FEBS Letters 487 (2000) 56–60 FEBS 24373

Genomic Exploration of the Hemiascomycetous Yeasts: 9. *Saccharomyces kluyveri*

Cécile Neuvéglise[a,*], Elisabeth Bon[a], Andrée Lépingle[a], Patrick Wincker[b], François Artiguenave[b], C. Gaillardin[a], Serge Casarégola[a]

[a] *Collection de Levures d'Intérêt Biotechnologique, Laboratoire de Génétique Moléculaire et Cellulaire, INRA UMR216, CNRS URA1925, INA-PG, P.O. Box 01, F-78850 Thiverval-Grignon, France*

[b] *Génoscope, Centre National de Séquençage, 2 rue Gaston Crémieux, P.O. Box 191, F-91057 Evry Cedex, France*

Received 3 November 2000; accepted 9 November 2000

First published online 27 November 2000

Edited by Horst Feldmann

Abstract The genome of *Saccharomyces kluyveri* was explored through 2528 random sequence tags with an average length of 981 bp. The complete nuclear ribosomal DNA unit was found to be 8656 bp in length. Sequences homologous to retroelements of the *gypsy* and *copia* types were identified as well as numerous solo long terminal repeats. We identified at least 1406 genes homologous to *Saccharomyces cerevisiae* open reading frames, with on average 58.1% and 72.4% amino acid identity and similarity, respectively. In addition, by comparison with completely sequenced genomes and the SwissProt database, we found 27 novel *S. kluyveri* genes. Most of these genes belong to pathways or have functions absent from *S. cerevisiae*, such as the catabolic pathway of purines or pyrimidines, melibiose fermentation, sorbitol utilization, or degradation of pollutants. The sequences are deposited in EMBL under the accession numbers AL404849–AL407376.

Key words: Yeast; Retrotransposon; Karyotype; tRNA gene

1. Introduction

Saccharomyces kluyveri forms a distinct subline from the highly heterogeneous *Saccharomyces* sensu lato group. It shows a close phylogenetic relationship with a variety of species of other genera including *Torulaspora, Kluyveromyces*, and *Zygosaccharomyces*. The few isolates of *S. kluyveri* (about 30 different strains are preserved in yeasts collections) come from Europe or the USA. They were mainly isolated from soil, tree exudate or *Drosophila*.

The genome of *S. kluyveri* is poorly studied. The strain used in this work (type strain CBS3082) is diploid. Petersen et al. [1] estimated that *S. kluyveri* contains seven chromosomes with a 10 Mb genome size, in agreement with Vaughan-Martini et al. [2]. Weinstock and Strathern [3], however, observed eight separate chromosomal bands by pulsed field gel electrophoresis of strain CBS4568, all larger than 0.9 Mb.

A 14.2 kb linear DNA plasmid, pSKL, was isolated from some strains of *S. kluyveri* [4]. Interestingly, it shows structural characteristics similar to those of adenoviruses or bacteriophage ϕ29 of *Bacillus subtilis*. The nucleotide sequence of pSKL is highly similar to pGKL2, the killer plasmid of *Kluyveromyces lactis*, though no associated killer function has been described in *S. kluyveri*. The mitochondrial genome of *S. kluyveri*, a petite-negative yeast, is 49 kb in length with an average G+C content of 22% [5].

In general, little is known about the genetic content of *Saccharomyces* species that do not belong to the sensu stricto group. *S. kluyveri* is among the most extensively studied species of this group. The pheromone system of *S. kluyveri* has been characterized [6], the α-factor receptor has been cloned and sequenced [7]. Physiological cross-reactions of pheromones between *S. kluyveri* and *Saccharomyces cerevisiae* have been observed, although the two species do not mate with each other [8]. Another interesting aspect of *S. kluyveri* is that, in contrast to the other species of the *Saccharomyces* genus, it can use pyrimidines and their degradation products as the only sources of nitrogen [9]. Recently, Gojkovic et al. [10] reported on the specific features of *PYD2* and its DPHase activity as an important regulatory checkpoint of the pyrimidine catabolic pathway in *S. kluyveri*.

