A role for Q/N-rich aggregation-prone regions in P-body localization

Martin A. M. Reijns*, Ross D. Alexander, Michael P. Spiller‡ and Jean D. Beggs§

Wellcome Trust Centre for Cell Biology, University of Edinburgh, King's Buildings, Mayfield Road, Edinburgh EH9 3JR, UK

*Present address: Medical Research Council, Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK

‡Present address: Faculty of Life Sciences, Michael Smith Building, University of Manchester, Oxford Road, Manchester M13 9PT, UK

§Author for correspondence (email: jbeggs@ed.ac.uk)

Accepted 14 May 2008

Journal of Cell Science 121, 2463-2472 Published by The Company of Biologists 2008

doi:10.1242/jcs.024976

Summary

P-bodies are cytoplasmic foci that are sites of mRNA degradation and translational repression. It is not known what causes the accumulation of RNA-degradation factors in P-bodies, although RNA is required. The yeast Lsm1-7p complex (comprising Lsm1p to Lsm7p) is recruited to P-bodies under certain stress conditions. It is required for efficient decapping and degradation of mRNAs, but not for the assembly of P-bodies. Here we show that the Lsm4p subunit and its asparagine-rich C-terminus are prone to aggregation, and that this tendency to aggregate promotes efficient accumulation of Lsm1-7p in P-bodies. The presence of glutamine- and/or asparagine-rich (Q/N-rich) regions in other P-body components suggests a more general role for aggregation-prone residues in P-body localization and assembly. This is supported by reduced P-body accumulation of Ccr4p, Pop2p and Dhh1p after deletion of these domains, and by the observed aggregation of the Q/N-rich region from Ccr4p.

Supplementary material available online at
http://jcs.biologists.org/cgi/content/full/121/15/2463/DC1

Key words: P-body localization, Protein aggregation, Q/N-rich domains, Stress

Introduction

Cytoplasmic mRNA-processing bodies (P-bodies) contain a variety of protein factors, some of which are involved directly in decapping (Dcp1p, Dcp2p), some in translational repression and/or activation of decapping (Pat1p, Dhh1p, Edc1p, Edc2p, Edc3p, Scd6p) and others are involved in deadenylation (Ccr4p, Pop2p, Not1p, Not2p, Not3p, Not4p, Not5p, Pan2p, Pan3p) or 5′ to 3′ degradation (Xrn1p). In addition, factors involved in nonsense-mediated decay and RNA interference (e.g. Ago1) are present in these foci (reviewed by Parker and Sheth, 2007). P-bodies in higher eukaryotes have also been called GW bodies after GW182 (also known as TNRC6A) (Eystathioy et al., 2003), a component that is required for their integrity (Liu et al., 2005), and which has a function in miRNA-mediated silencing (Jakymiw et al., 2005; Liu et al., 2005; Rehwinkel et al., 2005). In budding yeast, a cytoplasmic complex that consists of Lsm1p, Lsm2p, Lsm3p, Lsm4p, Lsm5p, Lsm6p and Lsm7p (hereafter referred to as Lsm1-7p), and which is involved in mRNA decapping and subsequent 5′ to 3′ decay (Bouveret et al., 2000; Tharun et al., 2000), localizes to P-bodies under certain stress conditions (Sheth and Parker, 2003; Teixeira et al., 2005). In higher eukaryotes, the equivalent of Lsm1-7p is present in similar foci, even under normal growth conditions (Ingelfinger et al., 2002; Eystathioy et al., 2003; Cougot et al., 2004). Lsm1-7p is thought to act as a chaperone, remodelling transcript-containing ribonucleoprotein particles (RNPs) at a step following deadenylation, thus promoting decapping (Tharun et al., 2000).

In yeast, no single protein component is responsible for P-body assembly but there is a level of interdependence in the recruitment of some of the components to these foci (Teixeira and Parker, 2007). By contrast, in human cells depletion of many components, with the notable exception of XRN1 and DCP2, affects localization of the others (reviewed by Jakymiw et al., 2007), suggesting that most components involved in early, but not late stages of mRNA decay are essential for P-body assembly. It is not known what makes any of these factors concentrate in cytoplasmic foci, although in yeast this seems to require RNA (Teixeira et al., 2005). More recently, various proteins in budding yeast have been implicated directly in P-body assembly, and the understanding of their physical and functional interactions is gathering pace. This includes Edc3p, Lsm4p (Decker et al., 2007), Pat1p (Pilkington and Parker, 2008), and Ded1p (Beckham et al., 2008).

LSM4 of *Saccharomyces cerevisiae* encodes an essential protein of 187 amino acids (aa). It is one of the seven subunits of the Lsm1-7p and Lsm2-8p (comprising Lsm2p to Lsm8p) complexes, the latter of which is needed for efficient pre-mRNA splicing through its role in U6 small nuclear RNA (snRNA) stability (Achsel et al., 1999; Mayes et al., 1999; Pannone et al., 1998; Salgado-Garrido et al., 1999) and localization (Spiller et al., 2007a) as well as U4/U6 di-snRNP formation (Verdone et al., 2004). The N-terminal 92 aa of Lsm4p include the Sm domain, which is involved in protein-protein and protein-RNA interactions within the Lsm complexes (Cooper et al., 1995; Hermann et al., 1995; Séraphin, 1995). This region is highly conserved between *Saccharomyces* species (Fig. 1A), and between budding yeast and humans (Fig. 1B). The C-teminal 95 aa are rich in asparagine (N; 36%) and serine (S; 17%), giving this region a highly hydrophylic character. It is less conserved than the N-terminus, however, homologues from various *Saccharomyces* species contain similar asparagine-rich stretches that vary in length and position. A notable exception is Lsm4p from *S. kluyveri* that has a glutamine (Q)-rich region (Fig. 1A). This N and/or Q-rich character of the Lsm4p C-terminus is conserved throughout the budding yeasts (supplementary material Fig. S1). By contrast, most Lsm4p homologues from higher organisms have an abundance of arginine and glycine residues in their C-termini, often in the form

Journal of Cell Science

Journal of Cell Science

of RG repeats (Fig. 1B and supplementary material Fig. S2), that are important for interactions with the SMN complex. Symmetrical dimethylation of the arginine residues is thought to be important for regulation of snRNP assembly (Brahms et al., 2001; Paushkin et al., 2002). *S. cerevisiae* does not have a known SMN complex equivalent, providing a possible explanation for the absence of RG repeats in yeast Lsm4p. Experiments with Lsm4p of *Kluyveromyces lactis* suggest that the Lsm4p C-terminus is needed for efficient RNA degradation (Mazzoni et al., 2003a; Mazzoni et al., 2003b).

Here we describe the role of yeast Lsm4p and its C-terminus in Lsm protein aggregation. We show that the ability of Lsm4p to aggregate, although not essential, promotes efficient accumulation of Lsm1-7p in P-bodies. Many other P-body components contain Q/N-rich regions suggestive of a more general role for such aggregation-prone residues in efficient accumulation of these factors in P-bodies. In support of this hypothesis we show that the Q/N-rich region of Ccr4p is prone to aggregation under normal growth conditions and shows increased focal localization under stress conditions. Furthermore, we show that the Q/N-rich region of Ccr4p is essential for its accumulation in microscopically visible P-bodies, whereas those of Pop2p and Dhh1p, although not essential for P-body localization, promote their efficient accumulation in these cytoplasmic foci.

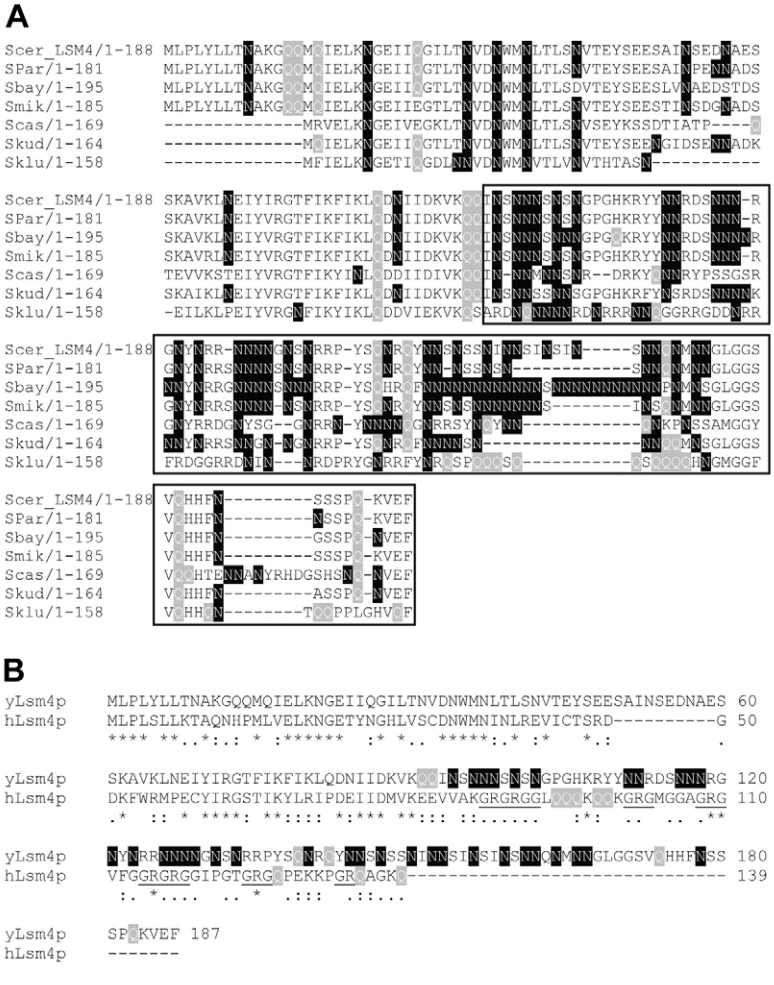


Fig. 1 Lsm4p has an N-rich C-terminal region. (A) The N-terminus of Lsm4p (aa 1-92) contains the Sm domain; the C-terminus (aa 93-187; boxed) contains N-rich (or Q-rich) stretches of variable lengths. Scer, *S. cerevisiae*; Spar, *S. paradoxus*; Sbay, *S. bayanus*; Smik, *S. mikatae*; Scas, *S. castellii*; Skud, *S. kudriavzevii*; Sklu, *S. kluyveri*. (B) Lsm4 protein alignment of the budding yeast (yLsm4p) and the human (hLsm4p) protein. Sequences were aligned using ClustalW. Q residues are highlighted in grey, N residues in black and RG (arginine-glycine) repeats are underlined.

Results and Discussion

Overproduced Lsm4p or its C-terminus accumulate in foci

While investigating the localization of various GFP-Lsm4p fusions, we observed that GFP fused to full-length Lsm4p (GFP-Lsm4) or to its C-terminus (GFP-Lsm4C; aa 92-187) accumulates in cytoplasmic foci as well as in larger aggregates, in a variable percentage of cells (10-60%), whereas GFP fused to the N-terminal half of Lsm4p (GFP-Lsm4ΔC; aa 1-92), containing the Sm domain, does not (Fig. 2A). The number of GFP-Lsm4 foci increases after hypo-osmotic shock, indicating that aggregation can be triggered by stress, and that these newly formed foci are probably P-bodies. This is confirmed by colocalization of Dcp2-RFP with GFP-Lsm4p in foci formed after stress. By contrast, Dcp2-RFP is not particularly enriched in the larger Lsm4p aggregates during log-phase growth, suggesting that these are probably not P-bodies (Fig. 2B). In cells expressing GFP-Lsm4ΔC as the only copy of Lsm4p, Dcp2-RFP localizes to foci after osmotic shock, showing that P-bodies are formed (Fig. 2B). However, GFP-Lsm4ΔC localizes throughout the cell, indicating its failure to accumulate in P-bodies even under stress conditions. The virtual absence of GFP-Lsm4ΔC in microscopically visible P-bodies is not due to reduced levels of this truncated protein, as shown by western analysis (see below).

Other Lsm proteins aggregate upon Lsm4p overproduction

To investigate a potential link between these Lsm4p aggregates and P-bodies, the presence of other proteins was examined. We observed accumulation of all tested GFP-tagged Lsm proteins (Lsm1p, Lsm2p, Lsm6p, Lsm7p and Lsm8p) in similar cytoplasmic aggregates during log-phase growth of cells overproducing non-tagged Lsm4p (P_{GAL}-*LSM4* strain grown on galactose; Fig. 2C and supplementary material Fig. S3). This was even the case with the normally nuclear Lsm8p, indicating that not only Lsm1-7p but also Lsm2-8p aggregates when Lsm4p is present at high levels. Aggregates were observed mostly in the cytoplasm and occasionally in the nucleolus, judging from colocalization with the nucleolar protein Nop1p (data not shown). By contrast, normal localization was observed for each of the Lsm proteins in the absence of excess Lsm4p (Fig. 2C), although each of these GFP-tagged proteins was moderately overexpressed from the *MET25* promoter. Lsm protein aggregates are therefore likely to be not physiologically significant, but simply the result of aggregation of Lsm4p-containing complexes when Lsm4p is present at higher than normal levels. As aggregation was observed with direct interaction partners of Lsm4p in the ring-shaped Lsm complex (Lsm1p, Lsm2p and Lsm8p) as well as with physically more distant subunits (Lsm6p and Lsm7p) it seems possible that overexpressed Lsm4p drives the aggregation of entire Lsm1-7p and Lsm2-8p complexes, probably via its C-terminus. This aggregation is specific to the Lsm proteins, as GFP alone localized throughout the cells regardless of Lsm4p levels (Fig. 2C), and the

Journal of Cell Science

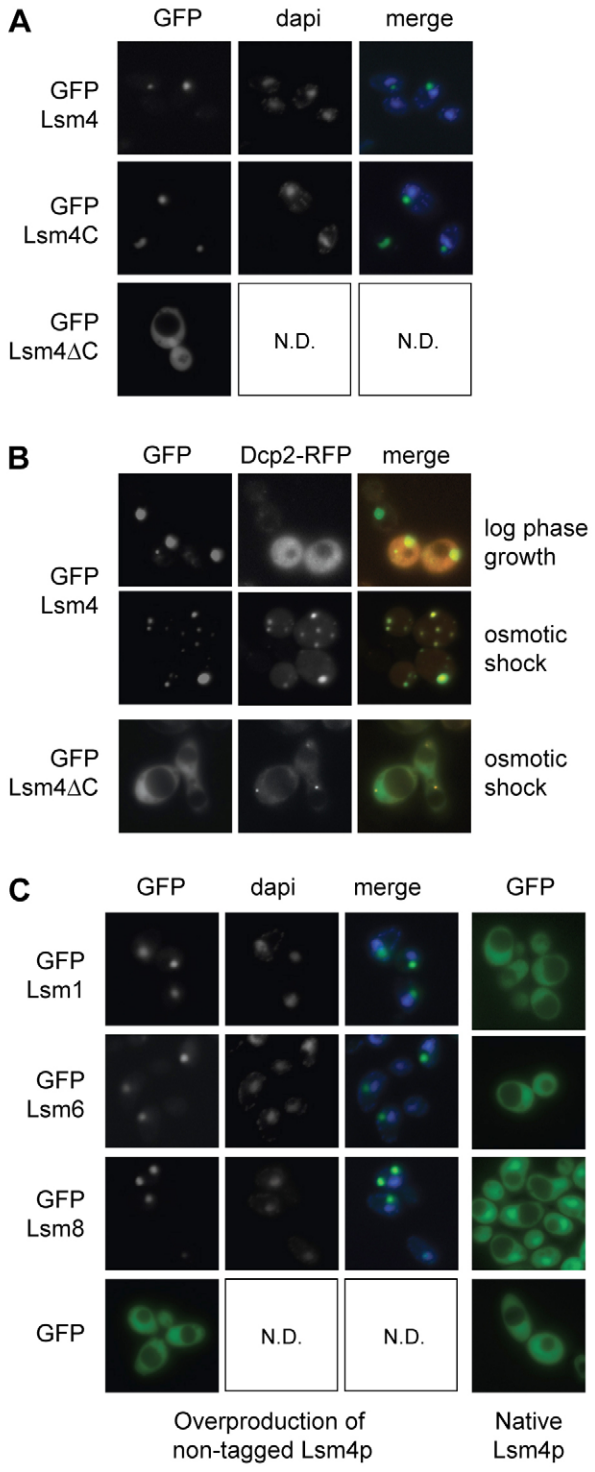


Fig. 2 Overproduction of Lsm4p or its C-terminus leads to aggregation in cytoplasmic foci. (A) GFP-Lsm4 (pMPSLsm4), GFP-Lsm4C (pMPSlsm4D2) and GFP-Lsm4ΔC (pMPSLsm4D1) were overexpressed from the *MET25* promoter in BY4741 cells grown in SD-Ura-Met. Localization was examined in cells during log-phase growth. (B) Colocalization of Lsm4p aggregates with Dcp2-RFP (pRP1155) was examined in BY4741 cells grown in SD-Ura-Leu-Met during log-phase growth or 20 minutes after hypo-osmotic shock. Dcp2-RFP (pMR171) localization was examined in P_{GAL}-*LSM4* cells expressing GFP-Lsm4ΔC grown in SD-Ura-His-Met (to prevent competition between GFP-Lsm4ΔC and endogenous Lsm4p for incorporation into Lsm1-7p), 20 minutes after hypo-osmotic shock. (C) Localization of GFP-Lsm1 (pGFP-N-Lsm1), GFP-Lsm6 (pMPSLsm6), Lsm8-GFP (pMR83) and GFP (pGFP-N-FUS) was examined in log-phase cells overproducing Lsm4p (P_{GAL}-*LSM4* cells grown in SDGal-Ura) and in cells with normal levels of Lsm4p (P_{GAL}-*LSM4* cells with pUSS1 grown in SD-Ura-Met). Nuclear DNA stained with DAPI is shown in blue.

exclusively nuclear Lhp1p fused with GFP remained nuclear under these conditions and did not form aggregates, although many cells showed abnormal nuclear morphology, which is a phenotype associated with Lsm4p overproduction (supplementary material Fig. S3).

A role for Lsm4p in Lsm1-7p P-body localization

The accumulation of GFP-Lsm4 and GFP-Lsm4C in foci when overexpressed, and failure of GFP-Lsm4ΔC to aggregate even under stress conditions, suggests a role for the Lsm4p C-terminus in targeting Lsm1-7p to P-bodies. As a complete Lsm1-7p complex is apparently needed for localization to P-bodies (Ingelfinger et al., 2002; Tharun et al., 2005) the C-terminal deletion is likely to affect localization of the entire Lsm1-7p complex. The localization of GFP-Lsm1 to P-bodies was therefore examined in cells producing either Lsm4p or Lsm4ΔCp (both non-tagged) from the native *LSM4* promoter (i.e. not overproduced). In comparison with the accumulation of GFP-Lsm1 in P-bodies (for colocalization of GFP-Lsm1 with Dcp2-RFP see supplementary material Fig. S4) following hypo-osmotic stress of *LSM4* cells, there was a reduction in the intensity of GFP-Lsm1 foci that formed in *lsm4ΔC* cells (Fig. 3A), as well as an apparent delay in their formation. To quantify this delay, the number of cells that displayed visible P-bodies 5 minutes and 1 hour after hypo-osmotic shock was counted in the *LSM4* and *lsm4ΔC* strains. Whereas 85% of *LSM4* cells displayed foci 5 minutes after hypo-osmotic shock, with only a small increase (to 90%) after 1 hour, only 4% of *lsm4ΔC* cells displayed foci after 5 minutes, increasing to 73% after 1 hour (Fig. 3B), with the majority of these foci still weaker than those observed in *LSM4* cells. The localization of GFP-Lsm2 and GFP-Lsm6 to P-bodies after hypo-osmotic stress was similarly reduced in the *lsm4ΔC* strain compared with the *LSM4* strain (Fig. 3C,D). Taken together, these data strongly suggest that the C-teminal domain of Lsm4p although not actually essential, is nevertheless important for efficient accumulation of Lsm1-7p in P-bodies under stress conditions. As the Lsm4p C-terminal domain seems to be important for efficient recruitment of Lsm1-7p to P-bodies, the C-terminal deletion might also have an effect on the accumulation of other proteins in P-bodies. However, no significant effect was seen on the localization of either Dcp1p or Dcp2p to P-bodies (Fig. 3E, and data not shown).

Detrimental effects of GFP-tagging Lsm4ΔCp

The complete absence of GFP-Lsm4ΔC from P-bodies after osmotic shock seems to contradict the mere reduction in P-body accumulation of GFP-Lsm1, Lsm2 and Lsm6 in the *lsm4ΔC* (non-tagged) strain. However, upon closer inspection, GFP-Lsm4ΔC was observed to localize weakly to P-bodies after hypo-osmotic shock in a small fraction of cells (<1%), and to accumulate in cytoplasmic foci in more than 90% of cells grown into late stationary phase (data not shown). Its reduced accumulation in P-bodies is likely to reflect negative effects of the GFP-tag in combination with the C-terminal deletion, possibly by reducing its incorporation into the Lsm1-7p complex. The negative effect

Journal of Cell Science

of the GFP-tag is emphasized by a slow growth defect of the *GFP-Lsm4ΔC* strain at all temperatures compared with the *lsm4ΔC* strain (with non-tagged protein expressed from its native promoter), which shows slower growth only at 37°C (supplementary material Fig. S5). We cannot formally rule out the possibility that the difference between the non-tagged *lsm4ΔC* and the *GFP-Lsm4ΔC* strains is caused by their different levels of expression (native promoter vs *MET25* promoter), although this is more likely to lead to the opposite of what we observe. While this manuscript was in preparation Mazzoni et al. reported that the asparagine-rich N-terminal region of the *K. lactis* Lsm4 protein, KlLsm4p, which is able to functionally replace its *S. cerevisiae* homologue, is essential for its own localization to P-bodies in budding yeast (Mazzoni et al., 2007). However, these authors only investigated the localization of a GFP-tagged version of this protein, and not the localization of other Lsm proteins in a strain expressing non-tagged klLsm4ΔCp. Thus the effect of the deletion may not have been distinguished from the additional, detrimental effect of the tag.

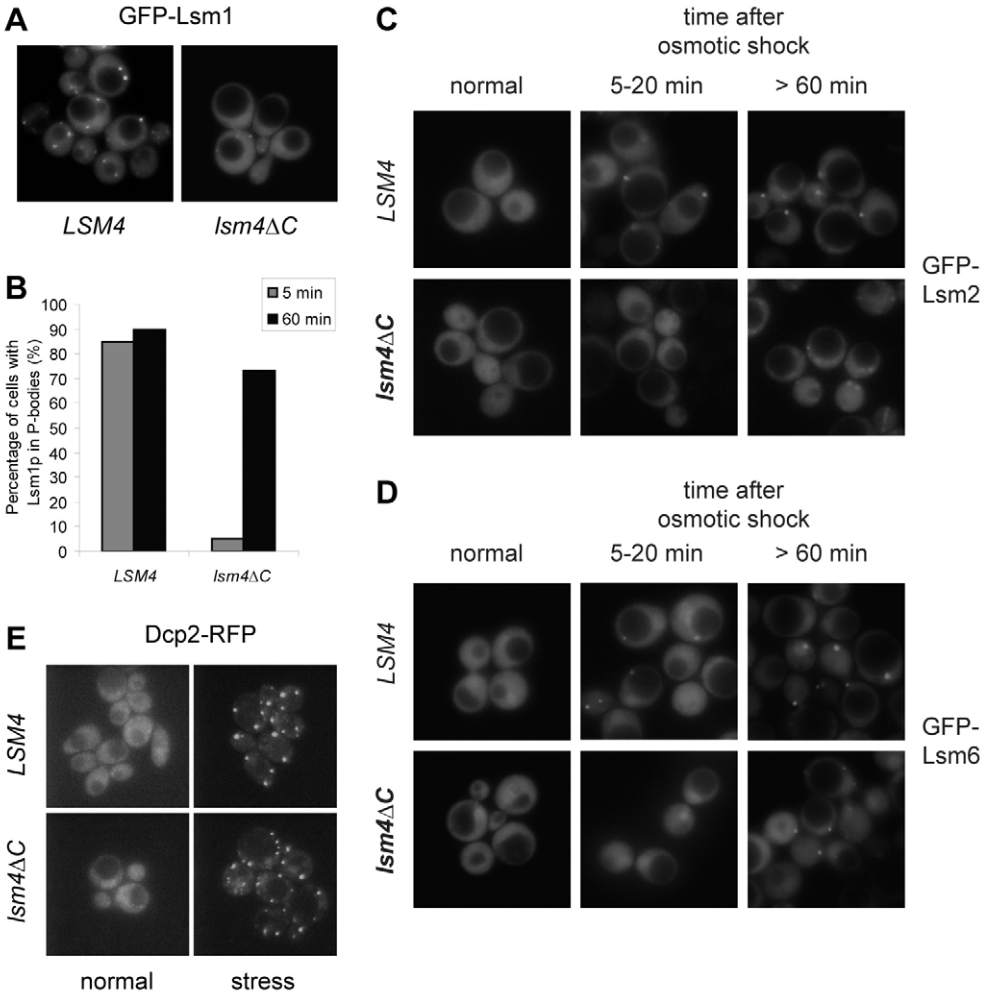


Fig. 3 The Lsm4 C-terminus is required for efficient localization of Lsm1p to P-bodies. (A) GFP-Lsm1 (pGFP-N-Lsm1) localization 20 minutes after hypo-osmotic shock in *LSM4* (MRY71) or *lsm4ΔC* (MRY73) cells. (B) Percentage of cells showing GFP-Lsm1 in foci 5 minutes or 1 hour after osmotic shock (n=100 cells per time point). (C,D) Localization of GFP-Lsm2 (pMPSLsm2) (C) and GFP-Lsm6 (pMPSLsm6) (D) in *LSM4* or *lsm4ΔC* cells before and after hypo-osmotic shock (E) Dcp2-RFP (pMR159) localization in log phase *LSM4* and *lsm4ΔC* cells grown in SD-His and 20 minutes after osmotic shock. All experiments in this figure were performed with strains expressing non-tagged Lsm4p or Lsm4ΔCp from the native *LSM4* promoter.

Absence of the Lsm4p C-terminus affects mRNA decay

To determine whether the Lsm4p C-terminal deletion affects mRNA decay, degradation of a PGK1pGmini reporter transcript (Mitchell and Tollervey, 2003) was investigated. This reporter is expressed from the *GAL1* promoter, allowing its transcription to be switched off by growth on glucose. The rate of subsequent disappearance of the reporter transcript is used as a measure of its 5′ to 3′ degradation through the major mRNA decay pathway. A small effect was observed, as the mRNA half life increased from 3.4±0.9 minutes in wild type to 4.7±1.1 minutes in the *lsm4ΔC* strain, on the basis of the quantitative reverse-transcriptase PCR (qRT-PCR) data presented in Fig. 4C. Half lives calculated using data obtained from the northern blot were slightly higher compared with those determined by qRT-PCR, but the relative difference between the two strains was similar. In addition, the steady-state level of this transcript appears to be about 40% higher in the *lsm4ΔC* cells compared with the *LSM4* cells (Fig. 4B). By contrast, no effect was observed on the splicing of pre-U3 RNA, compared with 12 hours of Lsm8p depletion (Fig. 4D), suggesting that Lsm4ΔCp does not detrimentally affect formation of Lsm2-8p or stability of U6 snRNA. It therefore seems unlikely that the stability or formation of Lsm1-7p is reduced because of this C-terminal deletion, unless the assembly requirements of these two complexes are significantly different. A similar effect on mRNA degradation was reported for klLsm4ΔC in *K. lactis* (Mazzoni et al., 2003a), whereas a seemingly stronger effect was observed for klLsm4ΔC in *S. cerevisiae* (Mazzoni et al., 2003b). The latter may reflect reduced incorporation of the mutant *K. lactis* Lsm4p into the *S. cerevisiae* Lsm1-7p complex. Decker et al. did not find a significant change in the half-lives of PGK1pG or MFA2pG reporter transcripts in the absence of the C-terminal 97 aa of Lsm4p, nor did they report on increased steady-state levels of these transcripts (Decker et al., 2007). The reason for this difference remains unclear; however, the strains used in these studies were constructed in different ways and in different genetic backgrounds. We cannot formally rule out that the effect we see on the PGK1pG half-life is caused by reduced expression and/or stability of Lsm4ΔCp, as we have no antibody to compare its level with that of full-length Lsm4p. However, the absence of an effect on splicing argues against this.

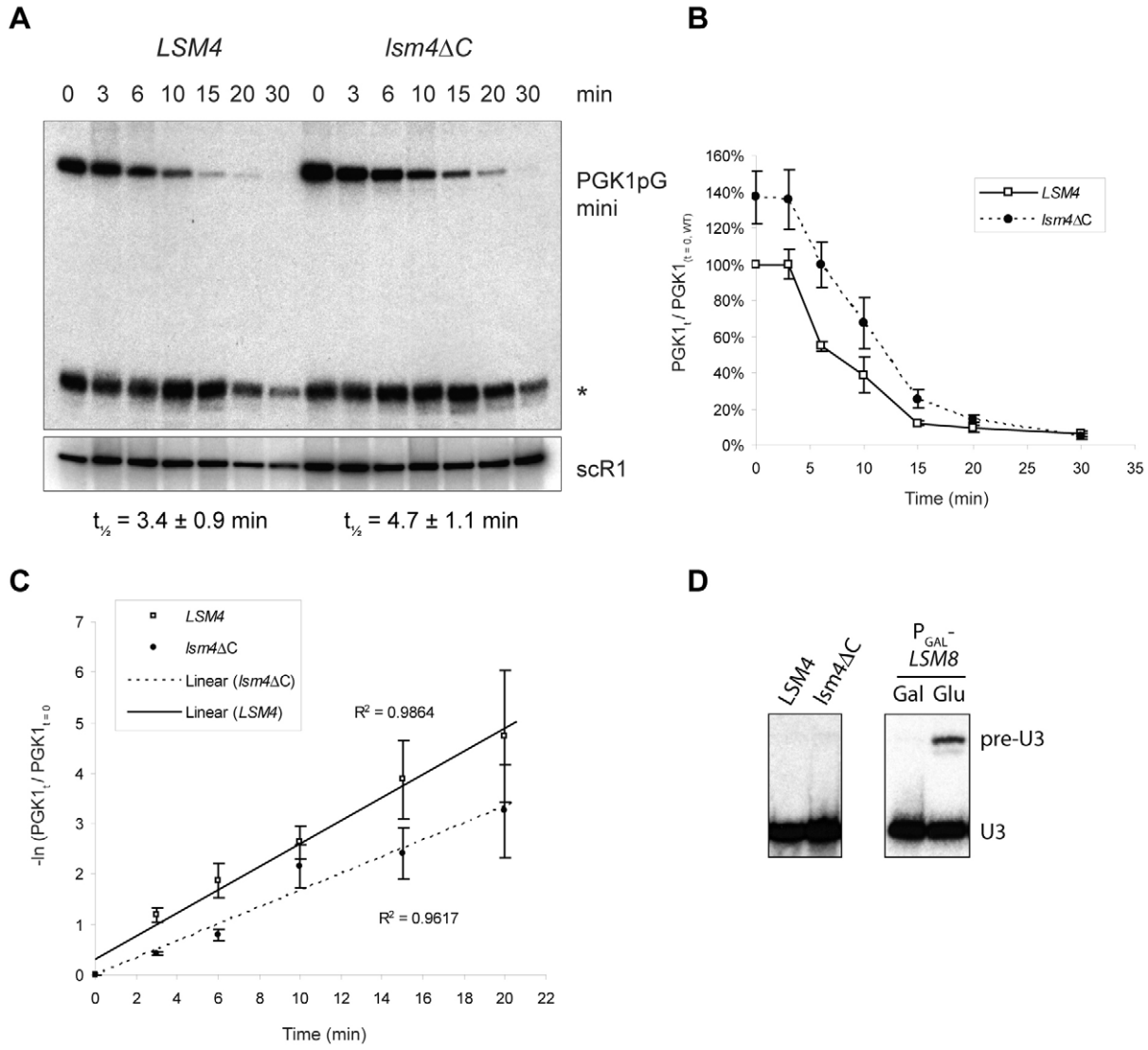


Fig. 4 The Lsm4 C-terminus is required for efficient mRNA degradation, but not splicing. (A) Degradation of PGK1pGmini reporter transcript in *LSM4* (MRY71) or *lsm4ΔC* (MRY73) strains grown in SDGal-Ura after addition of glucose to 4% (w/v). scR1 RNA was used as a loading control. The asterisk indicates a stable degradation fragment. (B) PGK1pGmini transcript levels over time as a percentage of the level in *LSM4* cells at *t*=0; averages of three northern blots with vertical bars indicating standard deviations. (C) Linearized degradation curves [–ln(PGK1$_t$/PGK1$_{t=0}$)] against time showing averages of qRT-PCR data of six RT repeats of two independent biological replicates; vertical bars indicate standard errors; half-lives indicated are based on the linearized qRT-PCR data (D) Northern blot detecting pre-U3 RNA and U3 RNA in *LSM4* and *lsm4ΔC* strains grown in YPDA, and in a P_{GAL}-*LSM8* strain (MPS7) before and after 12 hours of growth on glucose.

Q/N-rich regions in other P-body components

Investigation of the aa sequences of all core components of P-bodies in yeast (Parker and Sheth, 2007) reveals Q and/or N-rich stretches of varying length in many of them, most of which are conserved between various *Saccharomyces* species (Fig. 5, Table 1 and supplementary material Fig. S6). Some (Lsm4p, Ccr4p, Pop2p and Not1p) were previously found in a genome-wide screen looking for yeast proteins with Q/N-rich domains (Michelitsch and Weissman, 2000). Michelitsch and Weissman used an algorithm to count these residues in consecutive aa 80-mers for each of the predicted open reading frames, finding an average Q/N-content of 7.7 per 80-mer in the yeast proteome (Michelitsch and Weissman, 2000). We counted Q, N and P residues in a similar fashion in each of the P-body core components. Our results (Table 1) show that all of the 20 proteins tested score above average for Q/N content (Graphic representations are shown in Fig. 5B,C and supplementary material Fig. S6). Interestingly, some of the *Saccharomyces* homologues show further extensions of Q repeats, e.g. Edc3p, Not3p, Not4p and Not5p (supplementary material Fig. S6). In addition, many of these polypeptides contain high numbers of proline residues in or just downstream of these Q/N-rich regions (Table 1). This is a feature that is also found in other aggregation-prone proteins, e.g. huntingtin, aggregation of which causes Huntington disease (Michelitsch and Weissman, 2000). Proline-rich regions often form extended and flexible regions, in many proteins apparently reaching out to facilitate interactions with other proteins, with phosphorylation having a potential regulatory role. Binding via these proline-rich domains is generally not very specific, but can be both very rapid and strong (Williamson, 1994; Kay et al., 2000). Furthermore, proteins with Q/N-rich domains have

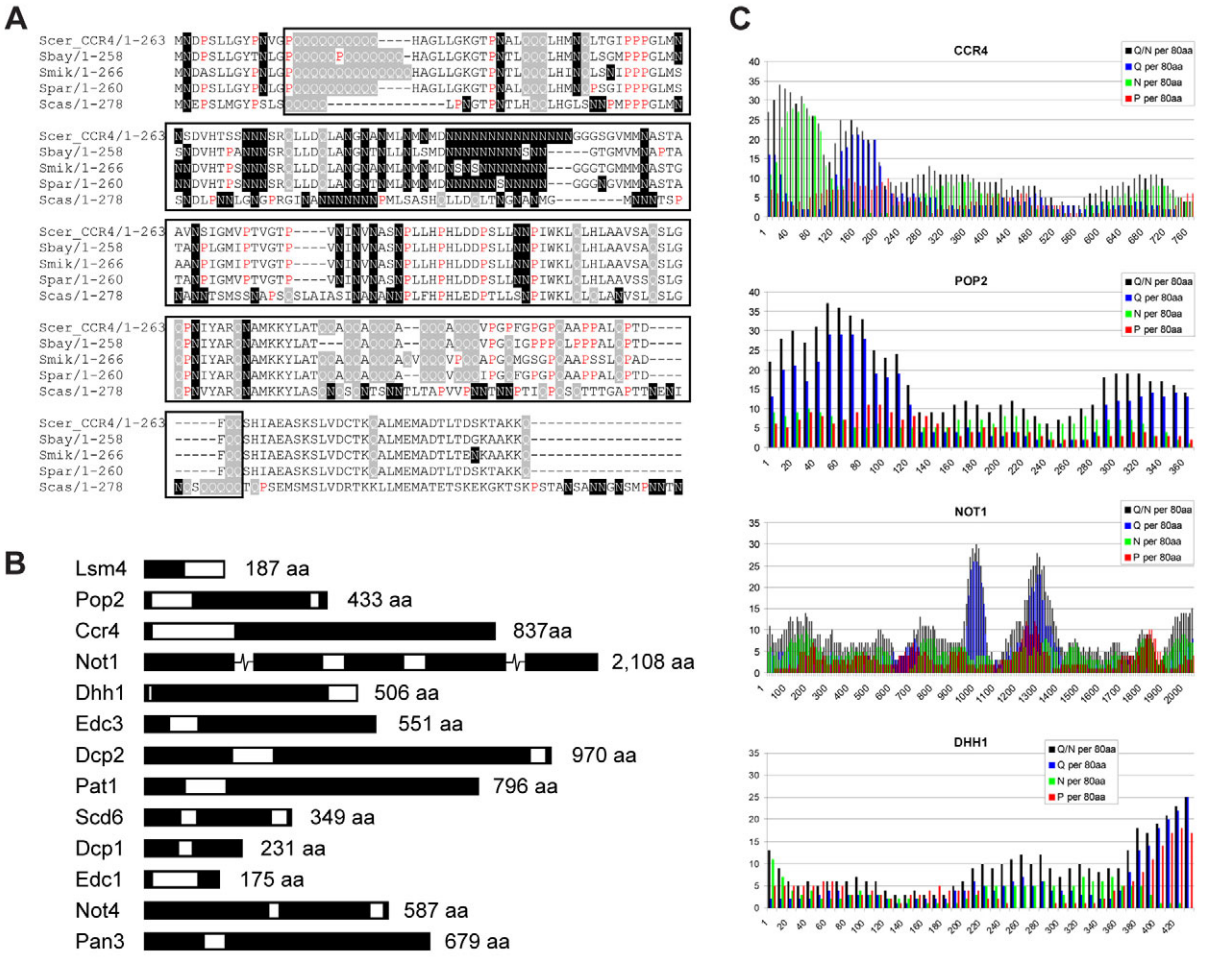


Fig. 5 Other P-body components contain Q/N-rich regions. (A) Alignment of the N-terminal region of *S. cerevisiae* Ccr4p with homologues from closely related *Saccharomyces* species (see Fig. 1 legend for abbreviations). The Q/N-rich region is boxed, with Q residues highlighted in grey, N residues highlighted in black and P residues in red. (B) Schematic representation of P-body components with areas rich in Q and/or N residues indicated in white (approximately to scale; Not1p is broken to fit; lengths are indicated in numbers of aa). (C) Q, N and P residues were counted in aa 80-mers of Ccr4p, Pop2p, Not1p and Dhh1p starting at position 1, shifting ten aa at a time.

