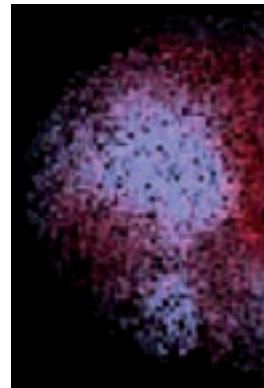


> La complexité de la composition du sérum est un obstacle majeur à son analyse protéomique. Il est possible d'éliminer les deux composantes protéiques les plus importantes du sérum, l'albumine et les immunoglobulines, en utilisant une procédure simple, ne comportant que deux étapes, qui permet aussi d'enrichir la fraction glycoprotéique des échantillons sériques. Cette procédure permet d'accroître la quantité de protéines susceptibles d'être chargées sur un gel et d'identifier des protéines quantitativement moins abondantes. L'application de cette procédure à des échantillons de sujets malades et de sujets sains révèle des profils protéiques très différents, suggérant que cette approche permettra la découverte de nouveaux biomarqueurs. Nous discutons aussi la mise en place de tests cliniquement utiles pour détecter les biomarqueurs ainsi identifiés. <

La déplétion des protéines abondantes à la recherche de biomarqueurs protéiques

Alex J. Rai



Department of Clinical Laboratories, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, Box 88, New York, NY 10021, États-Unis. raia@mskcc.org

accessible par des techniques classiques d'analyse protéomique faisant appel à l'électrophorèse 2D et à la spectrométrie de masse. Ces techniques sont cependant peu appropriées à la détection de biomarqueurs de maladie dans un échantillon de sérum ou de plasma brut car ces biomarqueurs sont souvent parmi les protéines de faible abondance [3]. En effet, on recherche dans le liquide biologique, généralement le plasma, la marque d'un processus cellulaire ou métabolique qui se produit dans une partie peu accessible du corps humain. Lorsque ce processus conduit à la modification d'une protéine abondante du plasma pour produire un produit spécifique, par exemple un fragment de clivage, l'identification de cette modification peut être relativement aisée. Si ce processus touche une protéine de moindre importance, sa détection devient difficile.

Pour contourner cette difficulté, il est possible de réduire la complexité des échantillons protéiques avant l'analyse, de réduire son hétérogénéité et de l'enrichir en un sous-groupe pré-défini de protéines. Cela peut être fait par différentes approches, la plupart permettant d'isoler des classes distinctes de sous-protéomes (phospho-protéomes, glyco-protéomes) à l'intérieur du protéome total.

L'identification de biomarqueurs protéiques dans un échantillon sérique ou plasmatique se heurte à un certain nombre d'obstacles. L'un d'entre eux est la très grande hétérogénéité de la composition protéique du sérum ou du plasma, l'ordre de grandeur entre les protéines les plus abondantes et les protéines les moins abondantes étant d'environ $10 \log$ [1, 2]. Du fait de l'abondance des protéines et de cette hétérogénéité, il est essentiel de savoir quelle partie du protéome est analysée par une méthode donnée. La détection de marqueurs biologiques de maladie fait habituellement appel à la comparaison de profils d'expression entre un état pathologique et l'état sain. Ces profils sont établis par des méthodes d'analyse directe en spectrométrie de masse après sélection de classes de protéines sur des surfaces fonctionnalisées ou sur des billes magnétiques. Une alternative est la comparaison de profils protéiques qui est depuis longtemps

Une stratégie d'enrichissement de l'échantillon en deux étapes

Différentes stratégies ont été imaginées pour enrichir un échantillon de liquide biologique. À titre d'exemple, nous décrirons l'un des schémas que l'on peut suivre pour sélectionner un sous groupe de glycoprotéines dans le sérum. Ce schéma simple combine deux étapes complémentaires : la première étape est une incubation de l'échantillon en présence de protéine G agarose. La seconde étape est une incubation du sérum en présence de concanavaline A sépharose. Les protéines éluées sont ensuite étudiées par électrophorèse (Figure 1).

