The yeast *genome* directory

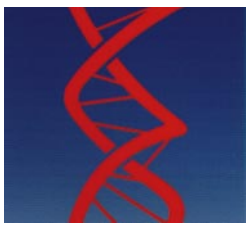
nature 387 **suppl.**
(29 May 1997)

nature

Supplement to Nature 29 May 1997

Volume 387 | issue no. 6632S

Nature **gratefully acknowledges the sponsorship from the European Commission, DG XII – Life Sciences, without which this supplement and its associated CD-ROM could not have been produced.**



introduction

maps

letters

Residents of a particular city notoriously have never visited most sights shown to tourists. In the same way, the complete genomic sequence of the brewer's yeast, *Saccharomyces cerevisiae*, has now been available for over a year, proving its usefulness to molecular biologists in a hundred quiet ways, without attracting the concerted attention that is its due. This *Directory* is intended to remedy that situation, by presenting papers on all the hitherto undescribed chromosomal sequences alongside an overview of the entire genomic sequence.

Like many a monument ignored by locals, familiarity has obscured what an achievement it represents. For one thing, it is still almost three times as large as any other genome sequenced so far. Then, too, this size necessitated a unique international collaboration which will surely inspire related projects in the future. Most importantly, it represents the first complete sequence of a eukaryotic genome, the last of the three superkingdoms of life to have a fully sequenced representative.

As such, it will be a rich mine of information for everyone concerned with all kinds of eukaryotic genomes, up to and including those of human beings. The opportunity this represents is equalled only by the challenge: dauntingly, of the 5,800 open reading frames in the sequence, the function of 45% cannot even be guessed. For now, we can only be grateful to the European Commission (DG XII – Life Sciences) for initiating and supporting, since 1989, the international project whose success has allowed *Nature* to present a unified account of this vital resource, and look forward to the flood of information that its study will surely release. Nicholas Short

Porters South, 4 Crinan St,
London N1 9XW, UK
Tel 44 (0) 171 833 4000
Fax 44 (0) 171 843 4596/7
e-mail: nature@nature.com
http://www.nature.com

Macmillan Magazines Ltd

Nature® (ISSN 0028-0836) is published weekly on Thursday, except the last week in December, by Macmillan Magazines Ltd (Porters South, 4 Crinan Street, London N1 9XW). Registered as a newspaper at the British Post Office. Annual subscription for USA and Canada US$495 (institutional/corporate), US$145 (individual making personal payment). Canada residents please add 7% GST (No. 140911595). US and Canadian orders to: Nature, Subscription Dept, P. O. Box 5055, Brentwood, TN 37024-5055, USA. Other orders to Nature, Brunel Road, Basingstoke, Hants RG21 2XS, UK. Periodicals postage paid at New York, NY 10010-1707, and additional mailing offices. Authorization to photocopy material for internal or personal use, or internal or personal use of specific clients, is granted by Nature to libraries and others registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided the base fee of $12.00 an article (or $2.00 a page) is paid direct to CCC, 222 Rosewood Drive, Danvers, MA 01923, USA. Identification code for Nature: 0028-0836/97 $12.00+$2.00. US Postmaster send address changes to: Nature, 345 Park Avenue South, New York NY 10010-1707, US. Published in Japan by Nature Japan K.K., Shin-Mitsuke Bldg. 36 Ichigaya

The Yeast Genome Directory

This directory was made possible by a unique international collaboration between the 633 scientists whose names appear below. It represents both the first published description of the complete sequence of most chromsomes from *Saccharomyces cerevisiae*, and the first published overview of the entire sequence. As such, the authors would like future papers referring to the entire sequence and/or its contents to cite this directory; future papers referring to the sequence of individual chromosomes should refer to the papers listed at the head of page 9. The authors' affiliations appear in the papers describing the individual chromosomes.

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, J. Alt-Mörbe, B. André, S. Andrews, W. Ansorge, G. Antoine, R. Anwar, A. Aparicio, R. Araujo, J. Arino, F. Arnold, J. Arroyo, E. Aviles, U. Backes, M. C. Baclet, K. Badcock, A. Bahr, 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, 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, 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, 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, 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, C. 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, 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. Patricio, 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, S. Rechmann, M. Remacha, J. L. Revuelta, P. Rice, G-F. Richard, P. Richterich, M. Rieger, L. Rifken, L. Riles, T. Rinaldi, M. Rinke, A. B. Roberts, D. Roberts, F. Rodriguez, E. Rodriguez-Belmonte, C. Rodriguez-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. Shu, J. Skala, J. Skelton, P. P. Slonimski, P. H. M. Smit, 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, 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. Yamasaki, 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.

Overview of the yeast genome

H. W. Mewes, K. Albermann, M. Bähr, D. Frishman, A. Gleissner, J. Hani, K. Heumann, K. Kleine, A. Maierl, S. G. Oliver[1], F. Pfeiffer & A. Zollner

Max-Planck-Institut für Biochemie, D-82152 Martinsried, Germany
[1]University of Manchester Institute of Science And Technology (UMIST), Sackville Street, Manchester M60 1QD, UK

The collaboration of more than 600 scientists from over 100 laboratories to sequence the *Saccharomyces cerevisiae* genome was the largest decentralised experiment in modern molecular biology and resulted in a unique data resource representing the first complete set of genes from a eukaryotic organism. 12 million bases were sequenced in a truly international effort involving European, US, Canadian and Japanese laboratories. While the yeast genome represents only a small fraction of the information in today's public sequence databases, the complete, ordered and non-redundant sequence provides an invaluable resource for the detailed analysis of cellular gene function and genome architecture. In terms of throughput, completeness and information content, yeast has always been the lead eukaryotic organism in genomics; it is still the largest genome to be completely sequenced.

The *Yeast Genome Directory* presents the basic features of this sequence: the arrangement of the 6,000 genes on 16 chromosomes; a summary of the function of the encoded proteins; and a view of the genome's architecture, based on an exhaustive intra-genomic sequence comparison[1]. The complete yeast sequence can be retrieved from a number of public databases, as well as from specialized World-Wide Web sites, which provide sophisticated query interfaces (see Box). These data are maintained and updated continuously. Although the form of the genome directory shown in this printed volume must present static information, its intention is to document the interpretation of the yeast genome shortly after its completion. To present the genome in a printable form, we have had to include a very limited selection of fact-oriented data. The *Yeast Genome Directory* cannot answer the question "What's in the yeast genome?" exhaustively, but summarizes what is known to be in it.

The sequence The final sequence was assembled from roughly 300,000 independent sequence reads, with error rates from 0.5 to 1%, resulting in an estimated error rate of the final sequence of less than 3 errors in 10,000 bases (0.03%). For the European Union part of the sequencing effort, a central database and informatics used to assemble, verify and analyse contiguous sequences submitted by the participating laboratories was provided by the Martinsrieder Institut für Protein Sequenzen (MIPS). The complete sequence was made available to the public on 24 April 1996. The first map represents the open reading frames (ORFs) of each of the successive chromosomes, at a uniform scale of 5 mm per kilobase. ORFs are named according to the location of the gene, using the convention Y (for yeast) followed by a letter denoting the chromosome (A for I, B for II, and so forth), a letter denoting the arm (R or L), a three-digit code ordering the ORFs from the centromere, and a letter denoting the coding strand (w or c).

Duplications The availability of the complete sequence of yeast allows us for the first time to examine the evolution of a eukaryotic species in a truly comprehensive manner. The footprints that indicate the evolutionary path taken by the yeast genome may be recognized by internal similarities between distinct regions of the present-day genome. Our approach to the inspection of these relationships is based on an all-against-all comparison of the genomic sequence data, applying local sequence alignments. Investigation of the yeast genome involved more than 24,000 blocks of 500 nucleotides. For each block, the six-frame translation into protein sequences was also generated, to allow for the concurrent comparison of DNA and amino-acid sequences[2].

Once an all-against-all matching of the yeast genome had been accomplished, duplication patterns within the genome could be investigated in a systematic way. The frequent, collinear block duplications found by our method seem to be an important consequence of the evolutionary development of *S. cerevisiae*. We have systematically inspected the genome for clusters of genes that have been produced by local duplication events. This involved evaluating the parameters that describe a cluster: it's size; the degree of similarity of the duplicated ORFs; and the compactness of the cluster. Scanning this parameter-space at a sensitivity just above noise level (50% identity on the DNA level), we found that a window size of 25 kb and an enforced compactness of 10% of the coding region in that window generated an optimal representation of gene clusters. These criteria allowed us to identify 53 regions of clustered gene duplications, not including the well-known high level of similarity in the telomeric and subtelomeric regions. The second map shows the set of all collinear clusters of genes in the genome as a two-page fold-out. The significant number of gene duplications in yeast must reflect an evolutionarily successful strategy: gene duplications allow for evolutionary modifications in one of the copies without disturbing possibly vital functions of the other.

Gene function The computational analysis of the yeast genome is a challenging task[3]. The scientific assessment of raw sequence data aims to connect genetic entities to biological knowledge, either by computational analysis or by linking sequence-deduced information to other experimental evidence (such as a genetic locus). Sophisticated data modelling is required to reflect the correct relationship between the sequence and any associated information and to allow for the consistent integration of complex, heterogeneous biological data. For example, the precursor of the ADP/ATP carrier protein AAC2p is represented five times in the EMBL Nucleic Acid Data Library, although four of the coding sequences reported are identical. The current nucleic acid databases (EBI/NCBI/DDBJ) contain 2,678 entries covering the yeast genome with a high degree of redundancy and inconsistency. Thus the traditional model of collecting

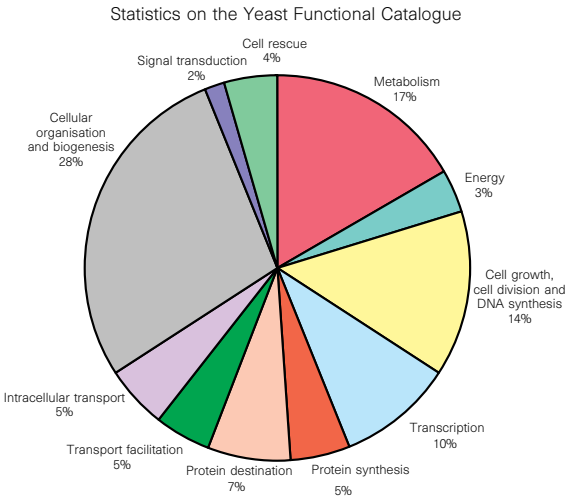


Figure 1 This shows the relative number of ORFs assigned to individual categories in the Gazetteer. There are eleven functional categories. Only proteins with a known function, or similarity or strong similarity to known proteins were assigned to one of the categories (similarities were measured by FASTA scores). In total, 3167 ORFs were assigned to at least one category. A single ORF can be assigned to more than one category.

database entries and presenting them as individual reports does not seem to be suitable for coping with the data analysis of complete genomes. Like the printed page, the traditional layout of the sequence databases is static, incorporating information from scientific annotation at the time of publication. Additional information extracted from the literature must be translated manually into the formal framework of a database entry. The knowledge of the functional properties of an uncharacterized gene may be published independently, and family members found in other organisms may allow for characterization by homology. Suitable models must thus be developed to cope with the data analysis of a complete genome, and to enable integrated views of the genomic sequence and associated information.

Only 43.3 % of the yeast genes are currently classified as 'functionally characterized', having experimentally well-investigated properties, being members of well-defined protein families, or displaying strong homology to proteins with known biochemical functions. The systematic functional analysis of these genes, currently in progress[4], will identify the functions of many of the uncharacterized 'orphans'. Various experimental methods, including improved '*in silico*' analysis, will also increase dramatically the information content of the biological databases. Previously, individual attempts have been made to analyse specific chromosomes systematically by sequence analysis[5,6], and the results of an automated software system, 'Genecrunch', to 'crunch' the complete yeast genome have been offered as an Internet service (http://genecrunch.sgi.com).

To provide information on the biochemical and physiological context of protein function, we have compiled a gazetteer listing all the ORFs that can be related to well-understood functions. Similarities between biological sequences were measured by FASTA scores[7]: a FASTA score between 100 and 200 was defined as a 'weak similarity'; between 200 and a third of the self-score (the score of the protein when aligned with itself) for the protein was defined as a 'similarity'; and higher than a third of the self-score was defined as a 'strong similarity'. Weak similarities were not considered. In addition to the similarity scores, we have used pattern data[8], including experimental data from the literature combined with genetic data, to characterize ORFs.

Each entry in the gazetteer lists the ORF name (defined as above), the gene name (if any), and the name or a short description of the protein. Entries are divided into 11 categories representing the cellular function of the individual ORFs (such as metabolism, energy or transcription), and each category is divided into sub-categories[9]. The yeast genome encodes about 5,800 proteins, less than half of which are 'known' in the sense that they have been genetically and biochemically well characterized. For about 20% of the remaining proteins, the experimental data are heterogeneous and provide only some indication of their functions *in vivo*. The remaining 38% either show similarities to other uncharacterized proteins or show no similarities at all. Defined categories provoke redundant entries, such as the classification of multifunctional enzymes. As gene names are used elsewhere in a very inconsistent way, we have used the name from the *Saccharomyces* genome database (see Box) register whenever possible. Gene names that are used twice for different ORF names are written in brackets.

Sequences of common evolutionary origin (homologous *sensu stricto*) reflect their relatedness by sequence similarities, and the organization of related primary, secondary and tertiary structures into groups ('families') remains the most powerful principle in sequence data analysis. We have used this principle to cluster proteins into families and superfamilies, following a previous scheme[10], allowing us to cope with the many taxonomic complications inherent in protein evolution.

The World-Wide Web site of the Martinsrieder Institut für Protein Sequenzen yeast resource combines information generated by automated procedures with the results of systematic analysis by yeast researchers. Users can: (1) visualize chromosomes and selected regions to inspect genetic elements, such as ORFs, Ty's, tRNAs etc.; (2) receive detailed information on a yeast gene by searching with accession numbers, systematic codes, or gene names; (3) browse yeast genes according to their functional classification; (4) search for human homologues; (5) obtain information on functional properties; (6) download nucleic-acid or protein sequence data; (7) inspect up-to-date sequence homologies and alignments (FASTA database); (8) browse the family and superfamily classification of yeast proteins; (9) search the yeast genome interactively for sequence patterns and sequence similarities; and (10) inspect the yeast genome for gene redundancy.

As a complement to the printed information in the *Yeast Genome Directory*, a CD-ROM compiled by Martinsrieder Institut für Protein Sequenzen, is available on request to subscribers of *Nature* and *Science* for the exploration of the yeast genome on a local, network-independent installation. The basic functionality of the retrieval software and the databases incorporated is equivalent to our World-Wide Web resource. The CD-ROM can be installed on Windows95, WindowsNT and Power-Macintosh operating systems, and is accompanied by a detailed description of its functionality, installation procedures and system requirements. □

1. Goffeau, A. *et al. Science* 274, 546–567 (1996).
2. Heumann, K., Harris, C. & Mewes, H. W. in *Proceedings of the Fourth Int. Conf. Intelligent Systems for Molecular Biology* 98–108(AAAI Press, Menlo Park, 1996). (1996).
3. Casari, G. *et al. Nature* 376, 647–648 (1995).
4. Oliver, S. G. *Trends Genet.* 12, 241–242 (1996).
5. Bork, P. *et al. Protein Sci.* 1, 1677–1690 (1992).
6. Koonin, E. V. *et al. EMBO J.* 13, 493–503 (1994).
7. Pearson, B. & Lipman, D. *Proc. Natl Acad Sci. USA* 85, 2444–2448 (1988).
8. Bairoch A., Bucher, P. & Hoffman K. *Nucleic Acids Res.* 24, 189–196 (1996).
9. Riley, M. *Microbiol. Rev.* 57, 862–869 (1993).
10. Dayhoff, M. O. *Fed. Proc.* 35, 2132–2138 (1976).

Acknowledgements. We thank A. Goffeau for support and encouragement; B. Jasny for suggestions; and Netscape Communications Inc. for allowing the limited distribution of Netscape 3.01. This work was supported the European Commission, the Max-Planck-Society, the Forschungszentrum für Umwelt und Gesundheit, and the Yeast Industry Platform.

Correspondence and requests for materials should be addressed to H.W.M. (e-mail: mewes@mips. embnet.org).

BOX Useful World-Wide Web addresses

Yeast databases

Munich Information Centre for Protein Sequences (MIPS)
http://www.mips.biochem.mpg.de/mips/yeast/

Yeast Protein Database (YPD)
http://quest7.proteome.com/YPDhome.html

Saccharomyces Genome Database (SGD)
http://genome-www.stanford.edu/Saccharomyces/

Specialized Yeast Databases

Sacch3D – Structural information for yeast proteins
http://genome-www.stanford.edu/Sacch3D/

Yeast Gene Duplications
http://acer.gen.tcd.ie/~khwolfe/yeast/topmenu.html

Related human disease genes (NIH XREFdb)
http://www.ncbi.nlm.nih.gov/XREFdb/

Genetic and physical maps (hyperlinked to biological information)
http://genome-www.stanford.edu/cgi-bin/SGD/pgMAP/pgMap

NIH yeast information page
http://www.ncbi.nlm.nih.gov/Yeast/budding.html

Schizosaccharomyces pombe
http://www2.bio.uva.nl/pombe/

Candida albicans information
http://alces.med.umn.edu/Candida.html

Common DNA and protein databases

European Bioinformatics Institute (EBI)
http://www.ebi.ac.uk/services/services.html

National Center for Biotechnology Information (NCBI)
http://www.ncbi.nlm.nih.gov/

DNA Data Bank of Japan (DDBJ)
http://www.ddbj.nig.ac.jp/

PIR-International
http://www.mips.biochem.mpg.de

SwissProt
http://expasy.hcuge.ch/sprot/sprot-top.html

References to be cited for the complete sequence and individual chromosomes

Complete sequence, The Yeast Genome Directory
Goffeau, A. *et al. Nature* 387 (suppl.), 1–105 (1997)

Chromosome 1, The nucleotide sequence of chromosome 1 from *Saccharomyces cerevisiae*
Bussey, H. *et al. Proc. Natl Acad. Sci. USA* 92, 3809–3813 (1995)
EMBL accession numbers: U12980, L20125, L05146, L22015, L28920

Chromosome 2, Complete DNA sequence of yeast chromosome II
Feldmann, H. *et al. EMBO J.* 13, 5795–5809 (1994)
EMBL accession numbers: Y08934, Z35762-Z35869, Z35870-Z36171

Chromosome 3, The complete DNA sequence of yeast chromosome III
Oliver, S. *et al. Nature* 357, 38–46 (1992)
EMBL accession number: X59720

Chromosome 4, The nucleotide sequence of *Saccharomyces cerevisiae* chromosome IV
Jacq, C. *et al. Nature* 387 (suppl.), 75–78 (1997)
EMBL accession numbers: Z74051-Z74296, Z48008, Z74305-Z74385, Z49770, Z47814, Z68196, Z54075, Z49812, Z49209, Z46796, Z50111, Z47746, Z48758, Z48179, Z54139, Z50046, Z47813, Z46727, Z48784, Z68194, Z68195, Z48612, Z49701, Z68329, Z70202, Z68290, U51030, U51031, U28374, U32517, U51032, U28372, U28373, U32274, U33050, U33057, Z74389

Chromosome 5, The nucleotide sequence of *Saccharomyces cerevisiae* chromosome V
Dietrich, F. S. *et al. Nature* 387 (suppl.), 78–81 (1997)
EMBL accession numbers: U73806, U18795, L10830, U18779, U18530, U18778, U18796, U18813, U18814, U18839, U18916, U18917, L10718, U18922

Chromosome 6, Analysis of the nucleotide sequence of chromosome VI from *Saccharomyces cerevisiae*
Murakami, Y. *et al. Nature Genet.* 10, 261–268 (1995)
EMBL accession number: D50617

Chromosome 7, The nucleotide sequence of *Saccharomyces cerevisiae* chromosome VII
Tettelin, H. *et al. Nature* 387 (suppl.), 81–84 (1997)
EMBL accession numbers: X94357, Z72785-Z73081

Chromosome 8, Complete nucleotide sequence of Saccharomyces cerevisiae Chromosome VIII
Johnston, M. *et al. Science* 265, 2077–2082 (1994)
EMBL accession numbers: U11583, U11582, U11581, U10555, U10400, U10399, U00062, U00061, U10556, U00060, U00059, U10398, U10397, U00027, U00028, U00030, U00029

Chromosome 9, The nucleotide sequence of *Saccharomyces cerevisiae* chromosome IX
Barrell, B. G. *et al. Nature* 387 (suppl.), 84–87 (1997)
EMBL accession number: Z47047

Chromosome 10, Complete nucleotide sequence of *Saccharomyces cerevisiae* Chromosome X
Galibert, F. *et al. EMBO J.* 15, 2031–2049 (1996)
EMBL accession numbers: Z34098, Z49276-Z49662

Chromosome 11, Complete DNA sequence of yeast chromosome XI
Dujon, B. *et al. Nature* 369, 371–378 (1994)
EMBL accession numbers: Z28001-Z28330

Chromosome 12, The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XII
Johnston, M. *et al. Nature* 387 (suppl.), 87–90 (1997)
EMBL accession numbers: Z73106-Z73327, U53879, U51921, U17246, U14913, U19027, U20865, U17244, U17245, U17243, U17247, U20618, U19028, U19102, U19103, U19104, U19729, U20162, U20939, U21094, U22382, U22383

Chromosome 13, The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XIII
Barrell, B. *et al. Nature* 387 (suppl.), 90–93 (1997)
EMBL accession numbers: Z50178, Z49218, Z49210, X80835, Z46660, Z46373, Z38114, Z46729, Z47816, Z48430, Z46659, Z49810, Z48613, Z49211, Z49213, Z48502, Z49703, Z48952, Z49259, Z49807, Z50179, Z49702, Z49273, Z48622, Z47071, Z49700, Z49808, Z47815, Z48755, Z49809, Z49939, Z48756, Z48639, Z49260, Z49704, X80836, Z49212, Z54141

Chromosome 14, The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XIV and its evolutionary implications
Philippsen, P. *et al. Nature* 387 (suppl.), 93–98 (1997)
EMBL accession numbers: Z71277-Z71692

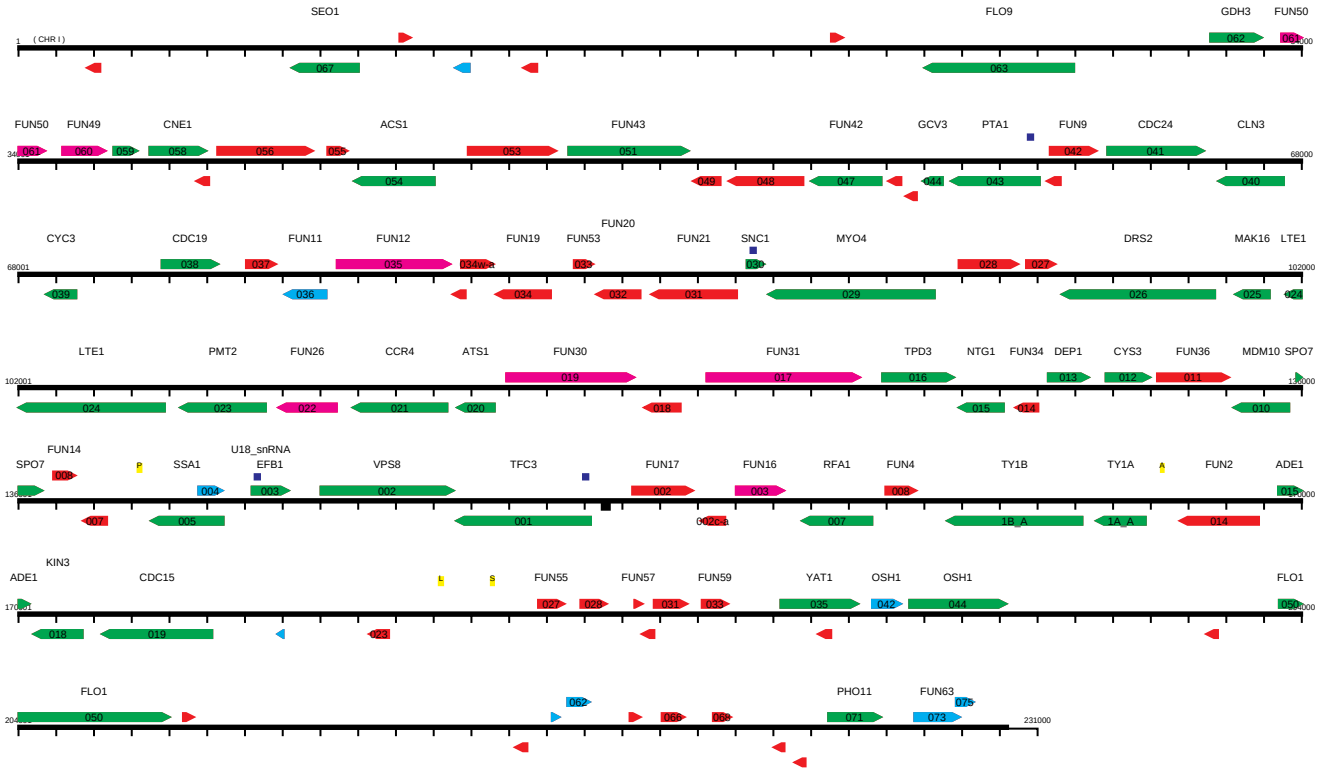
Chromosome 15, The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XV
Dujon, B. *et al. Nature* 387 (suppl.), 98–102 (1997)
EMBL accession numbers: Z74743-Z75302

Chromosome 16, The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XVI
Bussey, H. *et al. Nature* 387 (suppl.), 103–105 (1997)
EMBL accession numbers: Z73521, Z73499-Z73638, U43703, U43503, U43281, U41849, U39205, U44030, U36624, U33335, Z48951, Z49919, Z49274, Z68111, Z49219, U51033, U32445, U40828, U40829, U28371, U25840, U25842, U25841, Z73537, Z73541

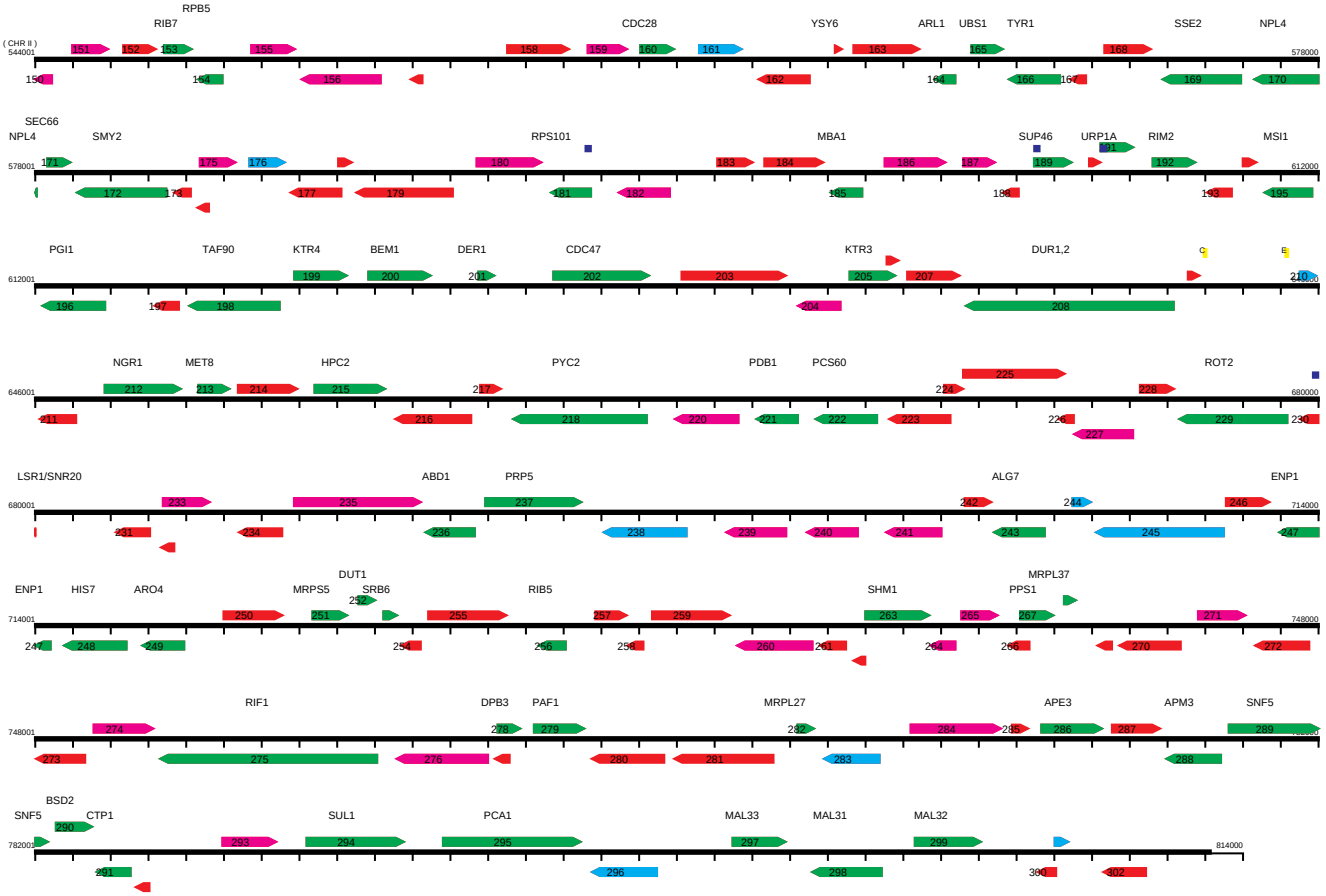
Legend of the chromosome displays

- ORF (Open Reading Frame)
  - proteins of known function
  - similarity to proteins of known function (more than one third of self-score)
  - weak similarity to proteins of known function (lower than one third of self-score)
- Intron
- tRNA
- Centromer

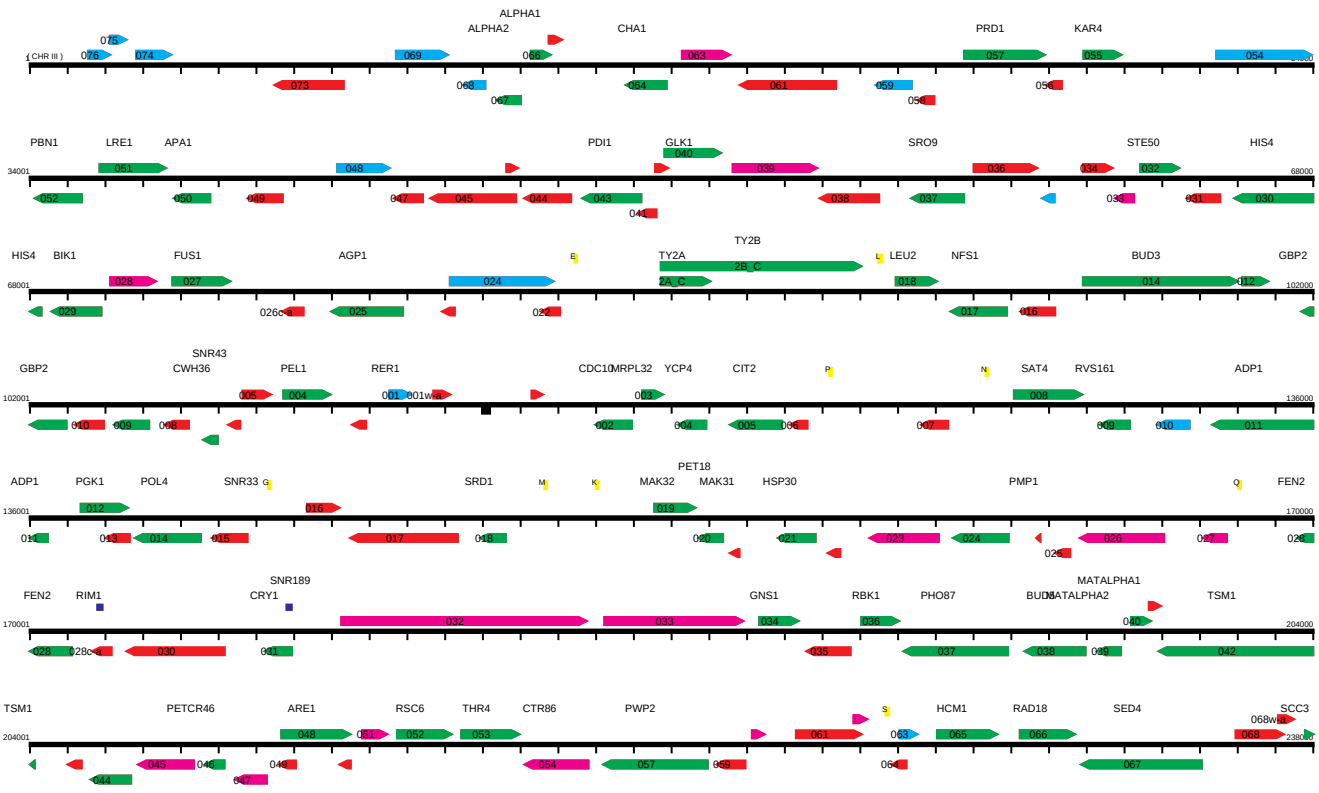
Chromosome 1

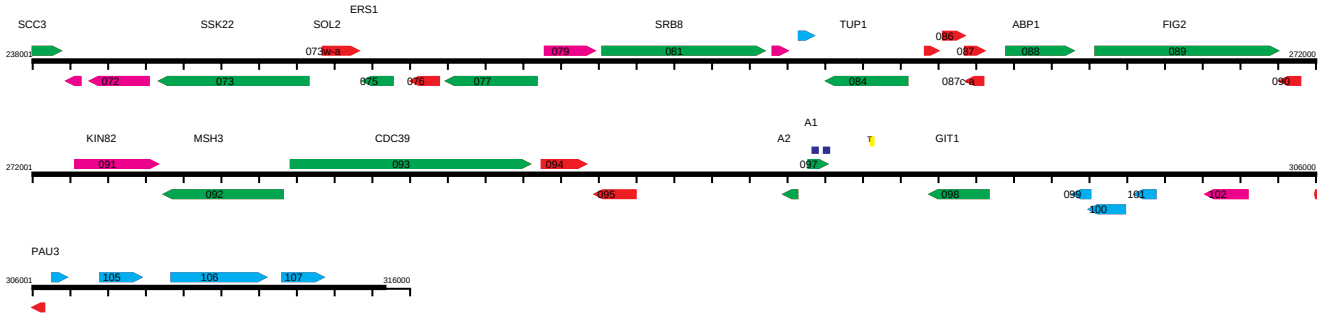


Chromosome 2



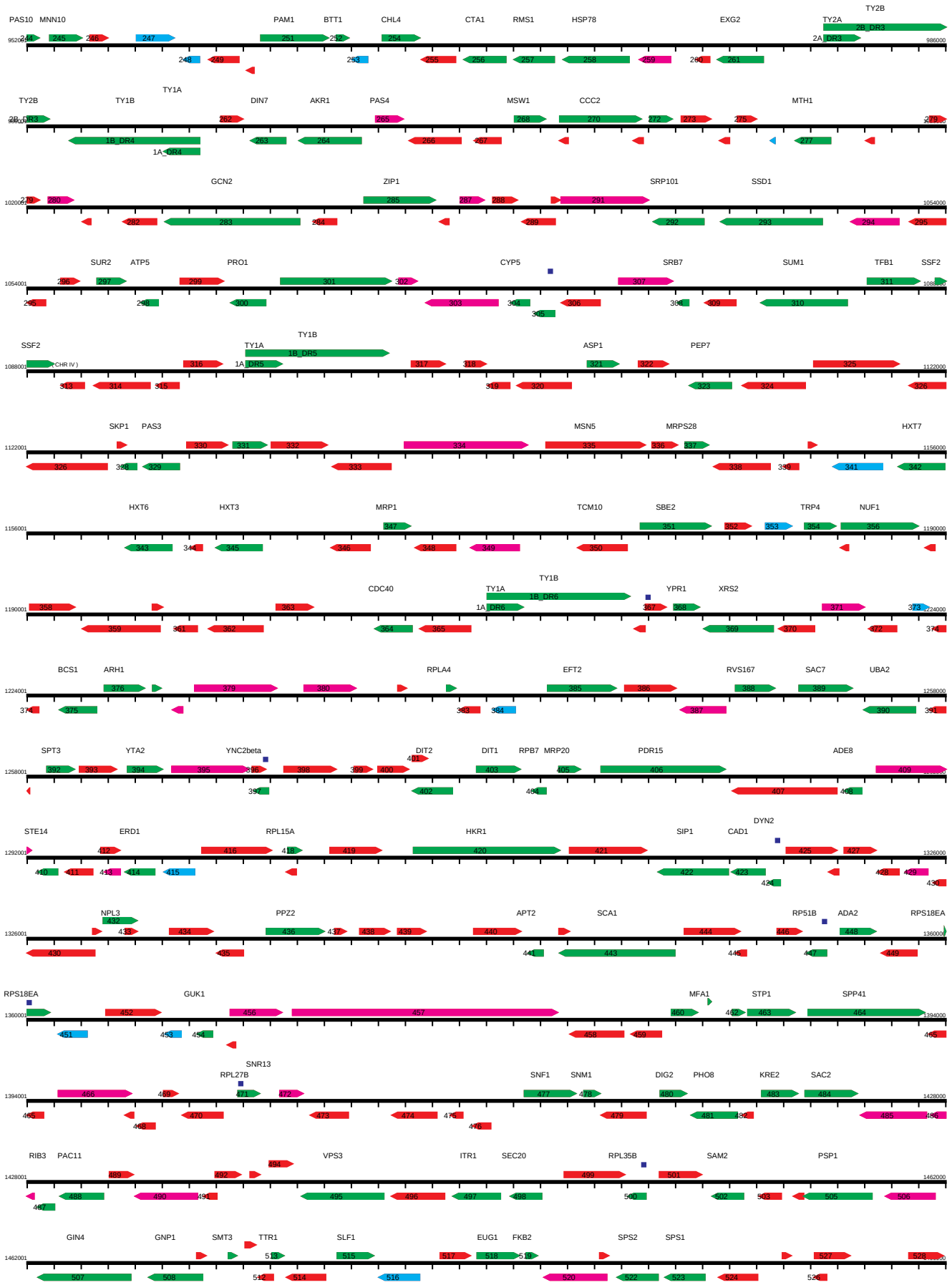
Chromosome 3

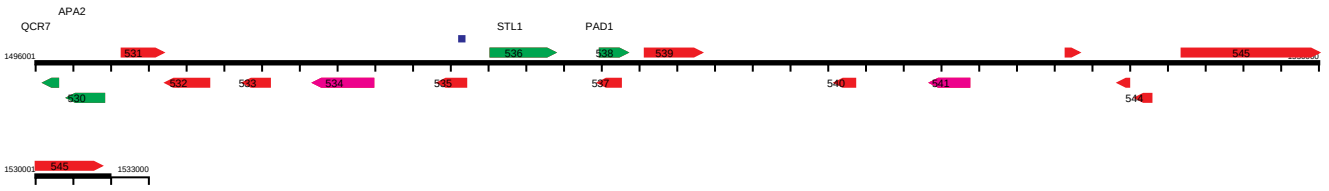




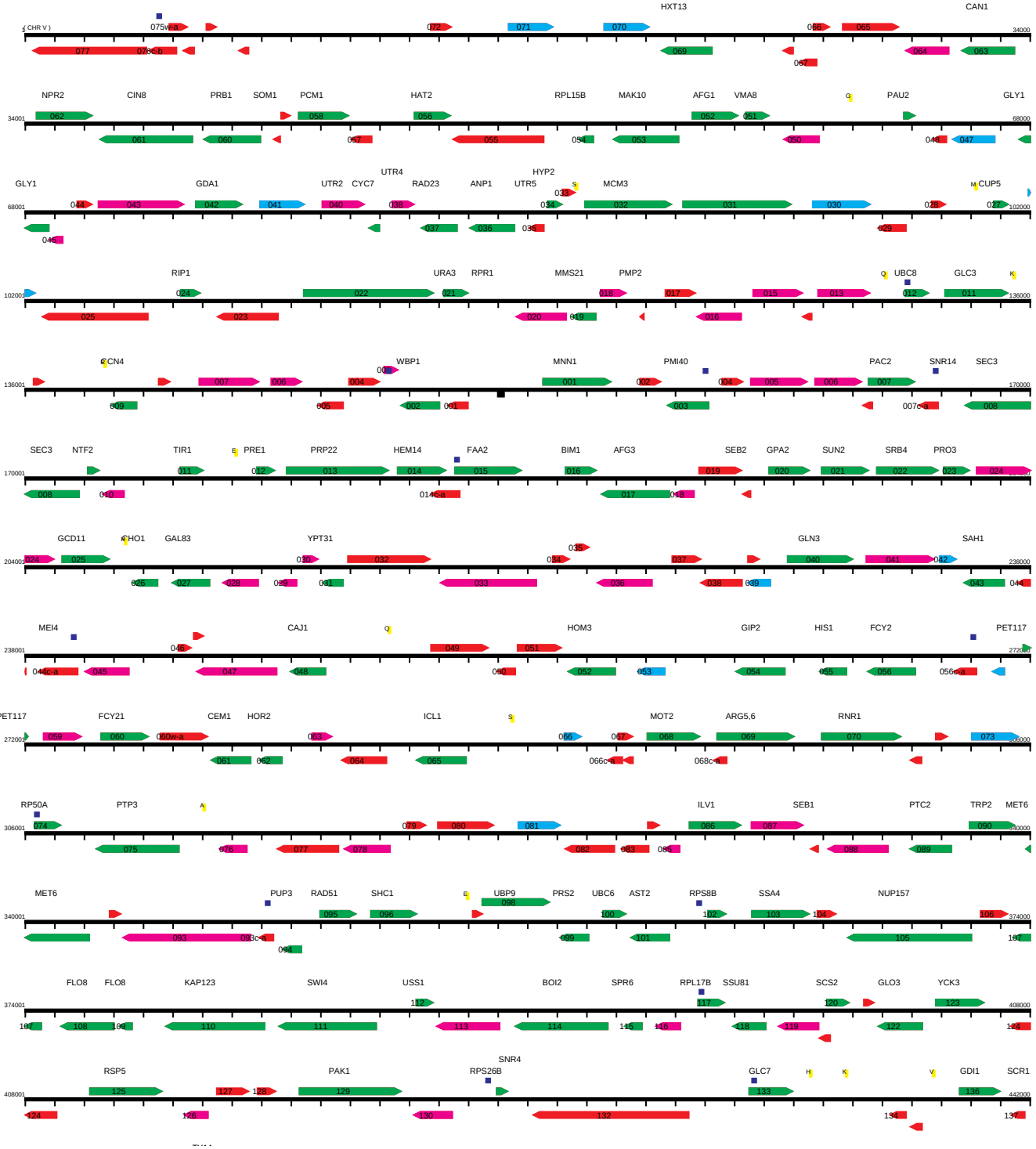
Chromosome 4

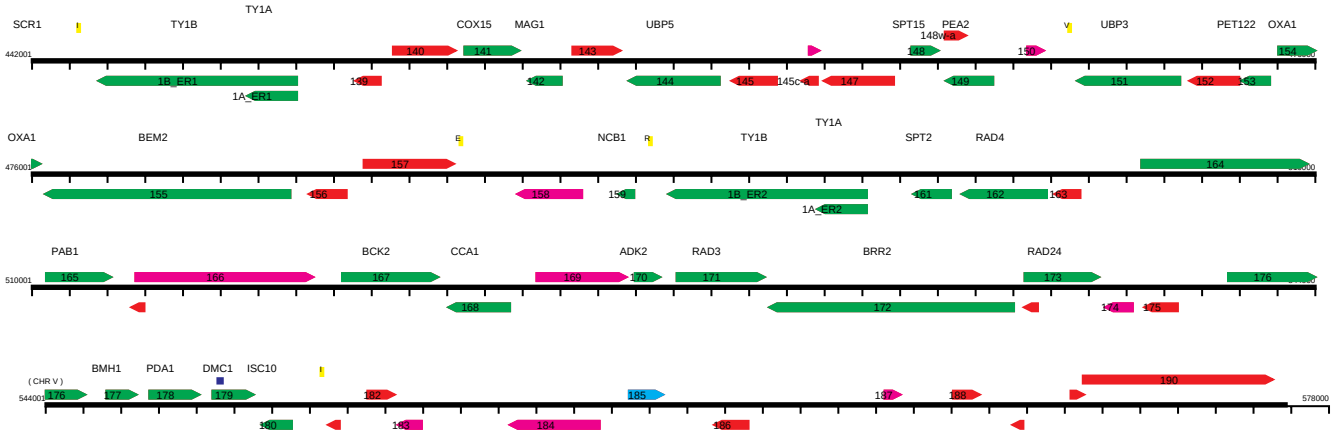

GPD1 GPM2 SON1 CDC7 NOP1 HEX3 YTA5 PTC1
ATP16 RHC21 NHP10 NTH1 YRB1 RAD57 MAF1 SOK1 TRP1 GAL3 SNQ2 RPL2B
KCS1 GCV1 SES1 RPS18A REG1 LYS14
LYS14 TY2B TY2A ARO3 KRS1 ENA5 ENA2 ENA1 SNR47
HEM13 (PAP1) HEM12 TPI1 DBF4 CDC34 TGL2 UBC5 LCB2
LCB2 YS15 DOS1 DOA4 SNF11 TPS2 PPH3 RAD55 SED1 PET100 PDC2
AFR1 SSS1 SLU7 UBC13 GIS1 MSH6 TY1B
TY1B TY1A BMH2 STE5 GSG1
PDS1 TRM1 KIN1 INO2 ARO1 SAC6
SAC6 YCF1 RGP1 HPR1 PAS7 SAN1 MKC7
SWI5 KGD2 NUM1 CTH1 CPH1 RPA14 HOM2 SAC3 NBP2
SEC1 SEC5 TAF25 CDC37 STB3 SEC7 TY1A TY1B HSP42 SUP35 ARG82 NGG1
UBC1 NGG1 SDH4 CDC1 BAT1 CCT6 SLY1 HST4 NUP42 MSS116 REF2
REF2 CBS2 UME6 MSS4 TY2A TY2B TY1B TY1A
GCD6 CCT1 ADR1 RAD9 SPR28 HTB1 HTA1 ADK1 SIR4
SIR4 HEM1 LYS4 SEC26 AMD2 PRP28 PAS10


PAS10 MNN10 PAM1 BTT1 CHL4 CTA1 RMS1 HSP78 EXG2 TY2A TY2B
TY1A TY2B TY1B DIN7 AKR1 PAS4 MSW1 CCC2 MTH1
GCN2 ZIP1 SRP101 SSD1
SUR2 ATP5 PRO1 CYP5 SRB7 SUM1 TFB1 SSF2
TY1B SSF2 TY1A ASP1 PEP7
SKP1 PAS3 MSN5 MRPS28 HXT7
HXT6 HXT3 MRP1 TCM10 SBE2 TRP4 NUF1
TY1B CDC40 TY1A YPR1 XRS2
BCS1 ARH1 RPLA4 EFT2 RVS167 SAC7 UBA2
SPT3 YTA2 YNC2beta DIT2 DIT1 RPB7 MRP20 PDR15 ADE8
DYN2 STE14 ERD1 RPL15A HKR1 SIP1 CAD1
NPL3 PPZ2 APT2 SCA1 RP51B ADA2 RPS18EA
RPS18EA GUK1 MFA1 STP1 SPP41
SNR13 RPL27B SNF1 SNM1 DIG2 PHO8 KRE2 SAC2
RIB3 PAC11 VPS3 ITR1 SEC20 RPL35B SAM2 PSP1
GIN4 GNP1 SMT3 TTR1 SLF1 EUG1 FKB2 SPS2 SPS1

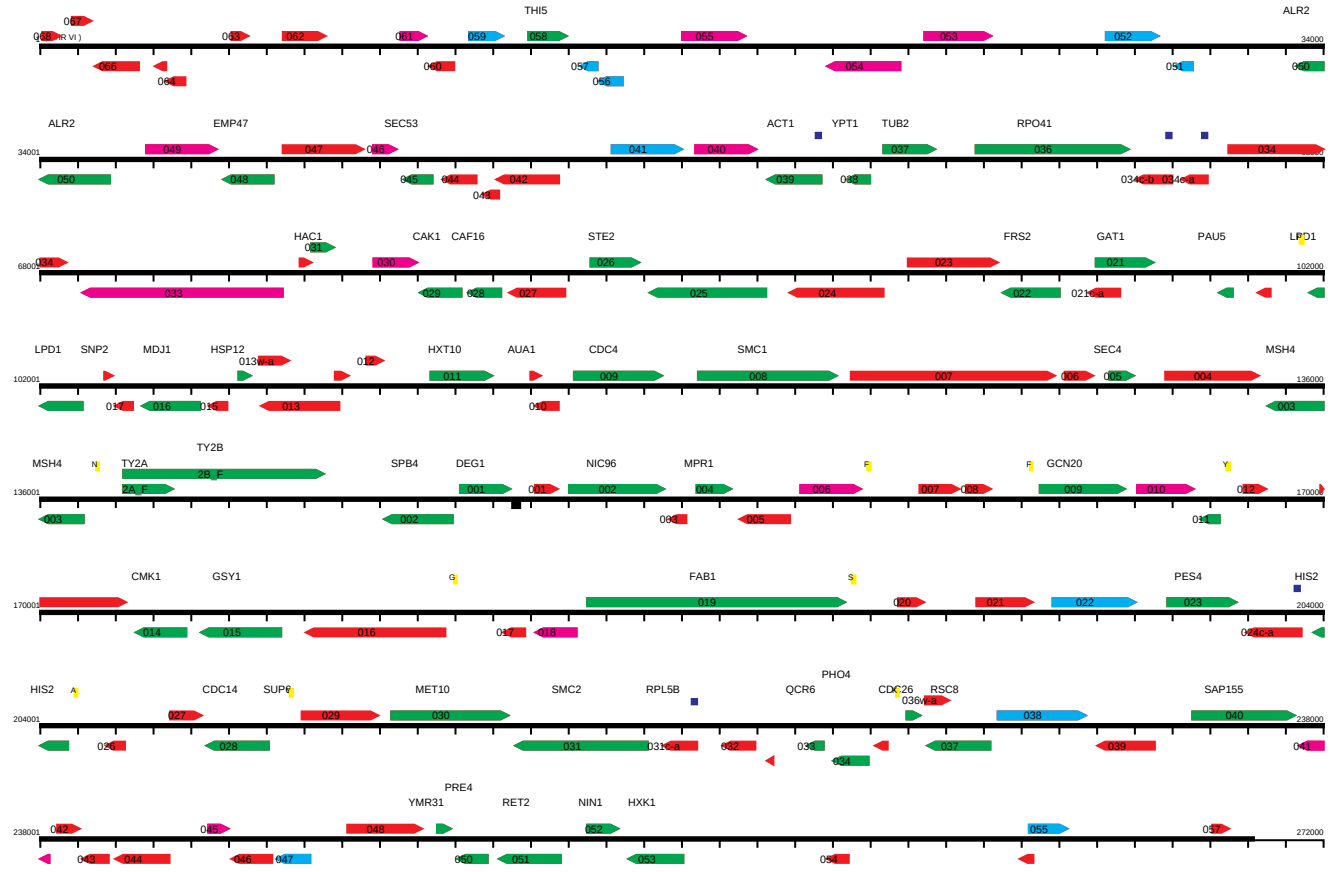


Chromosome 5

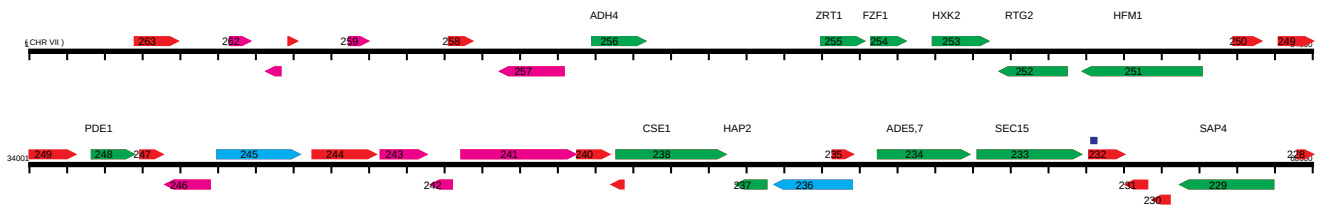


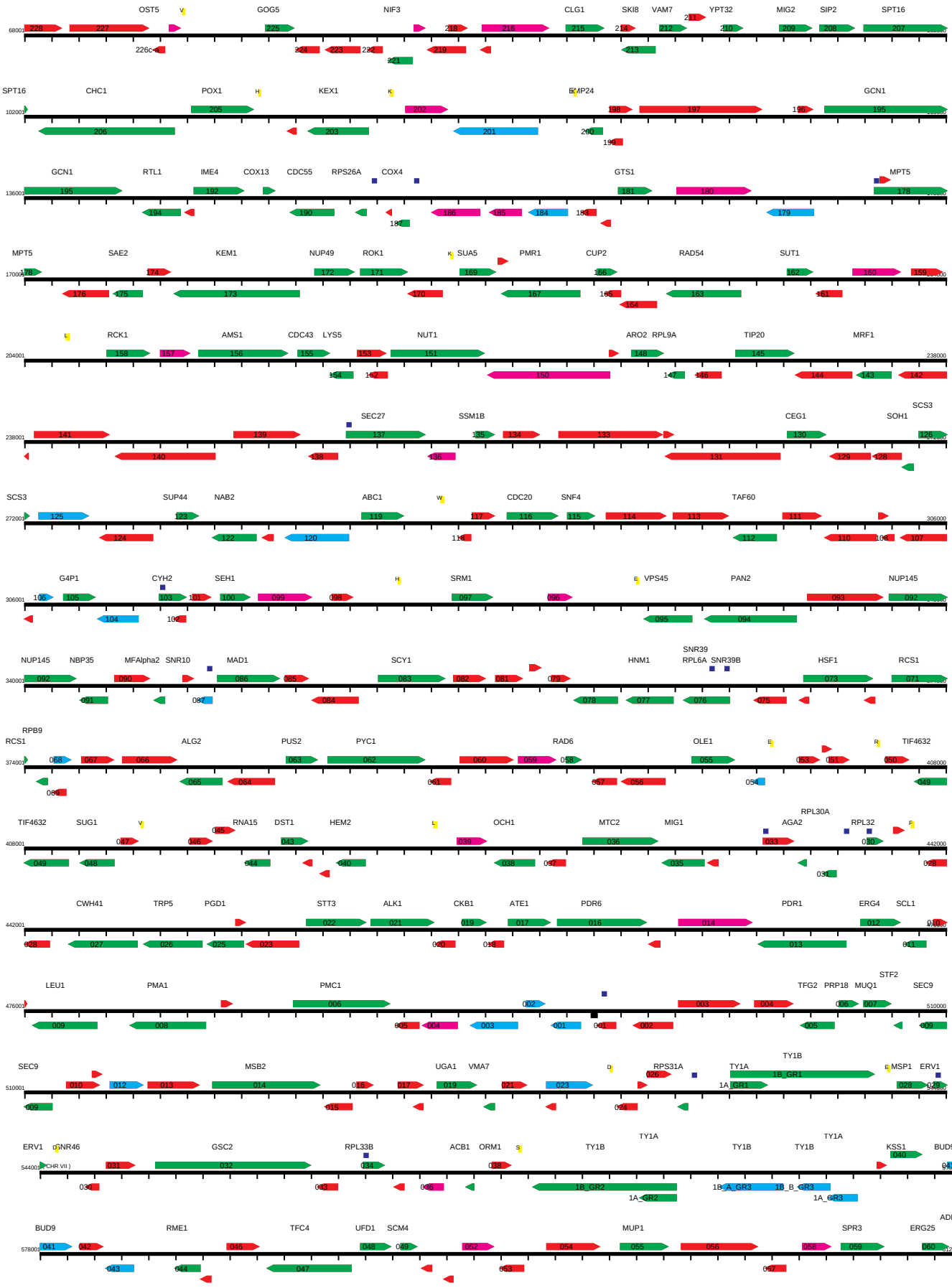


Chromosome 6



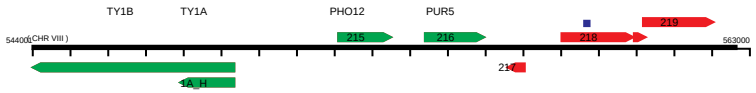
Chromosome 7




OST5
GOG5
NIF3
CLG1
SKI8
VAM7
YPT32
MIG2
SIP2
SPT16
CHC1
POX1
KEX1
ERP24
GCN1
RTL1
IME4
COX13
CDC55
RPS26A
COX4
GTS1
MPT5
SAE2
KEM1
NUP49
ROK1
SUA5
PMR1
CUP2
RAD54
SUT1
RCK1
AMS1
CDC43
LYS5
NUT1
ARO2
RPL9A
TIP20
MRF1
SEC27
SSM1B
CEG1
SOH1
SCS3
SUP44
NAB2
ABC1
CDC20
SNF4
TAF60
G4P1
CYH2
SEH1
SRM1
VPS45
PAN2
NUP145
NBP35
MFAlpha2
SNR10
MAD1
SCY1
HNM1
SNR39
RPL6A
SNR39B
HSF1
RCS1
RPB9
ALG2
PUS2
PYC1
RAD6
OLE1
TIF4632
SUG1
RNA15
DST1
HEM2
OCH1
MTC2
MIG1
RPL30A
AGA2
RPL32
CWH41
TRP5
PGD1
STT3
ALK1
CKB1
ATE1
PDR6
PDR1
ERG4
SCL1
LEU1
PMA1
PMC1
STF2
TFG2
PRP18
MUQ1
SEC9
MSB2
UGA1
VMA7
RPS31A
TY1B
TY1A
1B_GR1
1A_GR1
MSP1
ERV1
SNR46
GSC2
RPL33B
ACB1
ORM1
1B_GR2
1A_GR2
1B_A_GR3
1B_B_GR3
1A_GR3
KSS1
BUD9
RME1
TFC4
UFD1
SCM4
MUP1
SPR3
ERG25
ADE6

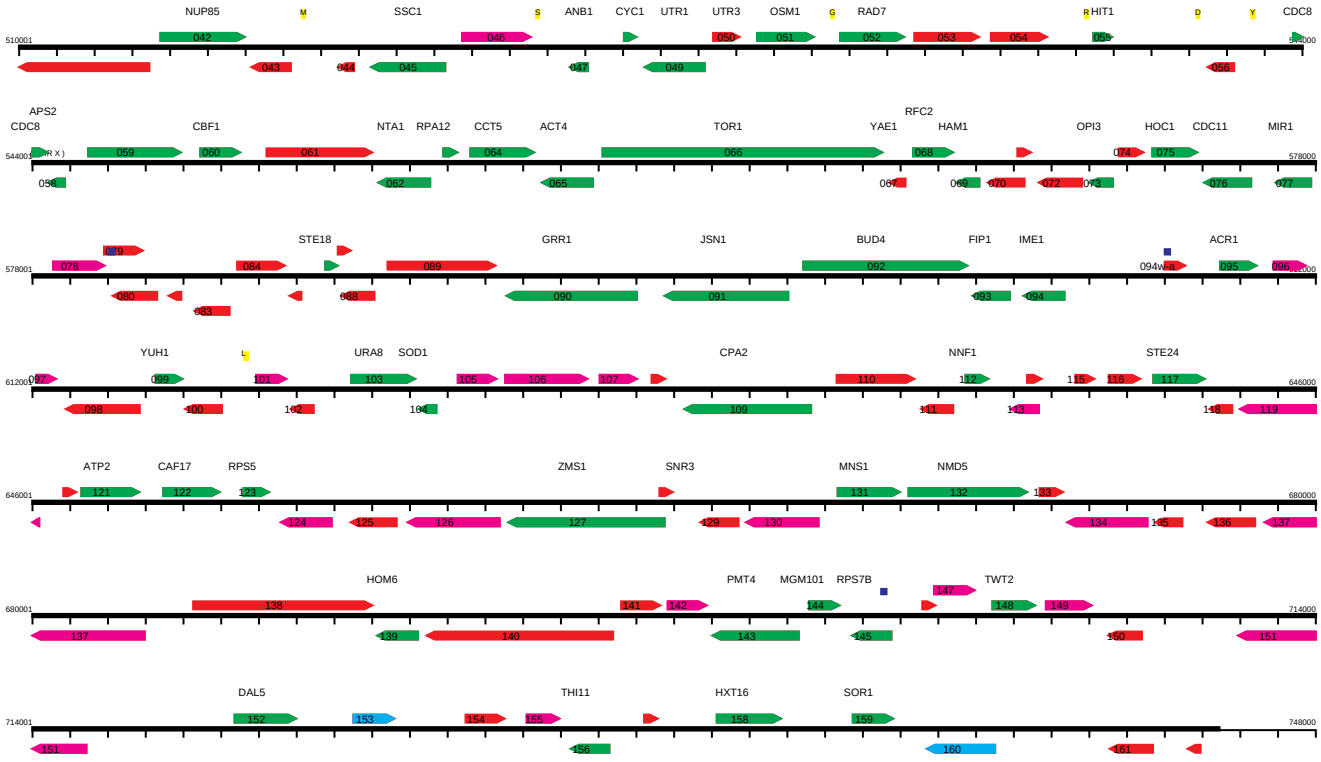

SPT4
ADE6
COX18
ROM1
PRP38
UPF3
SMD1
MRPL25
PAC10
TOM20
GCD2
GCD2
MRP13
RPL16B
PDC6
CTT1
PRP31
DBF2
VAS1
ASK10
ASK10
ESP1
TEL2
MIC1
VMA21
SRB5
CLB1
CLB6
TY3A
TY3B
SPT6
RPS28A
NUP57
MEP1
PPT1
ASN2
PHB1
PAS2
PRE9
CBF2
SKN1
THI4
NAT2RPL30B
RSR1
CYS4
CHO2
MTR3
NSR1
TY2A
TY2B
TY2B
TY1B
TY1A
TIF4631
KRE11
CLC1
PSD2
MSM1
YIP1
GSP4
ERG1
ATF2
ATF2
RNR4
QCR9
UBR1
TYS1
TFG1
HGH1
BUB1
HIP1
TDH3
PDX1
SNG1
PMT6
PCT1
ADE3
ETF-BETA
SER2
TRX2
RTA1
NAB1A
GPI1
CCH1
CRM1
MRPL9
SNR7-S
PET54
DIE2
KNR4
PHO81
YHB1
PFK1
BRF1
SOL4
SOL4
MGA1
GCN5
PUP2
ENO1
GND2
RAD2
YKS5
MES1
MES1
FOL2
YTA7
TAF145
YOR1
YOR1
BGL2
ZUO1
BIO2
MAL13
AGT1
MAL12
(CHR VII)

Chromosome 8

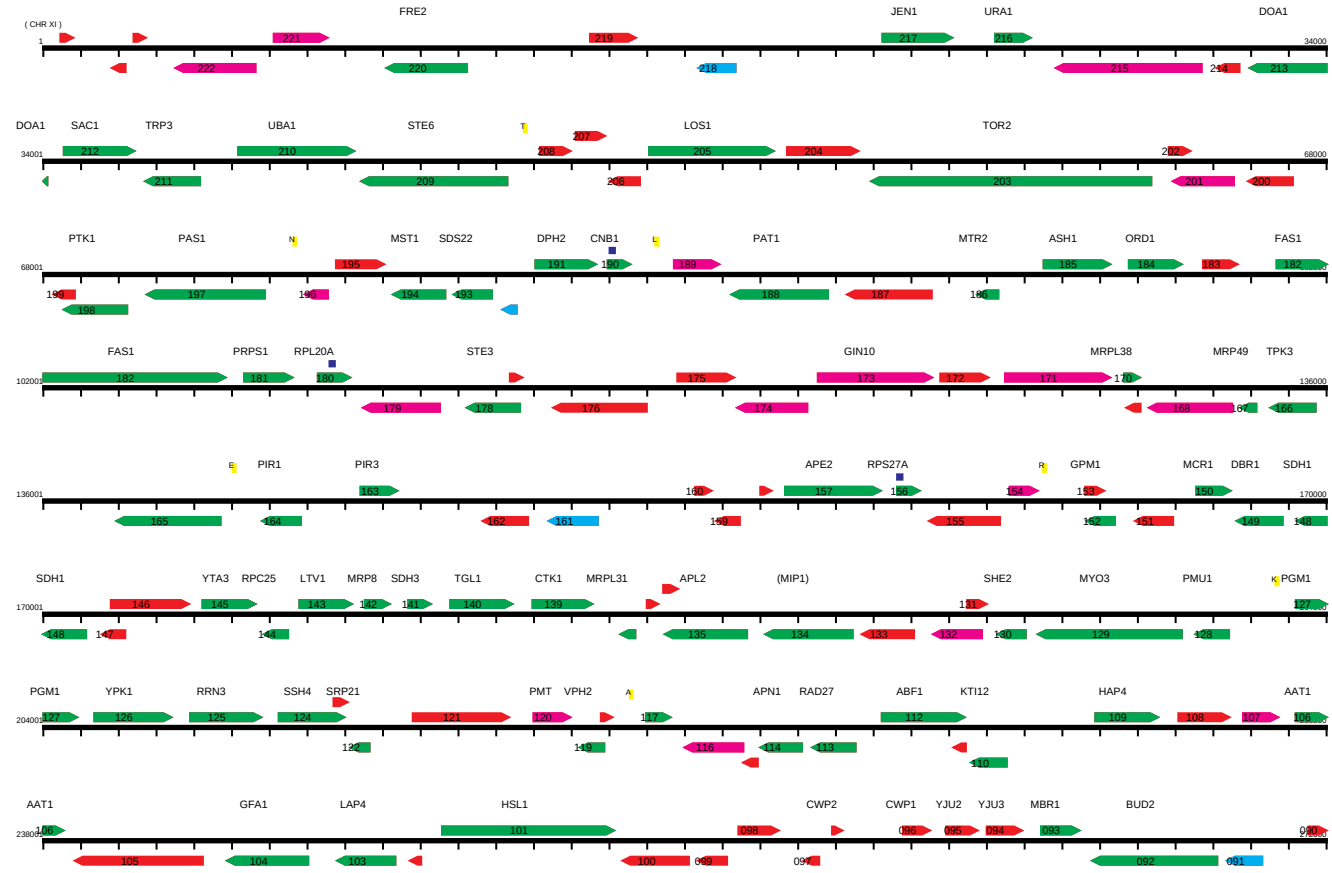


Chromosome 9

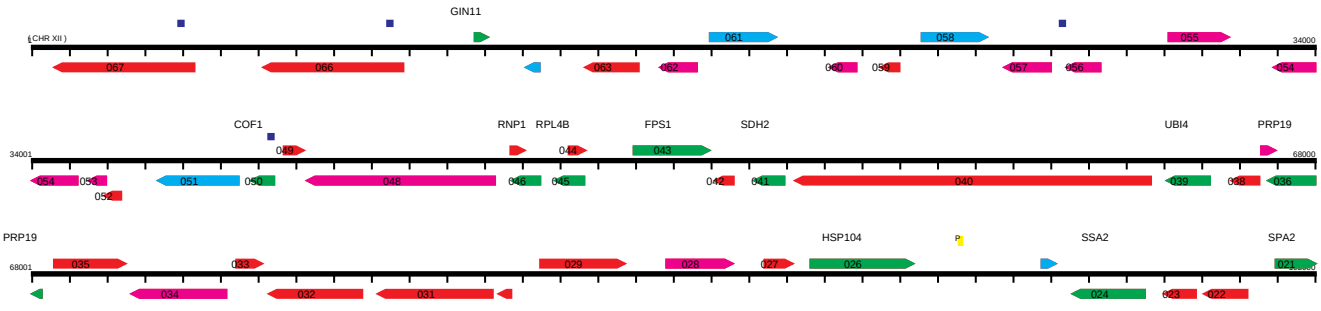
Chromosome 10

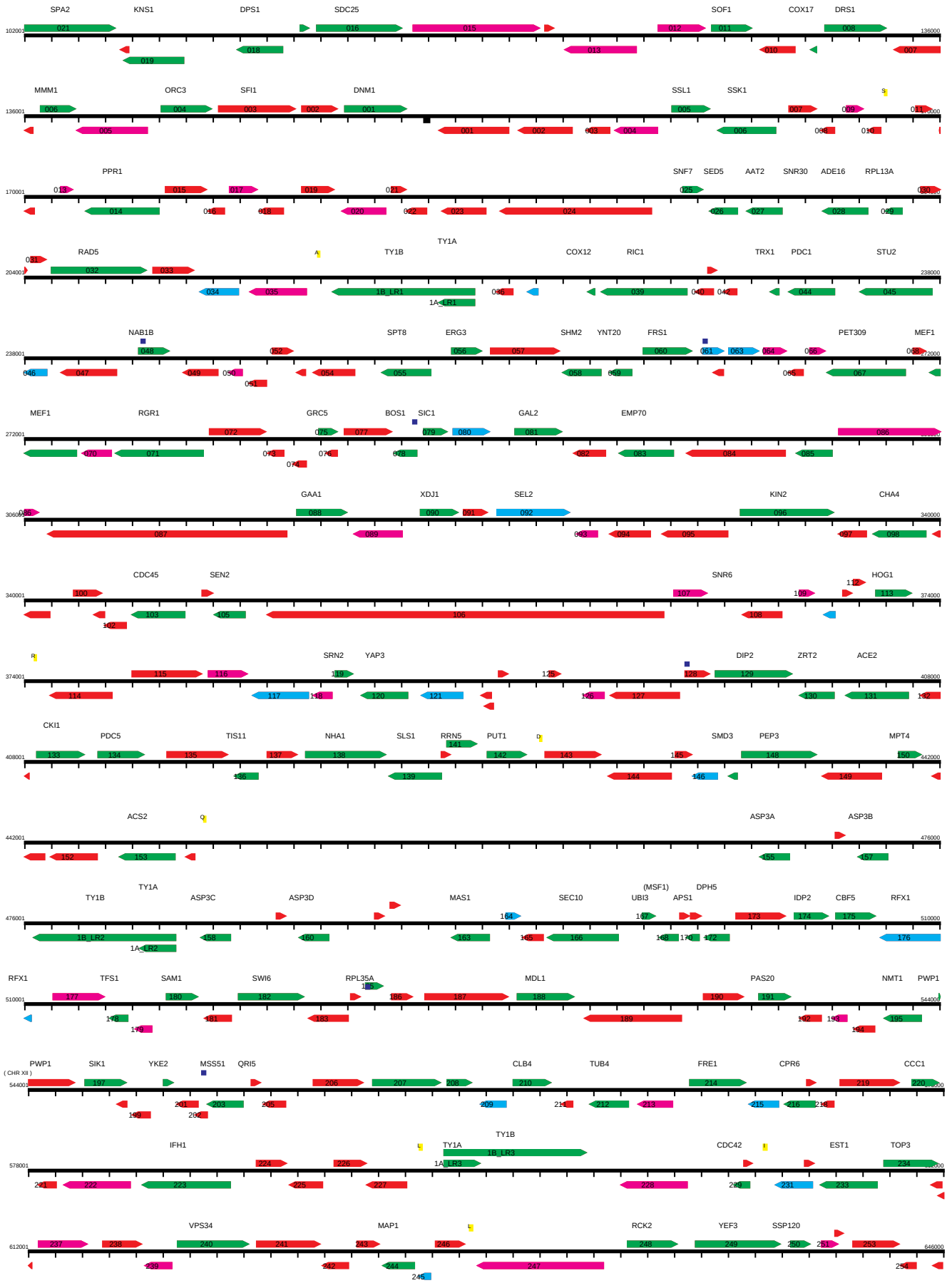


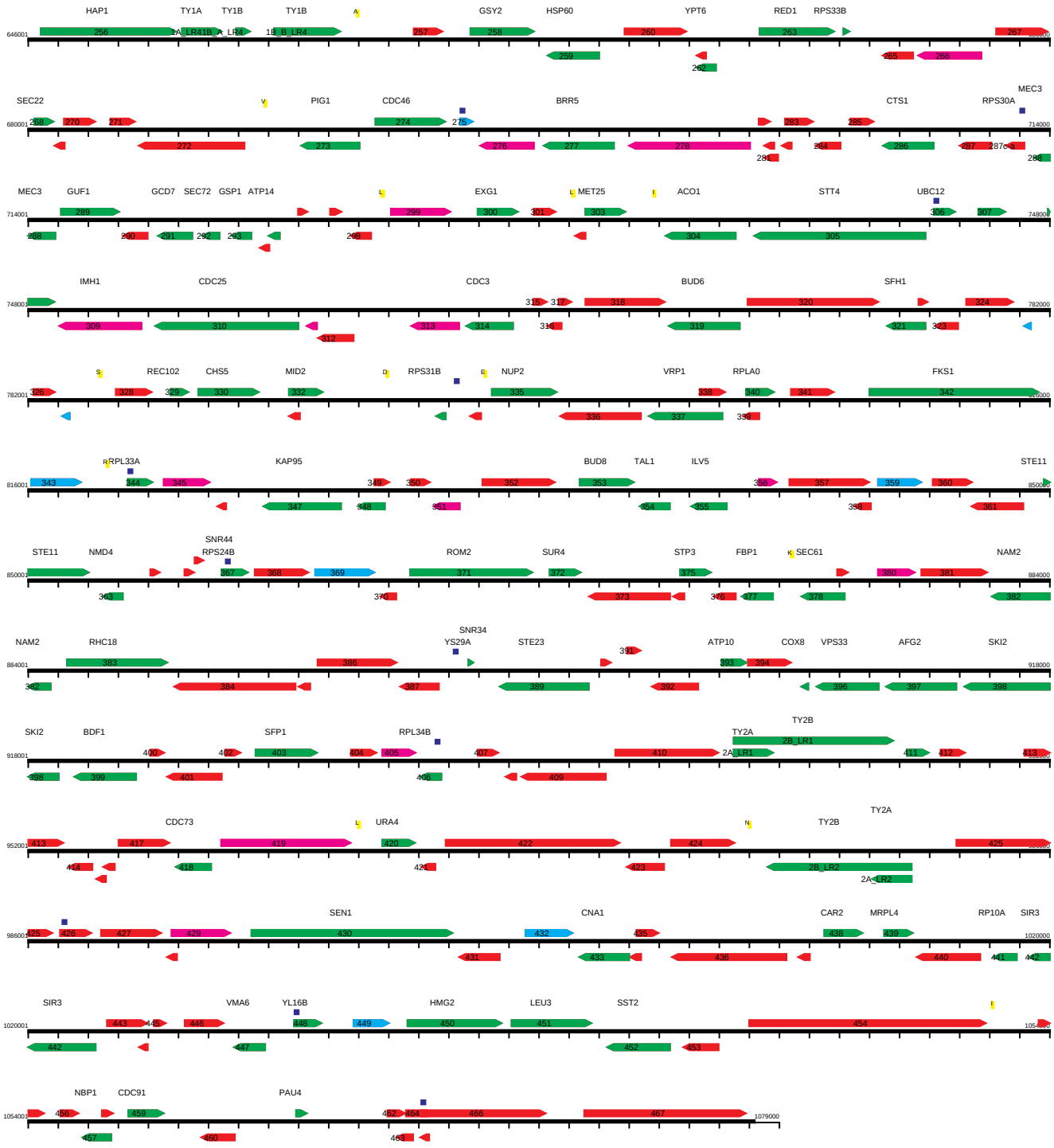
Chromosome 11



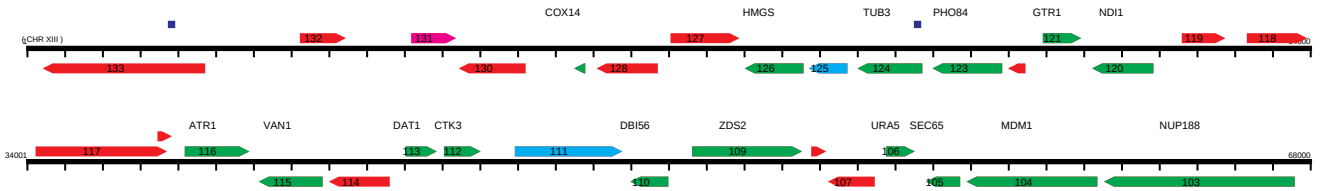
Chromosome 12




SPA2
KNS1
DPS1
SDC25
SOF1
COX17
DRS1
MMM1
ORC3
SFI1
DNM1
SSL1
SSK1
PPR1
SNF7
SED5
AAT2
SNR30
ADE16
RPL13A
TY1A
RAD5
TY1B
COX12
RIC1
TRX1
PDC1
STU2
NAB1B
SPT8
ERG3
SHM2
YNT20
FRS1
PET309
MEF1
MEF1
RGR1
GRC5
BOS1
SIC1
GAL2
EMP70
GAA1
XDJ1
SEL2
KIN2
CHA4
CDC45
SEN2
SNR6
HOG1
SRN2
YAP3
DIP2
ZRT2
ACE2
CKI1
PDC5
TIS11
NHA1
SLS1
RRN5
PUT1
SMD3
PEP3
MPT4
ACS2
ASP3A
ASP3B
TY1A
TY1B
ASP3C
ASP3D
MAS1
SEC10
(MSF1)
UBI3
DPH5
APS1
IDP2
CBF5
RFX1
1B_LR2
1A_LR2
RFX1
TFS1
SAM1
SWI6
RPL35A
MDL1
PAS20
NMT1
PWP1
PWP1
SIK1
YKE2
MSS51
QRI5
CLB4
TUB4
FRE1
CPR6
CCC1
(CHR XII)
TY1B
IFH1
TY1A
1B_LR3
1A_LR3
CDC42
EST1
TOP3
VPS34
MAP1
RCK2
YEF3
SSP120

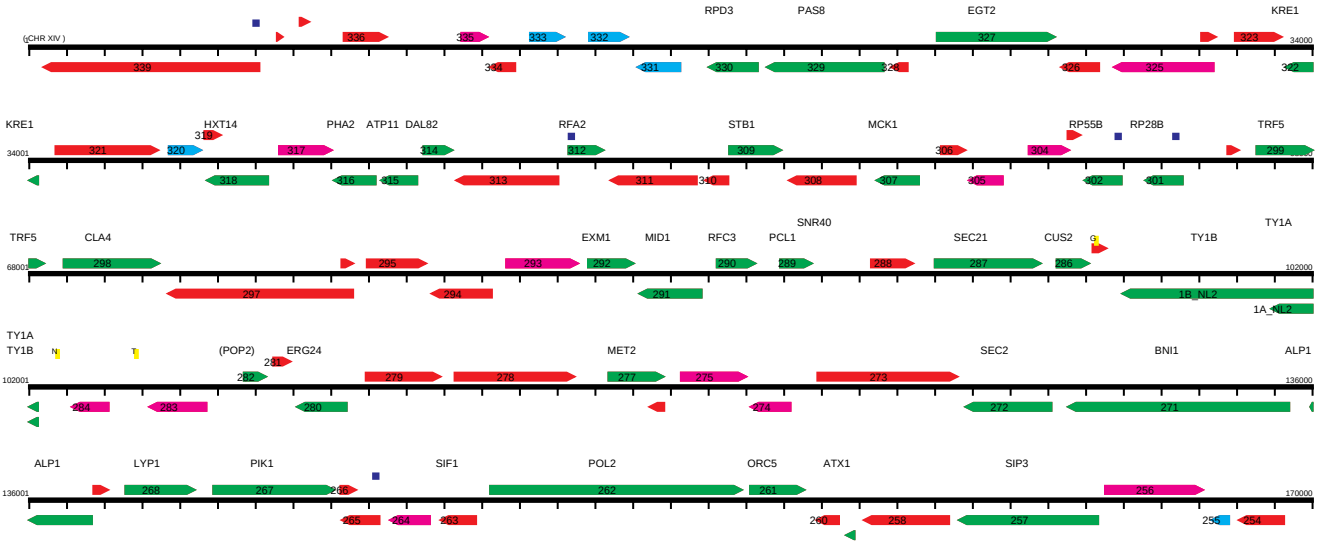


Chromosome 13

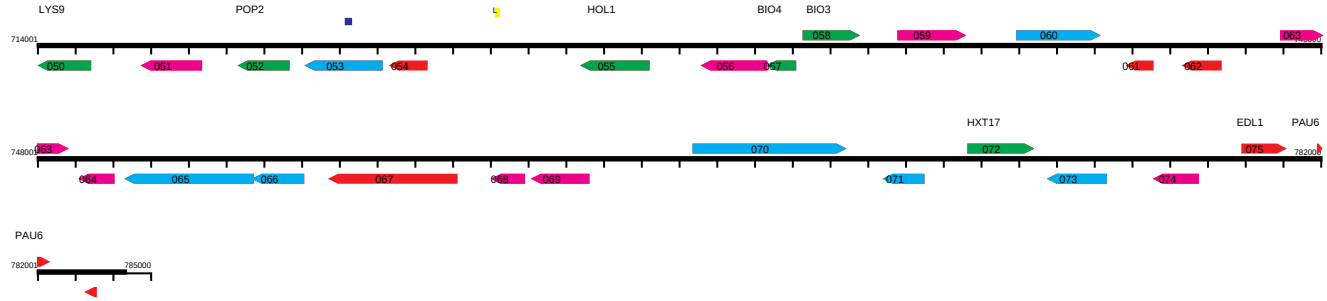



CAC2
TSL1
ARG81
FUN81
VPS9
RAD10
PRE8
RPM2
ALO
TUB1
CPR3
HMG1
NPI46
YL16A
POB3
ORC1
TEM1
RP10B
MFT1
PIF1
OGG1
CMP2
CYB2
GAL80
EFF2
PRP39
TY1A
1B_ML2
RRN11
CAT2
1A_ML2
1A_ML1
TY1B
AMD1
RAD52
NDC1
TSA1
YOX1
RP51A
APT1
UNG1
PSP2
PPZ1
ERV25
SPT5
MRPL39A
009wb
ERG6
YAP1
GLO1
YPT7
CDC5
MVP1
MPT1
PLB1
HXT2
SEC59
snoRNA
ERG5
SOK2
STB4
FMS1
MAC1
QRI8
MSS1
MRPL3
TAP42
IMP2
MIH1
MSN2
LYS7
SUB1
ARG80
MCM1
NUP116
1B_MR1
1A_MR1
1B_MR2
1A_MR2
TYR1
STB2
STV1
BUB2
AAC1
FET3
TOM37
RNA14
AEP1
HMS1
ABF2
PDS5
CHL12
SEC14
NAM7
ISF1
ADH3
YTA12
NPL6
AIP1
CTF13
YPK2
PGM2
HDF2
ILV2
MYO5
U24_snoRNA
RPL13B
PKR1
GCR3
SAS2
POM152
REC114
PSO2
CIN4
MDS1
RPS16A
TIF34
SWP1
IMP1
HLJ1
SMP2
MLH1
CEP3
ALD4
ALD5
DDR48
SIP18
PAI3

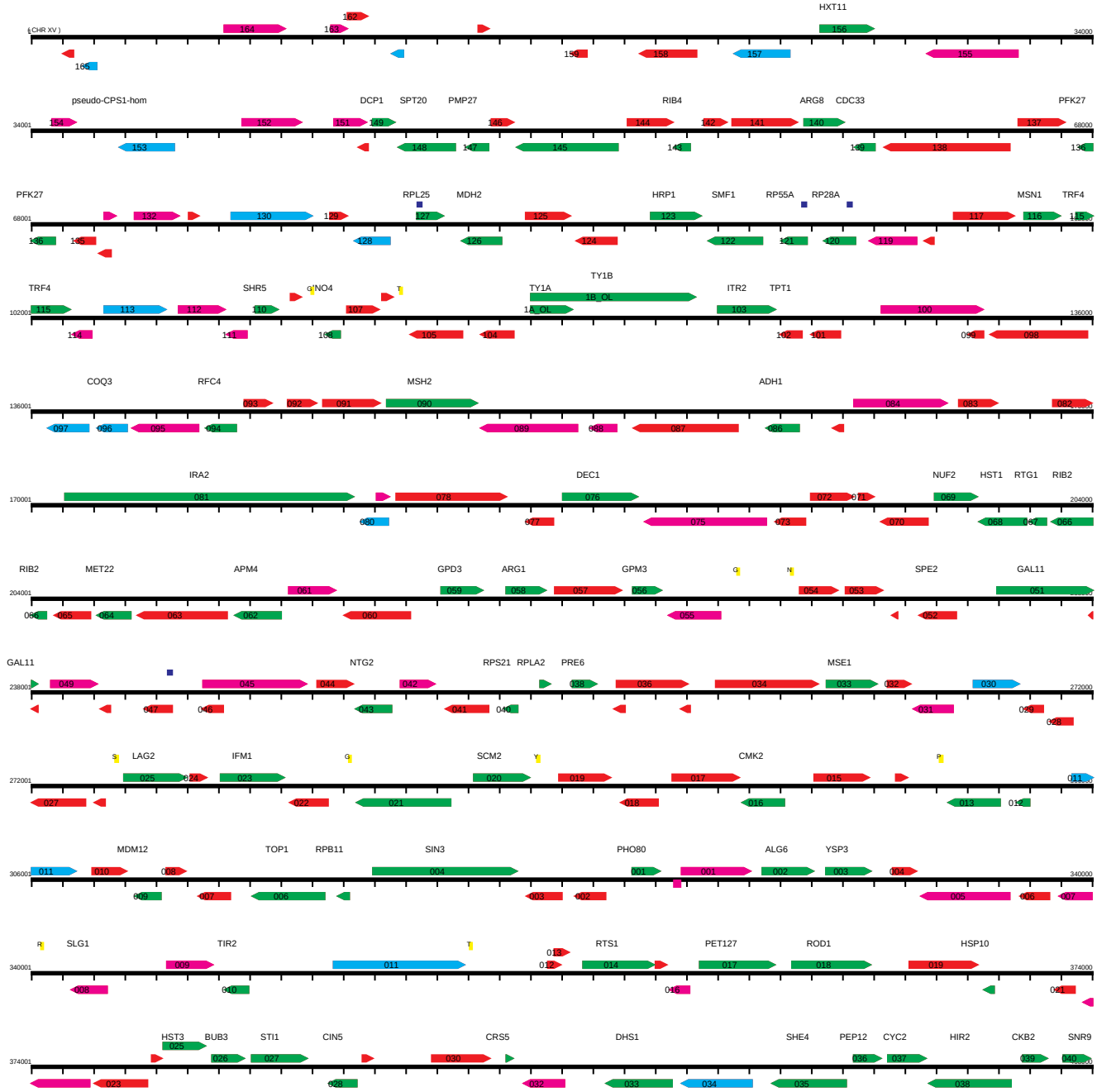
Chromosome 14



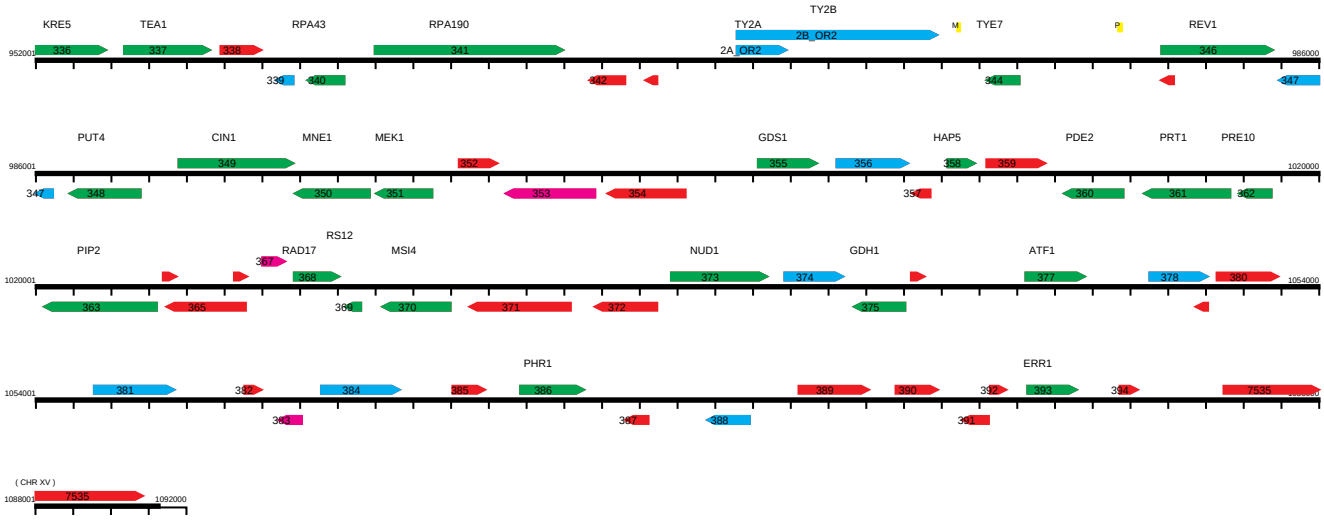

NRD1
RAD50
MPA43
RPA49
SUI1
SLA2
ZWF1
LAP3
KEX2
KEX2
YTP1
SIN4
URE2
SNR19
SSU72
POP1
ADE12
ALG9
RAP1
MER1
SSB2
SPS18
SPS19
GCR2
WHI3
CHS1
SRP1
KAR1
NPR1
RPS3
MDG1
APC1
PSD1
SKO1
RPL41A
YGP1
YCK2
RPC31
MFA2
MEP2
RLR1
SRV2
NAM9
FPR1
TOM22
CPT1
SPC98
TOM70
PSU1
MLS1
MLS1
RPC19
DBP2
CYB5
PIE3
LEU4
MET4
CDC17
RAS2
YPT53
RHO2
TOP2
MKT1
END3
PMS1
PMS1
TPM1
MKS1
YMK1
MSK1
LAT1
TOM7
RP23
FKH2
RPL9B
SUN4
YDJ1
GCD10
GCD10
NOP2
ARP5
POR1
TY1A
TY1B
1B_NL1
1A_NL1
MSG5
COX5A
ALG11
YIP3
(CHR XIV)
TFC5
IDH1
NCE3
TY2B
TY2A
2B_N
2A_N
HHT2
HHF2
KTR5
KTR5
SSN8
NME1
PUB1
PBI2
SPO1
SIS1
MRPL2
HRB1
PET8
RLP7
DOM34
CIT1
FUN34
RPC34
VPS27
CSE2
PRP2
PRP2
URK1
SMM1
ACC1
MAS6
ARE2
SNF12
SEC12
SSK2
PPG1
ABZ1
SOL1
COQ2
MVD1
AGA1
PET494
MSO1



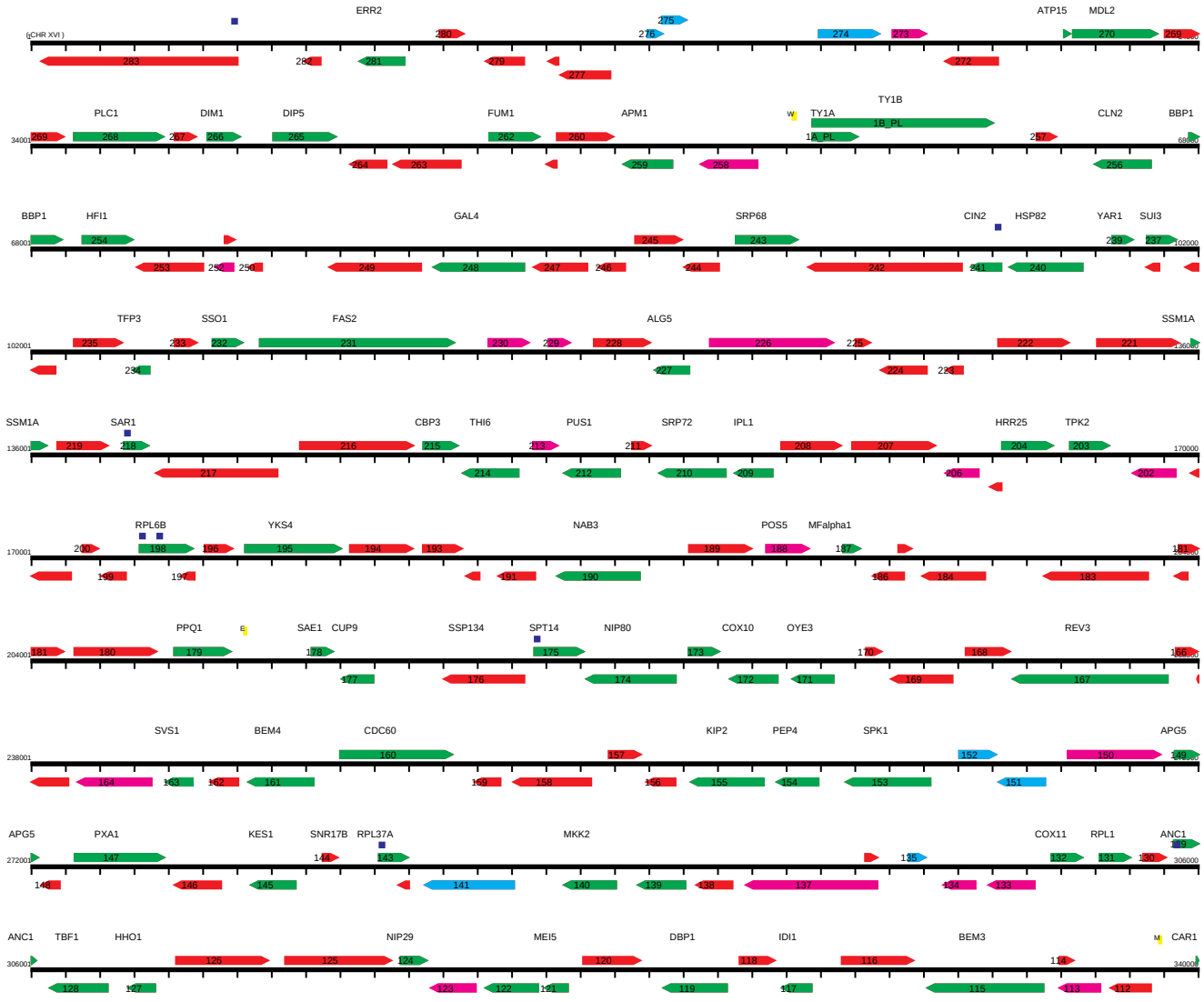
Chromosome 15



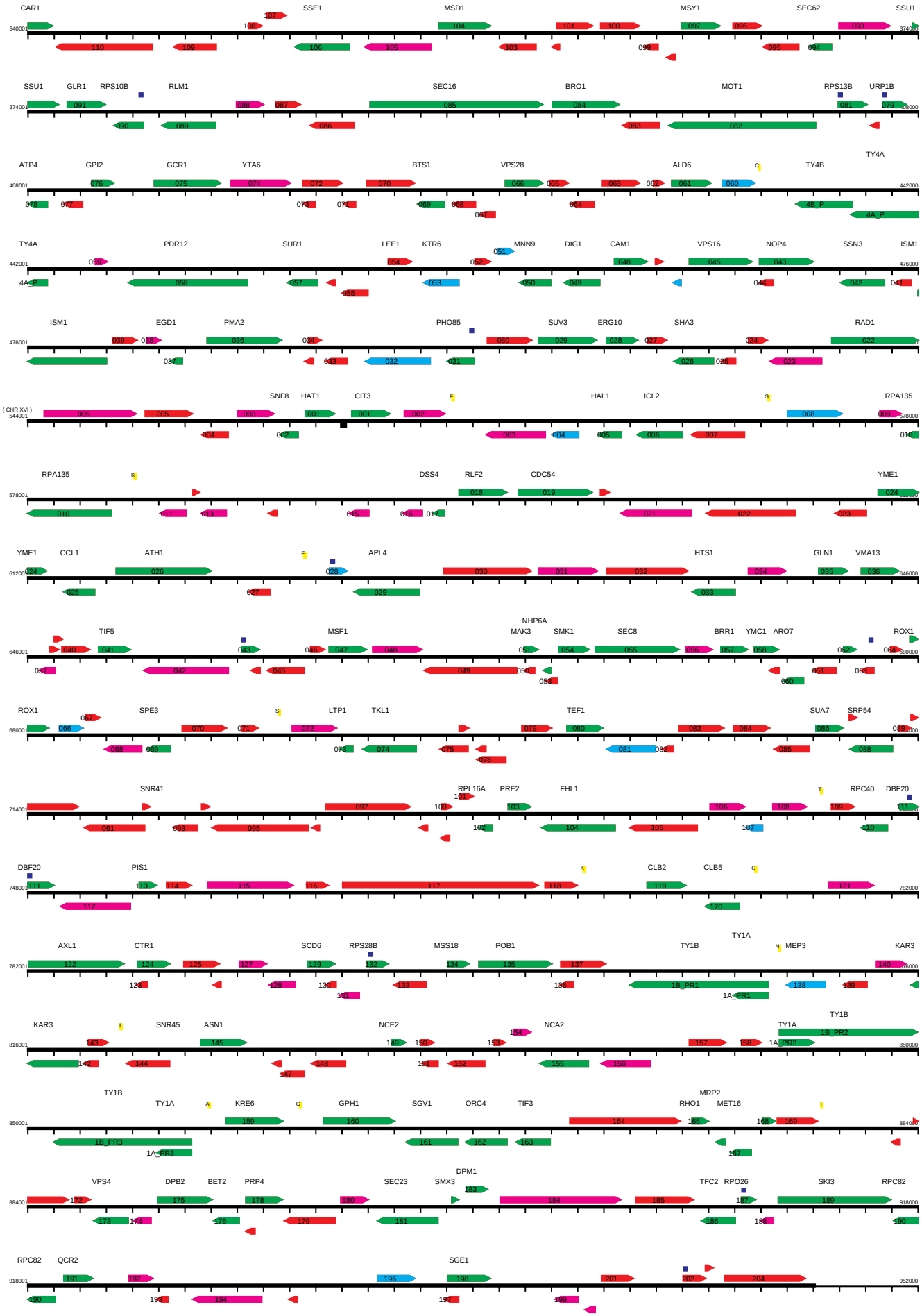

SNR9 WHI2 TOM6 DBP5 STD1 RAT1 ASE1 CKA2
CKA2 TCM1 CYT1 ALG8 CDC21 UFE1 RTS2 ATX2
OST3 VPS21 ARF3 RP30 NUP1
NUP1 KTR1 RAS1 OST2 VAM3 AZF1 RPO31
RPO31 YTA1 RIO1 GCY1 PFY1 LEO1 UBP2 CAT5 EST2 RGA1 ADE2 ARG11 VPS17 EFT1
(CHR XV)
BAG7 IDH2 SFL1 TY1B TY1A 1B_OR 1A_OR THI80 SPP2 SMP3
RPB2 PDR5 NFI1 PUP1 PET123 SME1 MTR10
RPS33A GLN4 HEM15 GAC1 LAS17 RPS30B SER1
SNR36 SER1 GSP2 TUF1 MSB1 SPR1 TY2B TY2A TOA1
2B_OR1
2A_OR1
LIP5 BFR1 PET56 HIS3 DED1 RPC128 PTP2 NPT1 RPB10 MGM1 STE4
RFC1 STE13 SNR35 RPB8 MKK1 MGE1 KIN4 RPL37B SNR17A DFR1 HES1
HES1 SEC63 CDC31 CRL13 GCD1
RBL2 PNT1 PAC1 VPH1 SNR8 MOD5 CAF20 SNR5 SNR31 HEM4
MPD1 SNF2 CPA1
CPA1 BAT2 SLY41 RPL18B SPS4 COT1 FAA1 HSH49 PMT3
PMT3 PRO2 MYO2 SNC2 PDR10 SCD5 MIP1 VMA4 MRS2 ALA1 KRE5



Chromosome 16

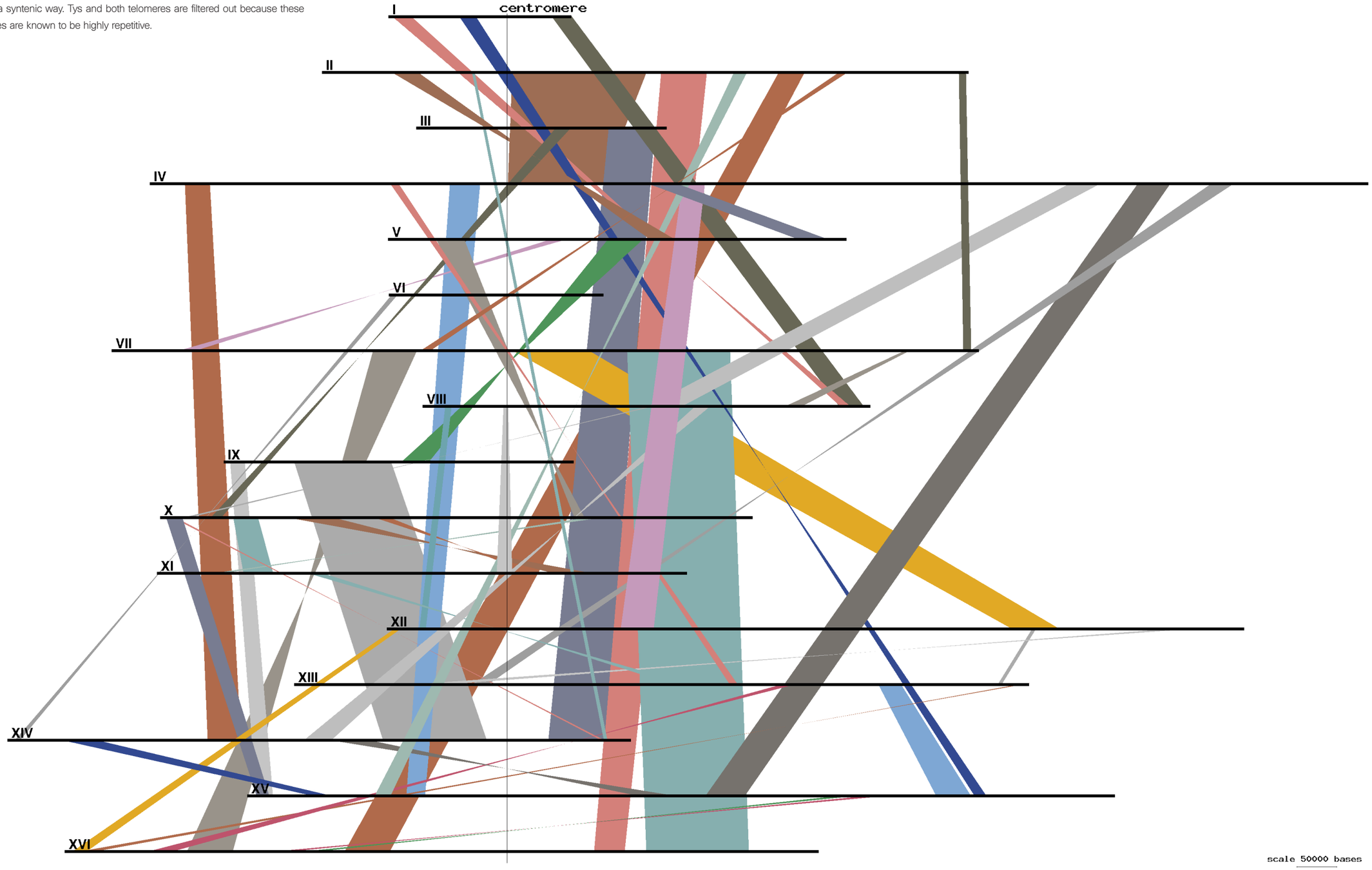


maps


CAR1
SSE1
MSD1
MSY1
SEC62
SSU1
SSU1
GLR1
RPS10B
RLM1
SEC16
BRO1
MOT1
RPS13B
URP1B
ATP4
GPI2
GCR1
YTA6
BTS1
VPS28
ALD6
TY4B
TY4A
TY4A
PDR12
SUR1
LEE1
KTR6
MNN9
DIG1
CAM1
VPS16
NOP4
SSN3
ISM1
ISM1
EGD1
PMA2
PHO85
SUV3
ERG10
SHA3
RAD1
(CHR XVI)
SNF8
HAT1
CIT3
HAL1
ICL2
RPA135
RPA135
DSS4
RLF2
CDC54
YME1
YME1
CCL1
ATH1
APL4
HTS1
GLN1
VMA13
TIF5
MSF1
NHP6A
MAK3
SMK1
SEC8
BRR1
YMC1
ARO7
ROX1
ROX1
SPE3
LTP1
TKL1
TEF1
SUA7
SRP54
SNR41
RPL16A
PRE2
FHL1
RPC40
DBF20
DBF20
PIS1
CLB2
CLB5
TY1A
AXL1
CTR1
SCD6
RPS28B
MSS18
POB1
TY1B
MEP3
KAR3
KAR3
SNR45
ASN1
NCE2
NCA2
TY1B
TY1A
TY1B
TY1A
KRE6
GPH1
SGV1
ORC4
TIF3
MRP2
RHO1
MET16
DPM1
VPS4
DPB2
BET2
PRP4
SEC23
SMX3
TFC2
RPO26
SKI3
RPC82
RPC82
QCR2
SGE1

duplication chart

View of 53 clustered gene duplications between the 16 chromosomes of yeast. Each chromosome is represented by a bold horizontal line. All chromosomes are aligned according to the centromere, shown as a vertical line. Gene clusters are represented by coloured polygons. A cluster is detected if, within a window of 25 kb, at least 5 blocks of coding DNA (block size 500 nucleotides) are duplicated in a syntenic way. Tys and both telomeres are filtered out because these sequences are known to be highly repetitive.



