(19) RÉPUBLIQUE FRANÇAISE

INSTITUT NATIONAL DE LA PROPRIÉTÉ INDUSTRIELLE

COURBEVOIE

(11) **N° de publication :** **3 029 671**

(à n'utiliser que pour les commandes de reproduction)

(21) **N° d'enregistrement national :** **15 61774**

(51) Int Cl[8] : ***G 06 T 7/00*** *(2016.01), G 06 F 19/26, G 01 N 27/62, 33/483, H 01 J 49/02*

(12) DEMANDE DE BREVET D'INVENTION A1

(22) **Date de dépôt :** 03.12.15.

(30) **Priorité :** 04.12.14 DE 102014224916.3.

(43) **Date de mise à la disposition du public de la demande :** 10.06.16 Bulletin 16/23.

(56) **Liste des documents cités dans le rapport de recherche préliminaire :** *Ce dernier n'a pas été établi à la date de publication de la demande.*

(60) **Références à d'autres documents nationaux apparentés :**

( ) **Demande(s) d'extension :**

(71) **Demandeur(s) :** *UNIVERSITAT BREMEN* — DE.

(72) **Inventeur(s) :** MAASS PETER, TREDE DENNIS, ALEKSANDROV FEDOR, OETJEN JANINA et HAUBERG-LOTTE LENA.

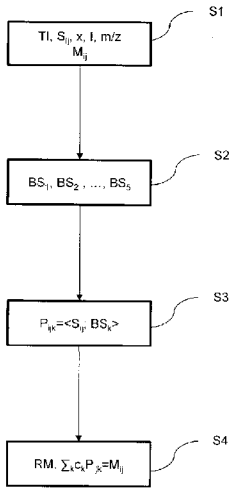
(73) **Titulaire(s) :** UNIVERSITAT BREMEN.

(74) **Mandataire(s) :** CABINET NUSS Société à responsabilité limitée.

(54) PROCEDE POUR L'ANALYSE ASSISTEE PAR ORDINATEUR D'UNE OU PLUSIEURS COUPES DE TISSU DU CORPS HUMAIN OU ANIMAL.

(57) L'invention concerne un procédé pour l'analyse assistée par ordinateur d'une ou plusieurs coupe(s) de tissu (TI) pour la préparation d'un test de coloration numérique, dans lequel, on détermine de manière assistée par ordinateur des zones de tissu qui présentent un symptôme clinique déterminé.

Dans le cadre du procédé, des données hyperspectrales à résolution spatiale sont traitées pour une pluralité de positions de mesure. On détermine dans ce cas, à partir des données hyperspectrales, plusieurs spectres de base (BS_1, BS_2, ..., BS_5) qui sont mis en corrélation avec les données hyperspectrales, moyennant quoi on obtient, pour chaque position de mesure (x), une pluralité de valeurs de corrélation (P_{ijk}) pour des spectres de base (BS_1, BS_2, ..., BS_5) respectifs. On détermine ensuite un modèle de régression (RM) lequel décrit un masque (Mij) de la ou des coupe(s) de tissu dans laquelle ou lesquelles des zones de tissu malades sont marquées, sur la base des valeurs de corrélation déterminées. Le modèle de régression (RM) et les spectres de base (BS_1, BS_2, ..., BS_5) qui ont été déterminés dans le cadre du procédé, peuvent ensuite être utilisés pour la réalisation du test de coloration numérique.



FR 3 029 671 - A1

DESCRIPTION

L'invention concerne un procédé pour l'analyse assistée par ordinateur d'une ou plusieurs coupe(s) de tissu du corps humain ou animal ainsi qu'un procédé pour la mise en œuvre assistée par ordinateur d'un test de coloration numérique.

En pathologie classique, on analyse des échantillons de tissus dans le but d'un diagnostic ou d'un pronostic d'indications médicales déterminées. Pour ce faire, des coupes de tissu sont traitées au moyen de procédés biochimiques, tels par exemple une coloration HE (coloration hématoxyline-éosine), des tests de coloration immunohistochimiques (également abrégés en tant que tests IHC) ou de procédés de test moléculaires. Avec ces procédés, des tissus malades sont marqués dans la coupe de tissu, de sorte que l'inspection visuelle et le diagnostic sont facilités. Dans le cadre des tests IHC par exemple, des anticorps pour le marquage visuel de biomarqueurs, lesquels sont spécifiques pour certains symptômes cliniques, sont appliqués sur une coupe de tissu.

L'état de la technique a par ailleurs fait connaître des méthodes de spectrométrie de masse pour l'identification de structures de tissu modifiées. A cet effet, on utilise souvent des procédés d'imagerie de spectrométrie de masse, par exemple DESI ou MALDI. Dans ce cas, pour chaque point d'image d'une coupe de tissu, un spectre complet de rapport masse/charge dans la plage de quelques 100 à environ 10 kDa est détecté. On obtient ainsi un spectre de masse présentant respectivement une coordonnée spatiale et une coordonnée spectrale de masse/charge. A partir de ces données spectrométriques de masse, il n'est à ce jour pas possible d'extraire les informations qui y sont contenues de manière à ce que l'analyse spectrométrique de masse puisse remplacer des tests de coloration histologiques classiques.

L'invention a pour but d'analyser les données hyperspectrales obtenues à partir de cette ou ces coupe(s) de tissu de manière à ce que ces données puissent être utilisées pour la réalisation de tests de coloration numériques ou assistés par ordinateur. L'invention a par ailleurs pour but de fournir un procédé pour la réalisation d'un test de coloration numérique.

Ce but est atteint grâce à un procédé pour l'analyse assistée par ordinateur d'une ou plusieurs coupe(s) de tissu du corps humain ou animal

pour la préparation d'un test de coloration numérique, dans lequel, dans le test de coloration numérique, on détermine de manière assistée par ordinateur des zones de tissu dans une coupe de tissu testée, lesquelles présentent un symptôme clinique prédéterminé, dans lequel :

a) des données hyperspectrales à résolution spatiale d'une ou plusieurs coupe(s) de tissu prédéterminée(s) à analyser qui présentent des zones de tissu avec le symptôme clinique prédéterminé sont importées, dans lequel les données hyperspectrales contiennent des spectres respectifs pour une pluralité de positions de mesure dans la ou les coupe(s) de tissu prédéterminée(s), dans lequel un spectre respectif représente la composition de la ou des coupe(s) de tissu prédéterminée(s) au niveau d'une position de mesure sous la forme d'un vecteur de valeurs d'intensité pour une pluralité de canaux, dans lequel une valeur d'intensité est une grandeur physique détectée pour le canal correspondant;

b) un masque de la ou des coupe(s) de tissu prédéterminée(s) étant importé, lequel contient des valeurs d'échelle respectives pour la pluralité des positions de mesure, dans lequel des zones de la ou des coupe(s) de tissu prédéterminée(s) avec des valeurs d'échelle divergentes par rapport à la zone restante de la ou des coupe(s) de tissu sont des zones de tissu avec le symptôme clinique prédéterminé;

c) grâce à une méthode pour la détermination de vecteurs de base, plusieurs spectres de base dans l'espace des vecteurs des données hyperspectrales ou dans un espace de vecteurs dérivés, lesquels sont calculés à partir des données hyperspectrales pour chaque position de mesure, étant déterminés ;

d) les spectres de base avec les vecteurs des données hyperspectrales ou les vecteurs dérivés étant mis en corrélation au niveau de chaque position de mesure, moyennant quoi on obtient, pour chaque position de mesure, une pluralité de valeurs de corrélation pour des spectres de base respectifs;

e) un modèle de régression étant déterminé à l'aide d'un procédé de régression, lequel décrit le masque sur la base de la pluralité des valeurs de corrélation pour les positions de mesure respectives, dans lequel le modèle de régression et les spectres de base sont enregistrés pour la réalisation du test de coloration numérique.

Selon une caractéristique de l'invention, les données hyperspectrales sont des données spectrométriques de masse, en particulier

des données spectrométriques de masse MALDI (Désorption/Ionisation laser assistée par matrice) où les canaux représentent des rapports masse/charge et la grandeur physique détectée pour le canal correspondant représentant le nombre de molécules détectées pour le rapport masse/charge correspondant.

En accord avec une autre caractéristique de l'invention, il peut être prévu que, à l'étape c), les spectres de base sont déterminés dans l'espace des vecteurs dérivés, dans lequel un vecteur dérivé est calculé pour chaque position de mesure de telle sorte que :

- pour le symptôme clinique prédéterminé, des données hyperspectrales pour une pluralité de compositions moléculaires sont prescrites, dans lequel les données hyperspectrales pour une composition moléculaire contiennent un vecteur sous la forme d'un spectre caractéristique de la composition moléculaire de valeurs d'intensité pour un nombre prescrit de canaux parmi une pluralité de canaux ;
- les spectres caractéristiques des compositions moléculaires étant mis en corrélation avec les spectres des données hyperspectrales importées au niveau des positions de mesure respectives, moyennant quoi on détermine, en guise de vecteur dérivé pour chaque position de mesure, un vecteur à partir de plusieurs coefficients de corrélation.

Le procédé selon l'invention peut en outre présenter éventuellement l'une ou plusieurs des caractéristiques suivantes :

- la ou les coupe(s) de tissu à analyser sont des coupes de tissu respectives avec digestion ;
- la méthode utilisée à l'étape c) pour la détermination de vecteurs de base est une factorisation matricielle non négative ou une analyse PCA (Analyse du composant principal) ;
- la factorisation matricielle non négative comprend la condition de l'occupation éparse des matrices déterminées par la factorisation matricielle non négative ;
- le procédé de régression utilisé à l'étape e) est un procédé de régression linéaire ou un procédé de régression SVM ;
- à l'étape b), le masque est déterminé à partir d'un test de coloration histologique, et en particulier immunohistochimique, de la ou des coupe(s) de tissu prédéterminée(s) et/ou

- le symptôme clinique prédéterminé représente la présence de cellules tumorales prédéterminées ou la présence d'îlots d'insuline dans la ou les coupe(s) de tissu prédéterminée(s).

