

2. **Ma, H., S. Kunes, P.J. Schatz, and D. Botstein.** 1987. Plasmid construction by homologous recombination in yeast. *Gene* 58:201-216.
3. **Miletti, K.E. and M.J. Leibowitz.** 2000. Construction of large plasmid lacking linearizing single restriction sites by simultaneous plasmid shuffling in yeast. *Yeast* 16:1527-1534.
4. **Raymond, C.K., T.A. Powder, and S.L. Sexson.** 1999. General method for plasmid construction using homologous recombination. *BioTechniques* 26:134-141.
5. **Tschumper, G. and J. Carbon.** 1980. Sequence of a yeast DNA fragment containing a chromosomal replicator and the TRP1 gene. *Gene* 10:157-166.
6. **Bautista, D.S., M. Hitt, J. McGrory, and F.L. Graham.** 1991. Isolation and characterization of insertion mutants in E1A of adenovirus type 5. *Virology* 182:578-596.
7. **Boeke, J.D., L. LaCroute, and G.R. Fink.** 1984. A positive selection for mutants lacking orotidine-5-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Mol. Gen. Genet.* 197:345-346.
8. **Becker, D.M. and L. Guarente.** 1991. High efficiency transformation of yeast by electroporation. *Methods Enzymol.* 194:182-187.
9. **Hoffman, C.S. and F. Winston.** 1987. A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. *Gene* 57:267-272.
10. **Graham, F.L., J. Smiley, W.C. Russell, and R. Narin.** 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.* 36:59-72.
11. **Parker, R.C., R.M. Watson, and J. Vinograd.** 1977. Mapping closed circular DNAs by cleavage with restriction endonucleases and calibration by agarose gel electrophoresis. *Proc. Natl. Acad. Sci. USA* 74:851-855.

Received 16 December 2002; accepted 8 July 2003.

Address correspondence to Peter Robert Anderson, University of Guelph, Department of Molecular Biology and Genetics, Guelph, ON, N1G 2W1, Canada. e-mail: panderso@uoguelph.ca

## A/T-rich inverted DNA repeats are destabilized by chaotrope-containing buffer during purification using silica gel membrane technology

Martin Přeborovský and František Půta  
Charles University, Prague, Czech Republic

*BioTechniques* 35:698-702 (October 2003)

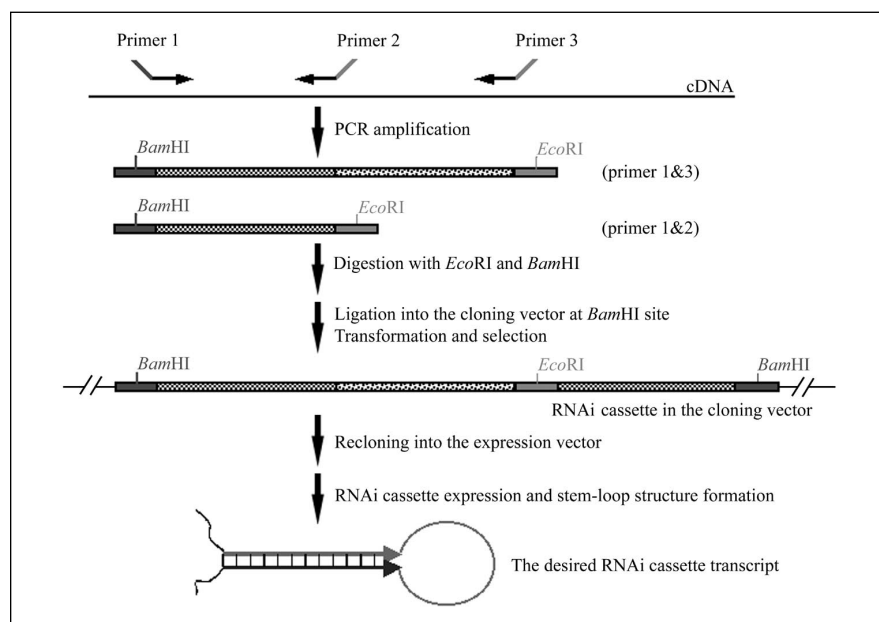
The recovery of DNA from agarose gel is a frequently used step in nearly all gene engineering procedures. Various methods have been developed for this purpose, and the use of silica gel or glass matrices seem to be the most efficient ones. Here we describe the difficulties encountered when purifying DNA fragments consisting of inverted repeat sequences.

In our study, we wanted to determine the knock-down phenotype of several novel *Dictyostelium discoideum* genes by using RNA interference (RNAi) (1), which is the method of choice when assaying gene function both in genetically well established and in genetically less well-characterized organisms (2). The RNAi method that we employed consisted of the introduction of double-stranded RNA (dsRNA) molecules (usually several hundred base pairs

long) that are processed into the effector short interfering RNAs (siRNAs; Reference 3), resulting in sequence-specific mRNA degradation (4,5).

Using the sequences of our candidate genes, we constructed an inverted repeat targeting cassette (RNAi cassette) for each gene, which should produce a hairpin RNA upon transcription. Figure 1 depicts the RNAi cassette construction. The identity of the constructed RNAi cassettes was verified by restriction analysis; Table 1 summarizes their properties.

To reclone the prepared RNAi cassettes from the cloning vector into the expression vector, we cut the RNAi cassettes using *Bam*HI, separated them from the vector DNA by electrophoresis on 1% agarose gel, and purified them by gel purification using the QIAquick™ Gel Extraction Kit (Qia-



**Figure 1.** The RNA interference (RNAi) cassette construction procedure and transcribed RNA structure.

**Table 1. Some Properties of the Analyzed RNA Interference (RNAi) Cassettes**

RNAi Cassette	Length (bp)	A/T Content (%)
1 DdHAT (SSF192 <sup>a</sup> )	1027	73.22
2 DdHDAC (FC-BE07 <sup>a</sup> )	982	65.88
3 <i>snwA</i> (U43887 <sup>b</sup> ) (6)	1412	67.35
4 DdARID (SSD413 <sup>a</sup> )	910	70.99
5 <i>cypE</i> (AF215865) (7)	746	69.98

<sup>a</sup>Tsukuba cDNA Sequencing Project (<http://www.csm.biol.tsukuba.ac.jp/cDNAproject.html>) (8).

<sup>b</sup>GenBank® accession nos. (National Center for Biotechnology Information, National Institutes of Health, Bethesda, MD, USA).

gen, Hilden, Germany), according to the manufacturer's instructions. To our surprise, the eluted DNA exerted some unexpected properties. The majority of DNA migrated as a band of approximately two-thirds of the expected size on 1% agarose gel (Figure 2A), and the ligation of the eluted DNA into the expression vector did not yield any transformants carrying the desired constructs (data not shown).

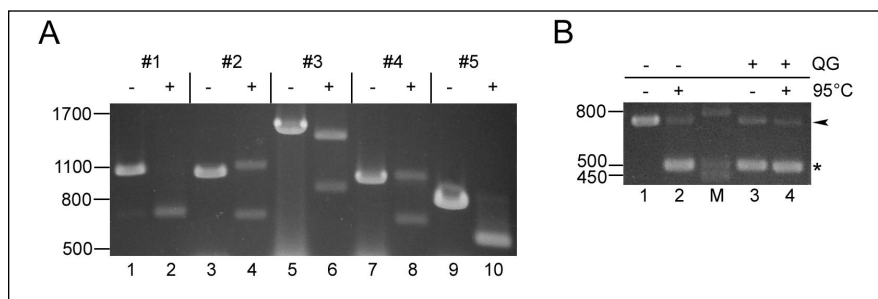
The QIAquick Gel Extraction Kit used for purification employs the chaotropic compound guanidine thiocyanate to dissolve the agarose. Because the *Dictyostelium* genome is known to be A/T rich, the base pairing in our RNAi

cassettes (>65% of A/T pairs; Table 1) could have been less stable and more susceptible to melting in the presence of chaotrope. We hypothesized that such duplex instability eventually lead to the formation of highly favored hairpin structures. The hairpin structures were expected to migrate faster than the original duplex molecules. Indeed, our observations (Figure 2A) were in agreement with this prediction.

We tested whether the chaotropic compound-containing QG buffer (part of the kit) was responsible for the DNA denaturation during the agarose-melting step. We used the RNAi cassette no. 5 for this analysis. For each treatment, a

sample containing approximately 2  $\mu$ g RNAi cassette DNA was ethanol-precipitated from the restriction reaction mixture using tRNA as a carrier and redissolved in 50  $\mu$ L EB buffer (part of the kit). The samples were left untreated as a native conformation control (sample 1), heat-denatured at 95°C for 10 min, and slowly cooled at room temperature to serve as a denaturation/renaturation control (sample 2), or subjected to the agarose-melting QG buffer treatment (sample 3). For sample 3, one-tenth of the volume of 10 $\times$  TAE (400 mM Tris, pH 8.5, 200 mM acetate, 20 mM EDTA) electrophoresis buffer and three volumes of QG buffer were added, and the mixture was incubated at 50°C for 10 min. This treatment recapitulates the procedure recommended by the manufacturer for dissolving a gel slice but does not involve contact of the DNA with the column matrix. The DNA was then ethanol-precipitated from the solution. To show that the effect of the QG buffer cannot be further potentiated by heat denaturation, sample 4 was subjected to the same treatment as in sample 3, followed by heat denaturation as in sample 2. Aliquots (approximately 100 ng of the RNAi cassette no. 5 per lane) were run on 1.5% agarose gel (Figure 2B). This analysis confirmed that the agarose-melting step of the QIAquick Gel Extraction Kit procedure (i.e., the treatment with the QG buffer) was responsible for the mobility shift of the RNAi cassette DNA. The vector shells migrated as expected, and there were no additional bands present.

We reported the problems with the purification of A/T-rich *D. discoideum* RNAi cassettes. We have shown that the agarose-melting buffer (QG buffer) caused destabilization or even denaturation of the short A/T-rich DNA fragments during gel extraction. We propose that this reflects the formation of the hairpin structures that were formed from the kit-destabilized DNA molecules after the denaturing conditions were alleviated. The hairpin structure formation would have rendered the DNA useless for subsequent cloning and would also interfere with the ligation of the few native DNA duplexes left in the solution because it would occupy the vector termini. Agarose electrophoresis followed by gel extraction is a very common method of DNA purification. We



**Figure 2. The agarose-melting QG buffer causes denaturation of short A/T-rich DNA fragments.** (A) The gel extraction procedure results in altered electrophoretic mobility of the purified DNA. Samples 1–5, containing approximately 2  $\mu$ g of five different A/T-rich *Dictyostelium* RNA interference (RNAi) cassettes (Table 1), were prepared by the digestion of the respective plasmids with *Bam*HI. One half of each sample was gel-purified using the QIAquick Gel Extraction Kit (+), and the other half was left untreated and used as a control (-). Aliquots corresponding to approximately 200 ng RNAi cassette DNA were adjusted to equal salt concentrations and run on 1% agarose gel. Molecular standards are indicated in base pairs. (B) The QG buffer treatment has the same effect on RNAi cassette DNA as heat denaturation. Four samples that contained equal amounts of the RNAi cassette no. 5 DNA were subjected to the treatments. Lane 3, DNA treated with the QG buffer co-migrated with (lane 2) the heat-denatured control. Lane 4, further heat denaturation of the QG buffer-treated sample did not cause any significant changes. Lane 1 shows an untreated native control. Molecular standards in base pairs are indicated on the left, the positions of the native and denatured RNAi cassette no. 5 DNA are indicated by the arrowhead and the asterisk, respectively. QG, QG buffer treatment; 95°C, heat denaturation. The equal ethidium bromide staining of approximate 750- and 500-bp bands in lanes 1 and 2, respectively, suggests that they have a double-stranded character. M, molecular weight marker (homemade;  $\lambda$  phage DNA digested with *Pst*I).

believe that our findings are of interest to researchers conducting RNAi experiments and/or working with *Plasmodium* and other unicellular pathogens characterized by A/T-rich genomes.

#### ACKNOWLEDGMENTS

*This work was supported by the grant no. 204/02/1512 of the Grant Agency of the Czech Republic (to F.P.).*

#### REFERENCES

1. Fire, A., S. Xu, M. K. Montgomery, S.A. Kostas, S.E. Driver, and C.C. Mello. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391:806-811.
2. Sharp, P.A. 2001. RNA interference—2001. *Genes Dev.* 15:485-490.
3. Elbashir, S.M., W. Lendeckel, and T. Tuschl. 2001. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* 15: 188-200.
4. Martens, H., J. Novotny, J. Oberstrass, T.L. Steck, P. Postlethwait, and W. Nellen. 2002. RNAi in *Dictyostelium*: the role of RNA-directed RNA polymerases and double-stranded RNase. *Mol. Biol. Cell* 13:445-453.
5. Malhotra, P., P.V. Dasaradhi, A. Kumar, A. Mohammed, N. Agrawal, R.K. Bhatnagar, and V.S. Chauhan. 2002. Double-stranded RNA-mediated gene silencing of cysteine proteases (falcipain-1 and -2) of *Plasmodium falciparum*. *Mol. Microbiol.* 45:1245-1254.
6. Folk, P., F. Puta, L. Krpejsova, A. Blahuskova, A. Markos, M. Rabino, and R.P. Dotin. 1996. The homolog of chromatin binding protein Bx42 identified in *Dictyostelium*. *Gene* 181:229-231.
7. Skruzny, M., M. Ambrozikova, I. Fukova, K. Martinkova, A. Blahuskova, L. Hamplova, F. Puta, and P. Folk. 2001. Cyclophilins of a novel subfamily interact with SNW/SKIP coregulator in *Dictyostelium discoideum* and *Schizosaccharomyces pombe*. *Biochim. Biophys. Acta* 1521:146-151.
8. Morio, T., H. Urushihara, T. Saito, Y. Uga-wa, H. Mizuno, M. Yoshida, R. Yoshino, B.N. Mitra, et al. 1998. The *Dictyostelium* developmental cDNA project: generation and analysis of expressed sequence tags from the first-finger stage of development. *DNA Res.* 5:335-340.

Received 9 June 2003; accepted 10 July 2003.

*Address correspondence to František Půta, Department of Physiology and Developmental Biology, Charles University, Viničná 7, Praha 2, 128 00, Czech Republic. e-mail: dicty@inbox.cesnet.cz*

## Complex DNA melting profiles of small PCR products revealed using SYBR<sup>®</sup> Green I

Weijie Li, Biao Xi, Wancai Yang, Meredith Hawkins, and Ulrich K. Schubart  
*Albert Einstein College of Medicine, Bronx, NY, USA*

*BioTechniques* 35:702-706 (October 2003)

The fluorescent dye SYBR<sup>®</sup> Green I (Roche Diagnostics, Mannheim, Germany) binds to the minor groove of the DNA double helix, and its fluorescence is greatly enhanced upon DNA binding. This property is used to monitor DNA amplification during real-time quantitative PCR in the LightCycler<sup>®</sup> (Roche Diagnostics), a microvolume multicycle fluorimeter with rapid temperature control (1). Products are identified and distinguished by their characteristic melting curves that are dependent on the GC content, length, and sequence of the product (2). However, distinct DNA fragments may have closely overlapping melting profiles and, therefore, a simple melting curve is not sufficient to conclude that a single DNA product is present. Conversely, it is known that longer DNA fragments may have internal melting domains that give rise to complex melting curves (3,4). Smaller DNA products generally melt with a single transition temperature peak (2). Here we show that small DNA fragments may also give rise to more complex melting curves and that the Poland algorithm (5) qualitatively predicts this melting pattern in the majority of cases.

Figure 1 shows two examples of melting curves exhibiting two or more temperature transition peaks and, for comparison, a more typical example, in which a single-peak melting curve is obtained with a single PCR product. The DNA melting patterns predicted by the Poland algorithm (4,5) resemble the observed melting curves, although it is evident in each case that the experimentally observed apparent melting temperature values were significantly higher than those predicted by the algorithm (Figure 1). This is not unexpected because it has been demonstrated that melting temperatures obtained using SYBR Green I fluorescence depend on dye concentration as well as the temperature transition rates (2). Furthermore, for

technical reasons, the solution conditions used by the algorithm could only be approximated to the experimental conditions employed in PCR (for experimental details, see Figure 1 legend). A complex melting curve was also observed when a 188-bp segment of the human  $\beta$ -actin cDNA was amplified (Figure 2). The melting profile was not significantly altered when a slower temperature transition rate (0.05°C/s) was used (data not shown). To gain insight into the origin of this complex melting behavior, we performed additional experiments. The presence of a single 188-bp DNA product was confirmed by agarose gel electrophoresis. To exclude the possibility of co-migrating unrelated DNA, the reaction product was purified and subjected to DNA sequencing, which confirmed the 188-bp sequence derived from human  $\beta$ -actin. It is of note that this DNA product (5'-TTGCGTTACACCCTTTCTTGACA-AAACCTAACTTGCGCAGAAAA-CAAGATGAGATTGGCATGGCTTT-ATTTGTTTTTTTTTTGTTTTGTT-TTGTTTTTTTTTTTTTTTTTTGGC-TTACTAGGATTTAAAACTG-GAACGGTGAAGGTGACAG-CAGTCGGTTGGAGCGAGCA-TCCCCAAAGTTCACAAT-3') contains a centrally located 44-bp A/T-rich region. To determine which subregion contributed to the complexity of the observed melting curve, we cloned the  $\beta$ -actin PCR product and used the resultant plasmid to amplify segments from within the 188-bp insert, as shown schematically in Figure 2A. Agarose gel electrophoresis of the resultant PCR products demonstrated the presence of a single band in each instance (Figure 2D). The predicted and observed melting curves for each of these products are also shown (Figure 2, B and C). As expected, the 67- and 61-bp products, corresponding to the regions 5' and 3' to the A/T-rich segment (primer pairs 1F-2R and 3F-1R,

Research article

Open Access

## Fungal CSL transcription factors

Martin Převorovský, František Půta and Petr Folk\*

Address: Department of Physiology and Developmental Biology, Faculty of Science, Charles University in Prague, Viničná 7, CZ 128 44, Praha 2, Czech Republic

Email: Martin Převorovský - [prevorov@natur.cuni.cz](mailto:prevorov@natur.cuni.cz); František Půta - [puta@natur.cuni.cz](mailto:puta@natur.cuni.cz); Petr Folk\* - [folk@natur.cuni.cz](mailto:folk@natur.cuni.cz)

\* Corresponding author

Published: 13 July 2007

Received: 12 February 2007

BMC Genomics 2007, 8:233 doi:10.1186/1471-2164-8-233

Accepted: 13 July 2007

This article is available from: <http://www.biomedcentral.com/1471-2164/8/233>

© 2007 Martin et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** The CSL (CBF1/RBP-Jκ/Suppressor of Hairless/LAG-1) transcription factor family members are well-known components of the transmembrane receptor Notch signaling pathway, which plays a critical role in metazoan development. They function as context-dependent activators or repressors of transcription of their responsive genes, the promoters of which harbor the GTG(G/A)GAA consensus elements. Recently, several studies described Notch-independent activities of the CSL proteins.

**Results:** We have identified putative CSL genes in several fungal species, showing that this family is not confined to metazoans. We have analyzed their sequence conservation and identified the presence of well-defined domains typical of genuine CSL proteins. Furthermore, we have shown that the candidate fungal protein sequences contain highly conserved regions known to be required for sequence-specific DNA binding in their metazoan counterparts. The phylogenetic analysis of the newly identified fungal CSL proteins revealed the existence of two distinct classes, both of which are present in all the species studied.

**Conclusion:** Our findings support the evolutionary origin of the CSL transcription factor family in the last common ancestor of fungi and metazoans. We hypothesize that the ancestral CSL function involved DNA binding and Notch-independent regulation of transcription and that this function may still be shared, to a certain degree, by the present CSL family members from both fungi and metazoans.

### Background

The CSL (CBF1/RBP-Jκ/Suppressor of Hairless/LAG-1) proteins compose a family of transcription factors essential for metazoan development [1,2]. They are present in all metazoan genomes studied and show remarkable sequence conservation across phylogeny. They localize predominantly or exclusively in the cell nucleus where they can either repress or activate transcription depending on the context and the presence of various coregulators.

CSL proteins recognize a very tightly defined consensus sequence GTG(G/A)GAA in target promoters. Their best characterized function relates to the signaling pathway of the transmembrane receptor Notch where they mediate the effector nuclear step – activation of Notch-responsive genes. The Notch pathway regulates metazoan embryonic development, cell fate decisions and tissue boundaries specifications [2,3]. Its deregulation is implicated in several diseases including cancer [4] and, in addition, several

viruses encode factors that misuse this pathway via interaction with CSL proteins [5].

CSL proteins are essential for the development of the organism as a whole, however, they are dispensable at the cellular level, because CSL knock-out cell lines can be established and do not show any obvious abnormalities. The mutant phenotypes of Notch and CSL genes do not fully overlap, as CSL mutants show more severe developmental perturbations [2,6]. Recently, several studies reported Notch-independent activities of CSL proteins indicative of their involvement in yet other signaling pathways [7-10]. In addition to the Notch pathway-dependent CSL proteins of the RBP- $\kappa$  type, at least in some metazoan species, CSL transcription factors called RBP-L can be found, which are only beginning to be characterized. They are highly similar to the RBP- $\kappa$  group but seem to act exclusively in a Notch-independent manner. Unlike the ubiquitous RBP- $\kappa$  type proteins the expression of RBP-L is confined to only a few tissue types [11,12].

In contrast to the generally accepted view, the presence of CSL proteins seems not to be confined to metazoan organisms and the Notch pathway. They are indeed absent from plants but there were indications of CSL proteins in one fungal species – the fission yeast *Schizosaccharomyces pombe* [13]. We have attempted to confirm the identity of CSL proteins in *S. pombe* and to further explore the distribution of this transcription factor family in fungi. We have documented the existence of fungal CSL proteins, which indicates that this family originated much earlier in evolution than previously appreciated. We hope that these findings will help to elucidate the CSL family ancestral function in cells and to better understand their complex engagements in metazoans.

## Results

### Identification of CSL genes in fungi

CSL transcription factors are generally considered a key part of the Notch signaling pathway and as such a hallmark of metazoan organisms [2]. However, it was noted earlier in the literature that distant CSL homologs may also be found in the genome of the fission yeast *Schizosaccharomyces pombe*, an organism that lacks the Notch pathway [13]. This raises interesting questions regarding the evolutionary origin as well as the ancestral function of the CSL family. We have therefore conducted exhaustive BLAST searches of publicly available sequence data (see Methods) to assess the presence and conservation of CSL family members in fungi. The results of these searches are summarized in Table 1 (the fungal taxonomical nomenclature used in this article was taken from [14]). Nineteen putative CSL genes were found in seven organisms, with *S. pombe* and *S. japonicus* belonging to the Taphrinomycotina basal subphylum of ascomycetes, *Rhizopus oryzae*

representing the zygomycetes and *Coprinus cinereus*, *Cryptococcus neoformans*, *Phanerochaete chrysosporium* and *Ustilago maydis* belonging to the basidiomycetes. Protein products of these genes contain motifs typical of the CSL family (see below). It is likely that more CSL genes will be found in these taxonomical groups as more genome sequences become available. In contrast, no CSL homologs could be found in either Saccharomycotina (including the budding yeast *Saccharomyces cerevisiae*) or Pezizomycotina, the later branching subphyla of ascomycetes.

Most of the candidates are hypothetical proteins with little or no annotation in the databases. Therefore, we have first verified the quality of each ORF prediction (see Methods). The confidence of exon-intron structure predictions in these less studied organisms is rather limited. Another obstacle is posed by the degree of divergence among the sequences together with the presence of multiple species- and protein-specific insertions. Nevertheless, we were able to construct three completely new gene predictions (designated SjCSL1 and SjCSL2 in *S. japonicus*, and PcCSL2 in *P. chrysosporium*) as well as to identify mispredictions and/or possible sequencing errors in other four genes (see Additional files 1 and 2 for a more detailed description). Our corrections comprised of intron inclusion/exclusion, different splice-site selection and exon addition. Some of the intron positions displayed inter-species conservation which supported our predictions (data not shown). We have also identified a less usual intron with a GC-AG boundary in the *R. oryzae* RO3G\_07636.1 gene. Such introns were found in other fungi as well [15] and are generally a problem for gene prediction algorithms.

Typically, there are two CSL paralogs per genome, differing considerably in length and each belonging to a different class (see below). A notable exception is the genome of *R. oryzae* which harbors seven CSL genes, three of them being class F1 and four of them belonging to class F2. Most candidate CSL proteins are predicted to be nuclear which supports their putative functioning as transcription factors (see below). SPCC736.08 of *S. pombe* is the only protein predicted to have exclusively non-nuclear subcellular localization but it was shown experimentally to be nuclear [16].

### Sequence conservation of fungal CSL proteins

According to the *C. elegans* LAG-1 protein crystal structure, the CSL fold is related to Rel-domain proteins, but is uniquely composed of three distinct domains [17]. The amino-terminal RHR-N (Rel-homology region) and central BTD (beta-trefoil domain) domains are involved in DNA-binding. BTD serves also as an interaction platform for Notch/SMRT coregulators. The carboxy-terminal RHR-C domain displays lower conservation in metazoans and

Table 1: Fungal CSL proteins

Organism	Protein	Accession number/Locus <sup>a</sup>	Length (aa)	Status <sup>b</sup>	Nuclear <sup>c</sup>	Source
<b>Ascomycota: Taphrinomycotina</b>						
<i>Schizosaccharomyces pombe</i>	SPCC1223.13	CAA20882	963	Exp <sup>d</sup>	+++	[40]
	SPCC736.08	<u>NP_587779</u>	613	Exp <sup>d</sup>	-	[40]
<i>Schizosaccharomyces japonicus</i>	SjCSL2	Supercontig 4 (bases 1104530–1107169)	879	Hyp	+++	[31] <sup>e</sup>
	SjCSL1	Supercontig 5 (bases 726033–727712)	559	Hyp	+++	[31] <sup>e</sup>
<b>Zygomycota</b>						
<i>Rhizopus oryzae</i>	RO3G_06481	RO3G_06481.1	694	Hyp	+++	[51]
	RO3G_07636	RO3G_07636.1	662	Hyp	++	[51] <sup>e</sup>
	RO3G_11583	RO3G_11583.1	764	Hyp	+++	[51] <sup>e</sup>
	RO3G_14587	RO3G_14587.1	886	Hyp	+++	[51] <sup>e</sup>
	RO3G_06953	RO3G_06953.1	449	Hyp	+++	[51]
	RO3G_08863	RO3G_08863.1	482	Hyp	+++	[51]
	RO3G_13784	RO3G_13784.1	478	Hyp	+++	[51]
<b>Basidiomycota</b>						
<i>Coprinus cinereus</i>	CCIG_03194	EAU91026	960	Hyp	+++	[40]
	CCIG_01706	CCIG_01706.1	803	Hyp	+	[52] <sup>e</sup>
<i>Cryptococcus neoformans</i>	CNBD3370	EAL21283	1015	Hyp	+++	[40]
	CNA01890	AAW40742	776	Exp	+++	[40]
<i>Phanerochaete chrysosporium</i>	PcCSL2	Scaffold 6 Contig 19 (bases 50978–54385)	1012	Hyp	++	[53] <sup>e</sup>
	Pc6518	protein id "6518"	960	Hyp	++	[53]
<i>Ustilago maydis</i>	UM06280	EAK82808	1482	Hyp	+++	[40]
	UM05862	EAK86807	1094	Hyp	+++	[40]

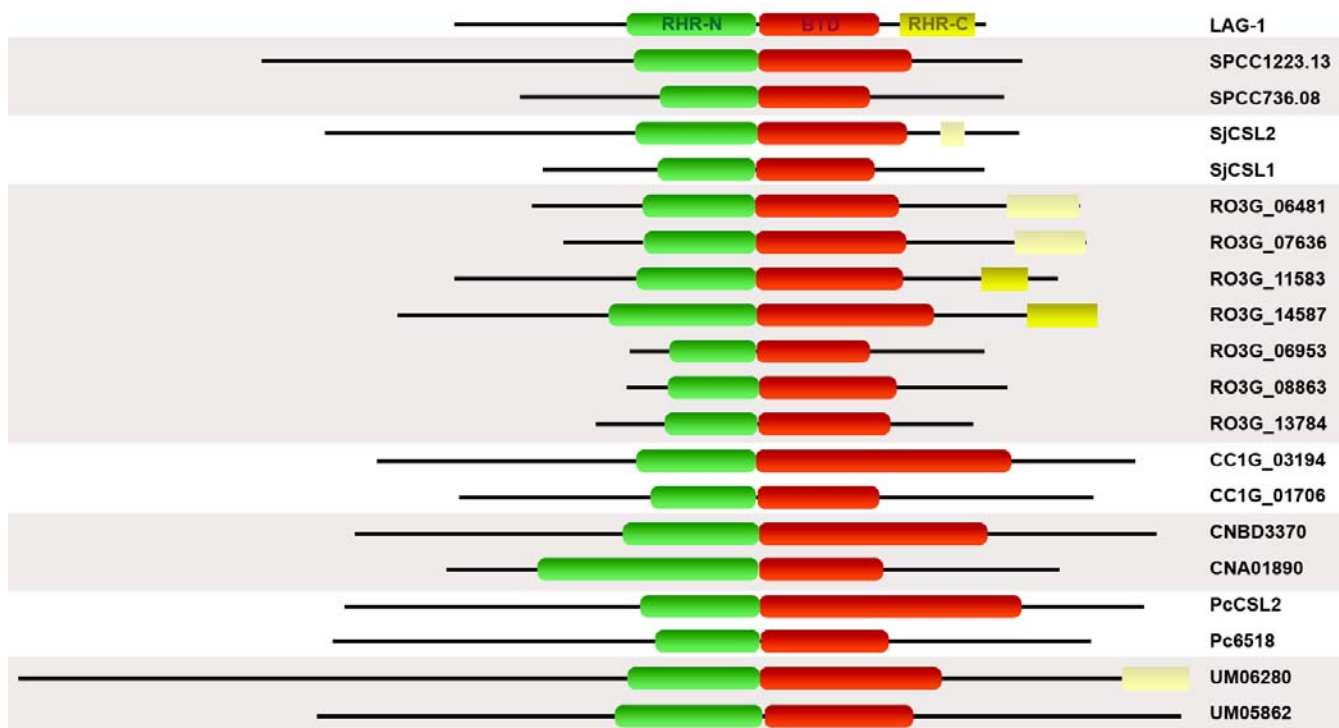
Notes: <sup>a</sup>refers to the respective database stated in the 'Source' column; <sup>b</sup>Exp – expressed protein, Hyp – hypothetical protein; <sup>c</sup>nuclear localization prediction score (see Methods); <sup>d</sup>MP et al., manuscript in preparation; <sup>e</sup>see Additional files 1 and 2 for new sequence predictions and corrections made to the original database sequences.

its function is not yet clear; one possibility is its participation in Notch-independent regulation of transcription [18].

We have used the Pfam protein domains database [19] to search for CSL-specific domains in all our candidate sequences and to identify any other known domains present. The results are schematized in Fig. 1. The RHR-N [Pfam:PF09271] and BTB [Pfam:PF09270] domains were identified in all fungal sequences with high significance, supporting the identity of our candidates as CSL family members. However, the RHR-C [Pfam:PF01833] domain could only be identified in RO3G\_11583 and RO3G\_14587 from *R. oryzae*. A rather divergent RHR-C domain was also found in *S. japonicus* SjCSL2 and two more *R. oryzae* proteins, RO3G\_06481 and RO3G\_07636. The lower degree of sequence conservation of RHR-C noted in metazoans is thus even more pronounced in fungi. No other conserved domains could be found, despite the fact that the putative fungal CSL proteins are typically significantly larger than their metazoan counterparts. The overall domain organization of the fungal proteins is the same as in metazoans. The increased size of the

fungal candidates was found to be caused by two factors. First, in some proteins, there are pronounced extensions of the amino-terminal part preceding the RHR-N domain. This region is about 200 amino acids long in *C. elegans* and gets much shorter in metazoan evolution. Its crystal structure is not known. Second, there are multiple amino acid insertions of varying length throughout the candidate sequences (see below).

To gain better insight into the specifics of the fungal CSL proteins, we have produced a multiple sequence alignment of all newly identified fungal sequences and selected metazoan family members (see Methods and Additional file 3). There are two sub-types of metazoan CSL proteins; one is represented by the Notch-pathway protein RBP-J $\kappa$  (CBF1, SuH, RBPSUH) and the other by the much less known transcription factor RBP-L, the function of which seems to be Notch-independent [11,12]. Both subtypes' representatives were included in the alignment. The most prominent feature of the resulting alignment is the presence of several highly conserved blocks of amino acids separated by species- and protein-specific insertions. These insertions are of considerable length in some cases



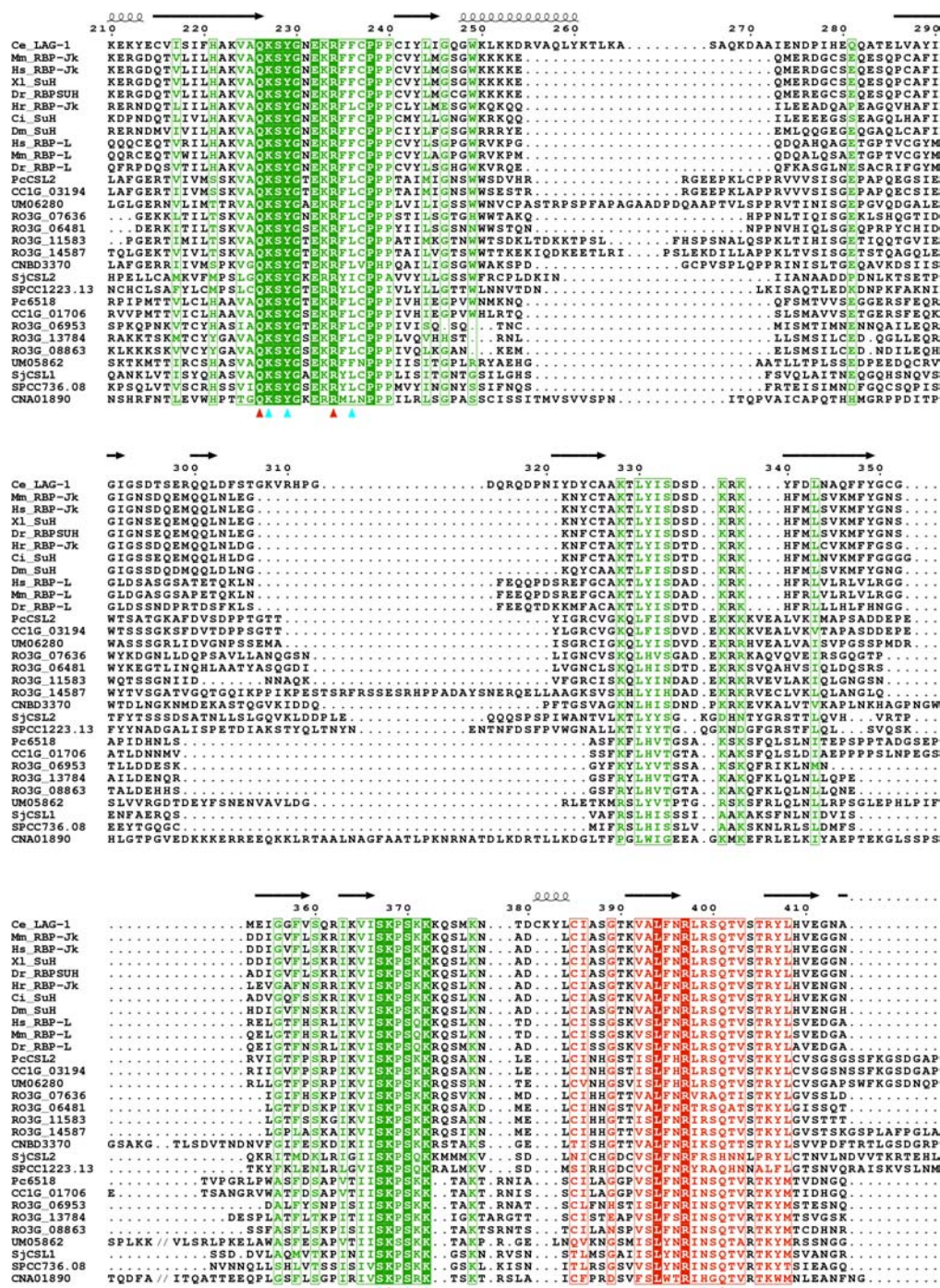
**Figure 1**  
**Fungal CSL proteins domain organization.** Black lines represent the respective CSL protein sequences (see Table 1 for details). The structure of *C. elegans* LAG-1 is shown at the top for comparison [17]. Recognized Pfam domains are indicated: RHR-N in green, BTD in red and RHR-C in yellow (light yellow for low significance). The proteins are drawn to scale.

and are more pronounced in the class F2 proteins. They are rich in amino acids proline, glycine, serine/threonine and lysine/arginine. Overall sequence conservation is highest in the RHR-N and BTD domains, including the immediately following long  $\beta$ -strand ( $\beta$ C4) that was shown to bridge all three CSL domains in the *C. elegans* LAG-1 [17]. The conservancy of the  $\beta$ C4 linker suggests that the CSL-specific arrangement between RHR-N and BTD is also likely preserved in fungi. The C-termini typically contain only 1–2 well-alignable stretches that can be identified as fragments of the RHR-C domain. The amino-terminal extensions preceding the RHR-N domain show little if any sequence conservation. As mentioned above, there are several regions located mostly in the RHR-N and BTD domains, that show very high or even absolute sequence conservation (see Fig. 2 and 3). It is notable that, according to the crystallography data, all these conserved blocks are involved in binding of the strictly defined CSL consensus site on DNA [17]. With the sole exception of the *S. japonicus* SjCSL2 protein (Q567H substitution corresponding to Q401 in *C. elegans* LAG-1, see Fig. 2), all residues required for sequence specific binding of the GTG(G/A)GAA response element are absolutely conserved in all fungal proteins, which strongly supports their inclusion in the CSL family. The interactions of CSL pro-

teins with their coactivators Notch/EBNA2 and corepressors SMRT/NCoR and CIR have been mapped to and around a hydrophobic pocket on the surface of BTD [17,20–22]. Not surprisingly, the residues mediating these interactions are generally not conserved in fungi, although some of them are found in class F2 fungal CSL proteins. However, the potential to form a hydrophobic pocket in BTD seems to be preserved (data not shown).

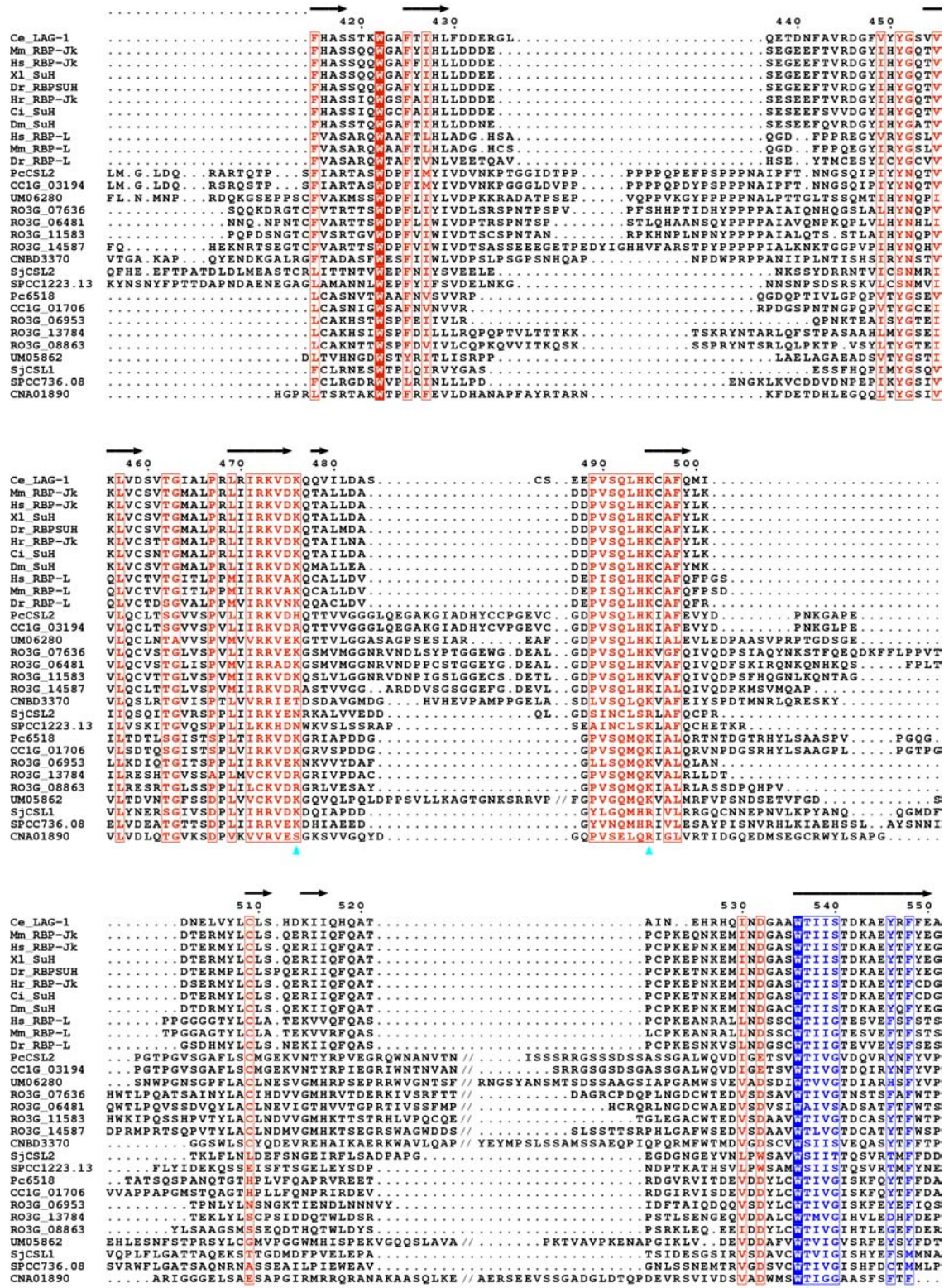
**Phylogenetic analysis of the CSL protein family**

As noted earlier, there are usually two fungal CSL paralogs per genome. We wanted to see whether these paralogs cluster to some well-defined groups and what their relationship to the metazoan CSL family members is. For this purpose, we have constructed an unrooted phylogenetic tree for the regions that could be aligned with confidence, that is, the RHR-N and BTD domains (see Methods and Fig. 4). As expected, the fungal CSL proteins form two distinct classes, designated class F1 and F2, with each class being represented in all fungal taxons included in the analysis. It should be noted at this point that the positions of *S. pombe* SPCC1223.13 and *S. japonicus* SjCSL2 proteins are slightly ambiguous, branching off either immediately before or after the class F2 core (data not shown). The intra-class branch topology roughly follows the taxonom-



**Figure 2**  
**Evolutionary conservation of the DNA-binding regions.** The alignment of fungal and selected metazoan CSL protein sequences (see Table 1 and Additional file 3 for details) shows high degree of conservation in regions responsible for DNA binding. Absolutely conserved residues are inverse-printed, positions with high residue similarity are boxed. Domain boundaries are indicated by color: green for RHR-N, red for BTB and blue for the  $\beta$ 4 linker connecting all three CSL domains. Red and cyan triangles below the alignment denote residues required for sequence specific and backbone DNA binding, respectively. The position numbering and secondary structures indicated above the sequences correspond to *C. elegans* LAG-1 [17]. The picture shows only a selected region of the whole alignment and, in order to save space, some parts of the long inserts are not shown (indicated by //). The picture was created using ESPript [50].