Les billes de protéine G agarose permettent d'éliminer les immunoglobulines tandis que la concanavaline A sépharose permet d'éliminer l'albumine et d'enrichir l'échantillon en glycoprotéines. Le résultat est la déplétion des deux composants les plus abondants du sérum, c'est-à-dire l'albumine et les immunoglobulines, et l'enrichissement de l'échantillon en protéines glycosylées contenant de l' α -mannose.

Comme il est observé dans la Figure 2, la fraction non liée est très similaire au sérum avant séparation parce que l'albumine, la plus abondante des protéines du sérum, n'est pas glycosylée. Il en résulte une réduction de plus de 50 % du contenu total en protéines, mesuré en poids. Cette réduction concerne au moins 75 % des protéines de grande abondance. La composition protéique de l'échantillon étant simplifiée après ces deux étapes de séparation, il est possible de charger 20 fois plus d'échantillon dans un gel d'électrophorèse.

Un groupe distinct de protéines est enrichi dans la fraction liée à la concanavaline A, ce qui montre que la procédure enrichit

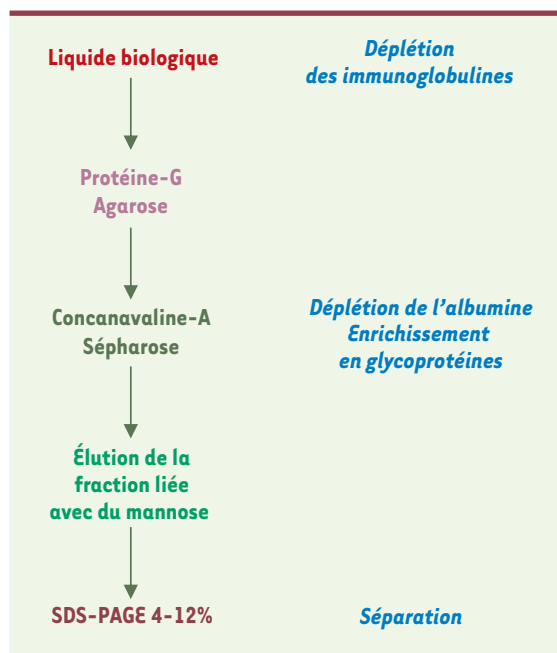


Figure 1. Schéma général de la méthodologie de déplétion et d'enrichissement des protéines.

fondées sur l'affinité des glycoprotéines pour la lectine [4]. Cette approche s'applique aussi bien aux échantillons issus de sujets normaux que ceux issus de malades et dans de nombreuses conditions différentes.

Pour tester directement l'efficacité de notre procédure et son application à la découverte de biomarqueurs, nous avons testé douze échantillons de sujets malades ($n = 6$) et non malades ($n = 6$) et la population enrichie de protéines montre effectivement des différences entre ces deux groupes de sujets, ce qui suggère que la procédure peut être utilisée pour identifier des protéines différentiellement exprimées. Les différences observées concernent notamment des protéines autour de 17 et de 40 kDa. Il y a également une hétérogénéité au niveau des protéines de faible poids moléculaire inférieur à 20 kDa. La nature et la valeur prédictive des marqueurs identifiés sont en cours d'analyse.

L'utilisation clinique des biomarqueurs identifiés

Lorsque les analyses protéomiques cliniques auront permis d'identifier un biomarqueur protéique ou une combinaison de biomarqueurs d'intérêt diagnostique ou pronostique, il faudra mettre au point un test de mesure de ce biomarqueur applicable en médecine quotidienne et gérable dans un laboratoire d'analyses médicales. Les approches technologiques couramment utilisées dans les laboratoires d'analyses médicales sont relativement peu nombreuses. Il s'agit de l'électrophorèse des protéines, des études enzymatiques et spectrophotométriques, des tests immunologiques et des méthodes fondées sur la néphélométrie/turbidimétrie.