Metabolism

amino-acid metabolism

amino-acid biosynthesis

[ORF]	[Gene]	[Encoded or related protein]
YHR037w	*PUT2*	1-pyrroline-5-carboxylate dehydrogenase
YDR035w	*ARO3*	2-dehydro-3-deoxyphosphoheptonate aldolase, phenylalanine-inhibited
YBR249c	*ARO4*	2-dehydro-3-deoxyphosphoheptonate aldolase, tyrosine-inhibited
YNL104c	*LEU4*	2-isopropylmalalate synthase
YGL009c	*LEU1*	3-isopropylmalate dehydratase
YPR167c	*MET16*	3'-phosphoadenylylsulphate reductase
YGR019w	*UGA1*	4-aminobutyrate aminotransferase (GABA transaminase)
YER091c	*MET6*	5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase
YMR108w	*ILV2*	acetolactate synthase
YCL009c	*ILV6*	acetolactate synthase, regulatory subunit
YER069w	*ARG5,6*	acetylglutamate kinase and acetylglutamyl-phosphate reductase
YJL071w	*ARG2*	acetylglutamate synthase
YOL140w	*ARG8*	acetylornithine aminotransferase
YDR234w	*LYS4*	aconitate hydratase
YER086w	*ILV1*	anabolic serine and threonine dehydratase
YDR354w	*TRP4*	anthranilate phosphoribosyltransferase
YER090w	*TRP2*	anthranilate synthase component I
YKL211c	*TRP3*	anthranilate synthase component II
YJR109c	*CPA2*	arginine-specific carbamoylphosphate synthase, large subunit
YOR303w	*CPA1*	arginine-specific carbamoylphosphate synthase, small subunit
YOL058w	*ARG1*	argininosuccinate synthetase
YHR018c	*ARG4*	arginosuccinate lyase
YDR127w	*ARO1*	arom pentafunctional enzyme
YPR145w	*ASN1*	asparagine synthetase
YGR124w	*ASN2*	asparagine synthetase
YLR027c	*AAT2*	aspartate aminotransferase, cytosolic
YKL106w	*AAT1*	aspartate transaminase, mitochondrial
YDR158w	*HOM2*	aspartate-semialdehyde dehydrogenase
YFR030w	*MET10*	assimilatory sulphite reductase flavin-binding subunit
YER055c	*HIS1*	ATP phosphoribosyltransferase
YCL018w	*LEU2*	β-isopropyl-malate dehydrogenase
YJR148w	*TWT2*	branched-chain amino acid aminotransferase, cytosolic
YHR208w	*TWT1*	branched-chain amino acid aminotransferase, mitochondrial
YGR204w	*ADE3*	C1-tetrahydrofolate synthase, cytoplasmic
YBR084w	*MIS1*	C1-tetrahydrofolate synthase, mitochondrial
YPR060c	*ARO7*	chorismate mutase
YGL148w	*ARO2*	chorismate synthase
YGR155w	*CYS4*	cystathionine β-synthase
YAL012w	*CYS3*	cystathionine γ-lyase
YER023w	*PRO3*	δ 1-pyrroline-5-carboxylate reductase
YHR068w	*DYS1*	deoxyhypusine synthase
YOR236w	*DFR1*	dihydrofolate reductase
YFL018c	*LPD1*	dihydrolipoamide dehydrogenase
YJR016c	*ILV3*	dihydroxy-acid dehydratase
YLR172c	*DPH5*	diphthamide methyltransferase
YKL191w	*DPH2*	diphtheria toxin resistance protein
YOR323c	*PRO2*	γ-glutamyl phosphate reductase
YDR300c	*PRO1*	glutamate 5-kinase
YDL171c	*GLT1*	glutamate synthase (NAPDPH) (GOGAT)
YPR035w	*GLN1*	glutamate-ammonia ligase
YBR248c	*HIS7*	glutamine amidotransferase/cyclase
YFR025c	*HIS2*	histidinol phosphatase
YIL116w	*HIS5*	histidinol-phosphate aminotransferase
YMR038c	*LYS7*	homocitrate dehydrogenase
YDL182w	*LYS20*	homocitrate synthase
YJR139c	*HOM6*	homoserine dehydrogenase
YHR025w	*THR1*	homoserine kinase
YNL277w	*MET2*	homoserine O-acetyltransferase
YOR202w	*HIS3*	imidazoleglycerol-phosphate dehydratase
YLR355c	*ILV5*	ketol-acid reducto-isomerase
YBR115c	*LYS2*	L-aminoadipate-semialdehyde dehydrogenase, large subunit
YGL154c	*LYS5*	L-aminoadipate-semialdehyde dehydrogenase, small subunit
YER052c	*HOM3*	L-aspartate 4-P-transferase
YLR303w	*MET25*	O-acetylhomoserine sulphydrylase
YLR438w	*CAR2*	ornithine aminotransferase
YJL088w	*ARG3*	ornithine carbamoyltransferase
YCL030c	*HIS4*	phosphoribosyl-AMP cyclohydrolase/phosphoribosyl-ATP pyrophosphatase/histidinol dehydrogenase
YDR007w	*TRP1*	phosphoribosylanthranilate isomerase
YGR208w	*SER2*	phosphoserine phosphatase
YOR184w	*SER1*	phosphoserine transaminase
YNL316c	*PHA2*	prephenate dehydratase
YBR166c	*TYR1*	prephenate dehydrogenase (NADP+)
YLR142w	*PUT1*	proline oxidase
YOL064c	*MET22*	protein ser/thr phosphatase
YCR054c	*CTR86*	putative threonine biosynthesis pathway protein
YEL046c	*GLY1*	required for glycine prototrophy in SHMT1 and SHMT2 double mutant
YIR034c	*LYS1*	saccharopine dehydrogenase
YNR050c	*LYS9*	saccharopine dehydrogenase (NADP+, L-glutamate forming)
YCL009c		similarity to acetolactate synthase III small subunit
YLR089c		similarity to alanine transaminases
YGL184c		similarity to cystathionine β-lyase
YHR112c		similarity to cystathionine γ-synthases
YIL074c		similarity to *E. coli* phosphoglycerate dehydrogenase
YGR012w		similarity to *E. nidulans* cysteine synthase
YMR250w		similarity to glutamate decarboxylases
YMR062c		similarity to glutamate N-acetyltransferase
YDL131w		similarity to homocitrate synthases and isopropylmalate synthases
YIL094c		similarity to isopropyl malate and tartrate dehydrogenases
YEL038w	*UTR4*	similarity to *K. oxytoca* enolase-phosphatase E-1
YML082w		similarity to *N. crassa* O-succinylhomoserine (thiol)-lyase
YJR130c		similarity to O-succinylhomoserine (thiol)-lyase
YKL215c		similarity to *P. aeruginosa* hyuA and hyuB
YOR280c		similarity to *S. pombe* dihydrofolate reductase
YFL030w		similarity to several transaminases
YOR108w		strong similarity to 2-isopropylmalate synthase
YAL004w		strong similarity to *A. klebsiana* glutamate dehydrogenase
YDR111c		strong similarity to alanine transaminase
YFR055w		strong similarity to β-cystathionases
YHR033w		strong similarity to glutamate 5-kinase
YHR070w		strong similarity to *N. crassa* met-10+ protein
YLL058w		strong similarity to *N. crassa* O-succinylhomoserine (thiol)-lyase
YER081w		strong similarity to phosphoglycerate dehydrogenases
YJR010w	*MET3*	sulphate adenylyltransferase
YCR053w	*THR4*	threonine synthase (o-p-homoserine p-lyase)
YPR074c	*TKL1*	transketolase 1
YBR117c	*TKL2*	transketolase 2
YGL026c	*TRP5*	tryptophan synthase
YML096w		weak similarity to asparagine synthases

regulation of amino-acid metabolism

YDR173c	*ARG82*	arginine metabolism transcription factor
YKL112w	*ABF1*	ARS-binding factor
YER055c	*HIS1*	ATP phosphoribosyltransferase
YBR253w	*SRB6*	DNA-directed RNA polymerase II suppressor protein
YNL236w	*SIN4*	global regulator protein

YFL010w-a	*AUA1*	involved in ammonia regulation of amino-acid transport
YIL046w	*MET30*	involved in regulation of sulphur assimilation genes
YJR060w	*CBF1*	kinetochore protein
YDR159w	*SAC3*	leucine permease transcriptional regulator
YDR207c	*UME6*	negative transcriptional regulator
YKL015w	*PUT3*	positive activator of the proline utilisation pathway
YCR028c	*FEN2*	similarity to allantoate permease transporter
YMR116c		strong similarity to *N. crassa* CPC2 protein
YKR099w	*BAS1*	transcription factor
YLR098c	*CHA4*	transcription factor
YLR451w	*LEU3*	transcription factor
YMR042w	*ARG80*	transcription factor involved in arginine metabolism
YML099c	*ARG81*	transcription factor involved in arginine metabolism
YMR043w	*MCM1*	transcription factor of the MADS box family
YIR023w	*DAL81*	transcriptional activator for allantoin and GABA catabolic genes
YEL009c	*GCN4*	transcriptional activator of amino acid biosynthetic genes
YDR034c	*LYS14*	transcriptional activator of lysine pathway genes
YIR017c	*MET28*	transcriptional activator of sulphur amino acid metabolism
YNL103w	*MET4*	transcriptional activator of sulphur metabolism

amino-acid transport

YBR068c	*BAP2*	amino-acid permease
YEL063c	*CAN1*	amino-acid permease
YBR069c	*VAP1*	amino-acid permease
YCL025c	*AGP1*	asparagine and glutamine permease
YPL265w	*DIP5*	dicarboxylic amino-acid permease
YKR039w	*GAP1*	general amino-acid permease
YDR508c	*GNP1*	high-affinity glutamine permease
YGR055w	*MUP1*	high-affinity methionine permease
YNL270c	*ALP1*	high-affinity permease for basic amino acids
YOL020w	*SCM2*	high-affinity tryptophan transport protein
YGR191w	*HIP1*	histidine permease
YLR375w	*STP3*	involved in pre-tRNA splicing and in uptake of branched-chain amino acids
YDL048c	*STP4*	involved in pre-tRNA splicing and in uptake of branched-chain amino acids
YNL268w	*LYP1*	lysine-specific high affinity permease
YOR130c	*ARG11*	member of the mitochondrial carrier family (MCF)
YHL036w	*MUP3*	methionine permease
YOR348c	*PUT4*	proline and γ-aminobutyrate permease
YFL055w		similarity to Gap1p and other amino-acid permeases
YDR160w		similarity to lysine transport protein LYP1
YLL061w		strong similarity to amino-acid transport protein Gap1p
YPL274w		strong similarity to amino-acid transport proteins
YDR046c	*(PAP1)*	strong similarity to amino-acid transport proteins
YBR132c		strong similarity to amino-acid permeases

amino-acid degradation

YPL111w	*CAR1*	arginase
YDR321w	*ASP1*	asparaginase
YOR375c	*GDH1*	glutamate dehydrogenase (NADP⁺)
YMR189w	*GSD2*	glycine decarboxylase subunit
YDR019c	*GCV1*	glycine decarboxylase T subunit
YDR272w	*GLO2*	glyoxalase II
YOR040w	*GLO4*	glyoxalase II
YLR155c	*ASP3A*	L-asparaginase II
YLR157c	*ASP3B*	L-asparaginase II
YLR158c	*ASP3C*	L-asparaginase II
YLR160c	*ASP3D*	L-asparaginase II
YCL064c	*CHA1*	L-serine/L-threonine deaminase
YAL062w	*GDH3*	NADP-glutamate dehydrogenase
YDL215c	*GDH2*	NAD-specific glutamate dehydrogenase (NAD)
YKL184w	*ORD1*	ornithine decarboxylase
YLR142w	*PUT1*	proline oxidase
YLR180w	*SAM1*	S-adenosylmethionine synthetase 1
YDR502c	*SAM2*	S-adenosylmethionine synthetase 2
YIL167w	*SDL1*	serine dehydratase
YDR294c		similarity to glutamate decarboxylases
YJR078w		similarity to indoleamine 2,3-dioxygenase
YIL042c		similarity to rat branched-chain α-ketoacid dehydrogenase kinase
YGL202w		similarity to rat kynurenine/α-aminoadipate aminotransferase
YHR137w		similarity to rat kynurenine/α-aminoadipate aminotransferase
YFL030w		similarity to several transaminases
YDR111c		strong similarity to alanine transaminase
YBR006w		strong similarity to *E. coli* succinate semialdehyde dehydrogenase
YAL044c	*GCV3*	strong similarity to human glycine cleavage system protein H
YIL168w	*SDL1*	strong similarity to L-serine dehydratase Cha1p
YLR231c		strong similarity to rat kynureninase
YKL218c		strong similarity to threonine dehydratase
YBR208c	*DUR1,2*	urea amidolyase

other amino-acid metabolism activities

YKL157w	*APE2*	aminopeptidase yscII
YGL017w	*ATE1*	arginyl tRNA transferase
YPR069c	*SPE3*	putrescine aminopropyltransferase (spermidine synthase)
YFR018c		similarity to human glutaminyl-peptide cyclotransferase

nitrogen and sulphur metabolism

nitrogen and sulphur utilization

YPR167c	*MET16*	3'-phosphoadenylylsulphate reductase
YIR027c	*DAL1*	allantoinase
YIR029w	*DAL2*	allantoinase
YDR242w	*AMD2*	amidase
YOL058w	*ARG1*	argininosuccinate synthetase
YLR027c	*AAT2*	aspartate aminotransferase, cytosolic
YFR030w	*MET10*	assimilatory sulphite reductase flavin-binding subunit
YKL001c	*MET14*	ATP adenosine-5'-phosphosulphate 3'-phosphotransferase
YHR176w	*FMO*	flavin-containing monooxygenase
YOR375c	*GDH1*	glutamate dehydrogenase (NADP⁺)
YDL171c	*GLT1*	glutamate synthase (NAPDPH) (GOGAT)
YPR035w	*GLN1*	glutamate-ammonia ligase
YJL172w	*CPS1*	Gly-X carboxypeptidase YSCS
YAL062w	*GDH3*	NADP-glutamate dehydrogenase
YDL215c	*GDH2*	NAD-specific glutamate dehydrogenase (NAD)
YLR438w	*CAR2*	ornithine aminotransferase
YJR149w		similarity to 2-nitropropane dioxygenase
YLR089c		similarity to alanine transaminases
YMR293c		similarity to amidases
YFL061w		similarity to *M. verrucaria* cyanamide hydratase
YFL030w		similarity to several transaminases
YOR251c		similarity to thiosulphate sulphurtransferases
YDR111c		strong similarity to alanine transaminase
YER057c		strong similarity to *Azotobacter* nitrogen fixation vnfA protein
YIL051c		strong similarity to *Azotobacter* nitrogen fixation vnfA protein
YJL060w		strong similarity to kynurenine aminotransferase
YPL135w		strong similarity to nitrogen fixation protein (nifU)
YOR226c		strong similarity to nitrogen fixation proteins
YJR010w	*MET3*	sulphate adenylyltransferase
YJR137c		sulphite reductase
YBR208c	*DUR1,2*	urea amidolyase
YIR032c	*DAL3*	ureidoglycolate hydrolase
YJL035c		weak similarity to *B. japonicum* nitrogen fixation protein
YKL040c		weak similarity to nitrogen fixation protein nifU

regulation of nitrogen and sulphur utilization

YGR019w	*UGA1*	4-aminobutyrate aminotransferase (GABA transaminase)
YPL111w	*CAR1*	arginase
YKL112w	*ABF1*	ARS-binding factor
YNL216w	*RAP1*	DNA-binding protein with repressor and activator activity
YFL010w-a	*AUA1*	involved in ammonia regulation of amino acid transport
YIR030c	*DCG1*	involved in nitrogen-catabolite metabolism
YBR213w	*MET8*	involved in the expression of PAPS reductase and sulphite reductase
YDR207c	*UME6*	negative transcriptional regulator
YNL229c	*URE2*	nitrogen catabolite repression regulator
YEL062w	*NPR2*	nitrogen permease regulator
YNL183c	*NPR1*	ser/thr protein kinase
YCR028c	*FEN2*	similarity to allantoate permease transporter
YLR013w		similarity to nitrogen regulatory proteins
YGL254w	*FZF1*	sulphite resistance protein
YFL021w	*GAT1*	transcription factor for nitrogen regulation
YER040w	*GLN3*	transcription factor for positive nitrogen regulation
YMR042w	*ARG80*	transcription factor involved in arginine metabolism
YML099c	*ARG81*	transcription factor involved in arginine metabolism
YMR043w	*MCM1*	transcription factor of the MADS box family
YIR023w	*DAL81*	transcriptional activator for allantoin and GABA catabolic genes
YDL170w	*UGA3*	transcriptional activator for GABA catabolic genes
YNL103w	*MET4*	transcriptional activator of sulphur metabolism
YKR034w	*DAL80*	transcriptional repressor for allantoin and GABA catabolic genes

nitrogen and sulphur transport

YGR121c	*MEP1*	ammonia permease of high capacity and moderate affinity
YNL142w	*MEP2*	high-affinity low-capacity ammonia permease
YBR294w	*SUL1*	high-affinity sulphate transport protein
YPR003c		similarity to sulphate transporter proteins
YPR138c	*MEP3*	strong similarity to ammonium transport proteins
YLR092w	*SEL2*	strong similarity to Sul1p
YHL016c	*DUR3*	urea transport protein

nucleotide metabolism

purine-ribonucleotide metabolism

YLR028c	*ADE16*	5-aminoimidazole-4-carboxamide ribotide transformylase
YGR061c	*ADE6*	5'-phosphoribosylformyl glycinamidine synthetase
YML022w	*APT1*	adenine phosphoribosyltransferase
YDR441c	*APT2*	adenine phosphoribosyltransferase
YNL220w	*ADE12*	adenylosuccinate synthetase
YIR027c	*DAL1*	allantoinase
YIR029w	*DAL2*	allantoinase
YMR300c	*ADE4*	amidophosphoribosyltransferase
YML035c	*AMD1*	AMP deaminase
YGR204w	*ADE3*	C1-tetrahydrofolate synthase, cytoplasmic
YBR084w	*MIS1*	C1-tetrahydrofolate synthase, mitochondrial
YOR236w	*DFR1*	dihydrofolate reductase
YMR217w	*GUA1*	GMP synthase (glutamine-hydrolyzing)
YHR216w	*PUR5*	IMP dehydrogenase
YIR031c	*DAL7*	malate synthase 2
YKR080w	*MTD1*	methylenetetrahydrofolate dehydrogenase (NAD⁺)
YAR015w	*ADE1*	phosphoribosylamidoimidazole-succinocarboxamide synthase
YGL234w	*ADE5,7*	phosphoribosylamine-glycine ligase and phosphoribosylformylglycinamidine cyclo-ligase
YOR128c	*ADE2*	phosphoribosylaminoimidazole carboxylase
YDR408c	*ADE8*	phosphoribosylglycinamide formyltransferase (GART)
YKL181w	*PRPS1*	ribose-phosphate pyrophosphokinase
YER099c	*PRS2*	ribose-phosphate pyrophosphokinase
YHL011c	*PRS3*	ribose-phosphate pyrophosphokinase
YBL068w	*PRS4*	ribose-phosphate pyrophosphokinase
YLR058c	*SHM2*	serine hydroxymethyltransferase, cytoplasmic
YBR263w	*SHM1*	serine hydroxymethyltransferase, mitochondrial
YNL141w		similarity to adenosine deaminase
YBR284w		similarity to AMP deaminase
YJL070c		similarity to AMP deaminases
YLR017w		similarity to human 5'-methylthioadenosine phosphorylase
YOL061w		similarity to ribose-phosphate pyrophosphokinases
YOR280c		similarity to *S. pombe* dihydrofolate reductase
YDR020c		similarity to uridine kinases and phosphoribulokinases
YLR359w		strong similarity to adenylosuccinate lyase
YMR120c		strong similarity to chicken purH bifunctional enzyme
YJR105w		strong similarity to human adenosine kinase
YAR075w		strong similarity to IMP dehydrogenases
YML056c		strong similarity to IMP dehydrogenases
YAR073w	*FUN63*	strong similarity to IMP dehydrogenases
YLR432w		strong similarity to IMP dehydrogenases, Pur5p and YML056c
YLR209c		strong similarity to purine-nucleoside phosphorylases
YBR208c	*DUR1,2*	urea amidolyase
YIR032c	*DAL3*	ureidoglycolate hydrolase

pyrimidine-ribonucleotide metabolism

YBL039c	*URA7*	CTP synthase 1
YJR103w	*URA8*	CTP synthase 2
YHR144c	*DCD1*	deoxycytidylate deaminase
YLR420w	*URA4*	dihydroorotase
YKL216w	*URA1*	dihydroorotate dehydrogenase
YJR057w	*CDC8*	dTMP kinase
YBR252w	*DUT1*	dUTP pyrophosphatase, mitochondrial
YJL130c	*URA2*	multifunctional pyrimidine biosynthesis protein
YML106w	*URA5*	orotate phosphoribosyltransferase 1
YMR271c	*URA10*	orotate phosphoribosyltransferase 2
YEL021w	*URA3*	orotidine-5'-phosphate decarboxylase
YFL058w	*THI5*	pyrimidine biosynthesis protein
YKL181w	*PRPS1*	ribose-phosphate pyrophosphokinase
YER099c	*PRS2*	ribose-phosphate pyrophosphokinase
YHL011c	*PRS3*	ribose-phosphate pyrophosphokinase
YBL068w	*PRS4*	ribose-phosphate pyrophosphokinase
YOL061w		similarity to ribose-phosphate pyrophosphokinases
YOR280c		similarity to *S. pombe* dihydrofolate reductase
YLR245c		strong similarity to *B. subtilis* cytidine deaminase
YNL332w		strong similarity to Thi5p, YJR156c, YDL244w and *A. parasiticus*, *S. pombe* Nmt1p
YDL244w		strong similarity to Thi5p, YJR156c, YNL332w and *A. parasiticus*, *S. pombe* nmt1 protein
YJR156c	*THI11*	thiamine regulated gene, homologous to nmt1a in *S. pombe*
YHR128w	*FUR1*	uracil phosphoribosyltransferase
YNR012w	*URK1*	uridine kinase
YKL024c	*URA6*	uridine-monophosphate kinase

deoxyribonucleotide metabolism

YDR513w	*TTR1*	glutaredoxin
YER070w	*RNR1*	ribonucleoside-diphosphate reductase, large subunit

YJL026w	*RNR2*	ribonucleoside-diphosphate reductase, small subunit
YGR180c	*RNR4*	ribonucleotide reductase small subunit
YIL066c	*RNR3*	ribonucleotide reductase, repair inducible large subunit
YBR014c		similarity to glutaredoxin
YPL059w		similarity to glutaredoxins
YOR269w	*PAC1*	similarity to human LIS-1 protein
YDL010w		similarity to hypothetical protein YBR014c and glutaredoxins
YCL035c		strong similarity to glutaredoxin
YDR353w		strong similarity to thioredoxin reductase (NADPH)
YHR106w		strong similarity to thioredoxin reductases
YOR074c	*CDC21*	thymidylate synthase

metabolism of cyclic and unusual nucleotides

YJL005w	*CYR1*	adenylate cyclase
YCL050c	*APA1*	ATP adenylyltransferase I
YDR530c	*APA2*	ATP adenylyltransferase II
YOL081w	*IRA2*	GTPase-activating protein for RAS proteins
YOR360c	*PDE2*	high affinity 3',5'-cyclic-nucleotide phosphodiesterase
YPL212c	*PUS1*	pseudouridine synthase 1
YGL063w	*PUS2*	pseudouridine synthase 2

regulation of nucleotide metabolism

YOL081w	*IRA2*	GTPase-activating protein for RAS proteins
YOR101w	*RAS1*	GTP-binding protein
YNL098c	*RAS2*	GTP-binding protein
YDL106c	*GRF10*	homeodomain protein
YGL248w	*PDE1*	low-affinity 3',5'-cyclic-nucleotide phosphodiesterase
YNL076w	*MKS1*	negative regulator of RAS-cAMP pathway
YJL096w		putative regulator of purine and/or pyrimidine biosynthesis
YOL110w	*SHR5*	RAS suppressor
YKR099w	*BAS1*	transcription factor
YLR014c	*PPR1*	transcription factor regulating pyrimidine pathway
YIR023w	*DAL81*	transcriptional activator for allantoin and GABA catabolic genes
YNL314w	*DAL82*	transcriptional activator for allantoin catabolic genes

polynucleotide degradation

YMR287c	*MSU1*	3'-5' exonuclease for RNA 3' ss-tail, mitochondrial
YPL029w	*SUV3*	ATP-dependent RNA helicase, mitochondrial
YOR033c	*DHS1*	exonuclease, interacting with Msh2p
YKL149c	*DBR1*	lariat-debranching enzyme
YGL173c	*KEM1*	multifunctional nuclease
YLR363c	*NMD4*	Nam7p/Upf1p-interacting protein
YJR132w	*NMD5*	Nam7p/Upf1p-interacting protein
YMR080c	*NAM7*	nonsense-mediated mRNA decay protein
YGR072w	*UPF3*	nonsense-mediated mRNA decay protein
YHR077c	*NMD2*	nonsense-mediated mRNA decay protein 2
YJL208c	*NUC1*	nuclease, mitochondrial
YMR234w	*RNH1*	ribonuclease H
YPL123c		similarity to ribonucleases
YGR195w		weak similarity to *P. aeruginosa* RNase PH

nucleotide transport

YMR056c	*AAC1*	ADP/ATP carrier protein (MCF)
YBL030c	*AAC2*	ADP/ATP carrier protein (MCF)
YBR085w	*AAC3*	ADP/ATP carrier protein (MCF)
YER056c	*FCY2*	purine-cytosine permease
YER060w	*FCY21*	purine-cytosine permease
YOR222w		similarity to ADP/ATP carrier proteins
YPR011c		similarity to ADP/ATP carrier proteins and Graves disease carrier protein
YOR071c		similarity to allantoin or uracil transport proteins
YOR192c		similarity to allantoin or uracil transport proteins
YLR237w		similarity to allantoin transport protein
YGR096w		similarity to bovine Graves disease carrier protein
YHR002w		similarity to bovine mitochondrial carrier protein/Grave's disease carrier protein
YGL186c		similarity to hypothetical protein YER060w and weak similarity to Fcy2p
YBL042c		strong similarity to allantoin and uracil transport proteins
YER060w-a	*FCY22*	strong similarity to Fcy2p
YBR021w	*FUR4*	uracil permease

other nucleotide-metabolism activities

YDR226w	*ADK1*	adenylate kinase, cytosolic
YER170w	*ADK2*	adenylate kinase, mitochondrial
YDR454c	*GUK1*	guanylate kinase
YKL067w	*YNK1*	nucleoside diphosphate kinase
YDL125c	*HNT1*	similarity to protein kinase C inhibitor-I
YDR305c	*HNT2*	strong similarity to *S. pombe* diadenosine 5',5'''-P1,P4-tetraphosphate asymmetrical hydrolase

phosphate metabolism

phosphate utilization

YHR044c	*DOG1*	2-deoxyglucose-6-phosphate phosphatase
YHR043c	*DOG2*	2-deoxyglucose-6-phosphate phosphatase
YMR282c	*AEP2*	2'-O-ribosyl phosphate transferase
YBR092c	*PHO3*	constitutive acid phosphatase
YHR201c	*PPX1*	exopolyphosphatase
YBR011c	*IPP1*	inorganic pyrophosphatase, cytoplasmic
YMR267w	*PPA2*	inorganic pyrophosphatase, mitochondrial
YBR093c	*PHO5*	repressible acid phosphatase
YAR071w	*PHO11*	secreted acid phosphatase
YHR215w	*PHO12*	secreted acid phosphatase
YDL024c		strong similarity to acid phosphatase
YNL330c	*RPD3*	transcription modifier protein
YBR243c	*ALG7*	UDP-N-acetylglucosamine-1-phosphate transferase

regulation of phosphate utilization

YOL001w	*PHO80*	cyclin
YGR233c	*PHO81*	cyclin-dependent kinase inhibitor
YPL031c	*PHO85*	cyclin-dependent protein kinase
YML121w	*GTR1*	GTP-binding protein
YDL106c	*GRF10*	homeodomain protein
YBR106w	*PHO88*	involved in phosphate transport
YFR034c	*PHO4*	transcription factor

phosphate transport

YLR348c		dicarboxylate carrier protein
YML123c	*PHO84*	high-affinity inorganic phosphate/H^+ symporter
YJL117w	*PHO86*	inorganic phospate transporter
YBR106w	*PHO88*	involved in phosphate transport
YCR037c	*PHO87*	member of the phosphate permease family
YJR077c	*MIR1*	phosphate transport protein, mitochondrial (MCF)
YNR013c		similarity to membrane protein Pho87p and hypothetical protein YJL198w
YCR098c	*GIT1*	similarity to phosphate transporter proteins
YER053c		strong similarity to mitochondrial phosphate carrier protein
YJL198w		strong similarity to Pho87p
YBR296c		strong similarity to phosphate-repressible phosphate permease

other phosphate-metabolism activities

YDL236w	*PHO13*	4-nitrophenylphosphatase
YOR008c	SLG1	weak similarity to *L. mexicana* lmsap2 gene

carbohydrate metabolism

carbohydrate utilization

YLR342w	*FKS1*	1,3—D-glucan synthase, catalytic subunit
YGR032w	*GSC2*	1,3—D-glucan synthase, subunit
YEL011w	*GLC3*	1,4-glucan branching enzyme (glycogen branching enzyme)
YHR044c	*DOG1*	2-deoxyglucose-6-phosphate phosphatase
YHR043c	*DOG2*	2-deoxyglucose-6-phosphate phosphatase
YIL125w	*KGD1*	2-oxoglutarate dehydrogenase complex E1 component
YDR148c	*KGD2*	2-oxoglutarate dehydrogenase complex E2 component
YGR240c	*PFK1*	6-phosphofructokinase, α subunit
YMR205c	*PFK2*	6-phosphofructokinase, β subunit
YIL107c	*PFK26*	6-phosphofructose-2-kinase, isoenzyme 1
YOL136c	*PFK27*	6-phosphofructose-2-kinase, isoenzyme 2
YHR183w	*GND1*	6-phosphogluconate dehydrogenase
YBR001c	*NTH2*	α,α-trehalase
YDR074w	*TPS2*	α,α-trehalose-phosphate synthase, 105K subunit
YMR261c	*TPS3*	α,α-trehalose-phosphate synthase, 115K subunit
YML100w	*TSL1*	α,α-trehalose-phosphate synthase, 123K subunit
YBR126c	*TPS1*	α,α-trehalose-phosphate synthase, 56K subunit
YJR131w	*MNS1*	α-1,2-mannosidase
YDR483w	*KRE2*	α-1,2-mannosyltransferase
YER001w	*MNN1*	α-1,3-mannosyltransferase
YGL038c	*OCH1*	α-1,6-mannosyltransferase
YAL054c	*ACS1*	acetyl-CoA synthetase
YLR153c	*ACS2*	acetyl-CoA synthetase
YPR026w	*ATH1*	acid trehalase, vacuolar
YBR299w	*MAL32*	α-glucosidase
YGR292w	*MAL12*	α-glucosidase of the *MAL1* locus
YOR377w	*ATF1*	alcohol acetyltransferase
YOL086c	*ADH1*	alcohol dehydrogenase I
YMR303c	*ADH2*	alcohol dehydrogenase II
YMR083w	*ADH3*	alcohol dehydrogenase III
YGL256w	*ADH4*	alcohol dehydrogenase IV
YBR145w	*ADH5*	alcohol dehydrogenase V
YGR177c	*ATF2*	alcohol O-acetyltransferase
YCL040w	*GLK1*	aldohexose specific glucokinase
YGL156w	*AMS1*	α-mannosidase
YGL027c	*CWH41*	β-1,6-glucan assembly protein
YHR101c	*BIG1*	big cells phenotype
YBR110w	*ALG1*	β-mannosyltransferase
YBR084w	*MIS1*	C1-tetrahydrofolate synthase, mitochondrial
YNL322c	*KRE1*	cell-wall protein
YJL174w	*KRE9*	cell-wall synthesis protein
YJL099w	*CHS6*	chitin biosynthesis protein
YNL192w	*CHS1*	chitin synthase I
YBR038w	*CHS2*	chitin synthase II
YBR023c	*CHS3*	chitin synthase III
YNR001c	*CIT1*	citrate (si)-synthase, mitochondrial
YPR001w	*CIT3*	citrate (si)-synthase, mitochondrial
YCR005c	*CIT2*	citrate (si)-synthase, peroxisomal
YML086c	*ALO*	D-arabinino-1,4-lactone oxidase
YNL071w	*LAT1*	dihydrolipoamide S-acetyltransferase
YDL174c	*DLD1*	D-lactate ferricytochrome c oxidoreductase (D-LCR)
YER062c	*HOR2*	DL-glycerol phosphatase
YIL053w	*RHR2*	DL-glycerol phosphatase
YPL227c	*ALG5*	dolichol-P-glucose synthetase
YPR183w	*DPM1*	dolichyl-phosphate β-D-mannosyltransferase
YJR143c	*PMT4*	dolichyl-phosphate-mannose-protein O-mannosyl transferase
YOR095c	*RKI1*	D-ribose-5-phosphate ketol-isomerase
YJL121c	*POS18*	D-ribulose-5-phosphate 3-epimerase
YGR282c	*BGL2*	endo-1,3—glucanase of the cell wall
YLR286c	*CTS1*	endochitinase
YGR254w	*ENO1*	enolase I (2-phosphoglycerate dehydratase)
YHR174w	*ENO2*	enolase II (2-phosphoglycerate dehydratase)
YOR393w	*ERR1*	enolase-related protein
YPL281c	*ERR2*	enolase-related protein
YOR190w	*SPR1*	exo-1,3—glucanase
YDR261c	*EXG2*	exo-1,3—glucanase minor isoform
YLR300w	*EXG1*	exo-1,3—glucanase(I/II), major isoform
YIR019c	*STA1*	extracellular α-1,4-glucan glucosidase
YLR377c	*FBP1*	fructose-1,6-bisphosphatase
YJL155c	*FBP26*	fructose-2,6-bisphosphatase
YKL060c	*FBA1*	fructose-bisphosphate aldolase
YDL049c	*KNH1*	functional homologue of Kre9p
YBR020w	*GAL1*	galactokinase
YDR009w	*GAL3*	galactokinase
YOR120w	*GCY1*	galactose-induced protein of aldo/keto reductase family
YPR159w	*KRE6*	glucan synthase subunit
YGR143w	*SKN1*	glucan synthase subunit
YKL104c	*GFA1*	glucosamine-fructose-6-phosphate transaminase
YNL241c	*ZWF1*	glucose-6-phosphate dehydrogenase
YBR196c	*PGI1*	glucose-6-phosphate isomerase
YBR229c	*ROT2*	glucosidase II, catalytic subunit
YOR002w	*ALG6*	glucosyltransferase
YOR067c	*ALG8*	glucosyltransferase
YJL052w	*TDH1*	glyceraldehyde-3-phosphate dehydrogenase 1
YJR009c	*TDH2*	glyceraldehyde-3-phosphate dehydrogenase 2
YGR192c	*TDH3*	glyceraldehyde-3-phosphate dehydrogenase 3
YHL032c	*GUT1*	glycerol kinase
YDL022w	*GPD1*	glycerol-3-phosphate dehydrogenase (NAD^+), cytoplasmic
YOL059w	*GPD3*	glycerol-3-phosphate dehydrogenase (NAD^+), mitochondrial
YIL155c	*GUT2*	glycerol-3-phosphate dehydrogenase, mitochondrial
YPR160w	*GPH1*	glycogen phosphorylase
YFL014w	*HSP12*	heat-shock protein
YFR053c	*HXK1*	hexokinase I
YGL253w	*HXK2*	hexokinase II
YDL013w	*HEX3*	hexose metabolism-related protein
YDL182w	*LYS20*	homocitrate synthase
YOR126c	*EST2*	isoamyl acetate hydrolytic enzyme
YNL037c	*IDH1*	isocitrate dehydrogenase (NAD^+) subunit 1, mitochondrial
YOR136w	*IDH2*	isocitrate dehydrogenase (NAD^+) subunit 2, mitochondrial
YDL066w	*IDP1*	isocitrate dehydrogenase ($NADP^+$), mitochondrial
YLR174w	*IDP2*	isocitrate dehydrogenase, cytosolic
YER065c	*ICL1*	isocitrate lyase
YOR336w	*KRE5*	killer toxin-resistance protein
YML054c	*CYB2*	lactate dehydrogenase cytochrome b2
YDL168w	*SFA1*	long-chain alcohol dehydrogenase
YOL126c	*MDH2*	malate dehydrogenase, cytoplasmic
YKL085w	*MDH1*	malate dehydrogenase, mitochondrial
YDL078c	*MDH3*	malate dehydrogenase, peroxisomal
YNL117w	*MLS1*	malate synthase 1
YIR031c	*DAL7*	malate synthase 2
YDL055c	*PSA1*	mannose-1-phosphate guanyltransferase
YER003c	*PMI40*	mannose-6-phosphate isomerase
YGL065c	*ALG2*	mannosyltransferase
YNL219c	*ALG9*	mannosyltransferase
YKR061w	*KTR2*	mannosyltransferase
YDL095w	*PMT1*	mannosyltransferase
YAL023c	*PMT2*	mannosyltransferase
YOR321w	*PMT3*	mannosyltransferase
YBL082c	*RHK1*	mannosyltransferase
YJL139c	*YUR1*	mannosyltransferase
YJL153c	*INO1*	myo-inositol-1-phosphate synthase
YPL175w	*SPT14*	N-acetylglucosaminyltransferase
YDR001c	*NTH1*	neutral trehalase (α,α-trehalase)
YPR006c	*ICL2*	non-functional isocitrate lyase
YJL002c	*OST1*	oligosaccharyltransferase α subunit
YEL002c	*WBP1*	oligosaccharyltransferase β subunit
YMR149w	*SWP1*	oligosaccharyltransferase δ subunit
YOR103c	*OST2*	oligosaccharyltransferase ε subunit
YOR085w	*OST3*	oligosaccharyltransferase γ subunit
YDL232w	*OST4*	oligosaccharyltransferase subunit
YGL022w	*STT3*	oligosaccharyltransferase subunit
YEL058w	*PCM1*	phosphoacetylglucosamine mutase
YKR097w	*PCK1*	phosphoenolpyruvate carboxykinase
YMR105c	*PGM2*	phosphoglucomutase, major isoform
YKL127w	*PGM1*	phosphoglucomutase, minor isoform
YGR256w	*GND2*	phosphogluconate dehydrogenase
YCR012w	*PGK1*	phosphoglycerate kinase
YKL152c	*GPM1*	phosphoglycerate mutase

YOL056w	*GPM3*	phosphoglycerate mutase
YFL045c	*SEC53*	phosphomannomutase
YCR034w	*GNS1*	probable 1,3–glucan synthase subunit
YIL045w	*PIG2*	Protein interacting with Gsy2p
YIL085c	*KTR7*	putative α-1,2-mannosyltransferase
YNL029c	*KTR5*	putative mannosyltransferase
YGR199w	*PMT6*	putative mannosyltransferase
YGL062w	*PYC1*	pyruvate carboxylase, isozyme 1
YBR218c	*PYC2*	pyruvate carboxylase, isozyme 2
YDR081c	*PDC2*	pyruvate decarboxylase regulatory protein
YLR044c	*PDC1*	pyruvate decarboxylase, isozyme 1
YLR134w	*PDC5*	pyruvate decarboxylase, isozyme 2
YGR087c	*PDC6*	pyruvate decarboxylase, isozyme 3
YER178w	*PDA1*	pyruvate dehydrogenase (lipoamide) α subunit
YBR221c	*PDB1*	pyruvate dehydrogenase (lipoamide) β subunit
YGR193c	*PDX1*	pyruvate dehydrogenase complex protein X
YAL038w	*CDC19*	pyruvate kinase
YNL048w	*ALG11*	required for asparagine-linked glycosylation
YCR036w	*RBK1*	ribokinase
YJL137c	*GLG2*	self-glucosylating initiator of glycogen synthesis
YMR306w		similarity to 1,3–glucan synthases
YAL060w	*FUN49*	similarity to alcohol/sorbitol dehydrogenase
YAL061w	*FUN50*	similarity to alcohol/sorbitol dehydrogenase
YMR318c		similarity to alcohol-dehydrogenase
YDL124w		similarity to aldose reductases
YHR204w		similarity to α-mannosidases
YPL088w		similarity to aryl-alcohol dehydrogenases
YML070w		similarity to *C. feundii* dihydroxyacetone kinase
YIL124w		similarity to *C. perfringens* nanH protein
YDR371w		similarity to chitinases
YIR036c		similarity to *E. coli* 3-ketoacyl-ACP reductase
YBR149w		similarity to Gcy1p and aldose reductases
YBR056w		similarity to glucan 1,3–glucosidase
YIL169c		similarity to glucan 1,4–glucosidase
YOL155c		similarity to glucan 1,4–glucosidase Mal5p
YNL274c		similarity to glycerate- and formate-dehydrogenases
YPL113c		similarity to glycerate dehydrogenases
YPR184w		similarity to glycogen debranching enzymes
YDL131w		similarity to homocitrate synthases and isopropylmalate synthases
YJR096w		similarity to *Leishmania* reductase
YKR096w		similarity to mitochondrial aldehyde dehydrogenase Ald1p
YGL257c		similarity to Mnn1p
YIL014w		similarity to Mnn1p
YEL020c		similarity to *O. formigenes* oxalyl-CoA decarboxylase
YDL093w	*PMT5*	similarity to O-mannosyltransferases Pmt1p-Pmt4p
YMR278w		similarity to phosphomannomutases
YDR307w		similarity to Pmt1p
YDR380w		similarity to pyruvate decarboxylases
YDR245w	*MNN10*	similarity to *S. pombe* galactosyltransferase
YLR070c		similarity to sugar dehydrogenases
YNR059w		similarity to to α-1,3-mannosyltransferase
YHL012w		similarity to UDP glucose pyrophosphorylase
YJR159w	*SOR1*	sorbitol dehydrogenase
YIL099w	*SGA1*	sporulation specific glucan 1,4–glucosidase
YER096w	*SHC1*	sporulation-specific protein
YLR307w	*CDA1*	sporulation-specific chitin deacetylase
YLR308w	*CDA2*	sporulation-specific chitin deacetylase
YNL331c		strong similarity to aryl-alcohol reductase
YCR107w		strong similarity to aryl-alcohol reductases
YPL017c		strong similarity to dihydrolipoamide dehydrogenases
YBR205w	*KTR3*	strong similarity to α-1,2-mannosyltransferase
YBR199w	*KTR4*	strong similarity to α-1,2-mannosyltransferase
YPL053c	*KTR6*	strong similarity to α-1,2-mannosyltransferase Kre2p
YJL221c	*FSP2*	strong similarity to α-D-glucosidase
YOL157c		strong similarity to α-glucosidases
YCR105w		strong similarity to alcohol dehydrogenases
YDL243c		strong similarity to aryl-alcohol dehydrogenase
YJR155w		strong similarity to aryl-alcohol dehydrogenase
YFL056c		strong similarity to aryl-alcohol dehydrogenases
YFL057c		strong similarity to aryl-alcohol dehydrogenases
YOL165c		strong similarity to aryl-alcohol dehydrogenases
YKR027w		strong similarity to Chs6p
YHR104w		strong similarity to D-xylose 1-dehydrogenase
YEL070w		strong similarity to *E. coli* D-mannonate oxidoreductase
YNR073c		strong similarity to *E. coli* D-mannonate oxidoreductase, identical to YEL070w
YJL218w		strong similarity to *E. coli* galactoside O-acetyltransferase
YDR248c		strong similarity to *E. coli* thermoresistant gluconokinase
YIL172c		strong similarity to Fsp2p
YDL037c		strong similarity to glucan 1,4–glucosidase
YDR516c		strong similarity to glucokinase
YOR299w		strong similarity to hypothetical protein YMR237w and similarity to Chs6p
YMR237w		strong similarity to hypothetical protein YOR299w and similarity to CHS6 protein
YNL009w		strong similarity to isocitrate dehydrogenase
YJL216c		strong similarity to Mal62p
YGR287c		strong similarity to maltase
YOR099w	*KTR1*	strong similarity to mannosyltransferases
YDR368w	*YPR1*	strong similarity to members of the aldo/keto reductase family
YDL021w	*GPM2*	strong similarity to phosphoglycerate mutase Gpm1p
YMR323w		strong similarity to phosphopyruvate hydratases
YJR153w		strong similarity to polygalacturonases
YDL080c		strong similarity to pyruvate decarboxylases
YOR347c		strong similarity to pyruvate kinase
YGR244c		strong similarity to rumen fungus β-succinyl-CoA synthetase
YKL029c		strong similarity to *S. pombe* malate oxireductase
YBL001c		strong similarity to *S. xylosus* glucose kinase
YLR164w		strong similarity to Sdh4p
YDL246c		strong similarity to Sor1p
YJL045w		strong similarity to succinate dehydrogenase flavoprotein
YGR043c		strong similarity to transaldolase
YNR071c		strong similarity to UDP-glucose 4-epimerase Gal10p
YDR384c		strong similarity to *Y. lipolytica GPR1* gene
YHR210c		strong similarity UDP-glucose-4-epimerase
YKL148c	*SDH1*	succinate dehydrogenase flavoprotein
YDR178w	*SDH4*	succinate dehydrogenase membrane anchor subunit for Sdh2p
YIL162w	*SUC2*	sucrose hydrolyzing enzyme
YJR075w	*HOC1*	suppressor of *pkc1*
YLR354c	*TAL1*	transaldolase
YPR074c	*TKL1*	transketolase 1
YBR117c	*TKL2*	transketolase 2
YDR050c	*TPI1*	triose-phosphate isomerase
YBR019c	*GAL10*	UDP-glucose 4-epimerase
YBR018c	*GAL7*	UDP-glucose-hexose-1-phosphate uridylyltransferase
YFR015c	*GSY1*	UDP-glucose-starch glucosyltransferase, isoform 1
YLR258w	*GSY2*	UDP-glucose-starch glucosyltransferase, isoform 2
YBR243c	*ALG7*	UDP-N-acetylglucosamine-1-phosphate transferase
YKL035w	*UGP1*	UTP-glucose-1-phosphate uridylyltransferase
YKR043c		weak similarity to phosphoglycerate mutase
YOR283w		weak similarity to phosphoglycerate mutases

regulation of carbohydrate utilization

YGR240c	*PFK1*	6-phosphofructokinase, α subunit
YMR205c	*PFK2*	6-phosphofructokinase, β subunit
YIL107c	*PFK26*	6-phosphofructose-2-kinase, isoenzyme 1
YOL136c	*PFK27*	6-phosphofructose-2-kinase, isoenzyme 2
YBR126c	*TPS1*	α,α-trehalose-phosphate synthase, 56K subunit
YHR047c	*AAP1*	alanine/arginine aminopeptidase
YGR229c	*KNR4*	β-1,3-glucan synthesis protein
YOL067c	*RTG1*	basic helix-loop-helix transcription factor
YOR344c	*TYE7*	basic helix-loop-helix transcription factor
YGR166w	*KRE11*	β-glucan synthesis-associated protein
YBL103c	*RTG3*	bHLH/zip transcription factor
YPL161c	*BEM4*	bud emergence protein
YGL209w	*MIG2*	C2H2 zinc-finger protein
YIL033c	*SRA1*	cAMP dependent protein kinase, regulatory subunit
YDR477w	*SNF1*	carbon catabolite derepressing ser/thr protein kinase
YCR002c	*CDC10*	cell division control protein
YLR330w	*CHS5*	chitin synthesis protein
YDR073w	*SNF11*	component of SWI/SNF transcription activator complex
YOR290c	*SNF2*	component of SWI/SNF transcription activator complex
YBR289w	*SNF5*	component of SWI/SNF transcription activator complex
YPL016w	*SWI1*	component of SWI/SNF transcription activator complex
YPL042c	*SSN3*	cyclin-dependent ser/thr protein kinase
YNL216w	*RAP1*	DNA-binding protein with repressor and activator activity
YOL051w	*GAL11*	DNA-directed RNA polymerase II holoenzyme and Kornberg's mediator (SRB) subcomplex subunit
YDR443c	*SCA1*	DNA-directed RNA polymerase II holoenzyme and Kornberg's mediator (SRB) subcomplex subunit
YCR081w	*SRB8*	DNA-directed RNA polymerase II holoenzyme and Kornberg's mediator (SRB) subcomplex subunit
YNL025c	*SSN8*	DNA-directed RNA polymerase II holoenzyme and Kornberg's mediator (SRB) subcomplex subunit, cyclin C homologue
YLR071c	*RGR1*	DNA-directed RNA polymerase II holoenzyme subunit
YBR253w	*SRB6*	DNA-directed RNA polymerase II suppressor protein
YOR047c	*STD1*	dosage-dependent modulator of glucose repression
YJL155c	*FBP26*	fructose-2,6-bisphosphatase
YPL037c	*EGD1*	GAL4 DNA-binding enhancer protein
YDR009w	*GAL3*	galactokinase
YLL016w	*SDC25*	GDP/GTP exchange factor
YLR310c	*CDC25*	GDP/GTP exchange factor for Ras1p and Ras2p
YGR070w	*ROM1*	GDP/GTP exchange protein for Rho1p
YLR371w	*ROM2*	GDP/GTP exchange protein for Rho1p
YDR176w	*NGG1*	general transcriptional adaptor or co-activator
YER054c	*GIP2*	Glc7p-interacting protein
YNL236w	*SIN4*	global regulator protein
YHL025w	*SNF6*	global transcription activator
YER027c	*GAL83*	glucose repression protein
YNL199c	*GCR2*	glycolytic genes transcriptional activator
YOR101w	*RAS1*	GTP-binding protein
YNL098c	*RAS2*	GTP-binding protein
YPR165w	*RHO1*	GTP-binding protein of the RHO subfamily of RAS-like proteins
YNL090w	*RHO2*	GTP-binding protein of the RHO subfamily of RAS-like proteins
YDR420w	*HKR1*	*Hansenula* MrakII k9 killer toxin-resistance protein
YGL253w	*HXK2*	hexokinase II
YPL002c	*SNF8*	involved in glucose derepression
YOR125c	*CAT5*	involved in glucose repression
YML048w	*EFF2*	involved in glucose repression
YNL201c		involved in regulation of carbon metabolism
YGR227w	*DIE2*	ITR1 expression promoting protein
YBR297w	*MAL33*	maltose fermentation regulatory protein
YGR288w	*MAL13*	maltose pathway regulatory protein
YLR131c	*ACE2*	metallothionein expression activator
YDR422c	*SIP1*	multicopy suppressor of *snf1*
YML051w	*GAL80*	negative regulator for expression of galactose-induced genes
YNL076w	*MKS1*	negative regulator of RAS-cAMP pathway
YIL119c	*RPI1*	negative regulator of RAS-cAMP pathway
YLR025w	*SNF7*	nuclear protein
YGL115w	*SNF4*	nuclear regulatory protein
YLR273c	*PIG1*	protein interacting with Gsy2p
YIL045w	*PIG2*	protein Interacting with Gsy2p
YDL006w	*PTC1*	protein ser/thr phosphatase 2c
YDL134c	*PPH21*	protein ser/thr phosphatase PP2A-1
YDL188c	*PPH22*	protein ser/thr phosphatase PP2A-2
YBL061c	*SKT5*	protoplast regeneration and killer toxin resistance protein
YOL110w	*SHR5*	RAS suppressor
YJR095w	*ACR1*	regulator of acetyl-CoA synthetase activity
YKL038w	*RGT1*	regulator of glucose induced genes
YDR028c	*REG1*	regulatory subunit for protein phosphatase Glc7p
YMR311c	*GLC8*	regulatory subunit for protein ser/thr phosphatase Glc7p
YBR050c	*REG2*	regulatory subunit of type I protein phosphatase
YDR277c	*MTH1*	repressor of hexose transport genes
YNR052c	*POP2*	required for glucose derepression
YJR090c	*GRR1*	required for glucose repression and for glucose and cation transport
YGL252c	*RTG2*	retrograde regulation protein
YJR076c	*CDC11*	septin
YHR107c	*CDC12*	septin
YER133w	*GLC7*	ser/thr phosphoprotein phosphatase 1, catalytic subunit
YOR178c	*GAC1*	ser/thr phosphoprotein phosphatase 1, regulatory subunit
YLR113w	*HOG1*	ser/thr protein kinase of MAP kinase (MAPK) family
YLR345w		similarity to 6-phosphofructo-2-kinase
YCR028c	*FEN2*	similarity to allantoate permease transporter
YFL053w		similarity to *C. freundii* dihydroxyacetone kinase
YDL225w		similarity to Cdc11p, Cdc3p and human CDC10 protein
YHR193c	*EGD2*	similarity to human α-NAC
YDL203c		similarity to Skt5p
YNL257c	*SIP3*	Snf1p protein kinase interacting protein
YLR150w	*MPT4*	specific affinity for guanine-rich quadruplex nucleic acids
YFL052w		strong similarity to Mal63p, Mal23p and Mal33p
YHR155w		strong similarity to Snf1p-interacting protein Sip3p
YNR002c	*FUN34*	strong similarity to *Y. lipolytica* glyoxylate pathway regulator GPR1
YCR010c		strong similarity to *Y. lipolytica* GPR1 protein and Fun34p
YIL154c	*(IMP2)*	sugar utilization regulatory protein
YPL026c	*SKS1*	suppressor kinase of *snf3*
YPL129w	*ANC1*	TFIIF subunit (transcription initiation factor), 30K
YPL248c	*GAL4*	transcription factor
YMR125w	*GCR3*	transcription factor for glycolytic genes

YMR280c	*CAT8*	transcription factor involved in gluconeogenesis
YMR043w	*MCM1*	transcription factor of the MADS box family
YJL176c	*SWI3*	transcription regulatory protein
YPL075w	*GCR1*	transcriptional activator
YOL116w	*MSN1*	transcriptional activator
YKL062w	*MSN4*	transcriptional activator
YGL035c	*MIG1*	transcriptional repressor
YIL147c	*SLN1*	two-component signal transducer
YLR006c	*SSK1*	two-component signal transducer
YJL128c	*PBS2*	tyrosine protein kinase of the MAP kinase kinase family
YLR258w	*GSY2*	UDP-glucose-starch glucosyltransferase, isoform 2
YDR216w	*ADR1*	zinc-finger transcription factor

carbohydrate transport

YCL040w	*GLK1*	aldohexose specific glucokinase
YKL217w	*JEN1*	carboxylic acid transporter protein
YBR291c	*CTP1*	citrate transport protein, mitochondrial (MCF)
YLR348c		dicarboxylate carrier protein
YLR081w	*GAL2*	galactose (and glucose) permease
YGR289c	*AGT1*	general α-glucoside permease
YLL043w	*FPS1*	glycerol channel protein
YNL318c	*HXT14*	hexose transport protein
YJL214w	*HXT8*	hexose transport protein
YJL219w	*HXT9*	hexose transport protein
YFL011w	*HXT10*	hexose transporter
YDL194w	*SNF3*	high-affinity glucose transporter
YEL069c	*HXT13*	high-affinity hexose transporter
YMR011w	*HXT2*	high-affinity hexose transporter
YDR343c	*HXT6*	high-affinity hexose transporter
YDR342c	*HXT7*	high-affinity hexose transporter
YOL156w	*HXT11*	low-affinity glucose transporter
YHR094c	*HXT1*	low-affinity hexose transporter
YDR345c	*HXT3*	low-affinity hexose transporter
YDR497c	*ITR1*	major myo-inositol permease
YBR298c	*MAL31*	maltose permease
YHR092c	*HXT4*	moderate- to low-affinity glucose transporter
YOL103w	*ITR2*	myo-inositol transport protein
YIR035c		similarity to *C. lanceolata* 3-oxoacyl-[acyl-carrier-protein] reductase
YBR241c		similarity to glucose transport proteins
YGL104c		similarity to glucose transport proteins
YPR021c		similarity to human citrate transporter protein
YDR387c		similarity to ltr1p and ltr2p
YDL199c		similarity to sugar transporter proteins
YFL040w		similarity to yeast glucose transport proteins
YHR096c	*HXT5*	strong similarity to hexose transporters
YDL245c	*HXT15*	strong similarity to Hxt17p and Haxt7p
YJR160c		strong similarity to Mal3Tp
YDR536w	*STL1*	strong similarity to members of the sugar permease family
YDL247w		strong similarity to sugar transport proteins
YIL170w	*HXT12*	strong similarity to sugar transport proteins
YIL171w	*HXT12*	strong similarity to sugar transport proteins
YJR158w	*HXT16*	strong similarity to sugar transport proteins
YNR072w	*HXT17*	sugar transport protein
YDL138w	*RGT2*	suppressor of *snf3* mutant

lipid, fatty-acid and sterol metabolism

lipid, fatty-acid and sterol biosynthesis

YML075c	*HMG1*	3-hydroxy-3-methylglutaryl-CoA reductase 1
YLR450w	*HMG2*	3-hydroxy-3-methylglutaryl-CoA reductase 2
YML126c	*HMGS*	3-hydroxy-3-methylglutaryl-CoA synthase
YPL028w	*ERG10*	acetyl-CoA C-acetyltransferase, cytosolic
YNR016c	*ACC1*	acetyl-CoA carboxylase
YCR048w	*ARE1*	acyl-CoA sterol acyltransferase
YNR019w	*ARE2*	acyl-CoA sterol acyltransferase
YIL009w	*FAA3*	acyl-CoA synthase
YGR037c	*ACB1*	acyl-CoA-binding protein (diazepam-binding inhibitor)
YBR222c	*PCS60*	AMP-binding protein, peroxisomal
YER061c	*CEM1*	β-keto-acyl-ACP synthase
YNL280c	*ERG24*	C-14 sterol reductase
YMR015c	*ERG5*	C-22 sterol desaturase
YGR060w	*ERG25*	C-4 sterol methyl oxidase
YLR056w	*ERG3*	C-5 sterol desaturase
YMR202w	*ERG2*	C-8 sterol isomerase
YML042w	*CAT2*	carnitine O-acetyltransferase
YPR113w	*PIS1*	CDP diacylglycerol-inositol 3-phosphatidyltransferase
YER026c	*CHO1*	CDP-diacylglycerol serine O-phosphatidyltransferase
YBR029c	*CDS1*	CDP-diacylglycerol synthase
YLR133w	*CKI1*	choline kinase
YGR007w	*MUQ1*	choline phosphate cytidylyltransferase
YGR202c	*PCT1*	cholinephosphate cytidylyltransferase
YNL111c	*CYB5*	cytochrome b5
YHR007c	*ERG11*	cytochrome P450 lanosterol 14a-demethylase
YHR190w	*ERG9*	farnesyl-diphosphate farnesyltransferase
YJL167w	*ERG20*	farnesyl-pyrophosphate synthetase
YDL052c	*SLC1*	fatty acyltransferase
YPL231w	*FAS2*	fatty-acyl-CoA synthase, α subunit
YKL182w	*FAS1*	fatty-acyl-CoA synthase, β subunit
YPL069c	*BTS1*	geranylgeranyl diphosphate synthase
YGL126w	*SCS3*	inositol phospholipid synthesis protein
YOR237w	*HES1*	involved in ergosterol biosynthesis
YPL145c	*KES1*	involved in ergosterol biosynthesis
YPL117c	*IDI1*	isopentenyl-diphosphate δ-isomerase
YHR072w	*ERG7*	lanosterol synthase
YMR246w	*FAA4*	long-chain-fatty-acid-CoA ligase
YNL219c	*ALG9*	mannosyltransferase
YJR073c	*OPI3*	methylene-fatty-acyl-phospholipid synthase
YNR043w	*MVD1*	mevalonate pyrophosphate decarboxylase
YPL076w	*GPI2*	N-acetylglucosaminyl-phosphatidylinositol biosynthetic protein
YPL175w	*SPT14*	N-acetylglucosaminyltransferase
YGR157w	*CHO2*	phosphatidylethanolamine N-methyltransferase
YDL142c	*PGS1*	phosphatidylglycerophosphate synthase
YNL169c	*PSD1*	phosphatidylserine decarboxylase 1
YGR170w	*PSD2*	phosphatidylserine decarboxylase 2
YCL004w	*PEL1*	phosphatidylserine synthase
YMR220w	*ERG8*	phosphomevalonate kinase
YCR034w	*GNS1*	probable 1,3—glucan synthase subunit
YGL001c		putative 3—hydroxysteroid dehydrogenase
YJL196c	*ELO1*	required for elongation of fatty acids
YGR216c	*GPI1*	required for N-acetylglucosaminyl phosphatidylinositol synthesis
YML008c	*ERG6*	S-adenosyl-methionine δ-24-sterol-c-methyltransferase
YMR296c	*LCB1*	serine C-palmitoyltransferase subunit
YDR062w	*LCB2*	serine C-palmitoyltransferase subunit
YLR058c	*SHM2*	serine hydroxymethyltransferase, cytoplasmic
YBR263w	*SHM1*	serine hydroxymethyltransferase, mitochondrial
YMR207c	*HFA1*	similarity to acetyl-CoA carboxylase
YDR147w		similarity to choline kinase
YBR159w		similarity to human 17—hydroxysteroid dehydrogenase
YDR376w	*ARH1*	similarity to human adrenodoxin reductase
YML131w		similarity to human leukotriene b4 12-hydroxydehydrogenase
YAR042w	*OSH1*	similarity to human oxysterol binding protein
YDR208w	*MSS4*	similarity to human PI P 5-kinase
YHR001w		similarity to Kes1p
YKR003w		similarity to Kes1p, Hes1p and Osh1p
YDL019c		similarity to Osh1p
YHR073w		similarity to Osh1p, Hes1p, Kes1p
YGR175c	*ERG1*	squalene monooxygenase
YGL055w	*OLE1*	stearoyl-CoA desaturase
YGL012w	*ERG4*	sterol C-24 reductase
YLR372w	*SUR4*	sterol isomerase
YKL192c		strong similarity to acyl-carrier proteins
YNL045w		strong similarity to human leukotriene-A4 hydrolase

breakdown of lipids and phospholipids

YPL268w	*PLC1*	1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase
YGL205w	*POX1*	acyl-CoA oxidase
YMR008c	*PLB1*	phospholipase B (lysophospholipase)
YKR031c	*SPO14*	phospholipase D
YLR020c		similarity to triacylglycerol lipase
YLL012w		similarity to triacylglycerol lipases
YJL068c		strong similarity to human esterase D
YOL011w		strong similarity to phospholipases
YMR006c		strong similarity to Plb1p
YKL140w	*TGL1*	triacylglycerol lipase
YDR058c	*TGL2*	triacylglycerol lipase
YJR107w		weak similarity to acylglycerol lipase
YBR204c		weak similarity to peroxisomal serine-active lipase

lipid, fatty-acid and sterol utilization

YIL160c	*POT1*	acetyl-CoA C-acyltransferase, peroxisomal
YIL009w	*FAA3*	acyl-CoA synthase
YMR013c	*SEC59*	dolichol kinase
YDR331w	*GPI8*	essential for GPI-anchor attachment
YDR410c	*STE14*	farnesyl cysteine carboxyl-methyltransferase
YPL172c	*COX10*	farnesyl transferase
YOR370c	*MSI4*	geranylgeranyltransferase regulatory subunit
YGL155w	*CDC43*	geranylgeranyltransferase type I β subunit
YPR176c	*BET2*	geranylgeranyltransferase type II β subunit
YJL031c	*BET4*	geranylgeranyltransferase, α subunit
YFL014w	*HSP12*	heat-shock protein
YBR003w	*COQ1*	hexaprenyl pyrophosphate synthetase
YKR009c	*FOX2*	hydratase-dehydrogenase-epimerase, peroxisomal
YOR317w	*FAA1*	long-chain-fatty-acid-CoA ligase
YER015w	*FAA2*	long-chain-fatty-acid-CoA ligase
YMR246w	*FAA4*	long-chain-fatty-acid-CoA ligase
YDL078c	*MDH3*	malate dehydrogenase, peroxisomal
YLR195c	*NMT1*	N-myristoyltransferase
YJR066w	*TOR1*	phosphatidylinositol 3-kinase
YKL019w	*RAM2*	protein farnesyltransferase, α subunit
YDL090c	*RAM1*	protein farnesyltransferase, β subunit
YHR013c	*ARD1*	protein N-acetyltransferase subunit
YDL040c	*NAT1*	protein N-acetyltransferase subunit

regulation of lipid, fatty-acid and sterol biosynthesis

YBL015w	*ACH1*	acetyl-CoA hydrolase
YDR123c	*INO2*	basic helix-loop-helix (BHLH) transcription factor
YOL108c	*INO4*	basic helix-loop-helix transcription factor
YGL162w	*SUT1*	hypoxic protein involved in sterol uptake
YMR208w	*ERG12*	mevalonate kinase
YHL020c	*OPI1*	negative regulator of phospholipid biosynthesis pathway
YDR207c	*UME6*	negative transcriptional regulator
YAL051w	*OAF1*	peroxisome proliferating transcription factor
YOR363c	*PIP2*	peroxisome proliferating transcription factor
YNL169c	*PSD1*	phosphatidylserine decarboxylase 1
YAL013w	*DEP1*	regulator of phospholipid metabolism
YMR207c	*HFA1*	similarity to acetyl-CoA carboxylase
YCR028c	*FEN2*	similarity to allantoate permease transporter
YMR043w	*MCM1*	transcription factor of the MADS box family
YOL004w	*SIN3*	transcription regulatory protein

lipid and fatty-acid transport

YGR037c	*ACB1*	acyl-CoA-binding protein (diazepam-binding inhibitor)
YNL130c	*CPT1*	diacylglycerol cholinephosphotransferase
YBR041w	*FAT1*	fatty-acid transporter
YKL188c	*PAT1*	long-chain-fatty-acid transporter
YPL147w	*PXA1*	long-chain-fatty-acid transporter
YOR317w	*FAA1*	long-chain-fatty-acid-CoA ligase
YER015w	*FAA2*	long-chain-fatty-acid-CoA ligase
YMR246w	*FAA4*	long-chain-fatty-acid-CoA ligase
YAR035w	*YAT1*	outer carnitine acetyltransferase, mitochondrial
YCL043c	*PDI1*	protein disulphide-isomerase
YER024w		similarity to carnitine O-acetyltransferase Yat1p
YKL174c		similarity to choline transport protein Ctr1p
YNR056c		similarity to choline transport protein Ctr1p
YHR123w	*EPT1*	sn-1,2-diacylglycerol ethanolamine- and cholinephosphotransferase

lipid and fatty-acid binding

YHR001w		similarity to Kes1p
YMR020w	*FMS1*	suppressor of fenpropimorph resistance mutation *fen2*

other lipid, fatty-acid and sterol metabolism activities

YBR036c	*CSG2*	calcium dependent regulatory protein
YMR079w	*SEC14*	phosphatidylinositol/phosphatidylcholine transfer protein
YNL264c		similarity to Sec14p
YKL091c		strong similarity to Sec14p
YIL002c	*SIH1*	synaptojanin homologue 1
YDR302w		weak similarity to human GPI-anchor biosynthesis protein

biosynthesis of vitamins, cofactors and prosthetic groups

metabolism of vitamins, cofactors and prosthetic groups

YOL096c	*COQ3*	3,4-dihydroxy-5-hexaprenylbenzoate methyltransferase
YDR232w	*HEM1*	5-aminolevulinate synthase
YOL143c	*RIB4*	6,7-dimethyl-8-ribityllumazine synthase
YGR286c	*BIO2*	biotin synthetase
YGR204w	*ADE3*	C1-tetrahydrofolate synthase, cytoplasmic
YBR084w	*MIS1*	C1-tetrahydrofolate synthase, mitochondrial
YDR044w	*HEM13*	coproporphyrinogen III oxidase
YPL132w	*COX11*	cytochrome c oxidase assembly protein
YNR058w	*BIO3*	DAPA aminotransferase
YDR487c	*RIB3*	DBP synthase
YOL066c	*RIB2*	DRAP deaminase
YPL172c	*COX10*	farnesyl transferase
YOR176w	*HEM15*	ferrochelatase
YDL045c	*FAD1*	flavin adenine dinucleotide (FAD) synthetase
YGR267c	*FOL2*	GTP cyclohydrolase I
YBL033c	*RIB1*	GTP cyclohydrolase II
YBR003w	*COQ1*	hexaprenyl pyrophosphate synthetase
YBR153w	*RIB7*	HTP reductase
YOR196c	*LIP5*	lipoic acid synthase
YHR042w	*NCP1*	NADPH-cytochrome P450 reductase
YOR209c	*NPT1*	nicotinate phosphoribosyltransferase
YNR033w	*ABZ1*	para-aminobenzoate synthase
YNR041c	*COQ2*	para-hydroxybenzoate-polyprenyltransferase
YOR184w	*SER1*	phosphoserine transaminase
YDL205c	*HEM3*	porphobilinogen deaminase
YGL040c	*HEM2*	porphobilinogen synthase
YER014w	*HEM14*	protoporphyrinogen oxidase, mitochondrial
YNR057c	*BIO4*	putative dethiobiotin synthetase
YBR035c	*PDX3*	pyridoxamine-phosphate oxidase
YBR256c	*RIB5*	riboflavin synthase, α subunit
YER043c	*SAH1*	S-adenosyl-L-homocysteine hydrolase
YNL256w		similarity to bacterial dihydropteroate synthase
YKL027w		similarity to *E. coli* molybdopterin-converting factor chlN
YIL145c		similarity to *E. coli* pantothenate synthetase
YGR255c		similarity to *E. coli* ubiH protein
YMR113w		similarity to folylpolyglutamate synthetases and strong similarity to hypothetical protein YKL132c
YGL125w		similarity to human methylenetetrahydrofolate reductase

YPL023c		similarity to human methylenetetrahydrofolate reductase
YFR047c		similarity to human quinolinate phosphoribosyltransferase
YHR111w		similarity to molybdopterin biosynthesis proteins
YOL151w		similarity to plant dihydroflavonol-4-reductases
YMR222c		similarity to *S. pombe* dihydrofolate reductase
YOR241w		similarity to tetrahydrofolylpolyglutamate synthase
YJR142w		similarity to thiamin pyrophosphokinase
YKR069w		similarity to uroporphyrinogen methyltransferases
YGL039w		similarity to *V. vinifera* dihydroflavonol reductase
YBR176w		strong similarity to *E. coli* 3-methyl-2-oxobutanoate hydroxymethyltransferase
YHR003c		strong similarity to molybdopterin-converting factor homologue YKL027w
YDL036c		strong similarity to Rib2p
YOL049w		strong similarity to *S. pombe* Gsa1p
YOR143c	*THI80*	thiamin pyrophosphokinase
YPL214c	*THI6*	thiamin-phosphate pyrophosphorylase and hydroxyethylthiazole kinase
YDR047w	*HEM12*	uroporphyrinogen decarboxylase
YOR278w	*HEM4*	uroporphyrinogen III synthase

utilization of vitamins, cofactors and prosthetic groups

YDL141w	*BPL1*	biotin holocarboxylase synthetase
YBR092c	*PHO3*	constitutive acid phosphatase
YAL039c	*CYC3*	holocytochrome c synthase (cytochrome c haem lyase)
YKL087c	*CYT2*	holocytochrome c1 synthase
YML120c	*NDI1*	NADH-ubiquinone-6 oxidoreductase
YBR093c	*PHO5*	repressible acid phosphatase
YKL132c		similarity to *B. subtilis* folC protein and strong similarity to hypothetical protein YMR113w

regulation of vitamins, cofactors and prosthetic groups

YPR065w	*ROX1*	haem-dependent transcriptional repressor of hypoxic genes
YOR125c	*CAT5*	involved in glucose repression

transport of vitamins, cofactors and prosthetic groups

YIL134w	*FLX1*	FAD carrier protein, mitochondrial (MCF)
YIL006w		similarity to Flx1p

other vitamin, cofactor and prosthetic-group activities

YDR541c		similarity to dihydroflavonol-4-reductases
YGL157w		similarity to *E. gunnii* cinnamyl alcohol dehydrogenase
YER183c		similarity to formyltetrahydrofolate cyclo-ligase
YHL018w		similarity to human pterin-4–-carbinolamine dehydratase
YGR169c		similarity to Rib2p
YGR144w	*THI4*	thiamine-repressed protein

ionic homeostasis

homeostasis of metal ions

YNL259c	*ATX1*	antioxidant protein and metal homeostasis factor
YJR049c	*UTR1*	associated with ferric reductase activity
YMR058w	*FET3*	cell-surface ferroxidase
YOR316c	*COT1*	cobalt accumulation protein
YPL177c	*CUP9*	copper homeostasis protein
YDR515w	*SLF1*	copper homeostasis protein
YPR124w	*CTR1*	copper transport protein
YHR175w	*CTR2*	copper transport protein
YLR411w	*CTR3*	copper transport protein
YGL166w	*CUP2*	copper-dependent transcription factor
YLR214w	*FRE1*	ferric (and cupric) reductase
YKL220c	*FRE2*	ferric (and cupric) reductase
YOR176w	*HEM15*	ferrochelatase
YGL255w	*ZRT1*	high-affinity zinc-transport protein
YLL009c	*COX17*	interacts genetically with *SCO1* and *SCO2* in cytochrome oxidase assembly
YBR037c	*SCO1*	involved in stabilization of Cox1p and Cox2p
YGL071w	*RCS1*	iron-regulated transcriptional repressor
YLR130c	*ZRT2*	low-affininty zinc transporter
YMR319c	*FET4*	low-affinity Fe(II) iron transport protein
YOL122c	*SMF1*	manganese transporter
YBR290w	*BSD2*	metal homeostasis protein
YKL064w	*MNR2*	overexpression overcomes manganese toxicity
YDR270w	*CCC2*	probable copper-transporting ATPase
YBR295w	*PCA1*	P-type Cu^{2+}-transporting ATPase
YGL096w		similarity to copper homeostasis protein Cup9p
YDR506c		similarity to Fet3p
YOL152w		similarity to Fre1p and Fre2p
YGL160w		similarity to hypothetical protein YLR047c and Fre2p
YJR126c		similarity to transferrin receptor protein
YFL041w		strong similarity to cell-surface ferroxidase Fet3p
YLL051c		strong similarity to ferric reductase Fre2p
YOR381w		strong similarity to ferric reductase Fre2p
YOR384w		strong similarity to ferric reductase Fre2p
YNR060w		strong similarity to Fre2p and hypothetical protein YOR381w, and similarity to Fre1p
YBR024w	*SCO2*	strong similarity to Sco1p
YOR031w	*CRS5*	suppressor of *cup1* deletion, metallothionein-like protein
YCL037c	*SRO9*	suppressor of *rho3*
YMR243c	*ZRC1*	zinc- and cadmium-resistance protein

homeostasis of other ions

YDL128w	*VCX1*	Ca^{2+}-transport (H^+/Ca^{2+} exchange) protein, vacuolar
YGL006w	*PMC1*	Ca^{2+}-transporting P-type ATPase
YGL167c	*PMR1*	Ca^{2+}-transporting P-type ATPase
YGR217w	*CCH1*	calcium channel protein
YLR348c		dicarboxylate carrier protein
YKL016c	*ATP7*	F1F0-ATPase complex, F0 D subunit
YBL099w	*ATP1*	F1F0-ATPase complex, F1 α subunit
YJR121w	*ATP2*	F1F0-ATPase complex, F1 β subunit
YDL004w	*ATP16*	F1F0-ATPase complex, F1 δ subunit
YPL078c	*ATP4*	F1F0-ATPase complex, F1 δ subunit
YPL271w	*ATP15*	F1F0-ATPase complex, F1 ε subunit
YBR039w	*ATP3*	F1F0-ATPase complex, F1 γ subunit
YDR298c	*ATP5*	F1F0-ATPase complex, OSCP subunit
YLR295c	*ATP14*	F1F0-ATPase complex, subunit h
YCR024c-a	*PMP1*	H^+-ATPase subunit, plasma membrane
YEL017c-a	*PMP2*	H^+-ATPase subunit, plasma membrane
YMR054w	*STV1*	H^+-ATPase V0 domain 102K subunit, vacuolar
YHR039c-a	*VMA10*	H^+-ATPase V0 domain 13K subunit, vacuolar
YEL027w	*CUP5*	H^+-ATPase V0 domain 17K subunit, vacuolar
YPL234c	*TFP3*	H^+-ATPase V0 domain 17K subunit, vacuolar
YLR447c	*VMA6*	H^+-ATPase V0 domain 36K subunit, vacuolar
YOR270c	*VPH1*	H^+-ATPase V0 domain 95K subunit, vacuolar
YGR020c	*VMA7*	H^+-ATPase V1 domain 14K subunit, vacuolar
YOR332w	*VMA4*	H^+-ATPase V1 domain 27K subunit, vacuolar
YEL051w	*VMA8*	H^+-ATPase V1 domain 32K subunit, vacuolar
YKL080w	*VMA5*	H^+-ATPase V1 domain 42K subunit, vacuolar
YPR036w	*VMA13*	H^+-ATPase V1 domain 54K subunit, vacuolar
YBR127c	*VMA2*	H^+-ATPase V1 domain 60K subunit, vacuolar
YDL185w	*TFP1*	H^+-ATPase V1 domain 69K subunit, vacuolar
YGL008c	*PMA1*	H^+-transporting P-type ATPase
YPL036w	*PMA2*	H^+-transporting P-type ATPase 2
YML123c	*PHO84*	high affinity inorganic phosphate/H^+ symporter
YJL129c	*TRK1*	high-affinity potassium transport protein
YBR294w	*SUL1*	high-affinity sulphate transport protein
YJL117w	*PHO86*	inorganic phospate transporter
YLR220w	*CCC1*	involved in calcium regulation
YCR037c	*PHO87*	member of the phosphate permease family
YJL093c	*TOK1*	outward-rectifier potassium channel
YJR077c	*MIR1*	phosphate transport protein, mitochondrial (MCF)
YEL031w	*SPF1*	P-type ATPase
YDR040c	*ENA1*	P-type ATPase involved in Na^+ and Li^+ efflux
YDR039c	*ENA2*	P-type ATPase involved in Na^+ efflux
YDR038c	*ENA5*	P-type ATPase involved in Na^+ efflux
YLR138w	*NHA1*	putative Na^+/H^+ antiporter
YBR235w		similarity to bumetanide-sensitive Na^+-K^+-Cl^- cotransport protein
YJL094c		similarity to *E. hirae* Na^+/H^+-antiporter NapA
YNR013c		similarity to membrane protein Pho87p and hypothetical protein YJL198w
YDR456w		similarity to NA^+/H^+ antiporters
YPR003c		similarity to sulphate transporter proteins
YER053c		strong similarity to mitochondrial phosphate carrier protein
YJL198w		strong similarity to Pho87p
YBR296c		strong similarity to phosphate-repressible phosphate permease
YLR092w	*SEL2*	strong similarity to Sul1p
YHL016c	*DUR3*	urea transport protein
YJR040w	*GEF1*	voltage-gated chloride channel protein

Energy

glycolysis

YGR240c	*PFK1*	6-phosphofructokinase, α subunit
YMR205c	*PFK2*	6-phosphofructokinase, β subunit
YIL107c	*PFK26*	6-phosphofructose-2-kinase, isoenzyme 1
YOL136c	*PFK27*	6-phosphofructose-2-kinase, isoenzyme 2
YNL071w	*LAT1*	dihydrolipoamide S-acetyltransferase
YGR254w	*ENO1*	enolase I (2-phosphoglycerate dehydratase)
YHR174w	*ENO2*	enolase II (2-phosphoglycerate dehydratase)
YOR393w	*ERR1*	enolase-related protein
YPL281c	*ERR2*	enolase-related protein
YKL060c	*FBA1*	fructose-bisphosphate aldolase
YBR196c	*PGI1*	glucose-6-phosphate isomerase
YJL052w	*TDH1*	glyceraldehyde-3-phosphate dehydrogenase 1
YJR009c	*TDH2*	glyceraldehyde-3-phosphate dehydrogenase 2
YGR192c	*TDH3*	glyceraldehyde-3-phosphate dehydrogenase 3
YOL059w	*GPD3*	glycerol-3-phosphate dehydrogenase (NAD^+), mitochondrial
YFR053c	*HXK1*	hexokinase I
YGL253w	*HXK2*	hexokinase II
YCR012w	*PGK1*	phosphoglycerate kinase
YKL152c	*GPM1*	phosphoglycerate mutase
YOL056w	*GPM3*	phosphoglycerate mutase
YER178w	*PDA1*	pyruvate dehydrogenase (lipoamide) α subunit
YBR221c	*PDB1*	pyruvate dehydrogenase (lipoamide) β subunit
YGR193c	*PDX1*	pyruvate dehydrogenase complex protein X
YAL038w	*CDC19*	pyruvate kinase
YLR345w		similarity to 6-phosphofructo-2-kinase
YDL021w	*GPM2*	strong similarity to phosphoglycerate mutase Gpm1p
YMR323w		strong similarity to phosphopyruvate hydratases
YOR347c		strong similarity to pyruvate kinase
YKL029c		strong similarity to *S. pombe* malate oxireductase
YMR125w	*GCR3*	transcription factor for glycolytic genes
YDR050c	*TPI1*	triose-phosphate isomerase
YKR043c		weak similarity to phosphoglycerate mutase
YOR283w		weak similarity to phosphoglycerate mutases

gluconeogenesis

YOR393w	*ERR1*	enolase-related protein
YPL281c	*ERR2*	enolase-related protein
YLR377c	*FBP1*	fructose-1,6-bisphosphatase
YBR196c	*PGI1*	glucose-6-phosphate isomerase
YER065c	*ICL1*	isocitrate lyase
YKR097w	*PCK1*	phosphoenolpyruvate carboxykinase
YCR012w	*PGK1*	phosphoglycerate kinase
YGL062w	*PYC1*	pyruvate carboxylase 1
YBR218c	*PYC2*	pyruvate carboxylase 2
YIL167w	*SDL1*	serine dehydratase
YOR347c		strong similarity to pyruvate kinase
YMR280c	*CAT8*	transcription factor involved in gluconeogenesis
YDR050c	*TPI1*	triose-phosphate isomerase

pentose-phosphate pathway

YHR183w	*GND1*	6-phosphogluconate dehydrogenase
YOR095c	*RKI1*	D-ribose-5-phosphate ketol-isomerase
YJL121c	*POS18*	D-ribulose-5-phosphate 3-epimerase
YNL241c	*ZWF1*	glucose-6-phosphate dehydrogenase
YGR256w	*GND2*	phosphogluconate dehydrogenase
YGR043c		strong similarity to transaldolase
YLR354c	*TAL1*	transaldolase
YPR074c	*TKL1*	transketolase 1
YBR117c	*TKL2*	transketolase 2

tricarboxylic-acid pathway

YIL125w	*KGD1*	2-oxoglutarate dehydrogenase complex E1 component
YDR148c	*KGD2*	2-oxoglutarate dehydrogenase complex E2 component
YLR304c	*ACO1*	aconitate hydratase
YNR001c	*CIT1*	citrate (si)-synthase, mitochondrial
YPR001w	*CIT3*	citrate (si)-synthase, mitochondrial
YFL018c	*LPD1*	dihydrolipoamide dehydrogenase
YPL262w	*FUM1*	fumarate hydratase
YNL009w		homology to isocitrate dehydrogenase
YNL037c	*IDH1*	isocitrate dehydrogenase (NAD^+) subunit 1, mitochondrial
YOR136w	*IDH2*	isocitrate dehydrogenase (NAD^+) subunit 2, mitochondrial
YLR174w	*IDP2*	isocitrate dehydrogenase, cytosolic
YOL126c	*MDH2*	malate dehydrogenase, cytoplasmic
YOR297c		similarity to SDH4 protein
YJL200c		strong similarity to aconitate hydratase
YGR244c		strong similarity to rumen fungus β-succinyl-CoA synthetase
YLR164w		strong similarity to Sdh4p
YMR118c		strong similarity to succinate dehydrogenase
YJL045w		strong similarity to succinate dehydrogenase flavoprotein
YOR142w		strong similarity to succinate-CoA ligase α subunit
YKL148c	*SDH1*	succinate dehydrogenase flavoprotein
YLL041c	*SDH2*	succinate dehydrogenase iron-sulphur protein subunit
YDR178w	*SDH4*	succinate dehydrogenase membrane anchor subunit for Sdh2p

respiration

YMR282c	*AEP2*	2'-O-ribosyl phosphate transferase
YGR008c	*STF2*	ATPase stabilizing factor
YDL130w-a	*STF1*	ATPase stabilizing factor, 10K
YDR377w	*ATP17*	ATPase synthase subunit f
YER061c	*CEM1*	β-keto-acyl-ACP synthase
YIL043c	*CBR1*	cytochrome b5 reductase
YKL150w	*MCR1*	cytochrome b5 reductase
YKL141w	*SDH3*	cytochrome b560 subunit of respiratory complex II
YJR048w	*CYC1*	cytochrome c isoform 1
YEL039c	*CYC7*	cytochrome c isoform 2

YML129c	*COX14*	cytochrome c oxidase assembly protein
YDR079w	*PET100*	cytochrome c oxidase assembly protein
YGL187c	*COX4*	cytochrome c oxidase subunit IV
YNL052w	*COX5A*	cytochrome c oxidase subunit V.A
YIL111w	*COX5B*	cytochrome c oxidase subunit Vb
YHR051w	*COX6*	cytochrome c oxidase subunit VI
YGL191w	*COX13*	cytochrome c oxidase subunit VIa
YLR038c	*COX12*	cytochrome c oxidase subunit VIB
YMR256c	*COX7*	cytochrome c oxidase subunit VII
YDL067c	*COX9*	cytochrome c oxidase subunit VIIA
YLR395c	*COX8*	cytochrome c oxidase subunit VIII
YOR065w	*CYT1*	cytochrome c1
YER154w	*OXA1*	cytochrome oxidase biogenesis protein
YGR207c	*ETF-β*	electron-transferring flavoprotein, β subunit
YKL016c	*ATP7*	F1F0-ATPase complex, F0 D subunit
YBL099w	*ATP1*	F1F0-ATPase complex, F1 α subunit
YJR121w	*ATP2*	F1F0-ATPase complex, F1 β subunit
YDL004w	*ATP16*	F1F0-ATPase complex, F1 δ subunit
YPL078c	*ATP4*	F1F0-ATPase complex, F1 δ subunit
YPL271w	*ATP15*	F1F0-ATPase complex, F1 ε subunit
YBR039w	*ATP3*	F1F0-ATPase complex, F1 γ subunit
YDR298c	*ATP5*	F1F0-ATPase complex, OSCP subunit
YLR295c	*ATP14*	F1F0-ATPase complex, subunit h
YEL053c	*MAK10*	glucose-repressible protein
YLL009c	*COX17*	interacts genetically with *SCO1* and *SCO2* in cytochrome oxidase assembly
YJR034w	*PET191*	involved in assembly of cytochrome oxidase
YMR165c	*SMP2*	involved in plasmid maintenance, respiration and cell proliferation
YML054c	*CYB2*	lactate dehydrogenase cytochrome b2
YBR192w	*RIM2*	mitochondrial carrier protein (MCF)
YMR089c	*YTA12*	protease of the *SEC18*/*CDC48*/*PAS1* family of ATPases (AAA)
YLR044c	*PDC1*	pyruvate decarboxylase, isozyme 1
YGR062c	*COX18*	required for activity of mitochondrial cytochrome oxidase
YBR185c	*MBA1*	respiratory chain assembly protein
YDL107w	*MSS2*	ser/thr protein kinase
YPR048w		similarity to *M. domestica* NADPH-ferrihemoprotein reductase and mammalian nitric-oxide synthases
YGL226w		similarity to *N. crassa* cytochrome c oxidase subunit V
YMR244w		similarity to NCA3 and SUN4 protein
YML087c		strong similarity to cytochrome b5- and nitrate reductases
YML125c		strong similarity to cytochrome b5- and nitrate reductases
YPR004c		strong similarity to electron transfer flavoproteins α subunit
YOR356w		strong similarity to human electron transfer flavoprotein-ubiquinone oxidoreductase
YDL085w		strong similarity to NADH dehydrogenase (ubiquinone)
YMR145c		strong similarity to NADH dehydrogenase (ubiquinone)
YDL080c		strong similarity to pyruvate decarboxylases
YDR178w	*SDH4*	succinate dehydrogenase membrane anchor subunit for Sdh2p
YNL118c	*PSU1*	suppressor protein of a yeast *pet* mutant
YFR033c	*QCR6*	ubiquinol-cytochrome c reductase 17K protein
YPR191w	*QCR2*	ubiquinol-cytochrome c reductase 40K subunit II
YBL045c	*COR1*	ubiquinol-cytochrome c reductase 44K core protein
YHR001w-a	*QCR10*	ubiquinol-cytochrome c reductase 8.5K subunit
YGR174c	*CBP4*	ubiquinol-cytochrome c reductase assembly factor
YGL119w	*ABC1*	ubiquinol-cytochrome c reductase complex assembly protein
YEL024w	*RIP1*	ubiquinol-cytochrome c reductase iron-sulphur protein
YDR529c	*QCR7*	ubiquinol-cytochrome c reductase subunit 7
YGR183c	*QCR9*	ubiquinol-cytochrome c reductase subunit 9
YJL166w	*QCR8*	ubiquinol-cytochrome c reductase subunit VIII
YMR073c		weak similarity to C-terminal part of cytochrome b5 and b2
YNL237w	*YTP1*	weak similarity to mitochondrial electron transport proteins

fermentation

YOL086c	*ADH1*	alcohol dehydrogenase I
YMR303c	*ADH2*	alcohol dehydrogenase II
YMR083w	*ADH3*	alcohol dehydrogenase III
YGL256w	*ADH4*	alcohol dehydrogenase IV
YBR145w	*ADH5*	alcohol dehydrogenase V
YER073w		aldehyde dehydrogenase (NAD^+)
YPL061w	*ALD6*	aldehyde dehydrogenase, cytosolic
YDL174c	*DLD1*	D-lactate ferricytochrome c oxidoreductase (D-LCR)
YML004c	*GLO1*	glyoxalase I
YOR126c	*EST2*	isoamyl acetate hydrolytic enzyme
YPL275w		putative formate dehydrogenase/putative pseudogene
YPL276w		putative formate dehydrogenase/putative pseudogene
YDR081c	*PDC2*	pyruvate decarboxylase regulatory protein
YLR044c	*PDC1*	pyruvate decarboxylase, isozyme 1
YLR134w	*PDC5*	pyruvate decarboxylase, isozyme 2
YGR087c	*PDC6*	pyruvate decarboxylase, isozyme3
YAL060w	*FUN49*	similarity to alcohol/sorbitol dehydrogenase
YAL061w	*FUN50*	similarity to alcohol/sorbitol dehydrogenase
YMR318c		similarity to alcohol-dehydrogenase
YHR039c		similarity to aldehyde dehydrogenases
YPL088w		similarity to aryl-alcohol dehydrogenases
YMR285c		similarity to Ccr4p
YHL008c		similarity to *M. formicicum* formate dehydrogenase
YKR096w		similarity to mitochondrial aldehyde dehydrogenase Ald1p
YDR380w		similarity to pyruvate decarboxylases
YNL331c		strong similarity aryl-alcohol reductase
YCR107w		strong similarity aryl-alcohol reductases
YCR105w		strong similarity to alcohol dehydrogenases
YMR169c	*ALD4*	strong similarity to aldehyde dehydrogenase
YDL243c		strong similarity to aryl-alcohol dehydrogenase
YJR155w		strong similarity to aryl-alcohol dehydrogenase
YFL056c		strong similarity to aryl-alcohol dehydrogenases
YFL057c		strong similarity to aryl-alcohol dehydrogenases
YOL165c		strong similarity to aryl-alcohol dehydrogenases
YOR388c		strong similarity to *H. polymorpha* formate dehydrogenase

metabolism of energy reserves (glycogen and trehalose)

YEL011w	*GLC3*	1,4-glucan branching enzyme (glycogen branching enzyme)
YBR001c	*NTH2*	α,α-trehalase
YDR074w	*TPS2*	α,α-trehalose-phosphate synthase, 105K subunit
YMR261c	*TPS3*	α,α-trehalose-phosphate synthase, 115K subunit
YML100w	*TSL1*	α,α-trehalose-phosphate synthase, 123K subunit
YBR126c	*TPS1*	α,α-trehalose-phosphate synthase, 56K subunit
YPR026w	*ATH1*	acid trehalase, vacuolar
YBR299w	*MAL32*	α-glucosidase
YHR047c	*AAP1*	alanine/arginine aminopeptidase
YPL031c	*PHO85*	cyclin-dependent protein kinase
YLR071c	*RGR1*	DNA-directed RNA polymerase II holoenzyme subunit
YER054c	*GIP2*	Glc7p-interacting protein
YBR229c	*ROT2*	glucosidase II, catalytic subunit
YPR160w	*GPH1*	glycogen phosphorylase
YPL240c	*HSP82*	heat-shock protein
YKL128c	*PMU1*	high copy suppressor of ts *tps2* mutant phenotype
YDR001c	*NTH1*	neutral trehalase (α,α-trehalase)
YMR105c	*PGM2*	phosphoglucomutase, major isoform
YKL127w	*PGM1*	phosphoglucomutase, minor isoform
YNR032w	*PPG1*	phosphoprotein phosphatase PPG catalytic subunit
YBL058w	*SHP1*	potential regulatory subunit for Glc7p
YLR273c	*PIG1*	protein interacting with Gsy2p
YIL045w	*PIG2*	protein interacting with Gsy2p
YJL137c	*GLG2*	self-glucosylating initiator of glycogen synthesis
YKR058w	*GLG1*	self-glycosylating initiator of glycogen synthesis
YER133w	*GLC7*	ser/thr phosphoprotein phosphatase 1, catalytic subunit
YOR178c	*GAC1*	ser/thr phosphoprotein phosphatase 1, regulatory subunit
YPR184w		similarity to glycogen debranching enzymes
YJL221c	*FSP2*	strong similarity to α-D-glucosidase
YOL157c		strong similarity to α-glucosidases
YIL172c		strong similarity to Fsp2p
YJL216c		strong similarity to Mal62p
YFR015c	*GSY1*	UDP glucose-starch glucosyltransferase, isoform 1
YLR258w	*GSY2*	UDP-glucose-starch glucosyltransferase, isoform 2
YKL035w	*UGP1*	UTP-glucose-1-phosphate uridylyltransferase

other energy-generation activities

YIL160c	*POT1*	acetyl-CoA C-acyltransferase, peroxisomal
YGL205w	*POX1*	acyl-CoA oxidase
YMR170c	*ALD5*	aldehyde dehydrogenase 2 (NAD^+), mitochondrial
YML035c	*AMD1*	AMP deaminase
YCR005c	*CIT2*	citrate (si)-synthase, peroxisomal
YKR009c	*FOX2*	hydratase-dehydrogenase-epimerase, peroxisomal
YMR267w	*PPA2*	inorganic pyrophosphatase, mitochondrial
YDL066w	*IDP1*	isocitrate dehydrogenase ($NADP^+$), mitochondrial
YER065c	*ICL1*	isocitrate lyase
YKL085w	*MDH1*	malate dehydrogenase, mitochondrial
YDL078c	*MDH3*	malate dehydrogenase, peroxisomal
YNL117w	*MLS1*	malate synthase 1
YIR031c	*DAL7*	malate synthase 2
YHR179w	*OYE2*	NAPDH dehydrogenase (old yellow enzyme), isoform 1
YPL171c	*OYE3*	NAPDH dehydrogenase (old yellow enzyme), isoform 3
YPR006c	*ICL2*	non-functional isocitrate lyase
YOL079w		similarity to NADH dehydrogenase
YEL020c		similarity to *O. formigenes* oxalyl-CoA decarboxylase
YDR384c		strong similarity to *Y. lipolytica GPR1* gene

Cell growth, cell division and DNA synthesis

cell growth

YPL268w	*PLC1*	1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase
YCR088w	*ABP1*	actin-binding protein
YDL029w	*ACT2*	actin-like protein
YOL052c	*SPE2*	adenosylmethionine decarboxylase
YNL138w	*SRV2*	adenylate cyclase-associated protein, 70K
YER170w	*ADK2*	adenylate kinase, mitochondrial
YBR109c	*CMD1*	calmodulin
YHR135c	*YCK1*	casein kinase I isoform
YNL154c	*YCK2*	casein kinase I isoform
YJL174w	*KRE9*	cell-wall synthesis protein
YNL327w	*EGT2*	cell-cycle regulation protein
YLR175w	*CBF5*	centromere/microtubule binding protein
YGR167w	*CLC1*	clathrin light chain
YDR251w	*PAM1*	coiled-coil protein multicopy suppressor of loss of PP2A
YBL007c	*SLA1*	cytoskeleton assembly control protein
YFL001w	*DEG1*	depressed growth-rate protein
YLR300w	*EXG1*	exo-1,3--glucanase (I/II), major isoform
YDR261c	*EXG2*	exo-1,3--glucanase minor isoform
YPR159w	*KRE6*	glucan synthase subunit
YBR212w	*NGR1*	glucose-repressible RNA-binding protein
YOR043w	*WHI2*	growth regulation protein
YKL021c	*MAK11*	involved in cell growth and replication of M1 dsRNA virus
YMR165c	*SMP2*	involved in plasmid maintenance, respiration and cell proliferation
YNL197c	*WHI3*	involved in regulation of cell size
YDR293c	*SSD1*	involved in the tolerance to high concentration of Ca^{2+}
YOR336w	*KRE5*	killer toxin-resistance protein
YDR480w	*DIG2*	MAP kinase-associated protein
YJR032w	*CPR7*	member of the cyclophilin family
YGR029w	*ERV1*	mitochondrial biogenesis and cell viability protein
YBR089c-a	*NHP6B*	nonhistone chromosomal protein
YPR052c	*NHP6A*	nonhistone chromosomal protein related to mammalian HMG1
YKL203c	*TOR2*	phosphatidylinositol 3-kinase
YCR034w	*GNS1*	probable 1,3--glucan synthase subunit
YOR149c	*SMP3*	protein kinase C pathway protein
YPL140c	*MKK2*	protein kinase of the MAP kinase kinase (MEK) family
YDR151c	*CTH1*	protein of the inducible CCCH zinc-finger family
YDR075w	*PPH3*	protein ser/thr phosphatase
YDL006w	*PTC1*	protein ser/thr phosphatase 2c
YDL134c	*PPH21*	protein ser/thr phosphatase PP2A-1
YDL188c	*PPH22*	protein ser/thr phosphatase PP2A-2
YDR137w	*RGP1*	reduced growth phenotype protein
YJR090c	*GRR1*	required for glucose repression and for glucose and cation transport
YGL190c	*CDC55*	ser/thr phosphatase 2A regulatory subunit B
YAL016w	*TPD3*	ser/thr phosphatase 2A, regulatory subunit A
YOR231w	*MKK1*	ser/thr protein kinase
YPR161c	*SGV1*	ser/thr protein kinase
YHR030c	*SLT2*	ser/thr protein kinase of MAP kinase family
YJL095w	*BCK1*	ser/thr protein kinase of the MEKK family
YER167w	*BCK2*	ser/thr protein kinase of the protein kinase C pathway
YNL180c		similarity to *S. pombe* Cdc42p and other GTP-binding proteins
YOR027w	*STI1*	stress-induced protein
YGL106w		strong similarity to calmodulins
YER177w	*BMH1*	strong similarity to mammalian 14-3-3 proteins
YPL032c		strong similarity to Pam1p
YKL161c		strong similarity to ser/thr-specific protein kinase Slt2p
YDL224c		strong similarity to WHI3 protein
YDR099w	*BMH2*	suppressor of clathrin deficiency
YJR075w	*HOC1*	suppressor of *pkc1*
YOR075w	*UFE1*	syntaxin (T-SNARE) of the ER
YHR206w	*SKN7*	transcription factor with similarity to Hsf1p
YNL079c	*TPM1*	tropomyosin 1
YLR337w	*VRP1*	verprolin
YLR403w	*SFP1*	zinc-finger protein

budding, cell polarity and filament formation

YFL039c	*ACT1*	actin
YDR129c	*SAC6*	actin filament bundling protein, fimbrin
YCR088w	*ABP1*	actin-binding protein
YDL029w	*ACT2*	actin-like protein
YNL138w	*SRV2*	adenylate cyclase-associated protein, 70K
YOR304c-a	*BAT2*	AIP3 binding protein

