(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 July 2008 (17.07.2008)

PCT

(10) International Publication Number
WO 2008/085024 A1

(51) International Patent Classification:
G01N 33/68 (2006.01) *G01N 33/574* (2006.01)

(21) International Application Number:
PCT/NL2007/050014

(22) International Filing Date: 12 January 2007 (12.01.2007)

(25) Filing Language: Dutch

(26) Publication Language: English

(71) Applicant *(for all designated States except US)*: **ERASMUS UNIVERSITY MEDICAL CENTER ROTTERDAM** [NL/NL]; Dr. Molewaterplein 50, NL-3015 GE Rotterdam (NL).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **LUIDER, Theo Marten** [NL/NL]; c/o Erasmus University Medical Center Rotterdam, Dr. Molewaterplein 50, NL-3015 GE Rotterdam (NL).

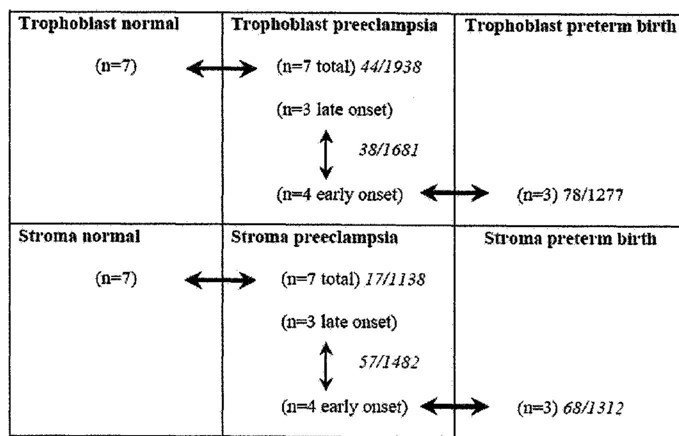
(74) Agent: **VAN LOON, C. J. J.**; Vereenigde, Johan de Wittlaan 7, NL-2517 JR Den Haag (NL).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(54) Title: IDENTIFICATION AND DETECTION OF PEPTIDES RELATING TO SPECIFIC DISORDERS



(57) Abstract: The present invention relates to a method for identification of disease-related peptides and/or proteins for use as markers in diagnosis, prognosis, or therapeutic monitoring of disease, said method comprising the steps of: (a) providing an optionally processed sample of a diseased body tissue or fluid as a test sample, and an optionally processed sample of a corresponding healthy body tissue or fluid as a reference sample, wherein said samples comprise peptides and/or proteins; (b) subjecting both test and reference sample to MALDI- FT-ICR mass spectrometry to generate mass spectra for individual peptides in each sample and to quantify the amount of individual peptides present in each sample; (c) comparing the amount of an individual peptide present in the test sample with the amount of a peptide having a corresponding mass spectrum in the reference sample to generate a list of peptides differentially expressed between test and reference sample, and (d) subjecting the test and/or reference sample of step (a) to tandem mass spectrometry (MS-MS), in order to identify the differentially expressed peptides and/or the proteins from which they derive thus providing a candidate marker protein or marker peptide.

WO 2008/085024 A1

Title: Identification and detection of peptides relating to specific disorders

FIELD OF THE INVENTION

The present invention is in the field of disease diagnostics. In particular, the invention relates to methods for identification of disease-related peptides and/or proteins for use as markers in diagnosis, prognosis, or
5 therapeutic monitoring of disease. The invention further relates to markers identified by the method of the invention and to methods for detection of early onset preeclampsia, glioma and/or multiple sclerosis. The invention further provides the use of choriomammotropin precursor, calcyclin and/or surfeit locus protein 4 for the diagnosis of preeclampsia; for the use of fibronectin,
10 fibrinogen, colligin 2 and/or acidic calponin 3 for the diagnosis of glioma; and for the use of chromogranin A, clusterin, complement C3, complement C4B, beta V spectrin, hypothetical protein XP_011125, apolipoprotein D, complement C4A, contactin 1, neuronal pentraxin receptor and/or RNA binding motif protein 7 for the diagnosis of multiple sclerosis.
15

BACKGROUND OF THE INVENTION

Gliomas are the most common primary brain tumors. The diagnosis of these tumors and the decisions regarding therapy is based almost exclusively on the tissue histopathology. Diffuse gliomas are highly infiltrative
20 and heterogeneous. Gliomas are among neoplasms with highest degree of vascularisation. The growth of gliomas largely depends on their blood supply. The elimination of the blood supply would result in the destruction of these tumors. Despite the elucidation of many genetic aberrations of gliomas over the last decades, only few useful biomarkers or therapeutic targets have been
25 identified so far. Despite the gradual unravelling of the roles of a large number of regulatory proteins in the process of tumor neovascularisation, no major steps forward in antiangiogenic therapies for gliomas have been recorded to

date. The identification of more tumor vasculature-related proteins may result in the finding of new targets of anti-angiogenic therapies and understanding of the formation of neovasculature in glioma.

Preeclampsia is a pregnancy specific syndrome that is diagnosed by the new appearance of increased blood pressure and proteinuria. It is a leading cause of maternal mortality in developed countries and increases perinatal mortality up to five-fold. Since its etiology is largly unknown, a panoply of pathophysiological abnormalities are described. However, it is evident that abnormal placentation plays an important role. In normal pregnancy spiral arteries undergo striking remodeling. They change from typical muscular arteries to flaccid tubes with no muscularis or elastic lamina with a diameter at least four times greater than that of non-pregnant vessels. Shortly after the invasion of trophoblasts into the superficial endometrium, the maternal erythrocytes can be observed within the precursors of the placental intervillous space. In women with preeclampsia endovascular remodelling and invasion of the spiral arteries is less prominent which is assumed to result in overwhelming placental oxidative stress and pregnancy failure. However, there are at present no diagnostic markers for detecting early onset preeclampsia in pragnant women.

Multiple sclerosis (MScl) is a complex disease of the central nervous system (CNS) with presumed autoimmune origin. Its pathology is characterized by a combination of inflammation, demyelination and axonal damage. These processes are not uniformly represented across patient populations but can predominate selectively in individual patients, contributing to the heterogeneity of phenotypic expression of the disease.

Due to the complex nature of the disease, there probably is no single protein or peptide that can serve as a biomarker for MScl in a clinically relevant way. However, individuals with MScl may possibly be differentiated from healthy individuals by a proteomic pattern consisting of a set of distinct individual proteins or peptides that are not independently useful. Because the

disease process in MScl is located in the CNS, cerebrospinal fluid (CSF) is a promising body fluid in which to search for biomarkers and disease-associated proteins and peptides. Additionally, the relative lower concentration of total protein in CSF compared to serum makes it an attractive bio-fluid to perform proteomic studies.

Hitherto no peptides and/or proteins have been identified that differ between patients having MScl or an MScl associated disease (i.e. demyelination) and normal non MScl patient groups, though such peptides and/or proteins would prove very valuable in diagnosis of the disease.

In fact, the three disease conditions described above are not unique in there lack of diagnostic markers. A large number of vary diverse disease conditions cannot be properly diagnosed because of a lack of biological diagnostic markers.

Rapid and major developments in proteomic technology and methodology over the last decade has opened a new stage in the identification of proteins. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOFMS) recently became available as a flexible tool in the search for disease markers [Marvin, L.F., M.A. Roberts, and L.B. Fay. Clin Chim Acta, 2003. 337(1-2): p. 11-21]. Moreover, the recently introduced technique of matrix-assisted laser desorption\ionization Fourier transform mass spectrometry (MALDI-FTMS) provides a powerful technique for accurate peptides mass measurements [Schmid, D.G., et al. Biotechnol Bioeng, 2000. 71(2): p. 149-61]. This technique has successfully been used for studies in protein interactions and post-translational modifications of proteins. However, none of these developments have hitherto resulted in the provision of valuable diagnostic markers.

It is an aim of the present invention to provide a method with which markers can be identified for a large number of diseases, thereby providing diagnostic tools for medical and veterinary diagnosis.

SUMMARY OF THE INVENTION

The present inventors have now found that a combination of different MALDI techniques provides for a very powerful method for marker detection.

5 In a first aspect, the present invention provides a method for identification of disease-related peptides and/or proteins for use as markers in diagnosis, prognosis, or therapeutic monitoring of disease, said method comprising the steps of:

(a) providing an optionally processed sample of a diseased body tissue or
10 fluid as a test sample, and an optionally processed sample of a corresponding healthy body tissue or fluid as a reference sample, wherein said samples comprise peptides and/or proteins;

(b) subjecting both test and reference sample to MALDI- FT-ICR mass spectrometry to generate mass spectra for individual peptides in each
15 sample and to quantify the amount of individual peptides present in each sample;

(c) comparing the amount of an individual peptide present in the test sample with the amount of a peptide having a corresponding mass spectrum in the reference sample to generate a list of peptides differentially expressed
20 between test and reference sample, and

(d) subjecting the test and/or reference sample of step (a) to tandem mass spectrometry (MS-MS), in order to identify the differentially expressed peptides and/or the proteins from which they derive thus providing a candidate marker protein or marker peptide.

25 Therapeutic monitoring of disease means monitoring of disease activity and treatment response.

The advantage of the the method of the present invention is that quantitative differences between diseased and healthy samples in the presence of specific peptide signatures can be detected with high accuracy. MALDI-FT-
30 ICR is capable of quantifying a peptide signature in a sample with a coefficient

of variation (CV), in a range of about 10 ± 5% in complex peptide mixture, compared to about 30 ± 10% for MALDI-TOF. However, the problem is that MALDI-FT-ICR is not suitable for identification of the protein or peptide itself, since de method effectively prevents tandem mass spectrometry and therefore
5 the sequencing of individual peptides. The reason for this is that only a single charge is provided to each peptide in a single MS round, and ionization fractionation of the peptide results in uncharged fragments that cannot be detected. The low reproducibility of MALDI-TOF (*i.e.* the high CV), on the other hand, prevents the accurate quantification of peptide levels and
10 therefore the detection of differentially expressed peptides between biological samples by MALDI-TOF. However, MALDI-TOF, in particular in tandem mass spectrometry (MS-MS) mode may suitably be used on a variety of biological samples to deduce the identity (*i.e.* the amino acid sequence) of peptides and proteins present therein, i.e. to provide peptide sequences information that
15 may lead to the identification of proteins from which they derive.

The present inventors have now found that the combined use of these two methods on the same sample provide sufficient quantitative information as well as sufficient information on the identity of the peptide in order to detect candidate marker peptides or proteins. When analyzing the
20 sample with e.g. MALDI-TOF according to methods of the present invention, one may analyze the same sample, or one may provide a body tissue or fluid sample of larger volume, i.e. including tissue elements not originally present in the sample subjected to MALDI-FT such as adjacent tissues, and still provide the necessary information on the identity of as many peptides and proteins as
25 possible that are present in the sample subjected to MALDI-FT. For instance, the peptide signature of tumor vasculature as analysed by MALDI-FT may be matched with the signature of peptides identified by analysing of a larger sized portion of the tumour that includes at least that vasculature.

In order to determine whether the candidate marker peptide or protein is indeed differentially expressed, high resolution MALDI FT may optionally used for confirmation purpose.

The present inventors have further found that the proposed combination of MALDI techniques can very suitably be used in combination with laser microdissection and/or nano-liquid chromatography prior to mass spectrometry. Although primary brain tumors have been subjected to direct-tissue profiling and imaging mass spectrometry techniques [Schwartz, S.A., et al. Cancer Res, 2005. **65**(17): p. 7674-81, Chaurand, P., et al. Toxicol Pathol, 2005. 33(1): p. 92-101], laser microdissection of brain blood vessels has never been used before in proteomic analysis.

It is known in the art of proteome analysis that factors such as sample stability and a low number of measurements per sample can cause difficulties regarding the reproducibility of proteomic profiling studies. Also, it is known that there is low reproducibility of peak height in MALDI-TOF MS. The method of the present invention overcomes these problems in several ways and is less affected by these variations. First, the samples are all handled in a standardized way. Secondly, the sample preparation method is uncomplicated and straightforward. Thirdly, the height of the peaks is not included in the analysis because quantitative measurements of peak heights with MALDI TOF MS are poorly reproducible, with standard deviations up to 30%. In the present method only the absence or presence of the peaks is scored. (see the Examples below for details)

The addition of a technique for pre-fractionation of test samples, such as nano-liquid chromatography prior to mass spectrometry, increases the number of identified proteins significantly.

Preferred embodiments of the method of the present invention include for instance a method further comprising the step of: (e) confirming the presence of said candidate marker protein or marker peptide in said test

sample by MALDI-FT mass spectrometry, MALDI Triple-quad mass spectrometry or an immunoassay, such as ELISA.

Samples used in aspects of the present invention may be unprocessed, or processed samples, meaning that the samples may or may not have been subjected to procedures wherein the biological, physical or chemical composition of the sample is altered. The samples may also be subjected to multiple processing steps.

In an alternative embodiment of a method of the invention, the optionally processed samples are body tissue samples processed by subjecting said samples to laser capture microdissection to provide collections of microdissected cells, said collections preferably amounting to about 200- 3,000 cells. Preferably, said collections of microdissected cells are provided in the form of pooled collections of microdissected cells.

In yet another alternative embodiment of a method of the invention the optionally processed samples are body tissue samples, body fluid samples, or collections of microdissected cells processed by subjection to protein digestion, preferably using trypsin, to provide processed samples comprising peptide fragments from the proteins in said samples. Thus, the method optionally comprises the step of cleaving the proteins in a sample (i.e. polypeptides in general) with a (optionally sequence specific) cleavage reagent to form peptide fragments, optionally followed by deactivating the cleavage reagent. A sequence specific cleavage agent in aspects of the present invention preferably cleaves the polypeptides on the C-terminal side of a lysine residue. The specific cleavage reagent preferably comprises Lys-C or Trypsin. The cleavage reagent is preferably trypsin. Polypeptide cleaving (e.g. trypsin digestion) is performed to provide peptide fragments sufficiently small to be analysed by MALDI analysis. However, some samples may comprise peptide fragments of sufficiently small size to allow direct MALDI analysis. Examples of peptides that can be detected or analyses in unprocessed samples include (neuro)peptides, hormones, etc.

In yet another alternative embodiment of a method of the invention, the method further comprises the step of performing an *in silico* digestion of the candidate marker protein or marker peptide and comparing the theoretical peptides generated by said digestion with the monoisotopic peaks obtained by 5 MALDI analysis of the peptides in step (b) or (d) of the above-described method.

In an alternative embodiment of a method of the invention step (d) comprises the steps of:

- comparing the mass spectrum of a differentially expressed peptide 10 identified by MALDI-FT in step (b) with the mass spectra of the peptides whose sequence is obtained by tandem mass spectrometry in step (d),
- deducing the sequence for the differentially expressed peptide identified in step (b) from the sequence of the corresponding peptide determined in step (d),
15 - matching the sequence deduced for the differentially expressed peptides with a protein database to identify the peptide or protein that is differentially expressed between said test and reference sample, and
- classifying said identified peptide or protein as a tentative marker peptide or protein for said disease;

20 In still a further alternative embodiment of a method of the invention the optionally processed sample is subjected to nano liquid chromatography (nano-LC) fractionation prior to the tandem mass spectrometry in step (d), in order to provide fractions comprising separated peptides or proteins.

25 In principle, any body tissue of a subject may be used in aspects of the invention. Suitably a body tissue is selected from the group consisting tissues of brain, lung, heart, prostate, esophagus, stomach, jejunum, ileum, caecum, colon, gall bladder, bile duct, breast, ovary, testicle, lymph node, thymus, kidney, liver, muscle, nerve, bone, bone marrow, and placenta.

The body fluid analysed in a method of the present invention may suitably be selected from the group consisting of blood, serum, cerebrospinal fluid (CSF), urine, saliva and semen.

Diseases for which marker peptides or proteins may be found by a method of the invention include for instance Alzheimer's disease, Parkinson's disease, frontotemporal dementia, Lewy body disease, corticobasal degeneration, progressive supranuclear palsy, multiple system atrophy, other neurodegenerative processes, multiple sclerosis, acute disseminating encephalomyelitis (ADEM), neuromyelitis optica (Devic's disease), other inflammatory diseases of the central nervous system, Guillain-Barre syndrome, other inflammatory diseases of the peripheral nervous system, preeclampsia, glioma, cancer, and inflammation.

Body fluid samples, when used in methods of the invention, may suitably be provided in sample volumes of between 0.01 and 100 μl. However, it is a particular advantage of the present invention that very small sample volumes will generally suffice. In step (b) therefore, an amount in a range from 0.1-25 μl, preferably in a range from 1-10 μl of optionally processed body fluid is generally subjected to MALDI- FT-ICR mass spectrometry. A suitable sample fluid preferably comprises about 0.05- 5 mg/ml of protein.

In methods of the present invention, the individual peptides subjected to MALDI analysis generally have a molecular mass in a range of 400-20,000, preferably in a range of 800 to 4,000 Da.

In another aspect, the present invention provides a candidate marker protein or marker peptide for a disease identified by the method according to the present invention.

In one embodiment, the marker protein or marker peptide is choriomammotropin precursor, calcyclin and/or surfeit locus protein 4 and the disease is early onset preeclampsia.

In an alternative embodiment the marker protein or marker peptide is fibronectin, fibrinogen, colligin 2 and/or acidic calponin 3, and the disease is glioma and wherein said marker protein.

In still another alternative embodiment the marker protein or marker peptide is chromogranin A, clusterin, complement C3, complement C4B, beta V spectrin, hypothetical protein XP_011125, apolipoprotein D, complement C4A, contactin 1, neuronal pentraxin receptor and/or RNA binding motif protein 7, and the disease is multiple sclerosis.

In another aspect, the present invention provides a method for predicting early onset preeclampsia comprising measuring the expression level of a marker protein selected from the group consisting of choriomammotropin precursor, calcyclin and surfeit locus protein 4 in villous trophoblast from placentas of pregnant female subjects.

Herein below, the terms "patient" and "subject" are used interchangeably to indicate animal subjects, including human and non-human subjects that are in need of disease diagnosis.

In yet another aspect, the present invention provides a method for detecting glioma, comprising measuring the expression level of a marker protein selected from the group consisting of fibronectin, fibrinogen, colligin 2 and acidic calponin 3 in blood, CSF or glioma vasculature samples of patients.

In still a further aspect, the present invention provides a method for detecting multiple sclerosis, comprising measuring the expression level of a marker protein selected from the group consisting of chromogranin A, clusterin, complement C3, complement C4B, beta V spectrin, hypothetical protein XP_011125, apolipoprotein D, complement C4A, contactin 1, neuronal pentraxin receptor and RNA binding motif protein 7 in CSF samples of patients.

In another aspect, the present invention provides a method for monitoring disease activity of glioma and/or the response to a treatment regimen, comprising measuring the expression level of fibronectin, fibrinogen,

colligin 2 and/or acidic calponin 3 in blood, CSF and/or glioma vasculature samples of patients.

In the various methods described in the present invention the step of detecting the marker peptide or marker protein in a sample may suitably be performed by MALDI Triple-quad analysis of proteins and peptides in a diseased tissue sample to quantify said marker protein or marker peptide indicative for a specific disease in suspect diseased tissue samples of subjects.

SHORT DESCRIPTION OF THE DRAWINGS

Fig 1. represents a chart that displays the various comparisons made between trophoblast versus stroma cells in placenta of controls, women with preeclampsia and preterm delivery. Significantly differentially expressed peptides ($p<0.01$) and total number of peptides detected are described in italics, respectively.

Fig 2. provides a gel view presentation of the comparison of controls (n=7) versus cases (preeclampsia) (n=7) as obtained in ClinProTools software (Bruker Daltonics, Germany). Average mass spectra of trophoblast cases (blue) and control trophoblasts (red) were compared (top panel). The bottom panel shows gel views of specific areas of the mass spectrum that display for instance significant differential expressed peptides at 1250 (bottom left) and 1464 (bottom right). The arrows and graphs adjacent the gel views indicate the expression levels of the various samples for 1250 and 1464 respectively, which were primarily observed in controls (A) and not in preeclampsia (B) in the various samples.

Fig 3. depicts a representation of the discriminant analysis of controls (n=7) versus cases (preeclampsia) (n=7). DF1 describes the difference between trophoblasts and stroma (1.22 % variance) and DF2 describes the difference between control and case (7.96 % variance). A total of 255 spectra was included in this analysis. Note that the ratio of the between/within group variance (B/W) is > 1 for the 4 groups considered.

Fig 4. depicts a representation of the frequency of differential expressed peptides between control (n=7) versus case trophoblasts (preeclampsia) (n=7) (A) and stroma cells (B), ($p<0.01$). Peptides that are significantly expressed will have a contribution from both differential
5 intensities but also the frequency of occurrence in the two groups compared. Some peptide peaks are exclusively found in control or case trophoblast and stroma (arrows correspond to peptides 1250 and 1464 in Figure 2).

Fig 5. depicts a representation of the frequency of differential expressed peptides between early onset preeclampsia (n=4) versus late onset
10 preeclampsia (n=3) trophoblasts (A) and stroma cells (B). A series of peptides that were exclusively found in early onset preeclamptics are observed.

Fig 6. provides a gel view presentation of the comparison of early onset preeclampsia versus preterm controls as obtained in ClinProTools software (Bruker Daltonics, Germany). Average mass spectra of early onset
15 preeclampsia trophoblasts (blue) and preterm control trophoblasts (red) were compared (top panel). The bottom panel shows gel views of specific areas of the mass spectrum that display significant differential expressed peptides. The arrows illustrate peptides that were primarily observed in early onset preeclampsia (B) and not in preterm controls (A) (arrows correspond to
20 peptides 1505 and 2691).

Fig 7. illustrates immunohistochemistry of calcyclin in placenta of early onset preeclampsia (n=3) versus preterm and term controls (resp, n = 2 and n = 3). The arrows illustrate trophoblast cells. Preeclamptic trophoblast cells stain heavily with antibodies specific for calcyclin (S100A6) (panel A) in
25 contrast to preterm and term controls (panel B and C, respectively). Some staining is observed in cells within the stroma of term controls (arrow heads; magnification 100X).

Fig 8. depicts a representation of the frequency of differential expressed peptides between early onset preeclampsia (n=4) versus preterm
30 controls (n=3) trophoblasts (A) and stroma cells (B). The arrows illustrate

peptides, as an example, which were exclusively observed in early onset preeclampsia.

Figure 9: hypertrophied vessels in high-grade glioma. The counterstain of a glioma section shows the hypertrophied vessels in the sample (arrows). These vessels were our target to be microdissected.

Figure 10: heat map of unsupervised clustering of the four groups. The figure illustrate close up of an unsupervised clustering dendrogram besed on peptide masses and group of samples on spotfire. The cluster masses are displayed on the x-axis, whereas the y-asis represents the samples ordered by group. Gp. no. 1, glioma blood vessels, gp. no. 2, normal brain blood vessels, gp. no. 3, glioma surrounding tissue, gp. no. 4, normal brain surrounding tissue. Red blocks showed the presence of peptide in the spectrum of the sample. The unsupervised clustering of the samples resulted in clustering eight out of ten glioma blood vessels samples, group no.1. One of the two samples that did not cluster had a poor spectra, this one clustered with the other poor spectrum sample of normal surround tissue at the top of the heat map. The other glioma sample did not cluster. While clustering based on peptide masses showed a specific peptide pattern of glioma blood vessels group. Those peptides appeared exclusively in glioma blood vessels group.

Figure 11: immunohistochemistry for fibronectin in glioma and normal brain samples. A: the strong positive staining of fibronectin protein in the hypertrophied vessels of glioma sample. B: the negative staining of fibronectin protein in normal brain vessels. C: some of the normal brain vessels showed a very faint staining for fibronectin.

Figure 12: immunohistochemistry for colligin 2 protein in glioma and normal brain samples. A: the strong positive staining of colligin 2 protein in the hypertrophied vessels of glioma sample. B: the negative staining of colligin 2 protein in normal brain vessels.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention now provides for the use of combined MALDI mass spectrometry techniques, wherein a sample is first analysed using MALDI-FTICR MS to identify differentially expressed peptides and then by any suitable MALDI technique to identify as many peptides in a sample as possible to ultimately reveal the identity of the differentially expressed peptides. These differentially expressed peptides, or the proteins from which they derive (*i.e.* as a result of polypeptide cleavage), are earmarked as candidate markers for the disease associated with the differential expression of that peptide or protein. Subsequently, the presence or level of expression of the peptide or protein can be analysed in a sample of interest using MALDI MS analysis or immunoassay techniques for diagnostic, prognostic, or therapeutic monitoring purposes.

Mass Spectrometry

The essential components of a mass spectrometer are (in order) an inlet system, an ion source, a mass analyser, an ion detector and a data capture system. Generally, there are preferred inlet systems, ion sources and mass analysers for the purposes of analysing peptides.

Inlet Systems

A variety of mass spectrometry techniques are compatible with separation technologies particularly capillary zone electrophoresis and High Performance Liquid Chromatography (HPLC). The choice of ionisation source is limited to some extent if a separation is required as ionisation techniques such as MALDI which ablate material from a solid surface are less suited to chromatographic separations. For most purposes, it has been very costly to link a chromatographic separation in-line with MALDI mass spectrometric analysis.

Ionisation techniques

For many biological mass spectrometry applications so called 'soft' ionisation techniques are used. These allow large molecules such as proteins and nucleic acids to be ionised essentially intact. The liquid phase techniques allow large biomolecules to enter the mass spectrometer in solutions with mild pH and at low concentrations. A number of techniques are appropriate for use with this invention including but not limited to Electrospray Ionisation Mass Spectrometry (ESI-MS), Fast Atom Bombardment(FAB), Matrix Assisted Laser Desorption Ionisation Mass Spectrometry (MALDI MS) and Atmospheric Pressure Chemical Ionisation Mass Spectrometry (APCI-MS).

Matrix Assisted Laser Desorption Ionisation (MALDI)

MALDI requires that the biomolecule solution be embedded in a large molar excess of a photo-excitable 'matrix'. The application of laser light of the appropriate frequency results in the excitation of the matrix which in turn leads to rapid evaporation of the matrix along with its entrapped biomolecule. Proton transfer from the acidic matrix to the biomolecule gives rise to protonated forms of the biomolecule which can be detected by positive ion mass spectrometry, particularly by Time-Of-Flight (TOF) mass spectrometry. Suitable Matrix materials for MALDI include for instance alpha-Cyano-4-hydroxycinnamic acid (CHCA) and derivatives such as alpha-Cyano-4-hydroxycinnamic acid Diethylamine salt, alpha-Cyano-4-hydroxycinnamic acid Butylamine salt and Sinapic acid (SA); 2-(4-Hydroxyphenylazo) benzoic acid (HABA); 2-Mercapto-benzothiazole; Succinic acid; 2,6-Dihydroxy acetophenone; dihydroxybenzoic acid (DHB); Ferulic acid; Caffeic acid; Glyerol; 4-Nitroaniline.

Negative ion mass spectrometry is also possible by MALDI TOF. This technique imparts a significant quantity of translational energy to ions, but tends not to induce excessive fragmentation despite this. Accelerating voltages can again be used to control fragmentation with this technique

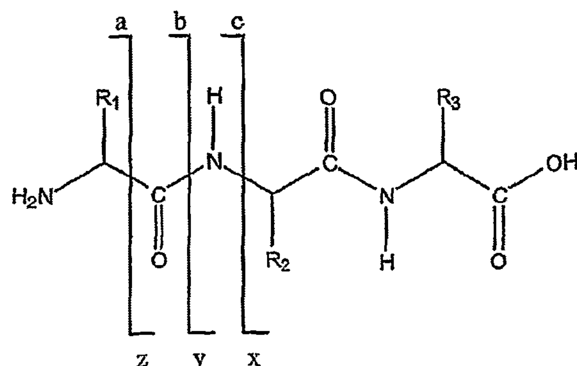
though. This technique is highly favoured for the determination of peptide mass fingerprints due to its large mass range, due to the prevalence of singly charged ions in its spectra and due to the ability to analyse multiple peptides simultaneously.

