Evolutionary relationships between *Saccharomyces cerevisiae* and other fungal species as determined from genome comparisons

Enrique Herrero

Departamento de Ciencias Médicas Básicas, Facultad de Medicina, Universitat de Lleida, Lleida, Spain

Summary

The increasing number of fungal genomes whose sequence has been completed permits their comparison both at the nucleotide and protein levels. The information thus obtained improves our knowledge on evolutionary relationships between fungi. Comparison of the *Saccharomyces cerevisiae* genome with other Hemiascomycetes genomes confirms that a whole-genome duplication occurred before the diversification between *Candida glabrata* and the *Saccharomyces sensu stricto* species and after separation from the branch leading to the other Hemiascomycetes. Duplication was followed by individual gene losses and rearrangements affecting extensive DNA regions. Although *S. cerevisiae* and *C. glabrata* are two closely related yeast species at an evolutionary scale, their different habitats and life styles correlate with specific gene differences and with more extensive gene loses having occurred in the parasitic *C. glabrata.* At a closer evolutive scale, diversification among the *sensu stricto* species began with nucleotide changes at the intergenic regions affecting sequences that are not relevant for gene regulation, together with more extensive genome rearrangements involving transposons and telomeric regions. One important characteristic of fungal genomes that is shared with other eukaryotes is the fusion of gene sequences coding for separate protein modules into a single open reading frame. This allows diversification of protein functions while saving gene information.

Key words

Saccharomyces cerevisiae, Candida, Hemiascomycetes, Comparative genomics, Genome duplication, Protein modules

Relaciones evolutivas entre *Saccharomyces cerevisiae* y otras especies fúngicas establecidas mediante comparaciones genómicas

Resumen

El creciente número de genomas fúngicos cuya secuencia se ha completado permite su comparación tanto a nivel de nucleótidos como de la proteína. La información obtenida de este modo mejora nuestro conocimiento sobre las relaciones evolutivas entre hongos. La comparación del genoma de *Saccharomyces cerevisiae* con el de otros Hemiascomicetes confirma que tuvo lugar una duplicación del genoma entero en un antecesor antes de la diversificación entre *Candida glabrata* y las especies *Saccharomyces sensu stricto,* y después de la separación respecto de la rama que condujo a otros Hemiascomicetes. La duplicación vino seguida de pérdidas de genes individuales así como de reordenaciones más extensas del DNA. Aunque *S. cerevisiae* y *C. glabrata* son dos especies de levaduras relativamente próximas a una escala evolutiva, sus diferentes hábitats y estilos de vida se correlacionan con diferencias genéticas específicas y con la existencia de pérdidas más numerosas de genes en la especie parásita *C. glabrata.* A una escala evolutiva más próxima, la diversificación entre las especies del grupo *sensu stricto* empezó con cambios nucleotídicos en las regiones intergénicas que afectarían secuencias no relevantes para la regulación génica, junto con reordenaciones más extensas que implicarían transposones y

Corresponding address:
Dr. Enrique Herrero
Departamento de Ciencias Médicas Básicas
Facultad de Medicina
Universitat de Lleida
Montserrat Roig 2
25008 Lleida, Spain
Tel.:+34 973 702 409
Fax: +34 973 702 426
E-mail: enric.herrero@cmb.udl.es

Apdo. 699, E-48080 Bilbao (Spain)
1130-1406/01/10.00 Euros

regiones teloméricas. Una característica importante de los genomas fúngicos que ocurre también en otros eucariotas es la fusión de secuencias génicas que codifican módulos proteicos individuales en una única pauta de lectura. Ello permite la diversificación de las funciones proteicas al mismo tiempo que se ahorra información genética.

Palabras clave *Saccharomyces cerevisiae, Candida,* Hemiascomycetes, Genómica comparada, Duplicación genómica, Módulos proteicos

Completion of the *Saccharomyces cerevisiae* genome sequence in 1996 [12] opened new approaches to the study of evolution of eukaryotic organisms, among other merits of such scientific achievement. Annotation of the genes from the DNA sequence revealed that the function of about 40% of them was totally or partially unknown at that time. Less than ten years later, much more is known on the function of the about 5,800 genes of *S. cerevisiae*, thanks to the focused work on individual genes and their products, but also on whole genome studies such as those on protein location [16], protein complexes [15,28], protein abundance [11] or gene expression in response to external stimuli [10]. These are only a few examples of large scale studies that have been possible only after sequencing and annotation of the whole *S. cerevisiae* genome, and that have opened the path to similar studies in other organisms. A new scientific area (systems biology) has emerged based on systemic experimental approaches as those initially employed with *S. cerevisiae.* A large amount of information on the above studies and on the function of the individual *S. cerevisiae* genes can be obtained in www.yeastgenome.org (Saccharomyces Genome Database). However, it should be kept in mind that in spite of such efforts, more than 1,000 genes of this organism still are classified in databases as being of unknown function.