**Figure 3**  
Evolutionary conservation of the DNA-binding regions – continued. The continuation of the alignment shown in Fig. 2.

ical relations [23] with the notable exception of the divergent *C. neoformans* CNA01890 and CNBD3370 proteins. It can be inferred from the branch lengths that the rate of divergence among the fungal protein sequences is much higher than in metazoa. Metazoan CSL proteins (designated class M) form a very coherent group that can be divided to RBP-J $\kappa$  and RBP-L subgroups. The RBP-J $\kappa$  subgroup displays an especially low extent of divergence, which may be due to their involvement in the developmentally critical Notch pathway. Of the two fungal CSL classes the class F2 proteins show higher similarity to the metazoan class M.

## Discussion

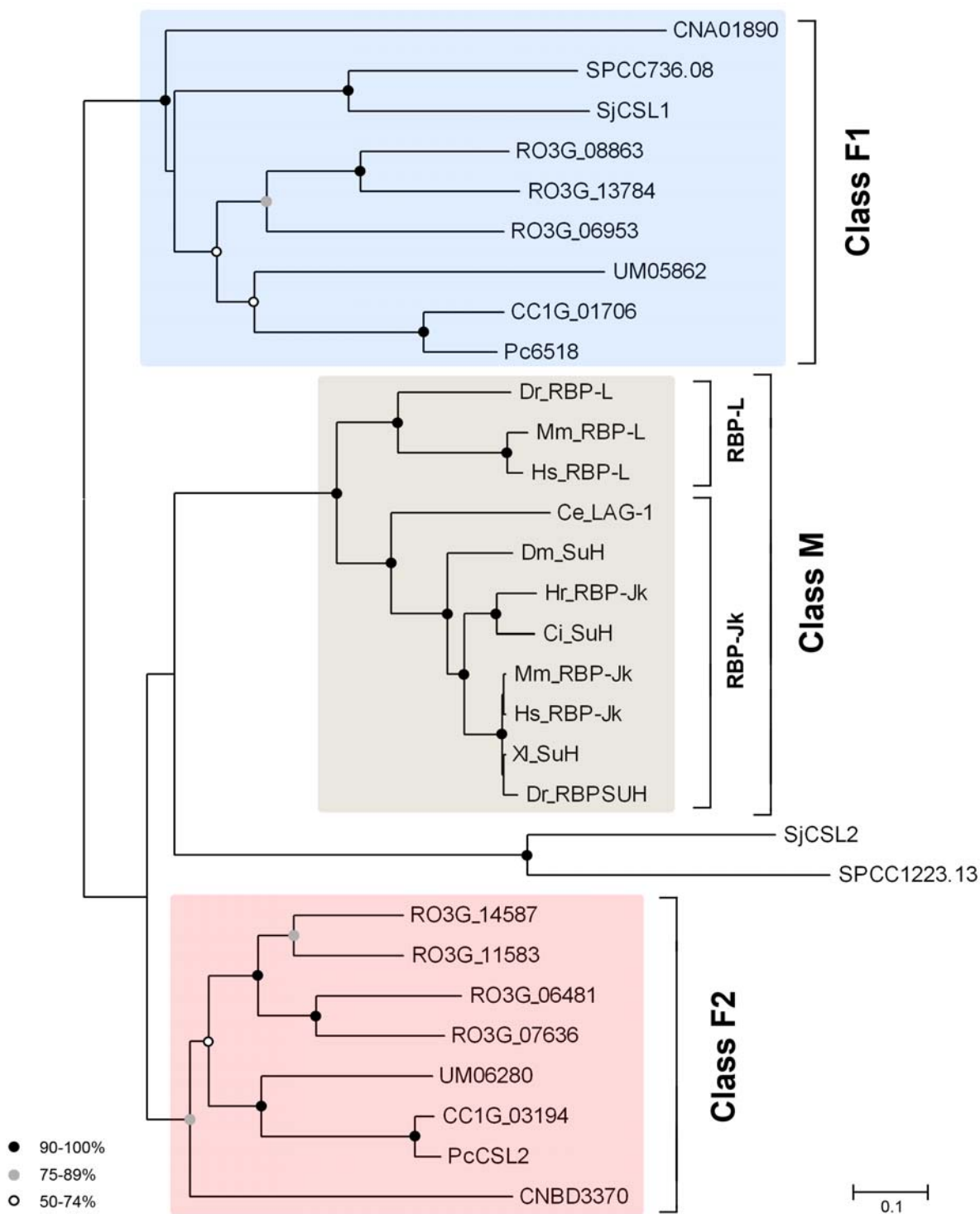
### The CSL family origin and distribution

To the best of our knowledge, there were only two brief notions of CSL proteins existence outside metazoans up to now. One paper showed Southern blot cross-hybridization of murine RBP-J $\kappa$  cDNA probe with *S. pombe* DNA [24]. The significance of these results is, however, questionable, as the hybridizing chromosomal DNA fragments had lengths differing from that expected for either of *S. pombe* CSL genes, SPCC736.08 and SPCC1223.13. Potential CSL homologs in *S. pombe* were also mentioned in the review of Lai [13], although no supporting evidence was presented.

We have rigorously searched for CSL proteins in eukaryotic genomes from all kingdoms of life to map their distribution. Apart from the known metazoan proteins, we have found no homologs in either plants or protozoa (data not shown), however, we have succeeded in finding CSL family members in several fungal species of the ascomycetes (the basal subphylum Taphrinomycotina), zygomycetes and basidiomycetes groups. These organisms range in complexity from the simple unicellular fission yeast to the macroscopic multicellular and highly differentiated *C. cinereus*. It is of notion that the presence of CSL homologs in fungi is not universal as there are no representatives found in either of the later branching ascomycetal groups, Saccharomycotina, including the important model organism *S. cerevisiae*, and Pezizomycotina. Our data support the idea that the ancestral CSL gene originated in the last common ancestor of animals and fungi, thus much earlier than previously assumed. This is in accord with the absence of CSL family in such large groups as plants and mycetozoa, that branched off earlier in evolution [25,26]. We hypothesize that the first CSL gene might have been created from a Rel-type transcription factor gene by the insertion of a beta-trefoil domain-encoding DNA sequence in between the amino- and carboxy-terminal Rel domains. Subsequently, a duplication event took place in the fungal lineage creating the two CSL classes we see today, class F2 being more alike the metazoan CSL proteins and class F1 being more fungi-specific

(see Fig. 4). We consider such explanation more likely than the alternative, where the ancestral CSL gene would both originate and undergo duplication in the common ancestor of metazoans and fungi and one copy would be soon lost again in the metazoan lineage.

Nevertheless, there have been independent losses of CSL genes in the fungal branch. First, we failed to find any CSL homologs in *Encephalitozoon cuniculi* (data not shown), a parasitic microsporidian and a representative of a group that is sister to fungi [25]. This fact is probably due to the parasitic lifestyle of these organisms, which often leads to pronounced gene eliminations [27]. Second, we have found no evidence of CSL genes in chytridiomycetes (data not shown), a likely polyphyletic group also basal to the fungal lineage [14]. Finally, the CSL family is apparently missing in the later branching ascomycetal fungi of the Saccharomycotina and Pezizomycotina groups [23], suggestive of another gene loss(es). The losses may have occurred during the transitions between saprophytic and parasitic nutritional modes [14], indicating that the CSL genes code for functions in fungi that are not universally required in their life cycles. On the other hand, there have been clade specific CSL genes multiplications in fungi illustrated by the three class F1 and four class F2 CSL genes of *Rhizopus oryzae*. Evolutionary pressure could have favored proliferation and diversification of the CSL family in this branch of zygomycetes, similarly to the expansions that were documented for other gene families and phyla, such as, e.g., nuclear hormone receptors and nematodes, or calmodulin-type proteins and dictyostelids, respectively [28,29]. A history of gene losses and duplications in the fungal lineage has also been described for proteins involved in various RNA silencing phenomena [30]. The metazoan CSL genes (class M) obviously underwent duplication too. It likely occurred in the common ancestor of all vertebrates and gave rise to the RBP-L type of proteins, in addition to the RBP-J $\kappa$  type universally present in both vertebrate and invertebrate animals. It should be noted in this regard, that the RBP-L type gene is present in zebrafish, but so far no homologs have been reported in the genetically rather complicated clawed frog *Xenopus laevis*. We have also failed to identify an RBP-L homolog in the more tractable species *X. tropicalis*, thus amphibians likely have developed ways to regulate all their CSL-responsive genes using the RBP-J $\kappa$  homolog only. In summary, we have found representatives of the important transcription factor family CSL, up to now generally considered metazoan-only, in several groups of fungi and showed that they are an ancient gene family that originated much earlier than their current metazoan affiliates like Notch or Mastermind [13].



**Figure 4**  
**Phylogenetic analysis of the CSL protein family.** An unrooted neighbor-joining phylogenetic tree of the region corresponding to RHR-N and BTD domains (see Methods). For protein descriptions see Table 1 and Additional file 3. For class F2 only the unambiguous core, not including the *S. pombe* SPCC1223.13 and *S. japonicus* SjCSL2, is indicated by shading. Symbols at nodes indicate percentual bootstrap values, no symbol means less than 50% node stability. The scale bar indicates the number of amino acid substitutions per site.

### **The conservation of fungal CSL proteins**

The degree of conservation of CSL proteins across phylogeny is remarkable, given the evolutionary distances, and points to an important role they likely play in cells [25]. The sequence similarity among metazoan CSL proteins is extremely high and does not allow for finding functionally important regions directly from sequence comparison. On the other hand, the distant CSL homologs from fungi may provide this information more readily. Indeed, we have found that the most prominent conservation can be found in the regions involved in DNA binding with the critical residues and several motifs being invariant in all proteins analyzed (see Fig. 2 and 3). As expected, when compared to metazoans, the rate of divergence has been much faster in fungi, especially in those having small genomes, i.e. *C. neoformans*, *S. pombe* and *S. japonicus* [31-33]. In fact, the *C. neoformans* CSL proteins are the most divergent ones among fungi and their position in our phylogenetic tree (Fig. 4) differs from that expected by looking at the fungal tree of life [23]. Such discrepancy has also been reported for other *C. neoformans* proteins [30] and it has been demonstrated for *S. pombe* that various types of proteins might produce inconsistent signals when used for phylogenetic analyses [34].

There are numerous insertions separating the above-mentioned conserved sequence stretches, but these insertions are often rich in amino acids that are likely to appear in loops and solvent-exposed regions [35]. In addition, such insertions are present, to a lesser degree, also in the *C. elegans* LAG-1, the most evolutionarily primitive CSL protein studied so far [17]. It may be argued that the fungal insertions could be an artifact produced by ORF misprediction. We cannot rule out this possibility completely as the tools for identifying exon-intron boundaries optimized for diverse fungal species are limited or lacking. However, many of these insertions are conserved among the classes of CSL proteins and their positions mostly correspond to the LAG-1 loops and regions exposed on the surface of the protein [17]. Thus the general CSL fold may be well preserved in fungi.

Furthermore, the splicing pattern of some fungal CSL genes is partially conserved among species (data not shown) and the ORF predictions used in this study are in good agreement with the multiple sequence alignment of the proteins they encode. Nevertheless, the prediction reliability of the non-conserved amino-terminal extensions found in some fungal CSL proteins remains questionable. The sequence similarity in the parts of the fungal proteins corresponding to known coregulator interaction sites in metazoans seems not to be significantly preserved. This is of no great surprise as these coregulators are frequently involved in the Notch signaling pathway, which is lacking in fungi, or are encoded by mammalian viruses

[5,13]. Also, the less-conserved metazoan RHR-C domain of yet unknown function is very loosely defined in fungi, as it was identified with confidence only in several class F2 members. Taken together, our data suggest that the fungal CSL proteins may adopt the CSL fold and we further show that these proteins possess notably conserved regions of functional significance related mostly to their ability to bind DNA in a sequence-specific manner.

### **The ancestral role of the CSL transcription factor family**

Our current knowledge of the CSL family derives exclusively from metazoan model organisms and is based mostly on studies concerning development and the Notch pathway [2,9,13]. It is now clear that this is not the whole picture as we have presented evidence of CSL proteins in several organisms that are evolutionarily distant to animals and lack the critical Notch pathway components. Moreover, recent reports on metazoan model organisms indicate, that there are yet unrecognized CSL activities in animals as well [7,8,10,11]. It is tempting to speculate that the CSL ancestral function is preserved in the fungal proteins of today and maybe even in metazoans, where it might be responsible for some of the Notch-independent activities observed. If this is the case we would have excellent models, e.g., the genetically tractable fission yeast *S. pombe*, to study it.

We hypothesize that the ancestral function is likely the regulation of gene expression, where other signals than Notch receptor activation are interpreted. Our first clue comes from the analysis of fungal CSL sequence conservation, which clearly indicates their potential to bind DNA. This includes not only DNA binding in general, but goes further to the ability to recognize the strict CSL consensus. The second clue derives from the lack of conservation of CSL interacting partners from metazoans. As stated above, the Notch receptor, its ligands and coactivators are not present in fungi. Finally, the metazoan CSL proteins are essential for embryonic development but dispensable in cultured cells [6]. Similarly, the deletion of either or both *S. pombe* CSL genes is viable (MP et al., manuscript in preparation; and [36,37]). This suggests, together with the secondary loss of CSL genes in some fungi (see above), that the proposed ancestral function in gene regulation is not essential.

We also have to account for the existence of two CSL classes in fungi. There is analogy to the metazoan class M sub-groups, the RBP-J $\kappa$  and RBP-L CSL types. Both are involved in transcription regulation, but differ in their interacting partners, their responsiveness to various signals, their expression profiles and their *in vivo* DNA-binding preferences [11,12]. The similar may be true for class F1 and class F2 fungal CSL proteins. They may all participate in transcription regulation, but have either distinct or

only partially overlapping target gene sets. Alternatively, they may differentially regulate the same genes, with the outcome depending on, e.g., environmental conditions. It was indeed found by whole-genome microarray experiments, that the *S. pombe* CSL genes display differential expression during sexual differentiation and under various stress conditions [38,39]. In conclusion, the CSL gene family encodes proteins that are likely universally involved in the regulation of transcription both in animals and fungi.

## Conclusion

We have shown the existence of CSL transcription factor family, known from studies of the metazoan Notch signaling pathway, in several fungal species. We have described conserved features of the fungal proteins supporting their identity as true CSL family members. These findings put the CSL family origin further back in evolution, deeper than currently understood. We have mapped the history of CSL gene duplication and gene loss events in the fungal lineage, showing the existence of two well-defined CSL classes, class F1 and class F2, respectively, with the second class being more similar to the metazoan class M proteins. We hypothesize that the ancestral CSL function involved DNA binding and Notch-independent regulation of transcription and that this function may still be shared, to a certain degree, by the present CSL family members from both fungi and metazoans. If true, that would allow for exploiting the simple fungal models to analyze this function. We are currently studying the CSL proteins role in *S. pombe* and experiments are underway to identify the sets of genes and processes they regulate.

## Methods

### Database searches for CSL genes

We have searched multiple publicly available fungal genome and protein databases (including NCBI [40] and UniProt [41]) using the appropriate BLAST algorithm with default settings and with the mouse CBF1 protein [GenBank:NP\_033061] as a query. Candidate hits containing at least one of the conserved CSL motifs (see Results) were considered and used for further analyses. The BLAST searches were then repeated with all the newly identified CSL sequences as queries until no more new hits were found. In cases where two or more nearly-identical candidate sequences, coming from independent sources and obviously representing a single gene, were found, the sequence showing the highest degree of similarity to the fungal CSL consensus was chosen. The final searches were performed between November 24, 2006 and November 30, 2006.

### Gene models prediction and verification

All candidate fungal CSL proteins were checked for the quality of their ORF prediction. We compared each data-

base gene model with GenScan [42] and/or WebGene [43] predictions. The models were also compared to a multiple sequence alignment of other CSL proteins. In some cases, the splicing pattern was corrected manually using the Gene Runner 3.05 software (Hastings Software, Inc.) in order to restore a highly conserved region (see Results and Additional files 1 and 2).

### Conserved domain search and protein localization prediction

Known domains present in the fungal CSL proteins were searched for by the Search Pfam server [19]. Subcellular localization of each CSL protein was predicted by three independent algorithms, namely SubLoc v1.0 [44], CELLO v.2.5 [45] and PSORT II [46]. Each sequence received score ranging from '-' to '+++' depending on the number of times the protein was predicted to be nuclear (see Table 1).

### Sequence alignments and phylogenetic analyses

Alignments used during the sequence retrieval part of the study were performed using ClustalW [47]. The final alignment of all identified fungal and selected metazoan CSL proteins was based on a ClustalX output (Blosum matrix series) [48], which was then manually edited in BioEdit 7.0.5.3 to correct some obvious alignment errors and to account for the information from the *C. elegans* CSL protein crystal structure [17]. See Additional file 3 for the final alignment and the list of metazoan sequences used.

For tree construction all positions containing gaps were removed from the final sequence alignment. An unrooted phylogenetic tree was then generated for the region corresponding to RHR-N and BTM domains (from helix  $\alpha$ 2 just before the  $\beta$ C4 linker, residues 210–535 in the *C. elegans* LAG-1 reference protein, see [17]) using the neighbor-joining method in the MEGA 3.1 software package [49] with 2000 bootstrap replicates.

### Authors' contributions

MP designed the study, carried out database searches and phylogenetic analyses and drafted the manuscript. FP participated in the study design and final manuscript preparation. PF participated in the final manuscript preparation. All authors read and approved the final manuscript.

### Additional material

#### Additional file 1

New and corrected fungal CSL gene prediction models

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-8-233-S1.doc>]

**Additional file 2**

New and corrected fungal CSL protein prediction models

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-8-233-S2.doc>]

**Additional file 3**

CSL proteins sequence alignment used for the phylogenetic analyses

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-8-233-S3.txt>]

**Acknowledgements**

This work was supported by the Grant Agency of the Charles University grant no. I57/2005/B-BIO/PrF, the Czech Science Foundation grant no. 204/03/H066 and the Czech Ministry of Education, Youth and Sport grant no. MSM0021620858.

We would like to thank Marian Novotný and Fatima Cvrčková for their expert help and suggestions in the initial phase of this study.

Data for *P. chrysosporium* CSL gene model prediction has been provided freely by the JGI for use in this publication only.

Data for *R. oryzae*, *C. cinereus* and *S. japonicus* CSL gene model prediction were obtained from the *Rhizopus oryzae* Sequencing Project, *Coprinus cinereus* Sequencing Project and *Schizosaccharomyces japonicus* Sequencing Project, respectively. Broad Institute of Harvard and MIT <http://www.broad.mit.edu>.

**References**

- Pursglove SE, Mackay JP: **CSL: a notch above the rest.** *Int J Biochem Cell Biol* 2005, **37**:2472-2477.
- Artavanis-Tsakonas S, Rand MD, Lake RJ: **Notch signaling: cell fate control and signal integration in development.** *Science* 1999, **284**:770-776.
- Weinmaster G, Kintner C: **Modulation of Notch signaling during somitogenesis.** *Annu Rev Cell Dev Biol* 2003, **19**:367-395.
- Weng AP, Aster JC: **Multiple niches for Notch in cancer: context is everything.** *Curr Opin Genet Dev* 2004, **14**:48-54.
- Hayward SD: **Viral interactions with the Notch pathway.** *Semin Cancer Biol* 2004, **14**:387-396.
- Oka C, Nakano T, Wakeham A, de la Pompa JL, Mori C, Sakai T, Okazaki S, Kawaichi M, Shiota K, Mak TW, Honjo T: **Disruption of the mouse RBP-J kappa gene results in early embryonic death.** *Development* 1995, **121**:3291-3301.
- Koelzer S, Klein T: **A Notch-independent function of Suppressor of Hairless during the development of the bristle sensory organ precursor cell of Drosophila.** *Development* 2003, **130**:1973-1988.
- Kaspar M, Klein T: **Functional analysis of murine CBF1 during Drosophila development.** *Dev Dyn* 2006, **235**:918-927.
- Bray S, Furriols M: **Notch pathway: making sense of suppressor of hairless.** *Curr Biol* 2001, **11**:R217-R221.
- Barolo S, Walker RG, Polyakov AD, Freschi G, Keil T, Posakony JW: **A notch-independent activity of suppressor of hairless is required for normal mechanoreceptor physiology.** *Cell* 2000, **103**:957-969.
- Beres TM, Masui T, Swift GH, Shi L, Henke RM, MacDonald RJ: **PTF1 is an organ-specific and Notch-independent basic helix-loop-helix complex containing the mammalian Suppressor of Hairless (RBP-J) or its paralogue, RBP-L.** *Mol Cell Biol* 2006, **26**:117-130.
- Minoguchi S, Taniguchi Y, Kato H, Okazaki T, Strobl LJ, Zimmer-Strobl U, Bornkamm GW, Honjo T: **RBP-L, a transcription factor related to RBP-Jkappa.** *Mol Cell Biol* 1997, **17**:2679-2687.
- Lai EC: **Keeping a good pathway down: transcriptional repression of Notch pathway target genes by CSL proteins.** *EMBO Rep* 2002, **3**:840-845.
- James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, Celio G, Gueidan C, Fraker E, Miadlikowska J, Lumbsch HT, Rauhut A, Reeb V, Arnold AE, Amtoft A, Stajich JE, Hosaka K, Sung GH, Johnson D, O'Rourke B, Crockett M, Binder M, Curtis JM, Slot JC, Wang Z, Wilson AW, Schussler A, Longcore JE, O'Donnell K, Mozley-Standridge S, Porter D, Letcher PM, Powell MJ, Taylor JW, White MM, Griffith GW, Davies DR, Humber RA, Morton JB, Sugiyama J, Rossman AY, Rogers JD, Pfister DH, Hewitt D, Hansen K, Hambleton S, Shoemaker RA, Kohlmeyer J, Volkmann-Kohlmeyer B, Spotts RA, Serdani M, Crous PW, Hughes KW, Matsuura K, Langer E, Langer G, Unterwiesing WA, Lueking R, Budel B, Geiser DM, Aptroot A, Diederich P, Schmitt I, Schultz M, Yahr R, Hibbett DS, Lutzoni F, McLaughlin DJ, Spatafora JW, Vilgalys R: **Reconstructing the early evolution of Fungi using a six-gene phylogeny.** *Nature* 2006, **443**:818-822.
- Rep M, Duyvesteyn RG, Gale L, Usgaard T, Cornelissen BJ, Ma LJ, Ward TJ: **The presence of GC-AG introns in Neurospora crassa and other euascomycetes determined from analyses of complete genomes: implications for automated gene prediction.** *Genomics* 2006, **87**:338-347.
- Matsuyama A, Arai R, Yashiroda Y, Shirai A, Kamata A, Sekido S, Kobayashi Y, Hashimoto A, Hamamoto M, Hiraoka Y, Horinouchi S, Yoshida M: **ORFeome cloning and global analysis of protein localization in the fission yeast Schizosaccharomyces pombe.** *Nat Biotechnol* 2006, **24**:841-847.
- Kovall RA, Hendrickson WA: **Crystal structure of the nuclear effector of Notch signaling, CSL, bound to DNA.** *EMBO J* 2004, **23**:3441-3451.
- Tang Z, Kadesch T: **Identification of a novel activation domain in the Notch-responsive transcription factor CSL.** *Nucleic Acids Res* 2001, **29**:2284-2291.
- Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V, Lassmann T, Moxon S, Marshall M, Khanna A, Durbin R, Eddy SR, Sonnhammer EL, Bateman A: **Pfam: clans, web tools and services.** *Nucleic Acids Res* 2006, **34**:D247-D251.
- Sakai T, Taniguchi Y, Tamura K, Minoguchi S, Fukuhara T, Strobl LJ, Zimmer-Strobl U, Bornkamm GW, Honjo T: **Functional replacement of the intracellular region of the Notch1 receptor by Epstein-Barr virus nuclear antigen 2.** *J Virol* 1998, **72**:6034-6039.
- Fuchs KP, Bommer G, Dumont E, Christoph B, Vidal M, Kremmer E, Kempkes B: **Mutational analysis of the J recombination signal sequence binding protein (RBP-J)/Epstein-Barr virus nuclear antigen 2 (EBNA2) and RBP-J/Notch interaction.** *Eur J Biochem* 2001, **268**:4639-4646.
- Hsieh JJ, Zhou S, Chen L, Young DB, Hayward SD: **CIR, a corepressor linking the DNA binding factor CBF1 to the histone deacetylase complex.** *Proc Natl Acad Sci U S A* 1999, **96**:23-28.
- Kuramae EE, Robert V, Snel B, Weiss M, Boekhout T: **Phylogenomics reveal a robust fungal tree of life.** *FEMS Yeast Res* 2006, **6**:1213-1220.
- Furukawa T, Kawaichi M, Matsunami N, Ryo H, Nishida Y, Honjo T: **The Drosophila RBP-J kappa gene encodes the binding protein for the immunoglobulin J kappa recombination signal sequence.** *J Biol Chem* 1991, **266**:23334-23340.
- Hedges SB: **The origin and evolution of model organisms.** *Nat Rev Genet* 2002, **3**:838-849.
- Ciccarelli FD, Doerks T, von MC, Creevey CJ, Snel B, Bork P: **Toward automatic reconstruction of a highly resolved tree of life.** *Science* 2006, **311**:1283-1287.
- Katinka MD, Duprat S, Cornillot E, Metenier G, Thomarat F, Prensier G, Barbe V, Peyretailade E, Brottier P, Wincker P, Delbac F, El AH, Peyret P, Saurin W, Gouy M, Weissenbach J, Vivares CP: **Genome sequence and gene compaction of the eukaryote parasite Encephalitozoon cuniculi.** *Nature* 2001, **414**:450-453.
- Sluder AE, Maina CV: **Nuclear receptors in nematodes: themes and variations.** *Trends Genet* 2001, **17**:206-213.
- Rosel D, Puta F, Blahuskova A, Smykal P, Folk P: **Molecular characterization of a calmodulin-like dictyostelium protein CalB.** *FEBS Lett* 2000, **473**:323-327.

30. Nakayashiki H, Kadotani N, Mayama S: **Evolution and diversification of RNA silencing proteins in fungi.** *J Mol Evol* 2006, **63**:127-135.
31. **Schizosaccharomyces japonicus Sequencing Project. Broad Institute of MIT and Harvard** [[http://www.broad.mit.edu/annotation/genome/schizosaccharomyces\\_japonicus/](http://www.broad.mit.edu/annotation/genome/schizosaccharomyces_japonicus/)]
32. Wood V, Gwilliam R, Rajandream MA, Lyne M, Lyne R, Stewart A, Sgouros J, Peat N, Hayles J, Baker S, Basham D, Bowman S, Brooks K, Brown D, Brown S, Chillingworth T, Churcher C, Collins M, Connor R, Cronin A, Davis P, Feltwell T, Fraser A, Gentles S, Goble A, Hamlin N, Harris D, Hidalgo J, Hodgson G, Holroyd S, Hornsby T, Howarth S, Huckle EJ, Hunt S, Jagels K, James K, Jones L, Jones M, Leather S, McDonald S, McLean J, Mooney P, Moule S, Mungall K, Murphy L, Niblett D, Odell C, Oliver K, O'Neil S, Pearson D, Quail MA, Rabinowitz E, Rutherford K, Rutter S, Saunders D, Seeger K, Sharp S, Skelton J, Simmonds M, Squares R, Squares S, Stevens K, Taylor K, Taylor RG, Tivey A, Walsh S, Warren T, Whitehead S, Woodward J, Volckaert G, Aert R, Robben J, Grymonprez B, Weltjens I, Vanstreels E, Rieger M, Schafer M, Muller-Auer S, Gabel C, Fuchs M, Dusterhoft A, Fritz C, Holzer E, Moestl D, Hilbert H, Borzym K, Langer I, Beck A, Lehrach H, Reinhardt R, Pohl TM, Eger P, Zimmermann W, Wedler H, Wambutt R, Purnelle B, Goffeau A, Cadieu E, Dreano S, Gloux S, Lelaure V, Mottier S, Galibert F, Aves SJ, Xiang Z, Hunt C, Moore K, Hurst SM, Lucas M, Rochet M, Gaillardin C, Tallada VA, Garzon A, Thode G, Daga RR, Cruzado L, Jimenez J, Sanchez M, del Rey F, Benito J, Dominguez A, Revuelta JL, Moreno S, Armstrong J, Forsburg SL, Cerutti L, Lowe T, McCombie WR, Paulsen I, Potashkin J, Shpakovski GV, Ussery D, Barrell BG, Nurse P, Cerrutti L: **The genome sequence of Schizosaccharomyces pombe.** *Nature* 2002, **415**:871-880.
33. **Cryptococcus neoformans Sequencing Project. Broad Institute of MIT and Harvard** [[http://www.broad.mit.edu/annotation/genome/cryptococcus\\_neoformans/](http://www.broad.mit.edu/annotation/genome/cryptococcus_neoformans/)]
34. Kuramae EE, Robert V, Snel B, Boekhout T: **Conflicting phylogenetic position of Schizosaccharomyces pombe.** *Genomics* 2006, **88**:387-393.
35. Coeytaux K, Poupon A: **Prediction of unfolded segments in a protein sequence based on amino acid composition.** *Bioinformatics* 2005, **21**:1891-1900.
36. Gregan J, Rabitsch PK, Sakem B, Csutak O, Latypov V, Lehmann E, Kohli J, Nasmyth K: **Novel genes required for meiotic chromosome segregation are identified by a high-throughput knock-out screen in fission yeast.** *Curr Biol* 2005, **15**:1663-1669.
37. Decottignies A, Sanchez-Perez I, Nurse P: **Schizosaccharomyces pombe essential genes: a pilot study.** *Genome Res* 2003, **13**:399-406.
38. Chen D, Toone WM, Mata J, Lyne R, Burns G, Kivinen K, Brazma A, Jones N, Bahler J: **Global transcriptional responses of fission yeast to environmental stress.** *Mol Biol Cell* 2003, **14**:214-229.
39. Mata J, Lyne R, Burns G, Bahler J: **The transcriptional program of meiosis and sporulation in fission yeast.** *Nat Genet* 2002, **32**:143-147.
40. **NCBI Protein** [<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Protein&itool=toolbar>]
41. **UniProt Knowledgebase** [<http://www.expasy.org/tools/blast/>]
42. Burge C, Karlin S: **Prediction of complete gene structures in human genomic DNA.** *J Mol Biol* 1997, **268**:78-94.
43. Milanese L, D'Angelo D, Rogozin IB: **GeneBuilder: interactive in silico prediction of gene structure.** *Bioinformatics* 1999, **15**:612-621.
44. Hua S, Sun Z: **Support vector machine approach for protein subcellular localization prediction.** *Bioinformatics* 2001, **17**:721-728.
45. Yu CS, Chen YC, Lu CH, Hwang JK: **Prediction of protein subcellular localization.** *Proteins* 2006, **64**:643-651.
46. Nakai K, Horton P: **PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization.** *Trends Biochem Sci* 1999, **24**:34-36.
47. Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD: **Multiple sequence alignment with the Clustal series of programs.** *Nucleic Acids Res* 2003, **31**:3497-3500.
48. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG: **The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools.** *Nucleic Acids Res* 1997, **25**:4876-4882.
49. Kumar S, Tamura K, Nei M: **MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment.** *Brief Bioinform* 2004, **5**:150-163.
50. Gouet P, Courcelle E, Stuart DI, Metz F: **ESPrpt: analysis of multiple sequence alignments in PostScript.** *Bioinformatics* 1999, **15**:305-308.
51. **Rhizopus oryzae Sequencing Project. Broad Institute of MIT and Harvard** [[http://www.broad.mit.edu/annotation/genome/rhizopus\\_oryzae/](http://www.broad.mit.edu/annotation/genome/rhizopus_oryzae/)]
52. **Coprinus cinereus Sequencing Project. Broad Institute of MIT and Harvard** [[http://www.broad.mit.edu/annotation/genome/coprinus\\_cinereus/](http://www.broad.mit.edu/annotation/genome/coprinus_cinereus/)]
53. **JGI Phanerochaete chrysosporium v2.0** [[http://genome.jgi-psf.org/Phchr1/Phchr1\\_home.html](http://genome.jgi-psf.org/Phchr1/Phchr1_home.html)]

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)



## New and corrected fungal CSL gene prediction models

*Schizosaccharomyces japonicus*

### SjCSL1

[http://www.broad.mit.edu/annotation/genome/schizosaccharomyces\\_japonicus](http://www.broad.mit.edu/annotation/genome/schizosaccharomyces_japonicus)

GenScan prediction

Supercontig 5 (bases 726033-727712, complementary strand)

```
ATGGGCGCTTCATTCTCTACAAGATATGATAGAGCTTCACCTGAGTCGTCTATGCAGCTGTTTTTCGCTGTATGCTGCCGTC AATTCTGC ACTAA
GTTCCCTTAGAAAGAAAAGACCGACAAACCAATGCGATACAAGCATTCCCTGAAAGCGTAAGTGAAGCTTCTACGATTCAAGAAATGCAGCGGCA
GCAACAGCAGGTGCCCTCAAAGCGTTCCCTCTGCAGTCTTTTTCTCACCCTCTCAATACCGGAAAAATCCAAAGAAAAGAAAAGCAAGTTTTTCA
AATCACCAGGACCTTTGATTGACAAAATTCGAGAAAACAATTTTCGATCACTTAAATCTACTTTTAAAAAAGGGCAGTGCTACTTCCAATGCC
TCTGTAATCGTATACGGGATCCACTGAATCAGTACAACCAAGCCAATAAGCTGGTCACCATTCTCTATCAGCATGCCTCTGTTGCACAAAAGTC
CTATGGTGCAGAAAACGTTACCTCTGTCCGCTCCGCTCATTCTATAACTGGCAACACTGGATCAATATTGGGTCACCTCTTTTCAGTGCAG
CTAGCCATCACAATGAGCAGGGTCAACACTCGAATCAGGTGTCTGAAAACCTTTGCTGAACGTC AATCAGTTCGCTTTTCGTTCTTTACATTT
CCTCTTCCATTGCTGCCAAGGCAAAGAGTTTTAATCTCAACATAGATGTCATAAGTCTTTCAGACGATGTTTTGGCTCAAATGGTTACAAAACC
CATTAACATTATATCTAAACCATCAAAAAAGGGAAGCAAAAACCGTGTTC AATCTACTCTCATGTCTGGCGCCATAATCAGTTTATACAAC
AGAATCAATTCTCAGACTGTTCTGTACAAAATACATGTCTGTTGCCAATGGACGTTTCTGCTTGCAGGAAACGAATCTTGGACGCCCTTGCAATTC
GTGTGTACGGTGCAGTGAAGTTTCAATTCATCAACCCATCATGTATGGTTTACAGGTTGTTCTCTACAACGAGCGAAGCGGATTTGTGAGCGA
TCCGCTGTACATTCTCGGTTGACAAGGATCAAATTCACCTGATGATGGTTATCTTGGGCAATGCATCGTATCGTTCTTCTGCGGCCAA
TGCAACACAGAACCCCAATGTGTTGAAAACCATACGCTAATCAACAGGCATGGATTTTGTACAACCACTTTTCTCGGGGCAACAACAGCCCAAG
AAAAAGTACTACCGGCGACATGGACTTTCCCTCGAATCGAACCAGGCTACTTCAATTGACGAGAGTGGTTCAATCCGGGTCAGTGTGACGAT
TTGCTGGACCGTTATTGGTATTTCCCACTATGAGTTTTTCGATGATGAATGCTCTTAACGCTCGTATCACTCAACCTATTACACCGTTTCCAGTA
GTCGAAACATCCCCAAGTACATTTGGGAAACACATTCATCGAATTTGCTGGTGGTGGTGTGTACCCTAAAACACAAGTTTGGTGGGTCTT
ACGGGCGCTTGGCATATACCAGGCAAGAAAACAGAAAAGCAGACGAGAATCTGTCTATGTCTCTCTTCCATACATTGAAGTGAATACTACAAA
CCCGTGGACAACCTTTCCTTCTTCACTCGGCCAAACCGACTTATTTACGTGGGCAAGTGTGACGTTGTTGCTCCGCTAG
```

*Schizosaccharomyces japonicus*

### SjCSL2

[http://www.broad.mit.edu/annotation/genome/schizosaccharomyces\\_japonicus](http://www.broad.mit.edu/annotation/genome/schizosaccharomyces_japonicus)

GenScan prediction

Supercontig 4 (bases 1104530-1107169)

```
ATGTCCGGCTTAAAAATCATTCTCCGTACGGAATTGCTCCAAATACAGTAAATTCAGAAACGCAAAATGTTGCCCGGCCGCAAGATGAAAATT
CAAACCTGGAACATGGAAGTTCGCTTATGACCGACTGGAACAAGAAGCCCTTAATTTGCGCGGTTGAATACTGTGTGCTCTCTACATATATGAA
CCCGCAGGCACCTTGTCTACATACGACGATTCCAATATTTACTGAGCAGCAACGCTACTCCTCAGGCTTTGTGCTAGCAGTGCAGCCCTCA
CTCTATCCCTCACAATCCCTCGCGTGGGTTACTACGTGGATAGTGCCGATCCCTCCCTTATGCAGACGACACTGGAATTCGAACGCTGGCA
ATGAAGC CAAGCTTGGGATGGCTTCAGAGTTGCCCTCAATGGGTATGAACAATGATGCTAATTCGCTGTACGACAACAACAACAACAACA
ATCGCATCAACAGCAGCTACAGGCACTAACAACCAATCTGCCTTTAACCATATGGGTTGGGACGTAGCTGGTCACTTTCTGCCTGGTGAATCT
GGTCTTTTCCCGCCGTAACAGACGATTTTCGAGTCAATTCGTTCCCTTGAACATCCCAACACGATTAACAACCTTACCACGACTCCCTA
ATGCAGGTGCCTATCCTGAGGTTTATCCGGCCGACCCGACCTCGTCTGATTTCCGGCCATTCACTGATCAGCAAGGCAATTTCCGGCTCATCGTT
CGGTGACGTCGTGCCCGCGGAAACCCAAACATCTCGAATCACCAGCAGTGTAGAGCTGTGACGACACTCACCAGCCATCTCTACGGTGTCT
GGTGTGCTCCTCGTCTGCAACAACCTTACCAGCAACCCAGACAGCAGTTCATGTTTCTGACTACGGCCAGTTTACGTTCCAGCTGGAACCA
ACAACCCAAATGCTAATAATGGCTTTCGAGTTCCGTTTGGTGGGTTCCAGCCCGGCGAGGTTGAGTCAAACATCTTTCTATGACGGGCCAGA
GTTTCCAGCACCCGCTGAATCTTCACATGCTTCCCTTTTGGAGCCCTCTACGTCTCACTCAGACTCCATGGGCTTAAACGCCCCGACAGAAAC
AATCTTTTCAACGGGTTAAAAATGTTCTGCGACACCCAGAGCTGTTGTGCGCTATGAAAGTATTCATGCCCTCCCTGGGTCAGAAAGTGTACG
GTAAGGAGCGTCGCTATATATGCCCGCTGCGGTGCTTACTTGTCTCGGCAGCAGCTGGTTTCGATGCCACTCGACAAAATAAATATCATGTC
GAACCGGCTGATGACCTGACAACCTTAAACCGTCAGAGACGCCAACCTTTTACACAGCTCTTCCGACAGCGCCACAACCTTCTTTCGCTG
GGTCAGGTGAAACTCGACGATCCATTGGAGCAACAACAATCACCATCTCCCATATGGGCGAATACGGTGTGAAAACGCTTTATTTACTCCGTA
AGGGGATCACAATACGATGCGCGCTCAACTACCCTGACAGTGCATGTTCCGACCCCAAAAAGCGGATAACTATGGATAAGCTTCGTAATTGG
CATCATATCGAAGCCGAGCCAAAAAATGATGATGAAAGTTTCGGATTTGAATATCTGTACAGCGGATTTGTGTGAGTTTATTCATCGGTTT
CGCTCTCAACAACCTCCACGGTATTGTGCAAAAATGTCCTTAATGACGTCGTTTACAAGCGAACAACAACACTTACAATCCATGAGGAGT
TTACGCCGCCACAGACTTAGACTTAATGGAAGCGTCAACATGTCGATTAATTAACAACAACAACCGTTTTGGGAGCCATTCAATATATATTTCTGT
GGAAGAGCTTGAACAACAATCAAGTTATGATCGGCGTAACACAGTTATTTGCTCCAATATGCGCATTATAATCCAAGTCAAATACGGGTGTG
CGTTCACCTCCGCTCATTATTCGGAAGTACGAGAACCCTAAAGCACTTGTGGTGGAGGATGACCAGTTGGGCGACTCGATCAATTTGTTTCTC
GATTAGCTTCCAGTGTCTCGTACAAAGTTATTTTGAATTTGGACGAATTTAGCAACGGTGGATGAGATCCGATTTCTATCGGCTGACCCGGCTC
TGGAGAGGGTGTGGAATGGGGAATATGTGAATCTTCCGTTGCTGCAAGTGTGGTCTATCATTAACACTCAGTCCGTTTCAACAATGTTTTTT
GACGACTGTTCCGGCAATGATGGCATGCTCTCCATCCCTCAGACCAATCATCAAGTTCAATTCGCATGGATGAGAGCAATATGTTGCATGTGT
ATGGCGTGAACCTTACCAGCCGACGCCAAGTCTGGATTGGTGAATAATCATGTCAGACGATTTCCGGTGCAGATGTCGAAATCGACGAGCCCA
ACTGCTTCCGTTGACTTATTTCTACTTCCACAGTACCTCCGCGACTGCATGCATATCTTGCAGATTTGGCTGATATCATTTGTGACGCCCGGCG
CTAAGCGTAACTACATCTCCAGAGTTACCAGATACTCGTTTCCGAACAACAACAATCATTTTTCCACAGTGGTTTCACTGGCCAAATCCATCCACATA
CCCTGTGA
```



