ARTHRITIS & RHEUMATISM
Vol. 52, No. 12, December 2005, pp 3801–3812
DOI 10.1002/art.21607

Discovery of New Rheumatoid Arthritis Biomarkers Using the Surface-Enhanced Laser Desorption/Ionization Time-of-Flight Mass Spectrometry ProteinChip Approach

Dominique de Seny,[1] Marianne Fillet,[1] Marie-Alice Meuwis,[1] Pierre Geurts,[2] Laurence Lutteri,[1] Clio Ribbens,[1] Vincent Bours,[1] Louis Wehenkel,[2] Jacques Piette,[1] Michel Malaise,[1] and Marie-Paule Merville[1]

Objective. **To identify serum protein biomarkers specific for rheumatoid arthritis (RA), using surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF-MS) technology.**

Methods. **A total of 103 serum samples from patients and healthy controls were analyzed. Thirty-four of the patients had a diagnosis of RA, based on the American College of Rheumatology criteria. The inflammation control group comprised 20 patients with psoriatic arthritis (PsA), 9 with asthma, and 10 with Crohn's disease. The noninflammation control group comprised 14 patients with knee osteoarthritis and 16 healthy control subjects. Serum protein profiles were obtained by SELDI-TOF-MS and compared in order to identify new biomarkers specific for RA. Data were analyzed by a machine learning algorithm called decision tree boosting, according to different preprocessing steps.**

Results. **The most discriminative mass/charge (m/z) values serving as potential biomarkers for RA were identified on arrays for both patients with RA versus controls and patients with RA versus patients with PsA. From among several candidates, the following peaks were highlighted: m/z values of 2,924 (RA versus controls on H4 arrays), 10,832 and 11,632 (RA versus controls on CM10 arrays), 4,824 (RA versus PsA on H4 arrays), and 4,666 (RA versus PsA on CM10 arrays). Positive results of proteomic analysis were associated with positive results of the anti–cyclic citrullinated peptide test. Our observations suggested that the 10,832 peak could represent myeloid-related protein 8.**

Conclusion. **SELDI-TOF-MS technology allows rapid analysis of many serum samples, and use of decision tree boosting analysis as the main statistical method allowed us to propose a pattern of protein peaks specific for RA.**

Rheumatoid arthritis (RA) is a chronic autoimmune disease of unknown etiology. RA is characterized by the development of synovitis, which is directly responsible for cartilage and bone degradation in multiple joints. Early identification of pathologic synovitis is of major importance, because synovitis represents the primary location of the inflammatory process in the rheumatoid joint and is a target for therapy (1,2). Two different approaches to such early identification have recently been developed.

The aim of the first approach is to directly evaluate synovial inflammation by means of new morphologic and/or metabolic imaging techniques such as high-resolution ultrasonography, fat-suppressed gadolinium-enhanced magnetic resonance imaging, and positron emission tomography with ^{18}F-fluorodeoxyglucose (3). The aim of the second approach is to identify specific biomarkers of the disease in the synovial fluid or serum. The recent discovery of an antibody response to citrullinated peptides in the serum of patients with RA illustrates this concept. With a sensitivity of ~76% and

Supported by grants from the "Région Wallonne," the National Fund for Scientific Research, and the "Fonds d'Investissement pour la Recherche Scientifique," Centre Hospitalier Universitaire, Liege, Belgium.

[1]Dominique de Seny, PhD, Marianne Fillet, PhD, Marie-Alice Meuwis, Laurence Lutteri, Clio Ribbens, MD, PhD, Vincent Bours, MD, PhD, Jacques Piette, PhD, Michel Malaise, MD, PhD, Marie-Paule Merville, PhD: University of Liège, and Centre Hospitalier Universitaire, Liege, Belgium; [2]Pierre Geurts, PhD, Louis Wehenkel, PhD: University of Liège, Liege, Belgium.

Drs. de Seny and Fillet contributed equally to this work.

Address correspondence and reprint requests to Dominique de Seny, PhD, Laboratory of Clinical Chemistry and Rheumatology, Tour de Pathologie +3 (CHU), University of Liège, 4000 Liege, Belgium. E-mail: ddeseny@chu.ulg.ac.be.

Submitted for publication December 22, 2004; accepted in revised form August 18, 2005.

a specificity of nearly 96%, identification of antibodies against cyclic citrullinated peptide (CCP) has strong clinical relevance (4) in patients with established diseases but is also a prognostic factor for progression of undifferentiated arthritis into RA (5). In some individuals, however, anti-CCP positivity years before the occurrence of any clinical symptoms (6) suggests that it reflects both latent and patent forms of RA and, therefore, not simply the pathophysiologic process. Routine measurements of IgM, IgG, and IgA rheumatoid factor (RF) may also be of some help in the diagnosis of RA (7,8), but these markers classically appear late in the course of disease and are less specific. There remains a clear need for the identification of other specific biomarkers of RA that would allow an early and specific diagnosis of the disease.

Protein differential display techniques such as 2-dimensional (2-D) gel electrophoresis, 1-D or 2-D liquid chromatography mass spectometry (MS), or surface-enhanced laser desorption/ionization time-of-flight MS (SELDI-TOF-MS) have become increasingly useful to establish fingerprint profiles of both disease and control states in large numbers of samples (9). Serum, plasma, urine, intestinal fluid, cell lysates, or cellular secretion products contain several thousand proteins or peptides that regulate a vast number of physiologic functions and may be related to pathology. Identification of protein patterns in these body fluids could allow a valid clinical diagnosis to be made before the onset of symptoms. Separation and characterization of complex protein mixtures are certainly not easy tasks, but each technique seems to have its own potential for reaching this goal.

SELDI-TOF-MS is a rapid and sensitive method that allows a large number of samples to be measured simultaneously in a relatively short period of time (10–12). Small sample volume is required, and the ProteinChip system appears suitable for profiling low molecular weight proteins. This system is currently used to identify biomarkers in the blood of patients with various cancers at different stages (13–19).

The SELDI-TOF-MS approach uses a variety of selective chips composed of different chromatographic, chemically active surfaces (anionic, cationic, hydrophobic, hydrophilic, or metal ion) on which a biologic sample (such as serum) is applied. On ProteinChip array, proteins are captured by, for example, Lewis acid–basis interaction, charge, hydrophobicity, or chromatographic affinity. Hence, each surface preferentially binds a particular class of protein based on its physiochemical properties and gives rise to a specific pattern. After several washes to eliminate nonspecific interactions, proteins are cocrystallized with an excess of energy-absorbing matrix molecules. A laser desorbs and ionizes the proteins. Ions are detected and, on a typical spectrum, are displayed with the corresponding mass/charge (m/z) ratio as a peak, the amplitude or area of which is proportional to protein abundance.

Here, we describe how SELDI-TOF-MS, when applied to serum, may generate complex protein profiles and how appropriate biostatistical analysis may select marker profiles specific to RA. For that purpose, SELDI-TOF-MS analyses of RA sera were compared with analyses of sera obtained from normal healthy subjects and from patients with noninflammatory rheumatic disease, such as osteoporosis (OA) (negative controls). The analyses were also compared with those performed on sera from patients with psoriatic arthritis (PsA), another immune-mediated inflammatory rheumatic disease, and from patients with immune-mediated inflammatory nonrheumatic diseases such as asthma and Crohn's disease (positive controls).

