ORIGINAL ARTICLE

Identification of Protein Biomarkers in Dupuytren's Contracture using Surface Enhanced Laser Desorption Ionization Time-of-Flight Mass Spectrometry (SELDI-TOF-MS)

David O'Gorman, MSc, PhD[1,2,3,4,7],
Jeffrey C. Howard, PhD[1,2]
Vincenzo M. Varallo, MSc[1,2]
Peter Cadieux, BSc, MSc[3,7]
Erin Bowley, HBSc[1,2,3],
Kris McLean, HBMSc[1,2,3,5,7],
Brian J. Pak, MSc, PhD[8],
Bing Siang Gan, MD, PhD, FRCSC, FACS[1,2,3,4,5,6,7]

From the Cell & Molecular Biology Laboratory[1], Hand & Upper Limb Centre[2], Lawson Health Research Institute[3], Departments of Surgery[4], Physiology and Pharmacology[5], Medical Biophysics[6], University of Western Ontario[7], London ON, Canada, Ciphergen Biosystems Inc.,[8] Fremont CA, USA

Manuscript submitted 3rd October 2005
Accepted for publication 6th February 2006

Clin Invest Med 2006; 29 (3): 136–145.

Abstract

Background: To study the protein expression profiles associated with Dupuytren's contracture (DC) to identify potential disease protein biomarkers (PBM) using a proteomic technology - Surface Enhanced Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (SELDI-TOF-MS).
Methods: Normal and disease palmar fascia from DC patients were analyzed using Ciphergen's SELDI-TOF-MS Protein Biological System II (PBSII) ProteinChip® reader. Analysis of the resulting SELDI-TOF spectra was carried out using the peak cluster analysis program (BioMarker Wizard, Ciphergen). Common peak clusters were then filtered using a bootstrap algorithm called SAM (Significant Analysis of Microarrays) for increased fidelity in our analysis.
Results: Several differentially expressed low molecular weight (<20 kDa) tissue proteins were identified. Spectra generated using both $ZipTip_{C18}$ aided Au array and WCX2 array based SELDI-TOF-MS were reproducible, with an average peak cluster mass standard deviation for both methods of $<1.74 \times 10^{-4}$. Initial peak cluster analysis of the SELDI spectra identified both up-(14) and down-(3)regulated proteins associated with DC. Further analysis of the peak cluster data using the bootstrap algorithm SAM identified three disease-associated protein features (4600.8 Da, 10254.5 Da, and 11405.1 Da) that were elevated (5.45, 11.7, and 4.28 fold, respectively, with a 0% median false discovery rate).
Conclusion: SELDI-TOF-MS identified three potential low molecular weight tissue protein markers ($p4.6^{DC}$, $p10^{DC}$, $p11.7^{DC}$) for DC. The ability of SELDI-TOF-MS to resolve low molecular weight proteins suggests that the method may provide a means of deciphering the biomarker-rich low molecular weight region of the human proteome. Application of such novel technology may help clinicians to focus on specific molecular abnormalities in diseases with no known molecular pathogenesis, and uncover therapeutic and/or diagnostic targets.

Dupuytren's contracture (DC) is a benign and often familial fibro-proliferative disorder of the hand.[1-3] Despite its long clinical history and remarkably high incidence among Caucasian populations, the genetic etiology of DC is still unknown. This lack of basic

understanding of the molecular patho-physiology of DC has resulted in an absence of significant treatment advances over the last decennia and surgical resection remains the mainstay of therapy. To rationally explore alternative treatment modalities for this disease, it would be necessary to identify potential therapeutic targets, as well as diagnostic markers that may indicate disease progression of DC.

The recent advances in DNA informatics (e.g. the Human Genome Project) offer a unique opportunity to study human disease in an innovative fashion. In addition, the availability of new technology such as micro-fabricated DNA microarrays and genome-wide single-nucleotide polymorphisms (SNPs) markers, have now allowed researchers to monitor the expression levels of thousands of genes simultaneously or to explore genotype-phenotype relationships.[4] For DC, the application of these types of technologies has only very recently been reported. Pan and colleagues used a small gene microarray to look at changes in gene expression associated with DC.[5] Several genes were shown to be consistently elevated in DC patients, however, the limited sample size and lack of meaningful statistical quantitative information about the levels of expression of these genes have been raised as criticism of this work. Undoubtedly, this is related to the relatively high cost of using this type of DNA array technology. In contrast, Bayat and colleagues have carried out a number of genetic association studies by examining common SNPs for several disease candidate genes that are components of the transforming growth factor-β (TGF-β) signalling pathway (i.e. TGF-β1, TGF-β2, TGF-β receptors I, II, III, and Zf9) [6-8]. Using a PCR-based restriction fragment length polymorphism (PCR-RFLP) method, this group found differences in genotype and allele frequencies for Zf9, a transcription factor known to enhance the fibrogenic potential of TGF-β1.[9, 10] Although there is great promise in using SNPs genotyping to help pinpoint disease-associated genes, there are many confounding factors that can cause false-positive genetic associations (e.g. population artefacts, non-functional SNPs, etc.).