When we started this work, only 46 *S. kluyveri* DNA sequences were available in GenBank, most of them corresponding to ribosomal RNA sequences (16) and a variety of nuclear genes (20). Only four of these sequences correspond to mitochondrial genes (*COX3*, *COX2* and *ATP8*) and three to the pSKL plasmid. We now report a global analysis of the *S. kluyveri* genome from 2528 random sequence tags (RSTs) totalling 2.5 Mb and revealing at least 1406 novel genes.

2. Materials and methods

2.1. DNA library preparation

S. kluyveri strain CBS3082 was grown overnight at 28°C in 10 ml YPD medium (1% yeast extract, 1% peptone, 1% glucose). Total DNA was prepared according to Querol et al. [11]. Partial digestion of *S. kluyveri* DNA was performed with *Cvi*JI (CHIMERx, Madison, WI, USA). DNA fragments of 3.5–5 kb were purified using GeneClean II kit (Bio 101, USA). The pBAM3 vector derivative from pBluescript was digested with *Sma*I and dephosphorylated using alkaline phosphatase (Boehringer Mannheim, Germany). Ligations were performed with 8 ng *Sma*I vector and 15–30 ng purified inserts overnight at 5°C using T4 DNA ligase (Gibco BRL). *Escherichia coli* strain DH10B was transformed using the ligation mixture by electroporation (100 Ω, 25 μF, 1700 V/cm). Bacteria were plated on LB medium with ampicillin (100 μg/ml), IPTG (25 μg/ml) and X-gal (40 μg/ml). Plasmids from 24 white colonies were purified and cut with either *Mlu*I or *Eco*RI in order to check the presence and the size of plasmid inserts. After quality control, white colonies were

*Corresponding author.
E-mail: ncecile@grignon.inra.fr

PII: S0014-5793(00)02280-8

randomly selected, grown in triplicate in LB with ampicillin and glycerol (15%) and stored at −80°C in 96 well plates. The *S. kluyveri* library is made of 18 plates, i.e. 1728 different clones individually stored.

2.2. Sequencing strategy, RST analysis and annotation

Plasmid purification, sequencing reactions and sequence data processing were performed as described [12]. *S. kluyveri* RSTs were compared with *S. cerevisiae* predicted open reading frame (ORF) products, and with all available predicted proteomes as well as with a filtered SwissProt database as described [13]. Blast version 2.0.10 with Blosum62 substitution matrix was used for comparison with all databases. Contigs were constructed using the Phred/Phrap programs [14,15] and visualized using the *consed* interface [16]. Scripts and criteria used for homology validation are described in [13].

2.3. Properties of the DNA library

The average length of the 2528 RSTs analyzed is 981 ± 115 bp which represents a total of 2482 kb sequenced. Most inserts (1213) were sequenced for both ends [12]. The average size of the inserts is 4.21 ± 1.29 kb. In 25 cases, the insert sizes were short enough (1121–1938 bp) that the two RSTs partially overlapped. The G+C content of the *S. kluyveri* nuclear genome was estimated taking into account all RSTs except those corresponding to the mitochondrial RSTs. We found 41.3% G+C, which is in agreement with the estimation of 40% made by Meyer and Phaff [17].

3. Results and discussion

3.1. Repeated sequences and extrachromosomal sequences

The Phred/Phrap software was used to build contigs between the traces of the RSTs resulting in 451 contigs comprising two or more RSTs (1157 RSTs included in total) which were subsequently screened for extrachromosomal sequences (mitochondrial DNA, plasmids) and repeated sequences.

