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

Martin Prévorovský and František Puta
Charles University, Prague, Czech Republic


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 BamHI, 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-

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

Figure 1. The RNA interference (RNAi) cassette construction procedure and transcribed RNA structure.
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 μg RNAi cassette DNA was ethanol-precipitated from the restriction reaction mixture using tRNA as a carrier and redissolved in 50 μL 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× 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-denatured 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

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

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

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

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

![Figure 2](image)
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


Received 9 June 2003; accepted 10 July 2003.

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

Complex DNA melting profiles of small PCR products revealed using SYBR® Green I

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


The fluorescent dye SYBR® 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® (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 domains (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 1 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 β-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 β-actin. It is of note that this DNA product (5′-TTGGGTACACCCCTTCTGGACA-3′) contains a centrally located 44-bp A/T-rich segment (primer pairs 1F-2R and 3F-1R, respectively), which may contribute to the complexity of the observed melting curve, we cloned the β-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, respectively),
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; František Půta - puta@natur.cuni.cz; Petr Folk* - folk@natur.cuni.cz
* Corresponding author

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.
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-Jκ 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-Jκ group but seem to act exclusively in a Notch-independent manner. Unlike the ubiquitous RBP-Jκ 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 bellow). SPC736.08 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 (b-haemofil 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...
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 BTD [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-Jκ (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.
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 β-strand (βC4) that was shown to bridge all three CSL domains in the C. elegans LAG-1 [17]. The conservancy of the βC4 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 proteins 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-
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 BTD and blue for the βC4 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].
Evolutionary conservation of the DNA-binding regions – continued

Figure 3
Evolutionary conservation of the DNA-binding regions – continued. The continuation of the alignment shown in Fig. 2.
ichical 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-Jκ and RBP-L subgroups. The RBP-Jκ 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-Jκ 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), zygomyces 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 metazoa 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-Jκ 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-Jκ 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-Jk 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 database 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 v2.5 [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 BTD domains (from helix a2 just before the BC4 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]
Acknowledgements
This work was supported by the Grant Agency of the Charles University grant no. 157/2005/B-PBIO/Fr, 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 Cvrljíková for their expert help and suggestions in the initial phase of this study.

Data for \textit{P. chrysosporium} CSL protein prediction has been provided freely by the JGI for use in this publication only.

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

References


51. Rhizopus oryzae Sequencing Project. Broad Institute of MIT and Harvard [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]

53. 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
New and corrected fungal CSL gene prediction models

Schizosaccharomyces japonicus
SjCSL1
http://www.broad.mit.edu/annotation/genome/schizosaccharomyces_japonicus
GenScan prediction
Supercontig 5 (bases 72633-727712, complementary strand)

ATGGCGCTCCATCTCCCTCATCAGAAGATGATGAGCGTCACCCTGATGATGCTGTTTGCAGCTGAGCTGGGCTCACATTCTCGACCTAA
GTTCCCTGAGAAGAAAGACCGCAACAATATGGCAATACGGTTCATCCGGAATAGGATATTGTCTCCGGATAAAAGGTAATCTTGCGATTTA
GCAACAGAGCGGCTATATATGGGGAATGATGAGCGTCACCCTGATGATGCTGTTTGCAGCTGAGCTGGGCTCACATTCTCGACCTAA

schizosaccharomyces japonicus
SjCSL2
http://www.broad.mit.edu/annotation/genome/schizosaccharomyces_japonicus
GenScan prediction
Supercontig 4 (bases 1104530-1107169)

ATGGCGCTCCATCTCCCTCATCAGAAGATGATGAGCGTCACCCTGATGATGCTGTTTGCAGCTGAGCTGGGCTCACATTCTCGACCTAA
GTTCCCTGAGAAGAAAGACCGCAACAATATGGCAATACGGTTCATCCGGAATAGGATATTGTCTCCGGATAAAAGGTAATCTTGCGATTTA
GCAACAGAGCGGCTATATATGGGGAATGATGAGCGTCACCCTGATGATGCTGTTTGCAGCTGAGCTGGGCTCACATTCTCGACCTAA

schizosaccharomyces japonicus
SjCSL2
http://www.broad.mit.edu/annotation/genome/schizosaccharomyces_japonicus
GenScan prediction
Supercontig 4 (bases 1104530-1107169)
Phanerochaete chrysosporium
PeCSL2
http://genome.jgi-psf.org/Phchr1/Phchr1.home.html
Manually corrected (predicted introns are shown in red lowercase) with manual definition of the 4th intron according to a protein sequence alignment.

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

Rhizopus oryzae
RO3G_11583.1
http://www.broad.mit.edu/annotation/genome/rhizopus_oryzae/
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.

[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.]
**Rhizopus oryzae**

**RO3G_07636.1**

[Link](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.

**Rhizopus oryzae**

**RO3G_14587.1**

[Link](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 2nd 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.
Corrected gene model CC1G_01706.1. A predicted intron (underlined) was included in the coding region in order to restore a highly conserved block of amino acids. The second intron seems to be rather long but is supported by several software predictions and unfortunately occurs in a poorly conserved region, which hampers making decisions based on a protein sequence alignment.

Coprinus cinereus
CC1G_01706.1
http://www.broad.mit.edu/annotation/genome/coprinus_cinereus/

Manually corrected (predicted introns are shown in red lowercase)
New and corrected fungal CSL protein prediction models

**Schizosaccharomyces japonicus**

*SjCSL1*

http://www.broad.mit.edu/annotation/genome/schizosaccharomyces_japonicus

**Schizosaccharomyces japonicus**

*SjCSL2*

http://www.broad.mit.edu/annotation/genome/schizosaccharomyces_japonicus

**Phanerochaete chrysosporium**

*PcCSL2*

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

**Rhizopus oryzae**

*RO3G_11583*

http://www.broad.mit.edu/annotation/genome/rhizopus_oryzae/
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        [AA803859];
- 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          MFQPTWSQEHAQPPHPHQLHPQEGLSPKQMLPQIQYDPSQISHQAPVH
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
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
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
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
ATSSGYSSLPLSANAGAERRSMEHGLTAVSPADI LKPPSAPRSARHRS

CC1G_03194
APPQHSLPYQPSDLGAFNLFHSETVGVFPPPLRTDFPADTKYQS--

UM06280
LPPSMYDNQPMLQHSDSANSTTMHDAPASSGA LM PAFGTEASYFAQAS

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
MS--I DFTPTGPQSMFAQPDI DQPFSSLQGDLHSGSHMSVGGHPLD

CC1G_03194
UM06280
ALGSPALHQAPDHILNSPRLHQAESLAGVPVPERLGPAAHLGVA VSP

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_RBP-Jk
Hr_RBP-Jk
Ci_SuH
Dm_SuH
Ce_LAG-1
Hs_RBP-L
Mm_RBP-L
Dr_RBP-L

PcČSL2
CC1G_03194
UM06280
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_RBP-Jk
Hr_RBP-Jk
Ci_SuH
Dm_SuH
Ce_LAG-1
Hs_RBP-L
Mm_RBP-L
Dr_RBP-L

PcČSL2
CC1G_03194
UM06280
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_RBP-Jk
Hr_RBP-Jk
Ci_SuH
Dm_SuH
Ce_LAG-1
Hs_RBP-L
Mm_RBP-L
Dr_RBP-L

PcČSL2
CC1G_03194
UM06280
RO3G_07636
RO3G_06481
RO3G_11583

CC1G_03194       -----PYDMIN----TYSSGKVSPLTPSDPVGGLHHGP--------FP
UM06280          HSRFGYPYSGPGFQTPOQMQSFSSLPSPLTPAQYSQYPHFCGNSSGMIPFFS
RO3G_07636       --------------------------------------MNYPELLTN-----EVE
RO3G_06481       ---------------------------------------MNDESSLNDALVYSSD
RO3G_11583       -----MTETRKRKQDEMNTLPTNWTDFIYSTPSPSIDNLFDQHYS
RO3G_14587       -----TSSSSRKRKQDFQFTPDLHSPQYNAYEPETPPLAQHEKFIQSL
CNBD3370         GYQQNNMPPOFYQFRPALALDPPYGGYLAQPQPSSAYEMPSGLRPDLGSSN
SjCSL2           PGNEALGMAELPOQGMNNDANSVLRQQQEQSSHQQQLQALTNQSAFN
SPCC1223.13      DLSQGLTHQVPNHEDYAFSPYPYELQKRPLPAHPVSFRPSTALSATVNVN
Pc6518           LSLPYDDHLLTPYSEGSLPEDQEALEYRTYCELRAHTEDTAVDAAGRGVL
CC1G_01706       LSFGYRDHNDNDDDDDDILDDVEIVDGTRGFEPEIETWEDLYEAHRREMGS
RO3G_06953       ---------------------------------------MNYPELLTN-------EVE
RO3G_13784       ---------------------------------------MNDESSLNDALVYSSD
RO3G_08863       ---------------------------------------MNDESSLNDALVYSSD
UM05862          ANQLRPSSAEPQHTPRQDSIVNSPYDTPCTOQNGTPATAEDIASFETLJAS
SjCSL1           ----STRYDRASP-ESSMQLFSLYAVNSALSSLRKDRQTNAC-----
SPCC736.08       RNNERSVGGSGPSVSMQVLSVSAVNSATLKEE-----
CNA01890         ---------------------------------------MNDESSLNDALVYSSD

Mm_RBP-Jk
---------------------------------------MNDESSLNDALVYSSD
Hs_RBP-Jk
---------------------------------------MNDESSLNDALVYSSD
Xl_SuH
---------------------------------------MNDESSLNDALVYSSD
Dr_RBPSUH
---------------------------------------MNDESSLNDALVYSSD
Hr_RBP-Jk
---------------------------------------MNDESSLNDALVYSSD
Ci_SuH
---------------------------------------MNDESSLNDALVYSSD
Dm_SuH
---------------------------------------MNDESSLNDALVYSSD
Ce_LAG-1
HVAQFPLGEPARNLDFKVIPEAMFQSVSPLAGVAAAPAQIADLQIQAL
Hs_RBP-L
---------------------------------------MNDESSLNDALVYSSD
Mm_RBP-L
---------------------------------------MNDESSLNDALVYSSD
Dr_RBPSUH
---------------------------------------MNDESSLNDALVYSSD

PcCSL2          FSNGQSKDFPPHHSYHEPMLDRRISNASGYNSDFNDEFGMGVNHGGLN
CC1G_03194       PSVPSDKDFSP--STYQDMGDRRLGMNGNGYQHNYQDDYSGINGMNGLAFP
UM06280          PSMDTMQHFQASTGHPAESFDIAGFSEQRSLASANASIASKPPSATTH
RO3G_07636       SISSPSSSL---SSLPHVSLPSFPLETTLKLED-----ENDFNTIHP
RO3G_06481       SVPSSWSSFNYQLMNHSETPSTSFSSSSLTDLKQEDDFMHLNTSRTVHP
RO3G_11583       FDDSSGTNRSRHSAVGELYHSDLNSLLEELPKRAMSREDDLTAN
RO3G_14587       HPGDSVGENEMLMVNFQSNFPRPLTQPI DLDDLQROQRQAFQTWDASSS
CNBD3370         TDSDLVMPSYPSAVQYPYNSDFMTPLSVNTGMVNPGQTRGAEQPQOS
SjCSL2           HMGWVDVAGHLPGDSGRFPATDAFSQFVLQHPDHTNYPYHLPGNAAG
SPCC1223.13      SNVPSSDSVRNSSNPQYASTKSQISPSQSONLPKASVGLTVNNYRQ
Pc6518           DIGPSQVPGQPAPPTQDVLAAQSAQSIADAATQTASITSESEPSSTSADS
CC1G_01706       GKSPP---------------------------------------MNDESSLNDALVYSSD
RO3G_06953       ---------------------------------------MNDESSLNDALVYSSD
RO3G_13784       ---------------------------------------MNDESSLNDALVYSSD
RO3G_08863       ---------------------------------------MNDESSLNDALVYSSD
UM05862          LMPTQPGSPAPPTQDVLAAQSAQSIADAATQTASITSESEPSSTSADS
SjCSL1
SPCC736.08       ---------------------------------------MNDESSLNDALVYSSD
CNA01890         ---------------------------------------MNDESSLNDALVYSSD