By the time this review is written (November 2005), eight fungal genomes are defined as completely sequenced in the GeneBank server (www.ncbi.nlm.nih.gov), although the sequence of several other fungal species are essentially completed too. Overall, the information on the full sequence of more than 20 fungal genomes is accessible. They include members of the Archiascomycetae, Hemiascomycetae and Basidiomycetae classes, and therefore they cover an evolutionary range that could extend more than 1,000 million years [2,13]. This amount of information allows the comparison of fungal genomes covering short phylogenetic distances (for instance, within Hemiascomycetes) or larger distances that include the other two classes (Table). Comparative genomics has merged as a new discipline that in the case of fungi allows a better understanding of species evolution and helps to explain the different life styles that can be found among them, besides introducing new information useful from a medical or technological point of view. Comparison of the genomes of two species that have so much diverged along evolution as *S. cerevisiae* and *Cryptococcus neoformans* demonstrates that they share at least 65% of the genetic information [22]. Obviously, some genetic traits such as those coding for capsule synthesis and other possible virulence factors (as melanin formation) are specific of *C. neoformans*. This high level of conservation is remarkable for two species that diverged in the evolutionary tree about 1,000 million years ago [13].

The *S. cerevisiae* genome results from a massive genome duplication and extensive gene loss

Analysis of the sequence of the 16 chromosomes of the *S. cerevisiae* genome revealed the existence of numerous pairs of chromosomal homologous regions. On these basis it was postulated that a whole-genome duplication had occurred in an ancestry of *S. cerevisiae* [32]. This duplication event would have been followed by extensive gene loss, which in general would have affected only one of the two members of the generated gene pairs (Figure 1), as well as gene inversions and transpositions among other events affecting individual genes. The result would be the present *S. cerevisiae* genome, where individual genes without homologues in the own genome coexist with families formed by two paralogous genes that are located in sister chromosomal regions (large blocks of homologous genes) at two different chromosomes [33]. Evolution after the duplication event would explain why only a fraction of the *S. cerevisiae* genome shows the remnants of such duplication. However, 56 of such blocks still can be detected in the *S. cerevisiae* genome. More recently, it has been described the partial sequencing of the genome of 13 Hemyascomycetes, including several *Saccharomyces sensu estricto* species [25], and the complete sequencing of the genomes of *Eremothecium* (*Ashbya*) *gossypii* [6], *Kluyveromyces waltii* [19], *Candida glabrata, Kluyveromyces lactis, Debaryomyces hansenii* and *Yarrowia lipolytica* [8]. By comparing the genomes of these species representing a broad evolutionary range within Hemiascomycetes, this has allowed to confirm that existence of such ancient duplication, which occurred after separation of the *K. waltii* and *Saccharomyces-C. glabrata* branches, but before diversification of the *Saccharomyces sensu stricto* species from *C. glabrata* (Figure 2). In fact, duplication remnants are present in the *C. glabrata* genome and many block regions in the latter display homology with blocks in the *S. cerevisiae* genome [7]. However, the fact that the total number of blocks in *C. glabrata* (twenty) is lower than in *S. cerevisiae* indicates a more extensive gene loss affecting one of the two paralogues in *C. glabrata* [8], and is in accordance with the parasitic life style of this species. At this respect, the total gene number in *C. glabrata* is lower than in the other Hemiascomycetes, although this does not parallels a lower genome size or chromosome number (Table). On the contrary, the number of chromosomes in *S. cerevisiae* and *C. glabrata* is significantly larger than in the other species, which is probably also a reflection of the ancient duplication affecting these two species and their common evolutionary origin.

A consequence of gene duplication is the possibility for functional redundancy, with its implications on phenotypic stability. This is for instance the case of two of the G1 cyclins of the *S. cerevisiae* cell cycle, Cln1 and

Table. General characteristics of completely sequences fungal genomes .