We identified a single contig made of 58 RSTs corresponding to the ribosomal DNA (contig 517). The complete unit including 18S, ITS1, 5.8S, ITS2, 26S, and 5S, and the complete intergenic region is 8656 bp in length. The 18S rDNA shows 99% identity and the complete 26S rDNA 96.2% identity with *S. cerevisiae* rDNA. Contig 517 exhibits only one nucleotide difference with the 1764 bp of the 18S rDNA already published [18] and none with the D1/D2 domain [19]. One of the two regions flanking the ribosomal cluster was identified in clone AU0AA010G06 carrying NTS2 and the 5′ end of 18S rDNA at one extremity of the insert and a *YDR*174w (*HMO1*) homologue at the other extremity.

Whereas the mitochondrial DNA of *S. kluyveri* type strain is 49 kb [5], only four contigs totalling 13 RSTs were identified as part of the mitochondrial DNA, as well as three non-contigated RSTs. We found four RSTs corresponding to the 15S rRNA with a tRNA-Ser gene upstream of the 15S rRNA gene and eight RSTs corresponding to the 5′ part of the 21S rRNA with a tRNA-Trp gene upstream of it. In two clones we found the 21S rRNA at one extremity of the insert and the 15S rRNA at the other extremity. This shows that the 15S rRNA and 21S rRNA genes map close to each other (about 2 kb) on the mitochondrial map of *S. kluyveri*, in contrast to *S. cerevisiae*, in which at least 30.5 kb separate 15S from 21S rDNA [20]. The *COX1* gene was identified in three RSTs. The *S. kluyveri* putative *COX1* gene lacks several of the introns found in *S. cerevisiae*, such as *COX1*-ai1, *COX1*-ai2, *COX1*-ai5α, *COX1*-ai5β and *COX1*-ai5γ. We detected only one intron homologous to *COX1*-ai3 (I-SceIII). The structure of the *S. kluyveri* COX1 gene is close to that in *S. exiguus COX1* gene [21]. None of the three mitochondrial genes already known in *S. kluyveri* (*COX3*, *COX2*, *ATP8*) was found in our RST set.

Retrotransposons in *S. kluyveri* were searched for by homology to the five different families of retroelements (Ty) of

	All RST	RST with Primer D	RST with Primer T
N_s	2528	1233	1219
N_r	58	29	29
N_m	16	11	5
N_y	7	4	3
P	25	0	0
C	418	130	127
N_c	1050	264	271
N	2422	1189	1182
L (bp)	981.91	981.91	981.91
I	1790	1055	1038
G	8.814 Mbp	8.751 Mbp	8.750 Mbp

N_s: number of RSTs sequenced,
N_r, N_m, N_y: number of RSTs corresponding to rDNA, mitochondrial DNA, transposable elements,
P: number of inserts with overlapping RSTs
C: number of contigs
N_c: number of RSTs involved in contigs
$N = N_s - (N_r + N_m + N_y + P)$
I: number of "islands" $I = N - N_c + C$
G: estimation of genome size without repeated elements
In our study, $\sigma = 0.985$

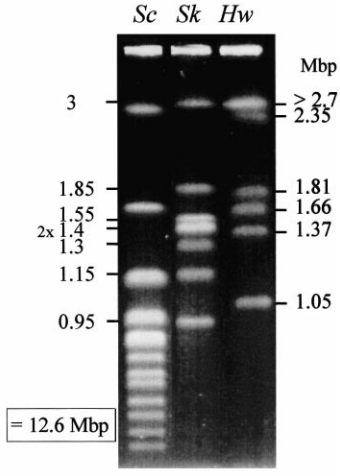

Fig. 1. Estimation of *S. kluyveri* genome size. The mathematical estimation using RST and resulting contigs is based on the Lander and Waterman [22] method. Pulsed field gel electrophoresis was carried out on a CHEF-DRII apparatus (Bio-Rad) at 13°C in a 1% agarose gel under the following conditions: 20 h, 480 s pulse time at 100 V; 20 h, 300 s pulse time at 130 V; and 12 h, 90 s pulse time at 200 V. *Sc*: *S. cerevisiae*; *Sk*: *S. kluyveri*; *Hw*: *Hansenula wingei.*

