

Protein-Protein-Interaktionen im Modell Hefe

Seit der Sequenzierung des Hefegenoms im Jahre 1996 werden weltweit Anstrengungen unternommen, die Funktionen der 6000 Hefeproteine zu verstehen. Alle Proteine wurden mehr oder weniger gründlich auf ihre Expression und Interaktionen untersucht. Die Interaktionsstudien haben bis heute rund 8000 einzelne Protein-Proteininteraktionen identifiziert. Daten dieser two-hybrid-Screens und anderer Experimente ergänzen dabei neuere massenspektrometrische Studien zu einem komplexen Bild der zellulären Architektur der Hefe.

Einleitung

Die Bäckerhefe *Saccharomyces cerevisiae* war 1996 der erste Eukaryont, dessen Genom komplett sequenziert wurde. Alleine dieser Durchbruch veranlasste viele Biologen – mich eingeschlossen – von anderen Modellorganismen

Keywords

Proteininteraktionen, Proteom, Genom, *Saccharomyces cerevisiae*

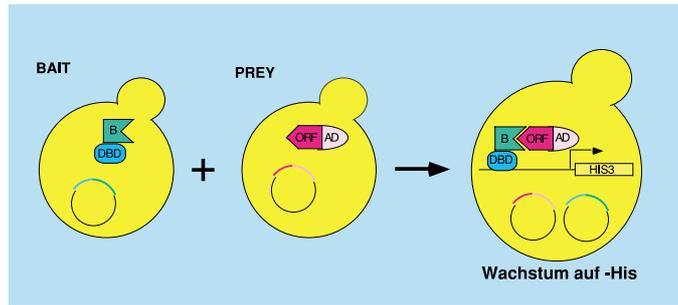


Abb. 1: Prinzip des two-hybrid-Systems

Das two-hybrid-System beruht auf der Expression zweier Fusionsproteine in einer Zelle, die in diesem Fall durch die Verpaarung von zwei haploiden Hefezellen erreicht wird, von denen jede ein anderes Fusionsprotein exprimiert. Eines der Fusionsproteine enthält eine DNA-bindende Domäne, die an den Promoter eines Reportergens (hier: His3) binden kann, und ein Protein "B" (für "bait" = Köder). Das zweite Fusionsprotein besteht aus einer Transkriptionsaktivierungsdomäne und einem Protein ORF (für open reading frame, ein beliebiges Protein). Interagiert B mit ORF wird ein Transkriptionsfaktor gebildet, der das Reportergen anschalten kann. Dadurch kann die Zelle auf Histidin-freiem Medium wachsen. Eine wachsende Hefekolonie zeigt damit an, ob die beiden exprimierten Proteine interagieren.

men auf Hefe umzusteigen. Die Sequenz als solches war an sich schon ein Meilenstein in der Geschichte der Genetik, kannte man dadurch doch zum ersten mal alle Gene einer höheren Zelle. Wichtiger war allerdings noch, dass die Sequenz das Startsignal für eine neue Ära in der biologischen Forschung gab – das Zeitalter der „postgenomischen Biologie“. Die Hefe war schon immer ein „Pioniermodell“ für neue Techniken oder

auch biologische Einsichten: sie war das erste Lebewesen, dessen Gene systematisch kloniert und deletiert, und deren Proteine schliesslich systematisch lokalisiert und analysiert wurden. In den fünf Jahren seit der Publikation der Hefesequenz hat die Gemeinde der Hefeforscher (ca. 1000 Labors mit ca. 5000 Mitarbeitern) vermutlich mehr Daten angehäuft, als in ihrer gesamten Geschichte zuvor. Trotzdem steht die Erforschung der Hefe immer noch am Anfang – auch wenn sie vermutlich der am besten verstandene Organismus ist – noch vor Bakterien wie *E. coli*.

Proteomics

Noch während der Sequenzierung des Hefegenoms wurde klar, dass die Zukunft der Proteinforschung gehört. Schließlich sind die Proteine die aktiven Bestandteile einer Zelle. Es ist auch schon lange klar, dass die meisten Proteine ihre Aufgabe nicht alleine, sondern in Zusammenarbeit mit anderen Proteinen erfüllen. Eine Beschreibung der Proteininteraktionen in einer Zelle ist schon für das Verständnis der Zellstruktur notwendig. Genauso wichtig sind diese Interaktionen aber auch für die dynamischen Prozesse in einer Zelle – man denke nur an die Proteinkomplexe, die an Replikation und Transkription oder auch an metabolischen Prozessen wie der Atmungskette beteiligt sind. Aus diesem Grund wurde auch gleich nach Abschluss der Hefesequenzierung begonnen, Proteininteraktionen in der Hefe systematisch zu untersuchen.