YBL085w	*BOI1*	BEM1 protein-binding protein
YBR200w	*BEM1*	bud emergence mediator
YPL161c	*BEM4*	bud emergence protein
YDR351w	*SBE2*	bud growth protein
YLR319c	*BUD6*	bud site selection protein
YNL271c	*BNI1*	budding protein
YER114c	*BOI2*	budding protein
YJR092w	*BUD4*	budding protein
YLR353w	*BUD8*	budding protein
YGR041w	*BUD9*	budding protein
YBR109c	*CMD1*	calmodulin
YHR135c	*YCK1*	casein kinase I isoform
YNL154c	*YCK2*	casein kinase I isoform
YPL255w	*BBP1*	cell division control protein
YDR182w	*CDC1*	cell division control protein
YLR314c	*CDC3*	cell division control protein
YFL009w	*CDC4*	cell division control protein
YJL174w	*KRE9*	cell-wall synthesis protein
YBR038w	*CHS2*	chitin synthase II
YBR023c	*CHS3*	chitin synthase III
YLR330w	*CHS5*	chitin synthesis protein
YLL050c	*COF1*	cofilin, actin binding and severing protein
YDR251w	*PAM1*	coiled-coil protein multicopy suppressor of loss of PP2A
YPL256c	*CLN2*	cyclin, G1/S specific
YBR160w	*CDC28*	cyclin-dependent protein kinase
YNL243w	*SLA2*	cytoskeleton assembly control protein
YLR071c	*RGR1*	DNA-directed RNA polymerase II holoenzyme subunit
YKL007w	*CAP1*	F-actin capping protein, α subunit
YIL034c	*CAP2*	F-actin capping protein, β subunit
YJL157c	*FAR1*	factor arrest protein
YAL041w	*CDC24*	GDP/GTP exchange factor for Cdc42p
YCR038c	*BUD5*	GDP/GTP exchange factor for Rsr1p/Bud1p
YGR070w	*ROM1*	GDP/GTP exchange protein for Rho1p
YLR371w	*ROM2*	GDP/GTP exchange protein for Rho1p
YGL155w	*CDC43*	geranylgeranyltransferase type I β subunit
YER155c	*BEM2*	GTPase-activating protein
YKL092c	*BUD2*	GTPase-activating protein for Bud1p/Rsr1p
YPL115c	*BEM3*	GTPase-activating protein for Cdc42p and Rho1p
YGR152c	*RSR1*	GTP-binding protein
YLR229c	*CDC42*	GTP-binding protein of RAS superfamily
YIL118w	*RHO3*	GTP-binding protein of the RHO family
YKR055w	*RHO4*	GTP-binding protein of the RHO family
YPR165w	*RHO1*	GTP-binding protein of the RHO subfamily of RAS-like proteins
YNL090w	*RHO2*	GTP-binding protein of the RHO subfamily of RAS-like proteins
YOR156c	*NFI1*	interacts with Cdc12p in 2-hybrid assay
YKR063c	*LAS1*	involved in cell morphogenesis, cytoskeletal regulation and bud formation
YLL021w	*SPA2*	involved in cell polarity
YDR085c	*AFR1*	involved in morphogenesis of the mating projection
YMR273c	*ZDS1*	involved in negative regulation of cell polarity
YER149c	*PEA2*	involved in oriented growth towards mating partner
YDR293c	*SSD1*	involved in the tolerance to high concentration of Ca^{2+}
YKL079w	*SMY1*	kinesin-related protein
YOR198c	*BFR1*	maintenance of normal ploidy
YJL158c		member of the Pir1p/Hsp150p/Pir3p family
YJL159w		member of the Pir1p/Hsp150p/Pir3p family
YJL160c		member of the Pir1p/Hsp150p/Pir3p family
YOR188w	*MSB1*	morphogenesis-related protein
YGR014w	*MSB2*	multicopy suppressor of a *cdc24* bud emergence defect
YHR023w	*MYO1*	myosin 1 isoform (type II myosin) heavy chain
YOR326w	*MYO2*	myosin heavy chain
YAL029c	*MYO4*	myosin heavy chain, unconventional (class V) isoform
YMR109w	*MYO5*	myosin I
YKL129c	*MYO3*	myosin type I
YBR089c-a	*NHP6B*	nonhistone chromosomal protein
YPR052c	*NHP6A*	nonhistone chromosomal protein related to mammalian HMG1
YCR057c	*PWP2*	periodic tryptophan protein
YBL058w	*SHP1*	potential regulatory subunit for Glc7p
YCR034w	*GNS1*	probable 1,3—glucan synthase subunit
YER108c	*FLO8*	probable transcriptional activator of Flo1p
YOR122c	*PFY1*	profilin
YPR122w	*AXL1*	protease
YOR149c	*SMP3*	protein kinase C pathway protein
YPL140c	*MKK2*	protein kinase of the MAP kinase kinase (MEK) family
YDL047w	*SIT4*	protein ser/thr phosphatase
YDL006w	*PTC1*	protein ser/thr phosphatase 2c
YDL134c	*PPH21*	protein ser/thr phosphatase PP2A-1
YDL188c	*PPH22*	protein ser/thr phosphatase PP2A-2
YDR388w	*RVS167*	reduced viability upon starvation protein
YMR016c	*SOK2*	regulatory protein in the PKA signal transduction pathway
YCL014w	*BUD3*	required for axial budding
YIL140w	*SRO4*	required for axial pattern of budding
YER109c	*FLO8*	required for diploid filamentous growth
YNL084c	*END3*	required for endocytosis and cytoskeletal organization
YJR090c	*GRR1*	required for glucose repression and for glucose and cation transport
YDL135c	*RDI1*	RHO GDP dissociation inhibitor with activity toward Rho1p
YOR127w	*RGA1*	RHO-type GTPase-activating protein for Cdc42p
YKL181w	*PRPS1*	ribose-phosphate pyrophosphokinase
YJR076c	*CDC11*	septin
YHR107c	*CDC12*	septin
YGL190c	*CDC55*	ser/thr phosphatase 2A regulatory subunit B
YAL016w	*TPD3*	ser/thr phosphatase 2A, regulatory subunit A
YER133w	*GLC7*	ser/thr phosphoprotein phosphatase 1, catalytic subunit
YOR178c	*GAC1*	ser/thr phosphoprotein phosphatase 1, regulatory subunit
YKL048c	*ELM1*	ser/thr protein kinase
YOR231w	*MKK1*	ser/thr protein kinase
YBL105c	*PKC1*	ser/thr protein kinase
YHR030c	*SLT2*	ser/thr protein kinase of MAP kinase family
YJL095w	*BCK1*	ser/thr protein kinase of the MEKK family
YLR362w	*STE11*	ser/thr protein kinase of the MEKK family
YER167w	*BCK2*	ser/thr protein kinase of the protein kinase C pathway
YDL159w	*STE7*	ser/thr/tyr protein kinase of MAP kinase kinase family
YDL225w		similarity to Cdc11p, Cdc3p and human CDC10 protein
YGL054c		similarity to *D. melanogaster* cni protein
YCR009c	*RVS161*	similarity to human amphiphysin and Rvs167p
YDR208w	*MSS4*	similarity to human PI P 5-kinase
YDR409w		similarity to Nfi1p
YDL125c	*YHI1*	similarity to protein kinase C inhibitor-I
YDR245w	*MNN10*	similarity to *S. pombe* galactosyltransferase
YFR040w	*SAP155*	Sit4p-associated protein
YKR028w	*SAP190*	Sit4p-associated protein
YJL098w	*SAP185*	Sit4p-associating protein
YLR372w	*SUR4*	sterol isomerase
YHR103w	*SBE22*	strong similarity to budding protein Sbe2p
YGL106w		strong similarity to calmodulins
YBR210w		strong similarity to *D. melanogaster* cornichon protein
YPL032c		strong similarity to Pam1p
YNL320w		strong similarity to *S. pombe* Bem46p
YKL161c		strong similarity to ser/thr-specific protein kinase Slt2p
YBR161w		strong similarity to Sur1p
YJR075w	*HOC1*	suppressor of *pkc1*
YDR297w	*SUR2*	suppressor of *rvs161* and *rvs167* mutations
YPL057c	*SUR1*	suppressor of r*vs161*, r*vs167*, and *cls2* mutations
YKL043w	*PHD1*	transcription factor
YGL181w	*GTS1*	transcription factor of the Gcs1p/Glo3p/Sps18p family
YHR084w	*STE12*	transcriptional activator
YNL079c	*TPM1*	tropomyosin 1
YIL138c	*TPM2*	tropomyosin 2
YBR083w	*TEC1*	Ty transcription activator
YGL095c	*VPS45*	vacuolar protein sorting-associated protein
YLR337w	*VRP1*	verprolin

pheromone response and mating-type determination

YJR004c	*SAG1*	α-agglutinin
YNR044w	*AGA1*	α-agglutinin anchor subunit
YGL032c	*AGA2*	α-agglutinin binding subunit
YDR264c	*AKR1*	ankyrin repeat-containing protein
YKL209c	*STE6*	ATP-binding cassette transporter protein
YIL015w	*BAR1*	barrier pepsin
YBR200w	*BEM1*	bud emergence mediator
YPL161c	*BEM4*	bud emergence protein
YML057w	*CMP2*	calcineurin B, catalytic subunit
YLR433c	*CNA1*	calcineurin B, catalytic subunit
YKL190w	*CNB1*	calcineurin B, regulatory subunit
YBR109c	*CMD1*	calmodulin
YCR002c	*CDC10*	cell division control protein
YLR314c	*CDC3*	cell division control protein
YNL188w	*KAR1*	cell division control protein
YCL027w	*FUS1*	cell fusion protein
YJL174w	*KRE9*	cell-wall synthesis protein
YNL192w	*CHS1*	chitin synthase I
YBR023c	*CHS3*	chitin synthase III
YLR330w	*CHS5*	chitin synthesis protein
YDR073w	*SNF11*	component of SWI/SNF transcription activator complex
YOR290c	*SNF2*	component of SWI/SNF transcription activator complex
YBR289w	*SNF5*	component of SWI/SNF transcription activator complex
YPL016w	*SWI1*	component of SWI/SNF transcription activator complex
YPL256c	*CLN2*	cyclin, G1/S specific
YNL243w	*SLA2*	cytoskeleton assembly control protein
YER095w	*RAD51*	DNA repair protein
YDR076w	*RAD55*	DNA repair protein
YDR004w	*RAD57*	DNA repair protein
YAR007c	*RFA1*	DNA replication factor A, 69K subunit
YOL051w	*GAL11*	DNA-directed RNA polymerase II holoenzyme and Kornberg's mediator (SRB) subcomplex subunit
YLR071c	*RGR1*	DNA-directed RNA polymerase II holoenzyme subunit
YBR253w	*SRB6*	DNA-directed RNA polymerase II suppressor protein
YNL053w	*MSG5*	dual-specifity protein phosphatase
YNL238w	*KEX2*	endoproteinase of late Golgi compartment
YLR300w	*EXG1*	exo-1,3—glucanase (I/II), major isoform
YDR261c	*EXG2*	exo-1,3—glucanase minor isoform
YKL007w	*CAP1*	F-actin capping protein, α subunit
YIL034c	*CAP2*	F-actin capping protein, β subunit
YJL157c	*FAR1*	factor arrest protein
YAL041w	*CDC24*	GDP/GTP exchange factor for Cdc42p
YGL097w	*SRM1*	GDP/GTP exchange factor for Gsp1p/Gsp2p
YNL236w	*SIN4*	global regulator protein
YHL025w	*SNF6*	global transcription activator
YKL104c	*GFA1*	glucosamine-fructose-6-phosphate transaminase
YPL115c	*BEM3*	GTPase-activating protein for Cdc42p and Rho1p
YHR005c	*GPA1*	GTP-binding protein α subunit of the pheromone pathway
YOR212w	*STE4*	GTP-binding protein β subunit of the pheromone pathway
YJR086w	*STE18*	GTP-binding protein γ subunit of the pheromone pathway
YLR229c	*CDC42*	GTP-binding protein of RAS superfamily
YNL173c	*MDG1*	GTP-binding protein of the pheromone pathway
YER020w	*GPA2*	guanine nucleotide-binding regulatory protein
YDL227c	*HO*	homothallic switching endonuclease
YNL291c	*MID1*	involved in Ca^{2+} influx during mating
YLL021w	*SPA2*	involved in cell polarity
YLR452c	*SST2*	involved in desensitization to α-factor pheromone
YDR085c	*AFR1*	involved in morphogenesis of the mating projection
YER149c	*PEA2*	involved in oriented growth toward mating partner
YMR052w	*FAR3*	involved in pheromone-mediated cell cycle arrest
YMR127c	*SAS2*	involved in silencing at HMR
YPR141c	*KAR3*	kinesin-related protein
YPL187w	*MFα1*	mating pheromone α-1 factor
YGL089c	*MFα2*	mating pheromone α-2 factor
YDR461w	*MFA1*	mating pheromone a-factor 1
YNL145w	*MFA2*	mating pheromone a-factor 2
YLR332w	*MID2*	mating process protein
YHR066w	*SSF1*	mating protein
YDR143c	*SAN1*	mating type regulatory protein
YCR040w	*MATα1*	mating type regulatory protein (expressed copy at MAT locus)
YCR039c	*MATα2*	mating type regulatory protein (expressed copy at MAT locus)
YCL066w	*α1*	mating type regulatory protein (silenced copy at HML locus)
YCL067c	*α2*	mating type regulatory protein (silenced copy at HML locus)
YCR097w	*A1*	mating type regulatory protein (silenced copy at HMR locus)
YCR096c	*A2*	mating type regulatory protein (silenced copy at HMR locus)
YIL047c	*SYG1*	member of the major facilitator superfamily
YBL016w	*FUS3*	mitogen-activated protein kinase (MAP kinase)
YGL151w	*NUT1*	negative regulator of HO endonuclease
YPR168w	*NUT2*	negative regulator of HO endonuclease
YKL185w	*ASH1*	negative regulator of HO expression
YCL029c	*BIK1*	nuclear fusion protein
YML065w	*ORC1*	origin recognition complex, 104K subunit
YNL261w	*ORC5*	origin recognition complex, 50K subunit
YHR118c	*ORC6*	origin recognition complex, 50K subunit
YPR162c	*ORC4*	origin recognition complex, 56K subunit
YLL004w	*ORC3*	origin recognition complex, 62K subunit
YBR060c	*RRR1*	origin recognition complex, 72K subunit
YFL026w	*STE2*	pheromone α-factor receptor
YKL178c	*STE3*	pheromone α-factor receptor
YCL032w	*STE50*	pheromone response pathway protein
YDR103w	*STE5*	pheromone signal transduction pathway protein
YCR034w	*GNS1*	probable 1,3—glucan synthase subunit
YLR389c	*STE23*	protease involved in α-factor processing
YML032c	*RAD52*	recombination and DNA repair protein
YDR392w	*SPT3*	regulatory protein
YCL055w	*KAR4*	regulatory protein required for pheromone induction of karyogamy genes
YMR232w	*FUS2*	required for cell fusion during mating
YBR040w	*FIG1*	required for efficient mating
YCR089w	*FIG2*	required for efficient mating
YNL084c	*END3*	required for endocytosis and cytoskeletal organization
YKL130c	*SHE2*	required for mother cell-specific expression of HO
YOR035c	*SHE4*	required for mother cell-specific gene expression
YOR127w	*RGA1*	RHO-type GTPase-activating protein for Cdc42p
YLR441c	*RP10A*	ribosomal protein S3a.e
YJR076c	*CDC11*	septin
YHR107c	*CDC12*	septin
YPR161c	*SGV1*	ser/thr protein kinase
YGR040w	*KSS1*	ser/thr protein kinase of the MAP kinase family
YLR362w	*STE11*	ser/thr protein kinase of the MEKK family
YHL007c	*STE20*	ser/thr protein kinase of the pheromone pathway

YDL159w	*STE7*	ser/thr/tyr protein kinase of MAP kinase kinase family
YOL068c	*HST1*	silencing protein
YBL052c	*SAS3*	silencing protein
YKR101w	*SIR1*	silencing regulatory protein
YDL042c	*SIR2*	silencing regulatory protein
YLR442c	*SIR3*	silencing regulatory protein
YDR227w	*SIR4*	silencing regulatory protein
YNL283c		similarity to a-agglutinin core protein Aga1p
YDL225w		similarity to Cdc11p, Cdc3p and human CDC10 protein
YDR191w	*HST4*	similarity to Hst3p,Hst1p and Sir2p
YGR023w		similarity to Mid2p
YHR146w		similarity to pheromone-response G-protein YNL173c
YKL117w		Ste5p-associated protein
YGL106w		strong similarity to calmodulins
YJL212c		strong similarity to *S. pombe* isp4 protein
YDR312w	*SSF2*	strong similarity to SSF1 protein
YPL129w	*ANC1*	TFIIF subunit (transcription initiation factor), 30K
YDR146c	*SWI5*	transcription factor
YMR043w	*MCM1*	transcription factor of the MADS box family
YNL330c	*RPD3*	transcription modifier protein
YOL004w	*SIN3*	transcription regulatory protein
YJL176c	*SWI3*	transcription regulatory protein
YHR084w	*STE12*	transcriptional activator
YER068w	*MOT2*	transcriptional repressor
YNL079c	*TPM1*	tropomyosin 1
YOR219c	*STE13*	type IV dipeptidyl aminopeptidase
YMR022w	*QRI8*	ubiquitin-conjugating enzyme
YBR082c	*UBC4*	ubiquitin-conjugating enzyme
YDR059c	*UBC5*	ubiquitin-conjugating enzyme
YER100w	*UBC6*	ubiquitin-conjugating enzyme
YJR117w	*STE24*	zinc metallo-protease

sporulation and germination

YGR032w	*GSC2*	1,3—D-glucan synthase, subunit
YPR103w	*PRE2*	26S proteasome subunit
YGR253c	*PUP2*	26S proteasome subunit
YBL015w	*ACH1*	acetyl-CoA hydrolase
YCR048w	*ARE1*	acyl-CoA sterol acyltransferase
YNR019w	*ARE2*	acyl-CoA sterol acyltransferase
YOL052c	*SPE2*	adenosylmethionine decarboxylase
YDR173c	*ARG82*	arginine metabolism transcription factor
YIL033c	*SRA1*	cAMP dependent protein kinase, regulatory subunit
YDR182w	*CDC1*	cell division control protein
YFL009w	*CDC4*	cell division control protein
YNL192w	*CHS1*	chitin synthase I
YBR023c	*CHS3*	chitin synthase III
YDR402c	*DIT2*	cytochrome P450 56
YML110c	*DBI56*	DBF2 interacting protein
YOL051w	*GAL11*	DNA-directed RNA polymerase II holoenzyme and Kornberg's mediator (SRB) subcomplex subunit
YBR253w	*SRB6*	DNA-directed RNA polymerase II suppressor protein
YMR013c	*SEC59*	dolichol kinase
YOR190w	*SPR1*	exo-1,3—glucanase
YLR300w	*EXG1*	exo-1,3—glucanase (I/II), major isoform
YDR261c	*EXG2*	exo-1,3—glucanase minor isoform
YBR045c	*GIP1*	Glc7p-interacting protein
YNL236w	*SIN4*	global regulator protein
YDL240w	*LRG1*	GTPase-activating protein of the RHO/RAC family
YNL098c	*RAS2*	GTP-binding protein
YPL240c	*HSP82*	heat-shock protein
YJL146w	*IDS2*	*IME2*-dependent signaling protein
YER044c-a	*MEI4*	meiosis-specific protein
YBR057c	*MUM2*	meiotic protein
YAL009w	*SPO7*	meiotic protein
YLR025w	*SNF7*	nuclear protein
YKR031c	*SPO14*	phospholipase D
YGL192w	*IME4*	positive transcription factor for *IME2*
YCR034w	*GNS1*	probable 1,3—glucan synthase subunit
YIR026c	*YVH1*	protein tyrosine phosphatase
YOR208w	*PTP2*	protein-tyrosine-phosphatase
YBL066c	*SEF1*	putative transcription factor
YNR052c	*POP2*	required for glucose derepression
YDR218c	*SPR28*	septin-related sporulation protein
YJL106w	*IME2*	ser/thr protein kinase
YDR523c	*SPS1*	ser/thr protein kinase
YNL307c	*MCK1*	ser/thr/tyr protein kinase
YDR403w	*DIT1*	spore-wall maturation protein
YBR148w	*YSW1*	spore-specific protein
YKL165c		sporulation protein
YLR399c	*BDF1*	sporulation protein
YHR152w	*SPO12*	sporulation protein
YHR153c	*SPO16*	sporulation protein
YIL099w	*SGA1*	sporulation specific glucan 1,4—glucosidase
YDR108w	*GSG1*	sporulation-specific protein
YER096w	*SHC1*	sporulation-specific protein
YLR307w	*CDA1*	sporulation-specific chitin deacetylase
YLR308w	*CDA2*	sporulation-specific chitin deacetylase
YPR054w	*SMK1*	sporulation-specific MAP kinase
YER115c	*SPR6*	sporulation-specific protein
YNL202w	*SPS19*	sporulation-specific protein
YDR522c	*SPS2*	sporulation-specific protein
YOR313c	*SPS4*	sporulation-specific protein
YGR059w	*SPR3*	sporulation-specific septin
YHR139c	*SPS100*	sporulation-specific wall maturation protein
YNL204c	*SPS18*	sporulation-specific zinc-finger protein
YBR078w		strong similarity to sporulation-specific Sps2p
YCL048w		strong similarity to sporulation-specific Sps2p
YDR055w		strong similarity to Sps2p
YOR075w	*UFE1*	syntaxin (T-SNARE) of the ER
YMR043w	*MCM1*	transcription factor of the MADS box family
YJR094c	*IME1*	transcription factor required for sporulation
YNL330c	*RPD3*	transcription modifier protein
YOL004w	*SIN3*	transcription regulatory protein
YNL012w	*SPO1*	transcriptional regulator involved in sporulation
YLL039c	*UBI4*	ubiquitin
YGL058w	*RAD6*	ubiquitin-conjugating enzyme
YDR177w	*UBC1*	ubiquitin-conjugating enzyme
YER125w	*RSP5*	ubiquitin-protein ligase
YML115c	*VAN1*	vanadate resistance protein

meiosis

YPR103w	*PRE2*	26S proteasome subunit
YGR253c	*PUP2*	26S proteasome subunit
YKL112w	*ABF1*	ARS-binding factor
YFL037w	*TUB2*	β-tubulin
YOR265w	*RBL2*	β-tubulin binding protein
YPL204w	*HRR25*	casein kinase I, ser/thr/tyr protein kinase
YBR136w	*ESR1*	cell cycle checkpoint protein
YPL255w	*BBP1*	cell division control protein
YDL220c	*CDC13*	cell division control protein
YPL042c	*SSN3*	cyclin-dependent ser/thr protein kinase
YOR368w	*RAD17*	DNA damage checkpoint control protein
YMR224c	*MRE11*	DNA repair and meiotic recombination protein
YNL250w	*RAD50*	DNA repair protein
YDR076w	*RAD55*	DNA repair protein
YDR004w	*RAD57*	DNA repair protein
YDR369c	*XRS2*	DNA repair protein
YLR234w	*TOP3*	DNA topoisomerase III
YGL163c	*RAD54*	DNA-dependent ATPase of the Snf2p family
YNL025c	*SSN8*	DNA-directed RNA polymerase II holoenzyme and Kornberg's mediator (SRB) subcomplex subunit, cyclin C homolog
YKL017c	*DIP1*	Dom34p Interacting Protein
YLR129w	*DIP2*	Dom34p-interacting protein
YBR045c	*GIP1*	Glc7-interacting protein
YPL240c	*HSP82*	heat-shock protein
YJL146w	*IDS2*	*IME2*-dependent signaling protein
YOR198c	*BFR1*	maintenance of normal ploidy
YHL022c	*SPO11*	meiosis-specific protein
YER179w	*DMC1*	meiosis-specific protein
YIL072w	*HOP1*	meiosis-specific protein
YER180c	*ISC10*	meiosis-specific protein
YER044c-a	*MEI4*	meiosis-specific protein
YDL154w	*MSH5*	meiosis-specific protein
YHR124w	*NDT80*	meiosis-specific protein
YHR157w	*REC104*	meiosis-specific protein
YLR263w	*RED1*	meiosis-specific protein
YHR079c-a	*SAE3*	meiosis-specific protein
YHR014w	*SPO13*	meiosis-specific protein
YPL121c	*MEI5*	meiotic protein
YBR057c	*MUM2*	meiotic protein
YAL009w	*SPO7*	meiotic protein
YNL210w	*MER1*	meiotic recombination protein
YHR086w	*NAM8*	meiotic recombination protein
YLR329w	*REC102*	meiotic recombination protein
YJR021c	*REC107*	meiotic recombination protein
YMR133w	*REC114*	meiotic recombination protein
YGL175c	*SAE2*	meiotic recombination protein
YHL027w	*RIM101*	meiotic regulatory protein
YDR207c	*UME6*	negative transcriptional regulator
YBL058w	*SHP1*	potential regulatory subunit for Glc7p
YMR076c	*PDS5*	precocious dissociation of sister chromatids
YDL017w	*CDC7*	protein kinase
YIR026c	*YVH1*	protein tyrosine phosphatase
YOR208w	*PTP2*	protein tyrosine phosphatase
YBR073w	*RDH54*	required for meiosis
YER133w	*GLC7*	ser/thr phosphoprotein phosphatase 1, catalytic subunit
YOR178c	*GAC1*	ser/thr phosphoprotein phosphatase 1, regulatory subunit
YJL106w	*IME2*	ser/thr protein kinase
YMR139w	*MDS1*	ser/thr protein kinase
YOR351c	*MEK1*	ser/thr protein kinase
YDR523c	*SPS1*	ser/thr protein kinase
YNL307c	*MCK1*	ser/thr/tyr protein kinase
YDL028c	*MPS1*	ser/thr/tyr protein kinase
YNL001w	*DOM34*	similarity to *D. melanogaster* pelota protein
YFL003c	*MSH4*	similarity to MSH proteins
YHR152w	*SPO12*	sporulation protein
YHR153c	*SPO16*	sporulation protein
YDR108w	*GSG1*	sporulation-specific protein
YNL202w	*SPS19*	sporulation-specific protein
YDR522c	*SPS2*	sporulation-specific protein
YNL204c	*SPS18*	sporulation-specific zinc-finger protein
YBR078w		strong similarity to sporulation-specific Sps2p
YCL048w		strong similarity to sporulation-specific Sps2p
YDR055w		strong similarity to Sps2p
YGR258c	*RAD2*	structure-specific nuclease of the nucleotide excision repairosome
YDR285w	*ZIP1*	synaptonemal complex protein
YJR094c	*IME1*	transcription factor required for sporulation
YGR044c	*RME1*	zinc-finger transcription factor

DNA synthesis and replication

YKL112w	*ABF1*	ARS-binding factor
YFL009w	*CDC4*	cell division control protein
YDR364c	*CDC40*	cell division control protein
YLR274w	*CDC46*	cell division control protein
YBR202w	*CDC47*	cell division control protein
YJL194w	*CDC6*	cell division control protein
YML102w	*CAC2*	chromatin assembly complex, subunit
YBR195c	*MSI1*	chromatin assembly complex, subunit p50
YPR018w	*RLF2*	chromatin assembly complex, subunit p90
YPR120c	*CLB5*	cyclin, B-type
YGR109c	*CLB6*	cyclin, B-type
YPL256c	*CLN2*	cyclin, G1/S specific
YBR160w	*CDC28*	cyclin-dependent protein kinase
YER176w		DNA dependent ATPase/DNA helicase B
YHR164c	*DNA2*	DNA helicase
YDL164c	*CDC9*	DNA ligase
YJL173c	*RFA3*	DNA replication factor A, 13K subunit
YNL312w	*RFA2*	DNA replication factor A, 36K subunit
YAR007c	*RFA1*	DNA replication factor A, 69K subunit
YOL094c	*RFC4*	DNA replication factor C, 37K subunit
YNL290w	*RFC3*	DNA replication factor C, 40K subunit
YBR087w	*RFC5*	DNA replication factor C, 40K subunit
YJR068w	*RFC2*	DNA replication factor C, 41K subunit
YOR217w	*RFC1*	DNA replication factor C, 95K subunit
YJR006w	*HUS2*	DNA replication protein
YOL006c	*TOP1*	DNA topoisomerase I
YNL088w	*TOP2*	DNA topoisomerase II (ATP-hydrolysing)
YNL102w	*CDC17*	DNA-directed DNA polymerase α, 180K subunit
YIR008c	*PRI1*	DNA-directed DNA polymerase α, 48K subunit (DNA primase)
YKL045w	*PRI2*	DNA-directed DNA polymerase α, 58K subunit (DNA primase)
YPR135w	*POB1*	DNA-directed DNA polymerase α-binding protein
YDL102w	*CDC2*	DNA-directed DNA polymerase δ, catalytic 125K subunit
YNL262w	*POL2*	DNA-directed DNA polymerase ε, catalytic subunit A
YPR175w	*DPB2*	DNA-directed DNA polymerase ε, subunit B
YBR278w	*DPB3*	DNA-directed DNA polymerase ε, subunit C
YOR330c	*MIP1*	DNA-directed DNA polymerase γ, catalytic subunit, mitochondrial
YPL167c	*REV3*	DNA-directed DNA polymerase ζ
YFL036w	*RPO41*	DNA-directed RNA polymerase, mitochondrial
YMR072w	*ABF2*	high-mobility group protein
YJL090c	*DPB11*	involved in DNA replication and S-phase checkpoint
YDR068w	*DOS1*	involved in genome stability
YMR001c	*CDC5*	involved in regulation of DNA replication
YPR019w	*CDC54*	member of the Cdc46p/Mcm2p/Mcm3p family
YBL023c	*MCM2*	member of the Mcm2p,Mcm3p,Cdc46p family
YML065w	*ORC1*	origin recognition complex, 104K subunit
YNL261w	*ORC5*	origin recognition complex, 50K subunit
YHR118c	*ORC6*	origin recognition complex, 50K subunit
YPR162c	*ORC4*	origin recognition complex, 56K subunit
YLL004w	*ORC3*	origin recognition complex, 62K subunit
YBR060c	*RRR1*	origin recognition complex, 72K subunit
YGR132c	*PHB1*	prohibitin, antiproliferative protein
YBR088c	*POL30*	proliferating cell nuclear antigen (PCNA)
YDL017w	*CDC7*	protein kinase
YFR028c	*CDC14*	protein-tyrosine-phosphatase
YDR052c	*DBF4*	regulatory subunit for Cdc7p protein kinase
YEL032w	*MCM3*	replication initiation protein
YIL139c	*REV7*	required for DNA damage-induced mutagenesis
YLR103c	*CDC45*	required for minichromosome maintenance and initiation of chromosomal DNA replication
YIL150c	*DNA43*	required for S-phase initiation or completion
YER070w	*RNR1*	ribonucleoside-diphosphate reductase, large subunit
YJL026w	*RNR2*	ribonucleoside-diphosphate reductase, small subunit
YIL066c	*RNR3*	ribonucleotide reductase, repair inducible
YJL065c		similarity to DNA-directed DNA polymerase II subunit C
YDR206w		similarity to EST1 protein
YOR005c		similarity to human mRNA for DNA ligase IV
YOL095c		similarity to *S. aureus* DNA helicase PCRA
YGL201c		similarity with rat intestinal DNA replication protein
YKL113c	*RAD27*	ssDNA endonuclease and 5'-3'exonuclease
YCR028c-a	*RIM1*	ssDNA-binding protein, mitochondrial
YGR231c		strong similarity to prohibitins
YBL035c	*POL12*	subunit of DNA polymerase α-primase complex
YLR233c	*EST1*	telomere elongation protein
YCR077c	*(PAT1)*	topoisomerase II-associated protein
YOL115w	*TRF4*	topoisomerase I-related protein

YLR182w	*SWI6*	transcription factor
YDR054c	*CDC34*	ubiquitin-conjugating enzyme
YGL058w	*RAD6*	ubiquitin-conjugating enzyme

recombination and DNA repair

YGL127c	*SOH1*	allows *hpr1* null mutant to grow at 37ºC
YBR136w	*ESR1*	cell-cycle checkpoint protein
YER173w	*RAD24*	cell-cycle checkpoint protein
YDR182w	*CDC1*	cell-division control protein
YMR106c	*HDF2*	component of DNA end-joining repair pathway
YKL011c	*CCE1*	cruciform-cutting endonuclease, mitochondrial
YER176w		DNA-dependent ATPase/DNA helicase B
YLR032w	*RAD5*	DNA helicase
YMR190c	*SGS1*	DNA helicase
YML061c	*PIF1*	DNA helicase involved in mitochondrial DNA repair and telomere length
YDL164c	*CDC9*	DNA ligase
YMR167w	*MLH1*	DNA mismatch repair protein
YOL090w	*MSH2*	DNA mismatch repair protein
YNL082w	*PMS1*	DNA mismatch repair protein
YHR120w	*MSH1*	DNA mismatch repair protein, mitochondrial
YCR014c	*POL4*	DNA polymerase
YMR224c	*MRE11*	DNA repair and meiotic recombination protein
YJR035w	*RAD26*	DNA repair and recombination protein
YEL019c	*MMS21*	DNA repair protein
YML095c	*RAD10*	DNA repair protein
YCR066w	*RAD18*	DNA repair protein
YNL250w	*RAD50*	DNA repair protein
YER095w	*RAD51*	DNA repair protein
YDR076w	*RAD55*	DNA repair protein
YDR004w	*RAD57*	DNA repair protein
YDR369c	*XRS2*	DNA repair protein
YAR007c	*RFA1*	DNA replication factor A, 69K subunit
YOL006c	*TOP1*	DNA topoisomerase I
YNL088w	*TOP2*	DNA topoisomerase II (ATP-hydrolysing)
YGL163c	*RAD54*	DNA-dependent ATPase of the Snf2p family
YFR023w	*PES4*	DNA-directed DNA polymerase ε suppressor
YPL167c	*REV3*	DNA-directed DNA polymerase ζ
YCR092c	*MSH3*	DNA-repair protein
YKR056w	*RNC1*	endo-exonuclease
YOR033c	*DHS1*	exonuclease, interacting with Msh2p
YLR288c	*MEC3*	G2-specific checkpoint protein
YMR284w	*HDF1*	high-affinity DNA-binding protein
YDR138w	*HPR1*	hyperrecombination protein related to Top1p
YHL022c	*SPO11*	meiosis-specific protein
YER179w	*DMC1*	meiosis-specific protein
YIL072w	*HOP1*	meiosis-specific protein
YHR157w	*REC104*	meiosis-specific protein
YLR263w	*RED1*	meiosis-specific protein
YPL121c	*MEI5*	meiotic protein
YNL210w	*MER1*	meiotic recombination protein
YHR086w	*NAM8*	meiotic recombination protein
YLR329w	*REC102*	meiotic recombination protein
YJR021c	*REC107*	meiotic recombination protein
YGL175c	*SAE2*	meiotic recombination protein
YBR114w	*RAD16*	nucleotide excision repair protein
YGL025c	*PGD1*	probable transcription factor
YBR088c	*POL30*	proliferating cell nuclear antigen (PCNA)
YML032c	*RAD52*	recombination and DNA repair protein
YDL059c	*RAD59*	recombination and DNA repair protein
YLR383w	*RHC18*	recombination repair protein
YBR073w	*RDH54*	required for meiosis
YIL066c	*RNR3*	ribonucleotide reductase, repair inducible
YMR228w	*MTF1*	RNA polymerase specific factor, mitochondrial
YOR351c	*MEK1*	ser/thr protein kinase
YPL164c		similarity to mismatch repair protein Mlh1p
YOR077w	*RTS2*	similarity to mouse KIN17 protein
YFL003c	*MSH4*	similarity to MSH proteins
YDR097c	*MSH6*	similarity to MSH proteins
YHR031c		similarity to Pif1p
YGL043w	*DST1*	TFIIS (transcription elongation factor)
YCR077c	*(PAT1)*	topoisomerase II-associated protein
YGR063c	*SPT4*	transcription initiation protein
YGR116w	*SPT6*	transcription initiation protein
YML021c	*UNG1*	uracil-DNA glycosylase

cell-cycle control and mitosis

YFR052w	*NIN1*	26S proteasome regulatory subunit
YPR103w	*PRE2*	26S proteasome subunit
YGL048c	*SUG1*	26S proteasome subunit
YKL145w	*YTA3*	26S proteasome subunit
YJL005w	*CYR1*	adenylate cyclase
YPL239w	*YAR1*	ankyrin repeat-containing protein
YHR101c	*BIG1*	big cells phenotype
YHR208w	*TWT1*	branched-chain amino-acid aminotransferase, mitochondrial
YFL037w	*TUB2*	β-tubulin
YBR200w	*BEM1*	bud emergence mediator
YBR109c	*CMD1*	calmodulin
YOR061w	*CKA2*	casein kinase II α′ subunit
YIL035c	*CKA1*	casein kinase II, catalytic α subunit
YER123w	*YCK3*	casein kinase, isoform 3
YLR178c	*TFS1*	cdc25-dependent nutrient- and ammonia-response cell-cycle regulator
YGR049w	*SCM4*	*cdc4* suppressor
YFL029c	*CAK1*	cdk-activating protein kinase
YMR055c	*BUB2*	cell-cycle arrest protein
YOR026w	*BUB3*	cell-cycle arrest protein
YPL178w	*SAE1*	cell-cycle block in meiotic prophase
YBR136w	*ESR1*	cell-cycle checkpoint protein
YPR111w	*DBF20*	cell-cycle protein kinase related to Dbf2p
YDR113c	*PDS1*	cell-cycle regulator
YBR215w	*HPC2*	cell-cycle regulatory protein
YOR373w	*NUD1*	cell-cycle regulatory protein
YPL255w	*BBP1*	cell division control protein
YDR182w	*CDC1*	cell division control protein
YCR002c	*CDC10*	cell division control protein
YDL220c	*CDC13*	cell division control protein
YGL116w	*CDC20*	cell division control protein
YLR314c	*CDC3*	cell division control protein
YDR168w	*CDC37*	cell division control protein
YFL009w	*CDC4*	cell division control protein
YDR364c	*CDC40*	cell division control protein
YLR274w	*CDC46*	cell division control protein
YBR202w	*CDC47*	cell division control protein
YJL194w	*CDC6*	cell division control protein
YNL188w	*KAR1*	cell division control protein
YDL226c	*GCS1*	cell proliferation zinc-finger protein
YNL327w	*EGT2*	cell-cycle regulation protein
YHR129c	*ACT5*	centractin
YLR175w	*CBF5*	centromere/microtubule binding protein
YDR254w	*CHL4*	chromosome segregation protein
YOR349w	*CIN1*	chromosome segregation protein
YFL008w	*SMC1*	chromosome segregation protein
YFR031c	*SMC2*	chromosome segregation protein
YMR028w	*TAP42*	component of the Tor signalling pathway
YIL106w	*MOB1*	conditional mutants arrest in late mitosis
YDL132w	*CDC53*	controls G1/S transition
YPR120c	*CLB5*	cyclin, B-type
YGR109c	*CLB6*	cyclin, B-type
YMR199w	*CLN1*	cyclin, G1/S specific
YPL256c	*CLN2*	cyclin, G1/S specific
YAL040c	*CLN3*	cyclin, G1/S specific
YNL289w	*PCL1*	cyclin, G1/S specific
YDL127w	*PCL2*	cyclin, G1/S specific
YGR108w	*CLB1*	cyclin, G2/M-specific
YPR119w	*CLB2*	cyclin, G2/M-specific
YDL155w	*CLB3*	cyclin, G2/M-specific
YLR210w	*CLB4*	cyclin, G2/M-specific
YBR135w	*CKS1*	cyclin-dependent kinase regulatory subunit
YBR160w	*CDC28*	cyclin-dependent protein kinase
YPL031c	*PHO85*	cyclin-dependent protein kinase
YDL108w	*KIN28*	cyclin-dependent ser/thr protein kinase
YGL215w	*CLG1*	cyclin-like protein
YOR368w	*RAD17*	DNA damage checkpoint control protein
YDR217c	*RAD9*	DNA repair checkpoint protein
YOR217w	*RFC1*	DNA replication factor C, 95K subunit
YLR234w	*TOP3*	DNA topoisomerase III
YKL017c	*DIP1*	Dom34p-interacting protein
YLR129w	*DIP2*	Dom34p-interacting protein
YNL053w	*MSG5*	dual-specifity protein phosphatase
YKR054c	*DYN1*	dynein heavy chain, cytosolic
YDR172w	*SUP35*	eukaryotic peptide chain release factor GTP-binding subunit
YNL292w	*EXM1*	exit from mitosis
YDL087c	*EXM2*	exit from mitosis
YJL157c	*FAR1*	factor arrest protein
YLR212c	*TUB4*	γ tubulin
YLR288c	*MEC3*	G2-specific checkpoint protein
YAL024c	*LTE1*	GDP/GTP exchange factor
YLL016w	*SDC25*	GDP/GTP exchange factor
YAL041w	*CDC24*	GDP/GTP exchange factor for Cdc42p
YLR310c	*CDC25*	GDP/GTP exchange factor for Ras1p and Ras2p
YGL207w	*SPT16*	general chromatin factor
YGL155w	*CDC43*	geranylgeranyltransferase type I β subunit
YER155c	*BEM2*	GTPase-activating protein
YMR138w	*CIN4*	GTP-binding protein
YNL098c	*RAS2*	GTP-binding protein
YLR229c	*CDC42*	GTP-binding protein of RAS superfamily
YML064c	*TEM1*	GTP-binding protein of the RAS superfamily
YNL007c	*SIS1*	heat-shock protein
YBR133c	*HSL7*	histone synthetic lethality
YJL080c	*SCP160*	histone-like protein
YOL076w	*DEC1*	interacts genetically with *CIN8*
YOR156c	*NFI1*	interacts with Cdc12p in 2-hybrid assay
YNR010w	*CSE2*	interacts with centromeric element CDEII
YML104c	*MDM1*	intermediate filament protein
YKR072c	*SIS2*	involved in cell cycle-specific gene expression
YKR063c	*LAS1*	involved in cell morphogenesis, cytoskeletal regulation and bud formation
YPL241c	*CIN2*	involved in chromosome segregation
YPL018w	*CTF19*	involved in chromosome segregation
YJL090c	*DPB11*	involved in DNA replication and S-phase checkpoint
YMR273c	*ZDS1*	involved in negative regulation of cell polarity
YMR052w	*FAR3*	involved in pheromone-mediated cell cycle arrest
YMR052w	*FAR3*	involved in pheromone-mediated cell cycle arrest
YMR001c	*CDC5*	involved in regulation of DNA replication
YMR127c	*SAS2*	involved in silencing at HMR
YKR042w	*UTH1*	involved in the aging process
YDR293c	*SSD1*	involved in the tolerance to high concentration of Ca^{2+}
YNL189w	*SRP1*	karyopherin-α or importin
YEL061c	*CIN8*	kinesin-related protein
YPR141c	*KAR3*	kinesin-related protein
YBL063w	*KIP1*	kinesin-related protein
YJR060w	*CBF1*	kinetochore protein
YGR140w	*CBF2*	kinetochore protein complex CBF3, 110K subunit
YMR094w	*CTF13*	kinetochore protein complex CBF3, 58K subunit
YMR168c	*CEP3*	kinetochore protein complex CBF3, 71K subunit
YDR328c	*SKP1*	kinetochore protein complex CBF3, subunit D
YOR198c	*BFR1*	maintenance of normal ploidy
YLR332w	*MID2*	mating process protein
YER179w	*DMC1*	meiosis-specific protein
YHR079c-a	*SAE3*	meiosis-specific protein
YPR019w	*CDC54*	member of the Cdc46p/Mcm2p/Mcm3p family
YDL126c	*CDC48*	microsomal protein of CDC48/PAS1/SEC18 family of ATPases
YOR058c	*ASE1*	microtubule-associated protein
YNL064c	*YDJ1*	mitochondrial and ER import protein
YGR029w	*ERV1*	mitochondrial biogenesis and regulation of cell cycle
YBL016w	*FUS3*	mitogen-activated protein kinase (MAP kinase)
YMR294w	*JNM1*	mitosis protein, involved in nuclear migration
YMR036c	*MIH1*	M-phase inducing protein tyrosine phosphatase
YGL178w	*MPT5*	multicopy suppressor of *pop2*
YML109w	*ZDS2*	multicopy suppressor of *sin4*
YGL173c	*KEM1*	multifunctional nuclease
YDR207c	*UME6*	negative transcriptional regulator
YBL020w	*RFT1*	nuclear division protein
YNL299w	*TRF5*	nuclear division protein
YML031w	*NDC1*	nuclear envelope protein
YJR112w	*NNF1*	nuclear envelope protein
YCL029c	*BIK1*	nuclear fusion protein
YJL034w	*KAR2*	nuclear fusion protein
YDR150w	*NUM1*	nuclear migration protein
YAL025c	*MAK16*	nuclear viral propagation protein
YKR048c	*NAP1*	nucleosome assembly protein I
YLR079w	*SIC1*	p40 inhibitor of Cdc28p-Clb protein kinase complex
YJR066w	*TOR1*	phosphatidylinositol 3-kinase
YKL203c	*TOR2*	phosphatidylinositol 3-kinase
YLR305c	*STT4*	phosphatidylinositol-4-kinase
YCL004w	*PEL1*	phosphatidylserine synthase
YBL058w	*SHP1*	potential regulatory subunit for Glc7p
YGL238w	*CSE1*	probable kinetochore protein
YGR132c	*PHB1*	prohibitin, antiproliferative protein
YDL017w	*CDC7*	protein kinase
YOR149c	*SMP3*	protein kinase C pathway protein
YPL140c	*MKK2*	protein kinase of the MAP kinase kinase (MEK) family
YAR019c	*CDC15*	protein kinase of the MAP kinase kinase kinase family
YHR013c	*ARD1*	protein N-acetyltransferase subunit
YPL008w	*CHL1*	protein of the DEAH box family
YDL047w	*SIT4*	protein ser/thr phosphatase
YDL134c	*PPH21*	protein ser/thr phosphatase PP2A-1
YDL188c	*PPH22*	protein ser/thr phosphatase PP2A-2
YBR267w	*PPS1*	protein tyrosine phosphatase
YFR028c	*CDC14*	protein-tyrosine-phosphatase
YOL021c	*DIS3*	Ran binding protein
YDR137w	*RGP1*	reduced growth phenotype protein
YDR052c	*DBF4*	regulatory subunit for Cdc7p protein kinase
YKL193c	*SDS22*	regulatory subunit for the mitotic function of type I protein phosphatase
YMR078c	*CHL12*	required for accurate chromosome transmission in mitosis and maintenance of normal telomere length
YOL145c	*CTR9*	required for G1 cyclin expression
YLR103c	*CDC45*	required for minichromosome maintenance and initiation of chromosomal DNA replication
YKL089w	*MIF2*	required for normal chromosome segregation and spindle integrity
YGR098c	*ESP1*	required for normal spindle structure
YIL150c	*DNA43*	required for S-phase initiation or completion
YJL074c	*SMC3*	required for structural maintenance of chromosomes
YGR078c	*PAC10*	required in the absence of Cin8p
YDR488c	*PAC11*	required in the absence of Cin8p
YLR075w	*GRC5*	ribosomal protein
YJR076c	*CDC11*	septin
YHR107c	*CDC12*	septin
YAL016w	*TPD3*	ser/thr phosphatase 2A, regulatory subunit A
YER133w	*GLC7*	ser/thr phosphoprotein phosphatase 1, catalytic subunit
YOR178c	*GAC1*	ser/thr phosphoprotein phosphatase 1, regulatory subunit
YGR188c	*BUB1*	ser/thr protein kinase
YKL048c	*ELM1*	ser/thr protein kinase
YJL106w	*IME2*	ser/thr protein kinase
YPL209c	*IPL1*	ser/thr protein kinase
YOR231w	*MKK1*	ser/thr protein kinase
YBL105c	*PKC1*	ser/thr protein kinase
YPR161c	*SGV1*	ser/thr protein kinase
YJL141c	*YAK1*	ser/thr protein kinase
YMR104c	*YPK2*	ser/thr protein kinase
YHR030c	*SLT2*	ser/thr protein kinase of MAP kinase family
YJL095w	*BCK1*	ser/thr protein kinase of the MEKK family
YER167w	*BCK2*	ser/thr protein kinase of the protein kinase C pathway
YGR092w	*DBF2*	ser/thr protein kinase related to Dbf20p

YKL101w	*HSL1*	ser/thr protein kinase that interacts genetically with histone mutations
YHR102w	*NRK1*	ser/thr protein kinase that interacts with Cdc31p
YNL307c	*MCK1*	ser/thr/tyr protein kinase
YDL028c	*MPS1*	ser/thr/tyr protein kinase
YPL153c	*SPK1*	ser/thr/tyr protein kinase
YJL187c	*SWE1*	ser/tyr dual-specifity protein kinase
YGL003c		similarity to *C. elegans* CDC20 protein and human p55CDC
YDL225w		similarity to Cdc11p, Cdc3p and human CDC10 protein
YDL179w		similarity to cyclin G1 homologue HCS26
YNL001w	*DOM34*	similarity to *D. melanogaster* pelota protein
YNL246w		similarity to *D. melanogaster* SET protein
YAR003w	*FUN16*	similarity to human RB protein binding protein
YKL179c		similarity to Nuf1p
YDL125c	*YHI1*	similarity to protein kinase C inhibitor-I
YMR262w		similarity to *S. pombe* scn1 protein
YLR356w		similarity to SCM4 protein
YIL095w	*PRK1*	similarity to ser/thr protein kinase
YLR179c		similarity to Tfs1p and Nsp1p
YOL080c		similarity to *X. laevis* XPMC2 protein
YFR040w	*SAP155*	Sit4p-associated protein
YKR028w	*SAP190*	Sit4p-associated protein
YGL229c	*SAP4*	Sit4p-associated protein
YJL098w	*SAP185*	Sit4p-associating protein
YGL086w	*MAD1*	spindle assembly checkpoint protein
YMR198w	*CIK1*	spindle pole body associated protein
YDR356w	*NUF1*	spindle pole body component
YKL042w	*SPC42*	spindle pole body component
YNL126w	*SPC98*	spindle pole body component
YOR257w	*CDC31*	spindle pole body component, centrin
YOL069w	*NUF2*	spindle pole body protein
YJL030w	*MAD2*	spindle-assembly checkpoint protein
YJL013c	*MAD3*	spindle-assembly checkpoint protein
YGL106w		strong similarity to calmodulins
YHR205w	*SCH9*	strong similarity to cAMP-dependent protein kinase
YER066w		strong similarity to cell division control protein Cdc4p
YLR117c		strong similarity to *D. melanogaster* putative cell-cycle control protein crn
YIR012w		strong similarity to Dbf8p
YBR238c		strong similarity to general chromatin factor Spt16p
YKL049c	*CSE4*	strong similarity to histone H3
YGR231c		strong similarity to prohibitins
YLR215c		strong similarity to rat cell cycle progression related D123 protein
YKL161c		strong similarity to ser/thr-specific protein kinase Slt2p
YNL172w	*APC1*	subunit of anaphase-promoting complex (cyclosome)
YKL022c	*CDC16*	subunit of anaphase-promoting complex (cyclosome)
YHR166c	*CDC23*	subunit of anaphase-promoting complex (cyclosome)
YFR036w	*CDC26*	subunit of anaphase-promoting complex (cyclosome)
YBL084c	*CDC27*	subunit of anaphase-promoting complex (cyclosome)
YOR213c		subunit of the RSC complex
YCR052w	*RSC6*	subunit of the RSC complex
YFR037c	*RSC8*	subunit of the RSC complex
YLR321c	*SFH1*	subunit of the RSC complex
YIL126w	*STH1*	subunit of the RSC complex
YJR091c	*JSN1*	suppresses the high-temperature lethality of tub2-150
YIL031w	*SMT4*	suppressor of *mif2* temperature-sensitive mutation
YDR285w	*ZIP1*	synaptonemal complex protein
YPL129w	*ANC1*	TFIIF subunit (transcription initiation factor), 30K
YPR025c	*CCL1*	TFIIH subunit (transcription initiation factor), cyclin C component
YLR043c	*TRX1*	thioredoxin I
YGR209c	*TRX2*	thioredoxin II
YOL115w	*TRF4*	topoisomerase I-related protein
YDL165w	*CDC36*	transcription factor
YER111c	*SWI4*	transcription factor
YLR182w	*SWI6*	transcription factor
YGL181w	*GTS1*	transcription factor of the Gcs1p/Glo3p/Sps18p family
YMR043w	*MCM1*	transcription factor of the MADS box family
YHR206w	*SKN7*	transcription factor with similarity to Hsf1p
YDL056w	*MBP1*	transcription factor, subunit of the MBF factor
YOR028c	*CIN5*	transcriptional activator
YNL012w	*SPO1*	transcriptional regulator involved in sporulation
YOR361c	*PRT1*	translation initiation factor eIF3 subunit
YHR165c	*PRP8*	U5 snRNP protein, pre-mRNA splicing factor
YDR054c	*CDC34*	ubiquitin-conjugating enzyme
YDL064w	*UBC9*	ubiquitin-conjugating enzyme
YMR276w	*DSK2*	ubiquitin-like protein
YML115c	*VAN1*	vanadate resistance protein
YHR071w		weak similarity to cyclin G1 homolog HCS26 protein
YER122c	*GLO3*	zinc-finger protein
YLR403w	*SFP1*	zinc-finger protein
YGR044c	*RME1*	zinc-finger transcription factor

cytokinesis

YDL029w	*ACT2*	actin-like protein
YPL255w	*BBP1*	cell division control protein
YCR002c	*CDC10*	cell division control protein
YLR314c	*CDC3*	cell division control protein
YNL327w	*EGT2*	cell-cycle regulation protein
YNL192w	*CHS1*	chitin synthase I
YLR286c	*CTS1*	endochitinase
YOR156c	*NFI1*	interacts with Cdc12p in 2-hybrid assay
YNL225c		involved in cytokinesis
YLR131c	*ACE2*	metallothionein expression activator
YAL029c	*MYO4*	myosin heavy chain, unconventional (class V) isoform
YHR023w	*MYO1*	myosin-1 isoform (type II myosin) heavy chain
YCR057c	*PWP2*	periodic tryptophan protein
YNL267w	*PIK1*	phosphatidylinositol 4-kinase
YJR076c	*CDC11*	septin
YHR107c	*CDC12*	septin
YGL190c	*CDC55*	ser/thr phosphatase 2A regulatory subunit B
YAL016w	*TPD3*	ser/thr phosphatase 2A, regulatory subunit A
YNL298w	*CLA4*	ser/thr protein kinase
YDL225w		similarity to Cdc11p, Cdc3p and human CDC10 protein
YMR197c		similarity to Nuf1p

other cell-growth, cell-division and DNA-synthesis activities

YLR248w	*RCK2*	Ca^{2+}/calmodulin-dependent ser/thr protein kinase
YOR101w	*RAS1*	GTP-binding protein
YNL098c	*RAS2*	GTP-binding protein
YBR138c	*HDR1*	high-dosage reductional segregation defective
YHL003c	*LAG1*	longevity-assurance protein
YDL003w	*MCD1*	mitotic chromosome determinant
YGL091c	*NBP35*	nucleotide-binding protein
YOL110w	*SHR5*	RAS suppressor
YGL158w	*RCK1*	ser/thr protein kinase
YKL189w		similarity to mouse MO25 gene

Transcription

rRNA synthesis

YJR063w	*RPA12*	DNA-directed RNA polymerase I, 13.7K
YPR010c	*RPA135*	DNA-directed RNA polymerase I, 135K subunit
YOR341w	*RPA190*	DNA-directed RNA polymerase I, 190K α subunit
YOR340c	*RPA43*	DNA-directed RNA polymerase I, 36K subunit
YNL248c	*RPA49*	DNA-directed RNA polymerase I, 49K α subunit
YDR156w	*RPA14*	DNA-directed RNA polymerase I, A14 subunit
YOR224c	*RPB8*	DNA-directed RNA polymerase I, II, III 16K subunit
YPR187w	*RPO26*	DNA-directed RNA polymerase I, II, III 18K subunit
YBR154c	*RPB5*	DNA-directed RNA polymerase I, II, III 25K subunit
YHR143w-a	*RPC10*	DNA-directed RNA polymerase I, II, III 7.7K subunit
YOR210w	*RPB10*	DNA-directed RNA polymerase I, II, III 8.3K subunit
YPR110c	*RPC40*	DNA-directed RNA polymerase I, III 40K subunit
YNL113w	*RPC19*	DNA-directed RNA polymerase I, III, 16K subunit
YOR207c	*RPC128*	DNA-directed RNA polymerase III, 130K subunit
YOR116c	*RPO31*	DNA-directed RNA polymerase III, 160K subunit
YKL144c	*RPC25*	DNA-directed RNA polymerase III, 25K subunit
YNL151c	*RPC31*	DNA-directed RNA polymerase III, 31K subunit
YNR003c	*RPC34*	DNA-directed RNA polymerase III, 34K subunit
YDL150w	*RPC53*	DNA-directed RNA polymerase III, 47K subunit
YPR190c	*RPC82*	DNA-directed RNA polymerase III, 82K subunit
YLR039c	*RIC1*	involved in transcription of ribosomal proteins and ribosomal RNA
YKL125w	*RRN3*	RNA polymerase I specific transcription factor
YBL025w	*RRN10*	RNA polymerase I specific transcription initiation factor
YML043c	*RRN11*	RNA polymerase I specific transcription initiation factor
YLR141w	*RRN5*	RNA polymerase I specific transcription initiation factor
YBL014c	*RRN6*	RNA polymerase I specific transcription initiation factor
YJL025w	*RRN7*	RNA polymerase I specific transcription initiation factor
YMR270c	*RRN9*	RNA polymerase I specific transcription initiation factor
YER148w	*SPT15*	TFIID and TFIIIB subunit
YPR186c	*TFC2*	TFIIIA subunit (transcription initiation factor)
YGR246c	*BRF1*	TFIIIB subunit, 70K
YNL039w	*TFC5*	TFIIIB subunit, 90K
YGR047c	*TFC4*	TFIIIC subunit (transcription initiation factor), 131K
YAL001c	*TFC3*	TFIIIC subunit (transcription initiation factor), 138K
YBR123c	*TFC1*	TFIIIC subunit (transcription initiation factor), 95K
YBR049c	*REB1*	transcription factor

rRNA processing

YOR048c	*RAT1*	5'-3' exoribonuclease
YBR142w	*MAK5*	ATP-dependent RNA helicase
YNL112w	*DBP2*	ATP-dependent RNA helicase of DEAD box family
YFL002c	*SPB4*	ATP-dependent RNA helicase of DEAH box family
YPL029w	*SUV3*	ATP-dependent RNA helicase, mitochondrial
YJL033w	*HCA4*	can suppress the U14 snoRNA rRNA processing function
YMR239c	*RNT1*	double-stranded ribonuclease
YDL014w	*NOP1*	fibrillarin
YGL097w	*SRM1*	GDP/GTP exchange factor for Gsp1p/Gsp2p
YMR235c	*RNA1*	GTPase-activating protein
YLL011w	*SOF1*	involved in 18S pre-rRNA production
YGR158c	*MTR3*	involved in mRNA transport
YGL173c	*KEM1*	multifunctional nuclease
YGR159c	*NSR1*	nuclear localization sequence binding protein
YPL043w	*NOP4*	nucleolar protein
YDR432w	*NPL3*	nucleolar protein
YCR018c	*SRD1*	nucleolar protein
YHR089c	*GAR1*	nucleolar rRNA processing protein
YLR223c	*IFH1*	pre-rRNA processing machinery control protein
YMR229c	*RRP5*	processing of pre-ribosomal RNA
YNL221c	*POP1*	protein component of ribonuclease P and ribonuclease MRP
YHR065c	*RRP3*	required for maturation of the 35S primary transcript
YLL008w	*DRS1*	RNA helicase of the DEAD box family
YPL266w	*DIM1*	rRNA (adenine-N6,N6-)-dimethyltransferase
YOR201c	*PET56*	rRNA (guanosine-2'-O-)-methyltransferase
YPL021w	*SRD2*	similarity to Srd1p
YHL034c	*SBP1*	single-strand nucleic acid binding protein
YLR059c	*YNT20*	suppressor of *rna12 yme2*

other rRNA-transcription activities

YNL216w	*RAP1*	DNA-binding protein with repressor and activator activity
YMR302c	*PRP12*	involved in early maturation of pre-rRNA

tRNA synthesis

YGL019w	*CKB1*	casein kinase II, β subunit
YOR224c	*RPB8*	DNA-directed RNA polymerase I, II, III 16K subunit
YPR187w	*RPO26*	DNA-directed RNA polymerase I, II, III 18K subunit
YBR154c	*RPB5*	DNA-directed RNA polymerase I, II, III 25K subunit
YHR143w-a	*RPC10*	DNA-directed RNA polymerase I, II, III 7.7K subunit
YOR210w	*RPB10*	DNA-directed RNA polymerase I, II, III 8.3K subunit
YPR110c	*RPC40*	DNA-directed RNA polymerase I, III 40K subunit
YNL113w	*RPC19*	DNA-directed RNA polymerase I, III, 16K subunit
YOR207c	*RPC128*	DNA-directed RNA polymerase III, 130K subunit
YOR116c	*RPO31*	DNA-directed RNA polymerase III, 160K subunit
YKL144c	*RPC25*	DNA-directed RNA polymerase III, 25K subunit
YNL151c	*RPC31*	DNA-directed RNA polymerase III, 31K subunit
YNR003c	*RPC34*	DNA-directed RNA polymerase III, 34K subunit
YDL150w	*RPC53*	DNA-directed RNA polymerase III, 47K subunit
YPR190c	*RPC82*	DNA-directed RNA polymerase III, 82K subunit
YER148w	*SPT15*	TFIID and TFIIIB subunit
YPR186c	*TFC2*	TFIIIA subunit (transcription initiation factor)
YGR246c	*BRF1*	TFIIIB subunit, 70K
YNL039w	*TFC5*	TFIIIB subunit, 90K
YGR047c	*TFC4*	TFIIIC subunit (transcription initiation factor), 131K
YAL001c	*TFC3*	TFIIIC subunit (transcription initiation factor), 138K
YBR123c	*TFC1*	TFIIIC subunit (transcription initiation factor), 95K

tRNA processing

YIL075c	*SEN3*	26S proteasome regulatory subunit
YGL097w	*SRM1*	GDP/GTP exchange factor for Gsp1p/Gsp2p
YMR235c	*RNA1*	GTPase activating protein

YHR006w	*STP2*	involved in pre-tRNA splicing
YLR375w	*STP3*	involved in pre-tRNA splicing and in uptake of branched-chain amino acids
YDL048c	*STP4*	involved in pre-tRNA splicing and in uptake of branched-chain amino acids
YCL017c	*NFS1*	involved in tRNA processing and mitochondrial metabolsim
YNR034w	*SOL1*	multicopy suppressor of *los1-1*
YCR073w-a	*SOL2*	multicopy suppressor of *los1-1*
YMR047c	*NUP116*	nuclear pore protein
YGL092w	*NUP145*	nuclear pore protein
YJR042w	*NUP85*	nuclear pore protein
YLR430w	*SEN1*	positive effector of tRNA-splicing endonuclease
YAL043c	*PTA1*	pre-tRNA processing protein
YKL205w	*LOS1*	pre-tRNA splicing protein
YDR463w	*STP1*	pre-tRNA splicing protein
YNL221c	*POP1*	protein component of ribonuclease P and ribonuclease MRP
YML091c	*RPM2*	ribonuclease P, mitochondrial
YFR004w	*MPR1*	strong similarity to *S. pombe* pad1 protein
YGR248w	*SOL4*	strong similarity to Sol3p
YJL087c	*TRL1*	tRNA ligase
YLR105c	*SEN2*	tRNA splicing endonuclease β subunit
YHR163w	*SOL3*	weak multicopy suppressor of *los1-1*

tRNA modification

YGR204w	*ADE3*	C1-tetrahydrofolate synthase, cytoplasmic
YMR283c	*RIT1*	initiator tRNA phosphoribosyl-transferase
YDR120c	*TRM1*	N2,N2-dimethylguanine tRNA methyltransferase
YGL105w	*G4P1*	protein with specific affinity for G4 quadruplex nucleic acids
YPL212c	*PUS1*	pseudouridine synthase 1
YGL063w	*PUS2*	pseudouridine synthase 2
YFR010w		similarity to *C. elegans* tRNA-guanine transglycosylase
YBL013w		similarity to methionyl-tRNA formyltransferase
YOR274w	*MOD5*	tRNA isopentenyltransferase
YER168c	*CCA1*	tRNA nucleotidyltransferase

other tRNA-transcription activities

YOR061w	*CKA2*	casein kinase II α' subunit
YOR039w	*CKB2*	casein kinase II β' subunit
YJL041w	*NSP1*	nuclear pore protein

mRNA synthesis

general transcription activities

YDL108w	*KIN28*	cyclin-dependent ser/thr protein kinase
YIL143c	*SSL2*	DNA helicase
YER171w	*RAD3*	DNA helicase/ATPase
YOR224c	*RPB8*	DNA-directed RNA polymerase I, II, III 16K subunit
YPR187w	*RPO26*	DNA-directed RNA polymerase I, II, III 18K subunit
YBR154c	*RPB5*	DNA-directed RNA polymerase I, II, III 25K subunit
YHR143w-a	*RPC10*	DNA-directed RNA polymerase I, II, III 7.7K subunit
YOR210w	*RPB10*	DNA-directed RNA polymerase I, II, III 8.3K subunit
YOL051w	*GAL11*	DNA-directed RNA polymerase II holoenzyme and Kornberg's mediator (SRB) subcomplex subunit
YDR443c	*SCA1*	DNA-directed RNA polymerase II holoenzyme and Kornberg's mediator (SRB) subcomplex subunit
YHR041c	*SRB2*	DNA-directed RNA polymerase II holoenzyme and Kornberg's mediator (SRB) subcomplex subunit
YER022w	*SRB4*	DNA-directed RNA polymerase II holoenzyme and Kornberg's mediator (SRB) subcomplex subunit
YGR104c	*SRB5*	DNA-directed RNA polymerase II holoenzyme and Kornberg's mediator (SRB) subcomplex subunit
YDR308c	*SRB7*	DNA-directed RNA polymerase II holoenzyme and Kornberg's mediator (SRB) subcomplex subunit
YCR081w	*SRB8*	DNA-directed RNA polymerase II holoenzyme and Kornberg's mediator (SRB) subcomplex subunit
YNL025c	*SSN8*	DNA-directed RNA polymerase II holoenzyme and Kornberg's mediator (SRB) subcomplex subunit, cyclin C homologue
YLR071c	*RGR1*	DNA-directed RNA polymerase II holoenzyme subunit
YBR253w	*SRB6*	DNA-directed RNA polymerase II suppressor protein
YOL005c	*RPB11*	DNA-directed RNA polymerase II, 13.6K subunit
YGL070c	*RPB9*	DNA-directed RNA polymerase II, 14.2K subunit
YOR151c	*RPB2*	DNA-directed RNA polymerase II, 140K subunit
YDR404c	*RPB7*	DNA-directed RNA polymerase II, 19K subunit
YDL140c	*RPO21*	DNA-directed RNA polymerase II, 215K subunit
YJL140w	*RPB4*	DNA-directed RNA polymerase II, 32K subunit
YFL036w	*RPO41*	DNA-directed RNA polymerase, mitochondrial
YIL021w	*RPB3*	DNA-directed RNA-polymerase II, 45K
YIL128w	*MMS19*	involved in repair and RNA polymerase transcription
YMR228w	*MTF1*	RNA polymerase specific factor, mitochondrial
YPR056w		similarity to human transcription factor BTF2/TFIIH subunit p34
YPL046c		strong similarity to human DNA-directed RNA polymerase II elongation factor SIII, p15
YDR045c		strong similarity to *S. acidocaldarius* transcription elongation factor tfs
YKL058w	*TOA2*	TFIIA subunit (transcription initiation factor), 13.5K
YOR194c	*TOA1*	TFIIA subunit (transcription initiation factor), 32K
YPR086w	*SUA7*	TFIIB subunit (transcription initiation factor), factor E
YER148w	*SPT15*	TFIID and TFIIIB subunit
YGR274c	*TAF145*	TFIID subunit (TBP-associated factor), 145K
YMR236w	*TAF17*	TFIID subunit (TBP-associated factor), 17K
YML098w	*TAF19*	TFIID subunit (TBP-associated factor), 19K
YDR167w	*TAF23*	TFIID subunit (TBP-associated factor), 23K
YML015c	*TAF40*	TFIID subunit (TBP-associated factor), 40K
YGL112c	*TAF60*	TFIID subunit (TBP-associated factor), 60K
YDR145w	*TAF61*	TFIID subunit (TBP-associated factor), 61K
YMR227c	*TAF67*	TFIID subunit (TBP-associated factor), 67K
YBR198c	*TAF90*	TFIID subunit (TBP-associated factor), 90K
YKR062w	*TFA2*	TFIIE subunit (transcription initiation factor), 43K
YKL028w	*TFA1*	TFIIE subunit (transcription initiation factor), 66K
YGR186w	*TFG1*	TFIIF subunit (transcription initiation factor), 105K
YPL129w	*ANC1*	TFIIF subunit (transcription initiation factor), 30K
YGR005c	*TFG2*	TFIIF subunit (transcription initiation factor), 54K
YDR311w	*TFB1*	TFIIH subunit (transcription initiation factor), 75K
YPR025c	*CCL1*	TFIIH subunit (transcription initiation factor), cyclin C component
YLR005w	*SSL1*	TFIIH subunit (transcription initiation factor), factor B
YPL122c	*TFB2*	TFIIH subunit (transcription/repair factor)
YDR460w	*TFB3*	TFIIH subunit (transcription/repair factor)
YGL043w	*DST1*	TFIIS (transcription elongation factor)
YGR063c	*SPT4*	transcription initiation protein
YML010w	*SPT5*	transcription initiation protein
YGR116w	*SPT6*	transcription initiation protein

transcriptional control

YOR259c	*CRL13*	26S proteasome subunit
YJL115w	*ASF1*	anti-silencing protein
YDL197c	*ASF2*	anti-silencing protein
YDR173c	*ARG82*	arginine metabolism transcription factor
YKL112w	*ABF1*	ARS-binding factor
YDR123c	*INO2*	basic helix-loop-helix (BHLH) transcription factor
YOL108c	*INO4*	basic helix-loop-helix transcription factor
YOL067c	*RTG1*	basic helix-loop-helix transcription factor
YOR344c	*TYE7*	basic helix-loop-helix transcription factor
YBL103c	*RTG3*	bHLH/zip transcription factor
YHL009c		bZip DNA binding protein
YGL209w	*MIG2*	C2H2 zinc-finger protein
YKL190w	*CNB1*	calcineurin B, regulatory subunit
YML112w	*CTK3*	carboxy terminal domain (CTD) kinase, γ subunit
YJL006c	*CTK2*	carboxy-terminal domain (CTD) kinase , β subunit
YKL139w	*CTK1*	carboxy-terminal domain (CTD) kinase, α subunit
YIL035c	*CKA1*	casein kinase II, catalytic α subunit
YGL237c	*HAP2*	CCAAT-binding factor subunit
YBL021c	*HAP3*	CCAAT-binding factor subunit
YKL109w	*HAP4*	CCAAT-binding factor subunit
YOR358w	*HAP5*	CCAAT-binding factor subunit
YJR122w	*CAF17*	*CCR4*-associated factor
YKR036c	*CAF4*	*CCR4*-associated factor
YBR215w	*HPC2*	cell-cycle regulatory protein
YDL220c	*CDC13*	cell division control protein
YDR073w	*SNF11*	component of SWI/SNF transcription activator complex
YNR023w	*SNF12*	component of SWI/SNF transcription activator complex
YOR290c	*SNF2*	component of SWI/SNF transcription activator complex
YBR289w	*SNF5*	component of SWI/SNF transcription activator complex
YPL016w	*SWI1*	component of SWI/SNF transcription activator complex
YCR042c	*TSM1*	component of TAF(II) complex
YPL177c	*CUP9*	copper homeostasis protein
YGL166w	*CUP2*	copper-dependent transcription factor
YNL167c	*SKO1*	cre-binding bzip protein
YPL042c	*SSN3*	cyclin-dependent ser/thr protein kinase
YLR176c	*RFX1*	DNA binding protein
YDR217c	*RAD9*	DNA repair checkpoint protein
YNL216w	*RAP1*	DNA-binding protein with repressor and activator activity
YLR418c	*CDC73*	DNA-directed RNA polymerase II accessory protein
YBR279w	*PAF1*	DNA-directed RNA polymerase II regulator
YGL208w	*SIP2*	dominant suppressor of some ts mutations *rpo21* and *prp4*
YOR047c	*STD1*	dosage-dependent modulator of glucose repression
YER159c	*NCB1*	functional homologue of human NC2α
YDR397c	*YNC2β*	functional homologue of human NC2β/Dr1
YPL037c	*EGD1*	*GAL4* DNA-binding enhancer protein
YJL110c	*GZF3*	GATA zinc-finger protein 3
YIL038c	*NOT3*	general negative regulator of transcription, subunit 3
YBR112c	*CYC8*	general repressor of transcription
YCR084c	*TUP1*	general transcription repressor
YDR176w	*NGG1*	general transcriptional adaptor or co-activator
YBR045c	*GIP1*	Glc7p-interacting protein
YNL236w	*SIN4*	global regulator protein
YHL025w	*SNF6*	global transcription activator
YER027c	*GAL83*	glucose repression protein
YNL199c	*GCR2*	glycolytic genes transcriptional activator
YGL073w	*HSF1*	heat shock transcription factor
YPR065w	*ROX1*	heme-dependent transcriptional repressor of hypoxic genes
YDR006c	*SOK1*	high copy suppressor of a cyclic AMP-dependent protein kinase mutant
YMR070w	*HMS1*	high-copy suppressor of *mot1 spt3* synthetic lethality
YGR252w	*GCN5*	histone acetyltransferase
YDR225w	*HTA1*	histone H2A
YBL003c	*HTA2*	histone H2A.2
YDR224c	*HTB1*	histone H2B
YBL002w	*HTB2*	histone H2B.2
YBR010w	*HHT1*	histone H3
YNL031c	*HHT2*	histone H3
YBR009c	*HHF1*	histone H4
YNL030w	*HHF2*	histone H4
YBL008w	*HIR1*	histone transcription regulator
YOR038c	*HIR2*	histone transcription regulator
YER161c	*SPT2*	HMG-like chromatin protein
YDL106c	*GRF10*	homeodomain protein
YDR138w	*HPR1*	hyperrecombination protein related to Top1p
YJL089w	*SIP4*	interacts with SNF1 protein kinase
YKL032c	*IXR1*	intrastrand crosslink recognition protein and transcription factor
YBR081c	*SPT7*	involved in alteration of transcription start site selection
YKR072c	*SIS2*	involved in cell cycle-specific gene expression
YNL251c	*NRD1*	involved in regulation of nuclear pre-mRNA abundance
YIL046w	*MET30*	involved in regulation of sulphur assimilation genes
YDL153c	*SAS10*	involved in silencing
YMR127c	*SAS2*	involved in silencing at HMR
YGR097w	*ASK10*	involved in Skn7p-dependent transcription
YLR039c	*RIC1*	involved in transcription of ribosomal proteins and ribosomal RNA
YGL071w	*RCS1*	iron-regulated transcriptional repressor
YJR060w	*CBF1*	kinetochore protein
YDR159w	*SAC3*	leucine permease transcriptional regulator
YGR100w	*MIC1*	Mac1p interacting protein
YBR297w	*MAL33*	maltose fermentation regulatory protein
YGR288w	*MAL13*	maltose pathway regulatory protein
YCR040w	*MATα1*	mating type regulatory protein (expressed copy at MAT locus)
YCR039c	*MATα2*	mating type regulatory protein (expressed copy at MAT locus)
YCL067c	*α2*	mating type regulatory protein (silenced copy at HML locus)
YHL027w	*RIM101*	meiotic regulatory protein
YLR216c	*CPR6*	member of the cyclophilin family
YJR032w	*CPR7*	member of the cyclophilin family
YLR136c	*TIS11*	member of the inducible ccch zinc-finger family
YOL148c	*SPT20*	member of the TBP class of SPT proteins that alter transcription site selection
YMR021c	*MAC1*	metal binding activator
YLR131c	*ACE2*	metallothionein expression activator
YBR026c	*(MRF1)*	mitochondrial respiratory function protein
YML051w	*GAL80*	negative regulator for expression of galactose-induced genes
YGL151w	*NUT1*	negative regulator of HO endonuclease
YPR168w	*NUT2*	negative regulator of HO endonuclease
YKL185w	*ASH1*	negative regulator of HO expression
YHL020c	*OPI1*	negative regulator of phospholipid biosynthesis pathway
YDR464w	*SPP41*	negative regulator of PRP3 and PRP4 gene expression
YNL076w	*MKS1*	negative regulator of RAS-cAMP pathway
YDR207c	*UME6*	negative transcriptional regulator
YGL221c	*NIF3*	Ngg1p-interacting factor 3
YGL115w	*SNF4*	nuclear regulatory protein
YCR093w	*CDC39*	nuclear transmembrane protein
YML065w	*ORC1*	origin recognition complex, 104K subunit
YNL261w	*ORC5*	origin recognition complex, 50K subunit
YHR118c	*ORC6*	origin recognition complex, 50K subunit
YPR162c	*ORC4*	origin recognition complex, 56K subunit
YLL004w	*ORC3*	origin recognition complex, 62K subunit
YBR060c	*RRR1*	origin recognition complex, 72K subunit
YAL051w	*OAF1*	peroxisome proliferating transcription factor
YOR363c	*PIP2*	peroxisome proliferating transcription factor
YBL005w	*PDR3*	pleiotropic drug resistance regulatory protein
YKL015w	*PUT3*	positive activator of the proline utilization pathway
YGL192w	*IME4*	positive transcription factor for *IME2*

YDR017c	*KCS1*	potential transcription factor of the BZIP type
YGL025c	*PGD1*	probable transcription factor
YER108c	*FLO8*	probable transcriptional activator of Flo1p
YGL194c	*RTL1*	putative deacetylase
YJL103c		putative regulatory protein
YBL066c	*SEF1*	putative transcription factor
YHR056c		putative transcription regulator
YBR275c	*RIF1*	Rap1p-interacting factor 1
YKL038w	*RGT1*	regulator of glucose induced genes
YDR257c	*RMS1*	regulatory protein
YDR392w	*SPT3*	regulatory protein
YMR016c	*SOK2*	regulatory protein in the PKA signal transduction pathway
YCL055w	*KAR4*	regulatory protein required for pheromone induction of karyogamy genes
YNR052c	*POP2*	required for glucose derepression
YMR179w	*SPT21*	required for normal transcription at a number of loci
YKL093w	*MBR1*	required for optimal growth on glycerol
YOL068c	*HST1*	silencing protein
YOR025w	*HST3*	silencing protein
YBL052c	*SAS3*	silencing protein
YKR101w	*SIR1*	silencing regulatory protein
YDL042c	*SIR2*	silencing regulatory protein
YLR442c	*SIR3*	silencing regulatory protein
YDR227w	*SIR4*	silencing regulatory protein
YKL070w		similarity to *B. subtilis* transcriptional regulatory protein
YDL070w		similarity to bromodomain protein BDF1
YIL131c	*FKH1*	similarity to *D. melanogaster* fork head protein
YJL056c		similarity to developmental control proteins
YBR061c		similarity to *E. coli* ftsJ protein
YJR147w		similarity to heat-shock transcription factors
YGR249w	*MGA1*	similarity to heat-shock transcription factors
YPL015c	*HST2*	similarity to Hst1p and Sir2p
YDR191w	*HST4*	similarity to Hst3p,Hst1p and Sir2p
YNL107w		similarity to human AF-9 protein
YHR193c	*EGD2*	similarity to human α-NAC
YAR003w	*FUN16*	similarity to human RB protein binding protein
YER169w		similarity to human retinoblastoma binding protein 2
YJR119c		similarity to human retinoblastoma binding protein 2
YLR116w		similarity to human ZFM1 protein and mouse CW17R protein
YJL206c		similarity to hypothetical protein YIL130p and Put3p
YOL133w		similarity to Lotus RING-finger protein
YER130c		similarity to Msn2p and Msn4p
YER184c		similarity to multidrug resistance protein PDR3
YIL050w		similarity to *N. crassa* regulatory protein preg(+)
YLR013w		similarity to nitrogen regulatory proteins
YDR334w		similarity to nuclear Sth1p, Snf2p and related proteins
YPR115w		similarity to probable transcription factor Ask10p, and to hypothetical proteins YNL047c and YIL105c
YIL130w		similarity to Put3p and to hypothetical protein YJL206c
YCR020c	*PET18*	similarity to regulatory protein
YBR182c		similarity to Rlm1p,Mcm1p, and hMEF2
YKR008w		similarity to *S. pombe* bromodomain protein
YGL150c		similarity to Snf2p and human SNF2α
YLR228c		similarity to transcription activator Lys14p
YLL054c		similarity to transcription factor Pip2p
YLR266c		similarity to transcription factors
YLR278c		similarity to transcription factors
YDR303c		similarity to transcriptional regulator proteins
YMR019w	*STB4*	SIN3 binding protein
YHR178w	*STB5*	SIN3 binding protein
YKL072w	*STB6*	SIN3 binding protein
YNL309w	*STB1*	Sin3p binding protein
YMR053c	*STB2*	Sin3p binding protein
YLR399c	*BDF1*	sporulation protein
YNL204c	*SPS18*	sporulation-specific zinc-finger protein
YMR037c		stress responsive regulatory protein
YDR026c		strong similarity to DNA-binding protein Reb1p
YDR252w	*BTT1*	strong similarity to Egd1p and to human Btf3 protein
YHR211w	*FLO5*	strong similarity to Flo1p
YOR304w		strong similarity to human Snf2p homologue
YOL055c		strong similarity to hypothetical proteins YPL258c, YPR121w, similarity to *B. subtilis* transcriptional activator tenA
YFL052w		strong similarity to Mal63p, Mal23p and Mal33p
YPR196w		strong similarity to regulatory protein Mal63p
YBR245c		strong similarity to SNF2/SWI2 DNA binding regulatory protein
YDR253c		strong similarity to zinc-finger proteins
YML081w		strong similarity to ZMS1 protein
YEL056w	*HAT2*	subunit of the major yeast histone acetyltransferase
YOR213c		subunit of the RSC complex
YCR052w	*RSC6*	subunit of the RSC complex
YFR037c	*RSC8*	subunit of the RSC complex
YLR321c	*SFH1*	subunit of the RSC complex
YIL126w	*STH1*	subunit of the RSC complex
YGL254w	*FZF1*	sulphite resistance protein
YNL222w	*SSU72*	suppressor of cs mutant *sua7*
YDR310c	*SUM1*	suppressor of *mar1-1*
YKR099w	*BAS1*	transcription factor
YDL165w	*CDC36*	transcription factor
YLR098c	*CHA4*	transcription factor
YPL248c	*GAL4*	transcription factor
YFL031w	*HAC1*	transcription factor
YLR256w	*HAP1*	transcription factor
YCR065w	*HCM1*	transcription factor
YLR451w	*LEU3*	transcription factor
YGL013c	*PDR1*	transcription factor
YKL043w	*PHD1*	transcription factor
YFR034c	*PHO4*	transcription factor
YBR049c	*REB1*	transcription factor
YBL093c	*ROX3*	transcription factor
YOR140w	*SFL1*	transcription factor
YLR055c	*SPT8*	transcription factor
YER111c	*SWI4*	transcription factor
YDR146c	*SWI5*	transcription factor
YLR182w	*SWI6*	transcription factor
YFL021w	*GAT1*	transcription factor for nitrogen regulation
YER040w	*GLN3*	transcription factor for positive nitrogen regulation
YMR042w	*ARG80*	transcription factor involved in arginine metabolism
YML099c	*ARG81*	transcription factor involved in arginine metabolism
YMR280c	*CAT8*	transcription factor involved in gluconeogenesis
YGL181w	*GTS1*	transcription factor of the Gcs1p/Glo3p/Sps18p family
YMR043w	*MCM1*	transcription factor of the MADS box family
YPL089c	*RLM1*	transcription factor of the MADS box family
YLR014c	*PPR1*	transcription factor regulating pyrimidine pathway
YJR094c	*IME1*	transcription factor required for sporulation
YHR206w	*SKN7*	transcription factor with similarity to Hsf1p
YDL056w	*MBP1*	transcription factor, subunit of the MBF factor
YNL330c	*RPD3*	transcription modifier protein
YOL004w	*SIN3*	transcription regulatory protein
YJL127c	*SPT10*	transcription regulatory protein
YJL176c	*SWI3*	transcription regulatory protein
YPL082c	*MOT1*	transcriptional accessory protein
YDR423c	*CAD1*	transcriptional activator
YOR028c	*CIN5*	transcriptional activator
YPL075w	*GCR1*	transcriptional activator
YOL116w	*MSN1*	transcriptional activator
YKL062w	*MSN4*	transcriptional activator
YHR084w	*STE12*	transcriptional activator
YIR023w	*DAL81*	transcriptional activator for allantoin and GABA catabolic genes
YNL314w	*DAL82*	transcriptional activator for allantoin catabolic genes
YDL170w	*UGA3*	transcriptional activator for GABA catabolic genes
YML007w	*YAP1*	transcriptional activator involved in oxidative stress response
YEL009c	*GCN4*	transcriptional activator of amino acid biosynthetic genes
YDR034c	*LYS14*	transcriptional activator of lysine pathway genes
YIR017c	*MET28*	transcriptional activator of sulphur amino acid metabolism
YNL103w	*MET4*	transcriptional activator of sulphur metabolism
YPR104c	*FHL1*	transcriptional activator of the forkhead/hnf3 family
YDR448w	*ADA2*	transcriptional adaptor
YMR039c	*SUB1*	transcriptional coactivator
YAL021c	*CCR4*	transcriptional regulator
YER164w	*CHD1*	transcriptional regulator
YIL084c	*SDS3*	transcriptional regulator
YNL012w	*SPO1*	transcriptional regulator involved in sporulation
YGL035c	*MIG1*	transcriptional repressor
YER068w	*MOT2*	transcriptional repressor
YKR034w	*DAL80*	transcriptional repressor for allantoin and GABA catabolic genes
YMR182c	*RGM1*	transcriptional repressor protein
YCL054w		transcriptional silencing protein
YBR083w	*TEC1*	Ty transcription activator
YOR337w	*TEA1*	TY1 enhancer activator
YER151c	*UBP3*	ubiquitin-specific proteinase
YER088c		weak similarity human transforming proteins (B-myb)
YBR240c		weak similarity to regulatory proteins
YML076c		weak similarity to transcription factor
YCR106w		weak similarity to transcription factor Pip2p
YBR033w		weak similarity to transcription factors
YKR064w		weak similarity to transcription factors
YOL089c		weak similarity to transcription factors
YKL222c		weak similarity to transcription factors, similarity to finger proteins YOR162c, YOR172w and YLR266c
YBL054w		weak similarity transforming protein (B-myb)
YER122c	*GLO3*	zinc-finger protein
YDR216w	*ADR1*	zinc-finger transcription factor
YGR044c	*RME1*	zinc-finger transcription factor

chromatin modification

YBR195c	*MSI1*	chromatin assembly complex, subunit p50
YML102w	*CAC2*	chromatin assembly complex, subunit p60
YPR018w	*RLF2*	chromatin assembly complex, subunit p90
YGL207w	*SPT16*	general chromatin factor
YGR252w	*GCN5*	histone acetyltransferase
YGL194c	*RTL1*	putative deacetylase
YHR119w	*YTX1*	regulatory protein
YBR238c		strong similarity to general chromatin factor Spt16p
YFR004w	*MPR1*	strong similarity to *S. pombe* pad1 protein
YOR213c		subunit of the RSC complex
YCR052w	*RSC6*	subunit of the RSC complex
YFR037c	*RSC8*	subunit of the RSC complex
YLR321c	*SFH1*	subunit of the RSC complex
YIL126w	*STH1*	subunit of the RSC complex
YGR274c	*TAF145*	TFIID subunit (TBP-associated factor), 145K
YGR063c	*SPT4*	transcription initiation protein
YML010w	*SPT5*	transcription initiation protein
YGR116w	*SPT6*	transcription initiation protein

mRNA processing (splicing)

YBL074c	*AAR2*	A1 cistron splicing factor
YHL038c	*CBP2*	apo-cytochrome b pre-mRNA processing protein 2
YPL119c	*DBP1*	ATP-dependent RNA helicase
YOR204w	*DED1*	ATP-dependent RNA helicase
YBR142w	*MAK5*	ATP-dependent RNA helicase
YPL029w	*SUV3*	ATP-dependent RNA helicase, mitochondrial
YNL286w	*CUS2*	cold sensitive U2 snRNA suppressor
YIR021w	*MRS1*	involved in mitochondrial RNA splicing of COB mRNA
YOR148c	*SPP2*	involved in pre-mRNA processing
YPR057w	*BRR1*	involved in snRNP biogenesis
YLR382c	*NAM2*	leucine-tRNA ligase, mitochondrial
YNL210w	*MER1*	meiotic recombination protein
YHR086w	*NAM8*	meiotic recombination protein
YMR023c	*MSS1*	mitochondrial GTPase involved in expression of *COX1*
YLL036c	*PRP19*	non-snRNP sliceosome component required for DNA repair
YBR237w	*PRP5*	pre-mRNA processing RNA-helicase
YKL074c	*MUD2*	pre-mRNA splicing factor
YDL043c	*PRP11*	pre-mRNA splicing factor
YJL203w	*PRP21*	pre-mRNA splicing factor
YER013w	*PRP22*	pre-mRNA splicing factor
YMR268c	*PRP24*	pre-mRNA splicing factor
YGR075c	*PRP38*	pre-mRNA splicing factor
YML046w	*PRP39*	pre-mRNA splicing factor
YDL030w	*PRP9*	pre-mRNA splicing factor (snRNA-associated protein)
YDR088c	*SLU7*	pre-mRNA splicing factor affecting 3' splice site choice
YDR243c	*PRP28*	pre-mRNA splicing factor RNA helicase of DEAD box family
YGR091w	*PRP31*	pre-mRNA splicing protein
YDL044c	*MTF2*	protein involved in mRNA splicing and protein synthesis, mitochondrial
YBR091c	*MRS5*	regulator of intron splicing, mitochondrial
YDR028c	*REG1*	regulatory subunit for protein phosphatase Glc7p
YDR478w	*SNM1*	RNA binding protein of RNase MRP
YDR194c	*MSS116*	RNA helicase of the DEAD box family, mitochondrial
YER172c	*BRR2*	RNA helicase-related protein
YOR334w	*MRS2*	RNA splicing protein and member of the mitochondrial carrier family (MCF)
YJL133w	*MRS3*	RNA splicing protein and member of the mitochondrial carrier family (MCF)
YKR052c	*MRS4*	RNA splicing protein and member of the mitochondrial carrier family (MCF)
YKR086w	*PRP16*	RNA-dependent ATPase
YNR011c	*PRP2*	RNA-dependent ATPase of DEAH box family
YHR004c-a		similarity to mitochondrial splicing protein Mrs5p
YKR024c		similarity to pre-mRNA processing protein Prp5p
YGR074w	*SMD1*	snRNA-associated protein
YPR182w	*SMX3*	snRNA-associated protein of the SM family
YBR055c	*PRP6*	snRNP(U4/U6)-associated splicing factor
YKL012w	*PRP40*	splicing factor
YPR134w	*MSS18*	splicing protein
YGR222w	*PET54*	splicing protein and translational activator, mitochondrial
YKL078w		strong similarity to ATP-dependent RNA helicases
YOR159c	*SME1*	strong similarity to human small nuclear ribonucleoprotein E
YLR275w		strong similarity to human snRNP subunit D2 involved in systemic lupus erythematosus
YPL060w		strong similarity to Mrs2p
YGL120c		strong similarity to Prp22p
YLR147c	*SMD3*	strong similarity to small nuclear ribonucleoprotein D3
YIR009w	*MSL1*	strong similarity to small nuclear ribonucleoproteins
YIL061c	*SNP1*	U1 small nuclear ribonucleoprotein
YBR119w	*MUD1*	U1 snRNP-specific A protein
YMR240c	*CUS1*	U2 snRNP protein

YPR178w	*PRP4*	U4/U6 snRNP 52K protein
YGR006w	*PRP18*	U5 snRNA-associated protein
YHR165c	*PRP8*	U5 snRNP protein, pre-mRNA splicing factor
YER112w	*USS1*	U6 snRNA associated protein

mRNA processing (5'-end, 3'-end processing and mRNA degradation)

YJL209w	*CBP1*	apo-cytochrome b pre-mRNA processing protein
YBR120c	*CBP6*	apo-cytochrome b pre-mRNA processing protein
YPL178w	*SAE1*	cell cycle block in meiotic prophase
YGL094c	*PAN2*	component of Pab1p-stimulated poly(A) ribonuclease
YMR061w	*RNA14*	component of pre-mRNA 3'-end processing factor
YGL044c	*RNA15*	component of pre-mRNA 3'-end processing factor
YJR093c	*FIP1*	component of pre-mRNA polyadenylation factor
YKL025c	*PAN3*	component of the Pab1p-dependent poly (A)
YOL149w	*DCP1*	component of the yeast decapping enzyme
YGL097w	*SRM1*	GDP/GTP exchange factor for Gsp1p/Gsp2p
YGR158c	*MTR3*	involved in mRNA transport
YBR236c	*ABD1*	methyltransferase
YGL130w	*CEG1*	mRNA guanylyltransferase (mRNA capping enzyme, α subunit)
YER165w	*PAB1*	mRNA polyadenylate-binding protein
YLR277c	*BRR5*	mRNA processing protein
YKR002w	*PAP1*	poly(A) polymerase
YGL122c	*NAB2*	poly(A)-binding protein
YPL190c	*NAB3*	polyadenylated RNA-binding protein
YDR301w	*CFT1*	pre-mRNA 3'-end processing factor
YJR017c	*ESS1*	processing/termination factor 1
YDR195w	*REF2*	RNA 3'-end formation protein
YOR179c		similarity to BRR5 protein
YOR319w	*HSH49*	similarity to human SAP49 and RNA-binding proteins
YHR015w		similarity to PES4 PAB-like protein
YOR159c	*SME1*	strong similarity to human small nuclear ribonucleoprotein E

other mRNA-transcription activities

YGL171w	*ROK1*	ATP-dependent RNA helicase
YOL006c	*TOP1*	DNA topoisomerase I
YGL251c	*HFM1*	DNA/RNA helicase
YNL037c	*IDH1*	isocitrate dehydrogenase (NAD$^+$) subunit 1, mitochondrial
YOR136w	*IDH2*	isocitrate dehydrogenase (NAD$^+$) subunit 2, mitochondrial
YNL016w	*PUB1*	major polyadenylated RNA-binding protein of nucleus and cytoplasm
YHR170w	*NMD3*	nonsense-mediated mRNA decay protein
YMR064w	*AEP1*	nuclear control of ATPase messenger RNA expression protein
YOL123w	*HRP1*	polyadenylated RNA-binding protein
YLR067c	*PET309*	required for stability and translation of COX1 mRNA
YAL016w	*TPD3*	ser/thr phosphatase 2A, regulatory subunit A
YOL042w		similarity to CCR4 protein
YER146w		similarity to human snRNP E
YER029c		similarity to human snRNP-associated protein B
YER028c		similarity to Mig1p
YDR429c		similarity to nuclear RNA binding proteins
YJR127c	*ZMS1*	similarity to regulatory protein ADR1
YDL031w		similarity to RNA helicases
YNL147w		similarity to snRNP proteins
YNL021w		similarity to transcription factor Rpd3p
YPR068c		similarity to transcription factor Rpd3p
YPL213w		similarity to U2 snRNP protein A'
YBL026w	*SNP3*	snRNP-related protein
YDL160c	*DHH1*	strong similarity to RNA helicases of the DEAD box family
YLR147c	*SMD3*	strong similarity to small nuclear ribonucleoprotein D3
YLR139c	*SLS1*	suppresses lethality of SSM4 deletion

RNA transport

YGL097w	*SRM1*	GDP/GTP exchange factor for Gsp1p/Gsp2p
YMR235c	*RNA1*	GTPase activating protein
YOR185c	*GSP2*	GTP-binding protein
YLR293c	*GSP1*	GTP-binding protein of the ras superfamily
YOR160w	*MTR10*	involved in mRNA transport
YGR158c	*MTR3*	involved in mRNA transport
YJL050w	*MTR4*	involved in nucleocytoplasmic transport of mRNA
YKL186c	*MTR2*	mRNA transport protein
YKL068w	*NUP100*	nuclear pore protein
YKL057c	*NUP120*	nuclear pore protein
YKR082w	*NUP133*	nuclear pore protein
YGL092w	*NUP145*	nuclear pore protein
YDR192c	*NUP42*	nuclear pore protein
YGL172w	*NUP49*	nuclear pore protein
YJL061w	*NUP82*	nuclear pore protein
YDR432w	*NPL3*	nucleolar protein
YGL122c	*NAB2*	poly(A)-binding protein
YKL205w	*LOS1*	pre-tRNA splicing protein
YER110c	*KAP123*	RAN-binding protein
YDR002w	*YRB1*	ran-specific GTPase-activating protein
YER107c	*GLE2*	required for nuclear pore complex structure and function
YIR011c	*STS1*	required for transport of Rna15p from the cytoplasm to the nucleus
YDL207w	*GLE1*	RNA export mediator
YLR119w	*SRN2*	suppressor of *rna1-1* mutation

other transcription activities

YGL127c	*SOH1*	allows *hpr1* null mutant to grow at 37°C
YOR113w	*AZF1*	asparagine-rich zinc-finger protein
YPR198w	*SGE1*	drug resistance protein
YBR212w	*NGR1*	glucose-repressible RNA-binding protein
YML027w	*YOX1*	homoeodomain protein
YNL068c	*FKH2*	homology to *D. melanogaster* forkhead protein
YIL030c	*SSM4*	involved in mRNA turnover
YNL282w	*(POP2)*	involved in processsing RNAs
YCR087c-a		nucleic acid-binding protein
YDL051w	*YLA1*	RNA binding protein
YMR164c		similarity to CYC8 protein
YGL014w		similarity to *D. melanogaster* pumilio protein and HTR1 protein
YIR001c		similarity to *D. melanogaster* RNA binding protein
YNL004w	*HRB1*	similarity to Gbp2p
YBR233w		similarity to human hnRNP-E1 protein
YPR031w		similarity to human zinc-finger protein BR140
YCL033c		similarity to *M. capricolum* transcription repressor
YIR005w		similarity to RNA-binding proteins
YMR213w		similarity to *S. pombe* putative transcription factor cdc5
YNL175c		similarity to *S. pombe* Rnp24p
YOR244w		similarity to SAS2 and SAS3 protein
YPR008w		similarity to transcription factor
YPL230w		similarity to transcription factors
YPR013c		similarity to transcription factors
YPR015c		similarity to transcription factors
YNL027w		similarity to zinc-finger proteins
YPL038w		similarity to zinc-finger proteins
YDR169c	*STB3*	SIN3 protein-binding protein
YPR107c		strong similarity to *D. melanogaster* zinc-finger protein
YIL105c		strong similarity to hypothetical protein YNL047c, similarity to hypothetical protein YPR115w and to probable transcription factor Ask10p
YNL255c		strong similarity to nucleic acid-binding proteins
YHR169w		strong similarity to RNA helicase
YGL078c		strong similarity to RNA helicase DBP2 protein
YCR004c	*YCP4*	strong similarity to *S. pombe* protein obr1
YFL017w-a	*SNP2*	strong similarity to snRNP E
YDR451c		strong similarity to Yox1p
YML017w	*PSP2*	suppressor of DNA polymerase α mutation
YKR092c	*SRP40*	suppressor of mutant AC40 of RNA polymerase I and III
YKL005c		weak similarity to *D. melanogaster* transcription elongation factor DmS-II
YNR063w		weak similarity to CYC1/CYP3 transcription activator
YJL124c		weak similarity to human Sm protein G
YDR043c		weak similarity to *K. marxianus* Mig1 and other regulatory proteins
YIR018w		weak similarity to transcription activator Pdr4p
YPR199c		weak similarity to transcription activator Yap1p
YBR239c		weak similarity to transcription factor Put3p
YBR150c		weak similarity to transcription factors
YGR067c		weak similarity to transcription factors
YPL133c		weak similarity to transcription factors
YER116c		zinc-finger protein

Protein synthesis

ribosomal proteins

YGR214w	*NAB1A*	40S ribosomal protein p40 homologue A
YLR048w	*NAB1B*	40S ribosomal protein p40 homologue B
YCR031c	*CRY1*	40S ribosomal protein S14.e
YGL189c	*RPS26A*	40S ribosomal protein S26e.c7
YML009c	*MRPL39A*	60S ribosomal protein, mitochondrial
YDL081c	*RPLA1*	acidic ribosomal protein a1
YLR340w	*RPLA0*	acidic ribosomal protein L10.e
YDL130w	*RPLA3*	acidic ribosomal protein L44prime
YDR382w	*RPLA4*	acidic ribosomal protein L45
YOL039w	*RPLA2*	acidic ribosomal protein P2.β
YOR369c	*RS12*	acidic ribosomal protein S12
YGL068w		probable ribosomal protein L12
YLR325c		putative ribosomal protein L38
YFL034c-a		ribosomal protein
YFR032c-a		ribosomal protein
YLR061w		ribosomal protein
YMR142c		ribosomal protein
YPL183w-a		ribosomal protein
YLR075w	*GRC5*	ribosomal protein
YIL133c	*RP22*	ribosomal protein
YNL069c	*RP23*	ribosomal protein
YOR096w	*RP30*	ribosomal protein
YKL006w	*RPL14A*	ribosomal protein
YHL001w	*RPL14B*	ribosomal protein
YMR242c	*RPL18A*	ribosomal protein
YOR312c	*RPL18B*	ribosomal protein
YDR471w	*RPL27B*	ribosomal protein
YLR344w	*RPL33A*	ribosomal protein
YMR194w	*RPL39A*	ribosomal protein
YDL184c	*RPL47A*	ribosomal protein
YDL133c-a	*RPL47B*	ribosomal protein
YPL198w	*RPL6B*	ribosomal protein
YOL040c	*RPS21*	ribosomal protein
YLR287c-a	*RPS30A*	ribosomal protein
YDL191w	*SOS1*	ribosomal protein
YDL136w	*SOS2*	ribosomal protein
YPL220w	*SSM1A*	ribosomal protein
YGL135w	*SSM1B*	ribosomal protein
YGL123w	*SUP44*	ribosomal protein
YIL148w	*UBI1*	ribosomal protein
YKR094c	*UBI2*	ribosomal protein
YHL015w	*URP2*	ribosomal protein
YML073c	*YL16A*	ribosomal protein
YLR448w	*YL16B*	ribosomal protein
YDR500c	*RPL35B*	ribosomal protein L.37.e
YPR102c	*RPL16A*	ribosomal protein L11.e
YDR418w	*RPL15A*	ribosomal protein L12.e
YEL054c	*RPL15B*	ribosomal protein L12.e
YDL082w		ribosomal protein L13
YKL170w	*MRPL38*	ribosomal protein L14, mitochondrial
YLR029c	*RPL13A*	ribosomal protein L15.e.c12
YMR121c	*RPL13B*	ribosomal protein L15.e.c13
YJL063c	*MRPL8*	ribosomal protein L17, mitochondrial
YKL180w	*RPL20A*	ribosomal protein L17.e
YJL177w	*RPL20B*	ribosomal protein L17.e
YNL301c	*RP28B*	ribosomal protein L18.e
YBL027w	*RPL19A*	ribosomal protein L19.e
YBR084c-a	*RPL19B*	ribosomal protein L19.e
YPL079w	*URP1B*	ribosomal protein L21
YBR191w	*URP1A*	ribosomal protein L21.e
YBL087c	*RPL17A*	ribosomal protein L23.e
YER117w	*RPL17B*	ribosomal protein L23.e
YOL127w	*RPL25*	ribosomal protein L23a.e
YGL031c	*RPL30A*	ribosomal protein L24.e.A
YGR148c	*RPL30B*	ribosomal protein L24.e.B
YGR034w	*RPL33B*	ribosomal protein L26
YHR010w	*RPL27A*	ribosomal protein L27.e
YGL103w	*CYH2*	ribosomal protein L27a.e
YBR031w	*RPL2A*	ribosomal protein L2A
YOR063w	*TCM1*	ribosomal protein L3.e
YGL030w	*RPL32*	ribosomal protein L30.e
YDL075w	*RPL43A*	ribosomal protein L31.e
YLR406c	*RPL34B*	ribosomal protein L31.e.c12
YBL092w		ribosomal protein L32.e
YIL052c		ribosomal protein L34.e
YOR234c	*RPL37B*	ribosomal protein L35a.e.c15
YPL143w	*RPL37A*	ribosomal protein L35a.e.c16
YNL162w	*RPL41A*	ribosomal protein L36a.e
YHR141c	*RPL41B*	ribosomal protein L36a.e
YLR185w	*RPL35A*	ribosomal protein L37.e
YPR043w		ribosomal protein L37a.e
YJL189w	*RPL46*	ribosomal protein L39.e
YDR012w	*RPL2B*	ribosomal protein L4.e.B
YPL131w	*RPL1*	ribosomal protein L5.e
YNL002c	*RLP7*	ribosomal protein L7.e
YGL076c	*RPL6A*	ribosomal protein L7e.A
YHL033c	*RPL4A*	ribosomal protein L7a.e.A
YLL045c	*RPL4B*	ribosomal protein L7a.e.B
YIL018w	*RPL5A*	ribosomal protein L8.e
YFR031c-a	*RPL5B*	ribosomal protein L8.e
YGL147c	*RPL9A*	ribosomal protein L9.e
YNL067w	*RPL9B*	ribosomal protein L9.e.c14
YGR076c	*MRPL25*	ribosomal protein m (YML25), mitochondrial
YJR094w-a		ribosomal protein of the large subunit
YOR293w		ribosomal protein S10.e
YDR025w	*RPS18A*	ribosomal protein S11.e
YBR048w	*RPS18B*	ribosomal protein S11.e.B
YPR166c	*MRP2*	ribosomal protein S14
YJL191w	*CRY2*	ribosomal protein S14.e.B
YJL190c	*RPS24A*	ribosomal protein S15a.e.c10
YLR367w	*RPS24B*	ribosomal protein S15a.e.c12
YPL013c		ribosomal protein S16, mitochondrial
YMR143w	*RPS16A*	ribosomal protein S16.e
YDL083c	*RPS16B*	ribosomal protein S16.e
YML024w	*RP51A*	ribosomal protein S17.e.A
YDR447c	*RP51B*	ribosomal protein S17.e.B
YOL120c	*RP28A*	ribosomal protein S18.e
YML026c	*RPS18EB*	ribosomal protein S18.e.c13
YDR450w	*RPS18EA*	ribosomal protein S18.e.c4
YOL121c	*RP55A*	ribosomal protein S19.e
YNL302c	*RP55B*	ribosomal protein S19.e
YKR057w	*RPS25A*	ribosomal protein S21.e
YJL136c	*RPS25B*	ribosomal protein S21.e
YGR118w	*RPS28A*	ribosomal protein S23.e
YPR132w	*RPS28B*	ribosomal protein S23.e
YER074w	*RP50A*	ribosomal protein S24.e
YIL069c	*RP50B*	ribosomal protein S24.e
YLR333c	*RPS31B*	ribosomal protein S25.e.c12
YGR027c	*RPS31A*	ribosomal protein S25.e.c7
YER131w	*RPS26B*	ribosomal protein S26e-c5
YKL156w	*RPS27A*	ribosomal protein S27.e
YHR021c	*RPS27B*	ribosomal protein S27.e
YDR064w	*YS15*	ribosomal protein S27A
YLR264w	*RPS33B*	ribosomal protein S28.e.c12
YOR167c	*RPS33A*	ribosomal protein S28.e.c15
YLR388w	*YS29A*	ribosomal protein S29.e.A
YDL061c	*YS29B*	ribosomal protein S29.e.B
YNL178w	*RPS3*	ribosomal protein S3.e

YLR441c	*RP10A*	ribosomal protein S3a.e
YML063w	*RP10B*	ribosomal protein S3a.e
YJR145c	*RPS7B*	ribosomal protein S4.e.c10
YHR203c	*RPS7A*	ribosomal protein S4.e.c8
YBR251w	*MRPS5*	ribosomal protein S5, mitochondrial
YJR123w	*RPS5*	ribosomal protein S5.e
YBR181c	*RPS101*	ribosomal protein S6.e
YPL090c	*RPS10B*	ribosomal protein S6.e
YBL072c	*RPS8A*	ribosomal protein S8.e
YER102w	*RPS8B*	ribosomal protein S8.e
YBR146w	*MRPS9*	ribosomal protein S9, mitochondrial
YPL081w	*RPS13B*	ribosomal protein S9.e.A
YBR189w	*SUP46*	ribosomal protein S9.e.B
YGR085c	*RPL16B*	ribosomal protein YL16.B
YKR006c	*MRPL13*	ribosomal protein YmL13, mitochondrial
YNL005c	*MRPL2*	ribosomal protein YmL2, mitochondrial
YKR085c	*MRPL20*	ribosomal protein YmL20, mitochondrial
YBR282w	*MRPL27*	ribosomal protein YmL27, mitochondrial
YDR462w		ribosomal protein YmL28, mitochondrial
YNL252c		ribosomal protein YmL30, mitochondrial
YKL138c	*MRPL31*	ribosomal protein YmL31, mitochondrial
YCR003w	*MRPL32*	ribosomal protein YmL32, mitochondrial
YBR122c	*MRPL36*	ribosomal protein YmL36, mitochondrial
YBR268w	*MRPL37*	ribosomal protein YmL37, mitochondrial
YPL173w		ribosomal protein YmL40, mitochondrial
YMR225c	*MRPL44*	ribosomal protein YmR44, mitochondrial
YPL249c-a	*RPL39B*	ribosomal protein, cytoplasmic
YML025c		ribosomal protein, mitochondrial
YDR347w	*MRP1*	ribosomal protein, mitochondrial
YGR084c	*MRP13*	ribosomal protein, mitochondrial
YKL003c	*MRP17*	ribosomal protein, mitochondrial
YDR405w	*MRP20*	ribosomal protein, mitochondrial
YHL004w	*MRP4*	ribosomal protein, mitochondrial
YKL167c	*MRP49*	ribosomal protein, mitochondrial
YKL142w	*MRP8*	ribosomal protein, mitochondrial
YLR312w-a	*MRPL15*	ribosomal protein, mitochondrial
YBL038w	*MRPL16*	ribosomal protein, mitochondrial
YMR193w	*MRPL24*	ribosomal protein, mitochondrial
YMR024w	*MRPL3*	ribosomal protein, mitochondrial
YMR286w	*MRPL33*	ribosomal protein, mitochondrial
YLR439w	*MRPL4*	ribosomal protein, mitochondrial
YHR147c	*MRPL6*	ribosomal protein, mitochondrial
YGR220c	*MRPL9*	ribosomal protein, mitochondrial
YDR337w	*MRPS28*	ribosomal protein, mitochondrial
YNL137c	*NAM9*	ribosomal protein, mitochondrial
YOR158w	*PET123*	ribosomal protein, mitochondrial
YCR046c	*PETCR46*	ribosomal protein, mitochondrial
YFR049w	*YMR31*	ribosomal protein, mitochondrial
YJR113c		similarity to mitochondrial ribosomal protein S7
YHR148w		similarity to ribosomal protein
YDR116c		similarity to ribosomal protein L1
YOR150w		similarity to ribosomal protein L13
YNL284c		similarity to ribosomal protein L15
YEL050c		similarity to ribosomal protein L2
YLR009w		similarity to ribosomal protein L24.e.B
YDR115w		similarity to ribosomal protein L34
YDR237w		similarity to ribosomal protein L5
YNL081c		similarity to ribosomal protein S13
YDR041w		similarity to ribosomal proteins
YKL009w		similarity to Rpl10p and S.solfataricus ribosomal protein L10
YDL208w	*NHP2*	strong similarity to high-mobility group (HMG) family
YEL026w		strong similarity to high-mobility group-like protein Nhp2p
YOR182c	*RPS30B*	strong similarity to human ubiquitin-like protein/ribosomal protein S30
YNR037c		strong similarity to *Mycoplasma* ribosomal protein S19
YNL185c		strong similarity to ribosomal protein L11
YER056c-a		strong similarity to ribosomal protein L34.e
YMR230w		strong similarity to ribosomal protein S10
YNR036c		strong similarity to ribosomal protein S12
YNL096c		strong similarity to ribosomal protein S7
YLR167w	*UBI3*	ubiquitin/ribosomal protein S27a
YMR188c		weak similarity to 30S ribosomal protein S17
YDL202w		weak similarity to ribosomal protein

translation (initiation, elongation and termination)