Mass Analysers

Fragmentation of peptides by collision induced dissociation, to determine their sequence, may be used in this invention to identify proteins, not identified by the pattern of masses of their digestion products. Various mass analyser geometries may be used to fragment peptides and to determine the mass of the fragments.

MS/MS and MS^n analysis of peptides

Tandem mass spectrometers allow ions with a pre-determined mass-to-charge ratio to be selected and fragmented by collision induced dissociation (CID). The fragments can then be detected providing structural information about the selected ion. When peptides are analysed by CID in a tandem mass spectrometer, characteristic cleavage patterns are observed, which allow the sequence of the peptide to be determined. Natural peptides typically fragment randomly at the amide bonds of the peptide backbone to give series of ions that are characteristic of the peptide. CID fragment series are denoted a_n, b_n, c_n, etc. for cleavage at thenth peptide bond where the charge of the ion is retained on the N-terminal fragment of the ion. Similarly, fragment series are denoted x_n y_n, z_n, etc. where the charge is retained on the C-terminal fragment of the ion.



Trypsin, Lys-C and thrombin are favoured cleavage agents for
5 tandem mass spectrometry as they produce peptides with basic groups at both ends of the molecule, i. e. the alpha-amino group at the N-terminus and lysine or arginine side-chains at the C-terminus. This favours the formation of doubly charged ions, in which the charged centres are at opposite termini of the molecule. These doubly charged ions produce both C-terminal and N-terminal
10 ion series after CID. This assists in determining the sequence of the peptide.

Generally speaking only one or two of the possible ion series are observed in the CID spectra of a given peptide. In low-energy collisions typical of quadrupole based instruments the b-series of N-terminal fragments or the y-series of C-terminal fragments predominate. If doubly charged ions are
15 analysed then both series are often detected. In general, the y-series ions predominate over the b-series. In general peptides fragment via a mechanism that involves protonation of the amide backbone follow by intramolecular nucleophilic attack leading to the formation of a 5membered oxazolone structure and cleavage of the amide linkage that was protonated (Schlosser A.
20 and Lehmann W.D. J. Mass Spectrom. 35 :1382-1390, "Five-membered ring formation in unimolecular reactions of peptides: a key structural element controlling low-energy collision induced dissociation", 2000).

A typical tandem mass spectrometer geometry is a triple quadrupole which comprises two quadrupole mass analysers separated by a collision
25 chamber, also a quadrupole. This collision quadrupole acts as an ion guide between the two mass analyser quadrupoles. A gas can be introduced into the collision quadrupole to allow collision with the ion stream from the first mass analyser. The first mass analyser selects ions on the basis of their mass/charge ration which pass through the collision cell where they fragment. The
30 fragment ions are separated and detected in the third quadrupole. Induced

cleavage can be performed in geometries other than tandem analysers. Ion trap mass spectrometers can promote fragmentation through introduction of a gas into the trap itself with which trapped ions will collide. Ion traps generally contain a bath gas, such as helium but addition of neon for example, promotes
5 fragmentation. Similarly photon induced fragmentation could be applied to trapped ions. Another favorable geometry is a Quadrupole/Orthogonal Time of Flight tandem instrument where the high scanning rate of a quadrupole is coupled to the greater sensitivity of a reflectron TOF mass analyser to identify the products of fragmentation.
10 Conventional 'sector' instruments are another common geometry used in tandem mass spectrometry. A sector mass analyser comprises two separate 'sectors', an electric sector which focuses an ion beam leaving a source into a stream of ions with the same kinetic energy using electric fields. The magnetic sector separates the ions on the basis of their mass to generate a
15 spectrum at a detector. For tandem mass spectrometry a two sector mass analyser of this kind can be used where the electric sector provide the first mass analyser stage, the magnetic sector provides the second mass analyser, with a collision cell placed between the two sectors. Two complete sector mass analysers separated by a collision cell can also be used for analysis of mass
20 tagged peptides.

Ion Traps

Ion Trap mass analysers are related to the quadrupole mass analysers. The ion trap generally has a 3 electrode construction-a cylindrical
25 electrode with 'cap' electrodes at each end forming a cavity. A sinusoidal radio frequency potential is applied to the cylindrical electrode while the cap electrodes are biased with DC or AC potentials. Ions injected into the cavity are constrained to a stable circular trajectory by the oscillating electric field of the cylindrical electrode. However, for a given amplitude of the oscillating
30 potential, certain ions will have an unstable trajectory and will be ejected from

the trap. A sample of ions injected into the trap can be sequentially ejected from the trap according to their mass/charge ratio by altering the oscillating radio frequency potential. The ejected ions can then be detected allowing a mass spectrum to be produced.

5 Ion traps are generally operated with a small quantity of a 'bathgas', such as helium, present in the ion trap cavity. This increases both the resolution and the sensitivity of the device as the ions entering the trap are essentially cooled to the ambient temperature of the bath gas through collision with the bath gas. Collisions both increase ionisation when a sample is

10 introduced into the trap and dampen the amplitude and velocity of ion trajectories keeping them nearer the centre of the trap. This means that when the oscillating potential is changed, ions whose trajectories become unstable gain energy more rapidly, relative to the damped circulating ions and exit the trap in a tighter bunch giving a narrower larger peaks.

15 Ion traps can mimic tandem mass spectrometer geometries, in fact they can mimic multiple mass spectrometer geometries allowing complex analyses of trapped ions. A single mass species from a sample can be retained in a trap, i. e. all other species can be ejected and then the retained species can be carefully excited by super-imposing a second oscillating frequency on the

20 first. The excited ions will then collide with the bath gas and will fragment if sufficiently excited. The fragments can then be analysed further. It is possible to retain a fragment ion for further analysis by ejecting other ions and then exciting the fragment ion to fragment. This process can be repeated for as long as sufficient sample exists to permit further analysis. It should be noted that

25 these instruments generally retain a high proportion of fragment ions after induced fragmentation. These instruments and FTICR mass spectrometers (discussed below) represent a form of temporally resolved tandem mass spectrometry rather than spatially resolved tandem mass spectrometry which is found in linear mass spectrometers.

30

FourierTransform Ion Cyclotron Resonance Mass Spectrometry FTICR MS)

FTICR mass spectrometry has similar features to ion traps in that a sample of ions is retained within a cavity but in FTICR MS the ions are trapped in a high vacuum chamber by crossed electric and magnetic fields. The electric field is generated by a pair of plate electrodes that form two sides of a box. The box is contained in the field of a superconducting magnet which in conjunction with the two plates, the trapping plates, constrain injected ions to a circular trajectory between the trapping plates, perpendicular to the applied magnetic field. The ions are excited to larger orbits by applying a radiofrequency pulse to two 'transmitter plates' which form two further opposing sides of the box. The cycloidal motion of the ions generate corresponding electric fields in the remaining two opposing sides of the box which comprise the 'receiver plates'. The excitation pulses excite ions to larger orbits which decay as the coherent motions of the ions is lost through collisions. The corresponding signals detected by the receiver plates are converted to a mass spectrum by Fourier Transform (FT) analysis.

For induced fragmentation experiments these instruments can perform in a similar manner to an ion trap - all ions except a single species of interest can be ejected from the trap. A collision gas can be introduced into the trap and fragmentation can be induced. The fragment ions can be subsequently analysed. Generally fragmentation products and bath gas combine to give poor resolution if analysed by FT analysis of signals detected by the 'receiver plates', however the fragment ions can be ejected from the cavity and analysed in a tandem configuration with a quadrupole, for example.

A typical peptide mass fingerprinting protocol involves determining the mass of the unidentified protein followed by digestion of the protein with trypsin. Trypsin cleaves polypeptides selectively at arginine and lysine residues, leaving either arginine or lysine at the C-termini of the product peptides. The positions of lysine and arginine in the sequence of a polypeptide determine where the polypeptide is cut giving rise to a characteristic series of

peptides. The pattern of peptides can be easily detected by MALDI-TOF mass spectrometry. This mass spectrometric technique has a large mass range, can readily ionise large biomolecules, will preferentially produce singly charged ions.

5 For any given protein, there may be some peptides, which have the same mass as a peptide from another protein but it is very unlikely that two different proteins will give rise to peptides that all have identical masses. This means that the pattern of masses of the tryptic digest of a protein is a fairly unique identifier of that protein and is called a Peptide Mass Fingerprint

10 (PMF). The relative uniqueness of PMFs means that databases of predicted PMFs, determined from known protein sequences or sequences that have been predicted fromgenomic DNA or expressed sequence tags (ESTs), can be used to identify proteins in biological samples (Pappin DJC,Hojrup P and Bleasby AJ, Current Biology 3: 327-332 (1993); Mann M, Hojrup P, Roepstorff P. Biol Mass

15 Spectrom 22 (6): 338-345 (1993); Yates JR 3rd, Speicher S, Griffin PR, Hunkapiller T, Anal Biochem 214 (2):397-408 (1993). The PMF for an unknown protein can be compared with all of the PMFs in a database to find the best match, thereby identifying the protein. Searches of this kind can be constrained by determining the mass of the protein prior to digestion. In this

20 way the pattern of masses of an unidentified polypeptide can be related to its sequence, which in turn can help to determine the role of a protein in a particular sample.

Conventional techniques for determining the expression of proteins in biological samples depend on protein identification. The goal of protein

25 expression profiling is to identify as many proteins in a sample as possible and, preferably, to determine the quantity of the protein in the sample. A typical method of profiling a population of proteins is by two-dimensional electrophoresis (R. A. Van Bogelen., E. R. Olson, Biotechnol Annu Rev, 1: 69-103,1995).

In this method a protein sample extracted from a biological sample is separated by two independent electrophoretic procedures. This first separation usually separates proteins on the basis of their iso-electric point using a gel-filled capillary or gel strip along which a pH gradient exists.
5 Proteins migrate electrophoretically along the gradient until the pH is such that the protein has no net charge, referred to as the iso-electric point, from which the protein can migrate no further. After all of the proteins in the sample have reached their iso-electric point, the proteins are separated further using a second electrophoretic procedure. To perform the second procedure, the
10 entire iso-electric focussing gel strip is then laid against one edge of a rectangular gel. The separated proteins in the strip are then electrophoretically separated in the second gel on the basis of their size. The proteins are thus resolved into a 2-dimensional array of spots in a rectangular slab of acrylamide.

15 However, after separating the proteins in a sample from each other, there remains the problem of detecting and then identifying the proteins. The currently favoured approach to identify proteins is to analyse the protein in specific spots on the gel by peptide mass fingerprinting using MALDI-TOF mass spectrometry (Jungblut P, Thiede B. Mass Spectrom Rev. 16: 145-162,
20 1997). 2-DE technology is therefore limited by the detection capabilities of the peptide mass fingerprinting methods used in the identification of proteins in gel spots.

The existing technology cannot easily compare the expression levels of two or more samples and there are sensitivity problems with such a complex
25 process due to sample losses during the separation of the proteins and their subsequent recovery from the 2-D gel. In addition, proteins extracted from a 2-D gel are generally in buffers containing solutes that are incompatible with mass spectrometric analysis.

It is an aim of this invention to solve the problems associated with
30 the known methods described above.

The present invention provides an a highly preferred embodiment a method for identification of disease-related peptides and/or proteins for use as markers in diagnosis, prognosis, or therapeutic monitoring of disease. This highly preferred embodiment comprises the following steps, wherein (a1), (a2), etc.....refer to step (a) in the general method description above.

In one of the first steps (a1), a sample of a diseased body tissue is provided. This sample is processed by subjecting the sample to laser capture microdissection to provide a test sample consisting of a first collection of microdissected cells of about 200- 3,000 cells.

Laser capture microdissection (LCM) is a method for isolating pure cells of interest from specific microscopic regions of tissue sections. A transparent transfer film is applied to the surface of a tissue section. Under a microscope, the thin tissue section is viewed through the glass slide on which it is mounted and microscopic clusters of cells are selected for isolation. When the cells of choice are in the center of the field of view, the operator pushes a button which activates a near IR laser diode integral with the microscope optics. The pulsed laser beam activates a precise spot on the transfer film, fusing the film with the underlying cells of choice. The transfer film with the bonded cells is then lifted off the thin tissue section, leaving all unwanted cells behind. The laser capture microdissection process does not alter or damage the morphology and chemistry of the sample collected, nor the surrounding cells. For this reason, LCM is a useful method of collecting selected cells for DNA, RNA and/or protein analyses. LCM can be performed on a variety of tissue samples including blood smears, cytologic preparations, cell cultures and aliquots of solid tissue. Frozen and paraffin embedded archival tissue may also be used.

Prior, subsequently or simultaneous thereto, in another one of the first steps (a2), a sample of a corresponding healthy body tissue is provided. This sample is treated similarly and therefore also subjected the sample to

laser capture microdissection to provide a reference sample consisting of a second collection of microdissected cells of about 200- 3,000 cells.

In order to provide for effective comparison of expression levels of peptides and proteins in test and reference samples, the skilled person will understand that processing is preferably similar between samples. Most preferably, diseased and healthy state are treated in parallel.

In a subsequent step (a3), both the test and reference sample are subjected to trypsin digestion to provide processed samples comprising peptide fragments from the proteins in said samples in a mass range of 800 to 4,000 Da.

Sample processing is now complete and in a subsequent step, (b), the processed test and reference samples are subjected to MALDI- FT-ICR mass spectrometry to generate mass spectra for the individual peptides in each sample and to quantify the amount of individual peptides present in each sample.

The MALDI FT-ICR procedures are those commonly used in the field of protein analysis. Suitable procedures are provided in the Examples below.

In order to generate a list of peptides that are differentially expressed between test and reference sample, the method comprises as a subsequent step (c) the step of comparing the amount of an individual peptide(s) present in the test sample with the amount of the same peptide (*i.e.* a peptide having a corresponding mass spectrum) in the reference sample.

Such comparisons are suitably performed as follows. Mono-isotopic peaks with $S/N > 3$ are annotated with the SNAP algorithm using the pre-release version of DataAnalysis software package (v3.4, built169). The peak lists are then saved in a general text format, which can then be used as an input for a home made script in the R-program, (www.r-project.org). With this script a matrix file can be generated, which indicaties the presence or absence of each peptide mass in the various mass spectra. If a specific peptide appears

at least in 5 samples for the test group and never appears in the reference groups, then such a peptide can be was considered as a peptide that is specific for the disease. In this way, a list of differentially expressed peptides can be generated.

5 The next phase in the method of the invention (step (d)) is the identification of the differentially expressed peptides. For this, large sample sizes are preferred and pooled collections of microdissected cells of the test and/or reference sample may therefore be provided. Samples are preferably subjected to nano liquid chromatography (nano-LC) fractionation to provide 10 fractions comprising separated proteins. The proteins in the samples are optionally trypsin digested and the samples are in a then subjected to tandem mass spectrometry (MS-MS).

The skilled person will understand that essentially the same samples as provided in step (a) of the method are used, although they need not 15 be identical samples. The purpose is the provide signatures from peptide obtained from the same tissues that can be matched with those of the differentially expressed peptides. For this, the sequence of the differentially expressed peptides and/or that of the peptide obtained from the same tissue are obtained by MALDI-TOF/TOF.

20 One way of obtaining the sequence for the differentially expressed peptides and/or the proteins from which they derive is to compare the mass spectrum of a differentially expressed peptide identified by MALDI-FT in step (b) with the mass spectra of the peptides whose sequence is obtained by tandem mass spectrometry in step (d), and deducing the sequence for the 25 differentially expressed peptide identified in step (b) from the sequence of the corresponding peptide determined in step (d) by annotation. Next, the sequence deduced for the differentially expressed peptides can be matched with a protein database to identify the peptide or protein that is differentially expressed between said test and reference sample. The thus identified peptide

or protein can be classified as a candidate marker peptide or protein for the disease.

This particular embodiment should not be construed as limiting the scope of the present invention. In fact, the skilled person will readily
5 understand which steps in the procedure may optionally be reversed. For instance, deducing the sequence for the differentially expressed peptide identified in step (b) from the sequence of the corresponding peptide determined in step (d) by annotation, and matching the sequence deduced for the differentially expressed peptides with a protein database to identify the
10 peptide or protein that is differentially expressed between said test and reference sample, may also be reversed by matching the sequence of the corresponding peptide determined in step (d) with a protein database and annotating the identity to and the peptide or protein that is differentially expressed.

15 Further confirmatory test can be performed to validate the candidate marker. For instance, the candidate marker protein or marker peptide can be deigested *in silico* and the resulting theoretical peptides can be compared to the monoisotopic peaks obtained for peptides in the MALDI analyses in step (b) and/or (d).

20 A particular advantage of the method of the invention is that very small samples can be used to find differentially expressed peptides between samples. In body fluid samples, an amount as low as 1-10 µl of body fluid is sufficient for MALDI- FT-ICR mass spectrometris analysis. In order to provide for a sufficient number of peptide, the body fluid preferably comprising about
25 0.05- 5 mg/ml of total protein.

The diseases for which diagnosis, prognosis or therapeutic monitoring can be provided through provision of the markers obtained by methods of the present invention are not limited to any particular disease. Any patient condition is in principle eligible for obtaining markers be the present
30 method. A patient can be any animal, but is preferably a human patient.

Markers for predicting early onset preeclampsia

By using the methods of the present invention, the present inventors identified specific proteins of villous trophoblast and villous stroma in cells from placentas of women with pregnancies complicated by preeclampsia using tryptic peptide profiling by MALDI-TOF mass spectrometry and NanoLC of homogenised tissue to obtain sequencing information (MS/MS spectra). Subsequently, specific placental peptide patterns were identified as being associated with the occurrence of early onset preeclampsia and discriminating proteins including calcyclin, surfeit locus protein, and choriomammotropin A precursor, were identified as indicative for early onset preeclampsia. The expression of calcyclin in early onset preeclampsia was verified by immunohistochemistry of placental sections. Details of these experiments are provided in Example 1. These data suggest that in early onset preeclampsia trophoblastic choriomammotropin regulation is abnormal, possibly through abnormal calcyclin expression and regulation.

Thus, as a direct result of the novel and inventive method the inventors discovered proteinaceous markers the presence or expression level of which is indicative for early onset preeclampsia. These markers specifically included choriomammotropin precursor, calcyclin and surfeit locus protein 4. Suitable body tissue sample wherein the expression level of these markers is to be detected is villous trophoblast tissue from placentas of pregnant women.

Generally, the markers are detected in amounts of around 0.1-100 femtomole per volume of 100-200 cells, preferably 0.5-5 fmole/100-200 cells, and generally around about 1 fmole/100-200 cells. The skilled person will understand that the exact value will depend on the marker and on the normal values (reference values) measured in normal, healthy patients, not suffering from the disease. The skilled artisan is well aware of methods to obtain reference values for diagnostic markers. Generally, typical reference samples will be obtained from subjects that are clinically well documented and that are

free of the disease. In such samples, normal (reference) concentrations of the marker proteins can be determined, for instance by providing the average concentration over the reference population. In determining the reference concentration of the marker concentration a variety of considerations is taken
5 into regard. . Among such considerations are the type of marker (e. g., hormone peptide, structural protein, etc.), the type of disease to be diagnosed, the location of disease and the type of sample involved (e. g., tissue or CSF), the patient's age, weight, sex, general physical condition and the like. For instance, a group of at least 2 to preferably more than 3, preferably ranked
10 according to the above considerations, for instance from various age categories, are taken as reference group.

In general, a level in the concentration of the marker that is increased at least 1.5-10 times, preferably 2-5 times, but suitably about 3 times, relative to concentration of the reference value is indicative of the
15 disease.

Depending on the normal (healthy) status, a marker indicative of a disease or condition as referred to herein may be present in the diseased condition vs. absent in the normal condition. More often however, the level of expression of the marker will e altered, usually enhanced, so that elevated
20 levels of the marker indicate the presence of the disease or even the severity of the disease condition. Therefore, in some instances, quantitative detection of the marker and comparison with reference values is necessary in order to draw conclusions. The steps which must be taken in order for a diagnosis to be made are generally:
25 i) an examination phase involving the collection of data,
ii) a comparison of these data with standard values,
iii) a finding of any significant deviation during the comparison, and
iv) the attribution of the deviation to a particular clinical picture, i.e. the deductive medical or veterinary decision phase.

In methods of the present invention, step iv is generally excluded. The methods of the present invention in particular relate to the technical steps of providing samples and providing clinical data on marker concentrations, which steps proceed the deductive medical or veterinary decision phase.

5 Detection of the marker in a patient sample may be performed by any method available to the artisan. Generally, in order to detect the subtle concentration differences in the expression level of the marker, sophisticated methods are required. The skilled person is well acquainted with the various methods available, and these need not be described in great detail here.

10 In short, suitable methods include mass spectrometric methods such as those described and used herein, in particular in the Examples, and immunological detection methods.

Immunological detection methods (i.e. immunoassays) for determining the (quantitative) presence of a peptide or protein in a sample are

15 well known to those of skill in the art. The markers identified by methods of the present invention can be employed as immunogens for constructing antibodies immunoreactive to a protein of the present invention for such exemplary utilities as immunoassays or protein purification techniques.

In another aspect, the present invention provides for the use of a

20 disease marker, identified by a method for identifying a disease marker according to the invention, in diagnosis, prognosis, or therapeutic monitoring of disease.

In another aspect, the present invention provides for an isolated or synthetic disease marker identified by a method for identifying a disease

25 marker according to the invention.

In another aspect, the present invention provides for a specific binding partner for the disease marker of the present invention. The preferred binding partner is an antibody against the disease marker of the present invention. The term "antibody" includes reference to antigen binding forms of

30 antibodies (e. g., Fab, F(ab) 2). The term "antibody" frequently refers to a

polypeptide substantially encoded by an immunoglobulin gene or immunoglobulin genes, or fragments thereof which specifically bind and recognize an analyte (antigen). However, while various antibody fragments can be defined in terms of the digestion of an intact antibody, one of skill will
5 appreciate that such fragments may be synthesized de novo either chemically or by utilizing recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments such as single chain Fv, chimeric antibodies (i. e., comprising constant and variable regions from different species), humanized antibodies (i. e., comprising a complementarity
10 determining region (CDR) from a non-human source) and heteroconjugate antibodies (e. g., bispecific antibodies). The antibody can be monoclonal or polyclonal and can be prepared by techniques that are well known in the art such as immunization of a host and collection of sera (polyclonal) (see, e.g., Parker, Radioimmunoassay of Biologically Active Compounds, Prentice-Hall
15 (Englewood Cliffs, N.J., U.S., 1976), Butler, J. Immunol. Meth. 7: 1-24 (1975); Broughton and Strong, Clin. Chem. 22: 726-732 (1976); and Playfair, *et al.*, Br. Med. Bull. 30: 24-31 (1974)) or by preparing continuous hybrid cell lines and collecting the secreted protein (monoclonal) (see, *e.g.*, Khler and Milstein, Nature 265:495-497, 1975, Lymphocyte Hybridomas, ed. Melchers, *et al.*
20 Springer-Verlag (New York 1978), Nature 266: 495 (1977), Science 208: 692 (1980), and Methods of Enzymology 73 (Part B): 3-46 (1981)) or by cloning and expressing nucleotide sequences or mutagenized versions thereof coding at least for the amino acid sequences required for specific binding of natural antibodies. Antibodies may include a complete immunoglobulin or fragment
25 thereof, which immunoglobulins include the various classes and isotypes, such as IgA, IgD, IgE, lgG1, IgG2a, lgG2b and lgG3, IgM, etc. Fragments thereof may include Fab, Fv and F(ab')2, Fab', and the like. In addition, aggregates, polymers, and conjugates of immunoglobulins or their fragments can be used where appropriate so long as binding affinity for a particular molecule is
30 maintained.

Monoclonal antibodies can be used for detection purpose in the present invention, for example, in immunoassays in which they can be utilized in liquid phase or bound to a solid phase carrier. In addition, the monoclonal antibodies in these immunoassays can be detectably labeled in various ways. A variety of immunoassay formats may be used to select antibodies specifically reactive with a particular peptide or protein marker. For example, solid-phase ELISA immunoassays are routinely used to select monoclonal antibodies specifically immunoreactive with a protein. See Harlow and Lane, Antibodies, A Laboratory Manual, Cold Spring Harbor Publications, New York (1988), for a description of immunoassay formats and conditions that can be used to determine selective binding. Examples of types of immunoassays that can utilize monoclonal antibodies of the invention are competitive and non-competitive immunoassays in either a direct or indirect format. Examples of such immunoassays are the radioimmunoassay (RIA) and the sandwich (immunometric) assay.

Detection of the peptide or protein marker using an antibody can be done utilizing immunoassays that are run in either the forward, reverse, or simultaneous modes, including immunohistochemical assays on physiological samples. Those of skill in the art will know, or can readily discern, other immunoassay formats without undue experimentation.

Antibodies can be bound to many different carriers and used to detect the presence of the disease markers. Examples of well-known carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, agaroses and magnetite. The nature of the carrier can be either soluble or insoluble for purposes of the invention. Those skilled in the art will know of other suitable carriers for binding monoclonal antibodies, or will be able to ascertain such using routine experimentation.

The binding of the antibody to the disease marker can be detected in numerous ways that are well known in the art. Binding of the antibody and

disease marker forms an immune complex that can be detected directly or indirectly. The immune complexes are detected directly, for example, when the antibodies employed are conjugated to a label. The immune complex is detected indirectly by examining for the effect of immune complex formation in an assay medium on a signal producing system or by employing a labeled receptor that specifically binds to an antibody of the invention. Suitable detection techniques that may be applied in concert with the above techniques include autoradiographic detection techniques, detection techniques based on fluorescence, luminescence or phosphorescence or chromogenic detection techniques. These detection techniques are known in the art of detection of biomolecules.

Use may for instance be made of signal producing systems, involving one or more components, at least one component being a detectable label, which generate a detectable signal that relates to the amount of bound and/or unbound label, i.e. the amount of label bound or not bound to the compound being detected. The label is any molecule that produces or can be induced to produce a signal, and preferably is a fluorescer, radio-label, enzyme, chemiluminescer or photosensitizer. Thus, the signal is detected and/or measured by detecting fluorescence or luminescence, radioactivity, enzyme activity or light absorbance.

Suitable labels include, by way of illustration and not limitation, enzymes such as alkaline phosphatase, glucose-6-phosphate dehydrogenase ("G6PDH") and horseradish peroxidase; ribozyme; a substrate for a replicase such as QB replicase; promoters; dyes; fluorescers, such as fluorescein, rhodamine compounds, phycoerythrin, phycocyanin, allophycocyanin, o-phthaldehyde, and fluorescamine; chemiluminescers such as isoluminol; sensitizers; coenzymes; enzyme substrates; radiolabels such as $^{125}1$, ^{14}O, ^{3}H, ^{57}Co and ^{75}Se; particles such as latex or carbon particles; metal sol; crystallite; liposomes; cells, etc., which may be further labeled with a dye, catalyst or other detectable group. Suitable enzymes and coenzymes are disclosed in U.S.

Patent No. 4,275,149; U.S. Patent No. 4,318,980; suitable fluorescers and chemiluminescers are disclosed i.a. in U.S. Patent No. 4,275,149.

There are numerous methods by which the label can produce a signal detectable by external means, for example, desirably by visual examination or by electromagnetic radiation, heat, and chemical reagents. The label or other signal producing system component can also be bound to a specific binding partner, another molecule or to a support.

The label can directly produce a signal, and therefore, additional components are not required to produce a signal. Numerous organic molecules, for example fluorescers, are able to absorb ultraviolet and visible light, where the light absorption transfers energy to these molecules and elevates them to an excited energy state. This absorbed energy is then dissipated by emission of light at a second wavelength. Other labels that directly produce a signal include radioactive isotopes and dyes.