L'électrophorèse peut être appliquée à l'étude de n'importe quel liquide biologique. L'électrophorèse des protéines sériques distingue l'albumine des globulines réparties en quatre fractions, $\alpha 1$ -, $\alpha 2$ -, β - et γ -globulines. Dans ce système, les caractéristiques de migration des 22 protéines les plus abondantes du sérum, qui représentent plus de 99 % des protéines totales en poids, sont bien connues. L'électrophorèse des protéines sériques ou urinaires peut être complétée par une immunofixation détectant des protéines spécifiques à l'aide d'anticorps. L'avantage de l'approche électrophorétique pour rechercher une protéine est qu'elle permet de séparer et d'identifier les différentes isoformes de cette protéine. En outre, comme toute autre méthode immunologique, l'immunofixation a le mérite de la spécificité.

Les méthodes enzymatiques et spectrophotométriques sont les plus couramment utilisées dans les laboratoires d'analyses médicales. Ils permettent les analyses chimiques les plus courantes en mesurant le plus souvent les changements d'absorbance qui accompagnent la génération ou la déplétion de molécules comme le NAD ou le NADP au cours d'une réaction enzymatique. L'absorbance est corrélée, directement ou inversement, au niveau d'expression de la molécule d'intérêt. Cette approche a l'inconvénient de ne pas séparer les isoformes d'une même protéine.

Les méthodes immunologiques qui permettent de détecter et de quantifier un biomarqueur sont de nature variée. Les méthodes dites compétitives permettent de détecter des protéines de petite taille mais nécessitent de disposer d'une version purifiée ou recombinante de la molécule d'intérêt et la relation entre l'intensité du signal et la concentration de la molécule analysée est inverse. Les méthodes dites non compétitives peuvent détecter de plus grandes molécules mais nécessitent que la molécule ciblée contienne deux épitopes non chevauchant bien distincts et que des anticorps contre chaque épitope soient disponibles. La relation entre l'intensité du signal et la concentration de la molécule analysée est directe. Enfin, s'il

s'agit de la recherche d'un auto-anticorps, des *immuno-assays* peuvent être utilisés qui nécessitent une version purifiée ou recombinante de la protéine cible, généralement complexée avec une molécule génératrice d'un signal de fluorescence ou de chimioluminescence.

Enfin, la néphélométrie turbidimétrie est basée sur les changements des propriétés de réflexion de la lumière, laquelle est modifiée par la présence d'anticorps liés à leur antigène spécifique, ce qui permet d'établir des corrélations avec la quantité d'antigène.

Pour choisir le test approprié, on devra prendre en compte les caractéristiques de la cible d'intérêt, les avantages et les inconvénients des différentes méthodes, et la possibilité d'adapter les analyseurs à cette nouvelle mesure.

Conclusions

En conclusion, la déplétion des deux catégories de protéines les plus abondantes du sérum permet un enrichissement de l'échantillon analysé en glycoprotéines. C'est une procédure qui, comme toutes les autres procédures de séparation, permet d'analyser des sous classes du protéome humain plus facilement qu'en étudiant du sérum ou du plasma non déplété et facilite la découverte de nouveaux biomarqueurs. Lorsqu'un biomarqueur est identifié, son application au diagnostic et au suivi thérapeutique impose de choisir parmi les différentes méthodes utilisées couramment dans les laboratoires d'analyses médicales celle qui sera la plus appropriée à sa détection. ♦

SUMMARY

Abundant protein' depletion in search of biomarkers

Using a simple two step procedure, we simultaneously removed the two most abundant components of serum, albumin and immunoglobulins, and

