ELSEVIER

Journal of Chromatography B, 782 (2002) 151–163

JOURNAL OF CHROMATOGRAPHY B

www.elsevier.com/locate/chromb

Review

Gene function on a genomic scale

Lars M. Steinmetz[a,*,1], Adam M. Deutschbauer[b,1]

[a]*Department of Biochemistry, Stanford University School of Medicine, Stanford, CA 94305, USA*
[b]*Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA*

Abstract

The ability to obtain experimental measurements for thousands of genes has revolutionized our view of biological systems. While traditional studies of gene function evaluated many different properties for a single gene, genomic approaches can measure a single property for thousands of genes. Over the last years, genomic approaches have been developed to measure many different properties, including gene expression, deletion phenotype, and protein characteristics. The promise of integrating these datasets has made it attractive to test whether genomic approaches can determine gene function with accuracy comparable to single gene approaches.

Keywords: Reviews; Genomics

Contents

1. Introduction

Sequencing projects are yielding new gene sequences faster than biologists can discover their function. Four years ago with the sequence of the first eukaryote, the yeast *Saccharomyces cerevisiae* [1], came the startling discovery that the function of over half of the ~6200 reported open reading frames (ORFs) was unknown. Many other eukaryotic genomes have since been sequenced, including the worm *Caenorhabditis elegans* [2], the fly *Drosophila melanogaster* [3], the plant *Arabidopsis thaliana* [4], and recently our own human species [5,6]. These genomes contain more genes than in yeast, with over 30 000 genes in human. As in the case with yeast, about half the genes from each genome project are uncharacterized.

*Corresponding author. Tel.: +1-650-723-6287; fax: +1-650-725-6044.
E-mail address: larsms@stanford.edu (L.M. Steinmetz).
[1]The authors contributed equally to this work.

1570-0232/02/$ – see front matter
PII: S1570-0232(02)00549-4

An understanding of gene function does not derive from knowledge of a single variable, rather it becomes more complete with information about multiple variables including expression pattern, splice variants, mutant phenotype, protein-encoding ability, protein interaction partners and protein localization. In model organisms where experimental manipulation is feasible, individual, single-gene experimental approaches have traditionally been applied. However, single-gene approaches are insufficient to meet the challenges of the genome, therefore highly parallel and systematic strategies have been developed. They address gene function from three different levels: mRNA transcription, mutant phenotype, and protein characteristics. How has each approach been used and what are its limitations?

2. mRNA transcript analysis

The largest efforts to characterize gene function on a genome-wide scale have focused on measuring mRNA expression levels with hybridization-based approaches. High-density DNA arrays, also known as microarrays [7,8], enable mRNA levels to be measured on a massively parallel, high-throughput scale (Fig. 1). Expression analysis with arrays is easy to carry out and can be applied to practically

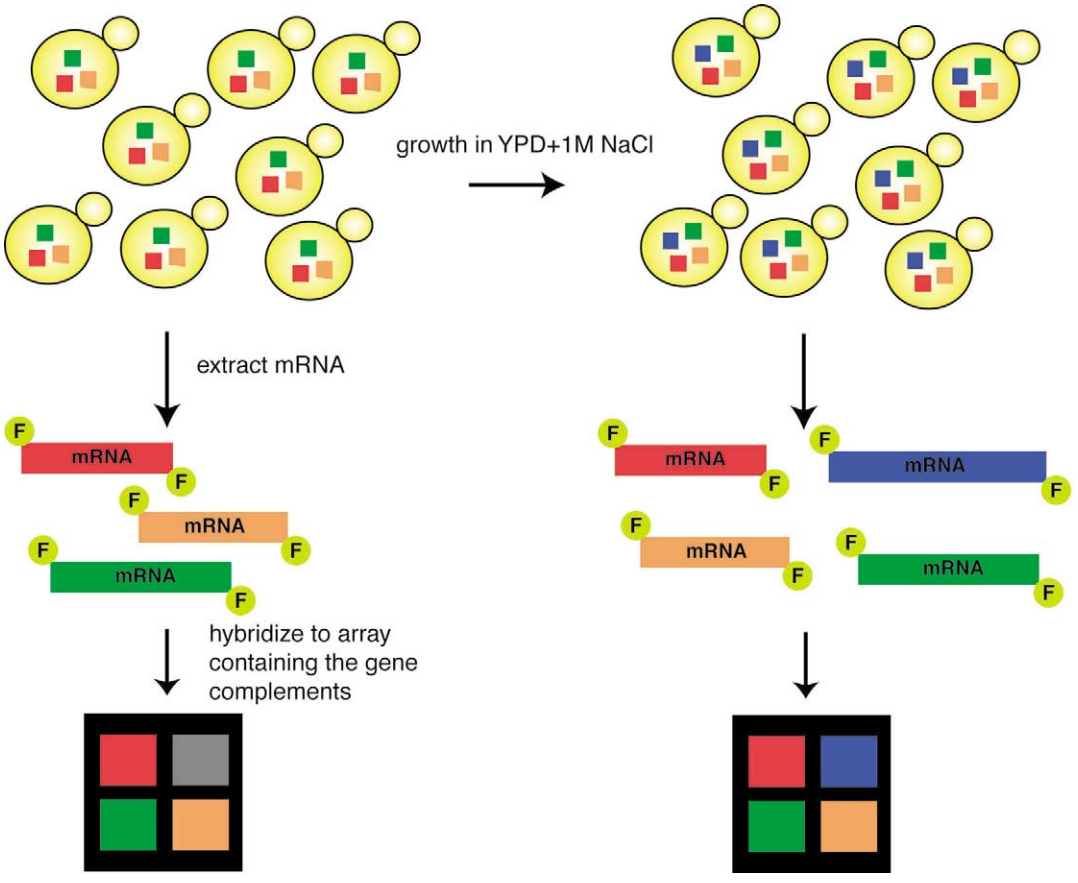


Fig. 1. mRNA transcript analysis in yeast. Cultures are grown in two different media conditions, depicted here for low and high salt. Cell samples are collected from each condition to isolate mRNA. Isolated mRNA is fluorescently labeled (shown here for four genes) and hybridized to a high-density array containing complementary sequences to each gene. Differences in mRNA abundance between the two conditions are inferred by differences in the hybridization signals on the arrays. In this example one mRNA (blue) is expressed only under high salt and shows no expression in the low salt condition.

any organism. This has made it attractive to test whether functional information can be gained from measuring the mRNA transcript level under a variety of conditions for every gene in the genome.

A tight correlation between transcription and function exists during development. For example, genes that function specifically in particular stages of the cell-cycle often show a cell-cycle dependent periodicity of expression [9–11]. In addition, experiments of meiosis in yeast [12,13], metamorphosis in flies [14], serum response in human [15], and a collection of conditions in worm development [16] are rich in examples where genes with similar function have a similar expression profile. However, a major drawback of mRNA expression analysis is that mRNAs are not functional entities themselves, but rather are transmitters of information from the genome to the proteome where function is enacted. The inference of function from mRNA expression is therefore indirect and rests on the assumptions that the execution of cellular processes and the activation of molecular pathways are tightly regulated and that evolutionary selection is likely to have limited transcription to times when protein products are needed [17]. Therefore, predicting function of uncharacterized genes based on similarity in expression to genes of known function demands an answer to two difficult questions: (1) if a gene is found to be expressed in a particular pathway, is it important for that pathway? And (2) if a gene is found not to be expressed, is it not important for that pathway?

Changes in transcript levels may not reflect changes in protein expression due to differences in translation, protein modification or degradation. Early comparisons of mRNA and protein levels have yielded contradictory results about their correlation [18,19]. A recent analysis of 289 genes showed a correlation of 0.61 between protein-abundance ratios and mRNA-level ratios and noted that 15 of 30 proteins with clear changes in protein abundance did not show any change in mRNA level [20]. Considering that genes compete for cellular resources such as polymerase enzymes and transcription factors, it is possible that an expression change in one gene will affect the transcript rates of all other genes with which it competes. As a result, a global understanding of gene regulation may be necessary before one can fully understand the significance of changes in expression [21]. Analysis of a transcription factor mutation in yeast attributed 96–99% of the initially measured expression changes to a secondary effect that turned out to be a change in the growth rate of yeast [22]. Likewise, the ongoing analysis of gene deletion phenotypes in yeast has revealed little correlation between fitness effect and expression level [23–26]: genes with a change in deletion phenotype between two conditions often did not show a significant change in mRNA expression level between the same two conditions and vice versa (Fig. 2). These measurements suggest that mRNA transcript analysis is a powerful tool for obtaining a rapid initial assessment of the activities in a genome but prediction of function requires caution.