*Phanerochaete chrysosporium*

**PcCSL2**

<http://genome.jgi-psf.org/Phchr1/Phchr1.home.html>

Manually corrected (predicted introns are shown in red lowercase)

GenScan prediction from Scaffold 6, Contig 19 (bases 50978-54385, complementary strand) with manual definition of the 4<sup>th</sup> intron according to a protein sequence alignment.

```
ATGGCGAAACATCTGCAACCTCCTCTGGATATTCGTCTGGCTGCCCTGTGCGCAAACGCCGGCGCTGAACGCCGATCAATGGgtgagtgct
tatagaatgattgctgagcacactcctcctcctcctcctcctccccgccccgccccgctcgagAGCACGGGCTGACGACGGCTGTGTCC
CCCGCAGACATCCTCAAGCCTCGTTCGGCCCCCGCAGCGCACGGGCACAGGAGCATGTCCATCGACTTCACCCCCACCGCCCCACCAGAGCA
TGTTCCGCCAGCCGGACATCGACCAGCCCTTCAGTCCCTTCAGGGGGGGACCTCGCCCACTCTGGCTCCACATGTCTGTTCGGCGGCCACCC
GCTCGACAGGAACGACCCCTCGACATCGACACCCATGGCAGCTACGACATCTTCTCCTCCTCCAGCGGCTCGCTCGCGTCCACAGAGGTACCCG
ACCAACCGGTCTCCTCCTCCTCCTCGGCCCCAACTACTCCTCGGGCTCGATCCCATCTACCAACAGTCTGCTCTCTGACAAATTTCTCT
CCTTCCACTCGTCAACTCGAACCCCTACGACCTCATCGGAGCTCTCTCCTCCTACAGCAGCGGCAAGCCATCGCCATCACCCCCAGCGA
TGTCAGCGCACTCCCGCACTCCTCGGGCTTCCCGTTCTCAACGGCCAGTCCAAGGACTTCCCGCCGACCACAGCTACCACGAGCCCATGTCTC
GACCGCCGCATCAGCAACGCGTTCGGGCTACTCGAACGACTTCAACGACGAGTTTCGGTTCATGGGCGTCAACCACGGTCTCGGGTCAACGGTT
TCCCTCCTCAGGACTTCCACCGTTCCCGGACCGCCTGGGACGCGTCCAGAATGAATCGCGTACCCCAACTCGACTGTGCCACCCCTCACAGC
CAGTCTCACAGTCCAGCTCATCAGAGGTTCAGAGGTCGCGCGCGCAGGACCCATCCTCCGCTGTTTTCAGCAGCATGAATTT
ATGGGCGCGAGCCGACCGTCTACGACCCGCGCTGCGTCTGCCCGCAGCGTCAACGACGACATGGCTCGTCTGCGCCTGCAGGGTTCTGGTG
ACCTGCAGACCTTTATTTCgtacgctcttctctgagcttagcagctttagaaccaagcctcatgatcctctccatttctagGCCATATCTGGATC
AGTACATTAGGACCCCAACCGTCTGGCGTTCGGCGAGCGCAGGTTATCGTATGTCTGTCFAAGGTCGCCCAGAAAGTCAATGGCACAGAAAA
GCGgtacgtaaacccaattttagcagaccgctgaaccagattgtgctcctactcctcagGTTCTCTGCCCTCCCGACCTTTCAGCgtac
ATTGGTAACTCGTGGTGGTCTGTATGTCCATCGACGTGGCGAGGAGCCGAAACTTTGCCCGCAAGAGTTGTCTCCATCTCCGGGAAACCTG
CACCGCAGGAGGTTTCGATCGAGTGGACGAGCGCCACGGGCAAAGCGTTTCGATGTGTCTGACCCGCCAACGGGTACCCTTACATTCGGGCGCTG
TGTCGGCAAGCAACTTTCATCTCCGACGTTGACGAGAAGAAGAAGGTCGAGGCACTGGTCAAGATCATGGCGCGCTCTGCCGACGACGAG
CCGGAACGTGTTATCGGCACGTTCCCCAGTCCGATCAAGGTCATCAGCAAACCGCAGCAAGAAGAGGCAAAGCGCAAGAACCTTTCAGCgtac
agtgctccatgcccacttctttttgagttcaacgtaataatgcccgttctgtggcaacagTCTGCATCAACCAGGTTCCACCATCTCGTGTTC
CGTCTCCGCTCCAAACCGTCTCCACAAAGTATTTGTGTGTCTGGGTCGGGCTCTTCTGTTCAAAGGTTCCGATGGGCTCCTCTCATGGG
CTGGATCAGCGCGCCGTACGACGACACCGTCTGTTATAGCGCGGACTGCTAGCTGGgtgagctctaacaccgctatttagcgcacagggcca
acttctaagcttgcacatttagATCCATTATTATGTACATCTGTCGACGTGAACAAGCCTACGGGAGGCATCGATACGCGCCACCCACCGCC
TCAGCCGAGTTCCTCCTCCGCTAACGCCATTCCCTTACCAACAACGGATCCAGATACCCATTACTACAATCAGACCGTCTGCTCCTT
CAGTGTCTGACTTCGGGCGTGGTTCAGCCCTGTCTCATCATCCGCAAGTTCGATCACCAGACCAGGTTAGTCGGCGGAGGTTCTCAGGAGGGTG
CCAAGGGTATCGCCGATCATTATTGTTGCCCTGGTGGGTTCTGCGGGGACCCGGTCTCGCAACTGCACAAGATCGCGTTTGGGTTCTACGACC
TAACAAGGGAGCCCGGAGCCCGGCGCACCCTGGGGTTAGCGCGCGTTCCTGTCTGATGGGCGAGAAGGTCAACACTACCGCCCGTGGAG
GGCCGCCAGTGGAACGCGAACGTTTACAAATACCGCCGAGGCGGACTCCCTATTGCTGCGCCCGCTCGCCGATCGTGGCACGCGGACTTCGG
CGAACGGCCACGGGACTTACTCGGTGGCAATGGCGGCTTCGGCAGCGCGGCAACTCCCTATCCCGACCGAGTTCCCTCAAGCGATGGCGG
CAAGGTCAAGAAGGGCAAGCGCGGCTCGTTCGAGTGGCGGCGCTCAGCAAGCCTGCCTCGCAGAAGGGTTCGCGACGGCCAACTCGCGGGGA
TCCCGCAGCGGCTCGATCTGTCGAGCGGCTTCGACGCGACTCGTCCGCTCCTCTGGAGCTCTTCGGCAGGTGGACATTGGCGAGA
CAAGCGTCTGGACGATCGTTGGTTCGgtcagataacttccccgttctgttcttaataatcattggtgacctcgttcgagACCAAGTCCGATA
CAACTTCTACGTCGCCCGTGTCTTTCGACAACCAGAGCGCGCCCGCAGACAGGTCGTTCCCATTCCTCAAACCGGTGACCCCTTCCCC
AACGTCGTCGAAGTACCTCCCGCCGATCGCGCTGCGGAGGCTCCCAAGCGTGCCCGAGTTCGCTGCGATGATGGCGAAGCCAAACCTTCAG
CGTCAAAAATGCTCAGGTCGATCGGCGAGAATTCGCAAGACGGATCTGTCACCGTGTCTTTGGCTCAGACCCGTCGCGTACGTTGAGT
CGGGTGACAGAGGTGCTCGGCTCCCGCTCCCGAGTGCAAACAACATGAAGCGCAGGCCCATCATCTCTGTTCCGCGACGACGGCGTCTGTTT
CCATCCAACACTACTACCTTGA
```

*Rhizopus oryzae*

**RO3G\_11583.1**

[http://www.broad.mit.edu/annotation/genome/rhizopus\\_oryzae/](http://www.broad.mit.edu/annotation/genome/rhizopus_oryzae/)

GenScan prediction (predicted introns are shown in red lowercase)

Corrected gene model RO3G\_11583.1. Two additional 5' exons were added to the annotated model based on the sequence alignment with other fungal CSL genes. Moreover, a possible frameshift was corrected by removing A1011 from a stretch of five adenines, which restored a conserved block of amino acids.

```
ATGTTTCAAACTGAAACAAGAAAAAGAAAAACAAGATGAAATGAATACAACCTTTGCCAACCAATTGGACAGATTTTATTTATTCTACACCTTCAT
CACCTTCGATAGATAACTTATTGATCAACATCTTATTATTATTGATTCAAGCAGTGGAAACAACAGTCTGTCATTCTGTTGCTGTTGGTGA
ATTAGATTATCATTCGTTGATTTAAATAGTTTATTAGAAGAAAGACCTTTACATAAAAGTGCATATTAAGAGAAGATGATTTGACTGCC
AATTTGTTTTCTTCTTATCTATTGATTTGGTTGATACAAGACCAAGAGAGTTATCGATGGATAGCAGTATTATATCTGACCTTTCTTTGAATG
ATTTATCAAAATAATAATCCTGACTGTGACAAGTTAAATACTTCTTTGGAAACTATTACTCTTCTGCTACTCTGACCAAGCAAAATCAATCAAT
GGCTGATTGGTTATTAGAAAATCCTCAAAGAGACCAAGAGCTTACTGATTACCCCTTGGTTCTTCTCTGATTCTTCTTCCACACCT
ATTACCCCATGAGCAGGTCCTCTGGGATTTGAACCTATTCAAGAAGAATGGGATCTTCAACCCCTTAATTCAAAATTTTACTTCAAAAAC
AATCAAGAGAAGACTATATCTGTTGAAAGAACAATCATGATTTACTAGTAAAGTAGCTCAAAGAGCTATGGTACCAGAAAAGgtaacac
cgaanaattagattagataaatgactctttgcttattcattcacagATTTTATGCCCTCCGCTGCAACCATGAAAGGTACCAATTTGGTGGAC
CTCGGACAAAATTGACCGATAAGAAGACGCTTCTTTATTTTCAATCTCCGTCAAACGCTTCTCAATCACCTAAACTCACGATACATATCTCGGGT
GAAACGATTCAGCAGACGGGTGTGATCGAGTGGCAGACATCTTCTGGTAAACATATCGATAAACAATGCACAAAAGGTATTCGGAAGATGATTT
CAAAGCAACTTTACATTAATGACGCGAGATGAAAAGAGAAAGCGTGTGGAGGCTTTGGCAAAGATCCAATTAGGCAATGGTTCCAATTTGGGTAC
```

TTTTCAAGTAAGGGTATCAAGGTGATCAGTAAACCTTCAAAAAAGAGACAGAGTGCAAAGAATATGGAATgtgagtgacttttgtttctattg  
tatgatgacttatttgttttagTATGTATTTCATCATGGTACAACCATCTCTTTATTCAATCGAATTCGGTCAAAAACCGTATCGACTAAATATC  
TTGGTGTATCAACCACCACACCTCAACCAGGATTCAAAATGGTACTTGCCTTGTCTCGAACCGGTGTTTGGGATCCATTTCGTATCTGGATCGT  
TGATACCTCTTGTTCACCCAACACTGCCAATCGACCTAAACATAATCCACTCAATCTTAATTATCCTCCTCCACCTGCCATCGCCTTACAGACC  
TCTTCTACCCCTAGCCATTCAATACAATCAACCCTCGTCTTCCAGTGTGTCCACTGGTCTGTGTGAGCCCTGTGATGATCATCCGTAAGTCCG  
ACAAAACAGAGCTTGGTCTTGGGTGGTAATCGTGTGGATAATCCGATCGGATCCCTGGGTGGTGAATGCAGTGTGAAACCTTGGGTGATCCTGT  
CTCTCAATTACACAAGGTGGCCCTTCAAAATCGTTCAGGATCCTTCATTTTCATCAAGGCAACCTGAAACAAAACACAGCTGGCCACTGGAAGATT  
CCTCAATCAAGTCATCCAGTGACCTATTTGGCCTGTCTGAATGATGTCTGCGGATGCACAAGACAACCTCCACCCGTACCTTGTCCCTCAGT  
GTCAAGAGAACACCTTTGGCGAAATCGCACAGGATCCTATTGTGCGTGCAGTACGGGTGAAATCAAACCTCGGGGAGTTTAGGAAAGGG  
ATCGTCTGCACCTGACCTACTGGGCTTGAAGGTGCTTGTGGACTGAAGATGTTTCAGATGCAGCTGTCTGGACGATTGTTGGCACGGATTGT  
GCTAGTTATACCTTCTGACACCGGATGAACGTACGATGCCACCGCACCTTTCCTGTCTTCATGAACCTCACCAAGAAAGGCAAGGATCGCT  
TGACACTGACCGGCGAGAATTTGAGTCCGGATATCGAAGTTTGGTTTGGTGTGTAATCGACCGAGACAGAATTTGTATCTCAAGATTCGGT  
ACACTGTAAAATTCCTTTTGTGATGTGCAAAATAGTACAACAATTGAACAAGAAAATGATCACAGAAGGATTCCTTTACTCTTGGTTCGTGTAA  
GGAATGTTTACAAAACGAACCTTTATTATATTTTATGA

*Rhizopus oryzae*

**RO3G\_07636.1**

[http://www.broad.mit.edu/annotation/genome/rhizopus\\_oryzae/](http://www.broad.mit.edu/annotation/genome/rhizopus_oryzae/)

Manually corrected (predicted introns are shown in red lowercase)

Corrected gene model RO3G\_07636.1. A different 5' splice site (GC) was chosen in order to restore an absolutely conserved part of the DNA-binding region. In addition a similar splicing pattern is found in other *R. oryzae* and *P. chrysosporium* genes.

ATGAATTATCCAGAATTATTAACAAATGAAGTTGAATCTATCCCATCCTTGGTCTTCTAATTCATCTTTACCACATGTTTCTTATTCTCTC  
CAAGTTTTTTAGAAACATTTAAATTAGAAGATGAAAAATGATTTAATACAATACATCCTTCTGTATCTATCATCATTACCTCAAAACCAC  
TATTTCTTACCCTGATCATTCCAATCTATTTAATTTACAAGAAAATAAAATTTTCAATCATCAACAAAAGAAATCTCATTCAACATTATCTA  
TCACCAAAACAGGAGAAAAGAACTGACGATTTTGACGAGTAAAGTAGCTCAAAAAGAGTTATGGAAATGAAAAAGgcaagtgggtaaataac  
tcaattctcattaaactcattcttcccatcagATTCTTGTGTCTCCTCTCTACTATCCTTCTGGTACGGGTCACTGGTGGACAGCAAAA  
CAACATCCACCGAATCTGACAATTCAAATCTCTGGTGAATAAAGTTCATCAAGGCACGATCGATTGGTACAAGGATGGCAATCTGTTGGATC  
AACCTTCTGCCGTCTTTTAGCAAAATCAAGGAAGTAATCTGATTGGTAATTTGTGTATCCAAGCAATTGCATGTGAGTGGCGCAGATGAAAAGCG  
AAGGAAAGCGCAAGTTCAAGTCGAGATTCGATCAGGTCAAGGTACACCGATCGGTATCTTTTCATTCAAAACCTATCAAAGTGTAGTAAACCA  
TCCAAGAAGAGACAGAGTGTAAAAATATGGATCTCTGTATTCATCATGGCAGCAGTGTGCTTTAATTCATCGTGTACGAGCTCAGACCATCT  
CCACAAAATATCTCGGTGTTTCTTCTCTTGGACAGTCAACAAAAGGACAGAGGTAATTTGTTGTCACCTCGAACCACTCTTGGGATCCATTCCT  
CATCTACATCGTGCATCTCTCTGTTCTCCAAACACACCTTCTCCAGTGCCATTCACTCATCATCCACCATCGATCACTATCCTCCTCCTCCT  
GCTATCGCCATTCAAAACCATCAGGGCTCACTTGCCTTTCATTAATCAACCTGTGCTATTACAGTGTGTCTCTACCGCCCTGGTCACTGTCGG  
TGTGATCATCCGTCGATGAGAGAAAGTAGCATGGTGTGATGGTGGTAAATCGTGTGAATGATCTTTCTTATCCTACAGGAGGTGAGTGGGGTGA  
TGAAGCCCTGGGTGATCTGTCTCAACTTCAAAAGTCGGATTTCAAATCGTTCAAAGATCCATCCATCGCCCAATACAAGTCTACTTTTT  
CAGGAACAGGACAAGTTCTTTTTACCCCTGTCACCACTGGACCTTGCCTCAGGCCACTTCTGCTATTAATTATCTCGCCTGTATTTCATGATG  
TCGTCCGAATGCATCGAGTGACGGATGAGAGAAAGATCGTGTGATTTACAACAGAAATCGAAGACATCAAGATGGCTGTTCCGTAACCGTCCG  
TCTGCTCATCCAGCAACACGACTACAGTCAAAGCTTAATCGCCGTCGAGTGAATCTTTAAATGATGAATTGCTATCAAGACATGTTGGT  
GGCAGCGTGGTCTGTTGCTGATCAACCTTAAATGATGATTGTTGAGTGAAGATGTTTCGGATTCCAGCCTGAGCAGATTGTCGGTACCA  
ATTCGACTAGCTTTGCCCTTCTGGACACCTCCTGATTATTCTAAACCATCTTTGATCTTCAAAGATTTCCCTATGTTGACTCTATCCAATCATT  
ATCCAGTACAGTCTTGTCTCTCGTAGGTGAACCTTTACATCAGACCTCACCGTTTGGTTTGGTGTGATGTGCTTCCATCCAGACGGAAATCAA  
TCAAGTCAACTTCTCCTGTACTGTGCTGAAACGACATGAATTATTGGATAGTTTGGCAACTCAGTTAGACCCCGACACAAGCAGGCATAAAA  
TACCTCTCTTTTGGTTCAAGAAGATGGCATCATTTATAAATCTACTACTGTTTATTTCATTTTAA

*Rhizopus oryzae*

**RO3G\_14587.1**

[http://www.broad.mit.edu/annotation/genome/rhizopus\\_oryzae/](http://www.broad.mit.edu/annotation/genome/rhizopus_oryzae/)

Manually corrected (predicted introns are shown in red lowercase)

Corrected gene model RO3G\_14587.1. A different 3' splice site was chosen for the 2<sup>nd</sup> intron in order to restore a highly conserved part of the DNA-binding region. In addition a similar splicing pattern is found in other *R. oryzae* and *P. chrysosporium* genes.

ATGACTGGTATCCCTCCTCAACATGATATCATAAAGCACAGTATCATATTTACAATAATGACACTAAGCCTCAACCTCCTTCTCCTGAAACAC  
CCACAAGCAGCAGCAGCAGTCAAAAACGTAAACAAGATTTTCAATTCCTCCTGATCACCCTTTCCTTTCTCAATACAATATGCGCAACCTGA  
AACACCTCTGGCCTTACAACACGAAAAATTCATTCAATCTTTACATCCAGATGGAAGCGTGGGAGAAAATGAAGAAGTGTGTCGTCACCTTT  
GATCAATCCAATCCATTCCAAGGCTAACTCAACCGATTGACCTGGATGACCTCCTTACGAAAGACAAGCCTTTCAAACCTGGGACCGTTCTT  
CTTCATCTCCTATTCAATCTCTACTCGTTATTTCCAGGAACACCAGGTTTTTTTACGCCAGGATTTTAGAGTTCGCTTCAAGAATCCAGT  
CTATAGCCCTGGTCTGTTGCTGATTAAGTTTCAATGATTTCAATGATTTGAGTGAAGATGTTTTCAGTCAAGCTTTTAGTCAAGCTTGAAGAACAAG  
aaattcaataaacaacaaactcatttcttaccagTCACCCGAAAAGAAATTTGGTATCACAAGTGGAGAAAGTGTACGCTTTTATTCCCTTC  
TGATCCTGCTTCCATAGTAAGACCCAATCAAGCGCACGACAGCCCATCCGTCGATCCTCCAGTACACGACGGCCAGCAGCCACACCTTCTT  
GTGCATCTCTGCCACTCAAGATCAAACCTTCACTCAGACTTACCTTGCACACGCCATCACCACCCGCGGACTCAGTGGGTGAGAAGA  
CGGTGATCGTGTGACGAGTAAAGTGGCCAAAAGAGTTATGGCACAGAGAAAAGgtaagcgatcggtgacaagagcccaaggctgacagc  
ccaagGTTCTCTGTCTCCTCCCGACAGCCATCTGGTGGCAGCAGCTGGTGGACGACGAAAGAAAAGATCCAAGACAAGGAGGAAACGTTA



## New and corrected fungal CSL protein prediction models

*Schizosaccharomyces japonicus*

### SjCSL1

[http://www.broad.mit.edu/annotation/genome/schizosaccharomyces\\_japonicus](http://www.broad.mit.edu/annotation/genome/schizosaccharomyces_japonicus)

MGLHSSTRYDRASPESMQLFSLYAAVNSALS SLEKDRQTNAIQAFPEVSEASTDSELQRQQQVVPQSVPPAVLFSPPSSIPENPKKRKASF  
NHQDPLIAQIRETIFDHLNLLLKKSATSALCNIRIRDP LNQYNQANKLV TISYQHASVAQKSYGAEKRYLCPPPLISITGNTG SILGHSFVSQ  
LAITNEQGQHSNQVSENFAERQSVAFRSLHSSIAAKAKSFNLDIVISSDDVLAQMVTKPINISKPSKKGSKNRVSNSTLMSGAIISLYN  
RINSQTVR TKYMSVANGRFCLRNESWTP LQIRVYGASESSFHQPIMYGSQVVLYNERSGIVSDPLYIHRVDKDIAPDDGYLGQMHRIVLRRGQ  
CNNENPV LKPYANQQGMDFVQPLFLGATTAQEKSTTGDMDFVPELEPATSIDESGSIRVSDAVCWTVIGISHYEF SMMNALNARITQPITPFPV  
VETSPKYIGETHSLELLVSGCTPKTQVWLGPYGLAYTRQETEKADENLLIVSLPIYEVNTTNPWTTLPLLFTRPNGLIYVGKCDVVVR\*

*Schizosaccharomyces japonicus*

### SjCSL2

[http://www.broad.mit.edu/annotation/genome/schizosaccharomyces\\_japonicus](http://www.broad.mit.edu/annotation/genome/schizosaccharomyces_japonicus)

MSGLKNHSPYGIAPNTVNSSETQNVARPDENSNNMEVAYDRLETRSP LIARLNTVCPPTMNPQAPCSTYDDSNITEQQRYSSGFVASSADPS  
LYPSQIPRVGYVYVDSADPSYADDTGIRTPGNEAKLGMASELPQMGMNDANSVLRQQQQQQSHQQQLQALTNQSAFNHMGWDVAGHFLPGDS  
GRFPAVTDAFSSQFVPLQHPQHDTPNYHGLPNAGAYPEVYPADPTSSYSGPFTDQQGNFGSSFGDVVPGGNPTSSNSPAVLELSGHSPTHLIGA  
GDAPRLQQLHQPRQQFHVSDYGFHVPA GTNNPNANNGFRVFPFGSSPGPELSQTSFYDGPFPAPAES SHASLLRPTSHSDSMGLNAPDRN  
NLFQRVKNVLRHPPELLCAMKVMP SLGQKSYGKERRYICPPAVVYLLGSSWFRCLDKINI IANAADDPDNLK TSETPTFYTSSSDSATNLLSL  
GQVKLDDPLEQQQSPSP IWANTVLKTL YSGKGDHNTYGRSTTLQVHVRTPKRITMDKLRIGIISKPSQKMMMKVSDLNICHGDCVSLFNRF  
RSHNNLPRYLCTNVLNDVVTKRTEHLQFHEEFTPATDLMEASTCR LITNTVWEPFNISVEELENKSSYDRRNTVICSNMRIIQSQITGV  
RSPPLIRKYENRKALVVEDDQLGDSINCLSR LAFQCPRTKFLNLDEF SNGEIRFLSADPAPGEGDNGEYVNL PWSAVWSIITQSVRTMFF  
DDCSGNDGMLSIPSAPI IKFIRMEDESNMLHVYGVNFTADAQVWIGENPCQTYSVT DVEIDEPTLLRGLISTSHVPPRLHAYLADLAD IICQPPA  
LSVTTSPPELPILVQSQNIIFHSGFTWPIHPHTL\*

*Phanerochaete chrysosporium*

### PcCSL2

<http://genome.jgi-psf.org/Phchr1/Phchr1.home.html>

MAQTSATSSGYSSWLPLSANAGAERRSM EHGHTTAVSPADILKPPSAPRSARHRSM SIDIPTPTGPHQSMFAQPDIDQPFSS LQGGDLAHS GSHM  
SVGGHPLDRNDPLDIDTHGSYDIFSSSSGSLASQR YRTNASSSSSLGPNYSLGVDP IYQQSSFDNIS SFHSSNSNPYDLIGLSSSYSSGKPS  
PITPSDVSALPHSSGFPSNGQSKDFP PHSYHEPMLDRRISNAGSYNDFNDFEFGSMGVNHGLGLNGFP PPSGLPFPFDR LGRVQNESRYPNST  
VPPLTATSHLTQSHSPELIRGVAPQATHLPSFDDMFMGASPTVYDPLRLPASVNDDMARLRLQGGSDLQTFIRPYLDQYIRTPNRLAFGERT  
VIVMSSKVAQKSYGTEKRF LCPPTAIMIGNSWSDVHRRGEEPKLCPRVVVSISGEPAPQEGSIEWTSATGKAFDVSDPPTGTTYIGRCVVK  
QLFISDVDEK KKKVEALVKIMAPSADDEPERVIGTFPSRP IKVISKPSKKRQS AKNLELCINHGSTISL FHRLRSQT VSTKYLCVSGSGSSFKG  
SDGAPLMGLDQRARTQPSF IARTASWDFIMYIVDVNKP TGGIDTTPPPPPQPEFPSPPPNAIPFTNNGSQIPIIYNNQTVV LQCLTSGVVSPV  
LIIRKVDHQTTVVGGLQEGAKGIADHYCCPGEVCGDPVSQLHKIAFEVYDPNKGAPEPGTPGVSGAF LSCMGEKVNTYR PVEGRQWNANVTNT  
AEADSPIAAPGSP IAGTPSANGHGDYFGGNGSGSAPNSPI PTEFPSSDGGKVKKGRGSSSAGGLSKPASQKRRRPT SAGSAGSISSSRR  
GSSSDSSASSGALWQVDIGETS VWTIVGVDQVRNYFVPPVLFNDQSA PQTGSPFIPSKPVTPFPNVVYLPDRAAEAPKPCPQSRAMMAKPN  
PHASKMLTVYGENFSKTDPVTVFVFGSDPSPYVEVRCTEVLGCLPPE SNMRRPIILVRHDGVVFP SNTYYP\*

*Rhizopus oryzae*

### RO3G\_11583

[http://www.broad.mit.edu/annotation/genome/rhizopus\\_oryzae/](http://www.broad.mit.edu/annotation/genome/rhizopus_oryzae/)

MFTTETRKRKQDEMNTTLP TNWTDFIYSTPSSPSIDNLF DQHSYFSDSSSGTNSRRHSVAVGELDYHSFDLNSLLEERPLHKRAMSLREDDLTA  
NLFSSYLFDLVDRPRELSMDSIISDLSLNDLSNNNDLYKFN TSLTITPSATLTNEINSMADW LLENPQKRPRRSTDSPLGSSSDSSSSPP  
ITPMQVSLGFEP IQEEDWLQPLIQNYLLQKQSR EYIPGERTIMILTSKVAQKSYGTEKRF LCPPTAIMKGTNWWTSKDLTKKTPSLFHSP  
SNALQSPKLTIIHSGETIQQTGVI EWQTSSGNIIDNNAQKVFGRCSKQLYINDADEKRRRVEVLAKIQ LGNGSNLGTFSKGIKVISKPSKRR  
QSAKNMELCIHHGTTISLFNRIRRSQTVSTKYLG VSTTTQPQDSNGTCFVSRTGVWDVFPV IWIIVDTSCSPNTANRPKHNLNPNYPPPPAIALQT  
SSTLAIHYNQPVV LQCVTTGLVSPVMIIRKVDKQSLV LGGNRVDNPIGSLGGECSDET LGDPVSQLHKVAFQIVQDPSF HQGNLQNTAGHWKI  
PQSSHPVTYLACLNDVVMGKHTTSTRHLVPPQCQENTFGEIAQDP IVRRLSTGEIKRGRSLGKGSALDPTGLEGACWTE DVSAAVWTIVGTD C  
ASYTFWTPDERMTPAPFV LHELTKKGDRLTLTGENLSPDIEVWFVDVKSTETEFVQSQDSVHCKIPFDVANSTTIEQENDHRRIP LLLVRGK  
GIVYKTNLYYIL\*

*Rhizopus oryzae*

**RO3G\_07636**

[http://www.broad.mit.edu/annotation/genome/rhizopus\\_oryzae/](http://www.broad.mit.edu/annotation/genome/rhizopus_oryzae/)

MNYPELLTNEVESIPSSWSSNSLPHVSLFSPFLETLEKLEDEDFNTIHPSVIVHHSPTTTISSPDHSNLFNLQENKIYSNHQQKNLIQHYL  
STKQGEKLTILTSTKVAQKSYGNEKRFLCPPPSTILSGTGHWWTAKQHPNLTIQISGEKLSHQGTIDWYKDGNDLDDQPSAVLLANQGSNLIGN  
CVSKQLHVSGADEKRRKAQVQVEIRSGQGTPIGIFHSKPIKVISKPSKRRQSVKNMDLCHHGTTVALFNRVRAQTISTKYLGVSSLDSSQKDR  
GTFCVTRTTSWDFLIYIVDLRSNPNTSPVPFSSHPTIDHYPPPPAIAIQNHQGSALALHYNQPVVLQCVSTGLVSPVLIIRRVEKGSMMVGGN  
RVNDLSYPTGGEWGDEALGDPVSQLHKVGFQIVQDPSIAQYNKSTFQEQDKFFLPPVTHWTLPAQATSAINYLACIHDVVGMHRVTDERKIVSRF  
TTEIEDIKMAVRKRRLSYQQHSTTVKSSNRRRVNSLNDELLSRHVGGDAGRCPPQPLNGDCWTEVDSDSAVWTIVGTNSTSFATWTPPDYKPF  
FDLQDFPYVDSIQSLSSVLSLVEHFTSDLTVWFGDVPISIQTEFKSSQLSCTVPERHELLEDSFATQLDPTSRHKIPLLLVQEDGIIYNLL  
FYSF\*

*Rhizopus oryzae*

**RO3G\_14587**

[http://www.broad.mit.edu/annotation/genome/rhizopus\\_oryzae/](http://www.broad.mit.edu/annotation/genome/rhizopus_oryzae/)

MTGIPPQHDIKHSIIFHNNDTKPQPPSPETPTSSSSSRKRKQDFQFTPDHLSFPQYNYAEPETPLALQHEKFIQSLHPDGSVGENEELMVVNF  
DQSNPFPRLTQPIDLDDLLQQRQAFQTDWASSSSPIQSPTRYSPGTPGFFTPGFLESLEHHPVYDHSLSIDYGSHHFNQYENPLLVKLEEQSPE  
KNLVSQSGESVTLFSPDPASIVRPNQAHDSPIRRSSHTTASSPHRLVHLSPLKIKPFIQTYLAHAITQPAATQLGEKTVIVLTSKVAQKSYG  
TEKRFLCPPPPTAILVGTSWWTTKEKIQDKEETLRIPSLEKDILLAPPKLTVVISGETSTQAGQLEWYTVSGATVGTGQIKPPIKPESTRFRS  
ESRHPADAYSNERQELLAAGKSVSKHLYIHDADEKRRKVECLVKLQLANGLQLGPLASKAIKVISKPSKRRQSIKNMELCIHHGTTVSLFNR  
IRSQTVSTKYLGVSTSKGSLAPFPLAFQHEKNRTSEGTCTFVARTTSWDFVVIWIVDTSASSEEEGETPEDYIGHHVAFRSTPYPPPPPIALKN  
KTGGPVIHYNQHVVLQCLTTGLVSPVMIIRKVDRASTVVGARDDDVSGSGGEGFGEVLGDPVSQLHKIALQIVQDPKMSVMQAPDPRMPTSQ  
PVTYLACLNDMVMGMHKTSEGRSWAGWDDSIQSQEGGIIRKRRVSTDVQPETLMSCMSLSDYPRRRVNSLEDPPAPYLARKSSVSSLSSTSRPH  
LGAFWSEVDVSDAAVWTLVGTDCATYTFWSPFLDDPSTPLSTGPFPAALSHFFSTSNKLDHERFLTMHGENFSRDLQVWFGDVKANHTEYRSRELI  
ICKVPPRHELMVKKVYGDLPILLVRGDGTICKTKCFSL\*

*Coprinus cinereus*

**CC1G\_01706**

[http://www.broad.mit.edu/annotation/genome/coprinus\\_cinereus/](http://www.broad.mit.edu/annotation/genome/coprinus_cinereus/)

MKPLTSTLATGPASPMNVQEDPFNLLTFFPPGIGRRDDWHWLRGEALAEVEVEASIDVEGDDGDSIPAIFVPTEDRHMSEAIKGEDKLGILSF  
GYRDSHNDDDDDDLDDEIVDETGRFEPETWEDLYEAHRRRRMGKSPDINPKRSPTPQVGGAPSPVVAQDPNLSIQSILAASKKDGPNTG  
TSQATWHPPPPATSTSTGKRKLDETESEDTSHKIRRVDRHVSQDPAVVPMPTVICLHAAVAQKSYGSEKRFCLPPPVIHIEGVPVWHLRTQS  
LSMAVSETGERSFEQKATLDNNMVSSFKFLHVTGTAKAKSFQLSLDIAEPPPPSLNPEGSETSANRGRVWATFDSAPVTIISKPSKKTAKTRNI  
SSCILAGPVSFLFNRINSQTVRKYMTIDHGQLCASNIGWSAFNVVRRRDPGSPNTNGPQPVTYGCEIVLSDTQSGISTSPVIRKVDKGRVS  
PDDGGPVSQMQLALQRVNPDGSRHYLSAAGPLPGTPGVVAPPAGMSTQAGTHPLLFQNPRIRDEVRDGIIRVISDEVDYLCWTVGISKFYQY  
TFFDAFGQNNKIPETPIPTFPPTLFTAPVYRAANNITELTVSNFFYAHPKTRMQTPLDVYLGSLGPLHTRVYQTSPPGPLTSISPFVPTVEVP  
PPAGDPNAPANRVYSTGLPHTIIVIVELPPLADVIKAMEDEPVPTASDVSGSKPHQESEGQEGVPPPPPPPPSMAGRSLPLLFIRSSDGVGYHSG  
RTIACEPIHYSLDLAAMAPNGPNDVDPQLAAQAQAAAADGNMQPWTLRVM\*

CSL proteins sequence alignment used for the phylogenetic analyses

The metazoan sequences used (GenBank accession numbers):

Mus musculus	Mm_RBP-Jk	[NP_033061];
Homo sapiens	Hs_RBP-Jk	[AAH64976];
Xenopus laevis	Xl_SuH	[Q91880];
Danio rerio	Dr_RBPSUH	[NP_942579];
Halocynthia roretzi	Hr_RBP-Jk	[BAA20141];
Ciona intestinalis	Ci_SuH	[O76808];
Drosophila melanogaster	Dm_SuH	[P28159];
Caenorhabditis elegans	Ce_LAG-1	[AAB03859];
Homo sapiens	Hs_RBP-L	[NP_055091];
Mus musculus	Mm_RBP-L	[NP_033062];
Danio rerio	Dr_RBP-L	[XP_697075];

Mm_RBP-Jk	-----
Hs_RBP-Jk	-----
Xl_SuH	-----
Dr_RBPSUH	-----
Hr_RBP-Jk	-----
Ci_SuH	-----
Dm_SuH	-----
Ce_LAG-1	-----
Hs_RBP-L	-----
Mm_RBP-L	-----
Dr_RBP-L	-----
PcCSL2	-----
CC1G_03194	-----
UM06280	MFQPTWSQEAHAQPPHQLHPOEGSSLPKDMQLPQQIQYDPSQISHQAPVH
RO3G_07636	-----
RO3G_06481	-----
RO3G_11583	-----
RO3G_14587	-----
CNBD3370	-----
SjCSL2	-----
SPCC1223.13	-----
Pc6518	-----
CC1G_01706	-----
RO3G_06953	-----
RO3G_13784	-----
RO3G_08863	-----
UM05862	-----
SjCSL1	-----
SPCC736.08	-----
CNA01890	-----
Mm_RBP-Jk	-----
Hs_RBP-Jk	-----
Xl_SuH	-----
Dr_RBPSUH	-----
Hr_RBP-Jk	-----
Ci_SuH	-----

Dm\_SuH -----  
 Ce\_LAG-1 -----  
 Hs\_RBP-L -----  
 Mm\_RBP-L -----  
 Dr\_RBP-L -----  
 PcCSL2 -----  
 CC1G\_03194 -----  
 UM06280 QLSEYDASISQLPNAAPPDPSAYPFSIQPQQHRDTHLAHNDPPYSIAQQH  
 RO3G\_07636 -----  
 RO3G\_06481 -----  
 RO3G\_11583 -----  
 RO3G\_14587 -----  
 CNBD3370 -----  
 SjCSL2 -----  
 SPCC1223.13 -----  
 Pc6518 -----  
 CC1G\_01706 -----  
 RO3G\_06953 -----  
 RO3G\_13784 -----  
 RO3G\_08863 -----  
 UM05862 -----  
 SjCSL1 -----  
 SPCC736.08 -----  
 CNA01890 -----

Mm\_RBP-Jk -----  
 Hs\_RBP-Jk -----  
 XI\_SuH -----  
 Dr\_RBPSUH -----  
 Hr\_RBP-Jk -----  
 Ci\_SuH -----  
 Dm\_SuH -----  
 Ce\_LAG-1 -----  
 Hs\_RBP-L -----  
 Mm\_RBP-L -----  
 Dr\_RBP-L -----  
 PcCSL2 -----  
 CC1G\_03194 -----  
 UM06280 NRFDDPMSASADLSAARPDSMQSVASLYLDGQTMSVDNASLDLGTSHHSN  
 RO3G\_07636 -----  
 RO3G\_06481 -----  
 RO3G\_11583 -----  
 RO3G\_14587 -----  
 CNBD3370 -----  
 SjCSL2 -----  
 SPCC1223.13 -----  
 Pc6518 -----  
 CC1G\_01706 -----  
 RO3G\_06953 -----  
 RO3G\_13784 -----  
 RO3G\_08863 -----  
 UM05862 -----  
 SjCSL1 -----  
 SPCC736.08 -----  
 CNA01890 -----

Mm\_RBP-Jk -----

Hs_RBP-Jk	-----	
XI_SuH	-----	
Dr_RBPSUH	-----	
Hr_RBP-Jk	-----	
Ci_SuH	-----	
Dm_SuH	-----	
Ce_LAG-1	-----	
Hs_RBP-L	-----	
Mm_RBP-L	-----	
Dr_RBP-L	-----	
PcCSL2	-----	
CC1G_03194	-----	
UM06280	-----	DYLAQHDDPSRSRPPTGSEQNWSHHTLPSNQAHHDASATGALAGPHRAAR
RO3G_07636	-----	
RO3G_06481	-----	
RO3G_11583	-----	
RO3G_14587	-----	
CNBD3370	-----	
SjCSL2	-----	
SPCC1223.13	-----	
Pc6518	-----	
CC1G_01706	-----	
RO3G_06953	-----	
RO3G_13784	-----	
RO3G_08863	-----	
UM05862	-----	
SjCSL1	-----	
SPCC736.08	-----	
CNA01890	-----	
Mm_RBP-Jk	-----	
Hs_RBP-Jk	-----	
XI_SuH	-----	
Dr_RBPSUH	-----	
Hr_RBP-Jk	-----	
Ci_SuH	-----	
Dm_SuH	-----	
Ce_LAG-1	-----	
Hs_RBP-L	-----	
Mm_RBP-L	-----	
Dr_RBP-L	-----	
PcCSL2	-----	MAQTS
CC1G_03194	-----	MASLDFAS
UM06280	-----	PAALKIDMPSVVSRRGVHDTTASTGTDPVSAEPGSATTPHALQTRFAPLPQ
RO3G_07636	-----	
RO3G_06481	-----	
RO3G_11583	-----	
RO3G_14587	-----	
CNBD3370	-----	
SjCSL2	-----	
SPCC1223.13	-----	
Pc6518	-----	
CC1G_01706	-----	
RO3G_06953	-----	
RO3G_13784	-----	
RO3G_08863	-----	
UM05862	-----	