Species	Genome size (Mb)	Number of chromosomes	Total ORFs	Total tRNA genes	Data from Ref.
S. cerevisiae	12.1	16	5807	274	8
C. glabrata	12.3	13	5283	207	8
K. waltii	10.7	8	5230	240	19
K. lactis	10.6	6	5329	162	8
E. gossypii	9.2	7	4718	199	6
D. hansenii	12.2	7	6906	205	8
C. albicans	14.8	8	6419	ND[1]	22
Y. lipolytica	20.5	6	6703	510	8
N. crassa	38.4	7	1082	424	22
S. pombe	12.6	3	4973	174	22
C. neoformans	19.0	14	6594	141	22

[1]ND: not determined

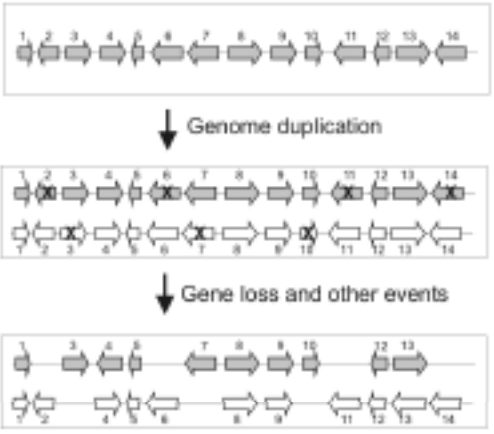

Figure 1. Scheme of the duplication of a genome sequence followed by loss of marked genes. Inversions and other possible events are not shown.

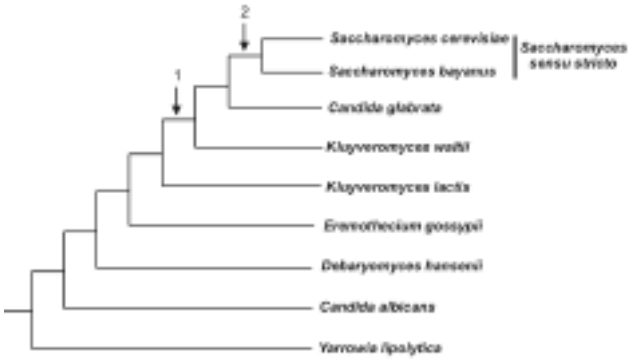

Figure 2. Phylogenetic tree of Hemiascomycetes, showing the branch points of representative species. The genomes of the indicated species have been totally or partially sequenced. Branch lengths are not proportional to evolutionary distances. 1: whole-genome duplication; 2: emergence of the *Saccharomyces sensu stricto* group with multiplication of sugar utilization genes. The tree is based on references 7 and 8.

Cln2, although even in this case evolution may have selected for some particularities in the function of each pair member [9]. In other cases, each family member has become functionally specialized by acting at a different cellular compartment from other members of the family. Enzymes involved in the defence against oxidative stress in *S. cerevisiae* offer a number of examples on this situation [27]. Thus, *S. cerevisiae* contains two cytosolic thioredoxins (Trx1, Trx2) plus one mitochondrial one (Trx3); all three proteins show significant homology in amino acid sequence. Of the two dithiol glutaredoxins, Grx1 is exclussively cytosolic while its homologue Grx2 shares a cytosolic and a mitochondrial location. Interestingly, *S. cerevisiae* has a peroxisomal omega class glutathione transferase (Gto1) that is induced under oxidative stress conditions, plus two homologues (Gto2, Gto3) located at the cytosol (our unpublished results). Among the fungal genomes sequenced to now, only the *S. cerevisiae*-closely related species *Saccharomyces paradoxus* has a predicted peroxisomal Gto1 orthologue; other fungal species contain a single Gto homologue, probably located at the cytosol. Gene duplication in the *Saccharomyces* evolutionary line has therefore led to a new enzyme activity at an organelle such as the peroxisome, as a distinctive trait of *S. cerevisiae* and close relatives in contrast to other fungi. Parallel loss of peroxisomal glutathione transferases in different evolutionary lines from a fungal ancestor that contained such enzyme could also have led to the present situation, but this seems a less plausible hypothesis, as it would require a larger number of independent genetic changes. The presence of a glutathione transferase protecting against reactive oxygen species generated at the peroxisome could explain the acquisition of some metabolic traits by *S. cerevisiae* peroxisomes compared to other fungal species, such as the participation in lysine metabolism [3].

Combination of protein modules leads to functional diversification

Evolution of new biological functions by new combinations of previously existing protein modules very probably has been a motor for evolution in eukaryotes [14]. *S. cerevisiae* offers a number of examples of this situation. We have studied a family of monothiol glutaredoxins that are characterized by the presence of a single cysteine residue at the active site, in contrast with classical glutaredoxins, which contain two cysteines at the active site [1,24,29]. *S. cerevisiae* has three monothiol glutaredoxins. One of them (Grx5) is localized at the mitochondrial matrix and is involved in the synthesis of iron/sulfur clusters, while Grx3 and Grx4 are nuclear [24]. Besides their differential location, these Grx molecules have another difference. While Grx5 contains a single glutaredoxin domain, Grx3 and Grx4 result from the fusion into a