Mm_RBP-Jk
---------------------------------------MNDESSLNDALVYSSD
Hs_RBP-Jk
---------------------------------------MNDESSLNDALVYSSD
Xl_SuH
---------------------------------------MNDESSLNDALVYSSD
Dr_RBPSUH
---------------------------------------MNDESSLNDALVYSSD
Hr_RBP-Jk
---------------------------------------MNDESSLNDALVYSSD
Ci_SuH
---------------------------------------MNDESSLNDALVYSSD
Dm_SuH
---------------------------------------MNDESSLNDALVYSSD
Ce_LAG-1         MTFQMQQNNLFPKIDTI-----------------------------SKSPT
Hs_RBP-L          ------QSMQPFQDRLG-----------------------------RYP
Mm_RBP-L          -----------------------------SPL
Dr_RBP-L          -----------------------------SMD
PcCSL2            GFPPSGLPPPFDRLG-----------------------------RVQ
CC1G_03194        ----QSMQPFQDRLG-----------------------------RYP
UM06280           QREGHSTPTIVEEEDVPQROTRATVLRTAHTDVRQEPFGOQSHAQHLQ
RO3G_07636        SVIP---HHSPTTTTS-----------------------------SPD
RO3G_06481        LTPDNNNHSPTNNIS-----------------------------SPL
RO3G_11583        LFSSYLDLVTSPRL-----------------------------SMD
RO3G_14587        SPIQSTPARYPQGFF-----------------------------TPG
CNBD3370          MLGTYQSTVMSQYYPIPHPMEMQPTQQQIEPSVQLGTMSPSELGQRQPL
SjCSL2            YPEVYAADPTSSYSPFQDQQGFG---QYFQKSYQMDPFYITG
CC1G_01706        -----------------------------MYHLNVT
RO3G_06953        -----------------------------MNRDVLESRI
RO3G_13784        ---------------------------------------MNRDVLESRI
RO3G_08863        -----------------------------MNRDVLESRI
UM05862           AAQNFVDKLLLREDNFPQSVQQQHQFHEFTPSNGVQDASIPATTETEPAK
SjCSL1            -----------------------------MNRDVLESRI
SPPC736.08        -----------------------------MNRDVLESRI
CNA01890          -----------------------------MNRDVLESRI

Mm_RBP-Jk         -----------------------------MP
Hs_RBP-Jk         -----------------------------MP
Xi_SuH            -----------------------------MP
Dr_RBPWSUH        -----------------------------MP
Hr_RBP-Jk         -----------------------------MP
Ci_SuH            -----------------------------MP
Dm_SuH            MKSYSQFNLNAAAPPAIAYETTVWNPGSPLDPHQQQQSQDMPFGGLP
Ce_LAG-1         PELASPSAKRMRLSPSTSSHSDVASTSKGTNGQDQSKNSPTNSDPFYTIG
Hs_RBP-L          -----------------------------SPL
Mm_RBP-L          -----------------------------SMD
Dr_RBP-L          -----------------------------SMD

PcCSL2            NESRYPNSTVPLTLATSHTLSQHSPHELVGAPQATHLPSFDMMN-----
CC1G_03194        NDR---FSHPQGGVQQVQISHISPAQSTDMMIRGPBTH5SFDRDPASAGYDS-
UM06280           HHHPLQOLQPLPHSMRQOHPPHPPLQPPLPGHPLPPALHRSHSDQQLSGLCMD
RO3G_07636        HSNLFNLQENKIYNSGQKN
RO3G_06481        NNNYDNTSFHAIKFKEKDKDN
RO3G_11583        NNNYDNTSFHAIKFKEKDKDN
RO3G_14587        FLESQOEHPYDHSLSIDYGSHHFQNOEYNPLLKLEQQSEPKNLSQSGE
CNBD3370          KPTKSFDSLMMGRASSSSSLSAQDVJPDGDWNGALEDWTRPLSRPSS
SjCSL2            HSPTHLYGADAPRLQQHHQPRQFHVSQEGFHPAGTNNPNANNFR
SPPC736.123.13    ELSYQQPSMMPPYDQVQYSQQLSRTDVPNQSFQYQPTTVDDDMIP
Pc6518            VHTPASSHPDVOQTADTTANRRPPSQTPAQTTISLTTATPEDTSCSAAP
CC1G_01706        -----DIINPKRSTPVQGAPPSAVQDPNLSISQIAAASKDQGTNTGS
RO3G_06953        -----------------------------MTPMBEILNAI
RO3G_13784        SEQQFGSRKQFIQLSKPPLAPQQQDKQMSIESLTTFSHPQSN-----
RO3G_08863        -----------------------------MSESLLTFSPHQS
UM05862           SSVGDDALASLEQFMASLQARVSSSLATEAATDTHDAKAKQAEESTVKS
SjCSL1            IQQAFPSEVSEITDSELQRRQQQPQSVPPAFLSPSSIPNPKKR
SPPC736.08        ------PFPWVKDEQSIIFPVAKASNSLDLAEQNIIPAPSPGPFRNKRK
CNA01890          -----------------------------MAEDHNEGIVNSTPLLLNIAIA

Mm_RBP-Jk         SGFPQSPRTSPRKPRTTRIGALMIDSEGLSAAERPAHAPSAGKGERP
Hs_RBP-Jk         -----------------------------MAWIKRKFGERP
<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>XI_SuH</td>
<td>MQPGLPKYTPSAIQLAPVVTGKFGERP</td>
</tr>
<tr>
<td>Dr_RBPSUH</td>
<td>MAPVVTGKFGERP</td>
</tr>
<tr>
<td>Hr_RB-L</td>
<td>MDPPAGAADPSVPNLPMTLQDSRSLQSEADRSSLPTGTWTRSSP</td>
</tr>
<tr>
<td>Mm_RB-L</td>
<td>MQDORKEDRTGDLVSSTDHLYTPDCAFTHLREDQ-SFS</td>
</tr>
<tr>
<td>PcCSL2</td>
<td>FMGAPSTYVDPPLRLPASVNDMMLRLQOG----SGDLQTF</td>
</tr>
<tr>
<td>CC1G_03194</td>
<td>-----MHYMNNPNPHQDMSLRMSTAVDETLLRMLKQVAP-PIGTTSSLQTF</td>
</tr>
<tr>
<td>UM06280</td>
<td>APPTFRQNQMSYPPHMRPSLYQVQVSAEMDQLRQHIHLGVTENFNGDCSAF</td>
</tr>
<tr>
<td>RO3G_07636</td>
<td>---LIQHY</td>
</tr>
<tr>
<td>RO3G_06481</td>
<td>---LNNHTCNOIGEDMKCSVHRY</td>
</tr>
<tr>
<td>RO3G_11583</td>
<td>QKRRRRTSDPLGSSSDSSSPPTPMQVQSLGFEPQEEWLQPLQNY</td>
</tr>
<tr>
<td>RO3G_14587</td>
<td>SVTSLFPSDPASVRPQNHADSPIRRSHTTASSPHRLVHSLKIPKF</td>
</tr>
<tr>
<td>CND03370</td>
<td>GTPVTSAGSRQPLSAIDSSLPTLKTSPASSLRNPSEFLNPSNASD</td>
</tr>
<tr>
<td>SiCSL2</td>
<td>--------FMGASPTVYDPPLRLPASVNDMMLRLQOG--SGDLQTF</td>
</tr>
<tr>
<td>UM06280</td>
<td>APPTFRQNQMSYPPHMRPSLYQVQVSAEMDQLRQHIHLGVTENFNGDCSAF</td>
</tr>
<tr>
<td>RO3G_07636</td>
<td>---LIQHY</td>
</tr>
<tr>
<td>SiCSL2</td>
<td>--------FMGASPTVYDPPLRLPASVNDMMLRLQOG--SGDLQTF</td>
</tr>
<tr>
<td>CC1G_03194</td>
<td>-----MHYMNNPNPHQDMSLRMSTAVDETLLRMLKQVAP-PIGTTSSLQTF</td>
</tr>
<tr>
<td>CC1G_01706</td>
<td>QAT--------WHPPPPATSTSTGKRKLDTESEDTHSK</td>
</tr>
<tr>
<td>RO3G_06953</td>
<td>--ITTKRRKLSMIES--</td>
</tr>
<tr>
<td>RO3G_13784</td>
<td>QATATNDITIDNLLTCD</td>
</tr>
<tr>
<td>RO3G_08863</td>
<td>QATATNDITIDNLLTCD</td>
</tr>
<tr>
<td>UM05862</td>
<td>KPSQKQKPKVKKKPKKLGGKRYEIPGMPRSVVEKALSEYASI DLTT</td>
</tr>
<tr>
<td>SiCSL1</td>
<td>KAS--------FSNHQDPLIAQI RETIFDHLNLLLKKSATSNA</td>
</tr>
<tr>
<td>SPC7376.08</td>
<td>TFDDESSVEMIRRAI SDHDLNINCCGIANL</td>
</tr>
<tr>
<td>CNA01890</td>
<td>DHTHEHDHDHSVLDHLSQROKDEDSQOHASSHHRARSQEVWSRQK</td>
</tr>
<tr>
<td>Mm_RB-P-Jk</td>
<td>PPKLRTREAMRNY---LKERGQTVLIHAKVAAQKSYGN</td>
</tr>
<tr>
<td>Hs_RB-Jk</td>
<td>PPKLRTREAMRNY---LKERGQTVLIHAKVAAQKSYGN</td>
</tr>
<tr>
<td>Xi_SuH</td>
<td>PPKLRTREAMRNY---LKERGQTVLIHAKVAAQKSYGN</td>
</tr>
<tr>
<td>Dr_RBPSUH</td>
<td>PPKLRTREAMRNY---LKERGQTVLIHAKVAAQKSYGN</td>
</tr>
<tr>
<td>Hr_RB-Jk</td>
<td>PPKLRTREAMRNY---LKERGQTVLIHAKVAAQKSYGN</td>
</tr>
<tr>
<td>Ci_SuH</td>
<td>PPKKTLRDMARRY---LKERGQTVLIHAKVAAQKSYGN</td>
</tr>
<tr>
<td>Dr_SuH</td>
<td>EEEKLRTDAMMKY---MRERNDVIMLHAKVAAQKSYGN</td>
</tr>
<tr>
<td>CeLAG-1</td>
<td>Q---SLTSDRMI DFLS---NKEKYECVIHFHAKVAAQKSYGN</td>
</tr>
<tr>
<td>Hs_RB-L</td>
<td>EHTTTLQGCRVRC---LQOQCEQTVRIHAKVAAQKSYGN</td>
</tr>
<tr>
<td>Mm_RB-L</td>
<td>EHAIILREGVRTC---LQOQCEQTVWIHAKVAAQKSYGN</td>
</tr>
<tr>
<td>Dr_RB-L</td>
<td>SLLKVTDRSVOY---LQFRPADSTLHAKVAAQKSYGN</td>
</tr>
<tr>
<td>PcCSL2</td>
<td>IRPYLDQY1RTPN--RLLAFERTIVMSKKVAQKSYGT</td>
</tr>
<tr>
<td>CC1G_03194</td>
<td>IRPFLEQYVRTP---RLAFERTIVMSKKVAQKSYGT</td>
</tr>
<tr>
<td>UM06280</td>
<td>IQDYVRQY1STSS--RLFLGERNLMTTTRVAAKSYGN</td>
</tr>
<tr>
<td>RO3G_07636</td>
<td>LSTKO---GEKLI LTSDKVAQKSYGN</td>
</tr>
<tr>
<td>RO3G_06481</td>
<td>LITKO---DERKIT LTSDKVAQKSYGN</td>
</tr>
<tr>
<td>RO3G_11583</td>
<td>LLQKQSSREDYI---GERIMT LTSDKVAQKSYGN</td>
</tr>
<tr>
<td>RO3G_14587</td>
<td>IOTYLAHIHOTA---ATOLGEKTVLTLSDKVAQKSYGN</td>
</tr>
<tr>
<td>CNBD3370</td>
<td>TDSIIKQY1RAPN--RLLAFERGTVIVMSKKVAQKSYGT</td>
</tr>
<tr>
<td>SjCSL2</td>
<td>DRNNLFQVRKVL---RHPELLCAMKVFMPSLQKSYGK</td>
</tr>
<tr>
<td>SPC7222.13</td>
<td>QPVITQERQIYQVL--SNOSLASYPCLMPLSQQKSYGT</td>
</tr>
<tr>
<td>Pc6518</td>
<td>I-RRRVDRHVSQHT---PRPITMPTTVLCHAAVAAQKSYGT</td>
</tr>
<tr>
<td>CC1G_01706</td>
<td>I-RRIVRDHVSQDP---ARVPMTTYIVLHAAVAAQKSYGS</td>
</tr>
<tr>
<td>RO3G_06953</td>
<td>-MI EKRQLNSSP---LSPKQPNKVTCYHAVIAAQS</td>
</tr>
<tr>
<td>RO3G_13784</td>
<td>VK------EPTR---KRAKKTSMVTCYYGAAAQKSYGS</td>
</tr>
<tr>
<td>RO3G_08863</td>
<td>MMRKWAESLSSV---KLLKKSKVCYGAQKSYGS</td>
</tr>
<tr>
<td>UM05862</td>
<td>KPSQKQKPKVKKKPKKLGGKRYEIPGMPRSVVEKALSEYASI DLTT</td>
</tr>
<tr>
<td>SjCSL1</td>
<td>LCNRIRDP-LNQY---NQANKLVTI SYQHASVAAQKSYGS</td>
</tr>
</tbody>
</table>
SPCC736.08       NEDSVHKISLTRS-------------GKPSQLVTVSCRHSSVIQKSYGS
CNA01890         VPKPLGSQAMFYPGFAPADKAPGLDENDNNSHFRNFTLEVHTPTTGQKSYGK

