Journal of proteome research

Article

pubs.acs.org/jpr

Verification of a Parkinson's Disease Protein Signature in T-Lymphocytes by Multiple Reaction Monitoring

Tiziana Alberio,*[,†,‡] Kelly McMahon,[§] Manuela Cuccurullo,[∥] Lee A. Gethings,[§] Craig Lawless,[⊥] Maurizio Zibetti,[#] Leonardo Lopiano,[#] Johannes P. C. Vissers,[§] and Mauro Fasano[†,‡]

[†]Division of Biomedical Sciences, Department of Theoretical and Applied Sciences, University of Insubria. Via Luciano Manara 7, Busto Arsizio I-21052, Italy

[‡]Center of Neuroscience, University of Insubria, Via Alberto da Giussano 12, Busto Arsizio I-21052, Italy

[§]Waters Corporation, Atlas Park, Simonsway, Manchester M22 5PP, United Kingdom

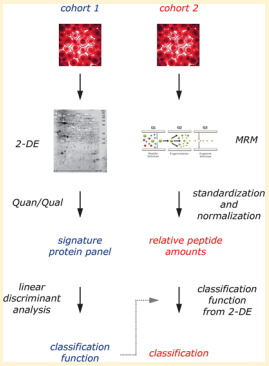
[∥]Waters Corporation, Viale dell'Innovazione 3, Milano I-20126, Italy

[⊥]Faculty of Life Sciences, University of Manchester, Manchester M13 9PT, United Kingdom

[#]Department of Neuroscience, University of Torino, Via Cherasco 15, Torino I-10126, Italy

Supporting Information

ABSTRACT: Diagnosis of Parkinson's disease, the second most common neurodegenerative disease, is based on the appearance of motor symptoms. A panel of protein biomarkers in the T-lymphocyte proteome was previously proposed as a Parkinson's disease signature. Here, we designed an LC−MS based method to quantitatively evaluate this protein signature by multiple reaction monitoring (MRM) in T-lymphocytes and peripheral blood mononuclear cells from a new cohort of nine patients with Parkinson's disease and nine unaffected subjects. Patients were classified using the discriminant function obtained from two-dimensional electrophoresis and protein amounts measured by MRM, thus assigning seven controls out of nine as true negatives and nine patients out of nine as true positives. A good discriminant power was obtained by selecting a subset of peptides from the protein signature, with an area under the receiver operating characteristic curve of 0.877. A similar result is achieved by evaluating all peptides of a selected panel of proteins (gelsolin, moesin, septin-6, twinfilin-2, lymphocyte-specific protein 1, vimentin, transaldolase), with an area under the curve of 0.840. Conversely, the signature was not able to classify the enrolled subjects when evaluated in whole mononuclear cells. Overall, this report shows the portability of the proposed method to a large-scale clinical validation study.



KEYWORDS: *biomarkers, multiple reaction monitoring, Parkinson's disease, peripheral blood lymphocytes*

INTRODUCTION

Parkinson's disease (PD) is the second most common age related neurodegenerative disease. It affects about six million people worldwide, with a prevalence raising rapidly with the age. The manifestation of Parkinson's disease usually starts with the appearance of cardinal motor symptoms, that is, tremor at rest, rigidity, and bradykinesia. The typical motor impairment is the result of a progressive degeneration of dopaminergic neurons of substantia nigra pars compacta, with a consequent dopamine loss in the striatum (caudate and putamen).[1]

Diagnosis of PD is currently assessed by the clinical evaluation of motor symptoms that appear when the degeneration of dopaminergic nigral neurons has raised over 70%.[2,3] As a consequence, therapeutic interventions may only reduce the symptoms and cannot stop or possibly revert the disease progression. Nonmotor symptoms, such as impaired olfaction, disordered sleep, and constipation, frequently precede the onset of PD, but they are unspecific predictive symptoms; instrumental investigations such as polysomnography or functional imaging are characterized by high costs and the use of radioactive tracers that hamper their application population-wide.[4−6] For these reasons, the identification of specific biomarkers would be critical for the early diagnosis of PD and also to monitor the progression of the disease. Additionally, the assessment of disease-modifying drugs requires the identification of early stage patients to be included in clinical studies.[4−8]

We recently reported a panel of 13 proteins as a discriminant signature for Parkinson's disease.[9,10] The discovery phase was performed by two-dimensional electrophoresis (2-DE) of proteins extracted from T-lymphocytes of 32 subjects, including PD patients, healthy subjects, and patients with atypical parkinsonisms. This signature appeared to be promising in terms of sensitivity and specificity by the leave-one-out cross-validation. However, a further verification with a higher-throughput approach was necessary to translate the proposed assay into the clinical practice.[11]

Received: November 19, 2013
Published: June 19, 2014

The aim of the present investigation was to verify that a molecular signature obtained by a discovery procedure based on a technique other than multiple reaction monitoring (MRM) retains its discriminant power when a new cohort is enrolled and classified by MRM. We enrolled therefore a new cohort of nine PD patients and nine control subjects to develop a method for the quantification of the protein signature by independent MRM measurement and to evaluate the discriminant power of the method when proteins are measured by an alternative technology instead of that used in the discovery phase (i.e., 2-DE).