Alternately, the label may need other components to produce a signal, and the signal producing system would then include all the components required to produce a measurable signal, which may include substrates, coenzymes, enhancers, additional enzymes, substances that react with enzymic products, catalysts, activators, cofactors, inhibitors, scavengers, metal ions, and a specific binding substance required for binding of signal generating substances. A detailed discussion of suitable signal producing systems can be found in U.S. Patent No. 5,185,243.

The label can be bound covalently to numerous specific binding partners: an antibody; a receptor for an antibody; a receptor that is capable of binding to a small molecule conjugated to an antibody; or a ligand analog. Bonding of the label to the specific binding partner may be accomplished by chemical reactions which result in replacing a hydrogen atom of the label with a bond to the specific binding partner member or may include a linking group between the label and the specific binding partner. Other signal producing system components may also be bound covalently to specific binding partners.

For example, two signal producing system components such as a fluorescer and quencher can each be bound to a different antibody that forms a specific complex with the analyte.

Formation of the complex brings the fluorescer and quencher in close proximity, thus permitting the quencher to interact with the fluorescer to produce a signal. Methods of conjugation are well known in the art. See for example, U.S. Patent No. 3,817,837. This invention also contemplates having an antibody bound to a first signal producing system component and a detectable label as the second signal producing system components. For example, when the detectable label is bound to a ligand analog, the extent of binding of the antibody to the analog can be measured by detecting the signal produced by the interaction of the signal producing system components.

Methods and means provided herein are particularly useful in a diagnostic kit for diagnosing a disease by immunological techniques. Such kits or assays may for example comprise one or more reference markers, one or more reference samples and/or one or more antibodies for any of the markers for the various disease conditions as described herein.

Markers for detecting glioma

The present inventors also set out to identify proteins that are specifically expressed in glioma vasculature, but not in the normal blood vessels of the brain. To this end, microdissected hypertrophied and normal blood vessels of the brain were used. The peptides of the enzymatically digested proteins derived from the small numbers of cells obtained by microdissection, were measured by MALDI-FT mass spectrometry. The identification of differentially expresses peptides was achieved by combining nano-LC fractionation of samples with offline MALDI-TOF/TOF and MALDI FTMS measurements. The findings were validated by using specific antibodies. Details of these experiments are described in the Examples.

Thus, as a direct result of the novel and inventive method, the inventors discovered proteinaceous markers the expression level of which is indicative for glioma. These markers specifically included fibronectin, fibrinogen, colligin 2 and acidic calponin 3. Suitable body fluid samples wherein the expression level of these markers is to be detected are blood or cerebrospinal fluid samples. Suitable body tissue sample wherein the expression level of these markers is to be detected is glioma vasculature tissue.

The markers of the present invention are very suitably used in a method for monitoring the disease activity of glioma or the response of the patient to treatment regimens. Such a method comprises the step of measuring the expression level of fibronectin, fibrinogen, colligin 2 and acidic calponin 3 in blood, CSF or glioma vasculature. Reference values for markers may be determined as described above and methods of diagnosis of glioma may be performed as described above for preeclampsia.

Markers for detecting multiple sclerosis

The inflammatory and neurodegenerative pathology of multiple sclerosis is localized in the central nervous system. Therefore cerebrospinal fluid (CSF) is a promising biofluid in the search for biomarkers and disease-associated proteins for multiple sclerosis. In this study we used peptide profiling by mass spectrometry to investigate if CSF peptides of multiple sclerosis patients differ significantly from CSF peptide profiles of control patients, followed by identification two independent mass spectrometry strategies (accurate mass determination and MS/MS).

As a direct result of the novel and inventive method, the inventors discovered proteinaceous markers the expression level of which is indicative for multiple sclerosis. These markers specifically included chromogranin A, clusterin, complement C3, complement C4B, beta V spectrin, hypothetical protein XP_011125, apolipoprotein D, complement C4A, contactin 1, neuronal pentraxin receptor, and RNA binding motif protein 7. Suitable body fluid

samples wherein the expression level of these markers is to be detected are cerebrospinal fluid samples.

The markers as described herein are very suitably used in a method for monitoring the disease activity of multiple sclerosis or the response of the patient to treatment regimens. Such a method preferably comprises the step of measuring the expression level of chromogranin A, clusterin, complement C3, complement C4B, beta V spectrin, hypothetical protein XP_011125, apolipoprotein D, complement C4A, contactin 1, neuronal pentraxin receptor, and/or RNA binding motif protein 7 in a patient sample and comparing the value obtained with normal (reference) values determined for healthy patients. Reference values for markers may be determined as described above and methods of diagnosis of glioma may be performed as described above for preeclampsia and glioma.

Methods for measuring the expression level of peptides or proteins by MALDI techniques as referred to herein are well known in the art and specific reference is made to the Experimental part described herein below.

The invention will now be illustrated by the following non-limiting Examples.

EXAMPLES

Example 1. Specific peptides identified by mass spectrometry in placental tissue from pregnancies complicated by early onset preeclampsia attained by laser capture dissection

1.1. Experimental Procedures and Methods

1.1.1. Placental samples

A total of 17 human placentas were obtained at the obstetrical wards of the departments of Obstetrics and Gynaecology of the Radboud University Nijmegen Medical Center, Nijmegen, and Erasmus MC Rotterdam, The Netherlands, after having given informed consent. Of these 17 placentas 7 were obtained from women after uncomplicated normotensive pregnancies, 7

from women with preeclampsia and 3 placentas from normotensive women after preterm delivery of unknown cause (excluding infection or systemic diseases). From the 7 women with preeclampsia, 4 experienced early onset preeclampsia (before 34 weeks gestation) and 3 late onset preeclampsia (Table 5 1).

Table 1. Clinical characteristics of pregnancies from studied placenta samples.

	Diagnosis	Maternal age (yr)	Parity	Blood pressure (mm Hg)*	Proteinuria (g/10 mmol)	Gestational age at delivery (days)	Birth weight (g)	Fetal sex
1	E-PE	31	0	110	47	239	1920	M
2	E-PE	28	0	115	6,63	219	909	F
3	E-PE	32	1	120	21,11	236	1856	M
4	E-PE	33	0	115	5,51	213	1053	F
5	PE	26	0	120	0,79	268	3045	M
6	PE	26	0	115	7,11	276	2580	M
7	PE	20	0	130	6,2	254	2343	M
8	NL	22	1	75	0	269	3025	F
9	NL	33	0	75	0	268	3010	F
10	NL	32	1	67	0	271	3220	M
11	NL	31	1	80	0	274	3370	M
12	NL	30	0	70	0	277	2335	M
13	NL	30	0	80	0	264	3150	F
14	NL	27	2	75	0	274	4110	M
15	Preterm	24	0	100	0	204	1110	M
16	Preterm	21	0	65	0	223	1750	M
17	Preterm	38	2	70	0	218	1695	F

Diagnosis: E-PE = early onset preeclampsia, PE = preeclampsia, NL = normal, Preterm = 10 preterm delivery, * Blood pressure = diastolic blood pressure at admission, proteinuria = protein/creatine ratio, normal 0.3 g/10 mmol, F = female, M = male

Women with early onset preeclampsia were matched for gestational age and parity with women with preterm delivery of unknown cause. One woman in the last mentioned group had hypertension at admission (#15, pregnancy induced hypertension at admission, no preeclampsia). Preeclampsia was defined according to the International Society for the Study of Hypertension in Pregnancy (ISSHP) definition (Am. J. Obstet. Gynecol. 183, S1-22 (2000)). The Medical Ethics Review Boards of the Radboud University Nijmegen Medical Center and the Erasmus MC, Rotterdam approved the protocol.

1.1.2. Laser microdissection

Placenta trophoblastic tissues were dissected as described recently (De Groot, C.J., *et al.*(2005) Proteomics 5, 597-607). In brief, placenta sections were directly frozen in liquid nitrogen and liquid iso-pentane (Brunschwig Chemie, Amsterdam, the Netherlands). The frozen tissue was embedded into Cryoblock (Klinipath BV, Duiven, Netherlands) tissue medium and subsequently frozen tissue sections (5 μm) were made by cryostat (type HM 500 Adamas, Rhenen, The Netherlands). The sections were mounted on a PEN (polyethylene naphthalate) membrane (1.35 μm) as recommended by the manufacturer (P.A.L.M. Microlaser Technologies AG, Bernried, Germany). After cutting, the sections were placed in 70 % alcohol (99,8 % purity) for 1 minute followed by 10 seconds in Milli-Q water, 1 minute in hematoxylin (2x repeated), 1 minute in 0.5% acetic acid, 1 minute in ammonia water, and eosine for 10 seconds, respectively. Subsequently, the slides were dehydrated in 96 % alcohol and 100 % alcohol each for 1 minute, respectively. Dehydration with 96 % alcohol and 100 % alcohol were repeated twice.

Trophoblast cells were captured with a P.A.L.M Microlaser capture microdissection (type P-MB, Bernried Germany). The microdissected cells were collected in 0.1 % Rapigest SF reagent (sodium 3-[(2-methyl-2-undecyl-1,3-dioxolan-4-yl)methoxy]-1-propanesulfonate) (Waters Corporation, Milford, MA,

USA) in 50 mM ammonium bicarbonate buffer for approximately 1.5 hour before digestion with trypsin.

Laser microdissected tissues were centrifuged at 10,000 rpm for 10 seconds and subsequently sonified by a Ultrasonic Disruptor Sonifier II Model
5 W-250/W-450 (Bransons Ultrasonics) equipped with a cuphorn for 1 minute at 70% amplitude. Samples were boiled for 5 minutes. After cooling to room temperature, trypsin (Promega Corporation Cat No.: V5280, U.S.A.) solutions were added routinely at a concentration of 0.1 ug/ uL. Digestion was performed for 1.5 hour at 37°C. Subsequently, the digestion reaction was stopped by
10 incubating at 50 mM HCl.

1.1.3. Mass spectrometry measurements

Mass spectrometry was automatically performed on an Ultraflex I MALDI-TOF mass spectrometer (Bruker Daltonics, Billerica, USA), with a
15 laser power of 50 - 60 % and 1,000 laser shots per single spectrum. Calibration occurred with the peptide calibration standard (Bruker Daltonics Part No.: 206195, Bremen, Germany). The control spectra, obtained from sole PEN membrane treated in the same way as microdissected tissue, were used as background spectra. Independently, microdissection was performed three
20 times and from each sample a triplicate mass spectrometry measurement was performed. Accordingly, 9 spectra were obtained from each of the 17 placentas. The reproducibility of the intensities of the measurements of the complex peptide mixtures ranged between 15% and 30%.

25 1.1.4. Data and Mass spectrometry analysis

Analysis of mass spectra was performed by using ClinProTools biosoftware (version 1, Bruker Daltonics, Leipzig, Germany) using comparisons illustrated in Figure 1. The proteome data of the trophoblast and stroma data were compared between placentas from women with
30 uncomplicated pregnancies (controls, n=7), women with pregnancies

complicated by late (n=3) and early onset preeclampsia (n=4), and placentas from women with unexplained preterm birth (n=3). Significant differences of monoisotopic peaks were defined as $p < 0.01$. The differentially expressed peaks excluding the background peaks of the PEN membrane were mass
5 annotated.

Discriminant analysis and Principal Component Analysis (PCA) of the MALDI-TOF data was performed using ChemoMetricks, an in-house developed statistical analysis toolbox. This toolbox was developed on the Matlab 5.2 (Mathworks, Nantucket MA) platform. All spectra were truncated
10 to 190700 points. Individual spectra were subsequently rebinned in segments of 100 points, yielding 122 spectra of 1907 variables for the preterm data and 255 spectra with 1907 variables in the 7 control versus 7 cases comparison. After reading in the full data the data was normalized and autoscaled prior to PCA. After computing the Principal Components (PC) and the Discriminant
15 functions (DF) all data was plotted using standard Matlab plotting routines.

Placental tissue sections of the samples (3, 10, 17, Table 1) representative for the three groups were separated on an Ultimate NanoLC system (Dionex, Sunnyvale, CA, USA) prior to MALDI-TOF analysis. Assuming that a cell size is 10 x 10 x 10 µm, the tissue represented
20 approximately 50.000 cells. Tissue samples were prepared in the same way as microdissected tissue. Subsequently, fractionation was performed on 5 µL tissue sample solutions, for 130 min with a 5 to 95% step wise gradient of solution A (100% H2O, 0.05% TFA) and solution B (20% H2O, 80% ACN, 0.04% TFA). A 15 cm C18 column with 75 µm ID and a guard column (Dionex,
25 Sunnyvale, CA, USA) were used. Separate fractions were collected using a robotic spotter (Probot Micro Fraction collector, Dionex, Sunnyvale, CA, USA) on a prespotted, with matrix of α-cyano-4-hydroxycinnimic acid, anchorchip plate (Bruker Daltonics Germany). One µL of each fraction was spotted. Mass spectrometry analyses were performed in reflectron mode using standard
30 settings for peptide measurements (Bruker default file "Proteomics_HPC.par")

in addition to the MS spectrum a MS/MS spectrum was obtained for each peak using the standard settings for a lift MS/MS measurement. LC Warp software (Bruker Daltonics) was used to combine MS/MS spectra of peptides of the same protein. The resulting spectra were analyzed and annotated in Flexanalyses 2.2 (Bruker Daltonics). In Biotools 2.2 (Bruker Daltonis) the peak lists from the MS spectra and the MS/MS were combined to create a MGF file. The MGF file was used as input for the Mascot search engine (Matrix Science, London, UK) to search the MSDB human database using a 100 ppm tolerance of the MS spectra and a 0.5 Da tolerance for the fragments in the MS/MS spectra. Identification was significant, if a Mowse score with a probability of $p<0.05$ was observed in the Mascot search engine.

1.1.5. MALDI FT ICR mass spectrometry

Prespotted anchorchip plates (Bruker Daltonics, Germany) used for the MS/MS sequencing in the MALDI TOF/TOF were spotted with DHB (10 mg/ mL 0.1%TFA water). Subsequently, for confirmation remains of the fractions were analyzed by MALDI FT ICR MS (APEX Q, Bruker Daltonics) for accurate mass annotation. Identification was confirmed if the calculated mass of the identified differential expressed peptide peaks observed in the MALDI-TOF/TOF analysis did not deviate more than 4 ppm in the MALDI FT-ICR MS analysis from the calculated mass of the peptide identified.

1.1.6. Validation of calcyclin and choriomammotropin by immunohistochemistry

Commercial available antibodies specific for calcyclin (P06703, Sigma Aldrich, Saint Louis, USA) and choriomammotropin (LCHUC, DakoCytomation, Glostrup, Denmark) were used for validation by immunohistochemistry: according to the recommendation of the manufacturers.

1.2. Results

1.2.1. Clinical data

Clinical and demographic characteristics are given in table 1. As defined women with preeclampsia (n=7; # 1 -7) had significantly higher blood pressure at admission and proteinuria. As expected these women also had on average a shorter gestational age ($p< 0.05$) and their newborns had a lower birth weight ($p<0.05$) compared to women with uncomplicated pregnancies (n=7; #8-14). No significant differences were found between women with preeclampsia and controls regarding maternal age and parity. For detailed analysis preeclampsia was stratified into early onset preeclampsia (n=4; # 1-4) and late onset preeclampsia (n=3; # 5-7). Between women with severe early onset preeclampsia (n=4) and women with preterm delivery of unknown cause (n=3) no differences were found for gestational age at delivery and birth weight. All but one woman were delivered by caesarean section (# 16).

Analysis of peptide spectra of placentas from controls and preeclamptics

In Figure 1 numbers of significantly differentially expressed peptides and total numbers of peptides for various comparisons are given for trophoblast and villous stroma cells of women with uncomplicated pregnancies, early- and late onset preeclamptic women and women with preterm delivery. In Figure 2 differences between peptide profiles of trophoblasts of placentas from women with uncomplicated pregnancy (n=7) and preeclampsia (n=7) are illustrated in different views. From each mass over charge value (m/z) a statistical evaluation is performed using ClinProTools software (Bruker Daltonics, Leipzig, Germany). In tables 2 and 3, the exact data of significantly expressed peptides ($p<0.01$) are described for both the comparison of trophoblast and stroma cells, respectively.

Table 2. Differences in peptides of trophoblast cells from placentas of preeclamptics (n=7) and controls (n=7).

mass	Ave control	Ave case	SD control	SD case	p-value	t-test	Spectra controls 1)	Controls 2)	Subsamples control	Spectra cases1)	Patient cases2)	Subsamples cases
263.082	16.3631	8.17539	21.1839	6.17205	0.00258327	3.10687	10	4	4/2/4	17	6	7/3/7
487.841	38.5705	20.8703	34.3616	11.5647	9.89E-05	4.07983	32	7	11/10/11	19	5	8/5/6
531.077	46.453	38.0458	16.8788	12.0523	0.00115908	3.32484	62	7	18/20/22	42	7	13/15/14
557.41	28.7307	21.4808	12.9487	8.67216	0.000207529	3.82398	11	4	4/6/1	4	2	2/1/1
570.232	71.805	86.1616	30.0831	31.0658	0.00820695	-2.6861	44	6	14/17/13	59	7	20/24/15
646.784	22.3257	28.7206	12.9758	12.4333	0.00456151	-2.8872	44	4	5/7/8	26	6	11/8/7
861.148	27.7616	20.388	13.1927	11.7102	0.000891299	3.40144	20	5	10/13/11	21	6	7/9/5
887.841	29.0736	14.299	23.8716	4.69182	2.39E-06	5.10167	4	2	1/1/2	1	1	1/0/0
901.488	20.4772	11.5813	17.8979	2.28037	9.96E-05	4.14907	15	4	4/4/7	1	1	0/1/0
917.284	25.6835	15.3069	19.8261	8.24842	0.00114068	4.02339	16	3	7/6/3	5	3	0/3/2
1119.29	11.995	8.94666	8.2868	2.96339	0.00480048	2.89186	6	2	1/2/3	0	0	0/0/0
1223.79	9.27621	7.10801	4.02103	2.26607	0.000174418	3.88226	4	1	0/3/0	0	0	0/0/0
1249.87	13.4275	7.46008	12.8263	3.3023	0.000302125	3.7725	4	1	1/2/1	0	0	0/0/0
1263.7	13.6549	9.11307	10.1092	3.28615	0.000580235	3.57236	13	4	2/5/6	3	1	0/0/3
1274.76	20.0197	40.2306	15.5879	28.6243	3.87E-06	-4.923	24	4	8/9/7	50	6	18/17/15
1463.9	9.67729	6.71123	7.91324	2.19671	0.00331065	3.02554	9	1	3/3/9	0	0	0/0/0
1529.8	13.9933	22.911	10.0955	15.1133	0.000161617	-3.9182	22	5	8/7/7	38	6	13/14/11
1744.91	13.3673	17.7288	8.52165	10.1513	0.009197	-2.6484	1	1	0/0/1	2	2	1/0/0
1750.98	6.7016	8.35852	2.10111	4.08152	0.00528678	-2.8614	0	0	0/0/0	3	1	2/1/0
1833.96	14.334	19.944	10.3843	11.3303	0.00383763	-2.9472	14	3	5/6/3	30	6	11/11/18
1922.99	11.6271	14.2023	5.33754	5.40957	0.00696017	-2.7436	16	4	7/6/3	19	5	7/7/5
1962.99	15.3871	18.0859	6.05165	5.36474	0.0075167	-2.7156	36	6	15/10/11	33	5	11/10/12
1969.04	8.7128	10.5099	3.37162	2.98614	0.00148358	-3.2471	2	1	0/0/2	3	1	1/2/0
1989.02	8.0858	9.24187	2.47274	2.15234	0.00476722	-2.8718	2	1	2/0/0	2	1	1/0/1
1993.04	6.4040	7.48408	2.04107	2.15957	0.00393622	-2.9381	0	0	1/0/3	2	1	1/0/1
2003.04	7.47808	8.89717	2.6828	3.13957	0.00655682	.2.7674	4	1	1/0/3	0	0	0/0/0
2065.08	6.4621	7.59931	1.62099	2.04548	0.000666571	-3.4994	2	1	1/1/0	2	1	1/1/0
2093.12	5.7865	7.99298	1.28468	5.5112	0.00323833	-3.0564	0	0	1/2/7	2	1	1/0/1
2149.11	6.1218	7.14392	1.31519	2.18626	0.00193221	-3.1891	10	3	1/2/7	11	4	6/1/4
2312.2	7.58108	10.0398	2.92523	6.18476	0.00561813	-2.8437	24	5	7/6/11	18	4	8/6/4
2391.28	7.53957	10.1508	3.0919	6.01408	0.00293531	-3.0613	6	2	3/3/0	10	2	5/2/3
2424.31	8.09969	9.53876	3.61635	2.6154	0.0092417	-2.6436	8	1	3/3/2	3	1	1/2/0
2435.5	7.48574	10.1287	3.02709	4.27602	0.000102918	-4.0361	4	1	2/2/0	0	0	0/0/0
2509.27	6.21672	7.64788	2.38112	2.53189	0.00115145	-3.3279	2	2	0/0/2	3	3	2/1/0
2676.35	8.50049	10.7379	3.54022	2.8841	0.000106862	-3.9974	30	5	9/11/10	34	6	13/12/9
2690.34	8.8189	10.9772	4.00463	3.41302	0.00108296	-3.343	44	5	15/15/14	45	6	17/16/12
2746.39	7.02622	8.58981	3.07062	2.45706	0.00148131	-3.2479	2	2	0/0/2	3	2	0/0/3
2869.51	6.93271	8.27684	2.82235	2.96351	0.00894567	-2.6557	2	1	2/0/0	1	1	1/0/0
2936.5	8.55322	10.1708	2.88537	2.47244	0.000709859	-3.4687	2	2	2/0/0	4	2	3/0/1
3092.57	6.26394	7.88215	2.83176	2.45235	0.000598814	-3.5184	4	2	2/0/2	3	1	1/1/1
3093.61	7.75564	9.646	3.48941	4.19775	0.00620724	-2.7871	2	1	1/0/1	6	5	3/1/2
3293.8	7.77449	9.39953	3.25616	2.6201	0.00210387	-3.1383	3	2	0/1/1	4	2	1/1/2
3294.8	8.70374	10.6677	3.78652	3.42341	0.00216987	-3.1286	3	3	0/1/2	4	3	1/1/2
3493.87	5.75585	6.93194	2.49604	205888	0.00359315	-2.9659	0	0	0/0/0	4	3	2/1/1

Ave= average; 1) = number of peptides observed in all spectra (the maximum is 63); 2) number of patients or controls that contributed; subsamples in x/x/x = measured in the first, the second
5 and the third measurement.

Table 3. Differences in peptides of stroma cells from placentas of preeclamptics (n=7) and controls (n=7).

Ave mass	Ave control	Ave case	SD control	SD case	p-value	t-test	Spectra controls [1]	Controls [2]	Subsamples control	Spectra cases [1]	Patient cases [2]	Subsamples cases
487.814	42.1178	23.1608	30.7113	11.0834	1.65E-05	-4.59059	20	6	8/6/6	11	5	4/4/3
490.309	25.0365	12.6763	20.8517	4.06296	1.85E-05	-4.61239	27	7	10/8/9	21	5	7/6/7
524.66	7.29572	8.82291	1.82001	2.07677	3.48E-05	4.30785	56	7	17/20/19	58	7	20/17/19
552.398	10.8553	7.8641	4.53703	3.11342	4.18E-05	-4.2687	6	4	4/2/0	18	5	7/7/4
566.2	10.5989	13.6552	3.63932	4.3106	5.05E-05	4.21418	2	2	1/0/1	7	4	3/3/1
570.173	29.8944	16.5442	24.467	7.38558	9.29E-05	-4.1344	26	6	8/8/9	48	7	18/14/15
856.671	10.2805	7.48291	5.34308	2.44612	0.000308589	-3.7567	26	5	11/6/9	41	7	15/14/12
886.954	63.3734	84.9042	32.2283	33.5808	0.00451973	3.60859	7	3	1/3/3	5	2	2/2/1
902.397	36.7834	48.4946	18.1009	19.6097	0.00085833	3.42113	4	3	3/0/1	0	0	0/0/0
917.846	11.6403	7.68683	8.68136	3.19104	0.00112956	-3.37908	19	6	6/6/7	10	4	5/3/2
1224.87	9.72086	7.31542	6.51566	2.53906	0.00801883	-2.71815	3	2	2/0/1	0	0	0/0/0
1249.83	34.3887	41.2978	11.8057	11.7529	0.00155927	3.23779	5	1	3/0/2	1	1	1/0/0
1263.82	28.3048	17.1258	25.4219	9.09092	0.00157863	-3.27395	6	3	1/0/5	0	0	0/0/0
2253.33	10.5989	13.6552	3.6392	4.31603	5.05E-05	4.21418	11	3	3/3/4	6	4	4/1/1
2261.21	10.8924	7.84545	7.56622	3.11725	0.004232	-2.94093	31	6	12/8/11	37	6	14/11/12
2997.68	8.36354	6.97422	3.37358	1.88571	0.00563323	-2.83051	2	1	0/0/2	5	2	1/4/0
3654.25	6.88494	8.51103	3.32037	3.23387	0.0070877	2.7398	2	2	1/0/3	0	0	0/0/0

Ave= average; [1] = number of peptides observed in all spectra; [2] number of patients or controls that contributed; subsamples in x/x/x = measured in the first, the second and the third
5 measurement.

Discriminant analysis revealed that the major spectral differences have to be attributed to patient variation. Approximately 17 % of the data variance is attributed to patient variation. The classification by cell-type and
10 control/case variance contributes only 11% to the full data set as can be seen in figure 3 where DF1 describes the difference between trophoblasts and stroma (1.22 % variance) and DF2 describes the difference between control and case (7.96 % variance). Figure 3 demonstrates that under these constraints the different groups of preeclampsia and controls can be clearly separated.
15 The PCA and DA of MALDI-TOF peptide patterns of trophoblasts as well as stroma cells of placentas both from women with preeclampsia as well as controls are separated significantly. Ten samples were misclassified in the trophoblast vs. stroma classification. The control vs. case separation contained 15 misclassified spectra out of the 255 total spectra, i.e. approximately 6%. In
20 Figure 4 the frequency of the significantly differentially expressed peptides (as function of presence and intensity) in trophoblasts and villous stroma cells in

placentas from women with uncomplicated pregnancies (n=7) and preeclampsia (n=7) are displayed.

1.2.2. Analysis of peptide spectra of placentas from early onset preeclamptics
5 and unexplained preterm delivery

Figure 5 illustrates the frequency of differentially expressed peptides in women with early compared to late onset preeclampsia. In order to determine whether the differences between differentially expressed peptides of women with early vs. late onset preeclampsia could be explained by the
10 maturity of the placenta, differences in peptide profiles of trophoblast cells between early onset preeclampsia and preterm delivery with unknown cause are illustrated in Figure 6.

We observed that the m/z values (e.g. 1269,76; 1505,81; 1910,99; 1969,01; 2044,01; 2691,3) were specific for early onset preeclampsia
15 trophoblast samples (Figure 8). In Tables 4 and 5 the data of significantly expressed peptides and their presence in trophoblast and villous stroma cells is described in detail.

Table 4. Significant differences in peptides of trophoblasts from placentas from early onset preeclampsia (n=4) and preterm delivery (n=3).