Protein-Interaktionen

Sofort nach der Publikation der Hefesequenz begann ein Wettlauf um die Klonierung und Analyse der Hefegene bzw. -proteine. Ich hatte das Glück, ab Mai 1997 im Labor von Stanley Fields (Seattle) an



diesem Wettlauf teilnehmen zu können. Im Fields-Labor konnte ich bis Ende 1997 einen Array aus 6000 Hefekolonien aufbauen, die jeweils ein anderes Hefeprotein exprimierten. Diese Fusionsproteine konnte man nun mit einem automatisierten two-hybrid-System auf Interaktionen testen, wobei die 6000 Kolonien jeweils mit einem Hefestamm gepaart wurden, der ein bestimmtes „Köder“-Protein exprimierte (Abb. 1). Bis Mitte 1999 hatten wir damit mehrere Hundert Hefeproteine auf ihre Interaktionen mit anderen Proteinen getestet. Die ersten Ergebnisse wurden Anfang 2000 veröffentlicht [1]. Ungefähr zur gleichen Zeit hatten Takashi Ito und seine Mitarbeiter in Japan ein ähnliches System entwickelt und fast zur gleichen Zeit publiziert, wobei die Zahl der untersuchten Proteine allerdings kleiner war. Im Frühjahr 2001 legten die Japaner dann eine weitere und umfangreichere Studie vor [2].

Obwohl two-hybrid-Screens rund zwei Drittel aller bekannten Proteininteraktionen identifiziert haben, lassen sie keine Schlüsse über die Stärke von Proteininteraktionen zu und

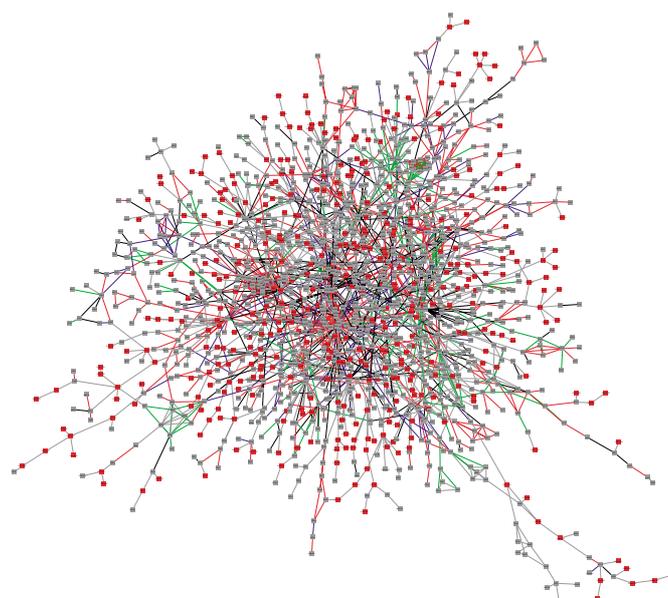


Abb. 2: Proteinnetzwerk in Hefe
 Ein Netzwerk verbindet 1548 Proteine über 2358 Interaktionen. Proteine bekannter Funktion sind hier in Grau, Proteine unbekannter Funktion in Rot dargestellt.

noch weniger, welche Proteine in stabilen Komplexen vorliegen. Genau das leistet aber die massenspektrometrische Analyse von Proteinkomplexen, die freilich zuerst gereinigt und in ihre Bestandteile aufgetrennt werden müssen. Allerdings wurden vor kurzem *high-throughput*-Verfahren entwickelt, um große Zahlen von Komplexen zu reinigen und dann auf ihre Zusammen-

setzung zu untersuchen [3]. Kombiniert man die Ergebnisse beider Methoden, ergibt sich ein zunehmend detaillierteres Bild der molekularen Anatomie einer Zelle.

Was lernen wir daraus?