S. cerevisiae. We identified two contigs as part of retroelements: contig 513 (seven RSTs) which shows a high similarity to integrase and protease of TY1B (75% amino acid identity) and contig 395 (three RSTs) which matches the 3′ end of TY2B (80% identity over 225 amino acids). In addition, one non-contigated RST matches TY1A over 48 amino acids at its 5′ extremity with 89% identity. Surprisingly, these data reveal that the *S. kluyveri* retroelements are very closely related to Ty1 and Ty2. This was unexpected since *S. kluyveri* is phylogenetically distant from *S. cerevisiae*. In contrast, we identified only one non-contigated RST showing only 37% identity over 102 amino acids with Ty3B, suggesting that Ty3-like elements are poorly conserved and represented in *S. kluyveri*. Numerous solo long terminal repeats (LTRs) were identified showing 91% identity over 126 bp with *S. cerevisiae* Ty1 δ LTRs. Altogether we found 57 RSTs, some of them possessing two or more contiguous LTRs, and some having 'overlapping' LTRs as in *S. cerevisiae*. The accumulation of LTRs and overlapping LTRs as retroelement remnants suggests that hotspots of transposition exist in *S. kluyveri,* as in *S. cerevisiae*. Among the 57 RSTs containing LTRs, 20 RSTs contained LTRs associated with tRNA genes, which suggests that the specificity of integration of the *S. kluyveri* retroelements is comparable to that of *S. cerevisiae*.

3.2. Estimation of genome size

Once repeated sequences and extrachromosomal sequences were identified, the genome size was estimated according to Lander and Waterman [22]. We made three calculations to estimate the *S. kluyveri* genome size. First, all RSTs were considered without taking into account the fact that the RSTs from the same insert are physically linked. Then, to ensure that no bias was introduced by the physical linkage of RSTs from the same insert, we estimated the genome size by considering on the one hand the RSTs sequenced with the forward primer [12] and on the other hand the RSTs sequenced with the reverse primer. The three different estimations gave a genome size of about 8.8 Mb without repeated elements (Fig. 1). This showed that there was no bias for genome size estimation due to the physical linkage of the RSTs. This is probably due to the insert size (4.21 ± 1.29 kb). Vaughan-Martini et al. [2] estimated that *S. castellii* had the smallest *Saccharomyces* genome, whereas more recently, Petersen et al. [1] suggested that *S. kluyveri* together with *S. castellii* genomes were the smallest ones. Petersen et al. [1] reported that *S. kluyveri* contained seven chromosomes with a genome size of 10 Mb. However, we recently determined precisely the number and size of *Saccharomyces* sensu lato strains by pulsed field gel electrophoresis and found eight chromosomes for *S. kluyveri*, with a genome size of 12.6 Mb (Fig. 1), which is quite different from the calculated size.

3.3. tRNAs genes

We found tRNAs genes in 53 RSTs, four of which corresponded to mitochondrial genes. We noted 44 different nuclear tRNAs genes. Considering that we have sequenced about 20% of *S. kluyveri* genome (2.482 Mb), this strain may contain a total of 220 tRNA genes which is comparable to that of *S. cerevisiae*. Two RSTs contained two consecutive tRNA genes homologous to *S. cerevisiae* tV_GTT_ER1_ tRNA and tH_CAC_ER2_tRNA genes which are also contiguous on the *S. cerevisiae* map. The sequences of *S. cerevisiae* and *S. kluyveri* tRNAs are highly conserved (91–100% nucleotide identity) and the anticodons are always conserved. The presence/absence of introns in tRNA genes is also conserved except in two cases. Both correspond to the presence of an additional intron in the *S. kluyveri* tS_TCT_AR1_tRNA homologue.