EXPERIMENTAL SECTION

Subjects

Eighteen subjects were enrolled by the Parkinson's Disease Center at the Department of Neuroscience, University of Torino. Every subject was associated with an alphanumeric code to ensure that his/her identity was not disclosed to investigators. Among them, nine subjects were PD patients, varied in terms of age, age at onset, pharmacological treatment, and familiarity, and nine subjects were clinically classified as non-PD subjects. Supporting Information Table 1 reports demographic and clinical data for all enrolled subjects. Gender and age distribution was similar in different groups. Absolute inclusion criteria for PD patients were as follows: idiopathic PD, absence of atypical signs, and a good response to L-DOPA. Supportive criteria were the following: asymmetry of symptoms or signs at onset, clinical course of more than five years without atypical signs, L-DOPA induced motor fluctuations, or dyskinesias. Exclusion criteria for patients and control subjects were as follows: use of neuroleptic drugs, focal cerebral lesions, and a history of encephalitis. Subjects suffering from inflammatory or infectious diseases and subjects that took drugs capable of interfering with T-lymphocytes at the time of enrollment were excluded from the study.[9] All patients signed an informed consent before being recruited for the present study, according to the guidelines of the Institutional Review Board of the University of Torino.

Isolation of PBMC and T-Lymphocytes

All subjects underwent a venous blood sampling (20 mL) from the antecubital vein, between 9 and 10 a.m., after an overnight fast. Whole blood was collected into vacuum tubes containing EDTA, diluted with 50 mL of phosphate-buffered saline (PBS), and stratified in two 50 mL tubes on top of 15 mL of Lympholyte-H (Cedarlane, Burlington, Canada) each. After centrifugation (800*g*, 20 min, 20 °C), peripheral blood mononuclear cells (PBMC) were removed and pelleted through centrifugation (400*g*, 15 min, 20 °C). Pellets were split, and one fraction stored at −80 °C. The second fraction was washed twice with 10 mL of magnetic-activated cell sorting (MACS) buffer (Miltenyi Biotec, Bergisch Gladbach, Germany). The isolation of T-lymphocytes was achieved by MACS with the Pan T-cell isolation kit using the manufacturer's protocol.[9]

Digestion Protocol

Cell pellets were resuspended in a solution of PBS and 1% RapiGest (Waters Corporation, Milford, MA) (1:1) and put in a ultrasonic bath for 15 min. Samples were then centrifuged (13 000*g*, 15 min, 4 °C) to pellet cell debris, and supernatant transferred to a new tube. Then 200 μL of 100 mM NH_4HCO_3 was added, and samples put at 80 °C for 45 min. Each sample was reduced and alkylated with 20 μL of 5 mM dithiothreitol (30 min, 60 °C) and 10 μL of 10 mM iodoacetamide (30 min, RT, in the dark). Afterward, 1:50 (w/w) trypsin (gold mass spectrometry grade, Promega, Madison, WI) was added to each sample and digestion allowed to proceed at 37 °C overnight. Eventually, samples were treated with CF_3COOH (0.5% v/v), at 37 °C for 45 min, and peptides collected in the supernatant. Prior to MRM measurement, the samples were diluted with an aqueous 0.1% formic acid solution and spiked with predigested chaperone protein ClpB from *Escherichia coli* (CLPB) standard (Waters Corporation) for single standard relative peptide quantification,[12] providing 100 ng of sample and 5 fmol of CLPB on-column amounts for MRM analysis by means of LC−MS, respectively.

Liquid Chromatography−Mass Spectrometry

All samples were analyzed by LC−MS using a nanoAcquity system (Waters Corporation). The samples (100 ng on-column protein digest equivalents) were injected onto a Symmetry C18, 180 μm × 20 mm trap column (Waters Corporation), using partial loop injection, for 1 min at a flow rate of 15 μL/min with 0.1% (v/v) formic acid. The sample was resolved on a BEH C18 75 μm × 150 mm 1.7 μm analytical column (Waters Corporation) using a gradient of 97% A (0.1% (v/v) formic acid), 3% B (99.9% acetonitrile 0.1% (v/v) formic acid) to 60% A, 40% B over 45 min at a flow rate of 300 nL/min. The LC system was coupled to a Xevo TQ-S tandem quadrupole mass spectrometer (Waters Corporation) operated in time-scheduled MRM mode with both quadrupole analyzers operating at unit resolution. The capillary voltage, cone voltage and source temperature were maintained at 3.50 kV, 35 V and 70 °C, respectively. The N_2 cone gas flow, nanoflow gas pressure, and nebulizer gas pressure were 35 L/h, 0.25, and 5 bar, respectively, whereas the Ar collision gas flow equaled 0.17 mL/min. Each peptide was targeted by at least two transitions with a minimum of 15 data points over a 15 s chromatographic peak (width at 10% of the peak height). Collision energies were calculated using the following regression equations: 0.034 times m/z (mass over charge) + 3.314 eV for double and 0.044 times m/z + 3.314 eV for triply charged peptides. Three technical LC−MS replicates per sample were acquired in randomized experiment order.