PATIENTS AND METHODS

Patients. This study was approved by the ethics committee of our academic hospital. A total of 103 serum samples from patients affected by various pathologies and from healthy controls were collected prospectively (from the beginning of 2002 until the end of 2003) into 10-ml Serum Separator Vacutainer tubes (Becton Dickinson, Franklin Lakes, NJ) and centrifuged at 3,000 rpm for 10 minutes. All sera were aliquoted and immediately frozen at −80°C until thawed specifically for SELDI-TOF-MS analysis. All of the individuals from whom serum samples were collected were Caucasian.

Thirty-four of the blood samples were obtained from patients fulfilling the 1987 American College of Rheumatology (formerly, the American Rheumatism Association) criteria for RA (20). Among the 34 patients with RA, 22 (65%) were women, the mean age was 52 years (range 26–79 years), and the mean disease duration was 8.8 years (range 0.5–22 years) (Table 1). IgM-RF positivity was observed in 26 (76%) of 34 patients. The mean C-reactive protein (CRP) serum level was 19 mg/liter (range 3–90 mg/liter). Concomitant treatments included methotrexate (MTX) at a mean dosage of 11.7 mg/week (range 5–20 mg/week) in 20 (59%) of 34 patients, and prednisolone at a mean dosage of 7 mg/day in 16 (47%) of 34 patients. The mean numbers of tender and swollen joints were 15 (range 6–32) and 11 (range 6–23), respectively. All patients were considered to have active disease because they exhibited at least 6 tender and swollen joints, which is the classically accepted criterion.

The inflammation control group comprised 20 patients with PsA, 9 with asthma, and 10 with Crohn's disease. Among the 20 patients with PsA, 10 (50%) were women, the mean age was 44 years (range 21–65 years), and the mean disease duration was 8.4 years (range 0.5–30 years). The mean CRP

Table 1. Epidemiologic characteristics of the patients and controls*

Characteristic	RA	PsA	Negative control
Age, years	52 (26–79)	44 (21–65)	45 (21–77)
% women	65	50	61
Disease duration, years	8.8 (0.5–22)	8.4 (0.5–30)	–
Serum CRP level, mg/liter	19 (3–90)	18 (0.7–69)	–
Concomitant therapy			
Methotrexate, mg/week†	11.7 (5–20)	15.6 (10–25)	–
Prednisolone, mg/day‡	7	5	–
No. of tender joints	15 (6–32)	15 (3–33)	–
No. of swollen joints	11 (6–23)	11 (5–18)	–
% anti-CCP2 antibody positive	82	0	3.30

* Except where indicated otherwise, values are the mean (range). Anti–cyclic citrullinated peptide 2 (anti-CCP2) antibodies were detected by enzyme-linked immunosorbent assay. RA = rheumatoid arthritis; PsA = psoriatic arthritis; CRP = C-reactive protein.
† Methotrexate was being received by 59% of the patients with RA and by 40% of the patients with PsA.
‡ Prednisolone was being received by 47% of the patients with RA and by 25% of the patients with PsA.

serum level was 18 mg/liter (range 0.7–69 mg/liter). Concomitant treatments included MTX at a mean dosage of 15.6 mg/week (range 10–25 mg/week) in 8 (40%) of 20 patients, and prednisolone at a mean dosage of 5 mg/day in 5 (25%) of 20 patients. The mean numbers of tender and swollen joints were 15 (range 3–33) and 11 (range 5–18), respectively. All patients were considered to have active disease because they exhibited at least 3 tender and swollen joints, which is the classically accepted criterion. Patients in the other positive control groups (asthma and Crohn's disease) were matched for age and sex with patients in the RA group.

Negative controls were 14 patients with knee OA and 16 unaffected healthy subjects, all of whom were age- and sex-matched with patients with RA.

The 103 serum samples were analyzed by an anti-CCP2 antibody enzyme-linked immunosorbent assay (ELISA; Euroimmun, Lubeck, Germany), according to manufacturer's instructions; the cutoff was 5 relative units/ml. Of the 34 patients with RA, 28 had positive results of the anti-CCP test. One serum sample (obtained from a healthy subject in the control group) was positive by anti-CCP testing. Thus, the sensitivity and specificity of this ELISA, based on the 103 serum samples, were calculated to be 82% and 98%, respectively.

A serum sample obtained from a healthy control subject was used as a quality control sample. It was used to determine reproducibility and as a control protein profile for each SELDI-TOF-MS experiment.

ProteinChip array preparation and analysis. Several chip arrays (Ciphergen Biosystems, Fremont, CA), including a strong anion-exchange surface (Q10), weak cation-exchange surface (CM10), or a hydrophobic surface (H4), were tested in order to determine the optimal profile in terms of the number and resolution of peaks. The pH (range 3–9) and salt concentration (range 30 m*M* to 1*M*) in washing buffers were optimized using ion-exchange arrays. The percentage of acetonitrile (ACN) (0–60%) was optimized for the H4 arrays. CM10 and H4 arrays were ultimately selected to give the best results. A pH value of 4 and 10% ACN were finally chosen as optimal conditions for serum analysis. The amount of proteins loaded onto the arrays was first optimized by diluting the serum from 1-fold to 70-fold in the corresponding binding buffer. Dilutions of 5-fold and 35-fold were selected as optimal conditions for CM10 and H4 arrays, respectively (data not shown).

Prior to sample loading, each spot of the H4 arrays was circled with a Pap pen (Zymed Laboratories, South San Francisco, CA). The CM10 and H4 arrays were activated with 10 μl of 10 m*M* HCl and 5 μl of ACN, respectively, and equilibrated with 10 μl of binding buffer (for CM10, 100 m*M* acetate, 30 m*M* NaCl [pH 4]; for H4, phosphate buffered saline [PBS], ACN 10%, and trifluoroacetic acid [TFA] 0.1%) for 5 minutes. Serum samples for SELDI-TOF-MS analysis were prepared by diluting 10 μl of serum with 40 μl of 100 m*M* acetate buffer (pH 4) for CM10 experiments and with 340 μl of PBS, ACN 10%, and TFA 0.1% for H4 experiments. Five microliters of each diluted serum mixture was applied, in duplicate, to a ProteinChip array and incubated for 1 hour at room temperature. After discarding the remaining sample, the CM10 and H4 arrays were washed 4 times and 2 times, respectively, with 10 μl of binding buffer for 5 minutes, followed by 2 (for CM10) and 4 (for H4) brief rinses with distilled water. The chips were air-dried and stored in the dark at room temperature until subjected to SELDI-TOF-MS analysis. A matrix solution (α-cyano-4-hydroxycinnamic acid [CHCA]) (Ciphergen Biosystems) was prepared according to the manufacturer's instructions in 50% volume/volume ACN and 0.5% TFA. Before SELDI-TOF-MS analysis, 1 μl of the saturated CHCA solution was applied onto each CM10 spot, and 1 μl of a 1:2 dilution was loaded twice on the H4 array and allowed to air dry.