single molecule of an N-terminal thioredoxin (Trx)-like domain plus a C-terminal glutaredoxin (Grx) domain (Figure 3a). In these molecules the Trx-like domain is important for nuclear localization but is does not display a thioredoxin enzyme activity, since it lacks one of the active site cysteine essential for such activity [24]. Therefore, the Grx3 and Grx4 molecules may have resulted from a fusion event of preexisting thioredoxin and glutaredoxins molecules followed by loss of the first enzyme activity. Such Trx-Grx molecules would have adopted biological functions different from their predecessors, which at least in *S. cerevisiae* could consist in the redox regulation of transcriptional factors at the nucleus [23]. The Grx3/Grx4 homologue in human cells is the PICOT protein, that could be a modulator of the protein kinase C activity [31]. The coexistence in a single organism of glutaredoxins with the single Grx structure and those with the mixed Trx-Grx structure is characteristic of eukaryotes but not of prokaryotes, where only single Grx domain monothiol glutaredoxins exist [29]. With respect to Grx5, its biological role in the formation of iron-sulfur clusters at the mitochondria seems to be evolutionarily conserved from bacteria to higher eukaryotes, as homologues from different origins are able to substitute for the Grx5 function when targeted to *S. cerevisiae* mitochondria [30, and our unpublished observations].

To determine the distribution of both Trx-Grx and single Grx monothiol glutaredoxins among fungal species, we made BLASTA searches (using *S. cerevisiae* Grx5 or Grx3 as respective queries) against all the protein sequences of fungal species whose genomes are considered as completely sequenced according to GeneBank. Among the eight genomes, only *D. hansenii* and *Encephalitozoon cuniculi* lacked Grx5 homologues, that is, monothiol glutaredoxins with a single Grx domain (Figure 3b). It is not easy to explain the case of the halophilic yeast, which probably has lost the Grx5 protein while its function in the mitochondrial formation of iron-sulfur clusters being substituted by some other thiol oxidoreductase activity. In the case of *E. cuniculi*, this is an obligate intracellular parasite that has experimented extensive genetic reduction (2.9 megabase genome, 1,997 potential protein-coding genes) and mitochondrial lost, as a general characteristic of fungi-derived microsporidia [18]. This genetic loss can explain the absence of genes coding for both Grx and Trx-Grx glutaredoxins in its genome. In fact, *E. cuniculi* is the only among the eight fungal genomes that lacks a gene for a Grx3 homologue (Figure 3b). As *S. cerevisiae* with Grx3 and Grx4, *C. glabrata* also has two Trx-Grx glutaredoxins, while the other species have a single one. The fact that glutaredoxins with the Trx-Grx structure are present both in Ascomycetae and Basidiomycetae (this study) but also in animals and plants [1,29] support the idea that this structure appeared in the eukaryotic lineage before the amimals-plants-fungi divergence, that is more than 1,600 million years ago [13].

These hybrid molecules offer the possibility to study the coevolution of their domains as compared to the evolution of the original independent molecules. To compare the monothiol glutaredoxins among fungi, we extended the study to *Neurospora crassa, Schizosaccharomyces pombe, Aspergillus nidulans* and *C. albicans* in addition to the species listed in figure 3b. The genomes of these four species code for both *S. cerevisiae* Grx5 and Grx3/Grx4 homologues. All of the ten molecules analyzed with a single Grx domain (including Grx5) have an N-terminal region compatible with a mitochondrial targeting sequence, as revealed by the application of localization prediction analysis programmes. It can therefore be concluded that the mitochondrial localization of *S. cerevisiae* Grx5 can be extended to its fungal homologues, and also to the correspondent higher eukaryotes molecules [30]. Multiple sequence alignment of Grx5 and its nine fungal homologues using ClustalW results in a phylogram tree (Figure 4a) that groups the sequences in a manner that basically parallels the evolutive relationships between the ten fungal species [7]. The only exception is the *C. albicans* Grx5 homologue, that appears well separated from the other sequences in the tree, not becoming associated in the same cluster with other Ascomycetous sequences. As expected from the proximity between *S. cerevisiae* and *C. glabrata*, the closest sequence to *S. cerevisiae* Grx5 is that of the *C. glabrata* orthologue.