The analysis of gene expression information may be aided by focusing on genes that are not differentially expressed under most conditions, except for one. Those genes that change expression only under one particular condition are most likely pathway-specific and biologically relevant. However, this extension also requires that one be able to clearly distinguish splice variants, a problem that has been largely overlooked. Although alternative splicing is not a problem in yeast where 70% of the genome is protein-coding sequence and only 238 spliced transcripts exist [27], it is a problem in higher eukaryotes. Only ~1.5% of the human genome is protein-coding and 99% of genes are spliced with an estimated median of seven exons per human gene [5]. Because protein isoforms carry out different cellular functions and an estimated 35–59% of human genes have at least two protein isoforms [28], distinguishing differently spliced transcripts is important. Array designs that verify expressed transcripts [29] and detect alternative splicing [30] need to be applied.

As demonstrated by studies of the cell cycle or multicellular eukaryotic development, genes with similar function are often transcriptionally co-regulated. Therefore knowing the temporal and spatial regulation of transcription may be a better indicator of gene function than changes in expression level alone. One key aspect of understanding transcriptional regulation is identifying the transcription factors and regulatory sequences that control gene expression. One powerful approach combines chromatin immunoprecipitation (ChIP) with DNA microarray analysis. This approach worked for finding the

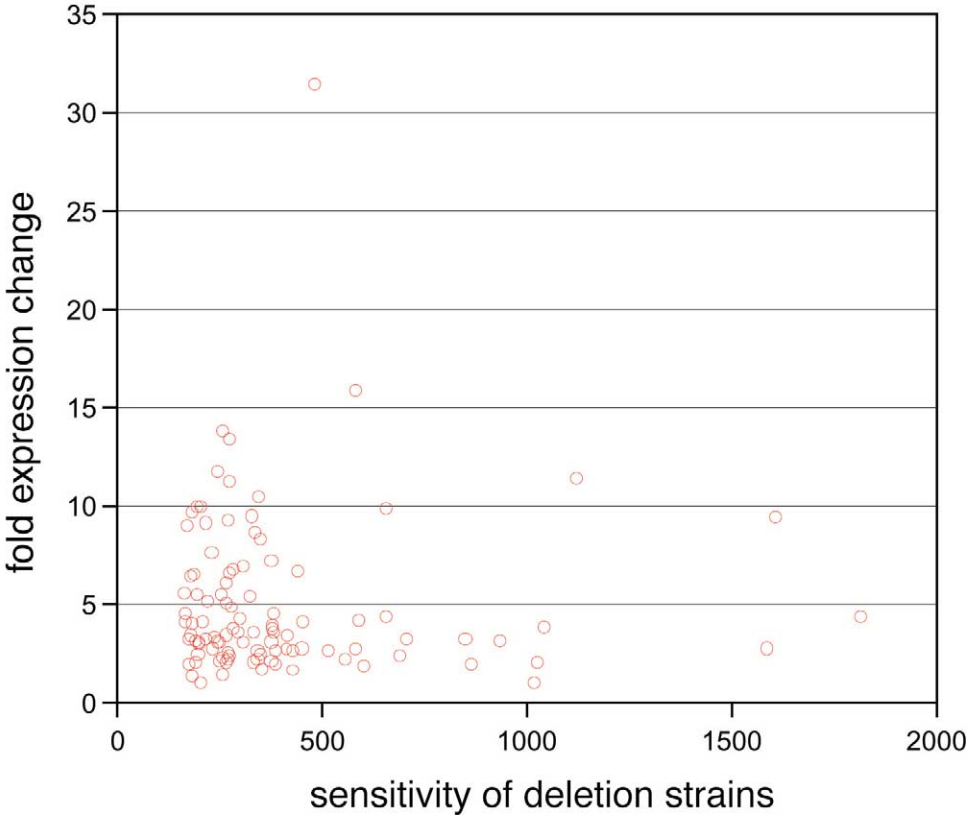


Fig. 2. The correlation between mutant phenotype and mRNA expression levels in the response of yeast to high salt. Fitness measures for all homozygous deletion strains (~4700 total) were assayed in parallel using a molecular bar-coding approach (see *mutant phenotype analysis* section and Fig. 3) to identify 103 deletion strains with a salt hypersensitive phenotype. The sensitivity value for each strain was calculated as the negative natural logarithm of the likelihood of observing the experimental values by chance (see Ref. [24] for more details on data analysis). Larger sensitivity values indicate greater fitness defects in high salt. The salt expression data were obtained from Ref. [72], from which 531 genes were identified with a 10-fold or greater fluctuation in expression during growth in high salt. A plot of the 103 salt hypersensitive deletion strains against their fold expression change shows that the majority of genes required for salt tolerance are not differentially expressed greater than 10-fold during salt treatment.

sequence vicinity of transcription factor binding sites with high accuracy [31,32]. An alternative approach measured the mRNA expression level of genes in strains containing transcription factor mutations [33]. The idea is that genes that are activated in expression by the transcription factor will not be transcribed in the mutant compared to a wild-type culture. However in a comparison, the ChIP technology has the advantage of overcoming the problem of having to distinguish primary from secondary effects. Similar advantages of ChIP exist over expression level studies that predict regulatory sequence motifs by aligning upstream sequences of genes that show similarities in expression profile. Although successful in isolated cases [9,10,34], this approach seems to depend heavily on the biological problem that is studied and how tight the co-expression is defined and measured. All three approaches confirm that transcription factor binding sites are found at many more places than are bound by transcription factor, illustrating the importance of combining computational analysis with experimentation.

A final application of mRNA transcript analysis to gene function uses expression levels as a detailed molecular phenotype or signature profile. Applied to group deletion mutants by similarity in their mRNA

transcription responses, this approach can be used to predict new gene function for uncharacterized gene products based on similarity in signature profile to characterized genes [35]. This approach requires one experiment per gene deletion mutant and scaling it to the genome is therefore a challenge.

3. Mutant phenotype analysis

Rather than expression level, the phenotypic consequence of mutations has more often been used as a critical determinant of gene function. Traditionally, geneticists have randomly mutagenized organisms with the goal of generating phenotypes of interest. Subsequent identification and cloning of the mutation strongly suggests a role for the corresponding gene. While this ''forward genetic'' approach has yielded functional information for many genes, it is too time consuming to analyze the thousands of uncharacterized genes identified by sequencing projects. The rise of functional genomics has ushered in a new paradigm for large-scale phenotypic analysis of gene mutations. For given organisms, genome-wide efforts are currently under way not only to generate mutations in all genes, but also to develop systematic approaches to examine mutant phenotypes in a high-throughput manner.

Methodologies for mutagenizing genomes can be separated into two general classes: random and direct. Random integration into the host genome using transposons or gene trap vectors remains a simple and cost-effective method for generating large numbers of mutations. Such an approach has been applied extensively to the genomes of Drosophila [36,37], yeast [38], bacteria [39], mouse [40], and plants [41]. While serving as a useful tool for inducing many mutations, approaches based upon random integration require sequencing [38] or pooling methods [42,43] to identify integration events in specific genes. In addition, all genes are not equally amenable to integration, therefore, other approaches are necessary to achieve full genome saturation.

Directed mutagenesis programs have the obvious advantage in that the user controls the precise nature of the mutation without prior knowledge of mutant phenotype. These ''reverse genetic'' approaches applied on a genome-wide scale will result in large collections of custom mutations and greatly facilitate biological discovery. In yeast, a consortium of laboratories used a PCR-based approach to generate ''knock-outs'' of all open reading frames in the yeast genome by homologous recombination [24]. Due to lower targeting efficiencies and practical considerations in maintaining deletion collections, genome-wide knock-outs have yet to be reported in the other organisms where the homologous recombination system is developed: mouse [44], Drosophila [45], and sheep [46]. An alternative approach is to use RNA molecules to selectively silence the function of a target gene [47]. The leading technology, applicable to a wide range of model systems, is RNA interference (RNAi) in which double-stranded RNA (dsRNA) corresponding to a particular gene selectively degrades its transcript thereby eliminating the function of that gene. The primary advantages of RNAi are the ease of dsRNA synthesis and the flexibility of inhibition. The user can spatially and temporally control the interference reaction. One key disadvantage of RNAi is that the effects are not permanently inheritable. In addition, injected dsRNA may be of limited use in whole organisms or tissues due to difficulty of targeting multiple cells at once.

For analyzing mutations in all genes in a genome, high-throughput methods of phenotypic analysis are necessary. Among these, the ones for microorganisms are most developed. One such approach, developed in yeast and termed genetic footprinting [48], takes advantage of gene-specific PCR primers to follow the fitness of thousands of transposon insertion mutants over time. Although genetic footprinting is applicable to any microorganism for which sequence data is known, this approach is limited by the fact that a specific PCR reaction needs to be carried out for each gene in the assay. A second approach is based upon tagging mutant strains with different DNA sequences and using hybridization to distinguish mutant variants. This technique was first developed for transposon mutagenesis in the bacteria *Salmonella typhimurium* and termed signature-tagged mutagenesis [49]. Although proven useful in a number of bacterial species [50], this approach typically involves random transposon mutagenesis, thereby potentially missing critical genes. More recently, a systematic approach using directed mutagenesis has been achieved in yeast

[23]. Molecular tags enable massively parallel analysis of a large number of strains to obtain a quantitative measurement for each individual strain (Fig. 3). Recent applications have focused on diverse processes such as sporulation [25], mitochondrial function [26], UV sensitivity [51], and the nonhomologous end joining pathway NHEJ [52].