Mm   RBP-Jk        EKRFFCPPPPCVYLMGSGWKKE-----------------------------
Hs   RBP-Jk        EKRFFCPPPPCVYLMGSGWKKE-----------------------------
Xi   SuH           EKRFFCPPPPCVYLMGSGWKKE-----------------------------
Dr   RBP5SUH       EKRFFCPPPPCVYLMGCGWKKE-----------------------------
Hr   RBP-Jk        EKRFLCPPPCYLVGLMSGWKQKOQ-----------------------------
Ci   SuH           EKRFFCPPPPCMYLLGNGWKRKQKOQ-----------------------------
Dm   SuH           EKRFFCPPPCLYLFGSGWRRYK-----------------------------
Ce   LAG-1         EKRFFCPPPICYLQGQWLKDKDRVATYQLKTLKAA----------SAQKDAA
Hs   RBP-L         EKRFFCPPPCVYLSPGWVRKPG-----------------------------
Mm   RBP-L         EKRFFCPPPPCVYLPGWVRKPM-----------------------------
Dr   RBP-L         EKRFFCPPPPCVYLGHWVRQOE-----------------------------
PcCSL2           EKRFLCPPPTAIIMNSWWSVDRH-------------RGEPEPCLPP
CC1G_03194       EKRFLCPPPTAIIMNSWWSVDESTR-------------RGEPEPCLPP
UM05280          EKRFLCPPPLVIILGSSWVNWCPATRPSFPAPAGADPDQAAPTNLSSPP
RO3G_07636       EKRFLCPPPSTILSGTGWVTAKQ-----------------------------HPP
RO3G_06468       EKRFLCPPPIILSSGNNSWSTON-----------------------------NPP
RO3G_11583       EKRFLCPPPATIMKGINTTSKTDKTTPSL------------FHPSNALSQP
RO3G_14587       EKRFLCPPPTAIILVGTSSWTVITKEIKDQIKEELRIDEPSLEKDI LLAPP
CNBD3370         EKRFLVHQPQAIISGGGSWAKPSG-----------------------------GCPVSPLQQP
SjCSL2           ERRYICPPAVVYLLGSSWFRCPLDKN-----------------------------
SPCC1223.13      ERYLCPPLVYLLGTWTLLNVTNDN-----------------------------LK
Pc6518           EKRFLCPPPIVHIEGPWLNKMQ-----------------------------
CC1G_01706       EKRFLCPPPIVHIEGPWHLRTQ-----------------------------
RO3G_06953       EKRFLCPPPIVSGSQ---TNC-----------------------------
RO3G_13784       EKRFLCPPLVQVHST--RNL-----------------------------
RO3G_08863       EKRFLCPPPQOKLQAN-----------------------------KEM
UM05862          EKRFFNPPIISITGGLRKYAEHG-----------------------------AA
SjCSL1           EKRYLCPPPLISITGNTGSILGHS-------------------------F
SPCC736.08       EKRYLCPPPMV3INGNYYSSIFNQO-----------------------------F
CNA01890         ERRMLNPPPIRLLSGPASSCISITMVSVSPN-----------------------------ITQPV

Mm   RBP-Jk        QMERDCSEQESQPACAFIIGNSQDMQQLNLEG-----------------------------
Hs   RBP-Jk        QMERDCSEQESQPACAFIIGNSQDMQQLNLEG-----------------------------
Xi   SuH           QMERDCSEQESQPACAFIIGNSQDMQQLNLEG-----------------------------
Dr   RBP5SUH       QMERDCSEQESQPACAFIIGNSQDMQQLNLEG-----------------------------
Hr   RBP-Jk        QMERDCSEQESQPACAFIIGNSQDMQQLNLEG-----------------------------
Ci   SuH           QMERDCSEQESQPACAFIIGNSQDMQQLNLEG-----------------------------
Dm   SuH           QMERDCSEQESQPACAFIIGNSQDMQQLNLEG-----------------------------
Ce   LAG-1         IENDPIHEQQATELVAYIGIGSDTSERQQLDFSTGKVRHPG-----------------------------
Hs   RBP-L         QDQAHQAGETGPTVCGYMGLDSASGSAPETQKLN-----------------------------
Mm   RBP-L         QDQAHQAGETGPTVCGYMGLDSASGSAPETQKLN-----------------------------
Dr   RBP-L         QDQAHQAGETGPTVCGYMGLDSASGSAPETQKLN-----------------------------
PcCSL2           RVVVSISGEPAPQEGSEQRAPIDHNS-----------------------------
CC1G_01706       RVVVSISGEPAPQEGSEQRAPIDHNS-----------------------------
UM05280          RVVVSISGEPAPQEGSEQRAPIDHNS-----------------------------
RO3G_07636       NLTIQSIGEKLQHQTIDWYKGDIWLDLOPSAVLQQNQSN-----------------------------
RO3G_06468       NVHIGSLQGQPRYCHIHWYEGITLNLQHALATYASQGI-----------------------------
RO3G_11583       KLTHISGETITQGTVIWSOTQSGNI-DDDNNNAQK-----------------------------
RO3G_14587       KLTVISGETSTOAGLQLEWTVSAGTVQGTQGIFKPPKTESTFRSSES
CNBD3370         RINISLTGQQAKIISWDTLNGKNDMDEKASTQGGKVDQDQ-----------------------------
SjCSL2           IIAADDDPPNLTSETPTFYTSSTSSDSATNLLSLQGQKLDDPL-----------------------------
PSC1232.13       IISAQTLDEKDNPPFAKNFFYNADGALISPETDIKSSTYQLTNN-----------------------------
Pc6518           QFSMTWSEGGERSFEQAPIIDHNL-----------------------------
CC1G_01706       SLSMAWSETGERSFQKATLDNNM-------------------------------
RO3G_06953       MISMTI MNENNQAI LEQRTLDDDESK------------------------
RO3G_13784       LLMSI LCED-QGLEQRAI LDENQR------------------------
RO3G_08863       ELMSI LCED-NDI LEQHTALDEHS------------------------
UM05862          TLTPSSEDPEEDOCRVLWVRGDTDYFSNENVAVLGD------------------------
SjCSL1           SVQLAITEQGQHSNQVSENFAERQS------------------------
SPCC736.08       RTEISI MNDFGQCSQPI SEETYTGQGC------------------------
CNA01890         AI CAPQTHHMGRPDPD TPHLGTPVGDVKK KERREEQK KRTALNAGFAA

Mm_RBP-Jk        -----------------KNYCTAKTLY1 SDSA--KRK----HFMLSVKMF
Hs_RBP-Jk        -----------------KNYCTAKTLY1 SDSA--KRK----HFMLSVKMF
Xl_SuH           -----------------KNYCTAKTLY1 SDSA--KRK----HFMLSVKMF
Dr_RBPSUH        -----------------KNFCTAKTLY1 SDSA--KRK----HFMLSVKMF
Hr_RBP-Jk        -----------------KNFCTAKTLY1 SDDT--KRK----HFMLCVMKF
Ci_SuH           -----------------KNFCTAKTLY1 SDDT--KRK----HFMLNVMKF
Dm_SuH           -----------------KQYCAAKTLFI SDSA--KRK----HFMLSVKMF
Ce_LAG-1         -----------------DQRQDPNI YDYCAA KTLY1 SDSA--KRK----YFDLNAQFF
Hs_RBP-L         -----------------FEQQPSREFGCAKTLY1 SDSA--KRK----HFRLVLRLV
Mm_RBP-L         -----------------FEQQPSREFGCAKTLY1 SDSA--KRK----HFRLVLRLV
Dr_RBP-L         -----------------FEQQPSREFGCAKTLY1 SDSA--KRK----HFRLVLRLV
PcCSL2           -----------------YI GRCGKQLF1 SDVD-EKKKV EALVKI MAPSA
CC1G_03194       -----------------YLGRCV GQKQL FI SDVD-EKKKVE ALVKVTAPAS
UM06280          -----------------I SGRCI GKQLY1 SDVD- EKRRHV EALV AI SVPGS
RO3G_07636       -----------------LIGNCV SKQL ALH VGGAD-EKRKRVEVLAKIQLNG
RO3G_06481       -----------------LVGNCSKQL HI SDDT- EKRKSVQAHVSI QLDOS
RO3G_11583       -----------------VFGRCI SKQLY1 NDAD- EKRRVEVLAKIQLNG
RO3G_14587       RHPPADAYSNERQELLAAGKSVKHLY1 HDAD- EKRRVECLVKQLANG
CNBD3370         -----------------PFTGVSA GKLNI SDDN- PKKEVKVALVT KAPL N
SjCSL2           -----------------QQQSPSPI WANTA KTLY1 YSG--KGDHNTYGRSTTLQVH--
SPCC1223.13      -----------------ENTNFDSFPWGNALLKTI YYTG--QKNDFGFRSTFLQ---
 Pc6518           -----------------ASFKFHLVTGSA--KSFSQLSLNI TEPS P
CC1G_01706       -----------------SSFKFHLVTGTA--KAKSFQLSDLIAEAPP
RO3G_06953       -----------------GYFKLYVTSSA--KSKQFRI KLMN------
RO3G_13784       -----------------GSFRYHLVTGTA--KAKQFKLQNLLOPQ--
RO3G_08863       -----------------GSFRYHLVTGTA--KAKQFKLQNLLOQNE--
UM05862          -----------------RLET KMRSYVTPTG-- RSKSFRLQNILRPSG
SjCSL1           -----------------VAFRSLHI SS SI--A AKA KSFNLNI DVI S--
SPCC736.08       -----------------MI FRSLHI SSSLV--AASKNRLRL SDLMFS--
CNA01890         TLPKRNNA TDLK DRTL LKDLGTFPGWLI GEEA- GKMKEF RLE KI YAEPT

Mm_RBP-Jk        YGN S------------------------
Hs_RBP-Jk        YGN S------------------------
Xl_SuH           YGN S------------------------
Dr_RBPSUH        YGN S------------------------
Hr_RBP-Jk        FGS G------------------------
Ci_SuH           FGGGG------------------------
Dm_SuH           YGN G------------------------
Ce_LAG-1         YGG C------------------------
Hs_RBP-L         LRGG------------------------
Mm_RBP-L         LRGG------------------------
Dr_RBP-L         HNG C------------------------
PcCSL2           DDEPE------------------------
CC1G_03194       DDEPE------------------------
UM06280          SPMDR------------------------
RO3G_07636       TP------------------------
RO3G_06481       RS------------------------
RO3G_11583       SN------------------------
RO3G_14587       LQ------------------------