SjCSL1 -----  
 SPCC736.08 -----  
 CNA01890 -----  
  
 Mm\_RBP-Jk -----  
 Hs\_RBP-Jk -----  
 XI\_SuH -----  
 Dr\_RBPSUH -----  
 Hr\_RBP-Jk -----  
 Ci\_SuH -----  
 Dm\_SuH -----  
 Ce\_LAG-1 -----  
 Hs\_RBP-L -----  
 Mm\_RBP-L -----  
 Dr\_RBP-L -----  
 PcCSL2 ATSSGYSSWLPLSANAGAERRSMEHGLTTAVSPADILKPPSAPRSARHRS  
 CC1G\_03194 APPPAQHSLYPQPSDLGDANFNLPHSETVGVFPPLRTDFPADTKYROS--  
 UM06280 LPPSMYDNQPMLQQHSDSANSTTMHDAPASSGSALMPAFGTESAYFAQAS  
 RO3G\_07636 -----  
 RO3G\_06481 -----  
 RO3G\_11583 -----  
 RO3G\_14587 -----  
 CNBD3370 -----  
 SjCSL2 -----  
 SPCC1223.13 -----M  
 Pc6518 -----  
 CC1G\_01706 -----  
 RO3G\_06953 -----  
 RO3G\_13784 -----  
 RO3G\_08863 -----  
 UM05862 -----  
 SjCSL1 -----  
 SPCC736.08 -----  
 CNA01890 -----  
  
 Mm\_RBP-Jk -----  
 Hs\_RBP-Jk -----  
 XI\_SuH -----  
 Dr\_RBPSUH -----  
 Hr\_RBP-Jk -----  
 Ci\_SuH -----  
 Dm\_SuH -----  
 Ce\_LAG-1 -----  
 Hs\_RBP-L -----  
 Mm\_RBP-L -----  
 Dr\_RBP-L -----  
 PcCSL2 MS---IDFTPTGPHQSMFAQPDIDQPFSSLQGGDLAHS GSHMSVGGHPLD  
 CC1G\_03194 -----  
 UM06280 ALGSPALAHQAPDHILNSPRLHQAESLAGVPGVPERLGPAAHLGVAVSP  
 RO3G\_07636 -----  
 RO3G\_06481 -----  
 RO3G\_11583 -----  
 RO3G\_14587 -----  
 CNBD3370 -----  
 SjCSL2 -----  
 SPCC1223.13 SPNVOKRPSEDIKTOEFYDSTRNIRRVATAIGSINANLESPOLYSLAKS  
 Pc6518 -----

CC1G\_01706 -----  
 RO3G\_06953 -----  
 RO3G\_13784 -----  
 RO3G\_08863 -----  
 UM05862 -----  
 SjCSL1 -----  
 SPCC736.08 -----  
 CNA01890 -----  
  
 Mm\_RBP-Jk -----  
 Hs\_RBP-Jk -----  
 XI\_SuH -----  
 Dr\_RBPSUH -----  
 Hr\_RBP-Jk -----  
 Ci\_SuH -----  
 Dm\_SuH -----  
 Ce\_LAG-1 -----  
 Hs\_RBP-L -----  
 Mm\_RBP-L -----  
 Dr\_RBP-L -----  
 PcCSL2 RNDPLDIDTHGSYDIFSSSSGSLASQRYRTNASS-----  
 CC1G\_03194 -DQLDTPDTHQNYDLFPNQQQHFSSQRYRTNASS-----  
 UM06280 SSPMMSQLSAGHYSPTSTLSGAFSSRPGSRPASRGLKSVASFTSSAGTGT  
 RO3G\_07636 -----  
 RO3G\_06481 -----  
 RO3G\_11583 -----  
 RO3G\_14587 -----  
 CNBD3370 -----  
 SjCSL2 -----MSGLNHSPYGIAPNTVNSETQN  
 SPCC1223.13 TSLQEPVRIYGDSVSPAISSSKAHSTSSVSPYYSEKNESQALNADGTAF  
 Pc6518 -----MSSIQRSRSAPSVASPOQFLASRSDNDGALSAHRPRIVRHKT  
 CC1G\_01706 -----  
 RO3G\_06953 -----  
 RO3G\_13784 -----  
 RO3G\_08863 -----  
 UM05862 -----MEAGESI  
 SjCSL1 -----  
 SPCC736.08 -----  
 CNA01890 -----  
  
 Mm\_RBP-Jk -----  
 Hs\_RBP-Jk -----  
 XI\_SuH -----  
 Dr\_RBPSUH -----  
 Hr\_RBP-Jk -----  
 Ci\_SuH -----  
 Dm\_SuH -----  
 Ce\_LAG-1 -----  
 Hs\_RBP-L -----  
 Mm\_RBP-L -----  
 Dr\_RBP-L -----  
 PcCSL2 -----SSSLGPNYSLGVDP-IYQSSFSDNISSFHSSNS  
 CC1G\_03194 -----SSSLGHPYAMNSDNGLYSHSGFADSVPSFGSSNS  
 UM06280 GTGTASGSEAHATGRPRSSSGGVAPRPISALTRTTSMQSDIASFQAASP  
 RO3G\_07636 -----  
 RO3G\_06481 -----  
 RO3G\_11583 -----

RO3G\_14587 -----MTGIPPQHDIKHSIIFHNNDTKPQPPSP  
 CNBD3370 -----MSYQWNHNYDTSTSHRDQEP  
 SjCSL2 VARQDENSNWNMEVAYDRLETRSPLIARLNTVCPPTYMNPQAPCSTYDD  
 SPCC1223.13 NPSFHSFGLPOEDSODNTQTYSTPYTTMNPSEMHPYPATFENNYSVLP  
 Pc6518 SLALDLDLSTPRARSQYQTPCGERIEDPFLGGFFPSHLAGTESPDTEWH  
 CC1G\_01706 -----MKPLTSTLATGPASPMNVQEDPFNLLTFFPPGIGRRD-----DWH  
 RO3G\_06953 -----  
 RO3G\_13784 -----  
 RO3G\_08863 -----  
 UM05862 DSAGASSAAANPPDOSSSKSGLESGEPQPADTGDNTHQNRRTSGFDTSEDR  
 SjCSL1 -----  
 SPCC736.08 -----  
 CNA01890 -----  
  
 Mm\_RBP-Jk -----  
 Hs\_RBP-Jk -----  
 XI\_SuH -----  
 Dr\_RBPSUH -----  
 Hr\_RBP-Jk -----  
 Ci\_SuH -----  
 Dm\_SuH -----  
 Ce\_LAG-1 -----  
 Hs\_RBP-L -----  
 Mm\_RBP-L -----  
 Dr\_RBP-L -----  
 PcCSL2 N-----  
 CC1G\_03194 N-----  
 UM06280 QPALTQSRSVQNFQGMQHPSPMPGPLSSPRPPAAAGGTEPWSPAHSMAQS  
 RO3G\_07636 -----  
 RO3G\_06481 -----  
 RO3G\_11583 -----  
 RO3G\_14587 ETP-----  
 CNBD3370 VYQDNDKQNVGYAPMQGQMLDLMGIPQPMFFPGIQTEFNNGRAPVSATG  
 SjCSL2 SNITEQQRYS SGFVASSADPSLYPSQIPRVGYVDSADPSYADDTGIRT  
 SPCC1223.13 DHSSQPNAYSFTGSNILPTQSPSLNQMODYQNLQONGSSNTTIPSFSSQH  
 Pc6518 WLR-----DDGLDAPDAAPEEAHDNKEPAYGKAEDEETSVVIRREDKLGV  
 CC1G\_01706 WLRGEALAEVEVEASIDVEGDDGGDSPAIFVPTEDRHMSEAIKGEDKLG I  
 RO3G\_06953 -----  
 RO3G\_13784 -----  
 RO3G\_08863 -----  
 UM05862 PSPDLTIMDGIPSLDTSNLVSLPAATAQDKTGEHEPSLEEQIQAF TAGL  
 SjCSL1 -----MGLHS-----  
 SPCC736.08 -----MGDYFAWDFANISGSNTSGSLNLDNLDNINNGLNQEDGAGG  
 CNA01890 -----  
  
 Mm\_RBP-Jk -----  
 Hs\_RBP-Jk -----  
 XI\_SuH -----  
 Dr\_RBPSUH -----  
 Hr\_RBP-Jk -----  
 Ci\_SuH -----  
 Dm\_SuH -----  
 Ce\_LAG-1 -----MPLAYSTHDNFYESP KTPQPTWD  
 Hs\_RBP-L -----  
 Mm\_RBP-L -----  
 Dr\_RBP-L -----  
 PcCSL2 -----PYDLIGSLSSSYSSGKPSITPSDVSALPHSSG-----FP

CC1G\_03194 -----PYDMIN----TYSSGKVSPLTPSDPVGGLHHGP-----FP  
 UM06280 HFRGPYSGPGFQTPQMQQFSPTLTSPAQSQYSPHTHFGSNSSGMIPFFS  
 RO3G\_07636 -----MNYPELLTN-----EVE  
 RO3G\_06481 -----MNDPELLSNDEADLVYSSD  
 RO3G\_11583 ----MFTTETRKRKQDEMNTTLPTNWTDFIYSTPSSPSIDNLFdqhsys  
 RO3G\_14587 ----TSSSSSRKRKQDFQFTPDHLSFPOQNYAEPETPLALQHEKFIQSL  
 CNBD3370 GYQNNMPPPQYQFRPALALDTTGGYLAPQPQPSSAYSEMPSGLRPDLGSN  
 SjCSL2 PGNEAKLGMASELPQMGMMNDANSVLRQOQQQQQSHOQQLOALTNOQSAFN  
 SPCC1223.13 DLSQGLTHQVPNHDEYAFSYPYELQRKPLIPAHVPVPSFRPTSALKVNMN  
 Pc6518 LSLPYDDHLLTPYSEGLPEDQEALYRXYCELRRRAHTEDTAVDAAGGRGLV  
 CC1G\_01706 LSGYRDSHNDDDDDEDDLDDVEIVDETGRFEPEIETWEDLYEAHRRRRMG  
 RO3G\_06953 -----  
 RO3G\_13784 -----  
 RO3G\_08863 -----  
 UM05862 ANQLRPSSAEPQHTPRQDSIVNSPYDTPTQANGTPATAEDIASFETLIAS  
 SjCSL1 ----STRYDRASP-ESSMQLFSLYAAVNSALSSLERKDRQTNA-----  
 SPCC736.08 RNENSERVGSGSPGSVSMQVLSLFSAVNSALATLEKSEE-----  
 CNA01890 -----

Mm\_RBP-Jk -----  
 Hs\_RBP-Jk -----  
 XI\_SuH -----  
 Dr\_RBPSUH -----  
 Hr\_RBP-Jk -----  
 Ci\_SuH -----  
 Dm\_SuH -----  
 Ce\_LAG-1 HVHAQFPLGEPARNLDKFI VPEAMFQSVSPLAGVAAAPSQIAALQOIQAL  
 Hs\_RBP-L -----  
 Mm\_RBP-L -----  
 Dr\_RBP-L -----  
 PcCSL2 FSNGQSKDFPPHHSYHEPMLDRRISNASGYSNDFNDEFGSMGVNHGLGLN  
 CC1G\_03194 PSVPSKDFSP--STYQDMGDRRLGMNGNGYQNDYQDDYSINGMNNGLAFP  
 UM06280 PSMDTMQHFQASTGHPAESFDIAGFSEORLSSALANSASISGAPPSATTH  
 RO3G\_07636 SIPSSWSSN-----SSLPHVSLFSPSFLETCLKLED----ENDFNTIHP  
 RO3G\_06481 SVPSSWSSFNDYQMLNHSETPTSFSSNLLDTLQEDDFMHLNTRSRTVHP  
 RO3G\_11583 FDSSSGTNSRRHSVAVGELDYHSFDLNSLLEERPLHKRAMSLREDDLTAN  
 RO3G\_14587 HPDGSVGENEELMVVNFQSNPFPRLTQPIDLDDLLOQRQAFQTDWASSS  
 CNBD3370 TDSDLVMPNSYPSAGVQYPPYNSDFMTSLSVNTGMVNPQGOTGRGIEQPQS  
 SjCSL2 HMGWDVAGHFLPGDSGRFPAVTDAFSSQFVPLQHPQHDTPYHGLPNAGA  
 SPCC1223.13 SNVPSSDSVRNSSPNQYYASTSKQSIQSQNLQPPQKASVLGTVNNYRQ  
 Pc6518 DIGPGTGAGDDLAVIHSLRRSSDALAPGLAYTEPAWERDTALMSTYPGFL  
 CC1G\_01706 GKSP-----  
 RO3G\_06953 -----  
 RO3G\_13784 -----  
 RO3G\_08863 -----  
 UM05862 LMPTPQPGSPAPPPTQDVLAQSASQSI PDAATQTASITSESEPSSTSADS  
 SjCSL1 -----  
 SPCC736.08 -----  
 CNA01890 -----

Mm\_RBP-Jk -----  
 Hs\_RBP-Jk -----  
 XI\_SuH -----  
 Dr\_RBPSUH -----  
 Hr\_RBP-Jk -----  
 Ci\_SuH -----  
 Dm\_SuH -----

Ce\_LAG-1 MTFQMQQNNLFPKIDTI-----SKSPT  
 Hs\_RBP-L -----  
 Mm\_RBP-L -----  
 Dr\_RBP-L -----  
 PcCSL2 GFPPSGLPPFPDRLG-----RVO  
 CC1G\_03194 ---QSMPOFPDRLG-----RYP  
 UM06280 QREGHSTPTIVEEEEDVPQRQTRATVLRTAHTDVRIEPGFQSHPAAHQLQ  
 RO3G\_07636 SVIY---HHSPQTTTIS-----SPD  
 RO3G\_06481 LTIPDNNHSPTRNIIS-----SPL  
 RO3G\_11583 LFSSYLFDLVDTRPREL-----SMD  
 RO3G\_14587 SPIOSPTRYSPGTPGFF-----TPG  
 CNBD3370 MLGTYQSTVMSQYGYYPHPPEMQQTQQPIEPSVOLGTISPSELGQROPL  
 SjCSL2 YPEVYPADPTSSYSGPFTDQQGNFGSSFGDVVPGGNPTSSNSPAVLELSG  
 SPCC1223.13 YQNSFISLNDYQAAQSNISSPSSRFPTPYSPSVFPGTYQEKEKSYSQDHA  
 Pc6518 SPSSPSSHLHFPMASIAVPSPADPFPAPETANPOELSRPSSSAQPTFDSL  
 CC1G\_01706 -----  
 RO3G\_06953 -----  
 RO3G\_13784 -----MNRDVLESRIM  
 RO3G\_08863 -----  
 UM05862 AAQNFVDKLLREANLFPQESVNQQQHEFTISNGVQDASIPTATTETEPKA  
 SjCSL1 -----  
 SPCC736.08 -----  
 CNA01890 -----  
  
 Mm\_RBP-Jk -----MP  
 Hs\_RBP-Jk -----  
 XI\_SuH -----  
 Dr\_RBPSUH -----  
 Hr\_RBP-Jk -----MYHLNVT  
 Ci\_SuH -----M  
 Dm\_SuH MKSYSQFNLNAAAPPAAIAYETTVVNPNGSPLDPHQQQQQSQDMPHFGLP  
 Ce\_LAG-1 PELASPSAKRMRLSPSTSSHSDVASTSKGTNGQDOSKNSPTNSDPFYTIG  
 Hs\_RBP-L -----  
 Mm\_RBP-L -----  
 Dr\_RBP-L -----  
 PcCSL2 NESRYPNSTVPPLTATSHLTQSHSPELIRGVAPQATHLPSFDDMN-----  
 CC1G\_03194 NDR--FSHPQGGQQVQSHISPAQSTDMMRGIPPHATHSFDRDPSAGYDS-  
 UM06280 HHHPNLQQLQHPHSMRQHQPPLPQLLPGHLPPALHRSHSDQLSGLCMD  
 RO3G\_07636 HSNLFNLQENKIYSNHQQKN-----  
 RO3G\_06481 NNNYDNTSLFHAIFKEDEKDKN-----  
 RO3G\_11583 SSIISDLSLNDLSNNNPDLKFNSTLETITPSATLTNEINSMADWLLNENP  
 RO3G\_14587 FLESLOEHPVYDHSLSIDYGSHHFNOEYNPLLVKLEEQSPEKNLVSQSGE  
 CNBD3370 KPTKSFSDLMMGSRASSSSSLASQDVPPDGWNGNALEDWTRPLSRALPSS  
 SjCSL2 HSPHTLYGAGDAPRLOQLHQQPROQFHVSDYGFHVPAGTNNPNANNGFR  
 SPCC1223.13 ELSYYQQSPSMMPPYDRSSVYFQOPLSRTDVPNOSFQQYPTTVDDGGSMIP  
 Pc6518 VHTPASSHPDDVQTADTTANRRPSSPTQPAQTIQSILTATTEPDTSCAAP  
 CC1G\_01706 -----DINPKRSPTPOVGGAPSPVVAQDPNLSIQSILAASKKDGPTNGTS  
 RO3G\_06953 -----MTPMESLLNAIDKVEP-----  
 RO3G\_13784 SEQIFGSSRKIQFIGLSKPPLAPAQQQKDQMSIESLLTFSHPRQS-----  
 RO3G\_08863 -----MSIESLLTFSHPRQS-----  
 UM05862 SSVGDDALTASLEQFMASLOARVSSSLATEAATDTDAAKAKQAEESTVKS  
 SjCSL1 ---IQAFPESVSEASTDSELORQQQVQSVPPAVLFSPPSIPENPKKR  
 SPCC736.08 -----FPSVVKDEQSIFPAVAKASNSLDELAQNIIPAPSPPGFNKRK  
 CNA01890 -----MAEDHNEGVSNTPLDLLINAIAGNHYPH  
  
 Mm\_RBP-Jk SGFQSPRTSPRARPKTRITGALPMDYSEGLSAEERPAHAPSAGKFGERP  
 Hs\_RBP-Jk -----MAWIKRKFGERP

XI\_SuH -----MQPGIPKYTPSAIQLAPVVTGKFGERP  
 Dr\_RBPSUH -----MAPVVTGKFGELP  
 Hr\_RBP-Jk PDKOTSPHGCLVSETDQQRSHVKERVNGTPNQNGGTSTSSKPRSVFENR  
 Ci\_SuH YHPHHLPAHGQVQSHOHREDAATSSRDVNGGLSVTESAIASFRLREKY  
 Dm\_SuH GPOPPSSQQQQQLOVHHQQQQQQQQQQQQQQHQQQMOMSLLPGPYRPHI  
 Ce\_LAG-1 QSTPNTSSFLDTSNSSFAGPSTVVANPMTNYQLAFOAKLGLSLHSLIGDSV  
 Hs\_RBP-L ---MDPAGAADPSVPPNPLTHLSLQDRSEMQLQSEADRRSLPGTWTRSSP  
 Mm\_RBP-L ---MDPRETTDPSLPPGPLTHLSLPSDSEVRLQSDGP--SLLGSWTRSP  
 Dr\_RBP-L -----MQDQRKEDRDTGLSIVDSTDSHLYTPDACFTHLREDQ-SFS  
 PcCSL2 -----FMGASPTVYDPPLRLPASVNDDMARLRLOG-----SGDLQTF  
 CC1G\_03194 -----MHYMNPNPHQDMSLRMSTAVDETLARMKLOAP--PIGTSSDLQTF  
 UM06280 APPTFRNQMSYPPHMRPSLYQQVSMAMDQLRQIHGLGVTERNFGDCSAF  
 RO3G\_07636 -----LIQHY  
 RO3G\_06481 -----LNHTCINQIGEDMKCSVHRY  
 RO3G\_11583 QKRPRRSTDSPLGSSSDSSSSPPITPMQOVSLGFEPQEEWDLOPLIQNY  
 RO3G\_14587 SVTSLFSPDPAIVRPNOAHDSPIRRSSHTTASSPHRLVHLSPLKIKPF  
 CNBD3370 GTPVSTSAGSRQLPSAADISSLPDLTKRPSFASSPLRNPSFSPLSPNASD  
 SjCSL2 VPFGGSSPGPELSQTSFYDGPFPAPAESSHASLLRPSTSHSDSMGLNAP  
 SPCC1223.13 NLYPTSAEQMGLYPODSONKDTYPKSLVNRPSAVCEPARNDSIPMMVYS  
 Pc6518 PAQPAWNPTANVGHPPRETHSALPGPPPAPPRTSTKRKLDASEADDPHTK  
 CC1G\_01706 QAT-----WHPPPPPATSTSTGKRKLDTESEEDTHSK  
 RO3G\_06953 -----ITKKRKLMSIES-----  
 RO3G\_13784 -----QASTNDITIDNLLTCD  
 RO3G\_08863 -----QASTNDITINNLLTCD  
 UM05862 KPSPQKQPKVVKKPKKVLGKRKYREIGPMERRSRVEKALSEYIASIDTLT  
 SjCSL1 KAS-----FSNHQDPLIAQIRETIFDHLNLLLKKSATSNA  
 SPCC736.08 -----TFDESSVEMIRRAISDHLDLLNCCIGIANL  
 CNA01890 DHTHEHDHHDVSLDHLLSQRKDSAEDSQQHASSHKRARSQEVVSSRQRK

Mm\_RBP-Jk PPKRLTREAMRNY-----LKERGDQTVLILHAKVAQKSYGN  
 Hs\_RBP-Jk PPKRLTREAMRNY-----LKERGDQTVLILHAKVAQKSYGN  
 XI\_SuH QPKRLTREAMRNY-----LKERGDQTVLILHAKVAQKSYGN  
 Dr\_RBPSUH QPKRLTREAMRNY-----LKERGDQTVLILHAKVAQKSYGN  
 Hr\_RBP-Jk PPQRLTREAMSKY-----LRERNDQTLIILHAKVAQKSYGN  
 Ci\_SuH PPKLTRDAMRRY-----LKDPNDQTLIVLHAKVAQKSYGN  
 Dm\_SuH EEKKLTRDAMEKY-----MRERNDMVIVILHAKVAQKSYGN  
 Ce\_LAG-1 Q--SLTSDRMIDFLS-----NKEYECVISIFHAKVAQKSYGN  
 Hs\_RBP-L EHTTILRGGVRRRC-----LQQQCEQTVRILHAKVAQKSYGN  
 Mm\_RBP-L EHAIILREGVRTC-----LQQRCEQTVWILHAKVAQKSYGN  
 Dr\_RBP-L SLLKVTRDSVRQY-----LQFRPDQSVTILHAKVAQKSYGN  
 PcCSL2 IRPYLDQYIRTPN-----RLAFGERTVIVMSSKVAQKSYGT  
 CC1G\_03194 IRPFLEQYVRTPN-----RLAFGERTIIVMSSKVAQKSYGT  
 UM06280 IQDYVRQYISTSS-----RLGLGERNVLIMTTRVAQKSYGA  
 RO3G\_07636 LSTKQ-----GEKKLITLSKVAQKSYGN  
 RO3G\_06481 LITKQ-----DERKITILTSKVAQKSYGS  
 RO3G\_11583 LLQKQSREDYI-----PGERTIMILTSKVAQKSYGT  
 RO3G\_14587 IQTYLAHAITQPA-----ATQLGEKTVIVLTSKVAQKSYGT  
 CNBD3370 TDSIHKQYIRAPN-----RLAFGERRIIVMSPKVGQKSYGT  
 SjCSL2 DRNNLFORVKNVL-----RHELLCAMKVFMPSLGQKSYGK  
 SPCC1223.13 QPVTIEQRIQYVL-----SNCHCLSAFYLCMPSLCQKSYGT  
 Pc6518 I-RRVVRDHVSQHT-----PRPIPMTTVLCLHAAVAQKSYGT  
 CC1G\_01706 I-RRIVRDHVSQDP-----ARVVPMTTVICLHAAVAQKSYGS  
 RO3G\_06953 --MIEKRQLNSPP-----LSPKQPNKVTCYHASIAQKSYGT  
 RO3G\_13784 VK-----EPTR-----KRAKTSKMTCYYGAVAQKSYGS  
 RO3G\_08863 MKRKWAEPSSLVR-----KLLKKKSKVVCYYGAVAQKSYGS  
 UM05862 KPNSVTGPIPAAD-----PSKTKMTTIRCSHASVAQKSYGS  
 SjCSL1 LCNRIIRDPLNQY-----NQANKLVTISYQHASVAQKSYGA

SPCC736.08 NEDSVHKISLTRS-----GKPSQLVTVSCRHSSVIQKSYGS  
CNA01890 VPKPLGSQAMFYPGFAPADKAPGSLDENNSHRFNTLEVWHPTTGQKSYGK

Mm\_RBP-Jk EKRFFCPPPVCVYLMGSGWKKKKE-----  
Hs\_RBP-Jk EKRFFCPPPVCVYLMGSGWKKKKE-----  
Xl\_SuH EKRFFCPPPVCVYLMGSGWKKKKE-----  
Dr\_RBPSUH EKRFFCPPPVCVYLMGCGWKKKKE-----  
Hr\_RBP-Jk EKRFLCPPCLYLMESGWKQKQO-----  
Ci\_SuH EKRFFCPPP MYLLGNGWKRKQO-----  
Dm\_SuH EKRFFCPPP CIYLF GSGWRRRYE-----  
Ce\_LAG-1 EKRFFCPPP CIYLI GQGWLK KDRVAQLYKTLKA-----SAQKDA  
Hs\_RBP-L EKRFFCPPP CVYLS GPGWRVKPG-----  
Mm\_RBP-L EKRFFCPPP CVYLA GPGWRVKPM-----  
Dr\_RBP-L EKRFFCPPP CVYLS GHGWKVRQE-----  
PcCSL2 EKRFLCPPPTAIMIGNSWWSDVHR-----RGE EPKLCP  
CC1G\_03194 EKRFLCPPPTAIMIGNSWWSESTR-----RGE EPKLAP  
UM06280 EKRFLCPPPLVILIGSSWWNVCPASTRPSFAPAGAADPDQAAPT  
RO3G\_07636 EKRFLCPPSTILSGTGHWWTAKO-----HPP  
RO3G\_06481 EKRFLCPPYIILSGSNNWWSTON-----NPP  
RO3G\_11583 EKRFLCPPPATIMKGTNWWTSDKLTDKKTPSL-----FHSP  
RO3G\_14587 EKRFLCPPTAILVGTSWWTTKEKI QDKEETLRI----PSLEK  
CNBD3370 EKRFLVPHPQAILIGSGWWAKSPD-----GCPVSPLOPP  
SjCSL2 ERRYICPPAVVYLLGSSWFRCPDKIN-----  
SPCC1223.13 ERRYLCPPIVLYLLGTTWLNNTDN-----LK  
Pc6518 EKRFLCPPPIVHIEGPVWNMKNQ-----  
CC1G\_01706 EKRFLCPPPIVHIEGPVWHLRTO-----  
RO3G\_06953 EKRFLCPPPIVISQ-SQ---TNC-----  
RO3G\_13784 EKRFLCPPPLVQVHHST--RNL-----  
RO3G\_08863 EKRFLCPPPIVOLKGAN--KEM-----  
UM05862 EKRFFNPPPIISITGPLRRYAEHG-----AA  
SjCSL1 EKRYLCPPLISITGNTGSILGHS-----F  
SPCC736.08 EKRYLCPPPMVIYINGNYSSIFNQS-----F  
CNA01890 ERRMLNPPPIRLRSLGPASSCISITMVS VVSPN-----ITQP

Mm\_RBP-Jk QMERDGCSEQESQPCAFI GIGNSDQEMQQLNLEG-----  
Hs\_RBP-Jk QMERDGCSEQESQPCAFI GIGNSDQEMQQLNLEG-----  
Xl\_SuH QMERDGCSEQESQPCAFI GIGNSEQEMQQLNLEG-----  
Dr\_RBPSUH QMEREGCSEQESQPCAFI GIGNSEQEMQQLNLEG-----  
Hr\_RBP-Jk ILEEADQAPEAGQVHAFI GIGSSDQEMQQLNLDG-----  
Ci\_SuH ILEEEGSSEAGQLHAFI GIGSSSEQEMQQLHLDG-----  
Dm\_SuH EMLOQGE GEGAQLCAF I GIGSSDQDMQQLDLNG-----  
Ce\_LAG-1 IENDPIHEQQATELVAYI GIGSDTSERQQLDFSTGKVRHPG-----  
Hs\_RBP-L QDOAHOAGETGPTVCGYMG LDSASGSATETQKLN-----  
Mm\_RBP-L QDOALQSAETGPTVCGYMG L DGASGSAPETQKLN-----  
Dr\_RBP-L QFKASGLNESACRIFGYMG L DSSNDPRTDSFKLS-----  
PcCSL2 RVVVSISGEPAPQEGSIEWTSATGKAFDVSDPPTGTT-----  
CC1G\_03194 RVVVSISGEPAPQEC SIEWTSSSGKSFVDTPPSGTT-----  
UM06280 RVTINISGEPGVODGALEWASSSGRLIDVGNPSEMA-----  
RO3G\_07636 NLTIQISGEKLSHQGTIDWYKDG NLLDQPSAVLLANQGSN-----  
RO3G\_06481 NVHIQLSGEQRPYCHIDWYKEGTLINOH LAATYASQGDI-----  
RO3G\_11583 KLTIHISGETIQOTGVIEWQTSSGNIID-----NNAQK-----  
RO3G\_14587 KLTVISGETSTQAGOLEWYTVSGATVGQTGQIKPPIKPESTSFRFSSES  
CNBD3370 RINISLTGEQAVKDSIISWTDLNGKNMDEKASTQGVKIDDO-----  
SjCSL2 IIANAADDPDNLKTSETPTFY TSSSDSATNLLSLGQVKLDDPLE-----  
SPCC1223.13 ISAQTLEDKDNPKFAKNIFYNADGALISPETDI AKSTYQLTNYN-----  
Pc6518 QFSMTVVSEGGERSFEQRAPIDHNL-----  
CC1G\_01706 SLSMAVSETGERSFEQKATLDN NMV-----

RO3G\_06953 MISMTIMNENNOAILEORTLLDDESK-----  
 RO3G\_13784 LLSMSILCED-QGLLEQRAILDENQR-----  
 RO3G\_08863 ELSMSILCED-NDILEQHTALDEHHS-----  
 UM05862 TLLTPLSSEDPEEDQCRVSLVVRGDTDEYFSNENVAVLGD-----  
 SjCSL1 SVOLAITNEOGQHSNOVSENFAERQS-----  
 SPCC736.08 RTEISIMNDFGQCSQPISSEYTGQGC-----  
 CNA01890 AICAPQTHHMGRPPDITPHLGTPGVEDKKKERREEQKLRRTAALNAGFAA

Mm\_RBP-Jk -----KNYCTAKTLYISDSD--KRK----HFMLSVKMF  
 Hs\_RBP-Jk -----KNYCTAKTLYISDSD--KRK----HFMLSVKMF  
 Xi\_SuH -----KNYCTAKTLYISDSD--KRK----HFMLSVKMF  
 Dr\_RBPSUH -----KNFCTAKTLYISDSD--KRK----HFMLSVKMF  
 Hr\_RBP-Jk -----KNFCTAKTLYISDSD--KRK----HFMLCVKMF  
 Ci\_SuH -----KNFCTAKTLYISDSD--KRK----HFMLNVKMF  
 Dm\_SuH -----KQYCAAKTLFISDSD--KRK----HFMLSVKMF  
 Ce\_LAG-1 -----DQRQDPNIYDYCAAKTLYISDSD--KRK----YFDLNAQFF  
 Hs\_RBP-L -----FEQOPDSREFGCAKTLYISDAD--KRK----HFRLVLRV  
 Mm\_RBP-L -----FEEQOPDSREFGCAKTLYISDAD--KRK----HFRLVLRV  
 Dr\_RBP-L -----FEEQTDKMMFACAKTLYISDAD--KRK----HFRLVLRV  
 PcCSL2 -----YIGRCVKGKQLFISDSD--KRK----HFMLNVKMF  
 CC1G\_03194 -----YLGRCVKGKQLFISDSD--KRK----HFMLNVKMF  
 UM06280 -----ISGRCIGKQLYISDSD--KRK----HFMLNVKMF  
 RO3G\_07636 -----LIGNCVSKQLHVSGAD-EKRRKAQVQVEIRSGQG  
 RO3G\_06481 -----LVGNCLSKQLHISDSD--KRK----HFRLVLRV  
 RO3G\_11583 -----VFGRCISKQLYINDAD-EKRRKVEVLAKIQLGNG  
 RO3G\_14587 RHPPADAYSNERQELLAAGKSVSKHLYIHDAD-EKRRKVECLVKLQLANG  
 CNBD3370 -----PFTGSVAGKNLHISDSD--KRK----HFRLVLRV  
 SjCSL2 -----QQQSPSPWIWANTVLTLYYSG--KGDHNTYGRSTTLQVH--  
 SPCC1223.13 -----ENTNFDSFPVWGNALLKTIYYTG--QGKNDGFRSTFLQL---  
 Pc6518 -----ASFKFLHVTGSA--KSKSFQLSLNITEPSP  
 CC1G\_01706 -----SSFKFLHVTGTA--KAKSFQLSLDIAEPPP  
 RO3G\_06953 -----GYFKYLYVTSSA--KSKQFRIKLNMMN----  
 RO3G\_13784 -----GSFRYLHVTGTA--KAKQFKLQLNLLQPE-  
 RO3G\_08863 -----GSFRYLHVTGTA--KAKQFKLQLNLLQNE-  
 UM05862 -----RLETKMRSLYVTPTG--RSKSFRLQLNLLRPSG  
 SjCSL1 -----VAFRSLHISSSI--AAKAKSFNLNIDVIS-  
 SPCC736.08 -----MIFRSLHISSLV--AAKSKNLRLSLDMFS-  
 CNA01890 TLPKNRNATDLKDRITLLKDGLTFPGLWIGEEA-GKMKEFRLELKIYAAPT

Mm\_RBP-Jk YGNS-----  
 Hs\_RBP-Jk YGNS-----  
 Xi\_SuH YGNS-----  
 Dr\_RBPSUH YGNS-----  
 Hr\_RBP-Jk FGSG-----  
 Ci\_SuH FGGGG-----  
 Dm\_SuH YGNG-----  
 Ce\_LAG-1 YGCG-----  
 Hs\_RBP-L LRGG-----  
 Mm\_RBP-L LRGG-----  
 Dr\_RBP-L HNGG-----  
 PcCSL2 DDEPE-----  
 CC1G\_03194 DDEPE-----  
 UM06280 SPMDR-----  
 RO3G\_07636 TP-----  
 RO3G\_06481 RS-----  
 RO3G\_11583 SN-----  
 RO3G\_14587 LQ-----



CNBD3370 KHAGPNGWWSAKG-----  
 SjCSL2 -VRTP-----  
 SPCC1223.13 SVQSK-----  
 Pc6518 PTADGSEP-----  
 CC1G\_01706 PSLNPEGSE-----  
 RO3G\_06953 -----  
 RO3G\_13784 -----  
 RO3G\_08863 -----  
 UM05862 LEPHLPIFSPLKKSRAADALSPAM-----  
 SjCSL1 -----  
 SPCC736.08 -----  
 CNA01890 EKGLSSPSTQDFAOSSYHLGNPDNDHETAHTFADMLEPPSMDALQPLAEA

Mm\_RBP-Jk -----  
 Hs\_RBP-Jk -----  
 XI\_SuH -----  
 Dr\_RBPSUH -----  
 Hr\_RBP-Jk -----  
 Ci\_SuH -----  
 Dm\_SuH -----  
 Ce\_LAG-1 -----  
 Hs\_RBP-L -----  
 Mm\_RBP-L -----  
 Dr\_RBP-L -----  
 PcCSL2 -----  
 CC1G\_03194 -----  
 UM06280 -----  
 RO3G\_07636 -----  
 RO3G\_06481 -----  
 RO3G\_11583 -----  
 RO3G\_14587 -----

CNBD3370 -----TLSDVTN  
 SjCSL2 -----  
 SPCC1223.13 -----  
 Pc6518 -----TVPG  
 CC1G\_01706 -----TSAN  
 RO3G\_06953 -----  
 RO3G\_13784 -----D  
 RO3G\_08863 -----  
 UM05862 -----SIGSPANAPIENGSEVLSRLPK  
 SjCSL1 -----SSD  
 SPCC736.08 -----NVN  
 CNA01890 VQOSSQOGHDTLSEHFPEVEPSLAATNLHNDNDAENPGNSPLPITQATTE

Mm\_RBP-Jk -DDIGVFLSKRIKVISKPSKKKQSLKN---AD---LCIASGTKVALFNRL  
 Hs\_RBP-Jk -DDIGVFLSKRIKVISKPSKKKQSLKN---AD---LCIASGTKVALFNRL  
 XI\_SuH -DDIGVFLSKRIKVISKPSKKKQSLKN---AD---LCIASGTKVALFNRL  
 Dr\_RBPSUH -ADIGVFLSKRIKVISKPSKKKQSLKN---AD---LCIASGTKVALFNRL  
 Hr\_RBP-Jk -LEVGFNSRRIKVISKPSKKKQSLKN---AD---LCIASGTKVALFNRL  
 Ci\_SuH -ADVGOQFSSKRIKVISKPSKKKQSLKN---AD---LCIASGTKVALFNRL  
 Dm\_SuH -HDIGVFNSKRIKVISKPSKKKQSLKN---AD---LCIASGTNVALFNRL  
 Ce\_LAG-1 -MEIGGFVSQRIKVISKPSKKKQSMKN---TDCKYLCLIASGTKVALFNRL  
 Hs\_RBP-L -RELGTFFHSRLIKVISKPSQKKQSLKN---TD---LCISSGSKVSLFNRL  
 Mm\_RBP-L -QELGTFFHSRLIKVISKPSQKKQSLKN---TD---LCISSGSKVSLFNRL  
 Dr\_RBP-L -QEIGTFNSRLIKVISKPSQKRQSMKN---AD---LCISSGSKVSLFNRL  
 PcCSL2 -RVIGTFPSRPIKVISKPSKKRQSAKN---LE---LCINHGSTISLFHRL  
 CC1G\_03194 -RIIGVFPSRPIKVISKPSKKRQSAKN---LE---LCINHGSTISLFHRL

UM06280 -RLLGTFPSRPIKVISKPSKKROSSRN---TE---LCVNHGVSISLFHRL  
 RO3G\_07636 ---IGIFHSKPIKVISKPSKKROSVKN---MD---LCIHHGTTVALFNRY  
 RO3G\_06481 ---LGTFDSPKPIKVISKPSKKROSAKD---ME---LCIHNGSTVALFNRT  
 RO3G\_11583 ---LGTFSKGIKVISKPSKKROSAKN---ME---LCIHHGTTISLFNRI  
 RO3G\_14587 ---LGPLASKAIKVISKPSKKROSIKN---ME---LCIHHGTTVSLFNRI  
 CNBD3370 DNVFGIFESKDIKISKPSKKRSTAKS---GE---LTISHGTTVALFNRI  
 SjCSL2 -QKRITMDKLRIGIISKPSQKKMMMKV---SD---LNICHGDCVSLFNRF  
 SPCC1223.13 -TKYFKLENLRLGVISKPSQKRALMKV---SD---MSIRHGDCVCLFNRY  
 Pc6518 RLPWASFDSAPVTIISKPSKK--TAKT-RNIA---SCILAGGPVSLFNRI  
 CC1G\_01706 GRVWATFDSAPVTIISKPSKK--TAKT-RNIS---SCILAGGPVSLFNRI  
 RO3G\_06953 ---DALFYSNPISIISKPSKK--TAKT-RNAT---SCLFNHSTISLFNRI  
 RO3G\_13784 ESPLATFLTKPITIISKPSKK--IGKTARGTT---SCISTEAPVSLFSRI  
 RO3G\_08863 -SSFASFLSKPISIISKPSKK--TAKTSRNTS---TCILANSPVSLFNRI  
 UM05862 ELAWASFESAPVTIISKSSK--TAKP-R-GE--LNQVKNGSMISLYNRI  
 SjCSL1 -DVL AQMVTKPINIISKPSKK--GSKN-RVSN---STLMSGAIISLYNRI  
 SPCC736.08 NQLLSHLVTSSISIVSKPSKK--GSKL-KISN---ITLRSGSVVSLYNRI  
 CNA01890 EQPLGSFLSGPIRIVSKPSRK--TSKT-RSLA---ICFPRDSVFLWTRI