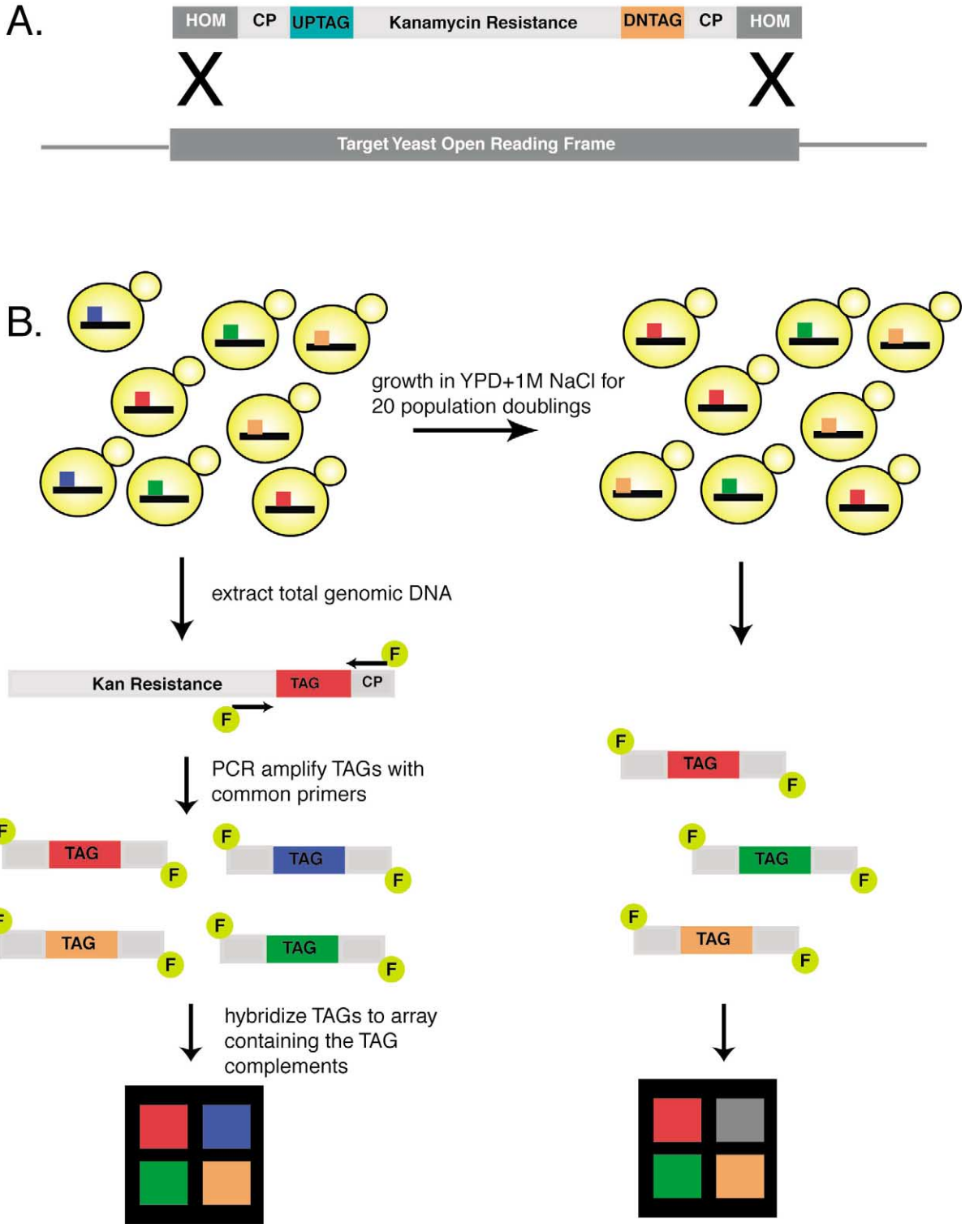
In multicellular eukaryotes, the ability to generate mutations currently outperforms our capacity to analyze their phenotypes in detail. One confounding issue is the complexity of phenotypes in higher organisms which are often more difficult to detect. Specific assays that screen through thousands of deletion lines to identify individual mutants for specific cellular processes are needed. This is difficult because whole multicellular organisms cannot be ''pooled'' as easily as microorganism. Despite these problems, efforts are currently under way to determine the loss of function phenotype of all worm genes using RNAi [53,54]. For example, to study function of essential genes in the early embryo, Gonczy et al. injected double-stranded RNA corresponding to 2174 genes of chromosome III into the germline to generate fertilized eggs lacking the targeted transcript. Time-lapse differential interference contrast microscopy was used to examine the first cell division of the zygote to identify 136 gene products that are required for specific aspects of the cell division. This result demonstrates that RNA silencing techniques can be used to systematically dissect early developmental systems using loss of function mutations that could not be propagated in the adult organism. Although not yet reported, RNA silencing methodologies can in theory be applied on a genomic scale for a number of organisms where the phenomenon is developed including flies [55], zebrafish [56], mice [57], and human cultured cells [58]. Studies of the latter and other cell lines may be aided by dsRNA arrays, in which cells are overlaid on a surface containing dsRNA at defined locations. The molecular and cellular consequences of a specific gene inhibition can be followed by microscopy-based assays [59].

As the techniques of homologous recombination and RNAi are optimized, the loss of function phenotypes for thousands of genes of common model organisms should be elucidated. However, as the results from yeast clearly demonstrate, the disruption of many genes will not result in a detectable phenotype due to redundancy at the individual gene or pathway level. Other techniques, including the functional analysis of genes at the protein level or systematic construction of double deletions [60] will be necessary to determine the function of these genes.

4. Protein analysis

The function of genes is mediated by the activities of their encoded proteins. Therefore, measuring mRNA transcript levels and analyzing mutant phenotypes are insufficient by themselves to understand the molecular function of proteins, and hence the genes which encode them. A global view of genome function is incomplete without an understanding of protein activity and dynamics. Similar to the revolution in genomics, systematic approaches

Fig. 3. Parallel phenotypic analysis using molecular bar-coding in yeast. (A) Each open reading frame (ORF) in the yeast genome was disrupted via homologous recombination of a targeting cassette. In addition to the regions of homology to the target locus, each targeting cassette was constructed with two unique 20 base pair sequences, or molecular tags, that serve as unique strain identifiers [73]. A universal selectable marker (Kanamycin resistance) and common PCR priming sites (CP) enable PCR amplification of tags in a complex mixture of DNA. (B) An example of parallel phenotypic analysis is illustrated here for a simplified pool of four deletion strains, each of which carries a single molecular tag. Prior to selection, total genomic DNA is extracted from the starting pool, all molecular tags amplified in a single PCR reaction, and the tags hybridized to an oligonucleotide array containing the tag complements. Visualization of bound product on the array is accomplished using biotinylated primers together with a fluorescent-streptavidin conjugate. Because all strains are present in the starting pool, all probes exhibit a hybridization signal (colored in the illustration for simplicity). The pool is then subjected to selection, in this case growth for 20 population doublings in media containing a high salt concentration. In this example, the deletion strain with a ''blue'' tag exhibits a relative fitness defect in salt media compared to the deletion pool average. Note the absence of blue signal on the array following hybridization of the molecular tags from the ''selected'' pool. By comparing tag hybridization signals before and after selection, strains with specific fitness defects are identified in parallel. Currently, this approach is being applied to deletion pools representing all nonessential ORFs (~4700) in the yeast genome.


A.
HOM
CP
UPTAG
Kanamycin Resistance
DNTAG
CP
HOM
Target Yeast Open Reading Frame
B.
growth in YPD+1M NaCl for
20 population doublings
extract total genomic DNA
Kan Resistance
TAG
CP
F
PCR amplify TAGs with
common primers
hybridize TAGs to array
containing the TAG
complements

are being developed to analyze the expression level, localization, biochemical activity, interaction partners and modifications of all proteins encoded by genomes. This emerging field of biological discovery, termed proteomics [61], is technically challenging due to diverse properties of individual proteins.

Proteomic strategies require the identification of proteins and their quantification. Identification is typically carried out by separating a protein sample by two-dimensional gel electrophoresis or chromatography followed by mass spectrometry of each protein in a fraction. The main problem for proteomic studies is, however, complexity. From crude cell extracts, measurements of only the most abundant proteins can be obtained. Affinity chromatography can enrich for specific lower abundance proteins but yields only a subset of the proteome. Of great promise are new developments in mass spectrometry that promise to extend the analysis of crude cell extracts to moderate and low abundance proteins by using more sensitive and precise detection mechanisms [62]. The use of chemically labeled protein samples has enabled two or more samples to be analyzed in parallel and has already been used to quantitate differences in the most abundant proteins between two crude cell extracts [63]. As an alternative, protein arrays offer a platform for measuring differences in abundance between specific proteins for which antibodies are available [64].