Table 1. Q, N and P residues counted per amino acid 80-mers in 20 different P-body components

Protein	Q/N*	Q	N	P in Q/N-rich	P (in best 80-mer)*	Percentage of P (in best 80-mer)	P-rich close to Q/N-rich
Lsm4	38	5	33	1	2	2.5	
Pop2	37	29	8	6	12	15.0	yes
Ccr4	34	11	23	4	10	12.5	yes
Not1	30	26	4	1	12	15.0	yes
Edc3	29	19	10	1	5	6.3	
Dhh1	25	25	0	17	19	23.8	yes
Dcp2	24	13	11	8	20	25.0	yes
Pat1	23	19	4	18	23	28.8	yes
Scd6	22	8	14	2	15	18.8	(yes)†
Dcp1	22	4	18	2	4	5.0	
Edc1	22	5	17	10	10	12.5	yes
Not4	21	11	8	9	10	12.5	yes
Pan3	20	0	20	6	15	18.8	yes
Not3	19	9	10	2	13	16.3	(yes)†
Not5	15	8	7	3	12	15.0	
Xrn1	15	5	10	1	24	30.0	(yes)†
Pan2	15	7	8	2	10	12.5	
Lsm1	14	8	6	1	2	2.5	
Not2	11	8	3	7	8	10.0	
Edc2	8	5	3	6	9	11.3	

*Values indicate the number of residues in the highest scoring 80-mer for that protein.
†P-rich region near Q/N-rich region scoring <20.

Journal of Cell Science

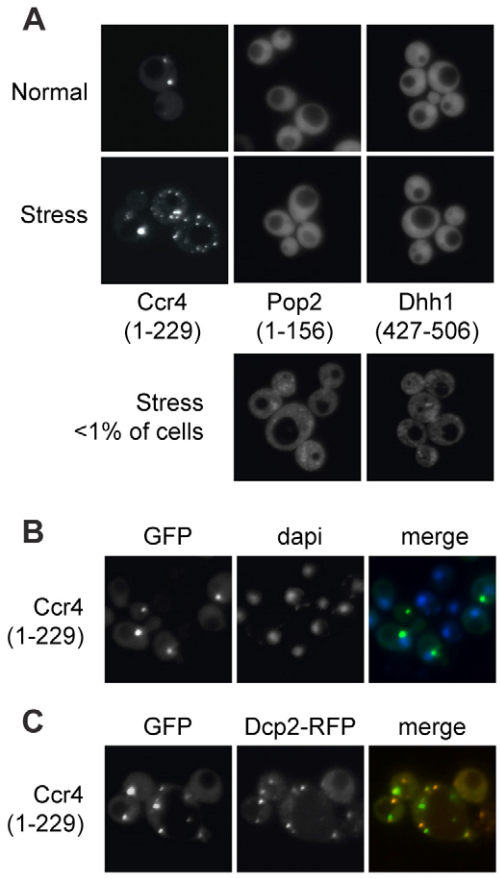


Fig. 6 The Q/N-rich region from Ccr4p aggregates in cytoplasmic foci and responds to stress. (A) Localization of GFP-tagged Q/N-rich regions of Ccr4p (aa 1-229; pMR202), Pop2p (aa 1-156; pMR203) and Dhh1p (aa 427-506; pMR204) before and after hypo-osmotic shock (B) GFP-Ccr4(1-229) aggregates localize to the cytoplasm as shown in these fixed cells with DAPI stained nuclear DNA (C) The majority of GFP-Ccr4(1-229) aggregates does not colocalize with Dcp2-RFP (pRP1155) foci after osmotic shock.

previously been shown to promote aggregation of heterologous proteins with similar domains (Derkatch et al., 2004). Indeed, Lsm4p was found as one of nine Q/N-rich proteins that, when overproduced, promote de novo appearance of [PSI^{+}], the prion-form of the Q/N-rich Sup35 protein (Derkatch et al., 2001). On the basis of this behaviour as well as its structural similarities to Sup35p, these authors proposed that Lsm4p itself is a prion protein. Furthermore, Decker et al. showed that the prion-like Q/N-rich domain of the Rnq1 prion protein can, at least in part, functionally replace the C-terminal prion-like domain of Lsm4p (Decker et al., 2007).

Q/N-rich regions affect P-body localization

It is plausible that Q, N and/or P-rich regions have a role in the accumulation of proteins in P-bodies. We tested Q/N-rich regions from Ccr4p, Pop2p and Dhh1p for their ability to aggregate and/or accumulate in P-bodies when fused to GFP. The Q/N-rich N-terminal region of Ccr4p fused to GFP [Ccr4(1-229)] aggregates in cytoplasmic foci under normal growth conditions (Fig. 6A,B) in ~20% of cells, and foci increase in numbers under stress conditions, with more than 50% of cells showing multiple foci per cell. Although the dynamics of increased focal accumulation resembled that of P-body formation, suggesting that the Q/N-rich N-terminus of Ccr4p is sufficient for P-body localization, we found that the majority did not colocalise with Dcp2-RFP (Fig. 6C). GFP-fusions of the Q/N-rich regions of Pop2p [Pop2(1-156)] and Dhh1p [Dhh1(427-506)], however, do not aggregate under normal growth conditions but show weak focal concentration in a low percentage of cells (<1%) when stressed, although the majority of cells do not show a change in localization (Fig. 6A). However, GFP-fusions of Pop2p and Dhh1p deleted for these domains [Pop2ΔN(147-433)] and [Dhh1ΔC(1-427)] do show decreased P-body localization compared to full-length Pop2p and Dhh1p (Fig. 7A), and Ccr4p deleted for 147 aa at its N-terminus [Ccr4ΔN(148-837)] completely fails to accumulate in cytoplasmic foci under stress conditions. We quantified the P-body localization of these proteins by counting the number of visible foci per cell at a set time after osmotic shock (Table 2). These numbers are an indication of the level of P-body localization, as a reduction in P-body accumulation will lead to a reduction in the number of visible P-bodies, which generally have variable sizes and/or intensities. Interestingly, deletion of a further 102 aa from the Ccr4p N-terminus [Ccr4ΔN2(250-837)] leads to exclusively nuclear localization (Fig. 7B). The latter suggests that Ccr4p normally shuttles between the nucleus and cytoplasm, and that its nuclear export depends on sequences within the N-terminal domain. The tendency for aggregation of these Q/N-rich regions is further emphasized by the fact that full-length Pop2p expressed from the *MET25* promoter aggregates in bright nuclear foci when tagged at the C-terminus (Pop2-GFP, Fig. 7C), at a much lower rate when tagged at the (Q/N-rich) N-terminus (data not shown) and not at all in the absence of this N-terminus (Pop2ΔN-GFP, Fig. 7C). As these experiments were performed in the presence of natively expressed non-tagged proteins, which may contribute to the observed absence of GFP-Ccr4ΔN concentration in P-bodies, we investigated the localization of this protein in *ccr4Δ* as well as *xrn1Δ* strains. Whereas GFP-Ccr4ΔN still failed to concentrate in microscopically visible foci in the absence of native Ccr4p (Fig. 7D), some weak foci were observed in the absence of Xrn1p (supplementary material Fig. S7), which generally leads to larger and more abundant P-bodies by preventing 5′-to-3′ degradation of transcripts. For Ccr4ΔNp and Pop2ΔNp the reduced P-body localization is not due to reduced levels of these truncated proteins as western analysis showed no difference between levels of full-length and mutant proteins (Fig. 7E). As the level of Dhh1ΔCp was only 60% of that of full-length Dhh1p we cannot rule out that its reduced P-body localization is, in part, due to the lower protein level.

In summary, although not absolutely essential, Q/N-rich sequences in Pop2p, Ccr4p and Dhh1p contribute to efficient accumulation of these proteins in P-bodies under stress conditions. This is most obvious for Ccr4p, which, in the absence of its N-terminal 147 aa is not microscopically detectable in P-bodies in otherwise normal cells. Increased focal accumulation under stress conditions of the N-terminal 229 aa fused to GFP, suggests that this region is capable of regulated aggregation in response to stress. The fact that the majority of these foci do not colocalize with Dcp2p, suggests that additional parts of Ccr4p are necessary for proper P-body localization, most probably through additional protein-protein interactions. It would therefore be interesting to further investigate the requirements of the Q/N-rich regions as well as other parts of these proteins for these interactions.

Is a mechanism for protein accumulation in P-bodies conserved?

As the C-terminal region of *S. cerevisiae* Lsm4p is semi-conserved between *Saccharomyces* species, at least in the high content of N

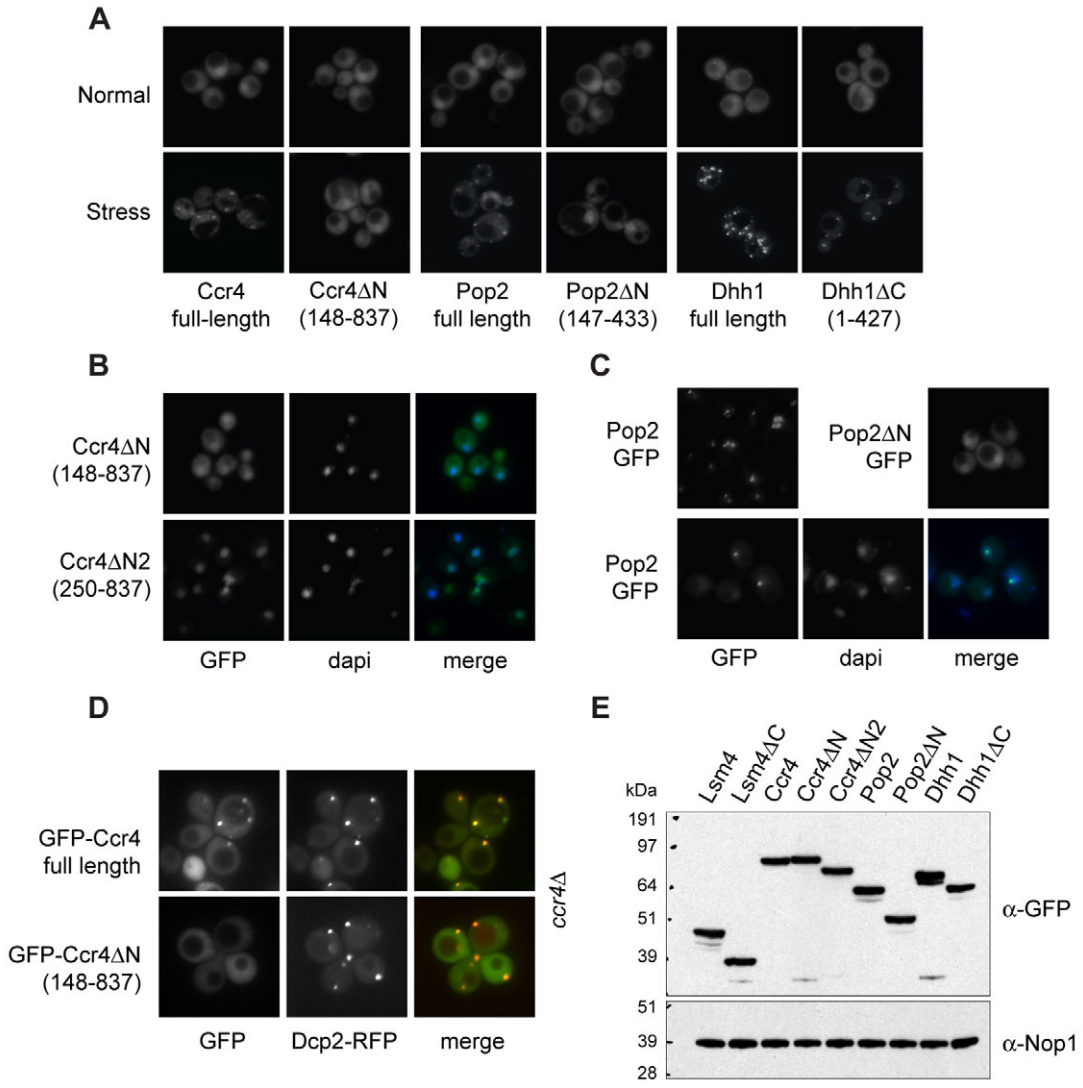


Fig. 7 Q/N-rich regions from Ccr4p, Pop2p and Dhh1p contribute to efficient accumulation of these proteins in P-bodies. (B) Localization of GFP-tagged full-length Ccr4p (pMR212), Pop2p (pMR214), Dhh1p (pMR210) or truncated versions of these proteins (Ccr4ΔN(148-837) from pMR218, Pop2ΔN(147-433) from pMR215, Dhh1ΔC(1-427) from pMR211) before and after osmotic shock. All GFP-fusions were expressed in BY4741 cells and localization was examined in cells during normal growth (normal) or 20-40 minutes after osmotic shock (stress). (B) Localization of GFP-tagged Ccr4ΔN(148-837) and Ccr4ΔN2 (aa 250-837; pMR213) in fixed cells with DAPI-stained nuclear DNA (C) Localization of C-terminally GFP-tagged full-length Pop2p (pMR216) or Pop2ΔNp (pMR217); DAPI-stained nuclear DNA in blue. (D) Localization of GFP-Ccr4, GFP-Ccr4ΔN and Dcp2-RFP in *ccr4Δ* cells (Y10387) 30 minutes after hypo-osmotic shock (E) Anti-GFP western blot analysis of full-length and truncated Lsm4, Ccr4, Pop2 and Dhh1 proteins. Curiously, GFP-Ccr4 (122 kDa) migrates faster than GFP-Ccr4ΔN (106 kDa), but slower than GFP-Ccr4ΔN2 (95 kDa; Fig. 7E), and all three GFP-Ccr4 proteins migrate faster than their predicted molecular weights. The presence or absence of the highly polar N-terminal region of Ccr4p causes a change in the effective charge of the entire protein (predicted charges at pH 7 are –5.8, –4.0 and –5.5 respectively) and might affect protein conformation resulting in unusual migration during SDS-PAGE. Nop1p was used as a loading control.

and/or Q residues (supplementary material Fig. S1), the ability to promote Lsm1-7p accumulation in P-bodies is likely to be conserved in these yeasts as well as in other budding yeasts. In fact, this was shown to be true for the budding yeast *K. lactis* Lsm4p produced in *S. cerevisiae* (Mazzoni et al., 2007). The human homologue, however, does not show a significant enrichment in Q or N residues, apart from a short stretch of five glutamines. Indeed, full-length human LSM4 fused to GFP did not aggregate when overexpressed in wild-type yeast cells, nor did it accumulate in foci under stress conditions. Surprisingly, it mostly accumulated in the nucleus instead (data not shown). As it was not able to support viability in the absence of native Lsm4p expression (data not shown), it might be unable to form a functional complex with yeast Lsm proteins. It is possible that residues in other human LSM1-7 complex (comprising LSM1 to LSM7) members normally contribute to its accumulation in P-bodies. Notably, the short N or C-terminal extensions of LSM1, LSM2, LSM3 and LSM7 proteins contain relatively high levels of glutamine residues. However, if, as we propose, Q/N-rich sequences contribute to a rapid response to stimuli in yeast, this may not be needed in human cells, as the LSM1-7 complex accumulates in P-bodies even under normal growth conditions.

Q/N-rich regions do not seem to be conserved in the human homologues of budding yeast Ccr4p, Pop2p and Dhh1p (CNOT6,

Table 2. Quantification of P-body localization

Protein	Foci per cell	Cells examined
Ccr4	2.7±0.5	138
Ccr4ΔN (aa148-837)	0	>1000
Pop2	4.8±1.6	80
Pop2ΔN (aa147-433)	1.9±0.6	93
Dhh1	10.7±5.3	35
Dhh1ΔC (aa1-427)	3.5±1.1	83

CNOT8 and DDX6, respectively) proteins either; they are significantly shorter, lacking the N-terminal and C-terminal Q/N-rich regions respectively (supplementary material Fig. S8). Perhaps the function of these protein domains has been replaced by alternative domains, possibly in other polypeptides with which they interact. For example, GW182 contains an internal Q/P-rich region that is essential, but not sufficient, for its own P-body localization and that of Ago1 (Behm-Ansmant et al., 2006). Another P-body component specific to higher eukaryotes, Ge-1/Hedls, contains a C-terminal repetitive sequence rich in hydrophobic residues that is essential for P-body localization and parts of which aggregate in cytoplasmic foci that are not P-bodies (Yu et al., 2005). In addition EDC4 (also known as Ge-1, Hedls), DCP2 and TNRC6B from humans as well as other higher eukaryotes contain high levels of Q and/or N residues (Decker et al., 2007). Thus, alternative aggregation-prone regions might have replaced some of the yeast Q/N-rich domains in higher eukaryotes, at least some of which are likely to have a role in P-body assembly.

Aggregation of P-body components through their Q/N-rich regions could promote efficient P-body formation. Whether this is really the case and, if so, whether this occurs through prion-like aggregation or through specific interactions via putative modular 'polar zipper' protein-protein interaction domains (Perutz et al., 1994; Michelitsch and Weissman, 2000) remains to be determined. The importance of the Q/N-rich protein Edc3p (also known as Lsm16p) in combination with Lsm4p in P-body assembly in yeast, which came to light while this manuscript was being revised (Decker et al., 2007), is in support of this hypothesis. An intriguing question is how Lsm4p aggregation, and that of other P-body components, is prevented under normal growth conditions. Post-translational modifications, e.g. phosphorylation of Lsm4p or other (Lsm) proteins, probably have a role. Such modifications could allow the cell to respond quickly and efficiently to changes in conditions, and might regulate the levels and intracellular localizations of Lsm1-7p and Lsm2-8p, in addition to promoting P-body localization. Such a mechanism could also regulate the competition between these two complexes that was observed by Spiller et al. (Spiller et al., 2007b). Similarly, post-translational modifications, e.g. of the N-terminal region of Ccr4p, could allow P-body localization of other proteins involved in RNA degradation. Q-rich regions in mouse TIA-1 and PUM2 have previously been shown to contribute to protein accumulation in stress granules (Gilks et al., 2004; Vessey et al., 2006). We now show that at least some of the Q/N-rich domains in P-body components have a role in the assembly of these RNA processing bodies. The presence of Q/N-rich regions in many other proteins that are involved in various aspects of RNA metabolism (Michelitsch and Weissman, 2000; Decker et al., 2007) hints at the possibility of a more general role for these prion-like domains in functional protein aggregation, in addition to stress-granule and P-body assembly.

Materials and Methods

Plasmids and strains

For a complete list of plasmids and strains used see supplementary material Tables S1 and S2.

Microscopy

Cells were grown at 30°C to mid-log phase in synthetic dropout (SD) medium. To stress cells, cultures were centrifuged and cells were resuspended in water. Live cells were placed on microscopy slides and examined by bright-field and/or fluorescence microscopy using a Leica FW4000 fluorescence microscope. Fixing of cells followed by DAPI staining was performed as previously described (Spiller et al., 2007b). Images were captured using LeicaFW4000 software (Scanalytics, Fairfax, VA) with a CH-250 16-bit, cooled CCD camera (Photometrics, Tucson, AZ).

RNA analyses

Cultures were grown at 30°C in synthetic dropout medium containing 2% (w/v) galactose. Transcription of the PGK1pGmini reporter gene was stopped by the addition of glucose to 4% (w/v) and 20 ml culture with an OD_{600} of 0.5 were snap-chilled at the indicated times after the addition of glucose. RNA extractions and northern blot analyses of 6% acrylamide/urea gels were as described (Mayes et al., 1999). The following oligonucleotide probes were used for northern hybridizations: to detect the PGK1pGmini reporter transcript 5′-AATTGATCTATCGAGGAATTCC-3′, to detect scR1 RNA 5′-ATCCCGGCCGCCTCCATCAC-3′ and to detect U3 RNA 5′-GGTTATGGGACTCATCA-3′. Northern blots were quantified using a STORM 860 PhosphorImager and ImageQuant software (Molecular Dynamics).

Quantitative reverse-transcriptase PCR

Ten µg of total RNA were treated with DNase1 (0.9 U RQ1, Promega) according to the manufacturer's instructions. cDNA was prepared from 5 µg of DNase-treated RNA in a 10 µl reaction: 1× first strand synthesis buffer, 2.5 mM DTT, 10 U RNase inhibitor (Roche), 0.75 mM dNTPs, 7.5 U ThermoScript RNaseH$^-$ (Invitrogen) and 500 nM of PGK1pGmini-specific primer (5′-AGCGTAAAGGATGGGGAAA-GAGAA-3′), according to the manufacturer's instructions. A negative control reaction was performed in the absence of reverse transcriptase (RT). Any remaining RNA was hydrolysed by incubating reactions for 1 hour at 37°C after addition of 15 µl of 0.1 mg/ml RNaseA (Roche). Quantitative PCRs (qPCRs) were performed with SYBR Green JumpStart Taq ReadyMix (Sigma) in a Stratagene MX3005P real-time PCR machine in 10 µl reactions: 6 µl containing 5 µl 2× SYBR Green ReadyMix, 300 nM of each primer (F: 5′-ATTGAAATGAAATGAAATCGAAGGAATTTGG-3′; R: 5′-AGCGTAAAGGATGGGGAAAGAGAA-3′) and 0.5× ROX, plus 4 µl of cDNA template (diluted 1 in 20 after RT-PCR). Cycling parameters were as follows: 2 minutes at 94°C, then 50 cycles of 10 seconds at 94°C, 10 seconds at 63°C and 20 seconds at 72°C. Each qPCR reaction was performed in triplicate for each repeat RT reaction.

Western analysis

For crude protein extracts (Volland et al., 1994), yeast cells were lysed in 0.5 ml of 0.2 M NaOH on ice for 10 minutes, followed by TCA precipitation (final 5% w/v) for 10 minutes on ice. After centrifugation, the pellet was resuspended in 35 µl of dissociation buffer (0.1 M Tris-HCl pH 6.8, 4 mM EDTA, 4% SDS, 20% (v/v) glycerol, 2% (v/v) β-mercaptoethanol, 0.02% (w/v) BPB) and 15 µl of 1 M Tris base. Samples were heated at 95°C for 10 min before separation by SDS-PAGE. Proteins were transferred to PVDF membrane and detected with mouse anti-GFP (BD Bioscience) or anti-Nop1p antibodies, and sheep anti-mouse IgG-HRP (Amersham Bioscience).

Polypeptide alignments

Amino acid sequences of P-body components were obtained from the *Saccharomyces* Genome Database (http://www.yeastgenome.org/) or the NCBI Entrez Protein database (http://www.ncbi.nlm.nih.gov/sites/entrez). Alignments were made using the ClustalW Multiple Sequence Alignment tool (Thompson et al., 1994) inside Jalview 2.2 (Clamp et al., 2004).

We thank Roy Parker and David Tollervey for reagents, Jon Houseley for the critical reading of the manuscript, and David Barrass, Daniela Hahn and Olivier Cordin for assistance. This work was funded by a Wellcome Trust Prize Studentship 71448 and Royal Society support to MAMR, a studentship from The Darwin Trust of Edinburgh to MPS and Wellcome Trust Grant 067311. J.D.B. is the Royal Society Darwin Trust Research Professor. M.A.M.R. designed and performed all experiments under the supervision of J.D.B., except the qPCR experiment which was performed by R.D.A.; M.P.S. first suggested a potential role for the Lsm4p C-terminus in P-body localization; M.A.M.R. and J.D.B. wrote the manuscript.

Journal of Cell Science

References

Achsel, T., Brahms, H., Kastner, B., Bachi, A., Wilm, M. and Lührmann, R. (1999). A doughnut-shaped heteromer of human Sm-like proteins binds to the 3′-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. *EMBO J.* **18**, 5789-5802.

Beckham, C., Hilliker, A., Cziko, A. M., Noueiry, A., Ramaswami, M. and Parker, R. (2008). The DEAD-Box RNA Helicase Ded1p Affects and Accumulates in Saccharomyces cerevisiae P-Bodies. *Mol. Biol. Cell* **19**, 984-993.

Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P. and Izaurralde, E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes Dev.* **20**, 1885-1898.

Bouveret, E., Rigaut, G., Shevchenko, A., Wilm, M. and Séraphin, B. (2000). A Sm-like protein complex that participates in mRNA degradation. *EMBO J.* **19**, 1661-1671.

Brachmann, C. B., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P. and Boeke, J. D. (1998). Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* **14**, 115-132.

Brahms, H., Meheus, L., de Brabandere, V, Fischer, U. and Lührmann, R. (2001). Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B′ and the Sm-like protein LSm4, and their interaction with the SMN protein. *RNA* **7**, 1531-1542.

Clamp, M., Cuff, J., Searle, S. M. and Barton, G. J. (2004). The Jalview Java alignment editor. *Bioinformatics* **20**, 426-427.

Cooper, M., Parkes, V., Johnston, L. H. and Beggs, J. D. (1995). Identification and characterization of Uss1p (Sdb23p): a novel U6 snRNA-associated protein with significant similarity to core proteins of small nuclear ribonucleoproteins. *EMBO J.* **14**, 2066-2075.

Cougot, N., Babajko, S. and Séraphin, B. (2004). Cytoplasmic foci are sites of mRNA decay in human cells. *J. Cell Biol.* **165**, 31-40.

Decker, C. J., Teixeira, D. and Parker, R. (2007). Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in Saccharomyces cerevisiae. *J. Cell Biol.* **179**, 437-449.

Derkatch, I. L., Bradley, M. E., Hong, J. Y. and Liebman, S. W. (2001). Prions affect the appearance of other prions: the story of [PIN(+)]. *Cell* **106**, 171-182.

Derkatch, I. L., Uptain, S. M., Outeiro, T. F., Krishnan, R., Lindquist, S. L. and Liebman, S. W. (2004). Effects of Q/N-rich, polyQ, and non-polyQ amyloids on the de novo formation of the [PSI+] prion in yeast and aggregation of Sup35 in vitro. *Proc. Natl. Acad. Sci. USA* **101**, 12934-12939.

Eystathioy, T., Jakymiw, A., Chan, E. K., Séraphin, B., Cougot, N. and Fritzler, M. J. (2003). The GW182 protein colocalizes with mRNA degradation associated proteins hDcp1 and hLSm4 in cytoplasmic GW bodies. *RNA* **9**, 1171-1173.

Gilks, N., Kedersha, N., Ayodele, M., Shen, L., Stoecklin, G., Dember, L. M. and Anderson, P. (2004). Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol. Biol. Cell* **15**, 5383-5398.

Hermann, H., Fabrizio, P., Raker, V. A., Foulaki, K., Hornig, H., Brahms, H. and Lührmann, R. (1995). snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions. *EMBO J.* **14**, 2076-2088.

Ingelfinger, D., Arndt-Jovin, D. J., Lührmann, R. and Achsel, T. (2002). The human LSm1-7 proteins colocalize with the mRNA-degrading enzymes Dcp1/2 and Xrnl in distinct cytoplasmic foci. *RNA* **8**, 1489-1501.

Jakymiw, A., Lian, S., Eystathioy, T., Li, S., Satoh, M., Hamel, J. C., Fritzler, M. J. and Chan, E. K. (2005). Disruption of GW bodies impairs mammalian RNA interference. *Nat. Cell Biol.* **7**, 1267-1274.

Jakymiw, A., Pauley, K. M., Li, S., Ikeda, K., Lian, S., Eystathioy, T., Satoh, M., Fritzler, M. J. and Chan, E. K. (2007). The role of GW/P-bodies in RNA processing and silencing. *J. Cell Sci.* **120**, 1317-1323.

Kay, B. K., Williamson, M. P. and Sudol, M. (2000). The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J.* **14**, 231-241.

Liu, J., Rivas, F. V., Wohlschlegel, J., Yates, J. R., III Parker, R. and Hannon, G. J. (2005). A role for the P-body component GW182 in microRNA function. *Nat. Cell Biol.* **7**, 1261-1266.

Mayes, A. E., Verdone, L., Legrain, P. and Beggs, J. D. (1999). Characterization of Sm-like proteins in yeast and their association with U6 snRNA. *EMBO J.* **18**, 4321-4331.

Mazzoni, C., Mancini, P., Madeo, F., Palermo, V. and Falcone, C. (2003a). A *Kluyveromyces lactis* mutant in the essential gene KlLSM4 shows phenotypic markers of apoptosis. *FEMS Yeast Res.* **4**, 29-35.

Mazzoni, C., Mancini, P., Verdone, L., Madeo, F., Serafini, A., Herker, E. and Falcone, C. (2003b). A truncated form of KlLsm4p and the absence of factors involved in mRNA decapping trigger apoptosis in yeast. *Mol. Biol. Cell* **14**, 721-729.

Mazzoni, C., D'Addario, I. and Falcone, C. (2007). The C-terminus of the yeast Lsm4p is required for the association to P-bodies. *FEBS Lett.* **581**, 4836-4840.

Michelitsch, M. D. and Weissman, J. S. (2000). A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions.*Proc. Natl. Acad. Sci. USA* **97**, 11910-11915.

Mitchell, P. and Tollervey, D. (2003). An NMD pathway in yeast involving accelerated deadenylation and exosome-mediated 3′ 5′ degradation. *Mol. Cell* **11**, 1405-1413.

Niedenthal, R. K., Riles, L., Johnston, M. and Hegemann, J. H. (1996). Green fluorescent protein as a marker for gene expression and subcellular localization in budding yeast. *Yeast* **12**, 773-786.

Pannone, B. K., Xue, D. and Wolin, S. L. (1998). A role for the yeast La protein in U6 snRNP assembly: evidence that the La protein is a molecular chaperone for RNA polymerase III transcripts. *EMBO J.* **17**, 7442-7453.

Parker, R. and Sheth, U. (2007). P bodies and the control of mRNA translation and degradation. *Mol. Cell* **25**, 635-646.

Paushkin, S., Gubitz, A. K., Massenet, S. and Dreyfuss, G. (2002). The SMN complex, an assemblyosome of ribonucleoproteins. *Curr. Opin. Cell Biol.* **14**, 305-312.

Perutz, M. F., Johnson, T., Suzuki, M. and Finch, J. T. (1994). Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. *Proc. Natl. Acad. Sci. USA* **91**, 5355-5358.

Pilkington, G. R. and Parker, R. (2008). Pat1 contains distinct functional domains that promote P-body assembly and activation of decapping. *Mol. Cell. Biol.* **28**, 1298-1312.

Rehwinkel, J., Behm-Ansmant, I., Gatfield, D. and Izaurralde, E. (2005). A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. *RNA* **11**, 1640-1647.

Salgado-Garrido, J., Bragado-Nilsson, E., Kandels-Lewis, S. and Séraphin, B. (1999). Sm and Sm-like proteins assemble in two related complexes of deep evolutionary origin. *EMBO J.* **18**, 3451-3462.

Séraphin, B. (1995). Sm and Sm-like proteins belong to a large family: identification of proteins of the U6 as well as the U1, U2, U4 and U5 snRNPs. *EMBO J.* **14**, 2089-2098.

Sheth, U. and Parker, R. (2003). Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science* **300**, 805-808.

Spiller, M. P., Boon, K. L., Reijns, M. A. M. and Beggs, J. D. (2007a). The Lsm2-8 complex determines nuclear localization of the spliceosomal U6 snRNA. *Nucleic Acids Res.* **35**, 923-929.

Spiller, M. P., Reijns, M. A. and Beggs, J. D. (2007b). Requirements for nuclear localization of the Lsm2-8p complex and competition between nuclear and cytoplasmic Lsm complexes. *J. Cell Sci.* **120**, 4310-4320.

Teixeira, D. and Parker, R. (2007). Analysis of P-body assembly in Saccharomyces cerevisiae. *Mol. Biol. Cell* **18**, 2274-2287.

Teixeira, D., Sheth, U., Valencia-Sanchez, M. A., Brengues, M. and Parker, R. (2005). Processing bodies require RNA for assembly and contain nontranslating mRNAs. *RNA* **11**, 371-382.

Tharun, S., He, W., Mayes, A. E., Lennertz, P., Beggs, J. D. and Parker, R. (2000). Yeast Sm-like proteins function in mRNA decapping and decay. *Nature* **404**, 515-518.

Tharun, S., Muhlrad, D., Chowdhury, A. and Parker, R. (2005). Mutations in the Saccharomyces cerevisiae LSM1 gene that affect mRNA decapping and 3′ end protection. *Genetics* **170**, 33-46.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680.

Verdone, L., Galardi, S., Page, D. and Beggs, J. D. (2004). Lsm proteins promote regeneration of pre-mRNA splicing activity. *Curr. Biol.* **14**, 1487-1491.

Vessey, J. P., Vaccani, A., Xie, Y., Dahm, R., Karra, D., Kiebler, M. A. and Macchi, P. (2006). Dendritic localization of the translational repressor Pumilio 2 and its contribution to dendritic stress granules. *J. Neurosci.* **26**, 6496-6508.

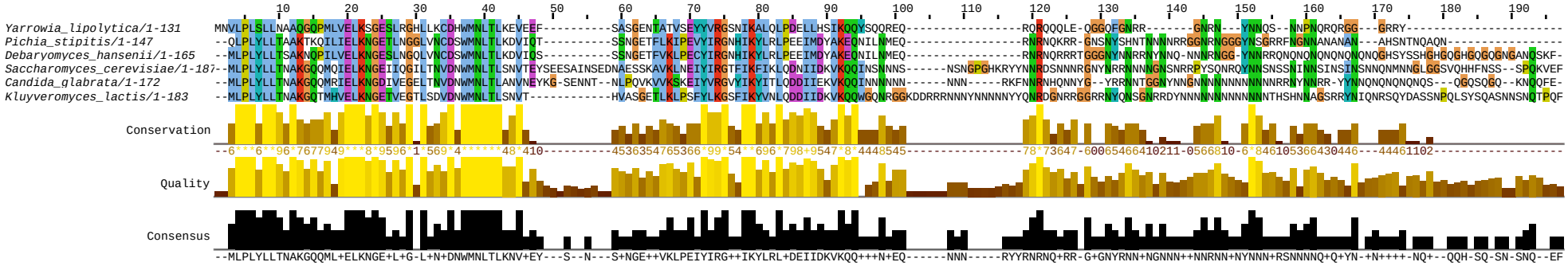
Volland, C., Urban-Grimal, D., Geraud, G. and Haguenauer-Tsapis, R. (1994). Endocytosis and degradation of the yeast uracil permease under adverse conditions. *J. Biol. Chem.* **269**, 9833-9841.

Williamson, M. P. (1994). The structure and function of proline-rich regions in proteins. *Biochem. J.* **297** (Pt 2), 249-260.

Yu, J. H., Yang, W. H., Gulick, T., Bloch, K. D. and Bloch, D. B. (2005). Ge-1 is a central component of the mammalian cytoplasmic mRNA processing body. *RNA* **11**, 1795-1802.