Le but précité est également atteint, selon un autre ou second aspect de l'invention, grâce à un procédé pour la réalisation assistée par ordinateur d'un test de coloration numérique en utilisant des spectres de base et un modèle de régression, lesquels ont été déterminés avec un procédé selon l'une des revendications précédentes pour un symptôme clinique prédéterminé et enregistrés numériquement, dans lequel :

i) des données hyperspectrales à résolution spatiale d'une coupe de tissu prescrite à tester sont importées, dans lequel les données hyperspectrales contiennent des spectres respectifs pour une pluralité de positions de mesure dans la coupe de tissu prescrite, dans lequel un spectre respectif représente la composition de la coupe de tissu prescrite au niveau d'une position de mesure sous la forme d'un vecteur de valeurs d'intensité pour des canaux parmi une pluralité de canaux, dans lequel une valeur d'intensité est une grandeur physique détectée pour le canal correspondant;

ii) les vecteurs des données hyperspectrales ou des vecteurs dérivés, lesquels ont été calculés à partir des données hyperspectrales, étant mis en corrélation au niveau de chaque position de mesure avec les spectres de base ayant été déterminés avec un procédé selon l'une des revendications précédentes, moyennant quoi on obtient, pour chaque position de mesure, une pluralité de valeurs de corrélation pour des spectres de base respectifs ;

iii) le modèle de régression, lequel a été déterminé avec un procédé selon l'une des revendications précédentes, étant appliqué aux valeurs de corrélation et un masque étant ainsi obtenu, lequel contient, pour la pluralité de positions de mesure, des valeurs d'échelle respectives, dans lequel des zones de la coupe de tissu prescrite avec des valeurs d'échelle divergentes par rapport à la coupe de tissu restante, sont des zones de tissu avec le symptôme clinique prédéterminé.

Selon une caractéristique additionnelle du procédé précité, une visualisation est générée sur un afficheur de telle sorte qu'une image est générée à partir de points d'image représentant les positions de mesure, dans lequel un point d'image respectif assure le codage de la valeur d'échelle du masque déterminé à l'étape iii) pour la position de mesure représentée par le point d'image respectif.

Conformément à une autre caractéristique possible des vecteurs dérivés à l'étape ii) sont mis en corrélation avec les spectres de base, dans lequel les vecteurs dérivés sont calculés de telle sorte que,

- pour le symptôme clinique prédéterminé, des données hyperspectrales pour une pluralité de compositions moléculaires sont prescrites, dans lequel les données hyperspectrales pour une composition moléculaire contiennent un vecteur sous la forme d'un spectre caractéristique pour la composition moléculaire de valeurs d'intensités pour un nombre prescrit de canaux parmi la pluralité de canaux;
- les spectres caractéristiques pour les compositions moléculaires étant mis en corrélation avec les spectres des données hyperspectrales importées à l'étape i) au niveau des positions de mesure respectives, moyennant quoi on détermine, en guise de vecteur dérivé pour chaque position de mesure, un vecteur à partir de plusieurs coefficients de corrélation.

L'invention concerne aussi un dispositif pour l'analyse assistée par ordinateur d'une ou plusieurs coupe(s) de tissu du corps humain ou animal pour la préparation d'un test de coloration numérique, dans lequel, dans le test de coloration numérique, on détermine de manière assistée par ordinateur des zones de tissu dans une coupe de tissu testée, lesquelles présentent un symptôme clinique prédéterminé, caractérisé en ce que le dispositif est conçu pour la réalisation d'un procédé tel que décrit précédemment en relation avec un premier aspect de l'invention.

L'invention porte en outre sur un dispositif pour la réalisation assistée par ordinateur d'un test de coloration numérique en utilisant des spectres de base et un modèle de régression, lesquels ont été déterminés avec un procédé tel qu'évoqué ci-dessus pour un symptôme clinique prédéterminé et enregistrés numériquement, caractérisé en ce que le dispositif est conçu pour la réalisation d'un procédé en accord avec un second aspect de l'invention également décrit précédemment.

Enfin l'invention comprend également un programme informatique avec un code de programme enregistré sur un support lisible par machine pour la réalisation d'un procédé précité lorsque le code de programme est exécuté sur un ordinateur, ainsi qu'un programme informatique avec un code de programme pour la réalisation d'un procédé précité lorsque le code de programme est exécuté sur un ordinateur.

Comme cela ressort de l'exposé qui précède, le procédé selon l'invention sert à l'analyse assistée par ordinateur d'une ou plusieurs coupe(s) de tissu du corps humain ou animal. Une coupe de tissu respective a été prélevée du corps au moyen d'une biopsie et est traitée en-dehors du corps sans que le tissu soit restitué au corps. Ici et dans ce qui suit, la notion de coupe de tissu s'entend au sens large. Une coupe de tissu peut en particulier également être un micro-réseau de tissu (Tissue-Microarray) où une pluralité d'échantillons de tissu (des dénommés Tissue-Cores) qui proviennent de préférence de patients différents sont appliqués sur un porte-échantillon. Grâce à l'analyse selon l'invention, un test de coloration numérique est préparé, avec lequel on détermine de manière assistée par ordinateur des zones de tissu dans une coupe de tissu testée, lesquelles présentent un symptôme clinique prédéfini.

On explique dans ce qui suit de manière plus détaillée les caractéristiques individuelles du procédé à l'aide des étapes a) à e). Ces étapes servent simplement à un meilleur référencement des caractéristiques du procédé et ne sont pas forcément en corrélation avec une succession chronologique de l'exécution des étapes.

Dans une étape a), des données hyperspectrales à résolution spatiale (numériques) d'une ou plusieurs coupe(s) de tissu prédéfinie(s) à analyser sont importées, la ou les coupe(s) de tissu présentant des zones de tissu avec un symptôme clinique prédéterminé. Les données hyperspectrales contiennent, pour une pluralité de positions de mesure (en particulier bidimensionnelles) dans la ou les coupe(s) de tissu prédéterminée(s), des spectres respectifs, un spectre respectif représentant la composition de la ou des coupe(s) de tissu prédéterminée(s) au niveau d'une position de mesure ou sous la forme d'un vecteur de valeurs d'intensité pour une pluralité de canaux, une valeur d'intensité étant une grandeur physique détectée pour le canal correspondant.

Dans le cas de l'analyse de plusieurs coupes de tissu prédéterminées, on peut différencier les positions de mesure par-delà les coupes de tissu. En d'autres termes, dans ce cas, le nombre global des positions de mesure correspond à la somme de toutes les positions de mesure contenues dans les coupes de tissu, même si des positions de mesure dans des coupes de tissu différentes présentent les mêmes coordonnées spatiales. Pour différencier des positions de coupes de tissu différentes, on peut utiliser un index qui spécifie la coupe de tissu respective. Les coupes

de tissu proviennent de préférence de différents patients (c'est-à-dire des humains ou animaux différents). Dans une variante préférée de l'invention, on analyse entre 50 et 100 coupes de tissu avec le procédé selon l'invention. Le test de coloration numérique est d'autant plus fiable que le nombre de coupes de tissu est important, lequel test est exécuté sur la base des résultats du procédé d'analyse.

Les données numériques hyperspectrales ci-dessus peuvent avoir été saisies à l'avance et être stockées dans une mémoire numérique à partir de laquelle elles sont extraites à l'étape a). Le procédé selon l'invention peut néanmoins également comprendre l'étape de la saisie des données hyperspectrales au moyen d'un spectromètre approprié.

Concernant les données hyperspectrales, il s'agit de préférence de données de spectrométrie de masse, en particulier de dénommées données de spectrométrie de masse MALDI (MALDI = Matrix-Assisted Laser Desorption/Ionization/Désorption/Ionisation laser assistée par une matrice). Pour des données de spectrométrie de masse, les canaux sont des rapports masse/charge, et la grandeur physique détectée pour le canal correspondant représente le nombre de molécules détectées pour le rapport de masse/charge correspondant. Les données hyperspectrales peuvent néanmoins le cas échéant également représenter d'autres données, telles par exemple des données issus de la spectroscopie Raman. Dans le cas de la spectroscopie Raman, les canaux représentent des fréquences de lumière réfléchie ou diffusée au niveau de l'échantillon de tissu, et la grandeur physique est l'intensité de cette lumière réfléchie ou diffusée.

Dans une étape b) du procédé selon l'invention, on importe un masque (numérique) de la ou des coupe(s) de tissu prédéterminée(s), lequel contient des valeurs d'échelle respectives pour la pluralité de positions de mesure, des zones de la ou des coupe(s) de tissu prédéterminée(s) avec des valeurs d'échelle divergentes (par exemple plus élevées) par rapport à la zone restante de la ou des coupe(s) de tissu étant des zones de tissu avec le symptôme clinique prédéterminé. Le masque numérique peut être obtenu à partir d'un test de coloration histologique classique, et en particulier d'un test de coloration immunohistochimique de la ou des coupe(s) de tissu à analyser. En cas d'analyse de plusieurs coupes de tissu, le masque comprend des masques partiels respectifs pour chaque coupe de tissu. Les positions de mesure dans les masques partiels sont en corrélation avec les positions de mesure correspondantes dans les coupes de tissu respectives.

Comme indiqué plus haut, les positions de mesure peuvent dans ce cas être différenciées par-delà les coupes de tissu, et donc également par-delà les masques partiels. Le masque numérique est en règle générale enregistré à l'avance dans une mémoire numérique à partir de laquelle les données sont extraites sous la forme des valeurs d'échelle correspondantes. Selon le mode de réalisation, les valeurs d'échelle peuvent provenir d'une plage de valeurs discrètes ou continues. Le cas échéant, les valeurs d'échelle peuvent également seulement adopter les valeurs 0 et 1, la valeur 0 représentant une zone de tissu sans symptôme clinique, et la valeur 1 représentant une zone de tissu avec symptôme clinique, ou vice versa.

Dans une étape c) du procédé selon l'invention, on détermine, au moyen d'une méthode pour la détermination de vecteurs de base, plusieurs spectres de base dans l'espace des vecteurs des données hyperspectrales ou dans un espace de vecteurs dérivés. Les vecteurs dérivés sont dans ce cas calculés à partir des données hyperspectrales pour chaque position de mesure. Des variantes spéciales du calcul des vecteurs dérivés sont décrites plus loin ou dans la revendication 3. Les spectres de base représentent les vecteurs de base. Le nombre de spectres de base est de préférence compris entre 5 et 12. On entend par méthode pour la détermination de vecteurs de base, un procédé qui détermine des vecteurs de base de telle sorte que les vecteurs des données hyperspectrales, ou les vecteurs dérivés, puissent être représentés par des combinaisons (en particulier des combinaisons linéaires) des vecteurs de base. Des variantes préférées de méthodes pour la détermination de vecteurs de base sont mentionnées plus loin.