YDL081c	*RPLA1*	acidic ribosomal protein a1
YMR309c	*NIP1*	associated with 40s ribosomal subunit
YGL094c	*PAN2*	component of Pab1p-stimulated poly(A) ribonuclease
YDR172w	*SUP35*	eukaryotic peptide chain release factor GTP-binding subunit
YNL007c	*SIS1*	heat-shock protein
YBL091c	*MAP2*	methionine aminopeptidase, isoform 2
YMR023c	*MSS1*	mitochondrial GTPase involved in expression of *COX1*
YGL143c	*MRF1*	mitochondrial peptide chain release factor
YGL049c	*TIF4632*	mRNA cap-binding protein (eIF4F), 130K subunit
YGR162w	*TIF4631*	mRNA cap-binding protein (eIF4F), 150K subunit
YOR276w	*CAF20*	mRNA CAP-binding protein (eIF4F), 20K subunit
YER165w	*PAB1*	mRNA polyadenylate-binding protein
YKL173w	*GIN10*	similarity to elongation factor 2 eFT1
YHR015w		similarity to PES4 PAB-like protein
YGR201c		similarity to Tef3p and Tef4p
YPL226w		similarity to translation elongation factor eEF3
YLR289w	*GUF1*	strong similarity to *E. coli* elongation factor-type GTP-binding protein lepa
YDL084w		strong similarity to nuclear RNA helicase (DEAD family)
YDR021w		strong similarity to translation initiation factor eIF4A
YKL010c	*SOS1*	suppressor of *sis1*
YLR005w	*SSL1*	TFIIH subunit (transcription initiation factor), factor B
YKR084c	*HBS1*	translation elongation factor eEF1 α subunit homologue
YPR080w	*TEF1*	translation elongation factor eEF1 α-A subunit, cytosolic
YBR118w	*TEF2*	translation elongation factor eEF1 α-A subunit, cytosolic
YPL048w	*CAM1*	translation elongation factor eEF1, α subunit
YKL081w	*TEF4*	translation elongation factor eEF1, γ subunit
YAL003w	*EFB1*	translation elongation factor eEF1β
YOR133w	*EFT1*	translation elongation factor eEF2
YDR385w	*EFT2*	translation elongation factor eEF2
YLR249w	*YEF3*	translation elongation factor eEF3
YNL014w		translation elongation factor eEF3 homologue
YNL163c		translation elongation factor eEF4
YLR069c	*MEF1*	translation elongation factor G, mitochondrial
YOR187w	*TUF1*	translation elongation factor TU, mitochondrial
YJL102w	*MEF2*	translation elongation factor, mitochondrial
YOL023w	*IFM1*	translation initiation factor 2, mitochondrial
YNL244c	*SUI1*	translation initiation factor 3 (eIF3)
YKR059w	*TIF1*	translation initiation factor 4A
YMR260c	*TIF11*	translation initiation factor eIF1A
YPL237w	*SUI3*	translation initiation factor eIF2 β subunit
YJR007w	*SUI2*	translation initiation factor eIF2, α subunit
YER025w	*GCD11*	translation initiation factor eIF2, γ subunit
YKR026c	*GCN3*	translation initiation factor eIF2B, 34K, α subunit
YLR291c	*GCD7*	translation initiation factor eIF2B, 43K subunit
YGR083c	*GCD2*	translation initiation factor eIF2B, 71K (δ) subunit
YDR211w	*GCD6*	translation initiation factor eIF2B, ε 81K subunit
YOR260w	*GCD1*	translation initiation factor eIF2B, γ subunit
YOR361c	*PRT1*	translation initiation factor eIF3 subunit
YMR146c	*TIF34*	translation initiation factor eIF3, P39 subunit
YNL062c	*GCD10*	translation initiation factor eIF3, RNA-binding subunit
YJL138c	*TIF2*	translation initiation factor eIF4A
YPR163c	*TIF3*	translation initiation factor eIF4B
YOL139c	*CDC33*	translation initiation factor eIF4E
YPR041w	*TIF5*	translation initiation factor eIF5
YEL034w	*HYP2*	translation initiation factor eIF5A.1
YJR047c	*ANB1*	translation initiation factor eIF5A.2
YGL169w	*SUA5*	translation initiation protein
YBR143c	*SUP45*	translational release factor

translational control

YMR282c	*AEP2*	2'-O-ribosyl phosphate transferase
YBR120c	*CBP6*	apo-cytochrome b pre-mRNA processing protein
YPL029w	*SUV3*	ATP-dependent RNA helicase, mitochondrial
YOR017w	*PET127*	component of mitochondrial translation system
YMR028w	*TAP42*	component of the Tor signalling pathway
YDR197w	*CBS2*	cytochrome b translational activator protein
YNL216w	*RAP1*	DNA-binding protein with repressor and activator activity
YER054c	*GIP2*	Glc7p-interacting protein
YMR080c	*NAM7*	nonsense-mediated mRNA decay protein
YGR072w	*UPF3*	nonsense-mediated mRNA decay protein
YHR077c	*NMD2*	nonsense-mediated mRNA decay protein 2
YMR064w	*AEP1*	nuclear control of ATPase messenger RNA expression protein
YPL179w	*PPQ1*	phosphoprotein phosphatase
YFR009w	*GCN20*	positive effector of Gcn2p
YLR203c	*MSS51*	possibly involved in translational activation of COX1 and COB mRNA
YBL058w	*SHP1*	potential regulatory subunit for Glc7p
YNL139c	*RLR1*	regulatory protein
YLR067c	*PET309*	required for stability and translation of COX1 mRNA
YMR257c	*PET111*	required for translation of COX2 mRNA
YER133w	*GLC7*	ser/thr phosphoprotein phosphatase 1, catalytic subunit
YOR178c	*GAC1*	ser/thr phosphoprotein phosphatase 1, regulatory subunit
YDR283c	*GCN2*	ser/thr protein kinase
YGR222w	*PET54*	splicing protein and translational activator, mitochondrial
YBR024w	*SCO2*	strong similarity to Sco1p
YGL195w	*GCN1*	translational activator
YDL069c	*CBS1*	translational activator of cob mRNA
YER153c	*PET122*	translational activator of cytochrome c oxidase subunit III
YNR045w	*PET494*	translational activator, mitochondrial
YJL125c	*GCD14*	translational repressor of GCN4

tRNA synthetases

YOR335c	*ALA1*	alanyl-tRNA synthetase, cytosolic
YHR091c	*MSR1*	arginyl-tRNA synthetase, mitochondrial
YCR024c		asn-tRNA synthetase, mitochondrial
YHR019c	*DED81*	asparaginyl-tRNA-synthetase
YPL104w	*MSD1*	aspartate-tRNA ligase, mitochondrial
YLL018c	*DPS1*	aspartyl-tRNA synthetase, cytosolic
YOR168w	*GLN4*	glutaminyl-tRNA synthetase
YOL033w	*MSE1*	glutamyl-tRNA synthetase, mitochondrial
YBR121c	*GRS1*	glycine-tRNA ligase
YPR033c	*HTS1*	histidine-tRNA ligase, mitochondrial
YPL040c	*ISM1*	isoleucine-tRNA ligase, mitochondrial
YBL076c	*ILS1*	isoleucyl-tRNA synthetase
YPL160w	*CDC60*	leucine-tRNA ligase, cytosolic
YLR382c	*NAM2*	leucine-tRNA ligase, mitochondrial
YDR037w	*KRS1*	lysyl-tRNA synthetase, cytosolic
YNL073w	*MSK1*	lysyl-tRNA synthetase, mitochondrial
YGR264c	*MES1*	methionyl-tRNA synthetase
YGR171c	*MSM1*	methionyl-tRNA synthetase
YPR047w	*MSF1*	phenylalanine-tRNA ligase α subunit, mitochondrial
YFL022c	*FRS2*	phenylalanine-tRNA ligase β subunit, cytosolic
YLR060w	*FRS1*	phenylalanyl-tRNA synthetase, α subunit, cytosolic
YHR011w		seryl-tRNA synthetase
YDR023w	*SES1*	seryl-tRNA synthetase, cytosolic
YNL247w		similarity to cysteinyl-tRNA synthetases
YER087w		similarity to *E. coli* prolyl-tRNA synthetase
YDR341c		strong similarity to arginine-tRNA ligase and mitochondrial arginyl-tRNA synthetase
YGL245w		strong similarity to glutamine-tRNA ligase
YPR081c		strong similarity to glycyl-tRNA synthetases
YHR020w		strong similarity to human glutamyl-prolyl-tRNA synthetase
YKL194c	*MST1*	threonine-tRNA ligase, mitochondrial
YIL078w	*THS1*	threonyl tRNA synthetase, cytosolic
YOL097c	*WRS1*	tryptophan-tRNA ligase
YDR268w	*MSW1*	tryptophanyl-tRNA synthetase, mitochondrial
YPL097w	*MSY1*	tyrosyl-tRNA synthetase
YGR185c	*TYS1*	tyrosyl-tRNA synthetase
YGR094w	*VAS1*	valyl-tRNA synthetase

other protein-synthesis activities

YDL229w	*SSB1*	heat-shock protein of HSP70 family
YNL209w	*SSB2*	heat-shock protein of HSP70 family, cytosolic
YGR147c	*NAT2*	N-acetyltransferase for N-terminal methionine
YJR066w	*TOR1*	phosphatidylinositol 3-kinase
YDL044c	*MTF2*	protein involved in mRNA splicing and protein synthesis, mitochondrial
YDL040c	*NAT1*	protein N-acetyltransferase subunit
YJL023c	*PET130*	protein synthesis protein, mitochondrial
YMR005w	*MPT1*	required for protein synthesis
YBL080c	*PET112*	required to maintain RHO^+ mitochondrial DNA
YHR189w		similarity to peptidyl-tRNA hydrolases
YLL039c	*UBI4*	ubiquitin

Protein destination

protein folding and stabilization

YEL030w		heat-shock protein of HSP70 family
YKL073w	*LHS1*	chaperone of the ER lumen
YIL142w	*CCT2*	chaperonin of the TCP1 ring complex, cytosolic
YJL014w	*CCT3*	chaperonin of the TCP1 ring complex, cytosolic
YDR212w	*CCT1*	component of chaperonin-containing T-complex
YDL143w	*CCT4*	component of chaperonin-containing T-complex
YJL111w	*CCT7*	component of chaperonin-containing T-complex
YJL008c	*CCT8*	component of chaperonin-containing T-complex
YDR188w	*CCT6*	component of chaperonin-containing T-complex (ζ subunit)
YDR155c	*CPH1*	cyclophilin (peptidylprolyl isomerase)
YML078w	*CPR3*	cyclophilin (peptidylprolyl isomerase), mitochondrial
YOR288c	*MPD1*	disulphide isomerase related protein
YER048c	*CAJ1*	dnaJ homologue
YDR519w	*FKB2*	FK506/rapamycin-binding protein of the ER
YFL016c	*MDJ1*	heat-shock protein - chaperone
YOR232w	*MGE1*	heat-shock protein - chaperone
YLR259c	*HSP60*	heat-shock protein - chaperone, mitochondrial
YJR045c	*SSC1*	heat-shock protein 70-related protein, mitochondrial
YAL005c	*SSA1*	heat-shock protein of HSP70 family
YLL024c	*SSA2*	heat-shock protein of HSP70 family, cytosolic
YDR171w	*HSP42*	heat-shock protein with similarity to Hsp26p
YLR216c	*CPR6*	member of the cyclophilin family
YJR032w	*CPR7*	member of the cyclophilin family
YJL034w	*KAR2*	nuclear fusion protein

YHR057c	*CYP2*	peptidyl-prolyl *cis-trans* isomerase
YCR069w	*SCC3*	peptidyl-prolyl *cis-trans* isomerase
YDR304c	*CYP5*	peptidyl-prolyl *cis-trans* isomerase D (cyclophilin D) of the ER
YNL135c	*FPR1*	peptidyl-prolyl *cis-trans* isomerase, FK506-binding protein
YML074c	*NPI46*	proline *cis-trans* isomerase
YDR518w	*EUG1*	protein disulphide isomerase
YCL043c	*PDI1*	protein disulphide-isomerase
YJR097w		similarity to Caj1p
YBR044c		similarity to chaperonin HSP60 proteins
YOL088c		similarity to disulphide isomerases and ER60 proteases
YNL077w		similarity to dnaJ protein homologue YDJ1
YMR161w	*HLJ1*	similarity to dnaJ proteins
YIR004w		similarity to DNAJ-like proteins
YLR090w	*XDJ1*	similarity to *E. coli* dnaJ
YJL073w		similarity to heat-shock proteins
YNR028w		similarity to peptidylprolyl isomerase Scc3p
YIL005w		similarity to protein disulphide isomerases
YLR369w		strong similarity to heat-shock protein 70-related proteins
YLR449w		strong similarity to peptidylprolyl isomerase FRP3
YJR064w	*CCT5*	T-complex protein 1, ε subunit
YFR041c		weak similarity to dnaJ-like heat-shock proteins
YNL227c		weak similarity to dnaJ-like proteins

protein targeting, sorting and translocation

YKL073w	*LHS1*	chaperone of the ER lumen
YGL206c	*CHC1*	clathrin heavy chain
YOL062c	*APM4*	clathrin-associate protein YAP54
YPL259c	*APM1*	clathrin-associated protein
YBR288c	*APM3*	clathrin-associated protein complex, medium subunit
YDL212w	*SHR3*	endoplasmic reticulum membrane protein
YBL040c	*ERD2*	ER lumen protein-retaining receptor
YER019c-a	*SEB2*	ER protein-translocation complex subunit
YLR378c	*SEC61*	ER protein-translocation complex subunit
YPL094c	*SEC62*	ER protein-translocation complex subunit
YOR254c	*SEC63*	ER protein-translocation complex subunit
YBR171w	*SEC66*	ER protein-translocation complex subunit
YLR292c	*SEC72*	ER protein-translocation complex subunit
YDR086c	*SSS1*	ER protein-translocation complex subunit
YOR089c	*VPS21*	GTP-binding protein
YKR014c	*YPT52*	GTP-binding protein of the RAB family
YNL093w	*YPT53*	GTP-binding protein of the RAB family (RAS superfamily)
YOR270c	*VPH1*	H^+-ATPase V0 domain 95K subunit, vacuolar
YAL005c	*SSA1*	heat-shock protein of HSP70 family
YCR075c	*ERS1*	intracellular protein transport protein
YGR028w	*MSP1*	intra-mitochondrial sorting protein
YOR069w	*VPS5*	involved in Golgi retention and vacuolar sorting
YLR347c	*KAP95*	karyopherin-β
YOL122c	*SMF1*	manganese transporter
YKR001c	*VPS1*	member of the dynamin family of GTPases
YOR020c	*HSP10*	mitochondrial chaperonin
YNR017w	*MAS6*	mitochondrial inner membrane import translocase subunit
YJL143w	*TIM17*	mitochondrial inner membrane import translocase subunit
YIL022w	*TIM44*	mitochondrial inner membrane import translocase subunit
YMR035w	*IMP2*	mitochondrial inner membrane protease subunit
YNL131w	*TOM22*	mitochondrial outer membrane import receptor subunit
YGR082w	*TOM20*	mitochondrial outer membrane import receptor subunit, 20K
YMR060c	*TOM37*	mitochondrial outer membrane import receptor subunit, 37K
YMR203w	*TOM40*	mitochondrial outer membrane import receptor subunit, 40K
YOR045w	*TOM6*	mitochondrial outer membrane import receptor subunit, 6K
YNL070w	*TOM7*	mitochondrial outer membrane import receptor subunit, 7K
YNL121c	*TOM70*	mitochondrial outer membrane specialized import receptor
YMR150c	*IMP1*	mitochondrial protease
YJL034w	*KAR2*	nuclear fusion protein
YMR091c	*NPL6*	nuclear protein localization factor
YBR170c	*NPL4*	nuclear protein localization factor and ER translocation component
YDR432w	*NPL3*	nucleolar protein
YLR191w	*PAS20*	peroxisomal protein involved in protein import
YPR047w	*MSF1*	phenylalanine-tRNA ligase α subunit, mitochondrial
YNL106c	*PIE3*	phosphatidylinositol phosphate phosphatase
YBL069w	*AST1*	PMA1 protein targeting protein
YLR168c	*(MSF1)*	probably involved in intramitochondrial protein sorting
YDR244w	*PAS10*	putative peroxisomal targeting signal receptor
YCL001w	*RER1*	required for correct localization of Sec12p
YDR414c	*ERD1*	required for retention of luminal ER proteins
YDR005c	*MAF1*	required for sorting of Mod5p
YMR004w	*MVP1*	required for vacuolar protein sorting
YDR498c	*SEC20*	secretory pathway protein
YBR162w-a	*YSY6*	secretory pathway protein
YBR097w	*VPS15*	ser/thr protein kinase
YPL243w	*SRP68*	signal recognition particle protein
YPL210c	*SRP72*	signal recognition particle protein
YDR292c	*SRP101*	signal recognition particle receptor, α subunit
YML105c	*SEC65*	signal recognition particle subunit
YDL092w	*SRP14*	signal recognition particle subunit
YKL122c	*SRP21*	signal recognition particle subunit
YPR088c	*SRP54*	signal recognition particle subunit
YOR285w		similarity to *D. melanogaster* heat-shock protein 67B2
YOR286w		similarity to *D. melanogaster* heat-shock protein 67B2
YHR110w		similarity to human gp25L2 protein
YAR002c-a		similarity to mammalian gp25L protein
YKL154w		similarity to mouse signal recognition particle receptor β subunit
YKL196c		similarity to Sec22p
YMR018w		similarity to tetratricopeptide-repeat protein PAS10
YMR214w	*SCJ1*	similarity to to *E. coli* dnaJ
YNL304w		similarity to Ypt1p and other GTP-binding proteins
YER101c	*AST2*	strong similarity to Ast1p
YOR016c		strong similarity to FUN54 protein, similarity to hamster COP-coated vesicle membrane protein
YGL002w		strong similarity to human gp25L2 protein
YER087c-a	*SEB1*	strong similarity to mammalian Sec61β subunit
YCR099c		strong similarity to Pep1p
YCR100c		strong similarity to Pep1p
YCR101c		strong similarity to Pep1p
YIL173w		strong similarity to Pep1p
YJL222w		strong similarity to Pep1p
YBR283c		strong similarity to Sec61p
YOR327c	*SNC2*	strong similarity to synaptobrevin
YOR329c	*SCD5*	suppressor of clathrin deficiency
YHR050w	*SMF2*	suppressor of mitochondrial matrix mutant
YOR036w	*PEP12*	syntaxin (T-SNARE)
YLR148w	*PEP3*	vacuolar membrane protein
YBL017c	*PEP1*	vacuolar protein sorting/targeting protein
YJL053w	*PEP8*	vacuolar protein sorting/targeting protein
YOR132w	*VPS17*	vacuolar protein sorting-associated protein
YNR006w	*VPS27*	vacuolar protein sorting-associated protein
YGL095c	*VPS45*	vacuolar protein sorting-associated protein
YJL154c	*VPS35*	vacuolar protein-sorting protein
YDR323c	*PEP7*	vacuolar segregation protein
YPL045w	*VPS16*	vacuolar sorting protein
YDR495c	*VPS3*	vacuolar sorting protein
YLR396c	*VPS33*	vacuolar sorting protein
YAL002w	*VPS8*	vacuolar sorting protein
YML097c	*VPS9*	vacuolar sorting protein

protein modification (glycosylation, acylation, myristylation, palmitylation, farnesylation and processing)

YJR131w	*MNS1*	α-1,2-mannosidase
YDR483w	*KRE2*	α-1,2-mannosyltransferase
YER001w	*MNN1*	α-1,3-mannosyltransferase
YGL038c	*OCH1*	α-1,6-mannosyltransferase
YBR164c	*ARL1*	ADP-ribosylation factor
YKL157w	*APE2*	aminopeptidase yscII
YJR062c	*NTA1*	amino-terminal amidase
YGL017w	*ATE1*	arginyl tRNA transferase
YPL154c	*PEP4*	aspartyl protease
YLR120c	*YAP3*	aspergillopepsin
YDL141w	*BPL1*	biotin holocarboxylase synthetase
YBR110w	*ALG1*	β-mannosyltransferase
YGL203c	*KEX1*	carboxypeptidase (YSC-α)
YPL227c	*ALG5*	dolichol-P-glucose synthetase
YJR143c	*PMT4*	dolichyl-phosphate-mannose-protein O-mannosyl transferase
YDL212w	*SHR3*	endoplasmic reticulum membrane protein
YNL238w	*KEX2*	endoproteinase of late Golgi compartment
YDR331w	*GPI8*	essential for GPI-anchor attachment
YDR410c	*STE14*	farnesyl cysteine carboxyl-methyltransferase
YOR370c	*MSI4*	geranylgeranyltransferase regulatory subunit
YGL155w	*CDC43*	geranylgeranyltransferase type I β subunit
YPR176c	*BET2*	geranylgeranyltransferase type II β subunit
YJL031c	*BET4*	geranylgeranyltransferase, α subunit
YOR002w	*ALG6*	glucosyltransferase
YOR067c	*ALG8*	glucosyltransferase
YEL042w	*GDA1*	guanosine diphosphatase
YGR252w	*GCN5*	histone acetyltransferase
YPL001w	*HAT1*	histone acetyltransferase subunit
YBR034c	*HMT1*	hnRNP methyltransferase
YAL039c	*CYC3*	holocytochrome c synthase (cytochrome c heme lyase)
YKL087c	*CYT2*	holocytochrome c1 synthase
YGL065c	*ALG2*	mannosyltransferase
YDL095w	*PMT1*	mannosyltransferase
YAL023c	*PMT2*	mannosyltransferase
YOR321w	*PMT3*	mannosyltransferase
YBL082c	*RHK1*	mannosyltransferase
YLR244c	*MAP1*	methionine aminopeptidase, isoform 1
YBL091c	*MAP2*	methionine aminopeptidase, isoform 2
YMR035w	*IMP2*	mitochondrial inner membrane protease subunit
YKL134c	*(MIP1)*	mitochondrial intermediate peptidase
YLR163c	*MAS1*	mitochondrial processing peptidase
YHR024c	*MAS2*	mitochondrial processing peptidase, catalytic 53K (α) subunit
YMR150c	*IMP1*	mitochondrial protease
YPR051w	*MAK3*	N-acetyltransferase
YGR147c	*NAT2*	N-acetyltransferase for N-terminal methionine
YBR247c	*ENP1*	N-glycosylation protein
YLR195c	*NMT1*	N-myristoyltransferase
YJL002c	*OST1*	oligosaccharyltransferase α subunit
YEL002c	*WBP1*	oligosaccharyltransferase β subunit
YMR149w	*SWP1*	oligosaccharyltransferase δ subunit
YOR103c	*OST2*	oligosaccharyltransferase ε subunit
YOR085w	*OST3*	oligosaccharyltransferase γ subunit
YDL232w	*OST4*	oligosaccharyltransferase subunit
YGL022w	*STT3*	oligosaccharyltransferase subunit
YFL045c	*SEC53*	phosphomannomutase
YPR122w	*AXL1*	protease
YEL060c	*PRB1*	protease B, vacuolar
YLR389c	*STE23*	protease involved in a-factor processing
YCL043c	*PDI1*	protein disulphide-isomerase
YKL019w	*RAM2*	protein farnesyltransferase, α subunit
YDL090c	*RAM1*	protein farnesyltransferase, β subunit
YHR013c	*ARD1*	protein N-acetyltransferase subunit
YDL040c	*NAT1*	protein N-acetyltransferase subunit
YIL085c	*KTR7*	putative α-1,2-mannosyltransferase
YGL194c	*RTL1*	putative deacetylase
YNL029c	*KTR5*	putative mannosyltransferase
YGR199w	*PMT6*	putative mannosyltransferase
YKL193c	*SDS22*	regulatory subunit for the mitotic function of type I protein phosphatase
YNL048w	*ALG11*	required for asparagine-linked glycosylation
YLR088w	*GAA1*	required for attachment of GPI anchor onto proteins
YPL050c	*MNN9*	required for complex N-glycosylation
YJR010c-a	*SPC1*	signal peptidase 10.8K subunit
YML055w	*SPC2*	signal peptidase 18K subunit
YIR022w	*SEC11*	signal sequence processing protein
YDL042c	*SIR2*	silencing regulatory protein
YER005w		similarity to Gda1p
YMR223w		similarity to human putative ubiquitin carboxyl-terminal hydrolase
YDR098c		similarity to *Legionella* glutaredoxin-like protein
YER174c		similarity to *Legionella* glutaredoxin-like protein
YGL257c		similarity to Mnn1p
YIL014w		similarity to Mnn1p
YPR131c		similarity to N-acetyltransferases
YDL093w	*PMT5*	similarity to O-mannosyltransferases Pmt1p-Pmt4p
YDR307w		similarity to Pmt1p
YDR245w	*MNN10*	similarity to *S. pombe* galactosyltransferase
YLR066w		similarity to signal peptidase
YNR059w		similarity to to α-1,3-mannosyltransferase
YPL003w		similarity to ubiquitin-activating enzymes
YPR180w		similarity to ubiquitin-activating enzymes
YGL087c		similarity to ubiquitin-protein ligase
YIR039c		similarity to Yap3p
YBR205w	*KTR3*	strong similarity to α-1,2-mannosyltransferase
YBR199w	*KTR4*	strong similarity to α-1,2-mannosyltransferase
YPL053c	*KTR6*	strong similarity to α-1,2-mannosyltransferase Kre2p
YPL051w		strong similarity to ADP-ribosylation factors
YLR121c		strong similarity to aspartylproteases
YOR099w	*KTR1*	strong similarity to mannosyltransferases
YCR083w		strong similarity to thioredoxin
YOR339c		strong similarity to ubiquitin conjugating enzymes
YPR066w		strong similarity to ubiquitin-activating enzymes
YML111w		strong similarity to ubiquitination protein Bul1p
YGL226c-a	*OST5*	subunit of N-oligosaccharyltransferase, ζ subunit
YEL056w	*HAT2*	subunit of the major yeast histone acetyltransferase
YJR075w	*HOC1*	suppressor of *pkc1*
YLR043c	*TRX1*	thioredoxin I
YGR209c	*TRX2*	thioredoxin II
YOR219c	*STE13*	type IV dipeptidyl aminopeptidase
YMR275c	*BUL1*	ubiquitination pathway protein
YDR054c	*CDC34*	ubiquitin-conjugating enzyme
YGR133w	*PAS2*	ubiquitin-conjugating enzyme
YMR022w	*QRI8*	ubiquitin-conjugating enzyme
YGL058w	*RAD6*	ubiquitin-conjugating enzyme
YDR177w	*UBC1*	ubiquitin-conjugating enzyme
YLR306w	*UBC12*	ubiquitin-conjugating enzyme
YDR092w	*UBC13*	ubiquitin-conjugating enzyme
YDR059c	*UBC5*	ubiquitin-conjugating enzyme
YER100w	*UBC6*	ubiquitin-conjugating enzyme
YEL012w	*UBC8*	ubiquitin-conjugating enzyme
YDL064w	*UBC9*	ubiquitin-conjugating enzyme
YDR139c		ubiquitin-like protein
YER125w	*RSP5*	ubiquitin-protein ligase
YKL210w	*UBA1*	ubiquitin-protein ligase
YJR099w	*YUH1*	ubiquitin-specific protease
YBR243c	*ALG7*	UDP-N-acetylglucosamine-1-phosphate transferase
YKL035w	*UGP1*	UTP-glucose-1-phosphate uridylyltransferase
YDR495c	*VPS3*	vacuolar sorting protein

YJR117w	*STE24*	zinc metallo-protease

assembly of protein complexes

YGR214w	*NAB1A*	40S ribosomal protein p40 homologue A
YLR048w	*NAB1B*	40S ribosomal protein p40 homologue B
YPL195w	*YKS4*	α- or γ- adaptin, large subunit of the clathrin-associated protein(AP) complex
YBL037w	*APL3*	α-adaptin, large subunit of the clathrin-associated protein(AP) complex
YOR094w	*ARF3*	ADP-ribosylation factor 3
YBL022c	*PIM1*	ATP-dependent protease, mitochondrial
YKL135c	*APL2*	β-adaptin
YGR261c	*YKS5*	β-adaptin, large subunit of the clathrin-associated protein(AP) complex
YBR195c	*MSI1*	chromatin assembly complex, subunit p50
YPR018w	*RLF2*	chromatin assembly complex, subunit p90
YLR170c	*APS1*	clathrin-associated protein (AP) complex, small subunit AP19
YJR058c	*APS2*	clathrin-associated protein 17, small subunit
YJR005w	*YAP80*	clathrin-associated protein complex, β subunit
YJL024c	*APS3*	clathrin-associated protein(AP) complex, small subunit
YER058w	*PET117*	cytochrome c oxidase assembly factor
YPL132w	*COX11*	cytochrome c oxidase assembly protein
YML129c	*COX14*	cytochrome c oxidase assembly protein
YDR079w	*PET100*	cytochrome c oxidase assembly protein
YLR038c	*COX12*	cytochrome c oxidase subunit VIB
YMR256c	*COX7*	cytochrome c oxidase subunit VII
YDL067c	*COX9*	cytochrome c oxidase subunit VIIA
YER141w	*COX15*	cytochrome oxidase assembly factor
YER154w	*OXA1*	cytochrome oxidase biogenesis protein
YBL007c	*SLA1*	cytoskeleton assembly control protein
YNL243w	*SLA2*	cytoskeleton assembly control protein
YLR393w	*ATP10*	F1F0 ATPase complex assembly protein
YNL315c	*ATP11*	F1F0-ATPase complex assembly protein
YJL180c	*ATP12*	F1F0-ATPase complex assembly protein
YPL172c	*COX10*	farnesyl transferase
YPL231w	*FAS2*	fatty-acyl-CoA synthase, α subunit
YPR029c	*APL4*	γ-adaptin, large subunit of the clathrin-associated protein(AP) complex
YDL192w	*ARF1*	GTP-binding protein of the ARF family
YDL137w	*ARF2*	GTP-binding protein of the ARF family
YKL119c	*VPH2*	H^+-ATPase assembly protein, vacuolar
YPL234c	*TFP3*	H^+-ATPase V0 domain 17K subunit, vacuolar
YLR447c	*VMA6*	H^+-ATPase V0 domain 36K subunit, vacuolar
YOR270c	*VPH1*	H^+-ATPase V0 domain 95K subunit, vacuolar
YOR332w	*VMA4*	H^+-ATPase V1 domain 27K subunit, vacuolar
YKL080w	*VMA5*	H^+-ATPase V1 domain 42K subunit, vacuolar
YLL009c	*COX17*	interacts genetically with *SCO1* and *SCO2* in cytochrome oxidase assembly
YJR034w	*PET191*	involved in assembly of cytochrome oxidase
YBR037c	*SCO1*	involved in stabilization of Cox1p and Cox2p
YDR375c	*BCS1*	mitochondrial protein of the CDC48/PAS1/SEC18 (AAA) family of ATPases
YPL085w	*SEC16*	multidomain vesicle coat protein
YKR048c	*NAP1*	nucleosome assembly protein I
YML102w	*CAC2*	p60 subunit of the chromatin assembly factor-I (CAF-I)
YKL197c	*PAS1*	peroxisomal assembly protein
YDR265w	*PAS4*	peroxisomal assembly protein
YNL329c	*PAS8*	peroxisomal assembly protein
YBR237w	*PRP5*	pre-mRNA processing RNA-helicase
YJL203w	*PRP21*	pre-mRNA splicing factor
YER017c	*AFG3*	protease of the *SEC18*/*CDC48*/*PAS1* family of ATPases (AAA)
YPR024w	*YME1*	protease of the *SEC18*/*CDC48*/*PAS1* family of ATPases (AAA)
YMR089c	*YTA12*	protease of the *SEC18*/*CDC48*/*PAS1* family of ATPases (AAA)
YPL215w	*CBP3*	required for assembly of cytochrome bc1 complex
YBR185c	*MBA1*	respiratory chain assembly protein
YLL008w	*DRS1*	RNA helicase of the DEAD box family
YBR227c		similarity to *E. coli* ATP-binding protein clpX
YGR241c		similarity to rat clathrin assembly protein
YHR161c		similarity to rat clathrin assembly protein AP180
YGR074w	*SMD1*	snRNA-associated protein
YBR024w	*SCO2*	strong similarity to Sco1p
YLR327c		strong similarity to STF2p
YOR134w	*BAG7*	structural homologue of Sac7p
YDR389w	*SAC7*	suppressor of actin mutation
YOR106w	*VAM3*	syntaxin related protein
YDR460w	*TFB3*	TFIIH subunit (transcription/repair factor)
YPR178w	*PRP4*	U4/U6 snRNP 52K protein
YGL119w	*ABC1*	ubiquinol cytochrome c reductase complex assembly protein
YPR191w	*QCR2*	ubiquinol-cytochrome c reductase 40K subunit II
YGR174c	*CBP4*	ubiquinol-cytochrome c reductase assembly factor
YLL039c	*UBI4*	ubiquitin
YGR133w	*PAS2*	ubiquitin-conjugating enzyme
YGR105w	*VMA21*	vacuolar ATPase assembly integral membrane protein
YHR060w	*VMA22*	vacuolar ATPase assembly protein
YDR495c	*VPS3*	vacuolar sorting protein

proteolysis

YPL149w	*APG5*	involved in autophagy and nutrient starvation
YOL013c	*HRD1*	involved in degradation of Hmg2p
YLR207w	*HRD3*	involved in degradation of Hmg2p
YMR174c	*PAI3*	protease A (ysca) inhibitor IA3
YNL015w	*PBI2*	proteinase B inhibitor 2
YCL052c	*PBN1*	required for Prb1p expression
YHR132c		similarity to carboxypeptidase
YNR069c		similarity to central part of Bul1p
YBR227c		similarity to *E. coli* ATP-binding protein clpX
YNL186w		similarity to human putative ubiquitin carboxy-terminal hydrolase
YIL137c		similarity to *M. musculus* aminopeptidase
YKR038c		similarity to Qri7p
YER047c		similarity to regulatory subunit Yta6p of 26S proteasome
YCR045c		similarity to serin proteases
YER078c		similarity to X-Pro aminopeptidase II
YBR139w		strong similarity to carboxypeptidase
YBL067c	*UBP13*	ubiquitin carboxyl-terminal hydrolase
YOL111c		weak similarity to human ubiquitin-like protein GDX

cytoplasmic degradation

YFR052w	*NIN1*	26S proteasome regulatory subunit
YIL075c	*SEN3*	26S proteasome regulatory subunit
YOR259c	*CRL13*	26S proteasome subunit
YPR103w	*PRE2*	26S proteasome subunit
YFR050c	*PRE4*	26S proteasome subunit
YOR157c	*PUP1*	26S proteasome subunit
YGR253c	*PUP2*	26S proteasome subunit
YER094c	*PUP3*	26S proteasome subunit
YGL048c	*SUG1*	26S proteasome subunit
YHR200w	*SUN1*	26S proteasome subunit
YER021w	*SUN2*	26S proteasome subunit
YOR117w	*YTA1*	26S proteasome subunit
YDR394w	*YTA2*	26S proteasome subunit
YKL145w	*YTA3*	26S proteasome subunit
YGR270w	*YTA7*	26S proteasome subunit
YOR362c	*PRE10*	26S proteasome subunit C1
YER012w	*PRE1*	26S proteasome subunit C11
YGR135w	*PRE9*	26S proteasome subunit Y13
YML092c	*PRE8*	26S proteasome subunit Y7
YGL011c	*SCL1*	26S proteasome subunit YC7α/Y8
YMR314w	*PRE5*	26S proteasome subunit, α-type
YNL239w	*LAP3*	aminopeptidase of cysteine protease family
YJR062c	*NTA1*	amino-terminal amidase
YGL017w	*ATE1*	arginyl tRNA transferase
YDL132w	*CDC53*	controls G1/S transition
YLR452c	*SST2*	involved in desensitization to α-factor pheromone
YKL213c	*DOA1*	involved in ubiquitin-dependent proteolysis
YDL126c	*CDC48*	microsomal protein of CDC48/PAS1/SEC18 family of ATPases
YJL001w	*PRE3*	multicatalytic endopeptidase complex subunit
YOL038w	*PRE6*	multicatalytic endopeptidase complex subunit
YBL041w	*PRE7*	multicatalytic endopeptidase complex subunit
YDL020c	*SON1*	nuclear protein
YBR165w	*UBS1*	positive regulator of Cdc34p
YDL007w	*YTA5*	probable component of 26S proteasome complex
YIL148w	*UBI1*	ribosomal protein
YKR094c	*UBI2*	ribosomal protein
YMR223w		similarity to human putative ubiquitin carboxyl-terminal hydrolase
YDR390c	*UBA2*	similarity to Uba1p
YPL003w		similarity to ubiquitin-activating enzymes
YPR180w		similarity to ubiquitin-activating enzymes
YGL087c		similarity to ubiquitin-protein ligase
YPL074w	*YTA6*	similarity to Vps4p and Yta4p
YOR261c		strong similarity to human 26S proteasome regulatory subunit, p40
YOR339c		strong similarity to ubiquitin conjugating enzymes
YPR066w		strong similarity to ubiquitin-activating enzymes
YML111w		strong similarity to ubiquitination protein Bul1p
YHR027c	*HRD2*	subunit of 26S proteasome
YKL022c	*CDC16*	subunit of anaphase-promoting complex (cyclosome)
YHR166c	*CDC23*	subunit of anaphase-promoting complex (cyclosome)
YBL084c	*CDC27*	subunit of anaphase-promoting complex (cyclosome)
YLL039c	*UBI4*	ubiquitin
YIL156w	*UBP7*	ubiquitin carboxy terminal hydrolase
YER098w	*UBP9*	ubiquitin carboxyl-terminal hydrolase
YKR098c	*UBP11*	ubiquitin C-terminal hydrolase
YJL197w	*UBP12*	ubiquitin C-terminal hydrolase
YGR048w	*UFD1*	ubiquitin fusion degradation protein
YDL190c	*UFD2*	ubiquitin fusion degradation protein
YBR058c	*UBP14*	ubiquitin specific protease
YLR167w	*UBI3*	ubiquitin/ribosomal protein S27a
YMR275c	*BUL1*	ubiquitination pathway protein
YDR054c	*CDC34*	ubiquitin-conjugating enzyme
YMR022w	*QRI8*	ubiquitin-conjugating enzyme
YDR177w	*UBC1*	ubiquitin-conjugating enzyme
YLR306w	*UBC12*	ubiquitin-conjugating enzyme
YDR092w	*UBC13*	ubiquitin-conjugating enzyme
YBR082c	*UBC4*	ubiquitin-conjugating enzyme
YDR059c	*UBC5*	ubiquitin-conjugating enzyme
YER100w	*UBC6*	ubiquitin-conjugating enzyme
YDL064w	*UBC9*	ubiquitin-conjugating enzyme
YDR139c		ubiquitin-like protein
YER125w	*RSP5*	ubiquitin-protein ligase
YKL210w	*UBA1*	ubiquitin-protein ligase
YGR184c	*UBR1*	ubiquitin-protein ligase
YDR069c	*DOA4*	ubiquitin-specific protease
YDL122w	*UBP1*	ubiquitin-specific protease
YER144c	*UBP5*	ubiquitin-specific protease
YJR099w	*YUH1*	ubiquitin-specific protease
YOR124c	*UBP2*	ubiquitin-specific proteinase
YER151c	*UBP3*	ubiquitin-specific proteinase

lysosomal and vacuolar degradation

YBR286w	*APE3*	aminopeptidase Y, vacuolar
YKL103c	*LAP4*	aminopeptidase yscl, vacuolar
YPL154c	*PEP4*	aspartyl protease
YMR297w	*PRC1*	carboxypeptidase y, serine-type protease
YHR028c	*DAP2*	dipeptidyl aminopeptidase B
YJL172w	*CPS1*	Gly-X carboxypeptidase YSCS
YEL060c	*PRB1*	protease B, vacuolar
YOR003w	*YSP3*	subtilisin-like protease III

other subcellular degradation

YDR144c	*MKC7*	aspartyl protease of the periplasmic space
YBL022c	*PIM1*	ATP-dependent protease, mitochondrial
YIL015w	*BAR1*	barrierpepsin
YBR201w	*DER1*	involved in degradation proteins in the ER
YER017c	*AFG3*	protease of the *SEC18*/*CDC48*/*PAS1* family of ATPases (AAA)
YPR024w	*YME1*	protease of the *SEC18*/*CDC48*/*PAS1* family of ATPases (AAA)
YMR089c	*YTA12*	protease of the *SEC18*/*CDC48*/*PAS1* family of ATPases (AAA)
YCL057w	*PRD1*	proteinase yscD
YHR113w		similarity to vacuolar aminopeptidase Lap4p/Ape1p
YFR006w		similarity to X-Pro dipeptidases

other protein-destination activities

YDR258c	*HSP78*	heat-shock protein of clpb family of ATP-dependent proteases, mitochondrial
YDL104c	*QRI7*	similarity to *H. influenzae* sialoglycoprotease (gcp)
YDR415c		strong similarity to bacterial leucyl aminopeptidase

Transport facilitation

ion channels

YGR217w	*CCH1*	calcium channel protein
YLL052c		member of mip family transmembrane channels
YJL093c	*TOK1*	outward-rectifier potassium channel
YPR192w		similarity to plasma membrane and water channel proteins
YLL053c		similarity to water channel proteins
YIL114c	*POR2*	voltage-dependent anion channel (YVDAC2)
YJR040w	*GEF1*	voltage-gated chloride channel protein

ion transporters

YNR055c	*HOL1*	member of major facilitator superfamily multidrug-resistance protein subfamily 1
YIL048w		similarity to amino-phospholipids-ATPase Drs2p

metal ion transporters (Cu, Fe, etc.)

YOR316c	*COT1*	cobalt accumulation protein
YPR124w	*CTR1*	copper transport protein
YHR175w	*CTR2*	copper transport protein
YLR411w	*CTR3*	copper transport protein
YGL255w	*ZRT1*	high-affinity zinc transport protein
YLR130c	*ZRT2*	low-affininty zinc transporter
YMR319c	*FET4*	low-affinity Fe(II) iron transport protein
YOL122c	*SMF1*	manganese transporter
YBR290w	*BSD2*	metal homeostasis protein
YDR270w	*CCC2*	probable copper-transporting ATPase
YBR295w	*PCA1*	P-type Cu^{2+}-transporting ATPase
YMR243c	*ZRC1*	zinc- and cadmium resistance protein

other cation transporters (Na, K, Ca, NH_4, etc.)

YGR121c	*MEP1*	ammonia permease of high capacity and moderate affinity
YDL128w	*VCX1*	Ca^{2+}-transport (H^+/Ca^{2+} exchange) protein, vacuolar
YGL006w	*PMC1*	Ca^{2+}-transporting P-type ATPase
YGL167c	*PMR1*	Ca^{2+}-transporting P-type ATPase
YCR024c-a	*PMP1*	H^+-ATPase subunit, plasma membrane
YEL017c-a	*PMP2*	H^+-ATPase subunit, plasma membrane
YMR054w	*STV1*	H^+-ATPase V0 domain 102K subunit, vacuolar
YHR039c-a	*VMA10*	H^+-ATPase V0 domain 13K subunit, vacuolar

YEL027w	*CUP5*	H^+-ATPase V0 domain 17K subunit, vacuolar
YPL234c	*TFP3*	H^+-ATPase V0 domain 17K subunit, vacuolar
YLR447c	*VMA6*	H^+-ATPase V0 domain 36K subunit, vacuolar
YOR270c	*VPH1*	H^+-ATPase V0 domain 95K subunit, vacuolar
YGR020c	*VMA7*	H^+-ATPase V1 domain 14K subunit, vacuolar
YOR332w	*VMA4*	H^+-ATPase V1 domain 27K subunit, vacuolar
YEL051w	*VMA8*	H^+-ATPase V1 domain 32K subunit, vacuolar
YKL080w	*VMA5*	H^+-ATPase V1 domain 42K subunit, vacuolar
YPR036w	*VMA13*	H^+-ATPase V1 domain 54K subunit, vacuolar
YBR127c	*VMA2*	H^+-ATPase V1 domain 60K subunit, vacuolar
YDL185w	*TFP1*	H^+-ATPase V1 domain 69K subunit, vacuolar
YGL008c	*PMA1*	H^+-transporting P-type ATPase
YPL036w	*PMA2*	H^+-transporting P-type ATPase 2
YNL142w	*MEP2*	high-affinity low-capacity ammonia permease
YJL129c	*TRK1*	high-affinity potassium transport protein
YAL026c	*DRS2*	membrane-spanning P-type amino-phospholipids-ATPase
YKR050w	*TRK2*	moderate-affinity potassium transport protein
YHR026w	*PPA1*	proteolipid protein of proton-transporting ATPase
YEL031w	*SPF1*	P-type ATPase
YLR138w	*NHA1*	putative Na^+/H^+ antiporter
YBR235w		similarity to bumetanide-sensitive Na^+-K^+-Cl^- cotransport protein
YJL094c		similarity to *E. hirae* Na^+/H^+-antiporter NapA
YDR456w		similarity to Na^+/H^+ antiporters
YPR138c	*MEP3*	strong similarity to ammonium transport proteins
YHL016c	*DUR3*	urea transport protein

anion transporters (Cl, SO_4 ,PO_4, etc)

YLR348c		dicarboxylate carrier protein
YML123c	*PHO84*	high-affinity inorganic phosphate/H^+ symporter
YBR294w	*SUL1*	high-affinity sulphate transport protein
YJL117w	*PHO86*	inorganic phospate transporter
YCR037c	*PHO87*	member of the phosphate permease family
YJR077c	*MIR1*	phosphate transport protein, mitochondrial (MCF)
YBR235w		similarity to bumetanide-sensitive Na^+-K^+-Cl^- cotransport protein
YJL094c		similarity to *E. hirae* Na^+/H^+-antiporter NapA
YNR013c		similarity to membrane protein Pho87p and hypothetical protein YJL198w
YCR098c	*GIT1*	similarity to phosphate transporter proteins
YPR003c		similarity to sulphate transporter proteins
YER053c		strong similarity to mitochondrial phosphate carrier protein
YJL198w		strong similarity to Pho87p
YBR296c		strong similarity to phosphate-repressible phosphate permease
YLR092w	*SEL2*	strong similarity to Sul1p

sugar and carbohydrate transporters

YPR021c		similarity to human citrate transporter protein
YKL217w	*JEN1*	carboxylic-acid transporter protein
YBR291c	*CTP1*	citrate transport protein, mitochondrial (MCF)
YLR348c		dicarboxylate carrier protein
YLR081w	*GAL2*	galactose (and glucose) permease
YGR289c	*AGT1*	general α-glucoside permease
YLL043w	*FPS1*	glycerol channel protein
YNL318c	*HXT14*	hexose transport protein
YJL214w	*HXT8*	hexose transport protein
YJL219w	*HXT9*	hexose transport protein
YFL011w	*HXT10*	hexose transporter
YDL194w	*SNF3*	high-affinity glucose transporter
YEL069c	*HXT13*	high-affinity hexose transporter
YMR011w	*HXT2*	high-affinity hexose transporter
YDR343c	*HXT6*	high-affinity hexose transporter
YDR342c	*HXT7*	high-affinity hexose transporter
YML123c	*PHO84*	high-affinity inorganic phosphate/H^+ symporter
YOL156w	*HXT11*	low-affinity glucose transporter
YHR094c	*HXT1*	low-affinity hexose transporter
YDR345c	*HXT3*	low-affinity hexose transporter
YDR497c	*ITR1*	major myo-inositol permease
YBR298c	*MAL31*	maltose permease
YDL198c	*YHM1*	member of the mitochondrial carrier family (MCF)
YHR092c	*HXT4*	moderate- to low-affinity glucose transporter
YOL103w	*ITR2*	myo-inositol transport protein
YBR241c		similarity to glucose transport proteins
YGL104c		similarity to glucose transport proteins
YDR387c		similarity to Itr1p and Itr2p
YFR045w		similarity to mitochondrial citrate transport proteins
YCR098c	*GIT1*	similarity to phosphate transporter proteins
YDL199c		similarity to sugar transporter proteins
YFL040w		similarity to yeast glucose transport proteins
YHR096c	*HXT5*	strong similarity to hexose transporters
YDL245c	*HXT15*	strong similarity to Hxt17p and Haxt7p
YJR160c		strong similarity to Mal3Tp
YDR536w	*STL1*	strong similarity to members of the sugar permease family
YOR271c		strong similarity to *Rattus* tricarboxylate carrier
YDL247w		strong similarity to sugar transport proteins
YIL171w		strong similarity to sugar transport proteins
YIL170w	*HXT12*	strong similarity to sugar transport proteins
YJR158w	*HXT16*	strong similarity to sugar transport proteins
YNR072w	*HXT17*	sugar transport protein
YDL138w	*RGT2*	suppressor of *snf3* mutant

amino-acid transporters

YBR068c	*BAP2*	amino-acid permease
YEL063c	*CAN1*	amino-acid permease
YBR069c	*VAP1*	amino-acid permease
YCL025c	*AGP1*	asparagine and glutamine permease
YGL077c	*HNM1*	choline permease
YPL265w	*DIP5*	dicarboxylic amino-acid permease
YDL210w	*UGA4*	GABA-specific high-affinity permease
YKR039w	*GAP1*	general amino-acid permease
YDR508c	*GNP1*	high-affinity glutamine permease
YGR055w	*MUP1*	high-affinity methionine permease
YNL270c	*ALP1*	high-affinity permease for basic amino acids
YOL020w	*SCM2*	high-affinity tryptophan transport protein
YGR191w	*HIP1*	histidine permease
YNL268w	*LYP1*	lysine-specific high-affinity permease
YOR130c	*ARG11*	member of the mitochondrial carrier family (MCF)
YHL036w	*MUP3*	methionine permease
YOR348c	*PUT4*	proline and γ-aminobutyrate permease
YKL174c		similarity to choline transport protein Ctr1p
YNR056c		similarity to choline transport protein Ctr1p
YFL055w		similarity to Gap1p and other amino-acid permeases
YDR160w		similarity to lysine transport protein LYP1
YLL061w		strong similarity to amino-acid transport protein Gap1p
YDR046c	*(PAP1)*	strong similarity to amino-acid transport proteins
YBR132c		strong similarity to amino-acid permeases
YPL274w		strong similarity to amino-acid transport proteins

lipid transporters

YBR041w	*FAT1*	fatty-acid transporter
YKL188c	*PAT1*	long-chain-fatty-acid transporter
YPL147w	*PXA1*	long-chain-fatty-acid transporter
YKL174c		similarity to choline transport protein Ctr1p
YNR056c		similarity to choline transport protein Ctr1p
YHR123w	*EPT1*	sn-1,2-diacylglycerol ethanolamine- and cholinephosphotransferase

purine and pyrimidine transporters

YMR056c	*AAC1*	ADP/ATP carrier protein (MCF)
YBL030c	*AAC2*	ADP/ATP carrier protein (MCF)
YBR085w	*AAC3*	ADP/ATP carrier protein (MCF)
YBR192w	*RIM2*	mitochondrial carrier protein (MCF)
YER056c	*FCY2*	purine-cytosine permease
YER060w	*FCY21*	purine-cytosine permease
YOR222w		similarity to ADP/ATP carrier proteins
YPL134c		similarity to ADP/ATP carrier proteins
YPR011c		similarity to ADP/ATP carrier proteins and Graves disease carrier protein
YOR071c		similarity to allantoin or uracil transport proteins
YOR192c		similarity to allantoin or uracil transport proteins
YGR096w		similarity to bovine Graves disease carrier protein
YHR002w		similarity to bovine mitochondrial carrier protein/Grave's disease carrier protein
YGL186c		similarity to hypothetical protein YER060w and weak similarity to FCY2 protein
YBL042c		strong similarity to allantoin and uracil transport proteins
YER060w-a	*FCY22*	strong similarity to Fcy2p
YBR021w	*FUR4*	uracil permease

allantoin and allantoate transporters

YJR152w	*DAL5*	allantoate permease
YIR028w	*DAL4*	allantoin permease
YIL166c		similarity to allantoate permease Dal5p
YCR028c	*FEN2*	similarity to allantoate permease transporter
YGR260w		similarity to allantoate transport protein
YLR004c		similarity to allantoate transport protein
YOR071c		similarity to allantoin or uracil transport proteins
YOR192c		similarity to allantoin or uracil transport proteins
YLR237w		similarity to allantoin transport protein
YLL055w		similarity to Dal5p
YBL042c		strong similarity to allantoin and uracil transport proteins
YAL067c	*SEO1*	suppressor of sulphoxyde ethionine resistance

transport ATPases

YGL006w	*PMC1*	Ca^{2+}-transporting P-type ATPase
YGL167c	*PMR1*	Ca^{2+}-transporting P-type ATPase
YKL016c	*ATP7*	F1F0-ATPase complex, F0 D subunit
YBL099w	*ATP1*	F1F0-ATPase complex, F1 α subunit
YJR121w	*ATP2*	F1F0-ATPase complex, F1 β subunit
YDL004w	*ATP16*	F1F0-ATPase complex, F1 δ subunit
YPL078c	*ATP4*	F1F0-ATPase complex, F1 δ subunit
YPL271w	*ATP15*	F1F0-ATPase complex, F1 ε subunit
YBR039w	*ATP3*	F1F0-ATPase complex, F1 γ subunit
YDR298c	*ATP5*	F1F0-ATPase complex, OSCP subunit
YLR295c	*ATP14*	F1F0-ATPase complex, subunit h
YCR024c-a	*PMP1*	H^+-ATPase subunit, plasma membrane
YEL017c-a	*PMP2*	H^+-ATPase subunit, plasma membrane
YMR054w	*STV1*	H^+-ATPase V0 domain 102K subunit, vacuolar
YHR039c-a	*VMA10*	H^+-ATPase V0 domain 13K subunit, vacuolar
YEL027w	*CUP5*	H^+-ATPase V0 domain 17K subunit, vacuolar
YPL234c	*TFP3*	H^+-ATPase V0 domain 17K subunit, vacuolar
YLR447c	*VMA6*	H^+-ATPase V0 domain 36K subunit, vacuolar
YOR270c	*VPH1*	H^+-ATPase V0 domain 95K subunit, vacuolar
YGR020c	*VMA7*	H^+-ATPase V1 domain 14K subunit, vacuolar
YOR332w	*VMA4*	H^+-ATPase V1 domain 27K subunit, vacuolar
YEL051w	*VMA8*	H^+-ATPase V1 domain 32K subunit, vacuolar
YKL080w	*VMA5*	H^+-ATPase V1 domain 42K subunit, vacuolar
YPR036w	*VMA13*	H^+-ATPase V1 domain 54K subunit, vacuolar
YBR127c	*VMA2*	H^+-ATPase V1 domain 60K subunit, vacuolar
YDL185w	*TFP1*	H^+-ATPase V1 domain 69K subunit, vacuolar
YGL008c	*PMA1*	H^+-transporting P-type ATPase
YPL036w	*PMA2*	H^+-transporting P-type ATPase 2
YAL026c	*DRS2*	membrane-spanning P-type amino-phospholipids-ATPase
YDR270w	*CCC2*	probable copper-transporting ATPase
YHR026w	*PPA1*	proteolipid protein of proton-transporting ATPase
YEL031w	*SPF1*	P-type ATPase
YDR040c	*ENA1*	P-type ATPase involved in Na^+ and Li^+ efflux
YDR039c	*ENA2*	P-type ATPase involved in Na^+ efflux
YDR038c	*ENA5*	P-type ATPase involved in Na^+ efflux
YBR295w	*PCA1*	P-type Cu^{2+}-transporting ATPase
YIL048w		similarity to amino-phospholipids-ATPase Drs2p
YER166w		similarity to ATPase *P. falciparum* ATPase 2
YMR162c		similarity to ATPases
YDR093w		similarity to *P. falciparum* ATPase 2

ABC transporters

YDR406w	*PDR15*	ATP-binding cassette protein family member
YLR188w	*MDL1*	ATP-binding cassette transporter family member
YPL270w	*MDL2*	ATP-binding cassette transporter family member
YKL209c	*STE6*	ATP-binding cassette transporter protein
YGR281w	*YOR1*	ATP-binding cassette transporter protein required for oligomycin resistance
YMR301c	*ATM1*	ATP-binding cassette transporter protein, mitochondrial
YCR011c	*ADP1*	ATP-dependent permease
YDR135c	*YCF1*	glutathione S-conjugate transporter, vacuolar
YKL188c	*PAT1*	long-chain-fatty-acid transporter
YPL147w	*PXA1*	long-chain-fatty-acid transporter
YDR011w	*SNQ2*	multidrug resistance protein
YPL058c	*PDR12*	multidrug resistance transporter
YOR153w	*PDR5*	pleiotropic drug resistance protein
YOL075c		similarity to *A. gambiae* ATP-binding-cassette protein
YER036c		similarity to members of the ABC transporter family
YLL015w		similarity to metal resistance proteins
YHL035c		similarity to multidrug resistance proteins
YKR103w		similarity to multidrug resistance proteins
YKR104w		similarity to multidrug resistance proteins
YLL048c		similarity to rat organic anion transporter
YIL013c	*PDR11*	similarity to Snq2p and other ATP-dependent permeases
YOR328w	*PDR10*	strong similarity to ABC transporter proteins
YOR011w		strong similarity to ATP-dependent permeases

YDR091c		strong similarity to human RNase L inhibitor and to *M. jannaschii* ABC transporter
YNR070w		strong similarity to Snq2p

drug transporters

YML116w	*ATR1*	aminotriazole and 4-nitroquinoline resistance protein
YGR281w	*YOR1*	ATP-binding cassette transporter protein required for oligomycin resistance
YGR197c	*SNG1*	involved in nitroguanidine resistance
YEL065w		probably multidrug resistance protein
YMR123w	*PKR1*	resistance against *Pichia farinosa* killer toxin (SMK toxin) when expressed by a multicopy plasmid
YMR088c		similarity multidrug resistance proteins
YDR119w		similarity to *B. subtilis* tetracyclin resistance
YBR293w		similarity to multidrug resistance proteins
YGR138c		similarity to multidrug resistance proteins
YHR048w		similarity to multidrug resistance proteins
YPR156c		similarity to multidrug resistance proteins
YNL065w		similarity to resistance proteins
YIL013c	*PDR11*	similarity to Snq2p and other ATP-dependent permeases
YMR279c		strong similarity to aminotriazole resistance protein
YGR224w		strong similarity to drug resistance protein SGE1
YBR052c		strong similarity to *S. pombe* brefeldin A resistance protein obr1
YAL067c	*SEO1*	suppressor of sulphoxyde ethionine resistance

other transport-facilitators

YFL028c	*CAF16*	ATP-binding cassette transporter protein family member
YER019c-a	*SEB2*	ER protein-translocation complex subunit
YLR378c	*SEC61*	ER protein-translocation complex subunit
YPL094c	*SEC62*	ER protein-translocation complex subunit
YOR254c	*SEC63*	ER protein-translocation complex subunit
YBR171w	*SEC66*	ER protein-translocation complex subunit
YLR292c	*SEC72*	ER protein-translocation complex subunit
YDR086c	*SSS1*	ER protein-translocation complex subunit
YIL134w	*FLX1*	FAD carrier protein, mitochondrial (MCF)
YCR023c		member of major facilitator superfamily multidrug-resistance protein family 2
YKL039w	PTM1	member of the major facilitator superfamily
YPR058w	*YMC1*	mitochondrial carrier protein (MCF)
YNR017w	*MAS6*	mitochondrial inner membrane import translocase subunit
YJL143w	*TIM17*	mitochondrial inner membrane import translocase subunit
YIL022w	*TIM44*	mitochondrial inner membrane import translocase subunit
YNL131w	*TOM22*	mitochondrial outer membrane import receptor subunit
YGR082w	*TOM20*	mitochondrial outer membrane import receptor subunit, 20K
YMR060c	*TOM37*	mitochondrial outer membrane import receptor subunit, 37K
YMR203w	*TOM40*	mitochondrial outer membrane import receptor subunit, 40K
YOR045w	*TOM6*	mitochondrial outer membrane import receptor subunit, 6K
YNL070w	*TOM7*	mitochondrial outer membrane import receptor subunit, 7K
YNL055c	*POR1*	mitochondrial outer membrane porin
YNL121c	*TOM70*	mitochondrial outer membrane specialized import receptor
YAR035w	*YAT1*	outer carnitine acetyltransferase, mitochondrial
YKR093w	*PTR2*	peptide transporter
YIL120w		similarity to antibiotic resistance proteins
YIL121w		similarity to antibiotic resistance proteins
YCL038c		similarity to bacterial membrane transporter
YBR043c		similarity to benomyl/methotrexate resistance protein
YHL040c		similarity to *C. carbonum* toxin pump
YHL047c		similarity to *C. carbonum* toxin pump
YER024w		similarity to carnitine O-acetyltransferase Yat1p
YOR291w		similarity to cation translocating ATPases
YFL054c		similarity to channel proteins
YBR180w		similarity to drug resistance proteins
YIL006w		similarity to Flx1p
YOR306c		similarity to human X-linked PEST-containing transporter
YIL056w		similarity to hypothetical protein YER064c
YNL003c	*PET8*	similarity to mitochondrial rat tricarboxylate transport protein
YOL119c		similarity to monocarboxylate transporter proteins
YGR065c		similarity to *P. putida* phthalate transporter
YOL163w		similarity to *P. putida* phthalate transporter
YEL006w		similarity to peroxisomal membrane and mitochondrial carrier proteins
YJR124c		similarity to *Staphylococcus* multidrug resistance protein
YNL125c		similarity to YKL221w and human X-linked PEST-containing transporter
YBR008c		strong similarity to benomyl/methotrexate resistance protein
YCL069w		strong similarity to drug resistance protein SGE1
YER087c-a	*SEB1*	strong similarity to mammalian Sec61β subunit
YKR067w		strong similarity to Sct1p
YBL011w	*SCT1*	suppresses a choline-transport mutant
YBL089w		weak similarity to *A. thaliana* aminoacid permease AAP3
YIL088c		weak similarity to *A. thaliana* aminoacid permease AAP4
YJR001w		weak similarity to *A. thaliana* aminoacid permease AAP4
YER119c		weak similarity to *E. herbicola* tyrosine permease
YKL221w		weak similarity to human X-linked PEST-containing transporter
YEL064c		weak similarity to members of the major facilitator superfamily
YJR106w		weak similarity to Na^+/H^+ antiporter
YDL206w		weak similarity to transporter proteins
YKL146w		weak similarity to transporter proteins

Intracellular transport

nuclear transport

YIL075c	*SEN3*	26S proteasome regulatory subunit
YOR048c	*RAT1*	5'-3' exoribonuclease
YGL097w	*SRM1*	GDP/GTP exchange factor for Gsp1p/Gsp2p
YOR185c	*GSP2*	GTP-binding protein
YLR293c	*GSP1*	GTP-binding protein of the RAS superfamily
YAL005c	*SSA1*	heat-shock protein of HSP70 family
YOR160w	*MTR10*	involved in mRNA transport
YJL050w	*MTR4*	involved in nucleocytoplasmic transport of mRNA
YBR017c	*KAP104*	karyopherin
YNL189w	*SRP1*	karyopherin-α or importin
YLR347c	*KAP95*	karyopherin-β
YKL186c	*MTR2*	mRNA transport protein
YPL124w	*NIP29*	nuclear import protein
YPL174c	*NIP80*	nuclear import protein
YMR129w	*POM152*	nuclear pore membrane glycoprotein
YFR002w	*NIC96*	nuclear pore protein
YJL041w	*NSP1*	nuclear pore protein
YOR098c	*NUP1*	nuclear pore protein
YKL068w	*NUP100*	nuclear pore protein
YMR047c	*NUP116*	nuclear pore protein
YKL057c	*NUP120*	nuclear pore protein
YKR082w	*NUP133*	nuclear pore protein
YGL092w	*NUP145*	nuclear pore protein
YER105c	*NUP157*	nuclear pore protein
YIL115c	*NUP159*	nuclear pore protein
YBL079w	*NUP170*	nuclear pore protein
YML103c	*NUP188*	nuclear pore protein
YLR335w	*NUP2*	nuclear pore protein
YDR192c	*NUP42*	nuclear pore protein
YGL172w	*NUP49*	nuclear pore protein
YGR119c	*NUP57*	nuclear pore protein
YJL061w	*NUP82*	nuclear pore protein
YDL116w	*NUP84*	nuclear pore protein
YJR042w	*NUP85*	nuclear pore protein
YGL100w	*SEH1*	nuclear pore protein
YBR170c	*NPL4*	nuclear protein localization factor and ER translocation component
YER009w	*NTF2*	nuclear transport factor
YDR432w	*NPL3*	nucleolar protein
YDR002w	*YRB1*	ran-specific GTPase-activating protein
YER107c	*GLE2*	required for nuclear pore complex structure and function
YIR011c	*STS1*	required for transport of Rna15p from the cytoplasm to the nucleus
YDL207w	*GLE1*	RNA export mediator
YIL063c	*YRB2*	similarity to Yrp1p and Nup2p
YLR119w	*SRN2*	suppressor of *rna1-1* mutation
YFL049w		weak similarity to Npl6p

mitochondrial transport

YMR056c	*AAC1*	ADP/ATP carrier protein (MCF)
YBL030c	*AAC2*	ADP/ATP carrier protein (MCF)
YBR085w	*AAC3*	ADP/ATP carrier protein (MCF)
YMR301c	*ATM1*	ATP-binding cassette transporter protein, mitochondrial
YML042w	*CAT2*	carnitine O-acetyltransferase
YBR291c	*CTP1*	citrate transport protein, mitochondrial (MCF)
YOR316c	*COT1*	cobalt accumulation protein
YOR037w	*CYC2*	cytochrome c mitochondrial import factor
YLR348c		dicarboxylate carrier protein
YKL016c	*ATP7*	F1F0-ATPase complex, F0 D subunit
YBL099w	*ATP1*	F1F0-ATPase complex, F1 α subunit
YJR121w	*ATP2*	F1F0-ATPase complex, F1 β subunit
YDL004w	*ATP16*	F1F0-ATPase complex, F1 δ subunit
YPL078c	*ATP4*	F1F0-ATPase complex, F1 δ subunit
YPL271w	*ATP15*	F1F0-ATPase complex, F1 ε subunit
YBR039w	*ATP3*	F1F0-ATPase complex, F1 γ subunit
YDR298c	*ATP5*	F1F0-ATPase complex, OSCP subunit
YLR295c	*ATP14*	F1F0-ATPase complex, subunit h
YIL134w	*FLX1*	FAD carrier protein, mitochondrial (MCF)
YOR232w	*MGE1*	heat-shock protein - chaperone
YLL024c	*SSA2*	heat-shock protein of HSP70 family, cytosolic
YGR028w	*MSP1*	intra-mitochondrial sorting protein
YOR130c	*ARG11*	member of the mitochondrial carrier family (MCF)
YDL198c	*YHM1*	member of the mitochondrial carrier family (MCF)
YNL064c	*YDJ1*	mitochondrial and ER import protein
YBR192w	*RIM2*	mitochondrial carrier protein (MCF)
YPR058w	*YMC1*	mitochondrial carrier protein (MCF)
YBR104w	*YMC2*	mitochondrial carrier protein (MCF)
YML062c	*MFT1*	mitochondrial fusion target protein
YNR017w	*MAS6*	mitochondrial inner membrane import translocase subunit
YJL143w	*TIM17*	mitochondrial inner membrane import translocase subunit
YIL022w	*TIM44*	mitochondrial inner membrane import translocase subunit
YNL131w	*TOM22*	mitochondrial outer membrane import receptor subunit
YGR082w	*TOM20*	mitochondrial outer membrane import receptor subunit, 20K
YMR060c	*TOM37*	mitochondrial outer membrane import receptor subunit, 37K
YMR203w	*TOM40*	mitochondrial outer membrane import receptor subunit, 40K
YOR045w	*TOM6*	mitochondrial outer membrane import receptor subunit, 6K
YNL070w	*TOM7*	mitochondrial outer membrane import receptor subunit, 7K
YNL055c	*POR1*	mitochondrial outer membrane porin
YNL121c	*TOM70*	mitochondrial outer membrane specialized import receptor
YAR035w	*YAT1*	outer carnitine acetyltransferase, mitochondrial
YJR077c	*MIR1*	phosphate transport protein, mitochondrial (MCF)
YBR091c	*MRS5*	regulator of mitochondrial intron splicing
YPL134c		similarity to ADP/ATP carrier proteins
YER024w		similarity to carnitine O-acetyltransferase Yat1p
YPR021c		similarity to human citrate transporter protein
YER053c		strong similarity to mitochondrial phosphate carrier protein
YLR034c		strong similarity to SMF2 protein
YHR117w	*TOM71*	strong similarity to Tom70p/Mas70p
YHR050w	*SMF2*	suppressor of mitochondrial matrix mutant
YIL114c	*POR2*	voltage dependent anion channel (YVDAC2)

vesicular transport (Golgi network, etc.)

YBR164c	*ARL1*	ADP-ribosylation factor
YOR094w	*ARF3*	ADP-ribosylation factor 3
YGL167c	*PMR1*	Ca^{2+}-transporting P-type ATPase
YGR167w	*CLC1*	clathrin light chain
YOL062c	*APM4*	clathrin-associate protein YAP54
YPL259c	*APM1*	clathrin-associated protein
YBR288c	*APM3*	clathrin-associated protein complex, medium subunit
YPL010w	*RET3*	coatomer complex ζ subunit
YDL145c	*RET1*	coatomer complex α subunit of secretory pathway vesicles
YGL137w	*SEC27*	coatomer complex β' subunit (β'-cop) of secretory pathway vesicles
YDR238c	*SEC26*	coatomer complex β subunit of secretory pathway vesicles
YFR051c	*RET2*	coatomer complex δ subunit
YNL287w	*SEC21*	coatomer complex γ subunit (γ-COP) of secretory pathway vesicles
YDR170c	*SEC7*	component of non-clathrin vesicle coat
YPR181c	*SEC23*	component of the COPII coat of ER-Golgi vesicles
YDL195w	*SEC31*	component of the COPII coat of ER-Golgi vesicles
YGL200c	*EMP24*	component of the COPII-coated vesicles, 24K
YML012w	*ERV25*	component of the COPII-coated vesicles, 25K
YMR017w	*DBI9*	Dbf2p interacting protein
YKR054c	*DYN1*	dynein heavy chain, cytosolic
YLR083c	*EMP70*	endosomal protein
YBL040c	*ERD2*	ER lumen protein-retaining receptor
YJR031c	*GEA1*	GDP/GTP exchange factor for ARF
YEL022w	*GEA2*	GDP/GTP exchange factor for ARF
YNR026c	*SEC12*	GDP/GTP exchange factor for Sar1p
YPR017c	*DSS4*	GDP/GTP exchange factor for Sec4p
YDL192w	*ARF1*	GTP-binding protein of the ARF family
YDL137w	*ARF2*	GTP-binding protein of the ARF family
YPL218w	*SAR1*	GTP-binding protein of the ARF family
YFL038c	*YPT1*	GTP-binding protein of the RAB family
YER031c	*YPT31*	GTP-binding protein of the RAB family
YLR262c	*YPT6*	GTP-binding protein of the RAB family
YML001w	*YPT7*	GTP-binding protein of the RAB family
YFL005w	*SEC4*	GTP-binding protein of the RAS superfamily
YLL024c	*SSA2*	heat-shock protein of HSP70 family, cytosolic
YLR268w	*SEC22*	high copy suppressor of *ypt1* null mutation
YDR189w	*SLY1*	hydrophilic suppressor of *ypt1* and member of the Sec1p family
YDL058w	*USO1*	intracellular protein transport protein
YPL145c	*KES1*	involved in ergosterol biosynthesis
YKR068c	*BET3*	involved in targeting and fusion of ER to Golgi transport vesicles
YMR183c	*SSO2*	involved in vesicle transport from Golgi to plasma membrane
YNL064c	*YDJ1*	mitochondrial and ER import protein

YJL004c	*SYS1*	multicopy suppressor of *ypt6*
YPL085w	*SEC16*	multidomain vesicle coat protein
YOR326w	*MYO2*	myosin heavy chain
YFL025c	*BST1*	negative regulator of COPII vesicle formation
YMR079w	*SEC14*	phosphatidylinositol/phosphatidylcholine transfer protein
YCR067c	*SED4*	protein of the endoplasmic reticulum
YIL004c	*BET1*	protein transport protein
YLR208w	*SEC13*	protein transport protein
YNL272c	*SEC2*	protein transport protein
YGR009c	*SEC9*	protein transport protein
YGL145w	*TIP20*	required for ER to Golgi transport
YDR498c	*SEC20*	secretory pathway protein
YBR097w	*VPS15*	ser/thr protein kinase
YBR264c		similarity to GTP-binding proteins
YKL196c		similarity to Sec22p
YJL193w		similarity to Sly41p
YGL210w	*YPT32*	small GTP-binding protein
YKL006c-a	*SFT1*	SNARE-like protein
YPL051w		strong similarity to ADP-ribosylation factors
YDR107c		strong similarity to EMP70 protein
YAL030w	*SNC1*	strong similarity to synaptic vesicle-associated membrane protein
YOR327c	*SNC2*	strong similarity to synaptobrevin
YER039c		strong similarity to vanadate resistance protein Van2p
YLR026c	*SED5*	syntaxin (T-SNARE)
YOR075w	*UFE1*	syntaxin (T-SNARE) of the ER
YPL232w	*SSO1*	syntaxin-related protein
YBL050w	*SEC17*	transport vesicle fusion protein
YLR396c	*VPS33*	vacuolar sorting protein
YML097c	*VPS9*	vacuolar sorting protein
YGL225w	*GOG5*	vanadate-resistance protein
YGL233w	*SEC15*	vesicular traffic control protein
YBR080c	*SEC18*	vesicular-fusion protein, functional homologue of NSF
YLR093c		weak similarity to synaptobrevin
YLR078c	*BOS1*	weak similarity to synaptobrevin (V-SNARE)

peroxisomal transport

YML042w	*CAT2*	carnitine O-acetyltransferase
YKL188c	*PAT1*	long-chain-fatty-acid transporter
YPL147w	*PXA1*	long-chain-fatty-acid transporter
YER015w	*FAA2*	long-chain-fatty-acid-CoA ligase
YDR329c	*PAS3*	peroxisomal assembly protein
YJL210w	*PAS5*	peroxisomal assembly protein
YDR142c	*PAS7*	peroxisomal import protein
YLR191w	*PAS20*	peroxisomal protein involved in protein import
YDR244w	*PAS10*	putative peroxisomal targeting signal receptor

vacuolar transport

YDL128w	*VCX1*	Ca^{2+}-transport (H^+/Ca^{2+} exchange) protein, vacuolar
YGL006w	*PMC1*	Ca^{2+}-transporting P-type ATPase
YDR135c	*YCF1*	glutathione S-conjugate transporter, vacuolar
YOR089c	*VPS21*	GTP-binding protein
YER031c	*YPT31*	GTP-binding protein of the RAB family
YKR014c	*YPT52*	GTP-binding protein of the RAB family
YNL093w	*YPT53*	GTP-binding protein of the RAB family (RAS superfamily)
YMR054w	*STV1*	H^+-ATPase V0 domain 102K subunit, vacuolar
YHR039c-a	*VMA10*	H^+-ATPase V0 domain 13K subunit, vacuolar
YEL027w	*CUP5*	H^+-ATPase V0 domain 17K subunit, vacuolar
YPL234c	*TFP3*	H^+-ATPase V0 domain 17K subunit, vacuolar
YLR447c	*VMA6*	H^+-ATPase V0 domain 36K subunit, vacuolar
YOR270c	*VPH1*	H^+-ATPase V0 domain 95K subunit, vacuolar
YGR020c	*VMA7*	H^+-ATPase V1 domain 14K subunit, vacuolar
YOR332w	*VMA4*	H^+-ATPase V1 domain 27K subunit, vacuolar
YEL051w	*VMA8*	H^+-ATPase V1 domain 32K subunit, vacuolar
YKL080w	*VMA5*	H^+-ATPase V1 domain 42K subunit, vacuolar
YPR036w	*VMA13*	H^+-ATPase V1 domain 54K subunit, vacuolar
YBR127c	*VMA2*	H^+-ATPase V1 domain 60K subunit, vacuolar
YDL185w	*TFP1*	H^+-ATPase V1 domain 69K subunit, vacuolar
YPL066w	*VPS28*	involved in vacuolar traffic
YMR004w	*MVP1*	required for vacuolar protein sorting
YBR097w	*VPS15*	ser/thr protein kinase
YJL222w		strong similarity to Pep1p
YOR036w	*PEP12*	syntaxin (T-SNARE)
YNR006w	*VPS27*	vacuolar protein sorting-associated protein
YGL095c	*VPS45*	vacuolar protein sorting-associated protein
YDR323c	*PEP7*	vacuolar segregation protein
YPL045w	*VPS16*	vacuolar sorting protein
YPR173c	*VPS4*	vacuolar sorting protein
YAL002w	*VPS8*	vacuolar sorting protein
YML097c	*VPS9*	vacuolar sorting protein

extracellular transport

YJL085w	*EXO70*	70K exocyst component protein
YDR129c	*SAC6*	actin filament bundling protein, fimbrin
YKL209c	*STE6*	ATP-binding cassette transporter protein
YDL226c	*GCS1*	cell proliferation zinc-finger protein
YER136w	*GDI1*	GDP dissociation inhibitor
YLL043w	*FPS1*	glycerol channel protein
YJL044c	*GYP6*	GTPase-activating protein
YLR262c	*YPT6*	GTP-binding protein of the RAB family
YJL205c-a	*NCE1*	involved in non-classical protein export pathway
YPR149w	*NCE2*	involved in non-classical protein export pathway
YNL036w	*NCE3*	involved in non-classical protein export pathway
YOR326w	*MYO2*	myosin heavy chain
YJL093c	*TOK1*	outward-rectifier potassium channel
YMR308c	*PSE1*	protein secretion enhancer
YDR164c	*SEC1*	protein transport protein
YNL272c	*SEC2*	protein transport protein
YIL068c	*SEC6*	protein transport protein
YPR055w	*SEC8*	protein transport protein
YGR009c	*SEC9*	protein transport protein
YKL212w	*SAC1*	recessive suppressor of secretory defect
YLR166c	*SEC10*	required for exocytosis
YDR166c	*SEC5*	required for exocytosis
YNR049c	*MSO1*	secretion protein, multicopy suppressor of *sec1*
YER008c	*SEC3*	secretory pathway protein
YOR307c	*SLY41*	secretory pathway protein
YBR162w-a	*YSY6*	secretory pathway protein
YLR250w	*SSP120*	secretory protein
YCR032w		similarity to human CDC4L protein
YER006w		similarity to *P. polycephalum* myosin-related protein mlpA
YGR131w		strong similarity to Nce2p
YNL325c		suppressor of *sac1* mutation
YPL232w	*SSO1*	syntaxin-related protein
YGL233w	*SEC15*	vesicular traffic control protein
YBR080c	*SEC18*	vesicular-fusion protein, functional homologue of NSF

cellular import

YPL195w	*YKS4*	α- or γ- adaptin, large subunit of the clathrin-associated protein(AP) complex
YBL037w	*APL3*	α-adaptin, large subunit of the clathrin-associated protein(AP) complex
YFL039c	*ACT1*	actin
YDR129c	*SAC6*	actin filament bundling protein, fimbrin
YDL029w	*ACT2*	actin-like protein
YCL040w	*GLK1*	aldohexose specific glucokinase
YJR152w	*DAL5*	allantoate permease
YIR028w	*DAL4*	allantoin permease
YBR068c	*BAP2*	amino-acid permease
YEL063c	*CAN1*	amino-acid permease
YBR069c	*VAP1*	amino-acid permease
YGR121c	*MEP1*	ammonia permease of high capacity and moderate affinity
YKL135c	*APL2*	β-adaptin
YGR261c	*YKS5*	β-adaptin, large subunit of the clathrin-associated protein(AP) complex
YBR109c	*CMD1*	calmodulin
YDL226c	*GCS1*	cell proliferation zinc-finger protein
YMR058w	*FET3*	cell-surface ferroxidase
YGL077c	*HNM1*	choline permease
YGL206c	*CHC1*	clathrin heavy chain
YGR167w	*CLC1*	clathrin light chain
YLR170c	*APS1*	clathrin-associated protein (AP) complex, small subunit AP19
YJR058c	*APS2*	clathrin-associated protein 17, small subunit
YJR005w	*YAP80*	clathrin-associated protein complex, β subunit
YJL024c	*APS3*	clathrin-associated protein(AP) complex, small subunit
YPR124w	*CTR1*	copper transport protein
YHR175w	*CTR2*	copper transport protein
YNL243w	*SLA2*	cytoskeleton assembly control protein
YLL001w	*DNM1*	dynamin-related protein
YDL210w	*UGA4*	GABA-specific high-affinity permease
YPR029c	*APL4*	γ-adaptin, large subunit of the clathrin-associated protein(AP) complex
YLR081w	*GAL2*	galactose (and glucose) permease
YNR026c	*SEC12*	GDP/GTP exchange factor for Sar1p
YKR039w	*GAP1*	general amino-acid permease
YLL043w	*FPS1*	glycerol channel protein
YOR089c	*VPS21*	GTP-binding protein
YER031c	*YPT31*	GTP-binding protein of the RAB family
YKR014c	*YPT52*	GTP-binding protein of the RAB family
YML001w	*YPT7*	GTP-binding protein of the RAB family
YNL093w	*YPT53*	GTP-binding protein of the RAB family (RAS superfamily)
YNL318c	*HXT14*	hexose transport protein
YJL214w	*HXT8*	hexose transport protein
YJL219w	*HXT9*	hexose transport protein
YFL011w	*HXT10*	hexose transporter
YDL194w	*SNF3*	high-affinity glucose transporter
YEL069c	*HXT13*	high-affinity hexose transporter
YMR011w	*HXT2*	high-affinity hexose transporter
YDR343c	*HXT6*	high-affinity hexose transporter
YDR342c	*HXT7*	high-affinity hexose transporter
YML123c	*PHO84*	high-affinity inorganic phosphate/H^+ symporter
YNL142w	*MEP2*	high-affinity low capacity ammonia permease
YGR055w	*MUP1*	high-affinity methionine permease
YNL270c	*ALP1*	high-affinity permease for basic amino acids
YJL129c	*TRK1*	high-affinity potassium transport protein
YBR294w	*SUL1*	high-affinity sulphate transport protein
YOL020w	*SCM2*	high-affinity tryptophan transport protein
YGR191w	*HIP1*	histidine permease
YMR319c	*FET4*	low-affinity Fe(II) iron transport protein
YOL156w	*HXT11*	low-affinity glucose transporter
YHR094c	*HXT1*	low-affinity hexose transporter
YDR345c	*HXT3*	low-affinity hexose transporter
YNL268w	*LYP1*	lysine-specific high-affinity permease
YDR497c	*ITR1*	major myo-inositol permease
YBR298c	*MAL31*	maltose permease
YOL122c	*SMF1*	manganese transporter
YKR050w	*TRK2*	moderate-affinity potassium transport protein
YHR092c	*HXT4*	moderate to low-affinity glucose transporter
YOL103w	*ITR2*	myo-inositol transport protein
YLR240w	*VPS34*	phosphatidylinositol 3-kinase
YOR348c	*PUT4*	proline and γ-aminobutyrate permease
YER056c	*FCY2*	purine-cytosine permease
YLR088w	*GAA1*	required for attachment of GPI anchor onto proteins
YNL084c	*END3*	required for endocytosis and cytoskeletal organization
YJR090c	*GRR1*	required for glucose repression and for glucose and cation transport
YCR009c	*RVS161*	similarity to human amphiphysin and Rvs167p
YDR046c	*(PAP1)*	strong similarity to amino-acid transport proteins
YHR096c	*HXT5*	strong similarity to hexose transporters
YIL170w	*HXT12*	strong similarity to sugar transport proteins
YJR158w	*HXT16*	strong similarity to sugar transport proteins
YNR072w	*HXT17*	sugar transport protein
YOR329c	*SCD5*	suppressor of clathrin deficiency
YPR129w	*SCD6*	suppressor of clathrin deficiency
YBR021w	*FUR4*	uracil permease
YMR231w	*PEP5*	vacuolar biogenesis protein
YLR337w	*VRP1*	verprolin
YBR080c	*SEC18*	vesicular-fusion protein, functional homologue of NSF
YJR040w	*GEF1*	voltage-gated chloride channel protein
YMR243c	*ZRC1*	zinc- and cadmium resistance protein

other intracellular-transport activities

YKR054c	*DYN1*	dynein heavy chain, cytosolic
YER019c-a	*SEB2*	ER protein-translocation complex subunit
YLR378c	*SEC61*	ER protein-translocation complex subunit
YPL094c	*SEC62*	ER protein-translocation complex subunit
YOR254c	*SEC63*	ER protein-translocation complex subunit
YBR171w	*SEC66*	ER protein-translocation complex subunit
YLR292c	*SEC72*	ER protein-translocation complex subunit
YDR086c	*SSS1*	ER protein-translocation complex subunit
YCR075c	*ERS1*	intracellular protein transport protein
YHL019c	*APM2*	involved in clathrin-dependent transport processes
YJL034w	*KAR2*	nuclear fusion protein
YBR170c	*NPL4*	nuclear protein localization factor and ER translocation component
YKR093w	*PTR2*	peptide transporter
YKL198c	*PTK1*	polyamine transport enhancing protein
YDR040c	*ENA1*	P-type ATPase involved in Na^+ and Li^+ efflux
YER060w	*FCY21*	purine-cytosine permease
YNL183c	*NPR1*	ser/thr protein kinase
YOR100c		similarity to mitochondrial carrier protein YMC1
YGL216w		similarity to mouse kinesin-related protein KIF3
YPR009w		similarity to sterol uptake protein Sut1p
YHR123w	*EPT1*	sn-1,2-diacylglycerol ethanolamine- and cholinephosphotransferase
YER060w-a	*FCY22*	strong similarity to Fcy2p
YER087c-a	*SEB1*	strong similarity to mammalian Sec61β subunit

Cellular organization and biogenesis

organization and biogenesis of cell wall and plasma membrane

YLR342w	*FKS1*	1,3--D-glucan synthase, catalytic subunit
YGR032w	*GSC2*	1,3--D-glucan synthase, subunit
YOR362c	*PRE10*	26S proteasome subunit C1
YJR004c	*SAG1*	α-agglutinin
YNR044w	*AGA1*	α-agglutinin anchor subunit
YGL032c	*AGA2*	α-agglutinin binding subunit
YDR077w	*SED1*	abundant cell-surface glycoprotein
YJL005w	*CYR1*	adenylate cyclase
YCL007c	*CWH36*	affects the mannoprotein layer of the cell wall
YOR335c	*ALA1*	alanyl-tRNA synthetase, cytosolic
YPL061w	*ALD6*	aldehyde dehydrogenase, cytosolic
YJR152w	*DAL5*	allantoate permease
YIR028w	*DAL4*	allantoin permease
YBR068c	*BAP2*	amino-acid permease
YEL063c	*CAN1*	amino-acid permease
YBR069c	*VAP1*	amino-acid permease

YML116w	*ATR1*	aminotriazole and 4-nitroquinoline resistance protein
YGR121c	*MEP1*	ammonia permease of high capacity and moderate affinity
YDR144c	*MKC7*	aspartyl protease of the periplasmic space
YLR120c	*YAP3*	aspergillopepsin
YKL209c	*STE6*	ATP-binding cassette transporter protein
YGR281w	*YOR1*	ATP-binding cassette transporter protein required for oligomycin resistance
YCR011c	*ADP1*	ATP-dependent permease
YER123w	*YCK3*	casein kinase, isoform 3
YCL027w	*FUS1*	cell fusion protein
YMR058w	*FET3*	cell-surface ferroxidase
YKL096w	*CWP1*	cell-wall mannoprotein
YKL096w-a	*CWP2*	cell-wall mannoprotein
YNL322c	*KRE1*	cell-wall protein
YAR050w	*FLO1*	cell-wall protein involved in flocculation
YNL192w	*CHS1*	chitin synthase I
YBR038w	*CHS2*	chitin synthase II
YBR023c	*CHS3*	chitin synthase III
YGL077c	*HNM1*	choline permease
YOR010c	*TIR2*	cold-shock induced protein
YER011w	*TIR1*	cold-shock induced protein of the Tir1p, Tip1p family
YPR124w	*CTR1*	copper transport protein
YGR282c	*BGL2*	endo-1,3—glucanase of the cell wall
YOR190w	*SPR1*	exo-1,3—glucanase
YDR261c	*EXG2*	exo-1,3—glucanase minor isoform
YLR214w	*FRE1*	ferric (and cupric) reductase
YKL220c	*FRE2*	ferric (and cupric) reductase
YDL210w	*UGA4*	GABA-specific high-affinity permease
YLR081w	*GAL2*	galactose (and glucose) permease
YLL016w	*SDC25*	GDP/GTP exchange factor
YLR310c	*CDC25*	GDP/GTP exchange factor for Ras1p and Ras2p
YCR038c	*BUD5*	GDP/GTP exchange factor for Rsr1p/Bud1p
YKR039w	*GAP1*	general amino-acid permease
YMR307w	*GAS1*	glycophospholipid-anchored surface glycoprotein
YOL081w	*IRA2*	GTPase-activating protein for RAS proteins
YOR101w	*RAS1*	GTP-binding protein
YNL098c	*RAS2*	GTP-binding protein
YHR005c	*GPA1*	GTP-binding protein α subunit of the pheromone pathway
YOR212w	*STE4*	GTP-binding protein β subunit of the pheromone pathway
YJR086w	*STE18*	GTP-binding protein γ subunit of the pheromone pathway
YLR229c	*CDC42*	GTP-binding protein of RAS superfamily
YCR024c-a	*PMP1*	H^+-ATPase subunit, plasma membrane
YEL017c-a	*PMP2*	H^+-ATPase subunit, plasma membrane
YGL008c	*PMA1*	H^+-transporting P-type ATPase
YPL036w	*PMA2*	H^+-transporting P-type ATPase 2
YDR420w	*HKR1*	*Hansenula* MrakII k9 killer toxin-resistance protein
YLL024c	*SSA2*	heat-shock protein of HSP70 family, cytosolic
YNL318c	*HXT14*	hexose transport protein
YJL214w	*HXT8*	hexose transport protein
YJL219w	*HXT9*	hexose transport protein
YFL011w	*HXT10*	hexose transporter
YDL194w	*SNF3*	high-affinity glucose transporter
YEL069c	*HXT13*	high-affinity hexose transporter
YMR011w	*HXT2*	high-affinity hexose transporter
YDR343c	*HXT6*	high-affinity hexose transporter
YDR342c	*HXT7*	high-affinity hexose transporter
YML123c	*PHO84*	high-affinity inorganic phosphate/H^+ symporter
YNL142w	*MEP2*	high-affinity low-capacity ammonia permease
YGR055w	*MUP1*	high-affinity methionine permease
YNL270c	*ALP1*	high-affinity permease for basic amino acids
YJL129c	*TRK1*	high-affinity potassium transport protein
YBR294w	*SUL1*	high-affinity sulphate transport protein
YOL020w	*SCM2*	high-affinity tryptophan transport protein
YGR191w	*HIP1*	histidine permease
YBR140c	*IRA1*	inhibitory regulator protein of the RAS-cyclic AMP pathway
YNL291c	*MID1*	involved in Ca^{2+} influx during mating
YHL019c	*APM2*	involved in clathrin-dependent transport processes
YLR452c	*SST2*	involved in desensitization to α-factor pheromone
YCL051w	*LRE1*	involved in laminarinase resistance
YER118c	*SSU81*	involved in the *HOG1* high-osmolarity signal transduction pathway
YMR183c	*SSO2*	involved in vesicle transport from Golgi to plasma membrane
YLR155c	*ASP3A*	L-asparaginase II
YLR157c	*ASP3B*	L-asparaginase II
YLR158c	*ASP3C*	L-asparaginase II
YLR160c	*ASP3D*	L-asparaginase II
YMR319c	*FET4*	low-affinity Fe(II) iron transport protein
YOL156w	*HXT11*	low-affinity glucose transporter
YHR094c	*HXT1*	low-affinity hexose transporter
YDR345c	*HXT3*	low-affinity hexose transporter
YNL268w	*LYP1*	lysine-specific high-affinity permease
YDR497c	*ITR1*	major myo-inositol permease
YBR298c	*MAL31*	maltose permease
YOL122c	*SMF1*	manganese transporter
YIL047c	*SYG1*	member of the major facilitator superfamily
YKR050w	*TRK2*	moderate-affinity potassium transport protein
YHR092c	*HXT4*	moderate- to low-affinity glucose transporter
YGR014w	*MSB2*	multicopy suppressor of a *cdc24* bud emergence defect
YDR011w	*SNQ2*	multidrug resistance protein
YOL103w	*ITR2*	myo-inositol transport protein
YJL093c	*TOK1*	outward-rectifier potassium channel
YKR093w	*PTR2*	peptide transporter
YFL026w	*STE2*	pheromone α-factor receptor
YKL178c	*STE3*	pheromone a-factor receptor
YMR008c	*PLB1*	phospholipase B (lysophospholipase)
YOR153w	*PDR5*	pleiotropic drug resistance protein
YCR034w	*GNS1*	probable 1,3—glucan synthase subunit
YOR348c	*PUT4*	proline and γ-aminobutyrate permease
YPL140c	*MKK2*	protein kinase of the MAP kinase kinase (MEK) family
YGR009c	*SEC9*	protein transport protein
YER056c	*FCY2*	purine-cytosine permease
YIL140w	*SRO4*	required for axial pattern of budding
YCR089w	*FIG2*	required for efficient mating
YDR122w	*KIN1*	ser/thr protein kinase
YLR096w	*KIN2*	ser/thr protein kinase
YPL176c		similarity to chinese hamster transferrin receptor protein
YKR102w		similarity to Flo1p
YHL028w		similarity to mucin and other ser-thr rich proteins
YJR151c		similarity to mucin proteins
YMR317w		similarity to mucins, glucan 1,4—glucosidase and exo—sialidase
YCR098c	*GIT1*	similarity to phosphate transporter proteins
YIL013c	*PDR11*	similarity to Snq2p and other ATP-dependent permeases
YIL063c	*YRB2*	similarity to Yrp1p and Nup2p
YNL126w	*SPC98*	spindle pole body component
YHR139c	*SPS100*	sporulation-specific wall maturation protein
YOR328w	*PDR10*	strong similarity to ABC transporter proteins
YDR046c	*(PAP1)*	strong similarity to amino-acid transport proteins
YDR134c		strong similarity to Flo1p
YAL063c	*FLO9*	strong similarity to Flo1p
YLR343w		strong similarity to Gas1p and *C. albicans* pH responsive protein
YOL030w		strong similarity to glycoprotein GAS1
YHR096c	*HXT5*	strong similarity to hexose transporters
YDR536w	*STL1*	strong similarity to members of the sugar permease family
YIL170w	*HXT12*	strong similarity to sugar transport proteins
YJR158w	*HXT16*	strong similarity to sugar transport proteins
YNR072w	*HXT17*	sugar transport protein
YPL232w	*SSO1*	syntaxin-related protein
YBR067c	*TIP1*	temp.shock induced protein of the Srp1p/Tip1p family
YJR007w	*SUI2*	translation initiation factor eIF2, α subunit
YIL147c	*SLN1*	two-component signal transducer
YER125w	*RSP5*	ubiquitin-protein ligase
YBR021w	*FUR4*	uracil permease
YHL016c	*DUR3*	urea transport protein
YGL233w	*SEC15*	vesicular traffic control protein
YJR040w	*GEF1*	voltage-gated chloride channel protein
YMR243c	*ZRC1*	zinc- and cadmium resistance protein

organization and biogenesis of cytoskeleton

YML085c	*TUB1*	α-1 tubulin
YML124c	*TUB3*	α-3 tubulin
YFL039c	*ACT1*	actin
YMR092c	*AIP1*	actin cytoskeleton component
YDR129c	*SAC6*	actin filament bundling protein, fimbrin
YDL178w	*AIP2*	actin interacting protein 2
YJR065c	*ACT4*	actin-related protein
YCR088w	*ABP1*	actin-binding protein
YDL029w	*ACT2*	actin-like protein
YDR106w	*ARP10*	actin-related protein
YNL059c	*ARP5*	actin-related protein
YLR085c	*ARP6*	actin-related protein
YOR141c	*ARP8*	actin-related protein
YNR033w	*ARP9*	actin-related protein
YNL138w	*SRV2*	adenylate cyclase-associated protein, 70K
YAL020c	*ATS1*	α-tubulin suppressor
YER016w	*BIM1*	binding to microtubules
YFL037w	*TUB2*	β-tubulin
YOR265w	*RBL2*	β-tubulin binding protein
YBR200w	*BEM1*	bud emergence mediator
YLR319c	*BUD6*	bud site selection protein
YNL271c	*BNI1*	budding protein
YBR109c	*CMD1*	calmodulin
YCR002c	*CDC10*	cell division control protein
YGL116w	*CDC20*	cell division control protein
YLR314c	*CDC3*	cell division control protein
YFL009w	*CDC4*	cell division control protein
YHR129c	*ACT5*	centractin
YLR175w	*CBF5*	centromere/microtubule binding protein
YOR349w	*CIN1*	chromosome segregation protein
YLL050c	*COF1*	cofilin, actin binding and severing protein
YBL007c	*SLA1*	cytoskeleton assembly control protein
YNL243w	*SLA2*	cytoskeleton assembly control protein
YLL001w	*DNM1*	dynamin-related protein
YKR054c	*DYN1*	dynein heavy chain, cytosolic
YDR424c	*DYN2*	dynein light chain 1, cytosolic
YKL007w	*CAP1*	F-actin capping protein, α subunit
YIL034c	*CAP2*	F-actin capping protein, β subunit
YLR212c	*TUB4*	γ tubulin
YMR138w	*CIN4*	GTP-binding protein
YOR156c	*NFI1*	interacts with Cdc12p in 2-hybrid assay
YML104c	*MDM1*	intermediate filament protein
YLL021w	*SPA2*	involved in cell polarity
YPL241c	*CIN2*	involved in chromosome segregation
YDR085c	*AFR1*	involved in morphogenesis of the mating projection
YEL061c	*CIN8*	kinesin-related protein
YPR141c	*KAR3*	kinesin-related protein
YBL063w	*KIP1*	kinesin-related protein
YPL155c	*KIP2*	kinesin-related protein
YKL079w	*SMY1*	kinesin-related protein
YBR172c	*SMY2*	kinesin-related protein
YOR058c	*ASE1*	microtubule-associated protein
YJL042w	*MHP1*	microtubule-associated protein
YMR294w	*JNM1*	mitosis protein, involved in nuclear migration
YBL034c	*STU1*	mitotic spindle protein
YOR326w	*MYO2*	myosin heavy chain
YAL029c	*MYO4*	myosin heavy chain, unconventional (class V) isoform
YMR109w	*MYO5*	myosin I
YKL129c	*MYO3*	myosin type I
YHR023w	*MYO1*	myosin-1 isoform (type II myosin) heavy chain
YCL029c	*BIK1*	nuclear fusion protein
YOR122c	*PFY1*	profilin
YDR164c	*SEC1*	protein transport protein
YDR388w	*RVS167*	reduced viability upon starvation protein
YNL084c	*END3*	required for endocytosis and cytoskeletal organization
YBR130c	*SHE3*	required for mother cell-specific expression of HO
YGR078c	*PAC10*	required in the absence of Cin8p
YDR488c	*PAC11*	required in the absence of Cin8p
YER007w	*PAC2*	required in the absence of Cin8p
YJR076c	*CDC11*	septin
YHR107c	*CDC12*	septin
YDR218c	*SPR28*	septin-related sporulation protein
YNL059c		similarity to actin
YLR085c		similarity to actin and actin-related proteins
YLR429w		similarity to actin binding protein coronin
YPR034w		similarity to actins
YGL003c		similarity to *C. elegans* CDC20 protein and human p55CDC
YCR009c	*RVS161*	similarity to human amphiphysin and Rvs167p
YER018c		similarity to lamins
YOR367w		similarity to mammalian smooth muscle protein SM22
YIL149c		similarity to Mlp1p and myosin heavy chains
YGL216w		similarity to mouse kinesin-related protein KIF3
YKL179c		similarity to Nuf1p
YJR134c		similarity to paramyosin, myosin
YPL105c		similarity to Smy2p
YDR356w	*NUF1*	spindle pole body component
YNL126w	*SPC98*	spindle pole body component
YOL069w	*NUF2*	spindle pole body protein
YGR059w	*SPR3*	sporulation-specific septin
YEL071w		strong similarity to Aip2p
YOR134w	*BAG7*	structural homologue of Sac7p
YAL047c	*SPI6*	Stu2p Interactant
YKL022c	*CDC16*	subunit of anaphase-promoting complex (cyclosome)
YHR166c	*CDC23*	subunit of anaphase-promoting complex (cyclosome)
YBL084c	*CDC27*	subunit of anaphase-promoting complex (cyclosome)
YDR484w	*SAC2*	suppressor of actin mutation
YDR389w	*SAC7*	suppressor of actin mutation
YNL079c	*TPM1*	tropomyosin 1
YIL138c	*TPM2*	tropomyosin 2
YPL045w	*VPS16*	vacuolar sorting protein
YLR337w	*VRP1*	verprolin

organization and biogenesis of endoplasmic reticulum and Golgi

YML075c	*HMG1*	3-hydroxy-3-methylglutaryl-CoA reductase 1
YLR450w	*HMG2*	3-hydroxy-3-methylglutaryl-CoA reductase 2
YJR131w	*MNS1*	α-1,2-mannosidase
YDR483w	*KRE2*	α-1,2-mannosyltransferase
YER001w	*MNN1*	α-1,3-mannosyltransferase
YGL038c	*OCH1*	α-1,6-mannosyltransferase
YGL027c	*CWH41*	β-1,6-glucan assembly protein
YBR110w	*ALG1*	β-mannosyltransferase
YMR015c	*ERG5*	C-22 sterol desaturase
YGR060w	*ERG25*	C-4 sterol methyl oxidase
YLR056w	*ERG3*	C-5 sterol desaturase
YMR202w	*ERG2*	C-8 sterol isomerase
YGL167c	*PMR1*	Ca^{2+}-transporting P-type ATPase
YBR036c	*CSG2*	calcium dependent regulatory protein
YGL203c	*KEX1*	carboxypeptidase (YSC-α)
YPR113w	*PIS1*	CDP diacylglycerol-inositol 3-phosphatidyltransferase
YER026c	*CHO1*	CDP-diacylglycerol serine O-phosphatidyltransferase
YBR029c	*CDS1*	CDP-diacylglycerol synthase
YKL073w	*LHS1*	chaperone of the ER lumen
YGR202c	*PCT1*	cholinephosphate cytidylyltransferase
YOL062c	*APM4*	clathrin-associate protein YAP54
YPL259c	*APM1*	clathrin-associated protein