Individual proteins typically mediate their function through their interaction with other proteins, either transiently through modifications such as phosphorylation or as a component of a multisubunit complex. Therefore, knowledge of the interaction partners of a given protein suggests function if prior information of one of the proteins has been obtained. The yeast two-hybrid approach is a flexible technique that takes advantage of the interaction of two fusion proteins to activate a reporter gene in yeast [65]. However, two genome-wide analyses of two-hybrid interactions in yeast revealed a lack of reproducibility and large percentage of false positive and false negative results [66,67]. Positive interactions identified by this approach are complicated by the possibility that under native conditions the proteins do not localize together or they are present at a low concentration and only interact under artificially high concentrations as used in the two-hybrid screen. A promising alternative to the two-hybrid approach involves the isolation of native protein complexes by affinity purification followed by individual identification of components by mass spectrometry (Fig. 4) [68]. Two groups have scaled this approach to a genomic level in yeast with encouraging results [69,70]. Gavin et al. purified 589 protein assemblies from yeast and could identify 232 distinct protein complexes containing 231 unclassified proteins, providing a higher-order map of the yeast proteome in addition to providing functional insight for unclassified proteins.

The large-scale analysis of proteins has also been applied to at least three other questions that are relevant to studies of gene function: sub-cellular localization, post-translational modifications, and biochemical activity. Transposons engineered with reporter genes such as the green fluorescent protein (GFP) were used to follow the localization of many fusion proteins in a single cell context [38]. High-density protein arrays, such as those described recently for the yeast proteome [64], were used to address not only protein activity and interaction partners, but also subtle modifications such as phosphorylation. GST fusions of all yeast proteins were isolated in a comprehensive approach to assay for new biochemical activities [71]. Although proteomics presents many technical challenges, overcoming these issues is important because proteomics can be applied to many different organisms to provide an informative measurement of gene function.

5. Conclusions

Each of the three approaches to gene function discussed here is insufficient by itself. Current results confirm expectations from single-gene analysis, that each approach measures a different variable and gene function cannot be determined with knowledge of only a single variable. A thorough knowledge of gene function requires an integration of approaches. This can only be achieved if the significance of each measurement is assessed with the same standards of rigor that have been applied to single gene studies. Only then can genome-wide data be applied to make specific and accurate discoveries for individual

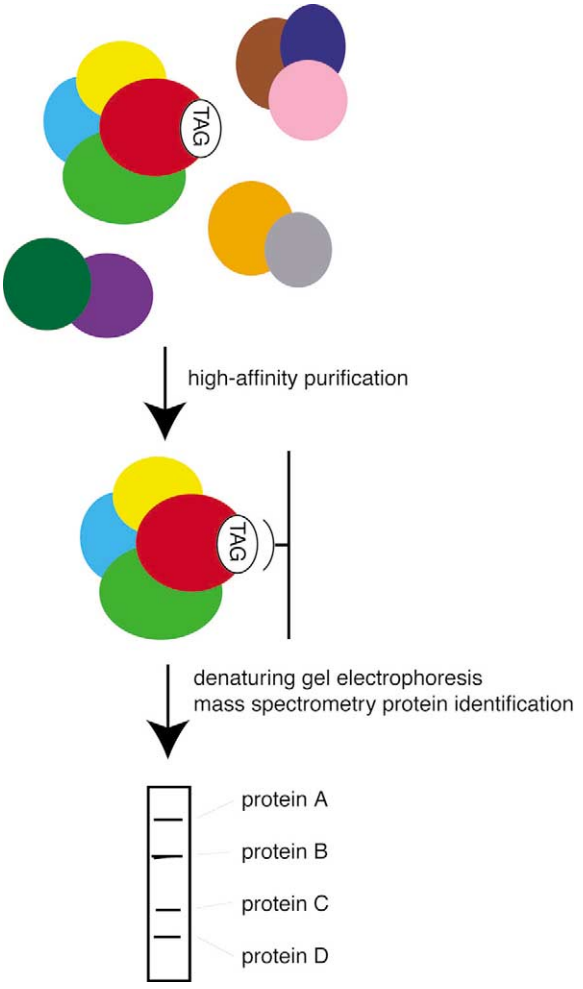


Fig. 4. Analysis of protein complexes by affinity purification and mass spectrometry. From crude cell extracts, an individual protein complex is purified by taking advantage of a fusion protein that has a protein-tag attached to it. This tagged protein serves as a bait to co-purify with it other proteins with which it interacts. During the affinity purification step other protein complexes that do not contain the tagged protein are washed away. The proteins in each complex are separated on a denaturing gel. The identity of each band is revealed by trypsin digestion and subsequent analysis by matrix-assisted laser desorption/ionization-time-of-flight (MALDI–TOF) mass spectrometry or electrospray tandem mass spectrometry. These processes measure the mass of each trypsinized protein fragment. The identity of the protein fragment is revealed by searching the mass against a database containing the calculated masses for all trypsinized protein fragments of an organism.

genes, which when integrated will achieve a detailed, but systems-level understanding of how cellular function and responses take place.

Acknowledgements

We thank Q. Wang, M. Mindrinos, and P. Oefner for helpful comments on the manuscript.

References

[1] A. Goffeau, R. Aert, M.L. Agostini-Carbone, A. Ahmed, M. Aigle, L. Alberghina, K. Albermann, M. Albers, M. Aldea, D. Alexandraki, G. Aljinovic, E. Allen, R. Altmann, J. Alt-Mörbe, B. André, S. Andrews, W. Ansorge, G. Antoine, R. Anwar, A. Aparicio, R. Araujo, J. Arino, W. Arnold, J. Arroyo, E. Aviles, U. Backes, M.C. Baclet, A.B.K. Badcock, V. Baladron, J. P.G. Ballesta, A.T. Bankier, A. Banrevi, M. Bargues, L. Baron, T. Barreiros, B.G. Barrell, C. Barthe, A.B. Barton, A. Baur, A.-M. Bécam, A. Becker, I. Becker, J. Beinhauer, V. Benes, P. Benit, G. Berben, E. Bergantino, P. Bergez, A. Berno, I. Bertani, N. Biteau, A. J. Bjourson, H. Blöcker, C. Blugeon, C. Bohn, E. Boles, P.A. Bolle, M. Bolotin-Fukuhara, R. Bordonné, J. Boskovic, P. Bossier, D. Botstein, G. Bou, S. Bowman, J. Boyer, P. Brandt, T. Brandt, M. Brendel, T. Brennan, R. Brinkman, A. Brown, A.J.P. Brown, D. Brown, M. Brückner, C.V. Bruschi, J. M. Buhler, M.J. Buitrago, F. Bussereau, H. Bussey, A. Camasses, C. Carcano, G. Carignani, J. Carpenter, A. Casamayor, C. Casas, L. Castagnoli, H. Cederberg, E. Cerdan, N. Chalwatzis, R. Chanet, E. Chen, G. Chéret, J.M. Cherry, T. Chillingworth, C. Christiansen, J.-C. Chuat, E. Chung, C. Churcher, C.M. Churcher, M.W. Clark, M.L. Clemente, A. Coblenz, M. Coglievina, E. Coissac, L. Colleaux, R. Connor, R. Contreras, J. Cooper, T. Copsey, F. Coster, R. Coster, J. Couch, M. Crouzet, C. Cziepluch, B. Daignan-Fornier, F. Dal Paro, D.V. Dang, M. D'Angelo, C.J. Davies, K. Davis, R.W. Davis, A. De Antoni, S. Dear, K. Dedman, E. Defoor, M. de Haan, Th. Delaveau, S. Del Bino, M. Delgado, H. Delius, D. Delneri, F. Del Rey, J. Demolder, N. Démolis, K. Devlin, P. de Wergifosse, F.S. Dietrich, H. Ding, C. Dion, T. Dipaolo, F. Doignon, C. Doira, H. Domdey, J. Dover, Z. Du, E. Dubois, B. Dujon, M. Duncan, P. Durand, A. Düsterhöft, S. Düsterhus, T. Eki, M. El Bakkoury, L.G. Eide, K.-D. Entian, P. Eraso, D. Erdmann, H. Erfle, V. Escribano, M. Esteban, L. Fabiani, F. Fabre, C. Fairhead, B. Fartmann, A. Favello, G. Faye, H. Feldmann, L. Fernandes, F. Feroli, M. Feuermann, T. Fiedler, W. Fiers, U.N. Fleig, M. Flöth, G.M. Fobo, N. Fortin, F. Foury, M.C. Francingues-Gaillard, L. Franco, A. Fraser, J.D. Friesen, C. Fritz, L. Frontali, H. Fukuhara, L. Fulton, L.J. Fuller, C. Gabel, C. Gaillardin, L. Gaillon, F. Galibert, F. Galisson, P. Galland, F.-J. Gamo, C. Gancedo,