A relatively novel alternative to DNA-based informatics is the use of proteomic-based technologies to study human health and disease, particularly mass spectrometry (MS). Advances in MS instrumentation, sample preparation and ionization techniques, and refinements in MS algorithms have now made it possible to analyze complex protein mixtures (e.g. serum) directly, on a large scale, and potentially in a high-throughput manner. Of all the recently developed MS techniques, Surface Enhanced Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (SELDI-TOF-MS) has perhaps garnered the most attention with respect to disease biomarker discovery (Figure 1). SELDI-TOF-MS is a highly sensitive form of Matrix-Assisted LDI (MALDI)-TOF technique, which uses special chromatographic-like probe surfaces (Protein Chip® arrays) to capture different classes of proteins.[11, 12] Unlike the inert metal probe surfaces used in MALDI-TOF, the SELDI Protein Chip® arrays provide specialized protein-

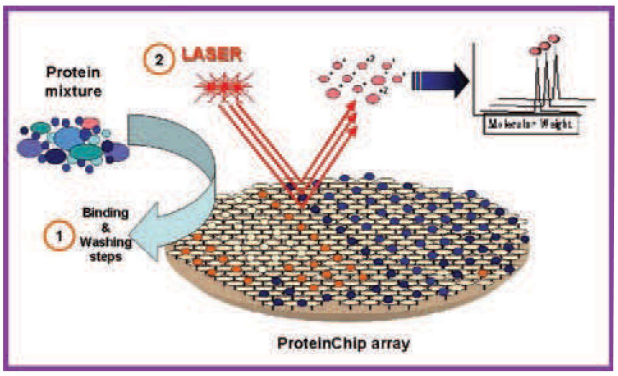


FIGURE 1 - Surface Enhanced Laser Desorption/Ionization Time-of-Flight Mass Spectrometry: General steps involved in the ProteinChip® based SELDI-TOF-MS. Unlike traditional MALDI-TOF-MS, which uses an inert metal probe surface to present the sample for MS-TOF analysis, ProteinChip® arrays play an active role in the TOF-MS process and permit direct analysis of crude complex protein mixtures (e.g. serum). **(1)** Bbriefly, small aliquots (1µl) of equivalent protein are spotted onto the active regions of 8-spot ProteinChip® arrays. Binding conditions are varied to reduce background and possibly facilitate sequential 'on-chip' protein purification. Following protein binding each spot or the entire array is washed using HPLC grade H_20 or more stringent buffers (e.g. PBS + 0.2M NaCl + 0.1%Tween). A laser energy-absorbing-matrix (EAM, e.g. 100% saturated solution of cyano-4-hydroxy cinnamic acid, or CHCA) is then applied to each spot and allowed to dry thoroughly. **(2)**The array is then placed in the SELDI ProteinChip® Biology System II (PBS II) reader for scanning using an automated data collection protocol. The PBS II system is configured such that the only variable determining its TOF within the drift tube is its mass-per-unit-charge (m/z). Thus, smaller and/or more highly charged ions will reach the detector first, with the abundance of any given ion being proportional to its peak amplitude or area.

binding surfaces (e.g. quaternary amine functionality) that are capable of capturing proteins with complementary physiochemical properties (e.g. anionic proteins). SELDI-TOF-MS is, therefore, particularly well-suited for clinical proteomic studies since it can rapidly detect disease-specific protein profiles in a largely unbiased manner without prior knowledge of the underlying genetic lesions, while using only small quantities of patient material.[13-15] While this method preferentially detects low molecular weight (<30kDa) proteins, more comprehensive proteomic profiling is possible when simple, high-throughput sample fractionation techniques are used in combination with SELDI-TOF-MS.[16] Because of these reasons, we have employed this TOF-MS technique to identify changes in protein expression associated with DC.

METHODS

Clinical specimen and primary cultures: DC patient specimens (normal and disease palmar fascia) were collected in compliance with the University Human Research Ethics Board. Areas of diseased fascia (disease) and adjacent, uninvolved normal palmar fascia (control) were collected from DC patients undergoing surgical resection of diseased palmar fascia tissue. Portions of the clinical specimens were stored at -80°C for protein extraction.

Preparation of protein extracts for SELDI-TOF analysis: Protein extracts were prepared using a commercially available tissue-protein extraction reagent T-PER (Pierce, Rockford, IL, USA), supplemented with Sigma protease inhibitor cocktail (2mM AEBSF, 1mM EDTA, 130µm Gestating, 14µm E-64, 1µm eupeptic, 0.3µm apportioning) + 1mM PMSF (paramethylsulfonyl fluoride, Sigma), and general phosphatase inhibitors (1mM Na_3VO_4, 1mM NaF, Sigma, St. Louis, MO, USA). Tissue samples were homogenized (Ultra-Turrax T8, IKA Labortechnik, Staufen, Germany) using short bursts (high setting) under ice to avoid mechanical heating of the samples. The resulting sample homogenates were briefly centrifuged (15min, 4°C, 12,000 x g) to remove insoluble debris. The resulting supernatant was then assayed for total protein content using a commercially available BCA protein assay kit (Pierce, Rockford IL, USA).

SELDI-TOF-MS analysis using WCX2 ProteinChip® arrays: For SELDI-TOF analysis using weak cation exchanger (WCX2) ProteinChip® arrays, equivalent protein (5µg) was suspended in a final 20µl volume containing 12.5mM sodium acetate (pH 5.0) and 0.1% triton X100. HPLC grade H_20 (Sigma) was used for all sample preparations. Prior to applying the samples onto the WCX2 arrays, the active spots of the array were pre-activated in accordance with the manufacturer's instructions. Briefly, the entire WCX2 chips were washed for five minutes in 0.01N HCl, followed by two quick washes in HPLC grade H_2O. A hydrophobic PAP pen was then used to create a containment barrier around each active spot of the WCX2 arrays. Once the hydrophobic barrier was dry the active spots of the array were washed 2x (10min washes, RT) with binding buffer (12.5mM sodium acetate, pH 5.0 + 0.1% triton X100). The samples (5µl containing 5µg total protein) were then applied to each pre-activated spot of the array and allowed to bind (20 min., 22°C, humidified chamber) to the array. Following the binding step, the entire WCX2 array was then washed (2x) with binding buffer (10 min, 22°C, rotator) and then once with HPLC H_2O. After briefly drying the arrays, 0.5µl of a saturated solution of 4-Hydroxy-3,5-dimethoxy-cinnamic acid (sinapinic acid or SPA, Sigma) dissolved in 50% (v/v) acetonitrile (ACN) and 0.5% (v/v) trifluoroacetic acid (TFA), was applied 2x to each of the active spots of the array and allowed to thoroughly dry. The WCX2 array was then placed in the SELDI ProteinChip® Biology System II reader (PBS II) for scanning using an automated data collection protocol, as we have previously described.[13, 17, 18]

The PBS II reader is a linear TOF-MS equipped with a 337nm nitrogen laser. It was operated in positive-ion mode with a uniform static electric field of +20kV. The resulting spectra were averaged from at least 65 collected transient laser shots that scan ~66% of the target area in linear sweeps (5 shots/region, 13 regions per spot). Sequential laser scans were carried out by incremental increasing the laser intensity (LI) with each subsequent scan to optimize data collection for both low and high molecular weight proteins (LI settings of 175-295, with a LI of 185 equivalent to ~5µJoules).