CNBD3370  KHAPOGWGASK---------------------------
SjCSL2    -VRTP-------------------------------
SPCC1223.13  SVOQK--------------------------
Pc6518    PTADGSEP---------------------------
CC1G_01706  PSLNPEGSE------------------------
RO3G_06953  ----------------------------------
RO3G_13784  ----------------------------------
RO3G_08863  ----------------------------------
UM05862    LEPILPSPLKKSRAADALSPAM-----------
SjCSL1    -----------------------------------
SPCC736.08  ----------------------------------
CNA01890  EKGLSSPSTQQYQLGNPNDHETAHTFADMLEPPSMDALQPLAEA

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  -----------------------------------
RO3G_07636  -----------------------------------
RO3G_06481  -----------------------------------
RO3G_11583  -----------------------------------
RO3G_14587  -----------------------------------
CNBD3370  -----------------------------------
SjCSL2    -----------------------------------
SPCC1223.13  -----------------------------------
Pc6518    -----------------------------------
CC1G_01706  -----------------------------------
RO3G_06953  -----------------------------------
RO3G_13784  -----------------------------------
RO3G_08863  -----------------------------------
UM05862    SIQPANEPIENGEVALUELPK-----------
SjCSL1    -----------------------------------
SPCC736.08  -----------------------------------
CNA01890  VQQSSQGHDTLSEHPEVPEASNLHNDNDAENPGNSPLITQATTE

Mm RBP-Jk  -DDIGVFLSKRIKVISKPSKKQSLKN---AD---LCIASGTKVALFNRL
Hs RBP-Jk  -DDIGVFLSKRIKVISKPSKKQSLKN---AD---LCIASGTKVALFNRL
Xl_SuH    -DDIGVFLSKRIKVISKPSKKQSLKN---AD---LCIASGTKVALFNRL
Dr_RBPSUH  -ADIGVFLSKRIKVISKPSKKQSLKN---AD---LCIASGTKVALFNRL
Hr RBP-Jk  -LEVGAFNSRRIKVISKPSKKQSLKN---AD---LCIASGTKVALFNRL
Ci_SuH    -ADVGQFSSKRIKVISKPSKKQSLKN---AD---LCIASGTKVALFNRL
Dm_SuH    -HDIGVFNSKRIKVISKPSKKQSLKN---AD---LCIASGTKVALFNRL
Ce_LAG-1  -MEIGVFNSQRIKVISKPSKKQSMKN---TDCKYLCIASGTKVALFNRL
Hs RBP-L  -RELGTFHSRLIKVISKPSQKKQSLKN---AD---LCIASGTKVALFNRL
Mm RBP-L  -QEIETFHSRLLIKVISKPSQKKQSLKN---AD---LCIASGTKVALFNRL
Dr RBP-L  -QEIETFNSRLLIKVISKPSQKKQSMKN---AD---LCIASGTKVALFNRL
PcCSL2    -RVIGTFPSRPIKVISKPSKKQSAKN---LE---LCINHGSTLHFR
CC1G_03194  -RIIGVFPSRPIKVISKPSKKQSAKN---LE---LCINHGSTLHFR
<table>
<thead>
<tr>
<th>Accession</th>
<th>Description</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>UM06280</td>
<td></td>
<td>-RLLGTFSPRPVI KVI SKPSKKQSSSRN---TE---LCVNHGSVI SLFHRL</td>
</tr>
<tr>
<td>RO3G_07636</td>
<td></td>
<td>---GI FIHFSPI KVI SKPSKKQSSVRK---MD---LCH HHTTValFNRV</td>
</tr>
<tr>
<td>RO3G_06481</td>
<td></td>
<td>-LGTDFSPI KVI SKPSKKQSAKD---ME---LCH HSTVALFNRT</td>
</tr>
<tr>
<td>RO3G_11583</td>
<td></td>
<td>-LGTFSSGKPVI KVI SKPSKKQSAKN---ME---LCH HHGTTLFNR</td>
</tr>
<tr>
<td>RO3G_14587</td>
<td></td>
<td>-LGPLASKAI KVI SKPSKKQSKQI KN---ME---LCH HHGTTSFNLRI</td>
</tr>
<tr>
<td>CNBD3370</td>
<td></td>
<td>DNVFGIFESKDI KI KI SKPSKKKSTAKS---GE---LTI SHHTTVALFNRI</td>
</tr>
<tr>
<td>SjCSL2</td>
<td></td>
<td>-QKRITMDKLRI GI KPSKKQSKMMF---SD---LNI CHGDCVSFLNRF</td>
</tr>
<tr>
<td>SPPC1223.13</td>
<td></td>
<td>-TKYFKLENLRGLVI KPSQKRAMLK---SD---MSI RHGDCVLFLNRY</td>
</tr>
<tr>
<td>Pc6518</td>
<td></td>
<td>RLPWASFDSAPVTI KPSKKK---TAKT-RNI A---SCI LAGPGVSLNRI</td>
</tr>
<tr>
<td>CC1G_01706</td>
<td></td>
<td>GRWATFSAPVTI KSKPSKK---TAKT-RNI S---SCI LAGPGVSLNRI</td>
</tr>
<tr>
<td>RO3G_06953</td>
<td></td>
<td>---DALLFYSNPI SII KPSKKK---TAKT-RNAT---SCLFNHSTI FNR</td>
</tr>
<tr>
<td>RO3G_13784</td>
<td></td>
<td>ESPLATFKTPITI KSKPSK---IGKTARGTT---SCI STEAVPLFSRI</td>
</tr>
<tr>
<td>RO3G_08863</td>
<td></td>
<td>-SSFASFLSKPI SII KPSKK---TAKTSRNTS---TCI LANSVPFLNRI</td>
</tr>
<tr>
<td>UM05862</td>
<td></td>
<td>ELAWASFESAPVTI KSKSSK---TAKP-R---GE---LNQVKNGSML SLYNRI</td>
</tr>
<tr>
<td>SjCSL1</td>
<td></td>
<td>-DVLAQMVTPKI NI SII KPSKKK---GSKN-RVS---STLMSGAI SLYNRI</td>
</tr>
<tr>
<td>SPPC736.08</td>
<td></td>
<td>NQLLSHLVTSSISI VSKPSKK---GSKL-KI SN---ITLRSGVSSLYNRI</td>
</tr>
<tr>
<td>CNA01890</td>
<td></td>
<td>EQLPGFLSGPVRPSK---TNTK-TSLA---ICFPRDSVFLWTRI</td>
</tr>
</tbody>
</table>