Ave mass	Ave control	Ave case	SD control	SD case	p-value	t-test	Spectra controls [1]	Controls [2]	Subsamples control	Spectra cases [1]	Patient cases [2]	Subsamples cases
262.729	2.54648	9.26491	1.21996	6.24575	4.63E-07	6.1268	0	0	0/0/0	4	3	3/1/0
515.873	9.44347	26.5593	3.64945	11.8849	8.65E-10	7.93921	22	3	9/6/7	29	4	11/10/8
537.536	376.572	169.167	347.463	83.4654	0.0047706	-3.06218	24	3	9/9/6	21	4	7/7/7
613.939	43.4994	23.1673	26.4162	9.18296	0.0004051	-3.9567	8	1	3/2/3	11	4	8/2/1
616.91	19.4343	54.8513	15.0946	63.2623	0.0031649	3.15334	8	3	0/6/2	14	3	7/4/3
622.257	16.8088	23.7614	3.40847	12.3725	0.0032935	3.1305	3	1	1/0/2	15	4	7/3/7
630.856	20.6938	16.6097	8.14696	3.73546	0.0042462	-3.06076	9	3	5/1/3	0	0	0/0/0
725.273	60.78	31.6	32.1626	19.6349	0.0001688	-4.14114	24	3	6/9/9	23	4	7/9/7
741.318	34.4079	17.8617	17.2802	7.70411	5.20E-05	-4.62387	17	3	3/7/7	7	3	2/4/2
788.429	30.2288	20.6407	11.4581	9.22807	0.0009037	-3.53249	17	3	8/6/3	13	4	5/4/4
796.39	24.214	17.448	8.95023	3.11057	0.0007235	-3.75225	13	3	6/4/3	13	4	5/4/4
825.082	99.2515	34.2368	85.5689	40.6163	0.0008758	-3.6361	23	3	8/6/9	21	4	8/6/7
842.46	911.88	452.619	506.004	236.448	0.0001111	-4.35376	27	3	8/6/9	34	4	12/11/11
864.614	86.9617	42.5941	49.9207	23.112	0.0001435	-4.26908	27	3	9/9/9	32	4	11/10/11
901.35	19.2208	13.1052	8.45677	2.58185	0.0010604	-3.62587	10	3	5/3/2	1	1	0/1/0
1045.54	227.097	83.5326	129.477	130.931	7.35E-05	-4.28007	27	3	9/9/9	34	4	12/11/11
1184.75	8.38233	10.3135	1.44016	3.11185	0.0023415	3.21142	0	0	0/0/0	4	1	1/2/1
1248.67	7.43521	9.84574	2.06567	4.25875	0.0055534	2.89884	0	0	0/0/0	5	1	1/2/2
1269.76	8.05761	9.74981	1.76648	2.90095	0.0068571	2.80831	0	0	0/0/0	4	2	1/0/3
1274.68	20.963	43.8466	9.21995	36.7108	0.0012068	3.49843	19	3	5/8/6	26	4	9/8/9
1277.73	15.6697	20.5643	5.53301	8.40005	0.008352	2.73224	18	3	7/6/5	22	4	7/6/9
1282.72	7.72289	10.8695	1.30683	5.53924	0.0027782	3.20201	0	0	0/0/0	7	1	2/3/2
1290.78	8.46801	11.2979	1.74642	5.38637	0.0062903	2.87885	0	0	0/0/0	8	3	2/4/2
1326.76	22.6535	16.2906	10.4958	9.84918	0.0071382	-2.79629	26	3	8/9/9	11	3	4/3/4
1396.81	7.30824	9.18133	1.28187	3.69646	0.0086366	2.75355	0	0	0/0/0	9	2	2/3/4
1442.81	12.8313	8.8864	3.54718	3.30942	4.43E-05	-4.4436	18	3	2/8/8	3	2	1/1/1
1466.82	11.0005	8.2812	3.53704	3.0265	0.0025214	-3.17666	5	2	0/1/4	1	1	0/1/0
1505.81	7.23864	14.6642	1.11111	14.4531	0.00527	2.9847	0	0	0/0/0	9	1	3/3/3
1566.8	7.91821	13.4852	2.10523	10.3143	0.0040552	3.0677	2	1	0/0/2	9	2	3/3/3
1833.97	11.2337	22.6854	4.75023	13.8808	5.40E-05	4.49079	9	3	3/3/3	27	4	8/11/8
1910.99	7.03505	8.44743	1.40697	2.34379	0.0051382	2.91422	0	0	0/0/0	7	1	2/3/2
1922.95	10.9151	15.7314	3.79846	6.30823	0.0005154	3.68876	11	3	2/4/5	29	4	9/10/10
1925.97	10.7102	13.9789	3.37766	4.95078	0.0033838	3.05683	6	3	1/0/5	24	4	8/10/6
1969.01	7.37333	11.0774	2.32176	3.34897	4.04E-06	6.09017	0	0	0/0/0	18	3	7/6/5
1987.02	10.8938	17.3113	4.40396	8.20355	0.0002686	3.90726	7	2	2/1/4	24	4	8/10/6
1989.07	9.38798	13.4327	3.11536	6.03838	0.0013886	3.38018	0	0	0/0/0	8	2	1/3/4
1993.03	6.25783	7.71587	1.15045	2.03977	0.0008823	3.52188	0	0	0/0/0	10	2	1/3/4
2003.01	6.90726	10.1168	1.30656	3.52314	1.34E-05	4.90422	0	0	0/0/0	7	2	4/2/1
2006.01	7.11286	11.6075	1.48333	6.64755	0.000486	3.82439	2	2	0/0/2	13	3	5/5/3
2044.01	7.25269	9.7212	1.53884	3.50695	0.0005917	3.68215	0	0	0/0/0	9	1	3/3/3
2065.08	6.82965	8.06093	1.25075	2.22014	0.0084678	2.7334	0	0	0/0/0	9	1	3/3/3
2121.04	10.4431	7.72746	2.68567	2.24851	0.0001042	-4.21115	7	3	1/4/2	0	0	0/0/0
2167.13	14.3241	9.78683	6.22877	3.63731	0.0017435	-3.35765	7	3	2/4/1	3	2	2/0/1
2195.14	24.3767	12.6862	8.19156	5.65711	1.12E-07	-6.31546	24	3	8/9/7	20	3	7/6/7
2225.14	229.975	134.612	109.304	109.152	0.0013014	-3.38666	27	3	9/9/9	29	4	11/9/9
2233.17	167.656	46.9804	86.3481	21.0183	9.23E-08	-7.09684	27	3	9/9/9	23	3	8/7/8
2246.17	16.4594	25.1637	7.02665	16.9516	0.0093021	2.71475	20	3	8/7/5	25	4	7/10/8
2255.14	42.1485	17.7824	19.5739	7.22485	7.49E-07	-6.14443	21	3	6/8/7	13	4	4/5/4
2269.2	26.3969	16.6138	10.0547	6.68881	8.18E-05	-4.349	7	3	4/3/0	8	2	2/3/3
2277.12	24.6783	13.1394	11.4348	4.47791	2.24E-05	-4.95064	7	3	3/1/3	3	2	2/0/1
2283.19	149.657	34.7215	75.3971	26.0205	1.58E-08	-7.5711	27	3	9/9/9	31	4	11/11/9
2297.2	92.3103	52.4811	46.6451	45.2308	0.0014327	-3.35714	27	3	9/9/9	27	4	9/8/10
2300.27	88.6015	50.1078	42.9161	29.2708	0.0002519	-3.98275	20	3	5/8/7	16	3	5/6/5
2305.2	38.2403	12.3083	18.0571	4.25516	5.59E-08	-7.30299	26	3	9/9/8	5	2	3/2/0
2319.23	23.4308	13.0863	10.4184	4.97491	3.36E-05	-4.74734	22	3	6/9/7	10	3	4/3/3
2419.32	9.552	14.9041	2.79127	5.01204	2.41E-06	5.28028	9	3	3/5/1	23	4	8/8/7
2678.33	9.44347	26.5593	3.64945	11.8849	8.65E-10	7.93921	13	3	5/5/3	6	3	1/1/4
2691.3	6.24636	36.9592	1.39736	27.8121	2.67E-07	6.4289	0	0	0/0/0	24	3	8/8/8
2706.3	9.28678	14.714	3.60576	8.01543	0.0008982	3.54426	6	3	0/2/4	9	3	2/4/3
2748.32	9.08826	28.9299	2.88675	21.1841	4.98E-06	5.3987	14	3	2/4/8	24	3	8/8/8
2766.06	9.14353	16.7684	2.21578	8.5206	1.99E-07	6.31633	2	2	0/1/1	3	1	1/0/2
2780.41	12.8917	46.8548	8.41174	38.3278	1.34E-05	5.01702	11	3	3/8/0	29	4	10/6/13
2792.76	15.9486	24.607	5.21167	12.0516	0.0004566	3.76888	13	3	1/6/6	5	2	1/2/2
2839.36	8.00334	11.4132	2.66019	4.62892	0.0006691	3.60983	3	1	2/1/0	7	2	3/3/1
2921.35	9.81981	24.5824	3.0698	25.0492	0.0017062	3.40439	16	3	5/5/6	20	3	8/7/5
2969.46	4.28688	9.68306	0.991791	6.60686	4.03E-05	4.69628	0	0	0/0/0	4	2	1/0/3
3076.56	8.94529	17.2618	5.03761	7.93412	6.52E-06	4.97174	14	3	6/5/3	23	3	8/8/7
3094.61	7.90327	16.9166	3.56821	5.65689	3.61E-10	7.5832	10	3	4/4/2	20	3	8/7/6
3223.15	6.11531	7.73976	1.2673	2.02518	0.0003274	3.82769	5	3	1/2/2	2	1	0/1/1
3346.68	16.7817	25.311	7.94749	11.7148	0.0013132	3.37793	14	3	6/5/3	26	4	9/8/8
3353.74	17.1876	26.6098	8.78887	12.3615	0.0009674	3.47591	4	2	2/1/1	7	3	1/1/5
3393.84	6.1185	10.3087	2.30376	4.05225	4.78E-06	5.08309	0	0	0/0/0	8	2	2/3/3
3493.82	4.12957	6.52261	1.42734	2.23361	4.45E-06	5.07675	0	0	0/0/0	14	2	4/5/5
3651.9	4.09495	6.62707	1.79813	2.85198	8.78E-05	4.2262	0	0	0/0/0	10	3	3/4/3
3795.87	14.9262	31.9072	9.1474	22.0212	0.0001793	4.07537	9	2		8	2	2/3/3
3809.85	17.1109	42.4032	11.9236	29.938	4.77E-05	4.49737	10	2	1/3/6	13	4	3/5/5
3970.95	2.56797	4.16099	1.10994	1.76872	7.00E-05	4.29388	0	0	0/0/0	2	1	0/0/2
4263.15	2.60469	4.53088	1.28554	1.54405	1.72E-06	5.31524	0	0	0/0/0	2	1	1/0/1

Ave= average; 1) = number of peptides observed in all spectra; 2) number of patients or controls that contributed; subsamples in x/x/x = measured in the first, the second and the third measurement.

Table 5. Differences in peptides of stroma from placentas from early onset preeclampsia (n=4) and preterm delivery (n=3).

Ave mass	Ave case	Ave control	SD case	SD control	p-value	t-test	Spectra controls 1)	Controls 2)	Subsamples control	Spectra cases 1)	Patient cases 2)	Subsamples cases
262.988	13.8738	2.71653	10.7176	1.15289	1.71E-06	5.84894	0	0	0/0/0	10	3	4/3/3
590.024	61.2175	41.2819	34.2643	22.5479	0.0098543	2.67554	7	3	2/4/1	8	3	2/4/2
614.051	25.1373	47.0355	9.44988	39.1856	0.0083272	-2.83506	20	3	8/7/3	7	4	5/2/0
620.296	31.9927	19.2892	16.6779	11.6128	0.0011367	3.43376	3	2	2/1/0	13	4	7/3/3
656.023	154.434	44.7987	139.774	35.0254	0.0001344	4.28068	24	3	9/6/9	32	4	11/11/10
659.167	36.4743	20.1289	16.5093	10.1047	2.21E-05	4.6608	4	2	2/2/0	14	3	4/6/4
665.986	55.5578	31.0713	42.3998	20.2903	0.0057428	2.89731	10	3	3/4/3	20	4	5/6/7
672.304	26.2096	17.3044	12.6069	6.54895	0.0010815	3.47817	0	0	0/0/0	4	3	1/2/1
681.122	103.104	32.7945	71.5925	15.777	4.93E-06	5.40218	19	3	8/4/7	31	4	11/10/9
706.05	61.8267	25.7221	35.7525	10.0457	3.44E-06	5.46272	13	3	8/2/3	29	4	10/10/9
713.813	25.6543	19.5254	10.5602	5.67433	0.0066606	2.83386	4	3	4/0/0	6	2	2/2/2
731.125	59.2565	37.6978	26.1338	20.8652	0.0008521	3.52206	19	3	8/6/5	28	4	10/9/9
741.223	18.6731	37.7905	7.80142	25.1625	0.0006595	-3.79684	15	3	5/3/7	3	2	1/1/1
776.333	20.9718	27.6288	6.0691	8.64863	0.0015776	-3.3617	4	2	4/0/0	1	1	0/2/0
796.324	19.0124	22.853	3.73905	6.1466	0.0070715	-2.83425	7	2	1/4/2	0	0	0/0/0
850.075	32.8469	18.0857	15.7007	8.54301	3.22E-05	4.5758	7	3	3/3/1	25	4	9/8/8
856.6	33.3832	20.4727	17.349	8.8306	0.0005897	3.65212	10	3	1/7/2	22	4	8/7/7
861.118	28.1182	20.055	11.5713	9.27223	0.0043501	2.97045	9	3	1/7/2	24	4	9/8/7
867.279	21.6574	13.013	12.5614	2.57365	0.0005718	3.79956	0	0	0/0/0	9	2	2/4/3
870.586	45.1729	29.866	22.3439	14.947	0.0027899	3.13259	21	3	7/8/6	32	4	12/11/19
877.092	32.0511	15.8613	24.0935	8.23402	0.0009815	3.56257	4	3	2/2/0	18	4	5/7/6
886.333	18.9041	13.8576	8.52964	3.77335	0.0042417	3.0154	1	1	0/0/1	5	2	2/2/1
892.292	20.971	11.4263	12.1759	2.54343	0.0001257	4.32401	0	0	0/0/0	9	3	3/4/2
895.302	17.1768	12.1348	6.14244	2.62382	0.0001263	4.20961	0	0	0/0/0	3	2	2/1/0
1004.6	17.8569	9.70526	14.7818	1.39701	0.0040114	3.10318	0	0	0/0/0	8	1	2/3/3
1045.59	62.1096	162.651	44.8452	107.291	6.75E-05	-4.5474	26	3	8/9/8	30	4	10/10/10
1282.63	10.1684	7.47772	3.3926	1.04673	0.0001333	4.25296	0	0	0/0/0	6	1	0/3/3
1247.85	9.95604	13.3618	3.33148	5.42244	0.0069078	-2.84225	11	3	2/4/5	2	2	0/1/1
1442.75	7.63388	12.3484	3.00514	3.74956	3.05E-06	-5.26141	9	2	1/5/3	0	0	0/0/0
1505.83	15.9917	7.28597	15.2648	1.46867	0.0030471	3.20885	9	1	3/3/3	0	0	0/0/0
1567.8	11.6434	7.17592	7.99041	1.60837	0.0039846	3.08944	9	1	3/3/3	0	0	0/0/0
1621.99	7.31068	10.5596	2.2698	3.85692	0.0004107	-3.85041	6	2	1/2/3	1	1	0/1/0
1774.89	7.15505	9.43972	2.7908	2.39548	0.0012998	-3.3836	3	2	0/1/2	0	0	0/0/0
1833.97	23.3924	12.94	11.2512	6.68514	5.22E-05	4.41255	11	3	3/2/6	17	4	6/7/4
1912.08	8.64263	6.98318	2.85834	1.50498	0.0064143	2.84927	0	0	0/0/0	5	1	0/3/2
1922.97	16.0702	11.9959	6.23042	4.46409	0.0050903	2.91685	11	3	4/3/4	25	4	8/9/7
1926.01	14.1119	11.0599	4.89031	3.79585	0.0092116	2.69649	3	2	0/2/1	3	3	0/2/1
1969.01	10.2268	7.99624	2.9029	2.67845	0.0033183	3.06668	0	0	0/0/0	10	2	4/4/2
1987.01	19.0747	11.1726	11.8143	4.73482	0.0012222	3.46788	8	1	2/3/3	15	3	5/4/6
2002.99	10.9827	7.22838	5.42922	1.97252	0.000863	3.60135	6	2	0/0/6	9	2	3/3/3
2009.04	9.62746	7.92852	2.40804	2.04619	0.0048719	2.92974	0	0	0/0/0	8	1	3/2/3
2121.02	7.42291	10.8348	2.18169	2.71602	3.14E-06	-5.25238	6	3	0/4/2	0	0	0/0/0
2167.17	8.49596	14.0443	2.96242	5.32158	2.17E-05	-4.82356	12	3	3/5/4	0	0	0/0/0
2195.16	11.3508	22.3006	4.78999	8.99456	1.54E-06	-5.68224	24	3	9/8/7	8	3	2/4/2
2225.11	101.838	194.542	72.3586	99.1678	0.0002009	-4.03502	25	3	9/7/9	23	3	8/8/7
2229.25	21.0833	34.21	11.9151	15.9005	0.0009247	-3.5335	26	3	9/8/8	30	4	9/11/10
2233.27	43.223	131.011	17.5229	86.3174	1.66E-05	-5.19514	26	3	9/8/9	22	3	8/8/6
2255.27	16.7614	34.234	7.18119	19.6326	0.0001188	-4.38364	23	3	8/8/7	6	3	3/1/2
2269.3	14.9051	24.0354	5.71852	12.022	0.0009136	-3.61617	19	3	5/7/7	13	3	5/6/2
2277.19	11.812	21.7456	4.6842	10.5847	7.00E-05	-4.51758	14	3	4/5/5	4	1	2/2/0
2283.24	26.9137	144.125	15.6262	80.565	4.52E-08	-7.44249	26	3	9/8/9	28	4	8/10/10
2297.37	39.8225	96.8226	28.4447	54.8085	1.99E-05	-4.878	25	3	9/8/8	30	4	9/11/10
2300.33	36.4582	85.3599	18.7976	48.0605	2.03E-05	-4.97572	17	3	7/7/3	8	3	2/2/4
2305.17	11.2787	34.437	4.34527	19.0142	1.04E-06	-6.19366	23	3	8/8/7	1	1	0/0/1
2319.29	13.3122	24.2634	5.48272	12.8219	0.0002226	-4.1308	18	3	8/6/4	5	2	2/2/1
2419.31	15.4153	11.3734	5.09731	4.78704	0.0027181	3.13646	9	2	2/4/3	17	4	7/5/5
2678.36	22.4772	10.4965	8.23497	4.66681	5.69E-09	7.0042	5	2	3/1/1	18	4	4/8/6
2691.33	28.6942	9.76988	20.2647	3.2826	1.02E-05	5.20241	0	0	0/0/0	21	3	9/6/6
2748.33	22.3689	10.5837	16.0865	4.5916	0.0003317	3.95762	13	3	2/5/6	20	3	9/5/6
2808.42	95.0629	187.118	82.2599	153.602	0.008087	-2.7943	25	3	8/8/9	24	3	9/8/7
2969.55	13.9863	5.86417	9.55017	1.84316	4.16E-05	4.70817	10	2	4/2/4	0	0	0/0/0
3094.68	18.1253	9.28129	8.22074	4.93244	4.96E-06	5.09509	8	3	3/3/2	14	3	7/2/5
3353.61	23.7099	16.6219	8.68048	9.25527	0.0039127	3.01484	3	2	1/2/0	17	3	5/7/5
3392.82	8.82184	6.04125	3.31654	2.25673	0.0003473	3.81597	0	0	0/0/0	7	1	3/1/3
3510.83	6.38991	3.92034	2.48897	0.98284	6.46E-06	5.1564	0	0	0/0/0	6	2	2/1/3
3651.99	6.42466	4.33235	2.10185	1.36311	2.63E-05	4.61142	3	2	3/3/2	10	3	4/3/3
3957.03	5.27971	3.26215	1.80184	1.05737	2.16E-06	5.33779	0	0	0/0/0	8	2	4/3/3
3972.03	6.70117	2.92008	3.90731	1.01906	6.69E-06	5.26596	0	0	0/0/0	10	2	2/3/3

Ave= average; 1) = number of peptides observed in all spectra; 2) number of patients or controls that contributed; subsamples in x/x/x = measured in the first, the second and the third measurement.

1.2.3. Off-line NanoLC MALDI measurements

Using the Ultimate NanoLC system and MALDI TOF/TOF measurements resulted in the identification of four masses of differentially expressed proteins that were confirmed by FT-MS (a tolerance of 4 ppm was allowed using external calibration). Subsequently, we focused on proteins specific and discriminating for early onset preeclampsia only (neither detectable in controls nor late onset preeclampsia). Significantly differentially expressed proteins include choriomammotropin A precursor (LCHUC), fragment OTTHUM (Q5T6S7), surfeit locus protein 4 (O15260) and calcyclin (P06703) (Table 6).

Table 6. Identified and confirmed significantly differentially expressed proteins in early onset preeclampsia placentas.

Ave Mass MALDI-Tof	Calc. mass	[prot_desc]	Score	Matched peptides	MW	FTMS	ppm
1269.760	1269.6194	choriomammotropin A precursor [validated] (LCHUC)	66	4	25004	1269.6142	4
1910.990	1910.7899	OTTHUMP00000018488 (Fragment) (Q5T6S7_HUMAN)	20	3	28076	1910.7928	1.6
2044.010	2043.9127	Surfeit locus protein 4 (O15260)	67	2	30374	2043.9056	3.5
891.410 915.560	891.4240 915.4894	Calcyclin (Prolactin receptor associated protein) (PRA) (Growth factor-inducible protein 2 (P06703))	51	2	10173	891.4233 915.4898	0.8 0.4

Calcyclin, was verified by immunostaining, for the other proteins no commercial antibodies were available yet. Immunohistochemistry showed for calcyclin positive staining for trophoblasts from placenta of women with early onset preeclampsia compared with controls (Fig. 7). Table 7 represents the expression levels of calcyclin in the early onset preeclampsia, control and preterm group. By immunohistochemistry, it was evident that choriomammotropin (using an antibody that is not specific for the precursor) is present both in trophoblast of preeclampsia and control placental tissue.

Table 7. The contents of the S100A6 (calcyclin) in the early onset preeclampsia, control and preterm group performed by immunohistochemical staining.

no.	study samples	expression level (S100A6) in trophoblast
1	early onset PE	++
2	early onset PE	+
3	early onset PE	+
4	control	-
5	control	+/-
6	control	+/-
7	preterm control	-
8	preterm control	-

Samples used for immunohistochemistry were not taken from the test cohort used in the mass spectrometry experiments. – no staining in trophoblasts; +/- faint cytoplasma staining in trophoblasts; + and ++ strong staining in trophoblast cells.

1.3. Conclusion

This Example describes different tryptic peptide patterns of villous trophoblast and stroma cells from placentas of women with pregnancies complicated by preeclampsia (cases) compared to uncomplicated pregnancies (controls) in a small number of 125 microdissected cells. We subsequently found that these peptide patterns are different in placentas from early onset preeclampsia (< 34 weeks gestation) compared to a late onset of the disease (> 34 weeks gestation. In order to determine whether this is an effect of preterm delivery per se, we compared the peptide patterns of women with early onset preeclampsia with women who had a pregnancy complicated by preterm delivery of unknown cause. These data confirmed that the peptides patterns are specific for early onset preeclampsia (as compared to late onset preeclampsia as well as controls) and are not the effect of preterm delivery. We identified as much as possible proteins in a representative tissue section of the three groups (preeclampsia, controls, preterm deliveries) by MSMS sequencing and confirmed the identification by accurate mass FT ICR MS measurement. Four proteins correspond to the differentially expressed peptides by accurate

FT-MS measurement. By immunohistochemistry we could confirm that calcyclin is significantly different (black white difference) in expression. We used 8 paraffine sections of the early onset preeclampsia (n=3), control (n=3) and preterm (n=2) group for the immunohistochemical staining.

5 Calcyclin, S100A6, is a member of the S-100 calcium-binding protein family (12). Recently, it has been described that calcyclin stimulates secretion of choriomammotropin (placental lactogen). A dose-dependent release in isolated mice trophoblasts of placental lactogen in response to calcyclin was described by independent researchers. Uterine natural killer (NK) cells were 10 the primary site of calcyclin RNA messengers and calcyclin is expressed in natural killer cells throughout pregnancy. Uterine natural killer cells have been implicated in trophoblast function. The present finding is further supported by the reported reduced natural killer cell proportion primarily in pregnancies complicated by impaired placental development (trophoblast 15 infiltration and placental growth) such as preeclampsia. Furthermore, calcyclin secretion is further controlled by cytokines such as interleukin-1. Interleukin-1 is spontaneously synthesized during preeclampsia. This implies a specific and different mechanism in early onset preeclampsia through calcyclin and choriomammotropin. Although, we observed by 20 imunohistochemistry that choriomammatropin is present in both the preeclamptic and control trophoblasts, the differences in the expression of the choriomammatropin precursor as judged by our proteomics experiment suggests some involvement in the pathophysiology of this metabolic pathway.

Surfeit locus protein 4 is another peptide which is differentially 25 expressed in placentas from women with early onset preeclampsia. The surfeit locus genes contain a tight cluster of mammalian genes. The surfeit locus 4 gene is encoding an integral membrane protein most likely associated with the endoplasmic reticulum. No association with placentation or human development has hitherto been described.

Recently, it was reported that elevated clusterin levels could be found in serum from women with preeclampsia. Clusterin was identified as over expressed spots on two-dimensional electrophoresis with MALDI-TOF-MS followed by Western blot analysis. Using placenta samples in line with the hypothesis that the responsible trigger for the disease is mostly likely located within the placenta, we could not confirm that clusterin is involved in the pathophysiology of preeclampsia.

Overall, early onset preeclampsia results more often in severe maternal and fetal complications although it is not an all or nothing phenomenon. Recent longitudinal studies of placental growth by ultrasound have revealed that in cases of early onset preeclampsia the placenta volume is reduced from as early as 12 weeks of gestation. Calcyclin has been detected also in early pregnancy during implantation in mice. Moreover, in the first trimester the endovascular trophoblast invasion - and subsequent conversion of the spiral arteries and release of the endovascular trophoblastic plugs allowing intervillous blood flow- is disturbed in early onset preeclampsia. By contrast, in cases of late onset preeclampsia, placental volume is slightly larger than normal at 12 weeks gestation and the placental insult seems less clear as villous growth is normal. This suggests a different underlying mechanism of early and late onset preeclampsia. In addition, we firstly found specific peptides in early onset preeclampsia in trophoblasts that were striking different from stroma cells. This suggests that abnormal placentation is not due to disturbed trophoblast metabolism only but that an abnormal interaction between villous trophoblast and villous stroma cells may be involved as well. In the near future, developments in clinical proteomics in terms of sensitivity, reproducibility, resolution and accuracy will help us to use this method for the analysis of small numbers of trophoblasts and stroma cells obtained in early pregnancy to predict subsequent preeclampsia in the second half of pregnancy.

In conclusion, in the present Example, specific peptides of placentas from women with pregnancies complicated by early onset preeclampsia are

described. Specific peptide patterns were found, using small numbers of cells, in villous trophoblast and their corresponding proteins could be identified being choriomammotropin precursor, calcyclin and surfeit locus protein 4.

5 **Example 2. Identification of glioma neovascularization-related proteins by using MALDI-FTMS and nano-LC fractionation to microdissected tumor vessels**

2.1. Material and methods

2.1.1. Sampling

10 Ten fresh-frozen samples of glioblastoma located in the cerebral hemispheres and 10 samples of normal control hemispheric brain were taken form the files of the Department of Pathology, Erasmus MC, Rotterdam (Table 8).

Table 8: Clinical data

Sample ID.	Sex	Age	Tumor location
G1	m	57	Ri F
G2	m	57	Le T
G3	m	55	Ri F
G4	m	51	Ri F
G5	m	51	Le T
G6	m	48	Le F
G7	m	47	Ri O
G8	m	36	Le P
G9	m	32	Bi F
G10	f	30	Ri F
Sample ID	**Sex**	**Age**	**Cause of death**
N1	f	76	pneumonia
N2	f	62	Cirrhosis + hepatocellular carcinoma
N3	m	62	Ischemic cardiac disease
N4	f	60	nasopharyngeal carcinoma
N5	m	48	SAB / aneurysm
N6	f	48	SAB / aneurysm
N7	f	39	SAB / aneurysm
N8	m	34	Brain stem abscess
N9	m	28	hypertensive stroke
N10	m	24 weeks	intra-uterine infection

G samples indicate Glioma patients
N samples indicate control patients

5 Sections of 5 mm from each sample were made, counterstained and examined by the neuropathologist (JMK) to verify the presence of proliferated tumor vessels (Figure 9). The control samples of normal brains were subjected to the same procedure for the identification of the blood vessels.