J.M. Garcia-Cantalejo, M.I. García-Gonzalez, J. J. Garcia-Ramirez, M. García-Saéz, H. Gassenhuber, M. Gatius, S. Gattung, C. Geisel, M.E. Gent, S. Gentles, M. Ghazvini, D. Gigot, V. Gilliquet, N. Glansdorff, A. Gómez-Peris, A. Gonzaléz, S.E. Goulding, C. Granotier, T. Greco, M. Grenson, P. Grisanti, L.A. Grivell, D. Grothues, U. Gueldener, P. Guerreiro, E. Guzman, M. Haasemann, B. Habbig, H. Hagiwara, J. Hall, K. Hallsworth, K. Hamberg, N. Hamlin, N.J. Hand, V. Hanemann, J. Hani, T. Hankeln, M. Hansen, D. Harris, D.E. Harris, G. Hartzell, D. Hatat, U. Hattenhorst, J. Hawkins, U. Hebling, J. Hegemann, C. Hein, A. Hennemann, K. Hennessy, C.J. Herbert, K. Hernandez, Y. Hernando, E. Herrero, K. Heumann, D. Heuss- Neitzel, N. Hewitt, R. Hiesel, H. Hilbert, F. Hilger, L. Hillier, K. Hinni, C. Ho, J. Hoenicka, B. Hofmann, J. Hoheisel, S. Hohmann, C.P. Hollenberg, K. Holmstrøm, O. Horaitis, T.S. Horsnell, M.-E. Huang, B. Hughes, S. Hunicke-Smith, S. Hunt, S.E. Hunt, K. Huse, R.W. Hyman, F. Iborra, K.J. Indge, I. Iraqui Houssaini, K. Isono, C. Jacq, M. Jacquet, A. Jacquier, K. Jagels, W. Jäger, C.M. James, J.C. Jauniaux, Y. Jia, M. Jier, A. Jimenez, D. Johnson, L. Johnston, M. Johnston, L. Jones, M. Jones, J.-L. Jonniaux, D.B. Kaback, T. Kallesøe, S. Kalman, A. Kalogeropoulos, L. Karpfinger-Hartl, D. Kashkari, C. Katsoulou, A. Kayser, A. Kelly, T. Keng, H. Keuchel, P. Kiesau, L. Kirchrath, J. Kirsten, K. Kleine, U. Kleinhans, R. Klima, C. Komp, E. Kordes, S. Korol, P. Kötter, C. Krämer, B. Kramer, W. Kramer, P. Kreisl, T. Kucaba, H. Kuester, O. Kurdi, P. Laamanen, M. J. Lafuente, O. Landt, G. Lanfranchi, Y. Langston, D. Lashkari, P. Latreille, G. Lauquin, T. Le, P. Legrain, Y. Legros, A. Lepingle, H. Lesveque, H. Leuther, H. Lew, C. Lewis, Z.Y. Li, S. Liebl, A. Lin, D. Lin, M. Logghe, A.J.E. Lohan, E. J. Louis, G. Lucchini, K. Lutzenkirchen, R. Lyck, G. Lye, A.C. Maarse, M. J. Maat, C. Macri, A. Madania, M. Maftahi, A. Maia e Silva, E. Maillier, L. Mallet, G. Mannhaupt, V. Manus, R. Marathe, C. Marck, A. Marconi, E. Mardis, E. Martegani, R. Martin, A. Mathieu, K.T.C. Maurer, M. J. Mazón, C. Mazzoni, D. McConnell, S. McDonald, R.A. McKee, A.D.K. McReynolds, P. Melchioretto, S. Menezes, F. Messenguy, H.W. Mewes, G. Michaux, N. Miller, O. Minenkova, T. Miosga, S. Mirtipati, S. Möller-Rieker, D. Möstl, F. Molemans, A. Monnet, A.-L. Monnier, M.A. Montague, M. Moro, D. Mosedale, D. Möstl, S. Moule, L. Mouser, Y. Murakami, S. Müller-Auer, J. Mulligan, L. Murphy, M. Muzi Falconi, M. Naitou, K. Nakahara, A. Namath, F. Nasr, L. Navas, A. Nawrocki, J. Nelson, U. Nentwich, P. Netter, R. Neu, C.S. Newlon, M. Nhan, D. Niblett, J.-M. Nicaud, R.K. Niedenthal, C. Nombela, D. Noone, R. Norgren, B. Nußbaumer, B. Obermaier, C. Odell, P. Öfner, C. Oh, K. Oliver, S.G. Oliver, B. F. Ouellette, M. Ozawa, V. Paces, C. Pallier, D. Pandolfo, L. Panzeri, S. Paoluzi, A.G. Parle-McDermott, S. Pascolo, N. Paricio, A. Pauley, L. Paulin, B.M. Pearson, D. Pearson, D. Peluso, J. Perea, M. Pérez-Alonso, J.E. Pérez-Ortin, A. Perrin, F.X. Petel, B. Pettersson, F. Pfeiffer, P. Philippsen, A. Piérard, E. Piravandi, R.J. Planta, P. Plevani, O. Poch, B. Poetsch, F.M. Pohl, T. M. Pohl, R. Pöhlmann, R. Poirey, D. Portetelle, F. Portillo, S. Potier, M. Proft, H. Prydz, A. Pujol, B. Purnelle, V. Puzos, M.A. Rajandream, M. Ramezani Rad, S.W. Rasmussen, A. Raynal, C. Rebischung, S. Rechmann, M. Remacha, J.L. Revuelta, P. Rice, G.-F. Richard, C. Richards, P. Richterich, M. Rieger, L. Rifkin, L. Riles, T. Rinaldi, M. Rinke, J. Robben, A.B. Roberts, D. Roberts, F. Rodriguez, E. Rodriguez-Belmonte, C. Rodrigues-Pousada, A.M. Rodriguez-Torres, M. Rose, R. Rossau, N. Rowley, T. Rupp, M. Ruzzi, W. Saeger, J.E. Saiz, M. Saliola, D. Salom, H.P. Saluz, M. Sánchez- Perez, M.A. Santos, E. Sanz, J.E. Sanz, A.-M. Saren, F. Sartorello, M. Sasanuma, S.-I. Sasanuma, T. Scarcez, I. Schaaf-Gerstenschläger, B. Schäfer, M. Schäfer, M. Scharfe, B. Scherens, N. Schroff, M. Sen-Gupta, T. Shibata, T. Schmidheini, E.R. Schmidt, C. Schneider, P. Scholler, S. Schramm, A. Schreer, M. Schröder, C. Schwager, S. Schwarz, C. Schwarzlose, B. Schweitzer, M. Schweizer, A.-M. Sdicu, P. Sehl, C. Sensen, J.G. Sgouros, T. Shogren, L. Shore, Y. Su, J. Skala, J. Skelton, P.P. Slonimski, P.H.M. Smits, V. Smith, H. Soares, E. Soeda, A. Soler-Mira, F. Sor, N. Soriano, J. L. Souciet, C. Soustelle, R. Spiegelberg, L. I. Stateva, H. Y. Steensma, J. Stegemann, S. Steiner, L. Stellyes, F. Sterky, R. K. Storms, H. St. Peter, R. Stucka, A. Taich, E. Talla, I. Tarassov, H. Tashiro, P. Taylor, C. Teodoru, H. Tettelin, A. Thierry, G. Thireos, E. Tobiasch, D. Tovan, E. Trevaskis, Y. Tsuchiya, M. Tzermia, M. Uhlen, A. Underwood, M. Unseld, J.H.M. Urbanus, A. Urrestarazu, S. Ushinsky, M. Valens, G. Valle, A. Van Broekhoven, M. Vandenbol, Q.J.M. Van der Aart, C.G. Van Der Linden, L. Van Dyck, M. Vanoni, J.C. van Vliet-Reedijk, A. Vassarotti, M. Vaudin, K. Vaughan, P. Verhasselt, I. Vetter, F. Vierendeels, D. Vignati, C. Vilela, S. Vissers, C. Vleck, D.T. Vo, D.H. Vo, M. Voet, G. Volckaert, D. Von Wettstein, H. Voss, P. Vreken, A. Wach, G. Wagner, S.V. Walsh, R. Wambutt, H. Wang, Y. Wang, J.R. Warmington, R. Waterston, M.D. Watson, N. Weber, E. Wedler, H. Wedler, Y. Wei, S. Whitehead, B.L. Wicksteed, S. Wiemann, L. Wilcox, C. Wilson, R. Wilson, A. Winant, E. Winnett, B. Winsor, P. Wipfli, S. Wölfl, P. Wohldman, K. Wolf, K.H. Wolfe, L.F. Wright, H. Wurst, G. Xu, M. Yamazaki, M.A. Yelton, K. Yokohama, A. Yoshikawa, S. Yuping, P. Zaccaria, M. Zagulski, F. K. Zimmermann, J. Zimmermann, M. Zimmermann, W.-W. Zhong, A. Zollner, E. Zumstein, Nature 387 (1997) 5.