Chips were read on a Protein Biological System II ProteinChip reader (Ciphergen Biosystems). All spectra were acquired in a positive mode and generated by averaging 130 laser shots at a laser intensity of 200 and 210, and a sensitivity of 8 and 9, for the CM10 and H4 arrays, respectively. The focus mass was set to 10,250 daltons.

Mass accuracy was calibrated externally using the All-in-1 Peptide Standard (Ciphergen Biosystems) complemented by myoglobin (MW 16,951.5) and cytochrome c (MW 12,360) in order to cover a larger range of mass (0–20,000 daltons). Calibration was carried out according to the manufacturer's instructions (21).

Preprocessing. Before data were analyzed, several processing steps were required, including baseline subtraction, normalization, or peak detection. Baseline subtraction was achieved by using a varying-width segmented convex hull algorithm that eliminates any baseline signal caused mostly by matrix distortions (21). All data were normalized according to the total ion current normalization function, by following the software instructions. Spectrum m/z values of <1,000 were not used for analysis, because the energy-absorbing matrix signal generally interfered with peak detection in this area. Due to biologic or technical reasons, there were further variations in the exact horizontal position of the same biologic peak from one spectrum to another. Thus, further preprocessing was necessary to reduce the dimensionality of the data and to take this noise into account. Two different approaches were considered.

In the first approach, peak detection was performed using ProteinChip Biomarker Wizard software version 3.0 (Ciphergen Biosystems). Peaks having an m/z ratio between 1,000 and 20,000 were autodetected with a signal-to-noise ratio of >3, and the peaks were clustered using second-pass peak selection with a signal-to-noise ratio of >2 and a 0.3% mass window.

In the second approach, the m/z axis was divided into nonoverlapping intervals whose sizes were increasing proportionally with the m/z values, and the intensity associated to each interval was taken as the sum of the intensities over the interval. The size of an interval starting at mass/charge m is computed as m.r, with r representing the parameter that determines the number of inputs that were used for the statistical analysis. Three r values were tested: 0.3%, 0.5%, and 1%. Unlike peak detection as carried out with ProteinChip Biomarker Wizard software, this second approach did not imply any filtering of the peaks; all m/z intervals were conserved as inputs for the statistical analysis.

Statistical analysis. *Decision tree boosting.* The data were analyzed by a machine learning algorithm called decision tree boosting. The decision tree method (22) is among the most popular of the learning algorithms, and it has already been used to analyze SELDI-TOF-MS measurements (23,24). A decision tree is a classification model represented by a tree in which each interior node is labeled with a test that compares an m/z value with an intensity threshold, and each terminal node is labeled with the name of a class. To retrieve the classification of a new patient, data generated by the (preprocessed) spectrum, which is related to his or her serum, are propagated into the tree by answering to the tests until a leaf node is reached, and the patient is classified according to the class label attached to that leaf. One drawback of this method is that it is highly unstable. A small modification of the set of patients may lead to a quite different tree. Hence, the prediction given by a single decision tree may not be very reliable. This instability translates into accuracy that usually is lower than that of other machine learning algorithms.

One very efficient way by which to circumvent this instability and improve decision tree accuracy is to use the ensemble method. According to this method, several trees are built instead of only 1, a classifier is defined by aggregating the classes predicted by these trees; the classification attributed to a new patient is represented by the majority class among classes predicted by all trees of the ensemble for this patient. Many tree-based ensemble methods exist. As we previously reported (25), trees with 4 different ensemble methods (i.e., bagging, boosting, random forests, and extra trees) have been tested for RA versus controls and RA versus PsA comparisons. In the present study, only decision tree boosting that gave competitive results with the other ensemble methods was considered (25). Boosting is a standard method (26) in which the ensemble of trees is built in sequence. Each tree of the sequence focuses on the data that are misclassified by the previous trees of the ensemble. More precisely, an AdaBoost algorithm was used (for review, see ref. 27) with classification and regression trees (28). Ensembles of 100 trees were constructed.

Evaluation of sensitivity and specificity. To obtain an unbiased estimate of the sensitivity and specificity of a diagnosis provided by the boosting algorithm, leave-one-out cross-validation was used. Using the leave-one-out approach, an unbiased diagnosis was obtained for each patient by removing all information concerning this patient (i.e., his or her 2 spectra) from the learning sample, building a model using the boosting algorithm from the remaining mass spectra, and then classifying the patient using the boosting model. Because a patient was described by 2 spectra, a diagnosis may be given in 2 ways using the boosting model: by classifying the 2 spectra independently from each other or by combining the classification of the 2 spectra. In the first case, sensitivity was estimated by the proportion of the 68 spectra from 34 patients with RA who were well classified by the boosting classifier, and the specificity was estimated by the proportion of the 138 spectra of 69 patients from the control group in whom RA was not diagnosed. In the second case, because the primary objective is to maximize sensitivity, a patient was diagnosed as having RA as soon as 1 of the spectra was classified as RA by the boosting classifier. Otherwise, the patient was classified as non-RA. The sensitivity achieved with the 2 combined spectra was then estimated by the proportion of RA patients who were well classified according to this rule, and the specificity was estimated by the proportion of patients from the control group in whom RA was not diagnosed.

Biomarker identification. As a first step to identify proteins that are potentially involved in RA, m/z peaks or intervals that are responsible for differentiating RA from control spectra must be identified. Biomarkers can be identified individually or by a multivariate analysis.

Univariate analysis. The classic statistical approach to determining the influence of the classification on the intensities of some m/z values is to use a statistical test to determine whether or not the distribution of the intensities at this position is significantly different in the RA group compared with the control group. The result of such an analysis is a *P* value that determines the probability of getting a more significant difference than that observed according to the statistical test. Hence, m/z values corresponding to small *P* values highlight significantly different protein concentrations between the 2 groups. By following the approach adopted (for review, see ref. 21), the discriminative power of peak values and m/z

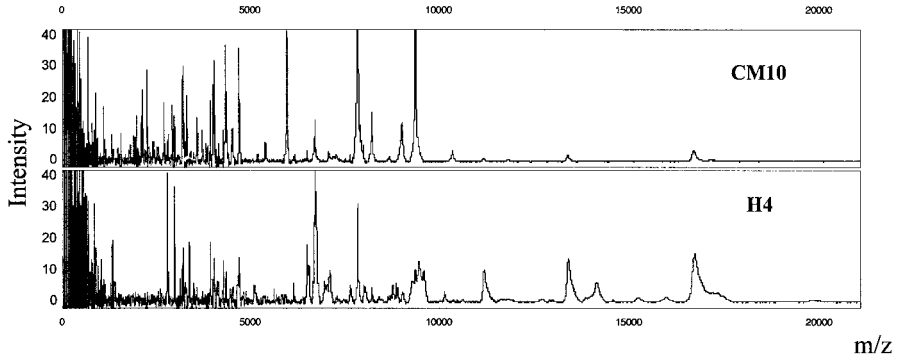

Figure 1. Protein mass spectra of a quality control serum sample from a healthy control subject, obtained on CM10 and H4 arrays. The intensity versus mass/charge (m/z) values are shown.

intervals was assessed according to a nonparametric Mann-Whitney test.