**Mm RBP-Jk**
- RSQTVSTRYLHVEGGN

**Hs RBP-Jk**
- RSQTVSTRYLHVEGGN

**Xl SuH**
- RSQTVSTRYLHVEGGN

**Dr RBPSUH**
- RSQTVSTRYLHVEGGN

**Hr RBP-Jk**
- RSQTVSTRYLHVEGGN

**Ci SuH**
- RSQTVSTRYLHVEGGN

**Dm SuH**
- RSQTVSTRYLHVEGGN

**Ce LAG-1**
- RSQTVSTRYLHVEGGN

**Hs RBP-L**
- RSQTVSTRYLHVEGGN

**Mm RBP-L**
- RSQTVSTRYLHVEGGN

**Dr RBP-L**
- RSQTVSTRYLHVEGGN

**PcCSL2**
- RSQTVSTKYLCVSGSSFKGDSGAPLM-G-LDQ---RARTQTP---SFI

**CC1G_03194**
- RSQTVSTKYLCVSGSSFKGDSGAPLM-G-LDQ---RDSQSTP---SFI

**UM06280**
- RSQTVSTKYLCVSGAPWSFKGDSNQPFL-N-MNP---RDQKGSERPSCF

**RO3G_07636**
- RAQTVSTKYLCVSGAPWSFKGDSNQPFL-N-MNP---RDQKGSERPSCF

**RO3G_06481**
- RSQTVSTKYLCVSGAPWSFKGDSNQPFL-N-MNP---RDQKGSERPSCF

**RO3G_11583**
- RSQTVSTKYLCVSGAPWSFKGDSNQPFL-N-MNP---RDQKGSERPSCF

**RO3G_14587**
- RSQTVSTKYLCVSGAPWSFKGDSNQPFL-N-MNP---RDQKGSERPSCF

**CNBD3370**
- KSQTVSTKYLCVSGAPWSFKGDSNQPFL-N-MNP---RDQKGSERPSCF

**SjCSL2**
- RSQTVSTKYLCVSGAPWSFKGDSNQPFL-N-MNP---RDQKGSERPSCF

**SPPC1223.13**
- RSQTVSTKYLCVSGAPWSFKGDSNQPFL-N-MNP---RDQKGSERPSCF

**SPPC736.08**
- RSQTVSTKYLCVSGAPWSFKGDSNQPFL-N-MNP---RDQKGSERPSCF

**CNA01890**
- RSQTVSTKYLCVSGAPWSFKGDSNQPFL-N-MNP---RDQKGSERPSCF

**Mm RBP-Jk**
- ASSQQWGAFYI HLLDDDDE---AG---SEG

**Hs RBP-Jk**
- ASSQQWGAFFI HLLDDDDE---AG---SEG

**Xl SuH**
- ASSQQWGAFFI HLLDDDDE---AG---SEG

**Dr RBPSUH**
- ASSQQWGAFFI HLLDDDDE---AG---SEG

**Hr RBP-Jk**
- ASSQQWGAFFI HLLDDDDE---AG---SEG

**Ci SuH**
- ASSQQWGAFFI HLLDDDDE---AG---SEG

**Dm SuH**
- ASSQQWGAFFI HLLDDDDE---AG---SEG

**Ce LAG-1**
- ASSQQWGAFFI HLLDDDDE---AG---SEG

---QET
Hs_RBP-L       ASARQWAAFTLHLADG -HSA---------------------------QGD
Mm_RBP-L       ASARQWAAFTLHLADG -HCS---------------------------QGD
Dr_RBP-L       ASARQWTAFTVNLVEETQAV---------------------------HSE
PcCSL2         ARTASWDPI MYI VDVKPRTG IDTPF----PPPQPEFSPSPNPAPF
CC1G_03194     ARTASWDPI MYI VDVKPRTG IDTPF----PPPQPEFSPSPNPAPF
UM06280        AKMSSWDPI IYLVDPPKRADIATPSP---VQPPVKGYPPPPPAPLP
RO3G_07636     TRTTSWDPFI YV DLRSPNTPSPV------PFSHPTDI DHPPPPAPAI
RO3G_06481     ARTTWSWDPI WVDPTRPSNTPS---STLQHANQYPPPPAPAI
RO3G_11583     SRTGWDPFVI WVDTSCSPNTAN------RPKHNLNPNYPPPAPAI
RO3G_14587     ARTTWSWDPI WVDTSASEEGETPEDI YGHVFARSTPPPPPAPL
CNBD3370       ADASFWESFI IWLVDPSLPSPSNHHAP------NPDPWPPPAPANI
SjCSL2         TTNTWEPFNI YSVEELE---------------------------NK
SPCC1223.13    MANNLWEPFYI FSVDELNKG---------------------------NNS
Pc6518         ASNVWAFNSVWRP---------------------------QGDO
CC1G_01706     ASNIGWASFNVNWV---------------------------RPDG
RO3G_06953     AKHSTWSPFEI IVLR---------------------------LQP
RO3G_13784     AKHISWPSFDI LLLRPQOQPTVLRTTTK------TSKRYNTRALQ
RO3G_08863     AKNTTSWPSFDI LWIVDPTRSPNTS---SPRNYNTSLQ
UM05862        VHSNGDSTYRI TLI SRPP---------------------------LA
SjCSL1         LRNESWTPLQI RYVGS---------------------------LA
SPCC736.08     LRGDRWVLRI NLLLD---------------------------ENGKLKV
CNA01890       SRTAKWTPRFEVLDHANAPFRTARN-----------------KFD
Mm_RBP-Jk      EEFTVIRDGYI HYGOTVKLVCSTGMLPRLI I RKVDKTALLDA------
Hs_RBP-Jk      EEFTVIRDGYI HYGOTVKLVCSTGMLPRLI I RKVDKTALLDA------
Xl_SuH         EEFTVIRDGYI HYGOTVKLVCSTGMLPRLI I RKVDKTALLDA------
Dr_RBP-Jk      EEFTVIRDGYI HYGOTVKLVCSTGMLPRLI I RKVDKTALLDA------
Gi_SuH         EEFSVIRDGYI HYGOTVKLVCSTGMLPRLI I RKVDKTALLDA------
Dm_SuH         EEFOVIRDGYI HYGATVKLVCSTGMLPRLI I RKVDKOMALLEA--
Ce_Lag-1       DNFARVDFWYI GYSVKLVDSTGALPRLI I RKVDKQVI DLAS------
Hs_RBP-L       --FFPREGYRVLQELCSTVGI TLPPMI I RKVAQCCALLDV------
Mm_RBP-L       --FFPREGYRVLQELCSTVGI TLPPMI I RKVAQCCALLDV------
Dr_RBP-L       --TMCSYFI CYGVQVLCTDGSQVA PPMPVI RKVNKOQACLDV------
PcCSL2         T-NNGSQIPI YY NOTQLCSTGQVPLVI I RKVDHTOO TTGGQLEGA------
CC1G_03194     T-NNGSQIPI YY NOTQLCSTGQVPLVI I RKVDHTQ TVVGGQLEGA------
UM06280        TGLTSSQMTI IYNQPI VLOCLNTA VWSVPLI V RKEGGTVLQGASPGS------
RO3G_07636     I QNHQGSLAI HYNQPVVLQCVSTGLVSPVI I RRVEKGS MV MGNNVRNDL------
RO3G_06481     VQNPKQVLHYLNHLI QLCOVSTGLPVI MMVRADKGKSVMGNNVRNDP------
RO3G_11583     LQTS-STLAI HYNQPVVLQCVTTGLVSPVI I RKVDQQLV LGGNVRNDP------
RO3G_14587     KNKTGGPVPI HYNQHVLQCLT QLQLGLVSPVI I RKVDRAST VG A---ARDD------
CNBD3370       I PNTISHI RNYSTVVLQSLRTGV I SPTLVWRRI ETDSDA VGMGDP-----
SjCSL2         SSSYDRTVNCISNR I I QSQTGVRSPPI I RYENKALVVEDD-----
SPCC1223.13    NPSDSSRKLCSNMMVI I LVSITGVQSPPLI L KKHDNW KVSLSRSPAP-----
Pc6518         PTVLQGVPQVTGVEI LDTDSLQI NTSPTLI RKVDGKRI APDDG-----
CC1G_01706     SPNTNQQPQVTYGEI VLDTSQSI STSPLI RKVDGKRSVDGDD-----
RO3G_06953     --QPNIKQEEYI TEIIQDGTI TSPIPLI RKVEKNNV YYDAF-----
RO3G_13784     FSTPSAALHYMGSEI I RRESHTGVSSAPLMV CKVDGRVIPDAC-----
RO3G_08863     LPKTP-VSYLTYGEI I RRESRTG SSLPPI LCKVDGRVLESAY-----
UM05862        ELAGAESVTDYSTI VLTDVDNQFSSDPVLVCKVDGQQLQPLDPPSV------
SjCSL1         --EHSFQIPI MYGOSQWLYNRSQG I VSPLI HRVDKDIQ APDD-----
SPCC736.08     CDDVDPENPI KGYI VE VDEATGTTSPLI I RRVEKDI AEEED-----
CNA01890       ETDHLEGQQTYGSI VLVLDQTVKSDPVKVRVESGKSVVGQYD-----
Dr_RBPSUH
Hr_RB-P-Jk
Ci_SuH
Dm_SuH
Ce_LAG-1
Hs_RB-P-L
Mm_RB-P-L
Dr_RB-P-L
PcSL2

KGIADHYCPGEVC

CC1G_03194

KGIADHYCPGEVC

UM06280

ESIAR----EAF

RO3G_07636

SYPTGGEWG-DEAL

RO3G_06481

PCSTGGEYG-DEAL

RO3G_11583

IGSLGGECSDETL

RO3G_14587

VSGSGGEGF-DEVL

CNBD3370

HVHEVPAMPPLGA

SjCSL2

QL

Pc6518

CC1G_01706

RO3G_06953

RO3G_13784

RO3G_08863

UM05862

LLKAGTGKSRVPI DPALRDKLATS VSVSHANPSTSLGAGASTSNHIE

SjCSL1

SPCC736.08

CNA01890

Mm_RB-P-Jk

Hs_RB-P-Jk

Xl_SuH

Dr_RBPSUH

Hr_RB-P-Jk

Ci_SuH

Dm_SuH

Ce_LAG-1

Hs_RB-P-L

Mm_RB-P-L

Dr_RB-P-L

PcSL2

CC1G_03194

UM06280

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

- GQPVS

<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mm_RBP-Jk</td>
<td>QLKHCAYFLK---------DTERMYL</td>
</tr>
<tr>
<td>Hs_RBP-Jk</td>
<td>QLKHCAYFLK---------DTERMYL</td>
</tr>
<tr>
<td>Xi_SuH</td>
<td>QLKHCAYFLK---------DTERMYL</td>
</tr>
<tr>
<td>Dr_RPBSUH</td>
<td>QLKHCAYFLK---------DTERMYL</td>
</tr>
<tr>
<td>Hr_RBP-Jk</td>
<td>QLKHCAYFLK---------DSERMYL</td>
</tr>
<tr>
<td>Ci_SuH</td>
<td>QLKHCAYFLK---------DTERMYL</td>
</tr>
<tr>
<td>Dm_SuH</td>
<td>QLKHCAYFLK---------DTERMPL</td>
</tr>
<tr>
<td>Ce_LAG-1</td>
<td>QLKHCAYQLMI---------DTERMYL</td>
</tr>
<tr>
<td>Hs_RBP-L</td>
<td>QLKHCAYQFPGS---------PPGGGTYL</td>
</tr>
<tr>
<td>Mm_RBP-L</td>
<td>QLKHCAYQFSD---------TPPGAGTYL</td>
</tr>
<tr>
<td>Dr_RBP-L</td>
<td>QLKHCAYQFQR---------GSDHMYL</td>
</tr>
<tr>
<td>PcCSL2</td>
<td>QLKHAIFEYV---------PNKGAFE---------PGTPGVSGAFLS</td>
</tr>
<tr>
<td>CC1G_03194</td>
<td>QLKHAIFEYV---------PNKGGLPE---------PGTPGVSGAFLS</td>
</tr>
<tr>
<td>UM06280</td>
<td>QLKHAILEVLEDPAASPRPGDSG----------SNWPGNSGPFLA</td>
</tr>
<tr>
<td>RO3G_07636</td>
<td>QLKHCAYFQIPSDF---------TPGGAGTYL</td>
</tr>
<tr>
<td>RO3G_06481</td>
<td>QLKHCAYFQFR---------GSDHMYL</td>
</tr>
<tr>
<td>RO3G_14587</td>
<td>QLKHCAYFQFQIR---------GSDHMYL</td>
</tr>
<tr>
<td>CNBD3370</td>
<td>QLKHAIFEYV---------PNKGAFE---------PGTPGVSGAFLS</td>
</tr>
<tr>
<td>SJCSL2</td>
<td>QLKHAIFEYV---------PNKGAFE---------PGTPGVSGAFLS</td>
</tr>
<tr>
<td>SPCCL2</td>
<td>QLKHAIFEYV---------PNKGAFE---------PGTPGVSGAFLS</td>
</tr>
<tr>
<td>PC6518</td>
<td>QLKHCAYFQIPSDF---------TPGGAGTYL</td>
</tr>
<tr>
<td>CC1G_01706</td>
<td>QLKHCAYFQIPSDF---------TPGGAGTYL</td>
</tr>
<tr>
<td>RO3G_06953</td>
<td>QLKHCAYFQIPSDF---------TPGGAGTYL</td>
</tr>
<tr>
<td>RO3G_1587</td>
<td>QLKHCAYFQIPSDF---------TPGGAGTYL</td>
</tr>
<tr>
<td>CNBD3370</td>
<td>QLKHCAYFQIPSDF---------TPGGAGTYL</td>
</tr>
</tbody>
</table>

- GQPVS

<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mm_RBP-Jk</td>
<td>QLKHCAYFLK---------DTERMYL</td>
</tr>
<tr>
<td>Hs_RBP-Jk</td>
<td>QLKHCAYFLK---------DTERMYL</td>
</tr>
<tr>
<td>Xi_SuH</td>
<td>QLKHCAYFLK---------DTERMYL</td>
</tr>
<tr>
<td>Dr_RPBSUH</td>
<td>QLKHCAYFLK---------DTERMYL</td>
</tr>
<tr>
<td>Hr_RBP-Jk</td>
<td>QLKHCAYFLK---------DTERMYL</td>
</tr>
<tr>
<td>Ci_SuH</td>
<td>QLKHCAYFLK---------DTERMYL</td>
</tr>
<tr>
<td>Dm_SuH</td>
<td>QLKHCAYFLK---------DTERMPL</td>
</tr>
<tr>
<td>Ce_LAG-1</td>
<td>QLKHCAYQLMI---------DTERMYL</td>
</tr>
<tr>
<td>Hs_RBP-L</td>
<td>QLKHCAYQFPGS---------PPGGGTYL</td>
</tr>
<tr>
<td>Mm_RBP-L</td>
<td>QLKHCAYQFSD---------TPPGAGTYL</td>
</tr>
<tr>
<td>Dr_RBP-L</td>
<td>QLKHCAYQFQR---------GSDHMYL</td>
</tr>
<tr>
<td>PcCSL2</td>
<td>QLKHAIFEYV---------PNKGAFE---------PGTPGVSGAFLS</td>
</tr>
<tr>
<td>CC1G_03194</td>
<td>QLKHAIFEYV---------PNKGGLPE---------PGTPGVSGAFLS</td>
</tr>
<tr>
<td>UM06280</td>
<td>QLKHAILEVLEDPAASPRPGDSG----------SNWPGNSGPFLA</td>
</tr>
<tr>
<td>RO3G_07636</td>
<td>QLKHCAYFQIPSDF---------TPGGAGTYL</td>
</tr>
<tr>
<td>RO3G_06481</td>
<td>QLKHCAYFQFR---------GSDHMYL</td>
</tr>
<tr>
<td>RO3G_14587</td>
<td>QLKHCAYFQFQIR---------GSDHMYL</td>
</tr>
<tr>
<td>CNBD3370</td>
<td>QLKHAIFEYV---------PNKGAFE---------PGTPGVSGAFLS</td>
</tr>
<tr>
<td>SJCSL2</td>
<td>QLKHAIFEYV---------PNKGAFE---------PGTPGVSGAFLS</td>
</tr>
<tr>
<td>SPCCL2</td>
<td>QLKHAIFEYV---------PNKGAFE---------PGTPGVSGAFLS</td>
</tr>
<tr>
<td>PC6518</td>
<td>QLKHCAYFQIPSDF---------TPGGAGTYL</td>
</tr>
<tr>
<td>CC1G_01706</td>
<td>QLKHCAYFQIPSDF---------TPGGAGTYL</td>
</tr>
<tr>
<td>RO3G_06953</td>
<td>QLKHCAYFQIPSDF---------TPGGAGTYL</td>
</tr>
<tr>
<td>RO3G_1587</td>
<td>QLKHCAYFQIPSDF---------TPGGAGTYL</td>
</tr>
<tr>
<td>CNBD3370</td>
<td>QLKHAIFEYV---------PNKGAFE---------PGTPGVSGAFLS</td>
</tr>
</tbody>
</table>
RO3G_13784       SCPSIDQTWLDSR------------------------
RO3G_08863       SSEQDHTQTWLDSY-------------------
UM05862          GMVPGGWMIHISPKVGQQLAVA-----FAQTHEEALGVQQLPAHTSDKH
SjCSL1           TTGMDFFPVELEPA-------------------
SPCC736.08       ASSEAILPIEWEAV-------------------
CNA01890         EASPRIQRRQRANAKAASQLKELNEERGKEKEIEGTSHDEPHLTSSFLQ