Selection of MRM Transitions

For proteins belonging to a signature able to discriminate PD patients from control subject,[9] MRM transitions were examined using Skyline[13] excluding C, M, and Q (N-terminal) containing peptides (Supporting Information Table 2). Missed cleaved and modified peptides were not allowed, and the minimum and maximum amino acid lengths were 8 and 25 amino acids, respectively. Both b and y product ions were considered with a maximum of six transitions per peptide for initial identification, and a lower limit m/z transition selection range of two product ions below precursor m/z and an upper limit of $n - 1$. VerifyE (Waters Corporation) was employed to exclude nonproteotypic, interfered transitions using 2 Thompson precursor and product ion tolerances and a 2 min retention time window from the data and information obtained from previous discovery experiments (Supporting Information Table 2). Basic local alignment searches[14] were conducted with the remaining peptides to confirm that the sequences were unique to the proteins of interest. A pooled sample comprising equal amounts of all patient and control subject samples was used to examine the usability of the final transitions in terms of response and

dx.doi.org/10.1021/pr401142p | J. Proteome Res. 2014, 13, 3554−3561

Table 1. Peptide (m/z) and MRM Transition (m/z) Overview, Transition Abundance and Precision, and Feature Scores

gene name	peptide m/z (charge)	MRM transitions	relative abundance transitions (standard deviation)[a]	coelution score[ab]	shape score[ac]
FGB	IRPFFPQQ	b5, b7, y3	0.56 (0.02), 0.18 (0.02), 0.26 (0.02)	1.320	0.956
	516.8 (2^+)	661.4, 886.5, 372.2			
	DNENVVNEYSSELEK	y9, y10	0.46 (0.06), 0.54 (0.06)	1.035	0.948
	884.9 (2^+)	1098.5, 1197.6			
	EEAPSLRPAPPPISGGGYR	y16, y17	0.68 (0.15), 0.32 (0.15)	1.483	0.878
	651.0 (3^+)	811.4, 847.0			
FLNA	YGGQPVPNFPSK	b4, y6, y8	0.13 (0.01), 0.48 (0.02), 0.39 (0.02)	1.520	0.922
	645.8 (2^+)	406.2, 689.4, 885.5			
	VTAQGPGLEPSGNIANK	y8, y13	0.73 (0.03), 0.27 (0.03)	1	0.978
	826.9 (2^+)	800.4, 1253.6			
GSN	YIETDPANR	y4, y7	0.70 (0.02), 0.30 (0.02)	1	0.996
	539.8 (2^+)	457.3, 802.4			
	HVVPNEVVVQR	b3, y8, y9	0.13 (0.03), 0.53 (0.04), 0.34 (0.04)	1.013	0.961
	638.4 (2^+)	336.2, 940.5, 1039.6			
	TPSAAYLWVGTGASEAEK	y10, y11	0.44 (0.06), 0.56 (0.06)	1.021	0.927
	919.5 (2^+)	948.5, 1134.5			
MSN	EDAVLEYLK	y4, y5	0.25 (0.02), 0.75 (0.02)	1	0.988
	540.3 (2^+)	552.3, 665.4			
	AQQELEEQTR	y6, y7, y8	0.31 (0.03), 0.28 (0.03), 0.41 (0.03)	1	0.984
	616.3 (2^+)	775.4, 904.4, 1032.5			
ARHGDIB	LNYKPPPQK	y4, y5, y7	0.23 (0.04), 0.70 (0.04), 0.07 (0.02)	1.907	0.865
	542.8 (2^+)	469.3, 566.3, 857.5			
	TLLGDGPVVTDPK	y10, y11	0.59 (0.03), 0.41 (0.03)	1	0.980
	656.4 (2^+)	984.5, 1097.6			
SEPT6	TVPLAGHVGFDSLPDQLVNK	y7, y18	0.78 (0.10), 0.22 (0.10)	1.139	0.884
	703.0 (3^+)	813.4, 954.0			
	SLDDEVNAFK	y7, y8	0.19 (0.12), 0.81 (0.12)	1.056	0.922
	569.3 (2^+)	822.4, 937.4			
TWF2	DDLSFAGYQK	y6, y7	0.34 (0.12), 0.66 (0.12)	1.132	0.901
	572.3 (2^+)	713.4, 800.4			
	HQTLQGLAFPLQPEAQR	b8, b9	0.63 (0.17), 0.37 (0.17)	2.189	0.801
	645.3 (3^+)	849.5, 996.5			
LSP1	EGPGPEDTVQDNLGAAGAEEEQEEHQK	y23, y25	0.14 (0.12), 0.86 (0.12)	2.340	0.731
	955.4 (3^+)	1262.6, 893.4			
VIM	ILLAELEQLK	b2, y7, y8	0.46 (0.02), 0.22 (0.01), 0.32 (0.02)	1	0.996
	585.4 (2^+)	227.2, 830.5, 943.5			
	NLQEAEEWYK	y6, y7, y8	0.33 (0.02), 0.41 (0.02), 0.26 (0.02)	1.046	0.971
	655.3 (2^+)	825.4, 954.4, 1082.5			
	EEAENTLQSFR	y4, y7, y9	0.36 (0.02), 0.38 (0.01), 0.26 (0.01)	1	0.989
	662.3 (2^+)	537.3, 865.5, 1065.5			
VCL	ELTPQVVSAAR	y6, y8, y9	0.32 (0.02), 0.61 (0.02), 0.07 (0.01)	1.013	0.964
	585.8 (2^+)	602.4, 827.5, 464.8			
	DPSASPGDAGEQAIR	b5, y10, y11	0.19 (0.03), 0.60 (0.04), 0.21 (0.03)	1.027	0.965
	735.8 (2^+)	458.2, 1013.5, 1100.5			
TALDO1	LSSTWEGIQAGK	y7, y10	0.66 (0.05), 0.34 (0.05)	1.007	0.957
	638.8 (2^+)	702.4, 1076.5			
	LLGELLQDNAK	y6, y7, y9	0.23 (0.01), 0.18 (0.01), 0.59 (0.02)	1	0.986
	607.3 (2^+)	688.4, 801.4, 987.5			
	SYEPLEDPGVK	y4, y8, y9	0.68 (0.02), 0.29 (0.02), 0.03 (0.01)	1	0.978
	617.3 (2^+)	400.3, 854.5, 983.5			
YWHAE	HLIPAANTGESK	b3, y9, y10	0.21 (0.06), 0.57 (0.08), 0.22 (0.06)	1.116	0.936
	619.3 (2^+)	364.2, 874.4, 987.5			
	YLAEFATGNDR	y6, y7, y9	0.30 (0.22), 0.26 (0.13), 0.44 (0.18)	1.836	0.858
	628.8 (2^+)	633.3, 780.4, 980.4			
TLN1	VLVQNAAGSQEK	y8, y9, y10	0.46 (0.02), 0.22 (0.03), 0.32 (0.03)	1	0.986
	622.3 (2^+)	804.4, 932.4, 1031.5			
	NGNLPEFGDAISTASK	b4, y12	0.31 (0.04), 0.69 (0.04)	1.042	0.946
	810.9 (2^+)	399.2, 1222.6			
ACTB	SYELPDGQVITIGNER	y12, y13, y12	0.69 (0.02), 0.10 (0.01), 0.21 (0.01)	1	0.935
	895.9 (2+)	1298.7, 706.4, 649.8			