3.4. Annotation of S. kluyveri ORFs

RSTs corresponding to rDNA, Ty elements and mtDNA were removed from the list. The remaining 2452 RSTs were systematically compared to the *S. cerevisiae* sequences database [13]. We obtained 1937 positive matches, 1716 of which were significant ('o') and 221 significant but ambiguous ('oo') because they correspond to gene families. The mean percentage of amino acid identity and similarity was computed for the matches 'o' and 'oo' taking into account the length of the

Table 1
S. kluyveri duplicated genes

S. cerevisiae homologue	*S. cerevisiae* gene family	Annotation[a]	Function
YLR130c	P2.263.f2.1	o	low affinity zinc transporter
YGL038c	P2.321.f2.1	o	α-1,6-mannosyltransferase
YBR015c	P2.329.f2.1	o	type II membrane protein
YGR231c	P2.384.f2.1	o	prohibitin
YHR179w	P2.54.f2.1	oo	NADPH dehydrogenase (old yellow enzyme), isoform 1
YGR192c	P3.99.f3.1	oo	glyceraldehyde-3-phosphate dehydrogenase 3
YNL125c	P4.39.f4.1	o	similarity to human X-linked PEST containing transporter
YOL119c	P4.39.f4.1	o/oo	similarity to monocarboxylate transporter proteins
YMR307w	P5.13.f5.1	o	glycophospholipid anchored surface glycoprotein
YDL137w[b]	P7.3.f6.1	oo	GTP binding protein of the ARF family
YBL017c	P8.2.f6.1	o/oo	vacuolar protein sorting/targeting protein
YDR483w	P9.2.f9.1	oo	α-1,2-mannosyltransferase
YDR011w	P10.1.f10.1	o/oo	multidrug resistance protein
YJR152w	P10.2.f3.1	o	allantoate permease
YDL210wi	P23.1.f3.1	o	GABA specific high affinity permease
YDL144c	singleton	o	hypothetical protein
YIL162w	singleton	o	invertase (sucrose hydrolyzing enzyme)
YIR002c	singleton	o	weak similarity to ATP dependent RNA helicases
YOL154w	singleton	o	similarity to *S. fumigata* Asp FII

[a]'o' indicates that both duplicated genes show a non-ambiguous match, 'oo' an ambiguous match and 'o/oo' indicates that the match was non-ambiguous for one copy and ambiguous for the other one.
[b]The two copies of this gene are tandemly duplicated in the same orientation.

match. We found 58.1% and 72.4% amino acid identity and similarity, respectively.

The 1937 significant annotations ('o'+'oo') match 1387 different *S. cerevisiae* ORFs. Indeed, 391 homologues of *S. cerevisiae* ORFs were present in more than one RST. We tried to estimate the number of genes duplicated but in many cases, it was impossible to come to a conclusion about the presence of one or more genes because the RSTs matched different parts of the same *S. cerevisiae* gene. Thus, we concluded that the minimal number of *S. kluyveri* genes homologous to *S. cerevisiae* genes was 1406 and the maximal number 1593. We found 19 duplicated genes (Table 1), one of which (YDL137w or ARF2, a GTP binding protein of the ARF family) is tandemly duplicated in the same direction. Only four out of the 19 duplicated genes are singletons in *S. cerevisiae*, the others belong to gene families. We enquired whether gene redundancy in *S. kluyveri* is comparable to that of *S. cerevisiae*. In *S. kluyveri*, the percentage of ORFs matching singletons is about 55% which is comparable to *S. cerevisiae* (60%). The mean number of ORFs per family identified in *S. kluyveri* is proportional to that of *S. cerevisiae* except for some families which have no homologues in *S. kluyveri* [23]. This includes only families of genes located in subtelomeric regions in *S. cerevisiae*, such as the *COS* gene family, the *PAU* gene family or the *FLO* gene family. The subtelomeric regions are known to exhibit significant gene redundancy and to be dynamic regions where recombination processes generate amplifications, deletions and translocations. In addition, most of these genes are thought to be duplicated in *S. cerevisiae* because of their role in fermentation processes [24].