$ZipTip_{C18}$ aided SELDI-TOF-MS analysis using Au ProteinChip® arrays: SELDI-TOF-MS analysis of $ZipTip_{C18}$ prepared surgical specimens was carried out using gold (Au) probe surfaces (Figure 2). $ZipTip_{C18}$ sample preparation, which concentrates and desalts the protein mixture, was carried according to the manufacturers instructions (Millipore Corporation, Billerica, MA USA). Briefly, 5µg protein (disease, or control tissue extract) was suspended in a final volume of 20µl HPLC grade H_2O (Sigma) containing 0.5% TFA (v/v). The acidic conditions (pH < 4) ensure

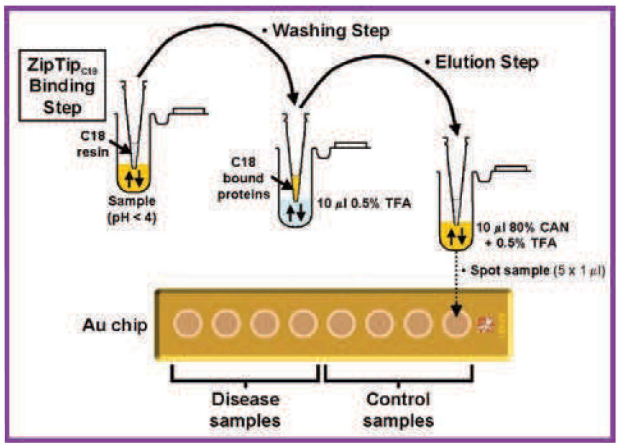


FIGURE 2 - SELDI-TOF-MS analysis using ZipTip$_{C18}$ and Au arrays: DC samples were processed using ZipTip$_{C18}$ prior to performing matrix assisted LDI-TOF-MS analysis using a simple gold (Au) chip probe. Depicted above is a single elution step approach to using the ZipTip$_{C18}$ pipette tips, which contains 0.6 µl of C18 reversed-phase media impregnated at its end. After ZipTip$_{C18}$ wetting (50% acetonitrile (ACN)) and equilibrating (0.5% TFA), sample binding was carried out by pipetting the sample up and down (10 cycles) through the fixed volume of equilibrated C18 resin. After binding the ZipTip$_{C18}$ was immediately washed (0.1% TFA, 5 pipette cycles) and the bound proteins eluted with a fixed volume (10µl) of 80% ACN + 0.5% TFA. Eluted fractions were then spotted (5 x 1µl) onto the Au probe surface and allowed to dry. Matrix (100% CHCA) was then applied to the sample spots on the Au probe surface and allowed to dry thoroughly. The Au probe was then placed in the SELDI ProteinChip® Biology System II (PBS II) reader for scanning using an automated data collection protocol.

maximum binding to the reversed-phase (0.6 µl bed volume of C_{18} media) ZipTip pipette tips. Following wetting and equilibriating of the ZipTip$_{C18}$ pippette tips with 50% (v/v) ACN and 0.1% (v/v) TFA, respectively, protein binding was carried out. Briefly, the protein mixture was cycled (aspirating and dispensing) 10x through the fixed, reverse phase bed volume of the ZipTip$_{C18}$ pippette tips. Following binding, the protein loaded ZipTip$_{C18}$ pipette tips were washed (three pipette cycles) with 0.1% (v/v) TFA. Elution of the ZipTip$_{C18}$ bound proteins was then carried out by cycling (10x) a single volume (15µl) of elution buffer (80% (v/v) ACN + 0.1% (v/v) TFA) up and down through the ZipTip$_{C18}$. The elution fraction (5 x 1µl volumes) was then applied to each spot of an Au ProteinChip® array and allowed to thoroughly dry at RT. Finally, 0.5µl of a saturated solution of α-cyano-4-hydroxy-cinnamic acid (CHCA, Sigma) in 50% (v/v) ACN + 0.5% (v/v) TFA was applied (2x) to each spot and allowing to thoroughly dry. The Au arrays were then placed in the SELDI ProteinChip® Biology System II reader (PBS II) for scanning using an automated data collection protocol as described above.

Biomarker Peak Cluster Analysis: SELDI ProteinChip® spectra were analyzed using the PBS II ProteinChip® software (v3.1.1). Briefly, calibrated spectra were initially normalized to total ion current (all peaks > 1500Da) to compensate for small variations in sample concentrations per spot or any m/z shifts in the spectrum. Peak cluster analysis was then performed using the Biomarker Wizard software (v3.1.1), which statistically evaluates clusters of peaks of similar MW from across sample groups of spectra. The initial peak detection scan was carried out using a signal-to-noise (S/N) ratio of ≥5. A second peak detection scan was then performed using a S/N ≥ 2 to ensure peak statistical analysis across all groups.

