

Proteomic Analysis of Maternal Serum in Down Syndrome: Identification of Novel Protein Biomarkers

Srinivasa R. Nagalla,^{*,‡,||} Jacob A. Canick,^{§,#} Thomas Jacob,^{||} Kimberly A. Schneider,^{||} Ashok P. Reddy,^{||} Archana Thomas,^{||} Surendra Dasari,[‡] Xinfang Lu,[‡] Jodi A. Lapidus,[†] Geralyn M. Lambert-Messerlian,^{§,#} Michael G. Gravett,[‡] Charles T. Roberts, Jr.,[‡] David Luthy,^{○,#} Fergal D. Malone,^{+,#} and Mary E. D'Alton^{+,#}

Departments of Pediatrics and Public Health and Preventive Medicine, Oregon Health and Science University, Portland, Oregon 97239, Brown University School of Medicine, Providence, Rhode Island 02903, ProteoGenix, Inc., Portland, Oregon 97213, Department of Obstetrics and Gynecology, University of Washington, Seattle, Washington 98101, Swedish Medical Center, Seattle, Washington 98122, and Columbia University College of Physicians and Surgeons, New York, New York 10032

Received October 13, 2006

Down syndrome (DS) is the most prevalent chromosomal disorder, accounting for significant morbidity and mortality. Definitive diagnosis requires invasive amniocentesis, and current maternal serum-based testing requires a false-positive rate of about 5% to detect 85% of affected pregnancies. We have performed a comprehensive proteomic analysis to identify potential serum biomarkers to detect DS. First- and second-trimester maternal serum samples of DS and gestational age-matched controls were analyzed using multiple, complementary proteomic approaches, including fluorescence 2-dimensional gel electrophoresis (2D-DIGE), 2-dimensional liquid chromatography-chromatofocusing (2D-CF), multidimensional protein identification technology (MudPIT; LC/LC-MS/MS), and MALDI-TOF-MS peptide profiling. In total, 28 and 26 proteins were differentially present in first- and second-trimester samples, respectively. Of these, 19 were specific for the first trimester and 16 for the second trimester, and 10 were differentially present in both trimesters. Analysis of MALDI-TOF-MS peptide profiles with pattern-recognition software also discriminated between DS and controls in both trimesters, with an average recognition capability approaching 96%. A majority of the biomarkers identified are serum glycoproteins that may play a role in cellular differentiation and growth of fetus. Further characterization and quantification of these markers in a larger cohort of subjects may provide the basis for new tests for improved DS screening.

Keywords: Aneuploidy • Prenatal Screening • Chrom • Pregnancy

Introduction

Down syndrome (DS) is the most frequent chromosomal disorder, with a prevalence of $1/500$ to $1/800$, depending on the maternal age distribution of the pregnant population.¹ The prevalence of DS in the second trimester is higher, because approximately one-quarter of the affected fetuses in the second trimester will not survive to term.² A majority of infants with DS have serious cardiac, gastrointestinal, or other abnormalities, and most have an IQ of less than 50, making this syndrome one of the leading causes of mental deficiency in the United States.

Definitive prenatal diagnosis of DS requires invasive testing by amniocentesis or chorionic villus sampling, techniques associated with a procedure-related risk of pregnancy loss, the most recent estimate being $\sim 1/1600$.³ Candidates for invasive testing must be at sufficient risk of having an affected pregnancy to warrant this still-significant procedure-related risk. Currently, there are a variety of prenatal screening tests that are used to identify those among all pregnant women who are at increased risk of DS. These tests use a combination of maternal serum markers, including pregnancy-associated plasma protein (PAPP-A), α -fetoprotein (AFP),⁴ human chorionic gonadotropin (hCG), unconjugated estriol (uE3), and inhibin-A.⁵ Current serum screening tests can identify 80–90% of affected pregnancies at an accompanying 5% false-positive rate.⁶ With the addition of the fetal ultrasound marker, nuchal translucency (NT), certain tests can detect 85–95% of cases at a 1–5% false-positive rate.⁷ DS screening has now been recommended for all pregnancies, regardless of maternal age.⁸

Recent advances in proteomics present new opportunities to examine the global distribution of proteins in tissues and

* To whom correspondence should be addressed. E-mail: nagallas@ohsu.edu; tel, (503)494-1928; fax, (503)494-4821.

[‡] Department of Pediatrics, Oregon Health and Science University.

^{||} ProteoGenix, Inc.

[§] Brown University School of Medicine.

[†] Department Public Health and Preventive Medicine, Oregon Health and Science University.

^{||} University of Washington.

[○] Swedish Medical Center.

⁺ Columbia University College of Physicians and Surgeons.

[#] First- and Second-Trimester Evaluation at Risk (FASTER) Research Consortium.

Table 1. Demographics of Patients with and without a DS Pregnancy at the Time of Screening^a

	first trimester		second trimester	
	DS	control	DS	control
Maternal Age (years)	36.9 ± 5.3	36.1 ± 5.1	37.3 ± 6.5	36.1 ± 5.1
Gestational Age (weeks)	12.4 ± 1.1	12.5 ± 1.1	16.1 ± 1.1	16.1 ± 0.6
Caucasian (%)	80%	70%	85%	80%
Abnormal Maternal Screen	70%	15%	70%	25%

^a A positive first-trimester maternal screen was defined as a risk at the end of pregnancy (40 weeks) of 1 in 150, and a positive maternal screen from second-trimester screening was defined as a risk at the end of pregnancy of 1 in 300¹⁰ (DS#1).

fluids. The proteins or peptides that are preferentially present and identified in a disease or pathologic state are well-suited for the development of convenient, rapid, sensitive, and specific diagnostic assays.⁹ However, to date, there has been no systematic proteomics-based attempt to identify unique biomarkers for DS in maternal serum.

In this work, we utilized four comprehensive and complementary proteomic analysis methods to characterize maternal serum proteins among a cohort of first- and second-trimester women with DS pregnancies and in gestational age (GA)-matched controls. Two global protein separation methods, fluorescence 2-dimensional gel electrophoresis (2D-DIGE) and two-dimensional liquid chromatography-chromatofocusing (2D-CF), coupled with tandem mass spectrometry, were used to detect protein abundance differences. In addition, we employed two independent methods to assess the ability to detect changes in global protein profiles through the monitoring of enzymatic digests of serum. Peptide-based analyses were chosen because the sensitivity of detection of peptides by current proteomic technology far exceeds that for the proteins from which they are derived. First, a comprehensive peptide identification strategy, multidimensional protein identification technology (MudPIT), was investigated to confirm that key biomarkers revealed from protein analyses are retained at the peptide level after tryptic digestion. In parallel, a high-throughput MALDI-MS peptide-profiling method was utilized to analyze cohorts of 20 pairs and 36 pairs of GA-matched control and DS samples from first trimester and second trimester, respectively. The resulting MALDI-MS profiles were analyzed with pattern-recognition software, and key classifiers were targeted for identification by MALDI-TOF/TOF-MS.

Materials and Methods

Samples. For this case-control proteomic analysis, sera from 56 women (including 20 in the first trimester of pregnancy at 11–13 weeks of gestation and 36 women in the second trimester of pregnancy at 15–18 weeks of gestation) carrying a DS fetus and 56 control women with non-DS pregnancies matched for GA, maternal age, and ethnicity were analyzed (Table 1). Maternal serum samples were obtained from the collection of the First- and Second-Trimester Evaluation of Risk (FASTER) clinical study¹⁰ and Women and Infants Hospital, Providence, RI. Women carrying a DS fetus, however, were more likely to have had a prior maternal screening test indicating an increased risk for DS (70% versus 20%, $p < 0.05$ by simple χ^2). Fetal DS status was confirmed either by amniocentesis, by neonatal cord blood sampling when amniocentesis was declined, or by tissue sampling in cases of miscarriage, termination, or stillbirth.¹⁰

Immunodepletion of Abundant Proteins in Human Serum.

Serum samples were depleted of six major proteins (albumin, IgG, IgA, anti-trypsin, transferrin, and haptoglobin) using the Multiple Affinity Removal System (Agilent Technologies, Inc., Palo Alto, CA). Serum (40 μ L) in Agilent buffer A was processed using Agilent immunoaffinity columns (4.6 \times 100 mm) attached to a Waters HPLC system (Waters, Milford, MA). Appropriate fractions were collected, buffer-exchanged with 10 mM Tris, pH 8.4, and concentrated using 5000 MWCO filters. Protein concentration was determined using the DC protein assay kit (Bio-Rad, Hercules, CA).

Fluorescence Two-Dimensional Differential in-Gel Electrophoresis (2D-DIGE). For this nested case-control proteomic analysis, two serum samples (one from late first trimester and one from early second trimester) from each of 20 women carrying DS fetuses and from each of 20 matched women carrying unaffected fetuses were used. Forty 2D gels were run, representing each sample pair of control and DS sera. Serum proteins (50 μ g) were labeled with CyDye DIGE Fluor minimal dye (GE Healthcare Bio-Sciences, Piscataway, NJ) at a concentration of 100–400 pmol of dye/50 μ g of protein. Samples were labeled with Cy3 (DS), Cy5 (control), or Cy2 (reference, DS + Control), and all three labeled samples were multiplexed and resolved in one gel. Labeled proteins were purified by acetone precipitation, dissolved in IEF buffer, and rehydrated onto a 24- or 13-cm IPG strip (pH 4–7) for 12 h at room temperature. The IPG strip was subjected to first-dimension electrophoresis at 65–70 kVh and then equilibrated with DTT and IAA equilibration buffers for 15 min sequentially. Second-dimension 8–16% SDS-PAGE was conducted at 80–90 V for 18 h.

Gels were scanned in a Typhoon 9400 scanner (GE) using appropriate lasers and filters with PMT voltage set between 550 and 600 V. Images in different channels (control and DS) were overlaid using pseudo-colors, and differences were visualized using ImageQuant software (GE Healthcare Bio-Sciences). 2D-DIGE gel image analysis to identify differentially abundant protein spots was performed using Phoretix 2D evolution, v.2005 (Non-Linear Dynamics, Durham, NC). A fixed area was selected from every gel, and a cross-stain analysis protocol was performed. Background subtraction was done using the ‘mode of non spot’ method, and images were wrapped to maximize the spot matching. A ratiometric normalization algorithm was applied to account for potential concentration differences in protein labeling. Normalized protein spots in the Cy5 and Cy3 channels were compared to the internal standard (Cy2) to generate a ratio of relative amount. The statistical significance of differences in the intensity of protein spots was determined using t tests on the averaged gels for each group. The protein spots with a relative ratio > 1.5 and a t -test value of < 0.05 were considered significant. For protein identification, serum proteins (1500 μ g) were subjected to 2-D GE and stained with Coomassie Blue R-250. Individual spots were cut from the gel, destained, and digested in-gel with trypsin for 24 h at 37 °C. Tryptic peptides were extracted with 0.1% TFA and purified using Zip Tip_{C18} pipet tips from Millipore (Billerica, MA).