3.5. Comparison with other genomes

The 2452 RSTs were then compared to the Gproteome database [13]. We identified matches to 27 additional genes of *S. kluyveri* without homologues in the *S. cerevisiae* sequenced strain (Table 2). Thus, including these 27 genes, the minimum and the maximum numbers of genes we identified in *S. kluyveri* are 1433 and 1620, respectively. We found two RSTs matching the *MEL* genes (α-galactosidase precursor) of *S. cerevisiae* (absent from the sequenced strain) as expected since the type strain of *S. kluyveri* ferments melibiose [25]. We also identified the homologue of the *SOU1* gene from *Candida albicans*, which enables sorbitol utilization. *S. kluyveri* is the only *Saccharomyces* species which is D-glucitol positive [26]. Homologues of the *SOU1* gene were also identified in *K. marxianus* [27] and *K. thermotolerans* [28], the closest species to *S. kluyveri*, which are also D-glucitol positive [29]. Homologues of three genes encoding enzymes for the conversion of 5-substituted hydantoins to corresponding L-amino acids were identified (*AMAB, ylbB*, and *HYUc*). These genes are involved in pyrimidine degradation, a pathway well documented in bacteria and eukaryotic organisms but not yet described in yeasts, probably because this pathway is absent from *S. cerevisiae* and *Schizosaccharomyces pombe*. Recently, Gojkovic et al. [9,10] reported on the presence of pyrimidine and purine

Table 2
Potential functions encoded by *S. kluyveri* ORF products having no validated homologue in the genome of *S. cerevisiae*

Organism	ORF name	Accession number in SwissProt	Putative function
Bacteria			
Bacillus stearothermophilus	*AMAB*	Q53389	*N*-carbamyl-L-amino amidohydrolase
Escherichia coli	ECb0516 (ylb B)	P77425	putative hydantoin utilization protein
Escherichia coli	yeiN	P33025	hypothetical protein
Haemophilus influenzae	HI0588	Q57051	hypothetical protein
Helicobacter pylori	HP1429		polysialic acid capsule expression protein
Mycobacterium tuberculosis	MTRv0154c		similar to the C-terminal region acyl-CoA dehydrogenase
Pseudomonas sp.	*HYUC*	Q01264	hydantoin utilization protein C
Rhodococcus sp.	*SOXA; DSZA*	P54995	dibenzothiophene desulfurization enzyme A
Ascomycetes			
Candida albicans	*SOU1*	P87219	sorbitol utilization protein
Emericella nidulans	*UAPA*	Q07307	uric acid xanthine permease
Pichia jadinii		P78609	uricase (urate oxidase)
S. cerevisiae	*MEL1, MEL2, MEL5, MEL6*[a]	P04824, P41945, P41946, P41947	α-galactosidase precursor
Schizosaccharomyces pombe	*MLO2*	Q09329	MLO2 protein
Schizosaccharomyces pombe	P78771		unknown
Schizosaccharomyces pombe	SPAC12B10.16C	Q10449	hypothetical protein
Schizosaccharomyces pombe	SPAC1D4.09C	Q10154	hypothetical protein
Schizosaccharomyces pombe	SPAC22H10.08	Q10301	hypothetical protein
Schizosaccharomyces pombe	SPAC2F3.16		hypothetical zinc finger protein
Schizosaccharomyces pombe	SPBC354.15		putative fructosyl amino acid oxidase
Schizosaccharomyces pombe	SPCC285.05		hypothetical protein
Schizosaccharomyces pombe	SPCC622.19		hypothetical protein
Schizosaccharomyces pombe	*SRP1*	Q10193	putative splicing protein, RNA binding
Other eukaryotes			
Arabidopsis thaliana	*PYRD*	P32746	dihydroorotate dehydrogenase precursor
Caenorhabditis elegans	K09H11.1		similar to acyl-CoA dehydrogenases and epoxide hydrolases
Caenorhabditis elegans	B0252.2		similar to sphingomyelin phosphodiesterase
Caenorhabditis elegans	ZK455.4		similar to sphingomyelin phosphodiesterase
Homo sapiens	*DIA4; NMOR1; NQO1*	P15559	NAD(P)H dehydrogenase [quinone]

[a]The *MEL* genes are not present in the *S. cerevisiae* strain sequenced (S288C). As the members of the *MEL* gene family are highly similar, the match was considered ambiguous ('oo').

