Time-course Reproducibility of Clinical Samples using Retentate Chromatography MALDI-TOF Mass Spectrometry

Sabine Jourdain[1], Amanda Bulman[2], Waltraud Passlack[3], Sonja Hartwig[3], Akos Czibere[4], Stefan Lehr[3], Fiona Plows[2] and Enrique Dalmasso[2]
[1]Bio-Rad Laboratories, Division Bio-Recherche, Marnes la Coquette, France; [2]Bio-Rad Laboratories, Inc., Hercules, CA, USA;
[3]Institute of Clinical Biochemistry and Pathobiochemistry, German Diabetes Center, Dusseldorf, Germany; [4]Harvard Stem Cell Institute, Boston, MA, USA

Life Science Group
2000 Alfred Nobel Drive
Hercules, CA 94547 USA

Introduction

Success for biomarker discovery projects in clinical proteomics relies primarily on analytical and long-term reproducibility. Sample collection, sample preparation and performance of analytical instruments can all contribute to output signal variability. In order to ensure that discovered differences of expression are physiological and not introduced by experimental variability, any biomarker discovery study should include a means to monitor reproducibility.

A biomarker discovery project was performed using the Lucid Proteomics System™, a top-down proteomic solution that combines the separation power of retentate chromatographic surfaces with high performance MALDI-TOF instrumentation. A total of 105 control and disease samples were split into 5 even groups such that their analysis on ProteinChip arrays was organized over 5 days to avoid sample freeze/thaw cycles and manage bench handling time. A reference sample was used to monitor inter-assay reproducibility within a day and day-to-day reproducibility. We show that sample preparation on chromatographic arrays following a simple but standardized protocol results in experimental variance coefficients below 15%.

Methods

- **Proteominer enrichment:**

The 105 patient plus one pooled reference serum samples (1 ml per patient) were enriched using ProteoMiner bead technology (Bio-Rad Laboratories) with a sample to bead ratio of 10:1 and a single elution (7 M urea, 2 M thiourea, 4% CHAPS, 25 mM Tris-HCl, pH>12). Resulting eluted fractions were frozen until array preparation.

- **Sample randomization:**

Five sets of 22 samples contained a random selection of an equal number of samples from each clinical group plus the reference sample. Information for patient samples (in duplicate) and the reference sample (in quadruplicate) were annotated and assigned on the ProteinChip arrays using the virtual notebook feature of Lucid™ Proteomics Software (Bio-Rad). Each set of samples was prepared on three different array surfaces [metal affinity (IMAC30-Cu^{2+}), cation exchange (CM10), and anion exchange (Q10)] over a five day period.

- **ProteinChip array preparation:**

ProteinChip arrays were pre-equilibrated with their corresponding binding and washing buffers: 0.1 M NaPhosphate + 0.5 M NaCl, pH 7.0 for IMAC-Cu arrays; 100 mM NaAcetate, pH 4.0 for CM10 arrays; and 50 mM Tris-HCl, pH 9.0 for Q10 arrays. ProteoMiner fractions were diluted 1:10 into appropriate buffers and allowed to bind for 1 hour at room temperature. Each spot was then washed three times with binding buffer and rinsed two times with water prior to addition of matrix (50% sinapinic acid).