dx.doi.org/10.1021/pr401142p | J. Proteome Res. 2014, 13, 3554−3561

Table 1. continued

gene name	peptide m/z (charge)	MRM transitions	relative abundance transitions (standard deviation)[a]	coelution score[ab]	shape score[ac]
	KDLYANTVLSGGTTMYPGIADR	y8, y9	0.35 (0.06), 0.65 (0.06)	1	0.992
	781.1 (2+)	922.4, 1023.5			
CLPB	VTDAEIAEVLAR	y9, y10, y11	0.10 (0.02), 0.62 (0.03), 0.28 (0.03)	1.805	0.908
	643.9 (2+)	971.5, 1086.6, 1187.6			

[a]Cohort average, n = 138 (3 technical replicates for 18 T-cell samples, 3 technical replicates for 18 PBMC samples, 12 matrix blanks, and 18 calibrations runs). [b]Coelution score = mean standard deviation between MRM transition peaks (1, 2, 3, ...). [c]Shape score = shape similarity between MRM transition peaks (0−1).

Table 2. Fold Change and Significance for Selected Peptides Used for Classification Following MRM Quantification

gene name	protein name	2-DE change	peptide	fold change	p[a]
FGB	fibrinogen beta chain	↓ PD	IRPFFPQQ	1.024	0.258
			DNENVVNEYSSELEK	0.878	0.796
			EEAPSLRPAPPPISGGGYR	1.289	0.340
FLNA	filamin-a	↓ PD	YGGQPVPNFPSK	1.091	0.340
			VTAQGPGLEPSGNIANK	1.070	0.297
GSN	gelsolin	↓ PD	YIETDPANR	1.202	0.007
			HVVPNEVVVQR	1.374	0.386
			TPSAAYLWVGTGASEAEK	1.252	0.258
MSN	moesin	↑ PD	EDAVLEYLK	1.271	0.113
			AQQELEEQTR	1.229	0.031
ARHGDIB	rho GDP-dissociation inhibitor 2	↑ PD	LNYKPPPQK	1.231	0.258
			TLLGDGPVVTDPK	1.102	0.258
SEPT6	septin-6	↑ PD	TVPLAGHVGFDSLPDQLVNK	1.310	0.050
			SLDDEVNAFK	1.069	0.297
TWF2	twinfilin-2	↑ PD	DDLSFAGYQK	1.065	0.297
			HQTLQGLAFPLQPEAQR	1.299	0.040
LSP1	lymphocyte-specific protein 1	↑ PD	EGPGPEDTVQDNLGAAGAEEEQEEHQK	1.159	0.024
VIM	vimentin	↑ PD	ILLAELEQLK	1.099	0.258
			NLQEAEEWYK	1.002	0.297
			EEAENTLQSFR	1.257	0.031
VCL	vinculin	↓ PD	ELTPQVVSAAR	1.135	0.136
			DPSASPGDAGEQAIR	1.358	0.545
TALDO1	transaldolase	↑ PD	LSSTWEGIQAGK	1.329	0.018
			LLGELLQDNAK	1.358	0.024
			SYEPLEDPGVK	1.282	0.024
YWHAE	14-3-3 protein epsilon	↓ PD	HLIPAANTGESK	1.166	0.545
			YLAEFATGNDR	0.801	0.666
TLN1	talin-1	↓ PD	VLVQNAAGSQEK	1.056	0.297
			NGNLPEFGDAISTASK	1.026	0.796

[a]Wilcoxon test after Log2 transformation. Benjamini-Hochberg corrected.

interference. The MRM transitions were manually inspected and curated using TargetLynx (Waters Corporation) and analyzed and scored with mProphet.[15] In summary, at least six identifiers per protein were required (Supporting Information Table 2) and at least two of them were selected for quantification. A total of 81 MRM transitions (73 analyte transitions, 5 housekeeping (ACTB) transitions, and 3 VTDAEIAEVLAR (CLPB) transitions for relative quantification) were monitored during an individual sample analysis (Table 1). A graphical representation of the selection process and validation of the MRM transitions is shown in Supporting Information Figure 1.