Multivariate analysis. One important characteristic of decision trees is that it is possible to compute from a tree the relative relevance or contribution of each variable to the classification. For each variable, this measure gives the percentage of information provided by the tree about the classification that can be attributed to this variable. The relative contribution of a variable to an ensemble of trees can then be obtained by averaging its relative contributions over all trees of the ensemble. Like the *P* value, this measure allows m/z values to be ranked according to their relevance for differentiating the disease and control groups. However, unlike the *P* value approach, which takes into account each variable individually, this approach considers all variables simultaneously, and hence it can take into account interactions among variables. Both approaches may thus provide substantially different results. The variable importance measure for a tree that we have used is the Shannon information measure (28). (See ref. 25 for a more detailed description of this measure in the context of tree ensembles.)

Western blotting. Myeloid-related protein 8 (MRP-8) was assessed by Western blot analysis. Briefly, 2 μl of serum was run on 12% Bis-Tris polyacrylamide gels (Invitrogen, Carlsbad, CA), transferred, and incubated with an anti–MRP-8 monoclonal antibody (1:500 dilution) (Biomedicals AG, Augst, Switzerland), followed by a mouse secondary antibody (1:5,000 dilution), and MRP-8 was revealed by the enhanced chemiluminescence detection method (ECL kit; Amersham Pharmacia Biotech, Little Chalfont, Buckinghamshire, UK).

Interaction discovery mapping (IDM) affinity beads. MRP-8 was depleted from RA and control serum samples using IDM affinity beads (Ciphergen Biosystems) coated with MRP-8 monoclonal antibody (Biomedicals AG). The experiment was performed according to the manufacturer's instructions.

RESULTS

Optimization of the experimental conditions and reproducibility evaluation. Several parameters have a large influence on the reproducibility and number of peaks detected in protein profiles and thus needed to be optimized. In order to simplify the procedure and to obtain good reproducibility, serum samples were not fractionated. In an effort to increase the size of the proteome examined and to enhance the chance of detecting protein biomarkers, 103 serum samples were analyzed in parallel on 2 types of ProteinChip arrays. The protein profiles obtained on CM10 (cation exchange) and H4 (hydrophobic) arrays were found to give the best results in terms of the number and resolution of peaks. These profiles, as obtained with the quality control serum sample, are shown in Figure 1.

Chips corresponding to the 103 serum samples were read over the course of a week in order to limit variability across time. Standardization of experimental conditions was carried out in an effort to minimize the effects of irrelevant sources of fluctuation, and coefficients of variation (CVs) were calculated to evaluate the reproducibility of experiments using the SELDI-TOF-MS approach. These CV values were obtained by adding a quality control serum sample on 8 spots of CM10 or H4 arrays according to the protocol described in Patients and Methods. The procedure was performed at the beginning of the study of the 103 serum samples and again 6 months later. CVs were calculated after the normalization process by comparing the intensity of 10 common peaks selected throughout the 8 spectra collected from the same array. CVs were also established by comparing interchip variation at an interval of 6 months. Intravariation of CM10 and H4 arrays were evaluated to be 9% and 16.6%, respectively, at the beginning of the study and 12% (CM10 array) and 20% (H4 array), respectively, 6 months later. Interchip variation across time was determined to be 20% and 25% for the CM10 and H4 arrays, respectively.

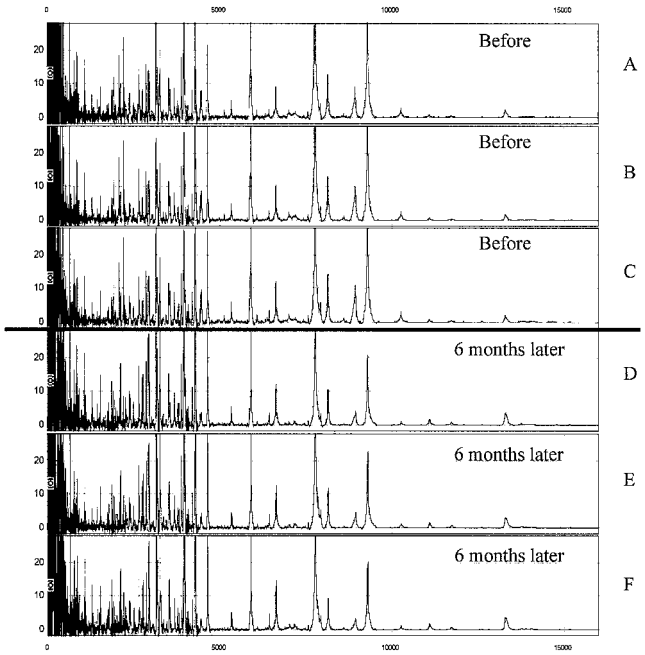

Figure 2. Reproducibility of the protein profile spectra of the quality control serum sample on the CM10 arrays, as demonstrated before starting the analysis of the 103 serum samples from patients and healthy controls (spectra A–C) and 6 months later (spectra D–F).

Figure 2 shows 3 spectra for a quality control serum sample on CM10 array, before starting the analysis and 6 months later. These spectra show similar protein profiles, which demonstrate the standardization of our experimental procedure over time.

At least 30 samples from each of the 3 classification groups (RA, inflammation control, and negative control) were profiled in a standardized manner. This number of samples was sufficient to give >90% statistical confidence in a single marker, with *P* values less than 0.01, and was also sufficient to allow use of different forms of multivariate analysis.

Data analysis. Two hundred six spectra (each serum sample was loaded in duplicate) from 103 serum samples were collected on each array (CM10 and H4). Peak detection and alignment resolved 140 peaks on CM10 arrays and 104 peaks on H4 arrays in the mass range of 1–20 kd. In contrast, the proportional integration of the mass range yielded 1,026, 628, and 319 mass intervals, respectively, for r = 0.3%, r = 0.5%, and r = 1%. This corresponded in each case to the number of input features provided to the boosting algorithm.

Sensitivity and specificity. RA spectra were first compared with control spectra (inflammation and noninflammation controls). Table 2 shows the sensitivity/specificity values estimated by decision tree boosting on both surfaces, with different r values and integrated peaks. Sensitivities of 69–78% and 81–85% for classifying individual spectra were obtained on CM10 and H4 arrays, respectively. Taking into account the 2 spectra corresponding to 1 patient, the sensitivity rose to a range of 79–91% on CM10 and 88–97% on H4. However, the specificities of 78–90% and 91–95% on CM10 and H4 arrays, respectively, were slightly decreased with 2 combined spectra, to 75–85.5% on CM10 array and 87–91% on H4 array.

Table 2. Sensitivities and specificities obtained by decision tree boosting analysis on CM10 and H4 arrays according to different preprocessing approaches, for patients with RA versus controls*

Preprocessing approach	2 independent spectra		2 combined spectra	
	Sensitivity	Specificity	Sensitivity	Specificity
CM10				
r = 0.3%	78	90	88	85.5
r = 0.5%	76.5	87	91	77
r = 1%	78	88	91	84
Integrated peaks	69	78	79	75
H4				
r = 0.3%	85	91	94	87
r = 0.5%	85	91	94	90
r = 1%	84	95	97	91
Integrated peaks	81	93	88	90

* Values are the percent. Two approaches were used, as follows: classifying 2 spectra from the same patient independently of each other, and combining classification of the 2 spectra. RA = rheumatoid arthritis.