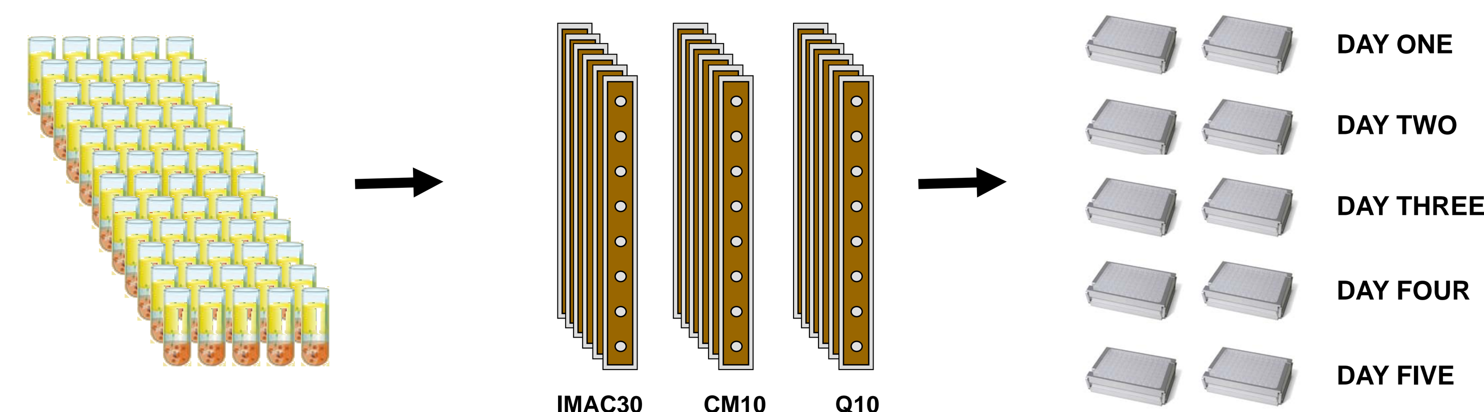
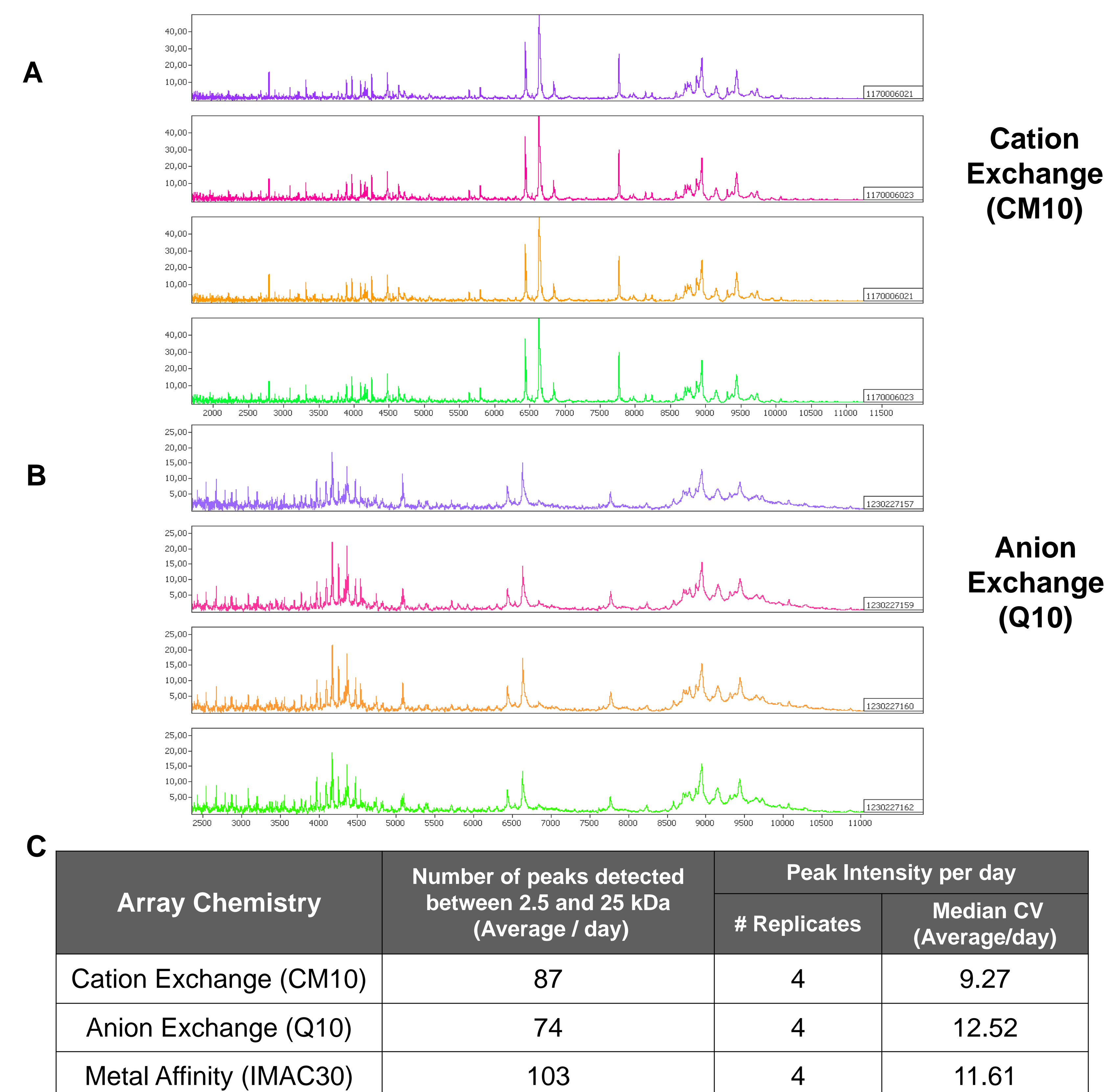


