

Comment enrichir le catalogue d'une banque de protéines sans investissements lourds ?

Dans un avenir maintenant proche, le génome de nombreux êtres vivants sera entièrement déchiffré. Cela ouvrira des perspectives qui donneront incontestablement un nouvel élan aux recherches dans bien des domaines, si toutefois l'on est capable d'assigner à chaque gène sa fonction, ce qui n'est tout de même pas une mince affaire ! La levure (*Saccharomyces cerevisiae*) peut nous y aider. En effet, le séquençage du génome de cet organisme unicellulaire eucaryote, très compact (13,5 Mb) et ne contenant que 6 000 à 7 000 gènes, vient d'être achevé. Or, il est frappant de constater que, chez cet organisme, de toutes les protéines dont la séquence peut être directement déduite du gène, 40 % ne présentent que peu ou pas de similitude avec les protéines déjà décrites et ne sont donc pas encore répertoriées dans le monde vivant. Comme, en outre, la plupart de ces gènes ne sont pas essentiels à la survie de la levure, il est très difficile d'en rechercher les fonctions. On conçoit donc que l'analyse fonctionnelle des génomes soit l'un des défis majeurs des années à venir.

L'électrophorèse bidimensionnelle appliquée à des extraits de levure, permet de séparer plus de 1 500 protéines solubles (figure 1). On peut ainsi avoir une vue globale des variations de synthèse, de dégradation et de modifications des protéines qui résultent de changements génétiques ou de l'environnement. Jusqu'ici, cependant, moins de 80 protéines ont pu être identifiées, car les techniques classiques d'identification sont longues et fastidieuses (microséquençage, spectrométrie de masse, immunodétection, techniques gé-

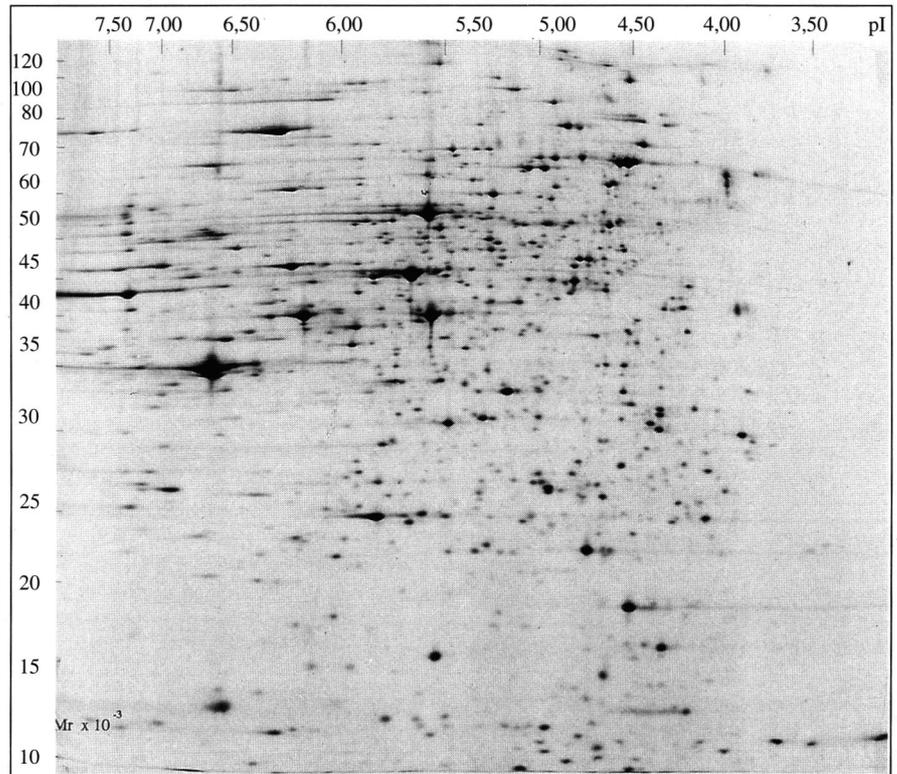


Figure 1. Carte bidimensionnelle de protéines de levures que l'on a fait pousser sur un milieu nutritif contenant un mélange d'acides aminés marqués par des radio-éléments. Cette carte révèle leur marquage au ^{14}C .

tiques de surexpression ou d'interruption de gènes...). Une équipe du CEA, en collaboration avec une Unité Cnrs de Bordeaux, propose une méthode simple et rapide d'identification de ces protéines.

La méthode consiste à analyser la composition en acides aminés des protéines extraites de cellules de levure que l'on a fait pousser sur des milieux nutritifs contenant deux acides aminés différents, marqués

au ^3H , au ^{14}C ou au ^{35}S . Les acides aminés sont choisis parmi ceux qui ne subiront pas d'interconversion ultérieure. Il en existe une bonne douzaine, qui seront donc incorporés comme tels dans les protéines. Sur la base de ce critère, on choisit divers couples d'acides aminés qui donneront lieu à autant de gels. A partir des séquences déduites du génome de la levure (la séquence des 7 000 gènes sera sous peu dispo-

nible) et, pour chaque protéine, à partir des rapports en acides aminés marqués pris deux à deux, déterminés pour chacune des taches obtenues sur les gels 2D, et de la masse moléculaire et du point isoélectrique évalués d'après la position de ces taches, on déduit la nature du gène à l'origine de la protéine. Cette méthode prédictive a été validée sur 45 protéines de référence: une seule erreur a été relevée. Elle a permis d'identifier 80 protéines supplémentaires, ce qui en double, d'un seul essai, la liste. Parmi celles-ci, on note de nombreuses protéines

appartenant aux grandes familles métaboliques (métabolisme des sucres, biosynthèse des acides aminés, choc thermique, etc.) et 14 phases de lecture ouverte dont on ignore la fonction. Certaines d'entre elles présentent de fortes homologues avec des protéines de mammifères. Avec la progression du séquençage du génome, on s'attend à de nouvelles identifications de protéines dont les fonctions sont encore mystérieuses, et ce sont naturellement elles qui suscitent le plus vif intérêt. Ainsi, cette méthode d'analyse est une première

approche rapide et globale d'identification des protéines sur gel 2D qui pourrait s'appliquer à d'autres organismes dont le génome est en cours de séquençage, tels que *Escherichia coli*, *Bacillus subtilis*, *Haemophilus influenzae* ou *Schizosaccharomyces pombe*... et pourquoi pas certaines lignées cellulaires d'eucaryotes supérieurs.

C.R.

I. Maillet I, Lagniel G, Perrot M, Boucherie H, Labarre J. Rapid identification of yeast proteins on two-dimensional gels. *J Biol Chem* 1996; 271: 10263-70.