YDL195w	*SEC31*	component of the COPII coat of ER-Golgi vesicles
YGL200c	*EMP24*	component of the COPII-coated vesicles, 24K
YML012w	*ERV25*	component of the COPII-coated vesicles, 25K
YHR007c	*ERG11*	cytochrome P450 lanosterol 14a-demethylase
YNL130c	*CPT1*	diacylglycerol cholinephosphotransferase
YMR013c	*SEC59*	dolichol kinase
YPL227c	*ALG5*	dolichol-P-glucose synthetase
YPR183w	*DPM1*	dolichyl-phosphate β-D-mannosyltransferase
YJR143c	*PMT4*	dolichyl-phosphate-mannose-protein O-mannosyl transferase
YDL212w	*SHR3*	endoplasmic reticulum membrane protein
YNL238w	*KEX2*	endoproteinase of late Golgi compartment
YBL040c	*ERD2*	ER lumen protein-retaining receptor
YLR378c	*SEC61*	ER protein-translocation complex subunit
YPL094c	*SEC62*	ER protein-translocation complex subunit
YOR254c	*SEC63*	ER protein-translocation complex subunit
YBR171w	*SEC66*	ER protein-translocation complex subunit
YLR292c	*SEC72*	ER protein-translocation complex subunit
YDR086c	*SSS1*	ER protein-translocation complex subunit
YDR331w	*GPI8*	essential for GPI-anchor attachment
YHR190w	*ERG9*	farnesyl-diphosphate farnesyltransferase
YDR519w	*FKB2*	FK506/rapamycin-binding protein of the ER
YJR031c	*GEA1*	GDP/GTP exchange factor for ARF
YEL022w	*GEA2*	GDP/GTP exchange factor for ARF
YNR026c	*SEC12*	GDP/GTP exchange factor for Sar1p
YPR159w	*KRE6*	glucan synthase subunit
YGR143w	*SKN1*	glucan synthase subunit
YOR067c	*ALG8*	glucosyltransferase
YFL048c	*EMP47*	Golgi membrane protein
YDL192w	*ARF1*	GTP-binding protein of the ARF family
YPL218w	*SAR1*	GTP-binding protein of the ARF family
YFL038c	*YPT1*	GTP-binding protein of the RAB family
YLR262c	*YPT6*	GTP-binding protein of the RAB family
YPR165w	*RHO1*	GTP-binding protein of the RHO subfamily of RAS-like proteins
YNL090w	*RHO2*	GTP-binding protein of the RHO subfamily of RAS-like proteins
YEL042w	*GDA1*	guanosine diphosphatase
YDR189w	*SLY1*	hydrophilic suppressor of *ypt1* and member of the Sec1p family
YCR075c	*ERS1*	intracellular protein transport protein
YDL058w	*USO1*	intracellular protein transport protein
YLR220w	*CCC1*	involved in calcium regulation
YBR201w	*DER1*	involved in degradation proteins in the ER
YOR069w	*VPS5*	involved in Golgi retention and vacuolar sorting
YOR336w	*KRE5*	killer toxin-resistance protein
YHR072w	*ERG7*	lanosterol synthase
YKR061w	*KTR2*	mannosyltransferase
YDL095w	*PMT1*	mannosyltransferase
YAL023c	*PMT2*	mannosyltransferase
YOR321w	*PMT3*	mannosyltransferase
YBL082c	*RHK1*	mannosyltransferase
YJL139c	*YUR1*	mannosyltransferase
YKR001c	*VPS1*	member of the dynamin family of GTPases
YJR073c	*OPI3*	methylene-fatty-acyl-phospholipid synthase
YNL064c	*YDJ1*	mitochondrial and ER import protein
YPL085w	*SEC16*	multidomain vesicle coat protein
YHR042w	*NCP1*	NADPH-cytochrome P450 reductase
YFL025c	*BST1*	negative regulator of COPII vesicle formation
YJL034w	*KAR2*	nuclear fusion protein
YJL002c	*OST1*	oligosaccharyltransferase α subunit
YEL002c	*WBP1*	oligosaccharyltransferase β subunit
YMR149w	*SWP1*	oligosaccharyltransferase δ subunit
YOR103c	*OST2*	oligosaccharyltransferase ε subunit
YOR085w	*OST3*	oligosaccharyltransferase γ subunit
YGL022w	*STT3*	oligosaccharyltransferase subunit
YHR057c	*CYP2*	peptidyl-prolyl *cis-trans* isomerase
YCR069w	*SCC3*	peptidyl-prolyl *cis-trans* isomerase
YDR304c	*CYP5*	peptidyl-prolyl *cis-trans* isomerase D (cyclophilin D) of the ER
YGR157w	*CHO2*	phosphatidylethanolamine N-methyltransferase
YGR170w	*PSD2*	phosphatidylserine decarboxylase 2
YDR518w	*EUG1*	protein disulphide isomerase
YCL043c	*PDI1*	protein disulphide-isomerase
YHR079c	*IRE1*	protein kinase
YEL036c	*ANP1*	protein of the endoplasmatic reticulum
YCR067c	*SED4*	protein of the endoplasmic reticulum
YIL004c	*BET1*	protein transport protein
YDR040c	*ENA1*	P-type ATPase involved in Na^+ and Li^+ efflux
YKL212w	*SAC1*	recessive suppressor of secretory defect
YLR088w	*GAA1*	required for attachment of GPI anchor onto proteins
YPL050c	*MNN9*	required for complex N-glycosylation
YCL001w	*RER1*	required for correct localization of Sec12p
YGL145w	*TIP20*	required for ER to Golgi transport
YDR414c	*ERD1*	required for retention of luminal ER proteins
YMR004w	*MVP1*	required for vacuolar protein sorting
YDR498c	*SEC20*	secretory pathway protein
YJR010c-a	*SPC1*	signal peptidase 10.8K subunit
YML055w	*SPC2*	signal peptidase 18K subunit
YDR292c	*SRP101*	signal recognition particle receptor, α subunit
YIR022w	*SEC11*	signal sequence processing protein
YAL058w	*CNE1*	similarity to calnexin
YDR245w	*MNN10*	similarity to *S. pombe* galactosyltransferase
YMR214w	*SCJ1*	similarity to to *E. coli* dnaJ
YHR123w	*EPT1*	sn-1,2-diacylglycerol ethanolamine- and cholinephosphotransferase
YKL006c-a	*SFT1*	SNARE-like protein
YGR175c	*ERG1*	squalene monooxygenase
YGL055w	*OLE1*	stearoyl-CoA desaturase
YLR080w		strong similarity to Emp47p
YER087c-a	*SEB1*	strong similarity to mammalian Sec61β subunit
YOR099w	*KTR1*	strong similarity to mannosyltransferases
YGL226c-a	*OST5*	subunit of N-oligosaccharyltransferase, ζ subunit
YDR297w	*SUR2*	suppressor of *rvs161* and *rvs167* mutations
YBL102w	*SFT2*	suppressor of *sed5* ts mutants
YLR026c	*SED5*	syntaxin (T-SNARE)
YOR075w	*UFE1*	syntaxin (T-SNARE) of the ER
YOR219c	*STE13*	type IV dipeptidyl aminopeptidase
YMR022w	*QRI8*	ubiquitin-conjugating enzyme
YER100w	*UBC6*	ubiquitin-conjugating enzyme
YBR243c	*ALG7*	UDP-N-acetylglucosamine-1-phosphate transferase
YGR105w	*VMA21*	vacuolar ATPase assembly integral membrane protein
YBL017c	*PEP1*	vacuolar protein sorting/targeting protein
YGL225w	*GOG5*	vanadate-resistance protein

organization and biogenesis of chromosome structure

YJL081c	*ACT3*	actin-related protein
YDL220c	*CDC13*	cell division control protein
YBR195c	*MSI1*	chromatin assembly complex, subunit p50
YML102w	*CAC2*	chromatin assembly complex, subunit p60
YPR018w	*RLF2*	chromatin assembly complex, subunit p90
YGR218w	*CRM1*	chromosome region maintenance protein
YNL216w	*RAP1*	DNA-binding protein with repressor and activator activity
YER159c	*NCB1*	functional homologue of human NC2α
YMR072w	*ABF2*	high-mobility group protein
YPL127c	*HHO1*	histone H1 protein
YDR225w	*HTA1*	histone H2A
YBL003c	*HTA2*	histone H2A.2
YDR224c	*HTB1*	histone H2B
YBL002w	*HTB2*	histone H2B.2
YBR010w	*HHT1*	histone H3
YNL031c	*HHT2*	histone H3
YBR009c	*HHF1*	histone H4
YNL030w	*HHF2*	histone H4
YER161c	*SPT2*	HMG-like chromatin protein
YPL254w	*HFI1*	interacts functionally with histone H2A
YNR010w	*CSE2*	interacts with centromeric element CDEII
YGR099w	*TEL2*	involved in controlling telomere length and position effect
YJR060w	*CBF1*	kinetochore protein
YGR140w	*CBF2*	kinetochore protein complex CBF3, 110K subunit
YMR094w	*CTF13*	kinetochore protein complex CBF3, 58K subunit
YMR168c	*CEP3*	kinetochore protein complex CBF3, 71K subunit
YDR328c	*SKP1*	kinetochore protein complex CBF3, subunit D
YIL072w	*HOP1*	meiosis-specific protein
YLR263w	*RED1*	meiosis-specific protein
YLR457c	*NBP1*	Nap1p-binding protein
YDR162c	*NBP2*	Nap1p-binding protein
YBR089c-a	*NHP6B*	nonhistone chromosomal protein
YPR052c	*NHP6A*	nonhistone chromosomal protein related to mammalian HMG1
YDR174w	*HMO1*	nonhistone protein
YDL002c	*HMO2*	nonhistone protein
YKR048c	*NAP1*	nucleosome assembly protein I
YCL011c	*GBP2*	potential telomere-associated protein
YGL238w	*CSE1*	probable kinetochore protein
YGL194c	*RTL1*	putative deacetylase
YBR275c	*RIF1*	Rap1p-interacting factor 1
YKL089w	*MIF2*	required for normal chromosome segregation and spindle integrity
YJL074c	*SMC3*	required for structural maintenance of chromosomes
YLR086w		similarity to chromosome condensation proteins
YML069w	*POB3*	similarity to HMG proteins
YOR053w		similarity to protamines
YNL206c		similarity to structure-specific recognition proteins
YDL208w	*NHP2*	strong similarity to high mobility group (HMG) family
YEL026w		strong similarity to high mobility group-like protein Nhp2p
YOL012c	*HTA3*	strong similarity to histone H2A protein
YKL049c	*CSE4*	strong similarity to histone H3
YOR213c		subunit of the RSC complex
YCR052w	*RSC6*	subunit of the RSC complex
YFR037c	*RSC8*	subunit of the RSC complex
YLR321c	*SFH1*	subunit of the RSC complex
YIL126w	*STH1*	subunit of the RSC complex
YDR285w	*ZIP1*	synaptonemal complex protein
YLR233c	*EST1*	telomere elongation protein
YBL088c	*TEL1*	telomere length control protein
YPL128c	*TBF1*	telomere TTAGGG repeat-binding factor 1
YGR274c	*TAF145*	TFIID subunit (TBP-associated factor), 145K
YGR187c	*HGH1*	weak similarity to human Hmg1p and Hmg2p
YGR285c	*ZUO1*	zuotin (Z-DNA binding protein)

mitochondrial organization and biogenesis

YHR037w	*PUT2*	1-pyrroline-5-carboxylate dehydrogenase
YNL104c	*LEU4*	2-isopropylmalalate synthase
YMR282c	*AEP2*	2'-O-ribosyl phosphate transferase
YIL125w	*KGD1*	2-oxoglutarate dehydrogenase complex E1 component
YDR148c	*KGD2*	2-oxoglutarate dehydrogenase complex E2 component
YOL096c	*COQ3*	3,4-dihydroxy-5-hexaprenylbenzoate methyltransferase
YMR287c	*MSU1*	3'-5' exonuclease for RNA 3' ss-tail, mitochondrial
YDR232w	*HEM1*	5-aminolevulinate synthase
YML009c	*MRPL39*	60S ribosomal protein, mitochondrial
YMR108w	*ILV2*	acetolactate synthase
YCL009c	*ILV6*	acetolactate synthase, regulatory subunit
YAL054c	*ACS1*	acetyl-CoA synthetase
YER069w	*ARG5,6*	acetylglutamate kinase and acetylglutamyl-phosphate reductase
YOL140w	*ARG8*	acetylornithine aminotransferase
YLR304c	*ACO1*	aconitate hydratase
YDR234w	*LYS4*	aconitate hydratase
YER170w	*ADK2*	adenylate kinase, mitochondrial
YMR056c	*AAC1*	ADP/ATP carrier protein (MCF)
YBL030c	*AAC2*	ADP/ATP carrier protein (MCF)
YBR085w	*AAC3*	ADP/ATP carrier protein (MCF)
YMR083w	*ADH3*	alcohol dehydrogenase III
YER073w		aldehyde dehydrogenase (NAD^+)
YMR170c	*ALD5*	aldehyde dehydrogenase 2 (NAD^+), mitochondrial
YER086w	*ILV1*	anabolic serine and threonine dehydratase
YJL209w	*CBP1*	apo-cytochrome b pre-mRNA processing protein
YBR120c	*CBP6*	apo-cytochrome b pre-mRNA processing protein
YHL038c	*CBP2*	apo-cytochrome b pre-mRNA processing protein 2
YHR091c	*MSR1*	arginyl-tRNA synthetase, mitochondrial
YCR024c		asn-tRNA synthetase, mitochondrial
YKL106w	*AAT1*	aspartate transaminase, mitochondrial
YPL104w	*MSD1*	aspartate-tRNA ligase, mitochondrial
YGR008c	*STF2*	ATPase stabilizing factor
YDL130w-a	*STF1*	ATPase stabilizing factor, 10K
YDR377w	*ATP17*	ATPase synthase subunit f
YMR301c	*ATM1*	ATP-binding cassette transporter protein, mitochondrial
YBL022c	*PIM1*	ATP-dependent protease, mitochondrial
YPL029w	*SUV3*	ATP-dependent RNA helicase, mitochondrial
YER061c	*CEM1*	β-keto-acyl-ACP synthase
YHR208w	*TWT1*	branched chain amino-acid aminotransferase, mitochondrial
YBR084w	*MIS1*	C1-tetrahydrofolate synthase, mitochondrial
YML042w	*CAT2*	carnitine O-acetyltransferase
YER026c	*CHO1*	CDP-diacylglycerol serine O-phosphatidyltransferase
YNR001c	*CIT1*	citrate (si)-synthase, mitochondrial
YPR001w	*CIT3*	citrate (si)-synthase, mitochondrial
YBR291c	*CTP1*	citrate transport protein, mitochondrial (MCF)
YOR316c	*COT1*	cobalt accumulation protein
YOR017w	*PET127*	component of mitochondrial translation system
YPR155c	*NCA2*	control of mitochondrial synthesis of Atp6p and Atp8p
YKL011c	*CCE1*	cruciform-cutting endonuclease, mitochondrial
YML078w	*CPR3*	cyclophilin (peptidylprolyl isomerase), mitochondrial
YDR197w	*CBS2*	cytochrome b translational activator protein
YIL043c	*CBR1*	cytochrome b5 reductase
YKL150w	*MCR1*	cytochrome b5 reductase
YKL141w	*SDH3*	cytochrome b560 subunit of respiratory complex II
YJR048w	*CYC1*	cytochrome c isoform 1
YEL039c	*CYC7*	cytochrome c isoform 2
YOR037w	*CYC2*	cytochrome c mitochondrial import factor
YER058w	*PET117*	cytochrome c oxidase assembly factor
YPL132w	*COX11*	cytochrome c oxidase assembly protein
YML129c	*COX14*	cytochrome c oxidase assembly protein
YDR079w	*PET100*	cytochrome c oxidase assembly protein
YGL187c	*COX4*	cytochrome c oxidase subunit IV
YNL052w	*COX5A*	cytochrome c oxidase subunit V.A
YIL111w	*COX5B*	cytochrome c oxidase subunit Vb
YHR051w	*COX6*	cytochrome c oxidase subunit VI
YGL191w	*COX13*	cytochrome c oxidase subunit VIa
YLR038c	*COX12*	cytochrome c oxidase subunit VIB
YMR256c	*COX7*	cytochrome c oxidase subunit VII
YDL067c	*COX9*	cytochrome c oxidase subunit VIIA
YLR395c	*COX8*	cytochrome c oxidase subunit VIII
YKR066c	*CCP1*	cytochrome c peroxidase
YOR065w	*CYT1*	cytochrome c1
YER141w	*COX15*	cytochrome oxidase assembly factor
YER154w	*OXA1*	cytochrome oxidase biogenesis protein
YOR386w	*PHR1*	deoxyribodipyrimidine photo-lyase
YLR348c		dicarboxylate carrier protein
YFL018c	*LPD1*	dihydrolipoamide dehydrogenase
YNL071w	*LAT1*	dihydrolipoamide S-acetyltransferase

YJR016c	*ILV3*	dihydroxy-acid dehydratase
YDL174c	*DLD1*	D-lactate ferricytochrome c oxidoreductase (D-LCR)
YML061c	*PIF1*	DNA helicase involved in mitochondrial DNA repair and telomere length
YHR120w	*MSH1*	DNA mismatch repair protein, mitochondrial
YOR330c	*MIP1*	DNA-directed DNA polymerase γ, catalytic subunit, mitochondrial
YFL036w	*RPO41*	DNA-directed RNA polymerase, mitochondrial
YBR252w	*DUT1*	dUTP pyrophosphatase, mitochondrial
YOR211c	*MGM1*	dynamin-like protein
YGR207c	*ETF-β*	electron-transferring flavoprotein, β subunit
YLR393w	*ATP10*	F1F0 ATPase complex assembly protein
YNL315c	*ATP11*	F1F0-ATPase complex assembly protein
YJL180c	*ATP12*	F1F0-ATPase complex assembly protein
YKL016c	*ATP7*	F1F0-ATPase complex, F0 D subunit
YBL099w	*ATP1*	F1F0-ATPase complex, F1 α subunit
YJR121w	*ATP2*	F1F0-ATPase complex, F1 β subunit
YDL004w	*ATP16*	F1F0-ATPase complex, F1 δ subunit
YPL078c	*ATP4*	F1F0-ATPase complex, F1 δ subunit
YPL271w	*ATP15*	F1F0-ATPase complex, F1 ε subunit
YBR039w	*ATP3*	F1F0-ATPase complex, F1 γ subunit
YDR298c	*ATP5*	F1F0-ATPase complex, OSCP subunit
YLR295c	*ATP14*	F1F0-ATPase complex, subunit h
YIL134w	*FLX1*	FAD carrier protein, mitochondrial (MCF)
YPL172c	*COX10*	farnesyl transferase
YOR176w	*HEM15*	ferrochelatase
YPL262w	*FUM1*	fumarate hydratase
YOL033w	*MSE1*	glutamyl-tRNA synthetase, mitochondrial
YOL059w	*GPD3*	glycerol-3-phosphate dehydrogenase (NAD^+), mitochondrial
YIL155c	*GUT2*	glycerol-3-phosphate dehydrogenase, mitochondrial
YMR189w	*GSD2*	glycine decarboxylase subunit
YDR019c	*GCV1*	glycine decarboxylase T subunit
YFL016c	*MDJ1*	heat-shock protein - chaperone
YOR232w	*MGE1*	heat-shock protein - chaperone
YLR259c	*HSP60*	heat-shock protein - chaperone, mitochondrial
YJR045c	*SSC1*	heat-shock protein 70-related protein, mitochondrial
YDR258c	*HSP78*	heat-shock protein of clpb family of ATP-dependent proteases, mitochondrial
YBR003w	*COQ1*	hexaprenyl pyrophosphate synthetase
YMR072w	*ABF2*	high mobility group protein
YAL039c	*CYC3*	holocytochrome c synthase (cytochrome c haem lyase)
YKL087c	*CYT2*	holocytochrome c1 synthase
YMR038c	*LYS7*	homocitrate dehydrogenase
YDL181w	*INH1*	inhibitor of mitochondrial ATPase
YMR267w	*PPA2*	inorganic pyrophosphatase, mitochondrial
YLL009c	*COX17*	interacts genetically with *SCO1* and *SCO2* in cytochrome oxidase assembly
YGR028w	*MSP1*	intra-mitochondrial sorting protein
YJR034w	*PET191*	involved in assembly of cytochrome oxidase
YMR302c	*PRP12*	involved in early maturation of pre-rRNA
YOL009c	*MDM12*	involved in mitochondrial inheritance
YAL010c	*MDM10*	involved in mitochondrial morphology and inheritance
YIR021w	*MRS1*	involved in mitochondrial RNA splicing of COB mRNA
YBR037c	*SCO1*	involved in stabilization of Cox1p and Cox2p
YNL066w	*SUN4*	involved in the aging process
YCL017c	*NFS1*	involved in tRNA processing and mitochondrial metabolsim
YNL037c	*IDH1*	isocitrate dehydrogenase (NAD^+) subunit 1, mitochondrial
YOR136w	*IDH2*	isocitrate dehydrogenase (NAD^+) subunit 2, mitochondrial
YDL066w	*IDP1*	isocitrate dehydrogenase ($NADP^+$), mitochondrial
YPL040c	*ISM1*	isoleucine-tRNA ligase, mitochondrial
YLR355c	*ILV5*	ketol-acid reducto-isomerase
YML054c	*CYB2*	lactate dehydrogenase cytochrome b2
YLR382c	*NAM2*	leucine-tRNA ligase, mitochondrial
YOR196c	*LIP5*	lipoic acid synthase
YNL073w	*MSK1*	lysyl-tRNA synthetase, mitochondrial
YKL085w	*MDH1*	malate dehydrogenase, mitochondrial
YOR130c	*ARG11*	member of the mitochondrial carrier family (MCF)
YDL198c	*YHM1*	member of the mitochondrial carrier family (MCF)
YGR171c	*MSM1*	methionyl-tRNA synthetase
YGR029w	*ERV1*	mitochondrial biogenesis and regulation of cell cycle
YBR192w	*RIM2*	mitochondrial carrier protein (MCF)
YPR058w	*YMC1*	mitochondrial carrier protein (MCF)
YBR104w	*YMC2*	mitochondrial carrier protein (MCF)
YOR020c	*HSP10*	mitochondrial chaperonin
YJR144w	*MGM101*	mitochondrial genome maintenance protein
YMR023c	*MSS1*	mitochondrial GTPase involved in expression of *COX1*
YNR017w	*MAS6*	mitochondrial inner membrane import translocase subunit
YJL143w	*TIM17*	mitochondrial inner membrane import translocase subunit
YIL022w	*TIM44*	mitochondrial inner membrane import translocase subunit
YMR035w	*IMP2*	mitochondrial inner membrane protease subunit
YKL134c	*(MIP1)*	mitochondrial intermediate peptidase
YNL131w	*TOM22*	mitochondrial outer membrane import receptor subunit
YGR082w	*TOM20*	mitochondrial outer membrane import receptor subunit, 20K
YMR060c	*TOM37*	mitochondrial outer membrane import receptor subunit, 37K
YMR203w	*TOM40*	mitochondrial outer membrane import receptor subunit, 40K
YOR045w	*TOM6*	mitochondrial outer membrane import receptor subunit, 6K
YNL070w	*TOM7*	mitochondrial outer membrane import receptor subunit, 7K
YNL055c	*POR1*	mitochondrial outer membrane porin
YNL121c	*TOM70*	mitochondrial outer membrane specialized import receptor
YGL143c	*MRF1*	mitochondrial peptide chain release factor
YLR163c	*MAS1*	mitochondrial processing peptidase
YHR024c	*MAS2*	mitochondrial processing peptidase, catalytic 53K (α) subunit
YMR150c	*IMP1*	mitochondrial protease
YDR375c	*BCS1*	mitochondrial protein of the CDC48/PAS1/SEC18 (AAA) family of ATPases
YDR120c	*TRM1*	N2,N2-dimethylguanine tRNA methyltransferase
YML120c	*NDI1*	NADH-ubiquinone-6 oxidoreductase
YOR355w	*GDS1*	*nam9-1* suppressor
YJL208c	*NUC1*	nuclease, mitochondrial
YAR035w	*YAT1*	outer carnitine acetyltransferase, mitochondrial
YNR041c	*COQ2*	para-hydroxybenzoate-polyprenyltransferase
YPR047w	*MSF1*	phenylalanine-tRNA ligase α subunit, mitochondrial
YJR077c	*MIR1*	phosphate transport protein, mitochondrial (MCF)
YNL169c	*PSD1*	phosphatidylserine decarboxylase 1
YLR203c	*MSS51*	possibly involved in translational activation of COX1 and COB mRNA
YGL068w		probable ribosomal protein L12
YLR168c	*(MSF1)*	probably involved in intramitochondrial protein sorting
YLR142w	*PUT1*	proline oxidase
YER017c	*AFG3*	protease of the *SEC18*/*CDC48*/*PAS1* family of ATPases (AAA)
YPR024w	*YME1*	protease of the *SEC18*/*CDC48*/*PAS1* family of ATPases (AAA)
YMR089c	*YTA12*	protease of the *SEC18*/*CDC48*/*PAS1* family of ATPases (AAA)
YDL044c	*MTF2*	protein involved in mRNA splicing and protein synthesis, mitochondrial
YIL136w	*OM45*	protein of the outer mitochondrial membrane
YAL011w	*FUN36*	protein of unknown function
YJL023c	*PET130*	protein synthesis protein, mitochondrial
YER014w	*HEM14*	protoporphyrinogen oxidase, mitochondrial
YER178w	*PDA1*	pyruvate dehydrogenase (lipoamide) α subunit
YBR221c	*PDB1*	pyruvate dehydrogenase (lipoamide) β subunit
YGR193c	*PDX1*	pyruvate dehydrogenase complex protein X
YJR095w	*ACR1*	regulator of acetyl-CoA synthetase activity
YBR091c	*MRS5*	regulator of mitochondrial intron splicing
YPL215w	*CBP3*	required for assembly of cytochrome bc1 complex
YEL059c-a	*SOM1*	required for mitochondrial Imp1 peptidase function
YLL006w	*MMM1*	required for mitochondrial shape and structure
YKL093w	*MBR1*	required for optimal growth on glycerol
YLR067c	*PET309*	required for stability and translation of *COX1* mRNA
YMR257c	*PET111*	required for translation of *COX2* mRNA
YBL080c	*PET112*	required to maintain RHO^+ mitochondrial DNA
YBR185c	*MBA1*	respiratory chain assembly protein
YML091c	*RPM2*	ribonuclease P, mitochondrial
YKL170w	*MRPL38*	ribosomal protein L14, mitochondrial
YJL063c	*MRPL8*	ribosomal protein L17, mitochondrial
YPR166c	*MRP2*	ribosomal protein S14
YPL013c		ribosomal protein S16, mitochondrial
YBR251w	*MRPS5*	ribosomal protein S5, mitochondrial
YBR146w	*MRPS9*	ribosomal protein S9, mitochondrial
YKR006c	*MRPL13*	ribosomal protein YmL13, mitochondrial
YNL005c	*MRPL2*	ribosomal protein YmL2, mitochondrial
YKR085c	*MRPL20*	ribosomal protein YmL20, mitochondrial
YBR282w	*MRPL27*	ribosomal protein YmL27, mitochondrial
YDR462w		ribosomal protein YmL28, mitochondrial
YNL252c		ribosomal protein YmL30, mitochondrial
YKL138c	*MRPL31*	ribosomal protein YmL31, mitochondrial
YCR003w	*MRPL32*	ribosomal protein YmL32, mitochondrial
YBR122c	*MRPL36*	ribosomal protein YmL36, mitochondrial
YBR268w	*MRPL37*	ribosomal protein YmL37, mitochondrial
YPL173w		ribosomal protein YmL40, mitochondrial
YGR076c	*MRPL25*	ribosomal protein YmR26 (YML25), mitochondrial
YMR225c	*MRPL44*	ribosomal protein YmR44, mitochondrial
YML025c		ribosomal protein, mitochondrial
YDR347w	*MRP1*	ribosomal protein, mitochondrial
YGR084c	*MRP13*	ribosomal protein, mitochondrial
YKL003c	*MRP17*	ribosomal protein, mitochondrial
YDR405w	*MRP20*	ribosomal protein, mitochondrial
YHL004w	*MRP4*	ribosomal protein, mitochondrial
YKL167c	*MRP49*	ribosomal protein, mitochondrial
YKL142w	*MRP8*	ribosomal protein, mitochondrial
YLR312w-a	*MRPL15*	ribosomal protein, mitochondrial
YBL038w	*MRPL16*	ribosomal protein, mitochondrial
YMR193w	*MRPL24*	ribosomal protein, mitochondrial
YMR024w	*MRPL3*	ribosomal protein, mitochondrial
YMR286w	*MRPL33*	ribosomal protein, mitochondrial
YLR439w	*MRPL4*	ribosomal protein, mitochondrial
YHR147c	*MRPL6*	ribosomal protein, mitochondrial
YGR220c	*MRPL9*	ribosomal protein, mitochondrial
YDR337w	*MRPS28*	ribosomal protein, mitochondrial
YNL137c	*NAM9*	ribosomal protein, mitochondrial
YOR158w	*PET123*	ribosomal protein, mitochondrial
YCR046c	*PETCR46*	ribosomal protein, mitochondrial
YFR049w	*YMR31*	ribosomal protein, mitochondrial
YDR194c	*MSS116*	RNA helicase of the DEAD box family, mitochondrial
YMR228w	*MTF1*	RNA polymerase specific factor, mitochondrial
YOR334w	*MRS2*	RNA splicing protein and member of the mitochondrial carrier family (MCF)
YJL133w	*MRS3*	RNA splicing protein and member of the mitochondrial carrier family (MCF)
YKR052c	*MRS4*	RNA splicing protein and member of the mitochondrial carrier family (MCF)
YOR201c	*PET56*	rRNA (guanosine-2'-O-)-methyltransferase
YBR263w	*SHM1*	serine hydroxymethyltransferase, mitochondrial
YPR037c		similarity to Erv1p and rat ALR protein
YDR376w	*ARH1*	similarity to human adrenodoxin reductase
YBL013w		similarity to methionyl-tRNA formyltransferase
YJR113c		similarity to mitochondrial ribosomal protein S7
YKL120w	*PMT*	similarity to mitochondrial uncoupling protein (MCF)
YMR244w		similarity to NCA3 and SUN4 protein
YOR150w		similarity to ribosomal protein L13
YPR134w	*MSS18*	splicing protein
YGR222w	*PET54*	splicing protein and translational activator, mitochondrial
YCR028c-a	*RIM1*	ssDNA-binding protein, mitochondrial
YKL192c		strong similarity to acyl-carrier proteins
YBR024w	*SCO2*	strong similarity to Sco1p
YHR117w	*TOM71*	strong similarity to Tom70p/Mas70p
YKL148c	*SDH1*	succinate dehydrogenase flavoprotein
YLL041c	*SDH2*	succinate dehydrogenase iron-sulphur protein subunit
YDR178w	*SDH4*	succinate dehydrogenase membrane anchor subunit for Sdh2p
YHR008c	*SOD2*	superoxide dismutase (Mn), mitochondrial
YHR050w	*SMF2*	suppressor of mitochondrial matrix mutant
YKL194c	*MST1*	threonine-tRNA ligase, mitochondrial
YLR069c	*MEF1*	translation elongation factor G, mitochondrial
YOR187w	*TUF1*	translation elongation factor TU, mitochondrial
YJL102w	*MEF2*	translation elongation factor, mitochondrial
YOL023w	*IFM1*	translation initiation factor 2, mitochondrial
YDL069c	*CBS1*	translational activator of cob mRNA
YER153c	*PET122*	translational activator of cytochrome c oxidase subunit III
YNR045w	*PET494*	translational activator, mitochondrial
YOR274w	*MOD5*	tRNA isopentenyltransferase
YER168c	*CCA1*	tRNA nucleotidyltransferase
YDR268w	*MSW1*	tryptophanyl-tRNA synthetase, mitochondrial
YPL097w	*MSY1*	tyrosyl-tRNA synthetase
YFR033c	*QCR6*	ubiquinol-cytochrome c reductase 17K protein
YPR191w	*QCR2*	ubiquinol-cytochrome c reductase 40K subunit II
YBL045c	*COR1*	ubiquinol-cytochrome c reductase 44K core protein
YHR001w-a	*QCR10*	ubiquinol-cytochrome c reductase 8.5K subunit
YGR174c	*CBP4*	ubiquinol-cytochrome c reductase assembly factor
YGL119w	*ABC1*	ubiquinol-cytochrome c reductase complex assembly protein
YEL024w	*RIP1*	ubiquinol-cytochrome c reductase iron-sulphur protein
YDR529c	*QCR7*	ubiquinol-cytochrome c reductase subunit 7
YGR183c	*QCR9*	ubiquinol-cytochrome c reductase subunit 9
YJL166w	*QCR8*	ubiquinol-cytochrome c reductase subunit VIII
YML021c	*UNG1*	uracil-DNA glycosylase
YGR094w	*VAS1*	valyl-tRNA synthetase
YIL114c	*POR2*	voltage-dependent anion channel (YVDAC2)
YNL237w	*YTP1*	weak similarity to mitochondrial electron transport proteins

peroxisomal organization and biogenesis

YIL160c	*POT1*	acetyl-CoA C-acyltransferase, peroxisomal
YAL054c	*ACS1*	acetyl-CoA synthetase
YGL205w	*POX1*	acyl-CoA oxidase
YBR222c	*PCS60*	AMP-binding protein, peroxisomal
YML042w	*CAT2*	carnitine O-acetyltransferase
YDR256c	*CTA1*	catalase A, peroxisomal
YCR005c	*CIT2*	citrate (si)-synthase, peroxisomal
YKR009c	*FOX2*	hydratase-dehydrogenase-epimerase, peroxisomal
YER065c	*ICL1*	isocitrate lyase
YKL188c	*PAT1*	long-chain-fatty-acid transporter

YPL147w	*PXA1*	long-chain-fatty-acid transporter
YER015w	*FAA2*	long-chain-fatty-acid-CoA ligase
YDL078c	*MDH3*	malate dehydrogenase, peroxisomal
YNL117w	*MLS1*	malate synthase 1
YIR031c	*DAL7*	malate synthase 2
YDR329c	*PAS3*	peroxisomal assembly protein
YDR265w	*PAS4*	peroxisomal assembly protein
YJL210w	*PAS5*	peroxisomal assembly protein
YNL329c	*PAS8*	peroxisomal assembly protein
YDR142c	*PAS7*	peroxisomal import protein
YOL147c	*PMP27*	peroxisomal membrane protein
YLR191w	*PAS20*	peroxisomal protein involved in protein import
YDR244w	*PAS10*	putative peroxisomal targeting signal receptor
YLR109w		similarity to *C. boidinii* peroxisomal membrane protein 20K A
YGR077c		similarity to peroxisomal matrix protein Per1p
YLR251w		similarity to peroxisomal rat membrane protein PMP22
YNL202w	*SPS19*	sporulation-specific protein
YGR133w	*PAS2*	ubiquitin-conjugating enzyme
YBR204c		weak similarity to peroxisomal serine-active lipase

endosomal organization and biogenesis

YKL135c	*APL2*	β-adaptin
YNL192w	*CHS1*	chitin synthase I
YBR023c	*CHS3*	chitin synthase III
YJR058c	*APS2*	clathrin-associated protein 17, small subunit
YJR005w	*APL1*	clathrin-associated protein complex, β subunit
YBR288c	*APM3*	clathrin-associated protein complex, medium subunit
YLR170c	*APS1*	clathrin-associated protein complex, small subunit AP19
YDL145c	*RET1*	coatomer complex α subunit
YDR238c	*SEC26*	coatomer complex β subunit of secretory pathway vesicles
YGL137w	*SEC27*	coatomer complex β′ subunit
YFR051c	*RET2*	coatomer complex δ subunit
YNL287w	*SEC21*	coatomer complex γ subunit
YPL010w	*RET3*	coatomer complex ζ subunit
YPR181c	*SEC23*	component of COPII coat of ER-golgi vesicles
YDR170c	*SEC7*	component of non-clathrin vesicle coat
YLR083c	*EMP70*	endosomal protein
YPR017c	*DSS4*	GDP/GTP exchange factor for Sec4p
YOR089c	*VPS21*	GTP-binding protein
YDL137w	*ARF2*	GTP-binding protein of the ARF family
YML001w	*YPT7*	GTP-binding protein of the RAB family
YFL005w	*SEC4*	GTP-binding protein of the RAS superfamily
YLR268w	*SEC22*	high copy suppressor of ypt1 null mutation
YPL085w	*SEC16*	multidomain vesicle protein
YLR208w	*SEC13*	protein transport protein
YKL196c		similarity to Sec22p
YLR080w		strong similarity to Emp47p
YDR107c		strong similarity to Emp70p
YAL030w	*SNC1*	strong similarity to synaptic vesicle-associated membrane protein
YOR327c	*SNC2*	strong similarity to synaptobrevin
YBL050w	*SEC17*	transport vesicle fusion protein
YNR006w	*VPS27*	vacuolar protein sorting-associated protein
YBR080c	*SEC18*	vesicular-fusion protein, functional homologue of NSF
YLR078c	*BOS1*	weak similarity to synaptobrevin (V-SNARE)

vacuolar and lysosomal organization and biogenesis

YGL156w	*AMS1*	α-mannosidase
YBR286w	*APE3*	aminopeptidase Y, vacuolar
YKL103c	*LAP4*	aminopeptidase yscl, vacuolar
YPL154c	*PEP4*	aspartyl protease
YDL128w	*VCX1*	Ca^{2+}-transport (H^+/Ca^{2+} exchange) protein, vacuolar
YGL006w	*PMC1*	Ca^{2+}-transporting P-type ATPase
YMR297w	*PRC1*	carboxypeptidase y, serine-type protease
YHR028c	*DAP2*	dipeptidyl aminopeptidase B
YDR135c	*YCF1*	glutathione S-conjugate transporter, vacuolar
YJL172w	*CPS1*	Gly-X carboxypeptidase YSCS
YKL119c	*VPH2*	H^+-ATPase assembly protein, vacuolar
YHR039c-a	*VMA10*	H^+-ATPase V0 domain 13K subunit, vacuolar
YEL027w	*CUP5*	H^+-ATPase V0 domain 17K subunit, vacuolar
YPL234c	*TFP3*	H^+-ATPase V0 domain 17K subunit, vacuolar
YLR447c	*VMA6*	H^+-ATPase V0 domain 36K subunit, vacuolar
YOR270c	*VPH1*	H^+-ATPase V0 domain 95K subunit, vacuolar
YGR020c	*VMA7*	H^+-ATPase V1 domain 14K subunit, vacuolar
YOR332w	*VMA4*	H^+-ATPase V1 domain 27K subunit, vacuolar
YEL051w	*VMA8*	H^+-ATPase V1 domain 32K subunit, vacuolar
YKL080w	*VMA5*	H^+-ATPase V1 domain 42K subunit, vacuolar
YPR036w	*VMA13*	H^+-ATPase V1 domain 54K subunit, vacuolar
YBR127c	*VMA2*	H^+-ATPase V1 domain 60K subunit, vacuolar
YDL185w	*TFP1*	H^+-ATPase V1 domain 69K subunit, vacuolar
YFR019w	*FAB1*	probable PI P 5-kinase
YEL060c	*PRB1*	protease B, vacuolar
YDR481c	*PHO8*	repressible vacuolar alkaline phosphatase
YIL099w	*SGA1*	sporulation specific glucan 1,4–-glucosidase
YOR036w	*PEP12*	syntaxin (T-SNARE)
YOR106w	*VAM3*	syntaxin related protein
YMR231w	*PEP5*	vacuolar biogenesis protein
YLR148w	*PEP3*	vacuolar membrane protein
YGL212w	*VAM7*	vacuolar morphogenesis protein
YJL053w	*PEP8*	vacuolar protein-sorting/targeting protein
YJL154c	*VPS35*	vacuolar protein-sorting protein
YDR323c	*PEP7*	vacuolar segregation protein
YPL045w	*VPS16*	vacuolar sorting protein
YDR495c	*VPS3*	vacuolar sorting protein
YLR396c	*VPS33*	vacuolar sorting protein
YPR173c	*VPS4*	vacuolar sorting protein

other cellular organization and biogenesis activities

YKL157w	*APE2*	aminopeptidase yscII
YIL015w	*BAR1*	barrier pepsin
YJL174w	*KRE9*	cell wall synthesis protein
YBR092c	*PHO3*	constitutive acid phosphatase
YLR286c	*CTS1*	endochitinase
YLR300w	*EXG1*	exo-1,3–glucanase (I/II), major isoform
YIR019c	*STA1*	extracellular α-1,4-glucan glucosidase
YPL187w	*MFα1*	mating pheromone α-1 factor
YGL089c	*MFα2*	mating pheromone α-2 factor
YDR461w	*MFA1*	mating pheromone a-factor 1
YNL145w	*MFA2*	mating pheromone a-factor 2
YKL163w	*PIR3*	member of the Pir1p/Pir2p/Pir3p family
YBR093c	*PHO5*	repressible acid phosphatase
YKL164c	*PIR1*	required for tolerance to heat-shock
YML008c	*ERG6*	S-adenosyl-methionine δ-24-sterol-c-methyltransferase
YAR071w	*PHO11*	secreted acid phosphatase
YHR215w	*PHO12*	secreted acid phosphatase
YNL160w	*YGP1*	secreted glycoprotein
YLR250w	*SSP120*	secretory protein
YBR046c	*ZTA1*	similarity to ζ-crystallin
YMR215w		similarity to GAS1 protein
YNL275w		similarity to human band 3 anion transport protein
YIL162w	*SUC2*	sucrose hydrolyzing enzyme

Signal transduction

pheromone response generation

YDR264c	*AKR1*	ankyrin repeat-containing protein
YPL161c	*BEM4*	bud emergence protein
YCL027w	*FUS1*	cell fusion protein
YNL053w	*MSG5*	dual-specifity protein phosphatase
YJL157c	*FAR1*	factor arrest protein
YAL041w	*CDC24*	GDP/GTP exchange factor for Cdc42p
YHR005c	*GPA1*	GTP-binding protein α subunit of the pheromone pathway
YOR212w	*STE4*	GTP-binding protein β subunit of the pheromone pathway
YJR086w	*STE18*	GTP-binding protein γ subunit of the pheromone pathway
YLR229c	*CDC42*	GTP-binding protein of RAS superfamily
YNL173c	*MDG1*	GTP-binding protein of the pheromone pathway
YER020w	*GPA2*	guanine nucleotide-binding regulatory protein
YLR452c	*SST2*	involved in desensitization to α-factor pheromone
YMR052w	*FAR3*	involved in pheromone-mediated cell cycle arrest
YDR461w	*MFA1*	mating pheromone a-factor 1
YIL047c	*SYG1*	member of the major facilitator superfamily
YBL016w	*FUS3*	mitogen-activated protein kinase (MAP kinase)
YFL026w	*STE2*	pheromone α-factor receptor
YKL178c	*STE3*	pheromone α-factor receptor
YCL032w	*STE50*	pheromone response pathway protein
YDR103w	*STE5*	pheromone signal transduction pathway protein
YGR040w	*KSS1*	ser/thr protein kinase of the MAP kinase family
YLR362w	*STE11*	ser/thr protein kinase of the MEKK family
YHL007c	*STE20*	ser/thr protein kinase of the pheromone pathway
YDL159w	*STE7*	ser/thr/tyr protein kinase of MAP kinase kinase family
YHR146w		similarity to pheromone-response G-protein YNL173c
YHR084w	*STE12*	transcriptional activator

morphogenesis

YBR200w	*BEM1*	bud emergence mediator
YPL161c	*BEM4*	bud emergence protein
YBR109c	*CMD1*	calmodulin
YAL041w	*CDC24*	GDP/GTP exchange factor for Cdc42p
YGR070w	*ROM1*	GDP/GTP exchange protein for Rho1p
YLR371w	*ROM2*	GDP/GTP exchange protein for Rho1p
YPL115c	*BEM3*	GTPase-activating protein for Cdc42p and Rho1p
YLR229c	*CDC42*	GTP-binding protein of RAS superfamily
YPR165w	*RHO1*	GTP-binding protein of the RHO subfamily of RAS-like proteins
YNL090w	*RHO2*	GTP-binding protein of the RHO subfamily of RAS-like proteins
YPL106c	*SSE1*	heat-shock protein of HSP70 family
YOR149c	*SMP3*	protein kinase C pathway protein
YPL140c	*MKK2*	protein kinase of the MAP kinase kinase (MEK) family
YDL135c	*RDI1*	RHO GDP dissociation inhibitor with activity towards Rho1p
YOR231w	*MKK1*	ser/thr protein kinase
YBL105c	*PKC1*	ser/thr protein kinase
YHR030c	*SLT2*	ser/thr protein kinase of MAP kinase family
YJL095w	*BCK1*	ser/thr protein kinase of the MEKK family
YER167w	*BCK2*	ser/thr protein kinase of the protein kinase C pathway
YOL113w	*SKM1*	Ste20p/PAK-like protein kinase
YGL106w		strong similarity to calmodulins
YKL161c		strong similarity to ser/thr-specific protein kinase Slt2p
YPL089c	*RLM1*	transcription factor of the MADS box family
YGL095c	*VPS45*	vacuolar protein sorting-associated protein

osmosensing

YDL022w	*GPD1*	glycerol-3-phosphate dehydrogenase (NAD^+), cytoplasmic
YFL014w	*HSP12*	heat-shock protein
YER118c	*SSU81*	involved in the *HOG1* high-osmolarity signal transduction pathway
YCR073c	*SSK22*	MAP kinase kinase kinase
YNR031c	*SSK2*	MAP kinase kinase kinase of the high-osmolarity signal transduction pathway
YDL006w	*PTC1*	protein ser/thr phosphatase 2c
YLR113w	*HOG1*	ser/thr protein kinase of MAP kinase (MAPK) family
YHR030c	*SLT2*	ser/thr protein kinase of MAP kinase family
YKL161c		strong similarity to ser/thr-specific protein kinase Slt2p
YIL147c	*SLN1*	two-component signal transducer
YLR006c	*SSK1*	two-component signal transducer
YJL128c	*PBS2*	tyrosine protein kinase of the MAP kinase kinase family
YDR069c	*DOA4*	ubiquitin-specific protease

nutritional response

YJL005w	*CYR1*	adenylate cyclase
YLR178c	*TFS1*	cdc25-dependent nutrient- and ammonia-response cell-cycle regulator
YBR195c	*MSI1*	chromatin assembly complex, subunit p50
YLL016w	*SDC25*	GDP/GTP exchange factor
YLR310c	*CDC25*	GDP/GTP exchange factor for Ras1p and Ras2p
YOR101w	*RAS1*	GTP-binding protein
YNL098c	*RAS2*	GTP-binding protein
YER020w	*GPA2*	guanine nucleotide-binding regulatory protein
YPL106c	*SSE1*	heat-shock protein of HSP70 family
YBR140c	*IRA1*	inhibitory regulator protein of the RAS-cyclic AMP pathway
YIL119c	*RPI1*	negative regulator of RAS-cAMP pathway
YIR026c	*YVH1*	protein tyrosine phosphatase
YOR208w	*PTP2*	protein-tyrosine-phosphatase
YOL110w	*SHR5*	RAS suppressor
YPL084w	*BRO1*	required for normal response to nutrient limitation
YLR362w	*STE11*	ser/thr protein kinase of the MEKK family
YDL159w	*STE7*	ser/thr/tyr protein kinase of MAP kinase kinase family
YHR084w	*STE12*	transcriptional activator

other signal-transduction activities

YPL268w	*PLC1*	1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase
YNL138w	*SRV2*	adenylate cyclase-associated protein, 70K
YFR014c	*CMK1*	Ca^{2+}/calmodulin-dependent ser/thr protein kinase, type I
YOL016c	*CMK2*	Ca^{2+}/calmodulin-dependent ser/thr protein kinase, type II
YPL203w	*TPK2*	cAMP-dependent protein kinase 2, catalytic subunit
YKL166c	*TPK3*	cAMP-dependent protein kinase 3, catalytic subunit
YJL164c	*SRA3*	cAMP-dependent protein kinase, catalytic subunit 1
YMR028w	*TAP42*	component of the Tor signalling pathway
YML064c	*TEM1*	GTP-binding protein of the RAS superfamily
YJL146w	*IDS2*	*IME2*-dependent signalling protein
YJR066w	*TOR1*	phosphatidylinositol 3-kinase
YKL203c	*TOR2*	phosphatidylinositol 3-kinase
YLR240w	*VPS34*	phosphatidylinositol 3-kinase
YNL267w	*PIK1*	phosphatidylinositol 4-kinase
YNL106c	*PIE3*	phosphatidylinositol phosphate phosphatase
YLR305c	*STT4*	phosphatidylinositol-4-kinase
YDL101c	*DUN1*	protein kinase

YHR079c	*IRE1*	protein kinase
YMR016c	*SOK2*	regulatory protein in the PKA signal transduction pathway
YGR216c	*GPI1*	required for N-acetylglucosaminyl phosphatidylinositol synthesis
YPL153c	*SPK1*	ser/thr/tyr protein kinase
YHR046c		similarity to bovine myo-inositol-1(or 4)-monophosphatase
YBR260c		similarity to *C. elegans* GTPase-activating protein
YGR136w		similarity to chicken growth factor receptor-binding protein GRB2 homologue
YDR379w		similarity to Dbm1p
YDR208w	*MSS4*	similarity to human PI P 5-kinase
YCR027c		similarity to human Ras-related GTP-binding protein
YDR287w		similarity to inositol monophosphatase
YLR150w	*MPT4*	specific affinity for guanine-rich quadruplex nucleic acids
YPR054w	*SMK1*	sporulation-specific MAP kinase
YNL132w		strong similarity to *A. ambisexualis* antheridiol steroid receptor
YHR206w	*SKN7*	transcription factor with similarity to Hsf1p

Cell rescue

stress response generation

YFR052w	*NIN1*	26S proteasome regulatory subunit
YIL075c	*SEN3*	26S proteasome regulatory subunit
YDR074w	*TPS2*	α,α-trehalose-phosphate synthase, 105K subunit
YBR126c	*TPS1*	α,α-trehalose-phosphate synthase, 56K subunit
YPR026w	*ATH1*	acid trehalase, vacuolar
YIL033c	*SRA1*	cAMP dependent protein kinase, regulatory subunit
YDR477w	*SNF1*	carbon catabolite derepressing ser/thr protein kinase
YHR135c	*YCK1*	casein kinase I isoform
YNL154c	*YCK2*	casein kinase I isoform
YGR088w	*CTT1*	catalase T, cytosolic
YDR251w	*PAM1*	coiled-coil protein multicopy suppressor of loss of PP2A
YOR010c	*TIR2*	cold shock induced protein
YER011w	*TIR1*	cold-shock induced protein of the Tir1p, Tip1p family
YDR155c	*CPH1*	cyclophilin (peptidylprolyl isomerase)
YEL039c	*CYC7*	cytochrome c isoform 2
YER062c	*HOR2*	DL-glycerol phosphatase
YDR263c	*DIN7*	DNA damage inducible protein
YGL021w	*ALK1*	DNA damage-responsive protein
YAL015c	*NTG1*	DNA repair protein
YDR519w	*FKB2*	FK506/rapamycin-binding protein of the ER
YGR234w	*YHB1*	flavohemoglobin
YIR037w	*HYR1*	glutathione peroxidase
YLL043w	*FPS1*	glycerol channel protein
YNL098c	*RAS2*	GTP-binding protein
YHR064c		heat-shock protein
YMR173w	*DDR48*	heat-shock protein
YOL052c-a	*DDRA2*	heat-shock protein
YMR186w	*HSC82*	heat-shock protein
YLL026w	*HSP104*	heat-shock protein
YFL014w	*HSP12*	heat-shock protein
YBR072w	*HSP26*	heat-shock protein
YCR021c	*HSP30*	heat-shock protein
YPL240c	*HSP82*	heat-shock protein
YFL016c	*MDJ1*	heat-shock protein - chaperone
YLR259c	*HSP60*	heat-shock protein - chaperone, mitochondrial
YJR045c	*SSC1*	heat-shock protein 70-related protein, mitochondrial
YDR258c	*HSP78*	heat-shock protein of clpb family of ATP-dependent proteases, mitochondrial
YEL030w		heat-shock protein of HSP70 family
YAL005c	*SSA1*	heat-shock protein of HSP70 family
YER103w	*SSA4*	heat-shock protein of HSP70 family
YPL106c	*SSE1*	heat-shock protein of HSP70 family
YLL024c	*SSA2*	heat-shock protein of HSP70 family, cytosolic
YBL075c	*SSA3*	heat-shock protein of HSP70 family, cytosolic
YBR169c	*SSE2*	heat-shock protein of the HSP70 family
YGL073w	*HSF1*	heat-shock transcription factor
YMR251w-a	*HOR7*	hyperosmolarity-responsive protein
YMR273c	*ZDS1*	involved in negative regulation of cell polarity
YER118c	*SSU81*	involved in the *HOG1* high-osmolarity signal transduction pathway
YDR293c	*SSD1*	involved in the tolerance to high concentration of Ca^{2+}
YKL143w	*LTV1*	low-temperature viability protein
YGR100w	*MIC1*	Mac1p interacting protein
YCR073c	*SSK22*	MAP kinase kinase kinase
YNR031c	*SSK2*	MAP kinase kinase kinase of the high-osmolarity signal transduction pathway
YJL158c		member of the Pir1p/Hsp150p/Pir3p family
YJL159w		member of the Pir1p/Hsp150p/Pir3p family
YJL160c		member of the Pir1p/Hsp150p/Pir3p family
YKL163w	*PIR3*	member of the Pir1p/Pir2p/Pir3p family
YMR021c	*MAC1*	metal binding activator
YNL064c	*YDJ1*	mitochondrial and ER import protein
YGL178w	*MPT5*	multicopy suppressor of *pop2*
YGR159c	*NSR1*	nuclear localization sequence binding protein
YGL115w	*SNF4*	nuclear regulatory protein
YJR051w	*OSM1*	osmotic growth protein
YMR175w	*SIP18*	osmotic stress protein
YBR070c	*SAT2*	osmotolerance protein
YHR057c	*CYP2*	peptidyl-prolyl *cis-trans* isomerase
YOR014w	*RTS1*	potential regulatory subunit of protein phosphatase 2A
YMR174c	*PAI3*	protease A (ysca) inhibitor IA3
YHR079c	*IRE1*	protein kinase
YPL140c	*MKK2*	protein kinase of the MAP kinase kinase (MEK) family
YOL064c	*MET22*	protein ser/thr phosphatase
YDL006w	*PTC1*	protein ser/thr phosphatase 2c
YDR436w	*PPZ2*	protein ser/thr phosphatase of the PP-1 family
YOR208w	*PTP2*	protein-tyrosine-phosphatase
YJR090c	*GRR1*	required for glucose repression and for glucose and cation transport
YPL084w	*BRO1*	required for normal response to nutrient limitation
YKL164c	*PIR1*	required for tolerance to heat-shock
YPR005c	*HAL1*	salt-induced protein
YNL160w	*YGP1*	secreted glycoprotein
YGL190c	*CDC55*	ser/thr phosphatase 2A regulatory subunit B
YML016c	*PPZ1*	ser/thr phosphatase required for normal osmoregulation
YJL165c	*HAL5*	ser/thr protein kinase
YOR231w	*MKK1*	ser/thr protein kinase
YBL105c	*PKC1*	ser/thr protein kinase
YLR113w	*HOG1*	ser/thr protein kinase of MAP kinase (MAPK) family
YHR030c	*SLT2*	ser/thr protein kinase of MAP kinase family
YDL025c		ser/thr protein kinase of the DEAD/DEAH box family
YJL095w	*BCK1*	ser/thr protein kinase of the MEKK family
YER167w	*BCK2*	ser/thr protein kinase of the protein kinase C pathway
YDR227w	*SIR4*	silencing regulatory protein
YKL088w		similarity to *C. tropicalis* hal3 protein, to C-term. of Sis2p and to hypothetical protein YOR054c
YBR044c		similarity to chaperonin HSP60 proteins
YJR147w		similarity to heat-shock transcription factors
YGR249w	*MGA1*	similarity to heat-shock transcription factors
YBR054w	*YRO2*	similarity to HSP30 heat-shock protein Yro1p
YAR020c		similarity to members of the Srp1p/Tip1p family
YCR008w	*SAT4*	similarity to Npr1p and Hal5p protein kinases
YOR009w		similarity to SRP1 and TIR2 proteins
YCR060w		similarity to stress inducible protein Sti1p
YMR037c		stress responsive regulatory protein
YOR027w	*STI1*	stress-induced protein
YBL009w		strong similarity to DNA damage responsive Alk1p
YLR369w		strong similarity to heat-shock protein 70-related proteins
YHL046c		strong similarity to members of the Srp1p/Tip1p family
YIL011w		strong similarity to members of the Srp1p/Tip1p family
YOL161c		strong similarity to members of the Srp1p/Tip1p family
YJL223c	*PAU1*	strong similarity to members of the Srp1p/Tip1p family
YEL049w	*PAU2*	strong similarity to members of the Srp1p/Tip1p family
YCR104w	*PAU3*	strong similarity to members of the Srp1p/Tip1p family
YFL020c	*PAU5*	strong similarity to members of the Srp1p/Tip1p family
YLR461w	*PAU4*	strong similarity to members of the Tir1p/Tip1p family
YNR076w	*PAU6*	strong similarity to members of the Tir1p/Tip1p family
YDR033w		strong similarity to putative heat-shock protein Yro2p
YKL161c		strong similarity to ser/thr-specific protein kinase Slt2p
YOR054c		strong similarity to SIS2 protein and *C. tropicalis* hal3 protein
YBR067c	*TIP1*	temp.shock induced protein of the Srp1p/Tip1p family
YGR144w	*THI4*	thiamine-repressed protein
YBL093c	*ROX3*	transcription factor
YGL181w	*GTS1*	transcription factor of the Gcs1p/Glo3p/Sps18p family
YMR043w	*MCM1*	transcription factor of the MADS box family
YHR206w	*SKN7*	transcription factor with similarity to Hsf1p
YKL062w	*MSN4*	transcriptional activator
YML007w	*YAP1*	transcriptional activator involved in oxidative stress response
YIL147c	*SLN1*	two-component signal transducer
YLR006c	*SSK1*	two-component signal transducer
YJL128c	*PBS2*	tyrosine protein kinase of the MAP kinase kinase family
YLL039c	*UBI4*	ubiquitin
YMR022w	*QRI8*	ubiquitin-conjugating enzyme
YBR082c	*UBC4*	ubiquitin-conjugating enzyme
YDR059c	*UBC5*	ubiquitin-conjugating enzyme
YER125w	*RSP5*	ubiquitin-protein ligase
YKL210w	*UBA1*	ubiquitin-protein ligase

DNA repair (direct repair, base excision repair and nucleotide excision repair)

YER142c	*MAG1*	3-methyladenine DNA glycosylase
YML060w	*OGG1*	8-oxoguanine DNA glycosylase
YKL114c	*APN1*	AP endonuclease
YJL092w	*HPR5*	ATP-dependent DNA helicase
YPL204w	*HRR25*	casein kinase I, ser/thr/tyr protein kinase
YPL022w	*RAD1*	component of the nucleotide excision repairosome
YDL108w	*KIN28*	cyclin-dependent ser/thr protein kinase
YOR386w	*PHR1*	deoxyribodipyrimidine photo-lyase
YOR368w	*RAD17*	DNA damage checkpoint control protein
YER176w		DNA dependent ATPase/DNA helicase B
YLR032w	*RAD5*	DNA helicase
YIL143c	*SSL2*	DNA helicase
YML061c	*PIF1*	DNA helicase involved in mitochondrial DNA repair and telomere length
YER171w	*RAD3*	DNA helicase/ATPase
YDL164c	*CDC9*	DNA ligase
YEL019c	*MMS21*	DNA repair protein
YAL015c	*NTG1*	DNA repair protein
YMR137c	*PSO2*	DNA repair protein
YML095c	*RAD10*	DNA repair protein
YCR066w	*RAD18*	DNA repair protein
YER095w	*RAD51*	DNA repair protein
YDR076w	*RAD55*	DNA repair protein
YDR004w	*RAD57*	DNA repair protein
YOR346w	*REV1*	DNA repair protein
YDR369c	*XRS2*	DNA repair protein
YGL163c	*RAD54*	DNA-dependent ATPase of the Snf2p family
YDL102w	*CDC2*	DNA-directed DNA polymerase δ, catalytic 125K subunit
YFR023w	*PES4*	DNA-directed DNA polymerase ε suppressor
YNL262w	*POL2*	DNA-directed DNA polymerase ε, catalytic subunit A
YPL167c	*REV3*	DNA-directed DNA polymerase ζ
YKR056w	*RNC1*	endo-exonuclease
YOL043c	*NTG2*	endonuclease III-like glycosylase 2
YER162c	*RAD4*	excision repair protein
YLR288c	*MEC3*	G2-specific checkpoint protein
YFL014w	*HSP12*	heat-shock protein
YIL128w	*MMS19*	involved in repair and RNA polymerase transcription
YMR035w	*IMP2*	mitochondrial inner membrane protease subunit
YKR095w	*MLP1*	myosin-like protein related to Uso1p
YMR201c	*RAD14*	nucleotide excision repair protein
YBR114w	*RAD16*	nucleotide excision repair protein
YEL037c	*RAD23*	nucleotide excision repair protein
YJR052w	*RAD7*	nucleotide excision repair protein
YDL200c	*MGT1*	O6-methylguanine DNA repair methyltransferase
YDL101c	*DUN1*	protein kinase
YML032c	*RAD52*	recombination and DNA repair protein
YIL139c	*REV7*	required for DNA damage induced mutagenesis
YBR073w	*RDH54*	required for meiosis
YIL066c	*RNR3*	ribonucleotide reductase, repair inducible large subunit
YPL153c	*SPK1*	ser/thr/tyr protein kinase
YDR061w		similarity to *E. coli* deoxyribodipyrimidine photolyase
YAL019w	*FUN30*	similarity to helicases of the Snf2/Rad54 family
YLR035c		similarity to human mutL protein homologue
YPR056w		similarity to human transcription factor BTF2/TFIIH subunit p34
YFR038w		strong similarity to mouse lymphocyte specific helicase
YOR206w		strong similarity to Rad4p
YGR258c	*RAD2*	structure-specific nuclease of the nucleotide excision repairosome
YDL088c	*ASM4*	suppressor of temperature-sensitive mutations in Pol3p
YDR311w	*TFB1*	TFIIH subunit (transcription initiation factor), 75K
YPR025c	*CCL1*	TFIIH subunit (transcription initiation factor), cyclin C component
YLR005w	*SSL1*	TFIIH subunit (transcription initiation factor), factor B
YPL122c	*TFB2*	TFIIH subunit (transcription/repair factor)
YDR460w	*TFB3*	TFIIH subunit (transcription/repair factor)
YGL058w	*RAD6*	ubiquitin-conjugating enzyme

detoxificaton

YOL052c	*SPE2*	adenosylmethionine decarboxylase
YFL050c	*ALR2*	aluminum resistance protein
YML116w	*ATR1*	aminotriazole and 4-nitroquinoline resistance protein
YNL259c	*ATX1*	antioxidant protein and metal homeostasis factor
YLR398c	*SKI2*	antiviral protein and putative helicase
YKL004w	*AUR1*	aureobasidin-resistance protein
YDR256c	*CTA1*	catalase A, peroxisomal
YGR088w	*CTT1*	catalase T, cytosolic
YJR104c	*SOD1*	copper-zinc superoxide dismutase
YKR066c	*CCP1*	cytochrome c peroxidase
YPR198w	*SGE1*	drug resistance protein
YJL101c	*GSH1*	glutamate-cysteine ligase

YDR513w	*TTR1*	glutaredoxin
YPL091w	*GLR1*	glutathione reductase (NADPH)
YDR135c	*YCF1*	glutathione S-conjugate transporter, vacuolar
YGR197c	*SNG1*	involved in nitroguanidine resistance
YDL168w	*SFA1*	long-chain alcohol dehydrogenase
YCR023c		member of major facilitator superfamily multidrug-resistance protein family 2
YNR055c	*HOL1*	member of major facilitator superfamily multidrug-resistance protein subfamily 1
YHR053c	*CUP1A*	metallothionein
YHR055c	*CUP1B*	metallothionein
YOR079c	*ATX2*	multicopy suppressor of SOD-linked defects
YDR011w	*SNQ2*	multidrug resistance protein
YHR042w	*NCP1*	NADPH-cytochrome P450 reductase
YOR018w	*ROD1*	O-dinitrobenzene,calcium and zinc resistance protein
YKL064w	*MNR2*	overexpression overcomes manganese toxicity
YOR266w	*PNT1*	pentamidine resistance factor
YDR538w	*PAD1*	phenylacrylic acid decarboxylase
YOR153w	*PDR5*	pleiotropic drug resistance protein
YBL005w	*PDR3*	pleiotropic drug resistance regulatory protein
YGL016w	*PDR6*	pleiotropic drug resistance regulatory protein
YIL120w		similarity to antibiotic resistance proteins
YIL121w		similarity to antibiotic resistance proteins
YLR299w		similarity to *B. subtilis* γ-glutamyltransferase
YHL040c		similarity to *C. carbonum* toxin pump
YHL047c		similarity to *C. carbonum* toxin pump
YBR180w		similarity to drug resistance proteins
YDL100c		similarity to *E. coli* arsenical pump-driving ATPase
YLL015w		similarity to metal resistance proteins
YBR293w		similarity to multidrug resistance proteins
YGR138c		similarity to multidrug resistance proteins
YHR048w		similarity to multidrug resistance proteins
YKR103w		similarity to multidrug resistance proteins
YKR104w		similarity to multidrug resistance proteins
YLL028w		similarity to multidrug resistance proteins
YPR156c		similarity to multidrug resistance proteins
YOR273c		similarity to resistance proteins
YOR251c		similarity to thiosulphate sulphurtransferases
YOR247w		similarity to vanadate-sensitive suppressor Svs1p
YJR025c		strong similarity to 3-hydroxyanthranilate 3,4-dioxygenase
YOL130w		strong similarity to Alr2p
YMR279c		strong similarity to aminotriazole resistance protein
YOR378w		strong similarity to aminotriazole resistance protein
YBR008c		strong similarity to benomyl/methotrexate resistance protein
YCL069w		strong similarity to drug resistance protein SGE1
YKL026c		strong similarity to glutathione peroxidase
YBR244w		strong similarity to glutathione peroxidases
YKL033w-a		strong similarity to holacid-halidohydrolase
YFR022w		strong similarity to Rod1p
YLR046c		strong similarity to Rta1p and Rtm1p protein
YER185w		strong similarity to Rtm1p
YKR105c		strong similarity to Sge1p and hypothetical protein YCL069w
YJR015w		strong similarity to Sng1p
YNR070w		strong similarity to Snq2p
YBL064c		strong similarity to thiol-specific antioxidant enzyme
YDR453c		strong similarity to thiol-specific antioxidant protein
YGL254w	*FZF1*	sulphite resistance protein
YPL092w	*SSU1*	sulphite sensitivity protein
YHR008c	*SOD2*	superoxide dismutase (Mn), mitochondrial
YOR031w	*CRS5*	suppressor of *cup1* deletion, metallothionein-like protein
YML028w	*TSA1*	thiol-specific antioxidant
YLR043c	*TRX1*	thioredoxin I
YGR209c	*TRX2*	thioredoxin II
YGL013c	*PDR1*	transcription factor
YMR043w	*MCM1*	transcription factor of the MADS box family
YDR423c	*CAD1*	transcriptional activator
YPL163c	*SVS1*	vanadate sensitive suppressor
YLL060c		weak similarity to glutathione transferase
YMR243c	*ZRC1*	zinc- and cadmium resistance protein

detoxification involving cytochrome P450

YMR015c	*ERG5*	C-22 sterol desaturase
YDR402c	*DIT2*	cytochrome P450 56
YHR007c	*ERG11*	cytochrome P450 lanosterol 14a-demethylase
YLL057c		similarity to *E. coli* dioxygenase
YDR403w	*DIT1*	spore wall maturation protein

cell death and ageing

YOR101w	*RAS1*	GTP-binding protein
YNL098c	*RAS2*	GTP-binding protein
YOL025w	*LAG2*	involved in determining longevity
YJL116c	*NCA3*	involved in regulation of synthesis of Atp6p and Atp8p
YNL066w	*SUN4*	involved in the ageing process
YKR042w	*UTH1*	involved in the ageing process
YHL003c	*LAG1*	longevity-assurance protein
YER167w	*BCK2*	ser/thr protein kinase of the protein kinase C pathway
YDR227w	*SIR4*	silencing regulatory protein
YKL008c		strong similarity to Lag1p
YIL123w		strong similarity to Sun4p,Nca3p

degradation of exogenous polynucleotides

YGL213c	*SKI8*	antiviral protein of the β-transducin (WD-40) repeat family

other cell-rescue activities

YPR189w	*SKI3*	antiviral protein
YJR069c	*HAM1*	controls 6-N-hydroxylaminopurine sensitivity and mutagenesis
YGR213c	*RTA1*	involved in 7-aminocholesterol resistance
YKL110c	*KTI12*	involved in resistance to *K. lactis* killer toxin
YER076c		similarity to killer toxin Khr1p
YER187w		similarity to killer toxin KHS

Unclassified proteins

YEL052w	*AFG1*	ATPase family gene
YDR184c	*BAT1*	binds Aip3p
YCR072c		β-transducin family (WD-40 repeat) protein
YLR459w	*CDC91*	cell division control protein
YJL079c	*PRY1*	contains homology to the plant PR-1 class of pathogen related proteins
YKR013w	*PRY2*	contains homology to the plant PR-1 class of pathogen related proteins
YJL078c	*PRY3*	contains homology to the plant PR-1 class of pathogen related proteins
YOR302w	*CPA1-l*	CPA1 leader peptide
YML113w	*DAT1*	datin, oligo(dA)/oligo(dT)-binding protein
YKL020c	*SPT23*	dosage-dependent suppressor of Ty-induced promotor mutations
YJL168c	*EZL1*	enhancer of zeste-like
YOR123c	*LEO1*	extremely hydrophilic protein
YLL065w	*GIN11*	growth inhibitory protein
YDL213c		has an RNA recognition domain in the N-terminal region
YDR505c	*PSP1*	high copy suppressor of ts of mutations in DNA polymerase α
YNL087w		homology to YOR086c
YLR161w		identical to hypothetical proteins YLR156w and YLR161w
YLR156w		identical to hypothetical proteins YLR159w and YLR161w
YLR159w		identical to hypothetical proteins YLR161w and YLR156w
YHR212c		identical with hypothetical protein YAR060c
YHR214w		identical with hypothetical protein YAR066w
YAR066w		identical with hypothetical protein YIL169c
YCR020c-a	*MAK31*	involved in stability of L-A dsRNA-containing particles
YPL049c	*DIG1*	MAP kinase-associated protein, down-regulator of invasive growth
YKL034w		member of Kazal serine protease inhibitors family
YLR397c	*AFG2*	member of the Sec18p, Pas1p, Cdc48p, TBP-1 family of ATPases
YCR007c		member of the YBR302c family
YGL036w	*MTC2*	Mtf1 two hybrid clone 2
YMR200w	*ROT1*	mutant suppresses *tor2* mutation
YAL059w	*SIM1*	mutants are hypersensitive to calcofluor white
YCR019w	*MAK32*	necessary for structural stability of L-A dsRNA-containing particles
YNL061w	*NOP2*	nucleolar protein
YFR011c		ochre suppressor tyr-tRNA
YLR196w	*PWP1*	periodic tryptophan protein
YDL235c	*YPD1*	phosphorelay intermediate between Sln1p and Ssk1p
YOR014w	*RTS1*	potential regulatory subunit of protein phosphatase 2A
YOR181w	*LAS17*	proline-rich protein
YER129w	*PAK1*	protein kinase
YOR223w		protein of unknown function
YHR093w	*AHT1*	protein of unknown function
YDL167c	*ARP1*	protein of unknown function
YHR191c	*CTF8*	protein of unknown function
YAL008w	*FUN14*	protein of unknown function
YAR002w	*FUN17*	protein of unknown function
YAR014c	*FUN2*	protein of unknown function
YAL031c	*FUN21*	protein of unknown function
YAL014c	*FUN34*	protein of unknown function
YAR008w	*FUN4*	protein of unknown function
YAL033w	*FUN53*	protein of unknown function
YDL234c	*GYP7*	protein of unknown function
YPL054w	*LEE1*	protein of unknown function
YDR335w	*MSN5*	protein of unknown function
YIR006c	*PAN1*	protein of unknown function
YDL105w	*QRI2*	protein of unknown function
YLR204w	*QRI5*	protein of unknown function
YLL003w	*SFI1*	protein of unknown function
YDR350c	*TCM10*	protein of unknown function
YDL169c	*UGX2*	protein of unknown function
YJR050w	*UTR3*	protein of unknown function
YEL035c	*UTR5*	protein of unknown function
YJR067c	*YAE1*	protein of unknown function
YGR172c	*YIP1*	protein of unknown function
YNL044w	*YIP3*	protein of unknown function
YBR111c	*YSA1*	protein of unknown function
YGR123c	*PPT1*	protein ser/thr phosphatase
YDL230w	*PTP1*	protein tyrosine phosphatase
YER075c	*PTP3*	protein tyrosine phosphatase
YPR073c	*LTP1*	protein-tyrosine-phosphatase
YNL128w		putative phosphatase
YCL039w		regulatory protein of the β-transducin family
YJR055w	*HIT1*	required for growth at high temperature
YNL085w	*MKT1*	required for propagation of M2 dsRNA satellite of L-A virus
YLL046c	*RNP1*	ribonucleoprotein
YOR046c	*DBP5*	RNA helicase
YDR466w		ser/thr protein kinase
YDR490c		ser/thr protein kinase
YGL179c		ser/thr protein kinase
YGL180w		ser/thr protein kinase
YKL168c		ser/thr protein kinase
YKL171w		ser/thr protein kinase
YLR063w		ser/thr protein kinase
YMR291w		ser/thr protein kinase
YNL161w		ser/thr protein kinase
YOL045w		ser/thr protein kinase
YOL100w		ser/thr protein kinase
YAR018c	*KIN3*	ser/thr protein kinase
YOR233w	*KIN4*	ser/thr protein kinase
YCR091w	*KIN82*	ser/thr protein kinase
YLL019c	*KNS1*	ser/thr protein kinase
YHR082c	*KSP1*	ser/thr protein kinase
YDL079c	*MRK1*	ser/thr protein kinase
YOR267c		ser/thr protein kinase with similarity to Npr1p
YMR216c		ser/thr protein kinase with similarity to *S. pombe* dsk1
YKL116c		ser/thr protein kinase with similarity to *S. pombe* nim1 protein
YBL056w	*PTC3*	ser/thr protein phosphatase PP2C
YDR507c	*GIN4*	ser/thr-protein kinase
YKL126w	*YPK1*	ser/thr-specific protein kinase
YOR007c		similarities to protein phosphatases
YOR119c	*RIO1*	similarity to a *C. elegans* ZK632.3 protein
YLR405w		similarity to *A. brasilense* nifR3 protein
YML079w		similarity to *A. brasilense* nifR3 protein
YML080w		similarity to *A. brasilense* nifR3 protein
YDR205w		similarity to *A. eutrophus* cation efflux system membrane protein czcD, rat zinc transport protein ZnT-1 and Cot1p
YDL228c		similarity to *A. klebsiana* glutamate dehydrogenase
YLL005c		similarity to *A. thaliana* hyp1 protein
YMR266w		similarity to *A. thaliana* hyp1 protein
YOL084w		similarity to *A. thaliana* hyp1 protein
YNL317w		similarity to *A. thaliana* PRL1 protein
YOL105c		similarity to a-agglutinin core protein AGA1 and mucin proteins
YPL252c		similarity to adrenodoxin and ferrodoxin
YMR110c		similarity to aldehyde dehydrogenase
YIL112w		similarity to ankyrin and coiled-coil proteins
YLR064w		similarity to *Anopheles* NADH-ubiquinone oxidoreductase, subunit 4
YMR153w		similarity to Asm4p
YDR349c		similarity to aspartyl proteases
YDR072c		similarity to Aur1p
YMR305c		similarity to *B. japonicum* putative β-(-6) glucan transferase
YDR291w		similarity to *B. subtilis* helicases
YPR201w		similarity to *B. subtilis* hypothetical protein
YPR002w		similarity to *B. subtilis* mmgE protein
YPL258c		similarity to *B. subtilis* transcriptional activator tenA, and strong similarity to hypothetical proteins YOL055c and YPR121w
YPR121w		similarity to *B. subtilis* transcriptional activator tenA, strong similarity to hypothetical proteins YPL258c and YOL055c
YIL159w		similarity to BNI1 protein
YDL119c		similarity to bovine Graves disease carrier protein
YNR068c		similarity to Bul1p
YNL218w		similarity to *C. burnetii* trxB, spoIIIE and serS genes
YCR102c		similarity to *C. carbonum* toxD gene
YNL134c		similarity to *C. carbonum* toxD gene
YJR013w		similarity to *C. elegans* B0491.1 protein
YPR040w		similarity to *C. elegans* C02C2.6 protein
YGR257c		similarity to *C. elegans* C16C10.1
YGR054w		similarity to *C. elegans* E04D5.1 protein
YMR012w		similarity to *C. elegans* hypothetical 139.9K protein F55H2.6
YLR022c		similarity to *C. elegans* hypothetical protein
YNL127w		similarity to *C. elegans* hypothetical protein
YKL095w	*YJU2*	similarity to *C. elegans* hypothetical protein
YML014w		similarity to *C. elegans* hypothetical protein C14B1.5
YKL099c		similarity to *C. elegans* hypothetical protein C16C10.2
YEL004w		similarity to *C. elegans* hypothetical protein C53B4.6
YHL030w		similarity to *C. elegans* hypothetical protein D2045.2

YOL093w similarity to *C. elegans* hypothetical protein F25H8.1
YDL008w similarity to *C. elegans* hypothetical protein F35G12.9
YKL151c similarity to *C. elegans* hypothetical protein R107.2
YDR196c similarity to *C. elegans* hypothetical protein T05G5.5
YJL029c similarity to *C. elegans* hypothetical protein T05G5.8
YNL223w similarity to *C. elegans* hypothetical protein ZK792.1
YOL003c similarity to *C. elegans* hypothetical protein, YDR126w, YNL326c and YLR246w
YDR049w similarity to *C. elegans* K06H7.3 protein
YGR278w similarity to *C. elegans* LET-858
YHR004c similarity to *C. elegans* SPAC2F7.02c protein
YML059c similarity to *C. elegans* ZK370.4 protein
YNL168c similarity to *C. elegans* ZK688.3 protein and *E. coli* hpcEp
YPR188c similarity to calmodulin and calmodulin-related proteins
YPR128c similarity to carrier protein FLX1
YPL202c similarity to cell size regulation protein Rcs1p
YPR154w similarity to chicken growth factor receptor-binding protein GRB2 homologue
YNL091w similarity to chicken h-caldesmon, Uso1p and hypothetical protein YKL201c
YPL176c similarity to chinese hamster transferrin receptor protein
YDR259c similarity to Cin5p
YGL241w similarity to Cse1p
YNL305c similarity to C-term. of *A. nidulans* regulatory protein (qutR)
YPR127w similarity to C-term. of *N. tabacum* auxin-induced protein
YDL112w similarity to C-terminus of human TRP-185 protein
YMR304w similarity to *D. melanogaster* fat facets gene
YGR003w similarity to *D. melanogaster* lin19 protein
YLL013c similarity to *D. melanogaster* pumilio protein
YNL023c similarity to *D. melanogaster* shuttle craft protein
YKL201c similarity to *D. melanogaster* sperm-tail-specific protein
YCL001w-a similarity to Dom34p
YIL103w similarity to Dph2 protein
YGL018c similarity to *E. coli* dnaJ homologue
YHL014c *YLF2* similarity to *E. coli* GTP-binding protein
YLR011w similarity to *E. coli* hypothetical 20.4K protein
YGL037c similarity to *E. coli* hypothetical 23K protein
YOR131c similarity to *E. coli* hypothetical 27K protein
YDR400w similarity to *E. coli* hypothetical 33.7K protein in nfo-frua intergenic region
YDR539w similarity to *E. coli* hypothetical 55.3K protein in rfah-rfe intergenic region
YDR332w similarity to *E. coli* hypothetical protein
YGR021w similarity to *E. coli* hypothetical protein
YJL046w similarity to *E. coli* lipoate-protein ligase A
YIL003w similarity to *E. coli* MRP protein
YOR165w similarity to *E. histolytica* surface lectin
YNR015w *SMM1* similarity to *E. histolytica* nitrogen fixation regulatory protein-3 homologue
YER113c similarity to Emp70p
YDR036c similarity to enoyl-CoA hydratase
YOR172w similarity to finger protein YKL222c, YOR162c and YLR266c, weak similarity to transcription factors
YOR162c similarity to finger proteins YKL222c, YOR162c and YLR266c, weak similarity to transcription factors
YOR338w similarity to FUN19 protein
YLR193c similarity to *G. gallus* px19
YKL054c similarity to glutenin, high molecular weight subunit
YOL132w similarity to glycophospholipid-anchored surface glycoprotein GAS1
YHR168w similarity to GTP-binding proteins
YDL033c similarity to *H. influenzae* hypothetical protein HI0174
YLL027w similarity to *H. influenzae* hypothetical protein HI0376
YLR419w similarity to helicases
YMR128w similarity to helicases
YGR271w similarity to Hfm1p
YDR080w similarity to human (clone S53) mRNA, 3' end of cds
YLR051c similarity to human acidic 82K protein
YCR026c similarity to human autotaxin
YJL059w *YHC3* similarity to human Batten disease-related protein CLN3
YHR099w similarity to human DNA-PK
YOL114c similarity to human DS-1 protein
YJR036c similarity to human E6-associated protein
YBR265w similarity to human FVT1 protein
YPL120w similarity to human GT197 (partial ORF)
YPR023c similarity to human hypothetical protein
YGL141w similarity to human hypothetical protein 1
YPL217c similarity to human hypothetical protein KIAA0187
YJR125c similarity to human KIAA0171 protein
YLR050c similarity to human MAC30 C-terminus
YCR071c similarity to human NOF1 protein
YOR001w similarity to human nucleolar 100K polymyositis-scleroderma protein
YEL016c similarity to human nucleotide pyrophosphatase
YIL153w similarity to human phosphotyrosyl phosphatase activator
YDR280w similarity to human PM-Scl-75 autoantigen
YPR108w similarity to human protein GPS1
YLR206w similarity to human protein KIAA0171 and hypothetical protein YDL161w
YNL265c similarity to human protein KIAA0174
YGR163w similarity to human ragA protein
YDR096w *GIS1* similarity to human retinoblastoma binding protein 2
YDR378c similarity to human Sm protein F
YGR112w similarity to human SURF-1 protein
YGR080w similarity to human tyrosine kinase A6
YDL060w similarity to hypothetical *C. elegans* protein
YDR183w similarity to hypothetical *C. elegans* protein
YDR427w similarity to hypothetical *C. elegans* protein
YDR473c similarity to hypothetical *C. elegans* protein
YER051w similarity to hypothetical *C. elegans* protein
YGR145w similarity to hypothetical *C. elegans* protein
YGR194c similarity to hypothetical *C. elegans* protein
YGR245c similarity to hypothetical *C. elegans* protein
YHL010c similarity to hypothetical *C. elegans* protein
YJR070c similarity to hypothetical *C. elegans* protein
YML094w similarity to hypothetical *C. elegans* protein
YMR002w similarity to hypothetical *C. elegans* protein
YJR072c similarity to hypothetical *C. elegans* protein and YLR243w
YOL060c similarity to hypothetical *C. elegans* protein M02F4.4
YDL097c similarity to hypothetical *C. elegans* proteins
YER004w similarity to hypothetical *E. coli* and *C. elegans* proteins
YDL238c similarity to hypothetical *E. coli* protein
YDR013w similarity to hypothetical human KIAA0186 protein
YDR496c similarity to hypothetical human protein
YDR524c similarity to hypothetical human protein
YHR098c similarity to hypothetical human protein
YOR294w similarity to hypothetical human protein
YOR215c similarity to hypothetical *M. xanthus* protein
YDR105c similarity to hypothetical mouse protein
YDL193w similarity to hypothetical *N. crassa* 32K protein
YOR324c similarity to hypothetical protein YAL028w
YOR371c similarity to hypothetical protein YAL056w
YJR115w similarity to hypothetical protein YBL043w
YER093c-a similarity to hypothetical protein YBL059w
YPR030w similarity to hypothetical protein YBL101c
YMR101c similarity to hypothetical protein YBR002c
YGR066c similarity to hypothetical protein YBR105c
YJL171c similarity to hypothetical protein YBR162c
YER145c similarity to hypothetical protein YBR207w
YLR387c similarity to hypothetical protein YBR267w
YJL048c similarity to hypothetical protein YBR273c
YFR044c similarity to hypothetical protein YBR281c
YDL177c similarity to hypothetical protein YCR059c
YDR282c similarity to hypothetical protein YDL001w, YFR048w and *S. pombe* hypothetical protein SPAC12G12.14
YDR425w similarity to hypothetical protein YDL113c
YLR426w similarity to hypothetical protein YDL114w
YJL151c similarity to hypothetical protein YDL123w
YDR233c similarity to hypothetical protein YDL204w
YER139c similarity to hypothetical protein YDR066c
YNL249c *MPA43* similarity to hypothetical protein YDR109c
YLR099c similarity to hypothetical protein YDR125c
YJL149w similarity to hypothetical protein YDR131c
YLR238w similarity to hypothetical protein YDR200c
YDL204w similarity to hypothetical protein YDR233c
YHR061c similarity to hypothetical protein YDR309c
YHR080c similarity to hypothetical protein YDR326c, YFL042c and YLR072w
YOR012w similarity to hypothetical protein YDR391c
YDL113c similarity to hypothetical protein YDR425w
YML018c similarity to hypothetical protein YDR438w
YML033w similarity to hypothetical protein YDR458c
YML034w similarity to hypothetical protein YDR458c
YOR019w similarity to hypothetical protein YDR474c
YLR183c similarity to hypothetical protein YDR501w
YCL036w similarity to hypothetical protein YDR514c
YNR027w similarity to hypothetical protein YEL029c
YGL197w similarity to hypothetical protein YER132c
YDR066c similarity to hypothetical protein YER139c
YBR207w similarity to hypothetical protein YER145c
YHR209w similarity to hypothetical protein YER175c
YGL262w similarity to hypothetical protein YER187w and killer toxin KHS
YPL100w similarity to hypothetical protein YFR021w
YGL228w similarity to hypothetical protein YFR039c
YBR281c similarity to hypothetical protein YFR044c
YDL001w similarity to hypothetical protein YFR048w, YDR282c and *S. pombe* hypothetical protein SPAC12G12.14
YPL216w similarity to hypothetical protein YGL133w
YPL219w similarity to hypothetical protein YGL134w
YDR444w similarity to hypothetical protein YGL144c and YDL1009c
YLR047c similarity to hypothetical protein YGL160w
YER132c similarity to hypothetical protein YGL197w
YFR039c similarity to hypothetical protein YGL228w
YHR036w similarity to hypothetical protein YGL247w
YGR031w similarity to hypothetical protein YGR015c
YGR015c similarity to hypothetical protein YGR031w
YBR105c similarity to hypothetical protein YGR066c
YLR373c similarity to hypothetical protein YGR071c
YPR158w similarity to hypothetical protein YGR142w
YHR149c similarity to hypothetical protein YGR221c
YHR158c similarity to hypothetical protein YGR238c
YMR295c similarity to hypothetical protein YGR273c
YMR310c similarity to hypothetical protein YGR283c
YDR338c similarity to hypothetical protein YHR032w
YGL247w similarity to hypothetical protein YHR036w
YDR309c similarity to hypothetical protein YHR061c
YNL075w similarity to hypothetical protein YHR088w
YDR348c similarity to hypothetical protein YHR097c
YNL144c similarity to hypothetical protein YHR131c
YNL156c similarity to hypothetical protein YHR133c
YGR221c similarity to hypothetical protein YHR149c
YOR147w similarity to hypothetical protein YHR194w
YER175c similarity to hypothetical protein YHR209w
YJR108w similarity to hypothetical protein YIL015c-a
YER064c similarity to hypothetical protein YIL056w
YLR036c similarity to hypothetical protein YIL089w
YFL034c-b similarity to hypothetical protein YIL106w
YNL058c similarity to hypothetical protein YIL117c
YNL074c *YMK1* similarity to hypothetical protein YIL135c
YKR100c similarity to hypothetical protein YIL158w
YKR015c similarity to hypothetical protein YJL043w
YBR273c similarity to hypothetical protein YJL048c
YLL031c similarity to hypothetical protein YJL062w
YKR010c similarity to hypothetical protein YJL076w
YKR019c similarity to hypothetical protein YJL083w
YKR029c similarity to hypothetical protein YJL105p
YDR131c similarity to hypothetical protein YJL149w
YDL123w similarity to hypothetical protein YJL151c
YBR162c similarity to hypothetical protein YJL171c
YJR030c similarity to hypothetical protein YJL181w
YJL181w similarity to hypothetical protein YJR030c
YPR114w similarity to hypothetical protein YJR116w
YKL002w similarity to hypothetical protein YKL041w
YMR031c similarity to hypothetical protein YKL050c and human restin
YMR086w similarity to hypothetical protein YKL105c
YMR171c similarity to hypothetical protein YKL124w
YMR115w similarity to hypothetical protein YKL133c
YNL101w similarity to hypothetical protein YKL146w
YJL076w similarity to hypothetical protein YKR010c
YJL043w similarity to hypothetical protein YKR015c
YJL083w similarity to hypothetical protein YKR019c
YJL105w similarity to hypothetical protein YKR029c
YIL158w similarity to hypothetical protein YKR100c
YLR019w similarity to hypothetical protein YLL010c
YNL328c similarity to hypothetical protein YLR008c
YPL063w similarity to hypothetical protein YLR019w and *S. pombe* hypothetical protein SPAC2F7.02c
YIL089w similarity to hypothetical protein YLR036c
YLL023c similarity to hypothetical protein YLR064w
YFL043c similarity to hypothetical protein YLR072w
YNR067c similarity to hypothetical protein YLR144c
YDR501w similarity to hypothetical protein YLR183c
YNL278w similarity to hypothetical protein YLR187w
YDR200c similarity to hypothetical protein YLR238w
YOR262w similarity to hypothetical protein YLR243w
YDR126w similarity to hypothetical protein YLR246w and YOL003c
YOR137c similarity to hypothetical protein YLR361c
YGR071c similarity to hypothetical protein YLR373c
YPR117w similarity to hypothetical protein YLR454w
YDR458c similarity to hypothetical protein YML034w
YJR054w similarity to hypothetical protein YML047c
YPL184c similarity to hypothetical protein YML117w
YKL050c similarity to hypothetical protein YMR031c
YKL105c similarity to hypothetical protein YMR086w
YKL133c similarity to hypothetical protein YMR115w
YNL008c similarity to hypothetical protein YMR119w
YLR031w similarity to hypothetical protein YMR124w
YPL228w similarity to hypothetical protein YMR180c
YML118w similarity to hypothetical protein YMR285c
YGR273c similarity to hypothetical protein YMR295c
YGR283c similarity to hypothetical protein YMR310c
YMR119w similarity to hypothetical protein YNL008c
YIL117c similarity to hypothetical protein YNL058c
YHR088w similarity to hypothetical protein YNL075w
YHR131c similarity to hypothetical protein YNL144c
YHR133c similarity to hypothetical protein YNL156c
YDL211c similarity to hypothetical protein YNL176c
YLR187w similarity to hypothetical protein YNL278w
YDR214w similarity to hypothetical protein YNL281w
YLR008c similarity to hypothetical protein YNL328c
YEL029c similarity to hypothetical protein YNR027w
YFR012w similarity to hypothetical protein YOL019w
YMR063w similarity to hypothetical protein YOL019w
YOL083w similarity to hypothetical protein YOL082w
YDR474c similarity to hypothetical protein YOR019w

YKR078w		similarity to hypothetical protein YOR069w
YHR194w		similarity to hypothetical protein YOR147w
YAL028w		similarity to hypothetical protein YOR324c
YLR361c		similarity to hypothetical protein YOR3329c
YAL034c	*FUN19*	similarity to hypothetical protein YOR338w
YAL056w		similarity to hypothetical protein YOR371c
YFR021w		similarity to hypothetical protein YPL100w
YML117w		similarity to hypothetical protein YPL184c
YGL133w		similarity to hypothetical protein YPL216w
YGL134w		similarity to hypothetical protein YPL219w
YMR180c		similarity to hypothetical protein YPL228w
YMR181c		similarity to hypothetical protein YPL229w
YJR116w		similarity to hypothetical protein YPR114w
YLR454w		similarity to hypothetical protein YPR117w
YOL027c		similarity to hypothetical protein YPR125w
YGR142w		similarity to hypothetical protein YPR158w
YOL048c		similarity to hypothetical proteins YAL018c and YOL047c
YOL047c		similarity to hypothetical proteins YAL018c and YOL048c
YMR210w		similarity to hypothetical proteins YBR177c and YPL095c
YML052w		similarity to hypothetical proteins YDL222c and YNL194c
YLR246w		similarity to hypothetical proteins YDR126w, YNL326c and YOL003c
YER072w		similarity to hypothetical proteins YFL004w and YPL019c
YLR072w		similarity to hypothetical proteins YFL042c, YFL043c, YDR326c and YHR080c
YGL028c		similarity to hypothetical proteins YGR279c and YMR305c
YFL042c		similarity to hypothetical proteins YHR080c, YDR326c and YLR072w
YNL097c		similarity to hypothetical proteins YHR090c and YHR090c
YJR061w		similarity to hypothetical proteins YKL200c and YKL201c
YMR313c		similarity to hypothetical proteins YKR089c and YOR081c
YOL019w		similarity to hypothetical proteins YMR063w and YFR012w
YPR027c		similarity to hypothetical proteins YNL019c and YNL033w
YNL056w		similarity to hypothetical proteins YNL032w and YNL099c
YNL099c		similarity to hypothetical proteins YNL032w, YNL056w and YDR067c
YNL032w		similarity to hypothetical proteins YNL099c, YNL056w and YDR067c
YOL101c		similarity to hypothetical proteins YOL002c and YDR492w
YNL326c		similarity to hypothetical proteins YOL003c, YLR246w and *C. elegans* hypothetical protein ZK757.1
YML072c		similarity to hypothetical proteins YOR3141c and YNL087w and weak similarity to synaptogamines
YLR152c		similarity to hypothetical proteins YOR3165w and YNL095c
YNL165w		similarity to hypothetical proteins YOR385w and YMR316w
YMR316w		similarity to hypothetical proteins YOR385w and YNL165w
YDR083w		similarity to hypothetical *S. pombe* protein
YDR346c		similarity to hypothetical *S. pombe* protein
YGR272c		similarity to hypothetical *S. pombe* protein
YLR241w		similarity to hypothetical *S. pombe* protein
YML005w		similarity to hypothetical *S. pombe* protein
YOL098c		similarity to hypothetical *S. pombe* protein
YOR091w		similarity to hypothetical *S. pombe* protein D83992_G
YOL071w		similarity to hypothetical *S. pombe* protein SPAC12B10.06c
YFR048w		similarity to hypothetical *S. pombe* protein SPAC12G12.14 and to YDL001w and YDR282c
YOR322c		similarity to hypothetical *S. pombe* protein SPAC1F12.05
YOR250c		similarity to hypothetical *S. pombe* protein SPAC22H10.05c
YER143w		similarity to hypothetical *S. pombe* protein SPAC56F8.08
YDR504c		similarity to hypothetical *T. brucei* protein
YDR223w		similarity to Ifh1p
YAL035w	*FUN12*	similarity to Ifm1p
YOR109w		similarity to inositol polyphosphate 5-phosphatases
YEL013w		similarity to intracellular attachement proteins
YPR042c		similarity to Jsn1p
YMR226c		similarity to ketoreductases
YGR232w		similarity to *L. mactans* α-latroinsectotoxin
YER010c		similarity to *L. pneumophila* dlpA protein
YGR210c		similarity to *M. capricolum* hypothetical protein SGC3
YMR095c		similarity to *M. leprae* hisH protein
YBR079c		similarity to *M. musculus* p162 protein
YNL335w		similarity to *M. verrucaria* cyanamide hydratase, identical to hypothetical protein YFL061w
YLL034c		similarity to mammalian valosin
YMR166c		similarity to members of the mitochondrial carrier protein family
YIR041w		similarity to members of the Srp1p/Tip1p family
YNL006w		similarity to Met30p
YNL293w		similarity to Mic1p and human transforming protein tre-2, and strong similarity to hypothetical protein YOL112w
YNR047w		similarity to microtubule-associated ser/thr protein kinases
YPL137c		similarity to microtubule-interacting protein Mhp1p and to hypothetical protein YOR227w
YIL151c		similarity to mitochondrial aldehyde dehydrogenase Ald1p
YGR058w		similarity to mouse calcium-binding protein
YBL047c		similarity to mouse eps15R protein
YLR200w	*YKE2*	similarity to mouse KE2 protein
YIL023c		similarity to mouse MHC H-2K/t-w5-linked ORF
YCR033w		similarity to mouse nuclear receptor co-repressor N-Cor
YER030w		similarity to mouse nucleolin
YDR060w		similarity to mouse putative CCAAT binding factor 1
YBR187w		similarity to mouse putative transmembrane protein FT27
YGR284c		similarity to mouse Surf-4 protein
YGR127w		similarity to mouse T10 protein
YMR192w		similarity to mouse Tbc1 protein
YPL249c		similarity to mouse Tbc1 protein
YHR181w		similarity to mouse TEG-261 protein
YDR109c		similarity to MPA43p
YDL038c		similarity to mucin proteins
YER085c		similarity to myosins
YJL112w		similarity to *N. crassa* sulphur controller-2
YDR395w		similarity to NMD and CSE1 proteins
YPR072w		similarity to N-terminal part of Cdc39p
YBL024w		similarity to nucleolar Nop2p
YDR071c		similarity to *O. aries* arylalkylamine N-acetyltransferase
YOL164w		similarity to *P. aeruginosa* alkyl sulphatase
YMR099c		similarity to *P. ciliare* possible apospory-associated protein
YNR029c		similarity to *P. denitrificans* cobW protein
YDR472w		similarity to *P. falciparum* 41-2 protein antigen
YNL136w		similarity to *P. falciparum* mature-parasite-infected erythrocyte surface antigen MESA
YIL010w		similarity to *P. falciparum* merozoite cap protein-1
YLR107w		similarity to *P. troglodytes* prot GOR
YOL065c		similarity to Pie3p and hypothetical proteins YIL002c
YNL047c		similarity to probable transcription factor Ask10p and hypothetical protein YPR115w, and strong similarity to hypothetical protein YIL105c
YBR125c		similarity to protein phosphatase 2C
YNR038w		similarity to Prp5p
YER150w		similarity to putative cell surface glycoprotein Sed1p
YGL099w		similarity to putative human GTP-binding protein HSR1
YPR062w		similarity to *R. corallinus* N-ethylammeline chlorohydrolase trzA
YJL055w		similarity to *R. fascians* hypothetical protein 6
YOR191w		similarity to RAD5 protein
YGL059w		similarity to rat branched-chain α-ketoacid dehydrogenase kinase
YBR053c		similarity to rat regucalcin
YDL015c		similarity to rat synaptic glycoprotein SC2
YDR457w		similarity to rat URE-B1
YBR028c		similarity to ribosomal protein kinases
YLR276c		similarity to RNA helicases
YPR112c		similarity to RNA-binding proteins
YPR016c		similarity to *S. acidocaldarius* ao2sac protein
YOL154w		similarity to *S. fumigata* Asp FII
YHR029c		similarity to *S. lincolnensis* lmbX protein
YFL033c		similarity to *S. pombe* cek1 serine/threonine protein kinase
YIL144w		similarity to *S. pombe* hypothetical protein
YKR051w		similarity to *S. pombe* hypothetical protein
YOL087c		similarity to *S. pombe* hypothetical protein
YBR004c		similarity to *S. pombe* hypothetical protein SPAC18B11.05
YKR079c		similarity to *S. pombe* hypothetical protein SPAC1D4.10
YNL308c		similarity to *S. pombe* hypothetical protein SPAC22G7.05
YAL042w	*FUN9*	similarity to *S. pombe* hypothetical protein SPAC24B11.08c
YNL310c		similarity to *S. pombe* hypothetical protein SPAC24H6.02c
YGR125w		similarity to *S. pombe* hypothetical protein SPAC24H6.11c
YMR075w		similarity to *S. pombe* hypothetical protein SPAC2F7.07c
YDR175c		similarity to *S. pombe* hypothetical protein SPAC2F7.15
YFL047w		similarity to *S. pombe* hypothetical protein SPAC2F7.18c
YLR023c		similarity to *S. pombe* hypothetical protein SPAC30D11.11
YDR180w		similarity to *S. pombe* hypothetical protein SPAC31A2.05c
YGL164c		similarity to *S. pombe* hypothetical protein SPAC31A2.10
YPL236c		similarity to *S. pombe* hypothetical protein SPAC3H1.13
YAL032c	*FUN20*	similarity to *S. pombe* hypothetical protein SPAC8A4.06
YEL007w		similarity to *S. pombe* pac2 protein
YLR247c		similarity to *S. pombe* rad8 protein
YJL062w		similarity to *S. pombe* SPAC13G6.3 protein
YHR009c		similarity to *S. pombe* YAM3 protein
YBL051c		similarity to *S. pombe* Z66568_C protein
YOR240w		similarity to *S. pombe* ZK1058.5 protein
YPL150w		similarity to ser/thr kinases
YGR052w		similarity to ser/thr protein kinases
YAL017w	*FUN31*	similarity to ser/thr protein kinases
YOR090c		similarity to ser/thr protein phosphatases
YBR059c		similarity to ser/thr-specific protein kinase Pak1p
YBR274w		similarity to ser/thr-specific protein kinases
YLR118c		similarity to several esterases
YMR077c		similarity to SNF7 protein
YLR313c		similarity to SPA2 protein
YIR033w		similarity to Spt23p
YBL109w		similarity to subtelomeric encoded proteins
YEL074w		similarity to subtelomeric encoded proteins
YGL263w		similarity to subtelomeric encoded proteins
YHL042w		similarity to subtelomeric encoded proteins
YHL043w		similarity to subtelomeric encoded proteins
YHL044w		similarity to subtelomeric encoded proteins
YIL177c		similarity to subtelomeric encoded proteins
YKL219w		similarity to subtelomeric encoded proteins
YMR326c		similarity to subtelomeric encoded proteins
YLR177w		similarity to suppressor protein Gin5p
YNL191w		similarity to *Synechocystis* hypothetical protein
YGR036c		similarity to *T. denticola* phosphatase
YDR485c		similarity to trichohyalin
YOR195w		similarity to USO1 protein
YPL188w	*POS5*	similarity to Utr1p and hypothetical protein YEL041w
YLR213c		similarity to UTR2 protein
YGR189c		similarity to Utr2p
YGL220w		similarity to *V. alginolyticus* bolA protein
YGR203w		similarity to *X. laevis* protein-tyrosin-phosphatase cdc homologue 2 and to hypothetical protein YPR200c
YDR284c		similarity to YDR503c
YEL003w		similarity to Yke2p
YIL135c		similarity to Ymk1p
YOL082w		similarity to YOL083w
YDR534c		similarity to YOR383c,Sta1p and pig mucin
YBL101c		similarity to YPR030w
YCR062w		similarity to Ytp1p protein
YFR024c-a		strong similarity hypothetical protein YHR016c
YPL151c		strong similarity to *A. thaliana* PRL1 and PRL2 proteins
YOR034c		strong similarity to Akr1p
YOR374w		strong similarity to aldehyde dehydrogenase
YIL113w		strong similarity to *C. albicans* dual-specificity phosphatase MSG5
YLR460c		strong similarity to *C. carbonum* toxD protein
YKL013c		strong similarity to *C. elegans* hypothetical protein
YNL288w		strong similarity to *C. elegans* hypothetical protein
YOL077c		strong similarity to *C. elegans* K12H4.3 protein
YGL080w		strong similarity to *C. elegans* R07E5.13 protein
YPR194c		strong similarity to C-term. of *S. pombe* isp4 protein
YAR023c		strong similarity to Fun55p, Fun59p, YGL051w, YCR007c, YGL053w, YAR031w and YAR028w
YAR028w		strong similarity to Fun55p, YGL053w, YCR007c, YAR031w, Fun59p and YGL051w
YGL051w		strong similarity to Fun59p
YGL236c		strong similarity to gidA *E. coli* protein
YAL036c	*FUN11*	strong similarity to GTP-binding proteins
YMR290c		strong similarity to helicases of the DEAD/DEAH box family
YDR276c		strong similarity to *Hordeum* blt101 protein
YDL103c	*QRI1*	strong similarity to human AgX-1 antigen
YDR373w		strong similarity to human BDR-1 protein and other calcium binding proteins
YNR053c		strong similarity to human breast tumor associated autoantigen
YDL120w		strong similarity to human frataxin (Friedreichs ataxia)
YGR173w		strong similarity to human GTP-binding protein
YKL056c		strong similarity to human IgE-dependent histamine-releasing factor

YPL152w strong similarity to human phosphotyrosyl phosphatase activator
YPR028w strong similarity to human protein TB2
YCL059c strong similarity to human Rev interacting protein Rip-1
YLR146c strong similarity to human spermidine synthase
YNL200c strong similarity to human TGR-CL10C
YBL036c strong similarity to hypothetical *C. elegans* protein
YBL078c strong similarity to hypothetical *C. elegans* protein
YMR292w strong similarity to hypothetical *C. elegans* protein
YDR430c strong similarity to hypothetical *C. perfringens* protein
YOR365c strong similarity to hypothetical protein YAL053w
YHR214w-a strong similarity to hypothetical protein YAR068w
YPR032w strong similarity to hypothetical protein YBL106c
YDR003w strong similarity to hypothetical protein YBR005w
YDR210w strong similarity to hypothetical protein YBR016w
YDL012c strong similarity to hypothetical protein YBR016w and YDR210w
YDR018c strong similarity to hypothetical protein YBR042c
YOL092w strong similarity to hypothetical protein YBR147w
YPL095c strong similarity to hypothetical protein YBR177c
YPL087w strong similarity to hypothetical protein YBR183w
YGL056c strong similarity to hypothetical protein YBR214w
YGL060w strong similarity to hypothetical protein YBR216c
YGL107c strong similarity to hypothetical protein YBR238c
YGL101w strong similarity to hypothetical protein YBR242w
YJL058c strong similarity to hypothetical protein YBR270c
YGR293c strong similarity to hypothetical protein YBR300c
YDR514c strong similarity to hypothetical protein YCL036w
YGL144c strong similarity to hypothetical protein YDL109c
YIL079c strong similarity to hypothetical protein YDL175c
YNL194c strong similarity to hypothetical protein YDL222c and similarity to hypothetical protein YML052w
YBR005w strong similarity to hypothetical protein YDR003w
YBR042c strong similarity to hypothetical protein YDR018c
YLR108c strong similarity to hypothetical protein YDR132c
YPL235w strong similarity to hypothetical protein YDR190c
YNL281w strong similarity to hypothetical protein YDR214w
YLR225c strong similarity to hypothetical protein YDR222w
YHR032w *ERC1* strong similarity to hypothetical protein YDR338c
YHR097c strong similarity to hypothetical protein YDR348c
YHR108w strong similarity to hypothetical protein YDR358w
YOR013w strong similarity to hypothetical protein YDR391c
YJR133w strong similarity to hypothetical protein YDR399w
YOL002c strong similarity to hypothetical protein YDR492w
YOR383c strong similarity to hypothetical protein YDR534c and similarity to *L. mexicana* secreted acid phosphatase 2
YPR193c strong similarity to hypothetical protein YEL066w
YGL224c strong similarity to hypothetical protein YER037w
YIL057c strong similarity to hypothetical protein YER067w
YPL019c strong similarity to hypothetical protein YFL004w
YMR096w strong similarity to hypothetical protein YFL059w, YNL333w, and Para rubber tree ethylene-responsive protein 1
YHR016c *YSC84* strong similarity to hypothetical protein YFR024c-a
YBR214w strong similarity to hypothetical protein YGL056c
YBR216c strong similarity to hypothetical protein YGL060w
YPL191c strong similarity to hypothetical protein YGL082w
YPL189w strong similarity to hypothetical protein YGL084c
YBR242w strong similarity to hypothetical protein YGL101w
YDL109c strong similarity to hypothetical protein YGL144c
YER037w strong similarity to hypothetical protein YGL224c
YOR387c strong similarity to hypothetical protein YGL258w
YLR324w strong similarity to hypothetical protein YGR004w
YLR328w strong similarity to hypothetical protein YGR010w
YLR350w strong similarity to hypothetical protein YGR038w
YLR357w strong similarity to hypothetical protein YGR056w
YPL004c strong similarity to hypothetical protein YGR086c
YPR157w strong similarity to hypothetical protein YGR141w
YHR162w strong similarity to hypothetical protein YGR243w
YBR300c strong similarity to hypothetical protein YGR293c
YHR054c strong similarity to hypothetical protein YHR056c
YDR358w strong similarity to hypothetical protein YHR108w
YNL116w strong similarity to hypothetical protein YHR115c
YGR238c strong similarity to hypothetical protein YHR158c
YGR243w strong similarity to hypothetical protein YHR162w
YHR199c strong similarity to hypothetical protein YHR198c
YHR198c strong similarity to hypothetical protein YHR199c
YAR060c strong similarity to hypothetical protein YHR212c
YAR068w strong similarity to hypothetical protein YHR214w-a
YPR071w strong similarity to hypothetical protein YIL029c
YER067w strong similarity to hypothetical protein YIL057c
YDL175c strong similarity to hypothetical protein YIL079c
YIL014c-a strong similarity to hypothetical protein YIL102c
YNL049c strong similarity to hypothetical protein YIL109c
YOL162w strong similarity to hypothetical protein YIL166c
YJL038c strong similarity to hypothetical protein YJL037w
YJL037w strong similarity to hypothetical protein YJL038c
YBR270c strong similarity to hypothetical protein YJL058c
YKR018c strong similarity to hypothetical protein YJL082w
YKR021w strong similarity to hypothetical protein YJL084c
YKR053c strong similarity to hypothetical protein YJL134w
YML047c strong similarity to hypothetical protein YJR054w
YKL200c strong similarity to hypothetical protein YJR061w
YDR399w strong similarity to hypothetical protein YJR133w
YMR238w strong similarity to hypothetical protein YKL046c
YMR040w strong similarity to hypothetical protein YKL065c
YMR102c strong similarity to hypothetical protein YKL121w
YLR413w strong similarity to hypothetical protein YKL187c
YJL082w strong similarity to hypothetical protein YKR018c
YJL084c strong similarity to hypothetical protein YKR021w
YJL134w strong similarity to hypothetical protein YKR053c
YOR081c strong similarity to hypothetical protein YKR089c
YIR013c strong similarity to hypothetical protein YLR013w
YLL010c strong similarity to hypothetical protein YLR019w
YDR125c strong similarity to hypothetical protein YLR099c
YDR132c strong similarity to hypothetical protein YLR108c
YDR185c strong similarity to hypothetical protein YLR168c
YDL161w strong similarity to hypothetical protein YLR206w and to human KIAA0171 protein
YDR222w strong similarity to hypothetical protein YLR225c
YDR213w strong similarity to hypothetical protein YLR228c
YOR171c strong similarity to hypothetical protein YLR260w
YOR173w strong similarity to hypothetical protein YLR270w
YOR180c strong similarity to hypothetical protein YLR284c
YGR004w strong similarity to hypothetical protein YLR324w
YGR010w strong similarity to hypothetical protein YLR328w
YGR038w *ORM1* strong similarity to hypothetical protein YLR350w
YGR056w strong similarity to hypothetical protein YLR357w
YKL187c strong similarity to hypothetical protein YLR413w
YPR172w strong similarity to hypothetical protein YLR456w
YDR438w strong similarity to hypothetical protein YML018c
YBR002c strong similarity to hypothetical protein YMR101c
YKL121w strong similarity to hypothetical protein YMR102c
YPL224c strong similarity to hypothetical protein YMR177w
YOR295w strong similarity to hypothetical protein YMR233w
YKL046c strong similarity to hypothetical protein YMR238w
YPL264c strong similarity to hypothetical protein YMR253c
YGR279c strong similarity to hypothetical protein YMR305c
YOR385w strong similarity to hypothetical protein YMR316w
YNL034w strong similarity to hypothetical protein YNL018c
YNL033w strong similarity to hypothetical protein YNL019c
YNL019c strong similarity to hypothetical protein YNL033w
YNL018c strong similarity to hypothetical protein YNL034w
YIL109c strong similarity to hypothetical protein YNL049c
YOR086c strong similarity to hypothetical protein YNL087w and weak similarity to synaptogamines
YOR092w strong similarity to hypothetical protein YNL095c
YOR110w strong similarity to hypothetical protein YNL108c
YHR115c strong similarity to hypothetical protein YNL116w
YDL222c strong similarity to hypothetical protein YNL194c and similarity to YML052w
YLR144c strong similarity to hypothetical protein YNR067c
YDR492w strong similarity to hypothetical protein YOL002c
YPR125w strong similarity to hypothetical protein YOL027c
YBR147w strong similarity to hypothetical protein YOL092w
YDR391c strong similarity to hypothetical protein YOR013w
YAL007c strong similarity to hypothetical protein YOR016c, similarity to hamster COP-coated vesicle membrane protein
YKR089c strong similarity to hypothetical protein YOR081c
YNL095c strong similarity to hypothetical protein YOR092w
YNL108c strong similarity to hypothetical protein YOR110w
YLR260w strong similarity to hypothetical protein YOR171c
YLR270w strong similarity to hypothetical protein YOR173w
YLR284c strong similarity to hypothetical protein YOR180c
YLR243w strong similarity to hypothetical protein YOR262w
YMR233w strong similarity to hypothetical protein YOR295w
YGL258w strong similarity to hypothetical protein YOR387c
YPL279c strong similarity to hypothetical protein YOR390w
YGR086c strong similarity to hypothetical protein YPL004c
YFL004w strong similarity to hypothetical protein YPL019c
YBR183w strong similarity to hypothetical protein YPL087w
YBR177c strong similarity to hypothetical protein YPL095c
YOR227w strong similarity to hypothetical protein YPL137c and to microtubule-interacting protein MHP1
YGL084c strong similarity to hypothetical protein YPL189w
YGL082w strong similarity to hypothetical protein YPL191c
YGL139w strong similarity to hypothetical protein YPL221w
YMR177w strong similarity to hypothetical protein YPL224c
YDR190c strong similarity to hypothetical protein YPL235w
YMR253c strong similarity to hypothetical protein YPL264c