Dans une étape d) du procédé selon l'invention, les spectres de base sont mis en corrélation avec les vecteurs des données hyperspectrales ou les vecteurs dérivés à chaque position de mesure, moyennant quoi l'on obtient, pour chaque position de mesure, une pluralité de valeurs de corrélation pour des spectres de base respectifs. C'est ainsi que pour chaque position de mesure et chaque spectre de base, il existe une valeur de corrélation. Des procédés pour la mise en corrélation de vecteurs sont bien connus. La valeur de corrélation est de préférence le produit scalaire de vecteurs normalisés, c'est-à-dire des vecteurs qui sont respectivement multipliés par la valeur inverse de leur montant. Dans ce cas, la valeur de corrélation est comprise entre -1 et 1. Le cas échéant, la valeur de corrélation peut également être le montant du produit scalaire. Dans ce cas,

la valeur de corrélation est comprise entre 0 et 1 et peut être interprétée en tant que degré de similitude entre les vecteurs, les vecteurs étant d'autant plus semblables que la valeur de corrélation est élevée.

Dans une étape e), on détermine finalement, au moyen du procédé de régression, un modèle de régression (à savoir des paramètres correspondants du modèle de régression). Le modèle de régression décrit le masque sur la base de la pluralité des valeurs de corrélation pour les positions de mesure respectives. Le modèle de régression avec ses paramètres et les vecteurs de base sont finalement enregistrés numériquement pour la réalisation ultérieure du test de coloration numérique.

Avec les résultats du procédé selon l'invention sous la forme des vecteurs de base et du modèle de régression, il est possible de réaliser de manière simple des tests de coloration assistés par ordinateur de nouvelles coupes de tissu afin de déterminer si dans ces coupes de tissu, il y a des zones de tissu avec le symptôme clinique prédéterminé. La notion de test de coloration doit être prise au sens large. Le test de coloration n'a en particulier pas forcément besoin de coder en couleur des zones de tissu avec le symptôme clinique prédéterminé, le codage peut au contraire généralement s'effectuer via les valeurs scalaires décrites plus haut.

Comme décrit à l'étape c) ci-dessus, dans une variante du procédé selon l'invention, la méthode pour la détermination de vecteurs de base peut être directement utilisée dans l'espace des vecteurs des données hyperspectrales. Dans ce cas, il est possible d'obtenir des résultats de manière simple sans informations spécifiques sur les biomarqueurs présents dans des zones de tissu malade, lesquels résultats permettent la réalisation de tests de coloration numériques. Les spectres de base peuvent néanmoins le cas échéant également être déterminés dans un espace de vecteurs dérivés, lesquels sont déterminés ou calculés à partir des données hyperspectrales pour chaque position de mesure. A cet effet, on utilise de préférence des informations sur les compositions moléculaires présentes dans le tissu malade, comme cela est le cas dans la variante expliquée dans ce qui suit du procédé selon l'invention.

Dans une variante spéciale du procédé selon l'invention, dans laquelle, à l'étape c), les spectres de base sont déterminés dans un espace de vecteurs dérivés, on détermine un vecteur dérivé pour chaque position de mesure de la manière qui suit.

Pour le symptôme clinique prédéterminé (déterminé théoriquement ou expérimentalement), des données hyperspectrales pour une pluralité de compositions moléculaires sont prescrites. Dans un mode de réalisation préféré, une composition moléculaire respective est représentée par un peptide prescrit avec ses isotopes. En cas de données de spectrométrie de masse, les données hyperspectrales d'un peptide sont représentées par le spectre de masse du peptide pour les différents rapports masse/charge de ses isotropes. Le cas échéant, une composition moléculaire peut également concerner plusieurs peptides ou d'autres molécules qui sont représenté(e)s par des données hyperspectrales et en particulier par un spectre de masse correspondant. Les compositions moléculaires et leurs données hyperspectrales ont été déterminées à l'avance avec des procédés adaptés, ces procédés ne faisant pas partie de l'invention et n'étant pas de ce fait expliqués plus en détails.

Les données hyperspectrales pour une composition moléculaire contiennent généralement un vecteur sous la forme d'un spectre de valeurs d'intensité caractéristique pour la composition moléculaire pour un nombre prescrit de canaux parmi la pluralité des canaux. Le nombre de canaux représente en règle générale une partie des canaux des données hyperspectrales importées à l'étape a) ci-dessus.

Dans une étape suivante, les spectres caractéristiques pour les compositions moléculaires sont mis en corrélation avec les spectres des données hyperspectrales importées au niveau des positions de mesure respectives, la corrélation pouvant s'effectuer de manière analogue à ce qui est décrit plus haut via le produit scalaire des vecteurs normalisés. Lors de la corrélation, on ne tient pas compte de tous les canaux, mais seulement du nombre prescrit de canaux. En guise de résultat de la corrélation on détermine, en guise de vecteur dérivé pour chaque position de mesure, un vecteur à partir de plusieurs coefficients de corrélation, sachant que pour une position de mesure respective, il existe un coefficient de corrélation pour chaque composition moléculaire. De manière préférée, il existe de 10 à 200 coefficients de corrélation pour une position de mesure respective.

Dans la variante du procédé selon l'invention décrite plus haut, on obtient ainsi, à partir des données hyperspectrales d'origine, des vecteurs à partir de coefficients de corrélation, lesquels décrivent la corrélation avec des compositions moléculaires connues à l'avance. Au lieu des données

hyperspectrales, ce sont alors ces vecteurs dérivés qui sont soumis à un traitement ultérieur dans les étapes c) à e).

Dans un variante particulièrement préférée, la ou les coupe(s) de tissu à analyser sont des coupes de tissu respectives avec digestion. La notion de digestion est bien connue en histologie. Dans ce cas, la coupe de tissu est soumise à un traitement préalable via des étapes biochimiques du procédé. Dans le cadre de la digestion, des molécules plus grandes sont en particulier décomposées en peptides par des enzymes.

A l'étape c), en guise de méthode pour la détermination de vecteurs de base, on peut utiliser des procédés connus de l'état de la technique. Dans un mode de réalisation préféré, on utilise une factorisation matricielle non négative. On peut également faire appel à une analyse PCA (PCA = Principal Component Analysis, en français : analyse du composant principal). Dans une variante particulièrement préférée on utilise, pour la détermination des vecteurs de base, une factorisation matricielle non négative avec pour contrainte l'occupation éparse des matrices déterminées par la factorisation matricielle non négative. On utilise en particulier la factorisation matricielle décrite dans les documents [1] et [2]. L'ensemble de la révélation de ces documents [1] et [2] est intégrée en tant que contenu à la présente demande en y faisant référence.

A l'étape d) également, il est possible d'utiliser des procédés de régression connus en tant que tels pour la détermination du modèle de régression. On peut par exemple utiliser un procédé de régression linéaire, lequel fixe des facteurs de pondération correspondants pour la combinaison linéaire de la pluralité des valeurs de corrélation pour les positions de mesure respectives. Dans une autre variante de l'invention, au lieu d'un procédé de régression linéaire, on utilise une régression SVM (SVM = Support Vector Machine-Machine à vecteurs de support) connue en tant que telle.

Le symptôme clinique prédéterminé, lequel est considéré dans le cadre du procédé décrit plus haut, signifie en particulier la présence de cellules tumorales déterminées. C'est ainsi que le symptôme clinique concerne une maladie tumorale. Le symptôme clinique peut tout aussi bien signifier, le cas échéant, la présence d'îlots d'insuline. Dans ce cas, le symptôme clinique concerne une maladie associée au diabète.

En plus du procédé décrit plus haut pour l'analyse assistée par ordinateur d'une ou plusieurs coupe(s) de tissu, l'invention concerne en

outre un procédé pour la réalisation assistée par ordinateur d'un test de coloration numérique en faisant appel à des spectres de base et à un modèle de régression, lesquels ont été déterminés avec le procédé d'analyse ci-dessus pour un symptôme clinique prédéterminé et enregistrés numériquement. Dans ce test de coloration numérique, on exécute les étapes i) à iii) expliquées dans ce qui suit, les étapes n'impliquant à nouveau pas forcément une suite chronologique pour leur exécution.

Dans une étape i), on importe des données hyperspectrales résolues spatialement d'une coupe de tissu prescrite à tester, les données hyperspectrales contenant des spectres respectifs pour une pluralité de positions de mesure dans la coupe de tissu prescrite, un spectre respectif représentant la composition de la coupe de tissu prescrite au niveau d'une position de mesure sous la forme d'un vecteur de valeurs d'intensité pour des canaux parmi une pluralité de canaux, une valeur d'intensité étant une grandeur physique détectée pour le canal correspondant. De manière préférée, les données hyperspectrales sont de nouveau les données de spectrométrie de masse définies plus haut, ou des données de spectrométrie de masse MALDI.

Dans une étape ii), les vecteurs des données hyperspectrales ou des vecteurs dérivés, lesquels sont calculés à partir des données hyperspectrales, sont mis en corrélation avec les spectres de base au niveau de chaque position de mesure, les spectres de base ayant été déterminés avec le procédé d'analyse décrit plus haut ou une variante préférée de ce procédé. De cette manière, on obtient pour chaque position de mesure une pluralité de valeurs de corrélation pour des spectres de base respectifs. La corrélation s'effectue de préférence de nouveau via un produit scalaire de vecteurs normalisés.

Finalement, dans une étape iii), le modèle de régression qui a été défini avec le procédé d'analyse ci-dessus ou une variante préférée de ce procédé, est appliqué aux valeurs de corrélation. Dans le cas d'un modèle de régression linéaire, on combine linéairement entre elles les valeurs de corrélation pour une position de mesure respective sur la base des facteurs de pondération du modèle. Grâce au modèle de régression, on obtient un masque qui contient des valeurs scalaires respectives pour la pluralité des positions de mesure, des zones de la coupe de tissu prescrite étant des zones de tissu avec le symptôme clinique prédéterminé avec des valeurs scalaires qui divergent par rapport au reste de la coupe de tissu.