Phylogram trees after ClustalW multiple alignments were generated from Grx3 and Grx4 and their fungal homologues, separately for their C-terminal Grx (Figure 4b) and N-terminal Trx-Grx (Figure 4c) domains. When the analysis was done on the whole amino acid sequence including both domains, the generated tree was similar to that based on the Grx domain alone (not shown). The tree resulting from the comparison of the Grx domains grouped the sequences in a manner that was also parallel to the evolutionary relationships among the compared species. The Grx domains of Grx3 and Grx4 group together and well separated from the other sequences, supporting that the two glutaredoxins are the result of the genome duplication occurring at the *Saccharomyces* line. Surprisingly, one of the two *C. glabrata* Grx sequences (corresponding to SwissProt entry Q6FKF5) does not position close to the other *C. glabrata* sequence and to Grx3/Grx4 (Figure 4b), as would have been expected from the evolutionary relationship between both species and the genome duplication having occurred in a common ancestor. This observation suggests that the gene coding

a

b

Species	Glutaredoxins with Grx structure	Glutaredoxins with Trx Grx structure
S. cerevisiae	Yes (Grx5)	Yes (Grx3, Grx4)
C. glabrata	Yes (Q6FJD0)	Yes (Q6FSS7, Q6FKF5)
D. hansenii	No	Yes (XP459996)
E. gossypii	Yes (Q75A65)	Yes (Q74ZT7)
K. lactis	Yes (Q6CVT2)	Yes (Q6CSU2)
Y. lipolytica	Yes (XP505699)	Yes (Q6CES4)
C. neoformans	Yes (Q5KLM7)	Yes (Q5KJR8)
E. cuniculi	No	No

Figure 3. Presence of monothiol glutaredoxins with the Grx and Trx-Grx structure among fungi. (a) General structure of both types of monothiol glutaredoxins, with the sequence reminiscent of thioredoxin active sites at the Trx domain and the glutaredoxin (thiol oxidireductase) active site at the Grx domain. (b) Monothiol glutaredoxins with Grx or Trx-Grx structure present in the eight fungal genomes totally sequenced as indicated at GeneBank (http://www.ncbi.nlm.nih.gov). Homology searches were done using BLASTP at the GeneBank server, with *S. cerevisiae* Grx5 (for Grx molecules) or Grx3 (for Trx-Grx molecules) as query. Only those proteins that resulted in significant alignments along at least 80% of query and subject sequences were considered positive (with E values higher than 10-5). SwissProt entries of the respective proteins are indicated into parenthesis, except for *D. hansenii* Trx-Grx protein and *Y. lipolytica* Grx protein, for which the GeneBank entry is indicated.

for the Q6FKF5 protein may have not resulted from such duplication but from a different genetic event, maybe from horizontal transfer of foreign DNA. Alignment of the Trx domains alone (Figure 4c) confirms the separation between the two *C. glabrata* Trx-Grx molecules. With this exception, the Trx regions also become grouped basically as expected from the phylogenetic relationships among the studied species, the Trx regions of the *C. neoformans, S. pombe, N. crassa* and *A. nidulans* molecules being the most distant from the *S. cerevisiae* homologues (Figure 4c). Branches lengths are larger for the Trx tree than for the Grx tree. Although these comparisons must be taken with caution, the fact may indicate that variation along evolution has affected more intensely to the Trx region. It must be considered that the enzymatic activity of the molecule resides in the Grx region, which could suppose an important restriction for amino acid changes, while the Trx region is necessary for nuclear targeting [24].

Summarizing, the example of the monothiol glutaredoxins indicates that these genomic comparisons may shed some light on evolution of multidomain molecules. In this particular case, our observations indicate that both Trx and Grx domains have evolved closely among them once the hybrid molecule was formed in an old eukaryotic ancestor, and that this evolution has occurred separately from the monothiol glutaredoxins with a single Grx universally present from bacteria to humans. This does not exclude the possibility of other events such as horizontal transfer as that suggested for *C. glabrata*.