Peptide Quantification

Twenty-nine peptides from the panel, two ACTB peptides as housekeeping (β-actin), and one CLPB peptide for internal calibration were quantified (Table 1). All peak area integration, regression analysis, and sample quantification were performed using TargetLynx. Relative in-run quantitation was achieved using the exogenous CLPB peptide mixture spike, applying linear regression analysis with $1/x$ weighting. At least two noninterfered, high precision MRM transitions per peptide were retained for final quantification. As a whole, 138 LC-MS quantification experiments were performed (i.e., three technical replicates for each of the 18 T-cell samples, three technical replicates for each of the 18 PBMC samples, 12 matrix blanks, and 18 calibrations runs). A summary of the MRM chromatograms is provided in Supporting Information Figure 2.

Statistical Analysis

After in-run calibration, MRM amount of each peptide considered was divided by the amount of the ACTB housekeeping peptide (SYELPDGQVITIGNER) of the same sample to correct loading differences. Moreover, a second peptide from ACTB, KDLYANTVLSGGTTMYPGIADR, was used as control. Indeed, it did not display significant differences

between the two groups (p = 0.73). To obtain fold change values, each peptide amount (already normalized to ACTB) was divided by the amount of the corresponding peptide in a control subject (TO107). This equalization procedure was iteratively performed against all subjects to evaluate a possible bias. The transformation to fold change values was necessary to make MRM measurements comparable with 2-DE quantifications.

Relative changes were analyzed after logarithmic transformation using the Wilcoxon test and p-values corrected according to Benjamini–Hochberg for multiple comparisons. A threshold of false discovery rate (FDR) smaller than 0.05 (5%) was considered as significant. Ten peptides, that is, those showing FDR < 5%, were selected for classification. Alternatively, all peptides of proteins having at least one peptide with FDR < 5% were selected for classification, thus including 16 peptides for 7 proteins of the signature (i.e., gelsolin, moesin, septin-6, twinfilin-2, lymphocyte-specific protein 1, vimentin, transaldolase). As a consequence, peptides from fibrinogen beta chain, filamin-A, Rho GDP-dissociation inhibitor 2, vinculin, 14-3-3 protein epsilon, and talin-1 were not taken into account here as they do not display sufficient discriminant power as taken stand alone.

Spot volumes corresponding to the same protein in 2-DE maps images of the discovery phase[9] were merged and all spot volumes normalized with respect to the median of the corresponding spot volume in the control group, so to obtain fold change values comparable to those from MRM measurements.[9,16] Coefficients for the classifying function were obtained by linear discriminant analysis of fold change values for the 13 proteins of the signature.

Classification of enrolled subjects was achieved by multiplying each coefficient by the corresponding relative MRM amount, so to obtain a PD likelihood score (PD score).[16] Scores were used as predictions to build a ROC curve. Subjects with a PD score greater than a threshold corresponding to the highest difference between false positive rate and true positive rate were classified as PD affected.

RESULTS

A preliminary pilot study to evaluate the in-solution digestion protocol generated all the peptides from the protein panel that was identified in the discovery phase.[9] Subsequently, all the candidates for all the peptides were imported into Skyline. Using this candidate selection process, six of the "best" transitions/peptide were retained and the transitions validated with a reference sample. Moreover, VerifyE was used to provide additional filtering with the removal of interferences. In this way, at least two quantifiers were retained for each protein with an adequate standard deviation (Supporting Information Table 2).

We obtained protein extracts from both T-cells and PBMC of the enrolled subjects (Supporting Information Table 1). The relative abundance, standard deviation, shape score, and coelution score of the MRM transitions of the analyzed peptides are summarized in Table 1. All proteins are referred to with their NCBI gene name in the text, with the corresponding official Uniprot protein name in Table 2. Selected peptides and transitions were used to quantify the proteins of the signature.

Relative quantification values for each peptide were normalized against ACTB, and all quantities from the same sample were equalized against a reference sample, to obtain fold change values comparable to those obtained with 2-DE in our previous report.[9] Briefly, the data set used for the discovery study was used to build a classification function that takes into account a signature of 13 proteins. Next, we measured the level of these proteins by MRM and classified the subjects according to them. Eventually, we refined the performance of the classification function by excluding those proteins that did not show significant changes in total level between the groups. A graphical representation of the workflow is summarized in Figure 1.We proceeded hereafter to identify the best performing peptides by ranked-sum Wilcoxon test with Benjamini–Hochberg correction for multiple comparisons. Nine peptides representative of seven proteins (GSN, MSN, SEPT6, TWF2, LSP1, VIM, TALDO1) were selected for classification. Eight peptides displayed a fold change concordant with the 2-DE discovery study (Table 2). Quantification of six proteins was not considered (FGB, FLNA, ARHGDIB, VCL, YWHAE, TLN1). This feature selection reflects the characteristics of a 2-DE experiment in that it can account for post-translational changes but to a lesser degree for total amount, which is typically measured by MRM-based protocols and immunochemical assays. Alternatively, classification of enrolled subjects was performed by taking into account all peptides for proteins showing at least one peptide with a significant fold change.