[2] The C. elegans Sequencing Consortium, Science 282 (1998) 2012.

[3] M.D. Adams, S.E. Celniker, R.A. Holt, C.A. Evans, J.D. Gocayne, P.G. Amanatides, S.E. Scherer, P.W. Li, R.A. Hoskins, R.F. Galle, R.A. George, S.E. Lewis, S. Richards, M. Ashburner, S.N. Henderson, G.G. Sutton, J.R. Wortman, M.D. Yandell, Q. Zhang, L.X. Chen, R.C. Brandon, Y.H. Rogers, R.G. Blazej, M. Champe, B.D. Pfeiffer, K.H. Wan, C. Doyle, E.G. Baxter, G. Helt, C.R. Nelson, G.L. Gabor, J.F. Abril, A. Agbayani, H.J. An, C. Andrews-Pfannkoch, D. Baldwin, R.M. Ballew, A. Basu, J. Baxendale, L. Bayraktaroglu, E.M. Beasley, K.Y. Beeson, P.V. Benos, B.P. Berman, D. Bhandari, S. Bolshakov, D. Borkova, M.R. Botchan, J. Bouck, P. Brokstein, P. Brottier, K.C. Burtis, D.A. Busam, H. Butler, E. Cadieu, A. Center, I. Chandra, J.M. Cherry, S. Cawley, C. Dahlke, L.B. Davenport, P. Davies, B. de Pablos, A. Delcher, Z. Deng, A.D. Mays, I. Dew, S.M. Dietz, K.

Dodson, L.E. Doup, M. Downes, S. Dugan-Rocha, B.C. Dunkov, P. Dunn, K.J. Durbin, C.C. Evangelista, C. Ferraz, S. Ferriera, W. Fleischmann, C. Fosler, A.E. Gabrielian, N.S. Garg, W.M. Gelbart, K. Glasser, A. Glodek, F. Gong, J.H. Gorrell, Z. Gu, P. Guan, M. Harris, N.L. Harris, D. Harvey, T.J. Heiman, J.R. Hernandez, J. Houck, D. Hostin, K.A. Houston, T.J. Howland, M.H. Wei, C. Ibegwam, M. Jalali, F. Kalush, G.H. Karpen, Z. Ke, J.A. Kennison, K.A. Ketchum, B.E. Kimmel, C.D. Kodira, C. Kraft, S. Kravitz, D. Kulp, Z. Lai, P. Lasko, Y. Lei, A.A. Levitsky, J. Li, Z. Li, Y. Liang, X. Lin, X. Liu, B. Mattei, T.C. McIntosh, M.P. McLeod, D. McPherson, G. Merkulov, N.V. Milshina, C. Mobarry, J. Morris, A. Moshrefi, S.M. Mount, M. Moy, B. Murphy, L. Murphy, D.M. Muzny, D.L. Nelson, D.R. Nelson, K.A. Nelson, K. Nixon, D.R. Nusskern, J.M. Pacleb, M. Palazzolo, G.S. Pittman, S. Pan, J. Pollard, V. Puri, M.G. Reese, K. Reinert, K. Remington, R.D. Saunders, F. Scheeler, H. Shen, B.C. Shue, I. Siden-Kiamos, M. Simpson, M.P. Skupski, T. Smith, E. Spier, A.C. Spradling, M. Stapleton, R. Strong, E. Sun, R. Svirskas, C. Tector, R. Turner, E. Venter, A.H. Wang, X. Wang, Z.Y. Wang, D.A. Wassarman, G.M. Weinstock, J. Weissenbach, S.M. Williams, T. Woodage, K.C. Worley, D. Wu, S. Yang, Q.A. Yao, J. Ye, R.F. Yeh, J.S. Zaveri, M. Zhan, G. Zhang, Q. Zhao, L. Zheng, X.H. Zheng, F.N. Zhong, W. Zhong, X. Zhou, S. Zhu, X. Zhu, H.O. Smith, R.A. Gibbs, E.W. Myers, G.M. Rubin, J.C. Venter, Science 287 (2000) 2185.

[4] The Arabidopsis Genome Initiative, Nature 408 (2000) 796.

[5] International Human Genome Sequencing Consortium, Nature 409 (2001) 860.

[6] J.C. Venter, M.D. Adams, E.W. Myers, P.W. Li, R.J. Mural, G.G. Sutton, H.O. Smith, M. Yandell, C.A. Evans, R.A. Holt, J.D. Gocayne, P. Amanatides, R.M. Ballew, D.H. Huson, J.R. Wortman, Q. Zhang, C.D. Kodira, X.H. Zheng, L. Chen, M. Skupski, G. Subramanian, P.D. Thomas, J. Zhang, G.L. Gabor Miklos, C. Nelson, S. Broder, A.G. Clark, J. Nadeau, V.A. McKusick, N. Zinder, A.J. Levine, R.J. Roberts, M. Simon, C. Slayman, M. Hunkapiller, R. Bolanos, A. Delcher, I. Dew, D. Fasulo, M. Flanigan, L. Florea, A. Halpern, S. Hannenhalli, S. Kravitz, S. Levy, C. Mobarry, K. Reinert, K. Remington, J. Abu-Threideh, E. Beasley, K. Biddick, V. Bonazzi, R. Brandon, M. Cargill, I. Chandramouliswaran, R. Charlab, K. Chaturvedi, Z. Deng, V. Di Francesco, P. Dunn, K. Eilbeck, C. Evangelista, A.E. Gabrielian, W. Gan, W. Ge, F. Gong, Z. Gu, P. Guan, T.J. Heiman, M.E. Higgins, R.R. Ji, Z. Ke, K.A. Ketchum, Z. Lai, Y. Lei, Z. Li, J. Li, Y. Liang, X. Lin, F. Lu, G.V. Merkulov, N. Milshina, H.M. Moore, A.K. Naik, V.A. Narayan, B. Neelam, D. Nusskern, D.B. Rusch, S. Salzberg, W. Shao, B. Shue, J. Sun, Z. Wang, A. Wang, X. Wang, J. Wang, M. Wei, R. Wides, C. Xiao, C. Yan, A. Yao, J. Ye, M. Zhan, W. Zhang, H. Zhang, Q. Zhao, L. Zheng, F. Zhong, W. Zhong, S. Zhu, S. Zhao, D. Gilbert, S. Baumhueter, G. Spier, C. Carter, A. Cravchik, T. Woodage, F. Ali, H. An, A. Awe, D. Baldwin, H. Baden, M. Barnstead, I. Barrow, K. Beeson, D. Busam, A. Carver, A. Center, M.L. Cheng, L. Curry, S. Danaher, L. Davenport, R. Desilets, S. Dietz, K. Dodson, L. Doup, S. Ferriera, N. Garg, A. Gluecksmann, B. Hart, J. Haynes, C. Haynes, C. Heiner, S. Hladun, D. Hostin, J. Houck, T. Howland, C. Ibegwam, J. Johnson, F. Kalush, L. Kline, S. Koduru, A. Love, F. Mann, D. May, S. McCawley, T. McIntosh, I. McMullen, M. Moy, L. Moy, B. Murphy, K. Nelson, C. Pfannkoch, E. Pratts, V. Puri, H. Qureshi, M. Reardon, R. Rodriguez, Y.H. Rogers, D. Romblad, B. Ruhfel, R. Scott, C. Sitter, M. Smallwood, E. Stewart, R. Strong, E. Suh, R. Thomas, N.N. Tint, S. Tse, C. Vech, G. Wang, J. Wetter, S. Williams, M. Williams, S. Windsor, E. Winn-Deen, K. Wolfe, J. Zaveri, K. Zaveri, J.F. Abril, R. Guigo, M.J. Campbell, K.V. Sjolander, B. Karlak, A. Kejariwal, H. Mi, B. Lazareva, T. Hatton, A. Narechania, K. Diemer, A. Muruganujan, N. Guo, S. Sato, V. Bafna, S. Istrail, R. Lippert, R. Schwartz, B. Walenz, S. Yooseph, D. Allen, A. Basu, J. Baxendale, L. Blick, M. Caminha, J. Carnes-Stine, P. Caulk, Y.H. Chiang, M. Coyne, C. Dahlke, A. Mays, M. Dombroski, M. Donnelly, D. Ely, S. Esparham, C. Fosler, H. Gire, S. Glanowski, K. Glasser, A. Glodek, M. Gorokhov, K. Graham, B. Gropman, M. Harris, J. Heil, S. Henderson, J. Hoover, D. Jennings, C. Jordan, J. Jordan, J. Kasha, L. Kagan, C. Kraft, A. Levitsky, M. Lewis, X. Liu, J. Lopez, D. Ma, W. Majoros, J. McDaniel, S. Murphy, M. Newman, T. Nguyen, N. Nguyen, M. Nodell, S. Pan, J. Peck, M. Peterson, W. Rowe, R. Sanders, J. Scott, M. Simpson, T. Smith, A. Sprague, T. Stockwell, R. Turner, E. Venter, M. Wang, M. Wen, D. Wu, M. Wu, A. Xia, A. Zandieh, X. Zhu, Science 291 (2001) 1304.