RA spectra were also compared with PsA spectra. As shown in Table 3, the sensitivities for classifying individual spectra were 84–90% on CM10 array and

Table 3. Sensitivities and specificities obtained by decision tree boosting analysis on CM10 and H4 arrays according to different preprocessing approaches, for patients with RA versus patients with PsA*

Preprocessing approach	2 independent spectra		2 combined spectra	
	Sensitivity	Specificity	Sensitivity	Specificity
CM10				
r = 0.3%	84	62	85	48
r = 0.5%	84	64	88	38
r = 1%	90	71	85	33
Integrated peaks	87	55	94	48
H4				
r = 0.3%	93	79	97	67
r = 0.5%	94	86	97	76
r = 1%	90	76	94	71
Integrated peaks	94	83	91	71

* Values are the percent. Two approaches were used, as follows: classifying 2 spectra from the same patient independently of each other, and combining classification of the 2 spectra. RA = rheumatoid arthritis; PsA = psoriatic arthritis.

90–94% on H4 array. Again, combining the duplicates improved sensitivity (85–94% on CM10 and 91–97% on H4). However, combining the duplicates decreased the specificity considerably.

Biomarker classification. *RA versus non-RA control group.* Table 4 presents the most discriminant m/z intervals or peaks provided by the boosting algorithm and by *P* value analysis for the 2 preprocessing approaches: mass intervals (r = 0.3% or r = 0.5%) and integrated peaks. For each m/z value, the first number represents the percentage of information attributed to this value (these numbers sum to 100% over all features) based on the multivariate analysis (decision tree boosting analysis). The second number is the rank of this m/z value when all m/z values are ordered according to the *P* value. It was observed that for a given preprocessing approach, the most discriminant m/z values provided by the multivariate analysis were not necessarily identical to those provided by the *P* value. This was particularly obvious on CM10 arrays, where the most discriminant mass range according to boosting (around 1,816 daltons) was not well ranked according to the *P* values (rank 594 with r = 0.3%, and rank 74 with r = 0.5%). The 2 preprocessing approaches may also highlight different m/z values for a given statistical approach. As an example, the m/z range around 1,816, which was ranked as the first value using the boosting analysis on CM10 with the mass interval preprocessing approach, was not ranked with peak integration analysis.

Nonetheless, a careful analysis of the generated data allowed us to focus our attention on peaks highlighted by the 2 preprocessing procedures, by the 2 statistical analyses, or by both approaches. For example, the m/z range around 2,924 daltons was considered as the most relevant value for discriminating RA versus controls on H4 arrays. Indeed, it was first according to multivariate analysis (% information) and *P* value (rank) analysis (r = 0.5% and integrated peaks), but also was second in percentage and the first with r = 0.3%. Concordance between boosting analysis and *P* value analysis was also found for m/z = 10,832 on CM10 arrays, using integrated peaks as the preprocessing approach. Figure 3 illustrates this potential biomarker.

According to the results obtained in previous studies (29,30), the 10,832 peak visualized by SELDI-TOF in serum from patients with RA was suspected to be MRP-8. In order to confirm this hypothesis, serum samples obtained from 1 healthy person and from 1 patient with RA were subjected to Western blotting. MRP-8 was detected (by MRP-8 monoclonal antibody)

Table 4. Most discriminant mass/charge (m/z) values obtained on CM10 and H4 arrays for patients with rheumatoid arthritis versus controls

	m/z	% information*	Rank†
CM10 array, r = 0.3%	**1,816–1,821**	10.90	594
	1,810–1,816	7.77	296
	15,764–15,812	3.11	11
	4,110–4,123	2.76	16
	11,628–11,663	2.39	1
	9,141–9,169	1.79	111
	1,939–1,945	1.65	45
	5,849–5,867	1.46	59
	3,111–3,121	1.43	103
	11,663–11,699	1.34	3
	10,825–10,859	1.28	4
CM10 array, r = 0.5%	**1,807–1,816**	15.69	74
	11,610–11,669	3.55	1
	4,084–4,105	2.72	20
	8,559–8,603	2.66	127
	11,260–11,316	2.46	8
	4,105–4,126	2.11	10
	1,940–1,950	1.83	86
	5,840–5,870	1.44	28
	1,467–1,475	1.33	19
	4,571–4,595	1.18	53
	15,109–15,186	1.16	62
CM10 array, integrated peaks	**10,833**	8.84	1
	1,944	4.68	19
	4,668	4.44	2
	11,632	4.14	3
	5,492	3.24	10
	8,563	2.95	35
	4,128	2.88	4
	11,706	2.72	5
	2,900	1.69	82
	2,847	1.46	49
	7,438	1.4	13
H4 array, r = 0.3%	8,052–8,076	4.19	43
	2,923–2,932	4.05	1
	1,057–1,061	3.51	52
	4,820–4,835	3.21	422
	6,181–6,200	2.27	165
	10,438–10,469	2.03	119
H4 array, r = 0.5%	**2,924–2,938**	6.15	1
	4,809–4,834	4.25	356
	2,223–2,234	3.93	2
	1,059–1,064	3.29	4
	1,092–1,097	2.99	23
	1,522–1,530	2.26	12
H4 array, integrated peaks	**2,924**	11.38	1
	4,538	7.37	44
	10,441	5.22	31
	4,825	4.42	92
	2,778	4.2	21
	5,686	3.11	5

* Relative contribution of a mass/charge value, by multivariate analysis.
† According to *P* value.