Avec le procédé venant d'être évoqué, il est donc possible de générer des masques numériques qui contiennent des informations du genre d'un test de coloration. Dans ce cas, on génère de préférence une visualisation sur un écran ou afficheur, de manière à générer une image à partir des points d'image représentant des positions de mesure, un point d'image respectif codant la valeur scalaire du masque déterminé à l'étape iii) pour la position de mesure représentée par la point d'image respectif. De cette manière, on obtient une représentation numérique qui correspond à un test de coloration classique.

Dans une variante du test de coloration venant d'être décrit, on utilise des résultats de celui des modes de réalisation du procédé d'analyse selon l'invention qui repose sur des vecteurs dérivés. Dans cette variante du test de coloration, à l'étape ii), des vecteurs dérivés sont mis en corrélation avec les spectres de base, les vecteurs dérivés étant déterminés de la manière qui suit.

Pour le symptôme clinique prédéterminé, des données hyperspectrales pour une pluralité de compositions moléculaires sont prescrites, les compositions moléculaires étant définies de la même manière que dans le procédé d'analyse ci-dessus. Les données hyperspectrales pour une composition moléculaire contiennent ainsi un vecteur sous la forme d'un spectre de valeurs d'intensité caractéristique pour la composition moléculaire pour un nombre prédéterminé de canaux parmi une pluralité de canaux. Les spectres caractéristiques pour la composition moléculaire sont mis en corrélation avec les spectres des données hyperspectrales importées à l'étape i) au niveau des positions de mesure respectives, moyennant quoi l'on détermine, en tant que vecteur dérivé pour chaque position de mesure, un vecteur à partir de plusieurs coefficients de corrélation, un coefficient de corrélation pour chaque composition moléculaire existant pour une position de mesure respective.

En plus du procédé décrit plus haut, l'invention concerne en outre un dispositif pour l'analyse assistée par ordinateur d'une ou plusieurs coupe(s) de tissu du corps humain ou animal pour la préparation d'un test de coloration numérique, le dispositif étant conçu pour la réalisation du procédé d'analyse décrit plus haut ou d'une ou plusieurs variantes préférées de ce processus d'analyse. En d'autres termes, le dispositif comprend des moyens pour la réalisation des étapes a) à e) du procédé d'analyse selon l'invention.

L'invention concerne en outre un dispositif pour la réalisation assistée par ordinateur d'un test de coloration numérique en utilisant des spectres de base et un modèle de régression, lesquels ont été déterminés avec le procédé d'analyse ci-dessus ou une ou plusieurs variantes préférées de ce procédé pour un symptôme clinique prédéterminé et ont été enregistrés numériquement. Dans ce cas, le dispositif est conçu pour la réalisation du test de coloration numérique ci-dessus ou d'une ou plusieurs variantes préférées de ce test de coloration. En d'autres termes, le dispositif comprend des moyens pour la réalisation des étapes i) à iii) du test de coloration selon l'invention.

L'invention concerne en outre un produit de programme informatique avec un code de programme enregistré sur un support lisible par machine pour la mise en œuvre d'un procédé pour l'analyse assistée par ordinateur d'une ou plusieurs coupe(s) de tissu ou la réalisation assistée par ordinateur d'un test de coloration numérique selon l'invention ou de variantes préférées de l'invention, lorsque le code de programme est exécuté sur un ordinateur.

L'invention concerne en outre un programme informatique avec un code de programme pour la mise en œuvre d'un procédé pour l'analyse assistée par ordinateur d'une ou plusieurs coupe(s) de tissu ou pour la réalisation assistée par ordinateur d'un test de coloration numérique selon l'invention ou de modes de réalisation préférés de l'invention, lorsque le code de programme est exécuté sur un ordinateur.

Des exemples de réalisation de l'invention sont décrits en détails dans ce qui suit à l'aide des figures annexées, sur lesquelles on peut voir :

sur la figure 1, un diagramme qui montre des étapes exécutées selon un mode de réalisation du procédé d'analyse selon l'invention ;

sur la figure 2, un diagramme qui montre les étapes exécutées selon un mode de réalisation du test de coloration numérique selon l'invention ;

sur la figure 3, une représentation qui reflète les spectres de base déterminés selon un mode de réalisation du procédé d'analyse selon l'invention ;

sur la figure 4, une représentation qui reflète la détermination d'un modèle de régression sur la base des spectres de base de la figure 3 ; et

sur la figure 5, une visualisation qui est générée conformément à un mode de réalisation du test de coloration numérique selon l'invention.

On explique dans ce qui suit des modes de réalisation de l'invention sur la base de données hyperspectrales sous la forme des données de spectrométrie de masse MALDI déjà mentionnées en préambule. Dans le cadre de cette spectrométrie, les molécules de la coupe de tissu à analyser sont recouvertes d'une solution matricielle, et un faisceau laser (par exemple un laser d'azote) est dirigé sur différentes positions de mesure du porte-échantillon correspondant. Par l'intermédiaire des impulsions laser du faisceau laser, il s'en suit un détachement de particules de molécules ionisées du tissu, dont le rapport masse/charge est détecté d'une manière connue par ailleurs avec un spectromètre de masse. Pour ce faire, on peut par exemple utiliser un spectromètre de masse TOF (TOF = Time of Flight/temps de vol) ou un spectromètre FTICR (FTICR = Fourier Transform Ion Cyclotron Resonance/résonnance cyclotonique d'ions à transformée de Fourier). Les procédés de l'invention décrits ici peuvent cependant également s'appliquer le cas échéant à d'autres données de spectrométrie de masse, dans la mesure où les données à résolution locale dans différentes positions de mesure contiennent des spectres lesquels, en fonction du rapport masse/charge, rendent compte du nombre d'ions détectés avec le spectromètre de masse. Le procédé de l'invention peut en outre également s'appliquer de manière générale à des données hyperspectrales (par exemple spectroscopie Raman) dans la mesure où les données résolues localement au niveau de différentes positions de mesure contiennent des vecteurs d'intensité avec des enregistrements pour différents canaux. Une intensité peut par exemple comprendre une fraction de lumière réfléchie ou diffusée d'une longueur d'onde déterminée, la longueur d'onde représentant le canal.

On explique tout d'abord dans ce qui suit un procédé pour l'analyse d'une coupe de tissu à l'aide de la figure 1, dont les résultats sont ensuite utilisés dans le cadre d'un test de coloration numérique, comme décrit ensuite à l'aide de la figure 2. Les figures 3 à 5 montrent des résultats du procédé d'analyse ou du test de coloration, lesquels ont été déterminés à l'aide d'une coupe de tissu du pancréas d'un rat. Les coupes de tissu correspondantes comprennent des îlots d'insuline dans le pancréas, de sorte que les coupes de tissu correspondent au symptôme clinique du diabète.

Dans le cadre du mode de réalisation du procédé d'analyse décrit ici, on considère une coupe de tissu T1 prédéterminée selon la figure 1, laquelle a été analysée au moyen de la spectrométrie MALDI ci-dessus. Il en résulte qu'il y a des spectres S_{ij} respectifs au niveau d'une pluralité de positions de mesure x bidimensionnelles de la coupe de tissu, i, j indiquant les coordonnées cartésiennes de la position de mesure respective. Un spectre individuel contient, pour une pluralité de canaux sous la forme de valeurs m/z (c.à.d. rapports masse/charge) des valeurs d'intensité I qui représentent respectivement le nombre de molécules détectées pour la valeur m/z correspondante. Le procédé décrit dans ce qui suit peut être transposé à l'analyse de plusieurs coupes de tissu, auquel cas les positions de mesure se poursuivent par-delà les coupes de tissu. Les positions de mesure de différentes coupes de tissu peuvent être différenciées les unes des autres, par exemple grâce un index pour la coupe de tissu respective.

Dans une variante préférée, avant l'exécution de la spectrométrie MALDI, la coupe de tissu T1 a été soumise à une digestion où les molécules de la coupe de tissu ont été décomposées de manière adéquate via des processus biochimiques. Pour ce faire, des enzymes sont appliquées sur la coupe de tissu, lesquelles découpent les molécules de la coupe de tissu en peptides. Par l'intermédiaire d'une digestion, il est plus facile de déterminer des symptômes cliniques dans une coupe de tissu correspondante à l'aide de données de spectrométrie de masse. Dans le cadre de l'analyse de la coupe de tissu T1, en plus des données de spectrométrie de masse S_{ij}, on a également besoin d'un masque (numérique) M_{ij} de la coupe de tissu correspondante. Ce masque indique, pour les positions de mesure x individuelles de la coupe de tissu, une valeur scalaire respective, par exemple comprise entre 0 et 1. La grandeur de la valeur scalaire indique la grandeur de la fraction des molécules correspondant au symptôme clinique prédéfini au niveau de la position de mesure respective. Le cas échéant, les valeurs scalaires peuvent également seulement adopter les valeurs 0 et 1, la valeur de 0 indiquant une zone de tissu sans symptôme clinique, et la valeur 1 indiquant une zone de tissu avec symptôme clinique. Le masque M_{ij} peut par exemple être obtenu à partir d'un test de coloration immunohistochimique classique où l'ajout d'anticorps dans la coupe de tissu permet de mettre en avant visuellement, grâce à une coloration du tissu, des biomarqueurs spécifiques pour le symptôme clinique prédéterminé, l'intensité de la coloration du tissu indiquant la fraction des

biomarqueurs au niveau de la position correspondante du tissu. Les résultats du test de coloration immunohistochimique sont ensuite convertis en données numériques du masque M_{ij}.

Comme on vient de l'évoquer, les données d'entrée du procédé d'analyse décrit ici sont les données de spectrométrie de masse S_{ij} ainsi que le masque M_{ij}, ce qui est indiqué dans l'étape S1 de la figure 1. Ces données sont importées sous forme numérique, l'étape de détermination des données et en particulier la détection spectrométrique de masse des spectres S_{ij} pouvant en particulier également faire partie du procédé d'analyse le cas échéant.