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       PGS---PMSPNGASAAYFQATGANENEPASPS-------A
UM06280          PTGGYLSMSPTSSAAAYAAAQTRYLSAGRAGATMMPSPASTGHSVEAME
RO3G_07636       --------------------------------------
RO3G_06481       --------------------------------------
RO3G_11583       --------------------------------------
RO3G_14587       --------------------------------------
CNBD3370         LPMTPHTN------------------NMNLSPSSPVSILSDYFNQ
SjCSL2           --------------------------------------
SPCC1223.13      --------------------------------------
Pc6518           --------------------------------------
CC1G_01706       --------------------------------------
RO3G_06953       --------------------------------------
RO3G_13784       --------------------------------------
RO3G_08863       --------------------------------------
UM05862          HTGETNSVVYAT------------------------
SjCSL1           --------------------------------------
SPCC736.08       --------------------------------------
CNA01890         DPQTLSAIENVSGSSTETNHQQVSLSIEILSPEGVHSADTFDHTDHPV

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       EFLSNDGGGRVKK-KRSTSSAGG----LSKAVGGKGRRPSSAGSVS---
UM06280          PPPSSDGKVVRPRVSSSVVQQKERAAAAAASKSRRGQSLMVQIN
RO3G_07636       --KMAVR-KRRLSYQOHSTTVK--------SSRRRVNSLNDDELSRH
RO3G_06481       --DAQIRSKRLSYQSSLTSK-------PSSRVRNSLNDDELSDD
RO3G_11583       --IAQDIPIVRLSTGEIKRGRGS-------LGKGSASSLDP
RO3G_14587       --GGK1IRKRVSTDQPETLMCSMRLSDQPRRVRVNSLNDPAPYLRK
CNBD3370         SQRSSWHHSMSVETALPSTDDGPIRRQRTSSMKGGLSRLPLHRKR
SjCSL2

SPCC1223.13

Pc6518

CC1G_01706

RO3G_06953

RO3G_13784

RO3G_08863

UM05862

SjCSL1

SPCC736.08

CNA01890

IDPSFNMAPLVDSEQQTKNPKGQSSFOQRNSNKTRRLALAAAVLAEDDDN

Mm RBP-Jk

Hs RBP-Jk

Xl SuH

Dr RBPSUH

Hr RBP-Jk

Ci 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

SjCSL2

SPCC1223.13

Pc6518

CC1G_01706

RO3G_06953

RO3G_13784

RO3G_08863

UM05862

SjCSL1

SPCC736.08

CNA01890

AVQTLLSWWKERSEEVSSGADGLDTPQDEVRSMIWIDSHWSMSTGGA
Mm_RBP-L     -DVEAETMYR----------------------------------------
Dr_RBP-L     -NMEAETMFR----------------------------------------
PcCSL2       SDPSPYVEVR----------------------------------------
CC1G_03194   SEPSPYVEVR----------------------------------------
UM06280      DWKSTHVDVR----------------------------------------
RO3G_07636   DVPSIQTEFK----------------------------------------
RO3G_06481   DQSPTTRFE----------------------------------------
RO3G_11583   DVKSTETEFV----------------------------------------
RO3G_14587   DVKANHTEYR----------------------------------------
CNBD3370     KHFVKNVDGR----------------------------------------
SjCSL2       ENPCQTYSVDVEI DEPTLLRGLI STS-----------------------
SPCC1223.13  YTRCEVKSI NAFLPDTPTLPPDI SDSR--------------------
Pc6518       NLGLPRLHRI NAPYPGGPLTSV/SPYVQAI GPDGVQREDPI GPGVSPLEKSP
CC1G_01706   SLGPLHTRVYQTSPPGLTSI SFVPVT---VEVPPPAGPNANR--
RO3G_06953   RQGPLKKK------------------------------------------
RO3G_13784   THGPLNTR------------------------------------------
RO3G_08863   THGPLPTR------------------------------------------
UM05862      SLGPLKVSHTN----------------------------------------
SjCSL1       PYGPLAYT------------------------------------------
SPCC736.08   VHGPLSYS------------------------------------------
CNA01890     --------------------------------------------------

Mm_RBP-Jk    ---------------CGESMLCVVPDI SAFREG-------------------
Hs_RBP-Jk    ---------------CGESMLCVVPDI SAFREG-------------------
Xl_SuH       ---------------CAESMLCVVPDI SAFREG-------------------
Dr_RBPSUH    ---------------CGESVLCVVPDI SAFREG-------------------
Hr_RBP-Jk    ---------------CEEGLLCVVPDI SAFREG-------------------
Gt_SuH       ---------------CAEGLLCVVPDI SAFREG-------------------
Dm_SuH       ---------------CTETLLCVVPEI SQFRGE-------------------
Ce_LAG-1     ---------------SEESLHCSI PPVSVRNEQTHW----------------
Hs_RBP-L     ---------------SPRSLCVVPDVAAFCD---------------------
Dr_RBP-L     ---------------SSRSLCCVPDVSLSGE---------------------
PcCSL2       ---------------CTEVLCCLPPES-NNMKRR-------------------
CC1G_03194   ---------------CTEVLCCLPPESGHNNKRR-------------------
UM06280      ---------------TRQTIVCAAAPPSFDQGLPRG------------------
RO3G_07636   ---------------SSQLLSCTVPERHELDS---------------------
RO3G_06481   ---------------SDKLSCI VPDYKELONQ---------------------
RO3G_11583   ---------------SQDSVHKCI PFDVAN----------------------
RO3G_14587   ---------------SRELICKVPPRHELMEVKKVYG-----------------
CNBD3370     ---------------AKHVMYGDEPARNHEVRCGEVMVASE-----------------
SjCSL2       ---------------HVPPRLHAYLADLADICQPP-------------------
SPCC1223.13  ---------------FSSRVYACTANLIELCEIP-------------------
Pc6518       ---------------YLAAOMPIHSTIVDLPMPADI LRSLOQEDLEPPRDGAAPQN---GT
CC1G_01706   ---------------YVSTGPLHTIVVELPPLABVI KAMEDEPVTASDVSGSKPHQSEGQEG
RO3G_06953   ---------------SLIELEPLDI QDI LNHHLNLSTQOQDS-------------------
RO3G_13784   ---------------LMQPN-------HWSIALP---------------------
RO3G_08863   ---------------IASSPETHTPHEI HWSIDLS---------------------
UM05862      ---------------APTSSTSTSTSTDPEIYTVTLPAI REMIKVALTA
SjCSL1       ---------------ROETEKADEN-LLIVSLPYEVNT------------------
SPCC736.08   ---------------FTAADTSTMGTIVTLGLSQI SYDP-------------------
CNA01890     -----------------------------------------------------------------

Mm_RBP-Jk    ---------------WRWVRQPVQVPTVLVRNDGI IYSTSLTFTYTEPGRPRPCHSAAG
Hs_RBP-Jk    ---------------WRWVRQPVQVPTVLVRNDGI IYSTSLTFTYTEPGRPRPCHSAAG
Xl_SuH       ---------------WRWVRQPVQVPTVLVRNDGI IYSTSLTFTYTEPGRPRPCHSAAG
Dr_RBPSUH    ---------------WRWVRQPVQVPTVLVRNDGI IYSTSLTFTYTEPGRPRPCHSAAG

Hr_RBP-Jk        WTWWKQSVPQPI NLVRHDGII YPTNLTFPTPEPGPRPTLRLL
Ci_SuH        WKKWKEVSPQPI NLVRNDGVII YPTNLTFPTPEPGPRQHCAAL
Dm_SuH        WLWRQPTQVPQI SLVRNGII YATGLTFYTPEPGPRPHCTQQA
Ce_LAG-1     MTFTNGDVEPI SLVRDGDGVYSSGTLFSYKSLERHGPCHIVSN
Hs_RBP-L       WRLVRPTPIT PI QMSLVRADGLFYPSAFSTTYPEYSVPVPHPGVP
Dr_RBP-L       WRTWQVQVVPQ PLRLDGLGILYRWSFTYPHEASOQSTSGSG
PcCSL2             PI I LRHDGVFSNVPYYP
CC1G_03194     PI I LRSDGFWFNSVMYP
UM06280        RVPILVRHDGVI FPTDFHYQC
RO3G_07636     -FAITQPDTSRHK IPLLLVQEDQI YNSSLFYSF
RO3G_06481     -PVQVNSSTSRQPI PLLLVRDGVYIYNFLFYSF
RO3G_11583     -STTIQENDHRPI PLLLVRGKGIVYKTNLYYI
RO3G_14587     -DLPILVLRGDTI CTKGFCSL
CNBD3370       -PEPSIKSDRKPI FLVREDKVI IPTSLHYPV
SjCSL2        ALSVTSPIPLVSSQNI I FHSGFTWPI HPHTL
SPCC1223.13    VCMFEPTVPSLPI LFLQYETLFHSGYKWPLESH
Pc6518         STDPKLNVAALNGRSLPLFI ARAVDGGYHVHSGRAI AVENIFQSLIE NGRT
CC1G_01706     VPPPPPQPMAGSLPLFI IRSDDGVGYSRGTIACEPHYSGLAMA
RO3G_06953     -RYLELPILLSPNTFIIYHSSKLSSFTLMHSPESSGQFS
RO3G_13784     -PLEQLVGPLPLLVMVR- QDGLVYHTGKTFSTNSCT
RO3G_08863     -SVKQOTEELPLLVLR- QDGLVYHTGKSSLKVJNEGEGWLC
UM05862        RRSSTPSAPI PPHVLLPLI FVNSDGTAYHFSGRHIVCEDLVQLMKATHGD
SjCSL1         -TPNWPTTLPLFTPRLPGLYVGKCDVWR
SPCC736.08     -SCAEQKYPLLLFI PGGIVI GKEICL LTSSAFGN
CNA01890       

Mm_RBP-Jk       AI LRASSQVPNSN TSEGNYNNSS---
Hs_RBP-Jk       AI LRASSQVPNSNTSEGNSYNSS---
Xl_SuH         AI LRASSLSASNEPNTSEGNSYNSS---
Dr_RBP-SUH     AI LRAANSSTSPPSSSASSLLQDGA---
Hr_RBP-Jk       YI SSTRP---DR---
Ci_SuH         NLHSGKRPASMPPTVSGEDSNPD---
Dm_SuH         EDVMARQNNNNNTISI SNNNSNSGL---
Ce_LAG-1       Y---
Hs_RBP-L       EPATDADLLESIQEFTRTNFHF------
Mm_RBP-L       EPAPDADTLLESIHHEFRTNFHF------
Dr_RBP-L       ERSADSDALITIHQEFTRTNFHF------
PcCSL2         
CC1G_03194     
UM06280        
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       

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

Martin Převorovský, Tomáš Groušl, Jana Staňurová, Jan Ryneš, Wolfgang Nellen, František Půta, Petr Folk*

*Department of Cell Biology, Faculty of Science, Charles University in Prague, Viničná 7, 128 43, Prague 2, Czech Republic

1Department of Genetics, Kassel University, Heinrich Plett Strasse 40, 34132 Kassel, Germany

1present address: Institute of Microbiology, Academy of Sciences of the Czech Republic, Vídeňská 1083, 142 20, Prague 4, Czech Republic

2present 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
Abstract

The CSL (CBF1/RBP-Jκ/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⁺ and cbf12⁺, 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⁺ and cbf12⁺ 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⁻/₀GAA in vitro. The deletion of cbf11⁺ 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-Jκ/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 (GTG³⁻/cGAA) 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-Jκ 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+ and cbf12+, 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%, H₂O₂ 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 place, “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$_2$ 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 $cbf11^+$ and $cbf12^+$ 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 $cbf11^+$ 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 $cbf11^+$ 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 BamHI site of the pACT2-AD, pAS2-1 and pLexA vectors) and the *S. pombe* expression vectors (primers mp20, mp25, mp26, if04; inserted using NdeI/BamHI into the pREP42-MHN and pREP42-EGFPN vectors) [20]. The $cbf12^+$ ORF was cloned using the primers if05 and if06, and inserted in between the NdeI/BamHI sites of pREP42-MHN, pREP42-EGFPN [20] and pAS2-1 vectors, and in between the EcoRI/BamHI sites of pLexA. The Δ$cbf11$ strain (CBF11-KO), obtained from Dr. Anabelle Decottignies [24], was crossed with the PN558 parent strain and sporulated to generate Δ$cbf11$ strains of the opposite mating type (MP05, MP06). The Δ$cbf12$ strains (MP01, MP03) were generated from the parent strains (PN558, PN559) by homologous recombination using the appropriate pCloneNat1 targeting vector obtained from Dr. Juraj Gregáň [25]. The Δ$cbf11$ Δ$cbf12$ 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 [γ-32P]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 MgCl2), 2 ng of radioactively labeled probe, up to 40 ng of unlabeled competitor, 1 μg (for purified proteins) or 20 μg (for cell lysates) of carrier sonicated salmon sperm DNA and either the *S. pombe* cell extract (up to 100 μg) 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× 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 (Δcbf11 strains) were removed by gating.