Western Blotting. A total of 50 μ g of serum protein was subjected to 2D-GE using 13-cm IPG strips (pH 4–7) and resolved using 8–16% SDS-PAGE. Proteins were transferred to PVDF membranes (Millipore) using a trans-blot cell (Bio-Rad). Membranes were blocked in 5% milk solution and probed with antibodies against Afamin (custom-made), α -1-acid glycoprotein, gelsolin, transferrin (Dako Cytomation, Fort Collins, CO), α -2-HS-glycoprotein, complement 4, serum amyloid A (Abcam,

Cambridge, MA), apolipoprotein A (Academy Biomedical Co., Houston, TX), apolipoprotein E (Biodesign International, Saco, ME), ceruloplasmin, complement factor H (Santa Cruz Biotechnology, Santa Cruz, CA), and clusterin (Chemicon, Temecula, CA), and blots were developed using Super-signal West pico-solution (Pierce, Rockford, IL).

Two-Dimensional Liquid Chromatography-Chromatofocusing (2D-CF). For comparative analysis of protein levels in control and DS sera, four sets of pooled maternal sera ($n = 5$ in each group) were prepared from first- and second-trimester samples. All sera were immunopurified (Agilent) and buffer-exchanged for CF compatibility. Total protein (5–7 mg) was used for 2D-CF analysis for each control/DS sample pair; first- and second-trimester sample pairs were analyzed independently.

2D-CF analysis was performed on a ProteomeLab PF2D system (Beckman-Coulter, Fullerton, CA). Briefly, serum protein was loaded onto the first-dimension CF anion-exchange column and eluted as 0.3 pH-unit fractions according to protein isoelectric point (pI/pH) using a descending linear pH gradient. Each pH fraction was then separated in the second dimension by protein hydrophobicity using a nonporous C18 RP-HPLC column (48 fractions from each pH fraction). In total, 864 fractions were collected from the RP-HPLC dimension (from each sample) for trypsin digestion prior to protein identification by mass spectrometry. A differential intensity cutoff of 20% of the higher-intensity peak from either the control or DS sample identified 95 bands in the first-trimester sample set and 80 bands in the second-trimester sample set. Differential intensities between control and DS fractions ranged from 0.004 to 0.638 AU. Bands in the difference map showing up-regulation in either DS or control serum were digested with trypsin and prepared for protein identification analysis using a hybrid quadrupole time-of-flight mass spectrometer (Q-ToF-2) (Waters).

Protein Identification. Spots from 2D preparative gels and 2D-CF fractions were digested with trypsin, and samples were analyzed on a Q-ToF-2 connected to a CapLC (Waters). A survey method was used to acquire MS and MS/MS spectra. Masses from m/z 400 to 1500 were scanned for MS survey, and masses from m/z 50 to 1900 were scanned for MS/MS. Data analysis was performed using ProteinLynx Global Server v.2.1 (Waters) and *de novo* sequencing using a PEAKS algorithm combined with the OpenSea alignment algorithm (v.1.3.1).¹¹ Peptides consisting of five or more amino acids were used and matched to either a nonredundant human IPI or the Swiss-Prot database to identify the corresponding proteins. Proteins with two or more peptides by both ProteinLynx (significance score >10.6) and OpenSea (significance score >100) scoring algorithms¹² were chosen.

Sample Digestion. MALDI-TOF-MS digest profiling was performed on 56 pairs of DS and control samples. A total of 12 μg of immunodepleted serum protein from each sample was resuspended in 6 μL of a buffer containing 25 mM NH_4HCO_3 (pH 8.0), 5 mM DTT, and 12% (w/v) RapiGest (Waters). After denaturation for 30 min at 65 °C, 2 μL of sequencing-grade trypsin (0.025 $\mu\text{g}/\mu\text{L}$ in 25 mM NH_4HCO_3) was added (Promega, Madison, WI). Samples were then incubated at 37 °C for 3 h. Following digestion, 2 μL of 40 mM HCl was added to stop digestion and cleave the RapiGest agent.

For serum MudPIT analysis, 200 μg from each of 5 immunodepleted second-trimester control and DS serum samples were pooled to create a 1-mg sample of each condition. Protein

was dissolved in 100 μL of digestion buffer containing 8 M urea, 1 M Tris base, 80 mM methylamine, and 8 mM CaCl_2 (pH 8.5). For reduction and alkylation of cysteine residues, samples were first incubated at 50 °C in 12.5 μL of 0.9 M DTT for 15 min and, then in 25 μL of 1.0 M iodoacetamide in the dark at room temperature for another 15 min. Before the addition of 40 μL of mass spectrometry-grade trypsin (1 $\mu\text{g}/\mu\text{L}$) (Promega), an additional 12.5 μL of 0.9 M DTT, along with 210 μL of water and 1 N NaOH, was added to adjust the solution to pH 8.5. Samples were then thoroughly mixed and incubated overnight at 37 °C. Digestion was halted by the addition of 40 μL of formic acid. Digests were desalted prior to MudPIT analysis using C18 SepPak cartridges (Waters).

Serum MudPIT Analysis. Desalted digests (1 mL) were injected onto a polysulfoethyl strong cation-exchange (SCX) column (2.1-mm i.d. \times 100 mm, 5- μm particles and 300-Å pore size, The Nest Group, Southborough, MA) and fractionated using an HPLC equipped with a UV detector and a fraction collector. Solvent A was 5.6 mM potassium phosphate (pH 3) with 25% acetonitrile (ACN), and Solvent B was 5.6 mM potassium phosphate (pH 3) and 350 mM KCl with 25% ACN. A 95-min gradient at a flow rate of 200 $\mu\text{L}/\text{min}$ was employed for fractionation of peptides: 100% A for 10 min, ramp to 50% B over 45 min, ramp to 100% B over 15 min, and ramp back to 100% A in 0.1 min, hold at 100% A for 20 min. In total, 80 fractions were collected and stored at -20 °C. The fractions were evaporated and resuspended in 100 μL of 0.1% TFA for desalting using a 96-well spin column, Vydac C18 silica (The Nest Group). After elution in 80% ACN/ 0.1% formic acid (FA), fractions were consolidated into 43 fractions, evaporated, and resuspended in 25 μL of 5% FA.

SCX fractions (5 μL each) were analyzed on a Q-ToF-2 mass spectrometer connected to a CapLC. The Q-ToF-2 was equipped with a nanospray source. Each SCX fraction was separated using a Nanoease C18 75- μm i.d. \times 15-cm fused silica capillary column (Waters) and a 95 min water/ACN gradient. The mass spectrometer was calibrated using Glu1Fibrinopeptide B. A survey method was used to acquire spectra. Masses from m/z 400 to 1500 were scanned for MS survey and masses from m/z 50 to 1900 for MS/MS. MS/MS spectra were processed with ProteinLynx Global Server v.2.1 (Waters).

MudPIT Protein Identification and Spectral Counting. In total, 10 417 MS/MS spectra from control and 10 286 MS/MS spectra from DS were searched against a combined database containing known contaminants and forward and reverse entries of Swiss-Prot human database (version 46.6) using three independent search engines: OpenSea,^{11,12} TurboSequest (ThermoFinnigan, Waltham, MA), and X!Tandem.¹³ PEAKS software (Bioinformatics Solutions, ON, Canada) was used to supply *de novo* sequences for the OpenSea search engine. Protein identifications from individual search engine results were combined using probabilistic protein identification algorithms implemented in Scaffold software (Proteome Software, Portland, OR). In total, 52% of the spectra from the control sample and 50% of the spectra from the DS sample were assigned to proteins that were defined by at least one confident peptide (probability ≥ 0.8). Proteins having at least two independent peptide identifications (probability ≥ 0.8) were considered as highly likely to be present in the sample.

A spectral counting method was used to determine the proteins that were differentially expressed between control and DS MudPIT samples. All proteins with greater than two peptides identified with high confidence were considered for

protein quantification using spectral counting. Identified protein lists were further curated by collapsing spectral counts for similar proteins (e.g., immunoglobulins, α -1-acid-glycoproteins 1 and 2, and pregnancy-specific glycoproteins) into a single entry. Curated protein lists for both samples were merged, and an independent $2 \times 2 \chi^2$ test on the spectral counts for each protein between the samples was used to find proteins that were differentially expressed between them. To reduce the false-positive rate of differentially abundant proteins, only proteins with at least two independent peptides matched to at least four MS/MS spectra (probability ≥ 0.8) in at least one of the samples were considered as statistically significant. Fold changes of proteins passing the above criteria were determined using a published formula for calculating spectral count ratios.¹⁴

MALDI-TOF-MS Peptide Profiling. Tryptic digests were spun for 5 min to separate decomposed RapiGest, and 1 μ L of digest solution was diluted with 5 μ L of 50% ACN/0.1% trifluoroacetic acid (TFA) and 6 μ L of matrix solution (saturated α -cyano-hydroxycinnamic acid (CHCA) in 60% ACN/0.1% TFA). Samples were spotted in quadruplicate, 1 μ L on each spot, onto a 382-well ground steel target. Samples were analyzed on a MALDI-TOF mass spectrometer (AutoFLEX II TOF/TOF, Bruker Daltonics, Billerica, MA) equipped with a pulsed-ion extraction source. The Autoflex was used in reflectron mode with an accelerating voltage of +19 kV. The pulsed-ion extraction drop voltage was 2100 V with a delay time of 140 ns. Matrix ions were suppressed up to 1000 Da by deflection. The sampling rate was 2.0 GHz, and each spectrum was a sum of 2000 laser shots fired at 20 different positions. A nitrogen laser ($l = 337$ nm) operating at 50 Hz was used to irradiate samples. The output energy of the laser was $\sim 110 \mu$ J attenuated with an offset of 62% and a range of 36%. Samples were irradiated at a laser power of 8% and standards at 6%.

Spectra were automatically collected using AutoExecute acquisition control at a fixed laser power. Low-intensity fuzzy logic was used for accumulation control to avoid the addition of zero-line spectra to the sum buffer; however, sample crystallization was carefully optimized to achieve a highly uniform crystallization. Spectra were calibrated by close external calibration using Peptide calibration standard II mixture (Bruker). The same peptide calibration mixture was used in spiking experiments; peptide calibration mixture was resuspended following manufacturer's instructions (125 μ L of 0.1% TFA). For spiking experiments, 0.5 μ L (0.5 \times -spiked), 1 μ L (1 \times -spiked), or 2 μ L (2 \times -spiked) of the mixture were spiked into a sample/matrix volume of 40 μ L.