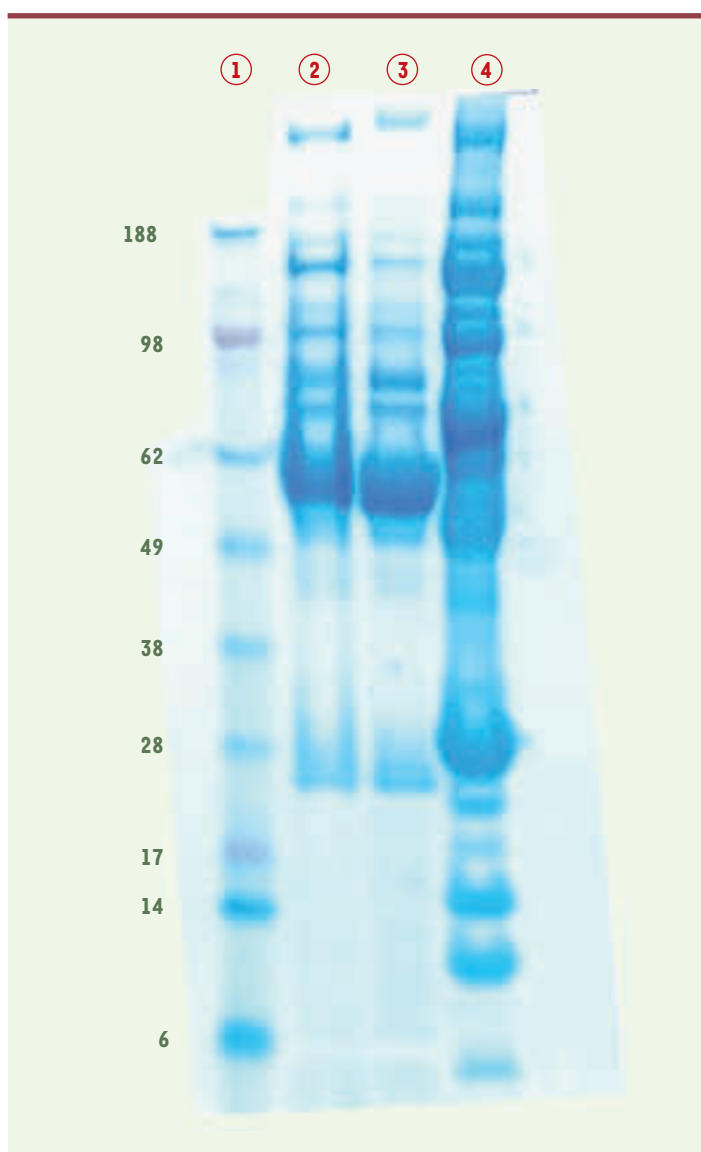


Figure 2. Gel d'électrophorèse du sérum. Ligne①: poids moléculaire; ligne②: sérum non traité; ligne③: fractions non liées à la concanaviline A; ligne④: fractions liées à la concanaviline A. La déplétion de l'albumine et des γ -globulines permet de charger 20 fois plus de sérum sur le gel.

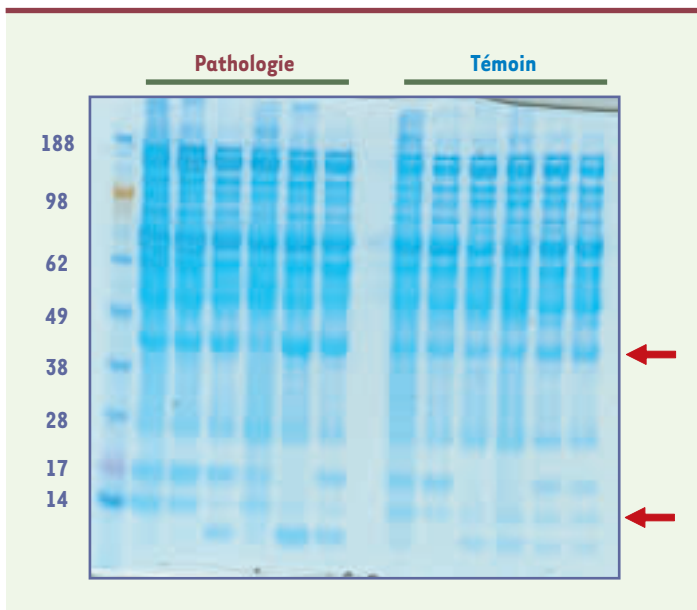


Figure 3. Comparaison de douze échantillons (6 malades et 6 témoins) par électrophorèse en gel à 4-12 % utilisant la procédure en deux étapes. Les flèches rouges indiquent les bandes dont l'intensité varie entre les deux groupes d'échantillons.