Results

\( cbf11^+ \) and \( cbf12^+ \), 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 \( cbf11^+ \) (SPCC736.08) and \( cbf12^+ \) (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 \( cbf11 \) mRNA being roughly 50\( \times \) less abundant, and the \( cbf12 \) mRNA being roughly 170\( \times \) less abundant than the \( act1 \) [GenBank:2540051] normalization control during the early logarithmic phase. There is an approximately two-fold downregulation of the \( act1 \) normalization control during sporulation [30]. For this reason, the sporulation samples are not directly comparable with the others, although comparisons between \( cbf11 \) and \( cbf12 \) mRNA levels under these conditions can still be made.

The expression of \( cbf11 \) 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 \( cbf12 \) mRNA has a more variable profile with a marked increase (up to the levels of \( cbf11 \)) 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 \( cbf12 \) mRNA levels was found in the sporulating cells. The \( cbf12 \) 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 \( \text{cbf12}^+ \), however, we repeatedly failed to get either an EGFP (kan\(^R\)) or EYFP (nat\(^R\)) KI strain for \( \text{cbf11}^+ \), 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 \( \text{cbf12}^+ \) 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 \( S. \text{ cerevisiae} \) 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 \textit{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 → 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 (∆cbf11 and ∆cbf11 ∆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 \textit{S. pombe} proteins/lysates for the RBP and KSHV probes, the ones that were bound most strongly by Cbf11 (data not shown). When \textit{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 \textit{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 \( \Delta \text{cbf}11 \) and \( \Delta \text{cbf}12 \) single and double deletion strains**

We have obtained the \( \Delta \text{cbf}11 \) 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 \( \Delta \text{cbf}12 \) 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 \( \Delta \text{cbf}11 \) and \( \Delta \text{cbf}12 \) 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 \( \Delta \text{cbf}11 \Delta \text{cbf}12 \) double KOs by crossing. These strains were also viable, thus, as it was found earlier, neither the \( \text{cbf}11^+ \) nor the \( \text{cbf}12^+ \) gene is essential under normal growth conditions, and \textit{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+ 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 Δcbf12 strain was indistinguishable from the WT control, the growth of the Δcbf11 cells was retarded as compared with WT. Intriguingly, there was further growth impairment when the cbf11+ and cbf12+ 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 Δcbf11-associated growth phenotype. WT and double KO strains were transformed with either an empty vector or plasmids encoding cbf11+ or cbf12+, respectively, and monocolonies were grown for 7 days on MB media containing phloxin B. The overexpression of cbf11+ 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+ 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+ in WT cells resulted in decreased cell viability as well (data not shown).

**Altered colony morphology of strains lacking cbf11+**

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+ when grown on the rich YES medium. Namely, when illuminated, the
surface of the spotted WT and Δchf12 giant colonies appeared dim, which was in sharp contrast to the Δchf11 and Δchf11 Δchf12 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 Δchf11-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 Δchf11 Δchf12 giant colonies could be rescued by mild overexpression of chf11+ but not chf12+ 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 chf11+ 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 Δchf11 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 Δchf12 cells formed regularly shaped, evenly stained round colonies, there were significant
morphology alterations in the case of the Δcbf11 single and Δcbf11 Δcbf12 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 Δcbf11 Δcbf12 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+ and cbf12+ 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 Δcbf11 spots remained attached to the agar surface, the Δcbf12 and Δcbf11 Δcbf12 spots were washed off completely. This observation was also documented microscopically (Fig. 5A). We performed a rescue experiment of this phenotype in the Δcbf11 Δcbf12 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 $cbf12^+$ 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 $cbf11^+$. 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 reintroduction of $cbf11^+$ 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 $cbf11^+$ abolished the flocculation (and the adhesion to the cultivation flasks) of the $\Delta cbf11$ strain. When $cbf12^+$ 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+ or overexpression of cbf12+ result in high-frequency spontaneous diploidization**

Next, we turned to the second prominent phenotype displayed by colonies lacking cbf11+, 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 G2 phase, with haploids showing a single 2C peak and diploids showing a single 4C peak on FACS histograms [29]. The profile of the haploid Δcbf12 culture was indistinguishable from WT. In contrast to that, the Δcbf11 and Δcbf11 Δcbf12 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 Δcbf11 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** or overexpression of cbf12 lead to multiple defects in cell and nuclear division

Importantly, all strains used in our analyses were heterothallic, either $h^+$ or $h^-$, 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 ∆cbf12 strain, the microscopy revealed a surprising range of defects exerted by either the cbf11$^+$-lacking or cbf12$^+$-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 ∆cbf11 and cbf12$^+$-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 ∆cbf11 and cbf12$^+$-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$^+$ was overexpressed in a ∆cbf11 strain (Fig. 8H). Finally, the deletion of cbf11$^+$ 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+ or excess
dose of cbf12+ 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+ (class F1) and cbf12+ (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 \textit{cbf12}+ 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 \textit{cbf12}+ 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 \textit{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 \textit{Δcbf11}-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 \textit{Δcbf11} 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 (\textit{Δcbf11} KO or \textit{cbf12}+ 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 \( \text{sep}^{10^+} \) gene encodes the Med31 subunit of the RNA polymerase II Mediator complex [49,50] and, interestingly, there are indications that the transcription of the \( \text{cbf}^{12^+} \) gene is Med31-dependent [51,52].

The most intriguing effect of the dysregulated CSL expression (again, \( \Delta\text{cbf}^{11} \) or \( \text{cbf}^{12^+} \) overexpression) was the emergence of stable diploid-like subpopulations. These manifested themselves as dark-staining clonal sectors in monocolonies 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 monocolonies as well, although their occurrence was about 1 sector in 100 colonies (data not shown). In contrast to that, for the \( \Delta\text{cbf}^{11} \) 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 $cbf11^+$, 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 $cbf12^+$ 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 Δ*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+ results in cold sensitivity and slow growth.** (A) When cultured on YES plates at 19°C for 5 days, the Δcbf11 and Δcbf11 Δcbf12 strains displayed marked growth impairment as compared with the Δcbf12 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 Δcbf12 strain and WT. In contrast, the deletion of cbf11+ was found to cause growth retardation, which phenotype is, intriguingly, further potentiated by the simultaneous deletion of cbf12+. (C) When overexpressed from a plasmid, cbf11+ but not cbf12+ rescues the growth defect in the Δcbf11 Δcbf12 strain. The monocolonies were cultured on inducing MB plates with phloxin B for 7 days. Notably, the overexpression of cbf12+ seems to be toxic for the cells as judged by the intense phloxin staining of the colony.

**Figure 4 – The deletion of cbf11+ results in altered colony morphology.** (A) Cells were spotted on a YES plate and incubated at 30°C for 7 days. The Δcbf11 and Δcbf11 Δcbf12 strains display a “shiny” phenotype, i.e., they appear glossy when illuminated, unlike the dim WT and Δcbf12 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\textsuperscript{+} but not cbf12\textsuperscript{+} rescues the shiny phenotype in the Δcbf11 Δcbf12 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 Δcbf11 and Δcbf11 Δcbf12 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 Δcbf12 monocolonies appear normal. (F) Overexpression of cbf11\textsuperscript{+} rescues the sectoring phenotype of the Δcbf11 Δcbf12 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\textsuperscript{+} and cbf12\textsuperscript{+} 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 Δcbf11 strains. By contrast, strains lacking cbf12\textsuperscript{+} were washed off completely. A microscopic picture of the washed surface is shown in the bottom panel; the bar represents 25 μm. (B) The defective adhesion of the Δcbf11 Δcbf12 strain can be rescued by overexpression of cbf12\textsuperscript{+} from a plasmid. Cells were grown on an inducing MB plate for 1 week and processed as in A. Note that overexpression of cbf11\textsuperscript{+} further decreased adhesion of the double KO strain. (C) The Δcbf11 single KO cells display higher adhesion than WT, which can be lowered again by the overexpression of cbf11\textsuperscript{+} 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 \( cbf11^+ \) or overexpression of \( cbf12^+ \) trigger flocculation.

Flocculation assays were performed as described under Materials and Methods; cultures transferred to Petri dishes were photographed from above. Areas of \( \sim 1.7 \text{ cm}^2 \) 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 \( cbf11^+ \) on a plasmid. (B) Cells transformed with either an empty vector or a plasmid encoding \( cbf12^+ \) 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 \( cbf11^+ \) or overexpression of \( cbf12^+ \) result in frequent diploidization.

Exponentially growing WT cells overexpressing \( cbf12^+ \) 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 G2-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 \( cbf12^+ \) 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 \( cbf11^+ \) or overexpression of \( cbf12^+ \) 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 \(\mu m\).

**Figure 9 – A proposed model of CSL functioning in \textit{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


    Wach, P. Philippsen, J.R. Pringle, Heterologous modules for efficient and versatile

    vectors for PCR-based gene tagging in *Saccharomyces cerevisiae* and
    *Schizosaccharomyces pombe* using nourseothricin resistance, Yeast 22 (2005) 1061-
    1068.

    Vectors for the expression of tagged proteins in *Schizosaccharomyces pombe*, Gene

    galactosylation is essential for nonsexual flocculation in *Schizosaccharomyces pombe*,

    Exp. (2006) 64.

    specific mutagenesis of DNA fragments: study of protein and DNA interactions,


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

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>PN558</td>
<td>$h^+ \text{ leu1-32 ura4-D18 ade6-M210}$</td>
<td>A. Decottignies [24]</td>
</tr>
<tr>
<td>PN559</td>
<td>$h^+ \text{ leu1-32 ura4-D18 ade6-M216}$</td>
<td>A. Decottignies [24]</td>
</tr>
<tr>
<td>MP01</td>
<td>$h^+ \text{ leu1-32 ura4-D18 ade6-M210 cbf12::pCloneNat1}$</td>
<td>This study</td>
</tr>
<tr>
<td>MP03</td>
<td>$h^+ \text{ leu1-32 ura4-D18 ade6-M216 cbf12::pCloneNat1}$</td>
<td>This study</td>
</tr>
<tr>
<td>MP05</td>
<td>$h^+ \text{ leu1-32 ura4-D18 ade6-M210 cbf11::kan''}$</td>
<td>This study</td>
</tr>
<tr>
<td>MP06</td>
<td>$h^+ \text{ leu1-32 ura4-D18 ade6-M210 cbf11::kan''}$</td>
<td>This study</td>
</tr>
<tr>
<td>CBF11-KO</td>
<td>$h^+ \text{ leu1-32 ura4-D18 ade6-M216 cbf11::kan''}$</td>
<td>A. Decottignies [24]</td>
</tr>
<tr>
<td>MP07</td>
<td>$h^+ \text{ leu1-32 ura4-D18 ade6-M210 cbf11::kan''}$</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>$\Delta cbf12::pCloneNat1$</td>
<td></td>
</tr>
<tr>
<td>MP09</td>
<td>$h^+ \text{ leu1-32 ura4-D18 ade6-M210 cbf11::kan''}$</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>$\Delta cbf12::pCloneNat1$</td>
<td></td>
</tr>
<tr>
<td>MP10</td>
<td>$h^+ \text{ leu1-32 ura4-D18 ade6-M216 cbf11::kan''}$</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>$\Delta cbf12::pCloneNat1$</td>
<td></td>
</tr>
<tr>
<td>MP12</td>
<td>$h^+ \text{ leu1-32 ura4-D18 ade6-M216 cbf12''::EGFP-kan''}$</td>
<td>This study</td>
</tr>
<tr>
<td>FY254</td>
<td>$h^+ \text{ leu1-32 ura4-D18 ade6-M210 can1-1}$</td>
<td>S. Forsburg</td>
</tr>
</tbody>
</table>
Table 2 – Oligonucleotide primers used in this study

