

PublisherInfo		
PublisherName	:	BioMed Central
PublisherLocation	:	London
PublisherImprintName	:	BioMed Central

## Finding the links within the yeast proteome

ArticleInfo		
ArticleID	:	4324
ArticleDOI	:	10.1186/gb-2002-3-3-reports0014
ArticleCitationID	:	reports0014
ArticleSequenceNumber	:	19
ArticleCategory	:	Paper report
ArticleFirstPage	:	1
ArticleLastPage	:	4
ArticleHistory	:	RegistrationDate : 2002-2-14 Received : 2002-2-14 OnlineDate : 2002-2-22
ArticleCopyright	:	BioMed Central Ltd2002
ArticleGrants	:	

Reiner Veitia

## Abstract

A generic method for purifying protein complexes coupled with mass spectroscopic identification has enabled part of the network of interactions within the yeast proteome to be unraveled

## Significance and context

The identification and analysis of components of multiprotein complexes is a straightforward way to understand how the cellular proteome is organized. It may also help to assign functions to gene products that remain to be annotated. Several large-scale approaches have been taken towards unraveling the functional relationships between cellular components. These include monitoring mRNA expression (chips and serial analysis of gene expression), loss-of-function approaches, extensive two-hybrid screens, protein chips, and *in silico* methods, such as the 'Rosetta stone' approach.

The only systematic protein-interaction studies described so far are based on *ex vivo* and *in vitro* systems, such as the two-hybrid system and protein chips. In addition, these approaches produce mainly binary data - that is, A interacts with B or A interacts with C, and so on. Here, however, Gavin *et al.* have carried out a comprehensive analysis of protein complexes in the baker's yeast *Saccharomyces cerevisiae*. They used tandem-affinity purification (TAP) combined with mass spectrometric analysis, which yields complex composition data - that is, A is part of a complex with B, C, D, and so on.

## Key results

Gavin *et al.* processed about 1,700 genes, including over 1,000 genes having metazoan orthologs (genes evolved by vertical descent from a common ancestor, and presumed to carry out the same function). A set of 600 non-orthologous genes was also studied and all analyses were performed in haploid cells. Protein complexes were isolated by TAP, a technique that combines a first high-affinity purification (elution using a site-specific protease) and a second affinity purification to obtain the protein complexes. After separation of these by denaturing gel electrophoresis, individual bands were trypsinized, analyzed by mass spectroscopy and identified by matching the results to a database. In all, about 600 tagged proteins (including over 400 orthologs and 40 membrane-associated proteins) were used to characterize protein complexes. Of the successfully tagged proteins (entry points), about four out of five had associated partners. In a third of the cases where no complexes were purified, the protein in

question was detected as part of other complexes (through other entry points). About 1,440 gene products (encoded by a quarter of the open reading frames in the genome) were shown to be 'interactors' and could be grouped into 232 distinct multiprotein complexes. The sizes of the complexes varied from 2 to 83 components, with an average of 12 components per complex. The authors cautiously excluded proteins present in more than 3.5% of the complexes (66 proteins), which were regarded as possible artifacts.

To understand the proteome circuitry, Gavin *et al.* then studied the relationships between complexes by linking those sharing polypeptides. This yielded a network whose connections reflect physical interactions but may also suggest common regulation, localization, turnover or architecture - 'guilt by association'. As a rule, the more connected a complex is, the more central its position in the network. Of the 232 TAP complexes, only 9% had no novel element, and roles could be proposed for 231 out of the 304 proteins for which there was no functional annotation.

Gavin *et al.* also carried out a parallel analysis of human and yeast complexes and found that orthologous proteins preferentially interact with complexes enriched with other orthologs. A propensity to associate among the products of essential genes was also observed (complexes containing orthologs and essential proteins overlap significantly). This suggests the existence of an 'orthologous proteome' representing core functions of the eukaryotic cell.

## Links

A software package is available at the [Yeast protein complex database](#) for navigation of the proteome map described in the paper.

## Conclusions

The TAP approach maintains protein concentration, localization and post-translational modifications in a way resembling normal physiology. TAP therefore allows isolation of complexes from different cellular compartments, efficient identification of low-abundance proteins (around 15 copies per cell) that are undetectable by other proteomic approaches, purification of very large complexes (more than 1-1.5 MDa), and the study not only of cohesive (tight) complexes but also of less stable ones. The system proved valuable for identifying proteins ranging from 6.6 kDa to 559 kDa in size and with pI values between 3.9 and 12.4.

The 20 kDa TAP tag may, however, interfere with the assembly of complexes, or with protein localization or function. Indeed, there is a clear technical bias towards tagging proteins smaller than 15 kDa. The method may fail to detect transient interactions, low-stoichiometric complexes and/or those interactions occurring only in specific physiological conditions. TAP-mass spectroscopy does not, for example, provide information on the orientation of components within a complex. Yeast two-hybrid analysis may prove complementary and relevant for the detection of pairwise and transient associations.

# Reporter's comments

This paper shows an elegant way to address the question of protein-protein interactions at a genomic scale and constitutes a paradigm that will certainly be applied to higher eukaryotes. Although proteins are believed to be the main functional players in the cell, there is overwhelming evidence implicating so-called 'non coding' RNAs (ncRNAs) as playing crucial roles in some cellular networking. When studying the interactomes, especially of higher eukaryotes, one should leave room for these. It will not be surprising if a great deal of ncRNAs prove to be 'missing links' in protein complexes.

## Table of links

[Nature](#)

[Yeast protein complex database](#)

## References

1. Gavin A-C, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick J M, Michon A-M, Cruciat C-M, et al: Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature*. 2001, 415: 141-147. 0028-0836