Fig. S1

A



B

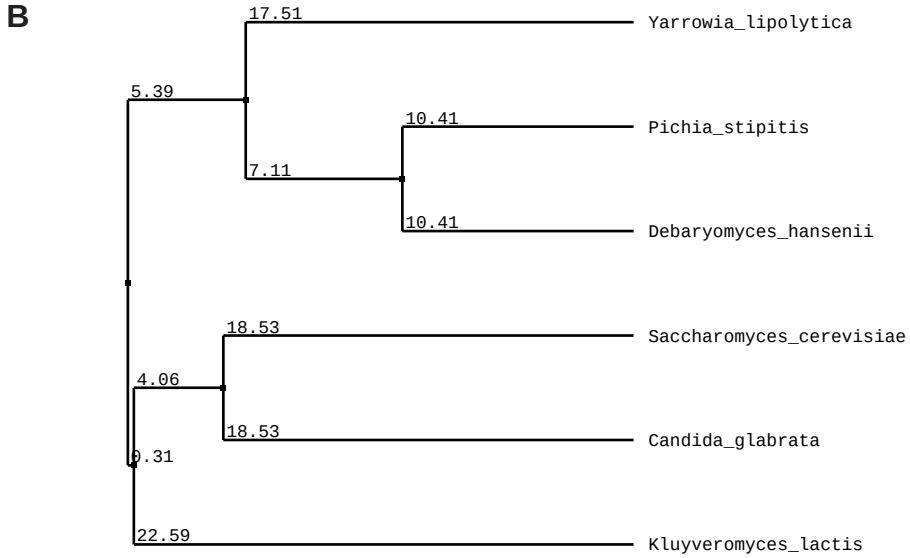


Fig. S1 The C-termini of budding yeast Lsm4 proteins contain high levels of N and/or Q residues.
(A) Alignment of Lsm4 proteins from budding yeasts
(B) Phylogenetic tree of budding yeast Lsm4 proteins

Table S2. Plasmids used in this study

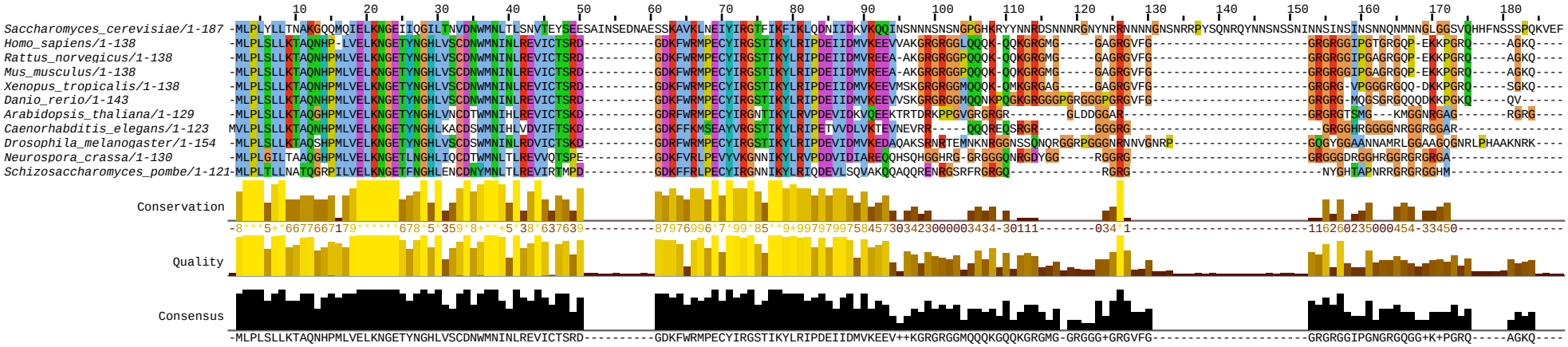
Plasmid	Description	Reference
pGFP-N-FUS	*CEN6*, *URA3* shuttle vector with MCS for N-terminal GFP tagging under control of *MET25* promoter	Niedenthal et al., 1996
pGFP-C-FUS	*CEN6*, *URA3* shuttle vector with MCS for C-terminal GFP tagging under control of *MET25* promoter	Niedenthal et al., 1996
pUSS1	1.7 kb *Hind*III genomic fragment with *LSM4* ORF, promoter and terminator sequences cloned into pFL39 (*CEN*6, *TRP1*)	Cooper et al., 1995
pMC12	Modified pFL39 (*CEN6*, *TRP1*) with a 1.36 kb *Hind*III/*Sal*I fragment containing truncated *LSM4* (aa 1-92) with *LSM4* promoter region	This work
pMC12 was cloned as follows: A 1.36 kb fragment containing the first 276 bp of the *LSM4* ORF and upstream sequences was generated by PCR using pBS-SDB23 (genomic *BamH*I restriction fragment cloned into pBluescript KS^+ cloned such that the *LSM4* gene faces in the same direction as the T7 promoter; described by Cooper *et al.* (1995)) as template and primers T7 (AATACGACTCACTATAG) and M0055 (GCAGTCGACTTGCTGCTTGACCTT). The resulting product was cut with *Hind*III and *Sal*I, and cloned between the *Hind*III and *Sal*I restriction sites of pFL39, upstream of a stop-codon linker (TCGATTAACTAACTAG and TCGACTAGTTAGTTAA) inserted into the *Sal*I restriction site of this vector. The resulting vector expresses the first 92 amino acids of Lsm4p from its native promoter, but lacks its normal 3' UTR.		
pMPSLsm4	pGFP-N-LSM4: *LSM4* coding region inserted in frame with *GFP* coding region in pGFP-N-FUS	This work
pMPSLsm4D1	pGFP-N-lsm4ΔC: *lsm4ΔC* coding region (aa 1-93) inserted in frame with *GFP* coding region	This work
pMPSLsm4D2	pGFP-N-lsm4C: *lsm4C* coding region (aa 92-187) inserted in frame with *GFP* coding region	This work
pGFP-N-LSM1	*LSM1* coding region in frame with N-terminal GFP tag	This work
pMPSLsm2	pGFP-N-LSM2: *LSM2* coding region in frame with N-	This work

	terminal GFP tag in pGFP-N-FUS	
pMPSLsm6	pGFP-N-LSM6: *LSM6* coding region in frame with N-terminal GFP tag in pGFP-N-FUS	This work
pMPSLsm7	pGFP-N-LSM7: *LSM7* coding region in frame with N-terminal GFP tag in pGFP-N-FUS	This work
pMPS2	pGFP-N-FUS, with *LHP1* coding region inserted in frame with *GFP* coding region	Spiller et al., 2007a
pMR83	pGFP-C-LSM8: *LSM8* coding region in frame with C-terminal GFP in pGFP-C-FUS	Spiller et al., 2007b
pRP1155	*DCP2* with C-terminal RFP tag in *CEN*, *LEU* shuttle vector (gift from R. Parker)	Teixeira et al., 2005
pMR171	Promoter-*DCP2-RFP-PGK1* terminator from pRP1155 cloned into pRS313 (*CEN6*, *HIS3*)	This work
pPM119	pUN50 (*CEN4*, *URA3*) with P_{GAL}-*PGKpGmini* reporter (gift from D. Tollervey)	Mitchell and Tollervey, 2003
pMR202	pGFP-C-CCR4(1-229): N-terminal 229 residues of Ccr4p in frame with GFP-tag in pGFP-C-FUS	This work
pMR203	pGFP-C-POP2(1-156): N-terminal 156 residues of Pop2p in frame with GFP-tag in pGFP-C-FUS	This work
pMR204	pGFP-N-DHH1(427-506): C-terminal 80 residues of Dhh1p in frame with GFP-tag in pGFP-N-FUS	This work
pMR210	pGFP-N-DHH1: *DHH1* coding sequence in frame with GFP-tag in pGFP-N-FUS	This work
pMR211	pGFP-N-DHH1ΔC: Dhh1p aa 1-427 in frame with GFP-tag in pGFP-N-FUS	This work
pMR212	pGFP-N-CCR4: *CCR4* coding sequence in frame with GFP-tag in pGFP-N-FUS	This work
pMR213	pGFP-N-CCR4ΔN2: Ccr4p aa 250-837 in frame with GFP-tag in pGFP-N-FUS	This work

pMR214	pGFP-N-POP2: *POP2* coding sequence in frame with GFP-tag in pGFP-N-FUS	This work
pMR215	pGFP-N-POP2ΔN: Pop2p aa 147-433 in frame with GFP-tag in pGFP-N-FUS	This work
pMR216	pGFP-C-POP2: *POP2* coding sequence in frame with GFP-tag in pGFP-C-FUS	This work
pMR217	pGFP-C-POP2ΔN: Pop2p aa 147-433 in frame with GFP-tag in pGFP-C-FUS	This work
pMR218	pGFP-C-CCR4ΔN: Ccr4p aa 148-837 in frame with GFP-tag in pGFP-C-FUS	This work

Fig. S2

A



B

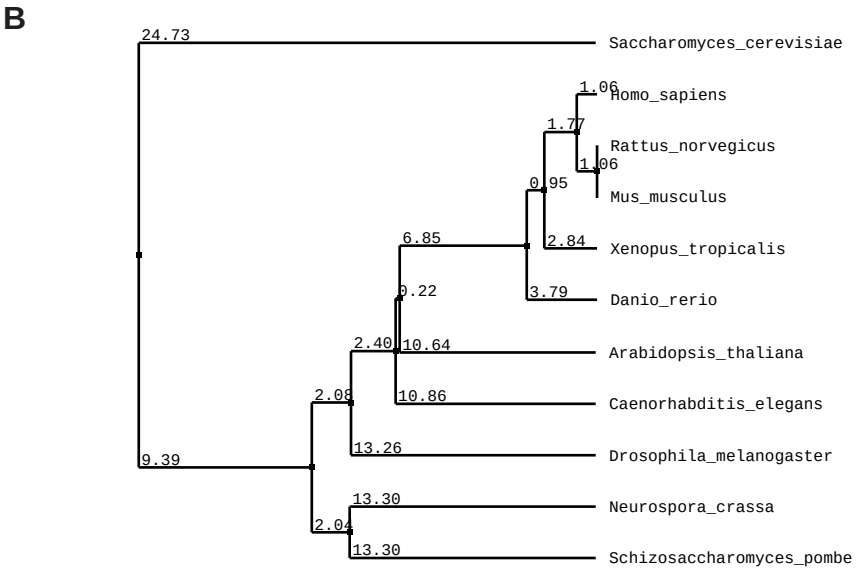


Fig. S2 The C-termini of Lsm4 proteins from most model organisms do not contain the high levels of N and/or Q residues seen in budding yeast Lsm4 proteins. Some contain GR repeats instead

(A) Alignment of Lsm4 proteins from model organisms

(B) Phylogenetic tree of Lsm4 proteins from model organisms.

Fig. S3

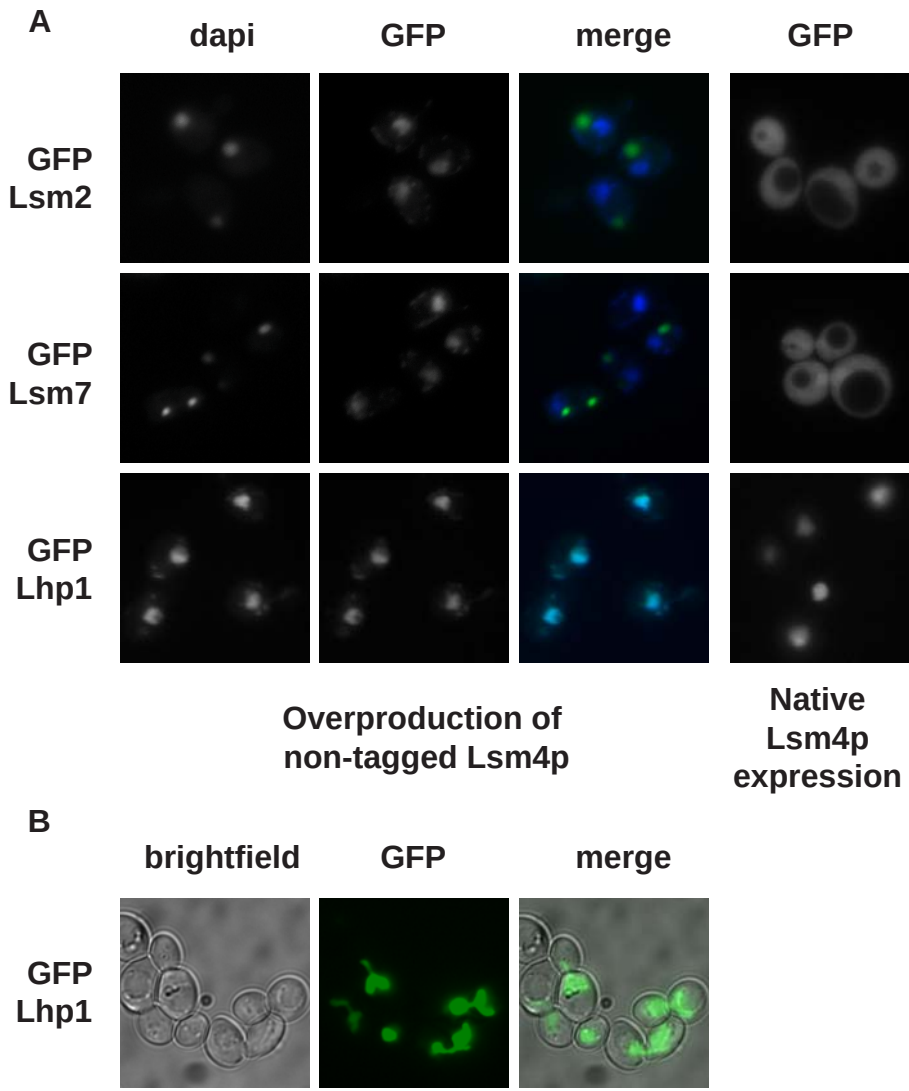


Fig. S3 Lsm proteins, but not Lhp1p aggregate when Lsm4p is overexpressed. **(A)** Localization of GFP-Lsm2 (pMPSLsm2), GFP-Lsm7 (pMPSLsm7) and GFP-Lhp1 (pMPS2) was examined in log phase cells over-producing Lsm4p (P_{GAL}-*LSM4* cells grown in SDGal-Ura) and in cells with normal levels of Lsm4p (P_{GAL}-*LSM4* cells with pUSS1 grown in SD-Ura-Met). Nuclear DNA stained with DAPI is shown in blue. **(B)** Nuclear Lhp1p localization shows abnormal nuclear morphology in many cells over-expressing Lsm4p.

Fig. S4

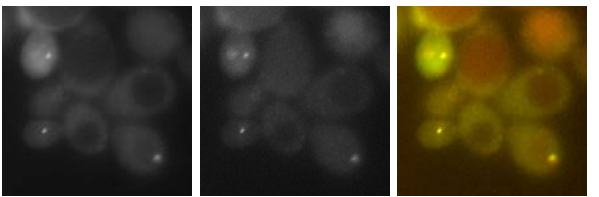


Fig. S4 Co-localization of GFP-Lsm1 (expressed from the *MET25* promoter in pGFP-N-Lsm1) and Dcp2-RFP (pRP1155) in P-bodies after hypo-osmotic shock. GFP-Lsm1 in green and Dcp2-RFP in red in merged image.

Fig. S5

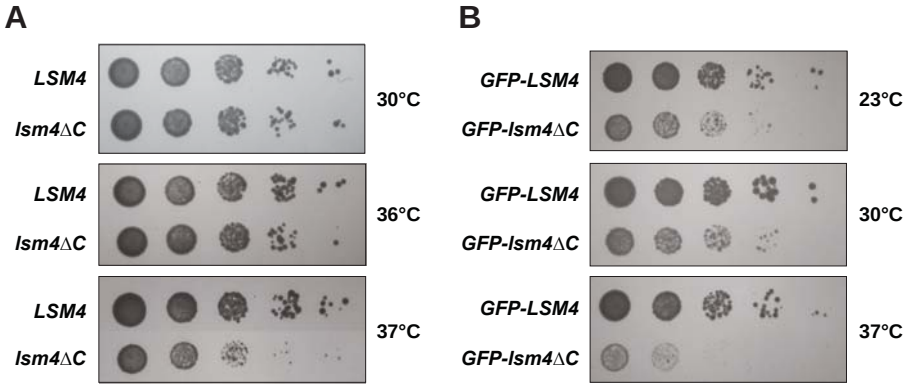
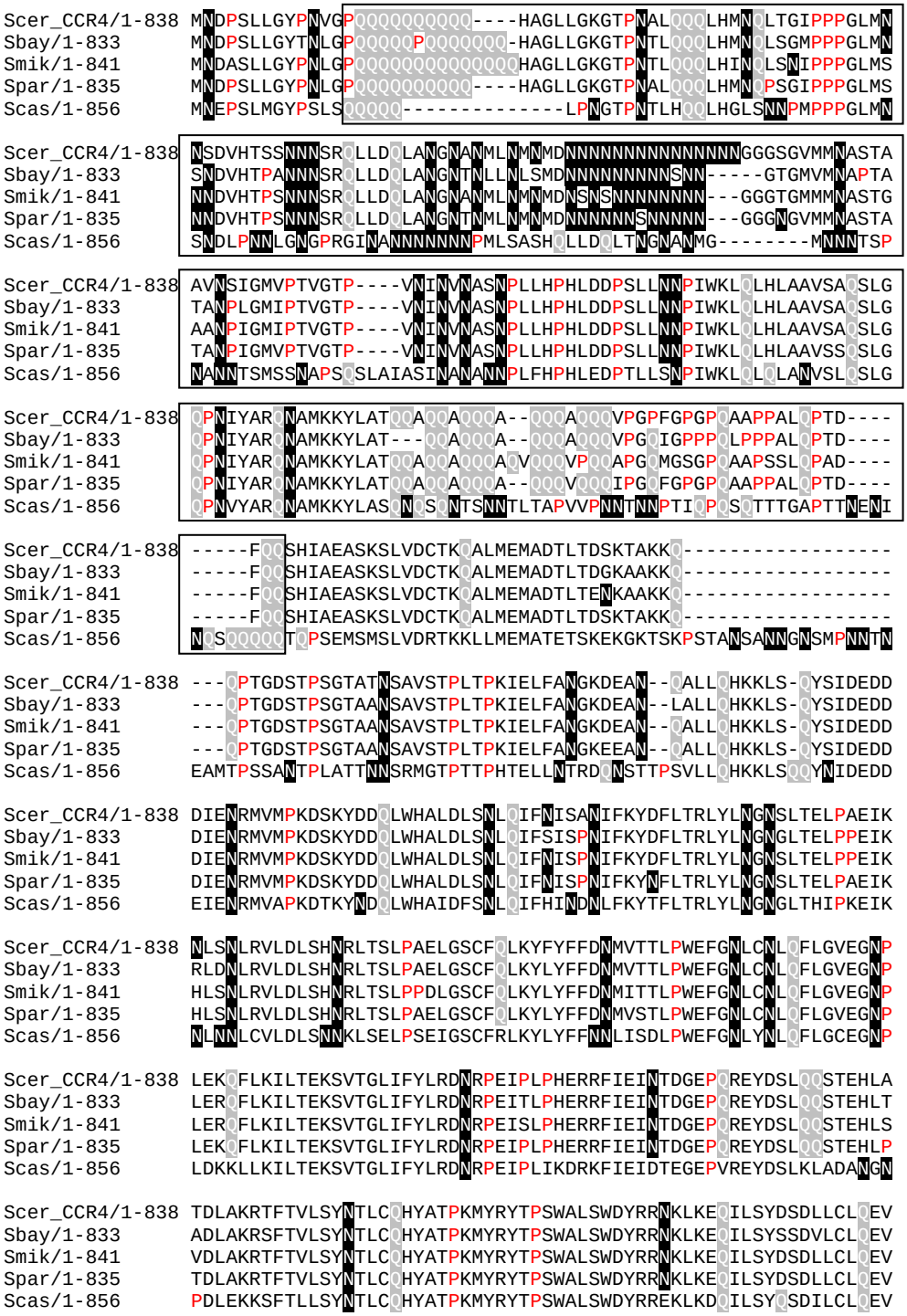


Fig. S5 An N-terminal GFP tag aggravates the slow growth phenotype of the Lsm4p C-terminal deletion strain. **(A)** MRY71 (*LSM4*) and MRY73 (*lsm4ΔC*) strains were grown overnight in YPDA at 30˚C, plated in ten-fold serial dilutions on YPDA and grown at 30°C, 36°C and 37°C. **(B)** P_{GAL}-*LSM4* strain (MCY4) with GFP-tagged Lsm4p (pMPSLsm4) or Lsm4ΔCp (pMPS Lsm4D1) was grown overnight in SD-Ura-Met at 30˚C, plated in ten-fold serial dilutions on SD-Ura-Met and grown at 23°C, 30°C and 37°C.

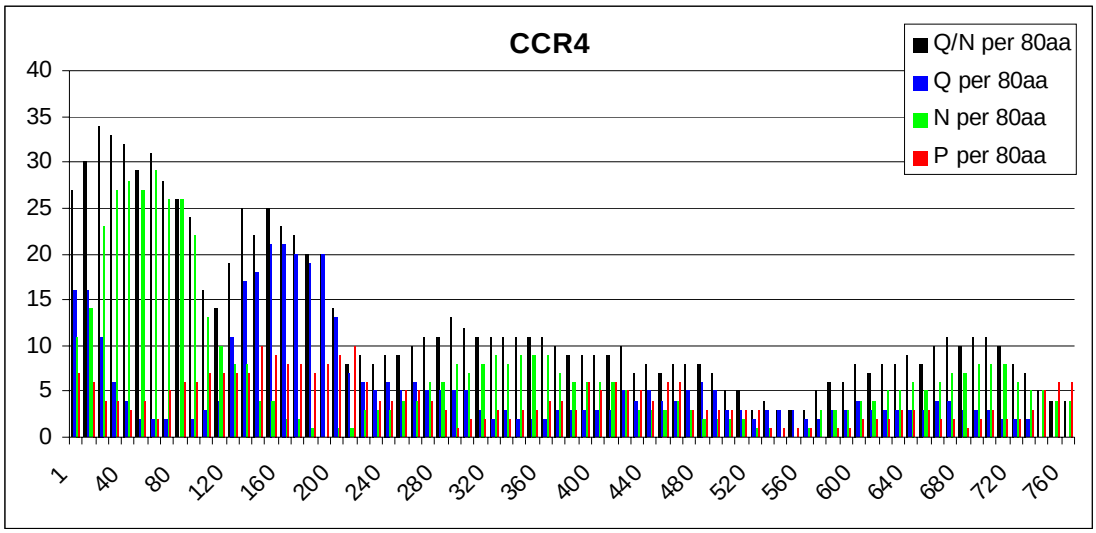
Fig. S6 Alignments of *Saccharomyces* P-body components. Ccr4p, Pop2p, Not1p, Edc3p, Dhh1p, Dcp2p, Pat1p, Scd6p, Dcp1p, Edc1p, Not4p, Pan3p, Not3p, Not5p, Xrn1p, Pan2p, Lsm1p, Not2p and Edc2p from *S. cerevisiae* and their closest homologs from other *Saccharomyces* species (see Fig. 1 legend for abbreviations) were aligned using ClustalW (Thompson et al., 1994). Visually detected Q, N and/or P-rich regions are boxed, with glutamine (Q) residues highlighted in grey, asparagine (N) residues highlighted in black and proline (P) residues in red. Q, N and P residues were counted in consecutive amino acid 80-mers (shifted 10 aa at a time) for each of the *S. cerevisiae* proteins. A visual representation of these results is shown for each protein.

Fig. S6 Alignments of Saccharomyces P-body components

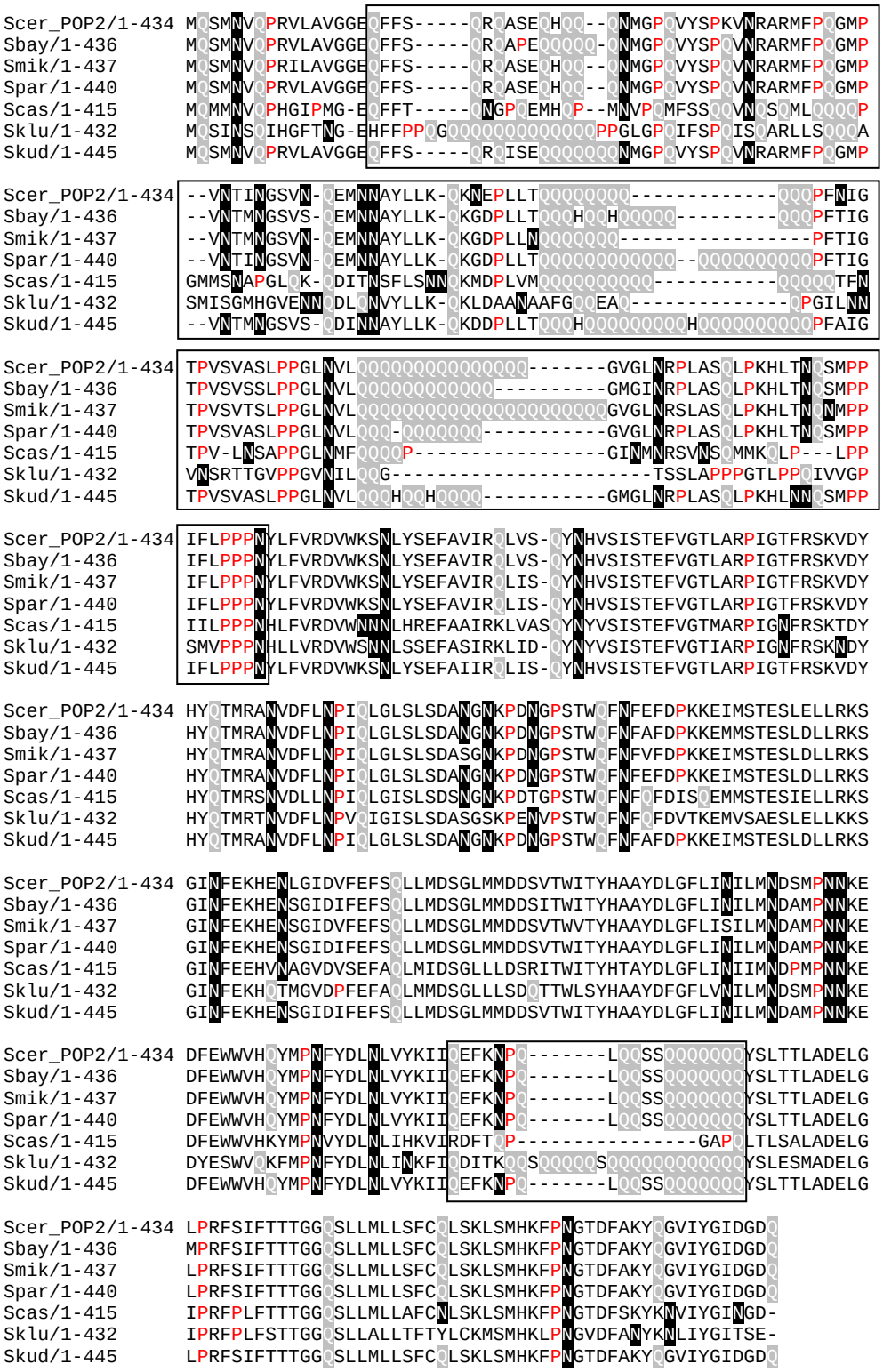
Ccr4

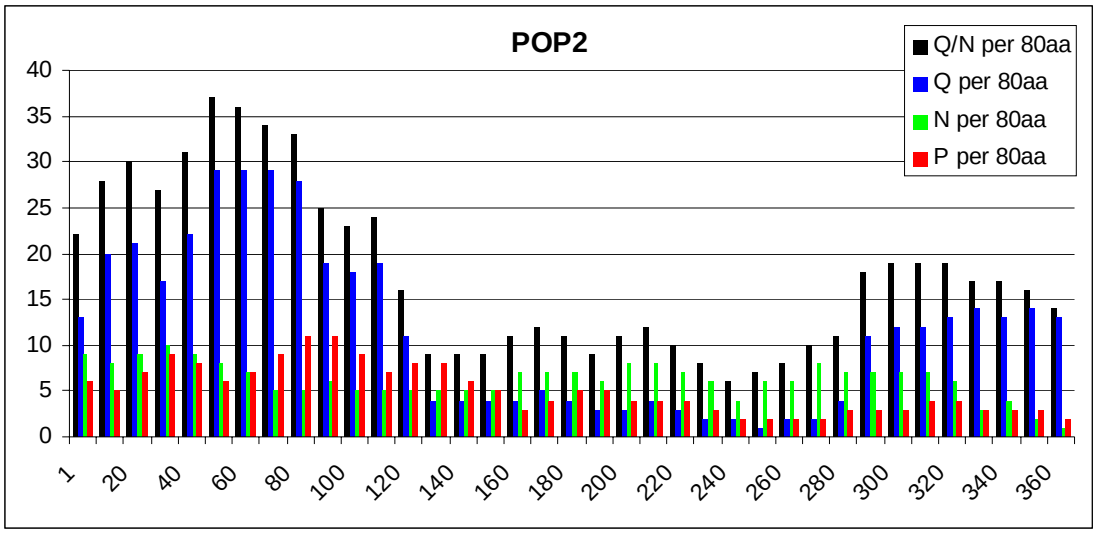


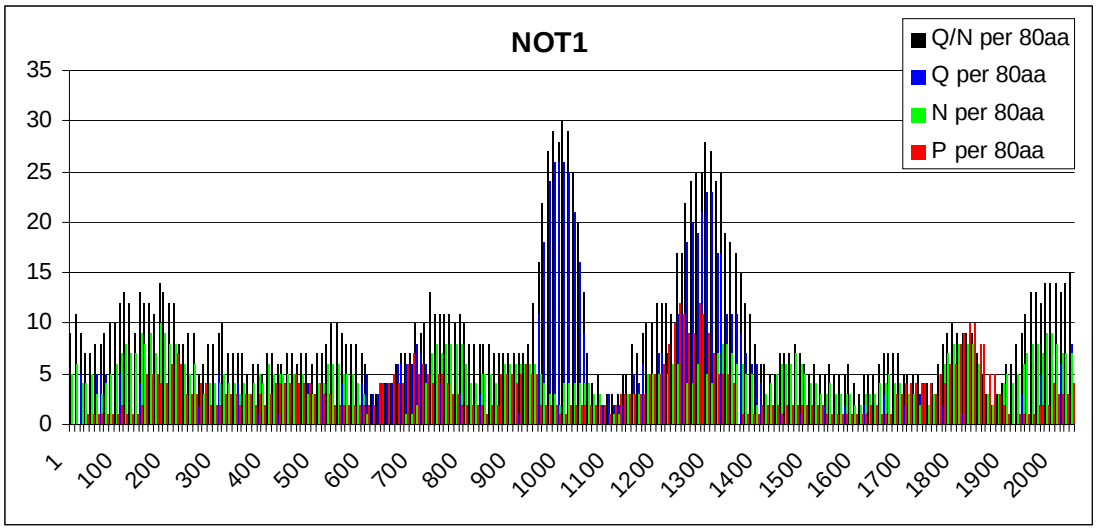
```
Scer_CCR4/1-838 ESKTFEEYWVPLLDKHGYTGIFHAKARAKTMHSKDSKKVDGCCIFFKRDQFKLITKDAMD
Sbay/1-833      ESKTFEDYWVPLLEKHGYTGVFHAKARAKTMHSKDSKKVDGCCIFFKRDQFKMVNRDSMD
Smik/1-841      ESKTFEDYWVPLLDKHGYTGIFHAKARAKTMHSKDSKKVDGCCIFFKKDQFKLITKDAMD
Spar/1-835      ESKTFEEYWVPLLDKHGYTGIFHAKARAKTMHSKDSKKVDGCCIFFKRDQFKLVTKDAMD
Sbay/1-833      ESKTFEDYWVPLLEKHGYTGVFHAKARAKTMHSKDSKKVDGCCIFFKRDQFKMVNRDSMD
Scas/1-856      ESKTFEEFWSPLLEKYDYQGIFHIKTRAKTMQSKDSKKVDGCCIFFKKSKFKLLFKEAMD

Scer_CCR4/1-838 FSGAWMKHKKFQRTEDYLNRAMNKDNVALFLKLQHIPSGDTIWAVTTHLHWDPKFNDVKT
Sbay/1-833      FSGAWMKHKKFQRTEDYLNRAMNKDNVALFLKLQHVSSGDTIWVVTTHLHWDPKFNDVKT
Smik/1-841      FSGAWMKHKKFQRTEDYLNRAMNKDNVALFLKLQHIPSGDTIWAVTTHLHWDPKFNDVKT
Spar/1-835      FSGAWMKHKKFQRTEDYLNRAMNKDNVALLLKLQHIPSGDTIWAVTTHLHWDPKFNDVKT
Sbay/1-833      FSGAWMKHKKFQRTEDYLNRAMNKDNVALFLKLQHVSSGDTIWVVTTHLHWDPKFNDVKT
Scas/1-856      FSGTWMKHKKFQRTEDYLNRAMNKDNVALYLKLQSLTSGESVWVVTTHLHWDPKFNDVKT

Scer_CCR4/1-838 FQVGVLLDHLETLLKEETSHNFRQDIKKFPVLICGDFNSYINSAVYELINTGRVQIHQEG
Sbay/1-833      FQVGVLLDHLETLLKEDTTHNSRQDIKKSPVLICGDFNSYINSAVYELISTGRVQAHQEG
Smik/1-841      FQVGVLLDHLETLLREETTHNFRQDIKKSPVLICGDFNSYINSAVYELISTGRVQINQEG
Spar/1-835      FQVGVLLDHLETLLKEETSHNFRQDIKKSPVLICGDFNSYINSAVYELISTGRVQMHQEG
Sbay/1-833      FQVGVLLDHLETLLKEDTTHNSRQDIKKSPVLICGDFNSYINSAVYELISTGRVQAHQEG
Scas/1-856      FQVGILLDHMEALLKEE---NPKQDVKKANVVICGDLNSYFDSAVYELLSTGRVVNHQDN

Scer_CCR4/1-838 NGRDFGYMSEKNFSHNLALKSSYNCIGELPFTNFTPSFTDVIDYIWFSTHALRVRGLLGE
Sbay/1-833      SSRDFGYMSEKNFSHNLALKSSYNCIGELPFTNFTPSFTDVIDYIWFSTHALRVRGLLGE
Smik/1-841      NGRDFGYMSEKNFSHNLALKSSYNCIGELPFTNFTPSFTDVIDYIWFSTHALRVRGLLGE
Spar/1-835      NGRDFGYMSERNFSHNLALKSSYNCIGELPFTNFTPSFTDVIDYIWFSTHALRVRGLLGE
Sbay/1-833      SSRDFGYMSEKNFSHNLALKSSYNCIGELPFTNFTPSFTDVIDYIWFSTHALRVRGLLGE
Scas/1-856      KGRDFGYMSQKNFAHNLSLRSSYDYIGELPFTNFTPSFTDVIDYIWFSTQSMRVRGLLGG

Scer_CCR4/1-838 VDPEYVSKFIGFPNDKFPSDHIPLLARFEFMKTNTGSKKV
Sbay/1-833      VDPEYVDKFIGFPNDKFPSDHIPLLARFEFMKTNTGSRKV
Smik/1-841      VDPDYVSKFIGFPNDKFPSDHIPLLARFEFMKTNTGSRKV
Spar/1-835      VDPEYVSKFIGFPNDKFPSDHIPLLARFEFMKTNTGSRKV
Sbay/1-833      VDPEYVDKFIGFPNDKFPSDHIPLLARFEFMKTNTGSRKV
Scas/1-856      VDQDYVSNFIGFPNDKFPSDHIPLLARFEFMKSSSGSRKI
```



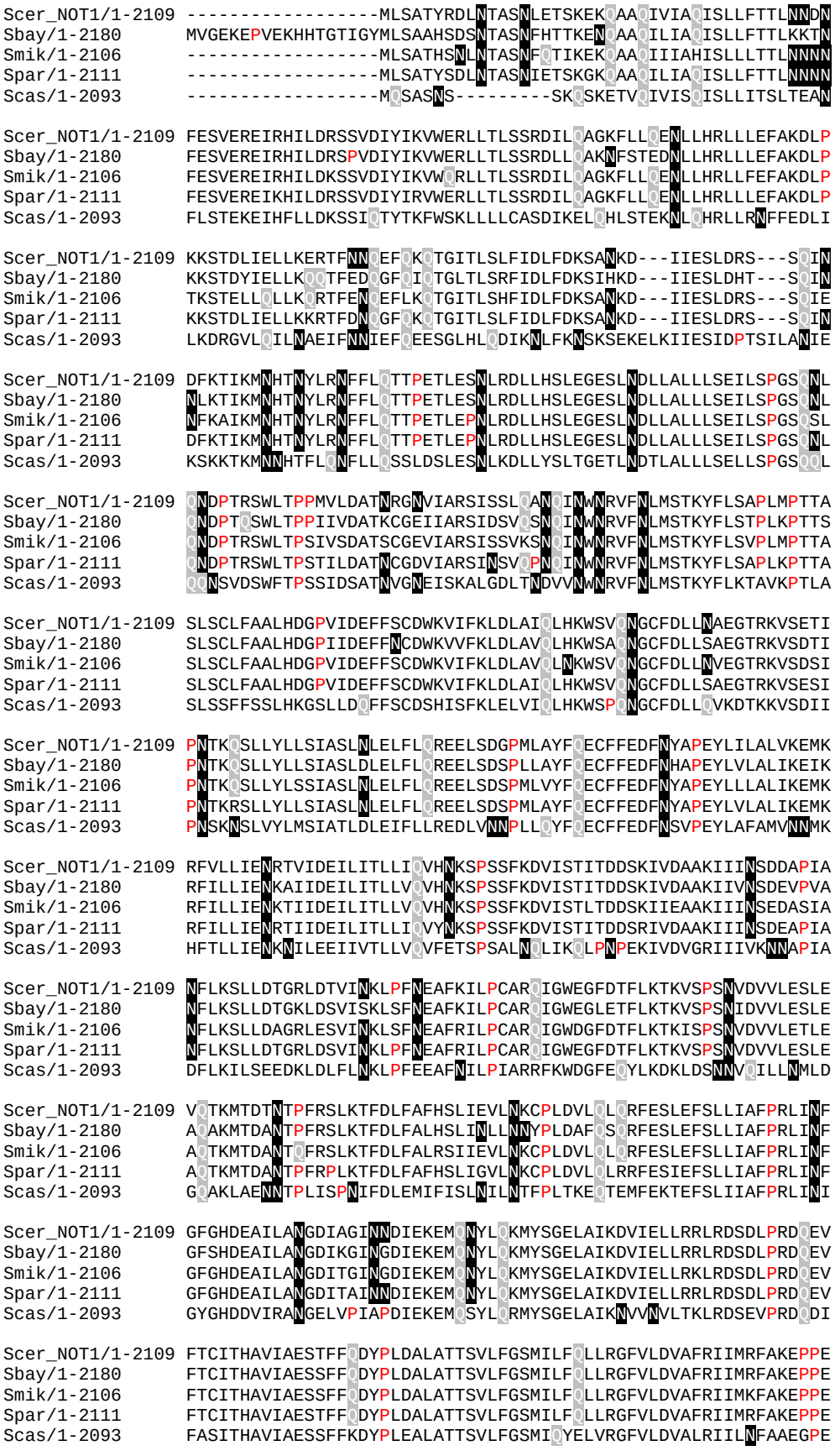
Pop2

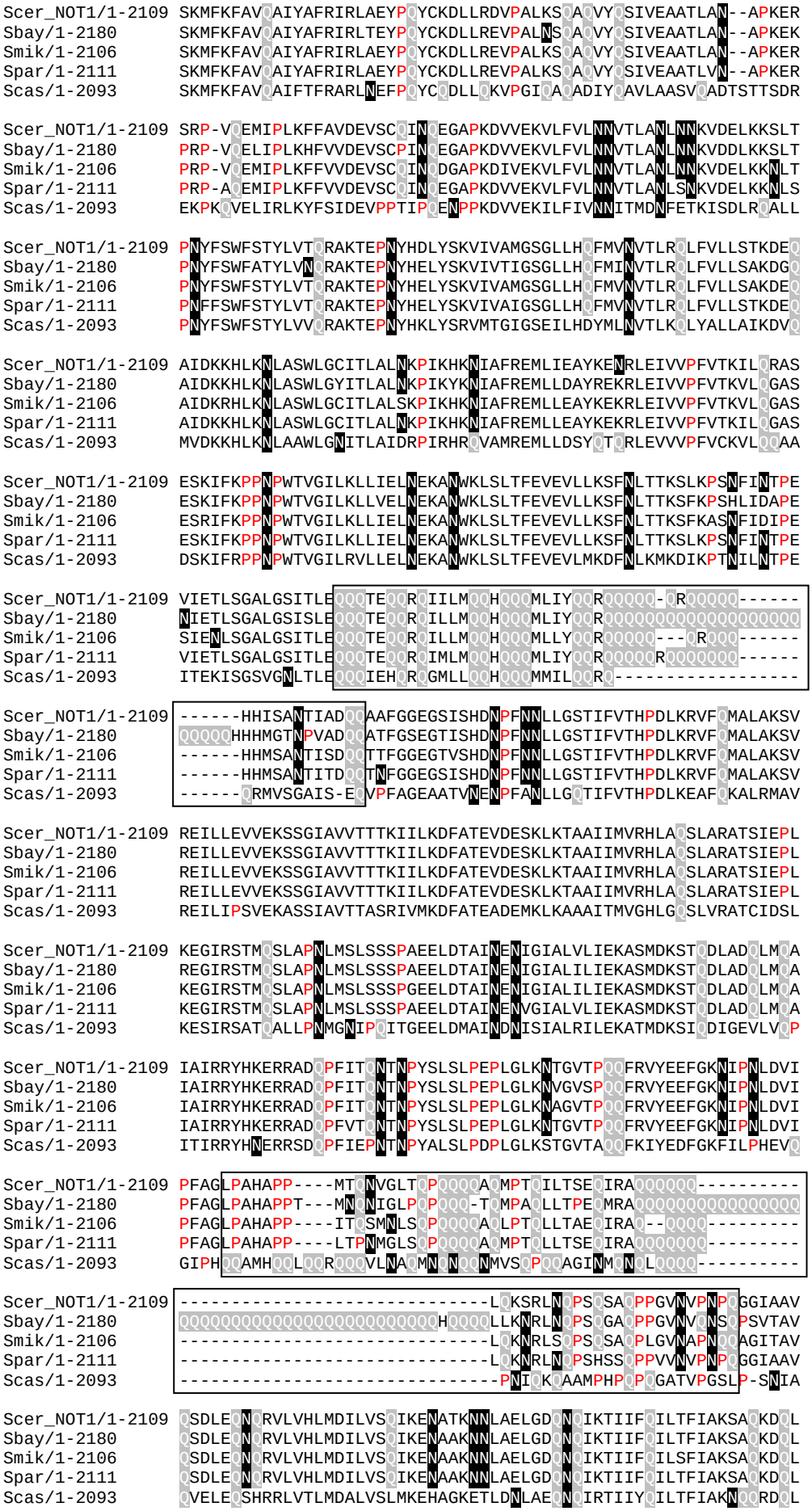

Scer_POP2/1-434
Sbay/1-436
Smik/1-437
Spar/1-440
Scas/1-415
Sklu/1-432
Skud/1-445