[7] S.P. Fodor, J.L. Read, M.C. Pirrung, L. Stryer, A.T. Lu, D. Solas, Science 251 (1991) 767.

[8] M. Schena, D. Shalon, R.W. Davis, P.O. Brown, Science 270 (1995) 467.

[9] R.J. Cho, M.J. Campbell, E.A. Winzeler, L. Steinmetz, A. Conway, L. Wodicka, T.G. Wolfsberg, A.E. Gabrielian, D. Landsman, D.J. Lockhart, R.W. Davis, Mol. Cell 2 (1998) 65.

[10] P.T. Spellman, G. Sherlock, M.Q. Zhang, V.R. Iyer, K. Anders, M.B. Eisen, P.O. Brown, D. Botstein, B. Futcher, Mol. Biol. Cell 9 (1998) 3273.

[11] R.J. Cho, M. Huang, M.J. Campbell, H. Dong, L. Steinmetz, L. Sapinoso, G. Hampton, S.J. Elledge, R.W. Davis, D.J. Lockhart, Nat. Genet. 27 (2001) 48.

[12] S. Chu, J. DeRisi, M. Eisen, J. Mulholland, D. Botstein, P.O. Brown, I. Herskowitz, Science 282 (1998) 699.

[13] M. Primig, R.M. Williams, E.A. Winzeler, G.G. Tevzadze, A.R. Conway, S.Y. Hwang, R.W. Davis, R.E. Esposito, Nat. Genet. 26 (2000) 415.

[14] K.P. White, S.A. Rifkin, P. Hurban, D.S. Hogness, Science 286 (1999) 2179.

[15] V.R. Iyer, M.B. Eisen, D.T. Ross, G. Schuler, T. Moore, J.C. Lee, J.M. Trent, L.M. Staudt, J. Hudson Jr., M.S. Boguski, D. Lashkari, D. Shalon, D. Botstein, P.O. Brown, Science 283 (1999) 83.

[16] S.K. Kim, J. Lund, M. Kiraly, K. Duke, M. Jiang, J.M. Stuart, A. Eizinger, B.N. Wylie, G.S. Davidson, Science 293 (2001) 2087.

[17] P.O. Brown, D. Botstein, Nat. Genet. 21 (1999) 33.

[18] S.P. Gygi, Y. Rochon, B.R. Franza, R. Aebersold, Mol. Cell Biol. 19 (1999) 1720.

[19] B. Futcher, G.I. Latter, P. Monardo, C.S. McLaughlin, J.I. Garrels, Mol. Cell Biol. 19 (1999) 7357.
[20] T. Ideker, V. Thorsson, J.A. Ranish, R. Christmas, J. Buhler, J.K. Eng, R. Bumgarner, D.R. Goodlett, R. Aebersold, L. Hood, Science 292 (2001) 929.
[21] S. Brenner, Curr. Biol. 9 (1999) R671.
[22] S.G. Oliver, Philos. Trans. R. Soc. Lond. B. Biol. Sci. 357 (2002) 17.
[23] E.A. Winzeler, D.D. Shoemaker, A. Astromoff, H. Liang, K. Anderson, B. Andre, R. Bangham, R. Benito, J.D. Boeke, H. Bussey, A.M. Chu, C. Connelly, K. Davis, F. Dietrich, S.W. Dow, M. El Bakkoury, F. Foury, S.H. Friend, E. Gentalen, G. Giaever, J.H. Hegemann, T. Jones, M. Laub, H. Liao, N. Liebundguth, D.J. Lockhart, A. Lucau-Danila, M. Lussier, N. M'Rabet, P. Menard, M. Mittmann, C. Pai, C. Rebischung, J.L. Revuelta, L. Riles, C.J. Roberts, P. Ross-MacDonald, B. Scherens, M. Snyder, S. Sookhai-Mahadeo, R.K. Storms, S. Véronneau, M. Voet, G. Volckaert, T.R. Ward, R. Wysocki, G.S. Yen, K. Yu, K. Zimmermann, P. Philippsen, M. Johnston, R.W. Davis, Science 285 (1999) 901.
[24] G. Giaever, A.M. Chu, L. Ni, C. Connelly, L. Riles, S. Veronneau, S. Dow, A. Lacau-Danila, K. Anderson, B. Andre, A.P. Arkin, A. Astromoff, M. El-Bakkoury, R. Bangham, R. Benito, S. Brachat, S. Campanaro, M. Curtiss, K. Davis, A. Deutschbauer, K.D. Entian, P. Flaherty, F. Foury, D.J. Garfinkel, M. Gerstein, D. Gotte, U. Guldener, J.H. Hegemann, S. Hempel, Z. Herman, D.F. Jaramillo, T. Jones, D.E. Kelly, S.L. Kelly, P. Kotter, D. LaBonte, D.C. Lamb, N. Lan, H. Liang, H. Liao, L. Liu, C. Luo, M. Lussier, R. Mao, P. Menard, S.L. Ooi, J.L. Revuelta, C.J. Roberts, M. Rose, P. Ross-Macdonald, B. Scherens, G. Schimmack, B. Shafer, D.D. Shoemaker, S. Sookhai-Mahadeo, R.K. Storms, J.N. Strathern, G. Valle, M. Voet, G. Volckaert, C.Y. Wang, T.R. Ward, J. Wilhelmy, E.A. Winzeler, Y. Yang, G. Yen, E. Youngman, K. Yu, H. Bussey, J.D. Boecke, M. Snyder, P. Philippsen, R.W. Davis, M. Johnston, Nature 418 (2002) 387.
[25] A.M. Deutschbauer, R.W. Williams, A.M. Chu, R.W. Davis, in preparation.
[26] L.M. Steinmetz, C. Scharfe, A.M. Deutschbauer, D. Mokranjac, Z.S. Herman, T. Jones, A.M. Chu, G. Giaever, H. Prokisch, P.J. Oefner, R.W. Davis, Nat. Genet. 31 (2002) 400.
[27] C.A. Davis, L. Grate, M. Spingola, M. Ares Jr., Nucleic Acids Res. 28 (2000) 1700.
[28] B. Modrek, C. Lee, Nat. Genet. 30 (2002) 13.
[29] D.D. Shoemaker, E.E. Schadt, C.D. Armour, Y.D. He, P. Garrett-Engele, P.D. McDonagh, P.M. Loerch, A. Leonardson, P.Y. Lum, G. Cavet, L.F. Wu, S.J. Altschuler, S. Edwards, J. King, J.S. Tsang, G. Schimmack, J.M. Schelter, J. Koch, M. Ziman, M.J. Marton, B. Li, P. Cundiff, T. Ward, J. Castle, M. Krolewski, M.R. Meyer, M. Mao, J. Burchard, M.J. Kidd, H. Dai, J.W. Phillips, P.S. Linsley, R. Stoughton, S. Scherer, M.S. Boguski, Nature 409 (2001) 922.
[30] G.K. Hu, S.J. Madore, B. Moldover, T. Jatkoe, D. Balaban, J. Thomas, Y. Wang, Genome Res. 11 (2001) 1237.
[31] B. Ren, F. Robert, J.J. Wyrick, O. Aparicio, E.G. Jennings, I. Simon, J. Zeitlinger, J. Schreiber, N. Hannett, E. Kanin, T.L. Volkert, C.J. Wilson, S.P. Bell, R.A. Young, Science 290 (2000) 2306.
[32] V.R. Iyer, C.E. Horak, C.S. Scafe, D. Botstein, M. Snyder, P.O. Brown, Nature 409 (2001) 533.
[33] F.C. Holstege, E.G. Jennings, J.J. Wyrick, T.I. Lee, C.J. Hengartner, M.R. Green, T.R. Golub, E.S. Lander, R.A. Young, Cell 95 (1998) 717.
[34] J.L. DeRisi, V.R. Iyer, P.O. Brown, Science 278 (1997) 680.
[35] T.R. Hughes, M.J. Marton, A.R. Jones, C.J. Roberts, R. Stoughton, C.D. Armour, H.A. Bennett, E. Coffey, H. Dai, Y.D. He, M.J. Kidd, A.M. King, M.R. Meyer, D. Slade, P.Y. Lum, S.B. Stepaniants, D.D. Shoemaker, D. Gachotte, K. Chakraburtty, J. Simon, M. Bard, S.H. Friend, Cell 102 (2000) 109.
[36] A.C. Spradling, D. Stern, A. Beaton, E.J. Rhem, T. Laverty, N. Mozden, S. Misra, G.M. Rubin, Genetics 153 (1999) 135.
[37] A. Peter, P. Schottler, M. Werner, N. Beinert, G. Dowe, P. Burkert, F. Mourkioti, L. Dentzer, Y. He, P. Deak, P.V. Benos, M.K. Gatt, L. Murphy, D. Harris, B. Barrell, C. Ferraz, S. Vidal, C. Brun, J. Demaille, E. Cadieu, S. Dreano, S. Gloux, V. Lelaure, S. Mottier, F. Galibert, D. Borkova, B. Minana, F.C. Kafatos, S. Bolshakov, I. Siden-Kiamos, G. Papagiannakis, L. Spanos, C. Louis, E. Madueno, B. de Pablos, J. Modolell, A. Bucheton, D. Callister, L. Campbell, N.S. Henderson, P.J. McMillan, C. Salles, E. Tait, P. Valenti, R.D. Saunders, A. Billaud, L. Pachter, R. Klapper, W. Janning, D.M. Glover, M. Ashburner, H.J. Bellen, H. Jackle, U. Schafer, EMBO Rep. 3 (2002) 34.
[38] P. Ross-Macdonald, P.S. Coelho, T. Roemer, S. Agarwal, A. Kumar, R. Jansen, K.H. Cheung, A. Sheehan, D. Symoniatis, L. Umansky, M. Heidtman, F.K. Nelson, H. Iwasaki, K. Hager, M. Gerstein, P. Miller, G.S. Roeder, M. Snyder, Nature 402 (1999) 413.
[39] N. Judson, J.J. Mekalanos, Trends Microbiol. 8 (2000) 521.
[40] B.P. Zambrowicz, G.A. Friedrich, E.C. Buxton, S.L. Lilleberg, C. Person, A.T. Sands, Nature 392 (1998) 608.
[41] S. Parinov, V. Sundaresan, Curr. Opin. Biotechnol. 11 (2000) 157.
[42] R.C. Meissner, H. Jin, E. Cominelli, M. Denekamp, A. Fuertes, R. Greco, H.D. Kranz, S. Penfield, K. Petroni, A. Urzainqui, C. Martin, J. Paz-Ares, S. Smeekens, C. Tonelli, B. Weisshaar, E. Baumann, V. Klimyuk, S. Marillonnet, K. Patel, E. Speulman, A.F. Tissier, D. Bouchez, J.J. Jones, A. Pereira, E. Wisman et al., Plant Cell 11 (1999) 1827.
[43] A.F. Tissier, S. Marillonnet, V. Klimyuk, K. Patel, M.A. Torres, G. Murphy, J.D. Jones, Plant Cell 11 (1999) 1841.
[44] K.R. Thomas, M.R. Capecchi, Cell 51 (1987) 503.
[45] Y.S. Rong, K.G. Golic, Science 288 (2000) 2013.
[46] K.J. McCreath, J. Howcroft, K.H. Campbell, A. Colman, A.E. Schnieke, A.J. Kind, Nature 405 (2000) 1066.
[47] P.D. Zamore, Nat. Struct. Biol. 8 (2001) 746.
[48] V. Smith, K.N. Chou, D. Lashkari, D. Botstein, P.O. Brown, Science 274 (1996) 2069.
[49] M. Hensel, J.E. Shea, C. Gleeson, M.D. Jones, E. Dalton, D.W. Holden, Science 269 (1995) 400.
[50] A.V. Karlyshev, P.C. Oyston, K. Williams, G.C. Clark, R.W. Titball, E.A. Winzeler, B.W. Wren, Infect. Immun. 69 (2001) 7810.