SAM Protein Filtering: Further statistical analysis of the peak cluster data sets was carried out using the bootstrap algorithm SAM (Significance Analysis of Microarrays, see: http://www-stat.stanford.edu/~tibs/SAM).[19] Although SAM was originally developed for Microarray gene expression analysis it has also been effectively used to filter SELDI-TOF data .[16] In this study, SAM was used to compare normalized peak cluster intensities between disease and control fascia groups (ProteinChip® BioMarker Wizard program, ProteinChip® Software v3.1.1, Ciphergen Bio systems Inc.). Data were exported to an Excel spreadsheet file using an Excel plug-in feature and labelled in accordance with SAM guidelines, with normalized peak cluster intensities values of the control group (n = 9) defined as group "1" and the disease group (n = 16) defined as group "2". Data analysis parameters included "two classes unpaired data" and 5000 permutations. An operator controlled (i.e. sliding) significance difference threshold, or delta value (Δ), permits the user to define the false discovery rate (FDR) (i.e. # falsely called proteins / # differentially expressed proteins) of the analysis. For each SAM analysis a table of Δ values was generated, which contained their associated FDR (median and 90th percentile).

RESULTS

Identification of DC protein biomarkers (PBM) using ZipTipC18 aided Au array SELDI-TOF-MS: Peak cluster analysis (Biomarker Wizard program,

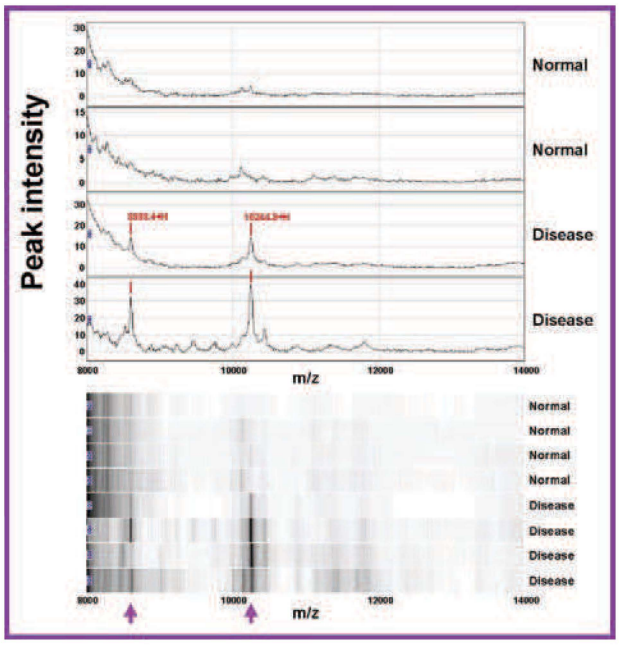
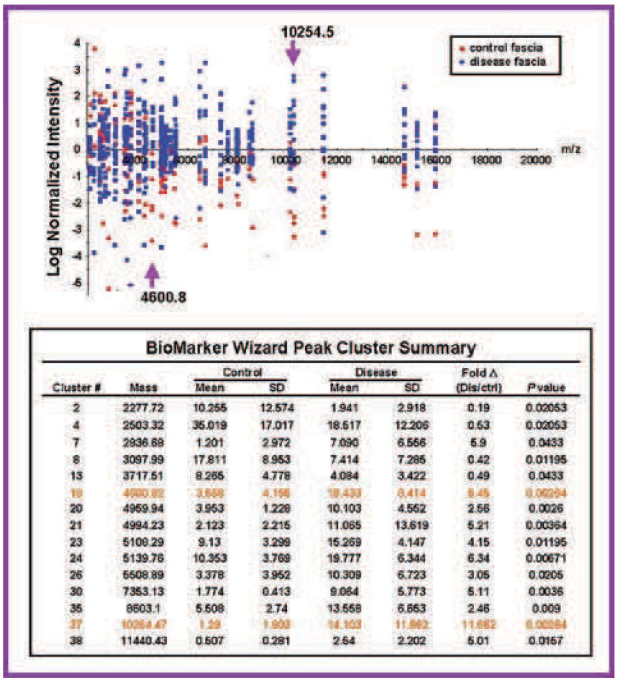


FIGURE 3 - ZipTip$_{C18}$ aided Au array SELDI-TOF-MS analysis of DC: Protein (5µg) extracted from normal and disease fascia was diluted in TFA (0.5%) to a final volume of 20µl. The protein samples were then allowed to bind to ZipTip$_{C18}$ pipette tips as described above (see Figure 2). ZipTip$_{C18}$ eluted fractions were spotted (5 x 1µl) onto Au chip probes and allowed to dry. EAM (2 x 1µl of CHCA) was applied to the gold chip spots, air dried and then placed in the PBS II reader for scanning using an automated collection protocol. The spectral data corresponding to peak intensity versus m/z were generated using 65 laser shots per laser intensity setting. Shown above are representative spectra (peak intensity, upper panel, and pseudo-gel view, lower panel) of 4 control and 4 disease patients. *Arrows show two protein peaks that were identified as significantly elevated in the disease tissue using the BioMarker Wizard peak cluster analysis program.