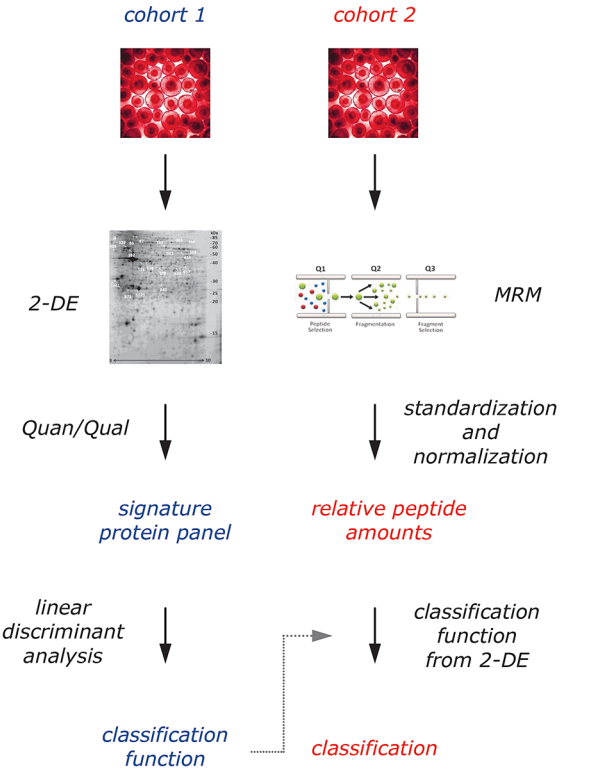


Figure 1. Protein panel verification workflow illustrating the discovery driven determination of a classification function (left) and the application of the function (right) to the analysis of a second Parkinson's disease cohort using MRM LC-MS technology.

To build a discriminant function to be applied to relative quantity changes, protein spots quantified in the previous work[9] and corresponding to the same protein were averaged so to obtain 13 features corresponding to all the proteins listed above. Each spot volume was normalized with respect to the median of the control group. Coefficients of the classification

dx.doi.org/10.1021/pr401142p | J. Proteome Res. 2014, 13, 3554–3561

function were obtained by linear discriminant analysis and are reported in Table 3.

Table 3. Linear Coefficients for the Classification of the Enrolled Subjects

gene ID	coefficient
FGB	−0.11442
FLNA	−1.22117
GSN	0.00921
MSN	0.02901
ARHGDIB	0.42035
SEPT6	0.47243
TWF2	0.13067
LSP1	1.03916
VIM	−0.38459
VCL	−1.31877
TALDO1	0.65897
YWHAE	−0.50467
TLN1	−0.03198

We classified enrolled subjects, either by including all features or by limiting to those nine showing a significant discriminant power, with FDR < 5%. Classification based on all features did not provide a satisfactory result (Figure 2, top panels), whereas the restriction of peptides to the described subset allowed us to classify correctly 16 subjects out of 18. Figure 2 (middle panels) shows the distribution of PD likelihood scores and its receiver operating characteristic (ROC) curve, with an area under the curve (AUC) of 0.877. By taking into account all the peptides for the mentioned proteins, thus including nonsignificant features, we obtained also a satisfactory confusion table. In this case, the model assigned correctly seven controls out of nine, and eight PD patients out of nine, with an AUC equal to 0.840 (Figure 2, bottom panels). In parallel, we classified enrolled subjects using protein features in PBMC, which are easier and faster to prepare from venous blood. However, the function was not able to correctly classify the subjects (Supporting Information Figure 3). Interestingly, the present finding supports the rationale for the use of T-lymphocytes as a source of biomarkers for PD.[17]

DISCUSSION

The biomarker pipeline is conventionally the result of three consecutive steps: discovery, verification, and validation.[11] The goal of the discovery phase is to identify biomarker protein candidates for a specific clinical need. This is not limited to single biomarkers that significantly classify the enrolled subjects. Rather, the biomarker may be the result of a linear combination of selected features that contribute at best to the correct classification of patients without having each feature a sufficient power.[18] In the subsequent verification and validation phases, these protein candidates are evaluated with regard to their detectability, abundance, and performance characteristics. A discovery study should deal with a small sample of