YOR390w		strong similarity to hypothetical protein YPL279c
YIL029c		strong similarity to hypothetical protein YPR071w
YDL242w		strong similarity to hypothetical protein YPR079w
YGR141w		strong similarity to hypothetical protein YPR157w
YLR456w		strong similarity to hypothetical protein YPR172w
YEL066w		strong similarity to hypothetical protein YPR193c
YGL053w		strong similarity to hypothetical proteins YAR031, YGL051w, YAR028w,Fun55p and YCR007c
YMR324c		strong similarity to hypothetical proteins YBL108w, YCR103c and YKL223w
YNR048w		strong similarity to hypothetical proteins YCR094w and YNL323w
YNL323w		strong similarity to hypothetical proteins YCR094w and YNR048w
YBR016w		strong similarity to hypothetical proteins YDL012c and YDR210w
YNL334c		strong similarity to hypothetical proteins YFL060c and YMR095c
YPL221w		strong similarity to hypothetical proteins YGL139w and YAL053w
YDR326c		strong similarity to hypothetical proteins YHR080c, YFL042c and YLR072w
YMR251w		strong similarity to hypothetical proteins YKR076w and YGR154c
YGR154c		strong similarity to hypothetical proteins YKR076w and YMR251w
YKR076w		strong similarity to hypothetical proteins YMR251w and YGR154c
YFL060c		strong similarity to hypothetical proteins YNL334c and YMR095c
YCR094w		strong similarity to hypothetical proteins YNR048w and YNL323w
YOR230w		strong similarity to hypothetical proteins YOR229w and YPL139c
YPL139c		strong similarity to hypothetical proteins YOR230w and YOR229w
YOR229w		strong similarity to hypothetical proteins YOR230w and YPL139c
YAL053w		strong similarity to hypothetical proteins YOR365c,YGL139w,YPL221w
YPL280w		strong similarity to hypothetical proteins YOR391c, YMR322c and YDR533c
YMR321c		strong similarity to hypothetical proteins YPL273w and YLL062c
YLL062c		strong similarity to hypothetical proteins YPL273w, weak similarity to *M. leprae* metH2 protein
YOR391c		strong similarity to hypothetical proteins YPL280w, YMR322c and YDR533c
YDR533c		strong similarity to hypothetical proteins YPL280w, YOR391c and YMR322c
YMR322c		strong similarity to hypothetical proteins YPL280w, YOR391c and YDR533c
YHR069c		strong similarity to hypothetical *S. pombe* and human proteins
YMR288w		strong similarity to hypothetical *S. pombe* protein
YER049w		strong similarity to hypothetical *S. pombe* protein YER049W
YNL240c		strong similarity to *K. marxianus* LET1 protein
YBR301w		strong similarity to members of the Srp1p/Tip1p family
YGL261c		strong similarity to members of the Srp1p/Tip1p family
YGR294w		strong similarity to members of the Srp1p/Tip1p family
YIL176c		strong similarity to members of the Srp1p/Tip1p family
YKL224c		strong similarity to members of the Srp1p/Tip1p family
YLL025w		strong similarity to members of the Srp1p/Tip1p family
YLL064c		strong similarity to members of the Srp1p/Tip1p family
YLR037c		strong similarity to members of the Srp1p/Tip1p family
YMR325w		strong similarity to members of the Srp1p/Tip1p family
YOR394w		strong similarity to members of the Srp1p/Tip1p family
YEL047c		strong similarity to Osm1p
YFL059w		strong similarity to Para rubber tree ethylene-responsive protein1
YNL333w		strong similarity to Para rubber tree ethylene-responsive protein 1 and identical to hypothetical protein YFL059w
YNR065c		strong similarity to Pep1p
YNR066c		strong similarity to Pep1p
YER089c	*PTC2*	strong similarity to phosphoprotein phosphatases
YPL141c		strong similarity to protein kinase Kin4p
YOL128c		strong similarity to protein kinase Mck1p
YNL020c		strong similarity to protein kinase PAK1
YJR150c		strong similarity to proteins of the Srp1p/Tip1p family
YDL214c		strong similarity to putative protein kinase NPR1
YEL077c		strong similarity to putative purine nucleotide-binding protein YIL177c
YHL017w		strong similarity to putative transmembrane protein PTM1
YHR017w	*YSC83*	strong similarity to *S. douglasii* YSD83
YDL219w		strong similarity to *S. equisimilis* hypothetical protein
YPL118w		strong similarity to *S. kluyveri* hypothetical protein
YNR046w		strong similarity to *S. pombe* hypothetical protein SPAC31A2.02
YNL072w		strong similarity to *S. pombe* hypothetical protein SPAC4G9.02
YDR032c		strong similarity to *S. pombe* obr1
YOL010w		strong similarity to *S. pombe* SPAC12G12.06c protein
YOR163w		strong similarity to *S. pombe* SPAC13G6.14 protein
YOR256c		strong similarity to secretory protein SSP134
YCL024w		strong similarity to ser/thr protein kinase GNP1
YOR310c		strong similarity to SIK1 protein
YDR247w		strong similarity to Sks1p
O7535		strong similarity to subtelomeric encoded proteins
YAL068c		strong similarity to subtelomeric encoded proteins
YBL108w		strong similarity to subtelomeric encoded proteins
YBL111c		strong similarity to subtelomeric encoded proteins
YBL112c		strong similarity to subtelomeric encoded proteins
YBL113c		strong similarity to subtelomeric encoded proteins
YBR302c		strong similarity to subtelomeric encoded proteins
YCL073c		strong similarity to subtelomeric encoded proteins
YCR103c		strong similarity to subtelomeric encoded proteins
YDL248w		strong similarity to subtelomeric encoded proteins
YDR542w		strong similarity to subtelomeric encoded proteins
YDR543c		strong similarity to subtelomeric encoded proteins
YDR544c		strong similarity to subtelomeric encoded proteins
YDR545w		strong similarity to subtelomeric encoded proteins
YEL075c		strong similarity to subtelomeric encoded proteins
YEL076c-b		strong similarity to subtelomeric encoded proteins
YER188c-a		strong similarity to subtelomeric encoded proteins
YER189w		strong similarity to subtelomeric encoded proteins
YER190w		strong similarity to subtelomeric encoded proteins
YFL062w		strong similarity to subtelomeric encoded proteins
YFL063w		strong similarity to subtelomeric encoded proteins
YFL064c		strong similarity to subtelomeric encoded proteins
YFL065c		strong similarity to subtelomeric encoded proteins
YFL066c		strong similarity to subtelomeric encoded proteins
YGL260w		strong similarity to subtelomeric encoded proteins
YGR295c		strong similarity to subtelomeric encoded proteins
YGR296w		strong similarity to subtelomeric encoded proteins
YHL045w		strong similarity to subtelomeric encoded proteins
YHL048w		strong similarity to subtelomeric encoded proteins
YHR217c		strong similarity to subtelomeric encoded proteins
YHR218w-a		strong similarity to subtelomeric encoded proteins
YIR040c		strong similarity to subtelomeric encoded proteins
YJL225c		strong similarity to subtelomeric encoded proteins
YJR161c		strong similarity to subtelomeric encoded proteins
YJR162c		strong similarity to subtelomeric encoded proteins
YKL223w		strong similarity to subtelomeric encoded proteins
YKL225w		strong similarity to subtelomeric encoded proteins
YKR106w		strong similarity to subtelomeric encoded proteins
YLL066c		strong similarity to subtelomeric encoded proteins
YLL067c		strong similarity to subtelomeric encoded proteins
YLR462w		strong similarity to subtelomeric encoded proteins
YLR463c		strong similarity to subtelomeric encoded proteins
YLR464w		strong similarity to subtelomeric encoded proteins
YLR466w		strong similarity to subtelomeric encoded proteins
YLR467w		strong similarity to subtelomeric encoded proteins
YML132w		strong similarity to subtelomeric encoded proteins
YNL336w		strong similarity to subtelomeric encoded proteins
YNL337w		strong similarity to subtelomeric encoded proteins
YNL338w		strong similarity to subtelomeric encoded proteins
YNL339c		strong similarity to subtelomeric encoded proteins
YNR077c		strong similarity to subtelomeric encoded proteins
YOL158c		strong similarity to subtelomeric encoded proteins
YPL282c		strong similarity to subtelomeric encoded proteins
YPL283c		strong similarity to subtelomeric encoded proteins
YPR202w		strong similarity to subtelomeric encoded proteins
YPR203w		strong similarity to subtelomeric encoded proteins
YPR204w		strong similarity to subtelomeric encoded proteins
YNR075w	*EDL1*	strong similarity to subtelomeric encoded proteins
YER042w		strong similarity to transcription factors and peptide methionine sulphoxide reductases
YJL186w		strong similarity to Ttp1p
YEL041w		strong similarity to Utr1p
YCR063w		strong similarity to *Xenopus* G10 and human edg-2 protein
YAR027w	*FUN55*	strong similarity to YAR028w, YCR007c, YGL053w, YAR031w, FUN59p and YGL051w
YAR029w	*FUN57*	strong similarity to YAR031w, YGL053w, Fun55p, Fun59p and YGL051w
YAR033w	*FUN59*	strong similarity to YGL051w, YGL053w, YAR031w, Fun55p, YAR028w and YCR007c
YAR031w		strong similarity to YGL053w, Fun59p, YGL051w, Fun55p, YAR028w and YCR007c
YIL102c		strong similarity to YIL014c-a
YBR025c		strong similarity to Ylf1p
YKL065c		strong similarity to YMR040w
YOL112w		strong similarity to YNL293w, similarity to Mic1p and human transforming protein tre-2
YBL106c		strong similarity to YPR032w
YLR045c	*STU2*	suppressor of a cs tubulin mutation
YER120w	*SCS2*	suppressor of an inositol auxotrophic and a choline sensitive mutant
YCR044c		suppressor of *cdc1-1* ts growth defect
YGL083w	*SCY1*	suppressor of GTPase mutant
YDR510w	*SMT3*	suppressor of *mif2* temperature-sensitive mutation
YKL124w	*SSH4*	suppressor of *shr3*
YLR197w	*SIK1*	suppressor of toxicity of Gal4-IKB
YOL102c	*TPT1*	tRNA 2'-phosphotransferase
YIL080w		Ty3-2 orf C fragment
YBR015c	*TTP1*	type II membrane protein
YBR066c		weak similarity to *A. niger* carbon catabolite repressor protein
YKL071w		weak similarity to *A. parasiticus* nor-1 protein
YJL126w		weak similarity to *A. thaliana* nitrilase 3
YCR079w		weak similarity to *A. thaliana* protein phosphatase 2C
YHR143w		weak similarity to a-agglutinin core protein AGA1
YOR353c		weak similarity to adenylate cyclases
YER158c		weak similarity to Afr1p
YBR074w		weak similarity to aminopeptidase Y
YCR051w		weak similarity to ankyrins
YNR039c		weak similarity to *Anopheles* mitochondrial NADH dehydrogenase subunit 2
YMR152w		weak similarity to AST1 and AST2 protein
YJL109c		weak similarity to ATPase Drs2p
YIR002c		weak similarity to ATP-dependent RNA helicases
YMR211w		weak similarity to β tubulins
YEL040w	*UTR2*	weak similarity to *B. subtilis* 1,3-1,4–glucanase
YNL203c		weak similarity to *B. subtilis* CDPdiacylglycerol-serine O-phosphatidyltransferase
YDR336w		weak similarity to *B. subtilis* hypothetical protein X
YOR111w		weak similarity to *B. subtilis* maf protein
YNR074c		weak similarity to *B. subtilis* nitrite reductase (nirB)
YNR004w		weak similarity to bovine interferon γ
YNL022c		weak similarity to *C. burnetii* FMU protein
YGL246c		weak similarity to *C. elegans* dom-3 protein
YKR071c		weak similarity to *C. elegans* hypothetical protein
YNL207w		weak similarity to *C. elegans* hypothetical protein ZK632.3
YLR242c		weak similarity to *C. elegans* R05H5.5 protein and *T. borreli* apocytochrome b

YKL037w — weak similarity to *C. elegans* ubc-2 protein
YHR012w *PEP11* — weak similarity to *C. elegans* Z47357_A ZK1128.1
YJR024c — weak similarity to *C. elegans* Z49131_E ZC373.5 protein
YBR086c — weak similarity to calcium and sodium channel proteins
YGR225w — weak similarity to Cdc20p
YNR051c — weak similarity to chicken nucleolin
YMR272c — weak similarity to cytochrome b5
YEL045c — weak similarity to cytochrome c oxidase III of *T. brucei* kinetoplast
YHR142w — weak similarity to cytochrome c oxidases
YDL035c — weak similarity to *D. discoideum* protein tyrosine phosphatase
YGL242c — weak similarity to *D. melanogaster* ANK protein
YPR200c — weak similarity to *D. melanogaster* cdc25 protein, and similarity to hypothetical protein YGR203w
YHL041w — weak similarity to *D. melanogaster* hypothetical protein 6
YOR022c — weak similarity to *D. melanogaster* probable Ca^{2+} transporter rdgB
YPL229w — weak similarity to *D. melanogaster* transcription factor shn
YGL185c — weak similarity to dehydrogenases
YLR222c — weak similarity to Dip2p
YER041w — weak similarity to DNA repair protein Rad2p
YOR032c — weak similarity to DNA-binding proteins
YJL162c — weak similarity to dnaJ proteins
YBR220c — weak similarity to *E. coli* ampG protein
YNL217w — weak similarity to *E. coli* bis(5'-nucleosyl)-tetraphosphatase
YER126c — weak similarity to *E. coli* colicin N
YGL136c — weak similarity to *E. coli* ftsJ protein
YKL094w *YJU3* — weak similarity to *E. coli* hypothetical protein
YMR155w — weak similarity to *E. coli* hypothetical protein f402
YER152c — weak similarity to *E. coli* hypothetical protein f470
YJR019c — weak similarity to *E. coli* thioesterase II
YPR067w — weak similarity to *F. alni* nitrogen fixation protein
YDR063w — weak similarity to glia maturation factor β
YCL028w — weak similarity to glutenins, high molecular weight subunit
YPL206c — weak similarity to glycerophosphoryl diester phosphodiesterases
YOR272w — weak similarity to GTP-binding protein β subunits
YAL048c — weak similarity to GTP-binding proteins
YBR175w — weak similarity to GTP-binding proteins
YPL093w — weak similarity to GTP-binding proteins
YGL067w — weak similarity to *H. influenzae* hypothetical protein
YLR165c — weak similarity to *H. influenzae* hypothetical protein HI0176
YLR239c — weak similarity to *H. influenzae* lipoate biosynthesis protein B
YNR062c — weak similarity to *H. influenzae* L-lactate permease (lctP) homologue
YNL176c — weak similarity to Hkr1p
YBL032w — weak similarity to hnRNP complex protein homologue YBR233w
YER033c — weak similarity to human BRCA2 early onset breast cancer gene
YIL036w — weak similarity to human cAMP responce element-binding protein
YJL148w — weak similarity to human chromatin assembly factor I
YDR030c — weak similarity to human CSA protein
YGL243w — weak similarity to human double-stranded RNA adenosine deaminase
YJL091c — weak similarity to human G protein-coupled receptor
YER063w — weak similarity to human heterogeneous ribonuclear particle protein U
YJR002w — weak similarity to human kinesin-related protein CENP-E
YMR029c — weak similarity to human nuclear autoantigen
YCL045c — weak similarity to human ORF
YJL057c — weak similarity to human P1/eIF-2a protein kinase
YNR008w — weak similarity to human phosphatidylcholine-sterol O-acyltransferase
YJL132w — weak similarity to human phospholipase D
YLL037w — weak similarity to human platelet-activating factor receptor
YHR090c — weak similarity to human retinoblastoma binding protein 2
YOR064c — weak similarity to human retinoblastoma binding protein 2
YMR131c — weak similarity to human retinoblastoma-binding protein
YLR272c — weak similarity to hypothetical human ORF
YDL111c — weak similarity to hypothetical human protein
YCR059c — weak similarity to hypothetical protein YDL177c
YNR061c — weak similarity to hypothetical protein YDL218w
YOR004w — weak similarity to hypothetical protein YDR339c
YMR010w — weak similarity to hypothetical protein YDR352w
YGR223c — weak similarity to hypothetical protein YFR021w
YGR095c — weak similarity to hypothetical protein YGR195w
YHR160c — weak similarity to hypothetical protein YGR239c
YGR239c — weak similarity to hypothetical protein YHR160c
YMR299c — weak similarity to hypothetical protein YJL062w
YKL041w — weak similarity to hypothetical protein YKL002w
YMR124w — weak similarity to hypothetical protein YLR031w
YPL109c — weak similarity to hypothetical protein YLR253w
YBR168w — weak similarity to hypothetical protein YLR324w
YNR014w — weak similarity to hypothetical protein YMR206w
YER186c — weak similarity to hypothetical protein YMR316w
YJR129c — weak similarity to hypothetical protein YNL024c
YJL016w — weak similarity to hypothetical protein YNL278w and YLR187w
YMR206w — weak similarity to hypothetical protein YNR014w
YDL218w — weak similarity to hypothetical protein YNR061c
YDR339c — weak similarity to hypothetical protein YOR004w
YLR253w — weak similarity to hypothetical protein YPL109c
YPR151c — weak similarity to hypothetical protein YPL159c
YPL159c — weak similarity to hypothetical protein YPR151c
YGR126w — weak similarity to hypothetical protein YPR156c
YLR391w — weak similarity to hypothetical proteins YAR068w and YHR214w-a
YNL024c — weak similarity to hypothetical proteins YBR271w and YJR129c
YAL018c — weak similarity to hypothetical proteins YOL047c and YOL048c
YDR352w — weak similarity to hypothetical proteins YOL092w, YBR147w and YMR010w
YDR517w — weak similarity to hypothetical *S. pombe* protein
YCR013c — weak similarity to *M. leprae* B1496_F1_41 protein
YPL273w — weak similarity to *M. leprae* metH2 protein, and strong similarity to hypothetical protein YLL062c
YEL043w — weak similarity to Mad1p
YBR186w — weak similarity to members of CDC48/PAS1/SEC18 family of ATPases
YFL046w — weak similarity to middle part of *C. elegans* myosin heavy chain A
YOR350c *MNE1* — weak similarity to mitochondrial *L. illustris* cytochrome oxidase I
YMR111c — weak similarity to MSN1 protein
YMR172w — weak similarity to MSN1 protein
YDL223c — weak similarity to mucin
YJL036w — weak similarity to Mvp1p
YNL063w — weak similarity to *Mycoplasma* protoporphyrinogen oxidase
YLR309c *IMH1* — weak similarity to myosin heavy chains
YBR156c — weak similarity to myosins
YAL022c *FUN26* — weak similarity to Na^+/H^+ antiporter
YDR413c — weak similarity to NADH dehydrogenase
YKR030w — weak similarity to NADH dehydrogenases
YPR174c — weak similarity to Nbp1p
YKR075c — weak similarity to negative regulator Srn1p/Hex2p
YCR047c — weak similarity to N-methyltransferases
YJL131c — weak similarity to non-epidermal *Xenopus* keratin, type I
YHL024w — weak similarity to nuclear protein NOP4
YLR126c — weak similarity to *P. aeruginosa* anthranilate synthase component II
YHL021c — weak similarity to *P. aeruginosa* γ-butyrobetaine hydroxylase
YMR009w — weak similarity to *P. aeruginosa* regulatory protein mmsR
YGR276c — weak similarity to *P. troglodytes* GOR protein
YER059w — weak similarity to Pho80p
YOR281c — weak similarity to phosducins
YMR221c — weak similarity to photosystem II protein D2
YBR094w — weak similarity to pig tubulin-tyrosine ligase
YOR287c — weak similarity to PITSLRE protein kinase isoforms
YBR151w — weak similarity to potato sucrose cleavage protein
YML050w — weak similarity to potato sucrose cleavage protein
YGR262c — weak similarity to protein kinases
YPR106w — weak similarity to protein kinases
YNR022c — weak similarity to protein phosphatases
YBR276c — weak similarity to protein-tyrosine-phosphatase
YMR241w — weak similarity to putative carrier protein RIM2
YJR044c — weak similarity to putative transport protein YKR103w
YBR103w — weak similarity to Pwp2p
YMR093w — weak similarity to Pwp2p
YNR064c — weak similarity to *R. capsulatus* bchO protein
YEL018w — weak similarity to Rad50p
YGL244w — weak similarity to Rad50p
YHR022c — weak similarity to RAS-related protein
YLR351c — weak similarity to rat β-alanine synthase
YCR082w — weak similarity to Rbk1p
YOR246c — weak similarity to reductases
YJL012c — weak similarity to regulatory protein PHO81
YDL114w — weak similarity to *Rhizobium* nodulation protein nodG
YGR068c — weak similarity to Rod1p
YKL107w — weak similarity to *S. antibioticus* probable oxidoreductase
YER093c — weak similarity to *S. epidermidis* PepB protein
YMR065w — weak similarity to *S. pombe* hypothetical protein SPAC13C5.03
YGR212w — weak similarity to *S. pombe* hypothetical protein SPAC18B11.03c
YER127w — weak similarity to *S. pombe* hypothetical protein SPAC18B11.06
YBR271w — weak similarity to *S. pombe* uvi22 protein and hypothetical protein YNL024c
YLR311c — weak similarity to *S. tarentolae* cryptogene protein G4
YLR380w — weak similarity to Sec14p
YNL231c — weak similarity to Sec14p
YJL123c — weak similarity to Sec7p
YKL055c — weak similarity to short-chain alcohol dehydrogenases
YNR030w — weak similarity to Smp3p
YDR486c — weak similarity to Snf7p
YEL015w — weak similarity to Spa2p
YHR177w — weak similarity to Spombe pac2 protein
YBR155w — weak similarity to stress-induced Sti1p
YJR101w — weak similarity to superoxide dismutases
YBR099c — weak similarity to *T. brucei* mitochondrion hypothetical protein 6
YJL145w — weak similarity to *T. pacificus* retinal-binding protein
YJL204c — weak similarity to Tor2p
YER045c — weak similarity to transcription factor Sko1p
YDR520c — weak similarity to transcription factors
YHR063c — weak similarity to translational activator *CBS2*
YGL004c — weak similarity to Tup1p
YJR046w — weak similarity to *Xenopus* vimentin 4
YOL031c — weak similarity to *Y. lipolytica* Sls1 protein
YLL056c — weak similarity to *Y. pseudotuberculosis* CDP-3,6-dideoxy-D-glycero-L-glycero-4-hexulose-5-epimerase
YCL063w — weak similarity to yeast translation regulator Gcd6p
YGR110w — weak similarity to YLR099c and YDR125c
YDR090c — weak similarity to Yro2p
YIL044c — weak similarity to zinc-finger protein GCS1
YLR040c — weak similarity to hypothetical protein YIL011w
YNL212w — weak simlarity to *C. cardunculus* cypro4 protein

Genetic and physical maps of *Saccharomyces cerevisiae*

J. Michael Cherry[1], Catherine Ball[1], Shuai Weng[1], Gail Juvik[1], Rita Schmidt[1], Caroline Adler[1], Barbara Dunn[1], Selina Dwight[1], Linda Riles[3], Robert K. Mortimer[2] & David Botstein[1]

[1]*Department of Genetics, Stanford University School of Medicine, Stanford, California 94305-5120, USA*
[2]*Department of Molecular and Cell Biology, University of California, Berkeley, California 94720, USA*
[3]*Department of Genetics, Washington University Medical School, St. Louis, Missouri 63110, USA*

Genetic and physical maps for the 16 chromosomes of *Saccharomyces cerevisiae* are presented. The genetic map is the result of 40 years of genetic analysis. The physical map was produced from the results of an international systematic sequencing effort. The data for the maps are accessible electronically from the *Saccharomyces* Genome Database (SGD: http://genome-www.stanford.edu/Saccharomyces/).

During the past 40 years, 11 compilations of mapping data for the yeast *Saccharomyces cerevisiae* have been made by R. K. Mortimer and colleagues[1–11]. The last such compilation[11] included mapping data to 1991, and contained for the first time the results of physical as well as genetic mapping methods. Here we present the twelfth, and probably the last, such compilation. These final maps are based on the genetic information accumulated over the years[1–11] and, for the physical mapping data, on an entirely new set of data: the complete genomic sequence of *S. cerevisiae*.

The genetic and physical maps were derived from two entirely different types of data. Genetic distances between genes were determined by tetrad analysis. Distances for gene–gene and gene–centromere linkages are expressed in centimorgans (cM) and were calculated using a maximum-likelihood equation[12] which yields values for map distance, an interference parameter, and error calculations for these two parameters. Mapping results on more than 2,600 named genes are presented. Physical distances are calculated directly from the complete DNA sequence. The precise values of all parameters (both tetrad analysis results and chromosomal base-pair coordinates) are available from the SGD.

Associations between open reading frames (ORFs) and corresponding mutations were made using a set of hybridization filters, originally produced by L. Riles and M. Olson[13], which are now available from the American Type Culture Collection (http://www.atcc.org/). Other such associations were made by complementation experiments using cloned DNA fragments and/or sequence analysis of mutants. The data for some of these associations are published, but the documentation for all of them can be found on SGD.

Now that the entire yeast genome sequence is available, most revisions of the map will consist of associations between a biological function and an ORF. These associations will often involve the study of mutants of the gene. In the past, such an association invariably resulted in the naming of the gene; this process is likely to continue until all of the genes have been associated with a function and have thereby acquired a name. Because the genetic and physical maps are unlikely to change significantly, we see no need for any future publications; rather, we expect the electronic version of the maps to evolve into increasingly accurate guides to *S. cerevisiae* biology.

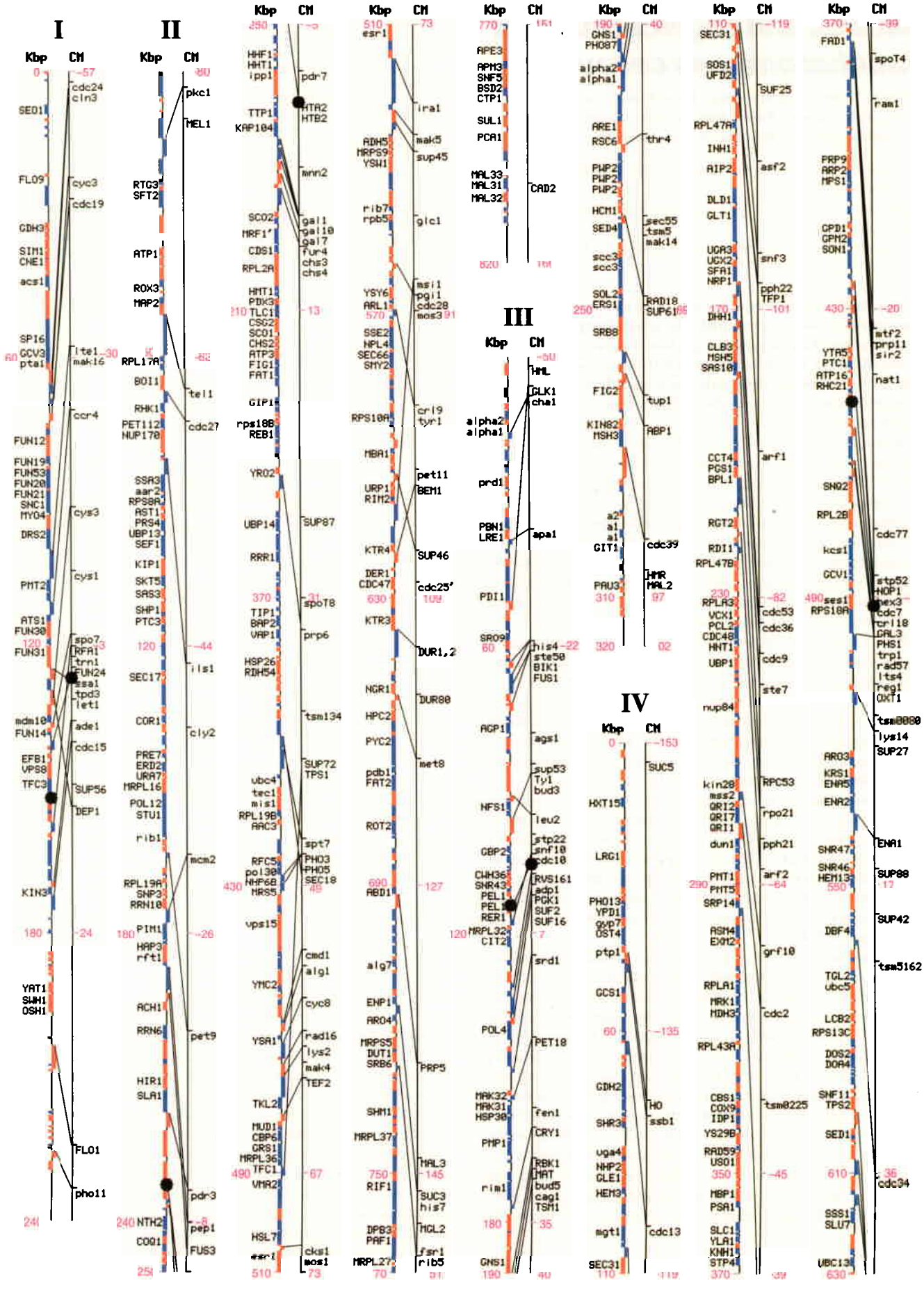
The maps shown here are also available in a continually updated electronic form from the SGD (http://genome-www.stanford.edu/Saccharomyces/), which will also provide directions to other useful information (gene names, aliases, phenotypes, mapping data, protein information, and curated compilations of published literature about genes). □

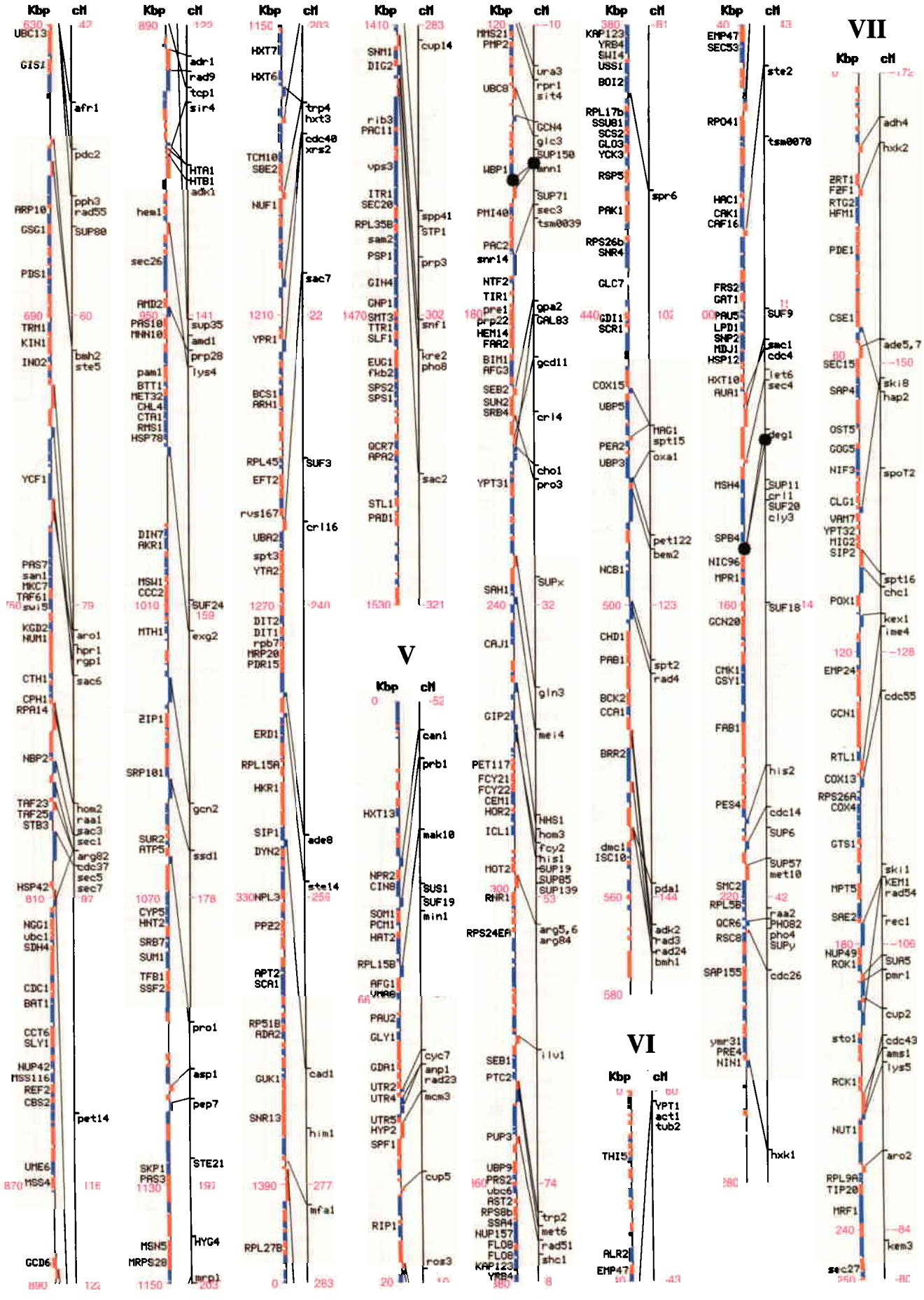
1. Hawthorne, D. C. & Mortimer, R. K. *Genetics* 45, 1085–1110 (1960).
2. Mortimer, R. K. & Hawthorne, D. C. *Genetics* 53, 165–173 (1966).
3. Hawthorne, D. C. & Mortimer, R. K. *Genetics* 60, 735–742 (1968).
4. Mortimer, R. K. & Hawthorne, D. C. *Genetics* 74, 33–54 (1973).
5. Mortimer, R. K. & Hawthorne, D. C. *Methods Cell Biol.* 11, 221–233 (1975).
6. Mortimer, R.K. & Schild, D. *Microbiol. Rev.* 44, 519–571 (1980).
7. Mortimer, R. K. & Schild, D. in *The Molecular Biology of the Yeast* Saccharomyces*: Metabolism and Gene Expression.* (eds J. N. Strathern, E. W. Jones, & J. R. Broach) 639–650 (Cold Spring Harbor Laboratory Press, NY, 1982).
8. Mortimer, R. K. & Schild, D. *Microbiol. Sci.* 1, 145–146 (1984).
9. Mortimer, R. K. & Schild, D. *Microbiol. Rev.* 49, 181–213 (1985).
10. Mortimer, R. K., Schild, D., Contopoulou, C. R. & Kans, J. *Yeast* 5, 321–404 (1989).
11. Mortimer, R. K., Contopoulou, C. R. & King, J. S. *Yeast* 8, 817–902 (1992).
12. King, J. S. & Mortimer, R. K. *Genetics* 129, 597–601(1991).
13. Riles, L. *et al. Genetics* 134, 81–150 (1993).

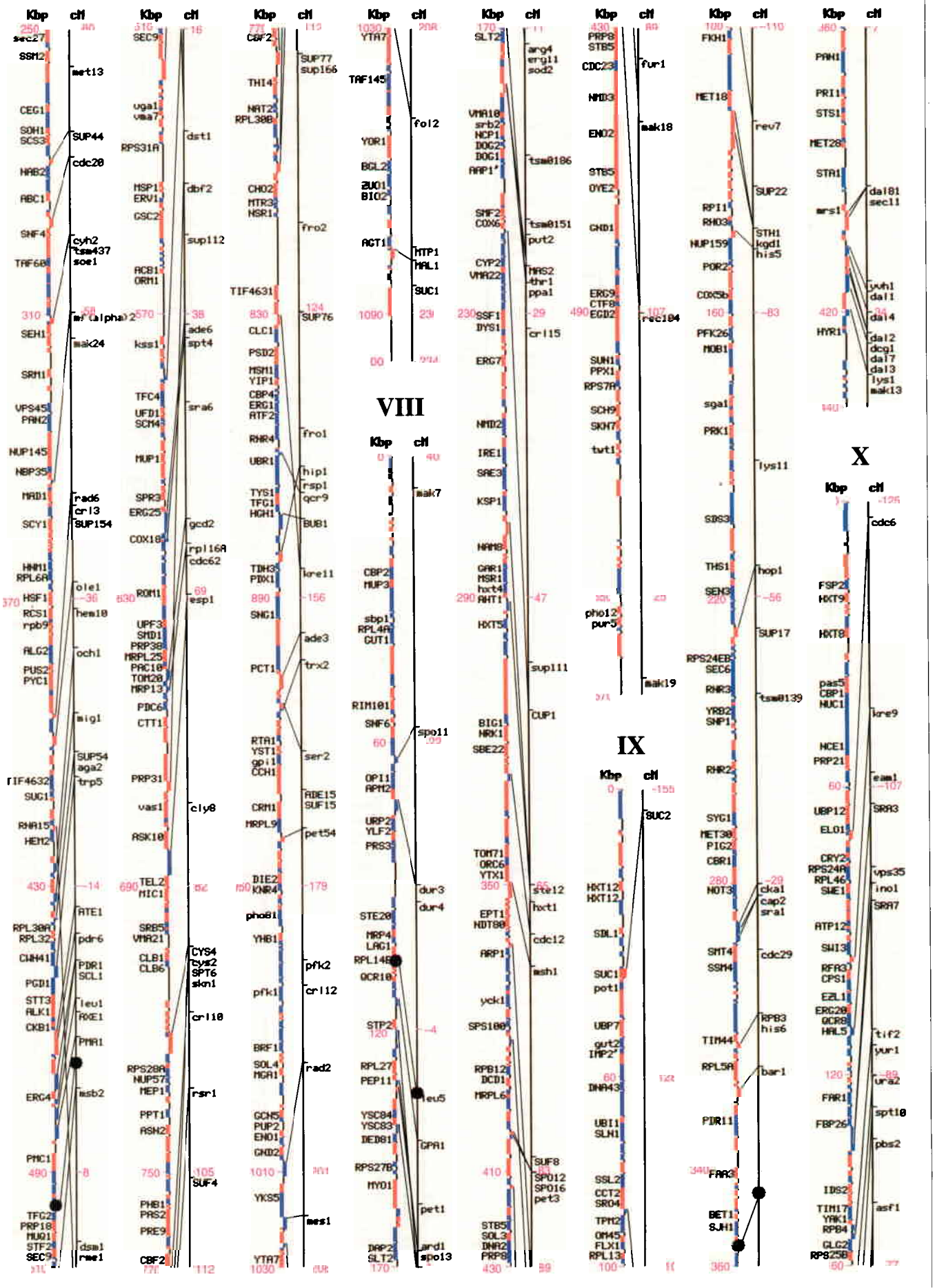
Acknowledgements. We thank all of the yeast researchers, who are too numerous to name individually; F. Dietrich, M. Johnston, E. W. Jones, M. Olson and B. F. F. Ouellette for criticism; A. Goffeau, for encouragement and advice and J. Garrels, D. Lipman, A. Bairoch and H. W. Mewes for their continuing collaboration. S. G. D. is funded by a grant from the National Center for Human Genome Research.

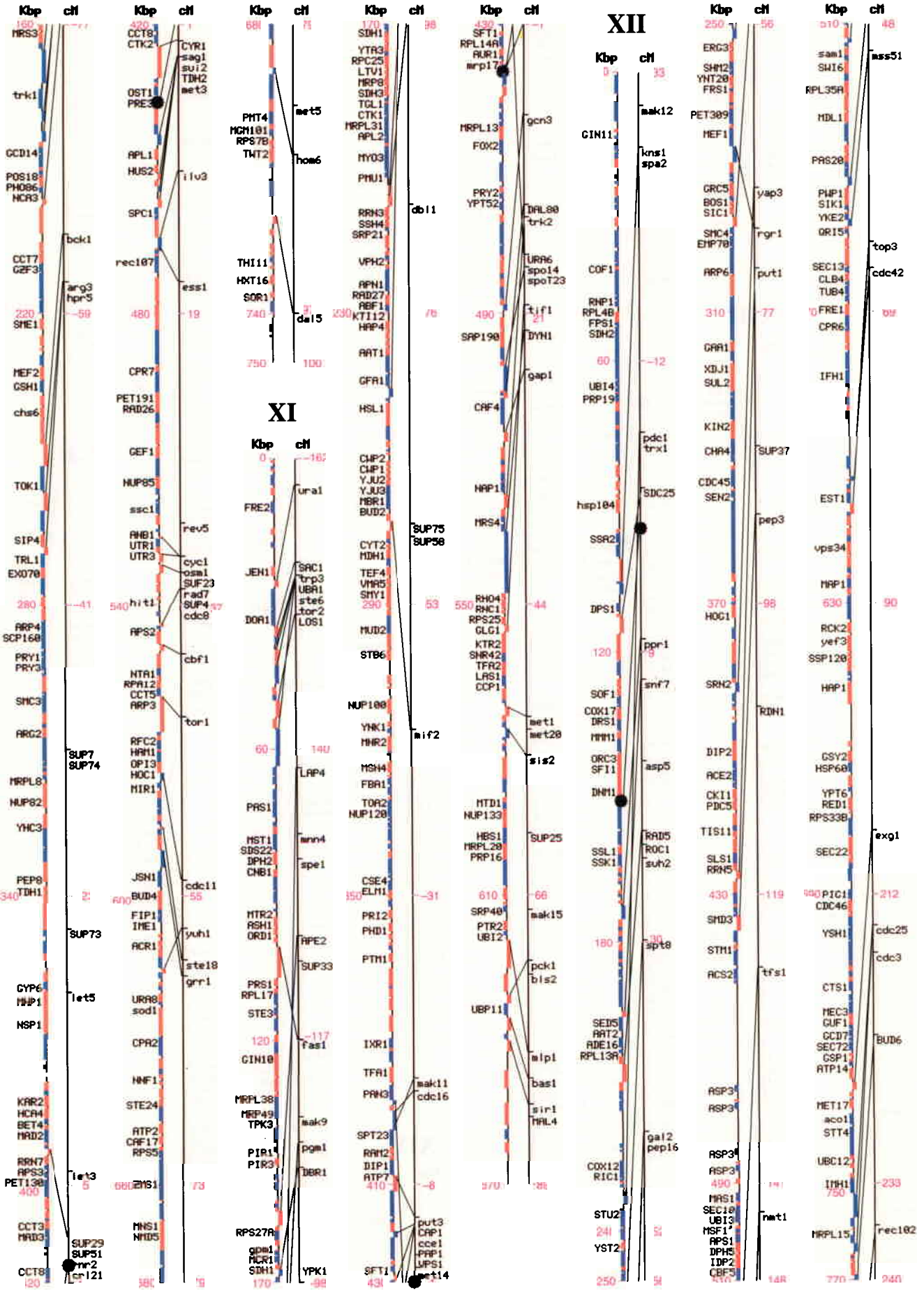
Correspondence should be addressed to J. M. C. (e-mail: cherry@genome.stanford.edu).

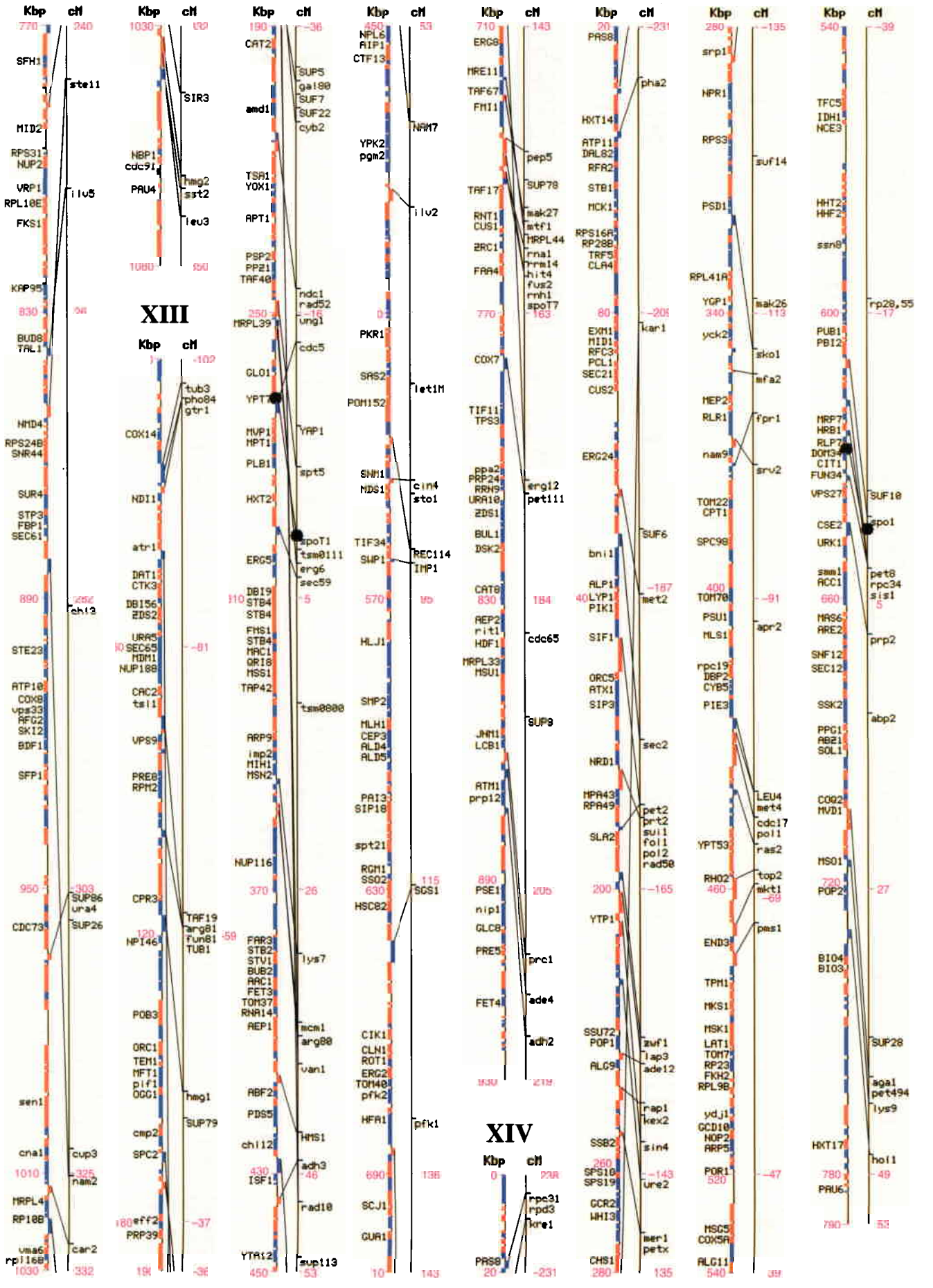
Figures I-XVI (overleaf, pages 68-73): Genetic and physical maps, and their correlations, of the 16 *Saccharomyces cerevisiae* chromosomes. A parallel comparison of the physical map (left, in kilobase pairs) and the genetic map (right, in centimorgans) of each of the 16 chromosomes is illustrated. The information in this figure is available on the *Saccharomyces* Genome Database (http://genome-www.stanford.edu/Saccharomyces/). The physical map consists of coloured boxes that indicate ORFs. ORFs on the Watson strand (left telomere is the 5′ end of this strand) are shown as red boxes, those on the Crick strand as blue boxes. Where it has been defined, the gene name of an ORF is indicated. The genetic map is based on data collected since 1991 by the SGD project, as well as on earlier data[1–11]. Horizontal tick marks on the right of the genetic map line indicate positions of genes. Lines connect genetically mapped genes with their ORF on the physical map. A single name is listed for known synonyms.

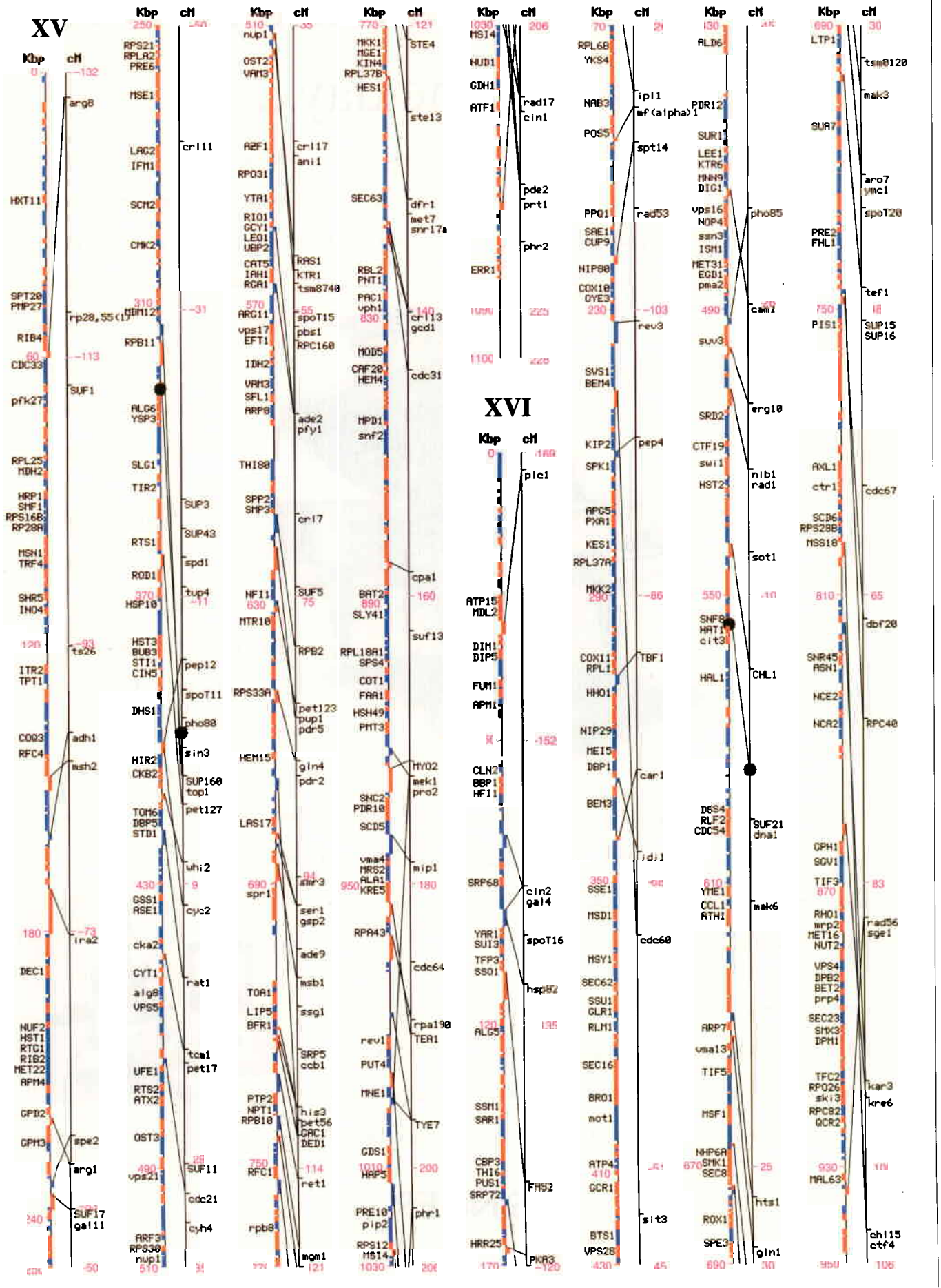

I
II
III
IV
Kbp
cM


Kbp
cM
VII
V
VI
UBC13
GIS1
afr1
pdc2
ARP10
pph3
rad55
GSG1
SUP80
PDS1
TRM1
KIN1
bmh2
ste5
IN02
YCF1
PAS7
san1
MKC7
TAF61
swi5
KGD2
NUM1
aro1
hpr1
rgp1
CTH1
sac6
CPH1
RPA14
NBP2
TAF23
TAF25
STB3
hom2
raa1
sac3
sec1
arg82
cdc37
sec5
sec7
HSP42
NGG1
ubc1
SDH4
CDC1
BAT1
CCT6
SLY1
NUP42
MSS116
REF2
CBS2
pet14
UME6
MSS4
GCD6
adr1
rad9
tcp1
sir4
HTA1
HTB1
adk1
hem1
sec26
AMD2
PAS10
MNN10
sup35
amd1
prp28
lys4
pam1
BTT1
MET32
CHL4
CTR1
RMS1
HSP78
DIN7
AKR1
MSW1
CCC2
SUF24
MTH1
exg2
ZIP1
SRP101
gcn2
SUR2
ATP5
ssd1
CYP5
HNT2
SRB7
SUM1
TFB1
SSF2
pro1
asp1
pep7
SKP1
PRS3
STE21
MSN5
MRPS28
HYG4
mrp1
HXT7
HXT6
trp4
hxt3
cdc48
xrs2
TCM10
SBE2
NUF1
sac7
YPR1
BCS1
RRH1
RPL45
EFT2
SUF3
rvs167
crl16
UBA2
spt3
YTA2
DIT2
DIT1
rpb7
MRP20
PDR15
ERD1
RPL15A
HKR1
SIP1
DYN2
ade8
PP22
ste14
APT2
SCA1
RPS1B
ADA2
GUK1
cad1
SNR13
him1
RPL27B
mfa1
cup14
SNM1
DIG2
rib3
PAC11
vps3
ITR1
SEC20
RPL35B
sam2
PSP1
GIH4
CNP1
SMT3
TTR1
SLF1
spp41
STP1
prp3
snf1
EUG1
fkb2
SPS2
SPS1
kre2
pho8
QCR7
APA2
STL1
PAD1
sac2
HXT13
can1
prb1
mak10
NPR2
CIN8
SUS1
SUF19
min1
SOM1
PCM1
HAT2
RPL15B
AFG1
UMR6
PAU2
GLY1
cyc7
anp1
rad23
GDA1
UTR2
UTR4
mcm3
UTR5
HYP2
SPF1
cup5
RIP1
ros3
MMS21
PMP2
ura3
rpr1
sit4
UBC8
GCN4
glc3
WBP1
SUP150
mnn1
SUP71
PMI40
sec3
tsm0039
PAC2
snr14
NTF2
TIR1
pre1
gpa2
prp22
GAL83
HEM14
FAA2
gcd11
BIM1
AFG3
SEB2
SUN2
SRB4
crl4
cho1
YPT31
pro3
SUPx
SAH1
gln3
CAJ1
GIP2
mei4
PET117
FCY21
FCY22
CEM1
HOR2
ICL1
HHS1
hom3
fcy2
his1
HOT2
SUP19
SUP85
RHR1
SUP139
arg5,6
RPS24EA
arg84
ilv1
SEB1
PTC2
PUP3
UBP9
PRS2
ubc6
RST2
RPS8b
SSA4
NUP157
FLO8
FLO8
KAP123
YRB4
shc1
KAP123
YRB4
SWI4
USS1
BOI2
RPL17b
SSU81
SCS2
GLO3
YCK3
RSP5
PAK1
spr6
RPS26b
SNR4
GLC7
GDI1
SCR1
COX15
UBP5
MAG1
spt15
PEA2
UBP3
oxa1
pet122
bem2
NCB1
spt2
CHD1
rad4
PAB1
BCK2
CCA1
BRR2
dmc1
ISC10
pda1
adk2
rad3
rad24
bmh1
YPT1
act1
tub2
THI5
ALR2
EMP47
EMP47
SEC53
ste2
RPO41
tsm0070
HAC1
CAK1
CAF16
FRS2
GAT1
PAU5
LPD1
SNP2
MDJ1
HSP12
SUF9
smc1
cdc4
HXT10
AUR1
let6
sec4
deg1
MSH4
SUP11
crl1
SUF20
cly3
SPB4
NIC96
MPR1
SUF18
GCN20
CNK1
GSY1
FAB1
his2
PES4
cdc14
SUP6
SUP57
met10
SMC2
RPL5B
QCR6
RSC8
raa2
PH082
pho4
SUPy
SAP155
cdc26
ymr31
PRE4
NIN1
hxk1
ADH4
hxk2
ZRT1
F2F1
RTG2
HFM1
PDE1
CSE1
SEC15
SAP4
OST5
GOG5
NIF3
CLG1
VAM7
YPT32
MIG2
SIP2
POX1
EMP24
GCN1
RTL1
COX13
RPS26A
COX4
GTS1
MPT5
SAE2
NUP49
ROK1
NUT1
RPL9A
TIP20
MRF1
sec27
ade5,7
ski8
hap2
spo12
spt16
chc1
kex1
ime4
cdc55
ski1
KEM1
rad54
rec1
SUA5
pmr1
cup2
sto1
cdc43
ams1
lys5
RCK1
aro2
kem3


VIII
IX
X


XI
XII
Kbp
chI


Kbp
cM
XIII
XIV
SFH1
MID2
RPS31
NUP2
VRP1
RPL10E
FKS1
KAP95
BUD8
TAL1
NMD4
RPS24B
SNR44
SUR4
STP3
FBP1
SEC61
STE23
ATP10
COX8
vps33
RFC2
SKI2
BDF1
SFP1
CDC73
sen1
cna1
MRPL4
RP10B
vma6
rpl16B
ste11
ilv5
chl3
SUP86
ura4
SUP26
cup3
nam2
car2
SIR3
NBP1
cdc9l
PAU4
hmg2
sst2
leu3
COX14
NDI1
atr1
DAT1
CTK3
DBI56
ZDS2
URA5
SEC65
MDM1
NUP188
CAC2
tsl1
VPS9
PRE8
RPN2
CPR3
NPI46
POB3
ORC1
TEM1
MFT1
pif1
OGG1
cmp2
SPC2
eff2
PRP39
tub3
pho84
gtr1
TAF19
arg81
fun81
TUB1
hmg1
SUP79
CAT2
amd1
TSA1
YOX1
APT1
PSP2
PPZ1
TAF40
MRPL39
GLO1
YPT7
MVP1
MPT1
PLB1
HXT2
ERG5
DBI9
STB4
STB4
FMS1
STB4
MAC1
QRI8
MSS1
TAP42
ARP9
imp2
MIH1
MSN2
NUP116
FAR3
STB2
STV1
BUB2
AAC1
FET3
TOM37
RNA14
AEP1
ABF2
PDS5
chl12
ISF1
YTA12
SUP5
gal80
SUF7
SUF22
cyb2
ndc1
rad52
ungl
cdc5
YAP1
spt5
spoT1
tsm0111
erg6
sec59
tsm0800
lys7
mcm1
arg80
van1
HMS1
adh3
rad10
sup113
NPL6
AIP1
CTF13
YPK2
pgm2
PKR1
SAS2
POM152
SNM1
MDS1
TIF34
SWP1
HLJ1
SMP2
MLH1
CEP3
ALD4
ALD5
PAI3
SIP18
spt21
RGM1
SSO2
HSC82
CIK1
CLN1
ROT1
ERG2
TOM40
pfk2
HFA1
SCJ1
GUA1
NAM7
ilv2
let1M
cin4
sto1
REC114
IMP1
SGS1
pfk1
ERG8
MRE11
TAF67
FMI1
TAF17
RNT1
CUS1
ZRC1
FAA4
COX7
TIF11
TPS3
ppa2
PRP24
RRN9
URA10
ZDS1
BUL1
DSK2
CAT8
REP2
rit1
HDF1
MRPL33
MSU1
JNM1
LCB1
ATM1
prp12
PSE1
nip1
GLC8
PRE5
FET4
pep5
SUP78
mak27
mtf1
MRPL44
rna1
rrm14
hit4
fus2
rnh1
spoT7
erg12
pet111
cdc65
SUP9
prc1
ade4
adh2
PAS8
HXT14
ATP11
DAL82
RFR2
STB1
MCK1
RPS16A
RP288
TRF5
CLA4
EXM1
MID1
RFC3
PCL1
SEC21
CUS2
ERG24
bni1
ALP1
LYP1
PIK1
SIF1
ORC5
RTX1
SIP3
NRD1
MPA43
RPA49
SLA2
YTP1
SSU72
POP1
ALG9
SSB2
SPS18
SPS19
GCR2
WHI3
CHS1
pha2
kar1
SUF6
met2
sec2
pet2
prt2
sui1
fol1
pol2
rad50
zwf1
lap3
ade12
rap1
kex2
sin4
ure2
mer1
petx
srp1
NPR1
RPS3
PSD1
RPL41A
YGP1
yck2
MEP2
RLR1
nam9
TOM22
CPT1
SPC98
TOM70
PSU1
MLS1
rpc19
DBP2
CYB5
PIE3
YPT53
RHO2
END3
TPM1
MKS1
MSK1
LAT1
TOM7
RP23
FKH2
RPL9B
ydj1
GCD10
NOP2
ARP5
POR1
MSC5
COX5A
ALG11
suf14
mak26
sko1
mfa2
fpr1
srv2
apr2
LEU4
met4
cdc17
pol1
ras2
top2
mkt1
pms1
TFC5
IDH1
NCE3
HHT2
HHF2
ssn8
PUB1
PBI2
MRP7
HRB1
RLP7
DOM34
CIT1
FUN34
VPS27
CSE2
URK1
smm1
ACC1
MRS6
ARE2
SNF12
SEC12
SSK2
PPG1
ABZ1
SOL1
COQ2
MVD1
MSO1
POP2
BIO4
BIO3
HXT17
PRU6
rp28.55
SUF10
spo1
pet8
rpc34
sis1
prp2
abp2
SUP28
aga1
pet494
lys9
hol1


XV
Kbp
cM
HXT11
SPT20
PMP27
RIB4
CDC33
pfk27
RPL25
MDH2
HRP1
SMF1
RPS16B
RP28A
MSN1
TRF4
SHR5
INO4
ITR2
TPT1
COQ3
RFC4
DEC1
NUF2
HST1
RTC1
RIB2
MET22
APM4
GPD2
GPM3
arg8
rp28,55(1)
SUF1
ts26
adh1
msh2
ira2
spe2
arg1
SUF17
gal11
RPS21
RPLA2
PRE6
MSE1
LAC2
IFM1
SCM2
CMK2
MDM12
RPB11
ALG6
YSP3
SLG1
TIR2
RTS1
ROD1
HSP10
HST3
BUB3
STI1
CIN5
DHS1
HIR2
CKB2
TOM6
DBP5
STD1
GSS1
ASE1
cka2
CYT1
alg8
VPS5
UFE1
RTS2
ATX2
OST3
vps21
ARF3
RPS30
nup1
crl11
SUP3
SUP43
spd1
tup4
pep12
spoT11
pho80
sin3
SUP160
top1
pet127
whi2
cyc2
rat1
tcm1
pet17
SUF11
cdc21
cyh4
nup1
OST2
VAM3
A2F1
RPO31
YTA1
RIO1
CCY1
LEO1
UBP2
CAT5
IAH1
RGA1
ARG11
vps17
EFT1
IDH2
VAM3
SFL1
ARP8
THI80
SPP2
SMP3
NF11
MTR10
RPS33A
HEM15
LAS17
spr1
TOA1
LIP5
BFR1
PTP2
NPT1
RPB10
RFC1
rpb8
crl17
ani1
RAS1
KTR1
tsm8740
spoT15
pbs1
RPC160
ade2
pfy1
crl7
SUF5
RPB2
pet123
pup1
pdr5
gln4
pdr2
smr3
ser1
gsp2
ade9
msb1
ssg1
SRP5
ccb1
his3
pet56
CAC1
DED1
ret1
mgm1
MKK1
MCE1
KIN4
RPL37B
HES1
SEC63
RBL2
PNT1
PAC1
vph1
MOD5
CAF20
HEM4
MPD1
snf2
BAT2
SLY41
RPL18A1
SPS4
COT1
FAR1
HSH49
PMT3
SNC2
PDR10
SCD5
vma4
MRS2
ALA1
KRE5
RPA43
rev1
PUT4
MNE1
GDS1
HAP5
PRE10
pip2
RPS12
MSI4
STE4
ste13
dfr1
met7
snr17a
crl13
gcd1
cdc31
cpa1
suf13
HYO2
mek1
pro2
mip1
cdc64
rpa190
TER1
TYE7
phr1
NUD1
GDH1
ATF1
rad17
cin1
pde2
prt1
phr2
XVI
plc1
ATP15
MDL2
DIM1
DIP5
FUM1
APM1
CLN2
BBP1
HFI1
SRP68
YAR1
SUI3
TFP3
SSO1
ALG5
SSM1
SAR1
CBP3
THI6
PUS1
SRP72
HRR25
cin2
gal4
spoT16
hsp82
FMS2
PKR3
RPL6B
YKS4
NAB3
POS5
PPQ1
SAE1
CUP9
NIP80
COX10
OYE3
SVS1
BEM4
KIP2
SPK1
APC5
PXR1
KES1
RPL37A
MKK2
COX11
RPL1
HHO1
NIP29
MEI5
DBP1
BEM3
SSE1
MSD1
MSY1
SEC62
SSU1
GLR1
RLM1
SEC16
BRO1
mot1
ATP4
GCR1
BTS1
VPS28
ipl1
mf(alpha)1
spt14
rad53
rev3
pep4
TBF1
car1
rdl1
cdc60
sit3
ALD6
PDR12
SUR1
LEE1
KTR6
MNN9
DIG1
vps16
NOP4
ssn3
ISM1
MET31
EGD1
pma2
suv3
SRD2
CTF19
swi1
HST2
SNF8
HAT1
cit3
HAL1
DES4
RLF2
CDC54
YME1
CCL1
ATH1
ARP7
vma13
TIF5
MSF1
NHP6A
SMK1
SEC8
ROX1
SPE3
pho85
cam1
erg10
nib1
rad1
sot1
CHL1
SUF21
dnaJ
mak6
hts1
gln1
LTP1
SUA7
PRE2
FHL1
PIS1
AXL1
ctr1
SCD6
RPS28B
MSS18
SNR45
ASN1
NCE2
NCR2
GPH1
SGV1
TIF3
RHO1
mrp2
MET16
NUT2
VPS4
DPB2
BET2
prp4
SEC23
SMX3
DPM1
TFC2
RPO26
ski3
RPC82
QCR2
MAL63
tsm8120
mak3
aro7
ymc1
spoT20
tef1
SUP15
SUP16
cdc67
dbf20
RPC40
rad56
sge1
kar3
kre6
chl15
ctf4

The nucleotide sequence of *Saccharomyces cerevisiae* chromosome IV

C. Jacq[1], J. Alt-Mörbe[2], B. Andre[3], W. Arnold[4], A. Bahr[5], J. P. G. Ballesta[6], M. Bargues[7], L. Baron[8], A. Becker[4], N. Biteau[8], H. Blöcker[9], C. Blugeon[1], J. Boskovic[6], P. Brandt[9], M. Brückner[10], M. J. Buitrago[11], F. Coster[12], T. Delaveau[1], F. del Rey[11], B. Dujon[13], L. G. Eide[14], J. M. Garcia-Cantalejo[6], A. Goffeau[12], A. Gomez-Peris[15], C. Granotier[8], V. Hanemann[16], T. Hankeln[5], J. D. Hoheisel[17], W. Jäger[9], A. Jimenez[6], J.-L. Jonniaux[12], C. Krämer[5], H. Küster[4], P. Laamanen[18], Y. Legros[8], E. Louis[19], S. Möller-Rieker[5], A. Monnet[8], M. Moro[20], S. Müller-Auer[10], B. Nußbaumer[4], N. Paricio[7], L. Paulin[18], J. Perea[1], M. Perez-Alonso[7], J. E. Perez-Ortin[15], T. M. Pohl[21], H. Prydz[14], B. Purnelle[12], S. W. Rasmussen[22], M. Remacha[6], J. L. Revuelta[11], M. Rieger[10], D. Salom[15], H. P. Saluz[16], J. E. Saiz[11], A.-M. Saren[18], M. Schäfer[10], M. Scharfe[23], E. R. Schmidt[5], C. Schneider[20], P. Scholler[17], S. Schwarz[17], A. Soler-Mira[6], L. A. Urrestarazu[3], P. Verhasselt[24], S. Vissers[3], M. Voet[24], G. Volckaert[24], G. Wagner[10], R. Wambutt[23], E. Wedler[23], H. Wedler[23], S. Wölfl[16], D. E. Harris[25], S. Bowman[25], D. Brown[25], C. M. Churcher[25], R. Connor[25], K. Dedman[25], S. Gentles[25], N. Hamlin[25], S. Hunt[25], L. Jones[25], S. McDonald[25], L. Murphy[25], D. Niblett[25], C. Odell[25], K. Oliver[25], M. A. Rajandream[25], C. Richards[25], L. Shore[25], S. V. Walsh[25], B. G. Barrell[25], F. S. Dietrich[26], J. Mulligan[26], E. Allen[26], R. Araujo[26], E. Aviles[26], A. Berno[26], J. Carpenter[26], E. Chen[26], J. M. Cherry[26], E. Chung[26], M. Duncan[26], S. Hunicke-Smith[26], R. Hyman[26], C. Komp[26], D. Lashkari[26], H. Lew[26], D. Lin[26], D. Mosedale[26], K. Nakahara[26], A. Namath[26], P. Oefner[26], C. Oh[26], F. X. Petel[26], D. Roberts[26], S. Schramm[26], M. Schroeder[26], T. Shogren[26], N. Shroff[26], A. Winant[26], M. Yelton[26], D. Botstein[26], R. W. Davis[26], M. Johnston[27], L. Hillier[27], L. Riles[27] and other members of the Genome Sequencing Center[27], K. Albermann[28], J. Hani[28], K. Heumann[28], K. Kleine[28], H. W. Mewes[28], A. Zollner[28] & P. Zaccaria[28]

[1]*Laboratoire de Génétique Moléculaire, URA 1302 du CNRS, Ecole Normale Supérieure, 46 rue d'Ulm 75230 Paris Cedex 05, France*
[2]*Labor für DNA-Analytik, Wipperstrasse 2, 79100 Freiburg, Germany*
[3]*Laboratoire de Physiologie Cellulaire et de Génétique des Levures, Campus Plaine CP244, Boulevard du Triomphe, 1050 Bruxelles, Belgium*
[4]*Lehrstuhl für Genetik, Fakultät für Biologie, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany*
[5]*Institut für Molekulargenetik, Universität Mainz, Becherweg 32, D-55099 Mainz, Germany*
[6]*Centro de Biologia Molecular, CSIC and UAM, Canto Blanco, 28049 Madrid, Spain*
[7]*Department of Genetics, University of Valencia, Campus of Burjasot, E-46100 Burjasot, Spain*
[8]*Pharmacia Biotech, Parc technologique, rue R. Razel, 91898 Orsay Cedex, France*
[9]*GB-Genome Analysis, Mascheroder Weg 1, D-38124 Braunschweig, Germany*
[10]*Genotype GmbH, Angelhofweg 39, D-69259 Wilhelmsfeld, Germany*
[11]*Departemento de Microbiologia y Genetica, Universidad de Salamanca, Avenida del Campo Charro E-37007 Salamanca, Spain*
[12]*Unité de Biochimie Physiologique, Université Catholique de Louvain, Place Croix du Sud, 2/20, 1348 Louvain-la-Neuve, Belgium*
[13]*Unité de Génétique Moléculaire des Levures, Département de Biotechnologies, Institut Pasteur, F-75724 Paris Cedex 15, France*
[14]*The Biotechnology Centre of Oslo, Ganstadalleen 21, 0371 Oslo, Norway*
[15]*Departamento de Bioquimica y Biologia Molecular, Facultad de Biologia, Universidad de Valencia, E-46100 Burjasot, Spain*
[16]*Hans-Knöll-Institut, Beutenbergstrasse 11, D-07745, Jena, Germany*
[17]*Molecular Genetic Genome Analysis, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, D-69210 Heidelberg, Germany*
[18]*DNA Syntheses and Sequencing Laboratory, Institute of Biotechnology, University of Helsinki, P.O. Box 56, Viikinkaari 9, FIN-00014, Helsinki, Finland*
[19]*Department of Yeast Genetics, Institute of Molecular Medecine, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK*
[20]*L.N.C.I.B., Area Science Park, Padriciano 99, I-34012 Trieste, Italy*
[21]*GATC GmbH, Fitz-Arnold-Strasse 23, 78467 Konstanz, Germany*
[22]*Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Copenhagen Valby, Denmark*
[23]*AGON GmbH, Glienicker Weg 185, D-12489 Berlin, Germany*
[24]*Katholieke Universiteit Leuven, Laboratory of Gene Technology, Willem de Croylaan, 42, B-3001 Leuven, Belgium*
[25]*The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK*
[26]*Department of Biochemistry, Stanford University, Beckman Center, Stanford CA 94305-5307, USA*
[27]*The Genome Sequencing Center, Department of Genetics, Washington University, School of Medicine, 630 S. Euclid Avenue, St Louis, Missouri 63110, USA*
[28]*Martinsrieder Institut für Protein Sequenzen, Max-Planck-Institut für Biochemie, D-82152 Martinsried bei München, Germany.*

The complete DNA sequence of the yeast *Saccharomyces cerevisiae* chromosome IV has been determined. Apart from chromosome XII, which contains the 1–2 Mb rDNA cluster, chromosome IV is the longest *S. cerevisiae* chromosome. It was split into three parts, which were sequenced by a consortium from the European Community, the Sanger Centre, and groups from St Louis and Stanford in the United States. The sequence of 1,531,974 base pairs contains 796 predicted or known genes, 318 (39.9%) of which have been previously identified. Of the 478 new genes, 225 (28.3%) are homologous to previously identified genes and 253 (32%) have unknown functions or correspond to spurious open reading frames (ORFs). On average there is one gene approximately every two kilobases. Superimposed on alternating regional variations in G+C composition, there is a large central domain with a lower G+C content that contains all the yeast transposon (Ty) elements and most of the tRNA genes. Chromosome IV shares with chromosomes II, V, XII, XIII and XV some long clustered duplications which partly explain its origin.

The technique of determining the DNA sequence of large genomes has been unchanged for 21 years[1]. Sequencing the yeast genome required considerable organization by the European Union, which initiated the grouping of 35 laboratories to sequence the first yeast chromosome[2] and coordinate an international effort to sequence the others. Chromosome IV had already been characterized both physically and genetically[3,4], and our sequence data presented are in good agreement with these preliminary data.

The average base composition of chromosome IV is 37.9% G+C, which is lower than for most of the yeast chromosomes (for example, 38.5% for chromosome III (ref. 2) and 38.3% for chromosome II (ref. 5)). Along the 1,513,914 base pairs of the chromosome there are alternating regions about 50 kilobases long of high and low G+C content (Fig. 1). This periodicity is not clearly associated with a variation in gene density, as has been observed for some other chromosomes[6,7]. The central domain of chromosome IV (coordinates 500,000 to 1,215,000) has a much lower G+C value (37.4%) than the two flanking regions (38.2%); a similar observation has been made for the much smaller chromosome VI(ref. 8).

The low G+C content of the central domain seems to be correlated with the presence of Ty elements. All nine Ty1 or Ty2 elements, including a truncated form of Ty1, are localized between coordinates 450,000 and 1,190,000 (Fig. 1). Yeast transposons seem to insert into specific chromosomal regions[9,10] where they are localized preferentially upstream of tRNA genes, as they might interact with the RNA polymerase III machinery[11]. The density of tRNA genes in the central domain of chromosome IV is twice that in the flanking regions in which no Ty elements are found. A total of 27 tRNA genes are localized on each strand of the chromosome, 17 of which are located in the central domain. Of the 27 tRNA genes, 18 are in the vicinity of long terminal repeats (LTRs). Thus most of the tRNA genes, LTRs and Ty elements,

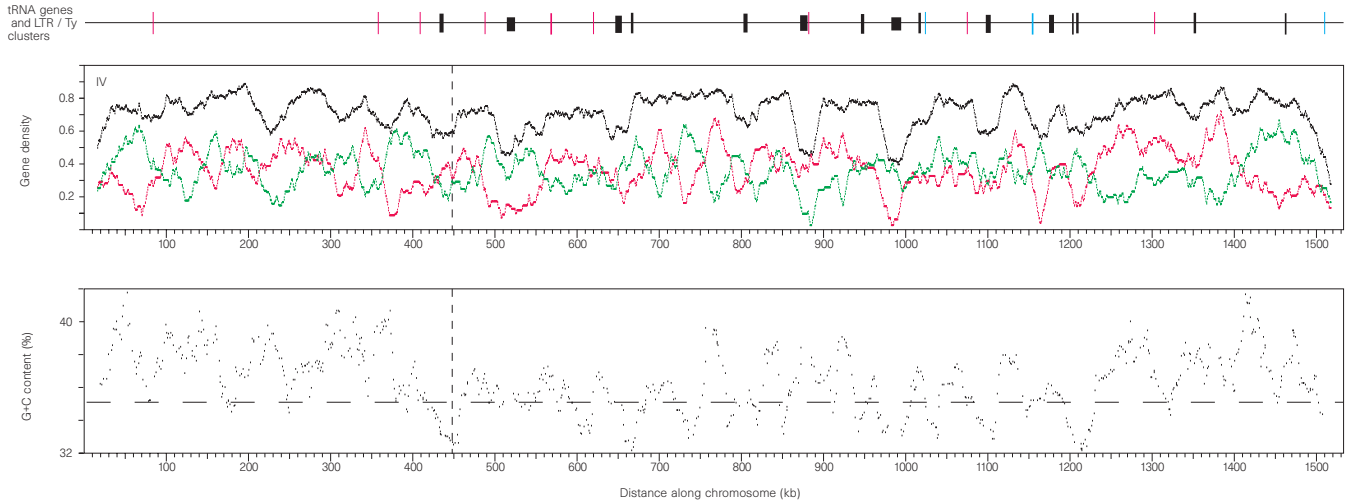


Figure 1 Overall molecular architecture of chromosome IV shows positions of tRNA genes, solo LTR or Ty elements (thin vertical lines), or clusters of them (thick vertical lines), along the chromosome map. Panels show variation of gene density (top) and base composition (bottom) along the sequence-based map of chromosome IV (scale in kilobases from the left telomere). Vertical broken lines indicate the centromere. Gene density is expressed as the probability for each nucleotide to be part of an ORF, andwas calculated using sliding windows of 30 kb (in steps of 0.5 kb) for the Watson strand alone (red line), the Crick strand alone (green line), and the sum of both (black line). G+C content was calculated from the silent positions of codons using a sliding window of 13 consecutive ORFs; the horizontal broken line indicates average G+C content (%) at silent positions of codons.

together with a lower G+C content, are found in this central domain. In chromosome II the 13 tRNA genes and three Ty elements are in AT-rich regions[5].

The left telomere of chromosome IV is very similar to other yeast telomeres. Adjacent to the $C_{1-3}A$ repeat are the usual STR-A, STR-B, STR-C, STR-D and the core X elements (435–904) shared by most of the telomeres[12]. The left end of chromosome IV shares with the right end of chromosome X a large, nearly identical block of sequence similarity more than 19 kilobases long. This duplication includes five ORFs, which code for almost identical products. Indeed the sequences are so similar that we needed to exclude the possibility of contamination of the cosmid contig of chromosome IV by DNA sequences from chromosome X. To confirm our data, we established the genomic sequence of the junction between the duplicated sequence and the rest of the chromosome. Such subtelomeric duplications have often been observed in the yeast genome, suggesting either recent or continuous exchange of genetic information[13]. The right telomere has a less conventional structure with an internal TG_{1-3} repeat.

Using the classical definition of ORFs (one ATG codon followed by at least 99 sense codons), 776 ORFs were recorded in the chromosome; there are also 20 ORFs shorter than 99 amino acids long, making a total of 796 ORFs. Small ORFs of between 25 and 99 codons were extracted and analysed for different properties (codon usage, homologies and ATG environment) to determine their function[7]; 15 had either a putative translation product that is homologous to proteins of other genes, or a codon adaptation index (CAI) greater than 0.2. Five other short ORFs, longer than 91 amino acids, are thought to be 'questionable'. These results support the choice of the threshold of more than 99 codons[14], but show that some short ORFs must be considered[15]. Moreover, this evaluation has to take into account that at least 7% of the ORFs can be considered 'questionable'[6]. Disregarding the retrotransposons, this corresponds to a gene density of one ORF per 2,000 base pairs. The G+C regions correlate roughly, although not precisely, with the regions of increased density. There are approximatively equal numbers of ORFs on the DNA strands (387 on the Watson and 409 on the Crick). However, the gene density is clearly not uniform on a given DNA strand (Fig. 1). This marked prefer ence for an arrangement in which genes are on the same strand has already been observed for other chromosomes[16,17] for which long runs of genes on one strand could be observed. Such gene-rich regions are mainly visible at the two ends of the chromosomes, with more uniform gene density in the central part of the chromosome (550,000 to 1,200,000). Although not as great as predicted by an approximate calculation[18], the high number of Ty and LTR elements in that region should give rise to large numbers of genome rearrangements (inversions, deletions and reciprocal translocations), which could explain this difference with the flanking regions. The construction of gene arrangement has led to general features that probably reflect important functional constraints. Thus the size of regions between ORFs is clearly dependent on the orientation of the flanking genes. In the case of divergent promoters, the mean size is 744 base pairs, whereas it is just 324 bp for convergent terminators. An intermediary situation (593 bp) is found for terminator–promoter combinations. This striking difference in inter-ORF size is probably due to the sequence requirements in the promoter regions for the regulation of gene expression. Nevertheless, the mean size of the inter-ORF region in the case of head-to-head gene orientations is small, and suggests that many divergent genes share common regulation signals. The *GAL1–GAL10* promoter was the first to be described[19], but many other candidates for common regulation have been revealed by the systematic genome sequencing.

Based on the canonical sequences known to control the splicing process, 30 introns can be identified in genes coding for proteins[20]. This represents 4% of genes having an intron in their ORF or in the 5′ untranslated region, a figure close to the situation for the genome as a whole. Of these intron-containing genes, 12 code for ribosomal proteins, three for proteins of the actin family, three for proteins involved in the ubiquitin-dependent protein degradation system, and the rest are distributed between genes that do not necessarily have a high CAI.

Although dependent on the criteria used to estimate the significance of sequence similarities, roughly 30% of the 796 ORFs of chromosome IV are orphans[21] with no sequence relatives in the available databases. This is one of the most exciting findings from the systematic sequencing approach of the yeast genome. Future work will tell us whether some of these genes are really 'yeast specific' and why they have escaped detection by the genetic approaches. The number of sequence orphans will no doubt decrease with the arrival of new sequence data[22]. As an example from chromosome IV, the ORF YDL120w, which had no relatives, is homologous to the human gene recently discovered to be involved in Friedreich's ataxia[23].

Chromosome IV is the longest chromosome in terms of coding sequences, and so might be expected to have features that are scattered on the other chromosomes. One of these features might be the non-uniform organization of the chromosome. The central domain (from 500,000 to

1,250,000) which makes up half of chromosome IV has several distinctive features. First, as with most of the chromosomes, it has more or less regularly spaced regions rich in G+C, but its central domain has a lower G+C content. Second, this central domain contains all of the Ty and most of the LTR elements found on the chromosome. Third, the central domain also contains 18 of the 27 tRNA genes, so its tRNA gene density is twice that of the rest of the chromosome. Finally, the DNA strand distribution of the ORFs is different in the central region when compared with that of the flanking regions. The ORF arrangement of this region might result from a greater genetic plasticity.

Analysis of structural relationships inside the yeast genome might provide an insight into eukaryotic genome organization and evolution. Redundancy is one of the most salient features of the yeast genome structure[24], and the DNA sequence of the whole yeast genome reveals several types of redundancy, probably originating from different biological processes. The most common form of redundancy involves individual genes that have a homologue in the genome; about 20% of the genes of chromosome IV are in this class. Second, there are clusters of very similar copies of a gene, often arranged in tandem; for example, there are five copies of the *ENA1* (or *PMR2*) gene on chromosome IV. Third, subtelomeric duplications are frequent and involve large regions of chromosomes that are very similar in both coding and non-coding regions[13]. Finally, clustered duplications are characterized by clusters of homologous genes in the same order, usually in the same orientation, and interspersed by long DNA fragments. Such paralogous regions have already been described between chromosomes III and XIV (ref. 25) and between chromosomes V and X (ref. 26). Only in the case of the duplication between chromosomes III and XIV is the gene order conserved. It is 15 kilobases long and contains four genes. The clustered duplications on chromosome IV are made up of at least 336 kilobase pairs, including 49 pairs of homologous genes. Chromosome IV shares large ordered cluster of homologous genes with chromosomes II, V, VIII, XII and XIII (ref. 27). A careful analysis of these duplications will no doubt tell us a great deal about the evolution of the yeast genome. In the largest interchromosomal clustered duplications, involving chromosome IV (coordinates 449,752–569,763) and chromosome II (238,164–407,122), the 18 gene pairs are all transcribed in the same direction. When known, most of the genes from a pair code for proteins with homologous but not identical functions (for example *GAL1* and *GAL3*)[19]. Homologous genes from a clustered duplication can also be completely identical or totally different in their function. An extreme case of divergence involves YDR037w, which codes for a lysyl tRNA synthetase, and YBR060w_A, which is part of chromosome II, in which many stop codons interrupt an ORF, of which parts are homologous to YDR037w. Such a pseudogene could not be detected by searching the DNA sequence of chromosome II, as it has very short ORFs and no ATG codon. To our knowledge, this is the most degenerated yeast pseudogene yet discovered. This observation suggests that similar degenerated pseudogenes may have escaped previous analyses, and hence that the total number of pseudogenes may be underestimated.

A pair of genes from a clustered duplication can also differ in the presence of an intron. Both YDR055w and YBRO78w are homologous to the gene *SPS2*, but only YBR078w has an intron, and the CAI of the two genes differs from 0.27 (YDR055w) to 0.61 (YBR078w), suggesting an unusual evolutionary process. The compared analysis of the interspersed DNA fragments is also very informative. For example, chromosome II has a Ty element where an LTR element is present at the equivalent position on chromosome IV, suggesting that the Ty element was lost from chromosome IV after the duplication process.

The greatly different degrees of similarities between the different gene pairs composing a duplicated region indicate that at least some of the duplications have evolved at very different rates, suggesting in some cases that gene conversion processes[28] have interfered with slower evolutionary processes. A careful quantitative analysis of the relative evolution rates of the different elements will be required to establish a chronological order of the different events. Nevertheless, evidence suggests that a first duplication event has been followed by the dispersal of the duplicated elements by the insertions of DNA fragments of various sizes and gene compositions. Most of these clustered duplications in chromosome IV are localized in the pericentromeric region. The centromere itself is included in the longest duplicated region, which occurs between chromosomes IV (coordinates 450,000–570,000) and II (238,000–407,000). Similar localizations of clusters have already been noticed on other chromosomes[7,17,25]. To explain the proximity of the centromere, it was suggested that the gene dispersion of the initial cluster of duplicated genes might be slower in the centromeric regions than nearer the telomere owing to the adverse effects of rearrangements on chromosome segregation[25]. Alternatively, the centromeric duplications might have been essential steps in the construction of the yeast genome[25]. In agreement with these ideas, the central domain of chromosome IV contains few traces of clustered redundancies, perhaps because of its genetic plasticity. These preliminary observations indicate that the availability of the complete sequence of the yeast genome will allow a greater understanding of the processes involved in creating the genome architecture. □

Methods

The sequence was assembled from a set of 44 partly overlapping cosmids and lambda phages from two independent contigs of chromosome IV. The 650-kb cosmid contig corresponding to the left part of chromosome IV was constructed mainly from a specific cosmid library obtained from a gel-purified chromosome portion (J. D. H. *et al.*, unpublished), and a few other cosmid clones from this contig came from a library[29]. The rest of the chromosome sequence was established from a cosmid-lambda phage library (L. Riles & M. Olson, unpublished, and ref. 30). The two cosmid contigs were made from two closely related yeast strains: AB972, derived from S288C[30], and FY1679, a diploid strain issued from the cross between FY23 and FY73, both of which are isogenic with S288C except for the markers indicated. Sequence analysis of a large overlapping fragment (of about 170 kb) confirmed that the strains AB972 and FY1679 are very similar, as the number of base differences was below the estimated error rate. Only the extreme left telomeric regions of the two strains clearly differ, probably in the number of their TG_{1-3} repeats (C. B. and C. J., unpublished). The telomeres were isolated independently and sequenced from a plasmid clone generated by integration at the TG_{1-3} repeats of the telomere, followed by excision of the plasmid and capture of the flanking sequences[13]. Two gaps in the 650-kb left cosmid contig (constructed from the strain FY1679) were filled with lambda clones from the library constructed from AB972. They correspond to the regions 9,756–11360 and 363,100–368,150 . The left 600-kb region was sequenced according to the rules followed by the European consortium and the 20 cosmids and phages were distributed to 18 contractors, whereas the central part was sequenced by the Sanger Centre (EMBL database SCCHRIV, accession no. Z71256); the rest of the chromosome was sequenced by groups from Washington University in St Louis and Stanford University in the United States.

There were very few base differences in the overlapping fragment sequenced in parallel by the Sanger Centre and by the European consortium, demonstrating that both approaches are reliable. However, a verification procedure was necessary because of the greater heterogeneity of the European approach. This was done on 25 regions of the left part of the chromosome, according to the protocol of G. V. (manuscript in preparation). This allows direct polymerase chain reaction (PCR) sequencing of a 300-bp region of the yeast genome limited by two previously designed oligonucleotides. We could thus correct a sequence in which a bacterial transposon had been inserted during the cloning process but no real sequence error could be detected at this final step of the sequencing project. Sequence errors could only be corrected after examination of the raw sequence data. From these data, the error rate of this part of the yeast chromosome IV sequence presented is less than four errors per 10 kb. In the central part of the chromosome the error rate is estimated as less than one error per 10 kb. Specific strategies were developed to sequence difficult parts of the chromosome. Thus, for example, to finish the regions between the two transposons located in cosmid 8142 (http://www.sanger.ac.uk/~yeastpub/svw/sequencing.html), a PCR product covering this region from strain MCYC2576 was sequenced. This strain, a gift from E. Louis, did not contain the transposons.

Received 24 July 1996; accepted 11 March 1997.

1. Sanger, F. & Coulson, A.R. *J. Mol. Biol.* 94, 441–447 (1975).
2. Oliver, S. *et al. Nature* 357, 38–46 (1992).
3. Mortimer, R. K. *et al.* http://genome-www.stanford.edu/sacchdb/edition 12.html (1995).
4. Olson, M. V. in *The Molecular and Cellular Biology of the Yeast* Saccharomyces: *Genome Dynamics, Protein Synthesis and Energetics* (eds Broach, J. R., Jones, E. W. and J. R. Pringle) 1–39 (Cold Spring

Harbor Laboratory Press, NY, 1991.
5. Feldmann, H. *et al. EMBO J.* 13, 5795–5809 (1994).
6. Dujon, B. *et al. Nature* 369, 371–378 (1994).
7. Galibert, F. *et al. EMBO J.* 15, 2031–2049 (1996).
8. Murakami, Y., Naitou, M., Hagiwarar, H. & Shibata, T. *Nature Genet.* 10, 261–268 (1995).
9. Ji, H. *et al. Cell* 73, 1007–1018 (1993).
10. Lochmüller, H., Stucka, R. & Feldmann, H. *Curr. Genet.* 16, 247–252 (1989).
11. Voytas, D .F. & Boeke, J. D. *Trends Genet.* 9, 421–427 (1993).
12. Pryde, F. E., Huckle, T. C. & Louis, E. J. *Yeast* 11, 371–382 (1995).
13. Louis, E. J. & Borts, R. H. *Genetics* 139, 125–136 (1995).
14. Termier, M. & Kalogeropoulos, A. *Yeast* 12, 369–384 (1996).
15. Navarre, C., Caty, P., Leterme, S., Dietrich, F. & Goffeau, A. *J. Biol. Chem.* 267, 21262–21268 (1994).
16. Bussey, H. *Proc. Natl Acad. Sci. USA* 92, 3809–3813 (1995).
17. Johnston, M. *et al. Science* 265, 2077–2081 (1994).
18. Kupiec, M. & Petes, T.D. *Mol. Cell. Biol.* 8, 2942–2952 (1988).
19. Johnston, M. & Davis, R.W. *Mol. Cell. Biol.* 4, 1440–1448 (1984).
20. Kalogeropoulos, A. *Yeast* 11, 555–565 (1995).
21. Dujon, B. *Trends Genet.* 12, 263–270 (1996).
22. Casari, G., De Daruvar, A., Sander, C. & Schneider, R. *Trends in Genet.* 12, 244–245 (1996).
23. Campuzano, V. *et al. Science* 271, 1423–1427 (1996).
24. Oliver, S.G. *Nature* 379, 597–600 (1996).
25. Lalo, D., Stettler, S., Mariotte, S., Slonimski, P. P. & Thuriaux, P. *C. R. Acad. Sci.* 316, 367–373 (1993).
26. Melnick, L.M. & Sherman, F. *J. Mol. Biol.* 233, 372–388 (1993).
27. Heumann, K. & Mewes, H. W. *Nature Genet.* (submitted).
28. Jinks-Robertson, S. & Petes, T.D. *Proc. Natl Acad. Sci. USA* 82, 3350–3354 (1985).
29. Thierry, A., Gaillon, L., Galibert, F. & Dujon, B. *Yeast* 11, 121–135 (1995).
30. Riles, L. *et al. Genetics* 134, 81–91 (1993).

Acknowledgements. The laboratory consortium operating under contracts with the European Commission was initiated and organized by A. G. A. Vassaroti and P. Mordant were in charge of the administrative coordination. This work was supported by the European Commission under the Biotech programmes, the Groupe de Recherche et d'Etudes sur le Génome, the Centre National de la Recherche Scientifique, the Wellcome Trust, the Région Wallone, the Belgian Federal Services for Science Policy, the Research Fund of the Katholieke Universiteit Leuven, the Région de Bruxelles Capitale, the Fundacion Ramon Aceres and Comision Interministerial de Ciencia y Tecnologia and the Bundesminister für Forschung und Technologie.

Correspondence and requests for materials should be addressed to C.J. (e-mail: jacq@biologie.ens.fr).

The nucleotide sequence of *Saccharomyces cerevisiae* chromosome V

F. S. Dietrich*, J. Mulligan*, K. Hennessy, M. A. Yelton*, E. Allen, R. Araujo, E. Aviles, A. Berno, T. Brennan, J. Carpenter, E. Chen, J. M. Cherry, E. Chung, M. Duncan, E. Guzman, G. Hartzell, S. Hunicke-Smith, R. W. Hyman, A. Kayser, C. Komp, D. Lashkari, H. Lew, D. Lin, D. Mosedale, K. Nakahara, A. Namath, R. Norgren, P. Oefner, C. Oh, F. X. Petel, D. Roberts, P. Sehl, S. Schramm, T. Shogren, V. Smith, P. Taylor, Y. Wei, D. Botstein & R. W. Davis

Stanford DNA Sequencing and Technology Center, 855 California Avenue, Palo Alto, California 94304 and Departments of Biochemistry and Genetics, Stanford University Medical School, Stanford, California 94305, USA
**These authors were in charge of the project at different times during the study.*

Here we report the sequence of 569,202 base pairs of *Saccharomyces cerevisiae* chromosome V. Analysis of the sequence revealed a centromere, two telomeres and 271 open reading frames (ORFs) plus 13 tRNAs and four small nuclear RNAs. There are two Ty1 transposable elements, each of which contains an ORF (included in the count of 271). Of the ORFs, 78 (29%) are new, 81 (30%) have potential homologues in the public databases, and 112 (41%) are previously characterized yeast genes.

As part of an international collaborative effort to sequence the total genome of the yeast *Saccharomyces cerevisiae*, we have deduced the DNA sequence of 569,202 base pairs of yeast chromosome V. We used an overlapping set of recombinant yeast cosmid and lambda clones that together cover the entire chromosome (except for the extreme ends of the telomeres). A line drawing of chromosome V and the identification of the recombinant DNAs sequenced are shown in Fig. 1. The sequence was broken arbitrarily into 11 slightly overlapping pieces for ease of handling and deposited in Genbank (see Fig. 1 for accession numbers).

Sequencing was accomplished in two phases: the 'shotgun' phase, using dye-primer chemistry, and the 'finishing' phase, using the polymerase chain reaction (PCR) and dye-terminator chemistry. There were no gaps in the sequence at the end of shotgun sequencing and assembly. The assembled, continuous sequence of chromosome V has 569,202 bp, starting from the guanine residue of the *Sau*3A site on the left vector boundary of the leftmost clone (1160 in Fig. 1). The 569-kilobase sequence is based on the results from 32,631 individual lanes of sequencing gels, or reads. The average depth of coverage was 12.5-fold. The minimum acceptable coverage was three, with at least one read from each strand.

After shotgun sequencing and assembly, problems remained in the sequence at a frequency of (roughly) two per kilobase and were of several types. They included the inability to count unambiguously the number of repeating units, such as poly (dA), and guanine compressions. There were also small regions in which only one of the two strands had been sequenced. These difficulties were resolved during the finishing phase.

After finishing, the 569-kb contig was checked against three external sets of data. First was the use of tetrad segregation data to derive a genetic map for yeast[1]. The chromosome V gene order based on DNA sequence was in complete agreement with the tetrad segregation data. There were two locations on the genetic map (*CEN*V at 151 kb and *PRO*3 at 200 kb) where closely spaced loci had been mapped against distant markers and not against each other, resulting in ambiguities of relative locus order[1], which were resolved using the DNA sequence. The gene order across the centromere is *GLC*3 tRNA-Arg *GCN*4 *CEN*V *MNN*1. In the region of *PRO*3, at 200 kb, the gene order is *PRO*3 *GPA*2 *GCD*11 *CHO*1 *GAL*83. Second, our sequence was compared to the *S. cerevisiae* sequences already deposited in Genbank, using both the FASTA and BLAST programs[2,3]. In the rare cases of sequence difference, we re-examined our trace files. Remaining ambiguities were resolved using the same methodology as finishing. Third, we checked our data against the primary *Eco*RI/*Hin*dIII double-digestion fragment maps of the recombinant yeast DNAs[4]. Our sequence was examined for *Eco*RI and *Hin*dIII cleavage sites. Of 534 mapped fragments, there were only five discrepancies, which is a tribute to the care taken in preparing the cleavage sites map[4]. The five apparent discrepancies between the double-digest map[4] and our sequence are: the map had doublets where the sequence predicts singlets after bases 272, 193; 280,936; and 441,102; the map has a fragment that was not found in the sequence after base 414,946; and the sequence is missing a cleavage site after base 506,807.

We examined all six possible reading frames of the 569-kb sequence for ORFs of at least 300 bp that began with a start codon and ended with a stop codon. As a special case, an ORF could be interrupted if there were yeast splice donor/acceptor/branchpoint sequences present at the appropriate intervals. The remaining sequence was examined using FASTA and BLAST for homology to sequences in the public databases. This enabled us to find small ORFs, as well as the centromere, 13 tRNAs, two Ty1 elements (which each contain an ORF), four small nuclear RNAs, many delta and delta-like elements, and the highly conserved X and Y′ sequences characteristic of yeast telomeres (see refs 5, 6) at the far left and right ends.

Initially, 271 ORFs were identified in the 569-kb sequence, although this number has changed as evaluation continued. The 271 ORFs make up roughly 70% of the sequence, with an average of 2.1 kb per ORF. The 'average' ORF (1.4 kb) encodes 475 amino acids. Of the ORFs, 112 (41%) have been characterized previously, 81 (30%) have apparent homologues in the public databases, and 78 (29%) are new; six (2%) are spliced. Of the 81 apparent homologues, 55 of these are to other *S. cerevisiae* sequences.

The fractional G+C content of the 569,202 bp of chromosome V is 0.384. The combined ORF DNAs have a fractional G+C content of 0.401, and the combined 'non-ORF' DNA has a G+C content of 0.351.

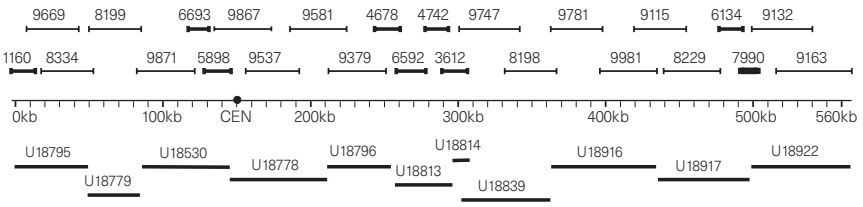


Figure 1 The central line representing *S. cerevisiae* chromosome V is marked in kilobase pairs, starts at the left at the guanine of the *Sau*3A site of the leftmost recombinant yeast DNA, and extends to the right to 570 kb. The centromere, *CEN*, is represented by a solid circle at ~151 kb. Above the line, placed across their map positions, are the individual recombinant yeast DNAs that were sequenced; 16 cosmids (thin lines), 8 lambdas (thick lines), and 1 plasmid (very thick line). Genbank accession numbers are placed below the line, above the bars indicating the corresponding map positions. From left to right, Genbank accession numbers are U18795, U18779, U18530, U18778, U18796, U18813, U18814, U18839, U18916, U18917 and U18922. There is some deliberate overlap between Genbank entries to maintain contiguity.

There are only two places on chromosome V where the quality of the sequence is not high. In the first case, at about 312 kb, there are ~50 bp of unique sequence bounded on both sides by poly (dA): poly (dT). *Taq* polymerase, and the other DNA polymerases tested, frequently terminated within the homopolymer, and seldom reached the short unique sequence. Therefore, we have only a few reads across the unique sequence. In the second case, at about 450 kb, there is a 5-kb stretch that contains many delta and delta-like sequences interspersed with a small amount of unique sequence. The clone containing this segment was shotgun sequenced to an average of 16-fold redundancy, yet there were relatively few reads in this region. Therefore, for PCR amplification, this 5-kb region was divided into many virtual parts, based on the positions of the unique sequences. Several custom primer pairs, and internal sequencing primers, were designed and synthesized for each part[7]. These were used in PCR amplification reactions with total yeast genomic DNA as the template. We have sequenced carefully across this region. For most bases, there is sequence from both strands.

There were three special cases that warrant further attention. First, a point mutation had occurred during either cloning or subsequent propagation in *Escherichia coli*. In an overlap region shared by two recombinant DNAs (lambda 5898 and cosmid 9867; Fig. 1), the sequence should be the same, but in this case there was one reproducible base difference. Lambda 5898 has a guanine residue where cosmid 9867 has an adenine residue. When total yeast genomic DNA was used as a template for PCR amplification, the product of which was used as a template for dye-terminator sequencing, the base at that position was an adenine. The traces showed no indication of a naturally occurring polymorphism. We therefore conclude that the guanine in lambda 5898 was the result of a mutation. In the second special case, we examined the ORFs for any apparently premature, in-frame stop codon, and found two that were puzzling. The first was the TGA stop codon at position (rounded-off) 352 kb. There are three rightward-reading frames, and this stop codon (TGA) is in the first of these. Following the TGA in this frame, there is a lysine codon (AAA) and a methionine codon (ATG) followed by a long ORF. In the second reading frame there are three stop codons: two (TGA, TAA) are next to each other, the third is five triplets further on; this frame is truly stopped. However, there is a long ORF starting with a methionine codon considerably upstream and ending at the double stop codons TGA, TAA. The third reading frame has many stop codons. The two ORFs share one base: A, the first base (ATG) of the first reading frame and the last base (AAA) of the second reading frame. In general, yeast ORFs are separated by several hundred bases. Except for a -1 frame shift in the first reading frame, there would be one ORF rather than two. However, despite sequencing through this position many times, the sequence, the TGA and the five A bases in a row remained invariant. The second apparently premature stop codon occurred in the ORF that corresponds to *FLO8* at ~375 kb. The TAG stop codon between YER108c and YER109c appears not just in the recombinant yeast DNA, but also in PCR amplifications from total yeast DNA. Thus it is possible that *FLO8* is not functional in yeast strain AB972. In the third special case, in the entire collection of recombinant yeast DNAs, only lambda 3612 (Fig. 1) covers this region. We found lambda 3612 to be highly unstable, giving rise to non-random DNA deletions at extremely high frequency. Starting with 30 individual plaques from a primary stock, only one yielded lambda 3612 DNA without a detectable deletion, as judged by *Eco*RI/*Hin*dIII double-digestion patterns, but even that gave rise to deleted DNAs upon subsequent growth. Therefore, all of the lambda 3612 sequence came (uncharacteristically) from just one preparation of DNA.