[51] G.W. Birrell, G. Giaever, A.M. Chu, R.W. Davis, J.M. Brown, Proc. Natl. Acad. Sci. USA 98 (2001) 12608.
[52] S.L. Ooi, D.D. Shoemaker, J.D. Boeke, Science 294 (2001) 2552.
[53] A.G. Fraser, R.S. Kamath, P. Zipperlen, M. Martinez-Campos, M. Sohrmann, J. Ahringer, Nature 408 (2000) 325.
[54] P. Gonczy, G. Echeverri, K. Oegema, A. Coulson, S.J. Jones, R.R. Copley, J. Duperon, J. Oegema, M. Brehm, E. Cassin, E. Hannak, M. Kirkham, S. Pichler, K. Flohrs, A. Goessen, S. Leidel, A.M. Alleaume, C. Martin, N. Ozlu, P. Bork, A.A. Hyman, Nature 408 (2000) 331.
[55] J.R. Kennerdell, R.W. Carthew, Cell 95 (1998) 1017.
[56] A. Nasevicius, S.C. Ekker, Nat. Genet. 26 (2000) 216.
[57] F. Wianny, M. Zernicka-Goetz, Nat. Cell Biol. 2 (2000) 70.
[58] S.M. Elbashir, J. Harborth, W. Lendeckel, A. Yalcin, K. Weber, T. Tuschl, Nature 411 (2001) 494.
[59] J. Ziauddin, D.M. Sabatini, Nature 411 (2001) 107.
[60] A.H. Tong, M. Evangelista, A.B. Parsons, H. Xu, G.D. Bader, N. Page, M. Robinson, S. Raghibizadeh, C.W. Hogue, H. Bussey, B. Andrews, M. Tyers, C. Boone, Science 294 (2001) 2364.
[61] A. Pandey, M. Mann, Nature 405 (2000) 837.
[62] P.K. Jensen, L. Pasa-Tolic, G.A. Anderson, J.A. Horner, M.S. Lipton, J.E. Bruce, R.D. Smith, Anal. Chem. 71 (1999) 2076.
[63] S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, R. Aebersold, Nat. Biotechnol. 17 (1999) 994.
[64] H. Zhu, M. Bilgin, R. Bangham, D. Hall, A. Casamayor, P. Bertone, N. Lan, R. Jansen, S. Bidlingmaier, T. Houfek, T. Mitchell, P. Miller, R.A. Dean, M. Gerstein, M. Snyder, Science 293 (2001) 2101.
[65] S. Fields, O. Song, Nature 340 (1989) 245.
[66] P. Uetz, L. Giot, G. Cagney, T.A. Mansfield, R.S. Judson, J.R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleisch, G. Vijayadamodar, M. Yang, M. Johnston, S. Fields, J.M. Rothberg, Nature 403 (2000) 623.
[67] T. Ito, T. Chiba, R. Ozawa, M. Yoshida, M. Hattori, Y. Sakaki, Proc. Natl. Acad. Sci. USA 98 (2001) 4569.
[68] G. Neubauer, A. King, J. Rappsilber, C. Calvio, M. Watson, P. Ajuh, J. Sleeman, A. Lamond, M. Mann, Nat. Genet. 20 (1998) 46.
[69] A.C. Gavin, M. Bosche, R. Krause, P. Grandi, M. Marzioch, A. Bauer, J. Schultz, J.M. Rick, A.M. Michon, C.M. Cruciat, M. Remor, C. Hofert, M. Schelder, M. Brajenovic, H. Ruffner, A. Merino, K. Klein, M. Hudak, D. Dickson, T. Rudi, V. Gnau, A. Bauch, S. Bastuck, B. Huhse, C. Leutwein, M.A. Heurtier, R.R. Copley, A. Edelmann, E. Querfurth, V. Rybin, G. Drewes, M. Raida, T. Bouwmeester, P. Bork, B. Seraphin, B. Kuster, G. Neubauer, G. Superti-Furga, Nature 415 (2002) 141.
[70] Y. Ho, A. Gruhler, A. Heilbut, G.D. Bader, L. Moore, S.L. Adams, A. Millar, P. Taylor, K. Bennett, K. Boutilier, L. Yang, C. Wolting, I. Donaldson, S. Schandorff, J. Shewnarane, M. Vo, J. Taggart, M. Goudreault, B. Muskat, C. Alfarano, D. Dewar, Z. Lin, K. Michalickova, A.R. Willems, H. Sassi, P.A. Nielsen, K.J. Rasmussen, J.R. Andersen, L.E. Johansen, L.H. Hansen, H. Jespersen, A. Podtelejnikov, E. Nielsen, J. Crawford, V. Poulsen, B.D. Sorensen, J. Matthiesen, R.C. Hendrickson, F. Gleeson, T. Pawson, M.F. Moran, D. Durocher, M. Mann, C.W. Hogue, D. Figeys, M. Tyers, Nature 415 (2002) 180.
[71] M.R. Martzen, S.M. McCraith, S.L. Spinelli, F.M. Torres, S. Fields, E.J. Grayhack, E.M. Phizicky, Science 286 (1999) 1153.
[72] H.C. Causton, B. Ren, S.S. Koh, C.T. Harbison, E. Kanin, E.G. Jennings, T.I. Lee, H.L. True, E.S. Lander, R.A. Young, Mol. Biol. Cell 12 (2001) 323.
[73] D.D. Shoemaker, D.A. Lashkari, D. Morris, M. Mittmann, R.W. Davis, Nat. Genet. 14 (1996) 450.