Mm\_RBP-Jk RSQTVSTRYLHVEGGN-----FH  
 Hs\_RBP-Jk RSQTVSTRYLHVEGGN-----FH  
 XI\_SuH RSQTVSTRYLHVEGGN-----FH  
 Dr\_RBPSUH RSQTVSTRYLHVEGGN-----FH  
 Hr\_RBP-Jk RSQTVSTRYLHVENGH-----FH  
 Ci\_SuH RSQTVSTRYLHVEKGN-----FH  
 Dm\_SuH RSQTVSTRYLHVENGH-----FH  
 Ce\_LAG-1 RSQTVSTRYLHVEGNA-----FH  
 Hs\_RBP-L RSQTVSTRYLSVEDGA-----FV  
 Mm\_RBP-L RSQTVSTRYLSVEDGA-----FV  
 Dr\_RBP-L RSQTVSTRYLAVEDGA-----FV  
 PcCSL2 RSQTVSTKYL CVSGSGSSFKGSDGAPLM-G-LDQ---RARTQTP---SFI  
 CC1G\_03194 RSQTVSTKYL CVSGSNSSFKGSDGAPLM-G-LDQ---RSRQSTP---SFI  
 UM06280 RSQTVSTRYLCVSGAPSWFKGSDNQPFL-N-MNP---RDQKGSEPPSCFV  
 RO3G\_07636 RAQTISTKYLGVSSLD-----SQQKDRGTFCV  
 RO3G\_06481 RSQATSTKYLGISSQT-----NNQ-NPNTCFV  
 RO3G\_11583 RSQTVSTKYLGVSTTT-----POPDSNGTCFV  
 RO3G\_14587 RSQTVSTKYLGVSTSKGSPLAFPGLAQ-----HEKNRTSEGTCFV  
 CNBD3370 KSQTTSTRYLSVVPDFTRTLGSDGRPVTGA-KAP---QYENDKGALRGFT  
 SjCSL2 RSHNNLPRYLCTNVLNDVVTKRTEHLQFHE-EFTPATDLDLMEASTCRLI  
 SPCC1223.13 RAQHNNALFLGTSNVQRAISKVSLNMKYNSNYFPTTDAPNDAENEGAGLA  
 Pc6518 NSQTVRRTKYMTVDNGQ-----LC  
 CC1G\_01706 NSQTVRRTKYMTIDHGO-----LC  
 RO3G\_06953 NSQTVRRTKYMSTESNQ-----LC  
 RO3G\_13784 NSQTVRRTKYMTSVGSK-----LC  
 RO3G\_08863 NSQTVRRTKYMTCDHNR-----LC  
 UM05862 NSQTARTKYMRSSNGG-----DLT  
 SjCSL1 NSQTVRRTKYMSVANGR-----FC  
 SPCC736.08 NSQTVRRTKYTSIEAGQ-----FC  
 CNA01890 HGQTVRRTKWMNLEANFNG-----HGPRLT

Mm\_RBP-Jk ASSQQWGAFYIHLLDDDE-----SEG  
 Hs\_RBP-Jk ASSQQWGAFFIHLLDDDE-----SEG  
 XI\_SuH ASSQQWGAFYIHLLDDEE-----SEG  
 Dr\_RBPSUH ASSQQWGAFYIHLLDDDE-----SEG  
 Hr\_RBP-Jk ASSIQWGSFAIHLLDDDE-----SES  
 Ci\_SuH ASSIQWGCFAIHLLDDDE-----SES  
 Dm\_SuH ASSTQWGAFTIHLLDDNE-----SES  
 Ce\_LAG-1 ASSTKWGAFTIHLFDDERGL-----QET

Hs\_RBP-L ASARQWAAFTLHLADG-HSA-----QGD  
 Mm\_RBP-L ASARQWAAFTLHLADG-HCS-----QGD  
 Dr\_RBP-L ASARQWTAFTVNLVEETOAV-----HSE  
 PcCSL2 ARTASWDPFIMYIVDVNKP TGGIDTPP-----PPPPQPEFSPPPNAIPF  
 CC1G\_03194 ARTASWDPFIMYIVDVNKP GGGLDVPP-----PPPPQPDYPSPPNAIPF  
 UM06280 AKMSSWDPFIIYLVDPKKRADATPSEP-----VQPPVKGYPPPPNALPT  
 RO3G\_07636 TRTTSWDPFLIYIVDLRSRPNTPSPV-----PFSHHPTIDHYPPPPAIA  
 RO3G\_06481 ARTTSWDPFLIWIWVDPTRSPNTSP-----STLQHAANSQYPPPPAIA  
 RO3G\_11583 SRTGVWDPFVIWIVDTS CSPNTAN-----RPKHNLNPNYPPPPAIA  
 RO3G\_14587 ARTTSWDPFVIWIVDTSASSEEEGETPEDYIGHHV FARSTPYPPPPPIAL  
 CNBD3370 ADASFWESEFIIWLVDPSLPSGSPSNHQAP-----NPDWPRPPANI  
 SjCSL2 TTNTVWEPFNIYSVEELE-----NK  
 SPCC1223.13 MANNLWEPFYIFSVDL NKG-----NNS  
 Pc6518 ASNVTWAAFNVSVVRP-----QGDO  
 CC1G\_01706 ASNIGWSAFNVNVVR-----RPDG  
 RO3G\_06953 AKHSTWSPFEIIVLR-----  
 RO3G\_13784 AKHSIWSPFDILLRQPQPTVLT TTKK-----TSKRYNTARLQ  
 RO3G\_08863 AKNTTWSPFDVIVLCQPKQV VITKQSK-----SSPRYNTSRLO  
 UM05862 VHNGDWSTYRITLISRPP-----LA  
 SjCSL1 LRNESWTPLOIRVYGAS-----  
 SPCC736.08 LRGDRWVPLRINLLLPD-----ENGKLV  
 CNA01890 SRTAKWTPFRFEVL DHANAPFAYRTARN-----KFD

Mm\_RBP-Jk EEFTVRDGYIHYGQTVKLVCSVTGMALPRLIIRKVDKQTALLDA-----  
 Hs\_RBP-Jk EEFTVRDGYIHYGQTVKLVCSVTGMALPRLIIRKVDKQTALLDA-----  
 XI\_SuH EEFTVRDGYIHYGQTVKLVCSVTGMALPRLIIRKVDKQTALLDA-----  
 Dr\_RBPSUH EEFTVRDGYIHYGQTVKLVCSVTGMALPRLIIRKVDKQTALMDA-----  
 Hr\_RBP-Jk EEFTVRDGYIHYGQTVKLVCSVTGMALPRLIIRKVDKQTAILNA-----  
 Ci\_SuH EEFSVVDGYIHYGQTVKLVCSNTGMALPRLIIRKVDKQTAILDA-----  
 Dm\_SuH EEFQVRDGYIHYGATV KLVCSVTGMALPRLIIRKVDKQ MALLEA-----  
 Ce\_LAG-1 DNFAVRDGFVYYGSVVKLVDSVTGIALPRLRIRKVDKQ QVILDAS-----  
 Hs\_RBP-L --FPPREGYVRYGSLVQLVCTVTGITLPPMIIRKVAKQCALLDV-----  
 Mm\_RBP-L --FPPQEGYIRYGSLVQLVCTVTGITLPPMIIRKVAKQCALLDV-----  
 Dr\_RBP-L --YTMCESYICYGCVVQLVCTDSGVALPPMIRKVNKQQA CLDV-----  
 PcCSL2 T-NNGSQIPIYYNQTVVLOCLTSGVVSPVLIIRKVDHOTT VVGGGLOEGA  
 CC1G\_03194 T-NNGSQIPIYYNQTVVLOCLTSGVVSPVLIIRKVDROTT VVGGGLOEGA  
 UM06280 TGLTSSQMTIHYNQPIVLQCLNTAVVSPVMVVRKVEK GTTVLGGASAGPS  
 RO3G\_07636 IQNHQGSALHYNQPVV LQCVSTGLVSPVLIIRRVEK GSMVMGGNRVNDL  
 RO3G\_06481 VQNPQPLV LHYNHLIVLQCVSTGLISPVMVIRRADK GSMVMGGNRVNDP  
 RO3G\_11583 LOTS-STLAIHYNQPVV LQCVTTGLVSPVMIIRKVDKQSLV LGGNRVDNP  
 RO3G\_14587 KNKTGGPVPIHYNQHV V LQCLTTGLVSPVMIIRKVDRASTV VGG--ARDD  
 CNBD3370 IPLNTISHSIRYNSTV VLQSLRTGVISPTLVVRR IETDSDAVGMDG----  
 SjCSL2 SSYDRRNTVICSNMRII IQSQITGVRSPPLIIRKYENRKALV VEDD----  
 SPCC1223.13 NPSDSRSKVLCSNMV IILVSKITGVOSPPLILKHDNWKVSL SSRAP---  
 Pc6518 PTIVLGPQPVTY GSEVILDTLSGISTSPLTIRKVDKGRIAPDDG----  
 CC1G\_01706 SPNTNGPQPVTY GCEIVLSDTQSGISTSPLVIRKVDKGRVSPDDG----  
 RO3G\_06953 --QPNKTEAISY GTEILLKDIQTGITSPLIIRKVEKNKV VYDAF-----  
 RO3G\_13784 FSTPASA AHLMYGSEIILRESHTGVSSAPLMVCKVDRGRIVPDAC-----  
 RO3G\_08863 LPKTP-VSYLTYGTEI ILRRESRTGLSSPPLILCKVDRGRLVESAY-----  
 UM05862 ELAGAEADSVTYG STIVLTDVNTGFSSDPLVVKVDKGOVQLPQLDPPSV  
 SjCSL1 --ESSFHOPIMYGSQV VLYNERSGIVSDPLYIHRVDKDIAPDD-----  
 SPCC736.08 CDDVDNPEPIKYGSIV ELVDEATGTTSDPLIIRRVEKD HIAEED-----  
 CNA01890 ETDHLEGOQLTYGSIV VLVDLQTVGKSDPVKVV RVESGKSVVGQYD----

Mm\_RBP-Jk -----  
 Hs\_RBP-Jk -----  
 XI\_SuH -----

Dr_RBPSUH	-----	
Hr_RBP-Jk	-----	
Ci_SuH	-----	
Dm_SuH	-----	
Ce_LAG-1	-----CS-----	
Hs_RBP-L	-----	
Mm_RBP-L	-----	
Dr_RBP-L	-----	
PcCSL2	KGIADHYCCPGEVC-----	
CC1G_03194	KGIADHYCVPGEVC-----	
UM06280	ESiAR-----EAF-----	
RO3G_07636	SYPTGGGEWG-DEAL-----	
RO3G_06481	PCSTGGGEYG-DEAL-----	
RO3G_11583	IGSLGGECS-DETL-----	
RO3G_14587	VSGSGGEFG-DEVL-----	
CNBD3370	HVHEVPAMPPGELA-----	
SjCSL2	-----QL-----	
SPCC1223.13	-----	
Pc6518	-----	
CC1G_01706	-----	
RO3G_06953	-----	
RO3G_13784	-----	
RO3G_08863	-----	
UM05862	LLKAGTGNKSRVPIDPALRDKLATVSKSVSHANPSTSLGAGASTSNHIE	
SjCSL1	-----	
SPCC736.08	-----	
CNA01890	-----	
Mm_RBP-Jk	-----DDPVS	
Hs_RBP-Jk	-----DDPVS	
XI_SuH	-----DDPVS	
Dr_RBPSUH	-----DDPVS	
Hr_RBP-Jk	-----DDPVS	
Ci_SuH	-----DDPVS	
Dm_SuH	-----DDPVS	
Ce_LAG-1	-----EEPVS	
Hs_RBP-L	-----DEPIS	
Mm_RBP-L	-----DEPIS	
Dr_RBP-L	-----DEPVS	
PcCSL2	-----GDPVS	
CC1G_03194	-----GDPVS	
UM06280	-----GDPVS	
RO3G_07636	-----GDPVS	
RO3G_06481	-----GDPVS	
RO3G_11583	-----GDPVS	
RO3G_14587	-----GDPVS	
CNBD3370	-----SDLVS	
SjCSL2	-----GDSIN	
SPCC1223.13	-----SEAIN	
Pc6518	-----GPVS	
CC1G_01706	-----GPVS	
RO3G_06953	-----GLLS	
RO3G_13784	-----GPVS	
RO3G_08863	-----GPVS	
UM05862	KPGSSVQSAIKIDGGETAVSSAPTGATSTSANVGSETKEKDDLNLFGPVG	
SjCSL1	-----GYLG	
SPCC736.08	-----GYVN	

CNA01890 -----GQPVS

Mm\_RBP-Jk QLHKCAFYLK-----DTERMYL  
Hs\_RBP-Jk QLHKCAFYLK-----DTERMYL  
XI\_SuH QLHKCAFYLK-----DTERMYL  
Dr\_RBPSUH QLHKCAFYLK-----DTERMPL  
Hr\_RBP-Jk QLHKCAFYLK-----DSERMYL  
Ci\_SuH QLHKCAFYLK-----DTERMYL  
Dm\_SuH QLHKCAFYMK-----DTDRMYL  
Ce\_LAG-1 QLHKCAFQMI-----DNELVYL  
Hs\_RBP-L QLHKCAFQFPGS-----PPGGGGTYL  
Mm\_RBP-L QLHKCAFQFSPD-----TPGGAGTYL  
Dr\_RBP-L QLHKCAFQFR-----GSDHMYL  
PcCSL2 QLHKIAFEVYD-----PNKGAPE-----PGTPGVSGAFLS  
CC1G\_03194 QLHKIAFEVYD-----PNKGLPE-----PGTPGVSGAFLS  
UM06280 QLHKIALEVLEDPAASVPRPTGDSGE-----SNWPGNSGPFLA  
RO3G\_07636 QLHKVGFQIVQDPSIAQYNKSTFQEQDKFFLPPVTHWTLPOATSAINYLA  
RO3G\_06481 QLHKVAFQIVQDFSKIRONKQNHKQS-----FPLTQWTLPOVSSDVQYLA  
RO3G\_11583 QLHKVAFQIVQDPSFHOGNLKQNTAG-----HWKIPOSSHPTVYLA  
RO3G\_14587 QLHKIALQIVQDPKMSVMQAP-----DPRMPRTSOPVYLA  
CNBD3370 QLQKVAFEIYSPDTMNRLORESKY-----GGSWLS  
SjCSL2 CLSRLAFQCPR-----TKLFLN  
SPCC1223.13 CLSKLAFQCHETKR-----FLYIDEKQSS  
Pc6518 QMOKIALQRTNTDGTTRHYLSAASPV-----PGQG---TATSQSPANQTGT  
CC1G\_01706 QMOKIALQRVNPDGSRHYLSAAGPL-----PGTPGVVAPPAPGMSTOAGT  
RO3G\_06953 QMOKVALQLAN-----TPNLYL  
RO3G\_13784 QMOKVALRLLDT-----TEKLYL  
RO3G\_08863 QMOKIALRLASSDPQHPV-----YLSAAGSM  
UM05862 QMOKVALMRFVPSNDSETVFGD-----SEHLESNFSTPRSYLC  
SjCSL1 QMHRIVLRRGQCNNNEPNVLKPYANO-----QGMDFVQPLFLGATTAQEKs  
SPCC736.08 QMHRIVLESAYPISNVRHLKIAEHSSL--AYSNNISVRWFLGATSAQNRN  
CNA01890 ELQRIGLVRTIDGQEDMSEGCRWYLSAPG-----ARIGGGELSA

Mm\_RBP-Jk CLS-QERIIQFQAT-----  
Hs\_RBP-Jk CLS-QERIIQFQAT-----  
XI\_SuH CLS-QERIIQFQAT-----  
Dr\_RBPSUH CLSPOERIIQFQAT-----  
Hr\_RBP-Jk CLS-QERIIQFQAT-----  
Ci\_SuH CLS-QERIIQFQAT-----  
Dm\_SuH CLS-QEKIIQFQAT-----  
Ce\_LAG-1 CLS-HDKIIQHQAT-----  
Hs\_RBP-L CLA-TEKVVFQFQAS-----  
Mm\_RBP-L CLA-TEKVVRVQFQAS-----  
Dr\_RBP-L CLS-NEKIIQFQAS-----  
PcCSL2 CMGEKVNTYRPEGVROWNANVTN----TAEAD-----SPIA  
CC1G\_03194 CMGEKVNTYRPIEGRIWNTNVAN----NNGSS-----SPVL  
UM06280 CLNESVGMHRPSEPRRWVGNSTF-----SVPSTPTTPVTSMNLAAMEGID  
RO3G\_07636 CIHDVVMHRVTDERKIVSRFTT-----EIEDI-----  
RO3G\_06481 CLNEVIGTHVVTGPRTIVSSFMP-----KLEQA-----  
RO3G\_11583 CLNDVVMHKTSTRHLVPOCQE-----NTFGE-----  
RO3G\_14587 CLNDMVMGMHKTSEGRSWAGWDDS-----ITSQE-----  
CNBD3370 CYQDEVREHAIKAERKWAVLQAP----AGSKPSTRPNSLPNTPQORFGV  
SjCSL2 LDEFSNGEIRFLSADPAPG-----  
SPCC1223.13 EISFTSGELEYS DP-----  
Pc6518 HPLVFOAPRVREET-----  
CC1G\_01706 HPLLFQNPRIRDEV-----  
RO3G\_06953 NSNGKTIENDLNNNVY-----

RO3G\_13784 SCPSIDDQTWLDSR-----  
 RO3G\_08863 SSEQDTHQTWLDYS-----  
 UM05862 GMVPGGWMHISPEKVGQOSLAVA-----FAQTHEEALGVGQLPAHTSDKH  
 SjCSL1 TTGDMDFPVELEPA-----  
 SPCC736.08 ASSEAILPIEWEAV-----  
 CNA01890 ESAPGIRMRRQRANAKAASQLKELNEERGKEKEIEGTSHDEPHLTSSFLQ

Mm\_RBP-Jk -----  
 Hs\_RBP-Jk -----  
 XI\_SuH -----  
 Dr\_RBPSUH -----  
 Hr\_RBP-Jk -----  
 Ci\_SuH -----  
 Dm\_SuH -----  
 Ce\_LAG-1 -----  
 Hs\_RBP-L -----  
 Mm\_RBP-L -----  
 Dr\_RBP-L -----  
 PcCSL2 APGSPIAGTPTSANGHGDYFGGNGGSG-SAPNSPIP-----T  
 CC1G\_03194 PGS---PMSPTNGSASAEYFGQATGANESASPSS-----A  
 UM06280 PTGGYLSMSPTSAAAYAAAQTRYSLSAGRAGATMMSPSASTGSHDVSAME  
 RO3G\_07636 -----  
 RO3G\_06481 -----  
 RO3G\_11583 -----  
 RO3G\_14587 -----  
 CNBD3370 LPMTPHTN-----NMNLPSTPSSPVSSSSSLDYFNYO  
 SjCSL2 -----  
 SPCC1223.13 -----  
 Pc6518 -----  
 CC1G\_01706 -----  
 RO3G\_06953 -----  
 RO3G\_13784 -----  
 RO3G\_08863 -----  
 UM05862 HTGETNSVVYAT-----  
 SjCSL1 -----  
 SPCC736.08 -----  
 CNA01890 DPQTLSAIENVSGSSTETNHNQOSVLSIELSLPEGVHSADTFDHTDHPV

Mm\_RBP-Jk -----  
 Hs\_RBP-Jk -----  
 XI\_SuH -----  
 Dr\_RBPSUH -----  
 Hr\_RBP-Jk -----  
 Ci\_SuH -----  
 Dm\_SuH -----  
 Ce\_LAG-1 -----  
 Hs\_RBP-L -----  
 Mm\_RBP-L -----  
 Dr\_RBP-L -----  
 PcCSL2 EFPSSDGGKVKKGRGSSSAGG-----LSKPASQKRRRPTSAGSASGS-  
 CC1G\_03194 EFLSNDGGRVKK-KRSTSSAGG-----LSKAVGGKRRRPSSAGSVS---  
 UM06280 PPPSSDGGKVKRPRRVSSSVVVQKERAASAAAASKSRRRGQSLSMVGION  
 RO3G\_07636 ---KMAVR-KRRLSYQQHSTTVK-----SSNRRRVNSLNDELLSRH  
 RO3G\_06481 ---DAQISRKRRLSYQYSSLTSK-----PSSRRRVNSLNDELLSSD  
 RO3G\_11583 ---IAODPIVRRLSTGEIKRRGS-----LGKGSSALDP-----  
 RO3G\_14587 ---GGKIIRKRRVSTDVQPETLMSCMSLSDYPRRRVNSLED PAPYLARKS  
 CNBD3370 SQRSSGHNHSLMSPGSVETALPSTDGGPIRRQRTSSMGKGPLSRPLHRKR

SjCSL2 -----  
 SPCC1223.13 -----  
 Pc6518 -----  
 CC1G\_01706 -----  
 RO3G\_06953 -----  
 RO3G\_13784 -----  
 RO3G\_08863 -----  
 UM05862 -----  
 SjCSL1 -----  
 SPCC736.08 -----  
 CNA01890 IDPSFENMAPLVDSQQQTKNPKGQSSFQRRSNKTRRLALAAAVLAEDDDN

Mm\_RBP-Jk -----PCPKEQNKEMINDGASWTIIST  
 Hs\_RBP-Jk -----PCPKEPNKEMINDGASWTIIST  
 XI\_SuH -----PCPKEPNKEMINDGASWTIIST  
 Dr\_RBPSUH -----PCPKETNKEMINDGASWTIIST  
 Hr\_RBP-Jk -----PCPKEPNKEMINDGASWTIIST  
 Ci\_SuH -----PCPKETNKEMINDGASWTIIST  
 Dm\_SuH -----PCPKEPNKEMINDGACWTIIST  
 Ce\_LAG-1 -----AIN--EHRHQINDGAAWTIIST  
 Hs\_RBP-L -----PCPKEANRALLNDSSCWTIIGT  
 Mm\_RBP-L -----LCPKEANRALLNDSSCWTIIGT  
 Dr\_RBP-L -----PCPKESNKVSLNDGSCWTIIGT  
 PcCSL2 -----ISSRRGSSSDSSASSGALWQVDIGETSVWTIVGV  
 CC1G\_03194 -----SRRGSGSDSGASSGALWQVDIGETSVWTIVGT  
 UM06280 OGLSEVEVLRNNGSYANSMTSDSSAAGSIAPGAMWSVEVADSDIWTVVGT  
 RO3G\_07636 VGG-----DAGRCPDQPLNGDCWTEVDVSDSAVWTIVGT  
 RO3G\_06481 TG-----HCRQRLNGDCWAEDVSDVSIWAIVSA  
 RO3G\_11583 -----TGLEGACWTEVDVSDAAVWTIVGT  
 RO3G\_14587 SVS-----SLSSTTSRPHLGAFWSEVDVSDAAVWTLVGT  
 CNBD3370 LSTDSTGNSYEMPSLSSAMSSAEQIQPQRMFWTMDVGDSCVWSIISV  
 SjCSL2 -----EGDNGEYVNLPWSAVWSIITT  
 SPCC1223.13 -----NDPTKATHSVLPWSAMWSIIST  
 Pc6518 -----RDGVRVITDEVDDYLCWTIVGI  
 CC1G\_01706 -----RDGIRVISDEVDDYLCWTIVGI  
 RO3G\_06953 -----IDFTAIQDQQVSDYLCWTIVGI  
 RO3G\_13784 -----PSTLSENGEQVDDALCWTMVGI  
 RO3G\_08863 -----PSRKLEQ-EEIDYLCWTIVGI  
 UM05862 -----PKTVAVPKENAPGIKLV--DEVDDAFVWTIVGV  
 SjCSL1 -----TSIDESGSIRVSDAVCWTIVGI  
 SPCC736.08 -----GNLSSNEMTRVGDSVCWTIVGI  
 CNA01890 AVQTLISWVKAERSEEVSSGADGLDTPDEVRVIVDSVADWMSWTIGGA

Mm\_RBP-Jk DKAEYTFYEG-----MGPVLAPVTPVPPVES-  
 Hs\_RBP-Jk DKAEYTFYEG-----MGPVLAPVTPVPPVES-  
 XI\_SuH DKAEYTFYEG-----MGPINAPVTPVPPVES-  
 Dr\_RBPSUH DKAEYTFYEG-----MGPVPSVTPVPPVVES  
 Hr\_RBP-Jk DKAEYTFCDG-----MGPTSKPVTPVPPVHS-  
 Ci\_SuH DKAEYTFCDG-----MGPTADPVTPVPPVHS-  
 Dm\_SuH DKAEYQFYEG-----MGPVASPVTPVPIVNS-  
 Ce\_LAG-1 DKAEYRFFEA-----MGOVANPISPCPVVGS-  
 Hs\_RBP-L ESVEFSFSTS-----LACTLEPVTPVPLIST-  
 Mm\_RBP-L ESVEFTFSTS-----LACTREPVTPVPLIST-  
 Dr\_RBP-L EVVEYSFSES-----LTGHPATISPVPVISG-  
 PcCSL2 DQVRYNFYVPPVFLDNQSAP-----QTGSFPIPSKPVTPFPNVVKY  
 CC1G\_03194 DQIRYNFYVPPVFLDNQHAQP-----SSLGAFPIPSKQVTPFPGVVKY  
 UM06280 DIARHSFYVPPQIVGGSQPQTNIITDSNIAHLITVPAPAQIPITPIISNM

RO3G\_07636 NSTSFAFWTPPDYS-----KPFDDLQDFPYVDSI  
 RO3G\_06481 DSATFTFWTSPCSIRS-----VTAPITPFPTIMNIKKISN  
 RO3G\_11583 DCASYTFWTPDERT-----MPTAPFPVLHEL  
 RO3G\_14587 DCATYTFWSPFLDDPS-----TPLSTGPFALSHFFTSTN  
 CNBD3370 EQASYTFFTPPITS-----FEFEPVAPFPVANRL  
 SjCSL2 QSVRTMFFDDCSGNDG-----MLSIPSAPIIKFIRMDES  
 SPCC1223.13 QSVRTMFYNEPIHQN-----AFHVVPSPMPFVKFIRLDEN  
 Pc6518 SKFOYTFFDAFG-----HTDPPEAPITPFPTLFTA  
 CC1G\_01706 SKFOYTFFDAFGQ-----NNKIPETPITPFPTLFTA  
 RO3G\_06953 SKFEYEFIQSPIE-----IQQPNSPPPSPRSIIPF  
 RO3G\_13784 HVLEDHFDEPP-----ARPYPV  
 RO3G\_08863 HTLEGEFDEROPT-----TPSSPPPEQPRLVNPYPY  
 UM05862 SRFEYSYFDTSATQAS-----KDEGPCEVPMTPFPLITSM  
 SjCSL1 SHYEFSMMNALNA-----RITQPITPFPPVETSPKY  
 SPCC736.08 SHFDCTMMLPFNQ-----NPVPTVTDYPYIEEPPEY  
 CNA01890 ASFSFTF-----KPTEHNEERVEA

Mm\_RBP-Jk LQLNGG-----GDVAMLELTGONFTPN--LRVWFG  
 Hs\_RBP-Jk LQLNGG-----GDVAMLELTGONFTPN--LRVWFG  
 XI\_SuH LQLNGG-----GDVAMLELTGONFTPN--LRVWFG  
 Dr\_RBPSUH LQLNGG-----GDVAMLELTGONFTPN--LRVWFG  
 Hr\_RBP-Jk LQLNGG-----GDVAMLEVNGENFSPQ--LKVWFG  
 Ci\_SuH LQLNGG-----GDVAMLEVNGECFTSN--LKVWFG  
 Dm\_SuH LNLNGG-----GDVAMLELSGDNFTPH--LQVWFG  
 Ce\_LAG-1 LEVDGH-----GEASRVELHGRDFKPN--LKVWFG  
 Hs\_RBP-L LELSGG-----GDVATLELHGENFHAG--LKVWFG  
 Mm\_RBP-L LELSGG-----GDVATLELHGENFHAG--LKVWFG  
 Dr\_RBP-L LELNGG-----GHVAMLELHGENFSPH--LRVWFG  
 PcCSL2 LPPDRAAEAPKP-CPOSRAMMAKPNPHASKMLTVYGENFSKTDPTVTVFFG  
 CC1G\_03194 LPPDRASEAPKSNCAQSRVLAQKPNPHAAKLLTVYGENFSKTDPTVHVFFG  
 UM06280 VAPQRD-----ARSQPMSAQAGEQSPGLVTILGENFHPN--LFIFFG  
 RO3G\_07636 QS-----LSSTVLSLVGEHFTSD--LTVWFG  
 RO3G\_06481 QS-----LTIFTLILSGENFKHD--VTWVFG  
 RO3G\_11583 TK-----KGKDRLTLTGENLSPD--IEVWFG  
 RO3G\_14587 KL-----DHERFLTMHGENFSRD--LQVWFG  
 CNBD3370 LPANLS-----AEVPSKYAHHYTTTTNMLPLVTLYG  
 SjCSL2 NML-----HVYGVNFTADAQVWIG  
 SPCC1223.13 SMF-----HIYGTGFANDVQIWMA  
 Pc6518 PALATGAGSEPC-----TLDLTVANFFYRHPHTNEQTPLSIYLG  
 CC1G\_01706 PVYRAANN-----TIELTVSNFFYAHPKTRMQTPLDVYLG  
 RO3G\_06953 PTIS-----CIQYIPSTHSIRVPSTEFNYLG  
 RO3G\_13784 VTCMDYQP-----NHTLELFGHNLHPS-----VQPWLG  
 RO3G\_08863 LTHLOQKP-----ETNTLEAIGQHLIQAAPVPRLLEPWLG  
 UM05862 PIYDQTR-----HTLVTTVTNFIVPAAITSDPIPLDVWAG  
 SjCSL1 IGETHS-----LELLVSGCTPKT-----QVWLG  
 SPCC736.08 LESSRS-----LQFKIGGYSVG-----LQIWLG  
 CNA01890 P-----

Mm\_RBP-Jk -DVEAETMYR-----  
 Hs\_RBP-Jk -DVEAETMYR-----  
 XI\_SuH -DVEAETMYR-----  
 Dr\_RBPSUH -DVEADTMYP-----  
 Hr\_RBP-Jk -EVEADTMYP-----  
 Ci\_SuH -EIEADTMFR-----  
 Dm\_SuH -DVEAETMYR-----  
 Ce\_LAG-1 -ATPVETTFR-----  
 Hs\_RBP-L -DVEAETMYR-----



Mm\_RBP-L -DVEAETMYR-----  
Dr\_RBP-L -NMEAETMFR-----  
PcCSL2 SDPSYVEVR-----  
CC1G\_03194 SEPSYVEVR-----  
UM06280 DWKSTHVDVR-----  
RO3G\_07636 DVPSIQTEFK-----  
RO3G\_06481 DIQSPTRFE-----  
RO3G\_11583 DVKSTETEFV-----  
RO3G\_14587 DVKANHTEYR-----  
CNBD3370 KHfVKNVDGR-----  
SjCSL2 ENPCQTYSVTDVEIDEPTLLRGLISTS-----  
SPCC1223.13 YTRCEVKSINAFKPDITLPPDIISDSR-----  
Pc6518 NLGPLRHRINAPYPPGPLTSVSPYVQAIGPDGVQRDEPIGPGVSPLERSP  
CC1G\_01706 SLGPLHTRVYQTSPPGPLTSISPFVPVTP---VEVPPPAGDPNAPANR--  
RO3G\_06953 ROGPLKKK-----  
RO3G\_13784 THGPLNTR-----  
RO3G\_08863 THGPLPTR-----  
UM05862 SLGPLKVSHSTN-----  
SjCSL1 PYGPLAYT-----  
SPCC736.08 VHGPLSYS-----  
CNA01890 -----  
  
Mm\_RBP-Jk -----CGESMLCVVPDISAFREG-----  
Hs\_RBP-Jk -----CGESMLCVVPDISAFREG-----  
XI\_SuH -----CAESMLCVVPDISAFREG-----  
Dr\_RBPSUH -----CGESVLCVVPDISAFREG-----  
Hr\_RBP-Jk -----CEGLLCVVPDISEFREG-----  
Ci\_SuH -----CAEGLLCVVPDISAFREG-----  
Dm\_SuH -----CTETLLCVVPEISQFRGE-----  
Ce\_LAG-1 -----SEESLHCSIPPVSQVRNEQTHW-----  
Hs\_RBP-L -----SPRSLVCVVPDVAAFCS-----  
Mm\_RBP-L -----SPRSLVCVVPDVAAFGSD-----  
Dr\_RBP-L -----SSRSLLCVVPDVSILSGE-----  
PcCSL2 -----CTEVLGCLPPES-NNMKRR-----  
CC1G\_03194 -----CTEVLGCLPPESGHNNKRR-----  
UM06280 -----TROTVCAAPPSFDOGLPRG-----  
RO3G\_07636 -----SSQLSCTVPERHELLDS-----  
RO3G\_06481 -----SDKLISCIVPDYKELQNS-----  
RO3G\_11583 -----SODSVHCKIPFDVAN-----  
RO3G\_14587 -----SRELIICKVPPRHELMVKKVYG-----  
CNBD3370 -----AKHVMYYGDEPARHNEVRCGEVMVASE-----  
SjCSL2 -----HVPPRLHAYLADLADIICQPP-----  
SPCC1223.13 -----FSSRVYACTANLIELICEIP-----  
Pc6518 YLAANPIHTSITVDLPPMADILRSLQEDLEPPRDGAAPQN-----GT  
CC1G\_01706 YVSTGPLHTIVIVELPPLADVIKAMEDEPVPTASDVSGSKPHQESEGOEG  
RO3G\_06953 -----SLIELELPDIQDILLNHHNLLSTQODGS-----  
RO3G\_13784 -----LMQPN-----HWSIALP-----  
RO3G\_08863 -----IASPPETHTPHEIHWSIDLS-----  
UM05862 -----APSTSTSTSTSTSTDPETYITVTLPAIREMIKVALTA  
SjCSL1 -----ROETEKADEN-LLIVSLPYIEVNT-----  
SPCC736.08 -----FTAAADTSTMGTVTLGLSQISYDP-----  
CNA01890 -----  
  
Mm\_RBP-Jk -----WRWVRQPVQVPVTLVRNDGVIYSTSLTFTYTPEPGPRPHCSAAG  
Hs\_RBP-Jk -----WRWVRQPVQVPVTLVRNDGIIYSTSLTFTYTPEPGPRPHCSAAG  
XI\_SuH -----WRWVRQPVQVPVTLVRNDGVIYSTSLTFTYTPEPGPRPHCSAAG  
Dr\_RBPSUH -----WRWVRQPVQVPVTLVRNDGIIYSTALTFTYTPEPGPRPHCSAAG

Hr\_RBP-Jk -----WTWVKQSVQVPINLVRHDGIIYPTNLTFTFTPEPGPRPTLARLL  
 Ci\_SuH -----WKWVKESVQVPINLVRNDGVIYPTNLTFTFTPEPGPRQHCPAAL  
 Dm\_SuH -----WLWVROPTQVPISLVRNDGIIYATGLTFTYTPEPGPRPHCNTQA  
 Ce\_LAG-1 ---MFTNRTTGDVEVPISLVRDDGVVYSSGLTFSYKSLERHGPCRIVSN  
 Hs\_RBP-L -----WRWLRAPITIPMSLVRADGLFYPSAFSFTYTPEYSVRPGHPGVP  
 Mm\_RBP-L -----WRWLRTPITVPVSLLRADALFYPSFSTYTPEYSALPRLPNAQ  
 Dr\_RBP-L -----WRWTRQPITVPLSLIRLDGLIYRSWFSFTYTPEHSASSQTSGSG  
 PcCSL2 -----PIILVRHDGVVFPSTYYP-----  
 CC1G\_03194 -----PIILLRSDGVVFPSTVMYP-----  
 UM06280 -----RVPILIVRHDGVIFPTDFHYQC-----  
 RO3G\_07636 --FATQLDPDTSRHKIPLLLVOEDGIIYNSLLFYSF-----  
 RO3G\_06481 --PVIQVNSSTSERQLPLLLVRTDGVYINPRLFYSF-----  
 RO3G\_11583 ---STTIEQENDHRRIPLLLVRGKGIYKTNLYYIL-----  
 RO3G\_14587 -----DLPILLVRGDGTICKTGKCFSL-----  
 CNBD3370 ----PEPSIKSDRKRPIFLVREDGKVIIPTSLHYPV-----  
 SjCSL2 -----ALSVTTSPPELILVSQQNIIFHSGFTWPIHPHTL-----  
 SPCC1223.13 -----VCMFEPTVELSPILLFOYETLFHSGYKWPLESH-----  
 Pc6518 STDPKLNVAALNGRSLPLLFIIRAADGVGYHSGRAIAVENIFQSLEINGRT  
 CC1G\_01706 VPPPPPPPPSMAGRSLPLLFISSDGVGYHSGRTIACEPIHYSLDLAAMA  
 RO3G\_06953 -----RYLELPILLSKPNDTFIYHSSKLLSFTFLMHSPESGQFS  
 RO3G\_13784 -----PLEQLVGELPLLMVR-QDGLVYHTGKTFSTNSCT-----  
 RO3G\_08863 -----SVKQOTEELPLLLVR-QDGLVYHTGKSLKWDQEGEGWLCK---  
 UM05862 RRSPTSAPIPPHVLLPLIFVNDSDGTAYHSGRHIVCEDLVQLMKATGHD  
 SjCSL1 -----TNPWTTLPLLFRPNGLIYVGKCDVVVR-----  
 SPCC736.08 -----SCAEQKYPLLFVIPGGIVIIKCEILLTSSAFGN-----  
 CNA01890 -----

Mm\_RBP-Jk AILRANSSQVPSNESNTNSEGNYNSS---  
 Hs\_RBP-Jk AILRANSSQVPPNESNTNSEGSYNSS---  
 XI\_SuH AILRANSSLLASNEPNTNSEGSYNSS---  
 Dr\_RBPSUH AILRANSNSSTSPSSSSSASSLLQDGA---  
 Hr\_RBP-Jk YISSTPR-----DR---  
 Ci\_SuH NILHGSKRPSASMPPTPVSGSEDSNDP---  
 Dm\_SuH EDVMRARQNNNNNNITSISNNNNNSGL---  
 Ce\_LAG-1 Y-----  
 Hs\_RBP-L EPATDADALLESIHQEFTRTNFHF-----  
 Mm\_RBP-L EPAPDADTLLESIHHEFTRTNFHF-----  
 Dr\_RBP-L ERSADSDALIDIHQEFTRTNFHF-----  
 PcCSL2 -----  
 CC1G\_03194 -----  
 UM06280 -----  
 RO3G\_07636 -----  
 RO3G\_06481 -----  
 RO3G\_11583 -----  
 RO3G\_14587 -----  
 CNBD3370 -----  
 SjCSL2 -----  
 SPCC1223.13 -----  
 Pc6518 -----DDQAWLAQAQGMQGLEAMS-----  
 CC1G\_01706 PNGPNGVDPOWLAAAQAAAAADGMT-----  
 RO3G\_06953 -----LT-----  
 RO3G\_13784 -----  
 RO3G\_08863 -----  
 UM05862 TTADVLKQLGFGGLGDLAGPPNDGAWSVRII  
 SjCSL1 -----  
 SPCC736.08 -----  
 CNA01890 -----



## **Cbf11 and Cbf12, the fission yeast CSL proteins, play opposing roles in cell adhesion and coordination of cell and nuclear division**

Martin Převorovský<sup>a</sup>, Tomáš Groušl<sup>a,1</sup>, Jana Staňurová<sup>a</sup>, Jan Rynesš<sup>a,2</sup>, Wolfgang Nellen<sup>b</sup>,  
František Půta<sup>a</sup>, Petr Folk<sup>a,\*</sup>

<sup>a</sup>Department of Cell Biology, Faculty of Science, Charles University in Prague, Viničná 7,  
128 43, Prague 2, Czech Republic

<sup>b</sup>Department of Genetics, Kassel University, Heinrich Plett Strasse 40, 34132 Kassel,  
Germany

<sup>1</sup>present address: Institute of Microbiology, Academy of Sciences of the Czech Republic,  
Vítězská 1083, 142 20, Prague 4, Czech Republic

<sup>2</sup>present address: Department of Genetics, Biology Center, Academy of Sciences of the Czech  
Republic, Branišovská 31, 370 05, České Budějovice, Czech Republic

\*Corresponding author: Petr Folk, Department of Cell Biology, Faculty of Science, Charles  
University in Prague, Viničná 7, 128 43, Prague 2, Czech Republic

phone: +420 22195 1765

fax: +420 22195 1758

e-mail: [folk@natur.cuni.cz](mailto:folk@natur.cuni.cz)

## Abstract

The CSL (CBF1/RBP-J $\kappa$ /Suppressor of Hairless/LAG-1) family is comprised of transcription factors essential for metazoan development, mostly due to their involvement in the Notch receptor signaling pathway. Recently, we identified two novel classes of CSL genes in the genomes of several fungal species, organisms lacking the Notch pathway. In this study, we characterized experimentally *cbf11*<sup>+</sup> and *cbf12*<sup>+</sup>, the two CSL genes of *Schizosaccharomyces pombe*, in order to elucidate the CSL function in fungi. We provide evidence supporting their identity as genuine CSL genes. Both *cbf11*<sup>+</sup> and *cbf12*<sup>+</sup> are non-essential; they have distinct expression profiles and code for nuclear proteins with transcription activation potential. Significantly, we demonstrated that Cbf11 recognizes specifically the canonical CSL response element GTG<sup>A</sup>/<sub>G</sub>GAA *in vitro*. The deletion of *cbf11*<sup>+</sup> is associated with growth phenotypes and altered colony morphology. Furthermore, we found that Cbf11 and Cbf12 play opposite roles in cell adhesion, nuclear and cell division and their coordination. Disturbed balance of the two CSL proteins leads to cell separation defects (sep phenotype), cut phenotype, and high-frequency diploidization in heterothallic strains. Our data show that CSL proteins operate in an organism predating the Notch pathway, which should be of relevance to the understanding of (Notch-independent) CSL functions in metazoans.

## Keywords

CSL family, fission yeast, adhesion, cell division, nuclear division, high-frequency diploidization

## Introduction

The CSL (CBF1/RBP-J $\kappa$ /Suppressor of Hairless/LAG-1) transcription factors function as key regulators of various aspects of metazoan developmental and cell fate decision processes, which include somitogenesis, neurogenesis, tissue boundaries specification, or blood cells differentiation. The members of this family, which displays a high degree of evolutionary conservation, specifically recognize and bind to a rather tightly defined response element (GTG<sup>A</sup>/<sub>G</sub>GAA) in the promoters of the genes they regulate [1]. Depending on the signaling of the transmembrane receptor Notch, they are capable of either repressing or activating transcription [2-5]. Deregulation of the Notch-driven CSL signaling results in severe developmental defects and several types of cancer [5-7]. There are examples of viruses encoding protein factors that make use of CSL proteins to promote viral replication [1].