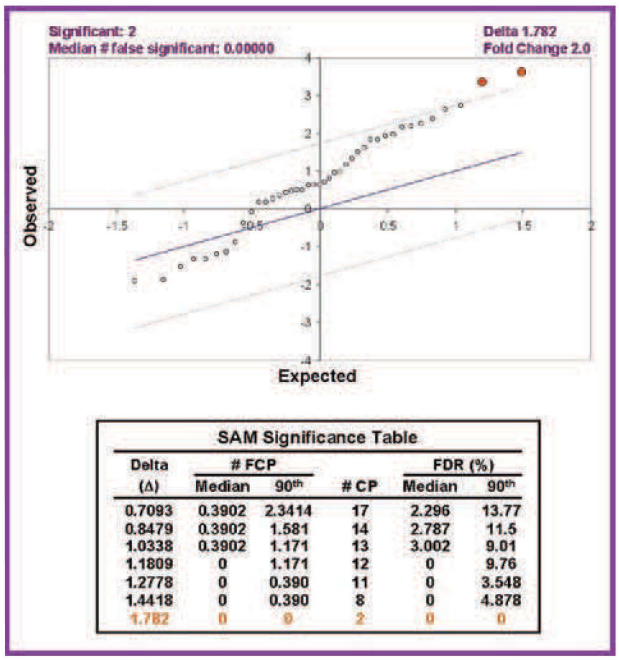


BioMarker Wizard Peak Cluster Summary

Cluster #	Mass	Control Mean	Control SD	Disease Mean	Disease SD	Fold Δ (Dis/ctrl)	*P* value
2	2277.72	10.255	12.574	1.941	2.918	0.19	0.02053
4	2503.32	35.019	17.017	18.517	12.206	0.53	0.02053
7	2836.69	1.201	2.972	7.090	6.556	5.9	0.0433
8	3097.99	17.811	8.953	7.414	7.285	0.42	0.01195
13	3717.51	8.265	4.778	4.084	3.422	0.49	0.0433
19	4600.82	3.658	4.156	19.439	8.414	5.45	0.00264
20	4959.94	3.953	1.228	10.103	4.562	2.56	0.0026
21	4994.23	2.123	2.215	11.065	13.619	5.21	0.00364
23	5100.29	9.13	3.299	15.269	4.147	4.15	0.01195
24	5139.76	10.353	3.769	19.777	6.344	6.34	0.00671
26	5508.89	3.378	3.952	10.309	6.723	3.05	0.0205
30	7353.13	1.774	0.413	9.054	5.773	5.11	0.0036
35	8603.1	5.508	2.74	13.558	6.853	2.46	0.009
37	10254.47	1.29	1.903	14.103	11.662	11.662	0.00264
38	11440.43	0.507	0.281	2.64	2.202	5.01	0.0167

FIGURE 4 - Peak cluster analysis of Au array SELDI-TOF-MS data: Peak cluster analysis (upper panel) was carried out using the Biomarker Wizard tool of the Ciphergen Biosystem software (v3.1.1) to identify common protein peaks across all spectra (total of n = 5 control, and n = 11 disease patients). All spectra were first normalized to total ion current. Peak cluster analysis was then performed for all m/z peaks between 2 and 20kDa with a S/N ≥ 2. The results are represented as a Biomarker Peak Cluster plot (Log normalized peak intensities versus m/z). The table in the lower panel lists the number of significant common peak clusters identified using the BioMarker Wizard program, and the associated mean and standard deviation of the mean (SDM) values for each peak cluster.

Ciphergen Bio systems Inc.) of surgical specimens (disease (DC lesions) and control (unaffected palmar fascia) from the same patient) using a ZipTip$_{C18}$ aided SELDI gold (Au) array approach (i.e. MALDI-TOF-MS, Fig. 2), uncovered several disease associated protein peaks (Fig. 3,4). In total, 15 of 42 identified common peak cluster intensities within the 2 - 20kDa range were identified as being either significantly higher (12 peaks) or lower (3 peaks) within the disease group (n = 13) compared with control (i.e. unaffected) palmar fascia (n = 5). The protein spectral patterns were very reproducible with an average peak cluster mass standard deviations of 1.73 x10^{-4}. Further data analysis was carried out by filtering the common peak cluster spectral data through the bootstrap statistical program called SAM (Significant Analysis of Microarrays).[19] This robust statistical method uses repeated permutations of the data to determine if the expression of any gene/protein are significantly related to a response variable(s), while permitting the user to define both a significant cut-off (Δ value, based on the false positive or discover rate, FDR) and a fold



SAM Significance Table

Delta	# FCP			FDR (%)	
(Δ)	Median	90th	# CP	Median	90th
0.7093	0.3902	2.3414	17	2.296	13.77
0.8479	0.3902	1.581	14	2.787	11.5
1.0338	0.3902	1.171	13	3.002	9.01
1.1809	0	1.171	12	0	9.76
1.2778	0	0.390	11	0	3.548
1.4418	0	0.390	8	0	4.878
1.782	0	0	2	0	0

FIGURE 5 - SAM analysis of Au array SELDI-TOF-MS data: Common peak cluster data were filtered using the SAM bootstrap algorithm (SAM plot, upper panel). SAM uses repeated permutations (n = 5000) of the data sets to determine whether the expression of any protein is significant. Parameters for analysis included a user-defined threshold for significance, or Δ value (Delta = 1.75), and a user-defined fold-change (± 2.0 fold). Using these parameters SAM identified two proteomic features (4600.8 and 10254.5 Da peaks) that are different at a median false significant value < 0.0001, with a corresponding % false discovery rate (FDR, # falsely called protein / # differentially proteins in original data set) of 0%. The table (lower panel) lists the number (median and 90th percentile) of falsely called proteins (# FCP), total number of called proteins (# CP) and the resulting FDR (%) associated with each chosen Δ value.

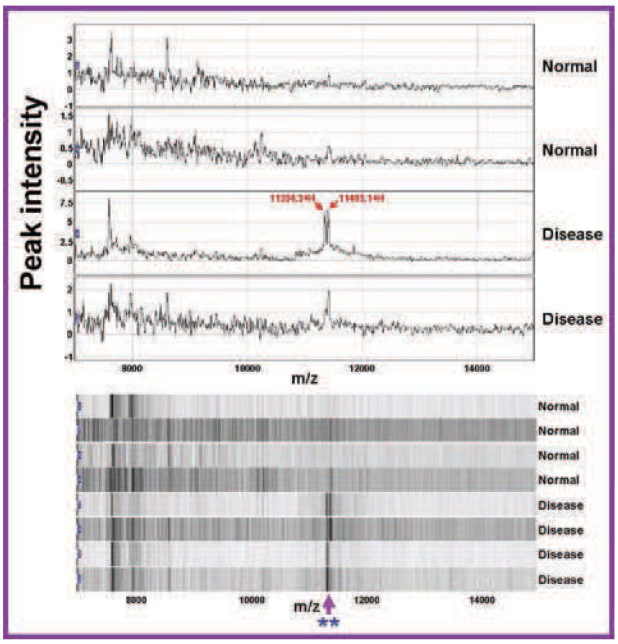


FIGURE 6 - SELDI-TOF-MS analysis of DC using WCX2 ProteinChip® arrays: Protein (5µg) extracted from normal and disease fascia was applied to the active spots of a pre-equilibrated weak cation exchange (WCX2) Ciphergen ProteinChip® array. The spectral data corresponding to peak intensity versus m/z were generated using 65 laser shots per laser intensity setting. Shown above are representative peak intensity spectra (upper panel) and corresponding pseudo-gel views (lower panel) of 4 control and 4 disease patients. *Arrows show two protein peaks that were identified as significantly elevated in the disease tissue using the BioMarker Wizard peak cluster analysis program.