MALDI-TOF-MS Pattern-Recognition Analysis. Statistical analyses of MALDI-TOF-MS spectral data were conducted using ClinProTools (CPT) v.2.0.319 (Bruker). Following automatic normalization and recalibration of all spectra, peak areas were computed at various signal-to-noise (S/N) ratios (2, 3, 4, and 5). Genetic algorithms were used to assess classification ability of peptide profiles. Briefly, the genetic algorithm starts by randomly selecting sets of peak areas within the training data for analysis. Each chosen set is tested to see how well it discriminates cases from controls in the training set via cluster analysis (k -nearest neighbor). Patterns that successfully discriminate are kept and recombined ("mated"), while those less successful at discrimination are discarded. This is a heuristic sampling-based approach which does not necessarily find unique or optimal solutions and depends on randomly selected starting point. However, optimal (or at least near-optimal) solutions are obtained after several iterations of the algorithm.

Then, the patterns obtained via training data can be reassessed via various techniques, such as cross- or external validation. In the current experiment, default parameter settings for crossover (0.5) and mutation rate (0.2) were utilized; however, other genetic-algorithm parameters were varied to achieve optimal classification, such as the maximum number of peaks used in a model (15–25) and the number of iterations for the algorithms to run (100–1000). A k -nearest neighbor (KNN = 5) classification algorithm within the genetic algorithm was used to perform the final classification for each set of parameter settings. This allowed assessment of the robustness of the classifications obtained from the spectral data.

Results

2D-DIGE Analysis of First- and Second-Trimester Maternal Serum in DS. Two-dimensional gel electrophoresis has been widely used to characterize the serum proteome¹⁵ to identify biomarkers for cancer and other diseases. To improve the detection of low-abundance proteins, the six most abundant serum proteins (albumin, IgG, IgA, transferrin, haptoglobin, and antitrypsin), which together constitute 85% of the total serum protein, were removed.

To enhance sensitivity, reproducibility, and detection over a wide dynamic range, a multiplex proteomic analysis approach to label proteins with fluorescent cyanine dyes (Cy3, Cy5, and Cy2; DS, control, and reference, respectively) was utilized. Twenty GA-matched controls, DS, and reference (DS + control) maternal sera from first- (GA 10–13 weeks) and second-trimester (GA 16–19 weeks) samples were labeled with Cy3, Cy5, and Cy2, respectively. Each labeled GA-matched sample pair of DS and control were resolved in the same gel. Pseudocolor visualization of protein patterns, with DS spots in green and control spots in red, showed a distinct pattern of up-regulation of proteins in DS (green), with similar patterns in both trimesters (Figure 1A,B).

To quantitatively compare DS and control samples, spot quantification and statistical analysis of first- and second-trimester gels was performed as individual groups. The detection protocol identified 1816 spots from the first-trimester gel set and 1842 spots from the second-trimester gel set. After background subtraction and ratiometric normalization, matched spots across all the gels were used for statistical analysis. When the criteria of differential abundance > 1.5 -fold change and p -value < 0.05 were used, 28 and 26 spots were significant in the first- and second-trimester samples, respectively (Figure 1C). Protein identification from the gel spots identified 19 and 16 individual proteins from first- and second-trimester spots, respectively (Table 2). The levels of 10 of these proteins were elevated in both trimesters.

Protein spot in-gel digestion and identification using tandem MS/MS revealed the presence of unique or multiple proteins in each protein spot. When multiple proteins were identified from a single spot, peptide counts representing each protein were used to determine the most significant protein. As commonly seen in 2D-gel analyses, multiple protein spots also revealed the presence of a single protein, as the given protein is resolved depending on its charge state and various modifications. 2D Western blot analyses performed with specific antibodies confirmed the distribution of each unique protein (Figure 1D). The protein profiles detected by 2D Western blots matched the distribution and differential patterns detected on 2D-DIGE (Figure 1A).

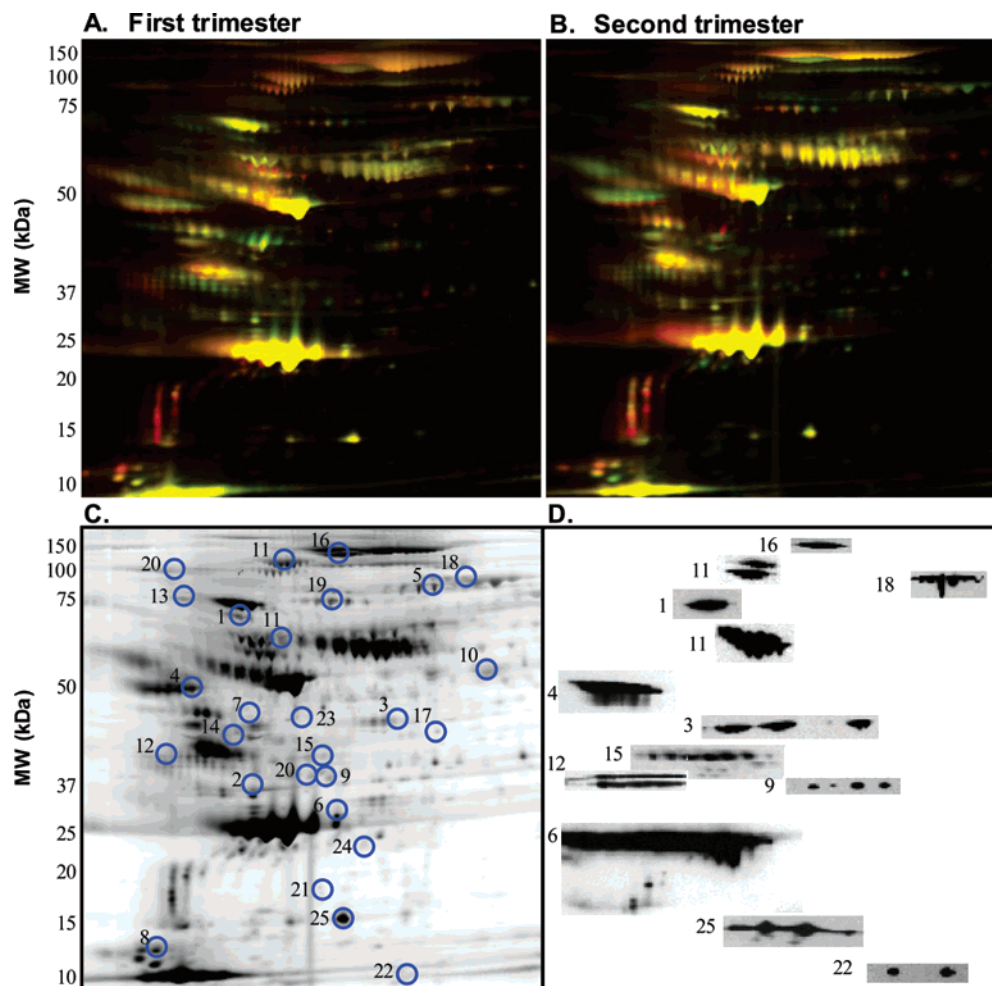


Figure 1. Proteomic analysis of DS maternal serum using 2D-DIGE. Equal amounts ($50 \mu\text{g}$) of labeled maternal serum proteins from DS and control following IEF were resolved on 8–20% linear gradient SDS-PAGE gels. (A and B) Differentially expressed proteins are shown as green (DS > control) and red (control > DS) spots for first and second trimester, respectively. (C) Image analysis and spot quantification of a first-trimester (A) 2D-DIGE gel. Spots showing statistically significant difference ($p < 0.05$) in abundance were circled and numbered. Proteins identified in these spots are listed in Table 2. (D) A composite of 2D Western blots showing the relative distribution pattern of specific proteins identified in significant spots (Table 2).

Two-Dimensional Liquid Chromatography-Chromatofocusing (2D-CF). As a complementary strategy to 2D-DIGE analyses of proteins from maternal control and DS sera, a 2D-CF method for separating intact proteins was employed. The 2D-CF methodology separates serum proteins by coupling chromatofocusing (CF) in the first dimension with reverse-phase high-performance liquid chromatography (RP-HPLC) in the second dimension.^{16,17} Chromatographic profiles resulting from separation of each first-dimension pH fraction are transformed and combined to produce virtual, gel-like, 2D maps that allow the comparison of differential protein levels between control and DS serum samples. The 2D-CF method enables the processing of higher quantities of serum protein (5–10 mg) that facilitates detection of low-abundance proteins and higher peptide recovery from non-gel embedded fractions.¹⁸

Figure 2 displays the protein expression map generated by the 2D-CF analysis of first-trimester maternal control versus DS serum. The 2D-CF map, generated using ProteoVue software (Beckman), shows the isoelectric point (pI) of the eluted protein from CF on the x -axis, with each lane representing a discrete pH fraction (typically 0.3 pH units) and the retention

time (ranging from 10 to 20 min), or hydrophobicity, of the eluted protein from RP-HPLC on the y -axis. The chromatographic traces on the far-left and far-right sides of the figure correspond to the UV (214 nm) traces from lane 8 (pH 5.2–5.5).

The difference map of the two samples shown in Figure 2 revealed the presence of multiple fractions with proteins up-regulated in the DS (green) and control (red) samples. Bands in the difference map showing up-regulation in either DS or control serum were targeted for analysis using MS/MS for the identification of the proteins found in each fraction. A differential intensity cutoff of 20% for the higher-intensity peak from either the control or DS sample resulted in 95 bands in the first-trimester sample set and 80 bands in the second-trimester sample set. Differential intensities between control and DS fractions ranged from 0.004 to 0.638 AU (limit of detection for MS analysis of fraction digests is conservatively ~ 0.05 AU; the AU scale for the second-dimension separations reaches a maximum of ~ 1.3 AU).