10 2.1.2. Laser Capture Microdissection

Cryosections of 8 mm were made from each sample and mounted on polyethylene naphthalate (PEN) covered glass slides (P.A.L.M. Microlaser Technologies AG, Bernried, Germany) as described previously [Umar, A., et

al., Proteomics, 2005. 5(10): p. 2680-8]. The slides were fixed in 70% ethanol and stored at -20°C for not more than 2 days. After fixation and immediately before microdissection, the slides were washed twice with Milli-Q water, stained for 10 seconds in haematoxylin, washed again twice with Milli-Q water
5 and subsequently dehydrated in a series of 50, 70, 95 and 100% ethanol solution and air dried. The P.A.L.M. laser microdissection and pressure catapulting (LMPC) device, type P-MB was used with PalmRobo v2.2 software at 40x magnification. Estimating that a cell has a volume of 10x10x10mm, we microdissected an area of about 190,000 mm2 of blood vessels and another
10 area of the same size of the surrounding tumor tissue from each sample, resulting in approximately 1,500 cells per sample. A total of 40 samples were collected, viz., 10 glioma vessels, 10 fields of glioma tissue surrounding the glioma vessels, 10 normal vessels and 10 fields of normal tissue surrounding the normal vessels. As a negative control, a corresponding area of the PEN
15 membrane only was microdissected and analysed in the same way as the other samples. This negative control experiment was performed in three folds.

The microdissected cells were collected in the caps of P.A.L.M. tubes in 5 ml of 0.1% RapiGest buffer (Waters, Milford, MA, USA). The caps were cut and placed onto 0.5 ml Eppendorf protein LoBind tubes (Eppendorf, Hamburg,
20 Germany). Subsequently, these tubes were centrifuged at 12,000 g for 5 minutes. To make sure that all the cells were covered with buffer, another 5 ml of RapiGest was added to the cells. After microdissection, all samples were stored at -80°C.

25 2.1.3. Sample Preparation

After thawing the samples, the cells were disrupted by external sonification for 1 minute at 70% amplitude at a maximum temperature of 25°C (Bransons Ultrasonics, Danbury, USA). The samples were incubated at 37°C and 100°C for 5 and 15 minutes respectively, for protein solubilization and
30 denaturation. To each sample, 1.5 ml of 100 ng/ ml gold grade trypsin

(Promega, Madison, WI, USA) in 3 mM Tris-HCL diluted 1:10 in 50 mM NH4HCO3 was added and incubated overnight at 37°C for protein digestion. To inactivate trypsin and to degrade the RapiGest, 2 ml of 500 mM HCL was added and incubated for 30 minutes at 37°C. Samples were dried in a speedvac
5 (Thermo Savant, Holbrook, NY, USA) and reconstituted in 5 ml of 50% acetonitrile (ACN) / 0.5% trifluoroacetic acid (TFA) / water prior to measurement. Samples were used for immediate measurements, or stored for a maximum of 10 days at 4°C.

10 2.1.4. MALDI-FTMS Measurements and Data Analysis

2.1.4.1. MALDI-FTMS measurements

Samples were spotted onto a 600/384 anchorchip target plate (Bruker Daltonics, Leipzig, Germany) in duplicate. Half a microliter of each sample was mixed on the spot with 1 ml of a 2,5-dihydroxybenzoic acid (DHB)
15 matrix solution (10 mg/mL in 0.1% TFA) / water and the mixture was allowed to dry at ambient temperature. The MALDI-FTMS measurements were performed on a Bruker Apex Q instrument with a 9.4 T magnet (Bruker Daltonics, Bremen, Germany). For each measurement, 450 scans of 10 shots each were accumulated with 60% laser power. Mass spectra were acquired in
20 the mass range of 800 to 4,000 Da. FTMS spectra were processed with a Gaussian filter and 2 zero fillings.

2.1.4.2. MALDI-FTMS external and internal calibration

A standard peptide calibration mix (Bruker Daltonics, Leipzig,
25 Germany) which contains angiotensin I and II, substance P, Bombesin, Renin Substrate, ACTH clip 1-17, ACTH clip 18-39 and Somatostatin 28 was used for external calibration. To obtain better mass accuracies, an additional post-acquisition internal calibration step in DataAnalysis v3.4, built 169 software (Bruker Daltonics, USA) was performed. Ubiquitous actin peptide masses (m/z
30 1198.70545, 1515.74913, 1790.89186, 2215.06990 and 3183.61423) were used

for internal calibration. To assess the accuracy of the measured masses, the peptides derived from keratin [Q8N175] present in the samples were compared to the calculated masses (1165.58475, 1234.67896, 1365.63930, 1381.64814, 1390.68085, 1707.77211, 1797.01161 and 2096.04673).

2.1.4.3. Data Analysis

Mono-isotopic peaks with S/N > 3 were annotated with the SNAP algorithm using the pre-release version of DataAnalysis software package (v3.4, built169). The peak lists were saved in a general text format, which was used as an input for a home made script in the R-program, (www.r-project.org). With this script a matrix file was generated, indicating the presence or absence of each peptide mass in the different mass spectra. If a specific peptide appeared at least in 5 samples for each group and never appeared in the other groups, it was considered as a group specific peptide. In this way, a list of differentially expressed peptides was generated. These masses of the differentially expressed peptides were submitted to the MASCOT search engine (Matrix Science, London, UK) using the SWISS-PROT (40.21) database, allowing 1 ppm peptide mass tolerance and one missed trypsin cleavage site. In addition, we performed Hierarchial Clustering based on masses and the group of samples using the matrix file in the Spotfire software (Spotfire, Somerville, MA, USA).

2.1.5. Sample Preparation for Nano-LC

Sample G8 was selected for fractionation (Table 8). One, 4 and 8 frozen sections were made, respectively. These sections from the entire tumor sample including the vessels were prepared as described above. Each section contained about 2,000,000 cells of which an estimated 10% were blood vessel derived cells. Twenty ml RapiGest buffer was added (Waters, Milford, MA, USA) to the frozen sections followed by 1 minute sonification, 5 minutes at 37°C and finally 15 minutes at 100°C. For each section 1 ml of 100 ng/ ml gold

grade trypsin (Promega, Madison, WI, USA) in 3 mM Tris-HCL was added and samples were incubated overnight at 37°C. Finally, 50 mM HCL was added. For comparison, 8 sections from normal brain sample N5 were prepared in exactly the same way.

5 In addition, an area of about 900,000 mm2 of blood vessels from each of the glioma samples and the normal control samples were microdissected and pooled, resulting in one sample of glioma blood vessels and one sample consisting of control blood vessels. Pooling of the samples was necessary because the nano-LC procedure requires far more tissue than obtained by 10 microdissection. Twenty ml RapiGest buffer was added and the samples were stored at about 80°C. All samples were subjected to the nano-LC fractionation immediately after preparation.

2.1.6. Fractionation by Nano-LC

15 Fractionation was performed using a C18 Pep Map column (75 mm i.d. x 150 mm, 3mm, Dionex, Sunnyvale, CA, USA). Five ml of the sample was loaded onto the trap column (300 mm i.d. x 5mm, 5mm, Dionex, Sunnyvale, CA, USA). Fractionation was performed for 130 minutes with a gradient of buffer A (100% H2O, 0.05% TFA) and buffer B (80% ACN, 20% H2O and 0.04% 20 TFA); 0 to 15 min, 0% buffer B, 15.1 min 15%, 75 min 40%, 90 min 70%, 90.1-100 min 95%, 100.1 min 0% and 130 min 0%. Fifteen second fractions of the sample were spotted automatically onto 384 prespotted anchorchip plates (Bruker Daltonics, USA) containing a- cyano-4-hydroxycinnamic acid (HCCA) matrix, using a robotic system (Probot Micro Fraction Collector, Dionex, 25 Sunnyvale, CA, USA). To each fraction, 1 ml water was added. Finally, we used a 10mM (NH)4H2PO4 in 0.1% TFA/water solution to wash the pre-spotted plate for 5 seconds to remove salts. The plates were subsequently measured by automated MALDI-TOF/TOF (Ultraflex, Bruker Daltonics, Germany) using WARLP-LC software. MS spectra of each individual spot were 30 obtained. Spots and peptide masses for performing MS/MS measurements

were determined automatically by the WARLP-LC software. A file containing the MS and the MS/MS peak lists was submitted to the MASCOT search engine (Matrix Science, London, UK) using the SWISS-PROT (40.21) database allowing 150 ppm parent mass tolerance, 0.5 Dalton fragments tolerance and
5 one missed trypsin cleavage site. In addition, identification was confirmed by exact mass measurements on the MALDI-FTMS, adding 1 mL DHB solution to the fractionated spot and allowed to dry.

2.1.7. Backward database searching

10 By in silico digestion of the identified proteins, theoretical peptides were generated which were sought for in the monoisotopic peaks of the MALDI-FTMS.

The (UniProtKB/Swiss-Prot) accession number for all of the identified proteins was entered into the peptide cutter program
15 (www.expasy.org/tools/peptidecutter), choosing trypsin as enzyme for digestion and allowing one trypsin missed cleavage site. All the possible tryptic fragments from each protein were compared with the peptide masses obtained by MALDI-FTMS within 0.5 ppm (the internal calibration). The distribution of the matched peptides over the four groups was checked manually.

20

2.1.8. Immunohistochemical staining:

Specific expression of fibronectin and colligin 2 proteins in glioma blood vessels was confirmed by immunohistochemistry, using specific antibodies against each protein on the paraffin sections of the same samples of
25 glioma and normal brain . To investigate the expression variation between the two groups, an additional six samples of glioma and four samples of normal brains were examined.

Immunohistochemical staining was performed following the manufacture procedure (alkaline phosphatase technique), using rabbit
30 polyclonal antibody for fibronectin at a 1:1,000 dilution (DakoCytomation,

Glostrup, Denmark) and mouse monoclonal antibody for colligin 2 at a 1:500 dilution (Stressgene, Victoria, B.C., Canada). Five mm paraffin sections were mounted to poly-L-lysin coated microslides, deparaffinized in xylene for 15 minutes and rehydrated through graded alcohol, then washed with water. The
5 sections were washed with phosphate-buffer saline (PBS) and incubated with the antibody for 30 minutes. After washing the sections with PBS, the corresponding antigen was added and incubated 30 minutes at room temperature. New Fuchsin Alkaline Phosphatase Substrate Solution was freshly prepared and the sections were incubated for about 30 minutes.
10 Afterwards, the sections were washed with tap water, counterstained and cover-slipped with permanent mounting medium.

2.2. Results

2.2.1 FTMS measurements

15 The MALDI-FTMS measurements of the microdissected samples yielded approximately 700 – 1,100 monoisotopic peaks for almost all spectra. Only one glioma vessel and one normal tissue sample contained less than 100 peaks. However, these spectra were not excluded from our analysis. An accuracy of 3 ppm was obtained by external calibration using a standard
20 peptide calibration mix. After internal calibration the accuracy increased below 0.5 ppm.

2.2.2. FTMS data analysis

Following our strict criteria, a list of 16 differentially expressed
25 peptides was obtained (Table 9). All 16 peptides were expressed in the glioma vessel group only. The MASCOT database search resulted in matching of four out of the 16 peptides to fibronectin precursor protein [P02751]. In order to exclude that matching of the four peptides to fibronectin was just by chance, the following database searches were performed. We added the integers 10, 11,
30 12, until 30 Daltons to the masses of the 16 peptides which were found for 20

additional searches. By this procedure no proteins were found to match by chance with four peptides. At maximum, only one peptide matched to one protein in the MASCOT database. This virtually ruled out the possibility of randomly finding fibronectin.

5

Table 9. List of differentially expressed peptides

Peptides (measured masses)	Number of samples in which these peptides were found:			
	Glioma vessels	Glioma surrounding tissue	Normal brain vessels	Normal brain surrounding tissue
1926.04620 *	8	0	0	0
2470.32072 *	6	0	0	0
1593.81172 *	5	0	0	0
1807.90584 *	5	0	0	0
1535.72354 *	5	0	0	0
2257.07971 *	5	0	0	0
1659.80041 *	5	0	0	0
1275.55961 *	5	0	0	0
1731.89535 *	5	0	0	0
1116.54323	5	0	0	0
1849.85488	5	0	0	0
2089.00769	5	0	0	0
2157.10653	5	0	0	0
2164.00992	5	0	0	0
2530.25829	5	0	0	0
2642.21770	5	0	0	0

* Peptides resulted in protein identification.

Figure 10 shows the result of the unsupervised cluster analysis in
10 two directions; peptide masses and groups of samples in the Spotfire program. A cluster of eight glioma vessel samples is observed. From the two samples which did not cluster, one had a poor spectrum (<100 peaks); this sample clustered with the sample from normal tissue at the top of the heat map which also displayed a poor spectrum. The other one did not cluster with any group.
15 Within the peptide masses, a specific pattern of glioma blood vessels is recognized.

2.2.3. Nano-LC Fractionation / MALDI-TOF-MS/MS

Pooling small number of cells collected by microdissection before
20 nano-LC fractionation resulted in the identification of some high abundant

proteins, among which fibronectin. To identify more proteins, we increased the number of cells by using whole sections of glioma and normal samples. The number of identified peptides was increased and the maximum was reached with the injection of eight sections (Table 10). The capacity of the nanoLC
5 column did not allow further expansion of the number of sections. Fractionation of eight sections led to the significant identification of 189 proteins, with a minimum mowse score of 24 for MS/MS.

Table 10. Results for the various numbers of sections used for fractionation in
10 the nano-LC:

Sample type	Glioma 1 section	Glioma 4 sections	Glioma 8 sections	Normal brain 8 sections	Glioma microdiss. cells[a]	Normal brain microdiss. cells[a]
No. of MS measurements	2307	3328	3383	2985	552	779
No. of MS/MS measurements	734	1194	2160	1752	368	416
No. of identified proteins	32	131	189	140	27	13

[a] 15,000 microdissected cells

The data obtained from MALDI-TOF/TOF after the fractionation procedure were compared to the MALDI-FTMS data, searching specifically for
15 the 16 differentially expressed peptides. Nine out of 16 peptides matched within 200 ppm. To obtain a higher mass accuracy for the peptides, the corresponding spots of these nine peptides were re-measured in the MALDI-FTMS. The exact mass of five out of nine peptides matched within 3 ppm (external calibration) with the masses originally obtained by FTMS. In order to
20 relate these peptides to proteins, the MS/MS data of these peptides were searched for in the database, resulting in a significant matching of four of them (sequence score > 24). Two peptides matched to fibrinogen beta chain precursor [p02675], one peptide to colligin 2 [P50454] and one peptide to acidic calponin 3 [Q15417]. In the MALDI-TOF data set more peptides belonging to
25 these proteins were sought and an additional three peptides belonging to

fibrinogen beta chain precursor, and two belonging to colligin 2 protein, were found. We also found an additional 17 peptides from fibronectin, of which nine had a significant MS/MS score.

5 2.2.4. Backward database searching

The search of the peak list obtained from the In silico digestion of fibronectin sequence in the FTMS data resulted in the finding of six extra peptides. Five peptides were found in the glioma vessels group only, and one was also seen in one sample of the normal brain blood vessels (Table 11). The 10 same search for the In silico digestion of fibrinogen yielded nine additional peptides of which three were exclusively found in the glioma vessels group and the others in one sample of the normal vessels (Table 12). Searching for the theoretical peptides of colligin 2 and acidic calponin3 did not result in the finding of any extra peptide.

15

Table 11. Differentially expressed Fibronectin precursor [P02751] peptides.

Fibronectin peptides found in FTMS spectra	Exact fibronectin peptide masses	Δ ppm	Number of samples in which these peptides were found:			
			Glioma vessels	Glioma surrounding tissue	Normal brain vessels	Normal brain surrounding tissue
1926.04620 [a]	1926.04833	1.11	8	0	0	0
2470.32072 [a]	2470.31874	0.80	6	0	0	0
1593.81172 [a]	1593.81188	0.05	5	0	0	0
1807.90584 [a]	1807.90471	0.63	5	0	0	0
1629.87232 [b]	1629.87070	0.99	4	1	0	0
2692.37550 [b]	2692.37292	0.97	4	0	0	0
1349.68509 [b]	1349.68481	0.21	3	0	0	0
1401.66582 [b]	1401.66582	0.01	3	0	0	0
2524.36562 [b]	2524.36567	0.03	3	0	0	0
3042.59234 [b]	3042.58942	0.96	3	0	0	0

[a] Peptides matching the criteria used in this Example.
[b] Peptides derived from *in silico* digestion.

Table 12. Peptides derived from in silico digestion of fibrinogen

Fibrinogen peptide masses found in the FTMS spectra	Exact fibrinogen masses derived from *in silico* digestion	Δ ppm	Number of samples in which these peptides were present in:			
			Glioma vessels	Glioma surrounding tissue	Normal brain vessels	Normal brain surrounding tissue
1032.56252	1032.5625	0.02	5	0	0	0
1239.51764	1239.5177	0.05	5	0	0	0
2385.17568	2385.1754	0.12	4	0	0	0
1275.55961	1275.5600	0.3	4	1	0	0
1544.69498	1544.6950	0.01	3	1	0	0
1668.71478	1668.7151	0.2	3	1	0	0
886.38736	886.3876	0.3	2	1	0	0
1951.00371	1951.0031	0.3	2	1	0	0

2.2.5. Immunohistochemistry:

The specific expression of the fibronectin and colligin 2 proteins in glioma blood vessels were confirmed by immunohistochemistry. For all the samples, the proliferated blood vessels present in the glioma samples were immunopositive for fibronectin and colligin 2 while the blood vessels in the control brain samples remained negative (Figures 11 and 12). In some capillaries of normal brain some fibronectin was expressed, but to a far lesser extent as the expression observed in the proliferated glioma vessels. The blood vessels for arachnoid were immunopositive for fibronectin, not for colligin 2.

2.3. Conclusion

In this Example it was attempted to identify angiogenesis-related proteins in glioma in the surgically removed specimens of patients suffering from glial tumors. To achieve this goal, relevant cell populations had to be targeted. Like all tissues, tumors consist of complex 3-dimensional structures of heterogeneous mixture of cell types. Laser microdissection provides an efficient and accurate method to obtain specific cell populations like glioma blood vessels in the present study. The hypertrophied vessel walls of glioma vasculature consist of endothelial cells, pericytes and cells expressing smooth

muscle actin. In addition, these vessels may also contain glial tumor cells (mosaic vessels). In order to eliminate proteins derived from these tumor cells, we also microdissected glial tumor tissue for comparison. Any peptide present in the blood vessels that was also found in the glioma tissue was eliminated
5 from the list of differentially expressed peptides. Therefore, comparison of the various microdissected tissues is essential for targeting structure-specific proteins.

Application of MALDI-FTMS holds significant advantages over that of other types of mass spectrometry. FTMS provides very high mass accuracies
10 and its ability to perform an internal calibration increases the accuracy considerably. In the present study we achieved an accuracy of ± 3 ppm by external calibration and up to ± 0.5 ppm by internal calibration. One of the advantages of MALDI-FTMS is the very high mass resolution, which in the present study generated relatively complex spectra, consisted 700-1,100 mono
15 isotopic peaks per spectrum. Yet, another advantage is the very high sensitivity of the FTMS, which is higher than any other mass spectrometric technique currently available. In addition, FTMS provides an excellent signal-to-noise ratio, since the source of noise in MALDI-FTMS is of physical origin and not a chemical based noise as generated in the MALDI-TOF. These
20 advantages allow studying very small numbers of targeted cells.

The MALDI-FTMS measurements of microdissected samples enabled us to detect a specific peptide pattern for the distinct targeted cell populations, but the results were not adequate to directly identify all of their related proteins. The chance of identifying a protein based on accurate peptide
25 masses rises by increasing the number of peptides generated and detected from that protein. The number detectable peptides per protein depend on some factors: the size of the protein, the chemical properties of both the protein and the derived peptides, the relative concentration of a protein and the enzyme used in digestion. Last but not least, protein identification by detection of
30 peptides relies highly on the accuracy and completeness of available databases.

In the present study we succeeded to identify the protein fibronectin based on the precise masses of four peptides generated by MALDI-FTMS.

The *in silico* digestion approach, appeared to be a valuable tool to confirm the presence of peptides derived from a specific proteins in the spectra obtained by MALDI-FTMS. The high peptide mass accuracy of MALDI-FTMS facilitates the match with the calculated masses generated by *in silico* digestion. Nevertheless, the nature of a protein, its concentration and its ionization ability still play major roles in the detection of peptides.

The complexity of the sample in combination with a relative low sensitive for MS/MS in FTMS on MALDI ions complicates the identification of peptides based on direct MS/MS measurements. To reduce those effects, we applied nano-LC fractionation prior to MALDI-TOF/TOF. Because the number of cells required for nano-LC fractionation is much higher than what is obtained from sample microdissection, we pooled the microdissected cells from all samples resulting in one sample of 15,000 cells. However the loss of sample during preparation steps and in the nano LC column is still considerable. In addition the overall sensitivity of MALDI-TOF measurements is considerably less compared to MALDI-FTMS. These factors together resulted in the identification of only the high abundant proteins of the pooled microdissected cells. The identification of lower abundant proteins can be achieved by using more cells however the microdissecting approach is then not longer feasible. The tryptic digest of whole sections allowed the identification of many more proteins in both glioma and normal brain samples, particularly when we used peptide concentrations close to the maximum capacity of the column (eight sections). Within the spectra that were generated by MALDI-TOF following nano-LC, we specifically sought the peaks that were previously identified by FTMS, i.e. the 16 differentially expressed peptides. The low percentage of vessels, which is at maximum 10% of the cells in a section, resulted in producing low number of peptides from their specific proteins. The detection of vessels specific peptides probably was masked by the detection of the high

percentage peptides derived from the surrounding tissue. For that reason not all the 16 differentially expressed peptides found in the MALDI-FTMS experiments were observed after fractionation followed by MALDI-TOF/TOF. Yet, MS/MS data of four peptides were obtained and their identification was
5 based on both, very accurate peptide masses and their significant MS/MS measurements. Importantly, fractionation also increased the number of peptides generated from a single protein, thus improving the confidence in the identification significantly (Table 13).

10 **Table 13.** Differentially expressed proteins identified by nano-LC fractionation

Identified protein (Accession no.)	FTMS mass[1]	Calculated mass	Δ ppm (score)	Sequence coverage[2]	Sequence[3]	Extra peptides identified[4]
Fibrinogen βchain (P02675)	1535.72354	1535.72366	0.13 (52)	13 %	AHYGGFTVQNEANK	5
	2257.07971	2257.08046	0.35 (48)		GGETSEMYLIQPDSSVKPYR	
Colligin 2 (P50454)	1659.80041	1659.80126	0.54 (39)	6 %	LYGPSSVSFADDFVR	1
Acidic calponin (Q15417)	1275.55961	1275.56000	0.31 (39)	3 %	YDHQAEEDLR	0

[1]Specific peptide masses by FTMS (pre-fractionation)
[2]Sequence coverage of protein
[3]Sequence obtained after nano-LC fractionation and MALDI-TOF/TOF measurements
15 [4]No. of extra peptides identified after nano-LC fractionation

The validation step performed for two of the three glioma vasculature specific proteins by immunohistochemistry showed this specific expression. Moreover, fibronectin antibody showed a very faint staining in
20 some of the normal brain blood vessels; this can explain the detection of one fibronectin peptide by mass spectrometry in the normal brain vessels. The colligin 2 antibody visualized only the glioma vessels. The very good correlation between a diagnostic technique such as immunohistochemistry and mass spectrometry highlight the sensitivity and the accuracy of mass
25 spectrometry and opens possibilities of using it in the diagnostic field.

In the present Example, fibronectin, fibrinogen, colligin 2 and acidic calponin 3 were identified as proteins which are specifically expressed in the glioma vasculature. Fibronectin is a high molecular weight, multifunctional matrix protein which binds to other extracellular matrix proteins such as
5 collagen, fibrin and heparin. Several studies addressed the relation between fibronectin and tumors among which breast cancer, melanoma [Berube, M., et al. Int J Oncol, 2005. 26(2): p. 405-13; Schor, S.L. and A.M. Schor. Breast Cancer Res, 2001. 3(6): p. 373-9], and also gliomas. Fibronectin was found to be differentially expressed in a study using suppression subtractive hybridization
10 in which pilocytic astrocytoma were compared to glioblastoma [Colin, C., et al. Oncogene, 2006. 25(19): p. 2818-26]. The glioblastomas expressed fibronectin while the pilocytic astrocytomas did not. In addition, overexpression of fibronectin in glioblastoma as detected by immunohistochemistry was reported previously [Caffo, M., et al. Acta Neurochir (Wien), 2004. 146(10): p. 1113-8].
15 The present finding of the expression of fibronectin by glioma blood vessels suggested that this protein plays a role in the development of glioma vasculature.

Colligin-2, also called Heat shock protein-47 is a collagen binding protein that is associated with an increase in the production of procollagen in
20 human vascular smooth muscle cells [Rocnik, E.F., et al. J Biol Chem, 2002. 277(41): p. 38571-8]. In the literature, colligin-2 was related to angiogenesis in oral squamous cell carcinomas [Nikitakis, N.G., et al. Am J Clin Pathol, 2003. 119(4): p. 574-86]. Acidic Calponin, also identified in this study, is a thin filament associated protein detected in a number of different cells and tissues.
25 It has been reported among the differentially expressed proteins in human glioblastoma cell lines and tumors [Zhang, R., et al. Glia, 2003. 42(2): p. 194-208]. Acidic calponin modulates the contraction of smooth muscle cells. Interestingly, the common function of the proteins found in the present study is their prominent role in cell motility. It may very well be that the
30 identification of these proteins is a reflection of their upregulation in glioma

vasculature. During neoplastic angiogenesis sprouting of pre-existent blood vessels may urge an increased motility of the activated endothelial cells involved in the process. Further, the putative influx of angiogenic precursor cells from the bone marrow into glioma may require the activation of motility even more. Further studies may detail the function and interaction of the proteins found in glioma vessels wall in this study.

Example 3. Multiple sclerosis related proteins identified in CSF by advanced mass spectrometry

3.1. Experimental procedures

3.1.1. Patient selection

All samples analyzed in this study were taken from patients that were followed prospectively by the Rotterdam Multiple Sclerosis Center and the department of Neurology at Erasmus University Medical Center (Rotterdam, the Netherlands). These samples were selected and classified into four groups by an experienced neurologist. The first designated group consisted of CSF samples from patients suffering from Multiple Sclerosis (MScl, n=44). The second group consisted of samples from patients who were diagnosed with a clinically isolated syndrome of demyelination (CIS, n=40), which is considered as a pre-stage of MScl. The third group of samples was taken from patients with another inflammatory neurological disease (OIND, n=26), to be used as controls. The fourth group, which was also a control group, consisted of samples from patients with another neurological disease (OND, n=54). All these diagnoses were based on data derived from the examination at the time the CSF sample was taken.

Immediately after sampling, the CSF samples were centrifuged to discard cellular elements (10 minutes at 3000 rpm). The samples were subsequently used for routine CSF diagnostics. This included quantification of total protein and high abundant protein (albumin and immunoglobulin G) concentration, assessment of the number of oligoclonal bands, as well as

quantification if the intrathecal cell count. The remaining volume of the samples was aliquoted and stored at -80°C, where they remained until sample preparation for this study.