change (ensure level of expression of the called genes/proteins changes a prescribed amount) parameter (Fig. 5). Using stringent screening parameters (0% median and 90th percentile FDR, and a fold change value of ±2 fold) SAM identified 2 significant protein peaks within our peak cluster data set (4600.8 Da, $p4.6^{DC}$; and 10254.5 Da, $p10^{DC}$). These two proteomic features were found to be elevated 5.45 and 11.7 fold, respectively, within the disease group using the Biomarker Wizard peak cluster analysis.

Identification of DC PBM using WCX2 array SELDI-TOF-MS analysis: Peak cluster analysis (Biomarker Wizard program, Ciphergen Biosystems Inc.) of disease and control surgical specimens using the WCX2 array SELDI-TOF-MS approach (Fig. 1) identified several disease associated protein peaks (Fig. 6,7). In total, 2 peaks (11358.2 and 11405.1 Da) out of a total of 19 common peak clusters within the 2 - 20kDa range were identified as being significantly higher within the disease group (n = 10) compared with controls (n = 4). The protein expression patterns were very reproducible with an average peak cluster mass standard deviations of 1.5982 $x10^{-4}$. More stringent bootstrap analysis of the common peak cluster data using SAM identified 1 proteomic feature

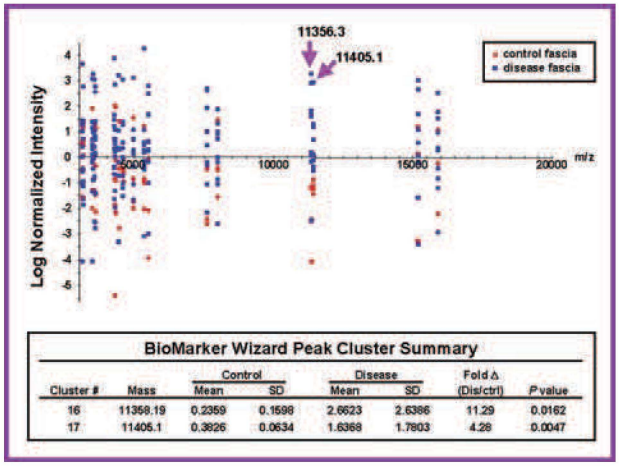


FIGURE 7 - Peak cluster analysis of WCX2 ProteinChip® SELDI-TOF-MS spectral data set: Peak cluster analysis (upper panel) was carried out using the Biomarker Wizard tool of the Ciphergen Biosystem software (v3.1.1) to identify common protein peaks across all spectra (total of n = 5 control, and n = 11 disease patients). All spectra were first normalized to total ion current. Peak cluster analysis was then performed for all m/z peaks between 2 and 20kDa with a S/N ≥ 2. The results are represented as a Biomarker Peak Cluster plot (Log normalized peak intensities versus m/z. The table l (lower panel) lists the number of significant common peak clusters identified using the BioMarker Wizard program, and the associated mean and SDM values for each peak cluster.

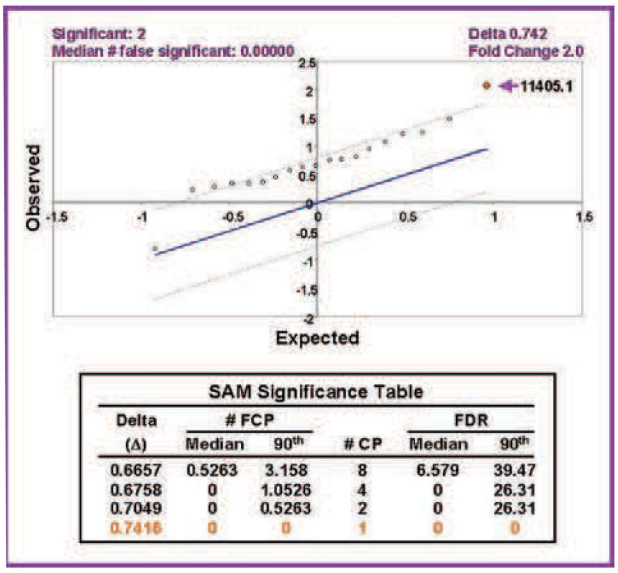


FIGURE 8 - SAM analysis of WCX2 ProteinChip® peak cluster data: Common peak cluster data generated using the Biomarker Wizard program was filtered through the SAM bootstrap statistical analysis program to generate a SAM plot (upper panel). Our chosen parameters were Δ = 0.742, 5000 permutations and a fold change of ± 2 Using these parameters SAM identified one proteomic feature (11405.1 Da peak) that was different at a median false significant value < 0.0001, with a corresponding FDR = 0%. The table (lower panel) lists the number (median and 90th percentile) of FCP, total number of CP and the resulting FDR (median and 90th percentile) for each user-defined Δ value.

(11405.1 Da, p11.4^{DC}) that was elevated (4.28 fold) within the disease group with a median and 90th percentile FDR of 0% (Fig. 8).