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Purpose&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>mp20</td>
<td>GTAGGATCCTATGGGGGACTATTTTG</td>
<td><em>cbf11&lt;sup&gt;+&lt;/sup&gt;</em> cDNA construction (exon 1, fwd, <em>S. pombe</em> plasmids)</td>
</tr>
<tr>
<td>if03</td>
<td>GTAGGATCCATATGGGGGACTATTTTG</td>
<td><em>cbf11&lt;sup&gt;+&lt;/sup&gt;</em> cDNA construction (exon 1, fwd, two-hybrid)</td>
</tr>
<tr>
<td>mp25</td>
<td>CCATTATTTATGGTTGTTCCAAGTCACTGA TTTAAATTTAAGG</td>
<td><em>cbf11&lt;sup&gt;+&lt;/sup&gt;</em> cDNA construction (exon 1, rev)</td>
</tr>
<tr>
<td>mp26</td>
<td>CCTAAATTTAAATGTTTCAAGTCACTGAACATAAATAATGG</td>
<td><em>cbf11&lt;sup&gt;+&lt;/sup&gt;</em> cDNA construction (exon 2, fwd)</td>
</tr>
<tr>
<td>if04</td>
<td>GAAGGATCCTCAGTTTCCAAAAGCAC</td>
<td><em>cbf11&lt;sup&gt;+&lt;/sup&gt;</em> cDNA construction (exon 2, rev)</td>
</tr>
<tr>
<td>if05</td>
<td>GTAGAATTCCATATGTCCCCAAACGTTC</td>
<td><em>cbf12&lt;sup&gt;+&lt;/sup&gt;</em> ORF (fwd)</td>
</tr>
<tr>
<td>if06</td>
<td>GCCAGATCTTTAGTCACTTTCAAAAGGG</td>
<td><em>cbf12&lt;sup&gt;+&lt;/sup&gt;</em> ORF (rev)</td>
</tr>
<tr>
<td>mp27</td>
<td>TATGCTGGACTATAGTGGGC</td>
<td><em>cbf11&lt;sup&gt;+&lt;/sup&gt;</em> KO verification (ORF, fwd)</td>
</tr>
<tr>
<td>mp28</td>
<td>GATACAGCAACTCCTCCCG</td>
<td><em>cbf11&lt;sup&gt;+&lt;/sup&gt;</em> KO verification (outer genomic, rev)</td>
</tr>
<tr>
<td>kan-rev</td>
<td>AATGCTGGTCGCTATACTGC</td>
<td><em>cbf11&lt;sup&gt;+&lt;/sup&gt;</em> KO, <em>cbf12&lt;sup&gt;+&lt;/sup&gt;</em> KI verification (kanMX6, fwd)</td>
</tr>
<tr>
<td>mp31</td>
<td>TGTGCAGATTTGGATGGGC</td>
<td><em>cbf12&lt;sup&gt;+&lt;/sup&gt;</em> KO, KI verification (ORF, fwd)</td>
</tr>
<tr>
<td>mp32</td>
<td>AAATCAATCCCTCCACG</td>
<td><em>cbf12&lt;sup&gt;+&lt;/sup&gt;</em> KO, KI verification</td>
</tr>
<tr>
<td></td>
<td>Sequence</td>
<td>Method</td>
</tr>
<tr>
<td>-----</td>
<td>--------------------------------------------------------------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>mp33</td>
<td>GCGCACGTCAAGACTGTC</td>
<td>(outer genomic, rev)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>chf12⁺ KO verification</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(natMX6, fwd)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mp43</td>
<td>CAGTGGAAATTATCTCCCATTTTATTATTTC</td>
<td>chf12⁺::EGFP knock in (fwd)</td>
</tr>
<tr>
<td></td>
<td>AATACGAGACACTCTTTCATCTGGATATA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AGTGGCCTTTGGAAAGTCACCGGATCCCA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GGGTTAATTAA</td>
<td></td>
</tr>
<tr>
<td>mp44</td>
<td>AAAACAAAAAAGAGTAATAATAAATATACT</td>
<td>chf12⁺::EGFP knock in (rev)</td>
</tr>
<tr>
<td></td>
<td>AATCCCTTTGCAAAAAACTTTTCAATAAAAAA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AAAGTAGTAAGACAAATAATGAATTGA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GCTCGTTTAAAC</td>
<td></td>
</tr>
<tr>
<td>mp35</td>
<td>ATTTGGCTAGGTGTTCATGG</td>
<td>chf11⁺, qRT-PCR (fwd)</td>
</tr>
<tr>
<td>mp45</td>
<td>TGACGTAGCAAAATCTCGC</td>
<td>chf11⁺, qRT-PCR (rev)</td>
</tr>
<tr>
<td>mp36</td>
<td>CATTCAAGCCTGATACGACG</td>
<td>chf12⁺, qRT-PCR (fwd)</td>
</tr>
<tr>
<td>mp46</td>
<td>AGTGACTTTTCAAGGCCAC</td>
<td>chf12⁺, qRT-PCR (rev)</td>
</tr>
<tr>
<td>mp37</td>
<td>GTAAACGATACCAGGTCCGC</td>
<td>act1⁺, qRT-PCR (fwd)</td>
</tr>
<tr>
<td>mp38</td>
<td>GGTAACCACATATGTATCCCCG</td>
<td>act1⁺, qRT-PCR (rev)</td>
</tr>
</tbody>
</table>

*The sequences are given in the 5’-3’ orientation.

fwd – forward primer, rev – reverse primer.
Table 3 – Frequencies of the sep and cut phenotypes in CSL mutant strains

<table>
<thead>
<tr>
<th>Strain(^a)</th>
<th>sep</th>
<th>cut</th>
<th>Medium</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>0.5%</td>
<td>0.0%</td>
<td>YES/MB</td>
</tr>
<tr>
<td>(\Delta cbf12)</td>
<td>0.7%</td>
<td>0.0%</td>
<td>YES</td>
</tr>
<tr>
<td>(\Delta cbf11)</td>
<td>4.2%</td>
<td>16.1%</td>
<td>YES</td>
</tr>
<tr>
<td>(\Delta cbf11\ \Delta cbf12)</td>
<td>4.3%</td>
<td>19.7%</td>
<td>YES</td>
</tr>
<tr>
<td>WT + (cbf12)(^+)\ OE</td>
<td>8.7%</td>
<td>1.2%</td>
<td>MB</td>
</tr>
<tr>
<td>(\Delta cbf11 + cbf12)(^+) OE</td>
<td>9.5%</td>
<td>4.1%</td>
<td>MB</td>
</tr>
</tbody>
</table>

\(^a\)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 4

Click here to download high resolution image
Figure 7

Click here to download high resolution image
Figure 9

- Logarithmic phase
- Stationary phase
- Meiosis
- Cbf11
- Cbf12

Extracellular material production
Adhesion
Coordination
Colony morphogenesis
Cell/nuclear division
Ploidy maintenance
(Meiosis ?)
High environmental iron concentrations stimulate adhesion and invasive growth of *Schizosaccharomyces pombe*

<table>
<thead>
<tr>
<th>Journal:</th>
<th><em>FEMS Microbiology Letters</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>FEMSLE-08-10-0946</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Research Letter</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>02-Oct-2008</td>
</tr>
</tbody>
</table>
| Complete List of Authors: | Prevorovsky, Martin; Faculty of Science, Charles University in Prague, Dept. of Cell Biology  
                      | Stanurova, Jana               |
|                   | Puta, 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, Schizosaccharomyces |
RESEARCH LETTER

High environmental iron concentrations stimulate adhesion and invasive growth of

Schizosaccharomyces pombe

RUNNING TITLE: Iron stimulates adhesion and invasion of fission yeast

Martin Převorovský, Jana Staňurová, František Půta, Petr Folk

Department of Cell Biology, Faculty of Science, Charles University in Prague, Czech Republic

CORRESPONDING AUTHOR:
Petr Folk
Department of Cell Biology, Faculty of Science, Charles University in Prague
Viničná 7, Prague 2, CZ128 43, Czech Republic
Tel.: +420 22195 1765
FAX: +420 22195 1758
E-mail: folk@natur.cuni.cz

KEYWORDS:
adhesion, fission yeast, iron, invasive growth, Schizosaccharomyces
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 $h^- leu1-32 ura4-D18 ade6-M216$; a kind gift of Dr. Anabelle Decottignies (Decottignies et al., 2003)) to
adhesivity/invasivity assays (based on (Guldal and Broach, 2006)), using the standard complex (YES) medium supplemented with various concentrations of FeCl₂. Exponentially growing PN559 cells were harvested from liquid culture by centrifugation, washed once in sterile deionised water, and the cell suspension (1×10⁵ cells in 5 µl) was plated. The YES complex medium (0.5% yeast extract, 3% glucose) with SP supplements (a standard mixture of adenine, leucine, uracil, histidine and lysine; Formedium) was solidified with 2% agar, and either 0, 1, 2, 3 or 4 mM FeCl₂ was added prior to autoclaving. The plates were then incubated at 30°C for up to 19 days in a wet chamber, and sample plates were subjected to adhesivity/invasivity assays on days 6, 14 and 19.

The adhesivity assay consisted of washing the plates evenly with a stream of water for 1 min. The spot cell mass remaining attached to the plate was then documented by photography, and the plate was subjected to the invasivity assay, in which the residual cells were washed off the surface completely by rubbing the plate carefully under running water. Under these conditions, only cells residing deep within the agar resist the washing. The spot location was photographed again to score for macroscopically visible invasive growth, and finally the invasive structures formed were investigated using an Olympus CK2 light microscope with an Olympus SP-350 digital camera attached to it.

Results

While conducting adhesivity tests with some of our adhesion-deficient fission yeast mutant strains we noticed that the wild-type control cells displayed markedly increased adhesion when cultured on a solid rich medium to which milimolar FeCl₂ was added. Furthermore, the cells seemed to have grown deep into the agar, a behavior that is normally not observed with *S. pombe*. As this could have had implications for a number of important pathogenic microorganisms, we decided to examine these observations further. We also
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$_2$. 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$_2$, 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
pseudo-mycelia we observed. However, these true mycelia are formed by long, septated hyphae, and, significantly, their formation was shown to be inhibited by the addition of nitrogen sources, such as auxotrophic supplements (lysine, adenine, uracil), to the cultivation media. In contrast to that, we observed invasive pseudo-mycelia formation on a rich medium with a standard nitrogen content and with the SP nutritional supplements added, the trigger being solely the abundant iron.

In summary, we have described in this study an unexpected link between the extracellular iron concentration and the fission yeast’s adhesive properties and its ability to invade solid media. We have shown that high iron strongly potentiates adhesion both to the agar surface and within the yeast colony. Furthermore, abundant environmental iron stimulates cells to invade solid substrates. *S. pombe* is a non-pathogenic species, but since adhesion and invasive growth are important virulence traits of numerous pathogenic microorganisms (Mendes-Giannini *et al.*, 2005; Saville *et al.*, 2003), we believe that our results set stage for studying this novel role of iron in the models of infection as well.

**Acknowledgements**

This work was supported by the research projects No. MSM0021620858 and LC07032 of the Czech Ministry of Education, Youth and Sports.
References


Vis. Exp., 64.

virulence factors in the AIDS-associated pathogen Cryptococcus neoformans. PLoS.
Biol. 4, e410.

pathogenic fungi with host cells: Molecular and cellular approaches. FEMS Immunol.


of cell morphology in vivo reveals distinct roles for yeast and filamentous forms of
**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$_2$ 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$_2$). (c) The structures are composed of rather elongated yeast cells. Bars represent 50 µm.