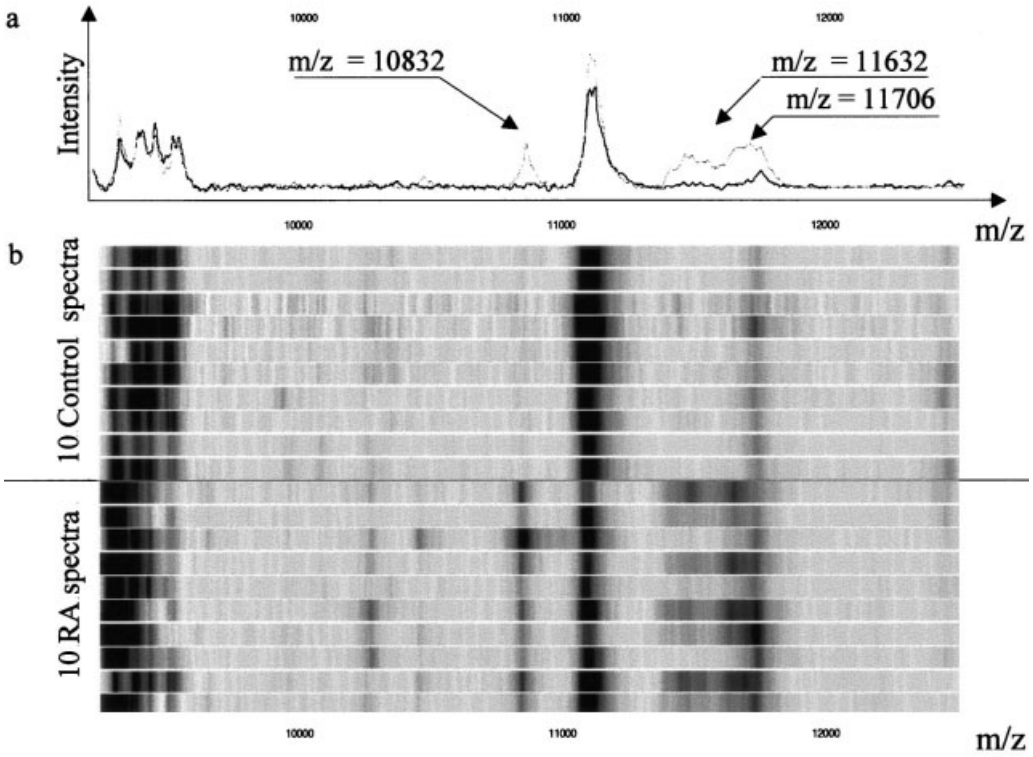

Figure 3. Potential biomarkers (m/z = 10,832, 11,632, and 11,706) on CM10 array, distinguishing patients with rheumatoid arthritis (RA) from controls. **a,** Mean values of all spectra. The black and grey lines represent values for controls and RA patients, respectively. **b,** Gel view of spectra for samples collected from 10 controls and 10 patients with RA.

in the serum from the patient with RA but not in the control (Figure 4a). The presence of MRP-8 was further confirmed in all other RA sera tested (n = 8), in 4 of the 8 PsA sera tested, and in none of the 8 negative control sera tested. This identification of MRP-8 was correlated with the presence of the 10,832 peak in the corresponding spectra obtained by SELDI-TOF.

In a second experiment, we carried out depletion of MRP-8 from RA serum by using IDM affinity beads coated with MRP-8 monoclonal antibody. As shown in Figure 4b, the peak intensity of 10,832 decreased after serum depletion (spectra C and D) compared with the crude RA serum sample (spectra A and B) and the quality control serum sample (E).

Our data suggest that the 10,832 peak could be MRP-8. However, definitive identification of this peak will be obtained only by MS/MS analysis. The m/z values around 11,630 (Figure 3) were also rather well classified when determining its relative position according to preprocessings and statistical analysis.

RA versus PsA group. Table 5 illustrates the most discriminative m/z intervals when comparing the RA group versus the PsA group. The m/z around 4,824 was perfectly classified according to the 2 preprocessing approaches and the 2 statistical analysis on H4 arrays, whereas the m/z around 4,666 was also of interest on CM10 arrays.

Association with anti-CCP positivity. Using multivariate analysis according to the 2 combined spectra approach, sensitivity ranged from 79% to 97% (Tables 2 and 3), with a mean of 91%. This value was therefore chosen to represent the sensitivity of the proteomic analysis of RA sera. Thus, 31 of 34 RA sera were considered positive. Among the RA sera, 28 of 34 were anti-CCP positive. Positivity by proteomic analysis was associated with positivity by anti-CCP testing, as shown by calculation of an odds ratio of 13.5 (95% confidence interval 1–185). In an attempt to study quantitative correlations between proteomic analysis and anti-CCP positivity, we plotted the intensity of the 10,832 peak of RA serum samples (n = 34), one of the major biomar-

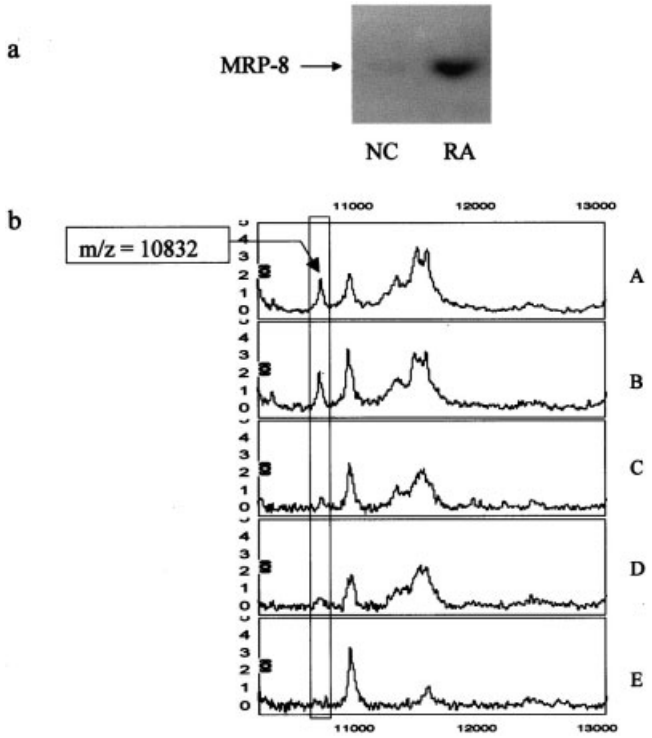

Figure 4. a, Detection of myeloid-related protein 8 (MRP-8) by Western blotting in serum from 1 normal control (NC) and 1 patient with rheumatoid arthritis (RA). **b,** Depletion of MRP-8 from RA serum by using interaction discovery mapping (IDM) affinity beads coated with MRP-8 monoclonal antibody. A and B show a crude serum sample from a patient with RA; C and D show an RA serum sample after MRP-8 depletion with IDM beads; E shows a crude serum sample from a healthy control.

kers, against the log of anti-CCP antibody concentrations (as detected by ELISA) and observed a significant positive linear correlation ($r = 0.435$; $P = 0.01$).

DISCUSSION

In this study, we used methodology based on SELDI-TOF-MS and decision tree boosting analysis, in adequately controlled patient groups, to select new biomarkers for RA. Based on our results, we propose a pattern of protein peaks that are specific to RA and independent of the inflammatory process and of putative concomitant therapies. Indeed, the RA and the PsA groups shared many epidemiologic characteristics, including similar disease duration, clinical disease activity, serum CRP levels, and a comparable percentage of patients receiving MTX.

SELDI-TOF-MS is a new proteomic approach that allows multiple serum samples to be analyzed in a relatively short time. This analysis is based on a comparison of protein profiles between 2 sample groups. Up-regulated or down-regulated proteins are underlined and characterized as potential biomarkers according to several statistical analyses.

However, special care must be applied in order to optimize the reliability and reproducibility of proteomic patterns obtained by SELDI-TOF-MS. Indeed, variations due to sample collection, data preprocessing, and statistical analysis can be problematic (31–33), giving

Table 5. Most discriminant mass/charge (m/z) values obtained on CM10 and H4 arrays for patients with rheumatoid arthritis versus patients with psoriatic arthritis

	m/z	% information*	Rank†
CM10 array, r = 0.3%	5,490–5,508	5.51	14
	11,031–11,066	4.23	4
	13,229–13,269	3.86	3
	4,653–4,668	3.86	2
	1,816–1,821	3.36	880
	4,110–4,123	2.42	126
CM10 array, r = 0.5%	2,835–2,850	6.25	14
	13,177–13,244	4.86	1
	1,807–1,816	4.8	319
	3,819–3,838	2.51	63
	13,515–13,583	2.25	17
	4,642–4,666	2.13	9
CM10 array, integrated peaks	**4,666**	8.84	1
	5,492	6.42	3
	4,647	3.75	6
	3,974	2.83	35
	2,261	2.59	75
	11,651	2.56	24
H4 array, r = 0.3%	**4,820–4,835**	13.13	1
	5,597–5,614	5	1,017
	4,534–4,549	4.72	10
	1,057–1,061	3.71	25
	6,181–6,200	2.6	380
	8,052–8,076	2.5	90
H4 array, r = 0.5%	**4,809–4,834**	19.27	1
	4,524–4,547	6.34	9
	1,059–1,064	4.75	3
	5,579–5,608	3.54	461
	1,444–1,452	3.32	5
	6,877–6,912	2.17	262
H4 array, integrated peaks	**4,824**	26.32	1
	5,684	8.2	54
	2,924	5.39	3
	10,525	4.2	42
	2,778	3.56	11
	10,441	2.93	38