A comparison of the DNA base sequence of chromosome V to that of the other *S. cerevisiae* chromosomes shows that there are two stretches that have similar genes in the same order on two other yeast chromosomes. A portion of the left arm of chromosome V, containing *CYC7* and *RAD*, shows the same relative gene order as chromosome X, but in the opposite orientation, as noted previously[8]. In addition, a 60-kb region of the left arm of chromosome IX contains nine genes or ORFs for which each has an apparent homologue within a 60-kb region of the right arm of chromosome V. The nine putative protein pairs and their calculated similarity (identicality)[9] are: (1) YIL045w/YER054c, 63 (44) %; (2) YIL050w/YER059w, 71 (52) %; (3) YIL051c/YER057c, 87 (70) %; (4) YIL053w/YER062c, 97 (92) %; (5) YIL056w/YER064c, 63 (47) %; (6) tRNA-ser, 100%; (7) YIL057c/YER067w, 85 (67) %; (8) *RNR*3/YER070w, 90 (82) %; and (9) YIL074c/YER081w, 95 (91) %. On chromosome V itself, the *FCY*2 protein product[10] and the putative YER060w translation product are 87 (75) % related.

When considering the accuracy of our 569,202-bp sequence of *S. cerevisiae* chromosome V, we must emphasize that (essentially) all of the sequence was determined from recombinant DNA propagated in *E. coli*. Even if our sequence of the recombinant DNAs were 100% accurate, there may be sequence differences between the recombinant DNAs and the yeast genome. We identified one apparent point mutation solely because it occurred within a region common to two recombinant DNAs. Other point mutations, occurring during cloning or propagation in *E. coli*, would probably not be detected.

There is a much more dramatic example of a discrepancy between the sequence of a recombinant DNA and the yeast genome. We deposited our chromosome V sequence in Genbank and SacchDB (http://genome-www.stanford.edu/) in December 1994. On 18 April 1996, we received an e-mail from J.-L. Souciet, J. de Montigny and S. Potier (CNRS, Strasbourg) politely telling us that ~2 kb were missing from our sequence. They had found that, in addition to *FCY2* (YER056c; encoding a purine-cytosine permease[10]) at ~267 kb and a closely related ORF (YER060w) at ~275 kb, there is also a third closely related ORF in this region (Genbank accession no. X97346) (Fig 2.). Our sequence across the apparent deletion came from two libraries made from lambda 6592. The sequence is 12-fold deep, and there are no traces that diverge from our Genbank sequence. The Genbank sequence is therefore an accurate sequence of this particular recombinant DNA: lambda 6592. We had intended to sequence cosmid 9380. Repeated attempts to prepare 9380 yielded only minute amounts of DNA. We abandoned cosmid 9380, and instead sequenced three lambda DNAs: 4678, 6592 and 4742. From the tiny amount of cosmid 9380 DNA, we constructed (and sequenced) a *Sau*3A library. The data (traces) from the library of *Sau*3A-cleaved 9380 DNA were used within the assemblies

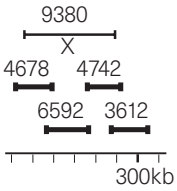


Figure 2 The bottom line is a schematic map of yeast chromosome V from 240 kb to 310 kb. Individual, sequenced recombinant DNAs are placed above the line and across the appropriate map positions. Cosmid 9380 covers map positions ~240 kb to 300 kb and overlaps lambda 3612. The latter is the only recombinant DNA in the entire Olson collection[4,14] that covers map positions 300 kb to 310 kb. Because 9380 could only be produced in minute amounts, three lambda DNAs (thick lines) were substituted: 4678, 6592 and 4742. The X marks the position of the ~2 kb deletion in 6592 DNA.

of the three lambda DNAs. There were some data from the 9380 library that were left over, mostly vector without insert, low-quality traces, *etc.*, which we put aside. We searched our 'left-over' 9380 traces for homology to unique sequences in Genbank X97346 and found one excellent (97.3% identity) match for 294 bp (bounded by *Sau*3A sites). We conclude that lambda 6592 has a deletion relative to the yeast genome and that one additional ORF should be added, bringing the ORF total to 272.

To complete the sequence of chromosome V, an insertion of 2,011 bp (Genbank accession number X97346) should be made at base 275,951 of our sequence as has been done in SacchDB. In addition, H. Wedler and R. Wambutt have sequenced the left (Genbank accession no. U73806) and right (Genbank accession no. U34775) telomeres of yeast chromosome V. Within SacchDB, 2,477 bp have been placed to the left of the leftmost base (the G of the leftmost *Sau*3A site) of our sequence. That G is no longer base 1 but base 2,478. Within the left telomere, there is an ORF, YEL077c, which brings the current ORF total to 273. Concomitantly in SacchDB, 3,181 bp have been placed to the right of our sequence.

Basically, there are two types of errors: random and systematic. If there is a random error in an individual sequence read, we will find and correct that error because we sequenced both strands to high redundancy (average of 12.5-fold). It is much more difficult to identify a systematic error that is inherent in, for example, the dye-primer chemistry, polyacrylamide gel electrophoresis, *Taq* polymerase, or base-calling software, that systematically misreads or deletes a base(s) within a particular sequence. *Taq* polymerase seemed to have systematic difficulties synthesizing across short repeating units; not only is the number of repeats often ambiguous, the sequence traces following a repeat are often diminished in signal quality. We believe that this observation reflects an inherent characteristic of *Taq* polymerase. Of the several DNA polymerases tested in an attempt to solve this problem, Amplitaq FS polymerase (Perkin-Elmer 402079) yielded the best number of good sequence calls, but did not solve the problem completely. However, the problem in counting the short repeating units unambiguously may not be a sequencing problem but in some cases may reflect true biological heterogeneity. A second possible systematic error arises from the well-known guanine compressions. Guanine compressions are usually identified when the base-calling software identifies fewer guanines on one strand than cytosines on the opposing strand. However, if two (or more) guanine compressions are positioned symmetrically on opposing strands, the compression on one strand is compensated by an analogous compression on the other strand. There are no 'extra' cytosines, and the existence of the compressions could be missed.

One reason for sequencing all of the *S. cerevisiae* DNA is that yeast is important as a model organism. A second reason is to test the approaches to, and develop technologies for, large-scale DNA sequencing in preparation for the sequencing of the human genome. In this regard, we would like to describe some important lessons learned during the sequencing of yeast chromosome V. First, 800 kb were shotgun sequenced to achieve 569,202 bp of contiguous sequence, an inefficiency of 40%. Considerable time and money would have been saved if the ends of the recombinant yeast DNAs had been mapped relative to each other (a 'sequence-ready' contig of cosmid DNAs). Second, a large amount of freezer space was used in archiving recombinant M13 DNAs, a small percentage of which were later used as templates for finishing. An important reason in the delay of finishing was the cost of oligonucleotide primers for PCR. Finishing has been made economical by the availability of low-cost oligonucleotides[7], so long-term storage of M13 DNAs is no longer necessary. Third, when the Yeast Genome Project was started, the conventional wisdom had that it was necessary to sequence a set of overlapping cosmids. However, we now know that the sequence of DNA as large as bacterial genomes can be assembled using a shotgun approach[11,12]. If we started again, we would purify *S. cerevisiae* chromosome V directly by pulse-field gel electrophoresis, hydrodynamically shear the DNA to an average size of 1 kb (ref 13) and shotgun clone the sheared DNA directly into the M13 sequencing vector. The yeast genome could probably have been sequenced by the direct shotgun cloning of total genomic DNA to generate one M13 sequencing library. Individual cosmid and lambda clones could have been used to fill holes and resolve ambiguities. □

Methods

All of the *S. cerevisiae* recombinant DNAs sequenced in this study were constructed in the laboratory of M. Olson[4,14]. With the exception of plasmid 7990, which was derived from two yeast strains[15], all of the recombinant DNAs were derived from yeast strain AB972. Those recombinant DNAs with number designations less than 8000 are lambdas (except for plasmid 7990), those with numbers over 8000 are cosmids. We sequenced 16 cosmids (8198, 8199, 8229, 8334, 9115, 9132, 9163, 9379, 9537, 9581, 9669, 9747, 9781, 9867, 9871 and 9981), eight lambdas (1160, 3612, 4678, 4742, 5898, 6134, 6592 and 6693), and one plasmid (7990) (Fig. 1). We also obtained some sequence from another five cosmids (8063, 9268, 9380, 9495 which contained a large deletion and 9675) and two lambdas (3955 and 6052). These *S. cerevisiae* recombinant DNAs (except 9495) are available from the American Type Culture Collection.

The 'shotgun' sequencing strategy was to reduce randomly the size of the yeast recombinant DNAs ('inserts') to approximately 1 kb. The inserts were ligated to the M13 sequencing vector by using a 'linker–adaptor' system, which minimizes the formation of chimaeric DNAs. The recombinant M13 'sequencing library' was electroporated into *E. coli* and plated. Individual M13 plaques were picked and grown, and recombinant M13 DNAs were purified. (Our detailed laboratory protocols are freely available on the World-Wide Web at http://sequence-www.stanford.edu.)

The shotgun sequencing used dye-primer chemistry in cycle sequencing reactions, followed by fluorescence detection using an ABI 373A automated sequencer. Most of the sequence data from individual lanes ('traces' or 'reads') were edited automatically using custom software, with borderline cases being edited manually using the TED software[16]. Individual sequence reads were assembled using the XBAP program[16]. The final sequence was determined by editing manually the assembled reads.

Where there seemed to be overlapping ORFs, in either the same or in the opposite direction, the conventional assumption was made that yeast seldom uses both overlapping frames. Three criteria were used to determine which was the most likely ORF. First, using both FASTA and BLAST programs, each of the overlapping reading frames was examined for homology to known genes in the public databases; an ORF with homology was chosen over one without. Second, each organism has its own distinctive preference for certain codons over others. This preference can be expressed in arithmetic terms, as it is within the GeneFinder program (L. Hiller and P. Green, 1990–1993; documentation, software and yeast codon usage data files. Genome Sequencing Center, Washington University School of Medicine, St Louis, MO 63108, USA). GeneFinder was used to compare the codon usage for each of the overlapping ORFs. The ORF that more closely matched yeast's codon usage was chosen; in almost all cases, this distinction was unequivocal. Third, the longer ORF was selected. The nomenclature for yeast ORFs is composed of five letter/number combinations: Y (for yeast), E (the fifth letter of the alphabet for V), L or R (for the left or right arm, as defined genetically), a number (counted sequentially from the centromere in both directions), and w or c (for the transcribed strand); for example, the *URA3* gene, encoding orotidine-5′-phosphate decarboxylase, is YEL021w.

Received 24 July 1996; accepted 11 March 1997.

1. Mortimer, R. K., Contopoulou, C. R. & King, J. S. *Yeast* 8, 817–902 (1992).
2. Pearson, W. R. & Lipman, D .J. *Proc. Natl Acad. Sci. USA* 85, 2444–2448 (1988).
3. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. *J. Mol. Biol.* 215, 403–410 (1990).

4. Riles, L. *et al. Genetics.* 134, 81–150 (1993).
5. Louis, E. J. & Haber, J. E. *Genetics* 131, 559–574 (1992).
6. Louis, E. J., Naumova, E. S., Lee, A., Naumov, G. & Haber, J. E. *Genetics* 136, 789–802 (1994).
7. Lashkari, D. A., Hunicke–Smith, S. P,. Norgren, R. M., Davis, R.W. & Brennan, T. *Proc. Natl Acad. Sci.* 92, 7912–7915 (1995).
8. Stiles, J. I., Friedman, L. R., Helms, C., Consaul, S. & Sherman, F. *J. Mol. Biol.* 148, 331–346 (1981).
9. Smith, T., Waterman, M. & Burks, C. *Nucleic Acids Res.* 13, 645–656 (1985).
10. Weber, E., Rodriguez, C., Chevallier, M. R. & Jund, R. *Mol. Microbiol.* 4, 585–596 (1990).
11. Fleischmann, R. D. *et al. Science* 269, 496–512 (1995).
12. Fraser, C. M. *et al. Science* 270, 397–403 (1995).
13. Oefner, P. S. *et al. Nucleic Acids Res.* 24, 3879–3886 (1996).
14. Olson, M. V. *et al. Proc. Natl. Acad. Sci. USA* 83, 7826–7830 (1986).
15. Fleer, R., Nicolet, C. M., Pure, G. A. & Friedberg, E. C. *Mol. Cell. Biol.* 7, 1180–1192 (1987).
16. Gleeson, T.J. & Staden, R. *Comp. Appl. Biosci.* 7, 398 (1991).

Acknowledgements. We thank the members of the informal, international group who sequenced the entire yeast genome for their generosity and cooperation that enabled the Yeast Genome Project to be completed a year earlier than scheduled and under budget. In particular, we thank L. Riles and M. Olson for sending us their set of recombinant yeast DNAs and unpublished mapping data; E. Gilbertson for assistance in preparing this manuscript; and N. Schroff and A. Wynant for technical assistance. This work was supported by grants from the NIH (U. S. Public Health Service).

Correspondence and requests for materials should be addressed to R. W. D. (e-mail: gilbert@genome.stanford.edu). The sequence of yeast chromosome V has been deposited in Genbank.

The nucleotide sequence of *Saccharomyces cerevisiae* chromosome VII

H. Tettelin[1], M. L. Agostoni Carbone[2], K. Albermann[3], M. Albers[4], J. Arroyo[5], U. Backes[4], T. Barreiros[6], I. Bertani[7], A. J. Bjourson[8], M. Brückner[9], C.V. Bruschi[7], G. Carignani[10], L. Castagnoli[11], E. Cerdan[12], M. L. Clemente[10], A. Coblenz[4], M. Coglievina[7], E. Coissac[13], E. Defoor14, S. Del Bino[1], H. Delius[15], D. Delneri[7], P. de Wergifosse[1], B. Dujon[16], P. Durand[17], K. D. Entian[18], P. Eraso[19], V. Escribano[19], L. Fabiani[20], B. Fartmann[21]*, F. Feroli[10], M. Feuermann[22], L. Frontali[20], M. García-Gonzalez[5]*, M. I. García-Sáez[5], A. Goffeau[1], P. Guerreiro[6], J. Hani[3], M. Hansen[4], U. Hebling[15], K. Hernandez[23], K. Heumann[3], F. Hilger[17], B. Hofmann[15], K. J. Indge[24], C.M. James[24], R. Klima[7], P. Kötter[18], B. Kramer[21]*, W. Kramer[21], G. Lauquin[25], H. Leuther[4], E. J. Louis[26], E. Maillier[13], A. Marconi[20], E. Martegani[27], M. J. Mazón[19], C. Mazzoni[20], A. D. K. McReynolds[8], P. Melchioretto[2], H. W. Mewes[3], O. Minenkova[11], S. Müller-Auer[9], A. Nawrocki[28], P. Netter[13], R. Neu[4], C. Nombela[5], S.G. Oliver[24], L. Panzeri[2], S. Paoluzi[11], P. Plevani[2], D. Portetelle[17], F. Portillo[19], S. Potier[22], B. Purnelle[1], M. Rieger[9], L. Riles[29], T. Rinaldi[20], J. Robben[14], C. Rodrigues-Pousada[6], E. Rodriguez-Belmonte[12], A. M. Rodriguez-Torres[12], M. Rose[18], M. Ruzzi[30], M. Saliola[20], M. Sánchez-Perez[5], B. Schäfer[4], M. Schäfer[9], M. Scharfe[31], T. Schmidheini[23], A. Schreer[4], J. Skala[28], J. L. Souciet[22], H.Y. Steensma[32,33], E. Talla[1], A. Thierry[16], M. Vandenbol[17], Q. J. M. van der Aart[32], L. Van Dyck[1], M. Vanoni[27], P. Verhasselt[14], M. Voet[14], G. Volckaert[14], R. Wambutt[31], M. D. Watson[34], N. Weber[23], E. Wedler[31], H. Wedler[31], P. Wipfli[23], K. Wolf[4], L. F. Wright[8], P. Zaccaria[7], M. Zimmermann[4], A. Zollner[3] & K. Kleine[3]

[1]Unité de Biochimie Physiologique, Université Catholique de Louvain, Place Croix-du-Sud 2/20, B-1348 Louvain-la-Neuve, Belgium
[2]Dipartimento di Genetica e di Biologia dei Microrganismi, Università di Milano, via Celoria 26, I-20133 Milano, Italy
[3]Martinsrieder Institut für Protein Sequenzen, Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, Germany
[4]Institut für Biologie IV, Mikrobiologie, Worringerweg, RWTH-Aachen, D-52056, Germany
[5]Departamento de Microbiología II and Centro de Secuenciación de DNA de la UCM, Faculdad de Farmacia, Universidad Complutense, E-28040 Madrid, Spain
[6]Laboratorio de Genetica Molecular, Instituto Gulbenkian de Ciencia, Ap 14, E-2781 Oeiras Codex, Portugal
[7]Microbiology Group, International Center for Genetic Engineering and Biotechnology, Padriciano 99, I-34012 Trieste, Italy;
[8]The Biotechnology Center for Animal and Plant Health, The Queens University of Belfast, Newforge Lane, Belfast, BT9 5PX, UK
[9]Genotype GmbH, Angelhofweg 39, D-69259 Wilhelmsfeld, Germany
[10]Dipartimento di Chimica biologica, Università di Padova, via Trieste 75, I-35121 Padova, Italy
[11]Dipartimento di Biologia, Università di Roma 'Tor Vergata', via della Ricerca Scientifica, I-00133 Roma, Italy
[12]Departamento de Biologia Celular y Molecular, Facultad de Ciencias, Universidad de La Coruna, Campus de La Zapateira s/n, La Coruna, Spain;
[13]Centre de Génétique Moléculaire, CNRS, Laboratoire Associé à l'Université Pierre et Marie Curie, F-91198 Gif-sur-Yvette Cedex, France
[14]Katholieke Universiteit Leuven, Laboratory of Gene Technology, Willem de Croylaan 42, B-3001 Leuven, Belgium
[15]Deutsches Krebsforschungszentrum, Department for Applied Virology, D-69120 Heidelberg, Germany
[16]Unité de Génétique Moléculaire des levures (URA 1149 of CNRS and UPR 927 of University P.M. Curie, Paris), Department of Biotechnologies, 25 rue du Dr. Roux, Institut Pasteur, F-75724 Paris Cedex 15, France
[17]Unité de Microbiologie, Faculté Universitaire des Sciences Agronomiques de Gembloux, Avenue Maréchal Juin 6, B-5030 Gembloux, Belgium
[18]Institut für Mikrobiologie der Johann Wolfgang Goethe-Universität Frankfurt /Main, Marie-Curie Strasse 9, D-60439 Frankfurt, Germany;
[19]Instituto de Investigaciones Biomédicas del C.S.I.C.y Departamento de Bioquímica, Facultad de Medicina de la U.A.M., E-28029 Madrid, Spain
[20]Dipartimento di Biologia Cellulare e dello Sviluppo, Università di Roma La Sapienza, P. le Aldo Moro 5, I-00185 Roma, Italy
[21]Institut für Molekulare Genetik, Georg-August-Universität, Grisebachstr. 8, D-37077 Göttingen, Germany
[22]Laboratoire de Microbiologie et Génétique URA 1481, Université Louis Pasteur/CNRS, Institut de Botanique, rue Goethe 28, F-67083 Strasbourg Cedex, France
[23]Microsynth GmbH, Schützenstrasse 15, CH-9436 Balgach, Switzerland
[24]Department of Biochemistry and Applied Molecular Biology, UMIST, PO Box 88, Sackville Street, Manchester M60 1QD, UK
[25]Institut de Biochimie Cellulaire, CNRS, rue Camille Saint-Saèns 1, F-33077 Bordeaux Cedex, France
[26]Yeast Genetics, Institute of Molecular Medicine, John Radcliffe Hospital, Oxford, OX3 9DU, UK
[27]Dipartimento di Biochimica e Fisiologia Generali, Università di Milano, via Celoria 26, I-20133 Milano, Italy
[28]Institute of Microbiology, Wroclaw University, Przybyszewskiego 63, P-51148 Wroclaw, Poland
[29]Genome Sequencing Center and Department of Genetics, Washington University School of Medicine, St Louis, Missouri 63110, USA
[30]Dipartimento di Agrobiologia e Agrochimica, Università della Tuscia, via S. Camillo de Lellis, I-01100 Viterbo, Italy
[31]AGON GmbH, Glienicker Weg 185, D-12489 Berlin, Germany
[32]Institute for Molecular Plant Sciences, Leiden University, Wassenaarseweg 64, NL-2333 AL Leiden, The Netherlands
[33]Department of Microbiology and Enzymology, Delft University of Technology, Julianalaan 67, NL-2628 BC Delft, The Netherlands
[34]Department of Biological Sciences, University of Durham, South Road, Durham, DH1 3LE, UK

*Present addresses: MWG-BIOTECH GmbH, Anzinger Strasse 7, D-85560 Ebersberg, Germany (B. F.); Centro de Biotecnología, Camagüey, Cuba (M. G.-G.); Abteilung Klinische Biochemie, Zentrum Innere Medizin, Georg-August-Universität, Robert-Koch-Strasse 40, D-37075 Göttingen, Germany, (B. K.).

The complete nucleotide sequence of *Saccharomyces cerevisiae* chromosome VII has 572 predicted open reading frames (ORFs), of which 341 are new. No correlation was found between G+C content and gene density along the chromosome, and their variations are random. Of the ORFs, 17% show high similarity to human proteins. Almost half of the ORFs could be classified in functional categories, and there is a slight increase in the number of transcription (7.0 %) and translation (5.2 %) factors when com-

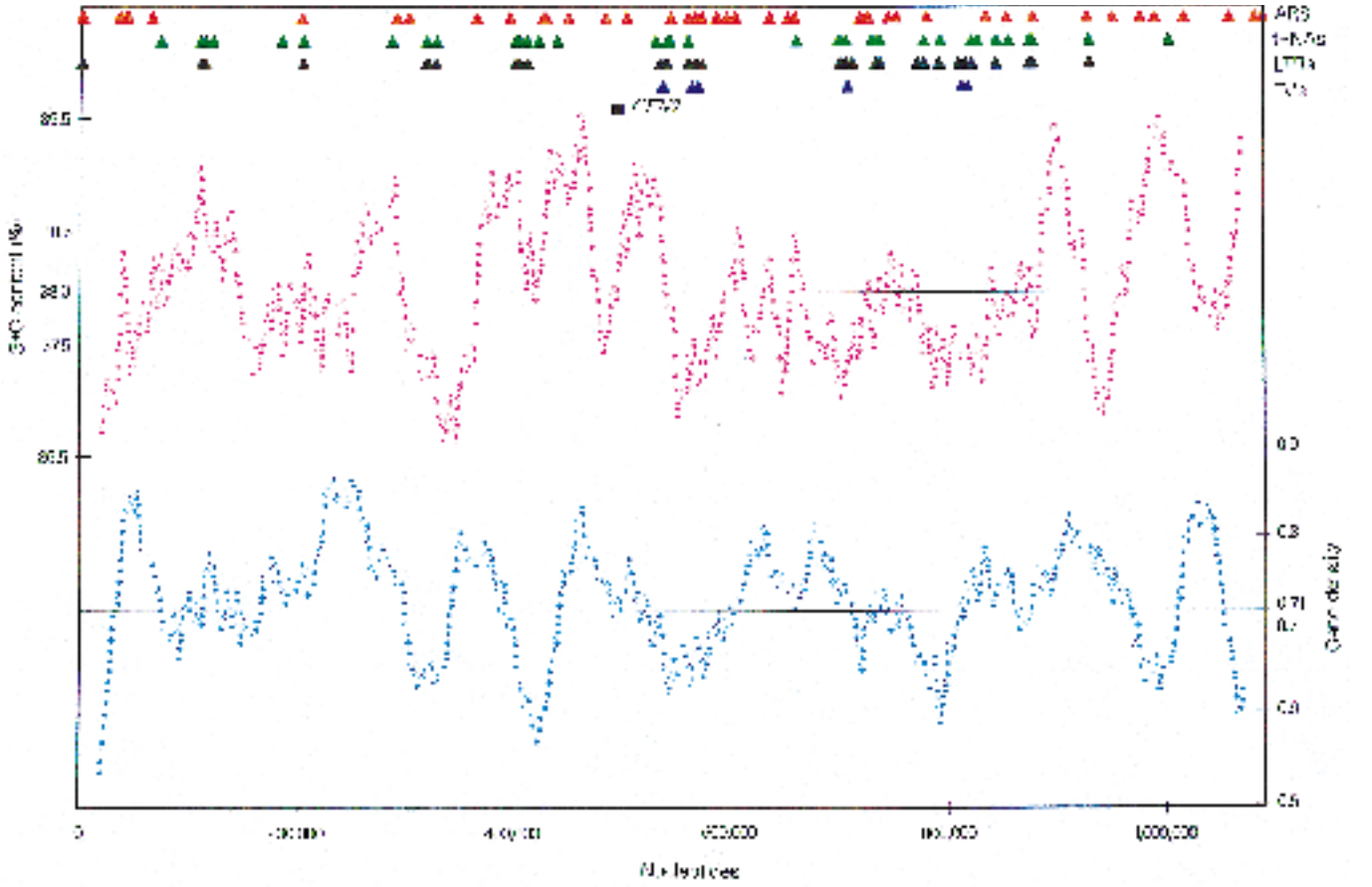


Figure 1 Top, position of genetic elements along chromosome VII. Middle, compositional variation curve; each point represents the average G+C content in a 40-kb sliding window (steps are 500 bp); the horizontal line represents the average G+C (38%). Bottom, gene density is expressed as the fraction of nucleotides within ORFs versus the total number of nucleotides in sliding windows of 40 kb (steps are 500 bp); the horizontal line represents the average gene density (0.71).

pared with the complete *S. cerevisiae* genome. Accurate verification procedures demonstrate that there are less than two errors per 10,000 base pairs in the published sequence.

Before the publication in 1992 of the yeast chromosome III sequence[1], the only available *S. cerevisiae* genome sequence of appreciable size was a contig of 24 kilobases (kb) from chromosome VII (refs 2, 3). A 60-kb physical map covering the left arm of this chromosome between *CEN7* and *TRP5* markers had been built, allowing sequencing and transcriptional mapping of nine ORFs located in the 24-kb region spanning the *PMA1* and *ATE1* loci. Analysis of these data led to the estimation of a minimum number of 5,300 expressed genes in yeast. In this centromeric region a recombination frequency of 1 cM corresponded to an average distance of 3.3 kb, compared with 2.9 kb for the complete chromosome. These extrapolations have been confirmed by the complete 1,090,936 nucleotide sequence of chromosome VII, the fourth longest in *S. cerevisiae*.

Chromosome VII contains 564 ORFs of more than 99 codons, plus eight smaller, previously identified, ORFs. Of these 572 ORFs, 19 are predicted to carry an intron at their extreme 5′ end. The *RPL6A* gene is interrupted by two introns, one of which codes for the small nuclear RNA39. Of the 572 ORFs, 152 (26.5 %) had previously been characterized biochemically, and an additional 79 (14 %) had been characterized phenotypically. Of these 231 known genes, disruption phenotypes have been reported in 140 cases, of which 37 are lethal. An additional 61 ORFs (11 %) show a high similarity (FASTA score above 300 or one third of self score) to another ORF of known function. However, if the threshold FASTA score is lowered to 150, which is a significant value in many cases, another 20 ORFs (3.5 %) could be envisaged to be of predictable function, raising the total number of ORFs of known or predictable function to 312 (54.6 %). Of the remaining ORFs, 74 (12.9 %) are similar to protein sequences of unknown function, and 186 ORFs (32.5 %) show weak or no significant similarity to any other protein sequence in the public data libraries. Finally, the expression of 63 ORFs (11 %) is questionable owing to their partial overlap with other ORFs (44) or to a combination of small size (less than 150 codons) and low codon adaptation index (CAI below 0.110)[4]. However, these criteria are not absolute as at least four expressed ORFs from chromosome VII do not meet them: *AGA2* (87 codons; CAI, 0.089), *SOH1* (127 codons; CAI, 0.096), *SPT4* (102 codons; CAI, 0.109) and *VMA21* (77 codons; CAI, 0.109). Almost 30 % of the ORFs from chromosome VII are redundant, as 166 show high similarity (FASTA score greater than 300 or one third of the self score) with other yeast genes.

The average ORF size is 468 codons, the longest (YGL195w, *GCN1*) being 2,672 codons. The average distance between ORFs located on the same strand is 514 base pairs when not containing tRNAs or long terminal repeats (LTRs) and six are longer than 2,000 bp. In the case of divergent promotors, the spacing is 553 bp, with two being longer than 2,000 bp. The mean size allocated to convergent terminators is only 304 bp, and one of these is longer than 2,000 bp. The mean G+C content of inter-ORFs regions is 33 %. All of these values are very similar to those found for the complete genome[5].

An attempt has been made to classify chromosome VII ORFs in functional categories (Table 1). Note that any given protein may belong to more than one category. The percentage of ORFs from chromosome VII in each functional category correlates with the functional distribution of ORFs within the complete genome, with a slightly higher content in the transcription (13 transcription factors) and protein synthesis (18 ribosomal proteins) categories.

The putative transmembrane spans have been computed with the KKD algorithm[6] using a rather low threshold[7] that takes into account not only the fully hydrophobic spans, but also the predicted amphipathic α-

Table 1 Functional categories of yeast ORFs from chromosome VII

Functional category	ORFs	
	Number	Percentage
Metabolism	55	9.0
Energy	19	3.1
DNA synthesis	10	1.6
Transcription	43	7.0
Protein synthesis	32	5.2
Protein destination	18	2.9
Transport facilitation (permeases)	18	2.9
Intracellular traffic	17	2.8
Cell structure	19	3.1
Organelle assembly	11	1.8
Signal transduction	5	0.8
Cell division	24	3.9
Cell rescue	12	1.9
Retrotransposons	13	2.1
Unclassified	318	51.8
Total	614	

Table 2 Similarity of yeast chromosome VII ORFs of unknown function with human proteins

Yeast ORF	Human protein
YGL150c	SNF2a transcription activator that cooperates with the oestrogen and retinoic acid receptors
YGL125w	methylenetetrahydrofolate reductase
YGL106w	calmodulin
YGL003c	probable cell-division control protein CDC 55
YGR034w	new ribosomal protein similar to the human ribosomal protein L26
YGR043c	transaldolase
YGR217w	first putative calcium channel in yeast similar to the human voltage-dependent L-type calcium channel a1 subunit
YGR231c	prohibitin, which inhibits DNA synthesis and regulates proliferation
YGR256w	phosphogluconate dehydrogenase

The similarity threshold is a FASTA score higher than 300 or higher than one third of self score.

helices which, when present in a bundle, can contribute to the formation of a polar channel within the lipid bilayer. Of the 572 ORFs, 359 (63 %) show no predicted transmembrane spans or are known to be soluble, 79 (14 %) carry at least three putative spans or are known to be membrane bound, and 134 (23 %) have one or two predicted hydrophobic α-helices, a feature which does not necessarily mean that they are membrane-bound. All ORFs have been submitted to PSORT analysis[8] to predict their subcellular localization. If we consider a high certainty score to be at least 0.8 for a given localization, and a low one to be less than 0.5 for all other possible localizations, only 76 ORFs match these criteria: 37 ORF products predicted to be in the nucleus, 16 in the endoplasmic reticulum, nine in the mitochondria, nine in the plasma membrane, two in the vacuole, two in the peroxisome, and one secreted outside the cell. For the 23 chromosome VII proteins of known subcellular localization, the PSORT prediction was correct in 14 cases (61 %). Note that three of the six ORFs that exhibited an erroneous nuclear localization were ribosomal proteins which are known to carry a nuclear localization signal to allow the biosynthesis of the ribosome particles[9].

By applying the program PYTHIA[10] to search for simple repeats within a gene, we detected at least 19 genes with regularly repeated nucleotides corresponding to repeated amino acids in the encoded proteins. Some of these regions are composed of single amino-acid repeats, including 21 aspartates (in YGL227w), 13 aspartates (YGL058w, *RAD6*), 11 and 8 glutamines (YGL066w), 18 asparagines (YGR233c), 15 asparagines (YGL014w), 10 asparagines (YGL013c, *PDR1*), 23 serines (YGR023w) or 16 serines (YGR130c). We also identified more complex repeats, such as $(T\ S/N\ ATT\ A/E\ S\ X_4)_{11}$ in YGR296w (=Y′), $(QQQP)_9$ in YGL122c (*NAB2*), $(DEEE)_3$ in YGL164w, $(AQ)_{14}$ in YGL181w (*GTS1*), $(TSSS)_9$ in YGL028c, or $(HN)_5$ in YGL178w (*MPT5*).

A systematic search for similarities with human proteins was performed for the 572 ORFs of yeast chromosome VII. A total of 95 ORFs (16.6 %) show a very significant similarity (FASTA score higher than 300 or higher than one third of self score) with human proteins, of which these 79 (13.8%) correspond to known yeast proteins whose function is often closely related to the function of the human homologue. Similarities of some of the 16 previously unknown ORFs (2.8 %) with human proteins are shown in Table 2.

Several chromosome VII ORFs show a high degree of similarity with interesting proteins from other organisms. These include: YGL236c, similar to the glucose-inhibited division (gidA) protein of the bacterium *Escherichia coli*; YGL201c, similar to the intestinal DNA replication protein of the rat *Rattus norvegicus*; and YGL054c, similar to the Cni protein necessary for anterior–posterior and dorsal–ventral patterning in the fruit fly *Drosophila melanogaster*.

Another interesting feature of *S. cerevisiae* chromosome VII is the existence of a pseudogene, which has been confirmed by direct polymerase chain reaction (PCR) sequencing on the yeast genome. This pseudogene, YGL259w, contains two frameshifts, and only one of the three ORFs is longer than 99 codons. However, all three parts show a high similarity with YIR039c, a hypothetical aspartyl proteinase. Another curious feature is the presence of three ORFs in the same frame separated by two stop codons. The rightmost one, YGL238w, corresponds to the *CSE1* chromosome segregation gene. The leftmost ORF, YGL241w, shows 17% identity over 1,053 amino acids with *CSE1*, whereas the central ORF shows no similarity with this protein. Finally, a possibly unique feature in the yeast genome is the tail-to-tail arrangement of the *SMD1* (YGR074w) and *PRP38* (YGR075c) genes, with their respective ORFs terminating on opposite strands without any intervening nucleotide between the stop codons[11]. This region, as well as the *CSE1* region, have been verified by direct PCR sequencing on the yeast genome.

Chromosome VII contains six yeast retrotransposons: three Ty1s, one Ty2, one Ty3 and a pseudo-Ty, which contains, in addition to the normal frameshift separating the Ty1A and Ty1B coding sequences, two frameshifts splitting both ORFs in two parts. Of 35 tRNA genes identified, eight are interrupted by introns. All LTRs on chromosome VII are associated with a tRNA (for review, see ref. 12), except for the two Ty5 LTRs located close to the left telomere; however, 12 tRNAs are not associated with LTRs. The positions of tRNAs, Tys, LTRs and putative ARS consensus have been compared to G+C content and gene density along the chromosome (see Fig. 1). No clear correlation could be identified regarding the location of these genetic elements. Furthermore, no statistical correlation exists between the G+C content and the gene density curves (correlation coefficient, 0.04) when highly overlapping, neighbouring sliding windows (98.7 % overlap) are used. These G+C content and gene density variations are not significant, as the same analysis performed on several random mixes of the original sequence yields similar G+C content and gene density variations. A similar graph was obtained using non-overlapping neighbouring windows along the sequence after removal of Ty and LTR elements. In this case, a good correlation was found between the G+C content of the genes and the gene density (correlation coefficient, 0.98), as well as between the G+C content of silent positions of codons in genes and the gene density (correlation coefficient, 0.66). Finally, there is no significant difference in the G+C content of the coding regions on each strand: 40.07 % G+C and 305 ORFs on the Watson strand, and 39.95 % G+C and 267 ORFs on the Crick strand.

The physical map of chromosome VII has been constructed independently from the genetic map using the meganuclease I-*Sce*I to produce *in vitro* nested fragmentations of the chromosome[13,14]. The cosmids chosen for sequencing were screened from two genomic libraries[14] and completed using a few cosmids from the physical map constructed by L. Riles (unpublished). The left telomere has been cloned[15] and the right telomeric sequence was obtained from a PCR fragment amplified from a strain carrying the pEL61 plasmid[15] integrated in the subtelomeric region.

The quality of the sequence of chromosome VII was assessed using dif-

ferent approaches. As well as partial overlaps between the regions sequenced by two laboratories, putative frameshift checking, alignment of the sequence with previously published data and a few random resequencing verifications were performed on cosmid subclones. A new method for verifying specific regions of the sequence was developed for this chromosome (G. V. *et al.*, manuscript in preparation) and applied to several other chromosomes. The extrapolation from the number of discrepancies observed in the overlaps (102,049 nucleotides, 9.4 %) to the whole sequence suggests that the nucleotide sequence of chromosome VII is 99.974 % accurate. The quality of the coding regions, where frameshifts are quite easy to check, is probably much higher than that of the intergenic regions. Indeed, the quality assessment procedure led to the correction of a total of 90 errors mainly located in the coding regions. A total of 56,344 bp (5.2 %) have been resequenced, and the comparison with the original data makes it possible to estimate that about 120 errors remain in the chromosome VII sequence. Parts of the sequence were published independently[16–30] before assembly of the contig and application of the final quality controls; several other manuscripts are in the press. □

Received 19 July 1996; accepted 11 March 1997.

1. Oliver, S.G. et al. *Nature* 357, 38-46 (1992).
2. Capieaux, E., Ulaszewski, S., Balzi, E. & Goffeau, A. *Yeast* 7, 275–280 (1991).
3. Chen, W. *et al. Yeast* 7, 287–299 (1991).
4. Dujon, B. *et al. Nature* 369, 371–378 (1994).
5. Dujon, B. *Trends Genet.* 12, 263–270 (1996).
6. Klein, P., Kanehisa, M. & Delisi, C. *Biochim. Biophys. Acta* 815, 468–476 (1985).
7. Goffeau, A., Slonimski, P., Nakai, K. & Risler, J.L. *Yeast* 9, 691–702 (1993).
8. Nakai, K. & Kanehisa, M. *Genomics* 14, 897–911 (1992).
9. Woolford, J. L. & Warner, J. R. in *The Molecular and Cellular Biology of the Yeast Saccharomyces* Vol.1 (ed. Jones, E., Pringle, J. & Broach, J.) 597–598 (Cold Spring Harbor Laboratory Press, NY, 1991).
10. Milosavljevic, A. & Jurka, J. *Comput. Adv. Biosci.* 9, 409–411 (1993).
11. Rymond, B. C. *Proc. Natl Acad. Sci. USA* 90, 848–852 (1993).
12. Olson, M. V. in *The Molecular and Cellular Biology of the Yeast Saccharomyces* Vol. 1 (ed. Jones, E., Pringle, J. & Broach, J.) 1–40 (Cold Spring Harbor Laboratory Press, NY, 1991).
13. Tettelin, H. *et al.* in *Methods in Molecular Genetics*, (ed. Adolph, K.W.) (Academic,London, 1995).
14. Thierry, A., Gaillon, L., Galibert, F. & Dujon, B. *Yeast* 11, 121–135 (1995).
15. Louis, E. & Borts, R. *Genetics* 139, 125–136 (1995).
16. Arroyo, J. *et al. Yeast* 11, 587–591 (1995).
17. Bertani, I. *et al. Yeast* 11, 1187–1194 (1995).
18. Coglievina, M. *et al. Yeast* 11, 767–774 (1995).
19. Coissac, E., Maillier, E., Robineau, S. & Netter, P. *Yeast* 12, 1555–1562 (1996).
20. Escribano, V., Eraso, P., Portillo, F. & Mazón, M.J. *Yeast* 12, 887–892 (1996).
21. Guerreiro, P. *et al. Yeast* 11, 1087–1091 (1995).
22. Guerreiro, P. *et al. Yeast* 12, 273–280 (1996).
23. Hansen, M. *et al. Yeast* 12, 1273–1277 (1996).
24. James, C.M., Indge, K.J. & Oliver, S.G. *Yeast* 11, 1413–1419 (1995).
25. Klima, R. *et al. Yeast* 12, 1033–1040 (1996).
26. Rodriguez–Belmonte, E. *et al. Yeast* 12, 145–148 (1996).
27. Skala, J., Nawrocki, A. & Goffeau, A. *Yeast* 11, 1421–1427 (1995).
28. Tizon, B. *et al. Yeast* 12, 1047–1051 (1996).
29. Vandenbol, M., Durand, P., Portetelle, D. & Hilger, F. *Yeast* 11, 1519–1523 (1995).
30. Van der Aart, Q.J.M., Kleine, K. & Steensma, H.Y. *Yeast* 12, 385–390 (1996).

Acknowledgements. This study is part of the third and final phase of the European Yeast Genome Sequencing Project carried out under the administrative coordination of A. Vassarotti and the Université Catholique de Louvain. We thank P. Mordant for accounting and help; F. Foury for support and advice; G. Gérard for statistical analyses; and our colleagues for help and discussion. We would also like to acknowledge the contribution of S. Kiefer who died in a car accident in 1994. This work was supported by the EU under the BIOTECH II programme and by the Services Fédéraux des Affaires Scientifiques, Techniques et Culturelles; the Pôles d'Attraction Inter-Universitaire; the Région Wallonne; the Fonds pour la Formation à la Recherche dans l'Industrie et l'Agriculture; the Belgian Federal Services for Science Policy; the Research Fund of the Katholieke Universiteit Leuven; the Schweizerisches Bundesamt für Bildung und Wissenschaft; the Fundaçao Calouste Gulbenkian; the International Centre for Genetic Engineering and Biotechnology of Trieste; the Bull HN Information Systems Italy; the Ministerio de Asuntos Exteriores from Spain; the Comisión Interministerial de Ciencia y Tecnología of Spain; and the Groupement de Recherche et d'Etude sur les Génomes.

Correspondence and requests for materials should be addressed to A. G. (e-mail: goffeau@fysa.ucl.ac.be).

The nucleotide sequence of *Saccharomyces cerevisiae* chromosome IX

C. Churcher[1], S. Bowman[1], K. Badcock[1], A. Bankier[2], D. Brown[1], T. Chillingworth[1], R. Connor[1], K. Devlin[1], S. Gentles[1], N. Hamlin[1], D. Harris[1], T. Horsnell[2], S. Hunt[1], K. Jagels[1], M. Jones[1], G. Lye[1], S. Moule[1], C. Odell[1], D. Pearson[1], M. Rajandream[1], P. Rice[1], N. Rowley[2], J. Skelton[1], V. Smith[2], S. Walsh[1], S. Whitehead[1] & B. Barrell[1]

[1]*The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK*
[2]*MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK*

Large-scale systematic sequencing has generally depended on the availability of an ordered library of large-insert bacterial or viral genomic clones for the organism under study. The generation of these large insert libraries, and the location of each clone on a genome map, is a laborious and time-consuming process. In an effort to overcome these problems, several groups have successfully demonstrated the viability of the whole-genome random 'shotgun' method in large-scale sequencing of both viruses and prokaryotes[1–5]. Here we report the sequence of S*accharomyces cerevisiae* chromosome IX, determined in part by a whole-chromosome 'shotgun', and describe the particular difficulties encountered in the random 'shotgun' sequencing of an entire eukaryotic chromosome. Analysis of this sequence shows that chromosome IX contains 221 open reading frames (ORFs), of which approximately 30% have been sequenced previously. This chromosome shows features typical of a small S*accharomyces cerevisiae* chromosome.

The sequence derived for chromosome IX is 439,886 nucleotides in length, and 71.6% codes for proteins or predicted proteins. There are 219 non-overlapping ORFs equal to or greater than 100 amino acids long, and a further two ORFs (YIL060W and YIL059C) that overlap; these are short, and both have a low codon adaptation index (CAI). Although it is unlikely that both are coding, one could not be selected above the other as more likely to encode a protein. A single Ty3-2 retrotransposon containing three ORFs is present on the left arm of chromosome IX (between bases 205,217 and 210,644), leaving 218 *S. cerevisiae*-derived ORFs encoded on this chromosome, of which 116 are on the Crick strand, and 102 (+ 3 transposon ORFs) are on the Watson strand. Of these, 66 (30.3%) have been sequenced previously. A further 68 (31.2%) have some similarity to genes in *S. cerevisiae* and other organisms for which some functional information is available. However, 74 (33.9%) of the predicted genes on this chromosome cannot be assigned even a putative function based on sequence similarity. These can be divided into two groups: those that show no similarity to current database entries (53, 24.3%), and those that are similar to predicted genes of unknown function (21, 9.6%). The remaining 10 (4.6%) are putative pseudogene ORFs.

The average length of a chromosome IX ORF is 476 codons, with an average of one ORF every 1,993 base pairs. The largest ORF on chromosome IX is YIL129C, which encodes a hypothetical protein of 2,376 amino acids. The YIL129C protein is similar to another hypothetical protein encoded on *Caenorhabditis elegans* chromosome III (EMBL database, accession numbers CEF21H11, U11279 and ORF F21H11.2) over a region of 2,009 amino acids. In total, 20 chromosome IX ORFs are longer than 1,000 codons. Short *S. cerevisiae* genes with no homology are difficult to detect[6]. On chromosome IX, five ORFs with less than 100 codons have been identified, but future analysis will probably reveal additional short coding regions. Less than 4% of the ORFs on chromosome IX are predicted to be spliced; eight ORFs contain introns. None of the tRNA genes on this chromosome are spliced.

Ten ORFs have been identified as contributing to five putative pseudogenes. These ORFs have very good homology to genes or predicted genes, but are separated from an adjacent ORF with homology to the same protein by internal stop codons or frameshifts. These areas have been sequenced on *S. cerevisiae* genomic DNA, and the frameshifts and stop codons confirmed. However, at least two pseudogene ORFs, YIL168W and YIL167W, probably constitute the single gene *SDL1*, which codes for a serine dehydratase[7]. This gene is not present elsewhere in the *S. cerevisiae* genome and is not essential in *S. cerevisiae*. A second putative pseudogene is highly similar to hexose transporter genes, and may also simply prove to be mutated in AB972 rather than being a true pseudogene. All putative pseudogenes are located near the telomeres of the chromosome.

For this and other *S. cerevisiae* chromosomes, the intergenic distance between two adjacent ORFs varies depending on their orientation with respect to each other. Of the adjacent ORFs on chromosome IX, 95 are arranged in tandem, 54 are divergent and 55 convergent. For ORFs arranged in tandem the intergenic distance averages 472 basepairs; ORFs with divergent promoters average 619 bp; and ORFs with convergent terminators average 421 bp. This is consistent with the greater information content required for transcription initiation and regulation than for transcription termination.

Chromosome IX also contains ten tRNAs, five solo delta elements, one solo sigma element, two sigma elements flanking a transposon, and a single solo tau element.

Reports describing the features of the other small yeast chromosomes suggest that they have used a variety of strategies to achieve a certain minimum length[8–10], and this seems to be the case for chromosome IX. Its

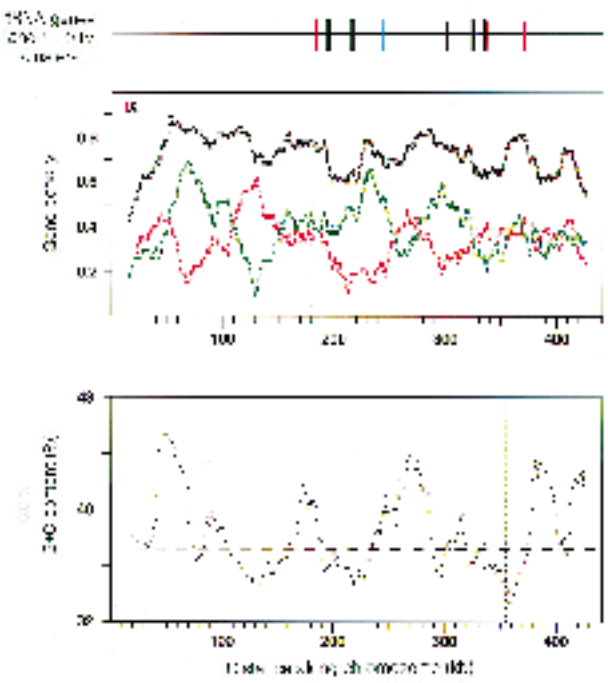


Figure 1 Overall molecular architecture of chromosome IX. The top line indicates positions of tRNA genes, solo long terminal repeat (LTR) or Ty elements (thin vertical lines) or clusters of them (thick vertical lines) along the chromosome map. The panels show variation in gene in gene density (top) and base composition (bottom) along the sequence-based map of chromosome IX (scale in kilobases from left telomere). Vertical broken lines indicate the centromere. Gene density is expressed as the probability for each nucleotide to be part of an ORF, and was calculated using sliding windows of 30 kb (steps of 0.5 kb) for the Watson strand alone (red line), the Crick strand alone (green line), and the sum of both (black line). G+C richness (%) was calculated from the silent positions of codons using a sliding window of 13 consecutive ORFs (horizontal broken line indicates average percentage G+C at silent positions of codons, 35.8%).

right telomeric region is gene poor, with only 25.3% of the sequence contributing to ORFs over approximately the last 15 kilobases of the chromosome. Chromosomes I and VI, the two smallest yeast chromosomes, also have a low coding density in their telomeric regions. Ten ORFs on chromosome IX are thought to contribute to five putative pseudogenes, all of which are located in the telomeric regions (four in the left telomeric region and one in the right). Four of these putative pseudogenes have a high degree of similarity to sequences repeated elsewhere in the yeast genome (the remaining putative pseudogene may be a mutated copy of *SDL1*, as discussed previously). Two of these occur in a 21-kb region of the chromosome IX left telomere, which is duplicated almost exactly on the chromosome X left telomere. Comparison of these two pseudogene ORFs with the equivalent regions on chromosome X shows that one region is interrupted by a frameshift on that chromosome, but the second is a single ORF. The putative pseudogene on the right arm of the chromosome is highly similar to a single ORF on chromosome XIII. The centromere[11] of chromosome IX is located towards the right end of the chromosome between bases 355,627 and 355,744.

Sequencing has revealed the definitive chromosomal position for all genes on chromosome IX. The resulting physical map correlates well with the genetic map[12] for this chromosome. Local variations in the ratio of cM to kb occur throughout this chromosome. For example, the region bounded by the genes *REV7* and *HOP1* has a significantly lower ratio of 0.3, indicating that recombination events in this area are less frequent than for the chromosome as a whole.

A cluster of genes on this chromosome occurs within its smaller, right arm between bases 399,775 and 415,615 (YIR023W to YIR032C). Six of the ten ORFs in this region are involved in the allantoin degradation pathway[13,14].

Evidence of several interchromosomal duplications occurs on chromosome IX. As well as the large region common to the left telomeres of chromosomes IX and X already described, a smaller region at the right telomere shows good homology with the telomeres of several other chromosomes. There are also several other internal chromosomal regions with long-range homology to other chromosomes. The largest of these is an area common to chromosomes IX and XIV, occurring at 89,233–186,363 and 478,568–616,076, respectively, and containing 15 homologous ORFs. Smaller interchromosomal duplications have also occurred, including a region at 230,272–258,279 repeated on chromosome V at 273,881–305,178 and a region at 46,201-69,525 repeated on chromosome XI at 617,636–639,600.

A further long-range feature previously observed for *S. cerevisiae* is a variation in the percentage G+C content along the length of its chromosomes[15]. In most cases the G+C composition for the third base of each codon has been analysed, as this is less influenced by biases in amino acid composition. For several chromosomes a periodic variation in G+C content has been observed. A plot of third-position G+C composition for chromosome IX (Fig. 1) shows that, as for other chromosomes, G+C content varies along its length. Consistent with previous analyses, the region around the centromere of the chromosome has low G+C content; the highest content occurs in peaks separated by about 100 kb.

The occurrence of first-, second- and third-codon position G+C composition was analysed for individual ORFs, as well as G+C composition in intragenic regions and the total G+C content of the chromosome (Fig. 2). This was superimposed on a chromosome map showing individual ORFs, as the usual method for plotting G+C composition uses a large window to calculate the average third-position G+C content of several adjacent ORFs, making it difficult to ascertain whether high G+C composition is a general trend in a particular region or derives from a single ORF. This method correctly predicted the peaks in third-position G+C content already described for chromosome III (data not shown).

In regions producing peaks of G+C content several adjacent reading frames show higher than average third-position G+C. Specific areas noted include 1–6, 52–67, 93–97, 175–180, 266–276, 308–313, 380–385 and 421–424 kilobases, with up to six ORFs contributing to each region. The areas of high G+C content are much more pronounced in chromosome III, with many more reading frames over a much longer distance contributing

to the two main peaks on each of the chromosome arms (results not shown).

As expected, the G+C content in intergenic regions is, on average, lower than that for coding DNA. However, it is not uniformly low, and local areas of high G+C content can be observed in some non-coding areas. The largest of these regions have been examined in detail for potential coding sequence. Although any ORFs present in these areas could conceivably be coding, they are all short, and standard methods for predicting coding sequences suggest that this is unlikely. A BLASTX search does not show any significant similarity to other *S. cerevisiae* ORFs, indicating that these areas are not remnants of ancient pseudogenes. Other possibilities have not yet been fully investigated.

Very few genes show high G+C content in position 2 of their codons. However, YIL169C, a putative glycoprotein, and YIR019C (*STA1*), a glucoamylase gene, show high second-position G+C content. Both genes code for proteins with a biased amino-acid composition. YIR019C is surrounded by regions low in G+C, but coincides with a high G+C peak on a plot of overall G+C content (results not shown), showing that a single gene can cause a peak on these plots. The role of these differences in base composition found in all yeast chromosomes has not been determined, although high G+C content might promote DNA replication or recombination.

The chromosome IX sequencing project served to highlight the difficulties involved in a whole-chromosome 'shotgun' project. This approach is described in detail in the Methods section. The purity of the starting chromosomal DNA is critical to avoid redundant sequencing effort and to minimize problems in assembly. Contamination of the chromosome IX preparation with DNA from other chromosomes resulted in a large number of single reads in the database that were not from that chromosome, but which increased the complexity of sequence assembly, slowing down both assembly and subsequent manipulation of contigs. Re-running the PFG-purified chromosome on a second pulsed-field gel has been shown to improve significantly the purity of the DNA preparation[16]. Data from the whole-chromosome 'shotgun' helped to fill the gaps in the cosmid and lambda libraries, but on their own were difficult to manage. The cosmid clones provided a framework on which to build data from the whole-chromosome shotgun. □

Methods

The chromosome IX sequencing project was initiated using a lambda clone library generated by M. Olsen and L. Riles[17]. This was an inefficient approach because of short insert sizes (20 kb) and relatively large overlaps (about 5 kb) between lambda clones. Few cosmids were available at the time, so the approach of randomly 'shotgun' sequencing the entire chromosome was attempted.

Yeast genomic DNA was electrophoresed on pulsed-field gels[18]. To increase the purity of the chromosomal preparation, yeast plugs were excised early in the run after chromosome IX had entered the gel, reducing the background DNA level. Degradation of the material of higher molecular weight occurs with prolonged running of the gel and is the likely cause of contamination of material of lower

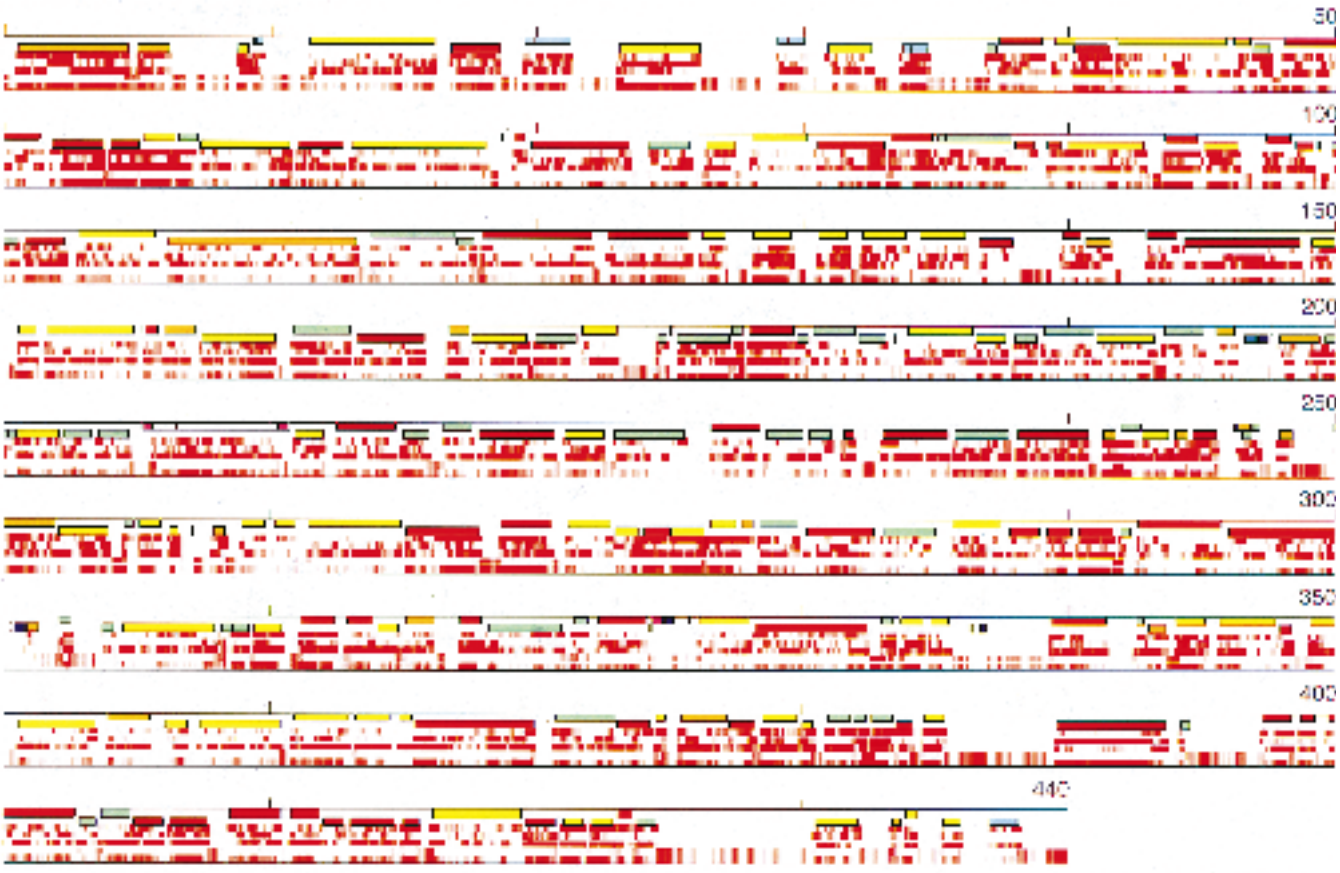


Figure 2 G+C composition of chromosome IX (drawn to scale). The top graduated line represents the chromosome split into 50-kb segments, with ORFs indicated below this as coloured boxes. ORFs located on the Watson strand are shown above those on the Crick strand. ORFs encoding previously identified genes are shown in red, those with similarities to known genes in yellow, those with similarities to hypothetical proteins in orange, and those with no significant similarities in green; pseudogene ORFs are shown in blue. tRNA genes are shown as white boxes, as are transposon-derived ORFs, with LTRs shown in dark blue (delta), turquoise (tau) and pink (sigma). Below this, variations in G+C composition (calculated using a sliding window of 200 bases) are shown as bars, with gradations of red varying from 35% to 45% (P.R., unpublished data); areas lower than 35% G+C are white, and those over 45% are red. Five bars of G+C variation are shown: the lowest bar shows total G+C content; the second shows G+C content in intergenic regions alone; and the G+C composition in each of the three bases of each codon are shown above this, with the central of the five bars representing first-position G+C, the next representing second-position G+C, and the top bar showing third-position G+C content.

molecular weight. The chromosome IX band was excised under long-wave ultraviolet transillumination to minimize DNA damage. Chromosomal DNA was purified by melting and phenol extraction, sonicated and end repaired. Two libraries were prepared: fragments 1.4–2 kb in length were cloned into M13mp18, and fragments 6–9 kb long were cloned into the phagemid vector pBS. Over 10,000 independent M13 clones were sequenced and assembled into a database using the program XBAP[19]. Sequencing strategy and methods used for sequence assembly are as described[19–25]. The lambda-clone consensus sequences were also entered into this database, which contained several thousand contigs at this stage, most of which contained a single gel read. We concluded that the chromosome IX DNA preparation was approximately 30% contaminated with DNA from other chromosomes, and that this was the source of most single-read contigs. This contamination, together with repetitive sequences in the database, caused many problems with the data assembly.

To overcome this problem, all single-read contigs were removed from the working copy of the chromosome IX database and collected in a secondary database. As further data was generated, the secondary database was periodically rescreened, and single reads were re-entered if they found matches in the primary database. This reduced the number of reads in the primary database to approximately 7,000, which represented coverage of the chromosome five times over. The database still contained several hundred contigs. At this stage a cosmid library covering most of chromosome IX became available[17]. The chromosomal 'shotgun' data were 'seeded' with reads from cosmid clones selected to give coverage over regions not previously sequenced by lambda clones. This approach also allowed the chromosome to be split up into manageable sections to solve double-stranding and compression problems. A minimal 'shotgun' of 300–500 reads was performed on each cosmid clone. Data from these cosmids were entered into the chromosome IX 'shotgun' database to contiguate the entire chromosome, and into separate cosmid databases for ease of handling, together with overlapping reads from the whole-chromosome shotgun. Each cosmid-sized project was contiguated, double stranded and all compressions were resolved.

Three regions of the chromosome remained unrepresented in either cosmid or lambda libraries: the left and right telomeres, and a region near the centre of the chromosome flanked by lambda clones 6569 and 3299. The right telomere was sequenced by primer walking using a plasmid clone[26]. The left telomere was finished using data from the whole-chromosome 'shotgun' and some primer walking from polymerase chain reaction (PCR) products generated from PFGE-purified chromosome IX DNA. The gap near the centre of the chromosome was filled using data from the whole-chromosome 'shotgun' and by sequencing a 1 kb fragment generated by PCR from genomic DNA. The gap between the lambda clones 6569 and 3299 was approximately 7 kb.

Received 24 July 1996; accepted 16 January 1997.

1. Davison, A. J. *DNA Sequence* 1, 389-394 (1991).
2. Telford, E.A.R. *et al. Virology* 189, 304-316 (1992).
3. Rawlinson, W.D. *et al. J. Virol* 70, 8833–8849 (1996).
4. Fleischmann, R. D. *et al. Science* 269, 496–512 (1995).
5. Fraser, C. M. *et al. Science* 270, 397–403 (1995).
6. Termier, M. & Kalogeropoulos, A. *Yeast* 12, 369–384 (1996).
7. Seufert, W. *Nucleic Acids Res.* 18, 3653 (1990).
8. Oliver, S. G. *et al. Nature* 357, 38–46 (1992).
9. Bussey, H. *et al. Proc. Natl Acad. Sci. USA* 92, 3809–3813 (1995).
10. Murakami, Y. *et al. Nature Genet.* 10, 261–268 (1995).
11. Fitzgerald-Hayes, M. *Yeast* 3, 187–200 (1987).
12. Mortimer, R. K. *et al.* http://genome.www.stanford.edu/saccdb/edition12.html (1995).
13. Cooper, T. G., Gorski , M. & Turoscy, V. *Genetics* 92, 383–396 (1979).
14. Turoscy, V., Chisholm, G. & Cooper, T. G. *Genetics* 108, 827–831 (1984).
15. Sharp, P. M. & Lloyd, A. T. *Nucleic Acids Res.* 21, 179–183 (1993).
16. Vaudin, M. *et al. Nucleic Acids Res.* 23, 670–674 (1995).
17. Riles, L. *et al. Genetics* 134, 81–150 (1993).
18. Schwartz, D. C. & Cantor, C. R. *Cell* 37, 67–75 (1984).
19. Dear, S. & Staden, R. *Nucleic Acids Res.* 19, 3907–3911 (1991).
20. Craxton, M. *Methods: A Companion to Methods in Enzymology* Vol. 3 (ed. Roe, B.) 20–26 (Academic, San Diego, 1991).
21. Halloran, N., Du, Z. & Wilson, R. K. in *Methods in Molecular Biology* Vol. 10, *DNA Sequencing: Laboratory Protocols* (eds Griffin, H. G. & Griffin, A. M.) 297–316 (Humana, Clifton, NJ, 1992).
22. Smith, V. *et al. Methods Enzymol.* 218, 173–187 (1993).
23. Hawkins, T. *DNA Sequence* 3, 65–69 (1992).
24. Staden, R. *Methods Mol. Biol.* 25, 27–36 (1994).
25. Staden, R. *Methods Mol. Biol.* 25, 37–67 (1994).
26. Louis, E. *Biochemica* 3, 25–26 (1995).

Acknowledgements. We thank A. Fraser for cosmid DNA preparation; the subcloning group for lambda and cosmid library preparation; the staff in the gel pouring and media kitchens for their help; the staff in the computer support and software development groups and R. Staden for software support; L. Riles and M. Olsen for lambda clones and cosmids; E. Louis for telomeric plasmids; B. Dujon for providing Fig. 1, which was prepared using unpublished software developed by C. Marck in collaboration with B. Dujon; and our colleagues for proof reading and criticism of this manuscript. This work was funded by the Medical Research Council and the Wellcome Trust.

Correspondence and requests for materials should be addressed to B.B. (e-mail: barrell@sanger.ac.uk). Clone accession numbers and other information can be found on http://www.sanger.ac.uk/yeast/home. html.

The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XII

M. Johnston[1], L. Hillier[1], L. Riles[1], other members of the Genome Sequencing Center[1], K. Albermann[2], B. André[3], W. Ansorge[4], V. Benes[4], M. Brückner[5], H. Delius[6], E. Dubois[7], A. Düsterhöft[8], K.-D. Entian[9], M. Floeth[8], A. Goffeau[10], U. Hebling[6], K. Heumann[2], D. Heuss-Neitzel[8], H. Hilbert[8], F. Hilger[11], K. Kleine[2], P. Kötter.[9], E. J. Louis[12], F. Messenguy[13], H. W. Mewes[2], T. Miosga[14], D. Möstl[8], S. Müller-Auer[2], U. Nentwich[4], B. Obermaier[15], E. Piravandi[15], T. M. Pohl[16], D. Portetelle[11], B. Purnelle[10], S. Rechmann[4], M. Rieger[5], M. Rinke[15], M. Rose[9], M. Scharfe[17], B. Scherens[18], P. Scholler[19], C. Schwager[4], S. Schwarz[19], A. P. Underwood[12], L. A. Urrestarazu[3], M. Vandenbol[11], P. Verhasselt[20], F. Vierendeels[13], M Voet[20], G. Volckaert[20], H. Voss[4], R. Wambutt[17], E. Wedler[17], H. Wedler[17], F. K. Zimmermann[14], A. Zollner[2], J. Hani[2] & J. D. Hoheisel[20]

[1] The Genome Sequencing Center, Department of Genetics, Washington University School of Medicine, 630 S. Euclid Avenue, St. Louis, Missouri 63110, USA
[2] Martinsrieder Institut für Protein Sequenzen, Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, Germany
[3] Laboratoire de Physiologie Cellulaire et de Génétique des Levures, Université Libre de Bruxelles - Campus Plaine CP 244 Boulevard du Triomphe, B-1050 Bruxelles, Belgium
[4] EMBL, Meyerhofstrasse 1, D-69117 Heidelberg, Germany
[5] Genotype, Klingenstrasse 35, D-69434 Hirschhorn and Angelhofweg 39, D-69259 Wilhelmsfeld, Germany
[6] DNA Sequencing, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 506, D-69120 Heidelberg, Germany
[7] Laboratoire de Microbiologie de l' Université Libre de Bruxelles, B-1070 Brussels, Belgium
[8] QIAGEN GmbH, Max-Volmer-Strasse 4, D-40724 Hilden, Germany
[9] Johann Wolfgang Goethe-Universität Frankfurt, Institut für Mikrobiologie, Marie-Curie-Strasse 9; Geb. N250, D-60439 Frankfurt/Main, Germany
[10] Unité de Biochimie Physiologique, Faculté des Sciences Agronomiques, Université Catholique de Louvain, Place Croix du Sud, 2-20, B-1348 Louvain-la-Neuve, Belgium.
[11] Faculté Universitaire des Sciences Agronomiques, Unité de Microbiologie, 6, avenue Maréchal Juin, 5030 Gembloux, Belgium
[12] Department of Yeast Genetics, Institute of Molecular Medicine, John Radcliffe Hospital, Oxford OX3 9DU, UK
[13] Research Institute of the CERIA-COOVI, B-1070 Brussels, Belgium
[14] Institut für Mikrobiologie und Genetik, Schnittspahnstrasse 10, D-64287 Darmstadt, Germany
[15] MediGene GmbH, Lochhamer Strasse 11, 82152 Martinsried, Germany
[16] GATC Gesellschaft für Analyse-Technik und Consulting mbH, Fritz-Arnold-Strasse 23, 78467 Konstanz, Germany
[17] AGON GmbH, Glienicker Weg 185, D-12489 Berlin, Germany
[18] Laboratorium voor Erfelijkheidsleer en Microbiologie van de Vrije Universiteit Brussel Vlaams Interuniversitair Instituut voor Biothechnologie, Departement Microbiologie, Avenue E. Gryson 1, B-1070 Brussels, Belgium
[19] Katholieke Universiteit Leuven, Laboratory of Gene Technology, Willem de Croylaan 42, B-3001 Leuven, Belgium
[20] Molecular Genetic Genome Analysis, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 506, D-69120 Heidelberg, Germany

The yeast *Saccharomyces cerevisiae* is the pre-eminent organism for the study of basic functions of eukaryotic cells[1]. All of the genes of this simple eukaryotic cell have recently been revealed by an international collaborative effort to determine the complete DNA sequence of its nuclear genome. Here we describe some of the features of chromosome XII.

The nucleotide composition of the chromosome, which is 38.48% G+C overall, and gene density vary across the chromosome. This has been observed for other yeast chromosomes[2–6] (Fig. 1). There are three main regions deficient in G+C, centred at approximately 150, 685 and 1,043 kilobases; one of these, as expected, coincides with the centromere. There is only one main peak of high G+C content, at approximately 473 kb, centred over the rDNA repeats. There does not seem to be any regularity in the variation in nucleotide composition, as may be the case for some other yeast chromosomes[3].

Like other yeast chromosomes, 72% of chromosome XII is predicted to code for protein (considering only two copies of the rDNA cluster). The sequence contains 534 open reading frames (ORFs) of 100 or more sense codons (excluding the 13 ORFs contained within yeast transposable elements), distributed roughly equally on the two strands (255 on the Watson (top) strand and 279 on the Crick (bottom) strand. The average ORF size is 485 codons. The largest gene in the chromosome, *YLR106c*, containing 4,910 codons, is the largest in the yeast genome. The average distance between ORFs is 545 base pairs for the 121 divergently transcribed genes (promoters abutting), 282 bp for the 120 convergently transcribed genes (terminators abutting), and 493 bp for the 208 genes that are transcribed in the same direction (promoter abutting terminator). Of the ORFs, 17 (3.2% of the total) contain introns, all of which are at the extreme 5′ end of the gene (except for *YLR464w*, a probable pseudogene). Two genes (*YLL057c* and *YLR388w*) may contain introns in the 5′-untranslated region of their mRNA. As expected[7], about half of the intron-containing genes (9) encode ribosomal proteins.

Only 170 (31.8%) of the genes were previously identified. Of the 364 newly identified genes, 34 (6.4% of the total genes) are obviously similar to proteins of known function, and 54 (10.1%) are weakly similar to proteins of known function. Thus a function is known or can be predicted for 48.3% of the encoded proteins. A further 69 genes (12.9%) encode proteins similar to proteins of unknown function; 207 (38.8%) of the predicted proteins are not similar to other proteins.

Included in the predicted ORFs are 55 that are 'questionable', that is, they consist of fewer than 150 codons and have a codon adaptation index (CAI)[8] of less than 0.110, or they overlap with another ORF. Of the 40 questionable ORFs that overlap with another ORF, the true gene can be predicted for 27 of these pairs, which include either a gene whose product is known (16 pairs) or whose predicted product is similar to another protein in the databases (11 pairs). There are therefore 13 overlapping ORFs that are suspect, although which of these ORFs is actually a gene awaits experimental determination.

Chromosome XII contains 22 tRNA genes, of which 7 are predicted to contain introns. Most of the tRNA genes are widely separated, although there are two clusters of three tRNA genes, each in a region of 9 kb to 13 kb (725,746–734,874 and 784,352–797,247). As expected[9], many (12) of the tRNA genes are near yeast retrotransposons (Ty elements) or their isolated long terminal repeats (LTRs). Three known small nuclear RNAs, *SNR6*, *SNR30* and *SNR34* are encoded on chromosome XII. Four of the six retrotransposons on chromosome XII are of the Ty1 type and two are Ty2 elements. There are several complete or partial 'solo' retrotransposon LTRs, including nine delta elements, four sigma elements, and a tau element.

The subtelomeric regions of chromosome XII are typical[10]. The left subtelomeric region contains a 'core X' element, and subtelomeric repeats STR-D, C, B and A, along with two tandem Y′ elements (short versions). The right subtelomeric region contains a core X element, the STR elements listed above, and 3–4 tandem Y′ (long version) elements. The sequence of the first 1.5 and the last Y′ elements were fused to give two copies of the Y′ element in the presented sequence. Proximal to the core X are shared homologies with several other telomere regions. As with several other chromosomes, both chromosome ends contain members of the *PAU/TIP/SRP* gene family[11].

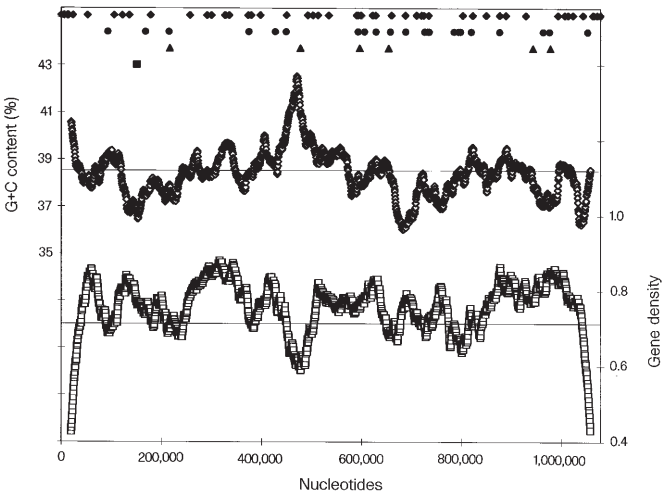


Figure 1 Top, non-coding elements of chromosome XII including autonomously replicating sequence (ARS) (filled diamonds); tRNA (filled circles); Ty element (filled triangles); centromere (filled square). Middle, G+C content as a percentage (open diamonds). Bottom, gene density (open squares).

Chromosome XII is estimated to contain 100–200 copies of the 9,137 base pair rDNA repeat[9,12–14]. Only one complete copy (the leftmost repeat in Fig. 2) and one nearly complete copy of the rDNA (the rightmost repeat in Fig. 2) are represented in the assembled sequence. In some strains these repeats seem to be interrupted by non-rDNA sequence[14].

The boundaries of the rDNA repeats are in non-transcribed regions downstream of the 35S rDNA (the left, or centromere-proximal, boundary) and 5S rDNA (the right, or centromere-distal, boundary)[15]. The structure of the left boundary of the rDNA (nucleotide 21,811 in Fig. 2) is straightforward; the right boundary[16] is more complicated. Immediately to the right of the rDNA repeats are several copies of a 3.6-kb repeat (one of which is interrupted by a Ty element) that includes the *ASP3* gene[17] and ends with a nearly complete 5S rDNA gene ($5S^{var}$ in Fig. 2). The precise number of copies of this 3.6-kb repeat in the genome is not known. The rightmost rDNA repeat ends in a 5S rDNA that adjoins a 3.6-kb repeat. Thus this rightmost rDNA repeat is lacking the 759 bp of sequence between the end of 5S rDNA and the end of the rDNA repeat (equivalent to nucleotides 30,186–30,947 in Fig. 2). The structure of the right rDNA junction differs in other yeast strains[15].

The 5S rDNA gene in the 3.6-kb repeats lacks the non-transcribed regions of the gene. It begins two nucleotides upstream of the 5′ end of 5S rRNA, and is missing the last four nucleotides of the 5S rRNA. These genes are labelled '$5S^{var}$' in Fig. 2 to indicate that they are incomplete. Immediately downstream of this gene is a run of 10 T residues that is reminiscent of the transcription termination sequence of RNA polymerase III (there are 29 T residues downstream of the 5S rDNA gene in the rDNA repeats). Because this gene seems to be missing the promoter and much of the terminator, it might represent a reverse-transcribed copy of the 5S rRNA that integrated into the genome. Nevertheless these genes produce 5S rRNA transcripts[18].

One possible explanation of the structure of the right junction is that a reverse-transcribed copy of 5S rRNA is inserted into the genome near the right border of the rDNA cluster. This gene could then have been part of a 3.6-kb duplication. It is then easy to imagine a recombination event between an intact 5S rDNA gene in one of the rDNA repeats and a $5S^{var}$ rDNA in one of the 3.6-kb repeats that generated the right rDNA junction we sequenced. Other explanations for the origin of this junction have been proffered[16].

To speed the completion of the sequence of this large chromosome, two groups collaborated on its sequencing. The rDNA repeats on chromosome XII served as a convenient point to divide the effort: the EU sequencing network[19] determined the sequence of the chromosome to the

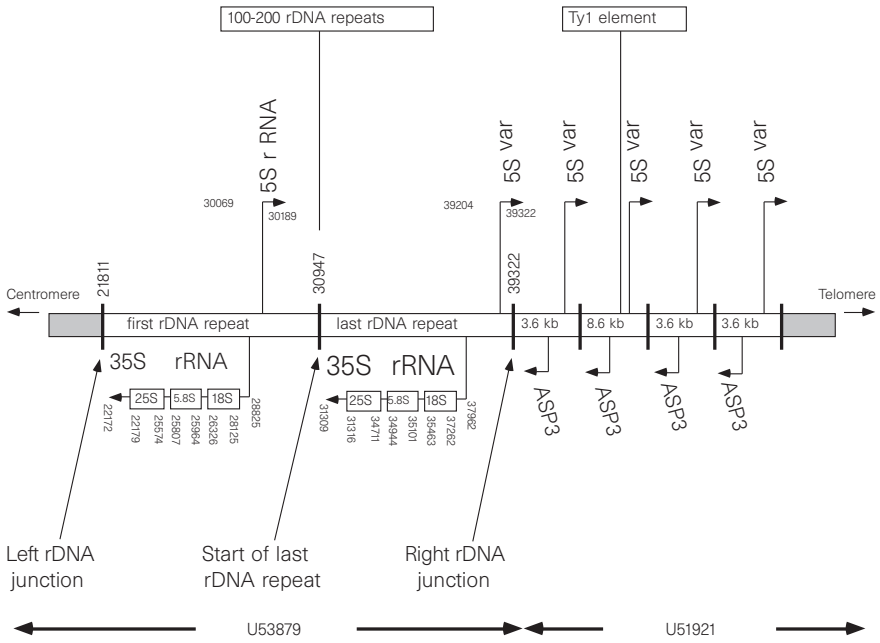


Figure 2 Diagram of the rDNA repeats and surrounding sequence, as assembled for cosmids YSCL9634 (left, GenBank accession no. U53879) and YSCL9362 (right, GenBank accession no. U51921). The numbers shown are the nucleotide coordinates for cosmid YSCL9634. The left rDNA junction (U53879 coordinate 21,811) begins at nucleotide 451,418 of chromosome XII; the right rDNA junction (U53879 coordinate 39,322) is at nucleotide 468,929. The sequence includes 1.92 rDNA repeats, representing the leftmost and rightmost copies in the genome. The remaining 100–200 rDNA repeats in the genome are represented as an insertion at coordinate 30,947. Only one complete 5S rDNA gene (in the left rDNA repeat) is included in this sequence (nucleotides 30,069–30,189); the 5S rDNA genes in the 3.6-kb repeats are variant genes. The 5S rDNA gene in the last rDNA repeat (nucleotides 39,204–39322) includes all 5′ non-translated sequences (like the normal 5S rDNA in the first repeat), but is missing sequences downstream of the 5S rRNA transcript (like the 5S rDNA genes in the 3.6-kb repeats).

left of the rDNA repeats; the sequence to the right was determined at the Washington University Genome Sequencing Center. The sequence of both strands of the entire 1,078,171 base-pair chromosome (but including only two copies of the rDNA repeats) was determined, nearly all the way to the telomeres.

The 781,865 nucleotides determined at Washington University came from 24 partly overlapping cosmid and two lambda clones[20]. The sequence of each clone was determined by a 'shotgun' strategy followed by directed sequencing[2]. The sequence of each clone was submitted to GenBank, and the entire non-overlapping sequence was assembled, analysed and annotated[2–5].

The 460,166 nucleotides to the left of the rDNA were determined by the EU network from a set of cosmid clones constructed from gel-purified chromosome XII DNA and mapped specifically for this purpose[21]. Sequencing was done by a directed approach that combines the advantages of primer walking (low redundancy) and 'shotgun' sequencing (use of a single primer)[22]. The sequence was determined from 'shotgun' sublibraries of 1-kb fragments of the cosmids that were then ordered by hybridization fingerprinting[22]. The sublibraries were arrayed on high-density filters, and sorted into smaller groups by hybridization with restriction fragments of the cosmids . Detailed mapping information was obtained by hybridizations with both oligodeoxynucleotides and pools of clone inserts amplified by the polymerase chain reaction (PCR).

The sequence of the left telomere region, including the TG_{1-3} sequence at the very end of the chromosome, was obtained from clones generated by integrating then excising a plasmid at the telomere, with capture of the flanking sequence[10]. The right telomere sequence was obtained by cycle sequencing of an anchored PCR product of the last Y′ element from a strain whose chromosome end was specifically marked by unique vector sequence[23]. The sequence of the very end of the Y′ element (about 130 bp short of the end of the chromosome) was not determined.

Only the sequence of the leftmost rDNA repeat (see Fig. 2) and about 300 nucleotides across the junction of the first and second repeat was determined. It was assembled appropriately to give the two rDNA repeats presented in Fig. 2 and in the database (GenBank accession no. U53879). The right junction sequence was not present in the cosmid closest to the rDNA on the right (YSCL9362; GenBank accession no. U51921), nor in two phage lambda clones that were mapped to this region. The structure of the junction was inferred from our ability to obtain a product of the expected size (the size of a 3.6-kb repeat) in PCR using an oligonucleotide primer in the 3.6-kb repeat (lying just to the right of $5S^{var}$) and a primer unique to the rDNA repeat (lying just to the left of 5S rDNA). Our sequence was assembled from these results, and found to match the sequence of the previously determined junction[16].

The complete, assembled, non-overlapping sequence of chromosome XII can be obtained at: http://speedy.mips.biochem.mpg.de/mips/yeast/ and http://genome-www.stanford.edu/Saccharomyces/.

Verification of 71,072 bp of sequence determined by the EU network (64,001 bp of overlaps between cosmids sequenced independently, and 7,071 bp of selected region that were resequenced) revealed five mistakes per 10 kb, but most errors were clustered in just a few regions. Only 14 differences were found in 175,891 nucleotides that were sequenced independently by both groups; six of these were sequencing errors, leading to an error frequency of only one mistake per 29 kb. The origins of the remaining eight discrepancies were determined by sequencing PCR products of the genome of the two strains used to generate the clones. Seven of the differences are due to changes that arose in the clones, presumably during propagation in *Escherichia coli*; only one results from differences between the two yeast strains (which are isogenic, but were propagated separately for many years) used to generate the two sets of clones. Thus the number of errors in the sequence is equivalent to the number of errors resulting from propagation of the DNA in *E. coli* and yeast. □

Received 31 July 1996; accepted 11 March 1997.

1. Jones, E. W., Pringle. J. R. & Broach J. R. *The Molecular and Cellular Biology of the Yeast* Saccharomyces, Vols 1–3, (Cold Spring Harbor Laboratory Press, NY, 1991–1996).
2. Johnston, M. *et al. Science* 265, 2077–2082 (1994).

3. Dujon, B. *et al. Nature* 369, 371–378 (1994).
4. Feldmann, H. *et al. EMBO J.* 13, 5795-5809 (1994).
5. Galibert, A. *et al. EMBO J.* 15, 2031–2049 (1996).
6. Sharp, P. & Lloyd, A. *Nucleic Acids Res.* 21, 179–183 (1993).
7. Rodriguez-Medina, J. R. & Rymond, B. C. *Mol. Gen. Genet.* 243, 532–539 (1994).
8. Sharp, P. M. & Li, W. H. *Nucleic Acids Res.* 15, 1281–1295 (1987).
9. Olson, M. V. in *The Molecular and Cellular Biology of the Yeast* Saccharomyces, Vol. 1 (ed. Broach J. R., Pringle. J. R. & Jones, E. W.) 1–39 (Cold Spring Harbor Laboratory Press, NY, 1991).
10. Louis, E. J. & Borts, R. *Genetics* 139, 125–136 (1995).
11. Viswanathan, M., Muthukumar, G., Lenard, J. & Cong, Y. S. *Gene* 148, 149–153 (1994).
12. Pasero, P. & Marilley, M. *Mol. Gen. Genet.* 236, 448–452 (1993).
13. Chindamporn, A., Iwaguchi, S., Nakagawa, Y., Homma, M. & Tanaka, K. *J. Gen. Microbiol.* 139, 1409–1415 (1993).
14. Rustchenko, E. P. & Sherman, F. *Yeast* 10, 1157–1171 (1994).
15. Zamb, T. & Petes, T. D. *Cell* 28, 355–364 (1982).
16. McMahon, M. E., Stamenkovich, D. & Petes, T. D. *Nucleic Acids Res.* 12, 8001–8016 (1984).
17. Kim, K. W., Kamerud, J. Q., Livingston, D. M. & Roon, R. J. *J. Biol. Chem.* 263, 11948–11953 (1988).
18. Piper, P. W., Lockheart, A. & Patel, N. *Nucleic Acids Res.* 12, 4083–4096 (1984).
19. Vassarotti, A. *et al. J. Biotechnol.* 41, 131–137 (1995).
20. Riles, L. *et al. Genetics* 134, 81–150 (1993).
21. Scholler, P., Schwarz, S. & Hoheisel, J. D. *Yeast* 11, 659–666 (1995).
22. Scholler, P. *et al. Nucleic Acids Res.* 23, 3842–3849 (1995).
23. Louis, E. J. *Biochemica* 3, 25–26 (1995).

Acknowledgements. We thank J. Warner for help with interpreting the rDNA sequence. This work was supported by the NIH National Center for Human Genome Research and the European Commision; the Belgian Federal Services for Science Policy (D.W.T.C.) and the Service Fédéraux des Affaires Scientifiques, Techniques et Culturelles, Pôles d'attraction Inter-universitaire, the Region de Bruxelles-Capitale, Region Wallonne, the Research Fund of the Katholieke Universiteit Leuven and The Wellcome Trust.

Correspondence and requests for materials should be addressed to M.J. (e-mail: mj@genetics.wustl.edu/)

The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XIII

S. Bowman, C. Churcher, K. Badcock, D. Brown, T. Chillingworth, R. Connor, K Dedman, K. Devlin, S. Gentles, N. Hamlin, S. Hunt, K. Jagels, G. Lye, S. Moule, C. Odell, D. Pearson, M. Rajandream, P. Rice, J. Skelton, S. Walsh, S. Whitehead & B. Barrell

The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK

Systematic sequencing of the genome of *Saccharomyces cerevisiae* has revealed thousands of new predicted genes and allowed analysis of long-range features of chromosomal organization. Generally, genes and predicted genes seem to be distributed evenly throughout the genome, having no overall preference for DNA strand. Apart from the smaller chromosomes, which can have substantially lower gene density in their telomeric regions[1–3], there is a consistent average of one open reading frame (ORF) approximately every two kilobases. However, one of the most surprising findings for a eukaryote with approximately 6,000 genes was the amount of apparent redundancy in its genome. This redundancy occurs both between individual ORFs and over more extensive chromosome regions, which have been duplicated preserving gene order and orientation[4–6]. Here we report the entire nucleotide sequence of chromosome XIII, the sixth-largest *S. cerevisiae* chromosome, and demonstrate that its features and organization are consistent with those observed for other *S. cerevisiae* chromosomes. Analysis revealed 459 ORFs, 284 have not been identified previously. Both intra- and interchromosomal duplications of regions of this chromosome have occurred.

Chromosome XIII of *S. cerevisiae* is 924,430 base pairs long, and contains 459 ORFs. Eight of these are TyA and TyB ORFs from four Ty1 retrotransposons present on chromosome XIII in strain AB972, two of which are located on each arm of the chromosome, and these are excluded from further analyses. The average gene density on this chromosome is one ORF for every 1,997 base pairs of DNA, which correlates well with that observed for other *S. cerevisiae* chromosomes, with 74.2% of DNA on this chromosome contributing to ORFs. An average chromosome XIII ORF is 494 codons long.

Of the 451 *S. cerevisiae* ORFs on chromosome XIII, 167 (37.0%) encode previously identified proteins. A further 281 (62.3%) predicted genes have not been previously sequenced; 121 (26.8%) of these ORFs have similarities to genes for which some biochemical information is available. However, several of this category of ORF have their best protein similarity to a protein of unknown function. A total of 160 ORFs (35.5%) encode predicted proteins that are not significantly similar to proteins of known function. Because of the rapidly advancing progress of other systematic sequencing projects, many of these ORFs have homology to hypothetical proteins both in yeasts and higher organisms. A total of 51 predicted genes have similarity only to predicted proteins of unknown function. Although the majority are most similar to another *S. cerevisiae* hypothetical protein (50.9%), several have their best homology to an ORF identified in systematic sequencing of the yeast *Schizosaccharomyces pombe*[7] (11.8%), or to predicted proteins in the nematode *Caenorhabditis elegans*[8] (17%). Thus they are members of gene families whose function is currently unknown. There were no significant protein sequence similarities for 109 ORFs, of which 11 are thought to be questionable ORFs based on their length, codon adaptation index (CAI) value and positional base preferences.

During the systematic sequencing of other chromosomes, several putative pseudogenes were identified[1,9]. These consisted of ORFs separated by a stop codon or frameshift from upstream or downstream sequences that shared a common homology to a single *S. cerevisiae* ORF. Most of these pseudogenes identified occur close to the telomeres of chromosomes. Three ORFs on chromosome XIII (YMR084W, YMR085W and YMR326C) have been classified as putative pseudogene ORFs. Of these, only YMR326C is located close to one of the chromosome telomeres; all three have strong similarity to sequences found elsewhere in the *S. cerevisiae* genome. These frameshifts have been confirmed by sequencing genomic DNA.

The average intergenic distance between adjacent ORFs depends on their relative orientation. This is certainly the case on chromosome XIII, in which 204 ORFs are arranged in tandem with an average intergenic distance of 450 base pairs. Of these, 110 are divergent and are an average 616 bp apart, and 111 are convergent and an average of 260 bp apart. This is consistent with a greater sequence requirement for the regulation of gene expression from promoter elements than for transcription termination.

Of the 451 ORFs on chromosome XIII, 24 (5.3%) are predicted to contain introns. There seems to be no preference for DNA strand, with 229 genes coded on the Watson strand and 222 on the Crick strand. There is no evidence of any significant clustering of related genes. However, there are several instances in which two very similar ORFs occur close to one another in tandem; for example, YMR169C and YMR170C/ALD2 (aldehyde dehydrogenases), and YMR006C and YMR008C/PLB1 (lysophospholipases).

The longest ORF on chromosome XIII is *HFA1*, (which is homologous to *FAS3*), a putative acetyl-CoA carboxylase that had been sequenced previously[10] (2,123 codons). A total of 39 ORFs on this chromosome are more than 1,000 codons in length. *S. cerevisiae* genes of less than 100 codons with no homology are difficult to detect[11]. On chromosome XIII, 10 ORFs shorter than 100 amino acids in length have been identified. The smallest of these is YMR248C, which is just 55 amino acids long, and may be spliced to a second small ORF immediately upstream. The smallest ORF on this chromosome that encodes a previously characterized protein is *COX7*, which is 59 amino acids long and encodes cytochrome oxidase polypeptide VII (ref.12).

Chromosome XIII encodes 21 predicted tRNA genes, of which six are spliced. In addition to the four Ty1 retrotransposons, several long terminal repeat (LTR) sequences are present, providing evidence of previous trans-

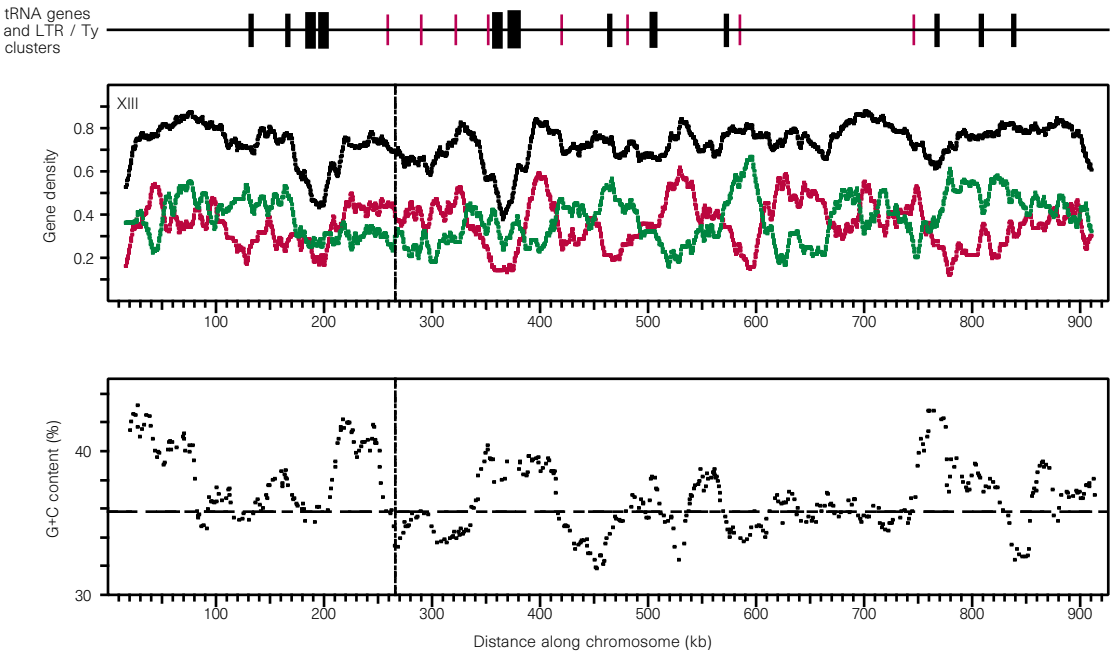


Figure 1 Overall molecular architecture of chromosome XIII. Variation in gene density (top) and base composition (bottom) along the sequence-based map of chromosome XIII (scale in kilobases from left telomere). Veritical broken lines indicate the position of the centromere. Gene density is expressed as the probability for each nucleotide to be part of an ORF. It was calculated using sliding windows of 30 kb (steps of 0.5 kb) for the Watson strand alone (red line), the Crick strand alone (green line), and the sum of both strands (black line). Percentage of G+C was calculated from the silent positions of codons using a sliding window of 13 consecutive ORFs (horizontal broken line indicates average percentage G+C% at silent positions of codons as 35.8). Top line, positions of tRNA genes, solo LTR or Ty elements (thin small vertical lines), or clusters of them (thick small vertical lines) along the chromosome map.

position events on this chromosome. Of the 30 LTRs (whole or partial) identified on this chromosome, 29 are in close proximity to tRNA genes. There are 24 delta elements, with eight of these flanking the four retrotransposons and 16 solo elements. A further five LTRs resemble tau elements, and there is evidence of a single sigma element.

The telomeres of chromosome XIII are highly similar to the telomeres of several other *S. cerevisiae* chromosomes. Adjacent to the terminal $C_{1-3}A$ telomeric repeat at the left telomere is a Y′ element, which is separated from a core X element by the subtelomeric repeats STR-D, STR-C, STR-B and STR-A[13]. The right telomere of chromosome XIII conforms to this structure but does not contain a Y′ element. The right telomere shares a region of 4 kb in which the sequence is almost identical to that found at the right telomere of chromosome XV (EMBL database, accession no. SC23472). The centromere of chromosome XIII is located between bases 268,031 and 268,150; it conforms to the consensus sequence derived from the centromeres of other *S. cerevisiae* chromosomes[14].

Comparison of the positions of genes on chromosome XIII with their corresponding genetically mapped loci[15] shows that, as for most *S. cerevisiae* chromosomes, the two are generally in agreement. Several genes are incorrectly positioned on the genetic map in comparison with the sequence-derived map; for example, *van1* is much closer to the left telomere of chromosome XIII than expected. Small inversions are evident, for example between *UPF1* and *ADH3*, but no gross discrepancies have occurred as observed for chromosome XI (ref. 16). *MEL6* has been mapped to this chromosome, but in the *S. cerevisiae* strain S288C, from which AB972 was derived, this locus is missing. Also, the locus *SUP8* has been only tentatively assigned to the tRNA(Tyr) at base 837928 as a second tRNA(Tyr) is present on the left arm of the chromosome.

Chromosome XIII contains both intra- and interchromosomal duplications of both single ORFs and more extensive regions. The largest intrachromosomal duplication consists of two regions approximately 40 kb to 50 kb in length containing six homologous genes in the same order and orientation with respect to each other; the duplication has occurred between the left and right arms of the chromosome. This encompasses the region from base 32,334 to 73,917 and 790,207 to 840,147. Smaller instances of tandem gene duplication have also occurred within chromosome XIII; for example, YML125C and YML124C (*TUB3*) are similar to YML087C and YML085C (*TUB1*) on the short arm of the chromosome.

The largest interchromosomal duplication observed has occurred between a 200 kb region on the right arm of chromosome XIII and an equivalent region on the left arm of chromosome XI (in the opposite orientation). On chromosome XIII the region spans bases 303,238 to 502,733 and is bounded by ORFs YMR016C and YMR118C, on chromosome XI the coordinates of this repeat occur between bases 179,672 and 357,489. Of a total of approximately 90 genes, 17 show significant similarity and are in the same order and orientation on each chromosome. Further evidence of large scale duplications exists, including two separate regions on both chromosome IV and XVI.

Chromosome XIII has been analysed for variations in base composition, as described in ref. 9. The percentage G+C content in the third position of each codon varies throughout the length of chromosome XIII (Fig. 1), as observed through long-range analysis of other *S. cerevisiae* chromosomes. Detailed analysis at the level of individual ORFs (Fig. 2) shows that, as for chromosome IX, regions of high third-position G+C are shorter and contain fewer ORFs that regions of chromosome III that are rich in G+C. Comparison with chromosome IX shows that regions of high G+C are slightly less evident on chromosome XIII. This is also seen when plotting total G+C content for these two chromosomes: peaks of high G+C are generally lower in magnitude for chromosome XIII (results not shown).

Local areas of high G+C composition have also been observed in several of the intergenic regions of this chromosome, for example regions at approximately 119, 306, 541 and 653 kb are G+C rich. It is possible that these intergenic regions of high G+C may contain previously unidentified, small *S. cerevisiae* genes. These areas do not appear to correlate with sequences of high coding potential. However, this does not exclude the

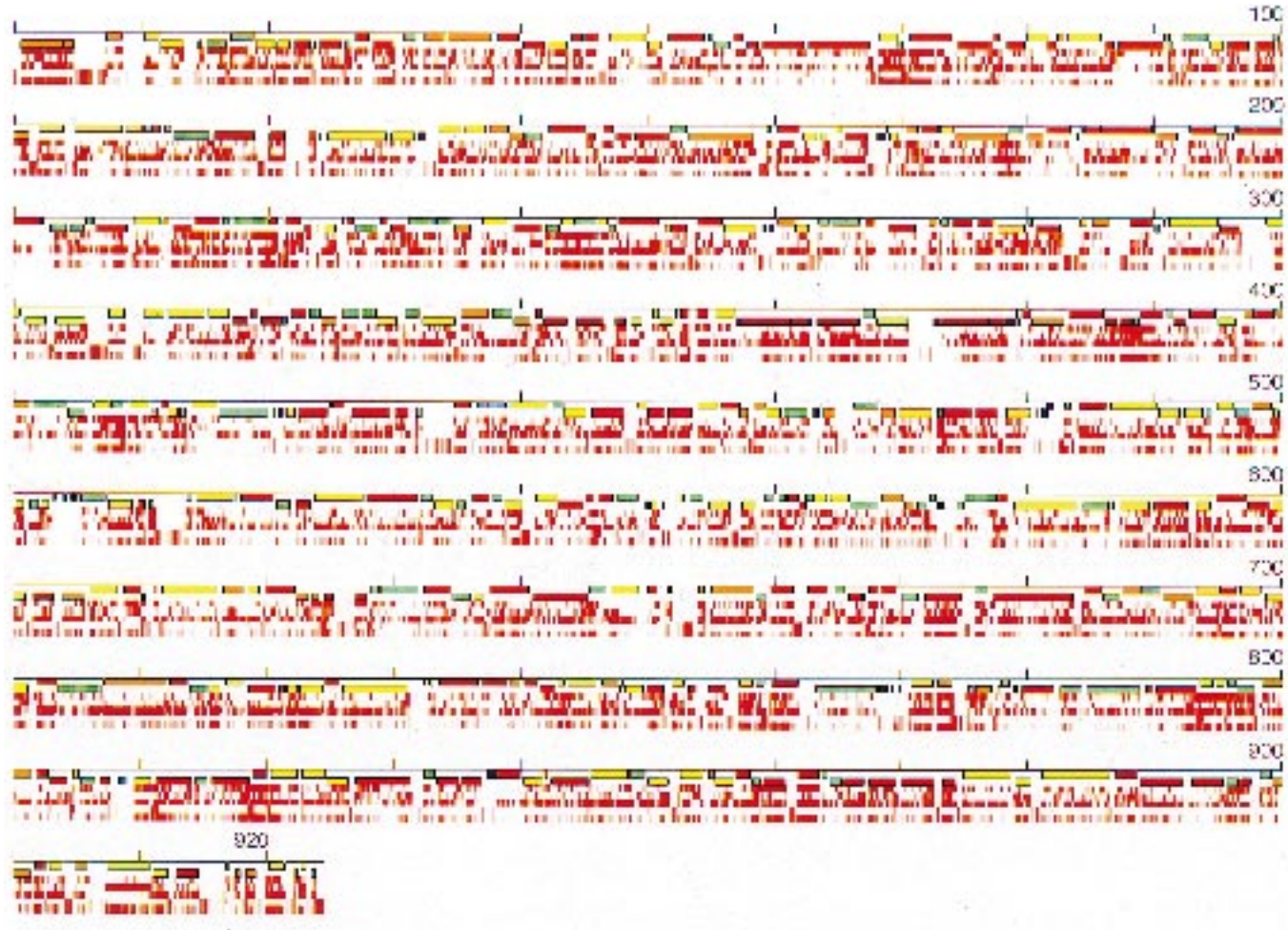


Figure 2 G+C composition of chromosome XIII. Chromosome XIII and its features are drawn to scale. The top graduated line represents the chromosome split into segments of 100 kb, with ORFs indicated below this as coloured boxes. ORFs located on the Watson strand are shown above those on the Crick strand. ORFs encoding previously identified genes are shown in red, those with similarities to hypothetical proteins in orange, and those with no significant similarities in green; pseudogene ORFs are shown in blue. tRNA genes and transposon-derived ORFS are shown in white boxes, with LTRs shown in dark blue (delta), turquoise (tau) and pink (sigma). Below this, variations in G+C composition (calculated using a sliding window of 200 bases) are shown as bars, with gradations of red varying from 35% to 45% G+C content (P. R., unpublished data). Areas of less than 35% are shown in white, with those over 45% red. Five bars of G+C variation are shown, the lowest bar shows total G+C content; above this a second shows G+C content in intergenic regions alone. The G+C composition in each of the three bases of each codon are shown above this, with the central of the five bars representing first position G+C, the next representing second position G+C, and the top bar showing third position G+C.

possibility that coding sequences exist in these areas. No similarities to other *S. cerevisiae* ORFs were found using the program BLASTX, indicating that these areas rich in G+C are not pseudogene remnants. Other possibilities for these G+C peaks in intergenic regions have not yet been fully investigated.

Few ORFs show a high incidence of high G+C content in the second position of their codons. The most pronounced example on chromosome XIII is YMR317W, which encodes a protein rich in serine and threonine residues, and is located close to the right telomere of the chromosome.

Analysis of other *S. cerevisiae* chromosomes has demonstrated that, in general, areas high in G+C content correlate with high gene density, and that regions around chromosome centromeres and telomeres appear to be G+C poor[4–6,16]. This is also the case for chromosome X111. However, the significance of these areas of high and low G+C content is not yet clear.

Two of the genes located on chromosome XIII have human homologues that have been implicated in certain forms of cancer. *MLH1* encodes a DNA mismatch repair protein that is very similar to a human protein defective in some forms of hereditary non-polyposis colon cancer[17]. In *S. cerevisiae*, null mutants of this gene are viable, but show an elevated rate of spontaneous mutations and an increased instability of simple repeat sequences[18,19]. A second gene, *SGS1*, is homologous to the human *BLM* gene, which is involved in Bloom's syndrome[20]. Both the human and yeast genes encode DNA helicases. Mutations in the human gene confer a predisposition to many types of cancer and also cause other clinical defects[20]. In *S. cerevisiae*, null mutants are viable, but again show genomic instability[21,22]. Future analysis of these and other *S. cerevisiae* genes should assist the understanding of the molecular mechanisms underlying many human diseases.