POP2
Q/N per 80aa
Q per 80aa
N per 80aa
P per 80aa
40
35
30
25
20
15
10
5
0
1
20
40
60
80
100
120
140
160
180
200
220
240
260
280
300
320
340
360


NOT1
Q/N per 80aa
Q per 80aa
N per 80aa
P per 80aa
35
30
25
20
15
10
5
0
1
100
200
300
400
500
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000

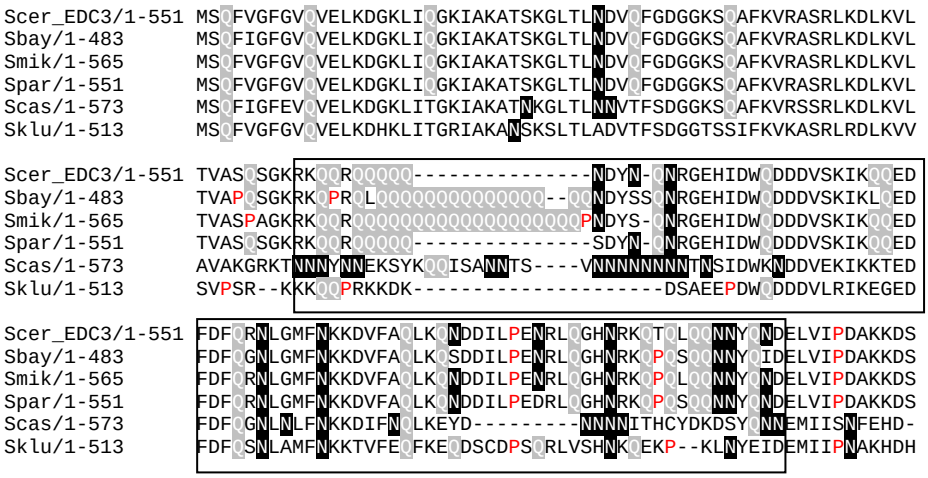
Not1

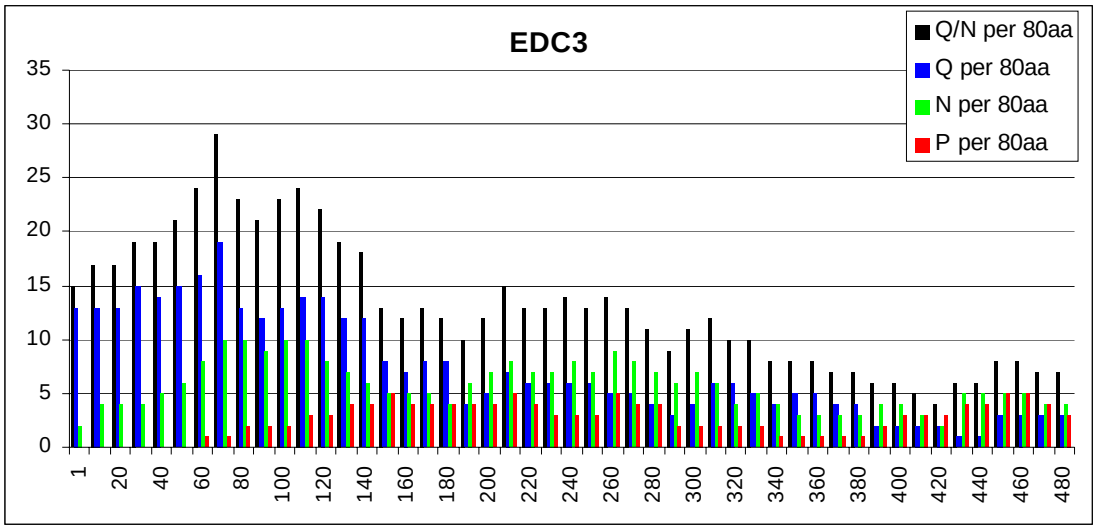

Scer_NOT1/1-2109 ----------MLSATYRDLNTASNLETSKEKQAAQIVIAQISLLFTTLNNDN
Sbay/1-2180 MVGEKEPVEKHHTGTIGYMLSAAHSDSNTASNFHTTKENQAAQILIAQISLLFTTLKKTN
Smik/1-2106 ----------MLSATHSNLNTASNFQTIKEKQAAQIIIAHISLLLTTLNNNN
Spar/1-2111 ----------MLSATYSDLNTASNIETSKGKQAAQILIAQISLLFTTLNNNN
Scas/1-2093 ----------MQSASNS---------SKQSKETVQIVISQISLLITSLTEAN
Scer_NOT1/1-2109 FESVEREIRHILDRSSVDIYIKVWERLLTLSSRDILQAGKFLLQENLLHRLLLEFAKDLP
Sbay/1-2180 FESVEREIRHILDRSPVDIYIKVWERLLTLSSRDLLQAKNFSTEDNLLHRLLLEFAKDLP
Smik/1-2106 FESVEREIRHILDKSSVDIYIKVWQRLLTLSSRDILQAGKFLLQENLLHRLLFEFAKDLP
Spar/1-2111 FESVEREIKHILDRSSVDIYIRVWERLLTLSSRDILQAGKFLLQENLLHRLLLEFAKDLP
Scas/1-2093 FLSTEKEIHFLLDKSSIQTYTKFWSKLLLLCASDIKELQHLSTEKNLQHRLLRNFFEDLI
Scer_NOT1/1-2109 KKSTDLIELLKERTFNNQEFQKQTGITLSLFIDLFDKSANKD---IIESLDRS---SQIN
Sbay/1-2180 KKSTDYIELLKQQTFEDQGFQIQTGLTLSRFIDLFDKSIHKD---IIESLDHT---SQIN
Smik/1-2106 TKSTELLQLLKQRTFENQEFLKQTGITLSHFIDLFDKSANKD---IIESLDRS---SQIE
Spar/1-2111 KKSTDLIELLKKRTFDNQGFQKQTGITLSLFIDLFDKSANKD---IIESLDRS---SQIN
Scas/1-2093 LKDRGVLQILNAEIFNNIEFQEESGLHLQDIKNLFKNSKSEKELKIIESIDPTSILANIE
Scer_NOT1/1-2109 DFKTIKMNHTNYLRNFFLQTTPETLESNLRDLLHSLEGESLNDLLALLLSEILSPGSQNL
Sbay/1-2180 NLKTIKMNHTNYLRNFFLQTTPETLESNLRDLLHSLEGESLNDLLALLLSEILSPGSQNL
Smik/1-2106 NFKAIKMNHTNYLRNFFLQTTPETLEPNLRDLLHSLEGESLNDLLALLLSEILSPGSQSL
Spar/1-2111 DFKTIKMNHTNYLRNFFLQTTPETLEPNLRDLLHSLEGESLNDLLALLLSEILSPGSQNL
Scas/1-2093 KSKKTKMNNHTFLQNFLLQSSLDSLESNLKDLLYSLTGETLNDTLALLLSELLSPGSQQL
Scer_NOT1/1-2109 QNDPTRSWLTPPMVLDATNRGNVIARSISSLQANQINWNRVFNLMSTKYFLSAPLMPTTA
Sbay/1-2180 QNDPTQSWLTPPIIVDATKCGEIIARSIDSVQSNQINWNRVFNLMSTKYFLSTPLKPTTS
Smik/1-2106 QNDPTRSWLTPSIVSDATSCGEVIARSISSVKSNQINWNRVFNLMSTKYFLSVPLMPTTA
Spar/1-2111 QNDPTRSWLTPSTILDATNCGDVIARSINSVQPNQINWNRVFNLMSTKYFLSAPLKPTTA
Scas/1-2093 QQNSVDSWFTPSSIDSATNVGNEISKALGDLTNDVVNWNRVFNLMSTKYFLKTAVKPTLA
Scer_NOT1/1-2109 SLSCLFAALHDGPVIDEFFSCDWKVIFKLDLAIQLHKWSVQNGCFDLLNAEGTRKVSETI
Sbay/1-2180 SLSCLFAALHDGPIIDEFFNCDWKVVFKLDLAVQLHKWSAQNGCFDLLSAEGTRKVSDTI
Smik/1-2106 SLSCLFAALHDGPVIDEFFSCDWKVIFKLDLAVQLNKWSVQNGCFDLLNVEGTRKVSDSI
Spar/1-2111 SLSCLFAALHDGPVIDEFFSCDWKVIFKLDLAIQLHKWSVQNGCFDLLSAEGTRKVSESI
Scas/1-2093 SLSSFFSSLHKGSLLDQFFSCDSHISFKLELVIQLHKWSPQNGCFDLLQVKDTKKVSDII
Scer_NOT1/1-2109 PNTKQSLLYLLSIASLNLELFLQREELSDGPMLAYFQECFFEDFNYAPEYLILALVKEMK
Sbay/1-2180 PNTKQSLLYLLSIASLDLELFLQREELSDSPLLAYFQECFFEDFNHAPEYLVLALIKEIK
Smik/1-2106 PNTKQSLLYLSSIASLNLELFLQREELSDSPMLVYFQECFFEDFNYAPEYLLLALIKEMK
Spar/1-2111 PNTKRSLLYLLSIASLNLELFLQREELSDSPMLAYFQECFFEDFNYAPEYLVLALIKEMK
Scas/1-2093 PNSKNSLVYLMSIATLDLEIFLLREDLVNNPLLQYFQECFFEDFNSVPEYLAFAMVNNMK
Scer_NOT1/1-2109 RFVLLIENRTVIDEILITLLIQVHNKSPSSFKDVISTITDDSKIVDAAKIIINSDDAPIA
Sbay/1-2180 RFILLIENKAIIDEILITLLVQVHNKSPSSFKDVISTITDDSKIVDAAKIIVNSDEVPVA
Smik/1-2106 RFILLIENKTIIDEILITLLVQVHNKSPSSFKDVISTLTDDSKIIEAAKIIINSEDASIA
Spar/1-2111 RFILLIENRTIIDEILITLLIQVYNKSPSSFKDVISTITDDSRIVDAAKIIINSDEAPIA
Scas/1-2093 HFTLLIENKNILEEIIVTLLVQVFETSPSALNQLIKQLPNPEKIVDVGRIIIVKNNAPIA
Scer_NOT1/1-2109 NFLKSLLDTGRLDTVINKLPFNEAFKILPCARQIGWEGFDTFLKTKVSPSNVDVVLESLE
Sbay/1-2180 NFLKSLLDTGKLDSVISKLSFNEAFKILPCARQIGWEGLETFLKTKVSPSNIDVVLESLE
Smik/1-2106 NFLKSLLDAGRLESVINKLSFNEAFRILPCARQIGWDGFDTFLKTKISPSNVDVVLETLE
Spar/1-2111 NFLKSLLDTGRLDSVINKLPFNEAFRILPCARQIGWEGFDTFLKTKVSPSNVDVVLESLE
Scas/1-2093 DFLKILSEEDKLDLFLNKLPFEEAFNILPIARRFKWDGFEQYLKDKLDSNNVQILLNMLD
Scer_NOT1/1-2109 VQTKMTDTNTPFRSLKTFDLFAFHSLIEVLNKCPLDVLQLQRFESLEFSLLIAFPRLINF
Sbay/1-2180 AQAKMTDANTPFRSLKTFDLFALHSLINLLNNYPLDAFQSQRFESLEFSLLIAFPRLINF
Smik/1-2106 AQTKMTDANTQFRSLKTFDLFALRSIIEVLNKCPLDVLQLQRFESLEFSLLIAFPRLINF
Spar/1-2111 AQTKMTDANTPFRPLKTFDLFAFHSLIGVLNKCPLDVLQLRRFESIEFSLLIAFPRLINF
Scas/1-2093 GQAKLAENNTPLISPNIFDLEMIFISLNILNTFPLTKEQTEMFEKTEFSLIIAFPRLINI
Scer_NOT1/1-2109 GFGHDEAILANGDIAGINNDIEKEMQNYLQKMYSGELAIKDVIELLRRLRDSDLPRDQEV
Sbay/1-2180 GFSHDEAILANGDIKGINGDIEKEMQNYLQKMYSGELAIKDVIELLRRLRDSDLPRDQEV
Smik/1-2106 GFGHDEAILANGDITGINGDIEKEMQNYLQKMYSGELAIKDVIELLRKLRDSDLPRDQEV
Spar/1-2111 GFGHDEAILANGDITAINNDIEKEMQNYLQKMYSGELAIKDVIELLRRLRDSDLPRDQEV
Scas/1-2093 GYGHDDVIRANGELVPIAPDIEKEMQSYLQRMYSGELAIKNVVNVLTKLRDSEVPRDQDI
Scer_NOT1/1-2109 FTCITHAVIAESTFFQDYPLDALATTSVLFGSMILFQLLRGFVLDVAFRIIMRFAKEPPE
Sbay/1-2180 FTCITHAVIAESSFFQDYPLDALATTSVLFGSMILFQLLRGFVLDVAFRIIMRFAKEPPE
Smik/1-2106 FTCITHAVIAESSFFQDYPLDALATTSVLFGSMILFQLLRGFVLDVAFRIIMKFAKEPPE
Spar/1-2111 FTCITHAVIAESTFFQDYPLDALATTSVLFGSMILFQLLRGFVLDVAFRIIMRFAKEPPE
Scas/1-2093 FASITHAVIAESSFFKDYPLEALATTSVLFGSMIQYELVRGFVLDVALRIILNFAAEGPE



```
Scer_NOT1/1-2109 ALKVSQAVVNSLFATSESPLCREVLSLLLEKLCSLSLVARKDVVWWLVYALDSRKFNVPV
Sbay/1-2180      ALKVSQAVVNSLFATSESSLCREVLSLLLEKLCSLSLVARKDVVWWLVYALDSRKFNVPV
Smik/1-2106      ALKVSQAVVNSLFATSESPLCREVLSLLLEKLCSLSLVARKDVVWWLVYALDSRKFNVPV
Spar/1-2111      ALKVSQAVVNSLFATSESPLCREVLSLLLEKLCSLSLVARKDVVWWLVYALDSRKFNVPV
Scas/1-2093      ALKVSQAVVNSLFGASDDVLCREVLSTLLEKLCSLSLVARKDVIWWLVYALDSRKFNVPV

Scer_NOT1/1-2109 IRSLLEVNLIDATELDNVLVTAMKNKMENSTEFAMKLIQNTVLSDDPILMRMDFIKTLEH
Sbay/1-2180      IRSLLEVNLIDATELDNVLVTAMKNSMENSTDFAMQLIQDSVLSDDPILMRMDFIKTLEH
Smik/1-2106      IRSLLEVNLIDATELDNVLVTAMKNRMENSTEFAMKLIENSVLSDDSILMRMDFIKTLEH
Spar/1-2111      IRSLLEVNLIDATELDNVLVTAMKNKMENSTEFAMKLIQNAVLSDDPILMRMDFIKTLEH
Scas/1-2093      IKSLLSVNLIDVSELDTVLVTAMENKMENATKFAIDLIKDTVLSDEPILMRMDFVKSLEF

Scer_NOT1/1-2109 LASSEDENVKKFIKEFEDTKIMPVRKGTKTTRTEKLYLVFTEWVKLLQRVENNDVITTVF
Sbay/1-2180      LGSLEDESVKKFIQDFEDTKIMPVKKGTKTTRTEKFYLVFTEWVKLLQRVECNDVITAVF
Smik/1-2106      LTSLEDENVKKFMKEFEDTKIMPVKKGTKTSRTEKFYLVFTEWVKLLQRVEKNDVITTVF
Spar/1-2111      LASLEDENVKKFIKEFEDTKIMPVRKGTKTTRTEKFYLVFTEWVKLLQRVENNDIITAVF
Scas/1-2093      LSSLDEEDVKNFFSEYESMKILPTSKNIETTSTERYYLVFTEWVRLLQRVTSDDKIIFVF

Scer_NOT1/1-2109 IKQLVEKGVISDTDNLLTFVKSSLELSVSSFKESDPTDEVFIAIDALGSLIIKLLILQGF
Sbay/1-2180      IKQLVEKGVISDTDNLITFVKSSLELSVSSFKESDPTDEVFIAIDALGKLIVKLLVLQDF
Smik/1-2106      IKQLVEKGVISDTDNLITFVKSSLELSVSSFKESDPTDEVFIAIDALGSLIVKLLILQDF
Spar/1-2111      IKQLVEKGVISDTDNLITFVKSSLELSVSSFKESDPTDEVFIAIDALGSLIVKLLILQDF
Scas/1-2093      IKQLMDKGVLSDSDNFIGFVKAALELSVYSFKESDPTGEVFTAIDALSKLLIKLFILQDF

Scer_NOT1/1-2109 KDDTRRDYINAIFSVIVLVFAKDHSQEGTTFNERPYFRLFSNILYEWATIRTHNFVRISD
Sbay/1-2180      KDGTRKDYINTIFSVILLVFAKDHSQKDTTFNERPYFRLFSNILYEWATIRTHNFVKISD
Smik/1-2106      KDTTRRDYINAIFSVIVLVFAKDHSQEGTTFNERPYFRLFSNILYEWATIRTHNFVRVSD
Spar/1-2111      KDDTRRDYINAIFSVIVLVFAKDHSQEGTTFNERPYFRLFSNILYEWATIRTHNFVRISD
Scas/1-2093      AGYSRQEYLNTVFSIILLVFSNDHEEDEATFNERPYFRLLSNFLCEWATLRGHNFIKVAD

Scer_NOT1/1-2109 SSTRQELIEFDSVFYNTFSGYLHALQPFAFPGFSFAWVTLLSHRMLLPIMLRLPNKIGWE
Sbay/1-2180      ARTREELIDFDSVFYNTFSGYLHSLQPFAFPGFSFAWVTLLSHRMLLPVMLRLPKKMGWE
Smik/1-2106      SVTRQELIKFDPVFYNTFSEYLHSLQPFAFPGFSFAWVTLLSHRMLLPVMLRLPKKMGWE
Spar/1-2111      SKTRQELIEFDSVFYNTFSGYLHALQPFAFPGFSFAWVTLLSHRMLLPIMLRLPNKIGWE
Scas/1-2093      QKTRKELLSFDAEFYNIFASYLHSFQPFAFPGFSFAWISLLSHRMFLPVMLRLPQKAGWE

Scer_NOT1/1-2109 KLMLLIIDLFKFLDQYTSKHAVSDAVSVVYKGTLRVILGISNDMPSFLIENHYELMNNLP
Sbay/1-2180      KLMLLIIDLFKFLDQYTSKHAVSDAISVVYKGTLRVILGISNDVPSFLIENHYELMNNLP
Smik/1-2106      KLMLLIIDLFNFLDQYTSKHSVSDAVSVVYKGTLRVILGISNDVPSFLIENHYELMNNLP
Spar/1-2111      KLMLLIIDLFKFLDQYTSKHAVSDAVSVVYKGTLRVILGISNDVPSFLIENHYELMNNLP
Scas/1-2093      KLMLLLIDLLKFLNQYTIKGKISDAVSVVYKGTLRIFLGISNDVPQFLIENHYELMNNLP

Scer_NOT1/1-2109 PTYFQLKNVILSAIPKNMTVPNPYDVDLNMEDIPACKELPEVFFDPVIDLHSLKKPVDNY
Sbay/1-2180      PTYFQLKNVILSAIPKHMTVPNPYDVDLSMDNIPSCKDLPEVFFDPVVDLHSLKKPVDNY
Smik/1-2106      PTYFQLKNVILSAIPKHMTVPNPYDVDISMENIPSCKELPEVFFDPVIDLHSLKKPVDNY
Spar/1-2111      PTYFQLKNVILSAIPKHMTVPNPYDVTLNMENIPSCKELPEVFFDPVIDLHSLKKPVDNY
Scas/1-2093      ISYFQLKNVILSAIPLKMLVPNPFDSDLALENITECQNPPVVFYDPVSDIQALKKPVDNY

Scer_NOT1/1-2109 LRIPSNSLLRTILSAIYKDTYDIKKGVGYDFLSVDSKLIRAIVLHVGIEAGIEYKRTSSN
Sbay/1-2180      LRIPSNSLLKTILNSIYRDTYDMKKGVGYDFLSVDSKLIRAIVLHVGIEAGIEYERTSSN
Smik/1-2106      LRIPSNSLLRTILNSVYKDTYDIKKGVGYEFLSVDSKLIRAIVLHVGIEAGIEYKRTSSN
Spar/1-2111      LRIPSNSLLRTILNSIYKDTYDIKKGVGYDFLSIDSKLVRAIVLHVGIEAGIEYKRTSSN
Scas/1-2093      LRIPSNSLLRTIINGLYLTEYDIKGGVGFDMLTTNNKLIRAIVLHVAVEAGLENGRTSSN

Scer_NOT1/1-2109 AVFNTKSSYYTLLFNLIQNGSIEMKYQIILSIVEQLRYPNIHTYWFSFVLMNMFKSDEWN
Sbay/1-2180      AVFNTKSSYYTLLFNLIQNGTIELKYQILLVIVEQLRYPNIHTYWFSFVLMNMFKSDEWD
Smik/1-2106      AVFNTKSSYYTLLFNLIQNGSIEIKHQIILAIVEQLRYPNIHTYWFSFVLMNMFKSEEWS
Spar/1-2111      AVFNTKSSYYTLLFNLIQNGNIEMKYQIILAIVEQLRYPNIHTYWFSFVLMNMFKSDEWN
Scas/1-2093      AVFNTKSSYYQLLFDLIHDGTIELKFQVIQVMIEQLRYPNIHTRWFIYVLRDMFVTEAWE

Scer_NOT1/1-2109 DQKLEVQEIILRNFLKRIIVNKPHTWGVSVFFTQLINNNDINLLDLPFVQSVPEIKLILQ
Sbay/1-2180      DQKLEVQEIILRHFLKRIIVNKPHTWGISVFFTQLINNDDLNLLDLPFVQNVPEIELILQ
Smik/1-2106      DQKLEVQELILRNFLKRIIVNKPHTWGVSVFFTQLINSDDINLLDLPFVQNVPEIKLILQ
Spar/1-2111      DQKLEVQEIILRNFLKRIIVNKPHTWGVSVFFTQLINSSDIHLLDLPFVQSVPEIKLILQ
Scas/1-2093      EQRTEVQEIILRSLLERVIVHNPHTWGVSVLFTQLLNSDEVNLLELDFINNIPEIKHMFV

Scer_NOT1/1-2109 QLVKYSKKYTTSEQDDQSATINRRQTPLQSNA
Sbay/1-2180      QLVKYSKKYTNCEQDNESLTLDGKQTPLQSNA
Smik/1-2106      QLVKYSKKYTNHEQDDKSATIDRGQTPLQSNA
Spar/1-2111      QLVKYSKKYTNCEQDDKSSTVDRRQTPLQSNA
Scas/1-2093      QLTKHTNKLTDKSPETNTASP----KPIPN--
```

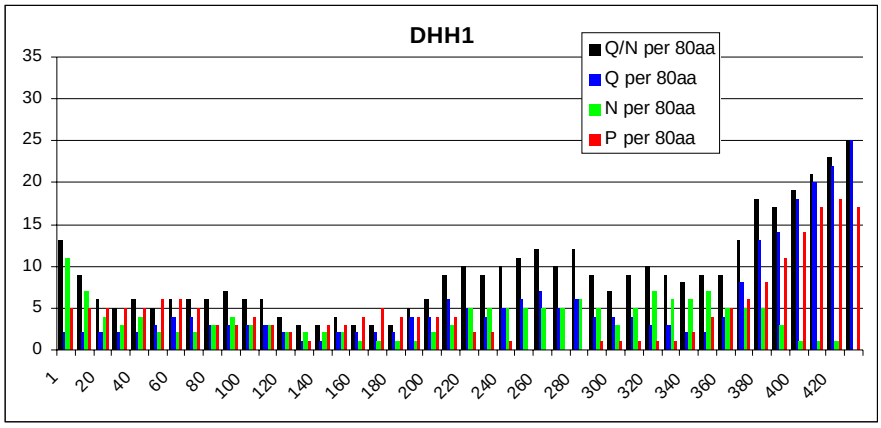
Edc3



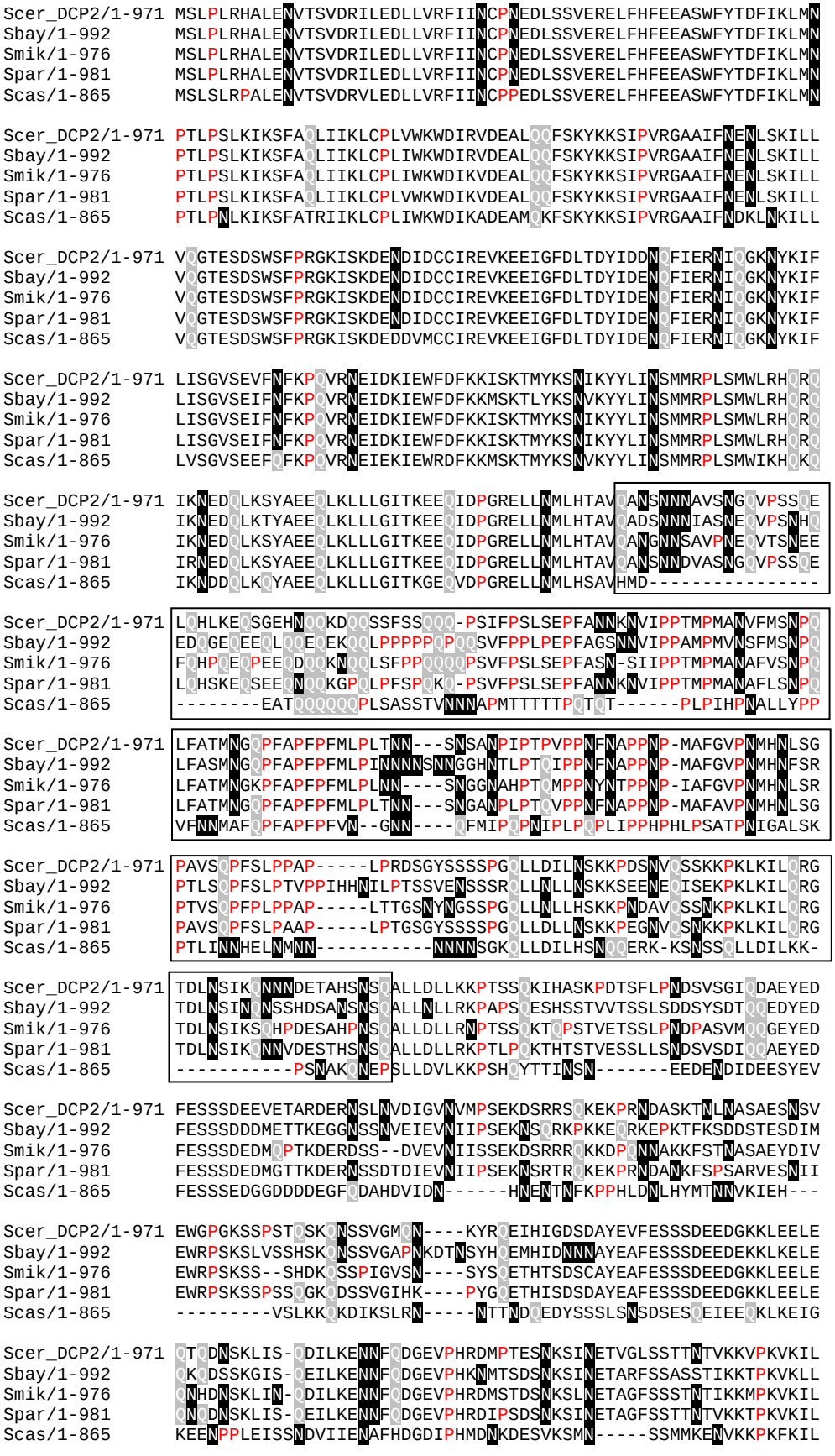

EDC3
Q/N per 80aa
Q per 80aa
N per 80aa
P per 80aa
35
30
25
20
15
10
5
0
1
20
40
60
80
100
120
140
160
180
200
220
240
260
280
300
320
340
360
380
400
420
440
460
480

Dhh1

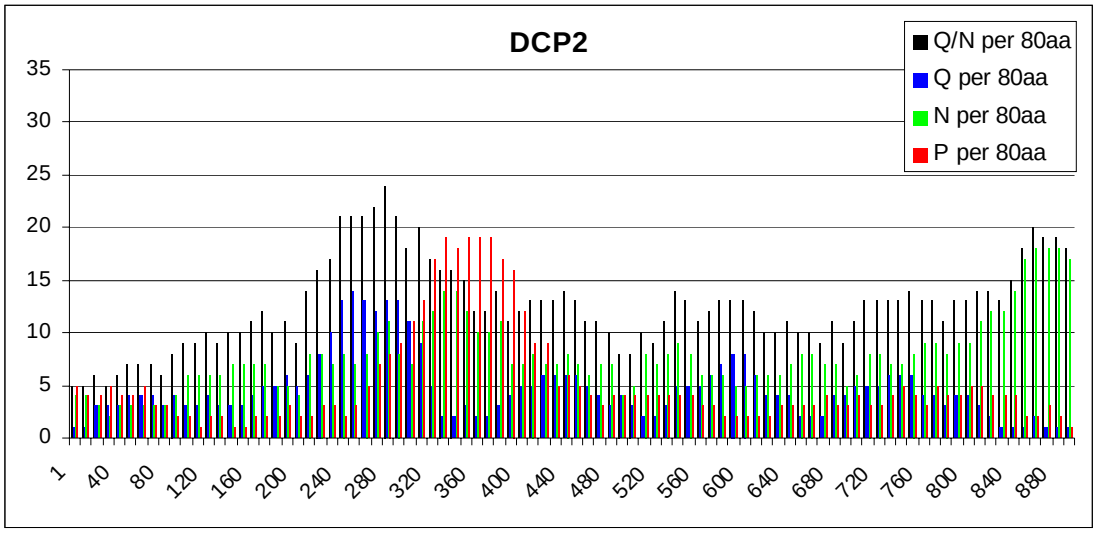
```
Scer_DHH1/1-507 MGSINNNFNTNNNSNTDLDRDWKTALNIPKKDTRPQTDDVLNTKGNTFEDFYLKRELLMG
Sbay/1-509      MSSTNNNFNTNNNSNTDLDLDWKTALNIPKKDTRPQTDDVLNTKGNTFEDFYLKRELLMG
Smik/1-507      MGSINNNFSTNSNNNTDLDLDWKTALSIPKKDTRPQTDDVLNTKGNTFEDFYLRRELLMG
Spar/1-507      MGSINNNFSTNNNSNTDLDRDWKTALNIPKKDTRPQTDDVLNTKGNTFEDFYLKRELLMG
Scas/1-511      ---MSSSTTTAAAATQQQGQDWKSTLNIPKKDTRPQTDDVLNTKGNTFEDFYLKRELLMG
Skud/1-508      MSSINKNFITNNNSNTDLDLDWKTALNIPKKDTRPQTDDVLNTKGNTFEDFYLKRELLMG

Scer_DHH1/1-507 IFEAGFEKPSPIQEEAIPVAITGRDILARAKNGTGKTAAFVIPTLEKVKPKLNKIQALIM
Sbay/1-509      IFEAGFEKPSPIQEEAIPVAITGRDILARAKNGTGKTAAFVIPTLEKVKPKLNKIQALIM
Smik/1-507      IFEAGFEKPSPIQEEAIPVAITGRDILARAKNGTGKTAAFVIPTLEKVKPKLNKIQALIM
Spar/1-507      IFEAGFEKPSPIQEEAIPVAITGRDILARAKNGTGKTAAFVIPTLEKVKPKLNKIQALIM
Scas/1-511      IFEAGFEKPSPIQEESIPIAITGRDILARAKNGTGKTAAFVIPTLEKIKPKLNKIQALIM
Skud/1-508      IFEAGFEKPSPIQEEAIPVAITGRDILARAKNGTGKTAAFVIPTLERVKPKLNKIQALIM

Scer_DHH1/1-507 VPTRELALQTSQVVRTLGKHCGISCMVTTGGTNLRDDILRLNETVHILVGTPGRVLDLAS
Sbay/1-509      VPTRELALQTSQVVRTLGKHCGISCMVTTGGTNLRDDILRLNETVHVLVGTPGRVLDLAS
Smik/1-507      VPTRELALQTSQVVRTLGKHCGISCMVTTGGTNLRDDILRLNETVHILVGTPGRVLDLAS
Spar/1-507      VPTRELALQTSQVVRTLGKHCGISCMVTTGGTNLRDDILRLNETVHILVGTPGRVLDLAS
Scas/1-511      VPTRELALQTSQVVRTLGKHCGISCMVTTGGTNLRDDILRLNESVHVLVGTPGRVLDLAS
Skud/1-508      VPTRELALQTSQVVRTLGKHCGISCMVTTGGTNLRDDILRLNETVHILVGTPGRVLDLAS

Scer_DHH1/1-507 RKVADLSDCSLFIMDEADKMLSRDFKTIIEQILSFLPPTHQSLLFSATFPLTVKEFMVKH
Sbay/1-509      RKVADLSDCSLFIMDEADKMLSRDFKTIIEQILSFLPPTHQSLLFSATFPLTVKEFMVKH
Smik/1-507      RKVADLSDCSLFIMDEADKMLSRDFKTIIEQILSFLPSSHQSLLFSATFPLTVKEFMVKH
Spar/1-507      RKVADLSDCSLFIMDEADKMLSRDFKTIIEQILSFLPPTHQSLLFSATFPLTVKEFMVKH
Scas/1-511      RKVADLSDCQLFIMDEADKMLSRDFKTIIEQILIFLPKKHQSLLFSATFPLTVKEFMVKH
Skud/1-508      RKVADLSDCSLFIMDEADKMLSRDFKTIIEQILSFLPTTHQSLLFSATFPLTVKEFMVKH

Scer_DHH1/1-507 LHKPYEINLMEELTLKGITQYYAFVEERQKLHCLNTLFSKLQINQAIIFCNSTNRVELLA
Sbay/1-509      LHKPYEINLMEELTLKGITQYYAFVEERQKLHCLNTLFSKLQINQAIIFCNSTNRVELLA
Smik/1-507      LHKPYEINLMEELTLKGITQYYAFVEERQKLHCLNTLFSKLQINQAIIFCNSTNRVELLA
Spar/1-507      LHKPYEINLMEELTLKGITQYYAFVEERQKLHCLNTLFSKLQINQAIIFCNSTNRVELLA
Scas/1-511      LTNPYEINLMDELTLKGITQYYAFVEERQKLHCLNTLFSKLQINQAIIFCNSTNRVELLA
Skud/1-508      LHKPYEINLMEELTLKGITQYYAFVEERQKLHCLNTLFSKLQINQAIIFCNSTNRVELLA

Scer_DHH1/1-507 KKITDLGYSCYYSHARMKQQERNKVFHEFRQGKVRTLVCSDLLTRGIDIQAVNVVINFDF
Sbay/1-509      KKITDLGYSCYYSHARMKQQERNKVFHEFRQGKVRTLVCSDLLTRGIDIQAVNVVINFDF
Smik/1-507      KKITDLGYSCYYSHARMKQQERNKVFHEFRQGKVRTLVCSDLLTRGIDIQAVNVVINFDF
Spar/1-507      KKITDLGYSCYYSHARMKQQERNKVFHEFRQGKVRTLVCSDLLTRGIDIQAVNVVINFDF
Scas/1-511      KKITDLGYSCYYSHARMKQQERNRVFHEFRQGKVRTLVCSDLLTRGIDIQAVNVVINFDF
Skud/1-508      KKITDLGYSCYYSHARMKQQERNKVFHEFRQGKVRTLVCSDLLTRGIDIQAVNVVINFDF

Scer_DHH1/1-507 PKTAETYLHRIGRSGRFGHLGLAINLINWNDRFNLYKIEQELGTEIAAIPATIDKSLYVA
Sbay/1-509      PKTAETYLHRIGRSGRFGHLGLAINLINWNDRFNLYKIEQELGTEIAAIPATIDKSLYVA
Smik/1-507      PKTAETYLHRIGRSGRFGHLGLAINLINWNDRFNLYKIEQELGTEIAAIPATIDKSLYVA
Spar/1-507      PKTAETYLHRIGRSGRFGHLGLAINLINWNDRFNLYKIEQELGTEIAAIPATIDKSLYVA
Scas/1-511      PKTSETYLHRIGRSGRFGHLGLAINLINWNDRFNLYKIEQELGTEIAAIPATIDKSLYVA
Skud/1-508      PKTAETYLHRIGRSGRFGHLGLAINLINWNDRFNLYKIEQELGTEIAAIPATIDKSLYVA

Scer_DHH1/1-507 ENDETVPVPFPIEQQSYHQQAIPQQQ----LPSQQQ---FAIPPQQHHPQFMVP-PS-HQ
Sbay/1-509      EDDETVPVPFPIEQQSYHQQATPQQQQ---LPPQQQ---FAIPPQQHHPQFMVP-PPQHH
Smik/1-507      ENDETVPVPFPIEQQSYHQQAVPQQQ----LPPQQQ---FALPPQQHHPQFVVP-PP-HQ
Spar/1-507      ENDETVPVPFPIEQQSYHQQAAPQQQ----LPSQQQ---FAIPPQQHHPQFMVP-PP-HQ
Scas/1-511      ENEANVPIPFPIEQQHFSNNNGAPMDPNFVPPPQQQNPQFAPQPLQNQQQFIPPQHPQYS
Skud/1-508      ENDETVPVPFPIEQRSYQPQAASQQQ----LPPQQQ---FAIPPQQHHPQFMVP-PPQHQ

Scer_DHH1/1-507 QQQAYPPPQMPSQQGYPPQQEHFMAMPPGQSQPQY
Sbay/1-509      QQQAYPPPQMPQQQGYPPQQEHFMAMPPGQPQPQF
Smik/1-507      QQQAYPPPQMPTQQGYPPQQEHFMAMPPGQSQPQY
Spar/1-507      QQQAYPPPQMPSQQGYPPQQEHFMAMPPGQSQPQY
Scas/1-511      QQQQQQQMPQIPQQQ-YAQQQQQQQQYNNNVQ-QPQF
Skud/1-508      QQQVYPPSQIPPQQGYPPQQEHFMAMPPGQPQPQF
```