Table 3 shows a list of identified proteins found in differential 2D-CF fractions of control versus DS serum. In total, 25 proteins were identified as differentially abundant in DS serum. Six

Table 2. Maternal Serum Proteins Differentially Present in First- and Second-Trimester DS Detected by 2D-DIGE

spot no.	Swiss-Prot ID	Swiss-Prot entry name	protein name	first trimester		second trimester		function
				fold change (DS/control)	p-value	fold change (DS/control)	p-value	
1	P43652	AFAM_Human	Afamin	6.84	0.000	1.89	0.010	Carrier activity
2	P02760	AMBP_Human	Alpha-1-microglobulin			3.2	0.000	Anti-inflammatory response/ cell adhesion
3	P02763	A1AG1_Human	Alpha-1-acid glycoprotein 1	1.92	0.010			Acute-phase response
4	P02765	FETUA_Human	Alpha-2-HS-glycoprotein	3.90	0.000	2.6	0.080	Acute-phase response
5	P01023	A2MG_Human	Alpha-2-macroglobulin			2.48	0.010	Protein homooligomerization
6	P02647	APOA1_Human	Apolipoprotein A-I	2.10	0.010			Cholesterol metabolism
7	P06727	APOA4_Human	Apolipoprotein A-IV	3.82	0.040			Lipid metabolism
8	P02655	APOC2_Human	Apolipoprotein C-II	-2.00	0.000	-2.76	0.000	Lipid metabolism
9	P02649	APOE_Human	Apolipoprotein E	2.36	0.000			Lipid metabolism
10	P02749	APOH_Human	Beta-2-glycoprotein 1	1.80	0.030	1.97	0.020	Heparin binding
11	P00450	CERU_Human	Ceruloplasmin	1.89	0.040	2.13	0.040	Iron homeostasis
12	P10909	CLUS_Human	Clusterin	2.51	0.020			Complement activation
13	P09871	C1S_Human	Complement C1S component	3.85	0.020			Complement component C1s activity
14	P01024	CO3_Human	Complement C3			2.39	0.010	Immune response
15	P01028	CO4A_Human	Complement C4	2.22	0.010			Complement activation
16	P08603	CFAH_Human	Complement factor H	1.80	0.090	1.57	0.030	Complement activation
17	O75636	FCN3_Human	Ficolin 3	-3.67	0.010			Sugar binding
18	P06396	GSN_Human	Gelsolin			1.53	0.040	Actin/calcium ion binding
19	P19823	ITIH2_Human	Inter-alpha-trypsin inhibitor heavy chain H2	2.33	0.000	1.59	0.030	Endopeptidase inhibitor
20	Q14624	ITIH4_Human	Inter-alpha-trypsin inhibitor heavy chain H4	2.67	0.010	1.69	0.040	Endopeptidase inhibitor
21	P36955	PEDF_Human	Pigment epithelium-derived factor	4.52	0.030	1.75	0.040	Cell proliferation
22	P02735	SAA_Human	Serum amyloid A			2.71	0.010	Acute-phase response
23	P04278	SHBG_Human	Sex hormone-binding globulin	2.69	0.020			Hormone transport
24	P05452	TETN_Human	Tetranectin			4.91	0.010	Skeletal development
25	P02766	TTHY_Human	Transthyretin	2.00	0.000	1.83	0.000	Transport

proteins in the first trimester and 6 proteins in the second trimester were uniquely present, and 13 proteins were detected in both trimesters. 2D-CF experiments yielded results complementary to the 2D-DIGE experiments. Additionally, 2D-CF identified smaller differentially expressed proteins (e.g., apolipoproteins) that are difficult to separate and visualize using 2D gel conditions. Furthermore, as a result of the all-liquid fractionation, better sequence coverage was obtained for many of the identified proteins in 2D-CF versus the 2D gel experiments. Proteins found in 2D-CF differential bands confirmed protein identifications from gel spots and areas of interest from the 2D-DIGE. Eleven proteins with differences in relative abundance in DS maternal serum were detected by both the 2D-DIGE and 2D-CF approaches. Three proteins (A2HS, CFH, and ITIH4) showed a difference in abundance in both trimesters and in both methodologies. The relative abundance of apolipoprotein CII was lower in DS by both 2D-2G and 2D-CF.

Serum MudPIT Analysis. Increasing confidence in mass spectrometry-based peptide identification and quantification methods has launched the development of extensive and varied multidimensional peptide separations coupled with MS/MS. Such “shotgun” peptide sequencing endeavors produce reliable protein identifications as well as relative quantitative information for comparing sets of samples analyzed in parallel. For the analysis of protein expression, immunodepleted control and DS sera (second trimester) were digested with trypsin and subjected to MudPIT analysis.

One hundred proteins with high-confidence identification (2 or more peptides/protein) derived from the MudPIT analysis were chosen to perform relative quantification using spectral counts. To decrease false-positive protein identification rates, MS/MS spectra were searched against a database containing known contaminants (i.e., trypsin, keratin, and serum albumin) and both forward and reverse peptide sequence entries from the Swiss-Prot human database using three independent search engines. A probability-based algorithm, Scaffold, was used to

combine results from the three search engines. The use of multiple searching algorithms increases the confidence in reported identifications by decreasing peptide identifications occurring by chance. A very high percentage of the acquired MS/MS spectra were assigned to proteins with at least one confident peptide identification (52% and 50% of control and DS spectra, respectively).

For quantitative comparison of control and DS samples, a spectral counting method was implemented. Spectral counting permits rapid detection of abundance differences between two sample pools without resorting to complicated differential labeling experimentation.¹⁹ Curated protein lists from control and DS were compiled, and independent χ^2 tests on the spectral counts of each protein were performed. All proteins with a calculated χ^2 value over 2.906 (90% confidence interval) are reported in Table 4. Included in the table are the spectral counts, the number of unique MS/MS peptide spectra matching to the given protein, for the control and DS samples. The fold change between control and DS for each of the significant proteins was calculated.¹⁴ Of the 14 proteins found to be differentially abundant between control and DS by spectral counting, 7 proteins had χ^2 values in the 99% confidence interval, and an additional 5 proteins had χ^2 values in the 95% confidence interval. When compared to comprehensive quantitative proteomic studies performed using 2D-DIGE and 2D-CF protein separations, 9 proteins found by spectral counting correspond to differential trends seen in either 2D-DIGE or 2D-CF experiments. The identification of potential lower-abundance serum protein markers is one of the benefits of MudPIT analysis. The multidimensional front-end peptide separations (SCX and RP-LC) permit the interrogation of a wider dynamic range of concentration than gel-based proteomic analyses or MALDI profiling technologies.

Proof-of-Principle Experiments for Complex Digest Peptide Profiling. Prior to the evaluation of the entire set of paired

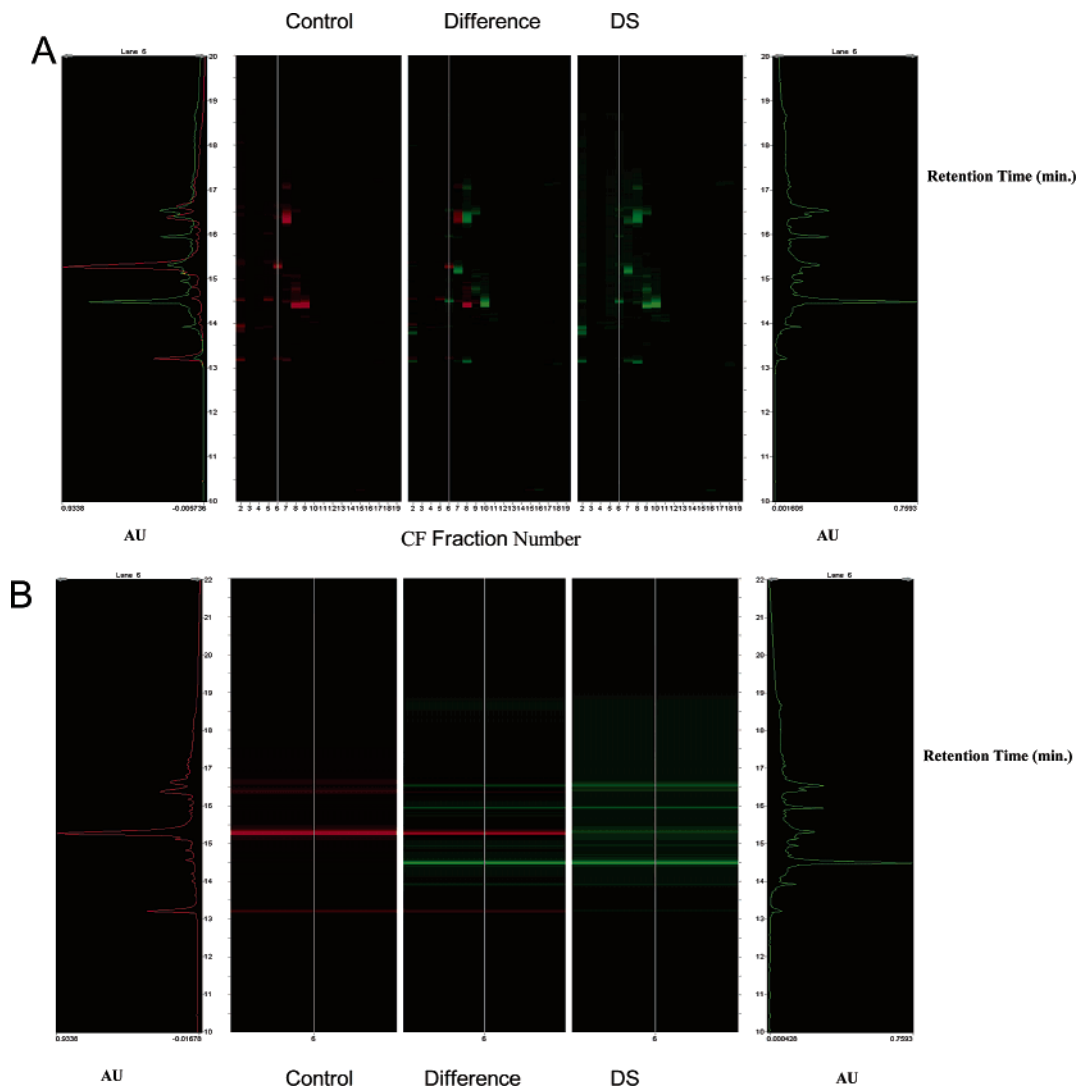


Figure 2. (A) Proteomic analyses of maternal serum in DS using 2D-CF. Protein elution maps of first-trimester control (red) and DS (green) are shown on the left and right, respectively. The combined map of control and DS showing the difference profile is shown in the center. First-dimension CF fraction number (lane 2 is a NaCl fraction, and lanes 3–19 are pH fractions of ~ 0.3 pH units from 3.9 to 8.8) is plotted on the x-axis against chromatographic retention time on the y-axis. Peak profiles on either end of the schematic are RP-HPLC chromatograms of the highlighted lane in the difference map (lane 6, pH 4.6–4.9). The left-most profile overlays the chromatograms from the control (red) and DS (green) for the pH fraction highlighted in the difference map. (B) Difference map of lane 6 (pH 4.6–4.9). Differentially abundant protein fractions were shown as red (control > DS) or green (DS > control).

control and DS sera, several proof-of-principle experiments were performed to evaluate the reproducibility of profiling complex MALDI digests. The reproducibility of MALDI spectra is highly dependent upon uniform matrix crystallization. Slight changes in matrix or sample concentration easily lead to the formation of sweet spots and irregular crystallization. Consequently, the amount of sample spotted on the target ($0.1 \mu\text{g}$), the solvent system for dilution, and the matrix concentration were thoroughly optimized to achieve the most reproducible profiles. Uniform crystallization allowed consistent total ion signal collection at a fixed laser power and a fixed number of positions per sample spot.