catabolic pathways in *S. kluyveri*. We also found the homologue of dihydroorotate dehydrogenase (PYRD) from *Arabidopsis thaliana*, another gene involved in pyrimidine catabolism, probably in the first step of pyrimidine degradation (Piskur, personal communication). Homologues of these genes were also identified in *K. thermotolerans* [28], *Y. lipolytica* [30], *K. lactis* [31] and *P. angusta* [32], but not in the other *Saccharomyces* species. The presence/absence of this pathway is consistent with the phylogenetic separation of the species. The *Saccharomyces* sensu lato and *Saccharomyces* sensu stricto phyla could have lost this pathway during evolution. Similarly, homologues of two genes involved in the purine pathway were identified: urate oxidase from *Pichia jadinii* and a purine permease from *Emericella nidulans*. Another interesting gene found in *S. kluyveri* and absent from *S. cerevisiae* is the homologue of *SOXA* from *Rhodococcus* sp. which is involved in dibenzothiophene desulfurization [33]. It would be interesting to know if *S. kluyveri* possesses a similar pathway, which has important applications in fossil fuel desulfurization and environmental depollution. Other genes currently present in yeasts other than *S. cerevisiae* were also identified: NAD(P)H dehydrogenase, sphingomyelin phosphodiesterase, fructosyl amino acid oxidase, Mlo2 protein (Table 2). If we consider that we identified about 24% of *S. kluyveri* genes, the complete *S. kluyveri* genome possesses about 112 genes absent from *S. cerevisiae*, most of them corresponding to pathways lost by *S. cerevisiae* or acquired by *S. kluyveri* during evolution.

3.6. Functional classification of S. kluyveri genes

Assuming that *S. kluyveri* and *S. cerevisiae* have approximately the same number of genes, the distribution of their functional categories [34] is very similar. Only few functional categories are overrepresented, clearly indicating the physiological adaptation of the species to its environment. In *S. kluyveri*, metabolism of phosphate (regulation of phosphate utilization and phosphate transport) and nitrogen and sulfur are overrepresented. We also noted an increase of amino acid metabolism and amino acid transporters (171% and 188%, respectively). Overall, the number of genes involved in transport is higher than expected, in particular drug transporters: we found at least 16 homologues of the 35 *S. cerevisiae* genes. Overrepresentation also concerned the category of cell organization, with a clear increase of genes involved in biogenesis and organization of intracellular transport vesicles (+196% and +168%).

Acknowledgements: We would like to thank Fredj Tekaia for his computing assistance throughout this project. Thanks are also expressed to Jure Piskur for helpful discussions about pyrimidine catabolism. This work was supported in part by BRG Grant (ressources génétiques des microorganismes No. 11-0926-99). E.B. was supported by EEC Scientific Research Grant QLRI-1999-01333.