DISCUSSION

Dupuytren's contracture (DC) is a debilitating disease of the hand that leads to finger contractures and loss of hand function and is considered one of the most common fibroproliferative connective tissue diseases. It has a high prevalence in Caucasians of Northern European descent, with a reported incidence ranging from 10-40% for males over the age of 60.[20] Unfortunately, at present, DC is incurable and the only (temporary) relief for patients is through surgical resection, an approach virtually unchanged since the middle of the 19th century.[21] Surgical resection, however, has a high recurrence rate , carries risks, is painful and requires prolonged post-operative rehabilitation.[22] DC has a number of remarkable characteristics, which suggest that it is part of continuum of disorders associated with dysregulated cell growth and proliferation. For example, patients with DC have an increased incidence of both benign and malignant neoplasia.[23-25] Also, histologically and biochemically, DC and wound granulation tissue share many similarities, leading to the common belief that DD may be an exaggerated, dysregulated fibrotic wound healing response.[26, 27] Therefore, elucidation of the molecular pathogenesis of DD may also have added clinical significance for the diagnosis and treatment of these other pathological conditions. Unfortunately for the clinician-investigator, it is an intimidating task to find a place to begin unravelling the molecular abnormalities of health and disease.

The advent of high-throughput analytical technology to pinpoint molecular abnormalities gives

the investigator the opportunity to rapidly focus on potential targets for further study. In this study, we have employed a highly sensitive TOF-MS technique to identify changes in protein expression associated with DC.

SELDI-based proteomic profiling of DC specimens resolved an average of 35 common low molecular weight (2 - 20kDa) tissue derived protein/peptide clusters using a $ZipTip_{C18}$ aided Au array based SELDI-TOF-MS approach (Fig. 2,3,4,5), and an average of 29 common low molecular weight (2 - 20kDa) protein/peptide clusters using WCX2 array based SELDI-TOF-MS (Fig. 6,7,8). The spectra generated using these two TOF-MS approaches were highly reproducible with average peak cluster mass standard deviations $<1.74 \times 10^{-4}$ for both methods. While stringent filtering of the protein cluster data using SAM identified only 3 up-regulated proteins (4600.8 Da, 5.45 fold, 10254.5 Da, 11.7 fold; and 11405.1 Da, 4.28 fold) associated with DC, our initial peak cluster analysis generated a more extensive list of both up- (14) and down-regulated (3) peptides and/or proteins (Fig. 4,7). The presence of numerous differentially expressed low molecular weight proteins is not difficult to comprehend if one considers the nature of the sample being examined in this study. DC lesions are very different structures compared to normal fascia tissue. Moreover, DC lesions are quite dynamic structures undergoing extensive remodelling as the disease progresses. Therefore, one would expect several protein and/or peptide by-products of the disease would be present within these active lesions, and perhaps even in the surrounding environment (e.g. blood). While we did not investigate patient serum, recent SELDI-based studies suggest that low molecular weight degradation products (i.e. peptides) can passively enter the blood stream and bind to carrier proteins such as serum albumin.[28] Peptide markers bound to these carrier proteins potentially possess a half-life many orders of magnitude higher than their free-phase counterparts, which are actively cleared by the kidneys, suggesting that serum albumin bound peptides may represent a 'treasure trove' of clinically useful diagnostic information.[29, 30] The rapidly growing list of low molecular weight (<20kDa) serum markers for various human diseases, including HBV-induced liver cirrhosis [15], prostate cancer [31], ovarian cancer [32], hepatocellular carcinoma [16] and head and neck cancers [14], perhaps best reflects the success of this SELDI-based MS approach. In those studies, the use of various algorithms to analyze the serum protein expression patterns has generated highly sensitive (83-100%) and specific (90-95%) proteomic disease signatures that may prove very useful for early disease detection. In our study, the use of the bootstrap algorithm SAM to filter the protein cluster data helped to further decipher the importance of the peak cluster generated by SELDI-TOF-MS. In particular, SAM permitted us to define both a fold-change parameter to ensure the level of expression of the called protein changes a prescribed amount, and a significance cut-off parameter (delta, Δ) based on the false discovery rate. Although further studies will be needed to identify and validate these potential protein markers, it appears that the use of these novel technologies can help us identifying elements of the human proteome that may be targets for the development of alternative methods for the detection and treatment of this disease.

CONCLUSIONS

Our analysis of DC patient material identified several disease-associated low molecular weight proteins. Peak cluster analysis of the SELDI-TOF spectra coupled with SAM bootstrap filtering of the common peak clusters identified three proteins (4600 Da, $p4.6^{DC}$; 10254.5 Da, $p10^{DC}$; and 11404.2 Da, $p11.4^{DC}$) that were expressed at higher levels within diseased palmar fascia tissue (5.45, 11.7, and 4.28 fold, respectively). The present study provides the first low molecular weight (<20 kDa) proteomic profiling of DC. Although the identity of the differentially expressed proteins and peptides is not known, further analysis using multidimensional SELDI-TOF based approaches, as we have employed before [16, 18] will permit more precise identification of these potential protein or peptide markers. Given the vast clinical potential of the low molecular weight region of the proteome and the inability of more conventional two-dimensional electrophoresis to detect them, it would seem that SELDI-TOF-MS provides a promising new avenue of deciphering the biomarker-rich low molecular weight region of the human proteome. Application of novel technology like this may help clinicians and clinician-scientists focus on specific molecular abnormalities in diseases that have no known molecular pathogenesis, thus potentially uncovering therapeutic and/or diagnostic targets.

Acknowledgements

Our lab is supported by the Canadian Institutes of Health Research, the US Plastic Surgery Education Foundation, the Lawson Health Research Institute and the Canadian Society for Surgery of the Hand.