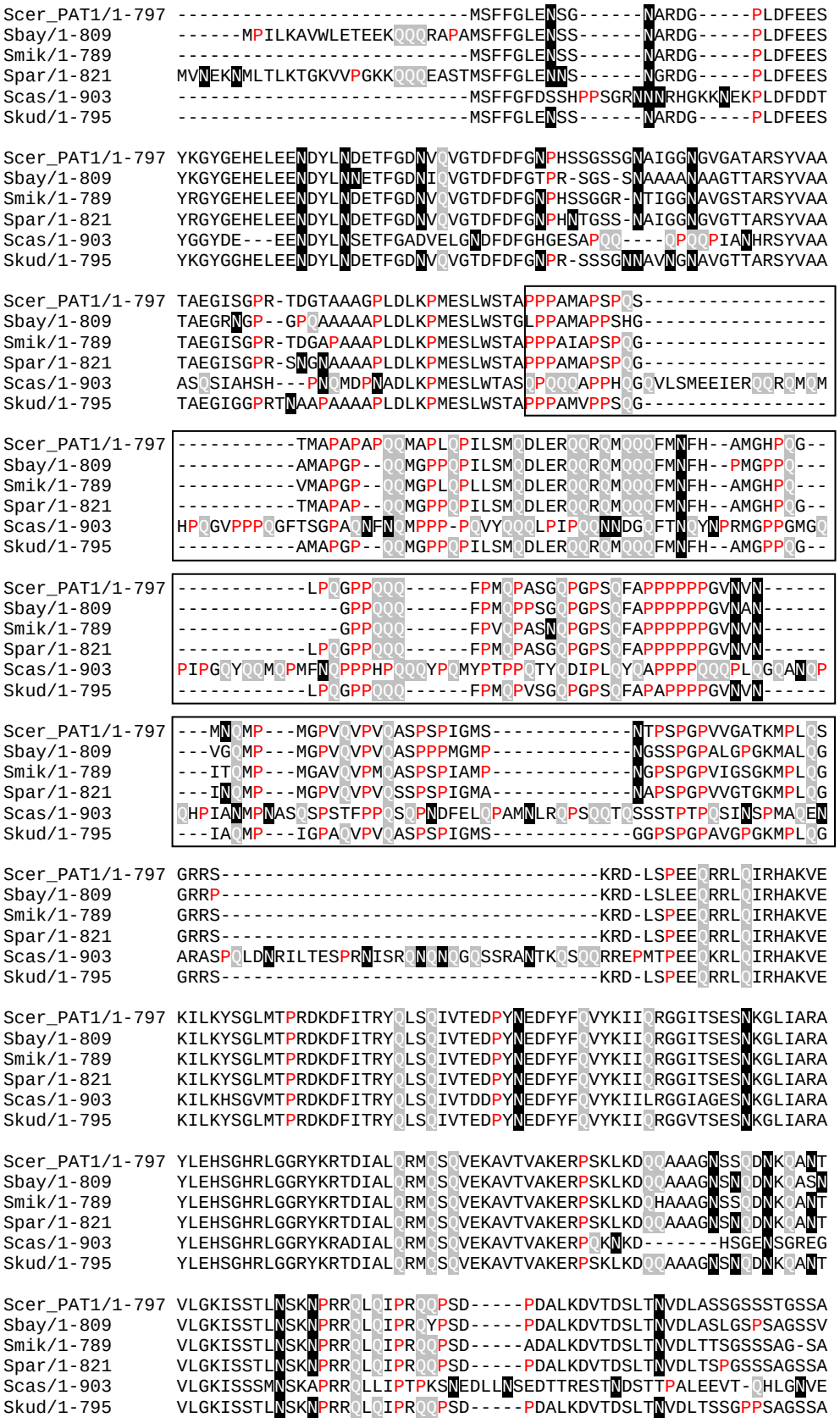
Dcp2


Scer_DCP2/1-971 MSLPLRHALENVTSVDRILEDLLVRFIINCPNEDLSSVERELFHFEEASWFYTDFIKLMN
Sbay/1-992 MSLPLRHALENVTSVDRILEDLLVRFIINCPNEDLSSVERELFHFEEASWFYTDFIKLMN
Smik/1-976 MSLPLRHALENVTSVDRILEDLLVRFIINCPNEDLSSVERELFHFEEASWFYTDFIKLMN
Spar/1-981 MSLPLRHALENVTSVDRILEDLLVRFIINCPNEDLSSVERELFHFEEASWFYTDFIKLMN
Scas/1-865 MSLSLRPALENVTSVDRVLEDLLVRFIINCPPEDLSSVERELFHFEEASWFYTDFIKLMN
Scer_DCP2/1-971 PTLPSLKIKSFAQLIIKLCPLVWKWDIRVDEALQQFSKYKKSIPVRGAAIFNENLSKILL
Sbay/1-992 PTLPSLKIKSFAQLIIKLCPLIWKWDIRVDEALQQFSKYKKSIPVRGAAIFNENLSKILL
Smik/1-976 PTLPSLKIKSFAQLIIKLCPLIWKWDIKVDEALQQFSKYKKSIPVRGAAIFNENLSKILL
Spar/1-981 PTLPSLKIKSFAQLIIKLCPLVWKWDIKVDEALQQFSKYKKSIPVRGAAIFNENLSKILL
Scas/1-865 PTLPNLKIKSFATRIIKLCPLIWKWDIKADEAMQKFSKYKKSIPVRGAAIFNDKLNKILL
Scer_DCP2/1-971 VQGTESDSWSFPRGKISKDENDIDCCIREVKEEIGFDLTDYIDDNQFIERNIQGKNYKIF
Sbay/1-992 VQGTESDSWSFPRGKISKDENDIDCCIREVKEEIGFDLTDYIDENQFIERNIQGKNYKIF
Smik/1-976 VQGTESDSWSFPRGKISKDENDIDCCIREVKEEIGFDLTDYIDENQFIERNIQGKNYKIF
Spar/1-981 VQGTESDSWSFPRGKISKDENDIDCCIREVKEEIGFDLTDYIDENQFIERNIQGKNYKIF
Scas/1-865 VQGTESDSWSFPRGKISKDEDDVMCCIREVKEEIGFDLTDYIDENQFIERNIQGKNYKIF
Scer_DCP2/1-971 LISGVSEVFNFKPQVRNEIDKIEWFDFKKISKTMYKSNIKYYLINSMMRPLSMWLRHQRQ
Sbay/1-992 LISGVSEIFNFKPQVRNEIDKIEWFDFKKMSKTLYKSNVKYYLINSMMRPLSMWLRHQRQ
Smik/1-976 LISGVSEIFNFKPQVRNEIDKIEWFDFKKISKTMYKSNIKYYLINSMMRPLSMWLRHQRQ
Spar/1-981 LISGVSEIFNFKPQVRNEIDKIEWFDFKKISKTMYKSNIKYYLINSMMRPLSMWLRHQRQ
Scas/1-865 LVSGVSEEFQFKPQVRNEIEKIEWRDFKKMSKTMYKSNVKYYLINSMMRPLSMWIKHQKQ
Scer_DCP2/1-971 IKNEDQLKSYAEEQLKLLLGITKEEQIDPGRELLNMLHTAVQANSNNNAVSNGQVPSSQE
Sbay/1-992 IKNEDQLKTYAEEQLKLLLGITKEEQIDPGRELLNMLHTAVQADSNNNIASNEQVPSNHQ
Smik/1-976 IKNEDQLKSYAEEQLKLLLGITKEEQIDPGRELLNMLHTAVQANGNNSAVPNEQVTSNEE
Spar/1-981 IRNEDQLKSYAEEQLKLLLGITKEEQIDPGRELLNMLHTAVQANSNNDVASNGQVPSSQE
Scas/1-865 IKNDDQLKQYAEEQLKLLLGITKGEQVDPGRELLNMLHSAVHMD----------------
Scer_DCP2/1-971 LQHLKEQSGEHNQQKDQQSSFSSQQQ-PSIFPSLSEPFANNKNVIPPTMPMANVFMSNPQ
Sbay/1-992 EDQGEQEEQLQQEQEKQQLPPPPPQPQQSVFPPLPEPFAGSNNVIPPAMPMVNSFMSNPQ
Smik/1-976 FQHPQEQPEEQDQQKNQQLSFPPQQQQQPSVFPSLSEPFASN-SIIPPTMPMANAFVSNPQ
Spar/1-981 LQHSKEQSEEQNQQKGPQLPFSPQKQ-PSVFPSLSEPFANNKNVIPPTMPMANAFLSNPQ
Scas/1-865 --------EATQQQQQQQPLSASSTVNNNAPMTTTTTPQTQT------PLPIHPNALLYPP
Scer_DCP2/1-971 LFATMNGQPFAPFPFMLPLTNN---SNSANPIPTPVPPNFNAPPNP-MAFGVPNMHNLSG
Sbay/1-992 LFASMNGQPFAPFPFMLPINNNNSNNGGHNTLPTQIPPNFNAPPNP-MAFGVPNMHNFSR
Smik/1-976 LFATMNGKPFAPFPFMLPLNN----SNGGNAHPTQMPPNYNTPPNP-IAFGVPNMHNLSR
Spar/1-981 LFATMNGQPFAPFPFMLPLTNN---SNGANPLPTQVPPNFNAPPNP-MAFAVPNMHNLSG
Scas/1-865 VFNNMAFQPFAPFPFVN--GNN----QFMIPQPNIPLPQPLIPPHPHLPSATPNIGALSK
Scer_DCP2/1-971 PAVSQPFSLPPAP-----LPRDSGYSSSSPGQLLDILNSKKPDSNVQSSKKPKLKILQRG
Sbay/1-992 PTLSQPFSLPTVPPIHHNILPTSSVENSSSRQLLNLLNSKKSEENEQISEKPKLKILQRG
Smik/1-976 PTVSQPFPLPPAP-----LTTGSNYNGSSPGQLLNLLHSKKPNDAVQSSNKPKLKILQRG
Spar/1-981 PAVSQPFSLPAAP-----LPTGSGYSSSSPGQLLDLLNSKKPEGNVQSNKKPKLKILQRG
Scas/1-865 PTLINNHELNMNN----------NNNNSGKQLLDILHSNQQERK-KSNSSQLLDILKK-
Scer_DCP2/1-971 TDLNSIKQNNNDETAHSNSQALLDLLKKPTSSQKIHASKPDTSFLPNDSVSGIQDAEYED
Sbay/1-992 TDLNSINQNSSHDSANSNSQALLNLLRKPAPSQESHSSTVVTSSLSDDSYSDTQQEDYED
Smik/1-976 TDLNSIKSQHPDESAHPNSQALLDLLRNPTSSQKTQPSTVETSSLPNDPASVMQQGEYED
Spar/1-981 TDLNSIKQNNVDESTHSNSQALLDLLRKPTLPQKTHTSTVESSLLSNDSVSDIQQAEYED
Scas/1-865 ----------PSNAKQNEPSLLDVLKKPSHQYTTINSN-------EEDENDIDEESYEV
Scer_DCP2/1-971 FESSSDEEVETARDERNSLNVDIGVNVMPSEKDSRRSQKEKPRNDASKTNLNASAESNSV
Sbay/1-992 FESSSDDDMETTKEGGNSSNVEIEVNIIPSEKNSQRKPKKEQRKEPKTFKSDDSTESDIM
Smik/1-976 FESSSDEDMQPTKDERDSS--DVEVNIISSEKDSRRRQKKDPQNNAKKFSTNASAEYDIV
Spar/1-981 FESSSDEDMGTTKDERNSSDTDIEVNIIPSEKNSRTRQKEKPRNDANKFSPSARVESNII
Scas/1-865 FESSSEDGGDDDDEGFQDAHDVIDN------HNENTNFKPPHLDNLHYMTNNVKIEH---
Scer_DCP2/1-971 EWGPGKSSPSTQSKQNSSVGMQN----KYRQEIHIGDSDAYEVFESSSDEEDGKKLEELE
Sbay/1-992 EWRPSKSLVSSHSKQNSSVGAPNKDTNSYHQEMHIDNNNAYEAFESSSDEEDEKKLKELE
Smik/1-976 EWRPSKSS--SHDKQSSPIGVSN----SYSQETHTSDSCAYEAFESSSDEEDGKKLEELE
Spar/1-981 EWRPSKSSPSSQGKQDSSVGIHK----PYGQETHISDSDAYEAFESSSDEEDGKKLEELE
Scas/1-865 ---------VSLKKQKDIKSLRN-----NTTNDQEDYSSSLSNSDSESQEIEEQKLKEIG
Scer_DCP2/1-971 QTQDNSKLIS-QDILKENNFQDGEVPHRDMPTESNKSINETVGLSSTTNTVKKVPKVKIL
Sbay/1-992 QKQDSSKGIS-QEILKENNFQDGEVPHKNMTSDSNKSINETARFSSASSTIKKTPKVKLL
Smik/1-976 QNHDNSKLIN-QDILKENNFQDGEVPHRDMSTDSNKSLNETAGFSSSTNTIKKMPKVKIL
Spar/1-981 QNQDNSKLIS-QEILKENNFQDGEVPHRDIPSDSNKSINETAGFSSTTNTVKKTPKVKIL
Scas/1-865 KEENPPLEISSSNDVIIENAFHDGDIPHMDNKDESVKSMN-----SSMMKENVKKPKFKIL

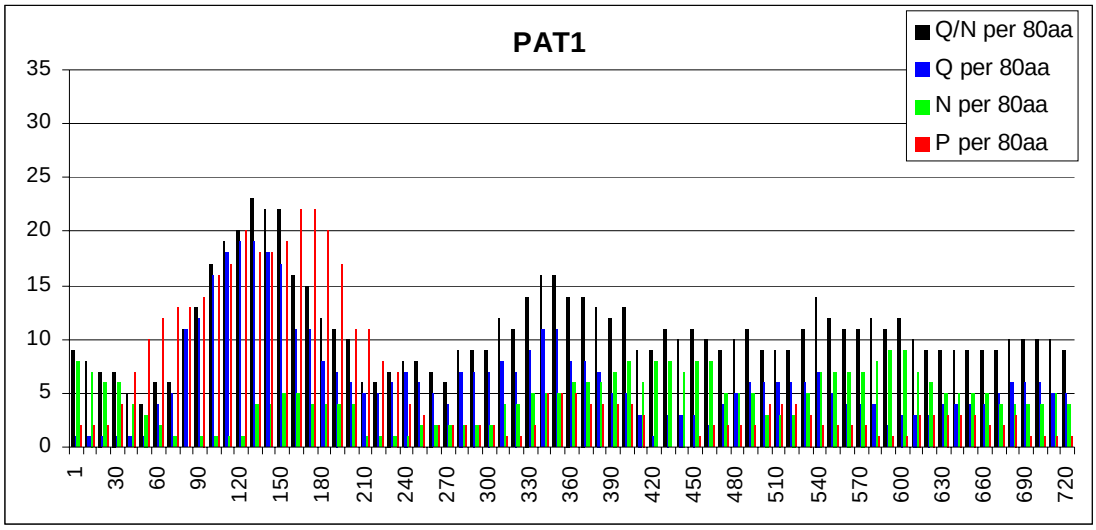
```
Scer_DCP2/1-971 KRGETFASLANDKKA---FDSSSNVSSSKDLLQMLRNPISSTVSSNQQSPKSQHLSGDEE
Sbay/1-992      KRGETIATLANENKT---LGSNSNQSSSKDLLQMLRNPISSTVSSNQQSPKSQHLDGDEE
Smik/1-976      KRGETFASLANDRKT---FDSSSNMSSSKDLLQMLRNPISSTVSSNQQSPKSQHPNGEDE
Spar/1-981      KRGETFASLANDKKN---FDSSSNISSSKDLLQMLRNPVSSTVSSNQQSPKSQHLSGDEE
Scas/1-865      KRGENISDIKLDNNEPTIAEEPTTNTDSKTLLNMLKTPK------EPEAPSSVLLN----

Scer_DCP2/1-971 IMMMLKRNSVSKPQNSEENASTSSIN-DANASELLGMLKQKEKDITAPKQP-YNVDSYSQ
Sbay/1-992      IMMMLKRNSVSGHQNIEESPSKLPKKNGANASELLGMLSQK--DTASPRQP-GNIDVNME
Smik/1-976      IMMMLKRNSVSKAENIEETASTSSNKNDGNASELLGILRPKGKDVTSSEHPGIKDVSIEE
Spar/1-981      IMMMLKRNSVSKPQNIEDNPSAFTNKSDANAGELLGMLKQKDKDIVSSKQP-YNVDSYME
Scas/1-865      ---ILKKPTEEQNDNAKDS-----------NILLNMLKKP-----STSPP--VANEEHT

Scer_DCP2/1-971 KNSAKGLLNILKKNDSTGYPRTEGGPSSEMSTSMKRNDATNNQELDKNSTELLNYLK-PK
Sbay/1-992      GNPAKGLLSMLEKNNVPKPPPTKDTPSNELFNLLKQNEITENQESNNNSSELLGFLK-PN
Smik/1-976      NNPAKGLLSILKNKDFERKPHTEDKQSNEFQ-ALKHEVVTENEKFNNTPPELLNFFN-PN
Spar/1-981      KNPAKGLLNILKKNDSTGQPRTEDRLSNELSTSMKYNGVTENQESNKSSTGLPNFLK-PK
Scas/1-865      QNSSTVLLNMLKKQSPE--SNTDAKGSEELMDMLKHP--HNGHMEQKVSPELLNADRRSS

Scer_DCP2/1-971 PLNDGYENISNKDSSHELLNILHGNKNSSAFNNNVYA---------TDGYSLASDNNENS
Sbay/1-992      PLNDRYPEVSSEDNSHELLNILHGNKSNNTFNNSIYTTPTATSTAAATGFPLMPTSHGNS
Smik/1-976      SPNSGYGKISSEDNSYKLPNILHDNKSSSTFNNSVYATPTEIPTATTGGYPMMPISNENS
Spar/1-981      PLNDGYGKTSDEDSSHELLNILHGNKNSSTFNSSVYSRPTEISTAGTGGYSMMPNNNENS
Scas/1-865      TLSPFSMEQSTSINTPNLGNTMQNTIPENLYDNQNRAQYS------TASNELLSMLHKKP

Scer_DCP2/1-971 SNKLLNMLQNRSSAINEPNFDVRSNGTSGSNELLSILHRK-
Sbay/1-992      SNELLGMLQNRPTATNAPTFDIRSNGNKESNELLSILHRK-
Smik/1-976      SNKLLSMLQNRPGTIHESNFDVRSNGTSGSNELLNILHRK-
Spar/1-981      SNKLLSMLQNRSSTISESNLDVRSNGTSGSNELLSILHGK-
Scas/1-865      SNEQQQANPESGLIDSFPSNSHIQNNFSASNELLNILHKRA
```



Pat1p

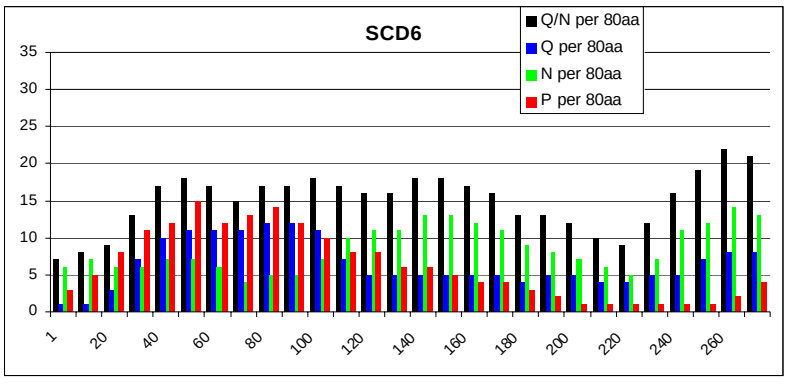

Scer_PAT1/1-797 ----------MSFFGLENSG------NARDG-----PLDFEES
Sbay/1-809 ------MPILKAVWLETEEKQQQRAPAMSFFGLENSS------NARDG-----PLDFEES
Smik/1-789 ----------MSFFGLENSS------NARDG-----PLDFEES
Spar/1-821 MVNEKNMLTLKTGKVVPGKKQQQEASTMSFFGLENNS------NGRDG-----PLDFEES
Scas/1-903 ----------MSFFGFDSSHPPSGRNNNRHGKKNEKPLDFDDT
Skud/1-795 ----------MSFFGLENSS------NARDG-----PLDFEES
Scer_PAT1/1-797 YKGYGEHELEENDYLNDETFGDNVQVGTDFDFGNPHSSGSSGNAIGGNGVGATARSYVAA
Sbay/1-809 YKGYGEHELEENDYLNNETFGDNIQVGTDFDFGTPR-SGS-SNAAAANAAGTTARSYVAA
Smik/1-789 YRGYGEHELEENDYLNDETFGDNVQVGTDFDFGNPHSSGGR-NTIGGNAVGSTARSYVAA
Spar/1-821 YRGYGEHELEENDYLNDETFGDNVQVGTDFDFGNPHNTGSS-NAIGGNGVGTTARSYVAA
Scas/1-903 YGGYDE---EENDYLNSETFGADVELGNDFDFGHGESAPQQ----QPQQPIANHRSYVAA
Skud/1-795 YKGYGGHELEENDYLNDETFGDNVQVGTDFDFGNPR-SSSGNNAVNGNAVGTTARSYVAA
Scer_PAT1/1-797 TAEGISGPR-TDGTAAAGPLDLKPMESLWSTAPPPAMAPSPQS----------
Sbay/1-809 TAEGRNGP--GPQAAAAAPLDLKPMESLWSTGLPPAMAPPSHG----------
Smik/1-789 TAEGISGPR-TDGAPAAAPLDLKPMESLWSTAPPPAIAPSPQG----------
Spar/1-821 TAEGISGPR-SNGNAAAAPLDLKPMESLWSTAPPPAMAPSPQG----------
Scas/1-903 ASQSIAHSH---PNQMDPNADLKPMESLWTASQPQQQAPPHQGQVLSMEEIERQQRQMQM
Skud/1-795 TAEGIGGPRTNAAPAAAAPLDLKPMESLWSTAPPPAMVPPSQG----------
Scer_PAT1/1-797 ----------TMAPAPAPQQMAPLQPILSMQDLERQQRQMQQQFMNFH--AMGHPQG--
Sbay/1-809 ----------AMAPGP--QQMGPPQPILSMQDLERQQRQMQQQFMNFH--PMGPPQ---
Smik/1-789 ----------VMAPGP--QQMGPLQPLLSMQDLERQQRQMQQQFMNFH--AMGHPQ---
Spar/1-821 ----------TMAPAP--QQMGPPQPILSMQDLERQQRQMQQQFMNFH--AMGHPQG--
Scas/1-903 HPQGVPPPQGFTSGPAQNFNQMPPP-PQVYQQQLPIPQQNNDGQFTNQYNPRMGPPGMGQ
Skud/1-795 ----------AMAPGP--QQMGPPQPILSMQDLERQQRQMQQQFMNFH--AMGPPQG--
Scer_PAT1/1-797 ----------LPQGPPQQQ------FPMQPASGQPGPSQFAPPPPPPGVNVN------
Sbay/1-809 ----------GPPQQQ------FPMQPPSGQPGPSQFAPPPPPPGVNAN------
Smik/1-789 ----------GPPQQQ------FPVQPASNQPGPSQFAPPPPPPGVNVN------
Spar/1-821 ----------LPQGPPQQQ------FPMQPASGQPGPSQFAPPPPPPGVNVN------
Scas/1-903 PIPGQYQQMQPMFNQPPPHPQQQYPQMYPTPPQTYQDIPLQYQAPPPPQQQPLQGQANQP
Skud/1-795 ----------LPQGPPQQQ------FPMQPVSGQPGPSQFAPAPPPPGVNVN------
Scer_PAT1/1-797 ---MNQMP---MGPVQVPVQASPSPIGMS----------NTPSPGPVVGATKMPLQS
Sbay/1-809 ---VGQMP---MGPVQVPVQASPPPMGMP----------NGSSPGPALGPGKMALQG
Smik/1-789 ---ITQMP---MGAVQVPMQASPSPIAMP----------NGPSPGPVIGSGKMPLQG
Spar/1-821 ---INQMP---MGPVQVPVQSSPSPIGMA----------NAPSPGPVVGTGKMPLQG
Scas/1-903 QHPIANMPNASQSPSTFPPQSQPNDFELQPAMNLRQPSQQTQSSSTPTPQSINSPMAQEN
Skud/1-795 ---IAQMP---IGPAQVPVQASPSPIGMS----------GGPSPGPAVGPGKMPLQG
Scer_PAT1/1-797 GRRS----------KRD-LSPEEQRRLQIRHAKVE
Sbay/1-809 GRRP----------KRD-LSLEEQRRLQIRHAKVE
Smik/1-789 GRRS----------KRD-LSPEEQRRLQIRHAKVE
Spar/1-821 GRRS----------KRD-LSPEEQRRLQIRHAKVE
Scas/1-903 ARASPQLDNRILTESPRNISRQNQNQGQSSRANTKQSQQRREPMTPEEQKRLQIRHAKVE
Skud/1-795 GRRS----------KRD-LSPEEQRRLQIRHAKVE
Scer_PAT1/1-797 KILKYSGLMTPRDKDFITRYQLSQIVTEDPYNEDFYFQVYKIIQRGGITSESNKGLIARA
Sbay/1-809 KILKYSGLMTPRDKDFITRYQLSQIVTEDPYNEDFYFQVYKIIQRGGITSESNKGLIARA
Smik/1-789 KILKYSGLMTPRDKDFITRYQLSQIVTEDPYNEDFYFQVYKIIQRGGITSESNKGLIARA
Spar/1-821 KILKYSGLMTPRDKDFITRYQLSQIVTEDPYNEDFYFQVYKIIQRGGITSESNKGLIARA
Scas/1-903 KILKHSGVMTPRDKDFITRYQLSQIVTDDPYNEDFYFQVYKIILRGGIAGESNKGLIARA
Skud/1-795 KILKYSGLMTPRDKDFITRYQLSQIVTEDPYNEDFYFQVYKIIQRGGVTSESNKGLIARA
Scer_PAT1/1-797 YLEHSGHRLGGRYKRTDIALQRMQSQVEKAVTVAKERPSKLKDQQAAAGNSSQDNKQANT
Sbay/1-809 YLEHSGHRLGGRYKRTDIALQRMQSQVEKAVTVAKERPSKLKDQQAAAGNSNQDNKQASN
Smik/1-789 YLEHSGHRLGGRYKRTDIALQRMQSQVEKAVTVAKERPSKLKDQHAAAGNSSQDNKQANT
Spar/1-821 YLEHSGHRLGGRYKRTDIALQRMQSQVEKAVTVAKERPSKLKDQQAAAGNSNQDNKQANT
Scas/1-903 YLEHSGHRLGGRYKRADIALQRMQSQVEKAVTVAKERPQKNKD-------HSGENSGREG
Skud/1-795 YLEHSGHRLGGRYKRTDIALQRMQSQVEKAVTVAKERPSKLKDQQAAAGNSNQDNKQANT
Scer_PAT1/1-797 VLGKISSTLNSKNPRRQLQIPRQQPSD-----PDALKDVTDSLTNVDLASSGSSSTGSSA
Sbay/1-809 VLGKISSTLNSKNPRRQLQIPRQYPSD-----PDALKDVTDSLTNVDLASLGSPSAGSSV
Smik/1-789 VLGKISSTLNSKNPRRQLQIPRQQPSD-----ADALKDVTDSLTNVDLTTSGSSSAG-SA
Spar/1-821 VLGKISSTLNSKNPRRQLQIPRQQPSD-----PDALKDVTDSLTNVDLTSPGSSSAGSSA
Scas/1-903 VLGKISSSMNSKAPRRQLLIPTPKSNEDLLNSEDTTRESTNDSTTPALEEVT-QHLGNVE
Skud/1-795 VLGKISSTLNSKNPRRQLQIPRQQPSD-----PDALKDVTDSLTNVDLTSSGPPSAGSSA

```
Scer_PAT1/1-797 AAVASKQRRRSSYAFNNGNGATNLNKSGGKKFILELIETVYEEILDLEANLRNGQQTDST
Sbay/1-809      AAAASKQRRRSSYAFNNGNGTTNLNKSGGKKFILELIETVYEEILDLEANLRNGQQTDST
Smik/1-789      AAAASKQRRRSSYAFNNGNGTTNLNKSGGKKFILELIETVYEEILDLEANLRNGQQTDST
Spar/1-821      AAAASKQRRRSSYAFNNGNGATNLNKSGGKKFILELIETVYEEILDLEANLRNGQQTDST
Scas/1-903      INQKSKTRRRSSYAFSSIDQNSVLSRSGGRKFVLSQIETVYEEVLELEANLRSGKEIDST
Skud/1-795      AAAASKQRRRSSYAFNNGNGTTNLNKSGGKKFILELIETVYEEILDLEANLRNGQQTDST

Scer_PAT1/1-797 AMWEALHIDDSSYDVNPFISMLSFDKGIKIMPRIFNFLDKQQKLKILQKIFNELSHLQII
Sbay/1-809      AMWDALHIDDTSYDVNPFISMLSFDKGIKIMPRIFNFLDKQQKLKILQKIFNELSHLQII
Smik/1-789      AMWDALHIDDCSYDVNPFISMLSFDKGIKIMPRIFNFLDKQQKLKILQKIFNELSHLQII
Spar/1-821      AMWDALHIDDCSYDVNPFISMLSFDKGIKIMPRIFNFLDKQQKLKILQKIFNELSHLQII
Scas/1-903      LLWEALHIDDDAYEVCPFISMLSFDKGVKIMPRIFNFLNKEQKLKLLQTFFSELSHLNII
Skud/1-795      AMWDALHIDDSSYDVNPFISMLSFDKGIKIMPRIFNFLDKQQKLKILQKIFNELSHLQII

Scer_PAT1/1-797 ILSSYKTTPKPTLTQLKKVDLFQMIILKIIVSFLSNNSNFIEIMGLLLQLIRNNNVSFLT
Sbay/1-809      ILSSYKTTPRPSLTQLKKVDLFQMIILKIIVSFLSNNSNFIEIMGLLLQLIKNNNVSFLT
Smik/1-789      ILSSYKITPKPTSPQLKKVDLFQMIILKIIVSFLSNNSNFIEIMGLLLQLIKNNNVSFLT
Spar/1-821      ILSSYKTTPKPTLTQLKKVDLFQMIILKIIVSFLSNNSNFIEIMGLLLQLIKNNNVSFLT
Scas/1-903      IISSYKTNPSPSDSQLKKIDLFQTVFLKIIVSFLSNNSNFVEIMGLLLHLIKNNNVSFIS
Skud/1-795      ILSSYKTTPRPTLMQLKKVDLFQMIILKIIVSFLSNNSNFIEIMGLLLHLIKNNNVSFLT

Scer_PAT1/1-797 TSKIGLNLITILISRAALIKQDSSRSNILSSPEISTWNEIYDKLFTSLESKIQLIFPPRE
Sbay/1-809      TSKIGLNLITILISRAALIKQDSSRSNILSSPEISTWNEIYDKLFTSLESKIQLIFPPRE
Smik/1-789      TSKIGLNLITILISRAALIKQDSSRSNILSSPEISTWNEIYDKLFTSLESKIQLIFPPRE
Spar/1-821      TSKIGLNLITILISRAALIKQDSSRSNILSSPEISTWNEIYDKLFTSLESKIQLIFPPRE
Scas/1-903      TSKIGLNLITVLISRAALIRQDSNRSNVLSSPEISAWNEIYDKLFTSLESKLSSIFPPPE
Skud/1-795      TSKIGLNLITILISRAALIKQDSSRSNILSSPEISTWNEIYDKLFTSLESKIQLIFPPRE

Scer_PAT1/1-797 YNVHIMRLQ---------NDKFMDEAYIWQFLASLALSGKLNHQRIIIDEVRDEIFATIN
Sbay/1-809      YNDYIMRLQ---------NEKFMDEAYIWQFLASLALSGKLNHQRIIIDEVRDEIFATIN
Smik/1-789      YNDHIMRLQ---------NEKFIDEAYIWQFLASLALSGKLNHQRIIIDEVRDEIFATIN
Spar/1-821      YNDHIMYLQ---------NDKFMDEAYIWQFLASLALSGKLNHQRIIIDEVRDEIFATIN
Scas/1-903      YTEKVVKIMVMEQHAGPIESIFYDQSYIWQFLASLALSGKLNHQRIIIDEVRDQIFGTIN
Skud/1-795      YNDYIMRLQ---------NEKFMDEAYIWQFLASLALSGKLNHQRIIIDEVRDEIFTTIN

Scer_PAT1/1-797 EAETLQKKEKELSVLPQRSQELDTELKSIIYNKEKLYQDLNLFLNVMGLVYRDGEISELK
Sbay/1-809      EAETLQKKETELSLLTQRSQDLEIEIKSIIYNKEKLYQDLNLFLNVMGLVYRDGEISELK
Smik/1-789      EAETLQKKEKELSILPQRSQELDTELKSIVYNKEKLYQDLNLFLNVMGLVYRDGEISELK
Spar/1-821      EAETLQKKEKELSVLPQRSQELDAELKSIIYNKEKLYQDLNLFLNVMGLVYRDGEISELK
Scas/1-903      VAEDLIKN------VPSDDK------SNAIYRREKLYQDLNLFLNVMGLVSRDGEISELK
Skud/1-795      EAETLQKKETELSTSPQRSPELETELKSITYNKEKLYQDLNLFLNVMGLVYRDGEISELK
```