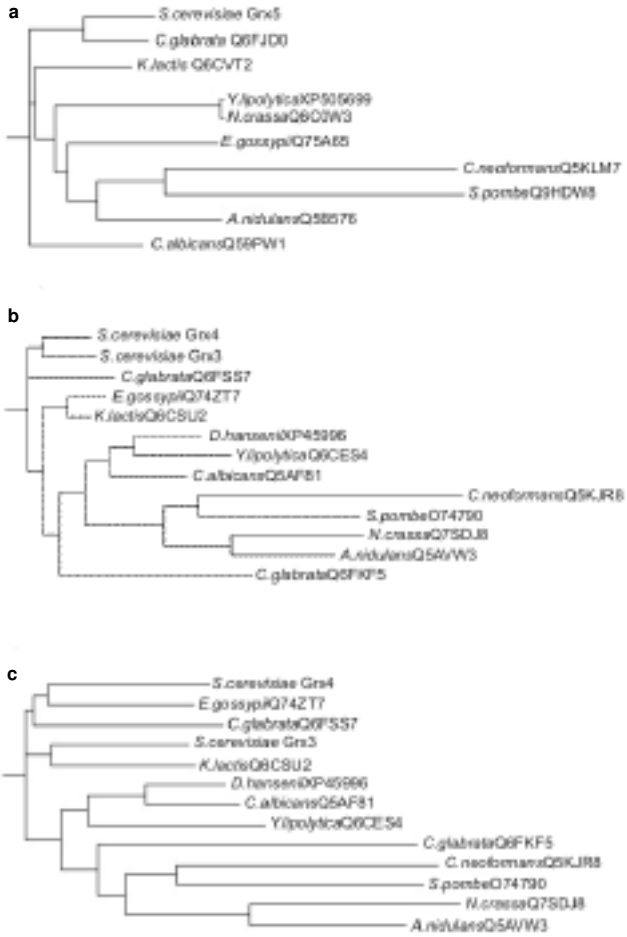

Figure 4. Philogram trees after ClustalW multiple alignments of (a) *S. cerevisiae* Grx5 and the indicated homologues, excluding the respective mitochondrial targeting sequences (first 29 amino acids in Grx5, Ref. 24), (b) the C-terminal Grx domains of the indicated proteins (from amino acid 199 to 283 in Grx3), and (c) the N-terminal Trx domains of the indicated proteins (from amino acid 38 to 133 in Grx3). Protein nomenclature is as in figure 3.

Evolution at a short-time scale within the genus *Saccharomyces*

The above considerations concern evolution within fungi at a long-time scale, that is, in a period of about 1,000 million years [13]. How is evolution operating at a shorter time scale, for instance within the genus *Saccharomyces*? Two recent studies [5,20] address this question within the *Saccharomyces sensu stricto* group, that accumulates an estimated 5-20 million years of separate evolution [4,21]. Partial sequencing of the genomes of the *S. paradoxus, Saccharomyces mikatae, Saccharomyces kudriavzevii* and *Saccharomyces bayanus* species and comparison with the *S. cerevisiae* genome sequence [5,20] allowed establishing that sinteny (that is, gene position and orientation relative to neighbour genes) is conserved for most of the compared *Saccharomyces* genomes. Nucleotide changes are more frequent in intergenic regions than inside genes (about twice the rate in the former), as expected from the fact that these intergenic regions have fewer restrictions for maintaining mutations than coding regions where many nucleotide changes would lead to non-functional amino acid changes. Even in intergenic regions, nucleotide changes have not occurred homogeneously. This has allowed characterizing conserved intergenic sequences that may correspond to promoter regulatory motifs less prone for evolutionary changes than other sequences, for instance those downstream of 3' ends of genes. The observation demonstrates the existence of a selective pressure that acts conservatively at gene promoters. With respect to coding sequences, there exists a high level of conservation among the *sensu stricto* yeast species and most changes at the nucleotide level have occurred at the third position among synonymous codons, that is, they are conservative [20]. More extensive rearrangements (reciprocal translocations, inversions, segment duplications) are much less frequent and have occurred preferentially at telomeric regions and near transposon-like (Ty) elements [20]. Transposons and telomeres therefore reveal as a leading cause of genome evolution. In spite of the high level of genome conservation between the *sensu stricto* species, there exist a low number of genes that are species-specific [4,20]. They may have resulted from recombination events involving more distant species. A large proportion of these species-specific genes have a metabolic function, particularly sugar metabolism. Remarkably, a few genes exist in the four analyzed species that are subjected to a change rate significantly higher or lower than the average; the results are protein products with a considerable number of amino acid positions changes or on the contrary, with extreme amino acid conservation [20]. An example of the latter is the case of the mating-type gene *MATa2*, which has perfect 100% amino acid and nucleotide conservation across all four species over the entire length (119 amino acids).

Although some of the above genome changes are still difficult to explain on a molecular basis, they shed new light on the mechanisms of genome evolution that may be applicable to other organisms. Again, genomic studies on *S. cerevisiae* and related species may open the way for analysis of the evolution of biological systems.