Despite the fact that CSL proteins are essential for proper development and functioning of metazoan organisms, they seem to be dispensable at the cellular level as CSL knock-out cell lines have been established and were not found to have any obvious phenotypes. Remarkably, at the organismal level, CSL deletion leads to pleiotropic defects that are more severe than those of Notch mutants [2,8]. This strongly suggests that CSL proteins engage in other processes and pathways beyond the Notch framework and, indeed, several recent studies have shown that the “classical” Notch-pathway CSL family members (RBP-J $\kappa$  type) may act independently of Notch [9-12]. Furthermore, most vertebrate genomes harbor a second CSL paralog, dubbed RBP-L, which has a distinct expression profile and likely functions in a completely Notch-independent manner [13,14]. It is worth noticing that studying the roles of CSL family in metazoa suffers from severe limitations posed by the extent and pleiotropy of defects elicited by manipulations of the CSL genes [5].

We have recently identified and described a number of novel putative CSL family members in the genomes of several species of fungi, bringing the evolutionary origin of this family back to the last common ancestor of metazoa and fungi [15]. We have shown *in silico* that, in addition to the CSL genes found in metazoa (class M), two fungi-specific classes, F1 and F2, exist in the genomes of *Schizosaccharomyces pombe* and other representatives of Taphrinomycotina, zygomycetes and basidiomycetes. It is likely that this list will expand as more fungal genomes are sequenced, though it is clear now that CSL genes were lost in the lineage of *Saccharomyces cerevisiae* and related yeast genera. Typically, there are two CSL paralogs per genome in the above-mentioned fungi, each belonging to a different class. The fungal CSL genes code for proteins with a domain organization similar to the metazoan family members. We have found a striking degree of amino acid sequence conservation of the novel CSL proteins, which is mostly confined to regions important for DNA binding [15,16]. Significantly, fungi lack the Notch pathway and most of the known metazoan CSL-interacting partners, thus the role of the CSL family in fungi is unclear.

In the present study, we have characterized experimentally *cbf11*<sup>+</sup> and *cbf12*<sup>+</sup>, the two novel CSL genes found in the model organism *S. pombe* [15]. We sought to prove their identity as genuine CSL family members, and to elucidate the CSL function in fungi.

## **Materials and methods**

### **Strains, media and cultivation**

The fission yeast strains used in this study are listed in Table 1. *S. pombe* was grown in the YES complex medium at 30°C in most experiments [17]. The selection of gene knock-out (KO)/knock-in (KI) cells was carried out using the appropriate antibiotics as described [18,19]. For *nmt41*-driven plasmid expression, the cells were grown in the MB minimal

medium (Formedium) with the appropriate auxotrophic selection, either with thiamine added (5 µg/ml; repressed state) or without thiamine (induction) [20]. Phloxin B, a red dye staining preferentially dead cells, was added to the media at 5 mg/l where required. This dye allows for discerning haploid and diploid colonies, as the latter contain more dead cells and stain darker. To screen for KO sensitivity/resistance phenotypes, YES medium with the addition of the following stressors was used: EGTA 5-40 mM, hydroxyurea 7.5-11.25 mM, KCl 0.4-1.2 M, sorbitol 1.2-2 M, SDS 0.003-0.01%, H<sub>2</sub>O<sub>2</sub> 1-10 mM, calcofluor 100-800 µg/ml, cycloheximide 10-45 µg/ml, latrunculin A 0.25-1 µM, carbendazim 10-25 µg/ml. We have also tested the KOs for growth on MB minimal plates and on YES at 19°C and 36°C, respectively. In this article, we discriminate between “giant colonies” or “spots”, inoculated by spotting cell suspension on a plate, “monocolonies”, originating from a single cell, and “patches”, which represent cell mass grown after patching the cells on a plate.

Cell transformations were carried out by the lithium-acetate method [18]. For diploid construction, the cells were mated on MES plates (3% malt extract, auxotrophic supplements, pH adjusted to 5.5), diploids were selected by adenine prototrophy on MB(w/o adenine) plates, verified on YES(w/o adenine) and selected white colonies were kept on YES.

Sporulation was induced either in liquid culture (cells were grown in YES to OD 0.5, washed twice in MB(w/o adenine and thiamine) and cultured further in MB(w/o adenine and thiamine)) or on MB(w/o adenine and thiamine) plates. For the yeast two-hybrid analysis two different systems, both by Clontech, were used according to the manufacturer’s instructions – the MATCHMAKER Two-Hybrid System 2 (Gal4-based; the CG-1945 *S. cerevisiae* reporter strain) and the MATCHMAKER LexA Two-Hybrid System (the EGY48 *S. cerevisiae* reporter strain).

### **Adhesion and flocculation**



Flocculation assays were performed as described in [21]. Cells from liquid culture were harvested, washed in 10 mM EDTA, and then in deionized water. Flocculation was initiated by the addition of 10 mM CaCl<sub>2</sub> in the presence or absence of 100 mM sugars. Cultures were transferred to a Petri dish for imaging. To test their adhesion, all strains were grown on a single agar plate for 7-14 days, and then washed carefully and evenly with a stream of water [22].

### **Constructs and plasmids**

Chromosomal DNA of the FY254 strain was used for cloning both *cbf11*<sup>+</sup> and *cbf12*<sup>+</sup> genes by PCR; all cloned PCR products were verified by sequencing (see Table 2 for a complete list of primers used in this study). Since *cbf11*<sup>+</sup> contains a short intron that is not spliced out effectively in *S. cerevisiae*, its cDNA was first constructed by PCR-mediated exon merging [23] and sequenced. The *cbf11*<sup>+</sup> cDNA was prepared with two primer sets in two different reading frames to fit into both the *S. cerevisiae* two-hybrid (primers if03, mp25, mp26, if04; inserted into the *Bam*HI site of the pACT2-AD, pAS2-1 and pLexA vectors) and the *S. pombe* expression vectors (primers mp20, mp25, mp26, if04; inserted using *Nde*I/*Bam*HI into the pREP42-MHN and pREP42-EGFPN vectors) [20]. The *cbf12*<sup>+</sup> ORF was cloned using the primers if05 and if06, and inserted in between the *Nde*I/*Bam*HI sites of pREP42-MHN, pREP42-EGFPN [20] and pAS2-1 vectors, and in between the *Eco*RI/*Bam*HI sites of pLexA. The  $\Delta$ *cbf11* strain (CBF11-KO), obtained from Dr. Anabelle Decottignies [24], was crossed with the PN558 parent strain and sporulated to generate  $\Delta$ *cbf11* strains of the opposite mating type (MP05, MP06). The  $\Delta$ *cbf12* strains (MP01, MP03) were generated from the parent strains (PN558, PN559) by homologous recombination using the appropriate pCloneNat1 targeting vector obtained from Dr. Juraj Gregaň [25]. The  $\Delta$ *cbf11*  $\Delta$ *cbf12* double KOs (MP07, MP09 and MP10) were constructed by crossing the CBF11-KO and MP01 single KOs, and

subsequent tetrad analysis. The *cbf12<sup>+</sup>::EGFP* KI strain (MP12) was prepared by PCR-mediated one-step gene-tagging as described [18] using the primers mp43 and mp44, and the pFA6a-GFP(S65T)-kanMX6 vector as template. All chromosomal integrations and deletions were verified by PCR (combined with restriction analysis) using a three-primer system based on [18]: the mp27, mp28 and kan-rev primers for the *cbf11<sup>+</sup>* KO, the primers mp31, mp32 and mp33 for the *cbf12<sup>+</sup>* KO, and the primers mp31, mp32 and kan-rev for the *cbf12<sup>+</sup>::EGFP* KI. The mating type of the newly constructed strains was determined by PCR as described [26].

### **RNA isolation and quantitative RT-PCR**

Total fission yeast RNA was extracted using the FastPrep instrument (Q-BIOgene) and the RNeasy Mini kit supplemented with the RNase-free DNase set (Qiagen). 2 µg of total RNA were reverse-transcribed with the RevertAid First Strand cDNA Synthesis Kit (Fermentas) and an oligo(dT) primer (total reaction volume 20 µl), 1 µl of cDNA was then used for amplification. Approximately 200 bp of the 3'-regions of *cbf11<sup>+</sup>* (mp35, mp45), *cbf12<sup>+</sup>* (mp36, mp46) and *act1<sup>+</sup>* (mp37, mp38; normalization control) were amplified in separate tubes using the primers indicated (see Table 2 for details). The efficiency and linear range of the quantitative real-time RT-PCR (qRT-PCR) was tested according to [27]. The qRT-PCR analysis was performed using the iQ SYBR Green Supermix (Bio-Rad) in the RotorGene 2000 system (Corbett Research) with 0.3 µM each primer and the following program: 95°C for 3 min, 40× (95°C for 30 sec; 53°C for 30 sec; 72°C for 30 sec). The reactions were run in triplicates and the data were analyzed in the Q-Gene software application [28].

### **Microscopy and imaging**

Fluorescent images were acquired using either a Leica TCS SP2 confocal microscope or an Olympus IX81 epifluorescence microscope. High-magnification colony images were taken by a Hitachi HV-C20 camera connected to a Navitar Zoom 6000 optical system and the NIS Elements imaging software (Nikon).

### **Gelshift experiments**

*S. pombe* lysates for gelshifts were prepared by vortexing the cells with HCl-washed glass beads in the lysis buffer (25 mM Hepes pH 7.6, 0.1 mM EDTA pH 8, 150 mM KCl, 0.1% Triton X-100, 25% glycerol, 2 mM DTT, 1/100 volume of Protease Inhibitor Mix FY (Serva)) at 4°C. The lysates were centrifuged for 20 min at 20 000 g, 4°C and the supernatants were used further. MycHis-tagged Cbf11 and Cbf12 proteins used for the shifts were purified from *S. pombe* by the TALON affinity chromatography (BD Biosciences) according to the manufacturer's instructions. A panel of double stranded DNA oligonucleotide probes containing either the CSL response element or a mutated/scrambled control (Fig. 2A) was synthesized and terminally labeled by incubation with [ $\gamma$ -<sup>32</sup>P]ATP and the T4 polynucleotide kinase (Fermentas). The shift reactions containing the shift buffer (25 mM Hepes pH 7.6, 34 mM KCl, 5 mM MgCl<sub>2</sub>), 2 ng of radioactively labeled probe, up to 40 ng of unlabeled competitor, 1  $\mu$ g (for purified proteins) or 20  $\mu$ g (for cell lysates) of carrier sonicated salmon sperm DNA and either the *S. pombe* cell extract (up to 100  $\mu$ g) or the purified Cbf11/Cbf12 proteins were prepared and incubated on ice for 20 min. The shift reactions were then resolved on a native 5% polyacrylamide gel in 0.5 $\times$  TBE at room temperature. The gels were visualized with a FUJIFILM BAS Reader Model 1500 imager.

### **Flow cytometry**

Exponentially growing cells were fixed with 70% ethanol, treated with RNase and stained with propidium iodide (4 µg/ml) as described [29]. DNA content was measured using the Becton Dickinson LSR II instrument; at least 20 000 cells were measured for each sample. Fluorescence intensity histograms were produced in the DiVa software. To assess the ploidy accurately, only cells having single nuclei (as judged by their signal amplitude/width ratio) were included in the final analysis. Thus, the confounding bi-nucleate cells undergoing cytokinesis, as well as spurious doublets and pseudohyphae ( $\Delta cbf11$  strains) were removed by gating.

## Results

### ***cbf11*<sup>+</sup> and *cbf12*<sup>+</sup>, and their expression profiles**

Using bioinformatics, we found putative CSL family members in several fungal species and defined two novel, fungi-specific classes of CSL genes [15]. We have decided to characterize experimentally in detail the two *S. pombe* CSL paralogs, the class F1 SPCC736.08 gene [GenBank:2539560] and the class F2 SPCC1223.13 gene [GenBank:2539119], in order to confirm their CSL identity and to gain insight into their roles in a unicellular organism. The products of both genes contain domains and sequence motifs typical of CSL proteins: the DNA-binding RHR-N (Rel-homology region) and BTD (beta-trefoil domain; appearing exclusively in CSL proteins) domains [Pfam:PF09271, Pfam:PF09270] [16] can be identified with confidence, and their organization follows that of metazoan CSL proteins (Fig. 1A). Because of their similarity to the mammalian CBF1 (C-promoter element-binding factor) family founding member, we have named the fission yeast CSL genes *cbf11*<sup>+</sup> (SPCC736.08) and *cbf12*<sup>+</sup> (SPCC1223.13), respectively.

First, we have used quantitative real-time RT-PCR (qRT-PCR) to map their expression levels at various stages of the fission yeast life cycle. We have analyzed RNA samples from haploid cells at various points of the growth curve, and from diploid cells under vegetative growth conditions or after induction of sporulation. The results are summarized in Fig. 1B. Both genes are expressed at low levels, with the *cbf11*<sup>+</sup> mRNA being roughly 50× less abundant, and the *cbf12*<sup>+</sup> mRNA being roughly 170× less abundant than the *act1*<sup>+</sup> [GenBank:2540051] normalization control during the early logarithmic phase. There is an approximately two-fold downregulation of the *act1*<sup>+</sup> normalization control during sporulation [30]. For this reason, the sporulation samples are not directly comparable with the others, although comparisons between *cbf11*<sup>+</sup> and *cbf12*<sup>+</sup> mRNA levels under these conditions can still be made.

The expression of *cbf11*<sup>+</sup> seems to be fairly constant throughout the growth phases of haploid cells and similar mRNA levels were found in vegetative diploid cells. The less abundant *cbf12*<sup>+</sup> mRNA has a more variable profile with a marked increase (up to the levels of *cbf11*<sup>+</sup>) as the cells enter the stationary phase, with a statistically significant peak at the late stationary phase (two-tailed independent Student's t-test, p = 0.05). A similar increase in the *cbf12*<sup>+</sup> mRNA levels was found in the sporulating cells. The *cbf12*<sup>+</sup> expression in vegetative diploid cells is similar to early-log phase haploids. The differences in expression patterns between the two CSL paralogs strongly suggest they play distinct roles in fission yeast.

### **Subcellular localization and transcription activation potential of Cbf11 and Cbf12**

We have carried out localization analyses with EGFP-tagged Cbf11 and Cbf12 proteins in both living and fixed *S. pombe* cells using confocal microscopy (Fig. 1C). First, we used plasmid expression of N-terminally tagged proteins driven by the medium-strength *nmt41* promoter. Both EGFP-Cbf11 and EGFP-Cbf12 displayed very similar diffuse nuclear localization, although some fine “structuring” could be seen on 3D reconstructions from Z-

axis optical sections (Fig. 1C and data not shown). Both CSL proteins seemed to be excluded from the nucleolus. An identical localization pattern was seen when fixed fission yeast cells were used (data not shown). To verify that the localization observed is not an artifact due to overexpression or N-terminal tagging, we attempted to construct knock-in (KI) strains to express C-terminally labeled fusion proteins at physiological levels. We have succeeded in the case of *cbf12*<sup>+</sup>, however, we repeatedly failed to get either an EGFP (kan<sup>R</sup>) or EYFP (nat<sup>R</sup>) KI strain for *cbf11*<sup>+</sup>, possibly due to locus-specific constraints. The subcellular distribution of Cbf12-EGFP in the KI strain was found to be the same as in the case of overexpression from a plasmid, although, as expected, the signal was considerably weaker (Fig. 1C, right panel). As *cbf12*<sup>+</sup> is more expressed at the stationary phase than at the log phase (Fig. 1B), we compared the Cbf12-EGFP localization pattern between these two stages but found no differences (data not shown).

While testing the usability of yeast two-hybrid systems for studying Cbf11 and Cbf12, we found that both fission yeast CSL proteins have a potential to activate transcription. In the two-hybrid systems used (LexA and Gal4-based) the CSL fusions either displayed toxicity in several *S. cerevisiae* reporter strains or triggered strong activation of the reporter genes without a requirement for other fusion-protein partners (data not shown). We judge this capacity as indicative of the presence of an intrinsic activation domain in both Cbf11 and Cbf12.

### **DNA binding properties of Cbf11**

We previously showed that the fungal CSL proteins contain all the amino acid motifs and residues necessary for sequence-specific binding of DNA [15,16]. We have therefore conducted gelshift experiments in order to characterize the DNA binding properties of Cbf11 and Cbf12 *in vitro*. First, we sought for a CSL-type DNA binding activity in the lysates of

wild-type (WT) *S. pombe*. We prepared a panel of double stranded DNA oligonucleotide probes containing CSL binding sites with their respective flanking regions from various mammalian, insect and viral CSL-responsive promoters (see Fig. 2A for details). We used two mutated versions of the Kaposi's Sarcoma-Associated Herpesvirus-derived probe as negative controls. The first control probe, termed MUT, contained a single G<sub>5</sub>→C substitution at a position that was shown to be critical for metazoan CSL binding [12]. The CSL response element was completely scrambled in the second control probe, DEL. As shown in Fig. 2B (a representative experiment with the RBP probe), we were indeed able to see a specific CSL-like DNA binding activity. This activity was concentration dependent and could be competed with a relatively low excess of unlabeled probe. This is indicative of a very specific binding, given the presence of great excess (10 000×) of carrier DNA in the gelshift reaction. Using protein extracts from deletion strains (see below) as controls, we have ascribed this DNA binding activity to Cbf11 (Fig. 2B, compare the left and the right halves of the gel). While we could observe the Cbf11 binding of varying strength to all RBP, KSHV, HES and m8 probes, there was clearly no binding to either the MUT or DEL control probes in the yeast lysates (data not shown), further confirming that Cbf11 recognizes and binds to the CSL response element in a highly specific manner. So far, we failed to see any DNA binding activity attributable to Cbf12, despite the fact that lysates prepared from cells grown under several different growth conditions and convenient deletion strain controls ( $\Delta cbf11$  and  $\Delta cbf11 \Delta cbf12$  KOs, see below) were used.

We have then repeated the gelshifts with MycHis-tagged Cbf11 protein affinity-purified from *S. pombe*. The results shown in Fig. 2C further supported our data from cell lysates and documented the varying affinity of Cbf11 for the respective probes. Again, there was no binding to either of the negative control probes. To demonstrate that the binding of Cbf11 to DNA is direct, we expressed the protein in *E. coli* and then carried out gelshifts with Cbf11-

containing bacterial lysates in settings similar to these in Fig. 2B. We were able to recapitulate the results obtained with *S. pombe* proteins/lysates for the RBP and KSHV probes, the ones that were bound most strongly by Cbf11 (data not shown). When *E. coli* cultures expressing Cbf12 were used, no binding could be detected, similar to the fission yeast lysates data. Taken together, at least one of the *S. pombe* CSL proteins, Cbf11, is capable of binding DNA in a highly specific way like the classical metazoan family members (class M) – a critical finding that establishes the fungi-specific CSL proteins as genuine family members.

### **Growth phenotypes of the $\Delta cbf11$ and $\Delta cbf12$ single and double deletion strains**

We have obtained the  $\Delta cbf11$  strain [24] and crossed the deletion into cells of the opposite mating type to allow for testing of any possible mating type-specific functions. The  $\Delta cbf12$  deletion strain had been prepared previously in  $h^{90}$  cells [25], and we used the available targeting vector to construct the KO in  $h^-$  and  $h^+$  cells to prevent possible phenotype ambiguities arising from mating-type switching. Both  $\Delta cbf11$  and  $\Delta cbf12$  strains were viable, which is in accordance with the available published data [24,25]. Despite their potentially different DNA-binding properties (see above), the fission yeast CSL paralogs could still have (partially) redundant functions and compensate for each other in the respective mutants. To account for this possibility, we have also constructed haploid  $\Delta cbf11 \Delta cbf12$  double KOs by crossing. These strains were also viable, thus, as it was found earlier, neither the  $cbf11^+$  nor the  $cbf12^+$  gene is essential under normal growth conditions, and *S. pombe* cells can sustain even the loss of the entire CSL gene family.

We then subjected our panel of KO strains (see Table 1) to a series of standard growth and sensitivity/resistance tests (see Materials and methods). Both single deletion mutants and the double KO were found to be capable of conjugation and spore formation (data not shown). All mutants exerted viability comparable to WT controls when serial dilutions were spotted



on minimal medium plates, rich medium plates under the conditions of heat, osmotic, salt and oxidative stress or in the presence of various substances affecting DNA replication, calcium signaling, translation, or damaging the plasma membrane, cell wall and cytoskeleton (data not shown; see Materials and methods for the list of treatments used).

However, the growth of the strains lacking *cbf11*<sup>+</sup> was found to be impaired on solid media at 19°C (Fig. 3A). To investigate the growth defect with a finer resolution, we measured growth curves in shaken cultures at 30°C. We found that while the  $\Delta$ *cbf12* strain was indistinguishable from the WT control, the growth of the  $\Delta$ *cbf11* cells was retarded as compared with WT. Intriguingly, there was further growth impairment when the *cbf11*<sup>+</sup> and *cbf12*<sup>+</sup> deletions were combined in one strain, a synthetic effect indicative of some crosstalk between Cbf11 and Cbf12 (Fig. 3B). We then performed a rescue experiment of the  $\Delta$ *cbf11*-associated growth phenotype. WT and double KO strains were transformed with either an empty vector or plasmids encoding *cbf11*<sup>+</sup> or *cbf12*<sup>+</sup>, respectively, and monoclonies were grown for 7 days on MB media containing phloxin B. The overexpression of *cbf11*<sup>+</sup> in the double KO background resulted in colony size and phloxin B staining (i.e., health status) similar to WT. In contrast, the high dosage of *cbf12*<sup>+</sup> seemed to be toxic for the cells, as the Cbf12-overproducing colonies were even darker than the double KO colonies transformed with the empty vector only (Fig. 3C). Indeed, when assayed by plating serial dilutions of cultures on inducing plates, the overexpression of *cbf12*<sup>+</sup> in WT cells resulted in decreased cell viability as well (data not shown).

### **Altered colony morphology of strains lacking *cbf11*<sup>+</sup>**

While carrying out the spot tests described above, we noticed that there were some macroscopically visible differences between the colonies of WT strains and strains carrying the deletion of *cbf11*<sup>+</sup> when grown on the rich YES medium. Namely, when illuminated, the

surface of the spotted WT and  $\Delta cbf12$  giant colonies appeared dim, which was in sharp contrast to the  $\Delta cbf11$  and  $\Delta cbf11 \Delta cbf12$  giant colonies, the surface of which reflected much more light, thus displaying a “shiny” phenotype (Fig. 4A). When the colonies were submerged gently in water, a thin floating layer of a highly reflective material came off the colony surface (Fig. 4B). Thus, it is likely that the  $\Delta cbf11$ -associated “shiny” appearance is caused by overproduction of some extracellular material that contains hydrophobic and reflective compounds. Interestingly, this phenotype was never detected on minimal media. When observed under higher magnification, the surface of the “shiny” spots appeared as a complicated network of irregular grooves. By contrast, the surface of the dim giant colonies was found to be smooth (Fig. 4C). We obtained similar results when we examined the surface morphology of cell patches (data not shown). The “shiny” phenotype in the  $\Delta cbf11 \Delta cbf12$  giant colonies could be rescued by mild overexpression of  $cbf11^+$  but not  $cbf12^+$  from a plasmid (Fig. 4D). Since the “shiny” phenotype does not appear on MB minimal media, the spots had to be grown on YES plates. This medium neither selects for the retention of the plasmids, nor supports high expression from the thiamine-repressible *nmt1*-derived promoters due to its natural thiamine content. Nevertheless, the majority of the  $cbf11^+$  expressing double KO colony surface was not “shiny”, and the mutant phenotype only manifested itself at the spot periphery, where the plasmid had likely been already lost.

We next analyzed the morphology of  $\Delta cbf11$  monocolonies. Since the grooves on giant colonies and patches formed in their older, central parts, containing an increased proportion of starving and old/dying cells, the monocolonies were grown for a prolonged period of time (14 days). To better visualize the physiological state of cells in different regions of the colonies, we added phloxin B into the medium. The phenotypes observed in these settings were in accord with the results obtained with giant colonies and patches (Fig. 4E). While the WT and  $\Delta cbf12$  cells formed regularly shaped, evenly stained round colonies, there were significant

morphology alterations in the case of the  $\Delta cbf11$  single and  $\Delta cbf11 \Delta cbf12$  double KO strains. These strains formed colonies of irregular shape with (sometimes protruding) sectors of darker-staining cell clones. Furthermore, the rim of these colonies was covered with a network of “shiny” grooves highly similar to those observed on the surface of the giant colonies. Again, the sectoring phenotype could be rescued in the  $\Delta cbf11 \Delta cbf12$  background by introducing functional  $cbf11^+$  on a plasmid, and also the overall colony morphology more closely resembled WT (Fig. 4F). Any potential cross-complementation of the sectoring by  $cbf12^+$  overexpression could not be determined as whole colonies stained intensely dark under these conditions, likely due to the toxic effect of excess Cbf12 (data not shown; see also Fig. 3C).

Taken together, the absence of  $cbf11^+$  leads to various marked changes in the morphology of multicellular structures formed by fission yeast cells (colonies, patches), suggesting possible involvement of Cbf11 in cell-cell contact formation.

### ***cbf11<sup>+</sup>* and *cbf12<sup>+</sup>* have opposing roles in cell adhesion**

Our data on Cbf11 affecting colony morphology prompted us to test the effects of fission yeast CSL genes deletion and overexpression on cell adhesive properties. First, we performed washing assays (see Materials and methods) with spotted giant colonies of the respective KO strains grown on YES plates. Significantly, cells lacking  $cbf12^+$  consistently showed decreased adhesion to the agar plate. While some cell mass of the WT and  $\Delta cbf11$  spots remained attached to the agar surface, the  $\Delta cbf12$  and  $\Delta cbf11 \Delta cbf12$  spots were washed off completely. This observation was also documented microscopically (Fig. 5A). We performed a rescue experiment of this phenotype in the  $\Delta cbf11 \Delta cbf12$  background. Cells were grown on inducing MB plates and the washing assay was carried out. On this media type, the overall cell adhesion was higher as compared with YES, and not even the double KO spots were

washed off completely. As expected, adhesion could be restored by the overexpression of *cbf12*<sup>+</sup> from a plasmid (Fig. 5B). Surprisingly, the residual adhesion displayed by the  $\Delta cbf11$   $\Delta cbf12$  cells on this type of media was abolished by the overexpression of *cbf11*<sup>+</sup>. Such a result suggested that the two CSL paralogs may have opposing functions in adhesion. Indeed, when we examined the adhesive properties of the  $\Delta cbf11$  single KO strain in more detail, we found that these cells displayed a higher degree of adhesion than the WT control, and the re-introduction of *cbf11*<sup>+</sup> neutralized this increase (Fig. 5C).

Next, we tested the CSL influence on adhesion in liquid culture. Unlike the WT control, the  $\Delta cbf11$  KO strain displayed flocculation (macroscopically visible cell aggregate formation; [31]) when grown to the stationary phase (Fig. 6A). There was also apparent a marked increase of adherence of these cells to the glass cultivation flask walls as compared with WT (data not shown). Significantly, plasmid-driven overexpression of *cbf11*<sup>+</sup> abolished the flocculation (and the adhesion to the cultivation flasks) of the  $\Delta cbf11$  strain. When *cbf12*<sup>+</sup> was overexpressed in a WT strain, it triggered a similar flocculation phenotype already in a logarithmic-phase culture (Fig. 6B). No flocculation was observed for the WT control under these conditions. As expected, the Cbf12-induced aggregation phenotype was strongly potentiated in the  $\Delta cbf11$  background (earlier onset, larger aggregates), lending further support to our hypothesis that Cbf11 and Cbf12 influence cell adhesion in an antagonistic manner. Thus, we conclude that Cbf11 behaves as a negative regulator and Cbf12 as a positive regulator of cell adhesion in fission yeast.

While flocculation is mainly mannose-dependent in *S. cerevisiae* [32], in *S. pombe* it was found to be mediated rather by cell-surface galactosyl residues [21]. In agreement with that, using a sugar competition assay we determined that Cbf12-induced flocculation could be abrogated by the addition of galactose but not mannose to the cell suspension (Fig. 6C). Since

agar is a galactose polymer, it is likely that both the cell-cell and cell-agar surface adhesion changes caused by CSL manipulation are of the same, galactose-dependent type.

### **Loss of *cbf11*<sup>+</sup> or overexpression of *cbf12*<sup>+</sup> result in high-frequency spontaneous diploidization**

Next, we turned to the second prominent phenotype displayed by colonies lacking *cbf11*<sup>+</sup>, that is, the dark sectoring of phloxin B staining (see Fig. 4E). In *S. pombe*, such dark-red staining is typical of diploid colonies [33], suggesting that the dark sectors might represent diploid clones arising in the mutant populations. We employed flow cytometry to determine the ploidy of the respective strains (see Materials and methods). Normally, most fission yeast cells in an exponentially growing liquid culture are in the G<sub>2</sub> phase, with haploids showing a single 2C peak and diploids showing a single 4C peak on FACS histograms [29]. The profile of the haploid  $\Delta cbf12$  culture was indistinguishable from WT. In contrast to that, the  $\Delta cbf11$  and  $\Delta cbf11 \Delta cbf12$  haploid strains (cultures inoculated from patches) contained a significant fraction of cells with >2C DNA content (Fig. 7). When cells were taken separately from the light and dark sectors of the phloxin B-stained monocolonies and cultured further, the former still comprised a mixture giving two peaks on FACS, while the latter sorted as diploids, suggesting that the shift to a higher DNA content is stable. As suspected, the overproduction of Cbf12 in a WT background also resulted in the appearance of a fraction of likely diploid cells. We did not see any further increase in the >2C cell fraction in a  $\Delta cbf11$  strain after transforming it with a plasmid encoding Cbf12 (data not shown). Nevertheless, the two paralogs again seem to act in opposite directions, this time affecting the maintenance of genome ploidy.

## **Loss of *cbf11*<sup>+</sup> or overexpression of *cbf12*<sup>+</sup> lead to multiple defects in cell and nuclear division**

Importantly, all strains used in our analyses were heterothallic, either *h*<sup>+</sup> or *h*<sup>-</sup>, unable to switch their mating types, and thus unable to self-conjugate. Nevertheless, stable, non-sporulating homozygous diploids can rarely appear even in WT strains as a result of failures at various points of cell cycle regulation [34,35]. Therefore, we performed microscopic analysis of the CSL mutant and overexpressor strains in search for hints as to the mechanism of the high-frequency diploid formation observed by FACS (see Fig. 7).

We analyzed both live and fixed cells from exponentially growing cultures. While there were virtually no aberrations present in the WT controls and very little in the  $\Delta cbf12$  strain, the microscopy revealed a surprising range of defects exerted by either the *cbf11*<sup>+</sup>-lacking or *cbf12*<sup>+</sup>-overexpressing cells (Fig. 8). The mutant phenotypes could generally be described as consequences of cell and/or nuclear division misregulation and their respective penetrances varied between particular genotypes, affecting usually only a minor fraction of the cells (Table 3). These phenotypes, generally more severe in the double KO, included a high degree of heterogeneity in both cell size and shape, with frequent large cells with diploid-like nuclei that presumably corresponded to the diploid fraction observed by FACS (both  $\Delta cbf11$  and *cbf12*<sup>+</sup>-overexpression). Extremely large cells with giant, likely polyploid nuclei were seen too, although rarely (Fig. 8E, right panel). We also noticed the so-called sep phenotype [36] (both  $\Delta cbf11$  and *cbf12*<sup>+</sup>-overexpression), i.e., various septation defects comprising the formation of multiple septa in a single cell, aberrant septum structure, a failure of daughter cells to separate after septum formation, and pseudohyphal (often multipolar) growth. Short filaments of up to six uninucleate compartments were often seen particularly when *cbf12*<sup>+</sup> was overexpressed in a  $\Delta cbf11$  strain (Fig. 8H). Finally, the deletion of *cbf11*<sup>+</sup> was found to be associated with the cut (cell untimely torn) phenotype (Fig 8D, E), which is a failure in

coordination of the nuclear and cell division [37]. In summary, the lack of *cbf11*<sup>+</sup> or excess dose of *cbf12*<sup>+</sup> both result in pleiotropic defects in cell and nuclear division, highlighting the requirement for proper balance of CSL proteins in these processes.

## Discussion

Current understanding of the CSL family function derives exclusively from metazoan model organisms and is based mostly on studies concerning the role of the Notch pathway in vertebrate ontogenesis or insect development [2,5,11,38]. It is now obvious that this view needs to be extended, as recent reports indicate that there are Notch-independent CSL activities in animals [9,10,12,14]. Intriguingly also, our bioinformatical data identified the CSL proteins in several organisms that are evolutionarily distant to animals and lack the Notch pathway components [15]. It is thus possible that the Notch-independent regulatory function of the metazoan CSL proteins is the evolutionarily more ancestral one and that at least some aspects of it are related to the functioning of the fungal factors.

To address the question of CSL genes function in fungi, we investigated the properties of this family members present in the fission yeast *S. pombe* – *cbf11*<sup>+</sup> (class F1) and *cbf12*<sup>+</sup> (class F2). We analyzed their expression profiles, studied the biochemical properties of their protein products, and collected evidence concerning the phenotypes associated with the manipulations of these two CSL genes.

We found that both Cbf11 and Cbf12, neither of which is an essential product, affect cell-cell and cell-surface adhesion of *S. pombe*. While Cbf11 behaves as a negative regulator of adhesion, Cbf12 acts to increase it. Under physiological conditions, one of the reasons cells aggregate (flocculate) is the exhaustion of nutrients coinciding with the stationary phase of

growth. It is a stress response with the aim of escaping the unfavorable environment either by sedimentation or by floating [39]. Notably, one of the *cbf12*<sup>+</sup> expression peaks occurs at the stationary phase, and we hypothesize that one of the Cbf12 functions may be to trigger the increase of adhesion at this stage, possibly by counteracting or replacing Cbf11 at the respective promoters.

In addition to that, since *cbf12*<sup>+</sup> is also upregulated during sexual differentiation, it might potentially participate in the regulation of meiotic transcription as well. Although, given the lack of CSL-associated meiotic phenotypes (this study and [25]), such a function would apparently be non-essential for the progress and completion of meiosis.

It was described that changes in adhesive properties of cells have profound impact on yeast colony morphology [40-42]. Also, extracellular matrix-like material was found to be produced by some *S. cerevisiae* strains, which forms a capsule around the colony and seems to serve as a scaffold for the cells within the colony, possibly influencing its shape [43]. It is thus possible that the  $\Delta$ *cbf11*-associated alterations of colony morphology we observed result from the increased adhesion and/or “shiny” material secretion of these strains. Alternatively, the aberrant colony morphology could be a consequence of the cell separation defects and pseudohyphal growth noticed for a fraction of cells in the  $\Delta$ *cbf11* colonies, as reported for other fungal species [41,44,45]. These two explanations, however, are not mutually exclusive, and the effects may combine.

When the expression of the fission yeast CSL genes was experimentally thrown off balance ( $\Delta$ *cbf11* KO or *cbf12*<sup>+</sup> overexpression), various cell separation defects appeared, although their penetrance was usually rather low. This may be explained as a result of an overall decrease in fidelity of a number of cellular processes that could have lead to stochastic manifestations in only a minor fraction of cells. These phenotypes (see Fig. 8), which were



virtually never found in the WT control cells, are highly reminiscent of the fission yeast *sep* mutants, which display very similar cell separation-related phenotypes, although often accompanied by sterility and impaired stress response to various substances [36,46]. Some of the *sep* mutations have already been cloned and found to reside in genes encoding (general) regulators of transcription, rather than factors involved directly in cell separation [47-49]. The *sep10*<sup>+</sup> gene encodes the Med31 subunit of the RNA polymerase II Mediator complex [49,50] and, interestingly, there are indications that the transcription of the *cbf12*<sup>+</sup> gene is Med31-dependent [51,52].

The most intriguing effect of the dysregulated CSL expression (again,  $\Delta cbf11$  or *cbf12*<sup>+</sup> overexpression) was the emergence of stable diploid-like subpopulations. These manifested themselves as dark-staining clonal sectors in monoclonies grown in the presence of phloxin B, as a >2C peak on FACS histograms of the mutant cultures, and as large cells with 2n-like nuclei observed under a microscope. Spontaneous diploidization occurs even in WT heterothallic strains, however, at a very low frequency [34]. Indeed, we have occasionally found dark sectors on WT monoclonies as well, although their occurrence was about 1 sector in 100 colonies (data not shown). In contrast to that, for the  $\Delta cbf11$  and double KO strains there were typically ~10 or more dark sectors in each colony. A number of mutations, mostly in genes involved in the cell cycle regulation and progression, were shown to increase the so-called endoreduplication or re-replication frequency, and the respective mutant populations contain varying fractions of diploid (or even polyploid) cells [35]. An incomplete M phase occurs in one type of these mutants, caused by the lack of coordination between the nuclear and cell division, resulting in the cut phenotype [37,53]. In this case, the septum develops prematurely and cuts through the nucleus before the latter completes its division. Although lethal in most cases, the septum occasionally forms without severing the nucleus, and, as a

result, a diploid and an anucleate compartment are formed. If the cell division proceeds further, a viable and stable (because homozygous, thus non-sporulating) diploid daughter cell is born and may start its clonal diploid subpopulation.

Remarkably, we have observed the cut phenotype in association with the deletion of *cbf11*<sup>+</sup>, and cut cells (both before and after daughter cell separation) were present in DAPI-stained  $\Delta cbf11$  and especially double KO cultures. Thus, the cut phenotype may contribute to the appearance of the diploid subpopulations we observed. However, other factors are likely involved as this explanation does not account for the diploidization observed when *cbf12*<sup>+</sup> was overexpressed. We did not see almost any cut cells under these conditions while the FACS profiles were highly similar to  $\Delta cbf11$ . The actual mechanism(s) of diploidization in CSL-manipulated strains will be addressed in future studies. Nevertheless, our experiments revealed an important requirement for the CSL family in the maintenance of the fission yeast genome ploidy.

The two fission yeast CSL proteins share the unique domain composition and important sequence motifs with the class M family members [15] and, similar to their metazoan counterparts, localize to the cell nucleus [3,54]. Both proteins have an ability to activate transcription in a heterologous reporter system, and one of them, Cbf11, was found to specifically recognize and bind directly to the canonical CSL response element on DNA [55]. One possibility is that the lack of binding observed for Cbf12 is due to, e.g., an inhibitory posttranslational modification, rather to its actual inability to bind DNA.

All clues taken together, we propose a role for Cbf11 and Cbf12 as novel transcription factors in *S. pombe* that regulate or fine-tune, in an antagonistic manner, a number of important processes (Fig. 9). Their regulatory engagements differ from those of the metazoan CSL family members, following a distinct logic of the unicellular organism stemming from the

vast evolutionary distance between these taxons [56]. Instead of embryonic development and cell fate decisions, the CSL paralogs in fission yeast seem to regulate cell adhesion, extracellular material production, colony morphology, septation and daughter cell separation, coordination of the nuclear and cell division, and the resulting proper maintenance of genome ploidy. Studying CSL proteins in an organism that does not have the Notch pathway may facilitate the understanding of their Notch-independent functions in metazoa.

## **Acknowledgements**

The authors would like to thank the members of the lab for their support, especially to Ondřej Tolde for his help with fluorescence microscopy, to the lab of Prof. Zdena Palková at the Faculty of Science, Charles University in Prague, for sharing their lab equipment and for their help with colony imaging, and to Drs. Simon Whitehall and Ramsay McFarlane for their helpful suggestions.

This work was supported by the Grant Agency of the Charles University grant no. 157/2005/B-BIO/PrF, the Czech Science Foundation grant no. 204/03/H066, and the research projects No. MSM0021620858 and LC07032 of the Czech Ministry of Education, Youth and Sports.

## Figure legends

**Figure 1 – The expression profiles of *cbf11*<sup>+</sup> and *cbf12*<sup>+</sup>, and the subcellular localization of their protein products.** (A) The domain structure of the two fission yeast CSL proteins.

The structure of *C. elegans* LAG-1 is shown at the top for comparison [16]. Recognized Pfam domains [57] are indicated: RHR-N, BTD and RHR-C. The proteins are drawn to scale. (B)

The expression of *cbf11*<sup>+</sup> and *cbf12*<sup>+</sup> mRNAs was measured under various conditions using quantitative RT-PCR. Samples were taken at different time points – the early, mid and late logarithmic phase (“log”), early and late stationary phase (“stat”), from diploid cells growing vegetatively (“diploid”) and 8 hours after induction of sporulation (“sex diff”). The mean values for 2-4 independent experiments are shown; standard errors are indicated where n > 2.

(C) Live log-phase cells expressing either EGFP-Cbf11 from a plasmid or Cbf12-EGFP from a chromosomally tagged allele were subjected to confocal microscopy. Both CSL proteins display rather diffuse nuclear localization but seem to be excluded from the nucleolus. Bars represent 5 μm.

**Figure 2 – Cbf11 binds to the CSL response element on DNA.** (A) The sequences (sense strands only) of the DNA probes used for gelshift experiments. The sequences were derived from the promoters of known CSL-responsive genes from various organisms. “m8” – drosophila m8 gene [58], “KSHV” – Kaposi’s Sarcoma-Associated Herpesvirus K14/vGPCR gene [59], “RBP” and “HES” – mouse RBP-Jκ and HES-1 genes, respectively [60]. The control probes were derived from the KSHV sequence by a point mutation of the critical G residue (“MUT”) or by a complete disruption of the CSL binding site (“DEL”). The CSL response element is printed in bold. (B) A gelshift experiment documenting the existence of a highly specific CSL response element-binding activity in WT *S. pombe* lysates (RBP probe). Significantly, the activity is absent from lysates prepared from  $\Delta cbf11$  cells. The faint

uppermost band (\*) corresponds to an unknown non-specific binding activity. (C)

Recombinant Cbf11 was affinity-purified from *S. pombe* lysates and subjected to gelshift experiments with the set of probes introduced in panel A. Cbf11 is capable of binding these probes with very high specificity as there is no binding either to the single point mutation-containing MUT or the scrambled DEL control probes.