Chromosome XIII is the largest *S. cerevisiae* chromosome to be sequenced by a single laboratory. Analysis shows that its features are typical of large *S. cerevisiae* chromosomes. □

Methods

Sequencing. Chromosome XIII was sequenced using a cosmid-based strategy that uses DNA isolated from the S288C-derived strain AB972. Methods used for sequence generation and assembly have been described in detail elsewhere[8,23]. Cosmids were chosen from the map generated by L. Riles and M. Olson (personal communication), giving rise to two large contigs with a central gap. This was filled using lambda clones L-6543 and L-6223. The two chromosome telomeres, pEL161H (left) and pEL175H (right), were provided as plasmid clones[24] by E. Louis. Two further lambda clones, L-4987 and L-7056, were sequenced to complete the left end of the chromosome. Sequence from pEL161H overlapped L-4987, so no further gap filling was required at the left telomere of chromosome XIII. However, at the right telomere a gap between cosmid clone 9924 and plasmid pEL175H remained, and

was filled using long-range polymerase chain reaction (PCR) from *S. cerevisiae* genomic DNA. The sequence generated for this chromosome extends into the $C_{1-3}A$ telomeric repeat sequences on both chromosome arms, although the exact number of these repeats has not been determined. Sequencing was considered to be finished when each base had been sequenced on both strands and all ambiguities had been resolved.

Analysis. For each completed clone, a consensus of the nucleotide sequence was generated in the Staden sequence assembly package XBAP[25], flanked by short regions of sequence overlapping neighbouring clones. This sequence was analysed primarily within the DIANA (Display and Analyse) package (T. Horsnell and B. B., unpublished), a sequence editor with a graphical interface. ORFs equal to or greater than 100 codons in length were marked and trimmed to their first methionine. Each ORF was screened against the SWIR database, a non-redundant compilation of the protein databases Swiss-Prot[26], TrEMBL[27] and WormPep, using the program FASTA[28] with limited optimization. The consensus sequence for each clone was screened against SWIR using BLASTX[29], and EMBL/EMNEW using BLASTN[29], to detect small ORFs less than 100 amino acids in length, other genome features, and local similarity. Some features were specifically identified; Prosite[30] amino-acid motifs (regular expression searching), transposon LTRs (GCG Wordsearch/Segments) and tRNAs (tRNA scan). Individual annotated clones were submitted to the EMBL database within days of being finished. The complete chromosomal sequence was built from overlapping clones and also submitted to the EMBL database as a single record (accession no. SCCHRXIII, Z271257).

Received 26 July 1996; accepted 11 March 1997.

1. Bussey, H. *et al. Proc. Natl Acad. Sci. USA* 92, 3809–3813 (1995).
2. Oliver, S. G. *et al . Nature* 357, 38–46 (1992).
3. Murakami, Y. *et al . Nature Genet.* 10, 261–268 (1995).
4. Feldmann, H. *et al. EMBO J.* 13, 5795–5809 (1994).
5. Johnston, M. *et al. Science* 265, 2077–2082 (1994).
6. Galibert, F. *et al. EMBO J.* 15, 2031–2049 (1996).
7. http://www.sanger.ac.uk/yeast/pombe.html
8. Wilson, R. *et al. Nature* 368, 32–38 (1994).
9. Churcher, C. *et al, Nature (*this issue).
10. Kearsey, S.E. *DNA Sequence* 4, 69–70 (1993)
11. Termier, M. & Kalogeropoulos, A. *Yeast*, 12, 369–384 (1996)
12. Calder, K. M. & McEwen, J. E. *Nucleic Acids Res.* 18, 1632 (1990).
13. Louis, E. J., Naumova, E. S., Lee, A., Naumov, G. & Haber, J. E. *Genetics* 136, 789–802 (1994).
14. Fitzgerald–Hayes, M. *Yeast* 3, 187–200 (1987).
15. Mortimer, R.K., Cherry, J.M., Dietrich, F.M., Riles, L., Olson, M.S. & Botstein, D. http.//genome.www.stanford.edu/saccdb/edition12.html (1995).
16. Dujon, B. *et al. Nature* 369, 371–378 (1994).
17. Papadopoulos, N. *et al. Science* 263, 1625–1629 (1994).
18. Strand, M., Earley, M.C., Crouse, G.F. & Petes, T.D. *Proc. Natl Acad. Sci. USA* 92, 10418–10421 (1995).
19. Prolla, T. A., Christie, D. M. & Liskay, R. M. *Mol. Cell. Biol.* 14, 407–415 (1994).
20. Ellis, N.A. *et al. Cell* 83, 655–666 (1995).
21. Gangloff, S., McDonald, J. P., Bendixen, C., Arthur, L. & Rothstein, R. *Mol. Cell. Biol.* 14, 8391–8398 (1994).
22. Watt, P.M., Louis, E.J., Borts, R.H. & Hickson, I.D. *Cell* 81, 253–260 (1995).
23. Smith, V. *et al. Methods Enzmol.* 218, 173–187 (1993).
24. Louis, E. J. *Biochemica* 3, 25–26 (1995).
25. Dear, S. & Staden, R. *Nucleic Acids Res.* 19, 3907–3911 (1991).
26. Bairoch, A. *Nucleic Acids Res.* 19, 2247–2249 (1991).
27. Bairoch, A. & Apweiler, R. *Nucleic Acids Res.* 24, 21–25 (1996).
28. Pearson, W.R. & Lipman, D.J. *Proc. Natl Acad. Sci. USA* 85, 2444–2448 (1988).
29. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D. J. *J. Mol. Biol.* 215 403–410 (1990).
30. Bairoch, A., Butcher, P. & Hoffman, K. *Nucleic Acids Res.* 24, 189–196 (1995).

Acknowledgements. We thank A. Fraser for cosmid DNA preparation; M. Jones and the subcloning group for library preparation; the staff in the gel-pouring and media kitchens for their help; the computer support and software development groups and R. Staden for software support; we thank L. Riles, M. Olsen and E. Louis for gifts of cosmid, lambda and plasmid clones; B. Dujon for providing Fig. 1 using unpublished software developed in collaboration with C. Marck, and D. Harris, J. Sulston and K. Plucknett for critical reading of the manuscript. This work was funded by the Wellcome Trust.

Correspondence and requests for materials should be addressed to B.B. (e-mail: barrell@sanger.ac.uk).
Clone accession numbers and other information can be found on http://www.sanger.ac.uk/yeast/home. html.

The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XIV and its evolutionary implications

P. Philippsen[1,2], K. Kleine[3], R. Pöhlmann[1], A. Düsterhöft[4], K. Hamberg[2], J. H. Hegemann[2], B. Obermaier[5,6], L. A. Urrestarazu[7], R. Aert[8], K. Albermann[3], R. Altmann[1], B. André[7], V. Baladron[9], J. P. G. Ballesta[10], A.-M. Bécam[11], J. Beinhauer[2], J. Boskovic[10], M. J. Buitrago[9], F. Bussereau[12], F. Coster[13], M. Crouzet[14], M. D'Angelo[15], F. Dal Pero[15], A. De Antoni[15], F. Del Rey[9], F. Doignon[14], H. Domdey[5], E. Dubois[16], T. Fiedler[2], U. Fleig[2], M. Floeth[4], C. Fritz[4], C. Gaillardin[17], J. M. Garcia-Cantalejo[10], N. N Glansdorff[16], A. Goffeau[13], U. Gueldener[2], C. Herbert[11], K. Heumann[3], D. Heuss-Neitzel[4], H. Hilbert[4], K. Hinni[1], I. Iraqui Houssaini[7], M. Jacquet[12], A. Jimenez[10], J.-L. Jonniaux[13], L. Karpfinger[3], G. Lanfranchi[15], A. Lepingle[17], H. Levesque[17], R. Lyck[2], M. Maftahi[17], L. Mallet[12], K. C. T. Maurer[18], F. Messenguy[16], H. W. Mewes[3], D. Möstl[4], F. Nasr[11], J.-M. Nicaud[17], R. K. Niedenthal[2], D. Pandolfo[15], A. Piérard[16], E. Piravandi[5], R. J. Planta[18], T. M. Pohl[19], B. Purnelle[13], C. Rebischung[1], M. Remacha[10], J. L. Revuelta[9], M. Rinke[5], J. E. Saiz[9], F. Sartorello[15], B. Scherens[16], M. Sen-Gupta[2], A. Soler-Mira[10], J. H. M. Urbanus[18], G. Valle[15], L. Van Dyck[13], P. Verhasselt[8], F. Vierendeels[16], S. Vissers[7], M. Voet[8], G. Volckaert[8], A. Wach[1], R. Wambutt[20], H. Wedler[20], A. Zollner[3] & J. Hani[3]

[1]Institute for Applied Microbiology, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland
[2]Justus-Liebig-Universität Giessen, Institut für Mikro- und Molekularbiologie, Frankfurter Strasse 107, D-35392 Giessen, Germany
[3]Martinsrieder Institut für Protein Sequenzen, Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, Germany
[4]QIAGEN GmbH, Max-Volmer-Strasse 4, D-40724 Hilden, Germany
[5]Laboratorium für Molekulare Biologie, Genzentrum der LMU München, Feodor-Lynen-Strasse 25, D-811377 München, Germany
[6]MediGene GmbH, Lochhamer Strasse 11, D-82152 Martinsried, Germany
[7]Université Libre de Bruxelles, Physiologie Cellulaire et Génétique des Levures, Boulevard du Triompe CP244, B-1050 Bruxelles, Belgium
[8]Katholieke Universiteit Leuven, Laboratory of Gene Technology, Willem de Croylaan 42, B-3001 Leuven, Belgium
[9]Departamento de Microbiologia y Genética, Universidad de Salamanca, Avenida del Campo Charro s/n, E-37007 Salamanca, Spain
[10]Centro de Biologia Molecular, CSIC & UAM, Cantoblanco, E-28049 Madrid, Spain
[11]Centre de Génétique Moléculaire, Laboratoire propre du CNRS associé à l'Université Pierre et Marie Curie, F-91198 Gif-sur-Yvette, France
[12]Université Paris-Sud, Institut de Génétique et Microbiologie, Laboratoire Information Génétique et Développement, Bât. 400, F-91405 Orsay Cedex, France
[13]Unité de Biochimie Physiologique, Université Catholique de Louvain, Place Croix du Sud 2/20, B-1348 Louvain-la-Neuve, Belgium
[14]LBMS, Université de Bordeaux 2, UPR CNRS 9026, BP 64, 146 rue Léo Saignat, F-33076 Bordeaux Cedex, France
[15]Department of Biology, CRIBI Biotechnology Centre, University of Padova, via Trieste, 75, I-35121 Padova, Italy
[16]CERIA-COOVI, Avenne E. Gryson 1, B-1070 Brussels, Belgium
[17]Institut National Agronomique Paris-Grignon, Laboratoire de Génétique Moléculaire et Cellulaire, Centre de Biotechnologies Agro-Industrielles, F-78850 Thiverval-Grignon, France
[18]Department of Biochemistry and Molecular Biology, IMBW, BioCentrum Amsterdam, Vrije Universiteit de Boelelaan 1083, NL-1081 HV Amsterdam, Netherlands
[19]GATC-Gesellschaft für Analyse-Technik und Consulting mbH, Fritz-Arnold-Strasse 23, D-78467 Konstanz, Germany
[20]AGON GmbH, Glienicker Weg 185, D-12489 Berlin, Germany

In 1992 we started assembling an ordered library of cosmid clones from chromosome XIV of the yeast *Saccharomyces cerevisiae*. At that time, only 49 genes were known to be located on this chromosome[1] and we estimated that 80% to 90% of its genes were yet to be discovered. In 1993, a team of 20 European laboratories began the systematic sequence analysis of chromosome XIV. The completed and intensively checked final sequence of 784,328 base pairs was released in April, 1996 (ref. 2). Substantial parts had been published before[3–22] or had previously been made available on request. The sequence contained 419 known or presumptive protein-coding genes, including two pseudogenes and three retrotransposons, 14 tRNA genes, and three small nuclear RNA genes. For 116 (30%) protein-coding sequences, one or more structural homologues were identified elsewhere in the yeast genome. Half of them belong to duplicated groups of 6–14 loosely linked genes, in most cases with conserved gene order and orientation (relaxed interchromosomal synteny). We have considered the possible evolutionary origins of this unexpected feature of yeast genome organization.

Figure 1 shows the map of cosmid, lambda and plasmid clones and of polymerase chain reaction (PCR) fragments from two unclonable regions which were used to determine the sequence of chromosome XIV. The final positions of genes listed in the 1992 map[1] are also presented changing the order of closely linked genes in only three regions. The assembled contig consists of 784,328 bp. The sequence of 180,983 bp (23%) was independently determined twice on both strands. These control sequences included 28 overlapping regions of cosmid and lambda clones (117,891 bp) as well as 108 selected regions, mainly at termini of open reading frames (ORFs), resequenced either on cosmids (54,540 bp) or by genomic PCR (8,552 bp). A total of 27 sequence mistakes were corrected. We estimate that the final sequence carries less than one error in every 10 kilobases, an estimate confirmed by a recent independent control analysis using 83 randomly picked genomic clones of chromosome XIV (G. Valle, unpublished data). Among the 40 kb sequenced, four deviations from our final sequence were noted: three single base-pair changes with neutral effects on coding regions (probably resulting from strain or clone differences), and only one confirmed sequence mistake. The left end of the chromosome carries telomeric repeat sequence (see below) and it is

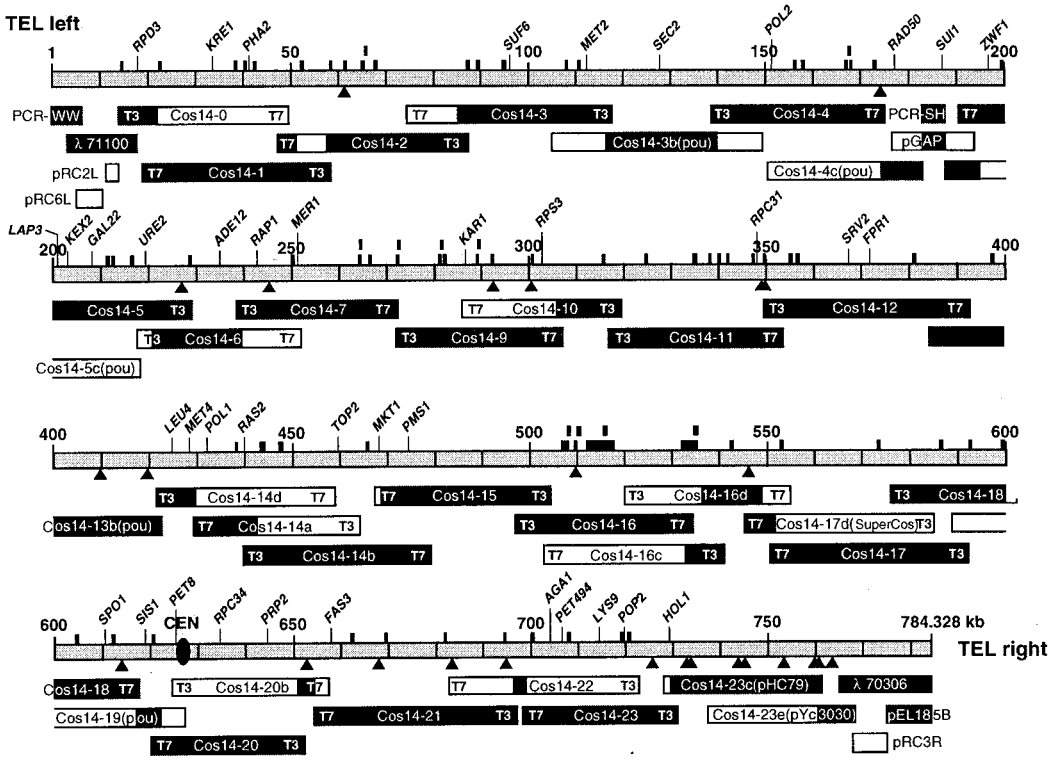


Figure 1 Physical map of subclones of chromosome XIV used for systematic DNA sequence analysis and final locations of originally genetically mapped genes[1]. Position of cosmid clones (cos), lambda clones (λ), plasmid clones (p) and genomic PCR fragments (PCR) are drawn as overlapping bars. Sequenced regions are shown in black. The 108 short regions selected for verification analyses are shown as small bars (resequenced cosmid clones) or triangles (sequenced genomic PCR fragments) along the contig. All clones were derived from S288C strains, except plasmid pGAP (carrying a spontaneus nonsense mutation in the toxic YNL247w), which originates from strain A364a[36]. Most cosmid clones with chromosome XIV DNA were isolated from cosmid libraries provided by B. Dujon[37] and R. Stucka[38], and mapped by a modified chromosome fragmentation approach[39,40]. Several clones extending into or bridging remaining gaps were isolated by colony screening using non-radioactively labelled restriction fragments as hybridization probes. Two cosmid clones (14–17d and 14–23c) and both lambda clones carrying telomere DNA were obtained from L. Riles[41] and the right telomere clone pEL185 was provided by E. Louis[42]. A more detailed description of the mapping strategy will be published elsewhere (K. Hamberg *et al.*, manuscript in preparation). The complete sequence can be retrieved from the EMBL database, accession nos Z71277–Z71692 or from the Martinsrieder website[2]. Different parts were sequenced in different laboratories: 1–6035 (R. Wambutt); 3,203–17,700 (B. Obermeier); 13,990–22,212 (A. Goffeau); 18,699–58,748 (C. Gaillardin); 47,022–51,246 and 57,523–87,525 (R. J. Planta); 85,152–132,424 (N. N. Glansdorf); 130,724–187,891 (J. H. Hegemann); 183,004–187900 (P. Philippsen); 187,809–192,153 (F. Del Rey); 192,154–195,234 (A. Jimenez); 190,506–229,360 (G. Valle); 220,854–239,907 (A. Düsterhöft); 238,582–273,742 (A. Goffeau); 271,932–319,898 (H. Domdey); 317,148–353,960 (C. Herbert); 349,559–393,039 (M. Jacquet); 384,059–421,858 (G. Valle); 421,188–443,100 (J.-L. Revuelta & F. Del Rey); 443,001–456,300 (A. Jimenez); 456,201–479,289 (J. P. G. Ballesta); 468,833–504,727 (P. Philippsen); 496,969–541,433 (M. Crouzet); 536,275–549,131 (C. Herbert); 545,180–592,214 (A. Düsterhöft); 575,858–617,912 (A. Urrestarazu); 617,105–622,324 (M. Crouzet); 620,016–652,539 (G. Volckaert.); 650,830–653,557 (R. J. Planta.); 651,995–654,446 (A. Düsterhöft); 654,389–731,357 (T. M. Pohl); 729,267–768,530 (A. Düsterhöft); 764,973–784,328 (A. Urrestarazu); 774,980–784,145 (C. Gaillardin).

Table 1 *S. cerevisiae* chromosome XIV ORFs and structurally homologous ORFs of other chromosomes

ORF Chr.XIV* 6–17.3 kb†	ORF Chr.VI* 6–15.5 kb	% identity/ stretch of amino acids	Biochemical or biological function (gene name) Chr. XIV ORF	Homologue in cluster duplication
YNL336w	YFL062w	94.2% overall	unknown	unknown
YNL335w	YFL061w	100 % overall	fungal cyanamide hydratase homologue	fungal cyanamide hydratase homologue
YNL334c	YFL060c	99.1% overall	unknown, probable membrane protein	unknown, probable membrane protein
YNL333w	YFL059w	99.7% overall	unknown	unknown
YNL332w	YFL058w	99.7% overall	thiamine regulated protein homologue	thiamine regulated protein (THI5)
YNL331c	YFL057–56c	87.0%/226 aa‡	probable aryl-alcohol reductase	probable aryl-alcohol reductase
ORF Chr.XIV 38.7–106.7 kb	ORF Chr.XV (A) 25.3–142.6 kb	% identity/ stretch of amino acids		
YNL318c	YOL156w	38.4%/510 aa	hexose transporter (HXT14)	glucose transporter (LGT3)
YNL307c	YOL128c	41.6%/316 aa	Ser/Thr/Tyr protein kinase (MCK1)	probable Ser/Thr protein kinase
YNL302c	YOL121c	99.3% overall	ribosomal protein (RPS16A)	ribosomal protein (RPS16B)
YNL301c	YOL120c	100% overall	ribosomal protein (RP28B)	ribosomal protein (RP28A)
YNL299w	YOL115w	54.1%/556 aa	topoismerase I related protein (TRF5)	topoisomeras I related protein (TRF4)
YNL298w	YOL113w	61.2%/317 aa	Ser/Thr protein kinase (CLA4)	probable Ser/Thr protein kinase
YNL293w	YOL112w	53.2%/417 aa	unknown	unknown
YNL290w	YOL094c	34.4%/317 aa	replication factor C subunit (RFC3)	replication factor C subunit (RFC4)
YNL283c	YOL105c	43.4%/302 aa	similarity to yeast chitinase	unknown
ORF Chr.XIV 252.1–307 kb	ORF Chr.IV 44.1–80.4 kb	% identity/ stretch of amino acids		
YNL209w	YDL229w	99.3% overall	heat shock protein (SSB2)	heat shock protein (SSB1)
YNL204c	YDL226c	30.5%/177 aa	sporul.spec.zinc finger protein (SPS18)	prolif.spec.zinc finger protein (GCS1)
YNL197c	YDL224c	36.5%/581 aa	regulator of cell size (WHI3)	unknown
YNL194c	YDL222c	52.0% overall	unknown, probable membrane protein	unknown, probable membrane protein
YNL183c	YDL214c	42.6%/479 aa	Ser/Thr protein kinase (NPR1)	probable Ser/Thr protein kinase
YNL176c	YDL211c	24.3%/292 aa	unknown, probable membrane protein	unknown, probable membrane protein
ORF Chr.XIV 309–410 kb	ORF Chr.VIII 390.3–341.4 kb	% identity/ stretch of amino acids		
YNL173c	YHR146w	27.4%/351 aa	pheromone-response G protein	unknown, probable G protein
YNL162w	YHR141c	100% overall	ribosomal protein (RPL41A)	ribosomal protein (RPL41A)
YNL160w	YHR139c	45.0%/307 aa	secreted glycoprotein (YGP1)	sporulat.spec. wall maturation (SPS100)
YNL156c	YHR133c	40.5%/205 aa	unknown	unknown
YNL154c	YHR135c	70.9%/499 aa	casein kinase I isoform (YCK2)	casein kinase I (YCK1)
YNL144c	YHR131c	36.2%/464 aa	unknown	unknown
YNL130c	YHR123w	53.8% overall	diacylglyc.choline-P transferase (CPT1)	ethanolamin P-tranferase (EPT1)
YNL121c	YHR117w	49.3%/651 aa§	import recept.mito outer memb.(TOM70)	mitochondrial outer membrane protein
YNL116w	YHR115c	55.4%/424 aa	unknown	unknown
ORF Chr.XIV 419–466 kb	ORF Chr.XV (B) 529–486.8 kb	% identity/ stretch of amino acids		
YNL108c	YOR110w	65.0%/273 aa	unknown	unknown
YNL106c	YOR109w	58.5%/979 aa	inositol phosphatase homologue	probable phosphatase
YNL104c	YOR108w	88.5%/601 aa	2-isopropyl malate synthase (LEU4)	2-isopropyl malate synthase homologue
YNL098c	YOR101w	55.8%/303 aa	GTP-binding protein (RAS2)	GTP-binding protein (RAS1)
YNL096c	YOR096w	87.9% overall	ribosomal protein S7 homologue	ribosomal protein (RP30)
YNL095c	YOR092w	55.3%/445 aa	unknown, probable membrane protein	unknown, probable membrane protein
YNL093w	YOR089c	56.9%/209 aa	GTP-binding protein (YPT53)	GTP-binding protein (VPS21)
YNL090w	YOR089c	56.9%/209 aa	GTP-binding protein (RHO2)	GTP-binding protein (VPS21)
YNL087w	YOR086c	54.5% overall	unknown, probable membrane protein	unknown, probable membrane protein
ORF Chr.XIV 478.6–597.6 kb	ORF Chr.IX 89.3–202.1 kb	% identity/ stretch of amino acids		
YNL079c	YIL138c	54.1%/159 aa	tropomyosin (TPM1)	tropomyosin (TPM2)
YNL074c	YIL135c	23.2%/375 aa	unknown	unknown
YNL069c	YIL133c	90.3% overall	ribosomal protein (RP23)	ribosomal protein (RP22)
YNL068c	YIL131c	52.1%/190 aa	unknown, fork head domain (FKH2)	unknown, fork head domain (FKH1)
YNL066w	YIL123w	62.9%/415 aa	ß-glucosidase homologue (SUN4)	homologue of aging gene UTH1
YNL065w	YIL121w	47.7%/342 aa	cycloheximid resist.protein homologue	antibiotic resitance protein homologue
YNL065w	YIL120w	43.0%/351 aa	cycloheximid resist.protein homologue	antibiotic resitance protein homologue
YNL058c	YIL117c	37.3%/126 aa	unknown	unknown
YNL055c	YIL114c	49.5% overall	outer mito membrane porin (OMP2)	OMP2 homologue
YNL053w	YIL113w	48.8%/162 aa	protein phosphatase (MSG5)	protein-Tyr phosphatase homologue
YNL052w	YIL111w	63.6% overall	cytochrome c oxidase (COX5A)	cytochrome c oxidase (COX5B)
YNL049c	YIL109c	61.8%/566 aa	unknown	unknown
YNL047c	YIL105c	54.3%/639 aa	unknown	unknown
YNL037c	YIL094c	35.8%/296 aa	isocitrate dehydrogenase (IDH1)	isopropyl malate dehydrog. homologue
YNL029c	YIL085c	55.7%/476 aa	mannosyl transferase homologue	mannosyl transferase homologue
YNL020c	YIL095w	41.2%/636 aa	probable Ser/Thr protein kinase	probable Ser/Thr protein kinase
ORF Chr.XIV 623.4–753.7 kb	ORF Chr.III 101.7–301.8 kb	% identity/ stretch of amino acids		
YNL004w	YCL011c	39.4%/409 aa	poly(A)bdg.protein homologue (TOM34)	probable TEL associated protein (GBP2)
YNR001c	YCR005c	81.4%/441 aa	citrate synthase (CIT1)	peroxysomal citrate synthase (CIT2)
YNR002c	YCR010c	77.7% overall	unknown, probable membrane protein	unknown, probable membrane protein
YNR013c	YCR037c	48.1%/489 aa	unknown, probable membrane protein	probable phophate transporter (PHO87)
YNR019w	YCR048w	52.5%/459 aa	sterol acyltransferase (SAT1)	cholesterol acyltransferase (ARE1)
YNR023w	YCR052w	29.5%/353 aa	unknown	unknown
YNR026c	YCR067c	45.4%/388 aa	GTP-GDP exchange factor (SEC12)	ER protein (SED4)
YNR028w	YCR069w	33.2%/304 aa	peptidyl-prolyl isomerase homologue	peptidyl-prolyl cis-trans isom.(SCC3)
YNR031c	YCR073c	53.6%/1172 aa	MAPKKK high osm.sign.transd. (SSK2)	MAP kinase kinase kinase (SSK22)
YNR034w	YCR073w-a	76.9% overall	multicopy sup of los1-1 (SOL1)	GlcN-6-P deaminase homologue (SOL2)
YNR047w	YCR091w	72.4%/424 aa	probable Ser/Thr protein kinase	probable Ser/Thr protein kinase (KIN82)
YNR048w	YCR094w	65.4% overall	unknown	unknown
YNR065-66c	YCR099-101c	64.8%/637 aa¶	peptidase Y sorting protein (pseudogene)	peptidase Y sorting (PEP1 homologue)

Degrees of homology were extracted from pairwise FASTA alignments of deduced protein sequences and are listed as percentage identity per stretch of amino acids.
*'Y' is included in Fig. 2 in the cluster duplication but is not listed in this table.
†Coordinates of clusters.
‡Overall homology between YNL331c and the sum of YFL056c and YFL057c (pseudogene in chromosomeVI?).
§Gaps introduced by the alignment algorithm may result in homology stretches slightly longer than the protein sequences.
¶Overall homology of the pseudogene (YNR065–YNR066c) to the sum of YCR099c, YCR100c and YCR101.

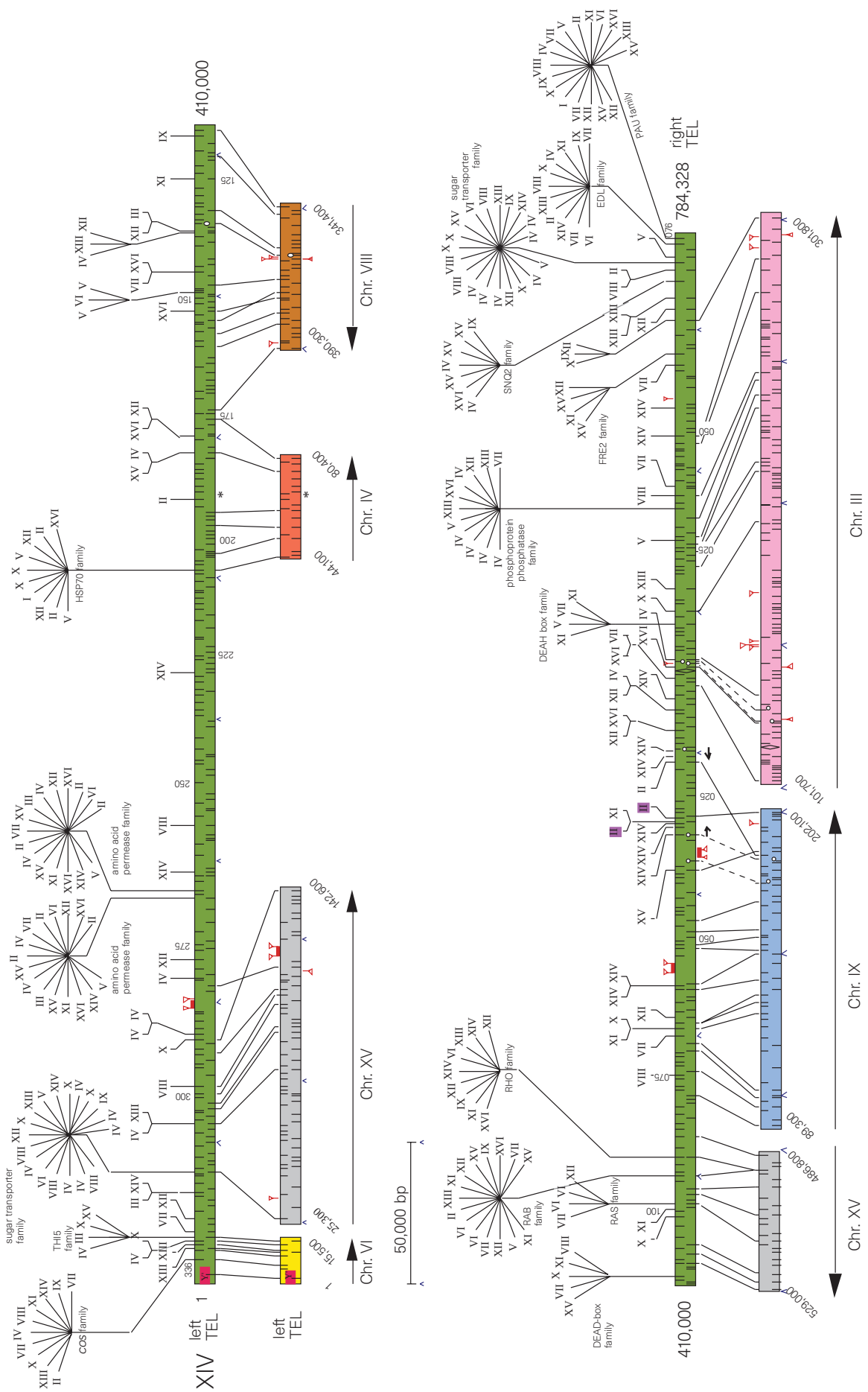

XIV
left TEL
410,000
Chr. VIII
Chr. IV
Chr. XV
Chr. VI
50,000 bp
HSP70 family
amino acid permease family
amino acid permease family
sugar transporter family
THI5 family
COS family
RHO family
RAB family
RAS family
DEAD-box family
phosphoprotein phosphatase family
DEAH box family
FRE2 family
SNO2 family
sugar transporter family
EDL family
PAU1 family
784,328
right TEL
Chr. III
Chr. IX
Chr. XV

Figure 2 Map of chromosome XIV ORFs that are members of either multigene families or of pairs, triplets or quadruplets of structurally related *S. cerevisiae* ORFs. The green bar represents both strands of chromosome XIV, with centres of all ORFs (excluding Ty and four short telomere ORFs) drawn as vertical lines; 215 are coded by the upper strand and 195 by the lower strand. Vertical lines with open circles mark selected tRNA genes. Three-digit numbers beneath or above the green bar refer to the systematic ORF nomenclature starting with 1 at either side of the centromere (white dot at 628 kb). Lines above or below the green bar indicate ORFs with structural homologues elsewhere in the genome (at least 30% identity in 150 amino acids, or, in a few cases, 25% in 300 amino acids). Lines extending into branches mark multigene families, with roman numbers indicating members on different chromosomes (order of decreasing homology to the chromosomes XIV ORF from left to right or clockwise, respectively). Coloured bars below chromosome XIV display seven syntenic or partly syntenic segments of other chromosomes with accumulations of ORFs structurally related to and arranged similarly to chromosome XIV ORFs. Broken lines in three of these clusters connect positions of pairs of functionally identical tRNA genes. The star at 280 kb indicates a functional ARS element on chromosome XIV (Ref. 36) which seems to be positionally conserved on chromosome IV. Further details of these cluster duplications are given in Table 1. The red bar of the left telomere represents the ubiquitous Y′ element found at many ends of *S. cerevisiae* chromosomes[43,44]. Red triangles mark positions of solo delta sequences (remnants of Ty elements) and red bars flanked by triangles indicate Ty elements. The two black arrows at 570 kb and 600 kb indicate an intrachromosomal highly conserved inverted repeat, involving in each repeat element one $tRNA^{Ile}$ gene and two new ORFs (YNL034w–YNL035w and YNL019c–YNL018c, respectively). The two marked chromosome II homologues and the corresponding chromosome XIV ORFs at 575 kb represent the two copies of the duplicated histone H3–H4 gene pair. As indicated there is an additional homologue to histone H3 on chromosome XI (*CSE4*), probably the yeast homologue of the human *CENP-A* gene[45]

possble that this end is a few hundred base pairs longer than indicated in Figure 1.

A systematic search of the chromosome XIV contig revealed 414 ORFs with 100 and more codons, including overlapping ORFs but excluding ORFs located within longer ORFs on either the same or the complementary strand. Chromosome XIV also has at least seven ORFs with less than 100 codons, of which four are known genes (*MFA2*, *TOM7*, *ATX1* and *PBI2*) and three show significant homology to known genes. A systematic nomenclature was given to all ORFs (excluding the six ORFs of the three Ty retrotransposons), indicating the organism (Y), the chromosome (N), the chromosome arm (L or R), the coding strand (Watson, w or Crick, c), and increasing numbers starting at the centromere; examples include YNL001w and YNR001c.

A simultaneous search for introns (using the EXPLORA program[23]) revealed 16 intron-carrying genes, in two of which, (YNL066w and YNL065w) the introns are located in the non-translated 5′ region[24,25]. EXPLORA failed to locate an additional, experimentally verified intron in ORF YNL044w, because it has an unusual 5′ splice sequence (EMBL database, accession nos X97400 and X97401).

Two pairs of adjacent ORFs (YNR065c and YNR066c; and YNR068c and YNR069c) were separated only by a stop codon; this was confirmed in both cases by genomic PCR. These pairs are rare examples of yeast pseudogenes, as highly conserved copies lacking internal stop codons are present on other chromosomes. Like their functional homologues these pseudogenes should be considered as single ORFs. Taking this into account, 419 ORFs are located on chromosome XIV, including six Ty-elements and 23 questionable ORFs (short ORFs overlapping longer ones). The ORF density, not counting questionable ORFs, is one ORF per 1.98 kb (a total of 396 ORFs in 784 kb), and the average ORF size is 1.5 kb. These numbers are very similar to corresponding numbers obtained with other *S. cerevisiae* chromosomes. The ORF density (the ratio of ORF nucleotides to total nucleotides) fluctuates between 0.6 and 0.9. These fluctuations do not correlate with fluctuations in G+C content; five of eight ORF density peaks coincide with regions of highest G+C content (39.4–40.0%) and the other three with regions of lowest G+C content (36.6–37.7%).

How many of the 396 non-questionable ORFs are new? Presently, functions are known, at least partly, for 149 ORFs (38%), based on detailed experiments or very high sequence homology to known genes[2,26] . Most of these are involved in metabolism, cell growth, cell division, translation, transcription and intracellular transport, with a few involved in energy production, metabolite transport, protein modification, signal transduction, and stress response. Of the 247 new ORFs, some functional predictions can be made for 43 (11%), owing to homologies to characterized genes in *S. cerevisiae* or other organisms. Presumptive products coded by these 43 ORFs include: a human breast cancer-associated autoantigen homologue; a genetically linked cluster of three proteins (transporter, epimerase and reductase) for potential utilization of an unidentified mono- or oligosaccharide; four proteins with homology to prokaryotic ribosomal proteins; three protein kinases; three GTP-binding proteins; two protein phosphatases; two translation factors; two drug- resistance proteins; one actin homologue; one zinc finger; one peptidyl-prolyl isomerase; and ten with presumptive metabolic activities, such as cyanamide hydratase, mannosyl transferase, isocitrate dehydrogenase and inositol phosphatase. Further details can be found on the Martinsrieder website[2].

The functions of the other 204 ORFs (51%) cannot yet be predicted. One third of these code for presumptive membrane proteins, and more than four transmembrane domains are predicted for 18. Of the 204 new ORFs, 12 have homology to human expressed sequence tags (EST)[27, 28] with FASTA scores of 200–760. Remarkably, two of the 23 questionable ORFs (YNL228w and YNL114c) also have significant homology to human EST sequences.

We used FASTA comparisons of all chromosome XIV ORFs (except the highly repetitious Y′ and Ty ORFs) to all *S. cerevisiae* ORFs in order to establish the extent of gene duplications, and found that 116 ORFs shared structural homology with one or more ORFs elsewhere in the genome. For this search, structural homology was defined as over 30% identity in a stretch of 150 amino acids (in some cases, 25% identity in a stretch of at least 300 amino acids). Of these 116 ORFs, 67 belong to pairs of homologues, 32 to groups consisting of three or four homologues, and 17 are members of multigene families. ORFs from all chromosomes contribute to this picture of sequence homology (Fig. 2). The list of homologies based on FASTA analyses also revealed several regions of chromosome XIV with accumulations of homologous ORFs originating from distinct regions of six other chromosomes, and showing, with only a few exceptions, conserved gene orders and gene orientations. One of these apparently ancient duplications, involving ORFs of the left arms of chromosomes IX and XIV, respectively, had previously been reported[19,29]. Duplications involving several genes had been described up to that time, mainly for relatively short subtelomeric and centromeric regions[30–34].

The extent of these types of duplications became apparent after the complete sequence information of the *S. cerevisiae* reference strain S288C was released[2] . With respect to chromosome XIV, so-called gene cluster duplications were found in seven regions of 17 kb to 130 kb. The precise locations of the 67 pairs of ORF homologues in these seven cluster duplications are shown in Fig. 2, together with all other chromosome XIV ORFs for which structural homologues were found; five pairs of positionally conserved duplicated tRNA genes are also indicated. Probably half of these structural homologies among different chromosomes would have remained undetected in classical DNA hybridization experiments.

Complementary to the graphical display of the seven cluster duplications, we have determined the degree of homologies for each ORF pair and, if known or predictable, their functions (Table 1). ORFs displayed from left to right in Fig. 2 are listed from top to bottom in the table. An automated means of finding and displaying structurally homologous segments in genomes several million base pairs long involves the screening of sliding windows of 500 bases between pairs of chromosomes[35]. This very efficient method was also applied to chromosome XIV, and most of the ORF pairs participating in cluster duplications were detected (K. Heumann, unpublished data). However, this automated approach still

requires manual editing to find all details of cluster duplications, such as multigene families, potentially inverted ORF members, more than averagely diverged ORFs, and tRNA genes.

The 17-kb subtelomeric cluster duplication between chromosomes XIV and VI (cluster duplications 14–6) consists entirely of highly conserved ORF pairs (average 96.6% amino-acid identity) and shows stringent synteny. The intergenic regions are also highly conserved, suggesting that the duplication of the six ORFs is a relatively recent event on an evolutionary timescale.

Most of the ORF pairs in the other six cluster duplications are much less conserved, and their promotor and terminator regions lack significant homologies, suggesting that they are ancient duplication events. Five of the highly conserved ORF pairs of these ancient duplications code for ribosomal proteins (average 95.3% amino-acid identity), one for two members of the 70K heat-shock protein family (99.3% amino-acid identity), one for two forms of iso-propyl malate synthase (88.5% amino-acid identity) and one for two forms of citrate synthase (81.4% amino-acid identity) (Table 1). Excluding these ORF pairs, which are apparently under high selection pressure to preserve their sequence information, the average homology of ORF pairs was determined for each of the cluster duplications. ORF pairs in cluster duplications CD14–15B and CD 14–3 (average 56% amino-acid identity) seem to be less diverged than ORF pairs in CD14–15A, CD14–8, CD14–9 (average 47.5% amino-acid identity) and CD14–4 (average 37% amino acid identity). However, there are too few ORF pairs to draw conclusions about different temporal orders for the cluster duplications involving chromosome XIV.

Could the six ancient cluster duplications, at the time of their creation, have looked similar to the recent cluster duplications between chromosomes XIV and VI, with perfect synteny of all ORFs? And could they have been shaped over evolutionary time by base-pair changes, insertions of new ORFs, deletions of some of the originally duplicated ORFs, inversions of single or groups of ORFs, and translocations to yield the present picture of 'relaxed synteny'? This is certainly possible if the now visible arrangements indeed originated from duplications of gene clusters, perhaps by long-range gene conversions or chromosome duplications. However, it remains possible that the evolutionary history of *S. cerevisiae* involved fusion of two ancient forms of yeast cells with smaller genomes already displaying sequence divergencies and some level of relaxed synteny and that, for most of the duplicated ORFs, one copy was lost over time because of a lack of selective advantage for *S. cerevisiae* to keep more than one copy. □

Received 22 July 1996; Accepted 11 March 1997.

1. Mortimer, R. K., Contopoulou, C.R. & King, J.S. *Yeast* 8, 817–902 (1992).
2. http://www.mips.biochem.mpg.de/mips/yeast/
3. Verhasselt, P., Aert, R., Voet, M. & Volckaert, G. *Yeast* 10, 945–951 (1994).
4. Verhasselt, P., Aert, R., Voet, M. & Volckaert, G. *Yeast* 10, 1355–1361 (1994).
5. Jonniaux, J. L., Coster, F., Purnelle, B. & Goffeau, A. *Yeast* 10, 1639–1645 (1994).
6. Kick, C. T., Maurer, J. H., Maurer, U. & Planta, R.J. *Yeast* 11, 1303–1310 (1995).
7. Mallet, L., Bussereau, F. & Jacquet, M. *Yeast* 11, 1195–1209 (1995).
8. Van Dyck, L., Pascual-Ahuir, A., Purnelle, B. & Goffeau, A. *Yeast* 11, 987–991 (1995).
9. Coster, F., van Dyck, L., Jonniaux, J.-L., Purnelle, B. & Goffeau, A. *Yeast* 11, 85–91 (1995).
10. Bergez, P., Doignon, F. & Crouzet, M. *Yeast* 11, 967–974 (1995).
11. Maftahi, M., Nicaud, J.-M., Levesque, H. & Gaillardin, C. *Yeast* 11, 567–572 (1995).
12. Maftahi, M., Nicaud, J.-M., Levesque, H. & Gaillardin, C. *Yeast* 11, 1077–1085 (1995).
13. Maurer, K. C., Urbanus, J. H. & Planta, R. J. *Yeast* 11, 1303–1310 (1995).
14. Soler-Mira, A., Saiz, J. E., Ballesta, J. P. G. & Remacha, M. *Yeast* 12, 485–491 (1996).
15. Levesque, H., Lepingle, A., Nicaud, J.-M. & Gaillardin, C. *Yeast* 12, 289–295 (1996).
16. Sen-Gupta, M., Lyck, R., Fleig, U., Niedenthal, R. K. & Hegemann, J. H. *Yeast* 12, 505–514 (1996).
17. Nasr, F., Bécam, A.-M. & Herbert, C.J. *Yeast* 12, 169–175 (1996).
18. Nasr, F., Bécam, A.-M. & Herbert, C.J. *Yeast* 12, 493–499 (1996).
19. Pöhlmann, R. & Philippsen, P. *Yeast* 12, 391-402 (1996).
20. Saiz, J. E., Buitrago, M. J., Soler-Mira, A., Del Rey, F. & Revuelta, J.L. *Yeast* 12, 403–409 (1996).
21. Garcia-Cantalejo, J. M., Boskovic, J. & Jimenez, A. *Yeast* 12, 599–608 (1996).
22. Pandolfo, D., De Antoni, A., Lanfranchi, G. & Valle, G. *Yeast* 12, 1071–1076 (1996).
23. Kalogeropoulos, A. *Yeast* 11, 555–565 (1995).
24. Logghe, M., Molemans, F., Fiers, W. & Contreras, R. *Yeast* 10, 1093–1100 (1994).
25. Rodriguez-Medina, J. R. & Raymond, B. C. *Mol. Gen. Genet.* 243, 532–539 (1994).
26. Garrels, J. I. *Nucleic Acids Res.* 24, 46–49 (1996).
27. Tugenreich, S., Boguski, M. S., Seldni, M. S. & Hieter, P. *Proc. Natl Acad. Sci. USA* 90, 10031–10035 (1993).
28. Tugenreich, S., Bassett, D. E., McKusick, V. A., Boguski, M. S. & Hieter, P. *Hum. Mol. Genet.* 3, 1509–1517 (1994).
29. Pöhlmann, R. & Philippsen, P. *Yeast* 11, 634 (1995).
30. Steensma, H.Y., de Jonge, P., Kaptein, A. & Kaback, D.B. *Curr. Genet.* 16, 131–137 (1989).
31. Lalo, D., Stettler, S., Mariotte, S., Slonimski, P. P. & Thuriaux, P. *C.R. Acad. Sci. Paris* 316, 367–373 (1993).
32. Johnston, M., *et al. Science* 265, 2077–2082 (1994).
33. Wolfe, K. H. & Lohan, A. J. *Yeast* 10, 41–46 (1994).
34. Melnick, L. & Sherman, F. *J. Mol. Biol.* 233, 372–388 (1993).
35. http://speedy.mips.biochem.mpg.de/programs/GENOME_BROWSER.html
36. Friedman, K. L. *et al. Genes Dev.* 10, 1595–1607 (1996).
37. Thierry, A., Gaillon, L., Galibert, F. & Dujon, B. *Yeast* 11, 121–135 (1995).
38. Stucka, R. & Feldmann, H. in *Molecular Genetics of Yeast* (ed. Johnston, J. R.) 49–64 (IRL Oxford, 1994).
39. Hamberg, K. *PhD-Thesis, Univ. Giessen* (1993).
40. Vollrath, D., Davis, W.D., Cornelly, C. & Hieter, P. *Proc. Natl. Acad. Sci. USA* 85, 6027–6031 (1988).
41. Riles, L. *et al. Genetics* 134, 81–150 (1993).
42. Louis, E.J. & Borts, R.H. *Genetics* 139, 125–136 (1995).
43. Chan, C.S.M. & Tye, B.-K. *Cell* 33, 563–573 (1983).
44. Louis, E.J. *Yeast* 11, 1553–1573 (1995).
45. Stoler, S., Keith, K.C., Curnick, K.E. & Fitzgerald-Hayes, M. *Genes Dev.* 9, 573–586 (1995).

Acknowledgements. We thank L. Riles, A. Thierry, B. Dujon, R. Stucka, H. Feldmann, E. Louis, K. Friedman and B. Brewer for clones and cosmid libraries; R. Spiegelberg, A. Thierry and D. Fischer for helping to isolate or characterize DNA clones; M. Johnston, T. Donahue, N. Pfanner, B. Winsor and D. Gallwitz for suggestions; and R. Niederhauser for secretarial help. The majority of funding was provided by the Biotech Programs of the European Commission. Additional financial support was contributed by the following national agencies: Groupement de Recherches et d'Etudes sur les Génomes du Ministre de la Recherche, France; Région de Bruxelles-Capital, Belgium; Belgian Federal Services for Science Policy (D.W.T.C.) ; Research Fund of the Katholieke Universiteit Leuven, Belgium; Services Fédéraux des Affaires Scientifiques, Techniques et Culturelles; Pôles d'attraction Inter-universitaire and Région Wallone, Belgium; Fundacion Ramon Areces and Comision Interministerial de Ciencia y Tecnologia, Spain. The participation of scientists from Switzerland was made possible by a grant from the Swiss Federal Agency for Education and Science.

The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XV

B. Dujon[1], K. Albermann[2], M. Aldea[3], D. Alexandraki[4,5], W. Ansorge[6], J. Arino[7], V. Benes[6], C. Bohn[8], M. Bolotin-Fukuhara[8], R. Bordonné[9], J. Boyer[1], A. Camasses[9], A. Casamayor[7], C. Casas[3], G. Chéret[10], C. Cziepluch[11], B. Daignan-Fornier[8], D. V. Dang[8], M. de Haan[12], H. Delius[13], P. Durand[14], C. Fairhead[1], H. Feldmann[15], L. Gaillon[1], F. Galisson[1], F.-J. Gamo[16], C. Gancedo[16], A. Goffeau[17], S. E. Goulding[18], L. A. Grivell[12], B. Habbig[19], N. J. Hand[18], J. Hani[2], U. Hattenhorst[19], U. Hebling[13], Y. Hernando[20], E. Herrero[3], K. Heumann[2], R. Hiesel[21], F. Hilger[14], B. Hofmann[13], C. P. Hollenberg[19], B. Hughes[22], J.-C. Jauniaux[11], A. Kalogeropoulos[8], C. Katsoulou[4], E. Kordes[11], M. J. Lafuente[16], O. Landt[23], E. J. Louis[24], A. C. Maarse[12], A. Madania[9], G. Mannhaupt[15], C. Marck[25], R. P. Martin[9], H. W. Mewes[2], G. Michaux[1], V. Paces[26], A. G. Parle-McDermott[18], B. M. Pearson[20], A. Perrin[1], B. Pettersson[27], O. Poch[9], T. M. Pohl[22], R. Poirey[11], D. Portetelle[14], A. Pujol[11], B. Purnelle[17], M. Ramezani Rad[19], S. Rechmann[6], C. Schwager[6], M. Schweizer[20], F. Sor[10], F. Sterky[27], I. A. Tarassov[9], C. Teodoru[6], H. Tettelin[1,17], A. Thierry[1], E. Tobiasch[11], M. Tzermia[4], M. Uhlen[27], M. Unseld[21], M. Valens[8], M. Vandenbol[14], I. Vetter[15], C. Vlcek[26], M. Voet[28], G. Volckaert[28], H. Voss[6], R. Wambutt[29], H. Wedler[29], S. Wiemann[6], B. Winsor[9], K. H. Wolfe[18], A. Zollner[2], E. Zumstein[20] & K. Kleine[2]

[1]Unité de Génétique Moléculaire des Levures (URA 1149 CNRS and UFR 927 Univ. P.M. Curie), Institut Pasteur, 25 Rue du Dr. Roux, F75724, Paris Cedex 15, France
[2]Martinsrieder Institut für Protein Sequenzen, Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, D-82152, Martinsried, Germany
[3]Department of Basic Medical Sciences, Faculty of Medicine, University of Lleida, E-25006, Lleida, Spain
[4]Foundation for Research and Technology-Hellas, IMBB, P.O. Box 1527, Heraklion 711 10 Crete, Greece
[5]Department of Biology, University of Crete, Heraklion 711 10 Crete, Greece
[6]Biochemical Instrumentation Program, EMBL, Meyerhofstrasse 1, D-69117, Heidelberg, Germany
[7]Departamento de Bioquimica y Biologia Molecular, Universidad Autonoma de Barcelona, Bellaterra, E-08193, Spain
[8]Institut de Génétique et Microbiologie, Bâtiment 400, Université Paris-Sud,

F-91405 Orsay Cedex, France
[9]UPR9005 (MMDCD) du CNRS, 15 rue René Descartes, F-67084, Strasbourg, France
[10]Institut Curie, Bâtiment 110, Centre Universitaire, F-91405, Orsay Cedex, France
[11]Tumorvirologie Abteilung 0610 and Virologie appliquée à l'oncologie (INSERM U375), Deutsches Krebsforschungszentrum, P.101949, D-69009, Heidelberg, Germany
[12]Section for Molecular biology, Department of Molecular Cell Biology, University of Amsterdam, Kruislaan 318, NL-1098 SM Amsterdam, The Netherlands
[13]Abteilung für Angewandte Tumorvirologie 0686, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 506, D-69120 Heidelberg, Germany
[14]Faculté Universitaire des Sciences Agronomiques, Unité de Microbiologie, 6 Avenue Maréchal Juin, B-5030 Gembloux, Belgium
[15]Institut für Physiologische Chemie, Physikalische Biochemie und Zellbiologie der Universität München, Goethestrasse 33, D-80336 München, Germany
[16]Instituto de Investigaciones Biomedicas, CSIC, Arturo Duperier 4, E-28029 Madrid, Spain
[17]Unité de Biochimie Physiologique, Université Catholique de Louvain, Place Croix du Sud 2-20, B-1348 Louvain-La-Neuve, Belgium
[18]Department of Genetics, Univeristy of Dublin, Trinity College, Dublin 2, Ireland
[19]Institut für Mikrobiologie der Heinrich-Heine-Universität Düsseldorf, Geb. 26.12, Universitätstrasse 1 D-40225 Düsseldorf, Germany
[20]Institute of Food Research, Genetics and Microbiology Department, Norwich Research Park, Colney, Norwich NR4 7UA, UK
[21]Institut für Genbiologische Forschung, Ihnestrasse 63, D-14195 Berlin, Germany
[22]GATC-Gesellschaft für Analyse-Technik und Consulting mbH, Fritz-Arnold-Strasse 23, Konstanz, D-78467, Germany
[23]TIB MolBiol, Tempelhofer Weg 11-12, D-10829 Berlin, Germany
[24]Yeast Genetics, Institute of Molecular Medicine, John Radcliffe Hospital, Oxford OX3 9DU, UK
[25]Service de Biochimie et de Génétique moléculaire, Département de Biologie Cellulaire et Moléculaire, DSV/CEA-Saclay, F-91191 Gif-sur-Yvette, France
[26]Institute of Molecular Genetics, Academy of Sciences, Prague, Czech Republic
[27]Department of Biochemistry and Biotechnology, Royal Institute of Technology (KTH), S10044, Stockholm, Sweden
[28]Laboratory of Gene Technology, Katholieke Universiteit Leuven, Willem de Croylaan, 42, B-3001, Leuven, Belgium
[29]AGON GmbH, Glienicker Weg 185, D-12489 Berlin, Germany

Chromosome XV was one of the last two chromosomes of *Saccharomyces cerevisiae* to be discovered[1]. It is the third-largest yeast chromosome after chromosomes XII and IV, and is very similar in size to chromosome VII. It alone represents 9% of the yeast genome (8% if ribosomal DNA is included). When systematic sequencing of chromosome XV was started, 93 genes or markers were identified, and most of them were mapped[2]. However, very little else was known about chromosome XV which, in contrast to shorter chromosomes, had not been the object of comprehensive genetic or molecular analysis. It was therefore decided to start sequencing chromosome XV only in the third phase of the European Yeast Genome Sequencing Programme, after experience was gained on chromosomes III, XI and II (refs 3–5). The sequence of chromosome XV has been determined from a set of partly overlapping cosmid clones derived from a unique yeast strain, and physically mapped at 3.3-kilobase resolution before sequencing. As well as numerous new open reading frames (ORFs) and genes encoding tRNA or small RNA molecules, the sequence of 1,091,283 base pairs confirms the high proportion of orphan genes and reveals a number of ancestral and successive duplications with other yeast chromosomes.

The DNA sequence of 1,091,283 nucleotides contains 560 ORFs, of at least 100 sense codons, that are not entirely included within a larger one (our standard basic definition; see ref. 4). If those corresponding to Ty or Y′ elements are excluded, and intron predictions are considered (see below), 551 different ORFs remain. To these were added eight known genes shorter than 100 codons (*BAT2, CRS5, RPB10, RPS30B, RPS33A, SME1, TOM6* and the *CPA1* leader), and a pseudogene 581 codons long (YOL153c) that contains two in-frame ochre codons. It is considered here because its putative translation product has significant homology with the Gly-X carboxypeptidase encoded by *CPS1*, and because, in another yeast strain, the two stop codons are replaced by two glutamine codons CAA[6]. Note that, in the present sequence, YOR031w (*CRS5*) also contains an in-frame ochre codon instead of the CAA codon found in other strains (the reality of the stop codon was verified by direct sequencing on yeast DNA). Other interesting but more complex cases of pseudogenes found in chromosome XV will be described elsewhere. Also note that YOL040c has been considered here (coordinates 253,147– 253,575) instead of the larger antisense ORF within which it is entirely included because it corresponds to a known gene (*RPS21*).

The sequence also reveals 526 short ORFs (from 50 to 99 sense codons) not entirely included within larger ones, six of which are already considered above. In this size range, it is obviously difficult to distinguish actual genes from random occurrences. Using high codon bias (CAI > 0.2) and absence of a partial overlap with larger ORFs as predictive criteria, only six candidates remained. One of these shows very significant homology with the 60 codon-long gene *HOR7* of chromosome XIII. This ORF, named YOL052ca, has been added to the above list, bringing the grand total to 561. Of these ORFs 33 (5.9%) are 'questionable', based on their short size and low CAI (see ref. 4), and 18 of them partly overlap other ORFs, increasing their suspicious character.

The longest ORF of chromosome XV is YOL081w (the *IRA2* gene) with 3,079 codons. The two shortest ORFs, if the 25 amino-acid leader peptide of the *CPA1* gene is ignored, are YOR045w (*TOM6*) and YOL052ca, with 61 codons each. The average size of chromosome XV ORFs is 457 codons, very close to the figure observed for the entire yeast genome[7]. In total, 33 pairs and 5 trios of partly overlapping ORFs are found; 25 pairs are antiparallel, excluding the possibility of sequencing errors. Among the parallel pairs, YOR012w and YOR013w are suggestive of a frameshift error, or a pseudogene, as their products share homology with the amino terminus and the carboxy terminus of the product of YDR391c, respectively.

The overall density for protein-coding genes is 70.6%, slightly lower than the average for the whole yeast genome subtracted from rDNA[7]. Variations are observed along the chromosome, with two short regions showing gene density above 85% (centred around ~200 kb and ~950 kb, respectively; see Fig. 1). These regions also correspond to areas of high G+C content. As is generally observed in yeast, the two subtelomeric regions show low gene density.

The entire chromosome shows no significant strand coding bias, but important local variations are observed, with seven short regions (of ~30–50 kb each) showing a clearcut excess of ORFs on the Watson strand; eight others have an excess on the Crick strand (Fig. 1). Orientation of neighboring ORFs is random for the whole chromosome, with 150 diverging pairs, 149 converging pairs, and 261 tandemly arranged pairs (123 on the Watson strand, and 138 on the Crick strand). The longest tandem array contains 11 successive ORFs (YOR104w to YOR114w, coordinates 517,639–538,451).

A total of 13 introns have been identified, most of which occur in short ORFs with high codon bias. Chromosome XV introns are short, as is typical for yeast, ranging in size from 135 to 527 nucleotides (average 334 nucleotides). They have an average G+C content of 34.9%, significantly lower than that of the entire chromosome (38.2%) or its ORFs, (39.5%). Note that the possible occurence of introns in the 5′ untranslated region of the pre-mRNA molecules has not been examined systematically for lack of discriminative criteria among the numerous occurrences of the intron consensus 5′-GYMHGH-N1-TACTAAC-N2-YAG-3′ in the sequence (294 occurences if N1 and N2 are set shorter than or equal to 400 and 50 nucleotides, respectively, or more than 600 for limits of 1,000 and 180 nucleotides).

The 20 tRNA genes recognized correspond to 12 different amino acids and 22 different codons. Six tRNA genes contain introns. All tRNA genes are significantly richer in G+C content (47–63%) than the average yeast genome sequence. The frequent duplication of tRNA genes in yeast is

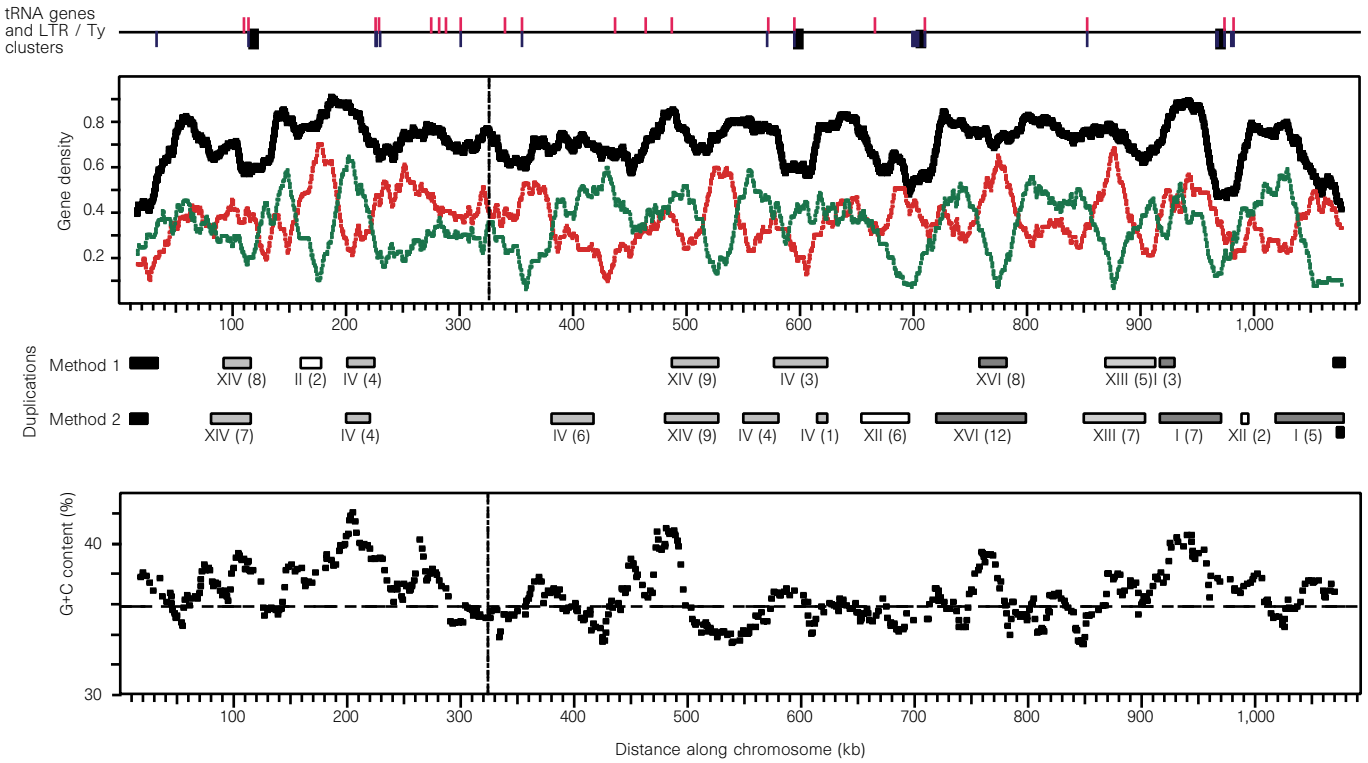


Figure 1 Variation in gene density (top) and base composition (bottom) along the sequence of chromosome XV (scale in kilobases from left telomere). Vertical broken lines indicate position of the centromere. Gene density is expressed as the probability that each nucleotide is part of an ORF, calculated using sliding windows of 30 kb (steps of 0.5 kb) for the Watson strand alone (red line), the Crick strand alone (green line), and their sum (black line). The Watson strand is oriented 5′ to 3′ from left to right on the chromosome map[4]. G+C content was calculated from silent codon positions using a sliding window of 13 consecutive ORFs (horizontal broken line indicates average G+C content at these positions; 35.7%). Top line, positions of tRNA genes (thin bars above line), solo LTR (thin bars below line), and complete Ty elements (thick bars below line). Middle, positions of major 'ancient' chromosomal duplications. Blocks represent extent of clusters (method 1), or arrays of ORF pairs (method 2), as defined in text, with indication of the matching chromosome and the number of ORFs involved (in parentheses). Note that several blocks corresponding to a given chromosome are often intermingled with those corresponding to another. Blocks in subtelomeric position (filled) match with several distinct chromosomes.

noticeable on chromosome XV alone, with four gly-tRNA genes (*tG(GGY)OL1* and *tG(GGY)OL2* are identical in sequences, and *tG(GGG)OR1* is similar to the previous two), three pro-tRNA genes (identical in sequences except for their introns), two thr-tRNA genes, and two asn-tRNA genes, respectively identical in sequences; there are also two met-tRNA genes, but they differ in sequence. Duplicated tRNA genes are always found within different sequence environments (including different associated long terminal repeats (LTRs); see below). The tRNA genes are distributed throughout the chromosome (Fig. 1), as is generally the case for yeast which, unlike other organisms, does not show large clusters of tRNA genes in its genome.

Seven genes encoding small known RNA molecules were recognized from the sequence. One of these, *snR17A*, which encodes the U3 snRNA, contains an intron and is duplicated elsewhere in the yeast genome[8]. The downstream part of the gene encoding snR35 partly overlaps the downstream part of an ORF of unknown function, YOR222w, the product of which shares similarity with ADP-ATP carrier proteins.

Four complete Ty elements (two Ty1 and two Ty2), 20 solo LTRs or remnants of them (12 delta, 6 sigma and 2 tau elements), and one Y′ element were found. Solo LTRs, or complete Ty elements, are almost always located immediately upstream of tRNA genes (Fig. 1). Only one solo LTR element and one Ty1 stand alone. Consistent with general trends[9], the two Ty2 elements are found at 'old' sites occupied by several solo LTRs associated with tRNA genes; the same is true for only one of the two Ty1 elements. Conversely, only 11 of the 20 tRNA genes have one or more LTR element upstream of their 5′ end, the closest element always being within 200 base pairs of the tRNA gene (sigma elements are only 16–18 nucleotides upstream of tRNA genes).

In total, RNA-coding genes and transposons occupy only 1% of the chromosome XV sequence.

There remain 309,627 base pairs (28.4% of the chromosome) that we will call here 'intergenic regions'. Such regions contain, for a small part, structural chromosomal elements such as the centromere (coordinates 326,592–326,706), the telomeric $(C_{1-3}A)_n$ repeats (coordinates 1–113 and 1,091,264–1,091,283) and two subtelomeric core X elements and their associated repeated elements (STR) (coordinates 114–847 and 1,083,914–1,084,611). But intergenic regions primarily contain promoters, terminators and transcriptional regulatory elements of the protein-coding genes, most of which have not yet been identified. Intergenic regions believed to contain promoter elements based on the orientation of flanking ORFs are noticeably longer (791 nucleotides on average between diverging ORFs) and are richer in G+C (36.2%) than intergenic regions containing putative terminators (421 nucleotides and 28.7% G+C). The presence of flanking RNA-coding genes does not alter this trend. The sequence also reveals 63 ARS consensus elements (5′-WTTTAYRTTTW-3′) the activity of which remains to be examined; 39 of these occur in intergenic regions.

Chromosome XV contains few simple sequence iterations. The longest dinucleotide repeat is an alternating poly(AT) stretch of 20-mers (coordinates 45,691–45,730) within the intergenic region that separates the converging ORFs YOL149w (*DCP1*) and YOL148c (*SPT20*). Only 19 other cases exist of either dinucleotide repeats of at least 10-mers (all are alternating poly(AT) or mononucleotide repeats of at least 20-mers (all poly(A) or poly(T)). Similarly, few trinucleotide repeats are found, the longest being a 20-mer of the triplet CAA occurring within an ORF of unknown function, YOR267c, and encoding a poly-glutamine stretch. Six

Table 1 Assembly of the chromosome XV sequence from individual submissions

Cosmids or DNA	Coordinates on final chromosome	Overlap	Strategy*	Reference
telomeric plasmid / pEOA363	1–32687		S, A	
pEOA179 / pEOA461	24486–97824	8202	N, A, M	6, 14, 16, 17, 18, 20
pEOA417	96924–140942	901	S, M	26
pEOA228	139031–178337	1912	S, A	28.29
pEOA1044	177014–210234	1324	N, W, A	in the press
pUOA1217	209185–235991	1050	SS, M	22
pUOA1344	222408–256233	13584	S, W, A	
pEOA321	253576–287613	2658	N, A	in the press
pEOA215	286637–323078	977	S, R	
pEOA156	321732–352202	1347	S, P, A	30
pEOA303 / pEOA270	350740–408356	1463	S, A	
pEOA272	391560–427841	16799	S, A	
pEOA213 / pEOA217	415169–477887	12673	SS, M	in the press
11 cosmids†	476475–606002	1413	S, W, A	27 and in the press
pEOA477 / pUOA1258	604545–660867	1458	SS, W, M	15, 21
pUOA533 / pEOA378 / pEOA241	655864–741096	5004	S, R	
pEOA423 / pEOA048	739215–799188	1882	S, A	8, 11
pUOA1302	795488–832262	3701	SS, W, M	in the press
pUOA1337	823760–861431	8503	S, A	19
pEOA487	859329–895527	2103	SS, W, M	in the press
pEOA284	892246–927955	3282	SS, A	24
pUOA502 / pEOA232	927738–957182	218	S, A	23
pEOA138	955761–996055	1422	S, M	25
5 cosmids‡	992145–1081258	3911	SS, W, M	
Right PCR	1080794–1091283	465	P	

*S, shotgun of cosmid or part thereof (SS); N, nested deletions; W, walking primer; P, PCR fragments; M, manual gels; A, automatic fluorescent; R, direct membrane blotting

†pEOA347, pUOA522, pEOA246, pEOA264, pEOA273, pEOA306, pEOA265, pEOA106, pEOA338, pEOA986 and pEOA1081.

‡pEOA387, pEOA360, pEOA343, pEOA434 and pEOA390.

other cases of trinucleotide repeats of at least 10-mers are found. Polymorphic variations in trinucleotide repeats have been described in yeast[10]. An example of such variations is given by two ORFs of unknown function, YOR229w and YOR230w, that represent an ancient and diverged tandem duplication (67% sequence identity), with an insertion of a long imperfect trinucleotide repeat in YOR229w that is absent from YOR230w[11].

Iterations of a few longer sequence motifs are also present. The clearest example is probably the near-perfect repeat of the 39 nucleotide-long unique sequence 5′-GAGCCTGATTCTGTGGCAGAAGATGAACCGGA-GACTGAT-3′, which occurs nine times between positions 30,935 and 31,309, at the beginning of YOL155c, an ORF of unknown function shows similarity to glucan-1,4-α-glucosidase. The repeats determine a serine-rich amino-acid sequence. At the end of the same ORF (coordinates 28,905–29,279), another near perfect repeat of a unique long sequence, 5′-CAGTAGTGTATGWYTTNGGRGAARCASTRGTTKCKG-3′, occurs four times. In both cases, degenerated copies of the unique sequence are also found beyond the main repeats.

Of the 561 identified ORFs, 212 (37.8%) correspond to known and functionally characterized genes; all of the others are new. By sequence comparison of their products with general databases, 34 of these (6.1% of the total) show significant homology to proteins of known biochemical and/or physiological function of either yeast or other organisms, and 69 others (12.3% of the total) show weak homology. There are 246 ORFs with products that have either no significant homologue (187 cases or 33.3% of total, among which 28 are questionable) or are homologous to proteins that are themselves of unknown function (59 cases or 10.5% of the total).

Using previously defined criteria[12], 239 ORF products (42.6% of total) are predicted to contain at least one transmembrane span, 170 of which are of unknown function. Two proteins have 12 predicted spans (SCM2, the product of YOL020w, and ALG8, the product of YOR067c), and 43 others have five predicted transmembrane spans or more (only 15 of them are functionally characterized).

Duplications are frequent in the yeast genome, and take several different forms that suggest distinct mechanisms of formation. Comparison of the chromosome XV sequence with the entire yeast genome (including chromosome XV), reveals 12 ‘clusters’ of duplicated sequences that may represent ancestral chromosomal duplications subsequentely modified. Such clusters range in size from 12 kb to 49.5 kb, and contain two to nine ORFs, making a total of 297.5 kb of the chromosome XV sequence (27% of total) that is duplicated on at least eight other chromosomes. Duplications with chromosomes IV and XIV are each represented by two clusters intermingled along the chromosome XV map, suggesting successive events of chromosome duplication or rearrangements. Among the four clusters that contain tRNA genes, there are two cases where homologous tRNA genes are conserved at equivalent positions on the other chromosome, further supporting, in such cases, the hypothesis of ancestral chromosomal duplications.

For chromosome XV, duplications were also examined using a second approach based on the systematic comparison of predicted translation products of all yeast ORFs against all others, followed by sequence alignments and estimation of their significance compared to randomizations, and to the overall distribution of similarity values for the entire yeast genome. A total of 193 different ORFs of chromosome XV (34.4%) were found to have at least one significant homologue in the yeast genome (including 28 pairs on the chromosome XV itself). Half of these have several homologues, forming gene families with various degrees of divergence. Among these, 35 ORFs, nearly all in a subtelomeric location, are members of large families with five partners or more (one of them has up to 26 partners). Large gene families include *HXT* genes, *PAU* genes and *RAS* genes, with more than 15 members each in the entire yeast genome. The chromosome XV ORF homologues include 426 different ORFs from the other chromosomes (7.8% of the yeast genome). Distribution of duplicated ORFs along chromosome XV using this method gives roughly the same results as the first method (Fig. 1).

Chromosome XV contains several local ORF duplications in tandem or inverted orientations. Those showing the highest degree of sequence

conservation are the tandem YOR229w and YOR230w already mentionned (63% amino-acid identity), and the inverted repeat YOR010c (*TIR2*) and YOR009w (62% identity). Other local duplications show significantly greater sequence divergence, suggesting more ancient events; they are represented by the tandems YOL083w and YOL082w (35% identity), YOR285w and YOR286w (37% identity), and YOL048c and YOL047c (38% identity, the latter ORF containing an intron). There also exists some 'local' duplications including pairs of ORFs, such as YOR162c and YOR172w (44% identity), or YOR381w and YOR384w (37% identity), that are separated by a few unrelated ORFs. They have also diverged, and may correspond to ancient local duplications that subsequentely received intervening DNA.

When this work started, 81 genes or markers were genetically mapped to chromosome XV, and 12 others were assigned to it but unmapped[2]. Seven of the unmapped genes and 55 of the mapped genes could be unambiguously assigned to ORFs or tRNA genes of the present sequence on the basis of previous partial sequence data, use of probes or gene function. Two genes, *ts26* and *PTP1*, originally mapped to chromosome XV, belong to chromosomes XII and IV (YLR268w and YDL230w, respectively). One gene, *TIR2* which corresponds to YOR010c, was originally named *SRP1*, creating confusion with YNL189w on chromosome XIV. Other than that, the original genetic map agrees fairly well with the present data, except for some local inversions of gene order, mostly around the centromere (*TOP1* and *SIN3* are on the left arm, *PEP12* is on the right) and in subtelomeric regions (*MEK1*, *PHR1* and *RAD17*), as is also observed for other chromosomes.

The chromosome XV sequence, like that of other yeast chromosomes, has been interpreted using criteria that are essentially predictive for ORFs, but are comparative with previously described sequences for other genetic features such as RNA-coding genes, Ty elements and the various chromosomal elements. It follows that the number of predicted ORFs is probably overestimated by a few percent compared with the number of actual protein-coding genes, whereas the identification of the other features should be considered as a minimum. Clear-cut identification of the questionable ORFs is not possible without independent experimental evidence, but it is suggested that they represent ~6% of all predicted ORFs. The *a posteriori* comparison of the predicted ORF products with general protein database entries forms the basis for the notion of 'orphan' genes[7]. Orphans are those protein-coding genes, predicted from the genomic sequence, that fail to show significant homology (at the chosen threshold value) when their translation products are compared to gene product sequences translated from all other genomic sequences present in public databases, whether they correspond to *S. cerevisiae* or any other organism. As is the case for other yeast chromosomes, a large fraction (33%) of chromosome XV ORFs are orphans without any significant structural homologue; while another 10.5% are orphans with structural homology to one or several other *S. cerevisiae* orphan(s). It is not surprising that, as with other genes, orphans can be duplicated in the yeast genome, or form diverged gene families. But what is more interesting is the significant deficit of these compared with the overall number of gene families in yeast. Of the 193 chromosome XV ORFs that are duplicated or parts of gene families, only 50 are orphans, compared with 83 predicted (43% of the 561 ORFs). The deficit of orphans among gene families (and the correlated excess of functionally characterized genes) is exactly opposite to the classical expectation from standard genetic screenings based on negative mutant phenotypes, which should tend to ignore isofunctional duplicated genes. One possible explanation for the bias observed may be that molecular methods, in contrast to classical genetic screenings, tend to facilitate the isolation of gene families that are structurally but not functionally related. Systematic functional analysis of the yeast genome, which is expected to follow the completion of the genomic sequence, should help to solve this important question. □

Methods

The sequence of chromosome XV was assembled from 46 cosmids covering the entire length of the chromosome except its left and right telomeres, which were sequenced from a rescued plasmid and a polymerase chain reaction (PCR) genomic product, respectively (Table 1). The cosmid map will be described separately. Because chromosome XV comigrates in PFGE with chromosome VII (which is only 348 bp shorter), the technique of chromosome fragmentation based on the insertion of unique artificial I-*Sce*I sites[13] played a key role in the physical mapping of these two chromosomes.

Each segment of the chromosome was sequenced on both strands using the methods and strategies indicated in Table 1. Overlaps between sequences submitted by different laboratories range from 218 bp to 16,799 bp (average 3,765 bp). In nearly all cases, overlapping sequences were also entirely determined on both strands by each laboratory and were found to be 100% identical. Only three differences were found in a total of 90,360 bp. After re-examination of the sequences, only one real divergence remained (an A to G transition), probably resulting from a mutation in one of the sequenced cosmids. The present chromosome sequence includes an A at position 400,735 but a G is equally probable. This uncertainty affects YOR036w (*PEP12*), by changing a codon CAG (Gln) to CGG (Arg).

After assembly, the entire sequence was verified as follows. A total of 201 short segments (259 bp–400 bp long) were selected after examination of the sequence using, as criteria, the possible occurrence of frameshifts, compressions (particularly in G+C-rich regions), and the presence of oligomeric stretches of mono- or dinucleotide repeats. Selected segments were attributed anonymously to four different laboratories and resequenced following the protocol of G. V. (unpublished). In total, 64,370 bp were verified revealing 21 original errors (13 nucleotides omitted, 3 nucleotides in excess, and 5 substitutions). Taken together with overlaps between different laboratories, 14.2% of the chromosome (154,730 bp) has thus been sequenced twice independently. The average sequence accuracy is 99.98%. This figure is probably an underestimate, however, as verifications were directed to suspicious regions. Parts of the present sequence were published independently by the sequencers before assembly of the contig and application of final quality controls[6,8,11,14–30]. Several other manuscripts are also in the press.

Received 23 July 1996; accepted 11 March 1997.

1. Hawthorne, D. C. & Mortimer, R. K. *Genetics* 60, 735–742 (1968).
2. Mortimer, R. K., Contopoulou, R. & King, J. S. *Yeast* 8, 817–902 (1992).
3. Oliver, S. G. *et al. Nature* 357, 38–46 (1992).
4. Dujon, B. *et al. Nature* 396, 371–378 (1994).
5. Feldmann, H. *et al. EMBO J.* 13, 5795–5809 (1994).
6. Gamo, F.-J. *et al. Yeast* 12, 709–714 (1996).
7. Dujon, B. *Trends Genet.* 12, 263–270 (1996).
8. Boyer, J. *et al. Yeast* 12, 1575–1586 (1996).
9. Goffeau A. *et al. Science* 274, 546–567 (1996).
10. Richard, G. F. & Dujon, B. *Gene* 174, 165–174 (1996).
11. Galisson, F. & Dujon, B. *Yeast* 12, 877–885 (1996).
12. Goffeau A. *et al. FEBS Lett.* 325, 112–117 (1993).
13. Thierry, A. & Dujon, B. *Nucleic Acids Res.* 20, 5625–5631 (1992).
14. Aldea, M. *et al. Yeast* 12, 1053–1058 (1996).
15. Bordonné, R. *et al. Yeast* 13, 73–83 (1997).
16. Casamayor, A. *et al. Yeast* 11, 1281–1288 (1995).
17. Casamayor, A. *et al. Yeast* 12, 1013–1020 (1996).
18. Casas, C. *et al. Yeast* 11, 1061–1068 (1995).
19. Chéret, G., Bernardi, A. & Sor, F. *Yeast* 12, 1059–1064 (1996).
20. Lafuente, M. J., Gamo, F-J. & Gancedo, C. *Yeast* 12, 1041–1045 (1996).
21. Madania, A. *et al. Yeast* 12, 1563–1573 (1996).
22. Mannhaupt G. *et al. Yeast* 12, 67–76 (1996).
23. Parle-McDermott, A. G. *et al. Yeast* 12, 999–1004 (1996).
24. Pearson, B. M. *et al. Yeast* 12, 1021–1031 (1996).
25. Purnelle, B. & Goffeau, A. *Yeast* 12, 1475–1481 (1996).
26. Vandenbol, M., Durand, P. Portetelle, D. & Hilger, F. *Yeast* 11, 1069–1075 (1995).
27. Wiemann, S. *et al. Yeast* 12, 281–288 (1996).
28. Zumstein, E., Griffin, H. & Schweizer, M. *Yeast* 10, 1383–1387 (1994).
29. Zumstein, E., Pearson, B. M., Kalogeropoulos, A. & Schweizer, M. *Yeast* 11, 975–986 (1995).
30. Sterky, F. *et al. Yeast* 12, 1091–1095 (1996).

Acknowledgements. This work is part of the third phase of the European Yeast Genome Sequencing Project carried out under the administrative coordination of A. Vassarotti (DG-XII) and the Université Catholique de Louvain. We thank P. Mordant for accounting and help; P. Jordan and S. Liebl for computer support; and S. Oliver and colleagues in the various laboratories for discussions. This work was supported by the European Commission under the Biotech II Program with additional contributions from national or local sources: Services Fédéraux des Affaires Scientifiques, Techniques et Culturelles, Pôles d'attraction inter-universitaire, Région Wallone and Katholieke Universiteit Leuven (Belgium); GREG, CNRS, Institut Pasteur and Institut Curie (France); DGICYT, CICYT and PFPI (Spain) and the Wellcome Trust (UK).

Correspondence and requests for materials should be addressed to B. D. (e-mail: bdujon@pasteur.fr).
Details of the sequence data set are available at http://speedy.mips.biochem.mpg.de/mips/yeast/

The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XVI

H. Bussey[1], R.K. Storms[2], A. Ahmed[3], K. Albermann[4], E. Allen[5], W. Ansorge[6], R. Araujo[5], A. Aparicio[5], B. Barrell[7], K. Badcock[7], V. Benes[6], D. Botstein[5], S. Bowman[7], M. Brückner[8], J. Carpenter[5], J.M. Cherry[5], E. Chung[5], C. Churcher[7], F. Coster[9], K. Davis[5], R.W. Davis[5], F.S. Dietrich[5], H. Delius[10], T. DiPaolo[1], E. Dubois[11,12], A. Düsterhöft[13], M. Duncan[5], M. Floeth[13], N. Fortin[1], J.D. Friesen[3], C. Fritz[13], A. Goffeau[9], J. Hall[1], U. Hebling[10], K. Heumann[4], H. Hilbert[13], L. Hillier[14] and other members of the Genome Sequencing Center[14] , S. Hunicke-Smith[5], R. Hyman[5], M. Johnston[14], S. Kalman[5], K. Kleine[4], C. Komp[5], O. Kurdi[5], D. Lashkari[5], H. Lew[5], A. Lin[5], D. Lin[5], E.J. Louis[15], R. Marathe[5], F. Messenguy[11], H.W. Mewes[4], S. Mirtipati[5], D. Moestl[13], S. Müller-Auer[8], A. Namath[5], U. Nentwich[6], P. Oefner[5], D. Pearson[7], F.X. Petel[5], T.M. Pohl[16], B. Purnelle[9], M.A. Rajandream[7], S. Rechmann[6], M. Rieger[8], L. Riles[14], D. Roberts[5], M. Schäfer[8], M. Scharfe[17], B. Scherens[18,19], S. Schramm[5], M. Schröder[5], A. M. Sdicu[1], H. Tettelin[9], L. A. Urrestarazu[20], S. Ushinsky[2], F. Vierendeels[11], S. Vissers[18], H. Voss[6], S.V. Walsh[7], R. Wambutt[17], Y. Wang[1], E. Wedler[17], H. Wedler[17], E. Winnett[1], W-W. Zhong[1], A. Zollner[4], D.H. Vo[1]* & J. Hani[4]*

[1]Department of Biology, McGill University, 1205 Dr Penfield Avenue, Montreal, H3A 1B1, Canada
[2]Department of Biology, Concordia University, Montreal, H3G 1M8, Canada
[3]Research Institute, Hospital for Sick Children, Toronto, M5G 1X8, Canada
[4]Martinsrieder Institut für Protein Sequenzen, Max-Planck-Institut für Biochemie, D-82152 Martinsried, Germany
[5]Department of Biochemistry, Stanford University, Stanford, California 94305, USA
[6]European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69117 Heidelberg, Germany
[7]The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
[8]Genotype GmbH, D-69259 Wilhelmsfeld, Germany
[9]Unité de Biochimie Physiologique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium
[10]DKFZ, im Neuenheimer Feld 506, D-69120 Heidelberg, Germany
[11]Research Institute of CERIA-COOVI, B-1070, Brussels, Belgium
[12]Laboratoire de Microbiologie de l'Université Libre de Bruxelles, B-1050, Brussels, Belgium
[13]QIAGEN GmbH, Max-Volmer-Strasse 4, D-40724 Hilden, Germany
[14]The Genome Sequencing Center, Department of Genetics, Washington University, School of Medicine, 630 S. Euclid Avenue, St Louis, Missouri 63110, USA
[15]Yeast Genetics, Institute of Molecular Medicine, John Radcliffe Hospital, Oxford OX3 9DU, UK
[16]Gesellschaft für Analyse Technik und Consulting mbH, Fritz-Arnold-Strasse 23, D-78467 Konstanz, Germany
[17]AGON, Gesellschaft für molekularbiologische Technologie mbH, Glienicker Weg 185, D-12489 Berlin, Germany
[18]Vlaams Interuniversitair Instituut voor Biothechnologie, Department Microbiologie, B-1070, Brussels, Belgium
[19]Laboratorium voor Erfelijkheidsleer en Microbiologie van de Vrije Universiteit, Brussels, Belgium
[20]Laboratoire de Physiologie Cellulaire et de Génétique des Levures, Université Libre de Bruxelles, B-1050, Brussels, Belgium
** These authors contributed equally to the work.*

The nucleotide sequence of the 948,061 base pairs of chromosome XVI has been determined, completing the sequence of the yeast genome. Chromosome XVI was the last yeast chromosome identified[1], and some of the genes mapped early to it, such as *GAL4*, *PEP4* and *RAD1* (ref. 2) have played important roles in the development of yeast biology. The architecture of this final chromosome seems to be typical of the large yeast chromosomes, and shows large duplications with other yeast chromosomes. Chromosome XVI contains 487 potential protein-encoding genes, 17 tRNA genes and two small nuclear RNA genes; 27% of the genes have significant similarities to human gene products, and 48% are new and of unknown biological function. Systematic efforts to explore gene function have begun.

There are 487 open reading frames (ORFs) on chromosome XVI, and 10 Ty-related ORFs. Of these ORFs, 17 have an intron, and ORF RPL6B (YPL198w), which encodes a ribosomal protein, has two introns. ORFs were identified using the working definition that they commence with an ATG and have at least a further 99 contiguous sense codons[3], and were analysed using established procedures[4,5]. Before systematic sequencing, there were 73 genes[6], with 47 genes and their relative positions defining the genetic map, and an additional 26 genes located on the physical map. An additional 92 genes have formal genetic names, some of which had been previously cloned but not mapped to this chromosome or have been studied following their identification by systematic sequencing, for a total of 165 known genes. Thus only 33% of the total ORFs found on chromosome XVI had been identified before the completion of the sequence. Other genetic elements include: 17 tRNAs (9 of which are within 500base pairs of a long terminal repeat (LTR) element of a retrotransposon; 5 Ty retrotransposons; 15 delta elements including partial elements; 4 sigma and 2 tau elements; and 2 snRNAs.

The number of ORFs of known function is 194 (40%), of which 76 are functionally characterized proteins; 88 are known proteins that are not fully characterized; 26 have similarity to proteins of known biochemical and physiological function; 6 are homologous to proteins of known biochemical function; and 55 have a weak homology to known proteins. These weak similarities alone are insufficient to confidently assign function. This leaves 236 ORFs of unknown function (48%), of which 50 have homologues of unknown function, and 186 have no similarity to known proteins. There are 36 questionable ORFs, all of which partly overlap another ORF. All have a low codon adaptation index (CAI) of not greater than 0.18, are short (with an average length of 132 codons), and have no known homology with other proteins or are associated with no known phenotype. For four of these ORFs in two pairs (YPL034c and YPL035c, and YPR038w and YPR039w) it is unclear which, if any, are biologically meaningful. There are few apparent pseudogenes but these include YPL276w and YPL275w, which occur together in the genome as a frameshifted pair, both with homology to a formate dehydrogenase. This arrangement has been confirmed by direct sequencing of genomic DNA. Whether this region represents a mutation specific to strain S288C awaits experimental determination.

Functional categories have been compiled for chromosome XVI ORFs: 214 (43%) are classified in some form and 283 remain unclassified. The chromosome is sufficiently large to have a broad representation of all of the predominant functional groups. Detailed global genome classification and statistics for these functional assignments are tabulated and have been discussed elsewhere[7,8].

A robust amino-acid sequence motif is the presence of a predicted transmembrane domain, a region of a protein that spans a lipid bilayer. There are 181 ORFs with one or more predicted transmembrane domains on chromosome XVI (37%)[9], a number close to that seen on chromosomes II and III (refs 3, 10). Of these, 68 are known functionally in some form, and this percentage of known membrane proteins (38%) is similar to that of all known ORFs. Most of these membrane-protein encoding ORFs contain one or two predicted transmembrane domains (99 and 39, respectively); 15 proteins have 7 or more predicted transmembrane domains, and Ypl006p has 13.

An immediate way to use the sequence information is to examine experimentally ORF function, and the Canadian group is compiling systematic transcript and gene-disruption data. For a section encoding 89 ORFs in the region spanned by YPL085w to YPR017c, 61 transcripts were detected in haploid cells grown on rich medium at 30°C. This level of ORF expression is similar to that seen on chromosomes I and III (refs 5, 11). Of the 117 genes on chromosome XVI that have been disrupted, 36

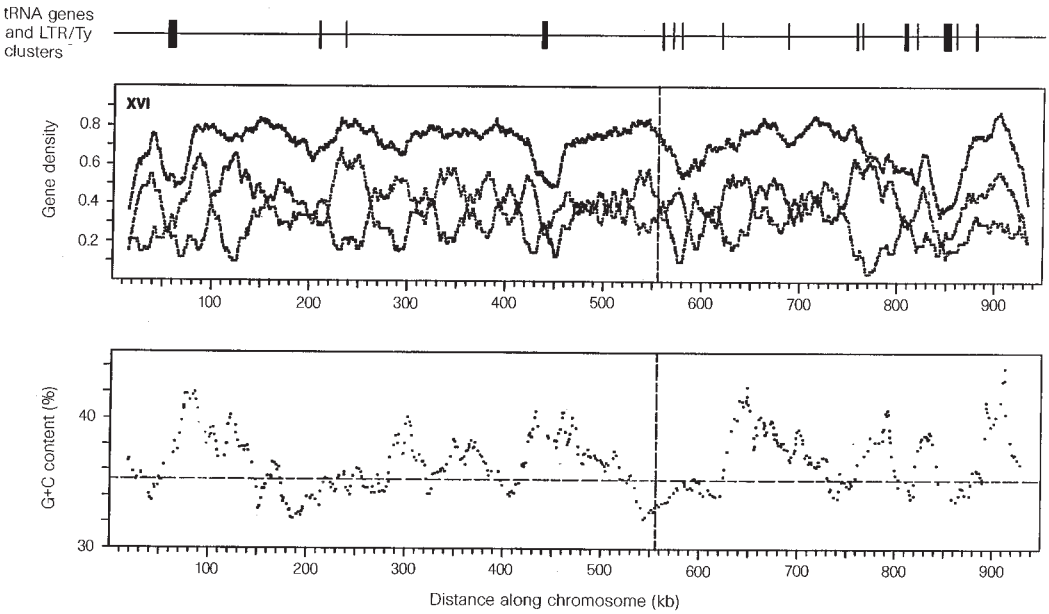


Figure 1 Molecular architecture of chromosome XVI. Top line shows positions of tRNA genes, solo LTR or Ty elements (thin vertical lines) or clusters of them (thick vertical lines) along the chromosome. Panels: variation of gene density (top) and base composition (bottom) along chromosome XVI (scale in kilobases from left telomere). Vertical broken lines indicate position of the centromere. Gene density determined as for chromosome XV is shown for the Watson (medium line) or the Crick (thin line) strands, and the sum of both (thick line). G+C content was calculated as for chromosome XV (horizontal broken line).

(31%) are 'essential' for vegetative growth. Although this fraction is higher than the approximately 12% estimated for the entire genome[12], or found on chromosome I (ref. 5), it still represents only 7% of the ORFs on the chromosome. The final proportion of essential genes on chromosome XVI awaits a complete disruption set. This information is displayed and will be updated (web sites URL http://www.mips.biochem.mpg.de and URL http://genome-www.stanford. edu).

Several ORFs deserve mention, some because of size, some because they were anticipated but not previously found, and others because they occur in metazoans with phenotypes or functions that appear characteristic of the larger multicellular eukaryotes, and whose unexpected presence in yeast affords some insight into function. There are 130 ORFs (27%) with significant similarity to products or predicted products of human genes. Four ORF products (*PAL1, PHO85, ROX1* and *RAD1*) are similar to products of the human genes *ALD, RET, SRY* and *XPF* (*ERCC4*) that, when altered by mutation, lead to X-linked adrenoleukodystrophy, multiple endocrine neoplasia 2A, gonadal dysgenesis, and xeroderma pigmentosum, respectively[13,14]. The largest ORF is YPR117w, specifying a 2,489-codon putative membrane protein of unknown function. Just one other ORF, *SEC18* (YPL085w), has more than 2,000 codons, and only 38 ORFs (8%) are larger than 1,000 codons. Identifying very small ORFs presents special problems[3,15] and just six identified ORFs on chromosome XVI are less than 100 codons. The smallest known functional ORF, *ATP15* (YPL271w), has 62 codons and encodes an F1-ATP synthase epsilon subunit[16]. Other small ORFs undoubtedly exist on the chromosome, and remain to be uncovered by functional analysis.

YPL127c encodes an apparent H1 histone, a protein previously not thought to be present in yeast. Plant and animal H1 histones are known to be involved in the higher-order assembly of nucleosomes but, despite considerable work, the precise role of H1 histone remains unclear[17,18]. Disruption of this single-copy yeast gene indicates that it is not essential for mitotic growth[32]. A detailed analysis of the possible role of this H1 histone in chromosome assembly, stability and the regulation of gene expression can now be explored.

YPR048w encodes a cytochrome P450 protein with similarity to human nitric oxide synthase, NOS. Human NOS is an amino-acid oxidoreductase that oxidizes the terminal N of arginine to give citrulline and nitric oxide. Nitric oxide (NO) in mammals behaves as a hormone, and is involved in many processes including vasodilation and neurological activities[19]. The involvement of NO in yeast metabolism had not previously been imagined. One possibility is that it acts to regulate cell signalling through interaction with small GTP-binding proteins[20].

We found that 74.57% of the chromosomal DNA is involved in the coding of ORFs. These 497 ORFs have an average size of 474 codons, close to the average values seen for other large chromosomes[3,4,15]. Of the 487 ORFs that are not Ty related, 251 are on the Watson strand and 236 are on the Crick strand, with no apparent strand bias[3]. There is, on average, an ORF every 1,908 bp over the chromosome. The positioning of ORFs relative to each other seems random: 111 ORFs are divergent, 118 are convergent, and 195 are tandemly arranged. Of the remaining ORFs, 38 are next to a non-ORF element, and in 12 cases two non-ORF elements are adjacent. With regard to intergenic spacing, the average length between tandemly arranged ORFs is 534 bp, for divergent ORFs it is 569 bp, and for convergent ORFs it is 340 bp. Only one region is apparently devoid of ORFs or other genetic entities for 3 kilobases or more: a subtelomeric 3,863 bp gap between YPL275w and YPL274w. There are no large non-coding regions comparable to those found in the subtelomeric regions of chromosomes I and VIII (refs 5,21).

There are no clustered gene families[5] on chromosome XVI and, as found with other yeast chromosomes, there is little apparent functional clustering of genes. There are two exceptions: a 'syntenic' pair, *CIT3* (YPR001w) and YPR002w, conserving the arrangement of the mmgD and mmgE genes in *Bacillus subtilis* (GenBank accession no. U29084); and two adjacent cyclin genes, *CLB2* (YPR119w) and *CLB5* (YPR120c), an arrangement that is duplicated on chromosome VII as *CLB1* (YGR108w) and *CLB6* (YGR109c), and which forms part of a larger duplication between these two chromosomes (see below).

The G+C periodicity on chromosome XVI varies about an average content of 38.1% (Fig. 1). Gene density also fluctuates along the chromosome, although it shows little correlation with G+C periodicity. Overall the G+C content of coding regions is 39.5%, and for non-coding regions is 33.29%. The centromere lies in a region of low G+C content, and the Ty and tRNA element clusters lie in regions of low gene density; such long-range compositional variations have been discussed elsewhere[22].

The centromere of chromosome XVI spans nucleotide residues 555,952–556,069, between *HAT1* and *CIT3*, making the chromosome

close to being metacentric. Both telomeres seem to be typical in structure. There are 47 potential origins of DNA replication that match the 11-bp ARS element consensus, although the actual autonomously replicating sequences (ARS) used in replicating the chromosome remain to be determined experimentally.

Like much of the yeast genome, regions of chromosome XVI are duplicated. There are some large-scale DNA duplications spanning 25 kb or more[23,24]. The largest of these is on a 129-kb section of the right arm of chromosome XVI, nucleotide coordinates 731,001–860,000, where regions are duplicated onto a 129.5-kb section on the right arm of chromosome VII (nucleotide coordinates 648,001–777,500). Although removed from the comparison that identified this duplication, the region on chromosome XVI is rich in repetitive elements and contains three Ty elements, five additional LTRs and six tRNA genes. Such DNA duplications form large regions of partial gene synteny between these two chromosomes. An example is a section from nucleotide 834,000 to 860,000 on chromosome XVI and from 762,422 to 777,500 on chromosome VII. In the chromosome XVI interval from YPR154 to YPR159, four of the six genes, two delta elements and two tRNA genes, are syntenic with their chromosome VII counterparts, YGR136 toYGR143, with the exception of a tandem Ty1 element inserted between YPR158 and YPR159 on chromosome XVI. It has been suggested[25] that the origin of the *KRE6* (YPR159w)/*SKN1*(YGR143w) pair in this region resulted from some event involving duplication through transposition of the retrotransposon or tRNA elements. The origin of these major cluster homology regions remains unclear but probably consisted of more than one event; the subject is discussed elsewhere at a global genome level[7]. Evidence for duplications can also be found at the ORF level; 125 of the ORFs on chromosome XVI have one or more counterpart with significant similarity in the yeast genome. In addition to duplications, chromosome XVI also contains members of larger gene families, with at least 38 ORFs having similarity to two or more yeast ORFs. An example is *KTR6* (YPL053c), a member of a family of nine mannosyltransferase-encoding genes located on eight different chromosomes[26].

Establishment of the genetic and physical maps were critical precursors to obtaining the nucleotide sequence of the chromosomes. The genetic map of chromosome XVI is 251 cM in length, giving an average cM/kb value of 0.26, the smallest seen for a yeast chromosome, and close to that previously reported[27]. It indicates that a chromosome XVI bivalent has, on average, ten crossovers per meiosis. Although the genetic map is generally correct, there are several discrepancies with the positions found on the sequenced chromosome. Of the 46 genes mapped genetically through the phenotypes of alleles, 14 (*spoT16*, *rad53*, *nib1*, *sot1*, *SUF21*, *dna1*, *mak6*, *tsm0120*, *ymc1*, *spoT20*, *SUP15*, *SUP16*, *cdc67* and *rad56*) remain to be identified on the sequenced chromosome.

Determination of the nucleotide sequence of chromosome XVI completes the sequence of the yeast genome, allowing a genome-wide analysis of a small and experimentally amenable eukaryotic organism. Such systematic studies should enhance our knowledge of cell function, and help us to understand the structure and function of eukaryotes with larger genomes. □

Methods

Chromosome XVI was sequenced using a set of overlapping cosmid and lambda clones based on a previous chromosome XVI physical map[28] (L. Riles and M.V. Olson, personal communication) from the S288C-derived strain AB972. Two gaps on the right arm and both telomeres were sequenced using polymerase chain reaction (PCR) products amplified from genomic DNA[29]. Individual cosmids representing the portion mapped by the EU group were sequenced by contracting laboratories using a variety of subcloning, sequencing and assembly methods[3,4]. Explanations of the cloning, sequencing, assembly and quality control methods used by the other groups have been described[21,30]. The telomeres were sequenced using specially devised procedures[31]. Assembly of the completed chromosome sequence was made in Martinsried or Montreal as described[3–5,15]. Determination of overlapping sequence between groups indicated that seven differences were found and resolved in 81 kb of sequence . This allowed us to estimate the accuracy of the sequence to be conservatively within the three errors per 10 kb average for the yeast genome[7].

Received 29 July 1997; accepted 11 March 1997.

1. Hawthorne, D.C. & Mortimer, R.K. *Genetics* 60, 735-742 (1968).
2. Broach, J.R. in *The Molecular Biology of the Yeast* Saccharomyces, *Life Cycle and Inheritance* 653–727 (Cold Spring Harbor Laboratory Press, NY, 1981).
3. Feldmann,H. *et al. EMBO J.* 13, 5795–5809 (1994).
4. Galibert, F. *et al. EMBO J.* 15, 2031–2049 (1996).
5. Bussey, H. *et al. Proc. Natl Acad. Sci. USA* 92, 3809–3813 (1995).
6. Mortimer, R.K, *et al.* http://genome.www.stanford.edu/saccdb/edition12.html (1995).
7. Goffeau, A. *et al. Science* 274, 546 (1996).
8. Goffeau, A. *et al. Science* 274, 563–567 (1996).
9. Klein, P., Kanehisa, M. & Delesi, C. *Biochem. Biophys. Acta* 815, 468–476 (1985).
10. Goffeau, A., Nakai, K., Slonimski, P.P. & Risler, J.L. *FEBS Lett.* 325, 112–117 (1993).
11. Tanaka, S. & Isono K. *Nucleic Acids Res.* 21, 1149–1153 (1992).
12. Goebl, M.E. & Petes, T.D. *Cell* 46, 983–992 (1986).
13. Bassett, D.E., Boguski, M. & Hieter, P. *Nature* 379, 589–590 (1996).
14. van Vuuren, A.J. *et al. EMBO J.* 12, 3693– 3701 (1993).
15. Dujon, B. *et al. Nature* 369, 371–378 (1994).
16. Guelin, E. *et al. J. Biol. Chem.* 268, 161–167 (1993).
17. Zlatanova, J. & Doenecke D. *FASEB J.* 8, 1260–1268, (1994).
18. Sirotkin, A.M.*et al. Proc. Natl Acad. Sci. USA* 92, 6434–6438 (1995).
19. Forstermann, U. *et al. Hypertension* 23, 1121–1131(1994).
20. Lander, H.M. *et al. Nature* 381, 380–381 (1996).
21. Johnston, M. *et al. Science* 265, 2077–2082 (1994).
22. Dujon, B. *et al. Trends Genet.* 12, 263–270 (1996).
23. Heumann, K. & Mewes, H.W. *Nature Genet.* (submitted).
24. Heumann, K., Harris, C. & Mewes, H.W. in *Proc. Fourth Int. Conf. Intelligent Systems for Mol. Biol.* (St Louis, MO, 1996).
25. Roemer, T., Fortin, N. & Bussey, H. *Yeast* 10, 1527–1530 (1994).
26. Lussier, M., *et al. Yeast* 13, 267–274 (1997).
27. Mortimer, R.K., Contopoulou, C.R. & King, J.S. *Yeast* 8, 817–902 (1992).
28. Riles, L. *et al. Genetics* 134, 81–150 (1993).
29. Louis, E.J. & Borts, R.H. *Genetics* 139, 125–136 (1995).
30. Wilson, R. *et al. Nature* 368, 32–38 (1994).
31. Louis, E.J. *Biochemica* 3, 25–26 (1995).
32. Ushinsky, S. C. *et al. Yeast* 13, 151–161 (1997).

Acknowledgements. We thank D. Oki for administrative assistance and manuscript preparation; C. Jallet, D. Jauniaux, M. Vanden Barre and A. Wong for technical help; M. Lussier, T. Ketela and B. Baum for input; and B. Dujon for preparation of the G+C periodicity figure, using unpublished software developed in collaboration with C. Marck. This work was supported by the Canadian Genome Analysis and Technology Program, European Yeast Genome Sequencing Project, Human Genome Organization, The Wellcome Trust, Région Wallone, and the Region of Brussels Capital.

Correspondence and requests for materials should be addressed to H.B. (e-mail: hbussey@ monod.biol.mcgill.ca)