References

[1] Petersen, R.F., Nilsson-Tillgren, T. and Piskur, J. (1999) Int. J. Syst. Bacteriol. 49, 1925–1931.
[2] Vaughan-Martini, A., Martini, A. and Cardinali, G. (1993) Antonie van Leeuwenhoek 63, 145–156.
[3] Weinstock, K.G. and Strathern, J.N. (1993) Yeast 9, 351–361.
[4] Hishinuma, F. and Hirai, K. (1991) Mol. Gen. Genet. 226, 97–106.
[5] Piskur, J., Smole, S., Groth, C., Petersen, R.F. and Pedersen, M.B. (1998) Int. J. Syst. Bacteriol. 48, 1015–1024.
[6] Sen, M., Shah, A. and Marsh, L. (1997) Curr. Genet. 31, 235–240.
[7] Marsh, L. and Herskowitz, I. (1988) Proc. Natl. Acad. Sci. USA 85, 3855–3859.
[8] McCullough, J. and Herskowitz, I. (1979) J. Bacteriol. 138, 146–154.
[9] Gojkovic, Z., Paracchini, S. and Piskur, J. (1998) Adv. Exp. Med. Biol. 431, 475–479.
[10] Gojkovic, Z., Jahnke, K., Schnackerz, K.D. and Piskur, J. (2000) J. Mol. Biol. 295, 1073–1087.
[11] Querol, A., Barrio, E. and Ramon, D. (1992) Syst. Appl. Microbiol. 15, 439–446.
[12] Artiguenave, F., Wincker, P., Brottier, P., Duprat, S., Jovelin, F. et al. (2000) FEBS Lett. 487, 13–16 (this issue).
[13] Tekaia, F., Blandin, G., Malpertuy, A., Llorente, B., Durrens, P. et al. (2000) FEBS Lett. 487, 17–30 (this issue).
[14] Ewing, B., Hillier, L., Wendl, M.C. and Green, P. (1998) Genome Res. 8, 175–185.
[15] Ewing, B. and Green, P. (1998) Genome Res. 8, 186–194.
[16] Gordon, D., Abajian, C. and Green, P. (1998) Genome Res. 8, 195–202.
[17] Meyer, S.A. and Phaff, H.J. (1969) J. Bacteriol. 97, 52–56.
[18] James, S.A., Cai, J., Roberts, I.N. and Collins, M.D. (1997) Int. J. Syst. Bacteriol. 47, 453–460.
[19] Kurtzman, C.P. and Robnett, C.J. (1998) Antonie van Leeuwenhoek 73, 331–371.
[20] Foury, F., Roganti, T., Lecrenier, N. and Purnelle, B. (1998) FEBS Lett. 440, 325–331.
[21] Bon, E., Neuvéglise, C., Lépingle, A., Wincker, P., Artiguenave, F. et al. (2000) FEBS Lett. 487, 42–46 (this issue).
[22] Lander, E.S. and Waterman, M.S. (1988) Genomics 2, 231–239.
[23] Llorente, B., Durrens, P., Malpertuy, A., Aigle, M., Artiguenave, F. et al. (2000) FEBS Lett. 487, 71–75 (this issue).
[24] Naumov, G.I., Naumova, E.S., Turakainen, H. and Korhola, M. (1996) Genet. Res. 67, 101–108.
[25] Barnett, J.A., Payne, R.W. and Yarrow, D. (1990) Yeasts: characteristics and identification. Cambridge University Press, Cambridge.
[26] Vaughan-Martini, A. and Martini, A. (1998) in: The Yeasts, A Taxonomic Study (Kurtzman, C.P. and Fell, J.W., Eds.), pp. 358–371 Elsevier Science, Amsterdam.
[27] Llorente, B., Malpertuy, A., Blandin, G., Wincker, P., Artiguenave, F., Wincker, P. and Dujon, B. (2000) FEBS Lett. 487, 101–112.
[28] Malpertuy, A., Llorente, A., Blandin, G., Artiguenave, F., Wincker, P. and Dujon, B. (2000) FEBS Lett.
[29] Lachance, M.A. (1998) in: The Yeasts, A Taxonomic Study (Kurtzman, C.P. and Fell, J.W., Eds.), pp. 227–247 Elsevier Science, Amsterdam.
[30] Casaregola, S., Neuvéglise, C., Lépingle, A., Bon, E., Feynerol, C. et al. (2000) FEBS Lett. 487, 95–100 (this issue).
[31] Bolotin-Fukuhara, M., Toffano-Nioche, C., Artiguenave, F., Duchateau-Nguyen, G., Lemaire, M. et al. (2000) FEBS Lett.
[32] Blandin, G., Llorente, B., Malpertuy, A., Wincker, P., Artiguenave, F. and Dujon B. et al. (2000) FEBS Lett. 487, 76–81 (this issue).
[33] Piddington, C.S., Kovacevich, B.R. and Rambosek, J. (1995) Appl. Environ. Microbiol. 61, 468–475.
[34] Gaillardin, C., Duchateau-Nguyen, G., Tekaia, F., Llorente, B., Casaregola, S. et al. (2000) FEBS Lett. 487, 134–149 (this issue).