References

1. Bellí G, Polaina J, Tamarit J, de la Torre MA, Rodríguez-Manzaneque MT, Ros J, Herrero E. Structure-function analysis of the yeast Grx5 monothiol glutaredoxin defines essential amino acids for the function of the protein. J Biol Chem 2002; 277: 37590-37596.
2. Berbee ML, Taylor JW. Systematics and evolution. In: McLaughlin DJ, McLaughlin EG (Eds) The Mycota VIIB. Berlin, Spinger, 2001: 229-245.
3. Breitling R, Sharif O, Hartman ML, Krisans SK. Loss of compartmentalization causes misregulation of lysine biosynthesis in peroxisome-deficient yeast cells. Eukaryot Cell 2002; 1: 978-986.
4. Cliften PF, Hillier LW, Fulton L, Graves T, Miner T, Gish WR, Waterston RH, Johnston M. Surveying *Saccharomyces* genomes to identify functional elements by comparative DNA sequence analysis. Genome Res 2001; 11: 1175-86.
5. Cliften P, Sudarsanam P, Desikan A, Fulton L, Fulton B, Majors J, Waterston R, Cohen BA, Johnston M. Finding functional features in *Saccharomyces* genomes by phylogenetic footprinting. Science 2003; 301: 71-76.
6. Dietrich FS, Voegeli S, Brachat S, Lerch A, Gates K, Steiner S, Mohr C, Pöhlmann R, Luedi P, Choi S, Wing RA, Flavier A, Gaffney TD, Philippsen P. The *Ashbia gossypii* genome asa tool for mapping the ancient *Saccharomyces cerevisiae* genome. Science 2004; 304: 304-307.
7. Dujon B. Hemiascomycetous yests at the forefront of comparative genomics. Curr Opin Gen Dev 2005; 15: 1-7.
8. Dujon B, Sherman D, Fischer G, Durrens P, Casaregola S, Lafontaine I, De Montigny J, Marck C, Neuveglise C, Talla E, Goffard N, Frangeul L, Aigle M, Anthouard V, Babour A, Barbe V, Barnay S, Blanchin S, Beckerich JM, Beyne E, Bleykasten C, Boisrame A, Boyer J, Cattolico L, Confanioleri F, De Daruvar A, Despons L, Fabre E, Fairhead C, Ferry-Dumazet H, Groppi A, Hantraye F, Hennequin C, Jauniaux N, Joyet P, Kachouri R, Kerrest A, Koszul R, Lemaire M, Lesur I, Ma L, Muller H, Nicaud JM, Nikolski M, Oztas S, Ozier-Kalogeropoulos O, Pellenz S, Potier S, Richard GF, Straub ML, Suleau A, Swennen D, Tekaia F, Wesolowski-Louvel M, Westhof E, Wirth B, Zeniou-Meyer M, Zivanovic I, Bolotin-Fukuhara M, Thierry A, Bouchier C, Caudron B, Scarpelli C, Gaillardin C, Weissenbach J, Wincker P, Souciet JL. Genome evolution in yeasts. Nature 2004; 430: 35-44.
9. Futcher B. Transcriptional regulatory networks and the yeast cell cycle. Curr Opin Cell Biol 2002; 14: 676-683.
10. Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO. Genomic expression programs in the response of yeast cells to environmental changes. Mol Biol Cell 2000; 11: 4241-57.
11. Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS. Global analysis of protein expression in yeast. Nature 2003; 425: 737-741.
12. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG. Life with 6000 genes. Science 1996; 274: 563-567.
13. Hedges SB, Blair JE, Venturi ML, Shoe JL. A molecular timescale of eukaryote evolution and the rise of complex multicellular life. BMC Evol Biol; 2004, 4:2
14. Herrero E, de la Torre MA, Valentín E. Comparative genomics of yeast species: new insights into their biology. Intl Microbiol 2003; 6: 183-190.
15. Ho Y, Gruhler A, Heilbut A, Bader GD, Moore L, Adams SL, Millar A, Taylor P, Bennett K, Boutilier K, Yang L, Wolting C, Donaldson I, Schandorff S, Shewnarane J, Vo M, Taggart J, Goudreault M, Muskat B, Alfarano C, Dewar D, Lin Z, Michalickova K, Willems AR, Sassi H, Nielsen PA, Rasmussen KJ, Andersen JR, Johansen LE, Hansen LH, Jespersen H, Podtelejnikov A, Nielsen E, Crawford J, Poulsen V, Sorensen BD, Matthiesen J, Hendrickson RC, Gleeson F, Pawson T, Moran MF, Durocher D, Mann M, Hogue CW, Figeys D, Tyers M. Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. Nature 2002; 415:180-183.
16. Huh WK, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weismann JS, O'Shea EK. Global analysis of protein localization in budding yeast. Nature 2003; 425: 686-691.
17. International Human Genome Consortium. Initial sequencing and analysis of the human genome. Nature 2001; 409: 860-921.