**Figure 3 – The deletion of *cbf11*<sup>+</sup> results in cold sensitivity and slow growth.** (A) When cultured on YES plates at 19°C for 5 days, the  $\Delta cbf11$  and  $\Delta cbf11 \Delta cbf12$  strains displayed marked growth impairment as compared with the  $\Delta cbf12$  and WT cells. (B) The cells were cultured at 30°C in YES and OD was measured every hour. There was no difference between the  $\Delta cbf12$  strain and WT. In contrast, the deletion of *cbf11*<sup>+</sup> was found to cause growth retardation, which phenotype is, intriguingly, further potentiated by the simultaneous deletion of *cbf12*<sup>+</sup>. (C) When overexpressed from a plasmid, *cbf11*<sup>+</sup> but not *cbf12*<sup>+</sup> rescues the growth defect in the  $\Delta cbf11 \Delta cbf12$  strain. The monoclonies were cultured on inducing MB plates with phloxin B for 7 days. Notably, the overexpression of *cbf12*<sup>+</sup> seems to be toxic for the cells as judged by the intense phloxin staining of the colony.

**Figure 4 – The deletion of *cbf11*<sup>+</sup> results in altered colony morphology.** (A) Cells were spotted on a YES plate and incubated at 30°C for 7 days. The  $\Delta cbf11$  and  $\Delta cbf11 \Delta cbf12$  strains display a “shiny” phenotype, i.e., they appear glossy when illuminated, unlike the dim WT and  $\Delta cbf12$  cells. (B) “Shiny” colonies (bottom left part of the image) on a YES plate were carefully submerged in water. A highly reflective thin layer detached from the “shiny” colonies and floated on the water surface (center of the image). This layer resembles the “shiny” spot texture presented in panels C and D. The plate was photographed from above. (C) The surface of the spots from panel A observed under higher magnification. The dim

giant colonies are smooth, in contrast to the wrinkled surface of the “shiny” giant colonies, which appears as a network of shallow grooves. The bars represent 1 mm. (D) Low plasmid-driven expression of *cbf11*<sup>+</sup> but not *cbf12*<sup>+</sup> rescues the shiny phenotype in the  $\Delta cbf11 \Delta cbf12$  strain. The cells were spotted on a YES plate and incubated at 30°C for 26 days. (E) Monocolonies were grown for 14 days on YES plates (non-inducing) containing phloxin B. The colonies of the  $\Delta cbf11$  and  $\Delta cbf11 \Delta cbf12$  strains are irregularly shaped, show sectoring of phloxin B staining and their rim is “shiny” and covered with similar grooves as observed for the spotted giant colonies. The  $\Delta cbf12$  monocolonies appear normal. (F) Overexpression of *cbf11*<sup>+</sup> rescues the sectoring phenotype of the  $\Delta cbf11 \Delta cbf12$  strain, and confers more WT-like colony morphology. The monocolonies were grown on inducing MB plates with phloxin B at 30°C for 4 weeks.

**Figure 5 – Opposing roles of *cbf11*<sup>+</sup> and *cbf12*<sup>+</sup> in cell adhesion.** (A) Cells were spotted on a YES plate, incubated at 30°C for 2 weeks, and then washed with a stream of water. A layer of cells remained adhering to the agar in the case of the WT and  $\Delta cbf11$  strains. By contrast, strains lacking *cbf12*<sup>+</sup> were washed off completely. A microscopic picture of the washed surface is shown in the bottom panel; the bar represents 25  $\mu$ m. (B) The defective adhesion of the  $\Delta cbf11 \Delta cbf12$  strain can be rescued by overexpression of *cbf12*<sup>+</sup> from a plasmid. Cells were grown on an inducing MB plate for 1 week and processed as in A. Note that overexpression of *cbf11*<sup>+</sup> further decreased adhesion of the double KO strain. (C) The  $\Delta cbf11$  single KO cells display higher adhesion than WT, which can be lowered again by the overexpression of *cbf11*<sup>+</sup> from a plasmid. Spots were grown on inducing MB plates for 6 and 19 days, respectively, and processed as in A.

**Figure 6 – Deletion of *cbf11*<sup>+</sup> or overexpression of *cbf12*<sup>+</sup> trigger flocculation.**

Flocculation assays were performed as described under Materials and Methods; cultures transferred to Petri dishes were photographed from above. Areas of ~1.7 cm<sup>2</sup> are shown for each culture. Flocculation manifests itself as the appearance of white aggregates. (A) Cultures were grown in parallel to stationary phase in liquid inducing MB medium at 30°C. The  $\Delta cbf11$  strain displayed strong flocculation that was completely abolished by the re-introduction of *cbf11*<sup>+</sup> on a plasmid. (B) Cells transformed with either an empty vector or a plasmid encoding *cbf12*<sup>+</sup> were grown to logarithmic phase under the conditions described in A. The overproduction of Cbf12 triggered flocculation. A synthetic hyper-flocculation phenotype was observed in the  $\Delta cbf11$  background. (C) The Cbf12-induced flocculation is galactose-dependent, as judged by competition assays.

**Figure 7 – Loss of *cbf11*<sup>+</sup> or overexpression of *cbf12*<sup>+</sup> result in frequent diploidization.**

Exponentially growing WT cells overexpressing *cbf12*<sup>+</sup> from a plasmid, and cells of the single and double CSL mutants were fixed, stained with propidium iodide, and their DNA content was analyzed by flow cytometry. The control haploid and diploid G<sub>2</sub>-phase peaks are indicated as 2C and 4C, respectively. While the  $\Delta cbf12$  strain appears wild-type, the  $\Delta cbf11$  and  $\Delta cbf11 \Delta cbf12$  KOs, and the *cbf12*<sup>+</sup> overexpressor strains contain a significant proportion of cells with a >2C DNA content. When cultures were inoculated from the light/dark sectors shown in Fig. 4E, the cells from the dark sectors sorted as diploids.

**Figure 8 – Loss of *cbf11*<sup>+</sup> or overexpression of *cbf12*<sup>+</sup> lead to multiple defects in cell and nuclear division.** Exponentially growing cells, either live (F) or fixed with ethanol and stained with DAPI (A-E, G-H), were observed under an epifluorescence/DIC microscope. WT haploid (A) and diploid (B) cells are shown for comparison. (C) No significant abnormalities

were found for the  $\Delta cbf12$  strain. By contrast, the cultures of the  $\Delta cbf11$  (D, F) and  $\Delta cbf11 \Delta cbf12$  (E) strains are heterogeneous in shape and size of both cells and nuclei. Infrequently, cells display the cut phenotype (asterisks), very large nuclei, pseudohyphal growth, multiple septa (arrowheads) or aberrantly thick septa (arrows). (G) When  $cbf12^+$  is overexpressed in a WT background, similar size/shape heterogeneity and septation defects can be seen (almost no cut, however). (H) In a  $\Delta cbf11$  background, the increased dosage of  $cbf12^+$  potentiates the septation defects, and a significant proportion of the cells grow as short unseparated filaments. The bar represents 5  $\mu\text{m}$ .

**Figure 9 – A proposed model of CSL functioning in *S. pombe*.** In log-phase vegetative cells the negative regulatory effects of Cbf11, the class F1 CSL representative, prevail. Cbf12, the class F2 paralog, is expressed at low levels under these conditions and cannot overcome the effect of Cbf11. Upon entry into the stationary phase of growth (high cell density, depleted nutrients) or during meiosis, the Cbf12 protein levels rise and trigger the activation of (a subset of) the CSL target genes, e.g., by canceling the effect of Cbf11 or by replacing Cbf11 at the respective promoters. The target genes might be organized into several subsets (e.g., meiotic, stationary-phase) differing in their responsiveness to Cbf11/12, and are involved in processes such as extracellular material production, cell adhesion, colony morphology establishment, cell and nuclear division and their mutual coordination, maintenance of genome ploidy, and potentially also meiosis. At all times, a proper balance between the Cbf11 and Cbf12 activities seems to be important for the above-mentioned processes not to be perturbed.



## References

- [1] S.D. Hayward, Viral interactions with the Notch pathway, *Semin. Cancer Biol.* 14 (2004) 387-396.
- [2] S. Artavanis-Tsakonas, M.D. Rand, R.J. Lake, Notch signaling: cell fate control and signal integration in development, *Science* 284 (1999) 770-776.
- [3] J.L. de la Pompa, A. Wakeham, K.M. Correia, E. Samper, S. Brown, R.J. Aguilera, T. Nakano, T. Honjo, T.W. Mak, J. Rossant, R.A. Conlon, Conservation of the Notch signalling pathway in mammalian neurogenesis, *Development* 124 (1997) 1139-1148.
- [4] G. Weinmaster, C. Kintner, Modulation of Notch signaling during somitogenesis, *Annu. Rev. Cell Dev. Biol.* 19 (2003) 367-395.
- [5] S.E. Pursglove, J.P. Mackay, CSL: a notch above the rest, *Int. J. Biochem. Cell Biol.* 37 (2005) 2472-2477.
- [6] A.P. Weng, J.C. Aster, Multiple niches for Notch in cancer: context is everything, *Curr. Opin. Genet. Dev.* 14 (2004) 48-54.
- [7] E.C. Lai, Notch signaling: control of cell communication and cell fate, *Development* 131 (2004) 965-973.
- [8] C. Oka, T. Nakano, A. Wakeham, J.L. de la Pompa, C. Mori, T. Sakai, S. Okazaki, M. Kawaichi, K. Shiota, T.W. Mak, T. Honjo, Disruption of the mouse RBP-J kappa gene results in early embryonic death, *Development* 121 (1995) 3291-3301.

- [9] S. Koelzer, T. Klein, A Notch-independent function of Suppressor of Hairless during the development of the bristle sensory organ precursor cell of *Drosophila*, *Development* 130 (2003) 1973-1988.
- [10] M. Kaspar, T. Klein, Functional analysis of murine CBF1 during *Drosophila* development, *Dev. Dyn.* 235 (2006) 918-927.
- [11] S. Bray, M. Furriols, Notch pathway: making sense of suppressor of hairless, *Curr. Biol.* 11 (2001) R217-R221.
- [12] S. Barolo, R.G. Walker, A.D. Polyanovsky, G. Freschi, T. Keil, J.W. Posakony, A notch-independent activity of suppressor of hairless is required for normal mechanoreceptor physiology, *Cell* 103 (2000) 957-969.
- [13] S. Minoguchi, Y. Taniguchi, H. Kato, T. Okazaki, L.J. Strobl, U. Zimmer-Strobl, G.W. Bornkamm, T. Honjo, RBP-L, a transcription factor related to RBP-Jkappa, *Mol. Cell Biol.* 17 (1997) 2679-2687.
- [14] T.M. Beres, T. Masui, G.H. Swift, L. Shi, R.M. Henke, R.J. MacDonald, PTF1 is an organ-specific and Notch-independent basic helix-loop-helix complex containing the mammalian Suppressor of Hairless (RBP-J) or its paralogue, RBP-L, *Mol. Cell Biol.* 26 (2006) 117-130.
- [15] M. Prevorovsky, F. Puta, P. Folk, Fungal CSL transcription factors, *BMC. Genomics* 8 (2007) 233.
- [16] R.A. Kovall, W.A. Hendrickson, Crystal structure of the nuclear effector of Notch signaling, CSL, bound to DNA, *EMBO J.* 23 (2004) 3441-3451.

- [17] S. Moreno, A. Klar, P. Nurse, Molecular genetic analysis of fission yeast *Schizosaccharomyces pombe*, *Methods Enzymol.* 194 (1991) 795-823.
- [18] J. Bahler, J.Q. Wu, M.S. Longtine, N.G. Shah, A. McKenzie, III, A.B. Steever, A. Wach, P. Philippsen, J.R. Pringle, Heterologous modules for efficient and versatile PCR-based gene targeting in *Schizosaccharomyces pombe*, *Yeast* 14 (1998) 943-951.
- [19] B. Van Driesche, L. Tafforeau, P. Hentges, A.M. Carr, J. Vandenhaute, Additional vectors for PCR-based gene tagging in *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* using nourseothricin resistance, *Yeast* 22 (2005) 1061-1068.
- [20] R.A. Craven, D.J. Griffiths, K.S. Sheldrick, R.E. Randall, I.M. Hagan, A.M. Carr, Vectors for the expression of tagged proteins in *Schizosaccharomyces pombe*, *Gene* 221 (1998) 59-68.
- [21] N. Tanaka, A. Awai, M.S. Bhuiyan, K. Fujita, H. Fukui, K. Takegawa, Cell surface galactosylation is essential for nonsexual flocculation in *Schizosaccharomyces pombe*, *J. Bacteriol.* 181 (1999) 1356-1359.
- [22] C.G. Guldal, J. Broach, Assay for adhesion and agar invasion in *S. cerevisiae*, *J. Vis. Exp.* (2006) 64.
- [23] R. Higuchi, B. Krummel, R.K. Saiki, A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions, *Nucleic Acids Res.* 16 (1988) 7351-7367.
- [24] A. Decottignies, I. Sanchez-Perez, P. Nurse, *Schizosaccharomyces pombe* essential genes: a pilot study, *Genome Res.* 13 (2003) 399-406.

- [25] J. Gregan, P.K. Rabitsch, B. Sakem, O. Csutak, V. Latypov, E. Lehmann, J. Kohli, K. Nasmyth, Novel genes required for meiotic chromosome segregation are identified by a high-throughput knockout screen in fission yeast, *Curr. Biol.* 15 (2005) 1663-1669.
- [26] C. D'Alessio, E.S. Trombetta, A.J. Parodi, Nucleoside diphosphatase and glycosyltransferase activities can localize to different subcellular compartments in *Schizosaccharomyces pombe*, *J. Biol. Chem.* 278 (2003) 22379-22387.
- [27] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta C(T)}$  Method, *Methods* 25 (2001) 402-408.
- [28] P.Y. Muller, H. Janovjak, A.R. Miserez, Z. Dobbie, Processing of gene expression data generated by quantitative real-time RT-PCR, *Biotechniques* 32 (2002) 1372-1379.
- [29] S. Sazer, S.W. Sherwood, Mitochondrial growth and DNA synthesis occur in the absence of nuclear DNA replication in fission yeast, *J. Cell Sci.* 97 (1990) 509-516.
- [30] J. Mata, R. Lyne, G. Burns, J. Bahler, The transcriptional program of meiosis and sporulation in fission yeast, *Nat. Genet.* 32 (2002) 143-147.
- [31] M.H. Straver, J.W. Kijne, G. Smit, Cause and control of flocculation in yeast, *Trends Biotechnol.* 11 (1993) 228-232.
- [32] B.L. Miki, N.H. Poon, A.P. James, V.L. Seligy, Possible mechanism for flocculation interactions governed by gene FLO1 in *Saccharomyces cerevisiae*, *J. Bacteriol.* 150 (1982) 878-889.
- [33] S.L. Forsburg, *S. pombe* strain maintenance and media, *Curr. Protoc. Mol. Biol.* Chapter 13 (2003) Unit 13.15.

- [34] H. Gutz, H. Heslot, U. Leupold, N. Loprieno, *Schizosaccharomyces pombe*, in: R.C. King (Ed.), Handbook of Genetics, Plenum Press, New York, 1974, pp. 395-446.
- [35] K. Kominami, T. Toda, Fission yeast WD-repeat protein pop1 regulates genome ploidy through ubiquitin-proteasome-mediated degradation of the CDK inhibitor Rum1 and the S-phase initiator Cdc18, *Genes Dev.* 11 (1997) 1548-1560.
- [36] A. Grallert, B. Grallert, E. Zilahi, Z. Szilagy, M. Sipiczki, Eleven novel sep genes of *Schizosaccharomyces pombe* required for efficient cell separation and sexual differentiation, *Yeast* 15 (1999) 669-686.
- [37] T. Hirano, S.I. Funahashi, T. Uemura, M. Yanagida, Isolation and characterization of *Schizosaccharomyces pombe* cut mutants that block nuclear division but not cytokinesis, *EMBO J.* 5 (1986) 2973-2979.
- [38] E.C. Lai, Keeping a good pathway down: transcriptional repression of Notch pathway target genes by CSL proteins, *EMBO Rep.* 3 (2002) 840-845.
- [39] K.J. Verstrepen, F.M. Klis, Flocculation, adhesion and biofilm formation in yeasts, *Mol. Microbiol.* 60 (2006) 5-15.
- [40] B. Nguyen, A. Upadhyaya, A. van Oudenaarden, M.P. Brenner, Elastic instability in growing yeast colonies, *Biophys. J.* 86 (2004) 2740-2747.
- [41] I. Vopalenska, M. Hulkova, B. Janderova, Z. Palkova, The morphology of *Saccharomyces cerevisiae* colonies is affected by cell adhesion and the budding pattern, *Res. Microbiol.* 156 (2005) 921-931.
- [42] T.B. Reynolds, G.R. Fink, Bakers' yeast, a model for fungal biofilm formation, *Science* 291 (2001) 878-881.

- [43] M. Kuthan, F. Devaux, B. Janderova, I. Slaninova, C. Jacq, Z. Palkova, Domestication of wild *Saccharomyces cerevisiae* is accompanied by changes in gene expression and colony morphology, *Mol. Microbiol.* 47 (2003) 745-754.
- [44] W.P. Voth, A.E. Olsen, M. Sbia, K.H. Freedman, D.J. Stillman, ACE2, CBK1, and BUD4 in budding and cell separation, *Eukaryot. Cell* 4 (2005) 1018-1028.
- [45] G. Weinzierl, L. Leveleki, A. Hassel, G. Kost, G. Wanner, M. Bolker, Regulation of cell separation in the dimorphic fungus *Ustilago maydis*, *Mol. Microbiol.* 45 (2002) 219-231.
- [46] M. Sipiczki, B. Grallert, I. Miklos, Mycelial and syncytial growth in *Schizosaccharomyces pombe* induced by novel septation mutations, *J. Cell Sci.* 104 (1993) 485-493.
- [47] B. Ribar, A. Banrevi, M. Sipiczki, sep1<sup>+</sup> encodes a transcription-factor homologue of the HNF-3/forkhead DNA-binding-domain family in *Schizosaccharomyces pombe*, *Gene* 202 (1997) 1-5.
- [48] E. Zilahi, I. Miklos, M. Sipiczki, The *Schizosaccharomyces pombe* sep15<sup>+</sup> gene encodes a protein homologous to the Med8 subunit of the *Saccharomyces cerevisiae* transcriptional mediator complex, *Curr. Genet.* 38 (2000) 227-232.
- [49] Z. Szilagyi, A. Grallert, N. Nemeth, M. Sipiczki, The *Schizosaccharomyces pombe* genes sep10 and sep11 encode putative general transcriptional regulators involved in multiple cellular processes, *Mol. Genet. Genomics* 268 (2002) 553-562.
- [50] S. Bjorklund, C.M. Gustafsson, The yeast Mediator complex and its regulation, *Trends Biochem. Sci.* 30 (2005) 240-244.

- [51] K.M. Lee, I. Miklos, H. Du, S. Watt, Z. Szilagyi, J.E. Saiz, R. Madabhushi, C.J. Penkett, M. Sipiczki, J. Bahler, R.P. Fisher, Impairment of the TFIIH-associated CDK-activating kinase selectively affects cell cycle-regulated gene expression in fission yeast, *Mol. Biol. Cell* 16 (2005) 2734-2745.
- [52] I. Miklos, Z. Szilagyi, S. Watt, E. Zilahi, G. Batta, Z. Antunovics, K. Enczi, J. Bahler, M. Sipiczki, Genomic expression patterns in cell separation mutants of *Schizosaccharomyces pombe* defective in the genes *sep10+* and *sep15+* coding for the Mediator subunits Med31 and Med8, *Mol. Genet. Genomics* 279 (2008) 225-238.
- [53] S. Saitoh, K. Takahashi, K. Nabeshima, Y. Yamashita, Y. Nakaseko, A. Hirata, M. Yanagida, Aberrant mitosis in fission yeast mutants defective in fatty acid synthetase and acetyl CoA carboxylase, *J. Cell Biol.* 134 (1996) 949-961.
- [54] Y. Chen, W.H. Fischer, G.N. Gill, Regulation of the ERBB-2 promoter by RBPJkappa and NOTCH, *J. Biol. Chem.* 272 (1997) 14110-14114.
- [55] T. Tun, Y. Hamaguchi, N. Matsunami, T. Furukawa, T. Honjo, M. Kawaichi, Recognition sequence of a highly conserved DNA binding protein RBP-J kappa, *Nucleic Acids Res.* 22 (1994) 965-971.
- [56] S.B. Hedges, The origin and evolution of model organisms, *Nat. Rev. Genet.* 3 (2002) 838-849.
- [57] R.D. Finn, J. Tate, J. Mistry, P.C. Coghill, S.J. Sammut, H.R. Hotz, G. Ceric, K. Forslund, S.R. Eddy, E.L. Sonnhammer, A. Bateman, The Pfam protein families database, *Nucleic Acids Res.* 36 (2008) D281-D288.

- [58] C.N. Chung, Y. Hamaguchi, T. Honjo, M. Kawaichi, Site-directed mutagenesis study on DNA binding regions of the mouse homologue of Suppressor of Hairless, RBP-J kappa, *Nucleic Acids Res.* 22 (1994) 2938-2944.
- [59] Y. Liang, D. Ganem, RBP-J (CSL) is essential for activation of the K14/vGPCR promoter of Kaposi's sarcoma-associated herpesvirus by the lytic switch protein RTA, *J. Virol.* 78 (2004) 6818-6826.
- [60] F. Oswald, S. Liptay, G. Adler, R.M. Schmid, NF-kappaB2 is a putative target gene of activated Notch-1 via RBP-Jkappa, *Mol. Cell Biol.* 18 (1998) 2077-2088.



Table 1 – *S. pombe* strains used in this study

Strain	Genotype	Source
PN558	<i>h<sup>+</sup> leu1-32 ura4-D18 ade6-M210</i>	A. Decottignies [24]
PN559	<i>h<sup>-</sup> leu1-32 ura4-D18 ade6-M216</i>	A. Decottignies [24]
MP01	<i>h<sup>+</sup> leu1-32 ura4-D18 ade6-M210 Δcbf12::pCloneNat1</i>	This study
MP03	<i>h<sup>-</sup> leu1-32 ura4-D18 ade6-M216 Δcbf12::pCloneNat1</i>	This study
MP05	<i>h<sup>+</sup> leu1-32 ura4-D18 ade6-M216 Δcbf11::kan<sup>r</sup></i>	This study
MP06	<i>h<sup>+</sup> leu1-32 ura4-D18 ade6-M210 Δcbf11::kan<sup>r</sup></i>	This study
CBF11-KO	<i>h<sup>-</sup> leu1-32 ura4-D18 ade6-M216 Δcbf11::kan<sup>r</sup></i>	A. Decottignies [24]
MP07	<i>h<sup>+</sup> leu1-32 ura4-D18 ade6-M210 Δcbf11::kan<sup>r</sup> Δcbf12::pCloneNat1</i>	This study
MP09	<i>h<sup>-</sup> leu1-32 ura4-D18 ade6-M210 Δcbf11::kan<sup>r</sup> Δcbf12::pCloneNat1</i>	This study
MP10	<i>h<sup>-</sup> leu1-32 ura4-D18 ade6-M216 Δcbf11::kan<sup>r</sup> Δcbf12::pCloneNat1</i>	This study
MP12	<i>h<sup>-</sup> leu1-32 ura4-D18 ade6-M216 cbf12<sup>+</sup>::EGFP-kan<sup>r</sup></i>	This study
FY254	<i>h<sup>-</sup> leu1-32 ura4-D18 ade6-M210 can1-1</i>	S. Forsburg

Table 2 – Oligonucleotide primers used in this study

Primer	Sequence <sup>a</sup>	Purpose <sup>b</sup>
mp20	GTAGGATCCTATGGGGGACTATTTTG	<i>cbf11</i> <sup>+</sup> cDNA construction (exon 1, fwd, <i>S. pombe</i> plasmids)
if03	GTAGGATCCATATGGGGGACTATTTTG	<i>cbf11</i> <sup>+</sup> cDNA construction (exon 1, fwd, two-hybrid)
mp25	CCATTATTTATGTTGTCCAAGTTCAACTGA TTTAAATTTAAGG	<i>cbf11</i> <sup>+</sup> cDNA construction (exon 1, rev)
mp26	CCTTAAATTTAAATCAGTTGAACTTGGACA ACATAAATAATGG	<i>cbf11</i> <sup>+</sup> cDNA construction (exon 2, fwd)
if04	GAAGGATCCTCAGTTTCCAAAAGCAC	<i>cbf11</i> <sup>+</sup> cDNA construction (exon 2, rev)
if05	GTAGAATTCATATGTCCCCAAACGTTC	<i>cbf12</i> <sup>+</sup> ORF (fwd)
if06	GCCAGATCTTAGTGACTTTCCAAAGG	<i>cbf12</i> <sup>+</sup> ORF (rev)
mp27	TATGCTGGACTATAGTGGGC	<i>cbf11</i> <sup>+</sup> KO verification (ORF, fwd)
mp28	GATACAGCAACTCCTCCCG	<i>cbf11</i> <sup>+</sup> KO verification (outer genomic, rev)
kan-rev	AATGCTGGTCGCTATACTGC	<i>cbf11</i> <sup>+</sup> KO, <i>cbf12</i> <sup>+</sup> KI verification (kanMX6, fwd)
mp31	TGTGCAGATTTGGATGGC	<i>cbf12</i> <sup>+</sup> KO, KI verification (ORF, fwd)
mp32	AAATCAATCCCCTCCACG	<i>cbf12</i> <sup>+</sup> KO, KI verification

		(outer genomic, rev)
mp33	GCGCACGTCAAGACTGTC	<i>cbf12</i> <sup>+</sup> KO verification (natMX6, fwd)
mp43	CAGTGGAATTATCTCCCATTTTATTATTTTC AATACGAGACACTCTTTCATTCTGGATATA AGTGGCCTTTGGAAAGTCACCGGATCCCC GGGTTAATTAA	<i>cbf12</i> <sup>+</sup> ::EGFP knock in (fwd)
mp44	AAAACAAAAAGAGTAAAAATAAATATACT AATCCCTTGCAAAAACCTTTTCAATAATAAA AAAGTAGTAAAGACAAATAATGAATTCGA GCTCGTTTAAAC	<i>cbf12</i> <sup>+</sup> ::EGFP knock in (rev)
mp35	ATTTGGCTAGGTGTTTCATGG	<i>cbf11</i> <sup>+</sup> , qRT-PCR (fwd)
mp45	TGACGTTAGCAAATCTCGC	<i>cbf11</i> <sup>+</sup> , qRT-PCR (rev)
mp36	CATTCAAGCCTGATACGACG	<i>cbf12</i> <sup>+</sup> , qRT-PCR (fwd)
mp46	AGTGACTTTCCAAAGGCCAC	<i>cbf12</i> <sup>+</sup> , qRT-PCR (rev)
mp37	GTAAACGATACCAGGTCCGC	<i>act1</i> <sup>+</sup> , qRT-PCR (fwd)
mp38	GGTACCACTATGTATCCCGG	<i>act1</i> <sup>+</sup> , qRT-PCR (rev)

<sup>a</sup>The sequences are given in the 5'-3' orientation.

<sup>b</sup>fwd – forward primer, rev – reverse primer.

**Table 3 – Frequencies of the sep and cut phenotypes in CSL mutant strains**

Strain <sup>a</sup>	sep	cut	Medium
WT	0.5%	0.0%	YES/MB
$\Delta cbf12$	0.7%	0.0%	YES
$\Delta cbf11$	4.2%	16.1%	YES
$\Delta cbf11 \Delta cbf12$	4.3%	19.7%	YES
WT + $cbf12^+$ OE	8.7%	1.2%	MB
$\Delta cbf11$ + $cbf12^+$ OE	9.5%	4.1%	MB

<sup>a</sup>OE – overexpression; n > 500 cells; note that the frequency of sep is underrated due to the fact that multi-compartment filaments were scored as a single entity. Both phenotypes may combine in a single cell.

Figure 1  
[Click here to download high resolution image](#)

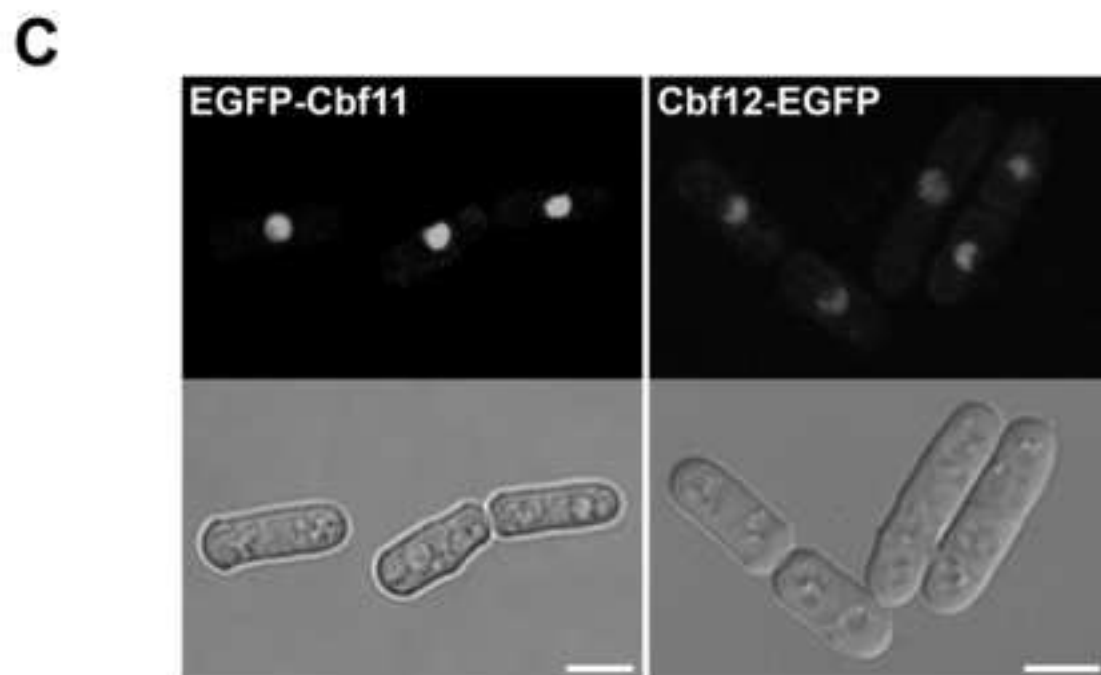
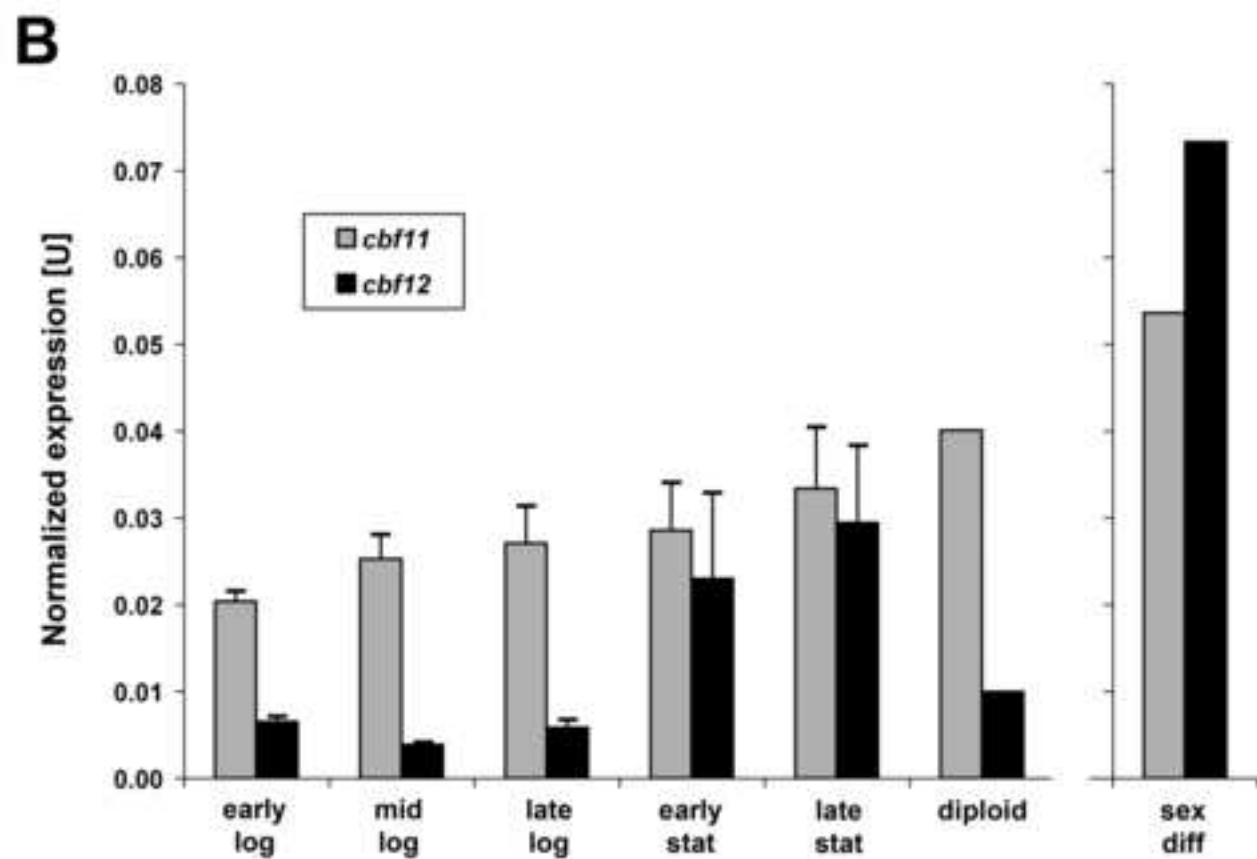
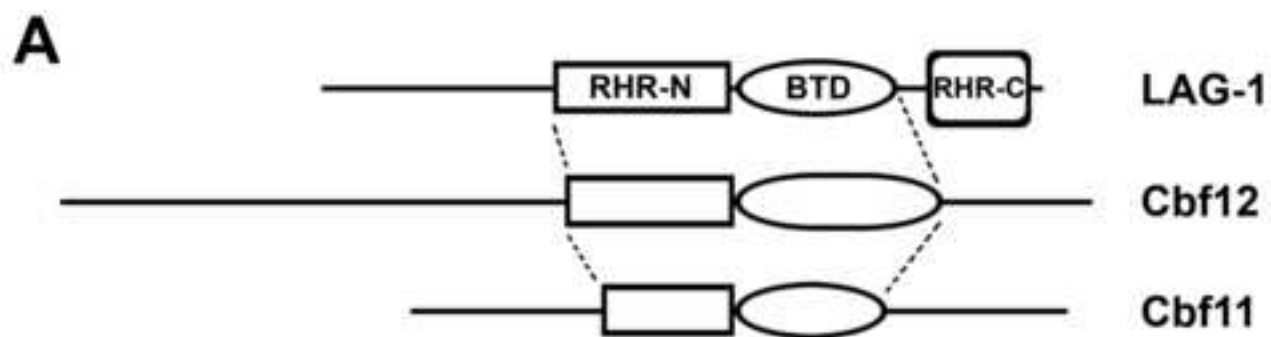


Figure 2

[Click here to download high resolution image](#)

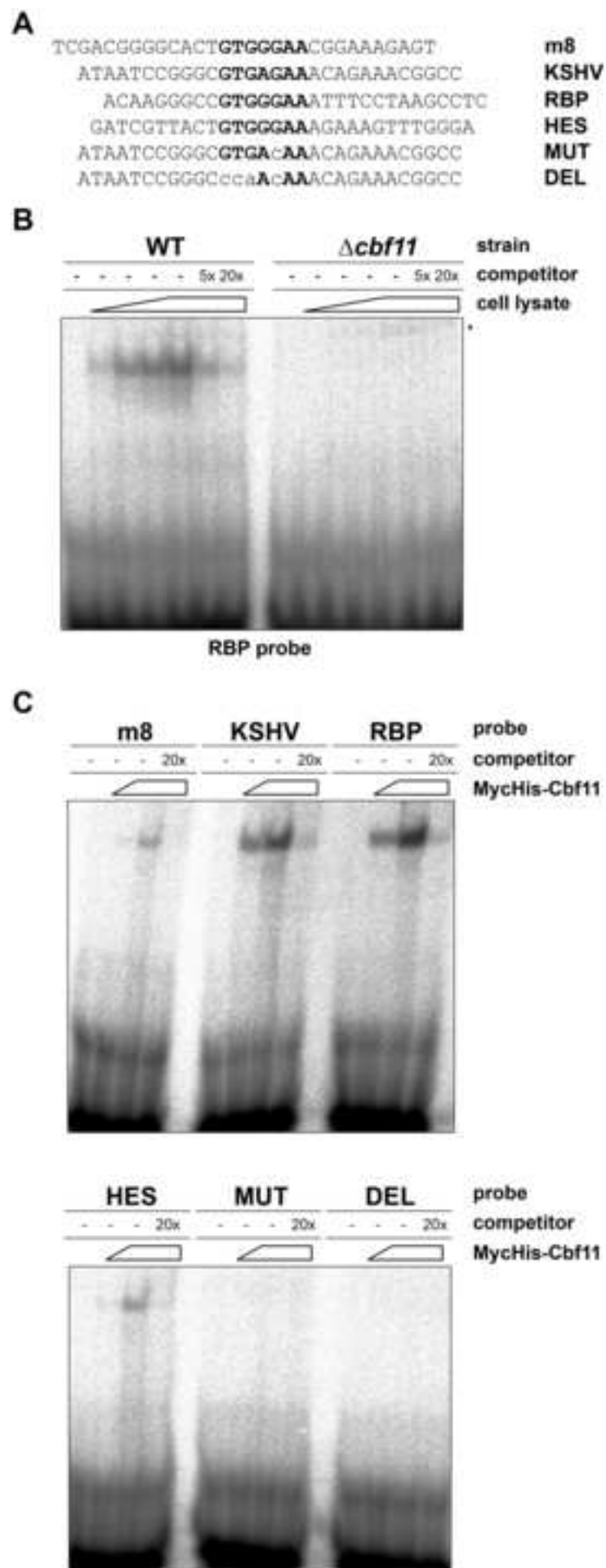


Figure 3  
[Click here to download high resolution image](#)

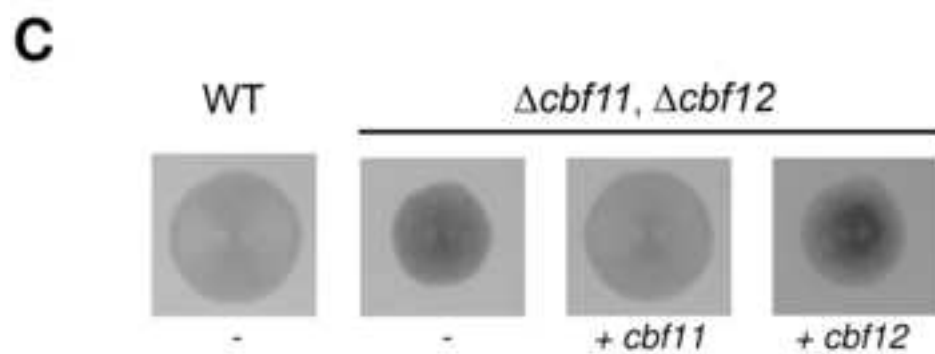
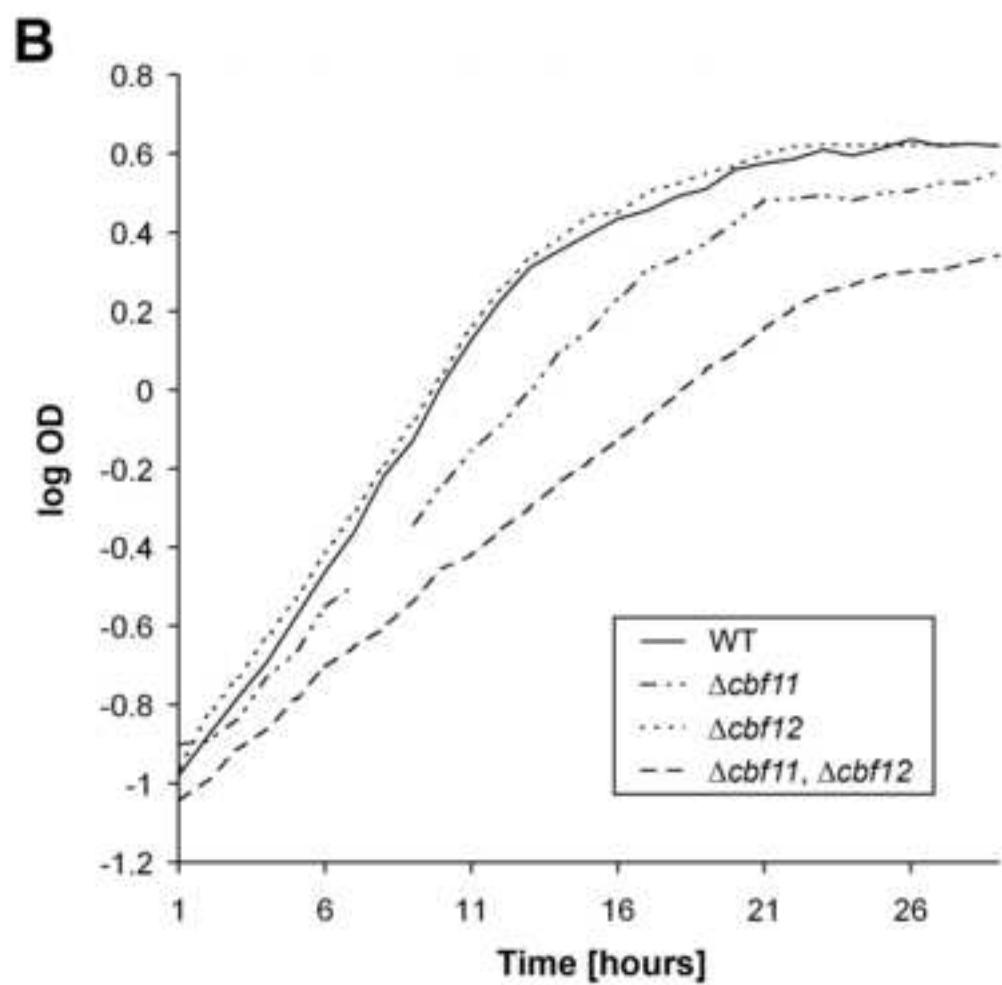
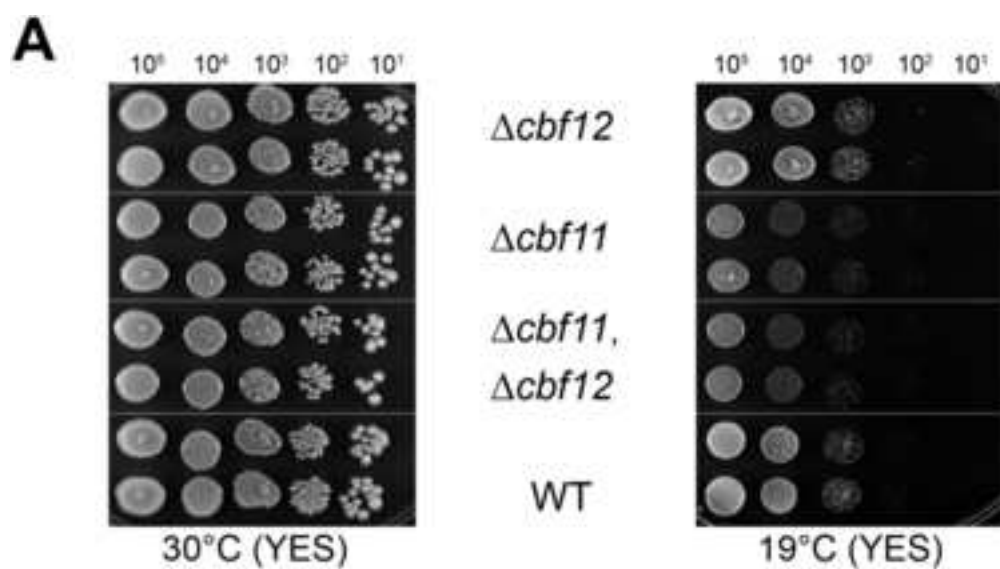