Figure 1. Sample and array preparation overview.

Methods, continued

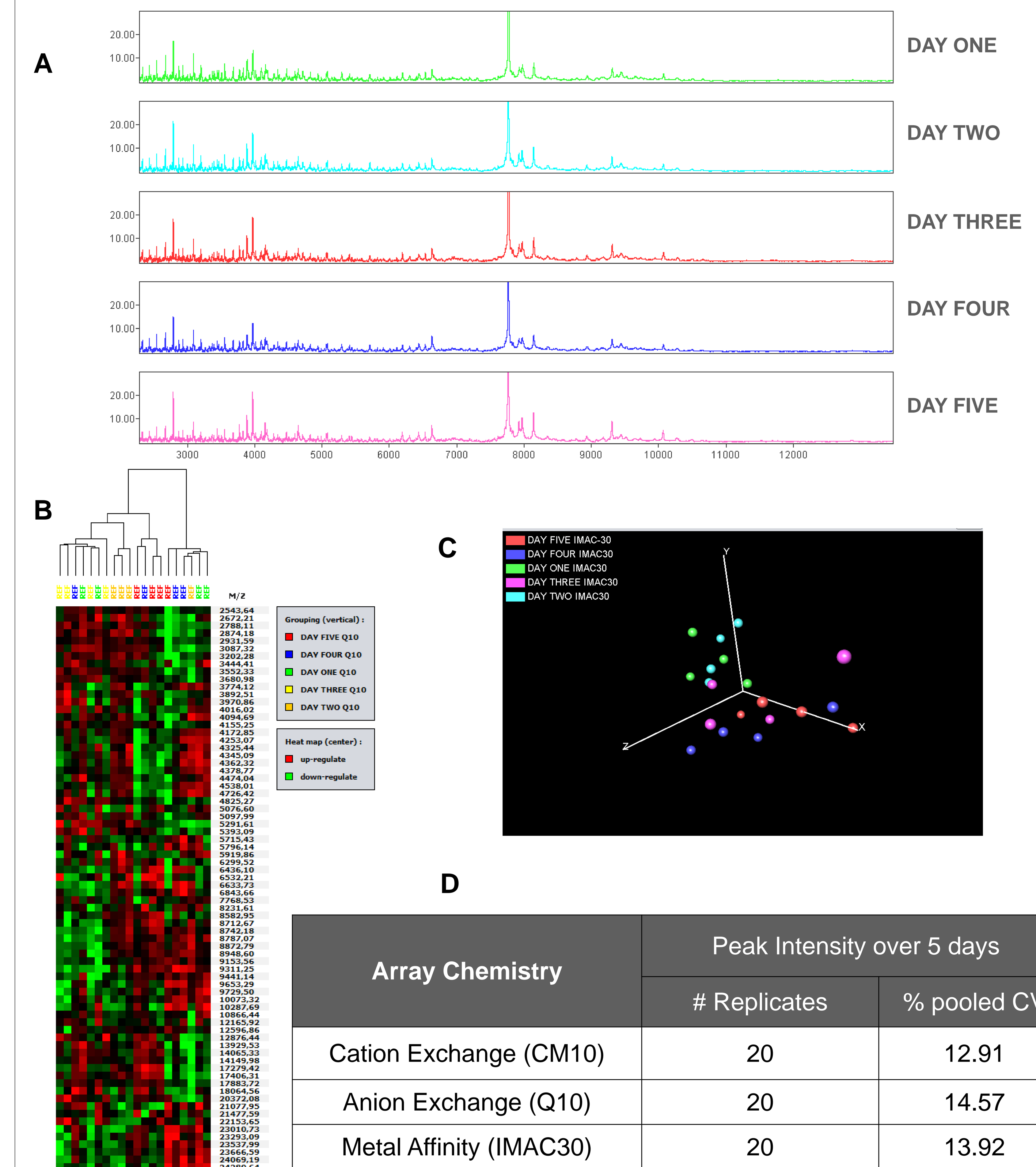
- Arrays were analyzed using an ultraflex MALDI-TOF/TOF mass spectrometer (Bruker Daltonics) in linear mode. System set-up, optimization and testing of MS parameters for reproducible protein profiling was performed prior to analysis using the Lucid System Qualification Kit (Bio-Rad). External calibration was performed with Bio-Rad's QC Peptide Array using calibrants from 1 to 12 kDa.
- Data was imported and analyzed using Lucid Proteomics Software. To calculate %CVs, automatic peak detection was performed using the Lucid Software Cluster Wizard. Within the mass range of 2.5-25 kDa, individual peaks were labeled across all spectra within a profiling condition (array type) and clustered together based on their observed m/z values. Peaks meeting the (user-defined) thresholds of minimum S/N ratio of 5.0 and valley depth of 3.0 were labeled automatically. For reproducibility, peak clusters were created when a given peak was detected in 100% of spectra. Peak intensity coefficients of variation (CVs) automatically calculated by the Lucid software for each peak cluster were exported and used to calculate the reported median CV values.