18. Katinka MD, Duprat S, Cornillot E, Metenier G, Thomarat F, Prensier G, Barbe V, Peyretaillade E, Brottier P, Wincker P, Delbac F, El Alaoui H, Peyret P, Saurin W, Gouy M, Weissenbach J, Vivares CP. Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. Nature 2001; 414: 401-402.
19. Kellis M, Birren BW, Lander ES. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. Nature 2004; 428: 617-624.
20. Kellis M, Patterson N, Endrizzi M, Lander ES. Sequencing and comparison of yeast species to identify genes and regulatory elements. Nature 2003; 423: 241-254.
21. Kurtzman CP, Blanz PA. Ribisomal RNA/DNA sequence comparisons for assessing phylogenetic relationships. In: Kutzman CP, Fell JW (Eds) The Yeasts: a Taxonomic Study, 4th edition. Amsterdam, Elsevier, 2000: 69-74.
22. Loftus BJ, Fung E, Roncaglia P, Rowley D, Amedeo P, Bruno D, Vamathevan J, Miranda M, Anderson IJ, Fraser JA, Allen JE, Bosdet IE, Brent MR, Chiu R, Doering TL, Donlin MJ, D'Souza CA, Fox DS, Grinberg V, Fu J, Fukushima M, Haas BJ, Huang JC, Janbon G, Jones SJ, Koo HL, Krzywinski MI, Kwon-Chung JK, Lengeler KB, Maiti R, Marra MA, Marra RE, Mathewson CA, Mitchell TG, Pertea M, Riggs FR, Salzberg SL, Schein JE, Shvartsbeyn A, Shin H, Shumway M, Specht CA, Suh BB, Tenney A, Utterback TR, Wickes BL, Wortman JR, Wye NH, Kronstad JW, Lodge JK, Heitman J, Davis RW, Fraser CM, Hyman RW. The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. Science 2005; 307: 1321-1324.
23. Lopreiato R, Facchin S, Sartori G, Arrigoni G, Casonato S, Ruzzene M, Pinna LA, Carignani G. Analysis of the interaction between piD261/Bud32, an evolutionarily conserved protein kinase of *Saccharomyces cerevisiae*, and the Grx4 glutaredoxin. Biochem J 2004; 377: 395-405.
24. Molina MM, Bellí G, de la Torre MA, Rodríguez-Manzaneque MT, Herrero E. Nuclear monothiol glutaredoxins of *Saccharomyces cerevisiae* can function as mitochondrial glutaredoxins. J Biol Chem 2004; 279: 51923-51930.
25. Souciet J, Aigle M, Artiguenave F, Blandin G, Bolotin-Fukuhara M, Bon E, Brottier P, Casaregola S, de Montigny J, Dujon B, Durrens P, Gaillardin C, Lepingle A, Llorente B, Malpertuy A, Neuveglise C, Ozier-Kalogeropoulos O, Potier S, Saurin W, Tekaia F, Toffano-Nioche C, Wesolowski-Louvel M, Wincker P, Weissenbach J. Genomic exploration of the hemiascomycetous yeasts. FEBS Lett 2000; 487: 3-147.
26. Thompson JD, Higgins DG, Gibson TJ. ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res 1994; 22: 4673-4680.
27. Toledano M, Delaunay A, Biteau B, Spector D, Azevedo D. Oxidative stress responses in yeast. In: Hohmann S, Mager PWH (Eds) Topics in Current Genetics, vol. 1, Yeast Stress Responses. Berlin, Springer, 2003: 241-303.
28. Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P, Qureshi-Emili A, Li Y, Godwin B, Conover D, Kalbfleisch T, Vijayadamodar G, Yang M, Johnston M, Fields S, Rothberg JM. A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. Nature 2000; 403: 623-627.
29. Vilella F, Alves R, Rodríguez-Manzaneque MT, Bellí G, Swaminathan S, Sunnerhagen P, Herrero E. Evolution and cellular function of monothiol glutaredoxins: involvement in iron-sulphur cluster assembly. Comp Funct Genom 2004; 5: 328-341.
30. Wingert RA, Galloway JL, Barut B, Foott H, Fraenkel P, Axe JL, Weber GJ, Dooley K, Davidson AJ, Schmid B, Paw BH, Shaw GC, Kingsley P, Palis J, Schubert H, Chen O, Kaplan J, Zon LI Tubingen 2000 Screen Consortium. Deficiency of glutaredoxin 5 reveals Fe-S clusters are required for vertebrate haem synthesis. Nature 2005; 436: 1035-1039.
31. Witte S, Villalba M, Bi K, Liu Y, Isakov N Altman A. Inhibition of the c-Jun N-terminal kinase/AP-1 and NF-*k*B pathways by PICOT, a novel protein kinase C-interacting protein with a thioredoxin homology domain. J Biol Chem 2000; 275: 1902-1909.
32. Wolfe KH, Shields DC. Molecular evidence for an ancient duplication of the entire yeast genome. Nature 1997; 387: 708-713.
33. Wong S, Butler G, Wolfe KH. Gene order evolution and paleopolyploidy in hemiascomycete yeasts. Proc Natl Acad Sci USA 2002; 99: 9272-9277.