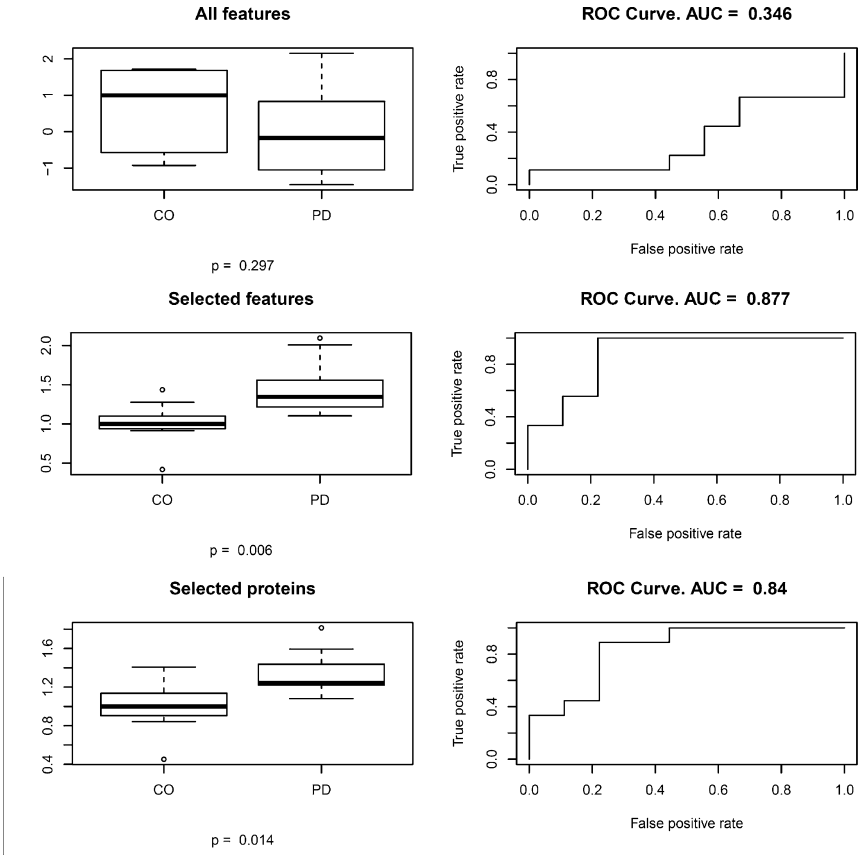


Figure 2. Classification of nine control subjects and nine patients with PD using the discriminant function obtained by 2-DE. Left: Distribution of PD likelihood scores. Right: receiver operating characteristic (ROC) curves. Top: all features were used for classification. Middle: only features with FDR < 5% were used for classification (Table 3). Bottom: all peptides for proteins that showed at least one significant peptide (GSN, MSN, LSP1, SEPT6, TALDO1, TWF2, and VIM) were taken into account. For further details, see text.

dx.doi.org/10.1021/pr401142p | J. Proteome Res. 2014, 13, 3554−3561

homogeneous subjects to be screened by an unbiased technique to generate a quite large number of candidates. Afterward, the verification step is aimed at the cross-validation of the candidate signature (by jackknifing or k-fold cross-validation)[16,18] and at the portability of the proposed assay to a high-throughput methodology suitable for the clinical validation. The first target is usually reached on the same cohort used in the discovery phase, whereas the second aim is fulfilled by enrolling a new small, homogeneous cohort.[11] Remarkably, it has been proposed that discovery should be decoupled from verification and validation in a phased approach, not necessarily performed by the same laboratory.[11]

According to the above-mentioned guidelines, we have enrolled a small cohort of nine PD patients and nine subjects diagnosed as not affected by PD (test set). Although the groups may appear to be relatively small when compared to clinical validation studies, it should be kept in mind that the purpose of the present investigation was to assess that the suggested protein signature discovered by differential 2-DE proteomics is effectively able to classify the test set when applied to protein amounts measured by MRM. Indeed, MRM provides a very fast, accurate, precise, high-throughput procedure to quantify a panel of proteins in a simultaneous way. We iteratively designed MRM transitions starting with in silico prediction, followed by experimental validation and comprehensive MRM testing of all of the candidate peptides. The resultant MRM chromatograms showed good chromatographic performance and separation, as well as adequate MRM quantitation response for the majority of the proteins from the signature panel with a mere 100 ng of material loaded on-column. Throughout the complete experiment, comprising 138 LC−MS runs in total, including blanks and calibration standard runs, the relative areas of the individual transitions were constant, as shown in Supporting Information Figures 1 and 2 and summarized in Table 1.

We processed again the data set used in the discovery step[9] to obtain a discriminant function after merging all 2-DE features (i.e., relative spot volumes) belonging to the same protein and transforming protein levels in fold-changes, so to have observables with comparable dimension with respect to those in the test set. By combining the selected fold-changes using the coefficients listed in Table 3, we obtained good performance values, that is, AUC values of 0.877 and 0.84 for the 10 peptide (selected features) and 16 peptide (selected proteins) panels, respectively (Figure 2). As it can be seen, scores do not significantly overlap when a filtered signature is taken into account, thus showing that MRM quantification of selected peptides provides a robust tool for the classification of subjects based on a protein signature obtained by 2-DE. On the other hand, we are not surprised that certain features obtained from 2-DE, a technique that can discriminate single protein modifications, cannot be applied to total protein amounts such as those obtained by MRM or ELISA, at least in the small cohort used in verification studies. Nevertheless, present findings demonstrate that MRM may be usefully applied in a further clinical validation study using the protein signature discovered by 2-DE.

Levels of some peptides do not correlate with the disease state. Actually, their fold change deviates from one with high *p*-values (0.258, 0.796, 0.340). As a consequence, these peptides were not used anyhow for final classification of the samples. Some type of post-translational modifications (including proteolytic processing) should cause a disagreement between the fold changes observed at the peptide level. As FGB is concerned, multiple spots were identified as fibrinogen in 2-DE maps, supporting the occurrence of post-translational modifications.

The verification procedure failed when the protein signature was applied on whole PBMC fractions, that are easier to obtain. However, different figures might arise from a large scale clinical validation study. Present results demonstrate that the proposed signature might still retain good sensitivity and specificity values when it would be measured by MRM, at least in T-cells. A further independent validation step conducted on a large cohort of highly varied subjects should therefore include both PBMC and T-cell fractions.

In conclusion, MRM quantification of proposed biomarkers in T-lymphocytes, but not necessarily in the whole PBMC fraction, may permit the classification of PD patients and it should be of great interest to validate the statistical power of the present method in a large-scale clinical study. In particular, the present verification study should reduce the risk associated with a clinical validation procedure, which is usually not performed in the absence of such supporting data, also by considering the high cost/high risk factor.[5]

■ ASSOCIATED CONTENT

Supporting Information

Tables including demographic and clinical data for the subjects enrolled in the study, and list of candidate peptides for each protein of the signature and the housekeeping *β*-actin; figures showing MRM quantification process illustrating the transition validation of an internal standard and target Parkinson's disease peptide, MRM chromatograms/transitions for the quantified peptides of the Parkinson's disease signature, and classification of nine control subjects and nine patients with PD based on the protein signature measured in PBMC by MRM using the discriminant function obtained by 2-DE. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: tiziana.alberio@uninsubria.it. Telephone: +390331339485.