* Relative contribution of a mass/charge value, by multivariate analysis.
† According to *P* value.

rise to artifacts. Hence, the impact of freeze–thaw cycles on protein profiles must be well considered, standard protocols must be developed to minimize unwanted fluctuation, and CVs between ProteinChips must be calculated by using common peaks across different spectra. Chip variability was controlled by using chips from the same lot, and the chemicals used during a given experiment were from the same batch. Matrix composition and instrument settings are extrinsic factors that also influence reproducibility and must be optimized. Calibration of the instrument must be performed frequently. Finally, normalization, baseline subtraction, and peak detection are processing steps that must be well achieved.

To avoid variation in the procedure, freshly collected sera were immediately aliquoted, stored at −80°C, and thawed only once. Use of quality control serum allowed detection of any unusual features during the process. Such precautions led to very good reproducibility of the protein peak patterns.

One of the challenges in the analysis of SELDI-MS–generated data is to avoid the false discovery of protein peaks with a discriminatory power due to random variation. A safeguard against this problem lies in the choice of the machine learning algorithm and the validation method. Several decision tree–based ensemble methods were tried, and boosting was found to be among the best alternatives for this kind of problem (25). The *k*-nearest neighbors method and support vector machines with linear kernels were also applied, but none of these methods was able to reach the same level of sensitivity and specificity as that associated with decision tree boosting. Furthermore, a very important advantage of boosting and other ensemble methods with trees is the ability to estimate from the trees the contribution of each variable to the classification. Although single decision trees are also able to select relevant variables, the ranking provided by an ensemble of trees is usually much more robust.

Special care was also taken in choosing the validation method. Leave-one-out cross-validation ensured an unbiased estimate of the sensitivity and specificity of our classifier. Hence, because boosting reaches high sensitivity and specificity values under leave-one-out, we can be quite confident in the best biomarkers that were proposed by this method.

Preprocessing was also an important step in avoiding detection of artifacts. We systematically compared 2 preprocessing procedures, mass intervals and integrated peaks, on 2 different arrays, CM10 and H4, and obtained good and comparable sensitivity and specificity values. In this study, all samples were examined in duplicate. The results with boosting showed that the combination of 2 spectra per patient at the time of diagnosis can significantly improve sensitivity without losing too much specificity.

We next conducted our analysis toward classification of the biomarkers by using 2 statistical approaches, a univariate (*P* value) and a multivariate (decision tree boosting) analysis on the 2 preprocessings. If a value of r = 1% in the mass interval method usually gives the best sensitivity and specificity values, these mass intervals are too high to identify a single biomarker with sufficient accuracy for further analysis. As a result, only the lower values of r = 0.3% and r = 0.5% were selected for this task.

Several m/z values were identified as biomarkers specific for RA. The likelihood for a peak to be a biomarker is increased if it has been well classified on the basis of the 2 statistical analyses and 2 preprocessings. This situation was encountered, and also confirmed, in the 2 comparisons (RA versus non-RA and RA versus PsA) with both arrays. However, variations in the biomarker classification may also be observed. Discrepancies linked to statistical analyses can be explained as follows. The comparison between boosting variable classification and *P* values shows the interest of a multivariate analysis in identifying biomarkers. Indeed, the discriminative power of some m/z values appears only when they are combined with each other. Variables that correspond to elevated *P* values can be identified only by multivariate analysis. Discrepancies linked to data processing can be explained by the fact that peak detection may filter out potentially important biomarkers, even with an optimal setting of the Biomarker Wizard software, as was observed for the m/z value of 1,816. This value was first classified on CM10 for discriminating RA versus control, whereas it was not selected from among the 140 peaks with peak detection.

At least 2 proteomic studies in RA have been previously published; 1 of these studies focused on serum (34), and the other focused on synovial fluid (29). In the first study, using a 2-D gel approach, α-enolase was proposed as a biomarker for RA (34). In the second study, which used the same SELDI-TOF-MS technology as that used in our study, several biomarkers discriminating between RA and osteoarthritis, particularly MRP-8, were identified (29). Recently, this MRP-8 was confirmed by MS/MS to be present in serum of patients with erosive RA (30).

Due to our experimental conditions with a cutoff

of 20,000 daltons, it is clear that our biomarker classification does not include α-enolase, whereas MRP-8 (MW 10,850) could have been detected. Preliminary results obtained by Western blotting and IDM affinity beads coated with MRP-8 antibody have suggested that the 10,832 peak observed in our RA patients could be the MRP-8 protein. MRP-8 is, however, not totally specific to RA, and is also observed in serum and synovial fluid from patients with active juvenile RA (30) as well as in the inflamed synovial tissue of patients with RA and patients with spondylarthropathy (35). It was also shown in our study that the 10,832 peak was not a discriminant value in the comparison by SELDI-TOF and boosting analysis of patients with RA versus patients with PsA. Because MRP-8 is specifically released by activated monocytes upon interaction with activated endothelial cells under inflammatory conditions (30), it could merely represent an excellent marker of the early inflammatory process. This finding highlights the necessity of using several biomarkers for optimizing a specific diagnosis of RA. Other purification procedures for identification of the most specific RA biomarkers are currently being investigated.

However, even without the purification of several protein peaks selected as specific biomarkers for RA, SELDI-TOF-MS technology allowed us to quickly analyze >100 serum samples, using decision tree boosting analysis as the main statistical method and a previously proposed pattern of protein peaks specific for RA. The potential clinical relevance of our results is further highlighted by the statistical correlations observed between the positivity of RA sera for the proteomic analysis and their positivity for the anti-CCP test, as well as the correlations between the intensity of the 10,832 peak and anti-CCP antibody concentrations. The biologic significance of such correlations remains unknown, however. The relevance of this proteomic technique for the diagnosis of RA remains investigational, but the existence of serum-based marker panels with sufficient sensitivity and specificity could facilitate the screening of individuals at high risk for developing RA, favor early identification of the disease, and allow us to predict the response to therapeutic agents. Validation of our first model with a complete independent set of samples collected in different institutions according to our standardized procedure is in progress. Moreover, for a better understanding of RA disease, identification of the most specific biomarkers described in this report would be of great interest.

ACKNOWLEDGMENTS

M. F. and M.-P. M. are Research Associates, P.G. is Senior Research Assistant, and J. P. is Research Director at the National Fund for Scientific Research. The authors thank E. Louis and R. Louis for providing serum samples from patients with Crohn's disease and patients with asthma, as well as Aline Desoroux, Simone Gaspar, and Delphine Bourgeois for providing expert technical assistance.