References:

1. Rayan GM. Clinical presentation and types of Dupuytren's disease. *Hand Clin.* 1999;15:87-96.
2. Gudmundsson KG, Arngrimsson R, Sigfusson N, Bjornsson A, Jonsson T. Epidemiology of Dupuytren's disease: clinical, serological, and social assessment. The Reykjavik Study. *J Clin Epidemiol.* 2000;53:291-6.
3. Yi IS, Johnson G, Moneim MS. Etiology of Dupuytren's disease. *Hand Clin.* 1999;15:43-51.
4. Sachidanandam R, Weissman D, Schmidt SC, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature.* 2001;409:928-33.
5. Pan D, Watson HK, Swigart C, Thomson JG, Honig SC, Narayan D. Microarray gene analysis and expression profiles of Dupuytren's contracture. *Ann Plast Surg.* 2003;50:618-22.
6. Bayat A, Watson JS, Stanley JK, et al. Genetic susceptibility in Dupuytren's disease. TGF-beta1 polymorphisms and Dupuytren's disease. *J Bone Joint Surg Br.* 2002;84:211-5.
7. Bayat A, Stanley JK, Watson JS, Ferguson MW, Ollier WE. Genetic susceptibility to Dupuytren's disease: transforming growth factor beta receptor (TGFbetaR) gene polymorphisms and Dupuytren's disease. *Br J Plast Surg* 2003;56:328-33.
8. Bayat A, Watson JS, Stanley JK, Ferguson MW, Ollier WE. Genetic susceptibility to dupuytren disease: association of Zf9 transcription factor gene. *Plast Reconstr Surg* 2003;111:2133-9.
9. Kim Y, Ratziu V, Choi SG, et al. Transcriptional activation of transforming growth factor beta1 and its receptors by the Kruppel-like factor Zf9/core promoter-binding protein and Sp1. Potential mechanisms for autocrine fibrogenesis in response to injury. *J Biol Chem* 1998;273:33750-8.
10. Ratziu V, Lalazar A, Wong L, et al. Zf9, a Kruppel-like transcription factor up-regulated in vivo during early hepatic fibrosis. *Proc Natl Acad Sci U S A* 1998;95:9500-5.
11. Merchant M, Weinberger SR. Recent advancements in surface-enhanced laser desorption/ionization-time of flight-mass spectrometry. *Electrophoresis* 2000;21:1164-77.
12. Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature* 2003;422:198-207.
13. Cadieux PA, Beiko DT, Watterson JD, et al. Surface-enhanced laser desorption/ionization-time of flight-mass spectrometry (SELDI-TOF-MS): a new proteomic urinary test for patients with urolithiasis. *J Clin Lab Anal* 2004;18:170-5.
14. Wadsworth JT, Somers KD, Stack BC, Jr., et al. Identification of patients with head and neck cancer using serum protein profiles. *Arch Otolaryngol Head Neck Surg* 2004;130:98-104.
15. Zhu XD, Zhang WH, Li CL, Xu Y, Liang WJ, Tien P. New serum biomarkers for detection of HBV-induced liver cirrhosis using SELDI protein chip technology. *World J Gastroenterol* 2004;10:2327-9.
16. Poon TC, Yip TT, Chan AT, et al. Comprehensive proteomic profiling identifies serum proteomic signatures for detection of hepatocellular carcinoma and its subtypes. *Clin Chem* 2003;49:752-60.
17. Howard JC, Heinemann C, Thatcher BJ, Martin B, Gan BS, Reid G. Identification of collagen-binding proteins in Lactobacillus spp. with surface-enhanced laser desorption/ionization-time of flight ProteinChip technology. *Appl Environ Microbiol* 2000;66:4396-400.
18. Reid G, Gan BS, She YM, Ens W, Weinberger S, Howard JC. Rapid identification of probiotic lactobacillus biosurfactant proteins by ProteinChip tandem mass spectrometry tryptic peptide sequencing. *Appl Environ Microbiol* 2002;68:977-80.
19. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 2001;98:5116-21.
20. Saar JD, Grothaus PC. Dupuytren's disease: an overview. *Plast Reconstr Surg.* 2000;106:125-134.
21. Elliot D. The early history of Dupuytren's disease. *Hand Clin.* 1999;15:1-19..
22. Jabaley ME. Surgical treatment of Dupuytren's disease. *Hand Clin.* 1999;15:109-26,
23. Nyberg LM, Jr., Bias WB, Hochberg MC, Walsh PC. Identification of an inherited form of Peyronie's disease with autosomal dominant inheritance and association with Dupuytren's contracture and histocompatibility B7 cross-reacting antigens. *J Uro.* 1982;128:48-51.
24. Wilbrand S, Ekbom A, Gerdin B. Dupuytren's contracture and sarcoma. *J Hand Surg* 2002;27:50-2.
25. Wilbrand S, Ekbom A, Gerdin B. Cancer incidence in patients treated surgically for Dupuytren's contracture. *J Hand Surg [Br].*2000;25:283-7.
26. Bains W. Vasoprotective VEGF as a candidate for prevention of recurrence of fibrotic diseases such as Dupuytren's contracture. *Med Hypotheses.* 2003;60:793-6.
27. Fitzgerald AM, Kirkpatrick JJ, Naylor IL. Dupuytren's Disease. The way forward? *Journal of Hand Surgery.* 1999;24B:395-9.
28. Mehta AI, Ross S, Lowenthal MS, et al. Biomarker amplification by serum carrier protein binding. *Dis*

Markers. 2003;19:-10.
29. Liotta LA, Ferrari M, Petricoin E. Clinical proteomics: written in blood. *Nature.* 2003;425:905.
30. Petricoin EF, Liotta LA. SELDI-TOF-based serum proteomic pattern diagnostics for early detection of cancer. *Curr Opin Biotechnol* 2004;15:24-30.
31. Adam BL, Qu Y, Davis JW, et al. Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. *Cancer Res.* 2002;62:3609-14.
32. Petricoin EF, Ardekani AM, Hitt BA, et al. Use of proteomic patterns in serum to identify ovarian cancer. *Lancet.* 2002;359:572-7.

Address Correspondence to:

Bing Siang Gan, MD, PhD, FRCSC, FACS
St. Joseph's Health Centre Suite L-021
268 Grosvenor Street
London, Ontario
CANADA N6A 4L6
E-mail: bsgan@lhrionhealth.ca