Notes

The authors declare the following competing financial interest(s): TA, MF and LL are listed as inventors in the patent WO/2013/098786 entitled "Method for the in vitro diagnosis of Parkinson's disease".

■ ACKNOWLEDGMENTS

The authors deeply appreciate those who have donated their blood for the present study. G.L. is supported by BBSRC (Grants BB/G009058/1 and BB/I000631/1).

■ ABBREVIATIONS

ARHGDIB, Rho GDP-dissociation inhibitor 2; AUC, area under the curve; CLPB, predigested chaperone protein ClpB from *Escherichia coli*; FDR, false discovery rate; FGB, fibrinogen beta chain; FLNA, filamin-A; GSN, gelsolin; LDA, linear discriminant analysis; LSP1, lymphocyte-specific protein 1; MRM, multiple reaction monitoring; MSN, moesin; PBL, peripheral blood lymphocytes; PBMC, peripheral blood mononuclear cells; PD, Parkinson's disease; ROC, receiver

operating characteristic; SEPT6, septin-6; TALDO1, transaldolase; TLN1, talin-1; TWF2, twinfilin-2; VCL, vinculin; VIM, vimentin; YWHAE, 14-3-3 protein epsilon

REFERENCES

(1) Sharma, S.; Moon, C. S.; Khogali, A.; Haidous, A.; Chabenne, A.; Ojo, C.; Jelebinkov, M.; Kurdi, Y.; Ebadi, M. Biomarkers in Parkinson's disease (recent update). *Neurochem. Int.* **2013**, *63* (3), 201−229.

(2) Jankovic, J. Parkinson's disease: clinical features and diagnosis. *J. Neurol. Neurosurg. Psychiatry* **2008**, *79* (4), 368−376.

(3) Shulman, J. M.; De Jager, P. L.; Feany, M. B. Parkinson's disease: genetics and pathogenesis. *Annu. Rev. Pathol.* **2011**, *6*, 193−222.

(4) Morgan, J. C.; Mehta, S. H.; Sethi, K. D. Biomarkers in Parkinson's disease. *Curr. Neurol. Neurosci. Rep.* **2010**, *10* (6), 423−430.

(5) Alberio, T.; Fasano, M. Proteomics in Parkinson's disease: An unbiased approach towards peripheral biomarkers and new therapies. *J. Biotechnol.* **2011**, *156* (4), 325−337.

(6) Gerlach, M.; Maetzler, W.; Broich, K.; Hampel, H.; Rems, L.; Reum, T.; Riederer, P.; Stöffler, A.; Streffer, J.; Berg, D. Biomarker candidates of neurodegeneration in Parkinson's disease for the evaluation of disease-modifying therapeutics. *J. Neural Transm.* **2012**, *119* (1), 39−52.

(7) Marek, K.; Jennings, D.; Tamagnan, G.; Seibyl, J. Biomarkers for Parkinson's disease: Tools to assess Parkinson's disease onset and progression. *Ann. Neurol.* **2008**, *64* (S2), S111−S121.

(8) Caudle, W. M.; Bammler, T. K.; Lin, Y.; Pan, S.; Zhang, J. Using 'omics' to define pathogenesis and biomarkers of Parkinson's disease. *Exp. Rev. Neurother.* **2010**, *10* (6), 925−942.

(9) Alberio, T.; Pippione, A. C.; Zibetti, M.; Olgiati, S.; Cecconi, D.; Comi, C.; Lopiano, L.; Fasano, M. Discovery and verification of panels of T-lymphocyte proteins as biomarkers of Parkinson's disease. *Sci. Rep.* **2012**, *2*, 953.

(10) Alberio, T.; Fasano, M.; Lopiano, L. Method for the in vitro diagnosis of Parkinson's disease. WIPO Patent No. 2013098786; World Intellectual Property Organization: Geneva, Switzerland, 2013.

(11) Surinova, S.; Schiess, R.; Hüttenhain, R.; Cerciello, F.; Wollscheid, B.; Aebersold, R. On the development of plasma protein biomarkers. *J. Proteome Res.* **2010**, *10* (1), 5−16.

(12) Hoofnagle, A. N.; Becker, J. O.; Oda, M. N.; Cavigiolio, G.; Mayer, P.; Vaisar, T. Multiple-Reaction Monitoring−Mass Spectrometric assays can accurately measure the relative protein abundance in complex mixtures. *Clin. Chem.* **2012**, *58* (4), 777−781.

(13) MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **2010**, *26* (7), 966−968.

(14) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215* (3), 403−410.

(15) Reiter, L.; Rinner, O.; Picotti, P.; Hüttenhain, R.; Beck, M.; Brusniak, M. Y.; Hengartner, M. O.; Aebersold, R. mProphet: automated data processing and statistical validation for large-scale SRM experiments. *Nat. Methods* **2011**, *8* (5), 430−435.

(16) Fasano, M.; Alberio, T. An automated procedure for the statistical analysis of two-dimensional electrophoresis gels for biomarkers discovery. *Protoc. Exch.* **2013**, DOI: 10.1038/protex.2013.001.

(17) Fasano, M.; Alberio, T.; Lopiano, L. Peripheral biomarkers of Parkinson's disease as early reporters of central neurodegeneration. *Biomarkers Med.* **2008**, *2* (5), 465−478.

(18) Squillario, M.; Barla, A. A computational procedure for functional characterization of potential marker genes from molecular data: Alzheimer's as a case study. *BMC Med. Genomics* **2011**, *4*, 55 DOI: 10.1186/1755-8794-4-55.