Dans une étape S2 du procédé d'analyse, on détermine ensuite, à partir des spectres S_{ij}, des signatures spectrales sous la forme de spectres de base pour différencier des structures métaboliquement pertinentes. Sur la figure 1, on montre en guise d'exemple les cinq spectres de base BS1, BS2, ..., BS5, plus ou moins de spectres de base pouvant être déterminés selon le mode de réalisation. Les spectres de base sont calculés avec un procédé connu par ailleurs pour la détermination de vecteurs de base dans l'espace vectoriel des données spectrales. Dans ce cas, un vecteur respectif dans l'espace vectoriel représente un spectre S_{ij} individuel dans une position de mesure correspondante, et les enregistrements du vecteur sont les valeurs d'intensité des valeurs m/z. Les spectres de base représentent les vecteurs de base déterminés avec la méthode. Les spectres de base effectuent en particulier l'approximation d'une base dans laquelle les spectres S_{ij} sont décrits en tant que combinaison linéaire des spectres de base. Dans le mode de réalisation décrit ici, on utilise une factorisation matricielle non négative avec la condition d'une occupation éparse des matrices qui en sont dérivées pour la détermination des vecteurs de base. Dans ce cas, la factorisation matricielle non négative des documents [1] et [2] est utilisée. Les éléments de la terminologie du document [1] correspondent aux vecteurs de base.

Dans une étape S3 suivante, les spectres S_{ij} d'origine sont mis en corrélation avec les spectres de base BS_k déterminés à l'étape S2 sur la base d'un produit scalaire. Cela signifie que des valeurs de corrélation P_{ij} correspondantes pour chaque spectre S_{ij} et chaque vecteur de base BS_k sont obtenues de la manière suivante :

$$P_{ijk} = \langle S_{ij}; BS_k \rangle.$$

La valeur de corrélation P_{ijk} est normalisée à 1 et est comprise entre -1 et 1. Elle correspond à la valeur 1 en cas de deux vecteurs identiques. Cela est obtenu en ce que les vecteurs S_{ij} et BS_k sont divisés par leur montant avant la formation du produit scalaire et sont donc normalisés de la sorte.

Dans une étape S4 finalement, on détermine, au moyen d'un procédé de régression, un modèle de régression RM qui décrit le masque M_{ij} d'origine sur la base des valeurs de corrélation P_{ijk}. Dans le mode de réalisation décrit ici, on utilise un procédé de régression linéaire où le masque M_{ij} est soumis à une approximation grâce à une combinaison linéaire des valeurs de corrélation, c'est-à-dire que ce qui suit s'applique :

$$\sum_k c_k P_{ijk} = M_{ij}.$$

Les facteurs de pondération c_k déterminés à l'aide du procédé de régression linéaire sont les paramètres du modèle de régression RM déterminé. Au lieu d'un procédé de régression linéaire, il est également possible d'utiliser d'autres procédés de régression au choix pour la détermination d'un modèle de régression qui décrit le masque en fonction des valeurs de corrélation. Dans un mode de réalisation particulièrement préféré, on utilise un procédé de régression SVM (SVM = Support Vector Machines/Machines à vecteurs support) connu par ailleurs.

En guise de résultat final du procédé de la figure 1, on obtient finalement le modèle de régression RM avec des paramètres c_k correspondants ainsi qu'une pluralité de spectres de base BS_k. Ces informations sont sauvegardées numériquement afin de pouvoir être ensuite utilisées dans un procédé pour la réalisation d'un test de coloration numérique. Ce procédé est décrit dans ce qui suit à l'aide de la figure 2.

Le point de départ du procédé de la figure 2 est l'étape S1' où l'on importe, pour une pluralité de positions de mesure bidimensionnelles y d'une coupe de tissu TI' à tester, des données spectrométriques de masse SP_{ij} avec des valeurs d'intensité I correspondantes de valeurs m/z respectives. Le cas échéant, la mesure spectrométrique de masse pour la détermination des spectres SP_{ij} peut elle-même faire partie du procédé. Les données peuvent néanmoins déjà être prescrites et extraites d'une mémoire numérique. A l'aide des seules données de spectrométrie de masse, il s'agit

de générer, pour la coupe de tissu TI', un masque MA_{ij} qui code, de manière analogue au masque Mij de la figure 1, via des valeurs scalaires correspondantes, la présence d'un symptôme clinique prédéterminé au niveau de positions de mesure correspondantes.

Dans une étape S2' du test de coloration numérique, on détermine alors, de manière similaire à l'étape S3 de la figure 1, des valeurs de corrélation qui ne reposent cependant plus sur le produit scalaire des spectres S_{ij} et des vecteurs de base BS_k, mais sur le produit scalaire des spectres SP_{ij} et les vecteurs de base BS_k. En d'autres termes, on obtient des valeurs de corrélation N_{ijk} correspondantes de la manière suivante :

$$N_{ijk} = <SP_{ij}; BS_k>.$$

Par analogie avec les valeurs de corrélation P_{ijk} de la figure 1, les valeurs de corrélation N_{ijk} sont également comprises entre -1 et 1. Les vecteurs de base BS_k traités à l'étape S2' et également le modèle de régression RM traité à l'étape S3' suivante sont extraits d'une mémoire numérique et ont été déterminés avec le procédé de la figure 1.

A l'étape S3' de la figure 2, on détermine ensuite, sur la base des valeurs de corrélation N_{ijk} en utilisant le modèle de régression RM ou les paramètres c_k de celui-ci, le masque MA_{ij} avec les valeurs scalaires de celui-ci, à savoir que l'on détermine de quelle manière les valeurs de corrélation des spectres de l'échantillon de tissu TI' à tester se superposent conformément au modèle de régression et forment ainsi les données de masque MA_{ij}. En d'autres termes, les données de masque sont déterminées de la manière suivante :

$$MA_{ij} = \sum_k c_k N_{ijk}$$

De cette manière, de manière assistée par ordinateur, en faisant appel aux résultats du procédé d'analyse de la figure 1, des zones de tissu avec un symptôme clinique prédéterminé sont marquées de manière correspondante. On visualise de préférence le masque MA_{ij} sous la forme d'une image numérique sur un afficheur. Cette image montre la coupe de tissu TI' sous la forme d'une pluralité de pixels qui correspondent respectivement à une position de mesure de la coupe de tissu. Dans ce cas,

chaque pixel d'image assure le codage, via sa valeur de gris ou le cas échéant sa valeur chromatique, de la valeur scalaire correspondante du masque MA_{ij}, une image de coloration étant ainsi obtenue, comme cela est déterminé habituellement au moyen de tests de coloration immunohistochimiques. Selon l'invention, il n'est cependant pas nécessaire d'exécuter des étapes biochimiques du procédé, et la valeur d'image de coloration est plutôt dérivée de manière purement assistée par ordinateur à partir de données de spectrométrie de masse.

Les procédés décrits à l'aide de la figure 1 et de la figure 2 constituent une dénommée approche bottom-up où l'on détermine, sans avoir connaissance des compositions moléculaires pertinentes dans l'échantillon de tissu, les spectres de base grâce à un pur procédé assisté par ordinateur. Dans une variante modifiée (approche top-down/haut-bas), il est également possible de déterminer les vecteurs de base avec la connaissance préalable sur les biomarqueurs ou des peptides correspondants en lesquels les biomarqueurs se décomposent par digestion. Dans ce cas, on a déterminé à l'avance ou simulé de quelle manière les biomarqueurs connus se décomposent en peptides sous l'effet de la digestion utilisée (par exemple trypsine). Pour la composition chimique d'un peptide respectif, on a par ailleurs déterminé un spectre de masse, lequel indique pour les valeurs m/z des isotopes du peptide, respectivement une valeur d'intensité sous la forme du nombre de molécules détectées. Des procédés pour la détermination des peptides correspondants et de leurs spectres sont connus en tant que tels. Etant donné que ces procédés ne font pas l'objet de l'invention et que l'on utilise simplement leurs résultats sous la forme de spectres de peptides, ils ne sont pas décrits en détail.

Conformément à la variante modifiée, les spectres individuels des peptides sont mis en corrélation avec les spectres S_{ij} des données de spectrométrie de masse de la figure 1 dans une étape de traitement préalable. La corrélation est de nouveau effectuée au moyen d'un produit scalaire, les valeurs m/z (normalisées) des spectres des peptides étant multipliées par les valeurs m/z (normalisées) correspondantes des spectres S_{ij} pour la formation du produit scalaire, et ces produits étant totalisés. Les valeurs de corrélation résultantes, lesquelles sont désignées comme étant des coefficients de corrélation pour établir la distinction par rapport aux valeurs de corrélation ci-dessus, sont à nouveau comprises entre 0 et 1. En guise de résultat, on obtient finalement, pour chaque position de mesure, un

nombre de coefficients de corrélation correspondant au nombre de peptides, lesquels constituent des vecteurs dérivés au sens des revendications. Pour ces vecteurs dérivés à partir de coefficients de corrélation on utilise, de la même manière que sur la figure 1, un procédé pour la détermination de vecteurs de base et en particulier une factorisation matricielle non négative, les vecteurs de base étant alors déterminés dans l'espace vectoriel des vecteurs dérivés. Ces vecteurs de base sont alors mis en corrélation avec les vecteurs dérivés au lieu des spectres d'origine, ce qui donne de nouveau des valeurs de corrélation lesquelles, de manière analogue à ce qui est décrit plus haut, sont déterminées d'une manière telle, qu'elles sont comprises entre 0 et 1. Par analogie avec l'étape S4 de la figure 1, on détermine ensuite un modèle de régression sur la base de ces valeurs de corrélation, lequel décrit le masque M_{ij}.

Si l'on utilise ensuite les résultats du procédé d'analyse du mode de réalisation modifié venant d'être décrit dans le cadre du test de coloration numérique, avant l'exécution de l'étape S2' de la figure 2, on détermine tout d'abord des vecteurs dérivés avec des coefficients de corrélation correspondants par mise en corrélation des spectres de peptides avec les spectres d'origine SP_{ij}. Ensuite, par analogie avec l'étape S2', il y a la détermination des valeurs de corrélation, les vecteurs dérivés étant cependant alors utilisés. Finalement, tout comme dans l'étape S3' de la figure 2, le masque MA_{ij} est déterminé avec le modèle de régression.