enriched for a subset of glycoproteins. Protein profiles analyzed from treated samples were different in composition and abundance when compared to unprocessed serum. The complexity of the protein sample was reduced and the problems associated with the dynamic range were mitigated, allowing for a greater protein load to be applied to the gel. This resulted in the detection of a greater number and variety of proteins. We subsequently tested the suitability of this procedure for biomarker discovery using 12 samples from diseased and non-diseased patients. We found that protein profiles from this enriched fraction were different between the two groups, suggesting that such an approach can assist with new biomarker discovery efforts and the identification of biomarker candidates. Finally, considerations for the translation of biomarkers, identified in such a manner, to a clinically useful assay, are discussed. ♦

REMERCIEMENTS

Je remercie M. German Echeverry pour sa remarquable assistance technique.

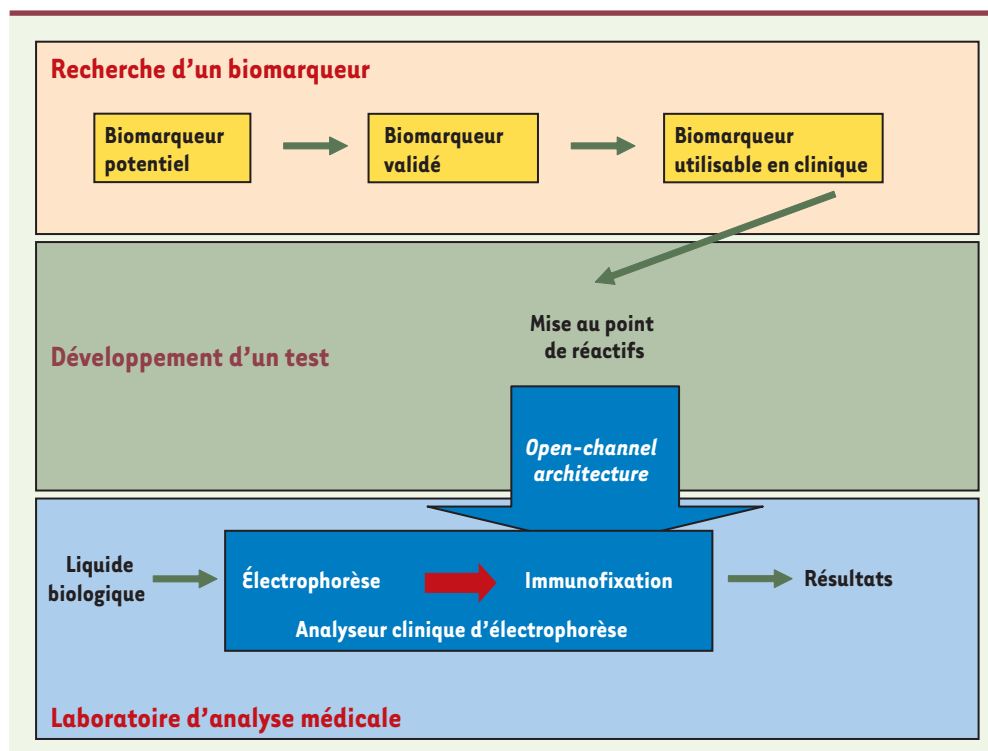


Figure 4. Les différentes phases de l'identification d'un biomarqueur à son utilisation en clinique. Ce schéma illustre la possibilité d'identifier des protéines par électrophorèse.

RÉFÉRENCES

1. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics* 2002 ; 1 : 845-67.
2. Omenn GS, States DJ, Adamski M, et al. Overview of the HUPO plasma proteome project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. *Proteomics* 2005 ; 5 : 3226-45.
3. Omenn GS. Advancement of biomarker discovery and validation through the HUPO plasma proteome project. *Dis Markers* 2004 ; 20 : 131-4.
4. Qiu R, Regnier FE. Use of multidimensional lectin affinity chromatography in differential glycoproteomics. *Anal Chem* 2005 ; 77 : 2802-9.

TIRÉS À PART

A.J. Rai