Figure 4  
[Click here to download high resolution image](#)

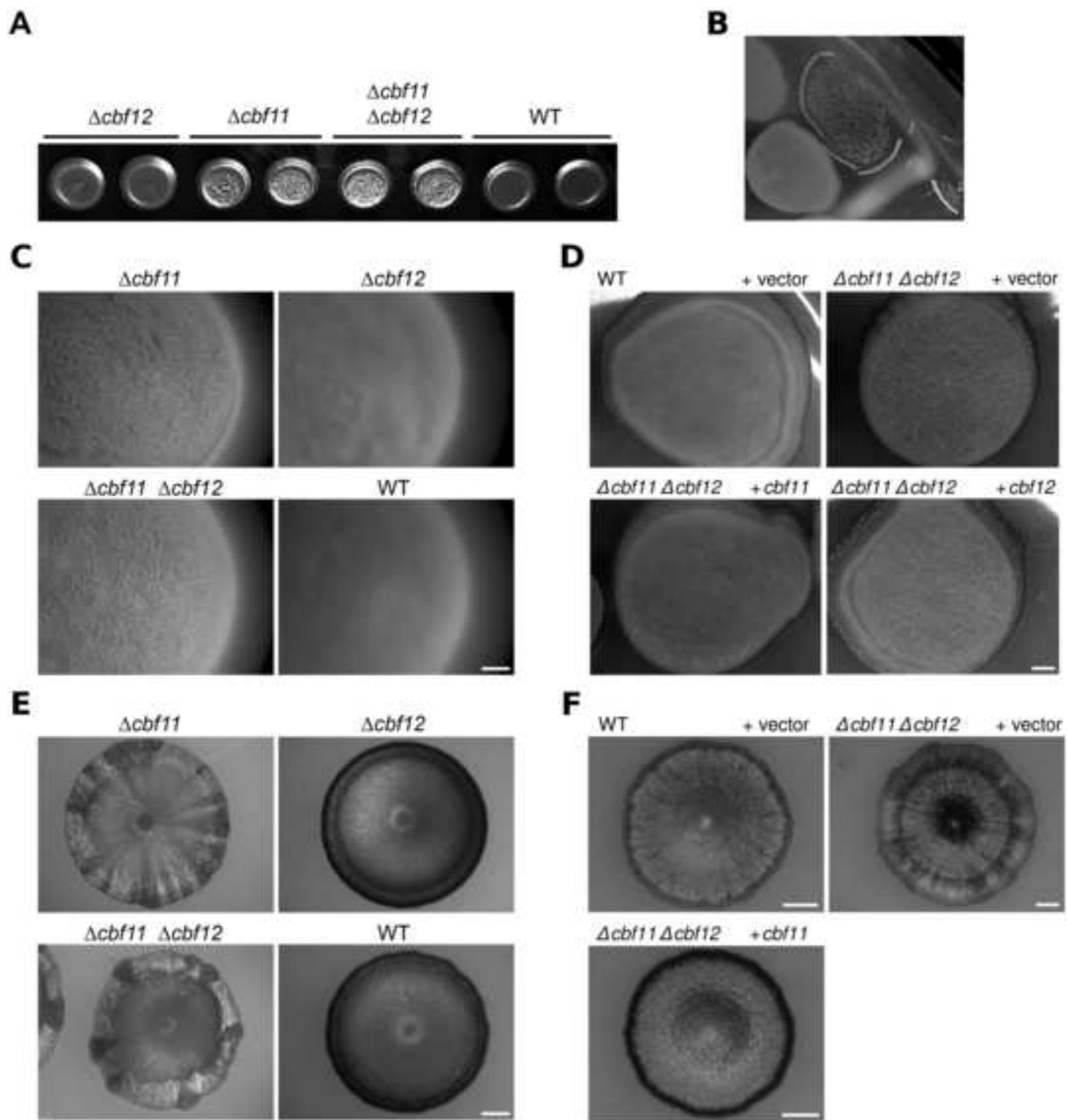
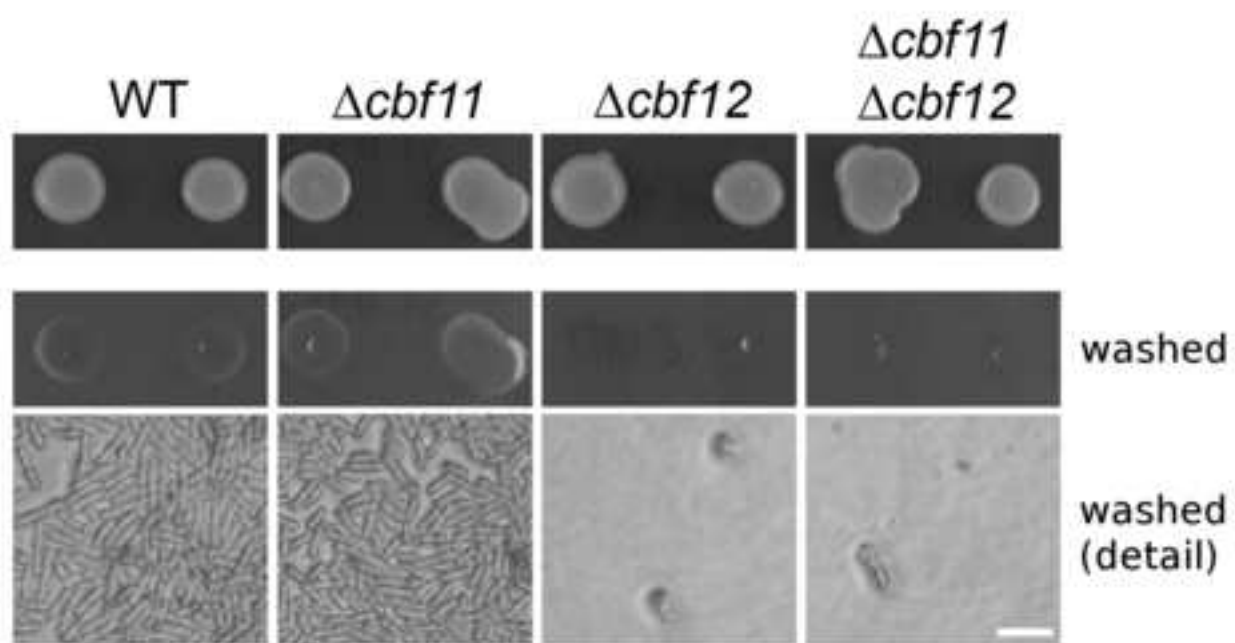


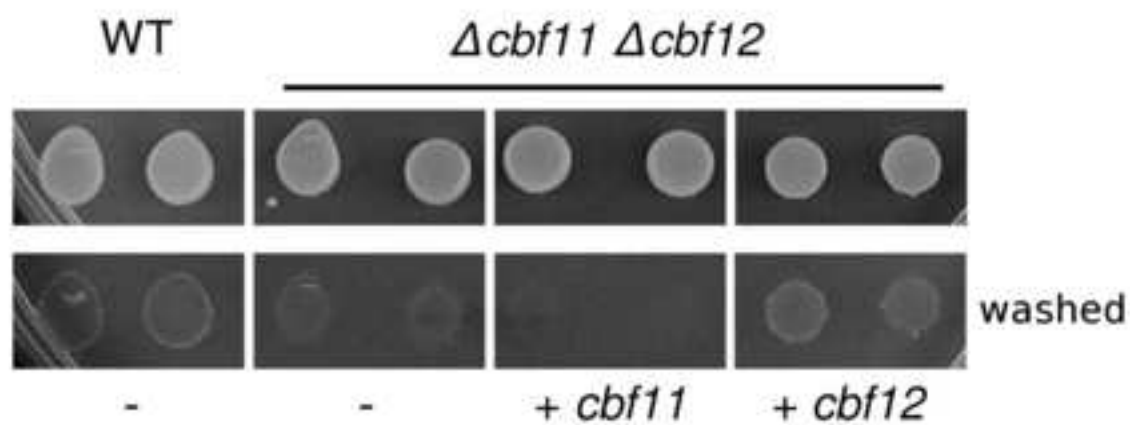


Figure 5  
[Click here to download high resolution image](#)

**A**



**B**



**C**

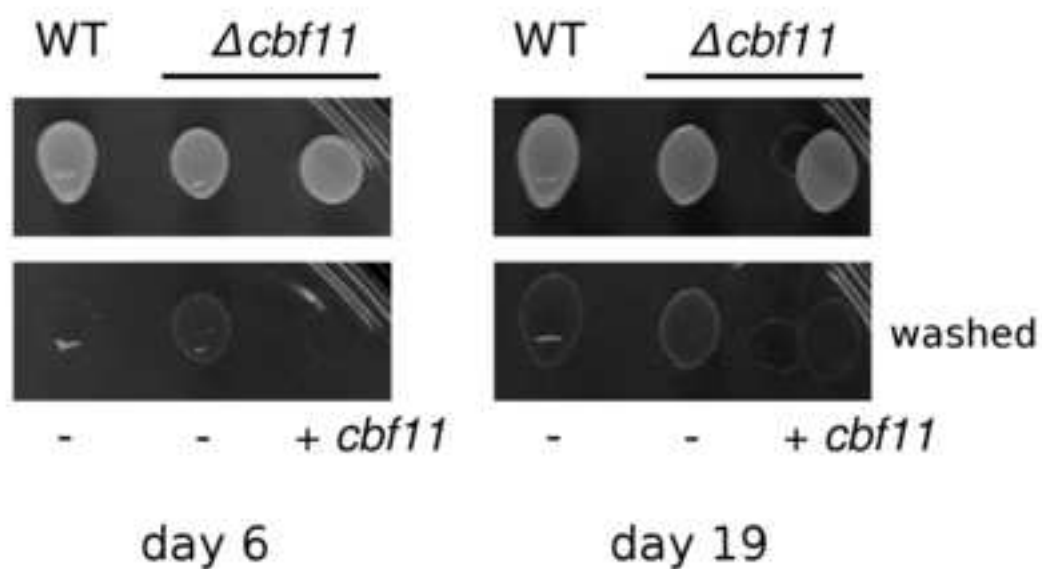


Figure 6  
[Click here to download high resolution image](#)

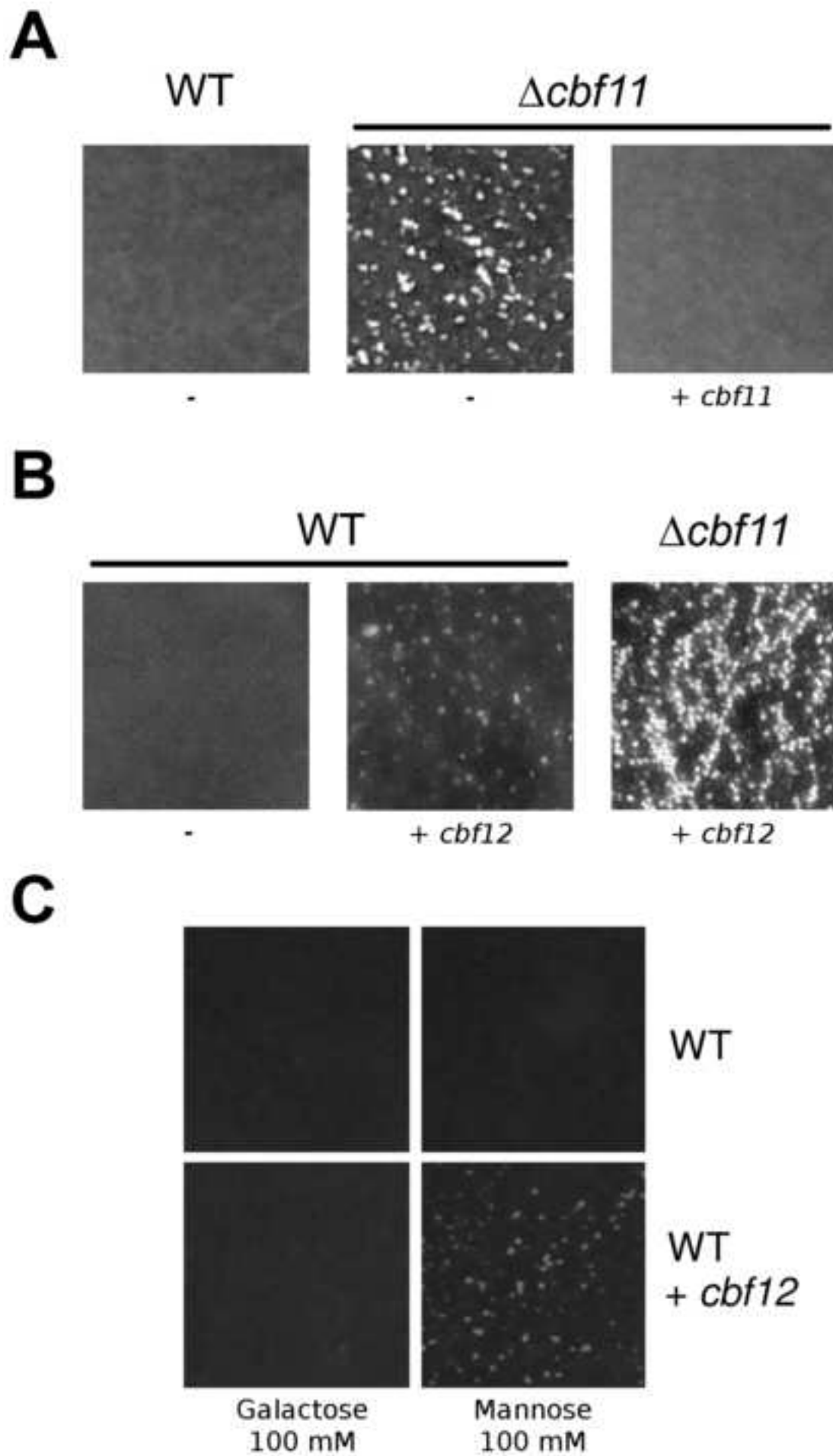


Figure 7  
[Click here to download high resolution image](#)

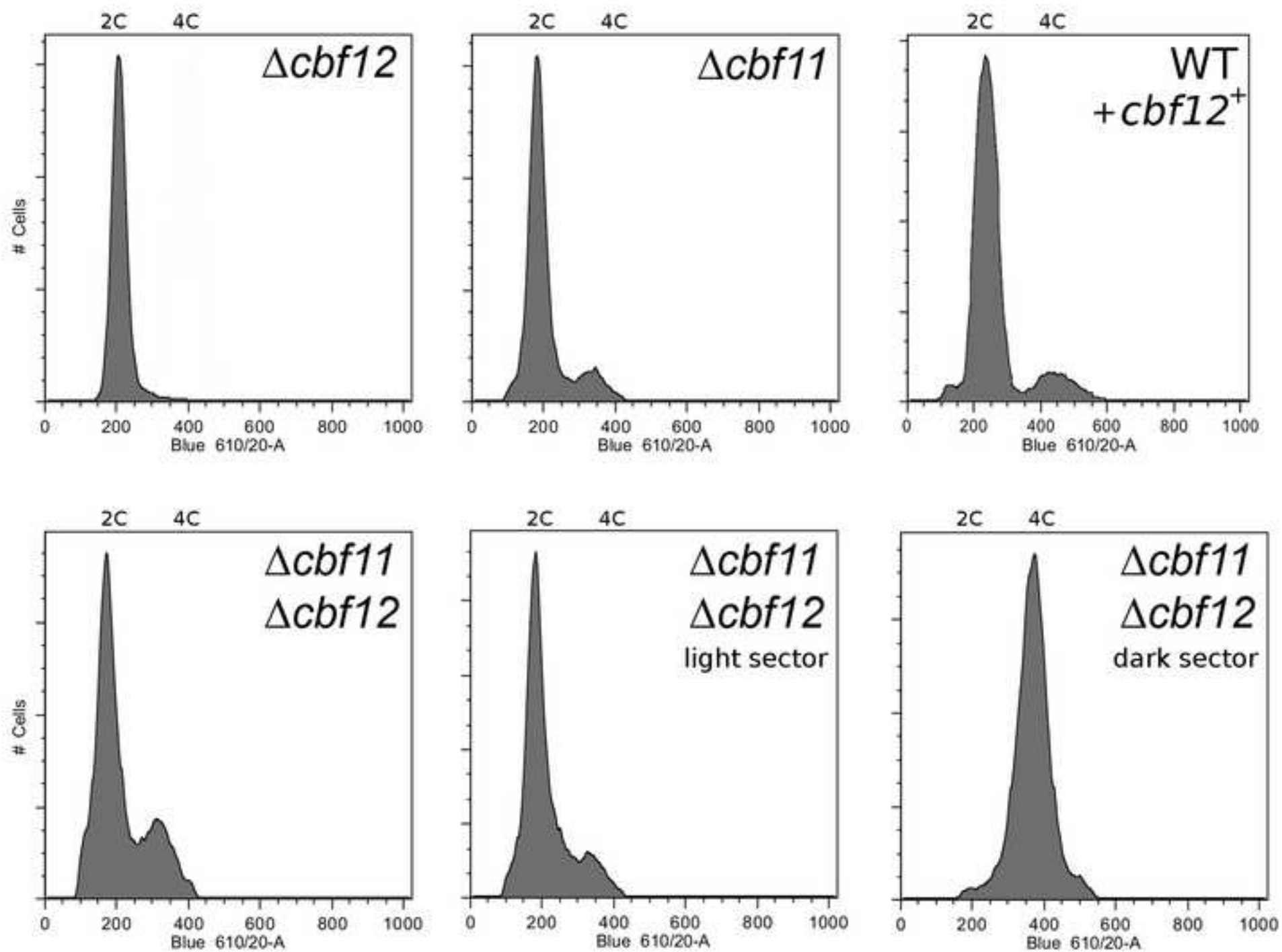


Figure 8  
[Click here to download high resolution image](#)

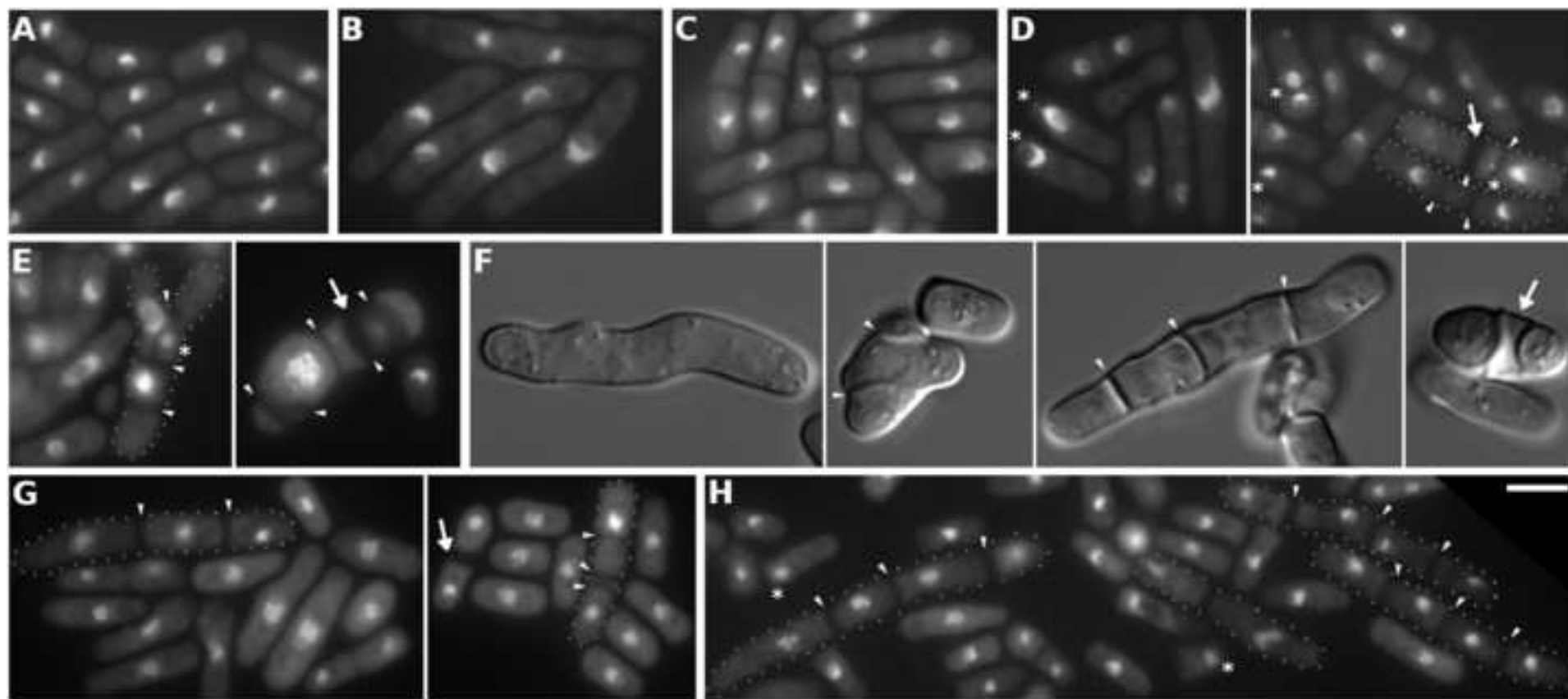
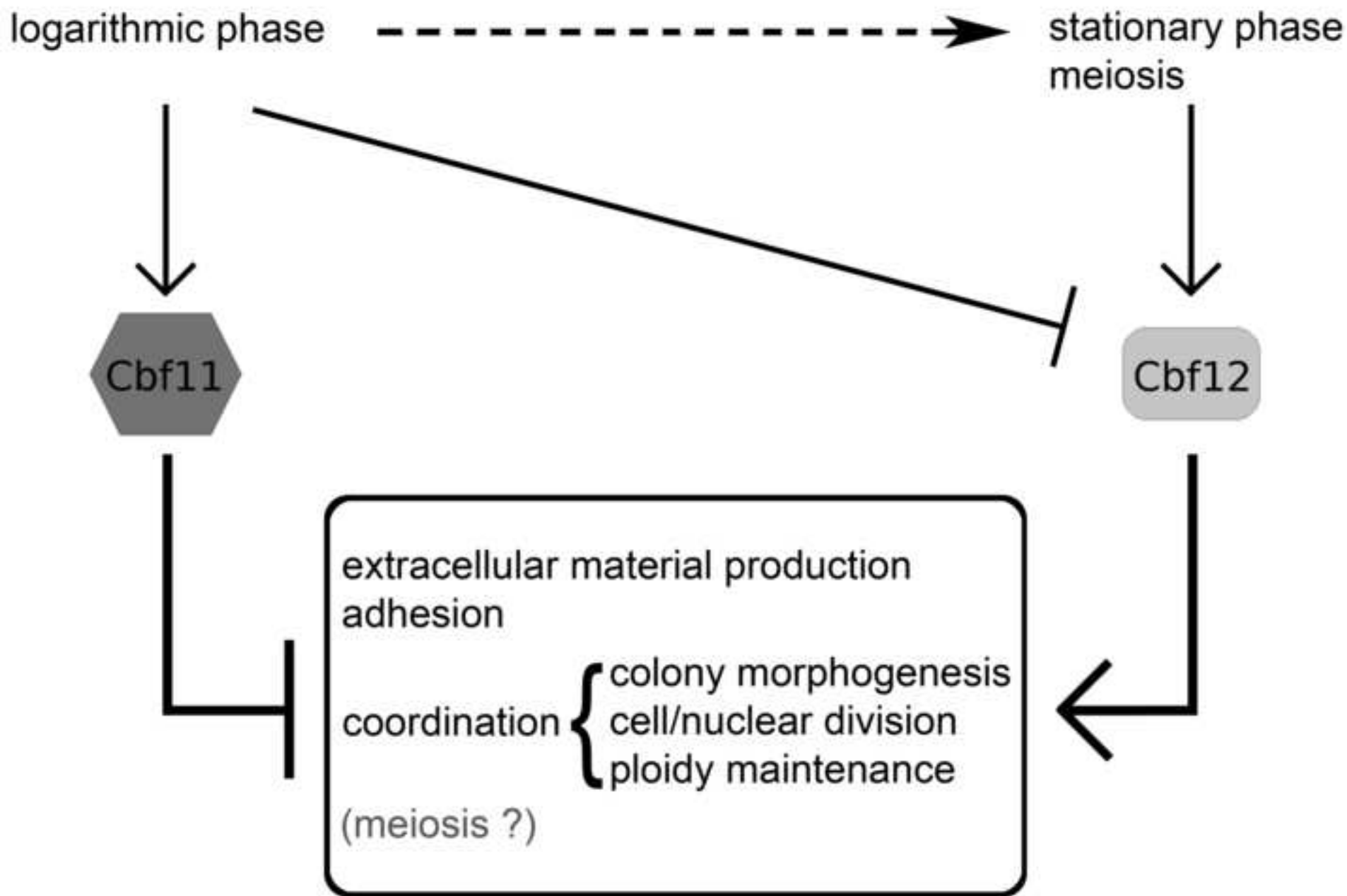


Figure 9  
[Click here to download high resolution image](#)





<http://mc.manuscriptcentral.com/fems>

**High environmental iron concentrations stimulate adhesion and invasive growth of *Schizosaccharomyces pombe***

Journal:	<i>FEMS Microbiology Letters</i>
Manuscript ID:	FEMSLE-08-10-0946
Manuscript Type:	Research Letter
Date Submitted by the Author:	02-Oct-2008
Complete List of Authors:	Prevorovsky, Martin; Faculty of Science, Charles University in Prague, Dept. of Cell Biology Stanurova, Jana Putá, Frantisek; Faculty of Science, Charles University in Prague, Dept. of Cell Biology Folk, Petr; Faculty of Science, Charles University in Prague, Dept. of Cell Biology
Keywords:	adhesion, fission yeast, iron, invasive growth, <i>Schizosaccharomyces</i>



1  
2  
3 **RESEARCH LETTER**  
4  
5

6  
7  
8 **High environmental iron concentrations stimulate adhesion and invasive growth of**  
9  
10 *Schizosaccharomyces pombe*  
11

12  
13  
14  
15 RUNNING TITLE: Iron stimulates adhesion and invasion of fission yeast  
16  
17

18  
19  
20 **Martin Převorovský, Jana Staňurová, František Půta, Petr Folk**  
21

22  
23  
24 Department of Cell Biology, Faculty of Science, Charles University in Prague, Czech  
25  
26 Republic  
27

28  
29  
30  
31 CORRESPONDING AUTHOR:  
32

33  
34 Petr Folk  
35

36 Department of Cell Biology, Faculty of Science, Charles University in Prague  
37

38 Viničná 7, Prague 2, CZ128 43, Czech Republic  
39

40  
41 Tel.: +420 22195 1765  
42

43 FAX: +420 22195 1758  
44

45  
46 E-mail: [folk@natur.cuni.cz](mailto:folk@natur.cuni.cz)  
47  
48  
49

50  
51 KEYWORDS:  
52

53 adhesion, fission yeast, iron, invasive growth, *Schizosaccharomyces*  
54  
55  
56  
57  
58  
59  
60

**Abstract**

We have found that high iron concentrations in solid complete cultivation medium potentiate cell-cell and cell-surface adhesion of the fission yeast *Schizosaccharomyces pombe*. Spotted giant colonies grown on iron-rich media were found to be more compact and more resistant to washing than those grown on plates with standard iron content. Furthermore, we have documented that excess environmental iron triggers invasive growth of *S. pombe*. Three-dimensional, branched, washing-resistant structures composed mostly of elongated but separate cells were formed within the solid agar medium. The degree of both adhesion and invasion displayed a gradual, iron concentration-dependent response. Our results suggest a novel link between iron availability and the intensively studied and important fungal virulence factors, adhesion and invasion.



## Introduction

Iron is an essential element required for numerous redox processes in virtually all organisms. The availability of iron in the environment is a limiting factor for both intra- and extracellular pathogens. As a result, their hosts have evolved ways to sequester iron from them, and many pathogenic microorganisms possess elaborate uptake systems to cope with this iron scarcity (for a review, see (Sutak *et al.*, 2008)). Perturbations of the iron homeostasis can exacerbate, for example, the severity of *Cryptococcus neoformans* infections, where iron availability was shown to regulate the major virulence factors of this organism (capsule formation, melanin production, growth at 37°C; (Jung *et al.*, 2006)).

The ability of cells to adhere to various substrates, such as, e.g., the host extracellular matrix, is considered a critical factor for many unicellular pathogenic organisms to establish infection. A number of adhesive proteins have been identified in several fungal pathogens, including *Candida albicans* and *Aspergillus fumigatus*, and their expression has been shown to represent an important virulence trait (for a review see (Mendes-Giannini *et al.*, 2005)).

Another such virulence factor is the cell's ability to invade host tissues. Several pathogenic fungal species are able to switch between yeast and invasive hyphal forms of growth, and infections by strains in which this switching is perturbed display significantly lower mortality rates (Nemecek *et al.*, 2006; Saville *et al.*, 2003).

In this article, we analyzed the relationship between the above-mentioned virulence-associated phenomena using a non-pathogenic model organism, the fission yeast *Schizosaccharomyces pombe*.

## Materials and methods

We subjected the wild-type PN559 fission yeast strain (genotype *h<sup>-</sup> leu1-32 ura4-D18 ade6-M216*; a kind gift of Dr. Anabelle Decottignies (Decottignies *et al.*, 2003)) to

1  
2  
3 adhesivity/invasivity assays (based on (Guldal and Broach, 2006)), using the standard  
4  
5 complex (YES) medium supplemented with various concentrations of FeCl<sub>2</sub>. Exponentially  
6  
7 growing PN559 cells were harvested from liquid culture by centrifugation, washed once in  
8  
9 sterile deionised water, and the cell suspension (1×10<sup>5</sup> cells in 5 µl) was plated. The YES  
10  
11 complex medium (0.5% yeast extract, 3% glucose) with SP supplements (a standard mixture  
12  
13 of adenine, leucine, uracil, histidine and lysine; Formedium) was solidified with 2% agar, and  
14  
15 either 0, 1, 2, 3 or 4 mM FeCl<sub>2</sub> was added prior to autoclaving. The plates were then  
16  
17 incubated at 30°C for up to 19 days in a wet chamber, and sample plates were subjected to  
18  
19 adhesivity/invasivity assays on days 6, 14 and 19.  
20  
21  
22  
23

24  
25 The adhesivity assay consisted of washing the plates evenly with a stream of water for  
26  
27 1 min. The spot cell mass remaining attached to the plate was then documented by  
28  
29 photography, and the plate was subjected to the invasivity assay, in which the residual cells  
30  
31 were washed off the surface completely by rubbing the plate carefully under running water.  
32  
33 Under these conditions, only cells residing deep within the agar resist the washing. The spot  
34  
35 location was photographed again to score for macroscopically visible invasive growth, and  
36  
37 finally the invasive structures formed were investigated using an Olympus CK2 light  
38  
39 microscope with an Olympus SP-350 digital camera attached to it.  
40  
41  
42  
43  
44

## 45 46 **Results**

47  
48 While conducting adhesivity tests with some of our adhesion-deficient fission yeast  
49  
50 mutant strains we noticed that the wild-type control cells displayed markedly increased  
51  
52 adhesion when cultured on a solid rich medium to which milimolar FeCl<sub>2</sub> was added.  
53  
54 Furthermore, the cells seemed to have grown deep into the agar, a behavior that is normally  
55  
56 not observed with *S. pombe*. As this could have had implications for a number of important  
57  
58 pathogenic microorganisms, we decided to examine these observations further. We also  
59  
60

1  
2  
3 wished to compare our findings with a recent work describing the ability of *S. pombe* to form  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
wished to compare our findings with a recent work describing the ability of *S. pombe* to form  
invasive filamentous structures within the agar under certain cultivation conditions (Amoah-  
Buahin *et al.*, 2005).

Fission yeast colonies were grown in the presence of various concentrations of iron  
and their adhesive and invasive properties were analyzed using the washing assays described  
under Materials and methods. There was a gradual increase in the adherent cell mass from  
nearly none, on plates to which no extra iron was added, up to a situation where almost the  
entire colony resisted washing on plates containing 4 mM FeCl<sub>2</sub>. Representative spots are  
shown in Fig. 1a. We conclude from these results that both cell-surface and cell-cell adhesion  
is potentiated by the presence of high iron concentrations in the environment.

A similar iron-dose dependency was documented for both the number and size of  
invasive structures formed within the agar (Fig. 1a, bottom row). On plates with 4 mM FeCl<sub>2</sub>,  
small invasive bodies could be observed microscopically as soon as on the day 6 post plating.  
The invasive growth occurred mostly at the periphery of the spot, forming a ring-like pattern.  
Under a microscope, the invasive bodies appeared as elaborately branched three-dimensional  
structures (Fig. 1b), composed of rather elongated but separate fission yeast cells (Fig. 1c) that  
grew in all directions within the solid agar medium. Thus, we classify these structures as  
invasive pseudomycelia.

## Discussion

Recently, invasive filamentous (mycelial) growth has been described for fission yeast.  
Under the conditions of low nitrogen and high glucose as a carbon source being available in  
the media (LNB medium), *S. pombe* can start to grow as long and branched invasive hyphae  
(Amoah-Buahin *et al.*, 2005). The mycelia also form mostly at the colony periphery (a ring-  
like pattern) and their three-dimensional structure is very similar to the iron-induced

1  
2  
3 pseudomycelia we observed. However, these true mycelia are formed by long, septated  
4  
5 hyphae, and, significantly, their formation was shown to be inhibited by the addition of  
6  
7 nitrogen sources, such as auxotrophic supplements (lysine, adenine, uracil), to the cultivation  
8  
9 media. In contrast to that, we observed invasive pseudomycelia formation on a rich medium  
10  
11 with a standard nitrogen content and with the SP nutritional supplements added, the trigger  
12  
13 being solely the abundant iron.  
14  
15

16  
17 In summary, we have described in this study an unexpected link between the  
18  
19 extracellular iron concentration and the fission yeast's adhesive properties and its ability to  
20  
21 invade solid media. We have shown that high iron strongly potentiates adhesion both to the  
22  
23 agar surface and within the yeast colony. Furthermore, abundant environmental iron  
24  
25 stimulates cells to invade solid substrates. *S. pombe* is a non-pathogenic species, but since  
26  
27 adhesion and invasive growth are important virulence traits of numerous pathogenic  
28  
29 microorganisms (Mendes-Giannini *et al.*, 2005; Saville *et al.*, 2003), we believe that our  
30  
31 results set stage for studying this novel role of iron in the models of infection as well.  
32  
33  
34  
35  
36  
37  
38

### 39 **Acknowledgements**

40  
41 This work was supported by the research projects No. MSM0021620858 and LC07032  
42  
43 of the Czech Ministry of Education, Youth and Sports.  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## References

- 1  
2  
3  
4  
5  
6  
7  
8 Amoah-Buahin, E., Bone, N. and Armstrong, J. (2005) Hyphal Growth in the Fission Yeast  
9  
10 *Schizosaccharomyces pombe*. *Eukaryot. Cell* **4**, 1287-1297.  
11  
12  
13  
14 Decottignies, A., Sanchez-Perez, I. and Nurse, P. (2003) *Schizosaccharomyces pombe*  
15  
16 essential genes: a pilot study. *Genome Res.* **13**, 399-406.  
17  
18  
19  
20  
21 Guldal, C.G. and Broach, J. (2006) Assay for adhesion and agar invasion in *S. cerevisiae*. *J.*  
22  
23 *Vis. Exp.*, **64**.  
24  
25  
26  
27  
28 Jung, W.H., Sham, A., White, R. and Kronstad, J.W. (2006) Iron regulation of the major  
29  
30 virulence factors in the AIDS-associated pathogen *Cryptococcus neoformans*. *PLoS.*  
31  
32 *Biol.* **4**, e410.  
33  
34  
35  
36  
37 Mendes-Giannini, M.J., Soares, C.P., da Silva, J.L. and Andreotti, P.F. (2005) Interaction of  
38  
39 pathogenic fungi with host cells: Molecular and cellular approaches. *FEMS Immunol.*  
40  
41 *Med. Microbiol.* **45**, 383-394.  
42  
43  
44  
45  
46 Nemecek, J.C., Wuthrich, M. and Klein, B.S. (2006) Global control of dimorphism and  
47  
48 virulence in fungi. *Science* **312**, 583-588.  
49  
50  
51  
52  
53 Saville, S.P., Lazzell, A.L., Montegudo, C. and Lopez-Ribot, J.L. (2003) Engineered control  
54  
55 of cell morphology in vivo reveals distinct roles for yeast and filamentous forms of  
56  
57 *Candida albicans* during infection. *Eukaryot. Cell* **2**, 1053-1060.  
58  
59  
60

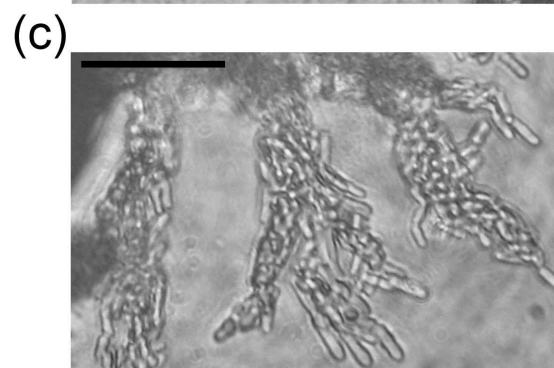
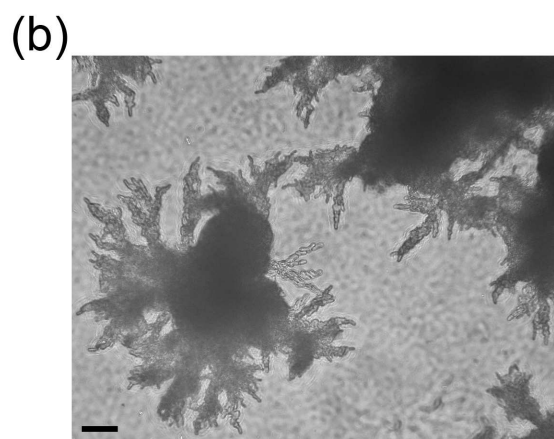
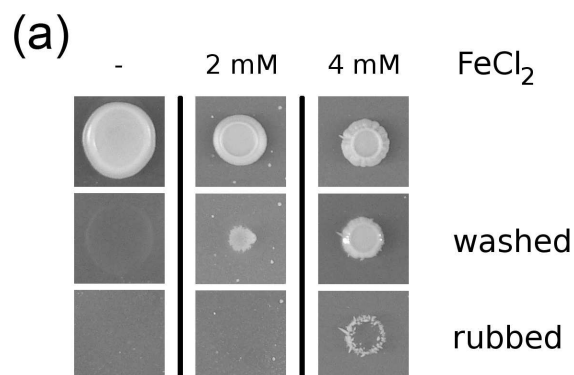
1  
2  
3 Sutak, R., Lesuisse, E., Tachezy, J. and Richardson, D.R. (2008) Crusade for iron: iron uptake  
4  
5 in unicellular eukaryotes and its significance for virulence. *Trends Microbiol.* **16**, 261-  
6  
7  
8 268.  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

For Peer Review

**Figure Legends**

Figure 1 – The effect of iron on adhesion and invasiveness of *Schizosaccharomyces pombe*.

(a) Cells were spotted on YES agar plates to which FeCl<sub>2</sub> was added to the final concentrations indicated. After 19 days the spots were washed with a stream of water to score for adhesion, and then rubbed carefully to test for invasive growth. The top left spot (no iron added) is about 1 cm in diameter; all spots were photographed at the same scale. (b) Branched invasive structures formed in the agar (YES, 4 mM FeCl<sub>2</sub>). (c) The structures are composed of rather elongated yeast cells. Bars represent 50 μm.



46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

76x159mm (400 x 400 DPI)