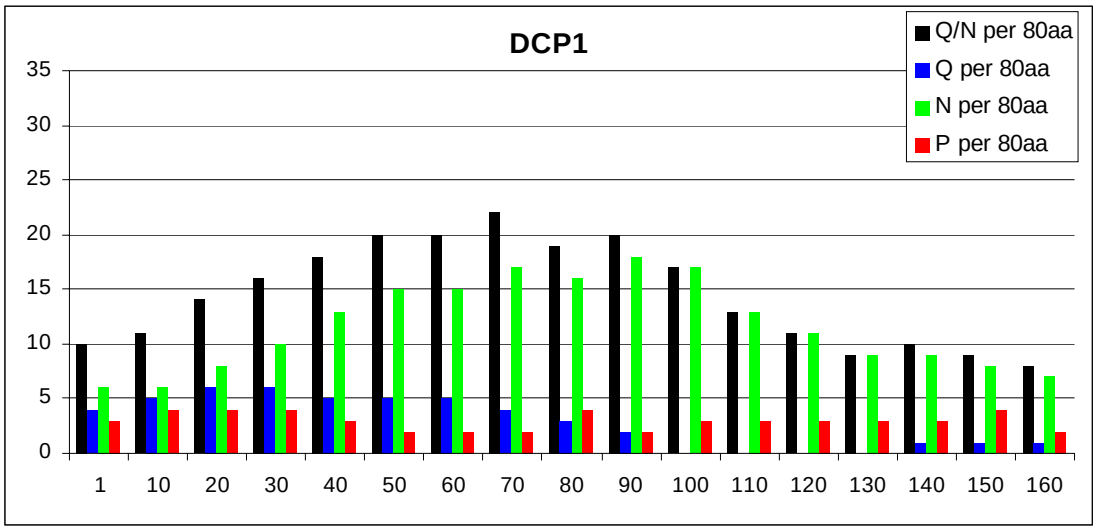
Scd6

```
Scer_SCD6/1-349 MSQYIGKTISLISVTDNRYVGLLEDIDSEKGTVTLKEVRCFGTEGRKNWGPEEIYPNPTV
Sbay/1-364      MSQYIGKTISLVSVTDNRYVGLLEDIDSEKGTVTLREVRCFGTEGRKNWGPEEIYPNPTV
Smik/1-352      MSQYIGKTISLISVTDNRYVGLLEDIDSEKGTVTLKEVRCFGTEGRKNWGPEEIYPNPTV
Spar/1-345      MSQYIGKTISLISVTDNRYVGLLEDIDSEKGTVTLKEVRCFGTEGRKNWGPEEIYPNPAV
Scas/1-370      MSEYIGKTISLISVTDNRYVGLLEGIDSDKGTVTLNEVRCFGTEGRKNWGPDEIYPQNTV
Sklu/1-270      MAQYIGKTISLISVTENRYVGLLEGIDSERGVVTLNQVRCFGTEGRKNWGPQEVYPNPTV
Skud/1-341      MSQYIGKTISLVSVTDNRYVGLLENIDSEKGTVTLKEVRCFGTEGRKNWGPEEIYPNPTV

Scer_SCD6/1-349 YNSVKFNGSEVKDLSILDANINDIQPVVPQMMPPASQFPP-QQAQSPPQAQAQAHVQTN-
Sbay/1-364      YNSVKFNGSEVKDLSILDANINDIQPVVPQMMPPATQASPSQQAQSPHQAQAQAYAQPQA
Smik/1-352      YNSVKFNGSEVKDLSILDANINDIQPVVPQTMPPASQFSSLQQVQSSPQTQTQTQVQAHS
Spar/1-345      YNSVKFNGSEVKDLSILDANINDIQPVVPQMMPPASQFPP-QQAQSSPPAQVQAQSQA--
Scas/1-370      YNSVKFNGNEVKDLSILEVKIEDVHPVLPPTATAVPMMPMQEQPRSQAPPSQQ-------
Sklu/1-270      YDTVTFNGNDVKDLNILDVKLEDVQPVLPPQV----------------------------
Skud/1-341      YNSVKFNGSEVKDLSILDANINDIQPVVPQMVPPTSQLSSSQQAQSPQEAQAS-------

Scer_SCD6/1-349 ----------------------PQVPKPESNVPAAVAGYGVYTPTSTETATASMNDKSTPQ
Sbay/1-364      QAQAQAPAPAQAPAPAQAPAPQPQATKPESNVPAAVAGYGVYTPSSTEASTAKK-DNAIP-
Smik/1-352      ----------------------PPQGAKPDSNVPAAVTGYGVYTPTSTETATTNIGDNTISE
Spar/1-345      ------------------------PKPESNVPAAVAGYGVYTPTSTETATANMNDKSLSQ
Scas/1-370      ------------------------SAPPQQVPSVVAGYGAYAPPPGSAASAPTATQTTTL
Sklu/1-270      ------------------------MMQPQQQVPPAMADYGVYAPS----------------
Skud/1-341      ----------------------PQAQKPVTNVPAAVAGYGVYTPTSTESTAVNKNDNAIPR

Scer_SCD6/1-349 DTNVNSQSRERGKNGE--------------NEPKYQRNKNRSSNRPPQSNRNFKVDIPNE
Sbay/1-364      ------QNRDERKDGE--------------HEQRYQRNK--STNRAPQSNRNHKVEIPNE
Smik/1-352      DVSVNSQSRNKRTDSE--------------HEQRYQRNKNRLTNRPPQSNRNFKVEIPNE
Spar/1-345      DANVNSQSRERGKNGE--------------NEQKYQRNKNRSTNRPPQSNRNFKVDIPNE
Scas/1-370      PSEPTTQPQQTTAFSSSTGESEFAGLQGEIQPPHDRGNRERSNQNRRRGSHSRKIEIPKS
Sklu/1-270      -------TEEKAEEPA--------------PPVRKQEPKNDEGVRRPSR----KTEIPNE
Skud/1-341      DVNKNSQSRDERKDGG--------------NEQKYQRNKNRSTNRPPQLNRNFKVEIPNE

Scer_SCD6/1-349 DFDFQSNNAKFTKGD-----------STDVE-----KEKELESAVHKQDESDEQFYNKKS
Sbay/1-364      DFDFQSNNAKFTKED-----------STDLE-----KEQELESAADKEDESDEAFYNKKS
Smik/1-352      DFDFESNNAKFTKED-----------SNDAE-----KEQELEATARKQDESDEAFYNKKS
Spar/1-345      DFDFQSNNAKFTKGD-----------SNDVE-----KEQELESASNKQDESDEPFYNKKS
Scas/1-370      DFDFQSSNAKFAKEAPDIANEHVESTSHDVPSSVHVNDGQETALAQNENATPDTFYNKKS
Sklu/1-270      DFDFESNNAKFNKPA-----------EDEEE-----QEEEQNTE---DDSSEEVFYDKKS
Skud/1-341      DFDFQSNNAKFTKED-----------SNDLE-----REQVLESAAHRQDESDEACYNKTS

Scer_SCD6/1-349 SFFDTISTSTETNTNMRWQEEKMLNVDTFGQASARPRFHSRGLGRGRGNYRGNRGN--RG
Sbay/1-364      SFFDTISTSTETNTNMRWQEEKVLNVDTFGQASARPRFHSRGRGRGNYRGNRGNT-RG
Smik/1-352      SFFDTISTSTETNTNMRWQEEKMLNVDTFGQASARPRFHSRGRGRGRGSFRGNRGG--RG
Spar/1-345      SFFDTISTSTETNTNMRWQEEKMLNVDTFGQASARPRFHSRGRGRGRGNFRGNRGN--RG
Scas/1-370      SFFDSISTSTEANTNMRWQEEKVLNLDTFGQTSARPNFHG---GRGRGRGRGGRGGRGRG
Sklu/1-270      SFFDTISTSTETNTNMRWQEERQLNMDTFGQVSARPR-----RGRG-GGFRGGRGRG-NG
Skud/1-341      SFFDTISTSTETNTNMRWQEEKVLNVDTFGQASARPRFHSRGRGRGRGNFRGNR-----G

Scer_SCD6/1-349 RGGQ--RGNYQNRNNYQNDSGAYQN-QNDSYSRPANQFSQPPSNVEF
Sbay/1-364      RGGQ--RGGYQNRNNYQNNRGGYQNYQNDSHDRSSNQFSQPPSNVEF
Smik/1-352      RGGH--RGNYQNRNSYQNNGGGFQN-QNDYHGRSTNQFSQSSSNVEF
Spar/1-345      RGGQ--RGNYQNRNNYQNNGAGFQD-QNDSYSRSANHFSQPPSNVEF
Scas/1-370      RGGRGGRGGYYNNNQHQGD---YQNRNNNNYNQHQQFSNDTTHQVEF
Sklu/1-270      RGGRGGRGGSRRGNNFRNN---FGN----------------DEKIEF
Skud/1-341      RGGQ--RGNYQNRNSYQNTRGGFQN-QNDSHGRSSNQFSQPSSNVEF
```



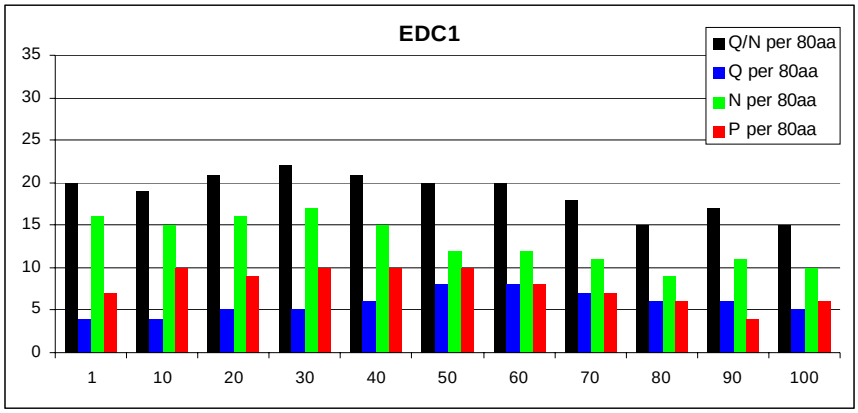
Dcp1

```
Scer_DCP1/1-232 MTG-AATAAENSATQLEFYRKALNFNVIGRYDPKIKQLLFHTPHASLYKWDFKKDEWNKL
Sbay/1-232      MTGTAATAAENSATQLEFYRKALNFNVIGRYDPKIKQLLFHTPHASLYKWDFKKDEWNKL
Smik/1-233      MTGAATTTAENSATPLEFYRKALNFNVIGRYDPKIKQLLFHTPHASLYKWDFKKDEWNKL
Spar/1-232      MTG-AATAAENSATQLEFYRKALNFNVIGRYDPKIKQLLFHTPHASLYKWDFKKDEWNKL
Sklu/1-212      MNVDSANTRASSDMTLEFYRKALNFNVIGRYDPKIKQLLFHTSHAAVYKWDCSTDGWSKL
Skud/1-233      MTETTATAAENSATQLEFYRKALNFNVIGRYDPKIKQLLFHTPHASLYKWDFKKDEWNKL

Scer_DCP1/1-232 EYQGVLAIYLRDVSQNTNLLPVSPQEVDIFDSQNGSNNIQVNNGSDNSNRNSSGNGNSYK
Sbay/1-232      EYQGVLAIYLRDISQNTDLLPVSPQEMDLFDSQNSSNNIQTSNGNNGNSSNS-GNGNGSK
Smik/1-233      EYQGVLAIYLRDVSQNINLLSIPPQETDIFDSQNGSNNIQTNNGADNSSRNSNGNGNGFK
Spar/1-232      EYQGVLAIYLRDVSQNTNLLPVSPQEVDIFDSQNGSNNIQTNNGAENNNRNSNGNGNGCK
Sklu/1-212      EYQGVLAIYLRDVSNHEDLLPVLG---------NNDNTVMTG------ATTTQQNGIGSE
Skud/1-233      EYQGVLAIYLRDVSQNTNLLSVSSQEMDMFDSQNGSNNIQTSISNNNNSSTSSGNGNGNK

Scer_DCP1/1-232 SNDSLTYNCGKTLSGKDIYNYGLIILNRINPDNFSMGIVPNSVVNKRKVFNAEEDTLNPL
Sbay/1-232      NNDSLAYNCGKTLSGKDIYNYGLIILNRINPDNFSMGIVPNSVVNKRKVFDAQEDAQNPI
Smik/1-233      NNDSLSYNCGKTLTGKDIYNYGLIILNRINPDNFSMGIVPNSVVNKRKVFNAQEDARNPL
Spar/1-232      NNDSLTYNCGKTLSGKDIYNYGLIILNRINPDNFSMGIVPNSVVNKRKVFNAEEDAHNPL
Sklu/1-212      SPSS------QLLTGHDIYNYGLIIINRINPDNFSLGIAPNSAINKRKIFNPEEDVRQPL
Skud/1-233      NNDSLADNCGKTLSGKDIYNYGLIILNRINPDIFSMGIVPNSVVNKRKVFDAQEDAKNPL

Scer_DCP1/1-232 ECMGVEVKDELVIIKNLKHEVYGIWIHTVSDRQNIYELIKYLLENEPKDSFA
Sbay/1-232      ECMGVEVKDELVIIKNLKHEVYGIWIHTVNDRQNIYELIKYLLENEPKDSFA
Smik/1-233      ECMGVEVKDELVIIKNLKHEVYGIWIHTVNDRKNIYDLIKYLLENEPKDSFA
Spar/1-232      ECMGVEVKDELVIIKNLKHEVYGIWIHTVNDRQNIYELIKYLLENEPKDSFA
Sklu/1-212      HTMGVEVKDDLLIIKSLTREVFGIWIHNESDRNNIYELIKYLLENEPKDSFT
Skud/1-233      ECMGVEVKDELVIIKNLKHEVYGIWIHTVNDRQNIYELIKYLLENEPKDSFA
```

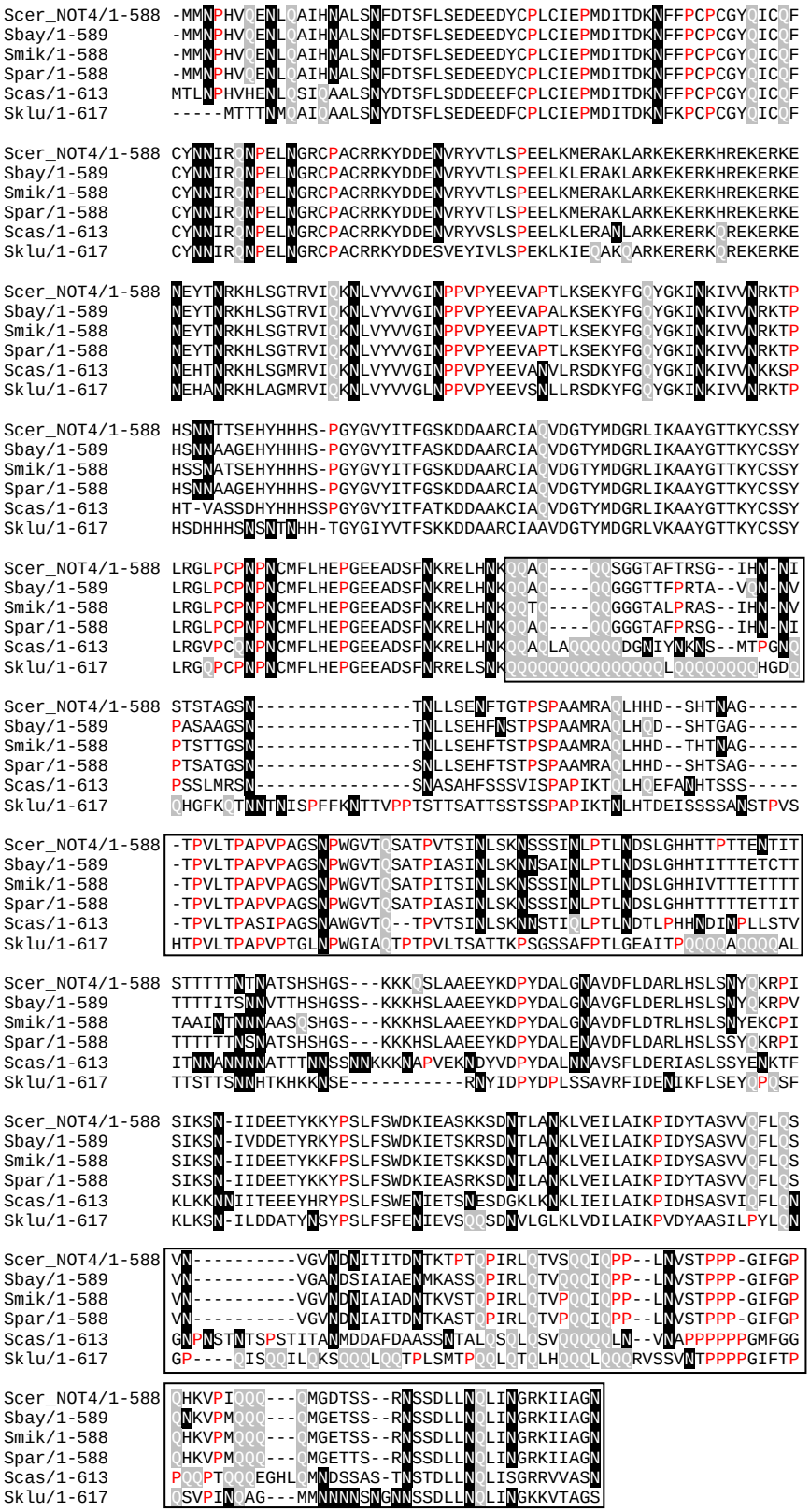


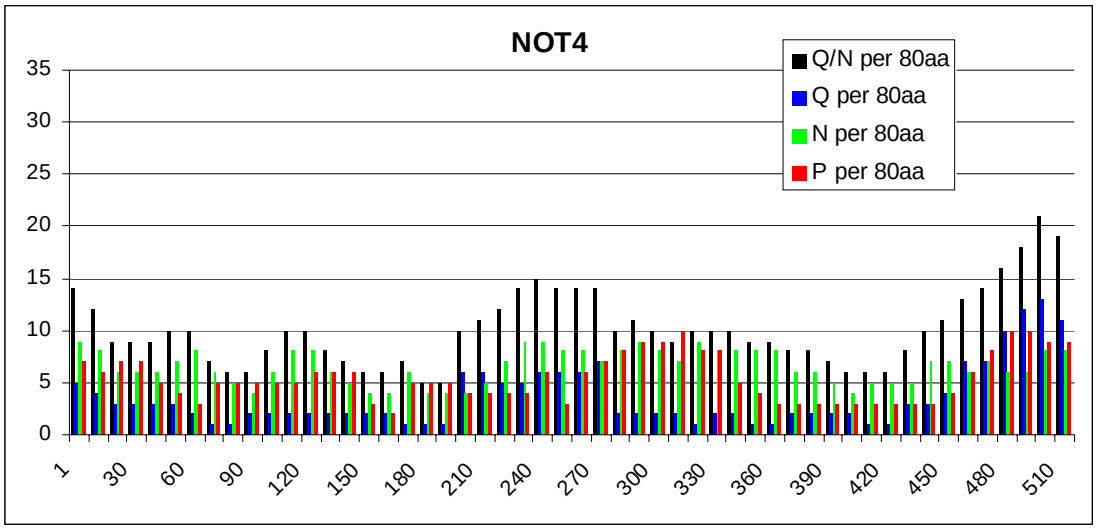
Edc1

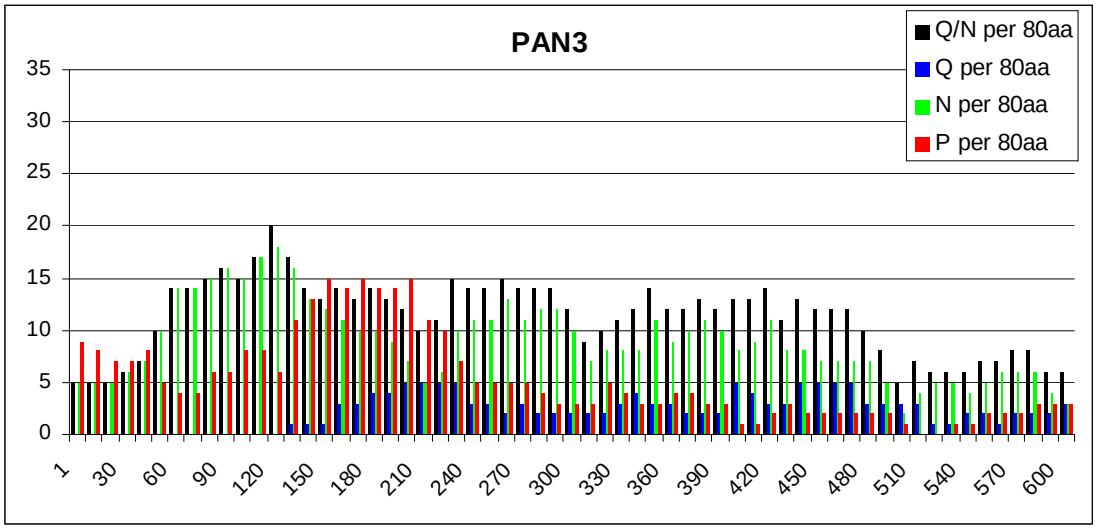
```
Scer_EDC1/1-176 MSTDTMYFNSSRLLPSAGRNKTNNLIKQKTRNNRARGNAAKNANNN---NYITDIPPPQT
Sbay/1-186      MSTDTMYFNSSRLLPSAGKNKTNNLIKHKPRKSRASENTAKNINNN---SFSTDIPPPQT
Smik/1-176      MSTDTMYFNSSRLLPSAGKNKTNNLIKQKTRNSRARGNAANSMNDN---SYITDIPPPQT
Spar/1-176      MSTDTMYFNSSRLLPSAGRNKTNNLIKQKNRNSRARGNTAKNVNNN---SHITDIPPPQT
Scas/1-371      MSTDTMYFNSSRLLPVHSKNKPISVPKVKNQPKKLNPTERKHKSKHDKKNKKTDVPEPQL
Sklu/1-360      MSTDTMYFNSSRLMSTAGKNKSNYLQKPDKKLSNTSNIVRKQHLKHD-VTVPETLPQPQA
Skud/1-176      MSTDTMYFNSSRLLPSAGKNKTNNLIKQKPRNNRASGNTAKNISNN---TSTTDIPPPQT

Scer_EDC1/1-176 LPNGQKPNFG--------HSSNKKPSFNQKKHSPP----------SSPSSTTT-------
Sbay/1-186      LPNGEKPDFG--------HSSTKKPSSRPKKHTSTPSSSSSSSSSSPSSTSS-------
Smik/1-176      LPNGEKPNFG--------HSSNKRLSFHQKKHSSP----------SSTSSTTS-------
Spar/1-176      LPNGEKPNFG--------HSSSKKPSFHQKKHSPP----------SSPSSTAA-------
Scas/1-371      LPNGEKPNFGNQNNDKSGSSSGKKPS-RPKKHSKQEKQFEQQSVSESTDRLTQTL-----
Sklu/1-360      LPNGEKPNFGN-------ASNKKSNSKGQYKKNRGQNYHQKKNEDSSADDLTNDLKQLLS
Skud/1-176      LPNGEKPNFG--------HSSNKKPSFYQKKHSSP----------SPPSSMSS-------

Scer_EDC1/1-176 -----------LGKKNRQNNKETPRQNNKDDTRLLSQN-------------LKNLLLN-
Sbay/1-186      -----------LGKKNRQYNTEGPRQNNGDETRLLSQN-------------LKNMLLN-
Smik/1-176      -----------LGKKNRQHNKETPRQGNKDETRLPSQN-------------LKNLLLN-
Spar/1-176      -----------LGKKNRQNNKEVPRQNNKDDTRLLSQN-------------LKNLLLN-
Scas/1-371      --------KDLLLKSQTSSSSASPMINNNNRVAVTSNNDIASTISPLNTPSTIPAALLNP
Sklu/1-360      VSEGEKEKARELQKEKRPKSNATQKKYPKEDKPIDTRPAATKVSSDSQSGIPKTQQLMSP
Skud/1-176      -----------PGKKNREHNKEGPRQSNRNESRLHSQN-------------LKNLLLN-

Scer_EDC1/1-176 ------------QKQSP-----------------------------------------
Sbay/1-186      ------------PIQAP-----------------------------------------
Smik/1-176      ------------QKKAP-----------------------------------------
Spar/1-176      ------------QKQSP-----------------------------------------
Scas/1-371      MGLSPIPQ----QQHQQPQLMTPPIMHPGLYPQQTLSPFAYQQQYQNSPQPPPLIHGNGS
Sklu/1-360      QVPSISPLPLHPGLYQQPPLNSTQQCAGYPYGNKSFVNQYQQPNYALTSNTLPHLTTPAA
Skud/1-176      ------------QKQTP-----------------------------------------

Scer_EDC1/1-176 ------------------------------------------------------HGSQG----
Sbay/1-186      ------------------------------------------------------HRPQG----
Smik/1-176      ------------------------------------------------------HNSYG----
Spar/1-176      ------------------------------------------------------HNPQG----
Scas/1-371      SIPGTGGGYPFQGYPYVSNVHYSLTSGTNNQMPNNAVSPTPAHMNLMFPYSNNNNGNNLS
Sklu/1-360      TLMN-FPLQPMQPAPPTPPMHYHPQMNMSPHPINVINGHIPHNYAPIYQPNYPTQVPMVM
Skud/1-176      ------------------------------------------------------HNPRG----

Scer_EDC1/1-176 ---IIPMGCNGSA-----------------------------------------------
Sbay/1-186      ---IIPMNCSGSA-----------------------------------------------
Smik/1-176      ---IIPMNSSGNG-----------------------------------------------
Spar/1-176      ---IIPMSCNGGA-----------------------------------------------
Scas/1-371      TMHVPPLSSSSSTSRQIQHSAPHTAMASPSTNSKPTSRSNSRPTSNASFNSRSPPKKNNV
Sklu/1-360      AVPGLPPTSSGSVASN-------------------SNASTTHATSKRNKRKPSGGGGGS
Skud/1-176      ---IIPMNCNGSA-----------------------------------------------

Scer_EDC1/1-176 KKLSHSYAGSTFATNGPREAKNLPKPSFL
Sbay/1-186      KKFSHSYAGSTFATNGPREAKSLPKPSFL
Smik/1-176      KKISHSYAGSTFATNGPREAKELPKPSFL
Spar/1-176      KKFSHSYAGSTFATNGPREAKNLPKPSFL
Scas/1-371      RRGSQSFAGASFATAIPQECN-LPKPSFT
Sklu/1-360      SSSSGGYAGASFTTHAPTITN-LPKPSFA
Skud/1-176      KKFSHSYAGSTFATNGPREAKSLPKPSFL
```



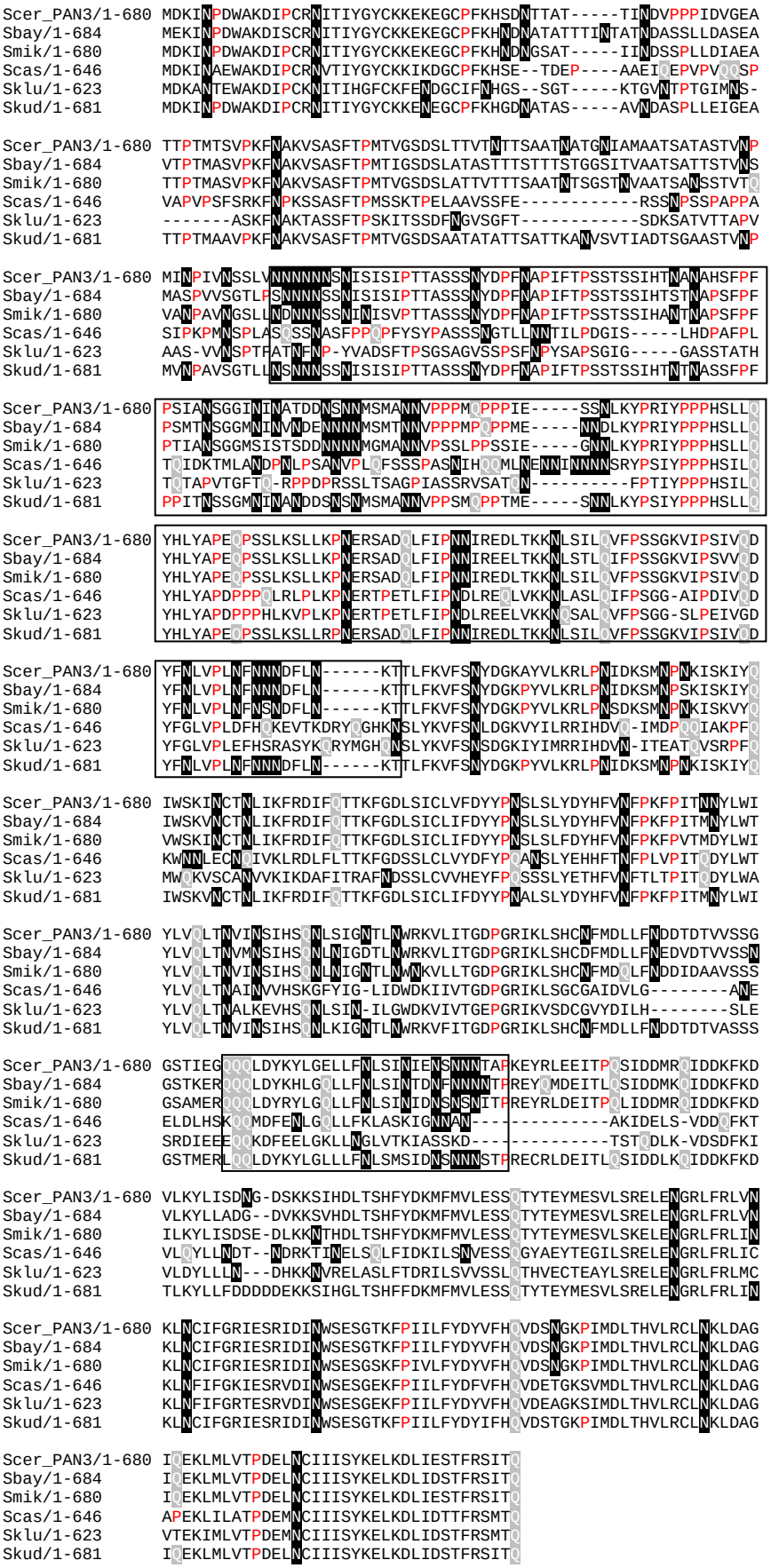
Not4


Scer_NOT4/1-588
Sbay/1-589
Smik/1-588
Spar/1-588
Scas/1-613
Sklu/1-617


NOT4
Q/N per 80aa
Q per 80aa
N per 80aa
P per 80aa
35
30
25
20
15
10
5
0
1
30
60
90
120
150
180
210
240
270
300
330
360
390
420
450
480
510


PAN3
Q/N per 80aa
Q per 80aa
N per 80aa
P per 80aa
35
30
25
20
15
10
5
0
1
30
60
90
120
150
180
210
240
270
300
330
360
390
420
450
480
510
540
570
600

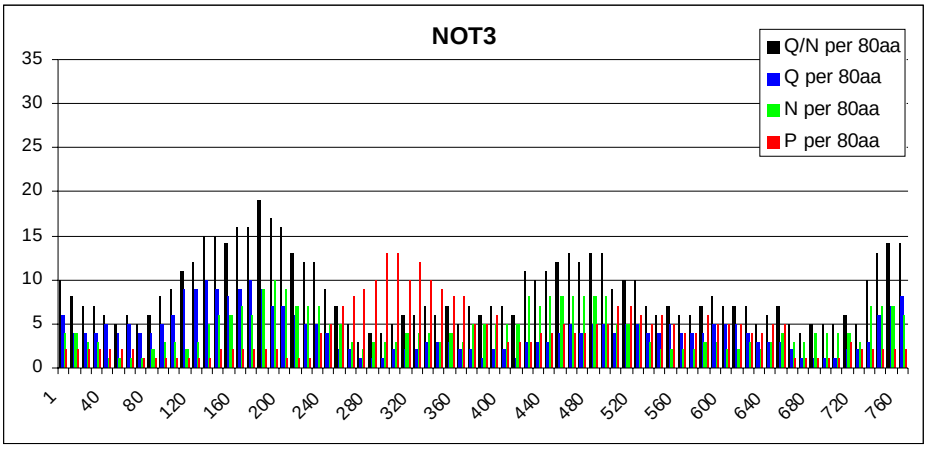
Pan3


Scer_PAN3/1-680
Sbay/1-684
Smik/1-680
Scas/1-646
Sklu/1-623
Skud/1-681

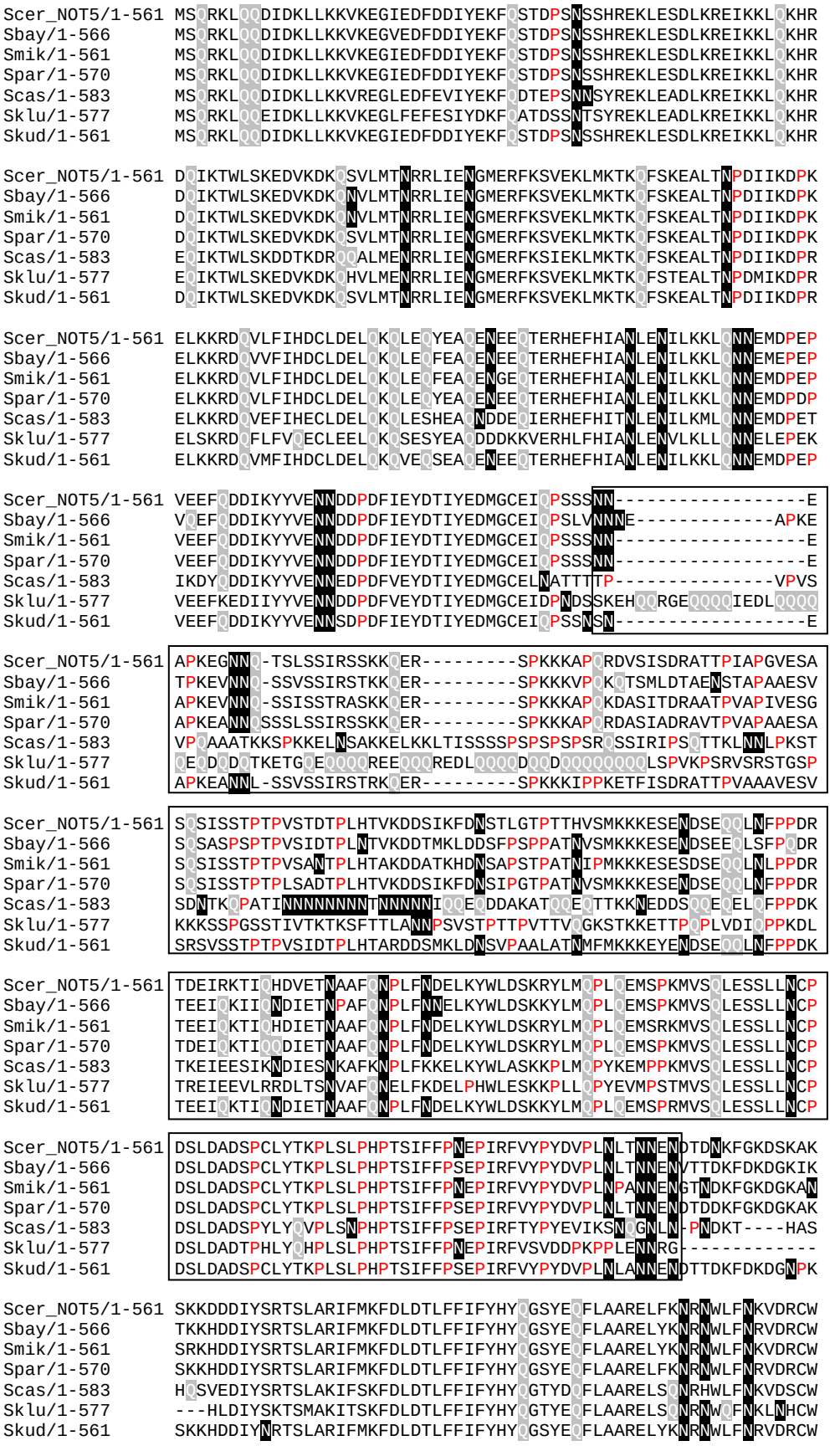
Not3

```
Scer_NOT3/1-837 MAHRKLQQEVDRVFKKINEGLEIFNSYYERHESCTNNPSQKDKLESDLKREVKKLQRLRE
Sbay/1-848      MAHRKLQQEVDRVFKKINEGLEIFNSYYERHESCTNNPSQKDKLESDLKREVKKLQRLRE
Smik/1-839      MAHRKLQQEVDRVFKKINEGLEIFNSYYERHESCTNNPSQKDKLESDLKREVKKLQRLRE
Spar/1-839      MAHRKLQQEVDRVFKKINEGLEIFNSYYERHESCTNNPSQKDKLESDLKREVKKLQRLRE
Scas/1-801      MAHRKLQQEVDRVFKKINEGLDIFNTYYERHESCNNNPSQKDKLESDLKREVKKLQRLRE
Sklu/1-830      MAHRKLQQEIDRVFKKINEGLEVFNTYYDRHEACTNNPSQKDKLESDLKREVKKLQRLRE
Skud/1-842      MAHRKLQQEVDRVFKKINEGLEIFNSYYERHESCTNNPSQKDKLESDLKREVKKLQRLRE

Scer_NOT3/1-837 QIKSWQSSPDIKDKDSLLDYRRSVEIAMEKYKAVEKASKEKAYSNISLKKSETLDPQERE
Sbay/1-848      QIKSWQSSPDIKDKDSLLDYRRSVEIAMEKYKAVEKASKEKAYSNISLKKSETLDPQERE
Smik/1-839      QIKSWQSSPDIKDKDSLLDYRRSVEIAMEKYKAVEKASKEKAYSNISLKKSETLDPQERE
Spar/1-839      QIKSWQSSPDIKDKDSLLDYRRSVEIAMEKYKAVEKASKEKAYSNISLKKSETLDPQERE
Scas/1-801      QIKSWQSSPDIKDKDSLLEYRRSVEIAMEKYKAVEKASKEKAYSNISLKKSEILEPEEQE
Sklu/1-830      QIKSWQSSPEVKDKDSLLDYRRSVEVAMEKYKAVEKASKEKAYSNISLKRSEVLDPQERE
Skud/1-842      QIKSWQSSPDIKDKDSLLDYRRSVEIAMEKYKAVEKASKEKAYSNISLKKSETLDPQERE

Scer_NOT3/1-837 RRDISEYLSQMIDELERQYDSLQVEIDKLLLLNKKKKTSSTTNDEKKEQYKRFQARYRWH
Sbay/1-848      RRDISEYLSQMIDELERQYDLLQVEIDKLLLLNKKKKTFSVSNEEKKEQFKHFQTRHRWH
Smik/1-839      RRDTSEYLSQMIDELERQYDSLQVEIDKLLLLNKKKKTSSTSNDEKKEQFRRFQARYRWH
Spar/1-839      RRDISEYLSQMIDELERQYDSLQVEIDKLLLLNKKKKTSSTTNDEKKEQYKRFQARYRWH
Scas/1-801      RRDASDYISSMIDELERQYEFLQVEIDKLLLLNKKKKTASQLNDEKIEHMKTLQLRYRWH
Sklu/1-830      RRIVADYLSQSIDELERQYDLLQVEVDRLLLLNKKKKTATSANEEKKKGLKELQARYRWH
Skud/1-842      RRDMSEYLSQMIDELERQYDSLQVEIDKLLLLNKKKKTSSTSNDEKKEQFKHFQGRYRWH

Scer_NOT3/1-837 QQQMELALRLLANEELDPQDVKNVQDDINYFVESNQDPDFVEDETIYDGLNLQSNEAIAH
Sbay/1-848      QQQMELALRLLANEELDPQDVKNVQDDINYFVESNQDTDFVEDETIYDALDLQSNEAIAH
Smik/1-839      QQQMELALRLLANEELNPQDVKNVQDDINYFVESNQEPDFVEDETIYDGLNLQSNEAIAH
Spar/1-839      QQQMELALRLLANEELDPQDVKNVQDDINYFVESNQDPDFVEDETIYDGLNLQSNEAIAH
Scas/1-801      QQQMELALRLLANEELDPQAVNDAKDDINYYVESNQDQDFIEDETIYDSLNLQSNEAIAH
Sklu/1-830      QQQMELALRLIANEELDPQLVKDIEDDINYFIESNQAPDFVEDDTIYEGLNLDANEAIAH
Skud/1-842      QQQMELGLRLLANEELDPQDVKNVQDDINYFVESNQDADFVEDETIYDGLNLQSNEAIAH

Scer_NOT3/1-837 EVAQYFASQNAEDNNTSDANESLQDISKLSKKEQRKLEREAKKAAKLAAKNAT------G
Sbay/1-848      EVAQYFASQKAEDNNTSDANDSLQDASRLSKKEQRKLEREAKKAAKMAAKSTTDTTTTIG
Smik/1-839      EVAQYFASQNAEDNNTSDANESLQDISKLSKKEQRKLEREAKKAAKLAAKNST------G
Spar/1-839      EVAQYFASQNAEDNNTSDANESLQDISKLSKKEQRKLEREAKKAAKLAAKNAT------G
Scas/1-801      EVAQYFASQSAADNED-NAEETNTDASKLSKKEQRKLEREAKKAAKLAAKSAP-------
Sklu/1-830      EVAASFAAQAS---DDAAEEEASKDTSKISKKEQRRLEREAKKAAKAAAKTTVDGAPSIA
Skud/1-842      EVAQYFASQNAEDNDTSDANESLQDISKLSRKEQRKLEREAKKAAKLLAKNAT------G

Scer_NOT3/1-837 AAIPVAGP----------------------SSTPSPVIPVADASKETERSPS--SSPIHNA
Sbay/1-848      STISAPDL----------------------SSTPSPVTPTANISRETERSPS--ASPIQSV
Smik/1-839      SAISAPGP----------------------SSTPSP-APIPDALKETERSLS--ASPIHNV
Spar/1-839      SAISAPGP----------------------SSTPSPVIPIADVSKETERSAS--ASPIHNG
Scas/1-801      TTASPSVK----------------------LEINKSATPSPVLKNESVELSS--MKSIPTS
Sklu/1-830      TTTTTTTTPPPPLETAISALTEASKGRPEASISTPAEPPKTATTNTSPSPSPMLASAPTS
Skud/1-842      SNISAPGP----------------------SNTPSPVTPFADISKETERSPS--ASPIHNM

Scer_NOT3/1-837 TKPEEAVKTS-IKSPRSSADNLLPSLQKSPSSATPETPTN--VHTHIHQTPNGITGATTL
Sbay/1-848      PKPEESIKPFSIKSPKPSTENLLQSLQKSPSSTTPGTPTN--VHTHIHQTPNGITGATTL
Smik/1-839      AKTEEPNKASSIKSPRPSTDNQLPSLQKSPSSATPETPTN--VHTNLHQTPNGITGATTL
Spar/1-839      VKPEEATKASSIKSPRPSADNLLPSLQKSPNSSTPETPTN--VHTHIHQTPNGNTGATTL
Scas/1-801      -----ASVPNTAKSP--NALPQLTSLKSSPPNEQHQDIQG--THTHIHQGQNGLTSSTIL
Sklu/1-830      IPSTANTTTATIASAASISKPQTPLSKPSPHNLTRTSSPKPQGHTHFHQSLNGTTTTSTL
Skud/1-842      VKPEGPTKASSIKSPKPSTDNLLPSLQKSPSSATPETPTS--VHTHIHQTPNGITGATTL

Scer_NOT3/1-837 KPATLPAKPAGELKWAVAASQAVEKDRKVTSASSTISNTSTKTPTTAAATTTSSNANSRI
Sbay/1-848      KPATLPVKPVGELKWAVAASQAIEKDRKVTSASSTISNTSTKTPTTAVATTTPSNANSRI
Smik/1-839      KPATLPAKPVGELKWAVAASQAVEKDRKVTSASSTISNTSTKTPTTAAATTTSSNANSRI
Spar/1-839      KPATLPAKPAGELKWAVAASQAVEKDRKITSASSTISNTSTKTPTTAAATTTSSNANSRI
Scas/1-801      KPATVPARPAGELKWSVAASMGLEKDKKST--STDISTPVSKP----------------
Sklu/1-830      KPATVPVRPAGELKWAVAAAQAAEKDKKHVSNQNPLTATSSATS-------------SRP
Skud/1-842      KPATLPAKPTGELKWAVAASQAIEKDRKVTSASSTISNTSTKTPTTAAPTTTSSSANSRI

Scer_NOT3/1-837 GSALNTPKLSTSSLSLQPDNTGASSS-AATAAAVLAAGAAAVHQNNQAFYRNMSS---SH
Sbay/1-848      SSTLNTPKLTTPSLSLQPENNGTSSSSAANAAAVLAAGAAAVNQNNQAFYRNMNS---SH
Smik/1-839      GSALNTPKISTSSLSLQPDNTGASSSSAATAAAVLAAGAAAVHQNNQAFYRNMNS---SH
Spar/1-839      GSALNTPKISTSSLSLQPDNTGASSSSAATAAAVLAAGAAAVHQNNQAFYRNMSS---SH
Scas/1-801      SSIVTTPRLGTP--VLEKSTITPTSSSSVTAAAVLAAGAAAVHQNNQQFHKGFENNVGSQ
Sklu/1-830      NSVVSTPMVSKSSLVTQEKPPTAASVLTAGVVATASPVAQAMGLSPQPQQQQQQQ---QQ
Skud/1-842      GSASNTPKISTSSLSLQPDNSGALCSSAATAAAVLAAGAAAVHQNNQAFYRNMNS---SH
```