Comme déjà évoqué en préambule, les figures 3 à 5 montrent des résultats d'un mode de réalisation du procédé d'analyse selon l'invention ou du test de coloration selon l'invention pour la coupe de tissu du pancréas d'un rat, le pancréas contenant des îlots d'insuline. Dans ce cas, la coupe de tissu a été analysée sur la base d'une approche bottom-up. On a déterminé un total de dix spectres de base à partir des spectres d'origine S_{ij}, cinq de ces spectres de base BS_1 à BS_5 étant montrés en guise d'exemple sur la figure 3. Pour les spectres de base individuels, on a déterminé les valeurs de corrélation correspondantes et sur ces bases, les coefficients c_k d'un modèle de régression linéaire. Cela est indiqué par la figure 4.

Sur le diagramme en haut à droite de la figure 4, sont représentés les dix facteurs de pondération c_k du modèle de régression pour les vecteurs de base $k = 1,\ldots,10$. On indique en outre sur le diagramme, en bas à droite, la superposition des spectres de base conformément aux facteurs de pondération ck, c.à.d. $\sum c_k BS_k$. On représente par ailleurs, sur

l'image numérique du côté gauche, un extrait de la coupe de tissu avec des pixels correspondants en tant que positions d'image x, les bordures blanches, lesquelles sont seulement partiellement indiquées par la lettre de référence L pour des raisons de visibilité, délimitant des zones dans lesquelles il y a des îlots d'insuline dans la coupe de tissu. Par l'intermédiaire des bordures L, on indique ainsi les données de masque d'origine, lesquelles adoptent seulement les valeurs d'échelle 0 et 1 dans le présent cas, la valeur d'échelle 0 représentant une zone sans insuline et la valeur d'échelle 1, une zone d'îlot d'insuline. Les valeurs de pixel individuelles de l'image montrée représentent les données de masque M_{ij} reconstituées d'après le modèle de régression déterminé, les valeurs d'échelle des pixels individuels étant codées en couleur, ce qui n'apparaît pas sur la représentation en noir et blanc de la figure 4. Des valeurs de pixels sombres représentent néanmoins des valeurs d'échelle élevées dans la plage de 1. C'est ainsi que l'on peut reconnaître que la reconstitution restitue bien les données de masque d'origine.

Le modèle de régression avec les facteurs de pondération c_k correspondants de la figure 4 a finalement été appliqué à une nouvelle coupe de tissu du pancréas du rat. En d'autres termes, on a déterminé des données spectrométriques de masse MALDI pour la nouvelle coupe de tissu et ensuite, sur la base des étapes S2' et S3' de la figure 2, les données de masque MA_{ij} ont été dérivées. Ces données de masque sont indiquées sur la figure 5 pour la nouvelle coupe de tissu, les positions de mesure y étant à nouveau des pixels individuels à codage en couleur, ce qui n'apparaît cependant pas d'après la représentation en noir et blanc de la figure 5. Des valeurs de pixel noires représentent cependant des valeurs d'échelle élevées. Sur la figure 5, on indique de nouveau grâce à des bordures blanches L correspondantes (seulement partiellement indiquées par cette référence) des îlots d'insuline tels qu'ils ont pu être effectivement identifiés dans la coupe de tissu. Comme on peut le reconnaître, les données de masque MA_{ij} montrées représentent très bien les îlots d'insuline effectivement présents dans la coupe de tissu.

Les modes de réalisation de l'invention venant d'être décrits présentent une série d'avantages. Pour la première fois, des coupes de tissu peuvent en particulier être analysées uniquement sur la base de méthodes de spectrométrie de masse pour en dériver ainsi des zones pathologique du tissu. C'est ainsi que le procédé peut remplacer, au moins en partie, des

tests immunohistochimiques ou d'autres procédés de tests moléculaires. Pour ce faire, il suffit simplement de déterminer, sur la base d'un procédé d'analyse, un modèle de régression et des spectres de base, lesquels sont ensuite utilisés lors de l'évaluation des données de spectrométrie de masse. De cette manière, on obtient un test de coloration numérique qui identifie des zones de tissu malade dans une coupe de tissu. On représente de préférence la coupe de tissu pour un pathologiste sous la forme d'une image sur laquelle des zones malades sont représentées sur la base d'un codage en couleur.

Dans la présente description, il est fait référence ci-dessus aux documents suivants :

[1] : DE 10 2013 207 402 A1

[2] : J.H. Kobarg et al., "Numerical experiments with MALDI Imaging data" (Expériences numériques avec des données images MALDI), Adv Comput Math (2014) 40, pages 667 à 682, novembre 2013

Bien entendu, l'invention n'est pas limitée aux modes de réalisation décrits et représentés aux dessins annexés. Des modifications restent possibles, notamment du point de vue de la constitution des divers éléments ou par substitution d'équivalents techniques, sans sortir pour autant du domaine de protection de l'invention.

REVENDICATIONS

1. Procédé pour l'analyse assistée par ordinateur d'une ou plusieurs coupe(s) de tissu (TI) du corps humain ou animal pour la préparation d'un test de coloration numérique, dans lequel, dans le test de coloration numérique, on détermine de manière assistée par ordinateur des zones de tissu dans une coupe de tissu (TI') testée, lesquelles présentent un symptôme clinique prédéterminé, dans lequel :

a) des données hyperspectrales à résolution spatiale d'une ou plusieurs coupe(s) de tissu (TI) prédéterminée(s) à analyser qui présentent des zones de tissu avec le symptôme clinique prédéterminé sont importées, dans lequel les données hyperspectrales contiennent des spectres (S_{ij}) respectifs pour une pluralité de positions de mesure (x) dans la ou les coupe(s) de tissu prédéterminée(s), dans lequel un spectre (S_{ij}) respectif représente la composition de la ou des coupe(s) de tissu prédéterminée(s) (TI) au niveau d'une position de mesure (x) sous la forme d'un vecteur de valeurs d'intensité (I) pour une pluralité de canaux (m/z), dans lequel une valeur d'intensité (I) est une grandeur physique détectée pour le canal (m/z) correspondant;

b) un masque (M_{ij}) de la ou des coupe(s) de tissu prédéterminée(s) (TI) étant importé, lequel contient des valeurs d'échelle respectives pour la pluralité des positions de mesure (x), dans lequel des zones de la ou des coupe(s) de tissu prédéterminée(s) (TI) avec des valeurs d'échelle divergentes par rapport à la zone restante de la ou des coupe(s) de tissu (TI) sont des zones de tissu avec le symptôme clinique prédéterminé;

c) grâce à une méthode pour la détermination de vecteurs de base, plusieurs spectres de base (BS_1, BS_2, ..., BS_5) dans l'espace des vecteurs des données hyperspectrales ou dans un espace de vecteurs dérivés, lesquels sont calculés à partir des données hyperspectrales pour chaque position de mesure (x), étant déterminés ;

d) les spectres de base (BS_1, BS_2, ..., BS_5) avec les vecteurs des données hyperspectrales ou les vecteurs dérivés étant mis en corrélation au niveau de chaque position de mesure (x), moyennant quoi on obtient, pour chaque position de mesure (x), une pluralité de valeurs de corrélation (P_{ijk}) pour des spectres de base (BS_1, BS_2, ..., BS_5) respectifs;

e) un modèle de régression (RM) étant déterminé à l'aide d'un procédé de régression, lequel décrit le masque (M_{ij}) sur la base de la pluralité des valeurs de corrélation (P_{ijk}) pour les positions de mesure (x) respectives, dans lequel le modèle de régression (RM) et les spectres de base (BS_1, BS_2, ..., BS_5) sont enregistrés pour la réalisation du test de coloration numérique.

2. Procédé selon la revendication 1, caractérisé en ce que les données hyperspectrales sont des données spectrométriques de masse, en particulier des données spectrométriques de masse MALDI où les canaux (m/z) représentent des rapports masse/charge et la grandeur physique détectée pour le canal (m/z) correspondant représentant le nombre de molécules détectées pour le rapport masse/charge correspondant.

3. Procédé selon l'une des revendications précédentes, caractérisé en ce que, à l'étape c), les spectres de base (BS_1, BS_2, ..., BS_5) sont déterminés dans l'espace des vecteurs dérivés, dans lequel un vecteur dérivé est calculé pour chaque position de mesure (x) de telle sorte que

- pour le symptôme clinique prédéterminé, des données hyperspectrales pour une pluralité de compositions moléculaires sont prescrites, dans lequel les données hyperspectrales pour une composition moléculaire contiennent un vecteur sous la forme d'un spectre caractéristique de la composition moléculaire de valeurs d'intensité (I) pour un nombre prescrit de canaux (m/z) parmi une pluralité de canaux (m/z);
- les spectres caractéristiques des compositions moléculaires étant mis en corrélation avec les spectres (S_{ij}) des données hyperspectrales importées au niveau des positions de mesure (x) respectives, moyennant quoi on détermine, en guise de vecteur dérivé pour chaque position de mesure (x), un vecteur à partir de plusieurs coefficients de corrélation.

4. Procédé selon l'une des revendications précédentes, caractérisé en ce que la ou les coupe(s) de tissu à analyser sont des coupes de tissu respectives avec digestion.

5. Procédé selon l'une des revendications précédentes, caractérisé en ce que la méthode utilisée à l'étape c) pour la détermination de vecteurs de base est une factorisation matricielle non négative ou une analyse PCA.

6. Procédé selon la revendication 5, caractérisé en ce que la factorisation matricielle non négative comprend la condition de

l'occupation éparse des matrices déterminées par la factorisation matricielle non négative.

7. Procédé selon l'une des revendications précédentes, caractérisé en ce que le procédé de régression utilisé à l'étape e) est un procédé de régression linéaire ou un procédé de régression SVM.

8. Procédé selon l'une des revendications précédentes, caractérisé en ce qu'à l'étape b), le masque (M_{ij}) est déterminé à partir d'un test de coloration histologique, et en particulier immunohistochimique, de la ou des coupe(s) de tissu prédéterminée(s).

9. Procédé selon l'une des revendications précédentes, caractérisé en ce que le symptôme clinique prédéterminé représente la présence de cellules tumorales prédéterminées ou la présence d'îlots d'insuline dans la ou les coupe(s) de tissu prédéterminée(s) (TI).