5 3.1.2. Sample preparation

Prior to all sample preparation procedures, all samples were blinded. They were subsequently measured and analyzed in a random order. Twenty µl of each CSF sample was put into a 96-microtiter well plate (Nunc. Low binding, VWR, the Netherlands), and an equal amount of 0.2% Rapigest
10 (Waters, USA) in 50 mM ammonium bicarbonate buffer was added to each well. Following a two-minute incubation period at 37°C, 4 µl 0.1 µg/µl gold grade trypsin (Promega, USA) / 3mM Tris HCl (pH 8.0) was added to each well. The samples were incubated at 37°C for two hours. To obtain a final concentration of 30-50 mM HCl (pH < 2), 2 µl of 500 mM HCl was added. The
15 samples were then incubated for 45 minutes at 37°C, which stopped the digestion reaction. This last step also broke down the detergent, which is important for preventing interference peaks in the mass spectrum.

Following the digestion procedure, the samples were desalted using 96-well zip C18 micro titer plates (Millipore, USA), which had been pre-wetted
20 and washed twice with 100 µl acetonitrile per well. The samples were centrifuged (Multifuge 3 S-R, Goffin Meyvis, the Netherlands) at 2000 rpm for five minutes. After the washing step, 3 µl acetonitrile were put on the C18 resin to prevent drying. Each trypsin digested CSF sample was mixed with 200 µl HPLC grade water / TFA 0.1%. The samples were subsequently put on the
25 washed and prewetted 96-well zip C18 plate, and 30 minutes of centrifugation at 1500 rpm was used to load the peptides onto the C18 material. The wells were washed twice with 100 µl 0.1% TFA (5 minutes, 1750 rpm). An elution volume of 15 µl 50% acetonitrile/ 0.1% TFA was used to elute the samples in a new 96-well plate. The centrifugation period for the elution step was 30

minutes at 1600 rpm. After this elution step, the samples were stored in 96-well plates, which were covered with aluminum seals, at 4°C.

3.1.3. Measurement

5 A matrix solution was made by dissolving 2 mg α-cyano-4 hydroxy-cinnamic acid (HCCA) in 1 ml acetonitrile, using an ultrasonic bath for 30 minutes. Two µl of elute of each sample was mixed with eight µl of the matrix solution, 0.5 µl of which was spotted onto a MALDI target (600/384 AnchorChipTM with transponder plate, Bruker Daltonics, Germany). All 10 samples were spotted in duplicate. The digestion step and the subsequent measurement in duplicate were performed three times for each sample, resulting in 6 spectra per sample. The samples were all measured using the automated measurement feature of the MALDI-TOF MS (Ultraflex, Bruker Daltonics, Germany). The standard method for peptide measurements 15 (Proteomics_HPC) was used on the MALDI-TOF MS, with the measurement range set to 500-3400 Dalton (Da).

The following settings were used for the automated measurements: the initial laser power was 25%, and the maximum was 45%. The peak with the highest intensity above the 750 Da had to have a signal-to-noise ratio of at 20 least 5 and a minimum resolution of 5000. Every 50 laser shots the sum spectrum was checked for these criteria. It was rejected if it did not meet these criteria. If 20 sum spectra of 50 laser shots met these criteria, they were combined and saved.

If 25 consecutive sum spectra of 50 shots had been rejected, the 25 measurement of the spot was ended and the measurement of the next spot commenced.

3.1.4. Analysis

The raw binary data files were first converted to ASCII files 30 containing the measured intensities for all channel indices of the spectra. All

spectrum files were designated a group number (1 – MScl, 2 – CIS, 3 – OIND, 4 – OND) and serial numbered. To calibrate the channel numbers to masses we used a quadratic fit with a number of internal calibrants, which were five omnipresent tryptic albumin peptides (927.4934, 1226.6051, 1467.8430, 1875.0156 and 2045.0953). The algorithm described by Dekker and co-workers (6) was used for peak detection, performed in the statistical language R (http://www.r-project.org). A percentile threshold of 96% was chosen, meaning that the intensity of the peak position must belong to the 4% highest intensity values of the spectrum. The mass window (minimum distance between two adjacent peak positions) was set at 0.5 Da. After the generation of a peak list for each spectrum using the peak detection algorithm, an analyses matrix was created in R. During this process all peak lists of every sample were checked for the presence of all peaks. Thus, a peak position that was present in all peak lists of one sample, was designated with a 6 and a peak position present in half of the peak lists was designated with a 3. In this matrix, all samples were tagged with their group number and with a number ranging from 1 to 6 for all separate peak positions. The matrix was subsequently used for statistical analysis of the data. Using a univariate analysis in R, a p-value was determined for every peak position. The Wilcoxon-Mann-Whitney test was used for comparison between the groups. A cross validation was performed on the same data by randomly assigning a group number to each CSF sample and then repeating the Wilcoxon-Mann-Whitney test. This scrambling procedure was subsequently repeated 10000 times, which gives an impression of the probability to find a significant differentially expressed peptide by chance.

3.1.5. Identification

The next step in the process was identification of the differentially expressed peptides. Due to the limitations of the MALDI-TOF strategy, the differentially expressed peptides could not be identified immediately from the complete CSF sample. The amount of peaks present was simply too much, so it

was not possible to find a mass window to do tandem mass spectrometry. So, for identification of the peptides alternative methods were needed. We used two different strategies. Firstly, we determined the accurate masses of the peptides with a 9.4 Tesla Fourier Transform Mass Spectrometer (FTMS, Bruker Daltonics, Germany). Due to the excellent mass accuracy and precision of this technique it is uniquely suited to determine the masses of the peptides up to an accuracy of better than 1 ppm. Because the HCCA matrix is not compatible with orthogonal MALDI (Apex I, Bruker Daltonics), we used another matrix molecule, 2,5-dihydroxy benzoic acid (DHB).

Secondly, we used a nano-LC system (Ultimate system, Dionex, USA) to separate the digested CSF peptides on a C18 reverse phase column. The digested CSF sample was injected and separated on this column during a 130 minute automated LC run, and subsequently spotted on a MALDI target plate, which was also done in an automated sequence. The MALDI target plate used in this nano-LC experiment was a pre-spotted target plate (PAC 384 plate, Bruker Daltonics, Germany). On this plate the matrix was already applied during the manufacturing process and only the sample has to be added. By dividing a single sample into 384 separate spots and corresponding spectra, it proved to be possible to acquire the required mass windows for several of the differentially expressed peptides. Measurement of the pre-spotted plate was done in an automated way, using the LC-warp software package (Bruker Daltonics, Germany), which first generated a mass spectrum for all spots and then proceeded to generate MS/MS spectra, in a data dependent manner, for peaks which were suitable for MS/MS by a series of pre-set conditions, such as, for example, a sufficiently high signal-to-noise ratio. Identifications were obtained by database searches of the MS/MS spectra using the mascot website and the SwissProt database. Identifications were confirmed by determination of accurate mass by MALDI-FTMS.

In a separate effort to identify the differentially expressed peptides we also used an Orbitrap mass spectrometer with electrospray ionization

(Thermo Electron, Bremen, Germany), using a C18 column nano-LC system online connected to the device. A small volume of the digested CSF was injected into the nano-LC and subsequently the digested peptides were separated on the C18 column. After ionization the peptides were measured in the orbitrap, using a data-dependent acquisition mode for the MS/MS identification step.

3.1.6. Immunoassay

The chromogranin A concentrations in CSF of MScl patients and controls was determined by external validation. To this end we used a separate group of MScl (n = 19) and control (OND) samples (n = 18) which included among others cancer, intracranial hypertension, Sneddon syndrome and headaches. We used a commercially available ELISA kit for Chromogranin A (DakoCytomation, Denmark) as specified by the manufacturer.

3.2. Results

3.2.1. Clinical information

Of the 44 patients in the MScl group, 30 patients had relapsing remitting (RR) MScl and 14 had primary progressive (PP) MScl (table 14). The mean age of the patients in this group was 42.1 ± 11.3 years, and the mean protein concentration of the CSF was 0.41 ± 0.13 g/l. The group contained 12 males and 32 females. The median of the time that these patients had been afflicted with Multiple Sclerosis was nearly two years.

Table 14. Clinical information, including routine protein quantifications, of all patients included in this study.

	MScl (n=44)	CIS (n=40)	OIND (n=26)	OND (n=54)
Males/Females	12/32	12/28	7/19	23/31
% of samples oligoclonal positive	72.7%	80.0%	7.7%	3.7%
Median time between first MScl or CIS related complaint and sampling in months (range)	22 (0-232)	3 (0-72)[#]		
Protein concentration (g/l). SD in brackets	0.408 (0.135)	0.358 (0.106)	0.435 (0.161)	0.422 (0.217)
Albumin concentration (g/l). SD in brackets	0.239 (0.087)	0.210 (0.080)	0.250 (0.160)	0.280 (0.190)
IgG concentration (g/l). SD in brackets	0.060 (0.033)	0.048 (0.029)	0.055 (0.045)	0.043 (0.032)
Diagnoses	- 30 RR MScl - 14 PP MScl	- 25 optic neuritis CIS - 7 myelitis CIS - 5 brainstem CIS - 3 other CIS	- 17 infections (bacterial and sterile) - 7 vasculitis - 2 Guillain Barré Syndrome	- 19 isolated headaches - 11 neurological degeneration - 10 non neurological disease - 8 infarct - 4 brain tumor - 2 other

Excepting two patients (72 and 27 months), all CIS samples were taken within twenty months after the first symptoms were observed.

- Concentrations: p>0.05 for all comparisons (two-tailed t-test), except total protein concentration for CIS-OIND (p=0.026) and IgG concentration for MScl-OND (p=0.035).

The CIS group (n=40) consisted of 12 males and 28 females, nearly 70 percent of whom had been diagnosed with optic neuritis. The other patients were diagnosed with one, or a combination, of the other possible localizations of a CIS, like for example brainstem syndromes or myelitis (7). The mean age of this second group was 33.7 ± 9.4 years, and the mean protein concentration of the CSF was 0.36 ± 0.11 g/l. The median time between the sampling of the CSF and the occurrence of the first symptoms in these patients was three months.

The OIND group (n=26) contained 7 males and 19 females, with a mean age of 49.6 ± 16.3 years. The mean CSF protein concentration was 0.43 ±

0.16 g/l. Diagnoses of the patients in this group were bacterial and sterile infections (n=17), vasculitis (n=7), and Guillain Barré Syndrome (n=2).

The 54 patients in the OND group had a mean age of 48.1 ± 16.7 years. This group contained 23 males and 31 females, and the CSF samples had a mean protein concentration of 0.42 ± 0.22 g/l. The diagnoses set for these patients were headaches (n=19), neurological degeneration (n=11), nonneurological disease (n=10), infarct (n=8), brain tumor (n=4), and other (n=2). The total protein concentrations of the CSF samples did not differ significantly between the groups when subjected to a two-tailed t-test. Also, the albumin and immunoglobulin G (IgG) concentrations of the CSF samples did not differ significantly between the groups.

3.2.2. Peak detection and data analysis

An average of 515 peaks was detected per spectrum. After spectrum conversion, the matrix was created, which consisted of a total of 1755 peaks. The significance of difference in distribution over the four groups was tested for each peak. The significant difference between the MScl group and the OND group ($p<0.05$: 132 peptide peaks; $p<0.01$: 44 peptide peaks) indicates profound differences in the CSF of MScl patients compared to that of the non-inflammatory control patients. The obvious skew towards the low p-values is clearly visible in the p-value histogram (figure 13). The largest statistical difference (figure 14) was observed in the comparison of the CIS group and the OND group ($p<0.05$: 161 peptide peaks; $p<0.01$: 60 peptide peaks).

The comparison of the MScl group with the OIND group (figure 15), and also that of the CIS group with the OIND group (figure 16), shows a number of peak positions with a statistically significant p-value ($p<0.05$: 127 peptide peaks; $p<0.01$:12 peptide peaks, and $p<0.05$: 134 peptide peaks; $p<0.01$: 27 peptide peaks, respectively). Both these p-value histograms are also skewed towards the lower p-values, and, in this region, the height of the bars

clearly exceeds the height of the red line, which reflects the expected statistical background.

Comparison of the MScl group with the CIS group (figure 17) shows few peaks with a significant p-value ($p<0.05$: 50 peptide peaks; $p<0.01$: 10 peptide peaks). The height of the bars in the low p-value area clearly lies below the red line generated by the randomized comparison, indicating that there is no significant difference between the CSF samples of the patients classified in the MScl group compared to the patients in the CIS group. The comparison between the two control groups (OIND vs. OND, figure 18) shows only a slightly higher number of statistically significant peak positions ($p<0.05$: 101 peptide peaks; $p<0.01$: 12 peptide peaks). The height of the bars in the low p-value area lies at about the same height as the red line generated by the randomized comparison, which indicates that these two groups are not statistically significantly different.

3.2.3. Identification

Using the combination of the FTMS and the nano-LC MALDI-TOF/TOF techniques we were able to determine the amino acid sequence of three of the differentially expressed peptides in the comparison between MScl and OND. These peptides were identified to be tryptic peptides of chromogranin A, clusterin and complement C3 by offline nano-LC MALDI-TOF/TOF. Determination of the accurate mass of these three tryptic peptides by FTMS resulted in parent ion masses differing 0.19, 0.27 and 0.67 ppm from the theoretical masses for those specific peptides of chromogranin A, clusterin and complement C3, respectively.

Using the sequencing capabilities of the Orbitrap mass spectrometer we were able to identify 10 peptides from the list of differentiating peptide peaks (table 15). The accurate masses of the parent ions of these identified peptides were further confirmed by FTMS. All showed accuracy below 1 ppm for FTMS.

Table 15. Identified differentially expressed proteins (P<0.01) amongst the patient groups tested.

					MALDI- FTMS		ESI-Orbitrap		
Comparison	Protein	Accession number	Peptide sequence	Theoretical mass	Measured mass	Δ (ppm)	Measured mass	Δ (ppm)	No. of pept.
MScl-OND	Chromogranin A	P10645	YPGPQAEGDSE GLSQGLVDR	2074.9675	2074.9679	0.193	2074.970	1.205	5
MScl-OND	Clusterin	P10909	TLLSNLEEAK	1117.6099	1117.6102	0.268	1117.610	0.089	9
MScl-OND	Complement C3	P01024	AGDFLEANYMN LQR	1641.7689	1641.7678	0.670	-	-	-
MScl-OIND	Complement C4B	P0C0L5	ASAGLLGAHAA AITAYALTLTK	2085.1703	2085.1684	0.911	2085.173	1.295	1
CIS-**OND**	Beta V spectrin	Q9NRC6	WINNVFQC#GQ AGIKIR	1903.9959	1903.9944	0.193	1904.000	2.153	1
CIS-**OND**	Hypothetical protein XP_011125	-	ISHELDSASSEVN	1387.6335	1387.6348	0.937	1387.635	1.081	2
CIS-OIND	Apolipoprotein D	P05090	VLNQELR	871.4996	871.4991	0.574	871.4994	0.229	1
CIS-OIND	Complement C4A	P0C0L4	GLQDEDGYR	1052.4643	1052.4641	0.190	1052.464	0.285	11
CIS-OIND	Contactin 1	Q12860	VQVTSQEYSAR	1267.6277	1267.6288	0.868	1267.628	0.237	4
CIS-OIND	Neuronal pentraxin receptor	O95502	QTALQQEAR	1044.5432	1044.5435	0.287	1044.543	0.191	2
CIS-**OIND**	RNA binding motif protein 7	Q9Y580	VTEELLFELFHQA GPVIKVK	2297.2903	2297.2881	0.958	2297.286	1.871	1

5 - Column 1: Bold denotes group with elevated expression for that specific protein

- 1 missed cleavage allowed

The ELISA for chromogranin A showed that a new set of MScl samples contained an average chromogranin A concentration of 240.4 ± 167.5 U/L, whereas the control samples contained a concentration of 139.8 ± 68.8 U/L. A significantly higher chromogranin A concentration was found in the CSF of MScl patients, as compared to controls (p=0.02, table 16).

Table 16. Chromogranin A concentration (ELISA results) is significantly higher in CSF of MScl patients as compared to controls (p=0.02)

	MScl	**Control (OND)**
Average concentration (in U/L)	240.44	139.87
SD (+/-)	167.53	68.80
Samples from original test set (average concentration in U/L)	193.78	83.39

Four(2 MScl and 2 OND) samples of the original test set were also
5 analysed in this ELISA, and the two MScl samples showed higher concentrations than the two control samples (MScl: 189.9 and 188.7, OND: 63.4 and 103.4).

A main observation in this study was that, on the basis of proteomic CSF analysis, the group of MScl patients could be clearly distinguished from
10 the noninflammatory neurological controls. However, as MScl is considered to be an inflammatory disease, it was imperative to include a group of controls covering other CNS inflammatory neurological diseases. It was also possible to clearly differentiate between the CSF proteomic profiles of the MScl patients and the inflammatory neurological controls. As MScl has both an
15 inflammatory and a neurodegenerative component, comparison with both control groups is essential to ensure that the control groups covered both the inflammatory and the non-inflammatory phase of the disease.

It was also possible to make similar distinctions in the proteomic CSF analyses of the CIS patients when compared to the non-inflammatory
20 controls, as well as in the comparison of the CIS patients with the inflammatory neurological controls.

Another indication of the similarities between the proteomic profiles of the CSF samples of the patients in the MScl group and those of the CIS

group is the overlap of the peptides found in the comparisons of MScl with the non-inflammatory controls (OND) and the comparison of CIS with OND. Eleven of the forty-four peptide peaks (25%) found to be significantly ($p<0.01$) different in the comparison of MScl and OND was also found to be significantly different in the comparison of CIS and OND. If the significance cut-off was set to $p<0.05$, this overlap increases to nearly 38%.

Two major distinct disease courses of MScl, PP MScl and RR MScl, were also compared. Specifically for this sub-comparison a relatively high number of PP MScl patients was included in this study (32% of the MScl samples (normal 10-15% of the total MScl population)). However, any significant differences between PP MScl and RR MScl was not observed (data not shown). Because pathology changes over time in MScl, one could imagine differences in CSF proteomic profiles as the disease duration progresses. The MScl samples were divided into a short and a long disease duration group, according to their position in relation to the median time between first symptoms and CSF sampling. No significant differences were observed in this comparison, compared to the statistical background (data not shown). It has been suggested by others that the course of the disease is more benign in MScl patients without oligoclonal Immunoglobulin G bands in CSF. One might therefore expect to find other differences in CSF composition when comparing MScl patients with oligoclonal Immunoglobulin G bands in the CSF versus MScl patients without these bands. However, thi was not observed. Finally, also no differences were observed between males and females, compared to the statistical background (data not shown). These sub-analyses lend support to the interpretation that common peptide patterns are shared within the whole group of MScl patients.

It is long known that factors such as sample stability and a low number of measurements per sample can cause difficulties regarding the reproducibility of proteomic profiling studies. The low reproducibility of peak height in MALDI-TOF MS was also previously reported. For several reasons it

is believes that the method described here is less affected by these variations. First, the samples were all handled in a standardized way, which included centrifugation immediately after sampling, to clear all cell debris, followed by instant freezing at -80°C. Secondly, the sample preparation method is uncomplicated and straightforward. Thirdly, the height of the peaks is not include in the analysis because quantitative measurements of peak heights with MALDI TOF MS are poorly reproducible, with standard deviations up to 30%. In this study only the absence or presence of the peaks was scored.

To examine the reproducibility of this study, we also separately analyzed the results of the three times the CSF samples were digested and measured in duplicate. Each comparison showed the same p-value diagram for each time we performed the experiment. Similarly, the differentially expressed peptides that were found in the total analysis also had low p-values in the separate analyses. Another beneficial factor is the lower protein load, especially of high abundant. proteins such as albumin in CSF, which allows CSF protein profiling without complicated sample pre-treatment steps (e.g. high-abundant protein depletion). Other studies have already delved into the CSF proteomic profile. Although these studies identified a number of interesting proteins that were present in CSF of MScl patients, these analyses were performed in a setting with limited numbers of patients. The method used in our study also differed from these studies in the fact that our sample pre-treatment was much less complex, as we did not perform a two-dimensional gel electrophoresis procedure prior to the trypsin digestion.

3.3. Conclusion

In conclusion, we identified eleven proteins in association with MScl (table 15). The MALDI-TOF analysis of tryptic digested CSF proteins showed significant differences between MScl patients and control patients (MScl vs. OND: 44 peptides with $p<0.01$), as well as significant differences between CIS patients and control patients (CIS vs. OND: 60 peptides with $p<0.01$). Roughly

10% of all differentially expressed proteins could be identified. Among these proteins were complement C3, complement C4A, complement C4B and apolipoprotein D. Also identified were clusterin, as well as chromogranin A, neuronal pentraxin receptor and contactin 1.

5

Claims

1. Method for identification of disease-related peptides and/or proteins for use as markers in diagnosis, prognosis, or therapeutic monitoring of disease, said method comprising the steps of:

(a) providing an optionally processed sample of a diseased body tissue or fluid as a test sample, and an optionally processed sample of a corresponding healthy body tissue or fluid as a reference sample, wherein said samples comprise peptides and/or proteins;

(b) subjecting both test and reference sample to MALDI- FT-ICR mass spectrometry to generate mass spectra for individual peptides in each sample and to quantify the amount of individual peptides present in each sample;

(c) comparing the amount of an individual peptide present in the test sample with the amount of a peptide having a corresponding mass in the reference sample to generate a list of peptides differentially expressed between test and reference sample, and

(d) subjecting the test and/or reference sample of step (a) to tandem mass spectrometry (MS-MS), in order to identify the differentially expressed peptides and/or the proteins from which they derive thus providing a candidate marker protein or marker peptide.

2. Method according to claim 1, wherein said method further comprises step of:

(e) confirming the presence of said candidate marker protein or marker peptide in said test sample by MALDI-FT mass spectrometry, MALDI Triple-quad mass spectrometry or immunoassay.

3. Method according to claim 1 or 2, wherein said optionally processed samples are body tissue samples processed by subjecting said samples to laser capture microdissection to provide collections of microdissected cells, said collections preferably amounting to about 200- 3,000 cells.

5

4. Method according to claim 3, wherein said collections of microdissected cells are provided in the form of pooled collections of microdissected cells.

5. Method according to any one of the preceding claims, wherein said optionally processed samples are body tissue samples, body fluid samples, or collections of microdissected cells processed by subjection to protein digestion, preferably using trypsin, to provide processed samples comprising peptide fragments from the proteins in said samples.

10

15 6. Method according to claim 5, wherein said method further comprises the step of performing an *in silico* digestion of said candidate marker protein or marker peptide for said disease and comparing the theoretical peptides generated with the monoisotopic peaks obtained for peptides in step (b) or (d).

20 7. Method according to any one of the preceding claims, wherein step (d) comprises the steps of:

- comparing the mass spectrum of a differentially expressed peptide identified by MALDI-FT in step (b) with the mass spectra of the peptides whose sequence is obtained by tandem mass spectrometry in step (d),

25 - deducing the sequence for the differentially expressed peptide identified in step (b) from the sequence of the corresponding peptide determined in step (d),

- matching the sequence deduced for the differentially expressed peptides with a protein database to identify the peptide or protein that is

30 differentially expressed between said test and reference sample, and

- classifying said identified peptide or protein as a tentative marker peptide or protein for said disease;

8. Method according to any one of the preceding claims, wherein prior to tandem mass spectrometry in step (d), said optionally processed sample is subjected to nano liquid chromatography (nano-LC) fractionation to provide fractions comprising separated proteins.

9. Method according to any one of the preceding claims, wherein said body tissue is selected from the group consisting tissues of brain, lung, heart, prostate, esophagus, stomach, jejunum, ileum, caecum, colon, gall bladder, bile duct, breast, ovary, testicle, lymph node, thymus, kidney, liver, muscle, nerve, bone, bone marrow, and placenta.

10. Method according to any one of the preceding claims, wherein said body fluid is selected from the group consisting of blood, serum, cerebrospinal fluid (CSF), urine, saliva and semen.

11. Method according to any one of the preceding claims, wherein said disease is selected from the group consisting of Alzheimer’s disease, Parkinson’s disease, frontotemporal dementia, Lewy body disease, corticobasal degeneration, progressive supranuclear palsy, multiple system atrophy, other neurodegenerative processes, multiple sclerosis, acute disseminating encephalomyelitis (ADEM), neuromyelitis optica (Devic’s disease), other inflammatory diseases of the central nervous system, Guillain-Barre syndrome, other inflammatory diseases of the peripheral nervous system, preeclampsia, glioma, cancer, and inflammation.

12. Method according to any one of the preceding claims, wherein said samples are body fluid samples, preferably a body fluid comprising about 0.05-

5 mg/ml of protein, and wherein in step (b) an amount of 1-10 μl of optionally processed body fluid is subjected to MALDI- FT-ICR mass spectrometry.

13. Method according to any one of the preceding claims, wherein in step
5 (d) said individual peptides are in a mass range of 800 to 4,000 Da.

14. A candidate marker protein or marker peptide for a disease identified by the method according to any one of claim 1-13.

10 15. Marker protein or marker peptide according to claim 14, wherein said disease is early onset preeclampsia and wherein said marker protein is choriomammotropin precursor, calcyclin and/or surfeit locus protein 4.

16. Marker protein or marker peptide according to claim 14, wherein said
15 disease is glioma and wherein said marker protein is fibronectin, fibrinogen, colligin 2 and/or acidic calponin 3.

17. Marker protein or marker peptide according to claim 14, wherein said disease is multiple sclerosis and wherein said marker protein is selected from
20 the group consisting of chromogranin A, clusterin, complement C3, complement C4B, beta V spectrin, hypothetical protein XP_011125, apolipoprotein D, complement C4A, contactin 1, neuronal pentraxin receptor, and RNA binding motif protein 7.

25 18. Method for predicting early onset preeclampsia comprising measuring the expression level of a marker protein selected from the group consisting of choriomammotropin precursor, calcyclin and surfeit locus protein 4 in villous trophoblast from placentas of pregnant female subjects.

19. Method for detecting glioma, comprising measuring the expression level of a marker protein selected from the group consisting of fibronectin, fibrinogen, colligin 2 and acidic calponin 3 in blood, CSF or glioma vasculature.

5 20. Method for detecting multiple sclerosis, comprising measuring the expression level of a marker protein selected from the group consisting of chromogranin A, clusterin, complement C3, complement C4B, beta V spectrin, hypothetical protein XP_011125, apolipoprotein D, complement C4A, contactin 1, neuronal pentraxin receptor and RNA binding motif protein 7 in
10 cerebrospinal fluid samples of patients.

21. Method for monitoring disease activity of glioma and/or the response to a treatment regimen, comprising measuring the expression level of fibronectin, fibrinogen, colligin 2 and/or acidic calponin 3 in blood, CSF and/or glioma
15 vasculature.

22. Method according to any one of claims 18-21, wherein said method comprises the step of performing MALDI Triple-quad analysis of proteins and peptides in a diseased tissue sample to quantify said tentative marker protein
20 or marker peptide for said disease in said diseased tissue sample.