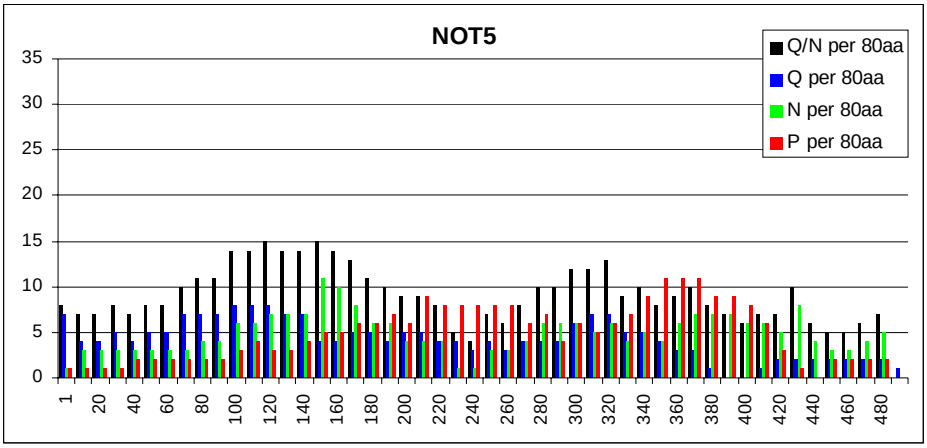
```
Scer_NOT3/1-837 HPLVSLATNPKSEHE--VATTVN-QNGPENTTKKVMEQKEEESPEERNKLQVPTFGVFDD
Sbay/1-848      HPLVSLATTTKTEHVSSISGTPLNEGAFENTTKPVSEQQEEAESTEEGQLQVPTFGVFED
Smik/1-839      HPLVSLATNIKSEYE--ITDTVANQNGTDNANKNVTEQKEEVPAMEVDQLEGPTFGVFDD
Spar/1-839      HPLVSLATNTKPEHE--VTTTVN-QNGPDNTSKKAMEQKEEESPTEGDELQVPNFGVFDD
Scas/1-801      TPSSILLNNVKQEGQ----------KDQESETNPVAETTNTVEP----EIESPSTPSLIE
Sklu/1-830      QQQQQHPASGSTEVG----------AASNNFTNSLLTSQVPEEE-------SHELELLID
Skud/1-842      HPLISLAANLKSERE--VTAPEA--PVIENASKQVMEQKEEELSTEGDQLQVPTFGVFDD

Scer_NOT3/1-837 DFESDRDSETEPEEEEQPSTPKYLSLEQREAKTNEIKKEFVSDFETLLLPSGVQEFIMSS
Sbay/1-848      DFESDRDLETESEDEEQPSTSIHLSSEQKEVRTNEIKNEFVSDFETLLLPSGVQEFIMSS
Smik/1-839      DFESDRNLETESEDEGQSSNPKYLSSEQRDAMTNEIKKEFLSDFETLLLPSGVQEFIMSS
Spar/1-839      DFESDRDSETEPEDEEQQSTPKYLSLEQREAKTNEIKKDFVSDFETLLLPSGVQEFIMSS
Scas/1-801      DYESDISDDDLGEDEPNLIPLTPDELKKNIIAHDHLHQEYMSDWGSLLLPSGIQEFIMGV
Sklu/1-830      DYNSDLS-DDELEEEPRELVMDEQELNRRRQVKKSLEESMCDNIGLLCLPSGIQDFIMCS
Skud/1-842      DFESDRDSETETEDEEQPSTPKILSSEQREARTSEIKKEFVSDFETLLLPSGVQEFIMSS

Scer_NOT3/1-837 ELYNS--QIESKIT----YKRSRDMCEISRLVEVPQGVNPPSPLDAFRSTQQWDVMRCSL
Sbay/1-848      ELYNN--QIESKIT----YKRSRDMCEISRLVEVPQGVNPPSPLDAFRSTQQWDVMRCSL
Smik/1-839      ELYNS--RIESKIT----YKRSRDMCEISRLVEVPQGVNPPSPLDAFRSTQQWDMMRCSL
Spar/1-839      ELYNS--QIESRIT----YKRSRDMCDISRLVEVPQGVNPPSPLDAFRSTQQWDVMRCSL
Scas/1-801      ELTKN--NLNSNNGRLGGYRRSIDLCEVNRLDPIPHGVNPPTPLDAFRSTQKWDIVRCSL
Sklu/1-830      VVSKNKLHINGKLG---GYRRPLDACKISRLNPIPLGVNPPSPLDALRSTNQWDLIRCTI
Skud/1-842      ELYNS--QIESKIT----YKRSRDMCEISRLVEVPQGVNPPSPLDAFRSTQQWDVMRCSL

Scer_NOT3/1-837 RDIIIGSERLKEDSSSIYAKILENFRTLEMFSLFYNYYFAITPLEREIAYKILNERDWKV
Sbay/1-848      RDVIIDSETQKESSSLVYDKILENFRTLEMFTLFYNYYFSITPLEREIAFKILNERDWKV
Smik/1-839      RDVIIGSERLKEDSSSIYAKILENFRTLEMFSLFYNYYFAITPLEREIAYKILNERDWKV
Spar/1-839      RDVIIGSERLREEPSTVYAKILENFRTLEMFSLFYNYYFAITPLEREIAYKILNERDWKV
Scas/1-801      RDVISKS----ESETEIYQEIIQRFRGLEMFTLFYNYYFAVTPLEKEISNVILNERSWRI
Sklu/1-830      FG-----MEPMENEISELDDILDRFRPLETFTLFYSYYYSITPLERKIAMALLKERNWKI
Skud/1-842      RDVIIDSERRNEDSSSVYAKIIKNFRTLEMFSLFYNYYFAITPLEREIAFKILNERDWKV

Scer_NOT3/1-837 SKDGTMWFLRQGEVKFFNEICEVGDYKIFKLDDWTVIDKINFRLDYSFLQPPVDTASEVR
Sbay/1-848      SKDGAMWFLRQGEIKFFNEICEVGDYKIFKLDDWTVIDKVNFRLDYSFLQPLVDSVPEAG
Smik/1-839      SKDGTMWFLRQGEVKFFNEVCEVGDYKIFKLDDWTVIDKINFRLDYSFLQPTVDTVSEVR
Spar/1-839      SKDGTMWFLRQGEVKFFNEICEVGDYKIFKLDDWTVIDKINFRLDYSFLQPPVHTASEVR
Scas/1-801      SKDETLWFLRQGSVKLQNEFCEIGDYKIFKLDDWTVIDKINFKLDYSNLKISSPTTTDSA
Sklu/1-830      SKGGNTWFLRQSATKFSNELCEVADYKIFKLDDWTVVDKLNFKLDYSVLEEIPLQATAMR
Skud/1-842      SKDGTMWFLRQGEIKFLNEICEVGDYKIFKLDDWTVIDKINFRLDYSFLQSSVDAVSEPR

Scer_NOT3/1-837 DVSVDN----NNVNDQSNVTLEQQKQEISHGKQLLKQLKQGKISV-----------
Sbay/1-848      DIDVNN----NDVNDQSNVTLEQQKQQIFHGKQLLNQLRQGKNNV-----------
Smik/1-839      DMNVGN----NDVNDQSNVTLEQQKQQISHGKQLLKQLRQGKISV-----------
Spar/1-839      DMNVDN----NNDNDQSNVTLEQQKQQISHGKQLLKQLRQGKIGV-----------
Scas/1-801      ADVIEE----SESINQETVSAESDK--LSHGQQLLQQLKQGKVGVPV---------
Sklu/1-830      ----------------------QSESSMSHGQQLLQQLKQGPIPNKFSSEGIVAQK
Skud/1-842      GMNVDNNADNNDVNDQSNVNLEQQRQQISHGKQLLKQLRQGKINV-----------
```



Not5


Scer_NOT5/1-561
Sbay/1-566
Smik/1-561
Spar/1-570
Scas/1-583
Sklu/1-577
Skud/1-561

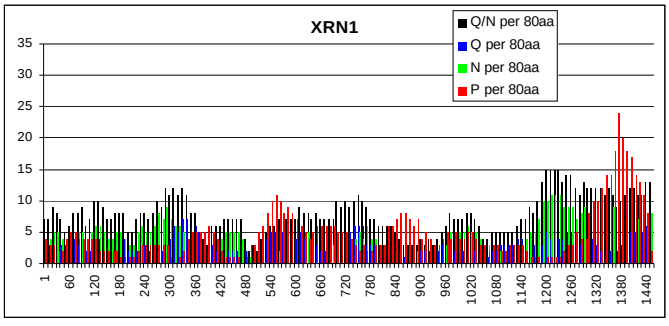
```
Scer_NOT5/1-561 YYKEIEKLPPGMGKS--------------EEESWRYFDYKKSWLARRCGNDFVYNEEDFE
Sbay/1-566      YYKEIEKLPPGMGKS--------------EEESWRYFDYKKSWLARRCGNDFVYNEEDFE
Smik/1-561      YYKEIEKLPPGMGKS--------------EEESWRYFDYKKSWLARRCGNDFVYDEEDFE
Spar/1-570      YYKEIEKLPPGMGKS--------------EEESWRYFDYKKSWLARRCGNDFLYNEKDFE
Scas/1-583      YFREIEKLPPTINMSNTIPNNKREDSNSDEEESWRYFDYKKSWLARRCGPDFVYHEEGFE
Sklu/1-577      FYKEVEKLPPGVVSSE----------NKQEEISWRYFDYQKSWLARRCGSDFVYREEEFE
Skud/1-561      YYKEIEKLPPGMGKS--------------EEESWRYFDYKKSWLARRCGNDFVYNEEDFE

Scer_NOT5/1-561 KL--------
Sbay/1-566      KL--------
Smik/1-561      KL--------
Spar/1-570      KLQTVYFLPI
Scas/1-583      KL--------
Sklu/1-577      KI--------
Skud/1-561      KL--------
```



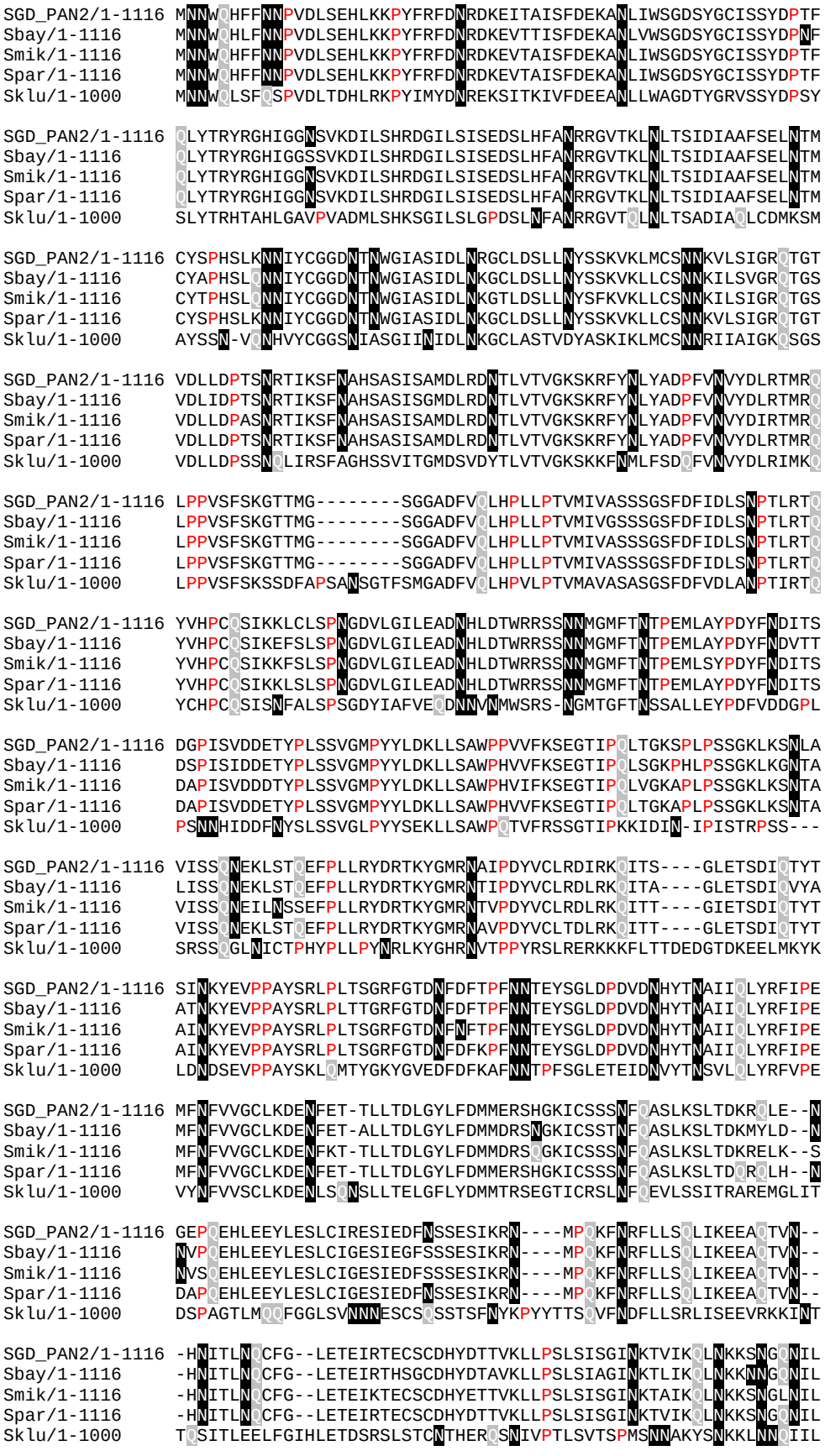
Xrn1

```
Scer_XRN1/1-1529 MGIPKFFRYISERWPMILQLIEGTQIPEFDNLYLDMNSILHNCTHGNDDDVTKRLTEEEV
Sbay/1-1516      MGIPKFFRYISERWPMILQLIEGTQIPEFDNLYLDMNSILHNCTHGNDDDVTKRLTEEEV
Smik/1-1528      MGIPKFFRYISERWPMILQLIEGTQIPEFDNLYLDMNSILHNCTHGNDDDVTKRLTEEEV
Scas/1-1489      MGIPKFFRYISERWPMILQLIEGTQIPEFDNLYLDMNSILHTCTHGNDDDVTKRMTEEEV
Skud/1-1527      MGIPKFFRYISERWPMILQLIEGTQIPEFDNLYLDMNSILHNCTHGNDDDVTKRLTEEEV

Scer_XRN1/1-1529 FAKICTYIDHLFQTIKPKKIFYMAIDGVAPRAKMNQQRARRFRTAMDAEKALKKAIENGD
Sbay/1-1516      FAKICTYIDHLFQTIKPKQIFYMAIDGVAPRAKMNQQRARRFRTAMDAEKAMKKAIENGD
Smik/1-1528      FAKICTYIDHLFQTIKPKKIFYMAIDGVAPRAKMNQQRARRFRTAMDAEKALRKAIENGE
Scas/1-1489      FAKIFTYIDHLFLTIKPKKTFYMAIDGVAPRAKMNQQRSRRFRTAMDAEHALQKAIDHGE
Skud/1-1527      FAKICTYIDHLFQTIKPNKIFYMAIDGVAPRAKMNQQRARRFRTAMDAEKAMKKAIENGD

Scer_XRN1/1-1529 EIPKGEPFDSNSITPGTEFMAKLTKNLQYFIHDKISNDSKWREVQIIFSGHEVPGEGEHK
Sbay/1-1516      EIPKGEPFDSNCITPGTEFMAKLTKNLQYFIHDKISNDSKWREVQIIFSGHEVPGEGEHK
Smik/1-1528      EIPKGEPFDSNSITPGTEFMAKLTKNLQYFIHDKISNDSKWREVQIIFSGHEVPGEGEHK
Scas/1-1489      EIPKGEPFDSNSITPGTEFMAKLTKNLKYFIHDKISNDAKWREIDIIFSGHEVPGEGEHK
Skud/1-1527      EIPKGEPFDSNSITPGTEFMAKLTKNLQYFIHDKISNDSKWREVQIIFSGHEVPGEGEHK

Scer_XRN1/1-1529 IMNFIRHLKSQKDFNQNTRHCIYGLDADLIMLGLSTHGPHFALLREEVTFGRRNSEK-KS
Sbay/1-1516      IMNFIRHLKSQKDFNQNTRHCIYGLDADLIMLGLSTHGPHFALLREEVTFGRRNSEK-KS
Smik/1-1528      IMNFIRHLKSQKDFNQNTRHCIYGLDADLIMLGLSTHGPHFALLREEVTFGRRNSEK-KS
Scas/1-1489      IMDFIRRITAEKDFDENTRHCIYGLDADLIILGLSTHAPHFALLREEVVFGRRNSNKVKT
Skud/1-1527      IMNFIRHLKSQKDFNQNTRHCIYGLDADLIMLGLSTHGPHFALLREEVTFGRRNSEK-KS

Scer_XRN1/1-1529 LEHQNFYLLHLSLLREYMELEFKEIADEMQFEYNFERILDDFILVMFVIGNDFLPNLPDL
Sbay/1-1516      LEHQNFFLLHLSLLREYMELEFKEIADEMQFEYDFERILDDFILVMFVIGNDFLPNLPDL
Smik/1-1528      LEHQNFYLLHLSLLREYMELEFKEIADEMQFEYDFERILDDFILVMFVIGNDFLPNLPDL
Scas/1-1489      LENQNFYLLHLSLLREYMELEFSEIADEMQFPFDFERVLDDFILVMFVIGNDFLPNLPDL
Skud/1-1527      LEHQNFYLLHLSLLREYMELEFKEIADEMRFEYDFERILDDFILVMFVIGNDFLPNLPDL

Scer_XRN1/1-1529 HLNKGAFPVLLQTFKEALLHTDGYINEHGKINLKRLGVWLNYLSQFELLNFEKDDIDVEW
Sbay/1-1516      HLNKGAFPVLLQTFKEALLHTDGYINEHGKINLRRLGVWLDYLSQFELLNFEKDDIDVEW
Smik/1-1528      HLNKGAFPVLLQTFKEALLHTDGYINEHGKINLKRLGVWLNYLSQFELLNFEKDDIDVEW
Scas/1-1489      HLNKGAFPVILQTFKEALLHLDGYINEHGKINLERLRVWFQYLSQFELLNFEKSDIDVEW
Skud/1-1527      HLNKGAFPVLLQTFKEALLHTDGYINEHGKINLKRLGVWLDYLSQFELLNFEKEDIDVEW

Scer_XRN1/1-1529 FNKQLENISLEGERKRQRVGKKLLVKQQKKLIGSIKPWLMEQLQEKLSPDLPDEEIPTLE
Sbay/1-1516      FNKQLENISLEGERKRQRVGKKLLIKQQKKLIGKIKPWLMEQLQEKLSPDLPDEKIPTLE
Smik/1-1528      FNKQLENISLEGERKRQRVGKKLLVKQQKKLIGSIKPWLMEQLQEKISPDLPDESIPTLE
Scas/1-1489      FNQQLENISLEGERKRARMGKKLLIKQQKKIVGMVKPWIMKELSVKIPTDLPEDQVPSLT
Skud/1-1527      FNKQLENISLDGERKRQRVGKKLLVKQQKKLIGSIKPWLMEQLQGKISPDLPDEKIPTLE

Scer_XRN1/1-1529 LPKDLDMKDHLEFLKEFAFDLGLFITHSKSKGSYSLKMDLDSINPDETEEEFQNRVNSIR
Sbay/1-1516      LPKDLNVKDHLEFLKEFAFDLGLFITHSKSKDTYSLKMDLDSISPDETEEEFQTRVNSIR
Smik/1-1528      LPKDLDIKDHLEFLKEFAFDLGLFITHSKSKNSYSLKMDLDSINPDETEEEFENRVNTIR
Scas/1-1489      IKQE-NIVENLDFLKKFAFDLGLFVVHSKSKDTYSLRVDLDSINPSETEEERQKRVAGLR
Skud/1-1527      LPRDLDIKDHLEFLKEFAFDLGLFITHSKSKNTYSLKLDLDSINPDETDEEFQNRVNSTR

Scer_XRN1/1-1529 KTIKKYQNAIIVEDKEELETEKTIYNERFERWKHEYYHDKLKFTTDSEEKVRDLAKDYVE
Sbay/1-1516      KTMKKYQNAIFVEDEEELDTEKTIYNDRFEKWKHEYYHDKLKFTSDKEEKVTDLAKDYVE
Smik/1-1528      KTIKKYQNAIIVEDKEELETEKTIYNERFEHWKHEYYHDKLKFTSNSEEKVMNLASDYVE
Scas/1-1489      KIIKSYQQAVLVEDADELEKEKGLYDERFENWKHDYYKDKLGFKTNQEESVVSLTNNYVE
Skud/1-1527      KTIKKYQNAIIVEDKEELETEKTIYNERFEHWKHEYYRDKLKYASDCEERVTDLSKDYVE

Scer_XRN1/1-1529 GLQWVLYYYYRGCPSWSWYYPHHYAPRISDLAKGLDQDIEFDLSKPFTPFQQLMAVLPER
Sbay/1-1516      GLQWVLYYYYRGCPSWSWYYPHHYSPRISDLVKGLDQDIEFDLSKPFTPFQQLMAVLPER
Smik/1-1528      GLQWVLYYYYRGCPSWSWYYPHHYAPRISDLVKGLDQDIEFDLSKPFTPFQQLMAVLPER
Scas/1-1489      GLQWVLYYYYRGCPSWGWYYHYHYAPRISDLEKGLDQIIKFDKGRPFTPFEQLMAVLPER
Skud/1-1527      GLQWVLYYYYRGCPSWSWYYPHHYAPRISDLVKGLDQDIEFDLSKPFTPFQQLMAVLPER

Scer_XRN1/1-1529 SKNLIPPAFRPLMYDEQSPIHDFYPAEVQLDKNGKTADWEAVVLISFVDEKRLIEAMQPY
Sbay/1-1516      SKNLIPPAFRPLMYDEQSPIHDFYPAEVQLDKNGKTADWEAVVLISFVDEERLIEAMQPY
Smik/1-1528      SKNLIPPAFRPLMYDEQSPIHDFYPAEVQLDKNGKTADWEAVVLISFVDEKRLIEAMQPY
Scas/1-1489      SKNLIPPALRPLMYDEKSPIIDFYPSEVKLDMNGKTADWEAVVLLSFVDEKRLVEAMKPY
Skud/1-1527      SKNLIPPAFRPLMYDEQSPIHDFYPAEVQLDKNGKTADWEAVVLISFVDEKRLIEAMQPY

Scer_XRN1/1-1529 LRKLSPEEKTRNQFGKDLIYSFNPQVDNLYKSPLGGIFSDIEHNHCVEKEYITIPLDSSE
Sbay/1-1516      LAKLSPEEKIRNQFGKDLIYSFNPQVDNIYKSPLGGIFSDIEHNHCVEKEYIAIPLDSSE
Smik/1-1528      LSKLSPEEKNRNQFGKDLIYSFNPQVDNIYKSPLGGIFSDIEHNHCVEKEYITIPLNSSE
Scas/1-1489      LSKFTPEEKLRNSFGTSLKFIFNQQVDTIYKSPLSGIFSDLEHDHCVEREFVEKHIDESE
Skud/1-1527      LSKLSPEEKTRNQFGKDLIYSFNPQVDVIYKSPLGGIFSDIEHNHCVEKEYVTVPLDSSD

Scer_XRN1/1-1529 IRYGLLPNAKLGAEMLAGFPTLLSLPFTSSLEYNETMVFQQPSKQQSMVLQITDIYKTNN
Sbay/1-1516      IRYGLLPDAKLGSEMLAGFPTLLSLPFTNSLEYNETMVFQQPSRQQSMVLKITDIYQTNN
Smik/1-1528      IRYGLLPNAKLGAEMLAGFPTLLSLPFTNSLEYNETMVFQQPSKQQSMVLQITDIYQTNN
Scas/1-1489      FRFGLLPHARTGTELQAGFPTLKSLPFKYQLEYNESLVFQQASRQQSMVLHMDDIYHENN
Skud/1-1527      IRYGLLPNAKLGAEMLAGFPTLLSLPFTNSLEYNETMVFQQPSRQQSMVLQITDIYQMNN

Scer_XRN1/1-1529 VTLEDFSKRHLNKVIYTRWPYLRESKLVSLTDGKTIYEYQESN---DKKKFGFITKPAET
Sbay/1-1516      VTLEDFSKRHLNKIIYTRWPYLRESKLVSLTDGQTVYEYQESN---DKKKFRFVSRPADN
Smik/1-1528      ITLEDFSKRHLNKVIYTRWPYLRESKLVSLTDGKTIYEYQESN---DKKRIRFVSKPTDA
Scas/1-1489      LTLDELSKRYLNQIIYTRWPYLRESKLISITDGETVYERKEHTGADGKVHFKRITRSADN
Skud/1-1527      VTLEDFSKRHLNKVIYTRWPCLRESKLVSLTDGKTIYEYQESN---DKKKTRFVSKPTDT
```

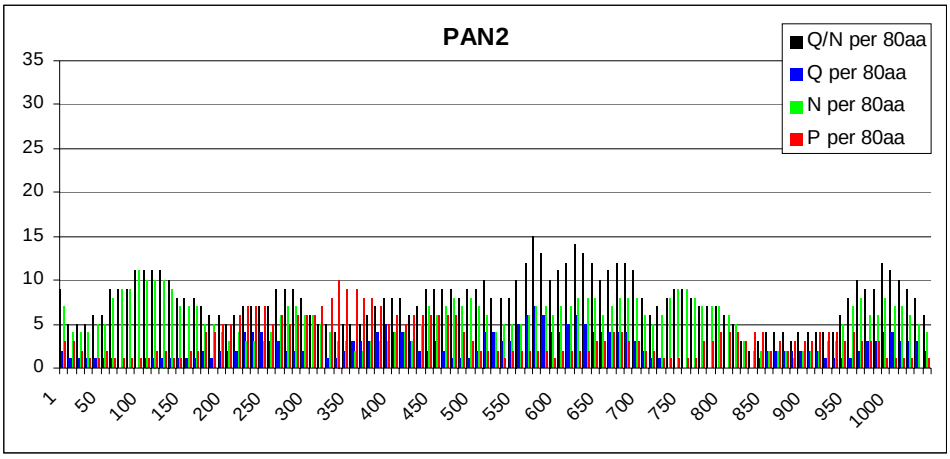
```
Scer_XRN1/1-1529 QDKKLFNSLKNSMLRMYAKQKAVKIGPMEAIATVFPVTGLVRDSDGGYIKTFSPTPDYYP
Sbay/1-1516      QDKKLFNSLKNSMLRNYAKQKAVKLDHVKAIATVCPVTGLARDPEGGYVKTFSCTPDHYP
Smik/1-1528      QDKKLFNSLKNSMLRMYAKQKAVKLGHVEAIATVFPVTGLVRDSDGGYIKTFSSTPDHYP
Scas/1-1489      HDRKQFNSLKHSMLRNYGKQKAVILDKVRAIIEVLPVVGLTRNDDGAYVKTFASQPDYYP
Skud/1-1527      QDKKLFNSLRNSMLRMYAKQKAVKLGPVEAIATVFPVTGLVRDHEGGYVKTFSSTPDYYP

Scer_XRN1/1-1529 LQLVVESVVNEDERYKERGPIPIEEEFPLNSKVIFLGDYAYGGETTIDGYSSDRRLKITV
Sbay/1-1516      LQLVVESIANEDERYKERGPIPIEEEFPLNSKVIFLGDYAYGGETTIDGYSSDRRLKITV
Smik/1-1528      LQLVVESVSNEDERYKERGPIPIEEEFPLKSKVIFLGDYAYGGETTIDGYSSDRRLKITV
Scas/1-1489      LQLMVESITNKDERYIEKPPLPIKEEFPVGSHSIFLGDYAYGGEAFVDGYNSEKRLKLTV
Skud/1-1527      LQLVVESVVNEDERYKERGPIPIEEEFPLDSKVIFLGDYAYGGETTIDGYSSDRRLKITV

Scer_XRN1/1-1529 EKKFLDSEPTIGKERLQMDHQAVKYYPSYIVSKNMHLHPLFLSKITSKFMITDATGKHIN
Sbay/1-1516      EKKFLDSEPTIGKERLQMDHQAVQYYPSYIVSKNMHLHPLFLSKITSKFMISDATGKHIN
Smik/1-1528      EKKFLDSEPTIGKERLQIDHQAVKYYPSYIVSKNLHLHPLFLSKITSKFMITDANGKHIN
Scas/1-1489      TKSFNKTEPTIGKERLAIDHRAVHYIPSFIIAKKLGLHPLFLSRITTKFMLADVSGRHIN
Skud/1-1527      EKKLLDSEPTIGKERLQMDHQAVQYFPSYIVSKNMHLHPLFLSKITSKFMITDATGKHIN

Scer_XRN1/1-1529 VGIPVKFEARHQKVLGYARRNPRGWEYSNLTLNLLKEYRQTFPDFFFRLSKVGNDIPVLE
Sbay/1-1516      VGIPVKFEARHQKVLGYARRNPRGWEYSNLTISVLKEYRQTFPEFFFRLSKVGNDIPALE
Smik/1-1528      VGIPVKFEARHQKVLGYARRNPRGWEYSNLTINVLKEYRQTFPEFFFRLSKVGNDIPVLE
Scas/1-1489      VGIPVKFEAKHEKVLGYARKGIKGWEYSNLTINLITEYKNKFPEFFNKLSKVGNGIPILE
Skud/1-1527      VGIPVKFEARHQKVLGYARRNPRGWEYSNLTINVLKEYRQTFPEFFYRLSKVGNDIPVLE

Scer_XRN1/1-1529 DLFPDTSTKDAMNLLDGIKQWLKYVSSKFIAVSLESDSLTKTSIAAVEDHIMKYAANIEG
Sbay/1-1516      DLYPNTSTKEAMNLLDGVKQWLKYVSSKFIAVSLESDSLTKTSIAAVEDHIMKYAANIEG
Smik/1-1528      DLFPNSSTKEAMNFLDGIKQWLKYVSSKFIAVSLESDSLTKTSIAAVEDHIMKYAANIEG
Scas/1-1489      DLYPELSSSNLAKFLDEIKKWLKFVTEKFVVVSIESDSLTKSSIGIVEEFVQEYTRKSYP
Skud/1-1527      DLFPNISTKEAMNSLDGVKQWLKYVSSKFIAVSLESDSLTKTSIAAVEEHIMKYAANIGG

Scer_XRN1/1-1529 HERKQLAKVPREAVLNPRSSFALLRSQKFDLGDRVVYIQDSGKVPIFSKGTVVGYTTLSS
Sbay/1-1516      HERKQLAKVPREAVLNPRSSFALLRSQKFDLGDRVVYIQDSGKVPIFSKGTVVGYTTLSS
Smik/1-1528      HEKKQLAKVPREAVLNPRSSFALLRSQKFDLGDRVVYIQDSGKVPIFSKGTVVGYTTLSS
Scas/1-1489      TEKKQLANVPRQAILNPRTSVGILRSQRFDLGDRVIYVQDSGKVPLFSRGTVVGYTTLGL
Skud/1-1527      HERKQLAKVPRKAVLNPRSSFALLRSQKFDLGDRVVYIQDSGKVPIFSKGTVVGYTTLSS

Scer_XRN1/1-1529 SLSIQVLFDHEIVAGNNFGGRLRTNRGLGLDASFLLNITNRQFIYHSKASKKALEKKKQS
Sbay/1-1516      SLSIQVLFDHEIVAGNNFGGRLNTNRGLGLDASFLLNITNRQFIYHSKASKKALEKKKQP
Smik/1-1528      SLSIQVLFDHEIVAGNNFGGRLHTNRGLGLDASFLLNVTNRQFIYHSKASKKALEKKKQP
Scas/1-1489      NLSVQVLFDNEIVAGNTFGGRLRTNRGLGLDASFLLNLTNKQFIYHSKASKKALHQKKK-
Skud/1-1527      SLSIQVLFDNEIVAGNDFGGRLRTNRGLGLDASFLLNITNRQFIYHSKASKKVLDKKKQP

Scer_XRN1/1-1529 NNRNNNTKTAHKTPSKQQSEEKLRKERAHDLLNFIKKDTNEKNSESVDNKSMGSQKDSKP
Sbay/1-1516      NNRNNNTAKTARKPPQKQQSEEKLRKERAHDLLNYIKKDTNEKHTEDEKEKKYPSQKNSKP
Smik/1-1528      NSRNNNSKAGHKNAPKLQPEEKLRKERAHDLLNYIKKDSNEKGSEGGDNKNVVPQGDSKP
Scas/1-1489      -APLSQETKQKIIASKREKITELRKIQAHDLLSHITQADNKIVHHN--------EEVDAE
Skud/1-1527      YNKTNNTKAAHKTQPKPQSEEKLRKERAHDLLNYIKKDTNEKEAESEKDKKMTSQKSFKP

Scer_XRN1/1-1529 AKKVLLKRPAQKSSE---NVQVDLANFEKAPLDNPTVAGSIFNAVANQYSDGIGSNLNIP
Sbay/1-1516      AKKTLLKRPSQQNTENKKQVDVNVASSEKPSPNGSTTAGSIFNAVANQYSDGMGNNLNIP
Smik/1-1528      VKKTILKRPAQKTSQG--QQQQVDVTISEKVPLNNPTVAGSIFNAVANQYSDGIGNNLNIP
Scas/1-1489      VKDSTKQTPSE-------GEQSSKAIEEITLPPNTQVANSIFNAVVNQYAT-NGNMEPIA
Skud/1-1527      AKKTLLKRPTQQNSPNSEIKQVDVISSEKASLNDPTVAGSIFNAVANQYSDGMGNHTNIP

Scer_XRN1/1-1529 TPPHPMNVVGGPIPGANDVADVGLPYNIPPGFMTHPNGLHPLHPHQMPYPNMNGMSIPPP
Sbay/1-1516      IPPH-------PLPGANDTGDVGLPYNIPPGFVAHPGGPHPLPLHQMPYPNMNGIPIP--
Smik/1-1528      VAPHLMSVAGGPIPGANDAVDVGLPYNIPPSFIAHPNGPHPLPLHQMPYPNMNGISIP--
Scas/1-1489      RPPN-----------NNGPVVAGLPYHIPAGHNGSIPPGLGTQQHPMGYPNMG-MPMP--
Skud/1-1527      IPPHPMAMVGGPLLGANDTPGVALPYNIPPNFIAHPNGPHPLTSQQMPYPNMNGVPIP--

Scer_XRN1/1-1529 APHGFG--------QPISFPPPPPMTNVSDQGSRIVVNEKESQDLKKFINGKQHSNGSTI
Sbay/1-1516      PPHGFG--------QPMPFPHPHAIVNNSDQGN-IAVNEKESQDLKEFINGKQNSNGSSV
Smik/1-1528      QSHGFS--------QPVPFPPPPPMTSFSEQGTQIVVNEKESQDLKNFITGKQNGNGSGI
Scas/1-1489      PPHPYNGNMVTVNQPPMMMQPPHVQHGHPNAHSNVKVDEKGSDALKQFVK----------
Skud/1-1527      PPRGFG--------QPIPFSHPHHMANLPDQGSRIVVNEKESQDLKKIINGAHDINGLNA

Scer_XRN1/1-1529 GGETKNSRKGEIKPSSGTNSTECQSPKSQSNAADRDNKKDEST
Sbay/1-1516      REETTGSRETEPNTTSNMSTTDSQFSKSGNDVVGRDI------
Smik/1-1528      SEQAKGSRKQETKPSSGTNTTECLSPKSGHNTAKLDSKKEESH
Scas/1-1489      ----DHSSDKELKSEKSKSSKAIKIKQRTTNAAPV--------
Skud/1-1527      DKIFKRSEQNETKTTSSAMAAEFQSPKPGNNGAEREKKSD---
```