10. Procédé pour la réalisation assistée par ordinateur d'un test de coloration numérique en utilisant des spectres de base (BS_1, BS_2, ..., BS_5) et un modèle de régression (RM), lesquels ont été déterminés avec un procédé selon l'une des revendications précédentes pour un symptôme clinique prédéterminé et enregistrés numériquement, dans lequel :

i) des données hyperspectrales à résolution spatiale d'une coupe de tissu (TI') prescrite à tester sont importées, dans lequel les données hyperspectrales contiennent des spectres (SP_{ij}) respectifs pour une pluralité de positions de mesure (y) dans la coupe de tissu (TI') prescrite, dans lequel un spectre (SP_{ij}) respectif représente la composition de la coupe de tissu (TI') prescrite au niveau d'une position de mesure (y) sous la forme d'un vecteur de valeurs d'intensité (I) pour des canaux (m/z) parmi une pluralité de canaux (m/z), dans lequel une valeur d'intensité (I) est une grandeur physique détectée pour le canal (m/z) correspondant;

ii) les vecteurs des données hyperspectrales ou des vecteurs dérivés, lesquels ont été calculés à partir des données hyperspectrales, étant mis en corrélation au niveau de chaque position de mesure (x) avec les spectres de base (BS_1, BS_2, ..., BS_5) ayant été déterminés avec un procédé selon l'une des revendications précédentes, moyennant quoi on obtient, pour chaque position de mesure (y), une pluralité de valeurs de corrélation (N_{ijk}) pour des spectres de base (BS_1, BS_2, ..., BS_5) respectifs;

iii) le modèle de régression (RM), lequel a été déterminé avec un procédé selon l'une des revendications précédentes, étant appliqué aux valeurs de corrélation (N_{ijk}) et un masque (MA_{ij}) étant ainsi obtenu, lequel

contient, pour la pluralité de positions de mesure (y), des valeurs d'échelle respectives, dans lequel des zones de la coupe de tissu (TI') prescrite avec des valeurs d'échelle divergentes par rapport à la coupe de tissu (TI') restante, sont des zones de tissu avec le symptôme clinique prédéterminé.

11. Procédé selon la revendication 10, caractérisé en ce qu'une visualisation est générée sur un afficheur de telle sorte qu'une image est générée à partir de points d'image représentant les positions de mesure (y), dans lequel un point d'image respectif assure le codage de la valeur d'échelle du masque (MA_{ij}) déterminé à l'étape iii) pour la position de mesure (y) représentée par le point d'image (BP) respectif.

12. Procédé selon la revendication 10 ou 11, caractérisé en ce que des vecteurs dérivés à l'étape ii) sont mis en corrélation avec les spectres de base (BS_1, BS_2, ..., BS_5), dans lequel les vecteurs dérivés sont calculés de telle sorte que,

- pour le symptôme clinique prédéterminé, des données hyperspectrales pour une pluralité de compositions moléculaires sont prescrites, dans lequel les données hyperspectrales pour une composition moléculaire contiennent un vecteur sous la forme d'un spectre caractéristique pour la composition moléculaire de valeurs d'intensités (I) pour un nombre prescrit de canaux (m/z) parmi la pluralité de canaux (m/z);
- les spectres caractéristiques pour les compositions moléculaires étant mis en corrélation avec les spectres (S_{ij}) des données hyperspectrales importées à l'étape i) au niveau des positions de mesure (y) respectives, moyennant quoi on détermine, en guise de vecteur dérivé pour chaque position de mesure (y), un vecteur à partir de plusieurs coefficients de corrélation.

13. Dispositif pour l'analyse assistée par ordinateur d'une ou plusieurs coupe(s) de tissu (TI) du corps humain ou animal pour la préparation d'un test de coloration numérique, dans lequel, dans le test de coloration numérique, on détermine de manière assistée par ordinateur des zones de tissu dans une coupe de tissu (TI') testée, lesquelles présentent un symptôme clinique prédéterminé, caractérisé en ce que le dispositif est conçu pour la réalisation d'un procédé selon l'une des revendications 1 à 9.

14. Dispositif pour la réalisation assistée par ordinateur d'un test de coloration numérique en utilisant des spectres de base (BS_1, BS_2, ..., BS_5) et un modèle de régression (RM), lesquels ont été déterminés avec un procédé selon l'une des revendications 1 à 9 pour un symptôme clinique

prédéterminé et enregistrés numériquement, caractérisé en ce que le dispositif est conçu pour la réalisation d'un procédé selon l'une des revendications 10 à 12.

15. Produit de programme informatique avec un code de programme enregistré sur un support lisible par machine pour la réalisation d'un procédé selon l'une des revendications 1 à 12 lorsque le code de programme est exécuté sur un ordinateur.

16. Programme informatique avec un code de programme pour la réalisation d'un procédé selon l'une des revendications 1 à 12 lorsque le code de programme est exécuté sur un ordinateur.

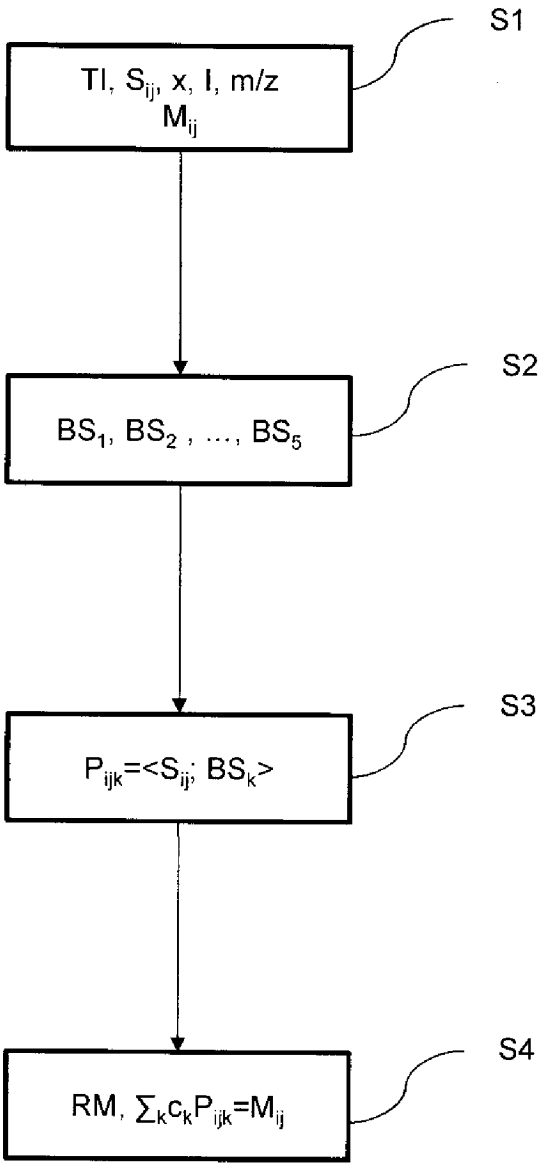

S1
TI, S_{ij}, x, I, m/z
M_{ij}
S2
BS_1, BS_2 , …, BS_5
S3
P_{ijk}=<S_{ij}; BS_k>
S4
RM, $\sum_k c_k P_{ijk}=M_{ij}$