C

Array Chemistry	Number of peaks detected between 2.5 and 25 kDa (Average / day)	Peak Intensity per day	
		# Replicates	Median CV (Average/day)
Cation Exchange (CM10)	87	4	9.27
Anion Exchange (Q10)	74	4	12.52
Metal Affinity (IMAC30)	103	4	11.61

Figure 2. Intra-assay reproducibility. The reference sample was spotted 4 times per day of experiment on 4 different arrays. High profile reproducibility is shown on cation exchange (A) and anion exchange (B) surfaces. For each experimental day, median peak intensity CVs were calculated for peaks detected between 2.5 and 25 kDa. Median intra-assay CVs ranged between 9 and 13% (C).



D

Array Chemistry	Peak Intensity over 5 days	
	# Replicates	% pooled CV
Cation Exchange (CM10)	20	12.91
Anion Exchange (Q10)	20	14.57
Metal Affinity (IMAC30)	20	13.92

Figure 3. Day-to-day reproducibility. Replicates of the reference and patient samples were randomized and profiled over 5 separate but non-consecutive days. Representative spectra obtained on IMAC30 are shown (A). Non-supervised multivariate analysis tools within Lucid Proteomics Software were used to detect potential bias linked to the day of array preparation but showed no significant data clustering with Hierarchical Clustering (B) or with Principal Component Analysis (C). Peak intensity pooled CVs were calculated for each tested array surface and ranged between 13 and 15% (D).

Conclusions

Good study design and implementation of standard operating procedures ensure the level of experimental reproducibility that is required to achieve successful clinical biomarker discovery.
In this study organized over 5 non-consecutive days, monitoring variability using a reference sample and the statistical tools provided in the Lucid Proteomics Software were imperative. Very acceptable intra-assay and day-to-day CVs of 10-15% were demonstrated and no bias was detected that could compromise the discovery of scientifically sound candidate markers.

Acknowledgement

Thanks to Martin Schuerenberg from Bruker Daltonics for his valuable contributions.

Copyright © 2010 Bio-Rad Laboratories, Inc. All rights reserved.