Figure 1

Trophoblast normal	Trophoblast preeclampsia	Trophoblast preterm birth
(n=7) ⟷	(n=7 total) *44/1938* (n=3 late onset) ↕ *38/1681* (n=4 early onset) ⟷	(n=3) 78/1277
Stroma normal	**Stroma preeclampsia**	**Stroma preterm birth**
(n=7) ⟷	(n=7 total) *17/1138* (n=3 late onset) ↕ *57/1482* (n=4 early onset) ⟷	(n=3) *68/1312*

Figure 2

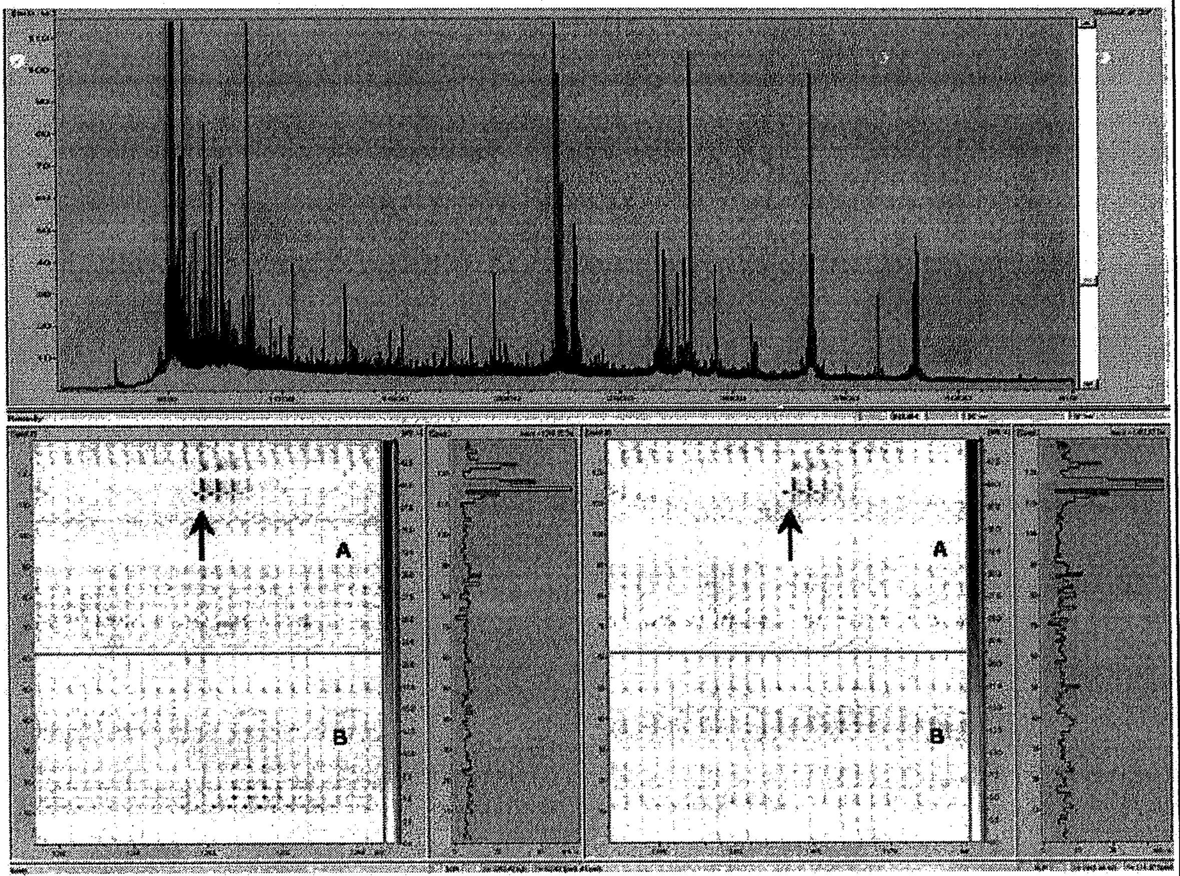


Figure 3

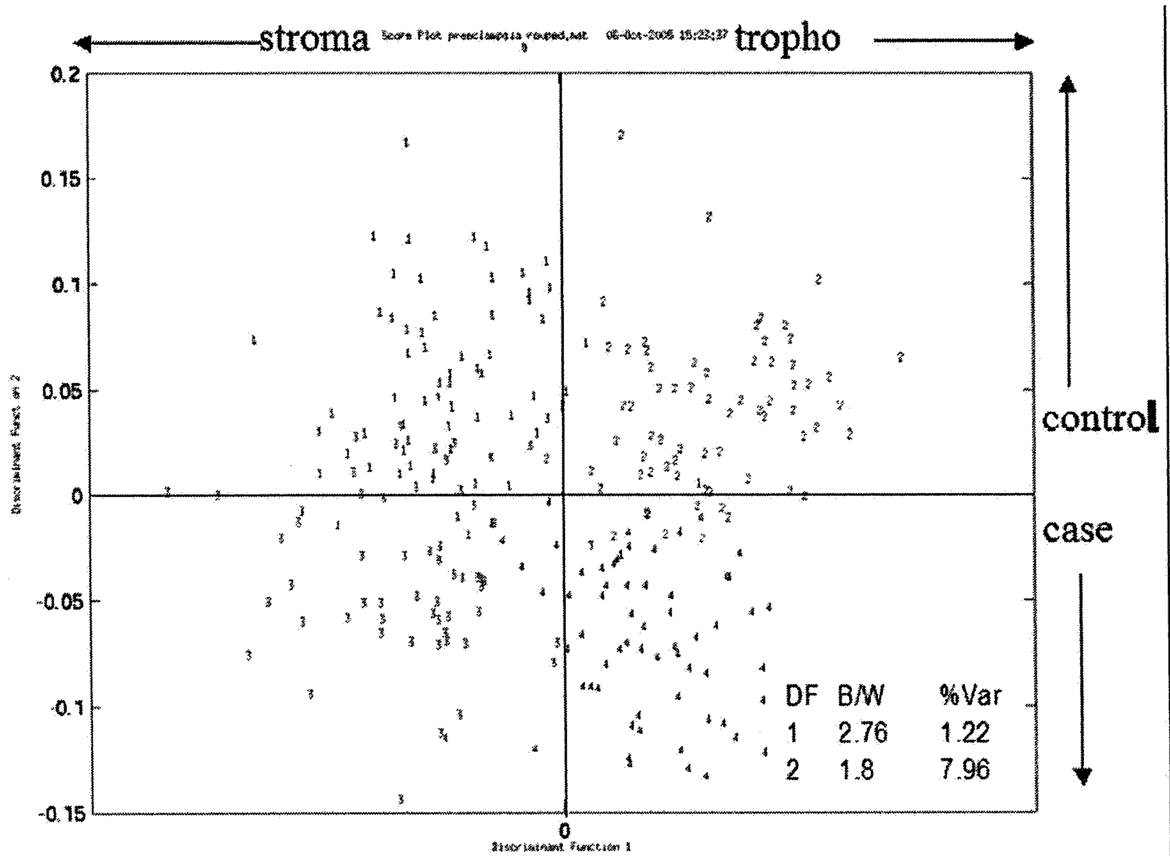
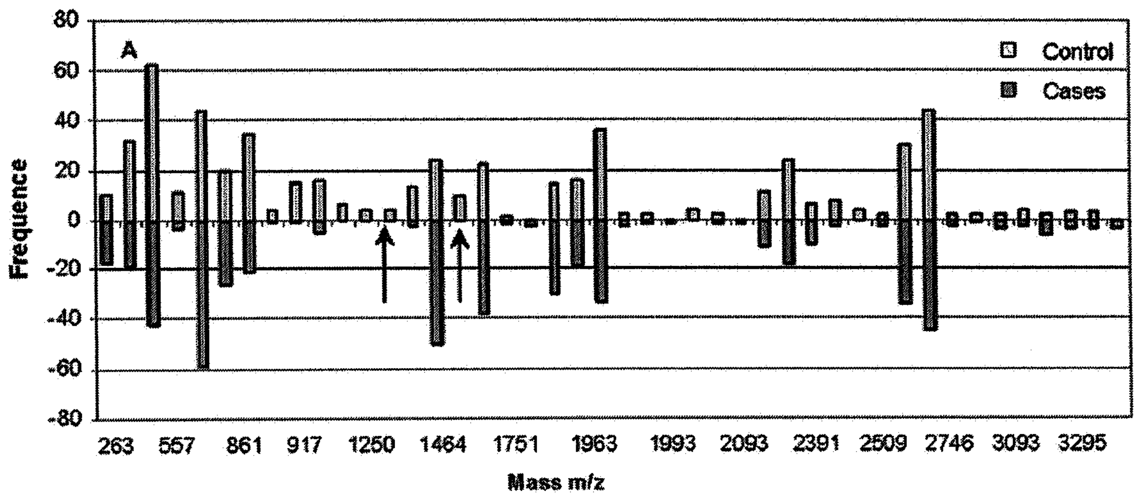


Figure 4



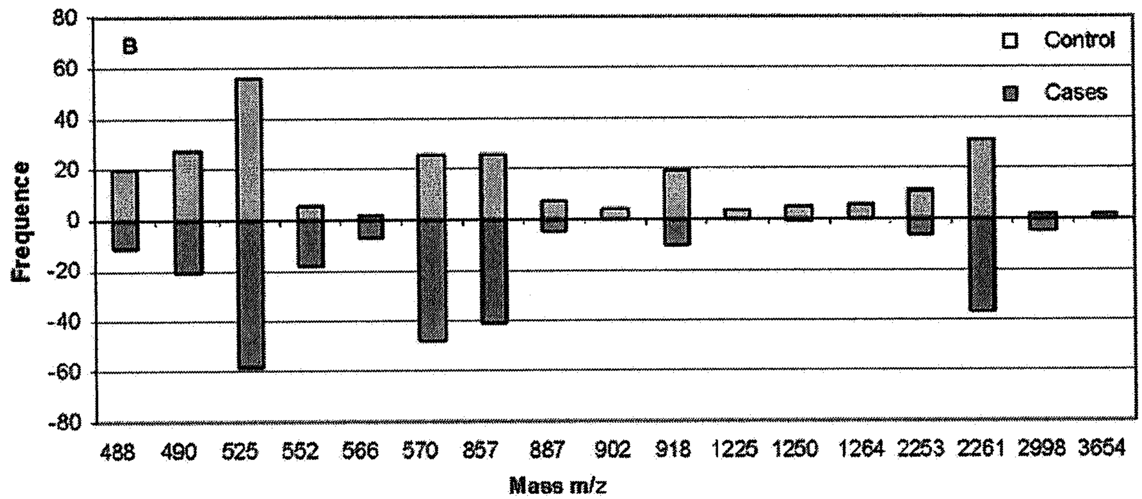


Figure 5

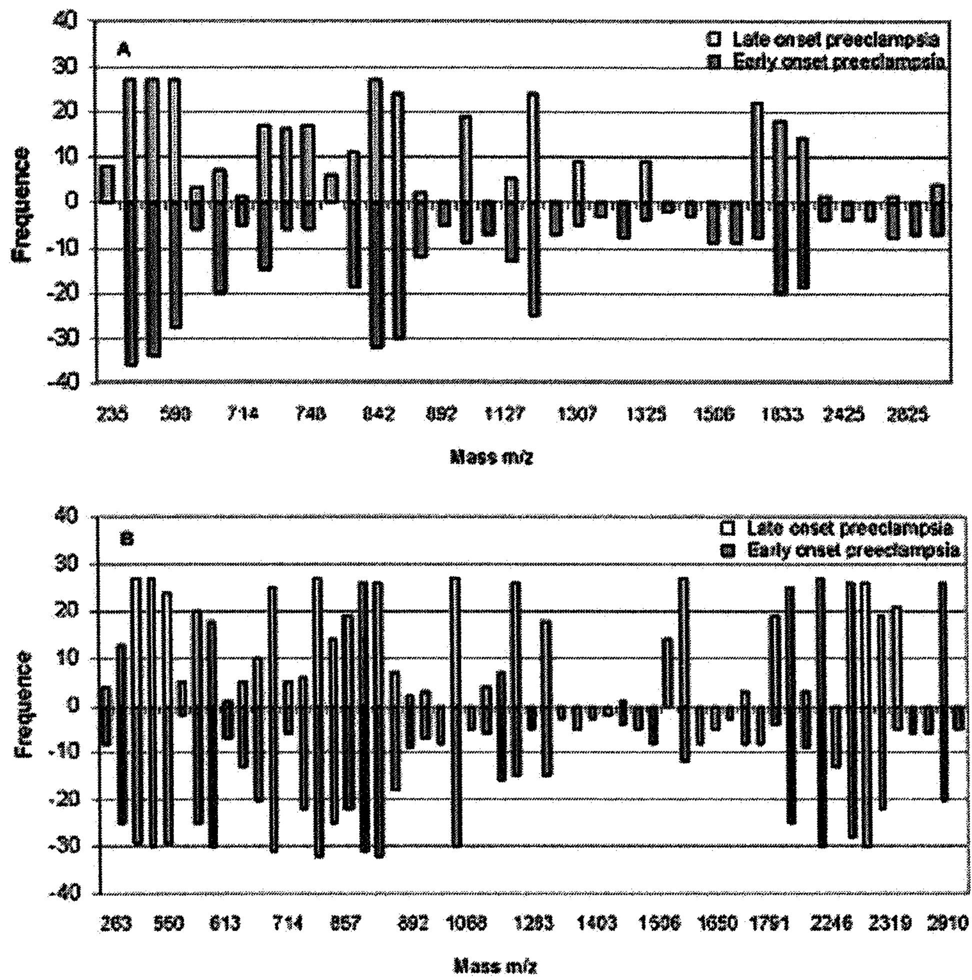


Figure 6

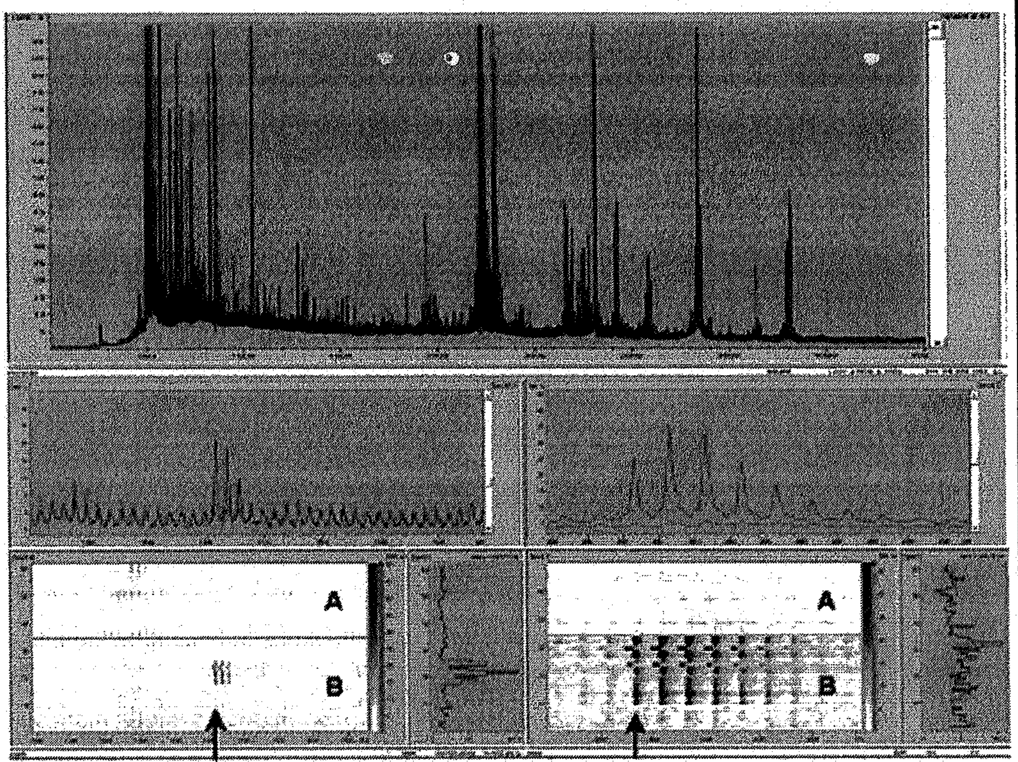


Figure 7

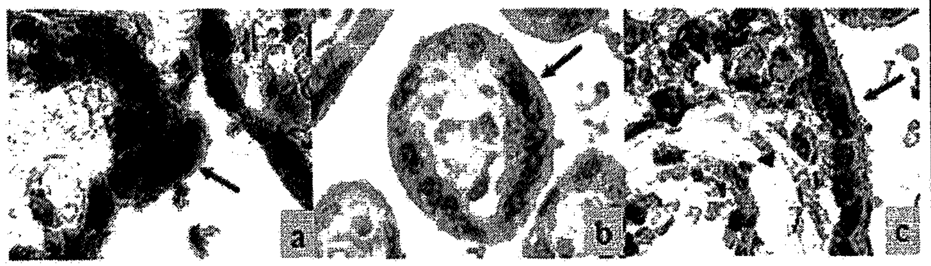


Figure 8

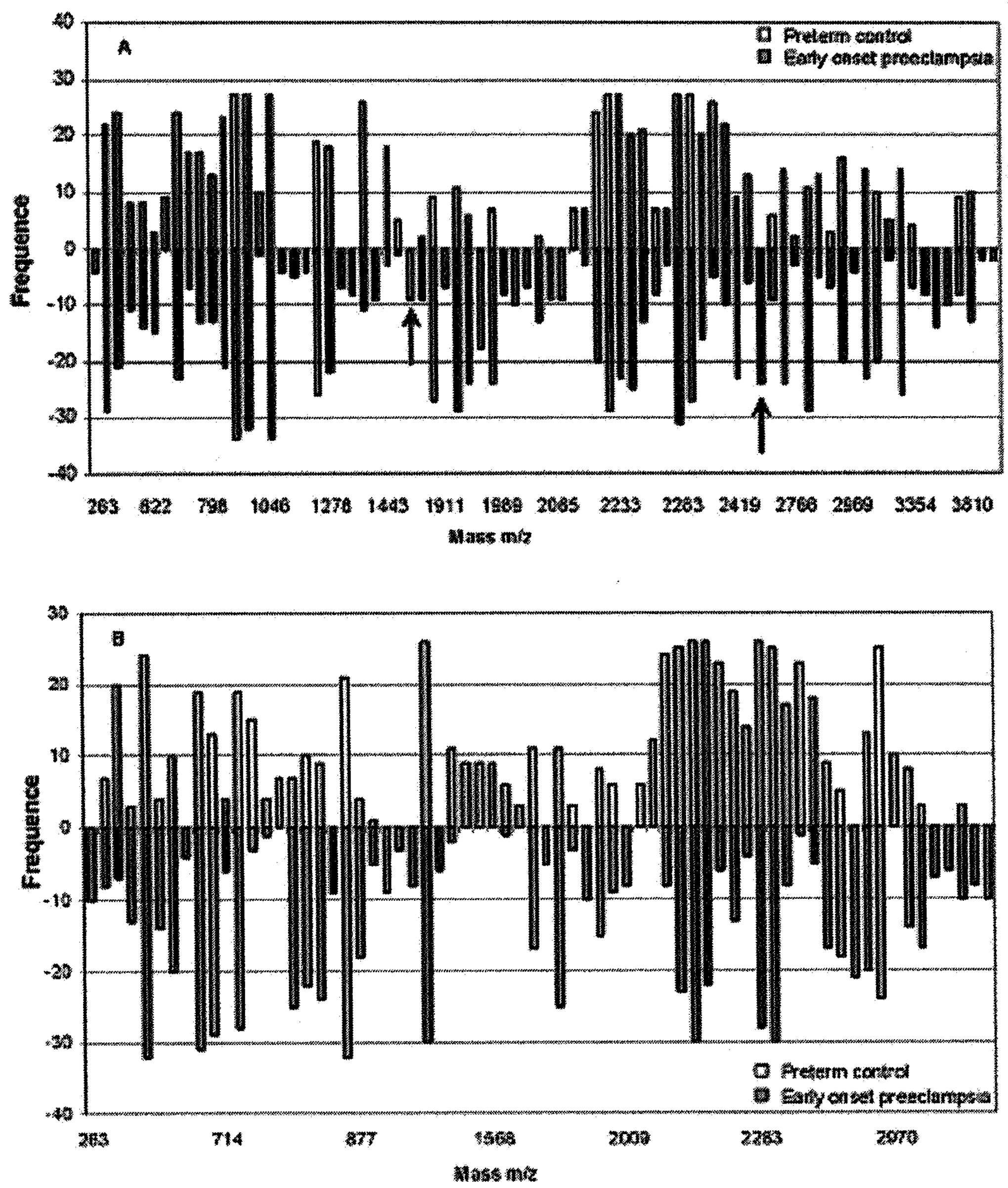
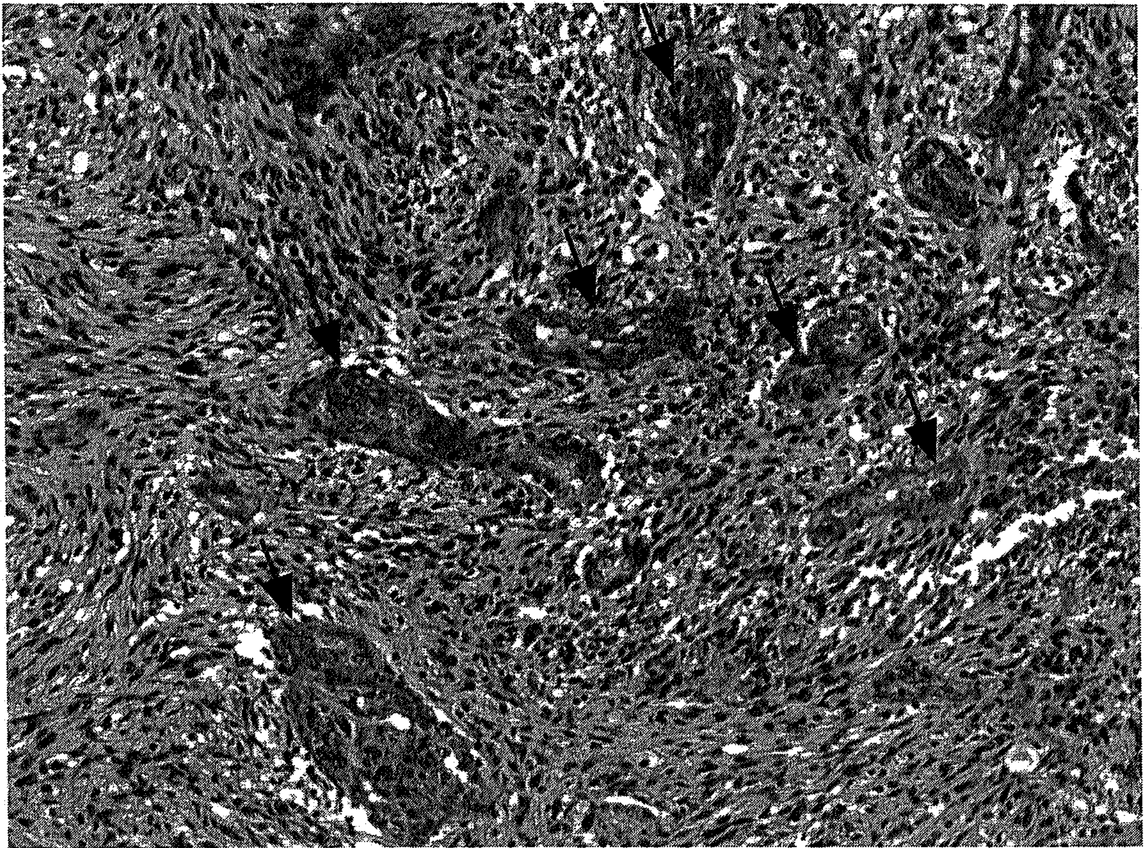