REFERENCES

1. Stenger AA, van Leeuwen MA, Houtman PM, Bruyn GA, Speerstra F, Barendsen BC, et al. Early effective suppression of inflammation in rheumatoid arthritis reduces radiographic progression. Br J Rheumatol 1998;37:1157–63.
2. Feldmann M, Maini RN. Anti-TNF α therapy of rheumatoid arthritis: what have we learned? Annu Rev Immunol 2001;19:163–96.
3. Beckers C, Ribbens C, Andre B, Marcelis S, Kaye O, Mathy L, et al. Assessment of disease activity in rheumatoid arthritis with (18)F-FDG PET. J Nucl Med 2004;45:956–64.
4. Schellekens GA, de Jong BA, van den Hoogen FH, van de Putte LB, van Venrooij WJ. Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies. J Clin Invest 1998;101:273–81.
5. Van Gaalen FA, Linn-Rasker SP, van Venrooij WJ, de Jong BA, Breedveld FC, Verweij CL, et al. Autoantibodies to cyclic citrullinated peptides predict progression to rheumatoid arthritis in patients with undifferentiated arthritis: a prospective cohort study. Arthritis Rheum 2004;50:709–15.
6. Nielen MM, van Schaardenburg D, Reesink HW, van de Stadt RJ, van der Horst-Bruinsma IE, de Koning MH, et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. Arthritis Rheum 2004;50:380–6.
7. Swedler W, Wallman J, Froelich CJ, Teodorescu M. Routine measurement of IgM, IgG, and IgA rheumatoid factors: high sensitivity, specificity, and predictive value for rheumatoid arthritis. J Rheumatol 1997;24:1037–44.
8. Jonsson T, Steinsson K, Jonsson H, Geirsson AJ, Thorsteinsson J, Valdimarsson H. Combined elevation of IgM and IgA rheumatoid factor has high diagnostic specificity for rheumatoid arthritis. Rheumatol Int 1998;18:119–22.
9. Rodland KD. Proteomics and cancer diagnosis: the potential of mass spectrometry. Clin Biochem 2004;37:579–83.
10. Weinberger SR, Morris TS, Pawlak M. Recent trends in protein biochip technology. Pharmacogenomics 2000;1:395–416.
11. Pusch W, Flocco MT, Leung SM, Thiele H, Kostrzewa M. Mass spectrometry-based clinical proteomics. Pharmacogenomics 2003;4:463–76.
12. Tang N, Tornatore P, Weinberger SR. Current developments in SELDI affinity technology. Mass Spectrom Rev 2004;23:34–44.
13. Wadsworth JT, Somers KD, Stack BC Jr, Cazares L, Malik G, Adam BL, et al. Identification of patients with head and neck cancer using serum protein profiles. Arch Otolaryngol Head Neck Surg 2004;130:98–104.
14. Koopmann J, Zhang Z, White N, Rosenzweig J, Fedarko N, Jagannath S, et al. Serum diagnosis of pancreatic adenocarcinoma using surface-enhanced laser desorption and ionization mass spectrometry. Clin Cancer Res 2004;10:860–8.
15. Zhukov TA, Johanson RA, Cantor AB, Clark RA, Tockman MS. Discovery of distinct protein profiles specific for lung tumors and pre-malignant lung lesions by SELDI mass spectrometry. Lung Cancer 2003;40:267–79.
16. Xiao XY, Tang Y, Wei XP, He DC. A preliminary analysis of

non-small cell lung cancer biomarkers in serum. Biomed Environ Sci 2003;16:140–8.
17. Won Y, Song HJ, Kang TW, Kim JJ, Han BD, Lee SW. Pattern analysis of serum proteome distinguishes renal cell carcinoma from other urologic diseases and healthy persons. Proteomics 2003;3:2310–6.
18. Vlahou A, Laronga C, Wilson L, Gregory B, Fournier K, McGaughey D, et al. A novel approach toward development of a rapid blood test for breast cancer. Clin Breast Cancer 2003;4: 203–9.
19. Lehrer S, Roboz J, Ding H, Zhao S, Diamond EJ, Holland JF, et al. Putative protein markers in the sera of men with prostatic neoplasms. BJU Int 2003;92:223–5.
20. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.
21. Fung ET, Enderwick C. ProteinChip clinical proteomics: computational challenges and solutions. Biotechniques 2002;Suppl:34–8.
22. Breiman L, Friedman JH, Olsen RA, Stone CJ. Classification and regression tree. New York: Chapman & Hall; 1984.
23. Rai AJ, Zhang Z, Rosenzweig J, Shih IeM, Pham T, Fung ET, et al. Proteomic approaches to tumor marker discovery. Arch Pathol Lab Med 2002;126:1518–26.
24. Adam BL, Qu Y, Davis JW, Ward MD, Clements MA, Cazares LH, 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.
25. Geurts P, Fillet M, de Seny D, Meuwis MA, Malaise M, Merville MP, et al. Proteomic mass spectra classification using decision tree based ensemble methods. Bioinformatics 2005;21:3138–45.
26. Hastie T, Tibshirani R, Friedman JH. The elements of statistical learning: data mining, inference, and prediction. New York: Springer. 2001.
27. Freund Y, Schapire RE, editors. A decision-theoretic generalization of on-line learning and an application to boosting. Proceedings of Second European Conference on Computational Learning Theory; 1995 March 13–15; Barcelona, Spain. p. 23–7.
28. Wehenkel L. Automatic learning techniques in power systems. Boston: Kluwer Academic; 1998.
29. Uchida T, Fukawa A, Uchida M, Fujita K, Saito K. Application of a novel protein biochip technology for detection and identification of rheumatoid arthritis biomarkers in synovial fluid. J Proteome Res 2002;1:495–9.
30. Liao H, Wu J, Kuhn E, Chin W, Chang B, Jones MD, et al. Use of mass spectrometry to identify protein biomarkers of disease severity in the synovial fluid and serum of patients with rheumatoid arthritis. Arthritis Rheum 2004;50:3792–803.
31. Baggerly KA, Morris JS, Coombes KR. Reproducibility of SELDI-TOF protein patterns in serum: comparing datasets from different experiments. Bioinformatics 2004;20:777–85.
32. Schaub S, Wilkins J, Weiler T, Sangster K, Rush D, Nickerson P. Urine protein profiling with surface-enhanced laser-desorption/ionization time-of-flight mass spectrometry. Kidney Int 2004;65: 323–32.
33. Diamandis EP. Mass spectrometry as a diagnostic and a cancer biomarker discovery tool: opportunities and potential limitations [review]. Mol Cell Proteomics 2004;3:367–78.
34. Saulot V, Vittecoq O, Charlionet R, Fardellone P, Lange C, Marvin L, et al. Presence of autoantibodies to the glycolytic enzyme α-enolase in sera from patients with early rheumatoid arthritis. Arthritis Rheum 2002;46:1196–201.
35. Tilleman K, van Beneden K, Dhondt A, Hoffman I, de Keyser F, Veys E, et al. Chronically inflamed synovium from spondyloarthropathy and rheumatoid arthritis investigated by protein expression profiling followed by tandem mass spectrometry. Proteomics 2005;5:2247–57.