Fig. 1

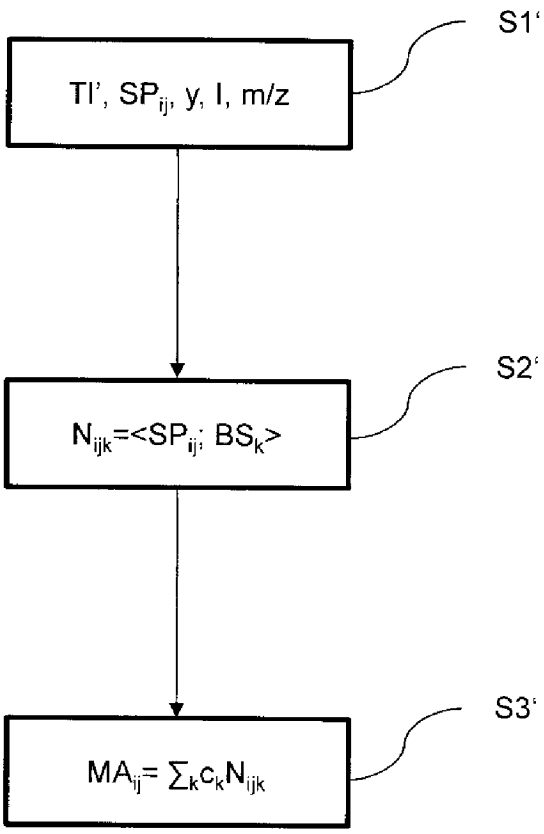


Fig. 2

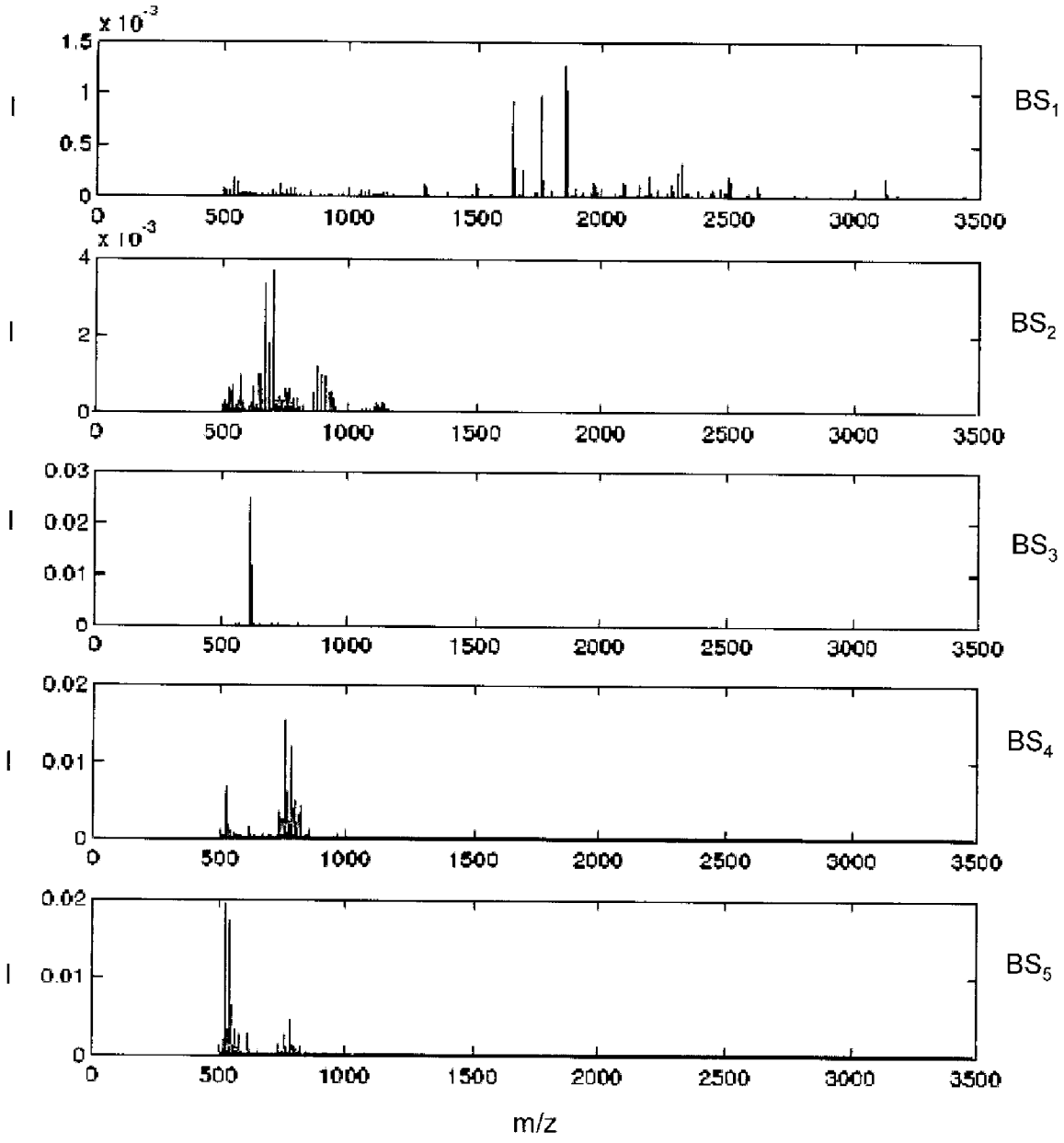

x 10^-3
1.5
1
0.5
0
I
0
500
1000
1500
2000
2500
3000
3500
BS1
x 10^-3
4
2
0
I
BS2
0.03
0.02
0.01
BS3
0.02
0.01
BS4
0.02
0.01
BS5
m/z

Fig. 3

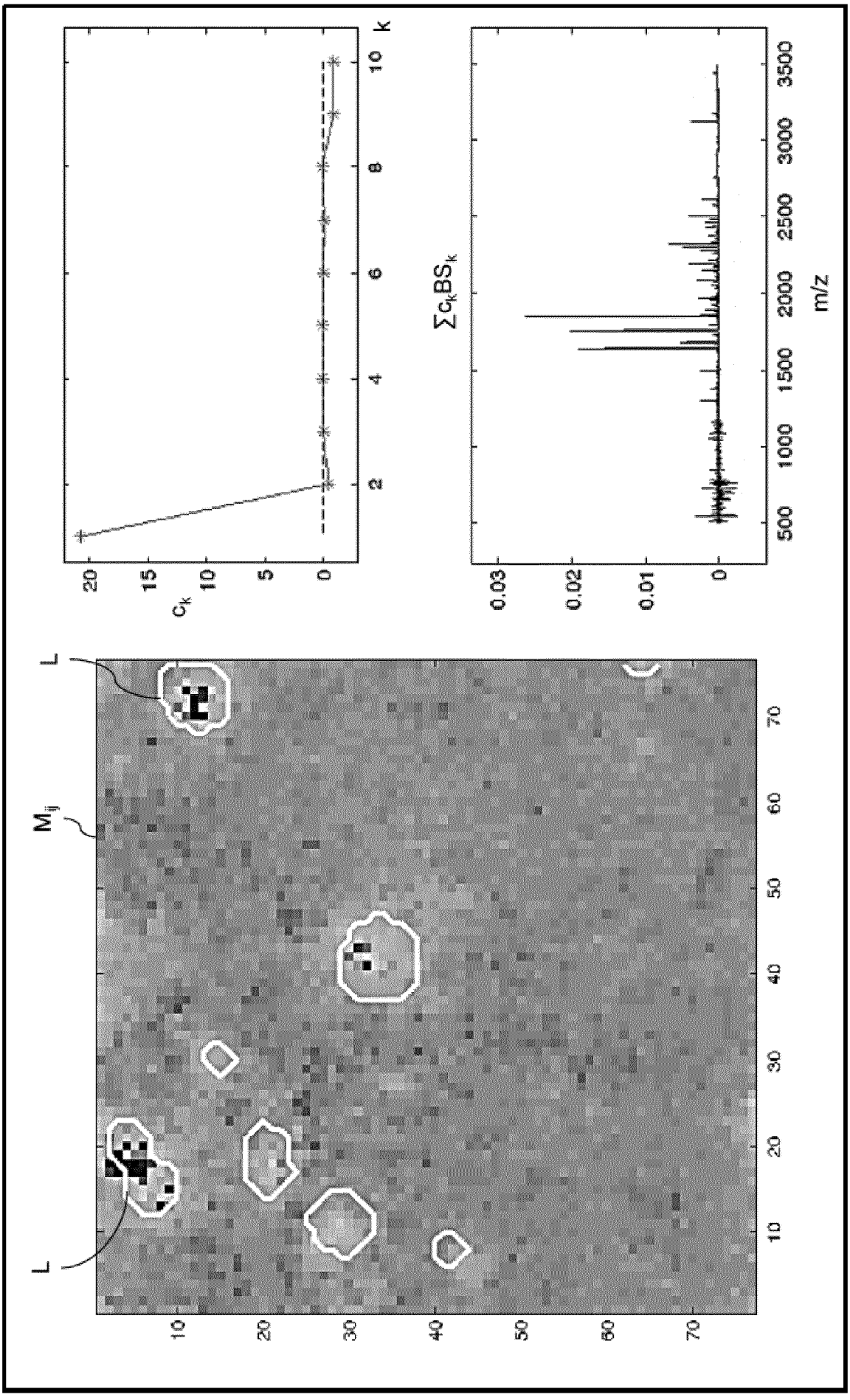

L
L
M_{ij}
10
20
30
40
50
60
70
c_k
20
15
10
5
0
2
4
6
8
10
k
$\sum c_k BS_k$
0.03
0.02
0.01
0
500
1000
1500
2000
2500
3000
3500
m/z

Fig. 4

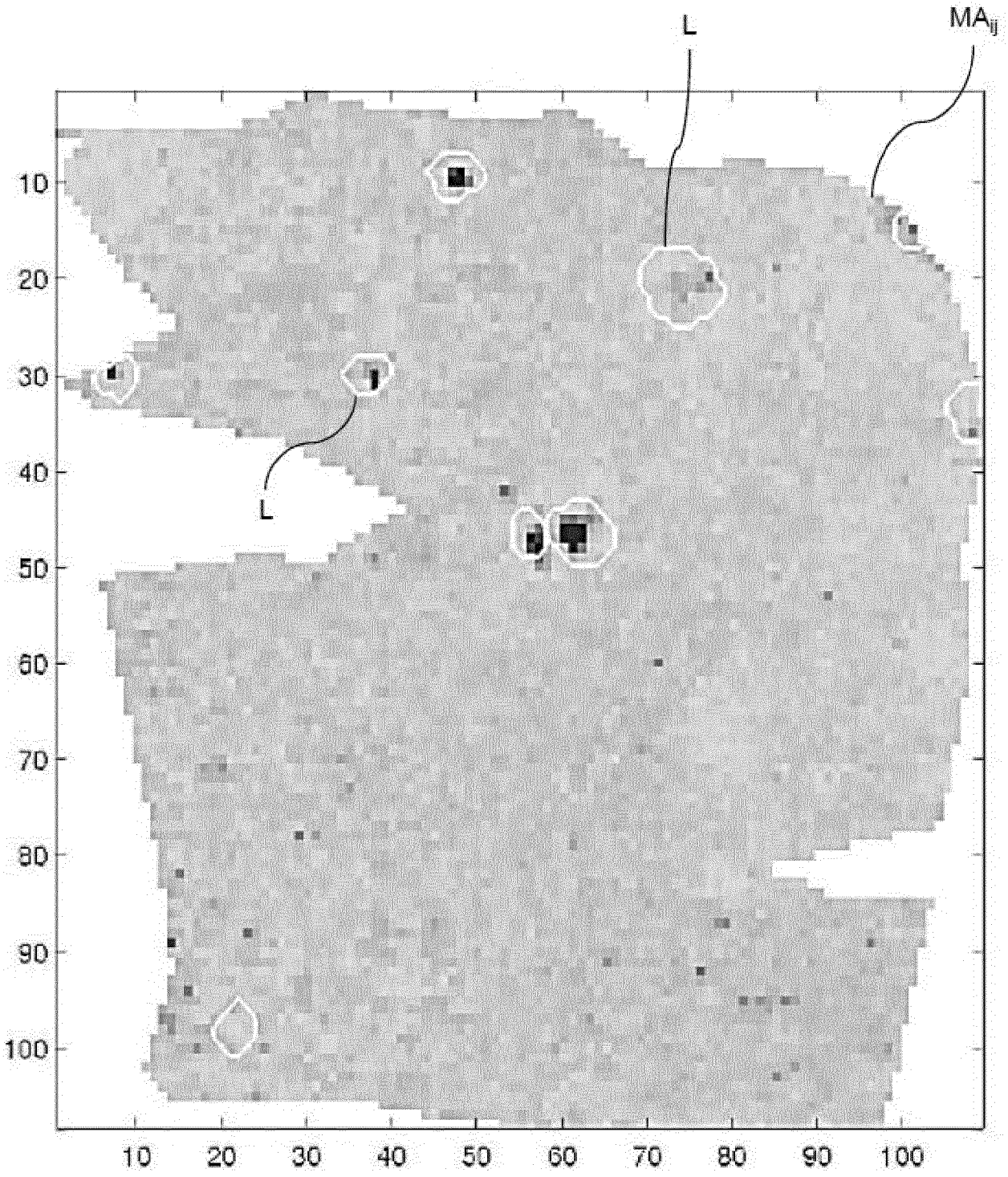


Fig. 5