The reproducibility of MALDI–MS measurement of the same sample was measured by calculating the coefficient of variance (CV) for the peak areas and intensities of selected peaks (denoted with asterisks in Figure 3) of varying intensities and signal-to-noise ratios (S/N) for four spotting replicates (Table 5). The CV for peak area ranged from 6% to 12% (average CV

of 9.8%) with better-resolved and more intense peaks displaying less variance. To assess the reproducibility of sample preparation and tryptic digestion, three aliquots of the same sample were digested. The average CV for peak area between digestions of the same sample was 12.9% (Table 5). This value is expected from the MALDI process and is acceptable for semiquantitative profiling analysis.

The tryptic digest profile of immunodepleted serum from m/z 1000–3000 is highly complex, with peptide signals encompassing nearly every available position (Figure 3A, pooled human serum control digest). To assess the ability to detect small changes in expression in an extremely complex digest, a classic spiking experiment was performed. A pooled control digest was created, and varying amounts ($0.5\times$, $1\times$, and $2\times$) of an 8-peptide standard mixture was incorporated to generate four different sample classes. In total, 32 replicates were generated for each class to permit analysis using CPT software.

Table 3. Maternal Serum Proteins Differentially Present in First- and Second-Trimester DS Detected by 2D-CF

Swiss-Prot ID	Swiss-Prot entry name	protein name	DS/control		function
			first trimester	second trimester	
P02763	A1AG1_Human	Alpha-1-acid glycoprotein 1	↑		Acute-phase response
P19652	A1AG2_Human	Alpha-1-acid glycoprotein 2	↑		Acute-phase response
P04217	A1BG_Human	Alpha-1B-glycoprotein	↑	↑	Unknown
P02750	A2GL_Human	Leucine-rich alpha-2glycoprotein		↑	Unknown
P43652	AFAM_Human	Afamin		↑	Carrier activity
P02647	APOA1_Human	Apolipoprotein A-I	↑	↑	Lipid metabolism
P06727	APOA4_Human	Apolipoprotein A-IV	↑	↑	Lipid metabolism
P02655	APOC2_Human	Apolipoprotein C-II		↓	Lipid metabolism
P02656	APOC3_Human	Apolipoprotein C-III	↑	↑	Lipid metabolism
P05090	APOD_Human	Apolipoprotein D	↑		Lipid metabolism
P02649	APOE_Human	Apolipoprotein E	↑		Lipid metabolism/Cholesterol homeostasis
P08603	CFAH_Human	Complement factor H	↑	↑	Complement activation
P05156	CFAI_Human	Complement factor I	↑		Complement activation
P10909	CLUS_Human	Clusterin	↑	↑	Complement activation/lipid metabolism
P01024	CO3_Human	Complement C3	↑	↑	Immune response/ signal transduction
P02765	FETUA_Human	Alpha-2-HS-glycoprotein	↑	↑	Acute-phase response
P04196	HRG_Human	Histidine-rich glycoprotein	↓	↓	Unknown
Q14624	ITIH4_Human	Inter-alpha-trypsin inhibitor heavy chain H4	↑	↑	Endopeptidase inhibitor
P01042	KNG1_Human	Kininogen 1	↑	↑	Inflammatory response
P00747	PLMN_Human	Plasminogen	↑		Negative regulation of cell proliferation
P11464	PSG1_Human	Pregnancy-specific beta-1 glycoprotein 1		↑	Pregnancy
P02753	RETBP_Human	Plasma retinol-binding protein		↑	Retinoid binding
P04278	SHBG_Human	Sex hormone-binding globulin		↑	Hormone transport
P05452	TETN_Human	Tetranectin	↑	↑	Skeletal development
P02766	TTHY_Human	Transthyretin	↑	↑	Transport

Table 4. Quantitative Analysis of MudPIT Data Using Spectral Counting

Swiss-Prot ID	protein symbol	protein name	spectral counts		χ^2	fold change
			DS	control		
P02763	A1AG	α -1-acid glycoprotein	167	97	23.2	1.6
P02735	SAA	Serum amyloid A	12	0	13.7	11.2
Q14624	ITIH4	Inter- α -trypsin inhibitor heavy chain H4	90	60	7.88	1.6
P01019	ANGT	Angiotensinogen	78	50	7.86	1.6
P02654	APOC1	Apolipoprotein C-I	21	8	6.59	2.5
P00915	CAH1	Carbonic anhydrase I	1	9	5.23	-4.3
P02751	FINC	Fibronectin	35	20	4.99	1.8
P02775	SCYB7	Platelet basic protein	6	1	4.65	3.4
P02766	TTHY	Transthyretin	12	5	3.74	2.2
P07358	CO8B	Complement component C8 beta chain	7	2	3.69	2.7
Q13046	PSG	Pregnancy-specific glycoproteins	7	2	3.69	2.7
P25311	ZA2G	Zinc-2-glycoprotein precursor	7	14	3.24	-1.8
P05090	APOD	Apolipoprotein D	5	1	3.02	2.9
P04278	SHBG	Sex hormone-binding globulin	13	6	2.99	2.1

Representative spectra for classes used in model generation are shown in Figure 3.

Using the CPT genetic algorithm, a model was created using the unspiked class versus the 1 \times -spiked class that achieves a recognition capability and cross-validation of 100% for both classes. Furthermore, the unspiked versus spiked model classifies both spiked classes of lower concentration (0.5 \times -spiked) and higher concentration (2 \times) as belonging to the spiked class. Thus, as expected, the algorithm is easily able to create a model discriminating between classes containing peptide markers that are present or absent in one class versus another. In addition,

this model demonstrates that variation in the concentration of the markers does not inhibit correct classification. For the assessment of the ability of the algorithm to detect small changes in concentration in a complex mixture not containing unique peptide markers, a CPT model was also generated using the 0.5 \times -spiked class versus the 1 \times -spiked class. Similarly, this model demonstrated 100% cross-validation and 100% recognition for both classes. This model classified the unspiked class as 0.5 \times -spiked and the 2 \times -spiked class as 1 \times -spiked, demonstrating the ability of the algorithm to discriminate classes based solely upon changes in concentration. Figure 3D displays CPT clustering diagrams of two peak masses utilized by the CPT genetic algorithm to discriminate between classes. The left panel displays the diagram for the unspiked versus spiked model, and the right panel displays the diagram for the 0.5 \times -spiked versus 1 \times -spiked model. For each sample spectrum, normalized peak area for m/z 1759 is plotted against the normalized peak area for m/z 2093. While the diagrams only depict two of the peptides used in model generation, the separation between classes is clearly evident.

MALDI-TOF-MS Profiling of Serum Digests for Discrimination between Control and DS. Since MALDI-TOF-MS peptide profiling is a high-throughput technique requiring small amounts of serum protein (~0.1 μ g for each MALDI spot), cohorts of 20 and 36 GA-matched immunodepleted and digested serum pairs were profiled from first and second trimesters, respectively. Spectral data from triplicate analyses for each digested sample were imported into CPT pattern-recognition software, and individual peptide peak statistics were generated at several S/N levels (2, 3, 4, and 5). Figure 4A displays the average class digest profile spectra for control (red) and DS (green) for both first and second trimesters.

Models were generated using the CPT genetic algorithm varying the maximal number of peaks permitted for model

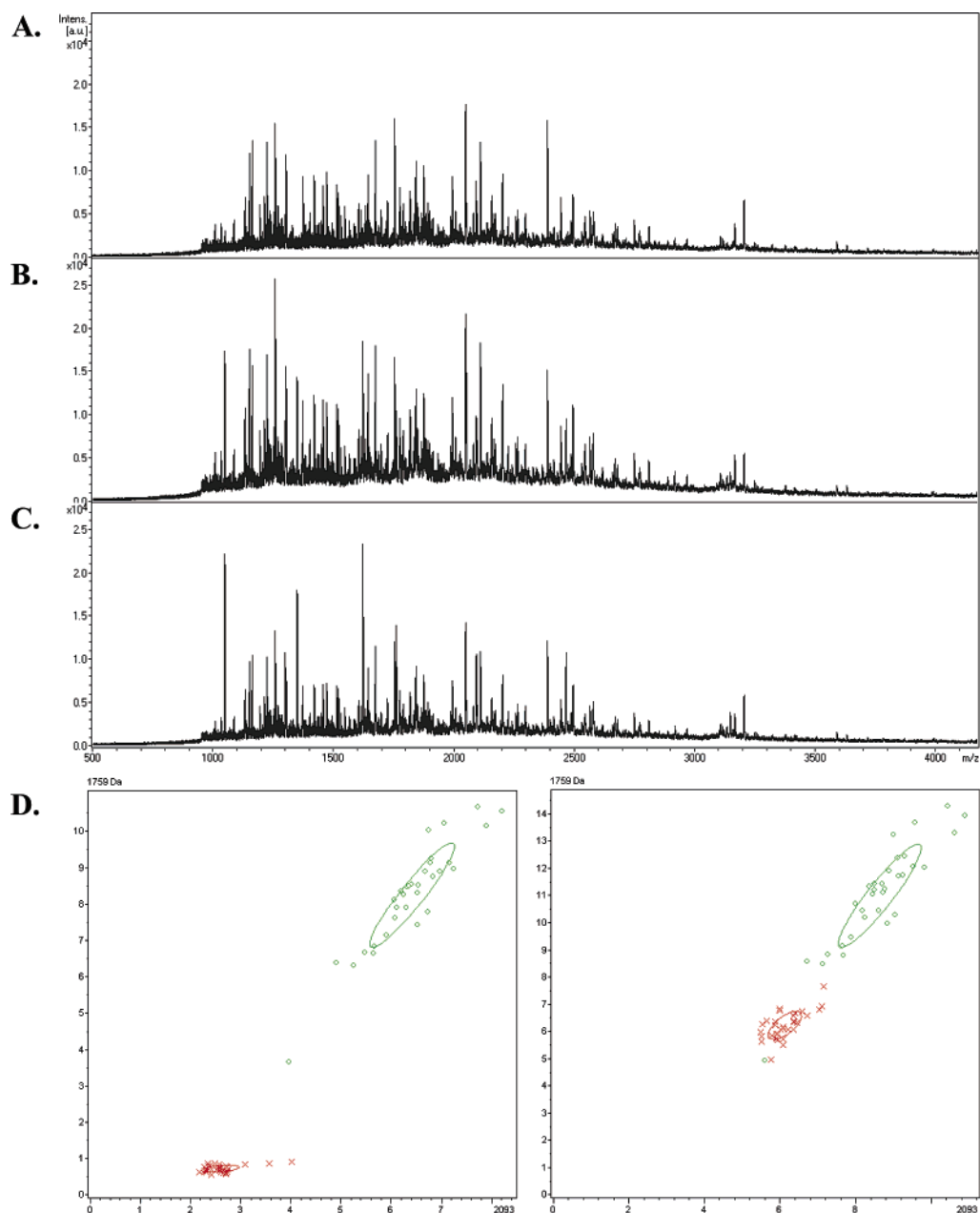


Figure 3. MALDI-TOF-MS profiling spectra of depleted serum tryptic digests. (A) Profile of unspiked human serum pooled control (HSC); (B) profile of HSC digest spiked with 0.5 \times peptide standards; (C) profile of HSC digest spiked with 1 \times peptide standards; (D) CPT cluster diagrams of two classifier peaks (m/z 1759 vs 2093) plotted against each other. The left panel is the cluster diagram created when plotting the unspiked class (A, in red) versus the 1 \times -spiked class (C, in green). The right panel is the cluster diagram created when plotting the 0.5 \times -spiked class (B, in red) versus the 1 \times -spiked class (C, in green). Circles surrounding the clusters show the standard deviation.

generation (15, 20, 25). Model generation for first and second trimesters was performed separately. In CPT, the discriminating power of calculated models is reported as the recognition capability. The recognition capability is the percentage of sample spectra used to generate the model that are classified as true. The average recognition capability over 12 calculated models was 96.9% for control and 95.5% for DS in first trimester, and 95.5% for control and 95.5% for DS in second trimester.