Figure 9



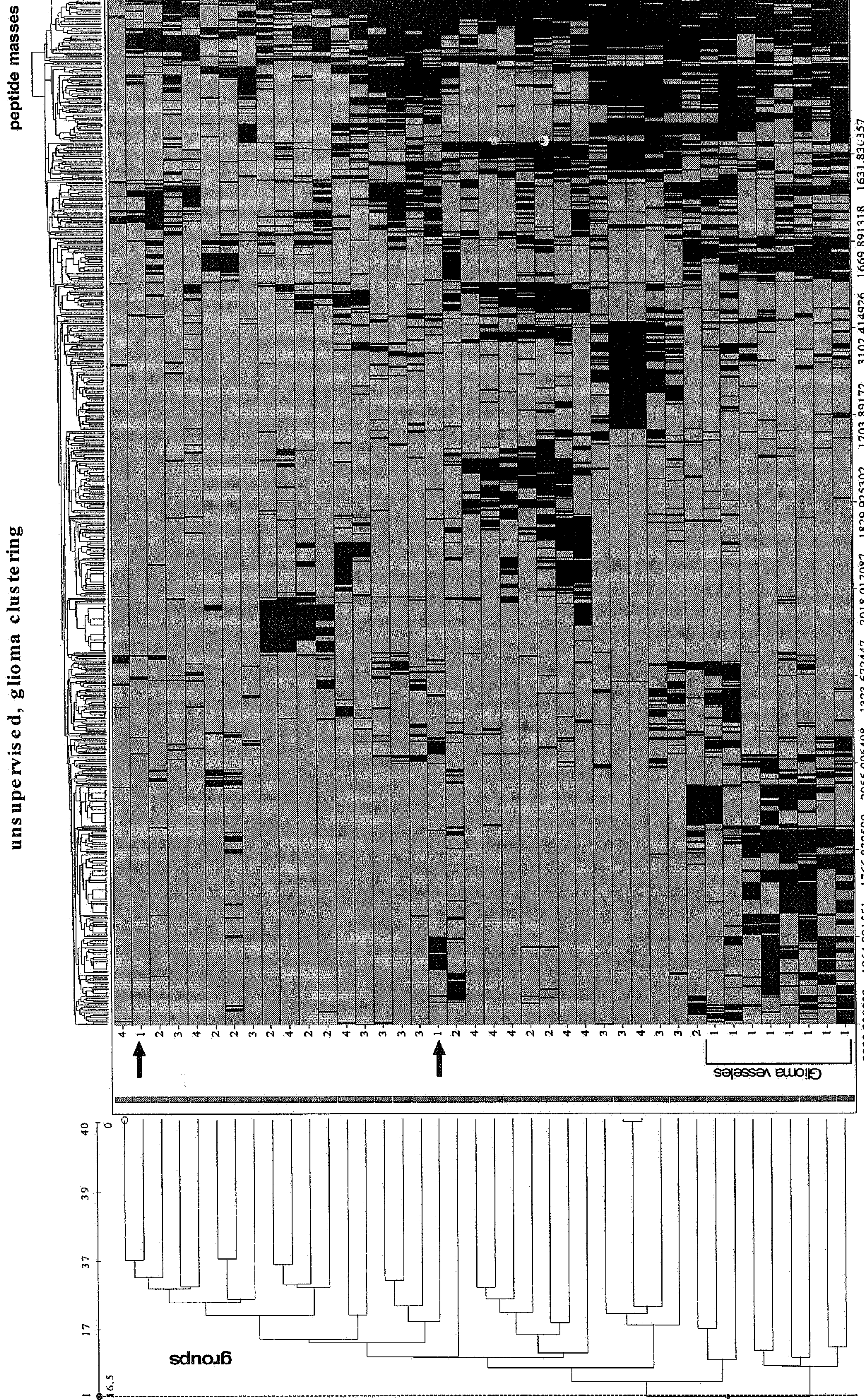


Fig. 10

Figure 11

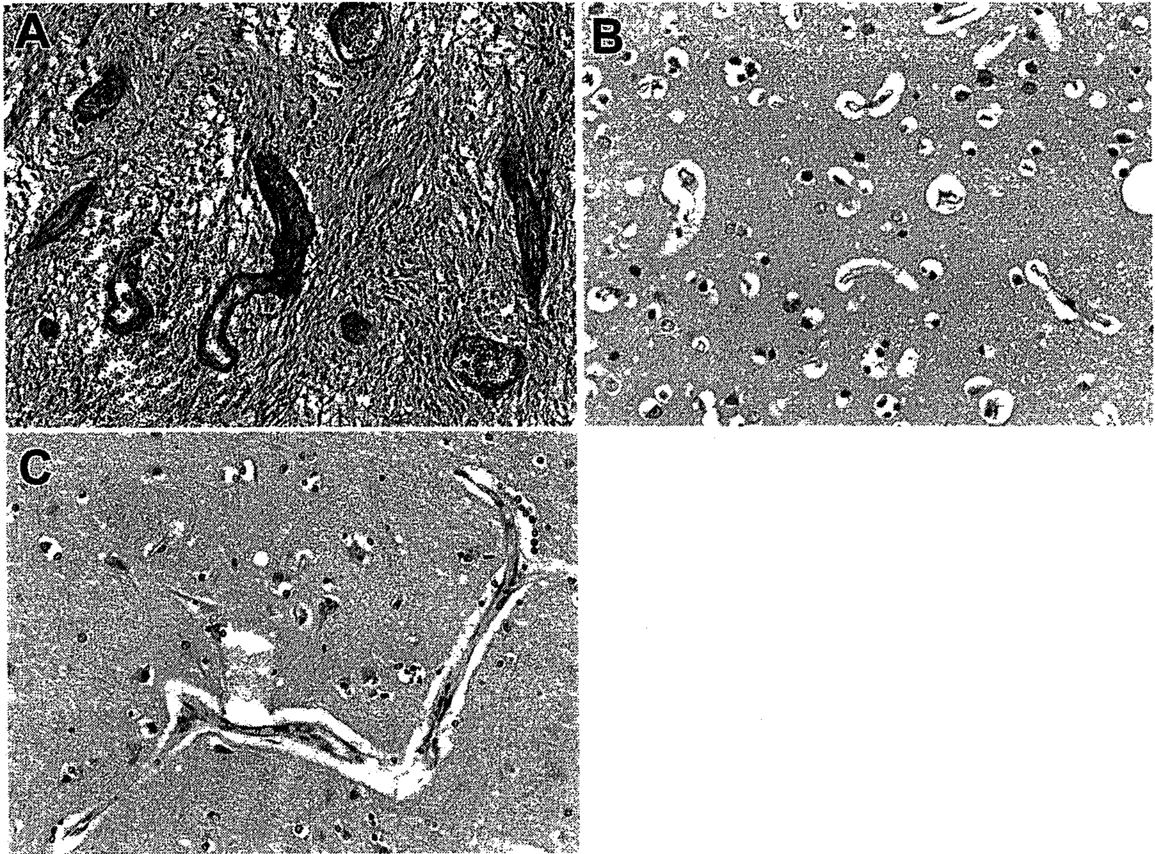


Figure 12

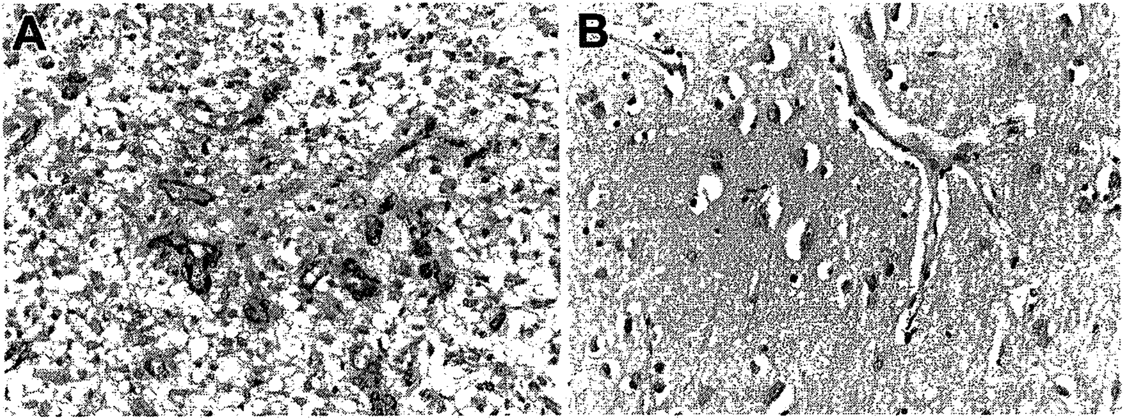


Figure 13

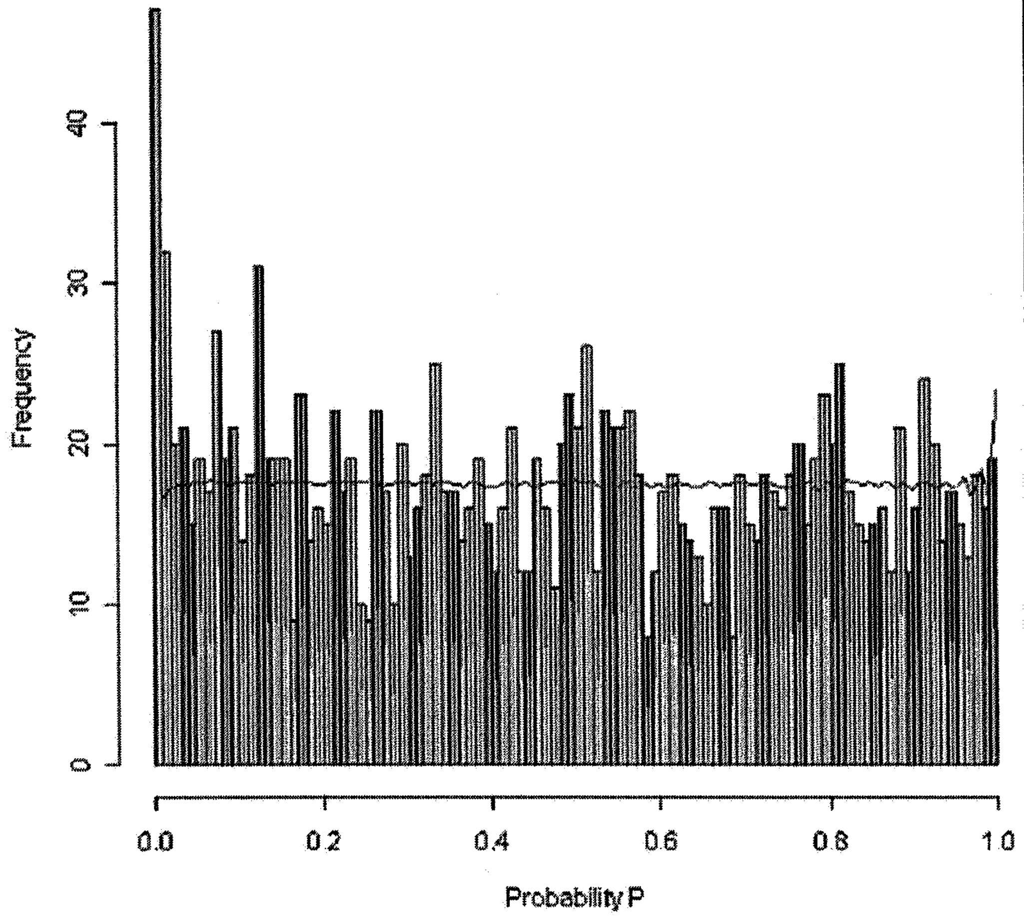


Figure 14

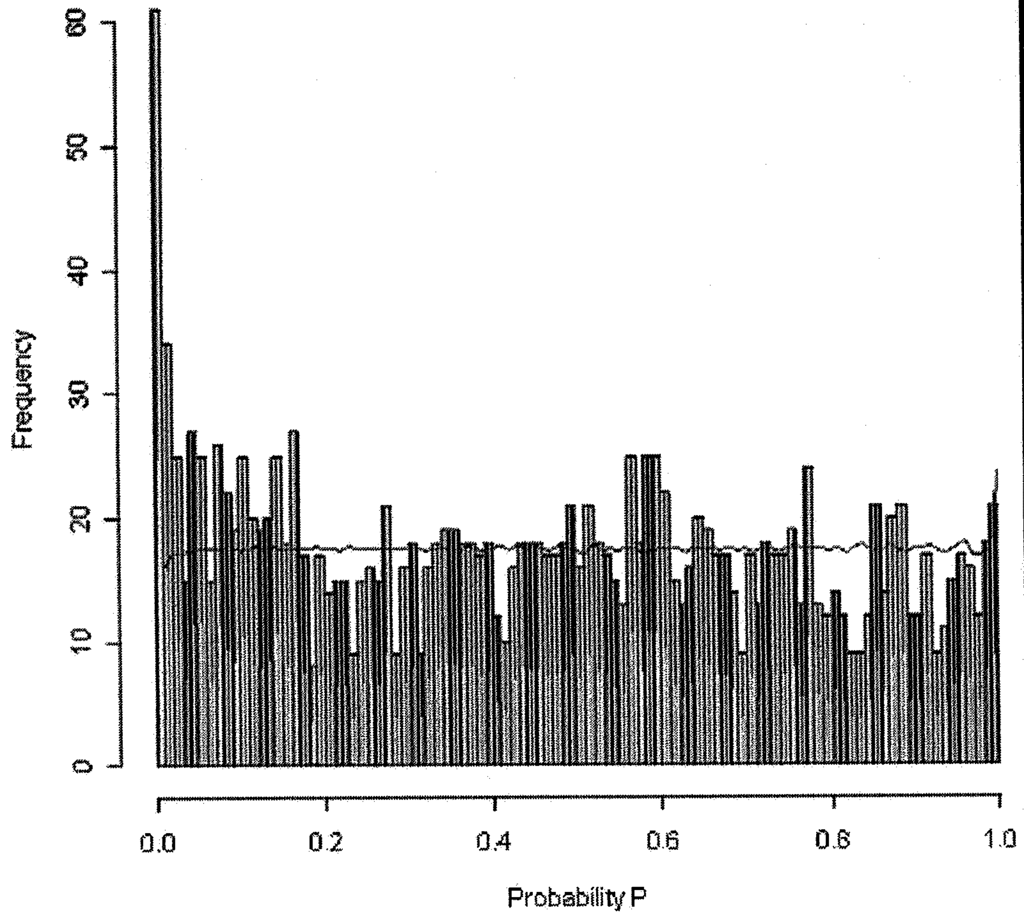


Figure 15

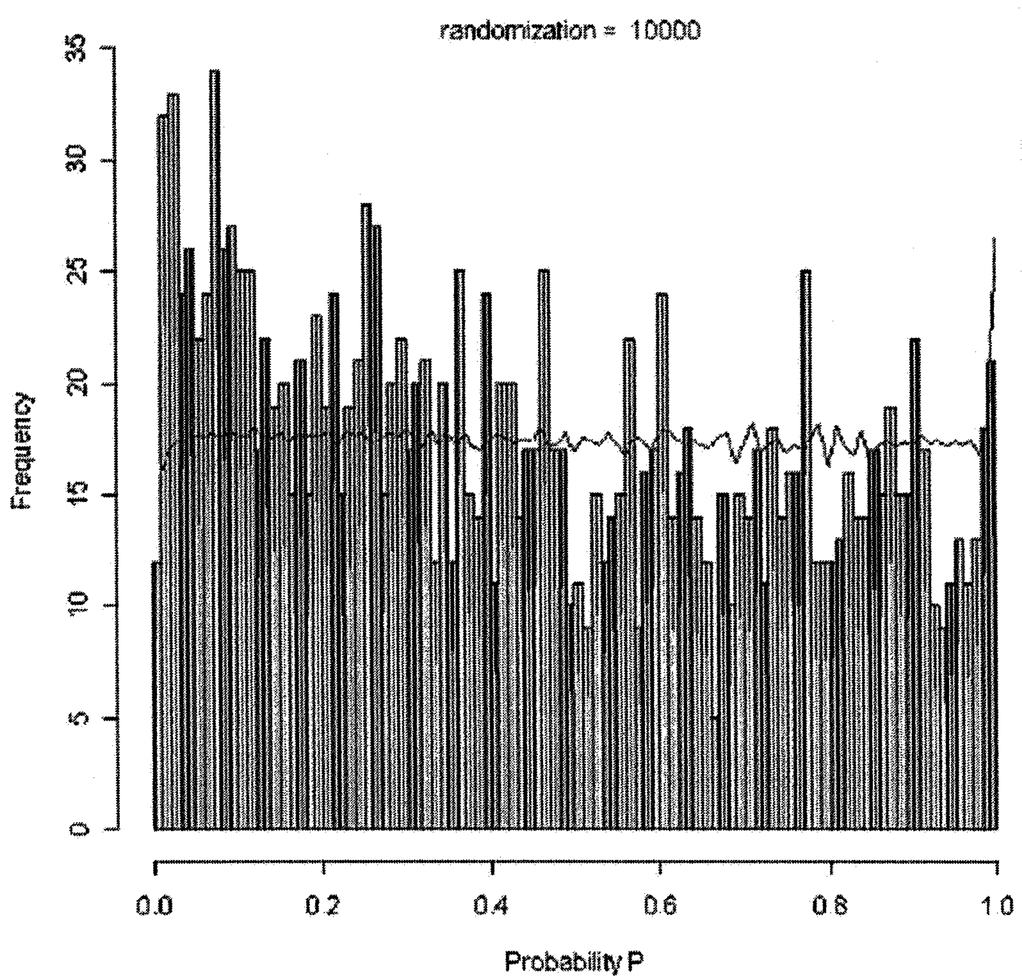


Figure 16

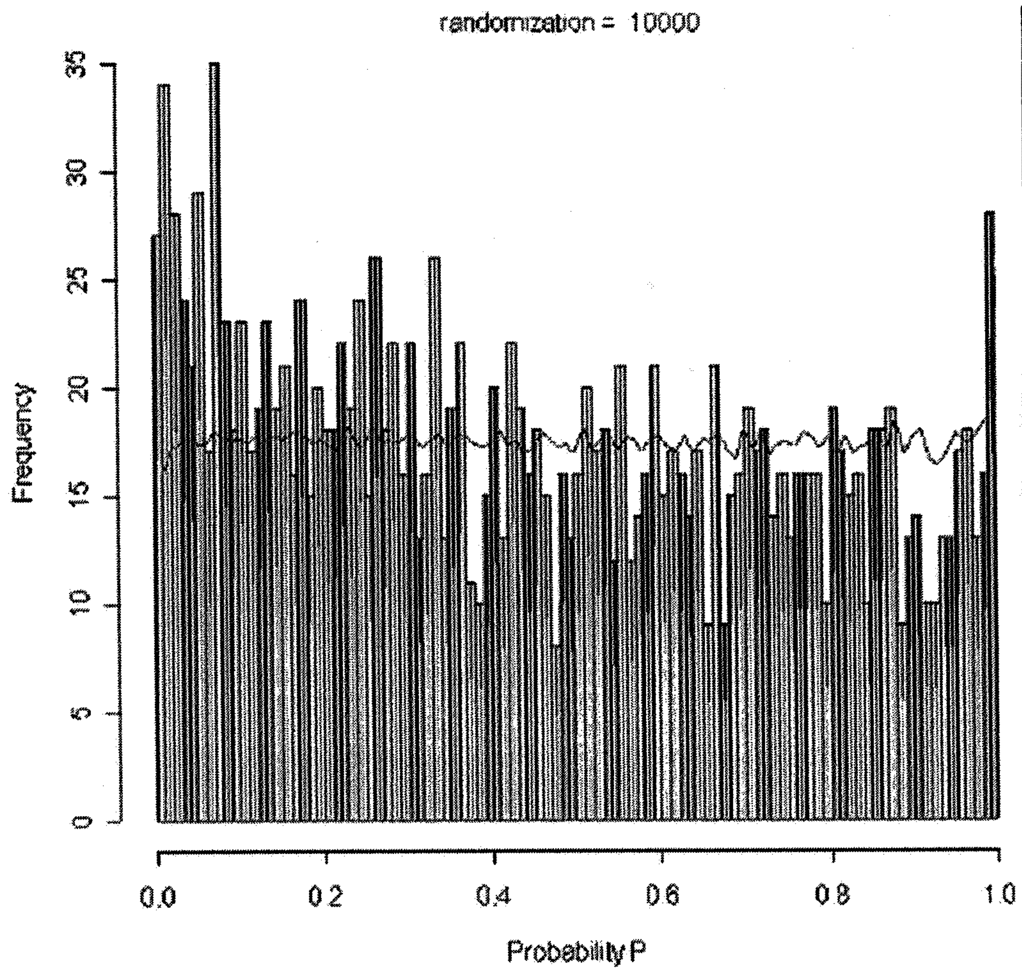


Figure 17

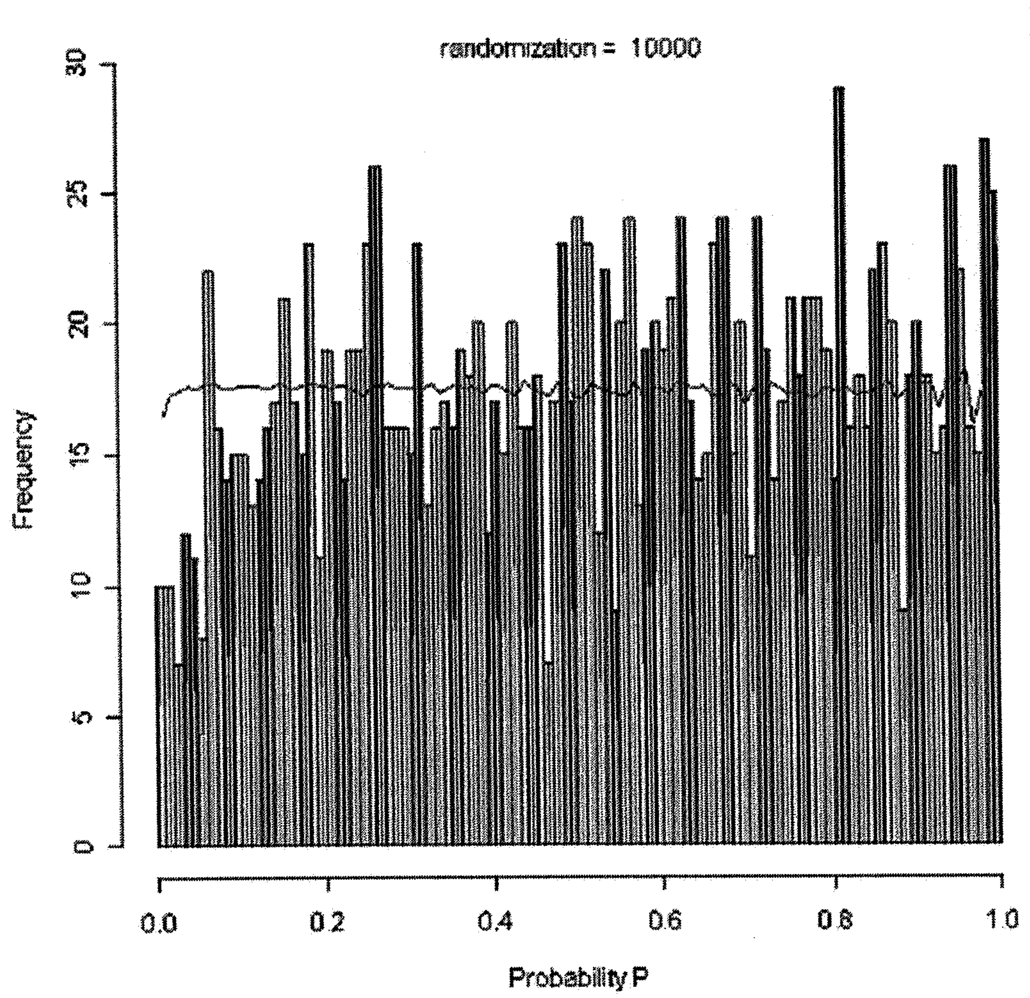
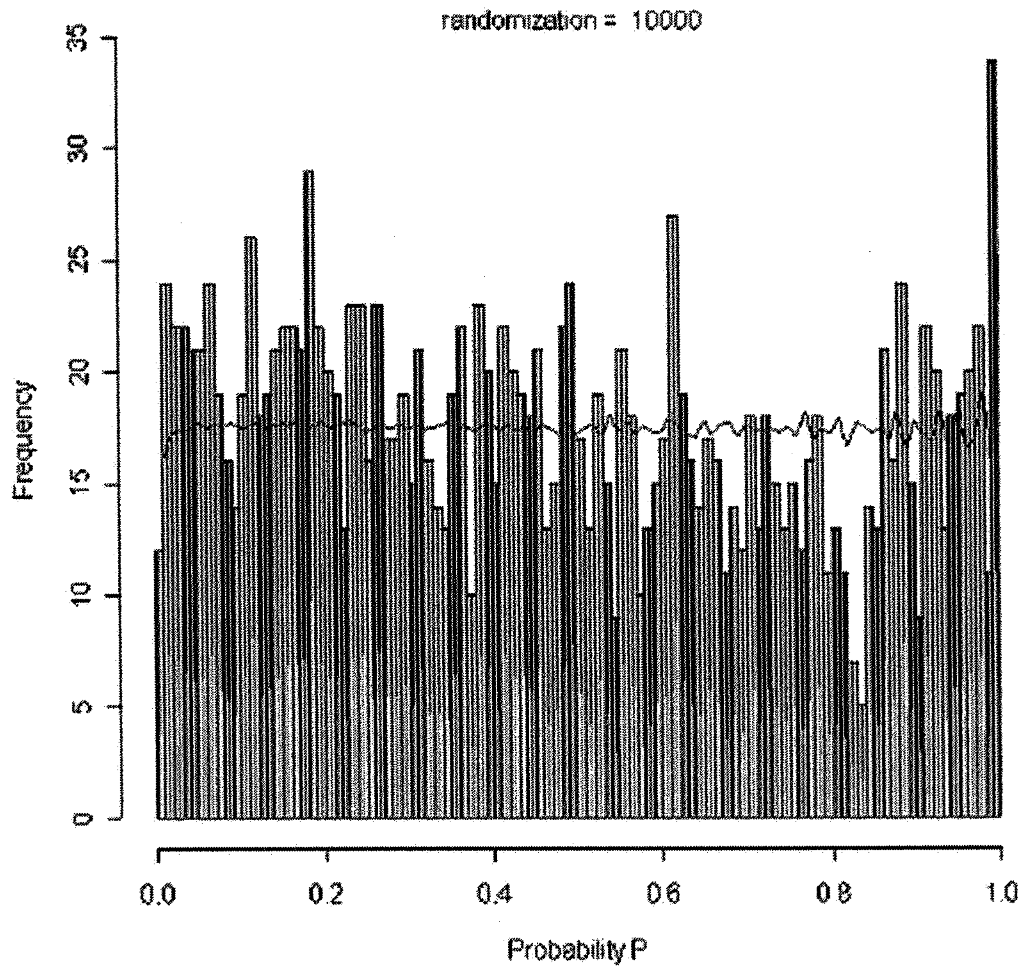


Figure 18



INTERNATIONAL SEARCH REPORT

International application No
PCT/NL2007/050014

A. CLASSIFICATION OF SUBJECT MATTER
INV. G01N33/68 G01N33/574

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, EMBASE, WPI Data, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DEKKER L J ET AL: "PEPTIDE PROFILING OF CEREBROSPINAL FLUID BY MASS SPECTROMETRY" EXPERT REVIEW OF PROTEOMICS, FUTURE DRUGS, LONDON, GB, vol. 3, June 2006 (2006-06), pages 297-309, XP008068285 ISSN: 1744-8387	1,2,5,7, 8,10-13
Y	abstract figures 2,4,5 page 299 - page 305 -----	3,4,6,9

-/--

[X] Further documents are listed in the continuation of Box C.

[ ] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
5 July 2007	30/11/2007

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Mulder, Lonneke

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SCHEEPERS ET AL: "Advances in the discovery of low-abundance proteins" TRAC, TRENDS IN ANALYTICAL CHEMISTRY, ELSEVIER, AMSTERDAM, NL, vol. 25, no. 9, October 2006 (2006-10), pages 841-847, XP005681788 ISSN: 0165-9936	14,18
Y	page 844, last paragraph - page 845, paragraph 1	3,4,6,9
A	ZHENG PING-PIN ET AL: "Low-molecular weight caldesmon as a potential serum marker for glioma." CLINICAL CANCER RESEARCH : AN OFFICIAL JOURNAL OF THE AMERICAN ASSOCIATION FOR CANCER RESEARCH 15 JUN 2005, vol. 11, no. 12, 15 June 2005 (2005-06-15), pages 4388-4392, XP002440128 ISSN: 1078-0432 abstract	19
A	FINEHOUT ERIN J ET AL: "Complement protein isoforms in CSF as possible biomarkers for neurodegenerative disease." DISEASE MARKERS 2005, vol. 21, no. 2, 2005, pages 93-101, XP008080716 ISSN: 0278-0240 abstract; figure 4	20
A	VAN GILS ET AL: "Innovations in Serum and Urine Markers in Prostate Cancer" EUROPEAN UROLOGY, S. KARGER AG., BASEL, CH, vol. 48, no. 6, December 2005 (2005-12), pages 1031-1041, XP005187116 ISSN: 0302-2838 page 1037 - page 1039	1-14
A	BISCHOFF R ET AL: "Methodological advances in the discovery of protein and peptide disease markers" JOURNAL OF CHROMATOGRAPHY B: BIOMEDICAL SCIENCES & APPLICATIONS, ELSEVIER, AMSTERDAM, NL, vol. 803, no. 1, 15 April 2004 (2004-04-15), pages 27-40, XP004495609 ISSN: 1570-0232 page 32 - page 35	1-14

-/--

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CAGNEY G ET AL: "Proteogest: A tool for facilitating proteomics using mass spectrometry" DRUG DISCOVERY TODAY: TARGETS 2004 UNITED KINGDOM, vol. 3, no. 2 SUPPL.1, 2004, pages S63-S65, XP002440975 ISSN: 1741-8372 abstract -----	1-14
O,X	BISCHOFF ET AL: "Biomarker discovery by mass spectrometry symposium, May 18-19, 2006" JOURNAL OF CHROMATOGRAPHY B: BIOMEDICAL SCIENCES & APPLICATIONS, ELSEVIER, AMSTERDAM, NL, vol. 847, no. 1, 18 May 2006 (2006-05-18), - 18 May 2006 (2006-05-18) pages 1-2, XP005881030 09-02-2007 ISSN: 1570-0232 abstract -----	15,18,22

INTERNATIONAL SEARCH REPORT

International application No.
PCT/NL2007/050014

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers allsearchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search reportcovers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-14 (complete)

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-14 (complete)

   Method and kits for identifying further biomarkers for diseases

   ---

2. claims: Claims 15, 18, 22 (partial)

   The biomarker for early onset preeclampsia is choriomammotropin precursor

   ---

3. claims: Claims 15, 18, 22 (partial)

   The biomarker for early onset preeclampsia is calcyclin

   ---

4. claims: Claims 15, 18, 22 (partial)

   The biomarker for early onset preeclampsia is sufeit locus protein 4

   ---

5. claims: Claims 16, 19, 21, 22 (partial)

   The biomarker for early onset preeclampsia is sufeit locus protein 4

   ---

6. claims: Claims 16, 19, 21, 22 (partial)

   The biomarker for glioma is fibrinogen

   ---

7. claims: Claims 16, 19, 21, 22 (partial)

   The biomarker for glioma is colligin 2

   ---

8. claims: Claims 16, 19, 21, 22 (partial)

   The biomarker for glioma is acidic calponin 3

   ---

9. claims: Claims 17, 20, 22 (partial)

   The biomarker for detecting multiple sclerosis is chromogranin A

   ---

10. claims: Claims 17, 20, 22 (partial)

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

The biomarker for detecting multiple sclerosis is clusterin

---

11. claims: Claims 17, 20, 22 (partial)

The biomarker for detecting multiple sclerosis is complement C3

---

12. claims: Claims 17, 20, 22 (partial)

complement C4

---

13. claims: Claims 17, 20, 22 (partial)

The biomarker for detecting multiple sclerosis is beta V spectrin

---

14. claims: Claims 17, 20, 22 (partial)

The biomarker for detecting multiple sclerosis is hypothetical protein XP_011125

---

15. claims: Claims 17, 20, 22 (partial)

The biomarker for detecting multiple sclerosis is apolipoprotein D

---

16. claims: Claims 17, 20, 22 (partial)

The biomarker for detecting multiple sclerosis is complement C4A

---

17. claims: Claims 17, 20, 22 (partial)

The biomarker for detecting multiple sclerosis is contactin 1

---

18. claims: Claims 17, 20, 22 (partial)

The biomarker for detecting multiple sclerosis is neuronal pentraxin receptor

---

19. claims: Claims 17, 20, 22 (partial)

The biomarker for detecting multiple sclerosis is RNA binding motif protein 7