High-throughput-Analysen ganzer Proteome generieren eine überwältigende Detailfülle, die unweigerlich den

„Nebeneffekt“ mit sich bringt, dass man vor lauter Bäumen den Wald nicht mehr sieht. Bereits nach den ersten systematischen two-hybrid-Screens in der Arbeitsgruppe von Pierre Legrain am Pasteur-Institut in Paris [4] wurde klar, dass sich einzelne Interaktionen zu komplexen Netzwerken verbinden lassen (Abb. 2). Um all die Daten in einen Sinnzusammenhang bringen zu können, benötigt man neue Methoden der Interpretation und Darstellung. Eine Möglichkeit, den Überblick zu behalten, ist die Zusammenfassung von vielen Daten in schematischen Darstellungen. Beispielsweise haben wir kürzlich gezeigt, dass sich eine Vielzahl von Proteininteraktionen als Beziehungen zwischen funktionellen Klassen darstellen lassen (Abb. 3). Man erkennt hier sofort, dass funktionelle Gruppen von Proteinen mit ganz bestimmten anderen Gruppen interagieren. So binden Rekombinationsproteine vor allem an Reparaturproteine (abgesehen natürlich von den Interaktionen innerhalb einer Klasse). Interessanterweise zeigt sich auch, dass Zellzyklusproteine mit auffällig vielen anderen Proteinklassen interagieren – schließlich müssen die vielfältigen Prozesse in einer Zelle im Verlauf der Zellteilung durch Interaktionen mit eben diesen Zellzyklusproteinen präzise koordiniert werden. Wenn dies nicht geschieht, kann es zu Fehlregulation und damit z.B. zu Krebs kommen!

Ausblick

Die meisten Daten aus dem Bereich *functional Genomics* kommen bisher aus DNA-Mikroarray-Experimenten, aber Expressionsmuster liefern nun mal keinen direkten Hinweis auf die Funktion der kodierten Proteine.

Es wird noch viele Jahre dauern, bis wir wirklich alle Proteinkomplexe und -interaktionen in einer Zelle kennen. Meine Arbeitsgruppe arbeitet z. T. daran, Interak-

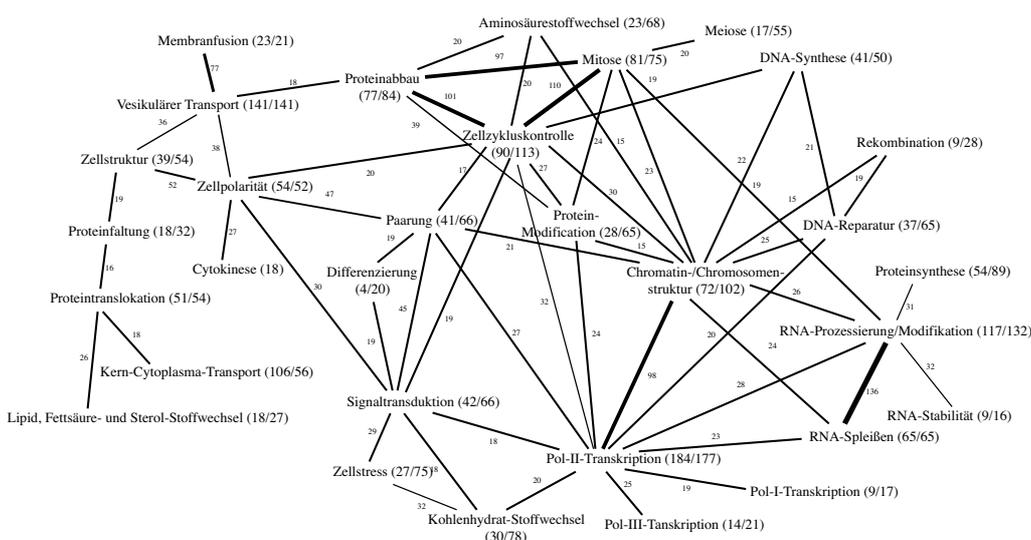


Abb. 3: Interaktionen zwischen Proteinklassen
 Zahlen in Klammern geben hierbei die Zahl der Interaktionen innerhalb einer Gruppe und die Zahl der Proteine in einer Gruppe an. Zahlen an den Verbindungslinien geben die Zahl der Interaktionen zwischen den verbundenen Gruppen an. Zum Beispiel kennt man 77 Interaktionen zwischen den 21 Membranfusionsproteinen und den 141 Proteinen des vesikulären Transports (links oben). 23 Interaktionen verbinden die 21 Membranfusionsproteine. Man beachte, dass nur Verbindungen mit mindestens 15 Interaktionen gezeigt werden; nur Proteine bekannter Funktion wurden hier berücksichtigt. Modifiziert nach Schwikowski et al. [5].