Table 6 displays a list of peptides commonly used as classifiers in models generated by CPT. LC-MS/MS analysis of pooled tryptic digests of control and DS serum identified 11 proteins corresponding to masses of CPT peptide classifiers.

Seven of the identified peptides correspond to proteins that were classified as differentially present in serum MudPIT analysis by spectral counting. As expected, MALDI-MS peptide profiling was useful at exploiting abundance differences in proteins that are themselves among the most abundant in serum. The seven proteins found using MALDI profiling correspond to proteins with the highest number of spectral counts in the MudPIT.

The protein with the greatest difference in normalized average area between control and DS in both trimesters by MALDI-MS profiling was α -1-acid glycoprotein (A1AG). Figure 4B displays the differential class intensities for peak 1160.6 Da (MH^+), which was identified by both MALDI-TOF/TOF-MS

Table 5. Assessment of Variability in MALDI-TOF–MS Measurement of Spotting Replicates and Tryptic Digest Preparation

<i>m/z</i> (SD, Da)	<i>n</i>	within-run		<i>n</i>	between-run	
		mean peak area	CV (%)		mean peak area	CV (%)
1148 (0.1)	4	11694	12.05	12	15216	11.20
1255 (0.1)		16587	11.84		23942	23.20
1511 (0.1)		9287	12.07		6115	15.43
1753 (0.1)		20269	6.98		12053	5.83
1872 (0.2)		13837	10.44		14604	11.98
2049 (0.2)		21784	8.25		32619	16.54
2202 (0.2)		11223	6.18		55043	9.54
2388 (0.4)		20867	10.09		7872	12.30
2578 (0.3)		7488	10.34		5211	10.22

(Figure 4C) and by LC–MS/MS as A1AG. Furthermore, A1AG was one of the proteins with one of the highest χ^2 expectation values ($\chi^2 = 23.2$) using MudPIT spectral counting.

Discussion

Prenatal DS screening has evolved over the last few decades through the sequential addition of maternal serum analytes (currently comprised of PAPP-A, β -HCG, AFP, uE3, and inhibin-A) and ultrasound-based procedures in order to reduce the need for invasive, but definitive, tests such as amniocentesis. However, because of the low occurrence of DS (13.65 per 10 000 births),²⁰ current screening procedures still report high numbers of false-positive cases relative to true-positive cases in order to attain high detection rates.^{10,21} Among women who screen positive for fetal DS, follow-up amniocentesis causes fetal loss in 1 out of every 1600 cases,³ a rate significantly lower than previous estimates,²² but one which still represents a serious concern. Consequently, the need for further development of noninvasive prenatal DS screening still exists.

One alternative to the use of invasive procedures is the analysis of fetal cells²³ or nucleic acids, including DNA and RNA,^{24,25} in the maternal circulation. While these approaches hold some promise, their utility is presently limited by a number of factors,²⁶ including the sporadic presence and/or nonproliferative nature of fetal cell types in maternal serum, the difficulty in distinguishing maternal and female fetal material, and the small amounts of fetal nucleic acids present in maternal serum.

Advances in proteomic separation technology over the past decade have made the discovery of proteins differentially expressed between samples almost commonplace. Comprehensive top-down (2D-DIGE and 2D-CF) and bottom-up (MudPIT and MALDI-TOF–MS peptide profiling) proteomic evaluation of control versus DS maternal serum has identified several potential new serum protein biomarkers differentially abundant in first-trimester and second-trimester (Figures 1, 2, and 4; Tables 2, 3, 4, and 6). Nine proteins were identified as potential candidate biomarkers by multiple (2 or 3) proteomic approaches used in all four studies.

A majority of the serum proteins differentially present in DS are glycoproteins. Altered protein glycosylation patterns in pregnancy have been noted for α 1-acid glycoprotein, α 2-HS glycoprotein,²⁷ corticosteroid binding globulin,²⁸ hCG,²⁹ and thyroxin binding globulin.³⁰ The candidate serum biomarker proteins detected can be categorized into three major functional groups: protease inhibitors, acute-phase response proteins, and serum carrier proteins. Many of the proteins

identified have been previously implicated as playing a role in developmental disorders and DS.

The protease inhibitors α -2-antiplasmin precursor (A2AP), antithrombin-3 (ANT3), and α -2-macroglobulin precursor (A2MG) are involved in the clotting mechanism and have been reported to have a role in preeclampsia.³¹ A2MG is a multifunctional protein that inhibits both self and foreign proteases³² and is part of the acute-phase response. A2MG, PAPP-A, and PZP act as carriers for IgG molecules, and these complexes may serve an immunoregulatory function and have been shown to be associated with familial Alzheimer's disease.^{33,34} The acute-phase response modulators inter- α trypsin inhibitor H4 (ITIH4),³⁵ α -1-acid glycoprotein (A1AG), and serum amyloid A (SAA) have been found to have a role in DS. A1AG expression is altered on the surface of T cells in DS patients,³⁶ and elevated serum SAA levels have been observed in DS.³⁷

Other acute-phase proteins identified belong to the complement system proteins B, C, and H of the alternate pathway. Complement system proteins and clusterin were found to colocalize with amyloid beta plaques in the brains of Alzheimer's and DS subjects, and studies indicate a role for acute-phase proteins in amyloid deposition.³⁸

In the serum carrier proteins group, afamin and vitamin D binding proteins are carriers for the fat-soluble vitamins E and D, respectively. Ceruloplasmin transports copper, and its impaired transport has been associated with poor outcome of pregnancy. The role of transthyretin in transport of thyroxine (T4) and its altered structure (misfolding) has been suggested to play a role in aggregation leading to the formation of amyloid plaques that contribute to senile systemic amyloidosis (SSA),³⁹ amyotrophic lateral sclerosis (ALS),⁴⁰ and familial amyloidotic polyneuropathy (FAP).⁴¹ Changes in apolipoprotein expression are known to affect pregnancy, and particular polymorphisms correlate with an exacerbation of neurodegenerative disease. Both apolipoprotein C and E levels are altered in preeclampsia. Many investigations have suggested a role for ApoE in Alzheimer's disease. The ApoE4 isoform, in particular, has been shown to promote the deposition of amyloid beta peptides.⁴² By analogy with AD, increased levels of acute-phase proteins may correlate with amyloid depositions that occur in DS.

It is important to note that the current serum markers used for screening were not detected as being differentially present in all four proteomic approaches. This could be for several reasons. The current serum markers were not discovered using a systematic proteomics-based investigation (i.e., any of the approaches used in our studies). They were, instead, developed based on the adoption of available immunoassays and the screening of large cohorts for the design of appropriate statistical algorithms to predict risk. Immunoassay quantification is relative, significantly depends on particular antibody characteristics, and may not reflect absolute differences in the levels of analytes in serum. In contrast, the four independent proteomic approaches used in this study (2D-DIGE, 2D-CF, 2D-LC–MS/MS, and MALDI-TOF–MS) represent a comprehensive analysis of DS serum that relies on the absolute quantities of the analytes. Furthermore, the global proteomic approaches used provide a comprehensive view, and significant biomarker candidates are determined based on the detected proteins (proteome coverage). Because of the diverse chemical and structural characteristics of serum proteins, some are more amenable than others for detection by mass spectrometry, and low-abundance proteins suffer from limited coverage and, therefore, seldom become statistically significant in global

