene exposure indicate involvement of oxidative stress in the molecular-response mechanism. Proteomics, 2009. **9**(21): p. 4920-33.
17. Benndorf, D., et al., Identification of spore allergens from the indoor mould Aspergillus versicolor. Allergy, 2008. **63**(4): p. 454-60.
18. Berth, M., et al., The state of the art in the analysis of two-dimensional gel electrophoresis images. Appl Microbiol Biotechnol, 2007. **76**(6): p. 1223-43.
19. Candiano, G., et al., Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis. Electrophoresis, 2004. **25**(9): p. 1327-33.
20. Georgieva, D., et al., Comparative analysis of the venom proteomes of Vipera ammodytes ammodytes and Vipera ammodytes meridionalis. J Proteome Res, 2008. **7**(3): p. 866-86.
21. Muller, S.A., et al., Optimization of parameters for coverage of

low molecular weight proteins. Anal Bioanal Chem.

22. Sorzano, C.O., et al., Elastic image registration of 2-D gels for differential and repeatability studies. Proteomics, 2008. **8**(1): p. 62-5.
23. Millioni, R., et al., Delta2D and Proteomweaver: Performance evaluation of two different approaches for 2-DE analysis. Electrophoresis. **31**(8): p. 1311-7.
24. Eravci, M., et al., Improved comparative proteome analysis based on two-dimensional gel electrophoresis. Proteomics, 2007. **7**(4): p. 513-23.
25. Engelen, K., et al., Alternative experimental design with an applied normalization scheme can improve statistical power in 2D-DIGE experiments. J Proteome Res. **9**(10): p. 4919-26.
26. Kang, Y., et al., Comparison of three commercially available DIGE analysis software packages: minimal user intervention in gel-based proteomics. J Proteome Res, 2009. **8**(2): p. 1077-84.
27. Karp, N.A. and K.S. Lilley, *Design and analysis issues in quantitative proteomics studies.* Proteomics, 2007. **7 Suppl 1**: p. 42-50.
28. McNamara, L.E., et al., Fluorescence two-dimensional difference gel electrophoresis for biomaterial applications. J R Soc Interface. **7 Suppl 1**: p. S107-18.
29. Benjamini, Y., et al., *Controlling the false discovery rate in behavior genetics research.* Behav Brain Res, 2001. **125**(1-2): p. 279-84.
30. Miller, I., J. Crawford, and E. Gianazza, *Protein stains for proteomic applications: which, when, why?* Proteomics, 2006. **6**(20): p. 5385-408.
31. Gharbi, S., et al., Evaluation of two-dimensional differential gel electrophoresis for proteomic expression analysis of a model breast cancer cell system. Mol Cell Proteomics, 2002. **1**(2): p. 91-8.
32. Tonge, R., et al., Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. Proteomics, 2001. **1**(3): p. 377-96.