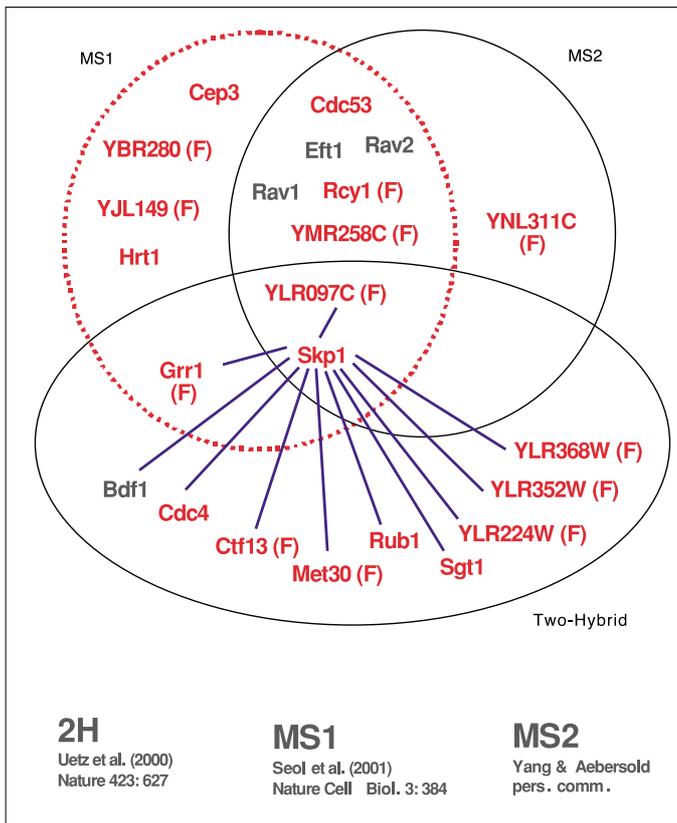


Abb. 4: Two-hybrid-System und Massenspektrometrie

Die Analyse von Proteinkomplexen per Massenspektrometrie zeigt mitunter überraschend große Unterschiede zu den Ergebnissen von two-hybrid-Screens, aber auch zu anderen massenspektrometrischen Untersuchungen. Im gezeigten Beispiel wurde der sogenannte SCF-Komplex mit epitop-markiertem Skp1-Protein aus Hefe isoliert und von 2 unabhängigen MS-Gruppen analysiert (MS1 und MS2, [6]). Unser two-hybrid-Screen mit Skp1 erbrachte Interaktionen mit unerwartet vielen anderen Proteinen. Interessanterweise sind fast alle gezeigten Proteine plausible Bestandteile des Komplexes: alle rot gefärbten Proteine wurden bereits zuvor mit der Funktion des SCF-Komplexes, nämlich dem Proteinabbau, in Verbindung gebracht.

tionen von Proteinen unbekannter Funktion zu finden, um damit Aufschlüsse über deren Funktion zu erhalten. Außerdem wollen wir die genauen interagierenden Sequenzen innerhalb der Proteine kartieren, nicht zuletzt um Vorhersagen machen zu können, welche Sequenzklassen an welche anderen binden. Das sollte uns schließlich auch erlauben, die Funktion homologer Sequenzen beim Menschen zu verstehen. Das langfristige Ziel ist nicht zuletzt eine 4-dimensionale „molekulare Anatomie“ der Zelle, die allerdings nur mit Hilfe röntgenkristallographischer und anderer Methoden erreicht werden kann.

Literatur

- [1] Uetz, P. et al.: Nature 403, 623-7 (2000)
- [2] Ito, T. et al.: Proc Natl Acad Sci U S A 98, 4569-74 (2001)
- [3] Gavin, A.-C. et al.: Nature, in press (2002)
- [4] Fromont-Racine, M. et al.: Nat Genet 16, 277-82 (1997)
- [5] Schwikowski, B. et al.: Nat Biotechnol 18, 1257-61 (2000)
- [6] Seol, J.H. et al.: Nat Cell Biol 3, 384-91 (2001)



Peter Uetz

Dr. Peter Uetz

Studium der Biologie in Stuttgart, Tübingen und Heidelberg; 1992 Diplomarbeit am Max-Planck-Institut für Medizinische Forschung, Heidelberg; 1993-96 Promotion am EMBL, Heidelberg; 1997-2001 Postdoktorand an der University of Washington, Seattle, USA; seit April 2001 Gruppenleiter am Institut für Toxikologie und Genetik, Forschungszentrum Karlsruhe

Institut für Toxikologie und Genetik
Forschungszentrum Karlsruhe
Postfach 3640
76021 Karlsruhe
E-mail: peter.uetz@itg.fzk.de
Internet:
<http://itgmv1.fzk.de/www/itg/uetz/>

Easy Info

000