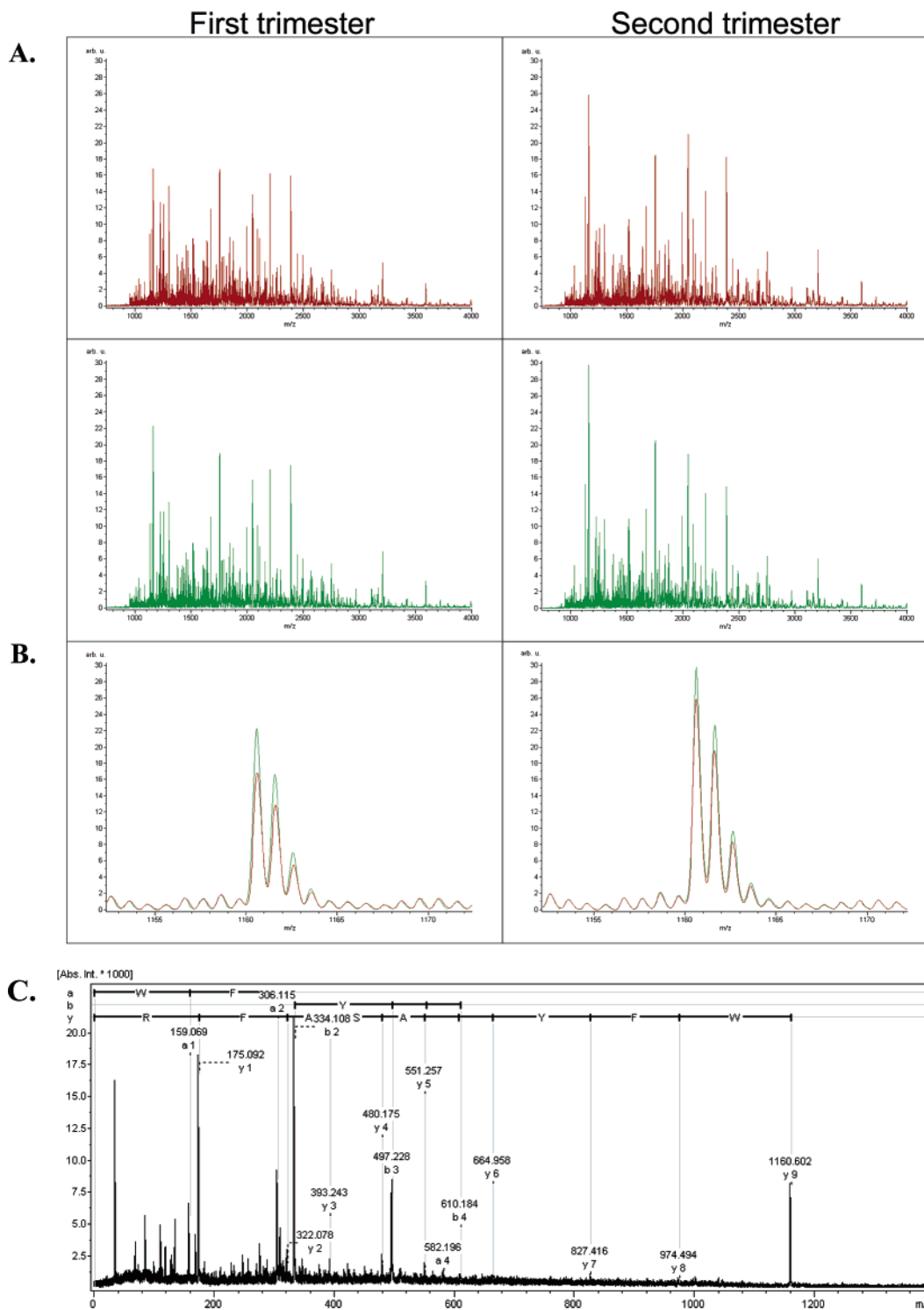


Figure 4. MALDI-TOF-MS profiling of digested immunodepleted maternal control and DS sera. (A) The average MALDI profiles for control serum (red) and DS serum (green) for both trimesters; (B) one of the peptides differentially present between control and DS profiles in both trimesters; (C) the MALDI-TOF/TOF spectrum for the same peptide (m/z 1160.6) identified as alpha-1-acid-glycoprotein.

proteome analytical approaches. One could envision further enrichment of the low-abundance serum proteome and alternative protein separation/proteomic approaches that could potentially detect other candidates in the subnanomolar category.

The identification of new DS markers by proteomic approaches offers the possibility of further improving the performance of DS screening. While the current best methods

using serum markers are a vast improvement over what was available 20 years ago, the ability to identify upward of 90% of affected pregnancies while further reducing the false-positive rate is an important and continuing goal in prenatal screening research. In the context of screening for DS, a false-positive rate as low as 5% still means that a substantial number of women with unaffected pregnancies will be offered invasive and risky diagnostic procedures. For example, if 10 000 women

Table 6. Identification of Peptides Used for MALDI-TOF–MS Peptide Profiling Classification^a

peptide mass (MH+), Da	peptide sequence	Swiss-Prot entry name	protein description
1160.6	WFYIASAFR	P02763	α-1-acid glyco-protein 1, 2
1753.1	YVGGQEHEFAHLLILR	P02763	α-1-acid glycoprotein 1
2487.1			
1129.6			
2659.2			
1301.7	THLAPYSDELRL	P02647	Apolipoprotein A-I
2749	YFKPGMPFDLMVF-VTNPDGSPAYR	P01024	Complement C3
1220.6	NFPSPVDAAFR	P02790	Hemopexin
2202.1	LREQLGPTQEFW-DNLEK	P02647	Apolipoprotein A-I
1529.8			
2225.1	ALEILQEEDLID-EDDIPVR	P01028	Complement C4
1234.7			
1646.9	DVWGIEGPIDAAFTR	P04004	Vitronectin
1497.7	AAMVGMLANFLGFR	P01019	Angiotensinogen
1519.8	ALYLQYTDETFR	P00450	Ceruloplasmin
1436.8			
1371.7	QYPLSIEPIGVR	P00450	Ceruloplasmin
1381.7	SFFPENWLWR	P01028	Complement C4
1030.5	YFIDFVAR	P01042	Kininogen
1641.8	AGDFLEANYMNLQR	P01024	Complement C3
2444.5	PWQDLVVLPLSITTD-FIPSR	P01024	Complement C3
1819			
1491.8	NQGNTWLTAFLVK	P01023	α-2-macroglobulin
1875			
1872.1			
1774	STVGNSNNYLHLSVLR	P01024	Complement C3
1478.7	YDQSATALHFLGR	P01019	Angiotensinogen
2296.2	DYNLNDILLQLGIEEA-FTSK	P01011	α-1-antichymotrypsin
1772.1	STVGNSNNYLHLSVLR	P01024	Complement C3
1454.8	FVELTMPYSVIR	P01023	α-2-macroglobulin
1604.7	IAQWQSFQLEGGLK	P01023	α-2-macroglobulin
1148.7	QGIPFFGQVR	P01023	α-2-macroglobulin
2110.1			
2049.1			
2045.1	LLLQVSLPELPG-EYSMK	P01023	α-2-macroglobulin
2388.5			

^a Peptide sequence and protein identifications are from LC–MS/MS analysis of second-trimester immunopurified digested control and/or DS sera.

were screened using the best current serum screening protocol, about 17 of the expected 20 DS pregnancies would be detected, but an additional 500 unaffected women would be found at increased risk and would be offered invasive testing. Clearly, there is room for improvement in serum screening.

The comprehensive proteomic analyses of maternal serum performed in this study provide a basis to further develop a global DS serum peptide database for targeted detection and monitoring of DS status. At present, the large dynamic range of serum proteins prevents the monitoring of all proteins utilizing a single proteomic strategy. However, with increasing confidence in mass-spectrometry-based peptide identification and quantification, the ability to monitor protein biomarkers of disease from mixtures of enzymatic serum digests is quickly reaching fruition. Presently, most proteomic mass spectrometric quantification is based upon relative quantitation of control versus perturbed sample sets. However, LC–MS/MS quantitative assays for proteins from digested plasma have been performed using a single digest peptide and an isotopically labeled counterpart.^{43,44} Furthermore, established MS methods using triple quadrupole MS (TQMS) offer great promise for multiplexing several peptide targets into a single assay.⁴⁵

In this report, numerous proteomic techniques identified several proteins as differentially abundant in control versus DS serum by relative protein and peptide quantitation. The further development and accurate quantitation of selected biomarkers identified in this study could lead to a more sensitive, noninvasive, and accurate test for detecting fetal DS using maternal serum.

Acknowledgment. This work has been supported by the National Institutes of Health (Grant No. R44HD046288 to S.R.N.) and, in part, by ProteoGenix, Inc. Oregon Health & Science University and Drs. Gravett, Roberts, and Nagalla all have a significant financial interest in ProteoGenix, Inc., a company that may have a commercial interest in the results of this research and technology. This potential conflict of interest has been reviewed and a management plan approved by the OHSU Conflict of Interest in Research Committee. Drs. Canick, Luthy, and D'Alton have received a one-time consulting fee from ProteoGenix, Inc., for expert advice on DS screening approaches.

References

- Egan, J. F.; Benn, P. A.; Zelop, C. M.; Bolnick, A.; Gianferrari, E.; Borgida, A. F. Down syndrome births in the United States from 1989 to 2001. *Am. J. Obstet. Gynecol.* **2004**, *191* (3), 1044–8.
- Cuckle, H. Epidemiology of down syndrome. In *Screening for Down Syndrome in the First Trimester*; Grudzinkas, J., Ward, R., Eds.; RCOG Press: London, U.K., 1997; pp 3–13.
- Eddleman, K. A.; Malone, F. D.; Sullivan, L.; Dukes, K.; Berkowitz, R. L.; Kharbutli, Y.; Porter, T. F.; Luthy, D. A.; Comstock, C. H.; Saade, G. R.; Klugman, S.; Dugoff, L.; Craigo, S. D.; Timor-Tritsch, I. E.; Carr, S. R.; Wolfe, H. M.; D'Alton, M. E. Pregnancy loss rates after midtrimester amniocentesis. *Obstet. Gynecol.* **2006**, *108* (5), 1067–72.
- Merkatz, I. R.; Nitowsky, H. M.; Macri, J. N.; Johnson, W. E. An association between low maternal serum alpha-fetoprotein and fetal chromosomal abnormalities. *Am. J. Obstet. Gynecol.* **1984**, *148* (7), 886–94.
- Wald, N. J.; Kennard, A.; Hackshaw, A.; McGuire, A. Antenatal screening for Down's syndrome [erratum appears in *J. Med. Screen.* **1998**, *5* (2), 110]. *J. Med. Screen.* **1997**, *4* (4), 181–246.
- Reddy, U. M.; Mennuti, M. T. Incorporating first-trimester Down syndrome studies into prenatal screening: Executive summary of the National Institute of Child Health and Human Development Workshop. *Obstet. Gynecol.* **2006**, *107*, 167–73.
- Rose, N. C. On the current dilemma of Down syndrome screening. *Obstet. Gynecol.* **2006**, *107*, 2–3.
- ACOG Practice Bulletin No. 77. Screening for fetal chromosomal abnormalities. *Obstet. Gynecol.* **2007**, *109* (1), 217–28.
- Gravett, M. G.; Novy, M. J.; Rosenfeld, R. G.; Reddy, A. P.; Jacob, T.; Turner, M.; McCormack, A.; Lapidus, J. A.; Hitti, J.; Eschenbach, D. A.; Roberts, C. T., Jr.; Nagalla, S. R. Diagnosis of intra-amniotic infection by proteomic profiling and identification of novel biomarkers. *JAMA, J. Am. Med. Assoc.* **2004**, *292* (4), 462–9.
- Malone, F. D.; Canick, J. A.; Ball, R. H.; Nyberg, D. A.; Comstock, C. H.; Bukowski, R.; Berkowitz, R. L.; Gross, S. J.; Dugoff, L.; Craigo, S. D.; Timor-Tritsch, I. E.; Carr, S. R.; Wolfe, H. M.; Dukes, K.; Bianchi, D. W.; Rudnicka, A. R.; Hackshaw, A. K.; Lambert-Messerlian, G.; Wald, N. J.; D'Alton, M. E.; First- and Second-Trimester Evaluation of Risk Research Consortium, First-trimester or second-trimester screening, or both, for Down's syndrome. *New Eng. J. Med.* **2005**, *353* (19), 2001–11.
- Searle, B. C.; Dasari, S.; Turner, M.; Reddy, A. P.; Choi, D.; Wilmarth, P. A.; McCormack, A. L.; David, L. L.; Nagalla, S. R. High-throughput identification of proteins and unanticipated sequence modifications using a mass-based alignment algorithm for MS/MS de novo sequencing results. *Anal. Chem.* **2004**, *76* (8), 2220–30.
- Searle, B. C.; Dasari, S.; Wilmarth, P. A.; Turner, M.; Reddy, A. P.; David, L. L.; Nagalla, S. R. Identification of protein modifications using MS/MS de novo sequencing and the OpenSea alignment algorithm. *J. Proteome Res.* **2005**, *4* (2), 546–54.
- Craig, R.; Beavis, R. C. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* **2004**, *20* (9), 1466–7.

- (14) Old, W. M.; Meyer-Arendt, K.; Aveline-Wolf, L.; Pierce, K. G.; Mendoza, A.; Sevinsky, J. R.; Resing, K. A.; Ahn, N. G. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol. Cell. Proteomics* **2005**, *4* (10), 1487–502.
- (15) Chromy, B. A.; Gonzales, A. D.; Perkins, J.; Choi, M. W.; Corzett, M. H.; Chang, B. C.; Corzett, C. H.; McCutchen-Maloney, S. L. Proteomic analysis of human serum by two-dimensional differential gel electrophoresis after depletion of high-abundant proteins. *J. Proteome Res.* **2004**, *3* (6), 1120–7.
- (16) Yan, F.; Sreekumar, A.; Laxman, B.; Chinnaiyan, A. M.; Lubman, D. M.; Barder, T. J. Protein microarrays using liquid phase fractionation of cell lysates. *Proteomics* **2003**, *3* (7), 1228–35.
- (17) Zheng, S.; Schneider, K. A.; Barder, T. J.; Lubman, D. M. Two-dimensional liquid chromatography protein expression mapping for differential proteomic analysis of normal and O157:H7 *Escherichia coli*. *BioTechniques* **2003**, *35* (6), 1202–12.
- (18) Qin, S.; Ferdinand, A. S.; Richie, J. P.; O'Leary, M. P.; Mok, S. C.; Liu, B. C. Chromatofocusing fractionation and two-dimensional difference gel electrophoresis for low abundance serum proteins. *Proteomics* **2005**, *5* (12), 3183–92.
- (19) Zybailov, B.; Coleman, M. K.; Florens, L.; Washburn, M. P. Correlation of relative abundance ratios derived from peptide ion chromatograms and spectrum counting for quantitative proteomic analysis using stable isotope labeling. *Anal. Chem.* **2005**, *77* (19), 6218–24.
- (20) CDC. Entry Into Prenatal Care—United States 1989–1997. In *Morbidity and Mortality Weekly Report*; U.S. Department of Health & Human Services, U.S. Government Printing Office: Washington, DC, 2000; Vol. 49, pp 393–397.
- (21) Wald, N. J.; Rodeck, C.; Hackshaw, A. K.; Walters, J.; Chitty, L.; Mackinson, A. M.; Group, S. R. First and second trimester antenatal screening for Down's syndrome: the results of the Serum, Urine and Ultrasound Screening Study (SURUSS). *Health Tech. Assess.* **2003**, *7* (11), 1–77.
- (22) Caughey, A. B.; Lyell, D. J.; Washington, A. E.; Filly, R. A.; Norton, M. E. Ultrasound screening of fetuses at increased risk for Down syndrome: how many missed diagnoses? *Prenatal Diagn.* **2006**, *26* (1), 22–7.
- (23) Guetta, E.; Simchen, M. J.; Mammon-Daviko, K.; Gordon, D.; Aviram-Goldring, A.; Rauchbach, N.; Barkai, G. Analysis of fetal blood cells in the maternal circulation: challenges, ongoing efforts, and potential solutions. *Stem Cells Dev.* **2004**, *13* (1), 93–9.
- (24) Lo, Y. M. Recent advances in fetal nucleic acids in maternal plasma. *J. Histochem. Cytochem.* **2005**, *53* (3), 293–6.
- (25) Tong, Y. K.; Lo, Y. M. Diagnostic developments involving cell-free (circulating) nucleic acids. *Clin. Chim. Acta* **2006**, *363* (1–2), 187–96.
- (26) Wilson, R. D. Cell-free fetal DNA in the maternal circulation and its future uses in obstetrics. *J. Obstet. Gynaecol. Can.* **2005**, *27* (1), 54–62.
- (27) Raynes, J. Variations in the relative proportions of microheterogeneous forms of plasma glycoproteins in pregnancy and disease. *Biomed. Pharmacother.* **1982**, *36* (2), 77–86.
- (28) Mitchell, E.; Torpy, D. J.; Bagley, C. J. Pregnancy-associated corticosteroid-binding globulin: high resolution separation of glycan isoforms. *Horm. Metab. Res.* **2004**, *36* (6), 357–9.
- (29) Sutton, J. M.; Cole, L. A. Sialic acid-deficient invasive trophoblast antigen (sd-ITA): a new urinary variant for gestational Down syndrome screening. *Prenatal Diagn.* **2004**, *24* (3), 194–7.
- (30) Strel'chyonok, O. A.; Avvakumov, G. V.; Akhrem, A. A. Pregnancy-associated molecular variants of human serum transcortin and thyroxine-binding globulin. *Carbohydr. Res.* **1984**, *134* (1), 133–40.
- (31) Shi, Q.; Chen, C.; Wang, X. F.; Wang, H. L. [Clinical study of hemostatic coagulation markers in prethrombosis state of pregnancy induced hypertension]. *Zhonghua Fuchanke Zazhi* **2004**, *39* (11), 733–6.
- (32) Borth, W. Alpha 2-macroglobulin, a multifunctional binding protein with targeting characteristics. *FASEB J.* **1992**, *6* (15), 3345–53.
- (33) Liao, A.; Nitsch, R. M.; Greenberg, S. M.; Finckh, U.; Blacker, D.; Albert, M.; Rebeck, G. W.; Gomez-Isla, T.; Clatworthy, A.; Binetti, G.; Hock, C.; Mueller-Thomsen, T.; Mann, U.; Zuchowski, K.; Beisiegel, U.; Staehelin, H.; Growdon, J. H.; Tanzi, R. E.; Hyman, B. T. Genetic association of an alpha2-macroglobulin (Val1000Ile) polymorphism and Alzheimer's disease. *Hum. Mol. Genet.* **1998**, *7* (12), 1953–6.
- (34) Zorina, V. V.; Levchenko, V. G.; Zorina, R. M.; Promzeleva, N. V.; Zorin, N. A.; Gorlina, N. K. Pregnancy-associated plasma protein-A, alpha-2-macroglobulin, pregnancy zone protein and their complexes with IgG in sera of healthy non-pregnant and pregnant women, and patients with breast cancer. *Russ. J. Immunol.* **2001**, *6* (1), 71–76.
- (35) Pineiro, M.; Alava, M. A.; Gonzalez-Ramon, N.; Osada, J.; Lasierra, P.; Larrad, L.; Pineiro, A.; Lampreave, F. ITH4 serum concentration increases during acute-phase processes in human patients and is up-regulated by interleukin-6 in hepatocarcinoma HepG2 cells. *Biochem. Biophys. Res. Commun.* **1999**, *263* (1), 224–9.
- (36) Singh, V. K.; Fudenberg, H. H.; Brown, F. R. 3rd Immunologic dysfunction: simultaneous study of Alzheimer's and older Down's patients. *Mech. Ageing Dev.* **1986**, *37* (3), 257–64.
- (37) Elovaara, I.; Maury, C. P.; Palo, J. Serum amyloid A protein, albumin and prealbumin in Alzheimer's disease and in demented patients with Down's syndrome. *Acta Neuro. Scand.* **1986**, *74* (3), 245–50.
- (38) Stoltzner, S. E.; Grenfell, T. J.; Mori, C.; Wisniewski, K. E.; Wisniewski, T. M.; Selkoe, D. J.; Lemere, C. A. Temporal accrual of complement proteins in amyloid plaques in Down's syndrome with Alzheimer's disease. *Am. J. Pathol.* **2000**, *156* (2), 489–99.
- (39) Bergstrom, J.; Gustavsson, A.; Hellman, U.; Sletten, K.; Murphy, C. L.; Weiss, D. T.; Solomon, A.; Olofsson, B. O.; Westermark, P. Amyloid deposits in transthyretin-derived amyloidosis: cleaved transthyretin is associated with distinct amyloid morphology. *J. Pathol.* **2005**, *206* (2), 224–32.
- (40) Ranganathan, S.; Williams, E.; Ganchev, P.; Gopalakrishnan, V.; Lacomis, D.; Urbinelli, L.; Newhall, K.; Cudkowicz, M. E.; Brown, R. H., Jr.; Bowser, R. Proteomic profiling of cerebrospinal fluid identifies biomarkers for amyotrophic lateral sclerosis. *J. Neurochem.* **2005**, *95* (5), 1461–71.
- (41) Koo, E. H.; Lansbury, P. T., Jr.; Kelly, J. W. Amyloid diseases: abnormal protein aggregation in neurodegeneration. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96* (18), 9989–90.
- (42) Buttini, M.; Yu, G. Q.; Shockley, K.; Huang, Y.; Jones, B.; Masliah, E.; Mallory, M.; Yeo, T.; Longo, F. M.; Mucke, L. Modulation of Alzheimer-like synaptic and cholinergic deficits in transgenic mice by human apolipoprotein E depends on isoform, aging, and overexpression of amyloid beta peptides but not on plaque formation. *J. Neurosci.* **2002**, *22* (24), 10539–48.
- (43) Barr, J. R.; Maggio, V. L.; Patterson, D. G., Jr.; Cooper, G. R.; Henderson, L. O.; Turner, W. E.; Smith, S. J.; Hannon, W. H.; Needham, L. L.; Sampson, E. J. Isotope dilution–mass spectrometric quantification of specific proteins: model application with apolipoprotein A-I. *Clin. Chem.* **1996**, *42* (10), 1676–82.
- (44) Jeppsson, J. O.; Kobold, U.; Barr, J.; Finke, A.; Hoelzel, W.; Hoshino, T.; Miedema, K.; Mosca, A.; Mauri, P.; Paroni, R.; Thienpont, L.; Umemoto, M.; Weykamp, C.; International Federation of Clinical, C.; Laboratory, M. Approved IFCC reference method for the measurement of HbA1c in human blood. *Clin. Chem. Lab. Med.* **2002**, *40* (1), 78–89.
- (45) Anderson, L. Candidate-based proteomics in the search for biomarkers of cardiovascular disease. *J. Physiol.* **2005**, *563* (Pt 1), 23–60.

PR060539H