Pan2

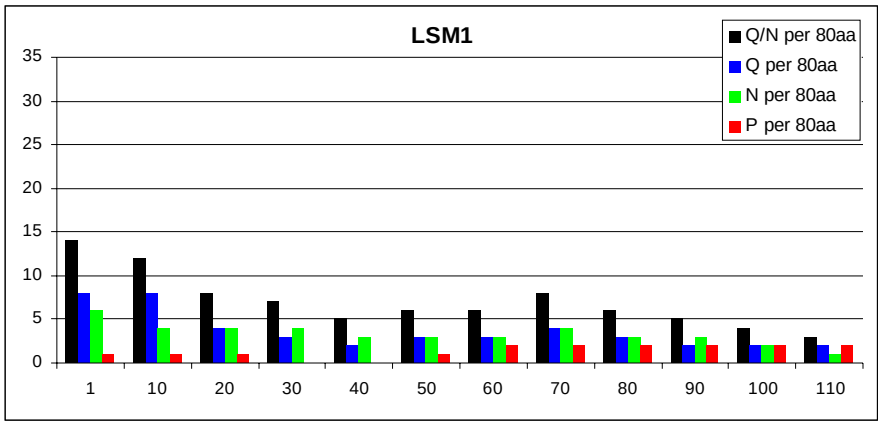

SGD_PAN2/1-1116 MNNWQHFFNNPVDLSEHLKKPYFRFDNRDKEITAISFDEKANLIWSGDSYGCISSYDPTF
Sbay/1-1116 MNNWQHLFNNPVDLSEHLKKPYFRFDNRDKEVTTISFDEKANLVWSGDSYGCISSYDPNF
Smik/1-1116 MNNWQHFFNNPVDLSEHLKKPYFRFDNRDKEVTAISFDEKANLIWSGDSYGCISSYDPTF
Spar/1-1116 MNNWQHFFNNPVDLSEHLKKPYFRFDNRDKEVTAISFDEKANLIWSGDSYGCISSYDPTF
Sklu/1-1000 MNNWQLSFQSPVDLTDHLRKPYIMYDNREKSITKIVFDEEANLLWAGDTYGRVSSYDPSY
SGD_PAN2/1-1116 QLYTRYRGHIGGNSVKDILSHRDGILSISEDSLHFANRRGVTKLNLTSIDIAAFSELNTM
Sbay/1-1116 QLYTRYRGHIGGSSVKDILSHRDGILSISEDSLHFANRRGVTKLNLTSIDIAAFSELNTM
Smik/1-1116 QLYTRYRGHIGGNSVKDILSHRDGILSISEDSLHFANRRGVTKLNLTSIDIAAFSELNTM
Spar/1-1116 QLYTRYRGHIGGNSVKDILSHRDGILSISEDSLHFANRRGVTKLNLTSIDIAAFSELNTM
Sklu/1-1000 SLYTRHTAHLGAVPVADMLSHKSGILSLGPDSLNFANRRGVTQLNLTSADIAQLCDMKSM
SGD_PAN2/1-1116 CYSPHSLKNNIYCGGDNTNWGIASIDLNRGCLDSLLNYSSKVKLMCSNNKVLSIGRQTGT
Sbay/1-1116 CYAPHSLQNNIYCGGDNTNWGIASIDLNKGCLDSLLNYSSKVKLLCSNNKILSVGRQTGS
Smik/1-1116 CYTPHSLQNNIYCGGDNTNWGIASIDLNKGTLDSLLNYSFKVKLLCSNNKILSIGRQTGS
Spar/1-1116 CYSPHSLKNNIYCGGDNTNWGIASIDLNKGCLDSLLNYSSKVKLLCSNNKVLSIGRQTGT
Sklu/1-1000 AYSSN-VQNHVYCGGSNIASGIINIDLNKGCLASTVDYASKIKLMCSNNRIIAIGKQSGS
SGD_PAN2/1-1116 VDLLDPTSNRTIKSFNAHSASISAMDLRDNTLVTVGKSKRFYNLYADPFVNVYDLRTMRQ
Sbay/1-1116 VDLIDPTSNRTIKSFNAHSASISGMDLRDNTLVTVGKSKRFYNLYADPFVNVYDLRTMRQ
Smik/1-1116 VDLLDPASNRTIKSFNAHSASISAMDLRDNTLVTVGKSKRFYNLYADPFVNVYDIRTMRQ
Spar/1-1116 VDLLDPTSNRTIKSFNAHSASISAMDLRDNTLVTVGKSKRFYNLYADPFVNVYDLRTMRQ
Sklu/1-1000 VDLLDPSSNQLIRSFAGHSSVITGMDSVDYTLVTVGKSKKFNMLFSDQFVNVYDLRIMKQ
SGD_PAN2/1-1116 LPPVSFSKGTTMG--------SGGADFVQLHPLLPTVMIVASSSGSFDFIDLSNPTLRTQ
Sbay/1-1116 LPPVSFSKGTTMG--------SGGADFVQLHPLLPTVMIVGSSSGSFDFIDLSNPTLRTQ
Smik/1-1116 LPPVSFSKGTTMG--------SGGADFVQLHPLLPTVMIVASSSGSFDFIDLSNPTLRTQ
Spar/1-1116 LPPVSFSKGTTMG--------SGGADFVQLHPLLPTVMIVASSSGSFDFIDLSNPTLRTQ
Sklu/1-1000 LPPVSFSKSSDFAPSANSGTFSMGADFVQLHPVLPTVMAVASASGSFDFVDLANPTIRTQ
SGD_PAN2/1-1116 YVHPCQSIKKLCLSPNGDVLGILEADNHLDTWRRSSNNMGMFTNTPEMLAYPDYFNDITS
Sbay/1-1116 YVHPCQSIKEFSLSPNGDVLGILEADNHLDTWRRSSNNMGMFTNTPEMLAYPDYFNDVTT
Smik/1-1116 YVHPCQSIKKFSLSPNGDVLGILEADNHLDTWRRSSNNMGMFTNTPEMLSYPDYFNDITS
Spar/1-1116 YVHPCQSIKKLSLSPNGDVLGILEADNHLDTWRRSSNNMGMFTNTPEMLAYPDYFNDITS
Sklu/1-1000 YCHPCQSISNFALSPSGDYIAFVEQDNNVNMWSRS-NGMTGFTNSSALLEYPDFVDDGPL
SGD_PAN2/1-1116 DGPISVDDETYPLSSVGMPYYLDKLLSAWPPVVFKSEGTIPQLTGKSPLPSSGKLKSNLA
Sbay/1-1116 DSPISIDDETYPLSSVGMPYYLDKLLSAWPHVVFKSEGTIPQLSGKPHLPSSGKLKGNTA
Smik/1-1116 DAPISVDDDTYPLSSVGMPYYLDKLLSAWPHVIFKSEGTIPQLVGKAPLPSSGKLKSNTA
Spar/1-1116 DAPISVDDETYPLSSVGMPYYLDKLLSAWPHVVFKSEGTIPQLTGKAPLPSSGKLKSNTA
Sklu/1-1000 PSNNHIDDFNYSLSSVGLPYYSEKLLSAWPQTVFRSSGTIPKKIDIN-IPISTRPSS---
SGD_PAN2/1-1116 VISSQNEKLSTQEFPLLRYDRTKYGMRNAIPDYVCLRDIRKQITS----GLETSDIQTYT
Sbay/1-1116 LISSQNEKLSTQEFPLLRYDRTKYGMRNTIPDYVCLRDLRKQITA----GLETSDIQVYA
Smik/1-1116 VISSQNEILNSSEFPLLRYDRTKYGMRNTVPDYVCLRDLRKQITT----GIETSDIQTYT
Spar/1-1116 VISSQNEKLSTQEFPLLRYDRTKYGMRNAVPDYVCLTDLRKQITT----GLETSDIQTYT
Sklu/1-1000 SRSSQGLNICTPHYPLLPYNRLKYGHRNVTPPYRSLRERKKKFLTTDEDGTDKEELMKYK
SGD_PAN2/1-1116 SINKYEVPPAYSRLPLTSGRFGTDNFDFTPFNNTEYSGLDPDVDNHYTNAIIQLYRFIPE
Sbay/1-1116 ATNKYEVPPAYSRLPLTTGRFGTDNFDFTPFNNTEYSGLDPDVDNHYTNAIIQLYRFIPE
Smik/1-1116 AINKYEVPPAYSRLPLTSGRFGTDNFNFTPFNNTEYSGLDPDVDNHYTNAIIQLYRFIPE
Spar/1-1116 AINKYEVPPAYSRLPLTSGRFGTDNFDFKPFNNTEYSGLDPDVDNHYTNAIIQLYRFIPE
Sklu/1-1000 LDNDSEVPPAYSKLQMTYGKYGVEDFDFKAFNNTPFSGLETEIDNVYTNSVLQLYRFVPE
SGD_PAN2/1-1116 MFNFVVGCLKDENFET-TLLTDLGYLFDMMERSHGKICSSSNFQASLKSLTDKRQLE--N
Sbay/1-1116 MFNFVVGCLKDENFET-ALLTDLGYLFDMMDRSNGKICSSTNFQASLKSLTDKMYLD--N
Smik/1-1116 MFNFVVGCLKDENFKT-TLLTDLGYLFDMMDRSQGKICSSSNFQASLKSLTDKRELK--S
Spar/1-1116 MFNFVVGCLKDENFET-TLLTDLGYLFDMMERSHGKICSSSNFQASLKSLTDQRQLH--N
Sklu/1-1000 VYNFVVSCLKDENLSQNSLLTELGFLYDMMTRSEGTICRSLNFQEVLSSITRAREMGLIT
SGD_PAN2/1-1116 GEPQEHLEEYLESLCIRESIEDFNSSESIKRN----MPQKFNRFLLSQLIKEEAQTVN--
Sbay/1-1116 NVPQEHLEEYLESLCIGESIEGFSSSESIKRN----MPQKFNRFLLSQLIKEEAQTVN--
Smik/1-1116 NVSQEHLEEYLESLCIGESIEDFSSSESIKRN----MPQKFNRFLLSQLIKEEAQTVN--
Spar/1-1116 DAPQEHLEEYLESLCIGESIEDFNSSESIKRN----MPQKFNRFLLSQLIKEEAQTVN--
Sklu/1-1000 DSPAGTLMQQFGGLSVNNNESCSQSSTSFNYKPYYTTSQVFNDFLLSRLISEEVRKKINT
SGD_PAN2/1-1116 -HNITLNQCFG--LETEIRTECSCDHYDTTVKLLPSLSISGINKTVIKQLNKKSNGQNIL
Sbay/1-1116 -HNITLNQCFG--LETEIRTHSGCDHYDTAVKLLPSLSIAGINKTLIKQLNKKNNGQNIL
Smik/1-1116 -HNITLNQCFG--LETEIKTECSCDHYETTVKLLPSLSISGINKTAIKQLNKKSNGLNIL
Spar/1-1116 -HNITLNQCFG--LETEIRTECSCDHYDTTVKLLPSLSISGINKTVIKQLNKKSNGQNIL
Sklu/1-1000 TQSITLEELFGIHLETDSRSLSTCNTHERQSNIVPTLSVTSPMSNNAKYSNKKLNNQIIL

```
SGD_PAN2/1-1116 PYIEYAMKNVTQKNSICPTCGKTETITQECTVKNLPSVLSLELSLLDTEFSNIRSSKNWL
Sbay/1-1116     PYIEYAMKNSIQKNNICPICGKNDIITQENTVKNLPSVLSLELSLLDTELSNIRSSKNWL
Smik/1-1116     PYIEYAMKSVIQKNSICPVCGKAETITQEHTVKNLPSVLSLELSLLDTELSNIRPSKNWL
Spar/1-1116     PYIEYAMKSVTQRNSICPTCGKTEIITQECTVKNLPSVLSLELSLLDAELSNIRSSKNWL
Sklu/1-1000     PYIESSMNRIKQKRTMCEKCHRYEIVESEKTVRNLPPLLSLNINLSSEEWIVAKTVRNWL

SGD_PAN2/1-1116 TSEFYGSIIKNKAVLRSTASELKGTSHIFKYELNGYVAKITDNNNETRLVTYVKKYNPKE
Sbay/1-1116     INEFYGSIIKNKAILRPTASELKGTGHIFKYELNGYVAKITDNNNETRLVTYVRKYDSRE
Smik/1-1116     ISEFYGSIIKNKAFLRASASELKGTGHIFKYELNGYVAKITDNNNETRLVTYIRKYDPNK
Spar/1-1116     TSEFYGSIIKNKAVLRSTASELKGTSHIFKYELNGYVAKITDNNNETRLVTYIRKYNPKE
Sklu/1-1000     VKEFYATISKDRPILKLQPTDLKTTNAIFKYELNAYVARISDDISESHLVTYAKVYDPKS

SGD_PAN2/1-1116 NCFKWLMFNDYLVVEITEEEALKMTYPWKTPEIIIYCDAEELRKPFFSVDTYSINYDILF
Sbay/1-1116     NNFKWLMFNDYLVVEISEEEALKMSYPWKTPEIIIYCDAEELRKPFFSVDTYSINYDILY
Smik/1-1116     NDFKWLMFNEYLVVEISEEEALKMTYPWKTPEIIIYCDAEELRKPFFSVDTYSINYDILF
Spar/1-1116     NGFKWLMFNDYLVVEISEEEALKMSYPWKTPEIIIYCDAEELRKPFFSVDTYSINYDILF
Sklu/1-1000     YTYKWYMFNDYLVAEVDEEEALNISYWWKTPEIIIYSDTEELRKPFISVSNYPISDSILY

SGD_PAN2/1-1116 RDYFANGIRDTARREYKLLTHDEAPKSGTLVAIDAEFVSLQSELCEIDHQGIRSIIRPKR
Sbay/1-1116     RDYFANGIRDSARSEYKLLTHDEAPKSGSLVAIDAEFVSLQSELCEIDHQGVRSVIRPKR
Smik/1-1116     RDYFANGIRDTARSEYKLLTHNEAPKSGTLVAIDAEFVSLQSELCEIDHQGIRSIVRPKR
Spar/1-1116     RDYFANGIRNTARREYKLLTHDEAPKSGTLVAIDAEFVSLQSELCEIDHQGIRSIIRPKR
Sklu/1-1000     RDYFANGIRANVKKEYELLADTEAPLAGSVVAIDAEFVVINEERVEISCKGIKTLIKPKK

SGD_PAN2/1-1116 TALARISIIRGEEGELYGVPFVDDYVVNTNHIEDYLTRYSGILPGDLDPEKSTKRLVRRN
Sbay/1-1116     TALARISIIRGEEGDLYGIPFVDDYVVNTNHIEDYLTRYSGILPGDLDPEKSTKRLVRRN
Smik/1-1116     TALARISIIRGEEGELYGVPFVDDYVVNTNHIEDYLTRYSGILPGDLDPTKSTKRLVERN
Spar/1-1116     TALARISIIRGEEGELYGVPFVDDYVVNTNHIEDYLTRYSGILPGDLDPEKSTKRLVKRN
Sklu/1-1000     TALARISVLRGEEGDKFGVPFIDDYVVNTNRIENYLTKYSGIEPRX--------------

SGD_PAN2/1-1116 VVYRKVWLLMQLGCVFVGHGLNNDFKHININVPRNQIRDTAIYFLQGKRYLSLRYLAYVL
Sbay/1-1116     VVYRKVWLLMQLGCIFVGHGLNNDFKHININVPRNQIRDTAIYFLQGKRYLSLRYLAYVL
Smik/1-1116     VVYRKVWLLMQLGCVFVGHGLNNDFKHININVPKNQIRDTAIYFLQGKRYLSLRYLAYVL
Spar/1-1116     VVYRKVWLLMQLGCVFVGHGLNNDFKHININVPRNQIRDTAIYFLQGKRYLSLRYLAYVL
Sklu/1-1000     ------------------------------------------------------------

SGD_PAN2/1-1116 LGMNIQEGNHDSIEDAHTALILYKKYLHLKEKAIFEKVLNSVYEEGRAHNFKVPETSKG
Sbay/1-1116     LGMNIQEGNHDSIEDAHTALILYKKYLDLKRKAVFEKVLNSVYEEGRAHNFKVPETPKT
Smik/1-1116     LGMNIQEGNHDSIEDAHTALILYKKYLDLKQKAIFEKVLNSVYEEGRAHNFKVPENLKG
Spar/1-1116     LGMNIQEGNHDSIEDAHTALILYRKYLDLKQKTIFDKVLNSVYEEGRAHNFKVPETVKG
Sklu/1-1000     ------------------------------------------------------------
```



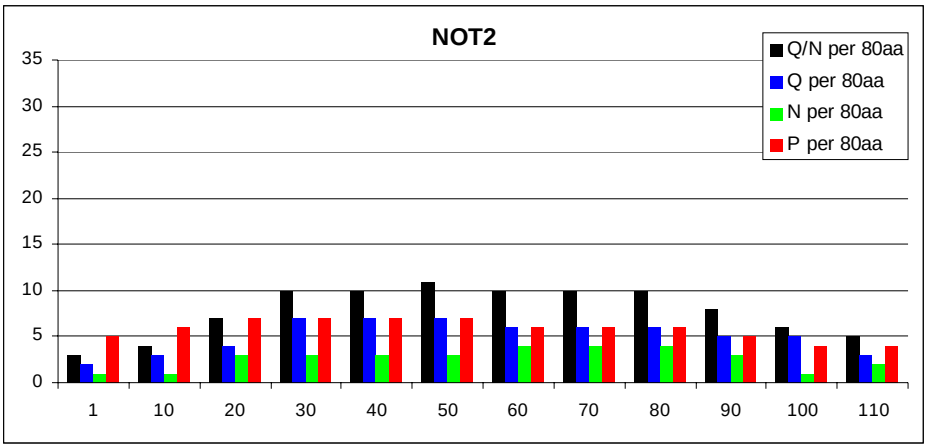
Lsm1

```
Scer_LSM1/1-173 MSANSKDRNQSNQDAKRQQQN--FPKKISEGEADLYLDQYNFTTTAAIVSSVDRKIFVL
Sbay/1-173      MPANTRERTQSGQDNKRQPQH--FSKKISEGEADLYLDQYNFTTTAAIVSSVDRKIFVL
Smik/1-173      MSANSKDRSQSNQDIKRQQQI--ASKKISEGEADLYLDQYNFTTTAAIVSSVDRKIFVL
Spar/1-173      MSANSKDRNQSNQDTKRQQQN--FSKKISEGEADLYLDQYNFTTTAAIVSSVDRKIFVL
Scas/1-176      MSSPAPKVPQQQQHAKPQGKQGSKASKKLSEGEADLYLDQYNFTTTAAIVSSVDRKIFVL
Sklu/1-159      MS----------QEKEPN-------AAKITEGEADLYLDQYNFTTTAAIVGSVDRKIFVL
Skud/1-173      MSTNSRERNQPNQDNKRQQQS--FSKKISEGEADLYLDQYNFTTTAAIVSSVDRKIFVL

Scer_LSM1/1-173 LRDGRMLFGVLRTFDQYANLILQDCVERIYFSEENKYAEEDRGIFMIRGENVVMLGEVDI
Sbay/1-173      LRDGRMLFGVLRTFDQYANLILQDCVERIYFSEENKYAEEDRGIFMIRGENVVMLGEVDI
Smik/1-173      LRDGRMLFGVLRTFDQYANLILQDCVERIYFSEVNKYAEEDRGIFMIRGENVVMLGEVDI
Spar/1-173      LRDGRMLFGVLRTFDQYANLILQDCVERIYFSKENKYAEEDRGIFMIRGENVVMLGEVDI
Scas/1-176      LRDGKLLFGILRTFDQYANLILQHCVERIYLTEENKYAENELGIFMIRGENVVMLGEVDI
Sklu/1-159      LRDGRMLFGVLRTFDQYANLILQHCVERIYITDENKYGECDRGVFMVRGENVVMLGEVDI
Skud/1-173      LRDGRMLFGVLRTFDQYANLILQDCVERIYFSEENKYAEEDRGIFMIRGENVVMLGEVDI

Scer_LSM1/1-173 DKEDQPLEAMERIPFKEAWLTKQKNDEKRFKEETHKGKKMARHGIVYDFHKSDMY
Sbay/1-173      DKEDQPLDAMERIPFKEAWQIKQQNDEKRFKEETHRGKEMARHGIVYDFHKSDMY
Smik/1-173      DKEDQPLEAMERISFKEAWQIKQKNDEKRFKEETHKGKKMARHGIVYDFHKSDMY
Spar/1-173      DKEDQPLEAMERIPFKEAWQIKQKNDEKRFKEETHKGKKMARHGIVYDFHKSDMY
Scas/1-176      DKEDKPLESMERIPFKDASRIKKLNDEKKFKTETKKGKEYARYGLIHDFHKSDMY
Sklu/1-159      DKEDKPLETMERVPFEEASLIKQQRDDGKFKRETRKGKEFAKLGLIHDFHKADMY
Skud/1-173      DKEDQPLEIMERIPFREAWQIKQQNDEKRFKEETHKGKEMARHGIVYDFHRSDMY
```



Not2

```
Scer_NOT2/1-192 ----------------------------MEKFGLKALVPLLKLEDKELSSTYDHSMTLG
Sbay/1-221      MKVKRNKHFQSLRTNSADKRRLKYITIYIMDKFGLQALVPLLKQEDKELSSTYDHSMTLG
Smik/1-192      ----------------------------MEKFGLQALVPLLKLEDKELSSTYDHSMTLG
Spar/1-192      ----------------------------MENFGLKALVPLLKLEDKELSSTYDHSMTLG
Scas/1-192      ----------------------------MDKFGLKALVPLLRQDDKELSKPYDHSMTLG
Sklu/1-192      ---------------------------MAVEKYGLKALLPLIKQETG--NAEYDPAMTMG
Skud/1-192      ----------------------------MEKFGLRALVPLLKLEDKELSSTYDHSMTLG

Scer_NOT2/1-192 ADLSSMLYSLGIPRDSQ-DHRVLDTFQSPWAETSRSEVEPRFFTPESFTNIPGVLQSTVT
Sbay/1-221      ADLSSMLYSLGIPRDSQ-DHRVLDTFQSPWAETSRSEVEPRFFTPESFTNIPSVLQSNVT
Smik/1-192      ADLSSMLYSLGIPRDSQ-DHRVLDTFQSPWAETSRSEVEPRFFTPESFTNISGVLQSNVT
Spar/1-192      ADLSSMLYSLGIPRDSQ-DHRVLDTFQSPWAETSRSEVEPRFFTPESFTNIPGVLQSNVT
Scas/1-192      ADLSSMIHSLGIPRDNT-RHRVLDTFQSPWAETSRSEVEPTFFIPDSFVGISGVLQSPVT
Sklu/1-192      MDLSSMLHSLGIDSNQEYKTRVLDTFPSPWVETSRSEVEPKFFTPESFVNIPNVLQSP-S
Skud/1-192      ADLSSMLYSLGIPRDSQ-DHRVLDTFQSPWAETSRSEVEPRFFTPESFTNIPGVLQSNVT

Scer_NOT2/1-192 PPCFNSIQNDQQRVALFQDETLFFLFYKHPGTVIQELTYLELRKRNWRYHKTLKAWLTKD
Sbay/1-221      PPCFNSIQNDQQRVALFQDETLFFLFYKHPGTVIQELTYLELRKRNWRYHKTLKAWLTKD
Smik/1-192      PPCFNSIQNDQQRVALFQDETLFFLFYKHPGTVIQELTYLELRKRNWRYHKTLKAWLTKD
Spar/1-192      PPCFNSIQNDQQRVALFQDETLFFLFYKHPGTVIQELTYLELRKRNWRYHKTLKAWLTKD
Scas/1-192      PPRFDSAQNDQQRVALFQDETLFYLFYKHPGTVLQELTYLELRNRNWRYHTTLRAWLTKD
Sklu/1-192      PPSFNSVQQDHPKIALFQDETLFYLFYKHPGTVVQEITYLELRKRNWRYHKTLKVWLTKD
Skud/1-192      PPCFNSIQNDQQRVALFQDETLFFLFYKHPGTVIQELTYLELRKRNWRYHKTLKAWLTKD

Scer_NOT2/1-192 PMMEPIVSADGLSERGSYVFFDPQRWEKCQRDFLLFYNAIM
Sbay/1-221      PMMEPIVSADGLSERGSYVFFDPQRWEKCQRDFLLFYNAIM
Smik/1-192      PMMEPIVSADGLSERGSYVFFDPQRWEKCQRDFLLFYNAIM
Spar/1-192      PMMEPIVSADGLSERGSYVFFDPQRWEKCQRDFLLFYNAIM
Scas/1-192      PLMEPEVSADGLSERGSYVFFDPQRWEKCQKDFLLFYNAIM
Sklu/1-192      PMMEPVVSQDGLSERGSYVFFDPQRWEKCQRDFVLFYNAIM
Skud/1-192      PMMEPIVSADGLSERGSYVFFDPQRWEKCQRDFLLFYNAIM
```



Edc2

```
Scer_EDC2/1-146 ----------------------------MGSETKHSAKVKIVTRESPPSAKEHMRPTKTQ
Sbay/1-147      ----------------------------MGPDTKHSAKAKILTKKSSPSTKEHIRPRKPQ
Smik/1-173      MSSRIITYHVRHSQTKTFLYNQFQISLKMISDTKHSSKAKIFTKKNLPPTMEYIRPSMPQ
Spar/1-145      ----------------------------MGPDTKNSATAKILTGESPPSAKEHMRPTKPQ
Skud/1-145      ----------------------------MGPDTKHFANIKILNKKSSPPTKEHIRPRKPQ

Scer_EDC2/1-146 ILVPPTQSLPNGKKPNFGKSTKQRREPRERTSKTGHEDDKAT-MVTVNIDAFLHDKAPKK
Sbay/1-147      VAVPPAQNLPNGDRPNFGSAIKQRREPRERATRVRHEDSEATTMVTVNIDGFLSDKTPKK
Smik/1-173      ISVPPAQNLPNGEKPNFGRSTKQRCEPRKGTSRTRRGDEGAT-MVTVNIDSFLCDKAPKK
Spar/1-145      VLVPPAQSLPNGEKPNFGKSTKQRREPRERTSKTRHGDDKAT-MVTVNIDAFLYDKAPKK
Skud/1-145      IIVPPAQNLPNGEKPNFGKSAKQHREPSERTSKARHDDAKAT-MVTVNIDAFLSDKSPKK

Scer_EDC2/1-146 KSCKYKKKKTRQYQDRAAASIDSKPHVAGHTAFAGASFTTDIPHEAALPKPSFV
Sbay/1-147      KSSKNKKKKSRHPQGAAAVTISSKPQEATHTVFAGASFTTDIPHETALPKPSFV
Smik/1-173      KSSKNKKKKSRHHQNTEAS-IDAKPHAVAHTVFAGASFTTDIPHEAALPKPSFV
Spar/1-145      KSCKYKKKKSRQYQDTAAP-IDSKPHAAGHTAFAGASFTTDIPHEAALPKPSFV
Skud/1-145      NSSKNKKKKSKHHHNAAEV-VDSKPHGAANAVFAGASFTTDIPHEAALPKPSFV
```

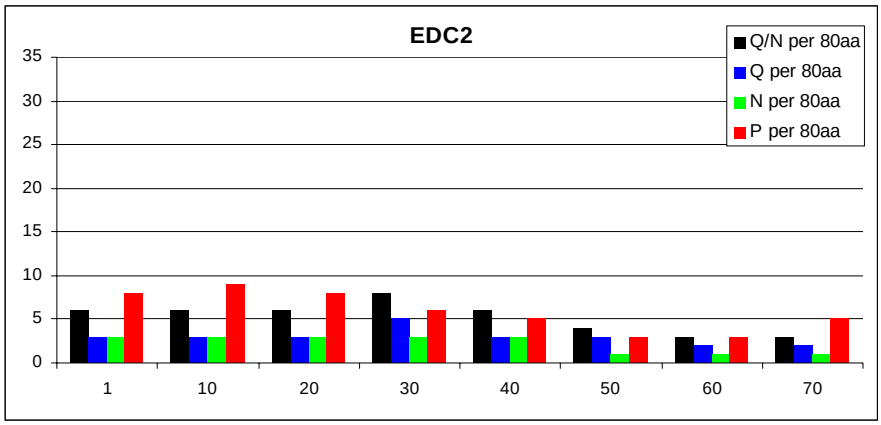


Fig. S7

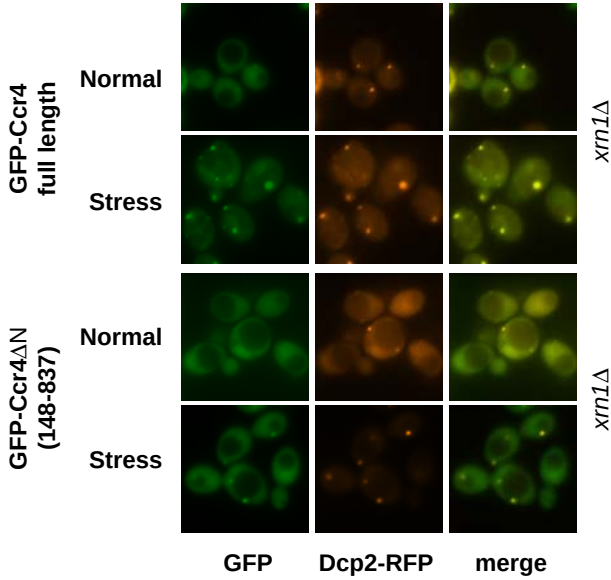


Fig. S7 Localization of GFP-Ccr4, GFP-Ccr4ΔN and Dcp2-RFP in *xrn1*Δ cells (Y14540) during log phase growth and 30 min after hypo-osmotic shock. GFP tagged proteins in green and Dcp2-RFP in red in merged images.

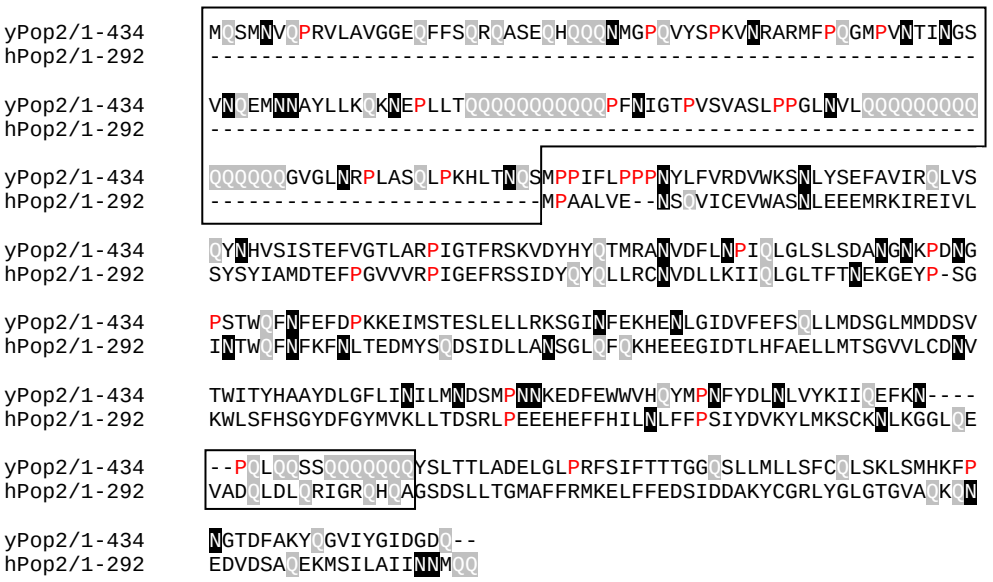
Fig. S8 Alignments of Ccr4p, Pop2p and Dhh1p with their human homologs

Alignments of *S. cerevisiae* Ccr4p **(A)**, Pop2p **(B)** and Dhh1p **(C)** with their human homologs show that the Q/N-rich regions in the N- and C-termini of the budding yeast proteins are absent in their human counterparts.

A. Budding yeast Ccr4p and human Ccr4 (CNOT6)

```
yCcr4/1-837     MNDPSLLGYPNVGPQQQQQQQQQQHAGLLGKGTPNALQQQLHMNQLTGIPPPGLMNNSDV
hCcr4/1-557     ------------------------------------------------------------

yCcr4/1-837     HTSSNNNSRQLLDQLANGNANMLNMNMDNNNNNNNNNNNNNNGGGSGVMMNASTAAVNS
hCcr4/1-557     ------------------------------------------------------------

yCcr4/1-837     IGMVPTVGTPVNINVNASNPLLHPHLDDPSLLNNPIWKLQLHLAAVSAQSLGQPNIYARQ
hCcr4/1-557     ------------------------------------------------------------

yCcr4/1-837     NAMKKYLATQQAQQAQQQAQQQAQQQVPGPFGPGPQAAPPALQPTDFQQSHIAEASKSLV
hCcr4/1-557     ------------------------------------------------------------

yCcr4/1-837     DCTKQALMEMADTLTDSKTAKKQQPTGDSTPSGTATNSAVSTPLTPKIELFANGKDEANQ
hCcr4/1-557     ------------------------------------------------------------

yCcr4/1-837     ALLQHKKLSQYSIDEDDDIENRMVMPKDSKYDDQLWHALDLSNLQIFNISANIFKYDFLT
hCcr4/1-557     -----MPKEKYEPPDPRRMYTIMSSEEAANGKKSHWAELEIS-GKVRSLSASLWSLTHLT

yCcr4/1-837     RLYLNGNSLTELPAEIKNLSNLRVLDLSHNRLTSLPAELGSCFQLKYFYFFDNMVTTLPW
hCcr4/1-557     ALHLSDNSLSRIPSDIAKLHNLVYLDLSSNKIRSLPAELGNMVSLRELHLNNNLLRVLPF

yCcr4/1-837     EFGNLCNLQFLGVEGNPLEKQFLKILTEKSVTG-LIFYLRDNRPEIPLPHERRFIEINTD
hCcr4/1-557     ELGKLFQLQTLGLKGNPLTQDILNLYQEPDGTRRLLNYLLDN-------LSGTAKRITTE

yCcr4/1-837     GEPQREYDSLQQSTEHLATDLAKRTFTVLSYNTLCQHYATPKMYRYTPSWALSWDYRRNK
hCcr4/1-557     QPPPRSWIMLQEPDRTRPTALFS----VMCYNVLCDKYATRQLYGYCPSWALNWDYRKKA

yCcr4/1-837     LKEQILSYDSDLLCLQEVESKTFEEYWVPLLDKHGYTGIFHAKARAKTMHSKDSKKVDGC
hCcr4/1-557     IIQEILSCNADIVSLQEVETEQYYSFFLVELKERGYNGFFSPKSRARTMSEQERKHVDGC

yCcr4/1-837     CIFFKRDQFKLITKDAMDFSGAWMKHKKFQRTEDYLNRAMNKDNVALFLKLQH------I
hCcr4/1-557     AIFFKTEKFTLVQKHTVEFN--QLAMANSEGSEAMLNRVMTKDNIGVAVLLELRKESIEM

yCcr4/1-837     PSG--------DTIWAVTTHLHWDPKFNDVKTFQVGVLLDHLETLLKEETS---HNFRQD
hCcr4/1-557     PSGKPHLGTEKQLILVANAHMHWDPEYSDVKLVQTMMFLSEVKNIIDKASRNLKSSVLGE

yCcr4/1-837     IKKFPVLICGDFNSYINSAVYELINTGRVQIHQEGNGRDFGYMSEKNFSHN---------
hCcr4/1-557     FGTIPLVLCADLNSLPDSGVVEYLSTGGVETNHKDFKELRYNESLTNFSCHGKNGTTNGR

yCcr4/1-837     ----LALKSSYNCIGELPFTNFTPSFTDVIDYIWFSTHALRVRGLLGEVDPEYVSK--FI
hCcr4/1-557     ITHGFKLQSAYESG-LMPYTNYTFDFKGIIDYIFYSKPQLNTLGILGPLDHHWLVENNIS

yCcr4/1-837     GFPNDKFPSDHIPLLARFEFMKTNTGSKKV-------
hCcr4/1-557     GCPHPLIPSDHFSLFAQLELLLPFLPQVNGIHLPGRR
```

B. Budding yeast Pop2p and human Pop2 (CNOT8)



C. Budding yeast Dhh1p and human Dhh1 (DDX6)

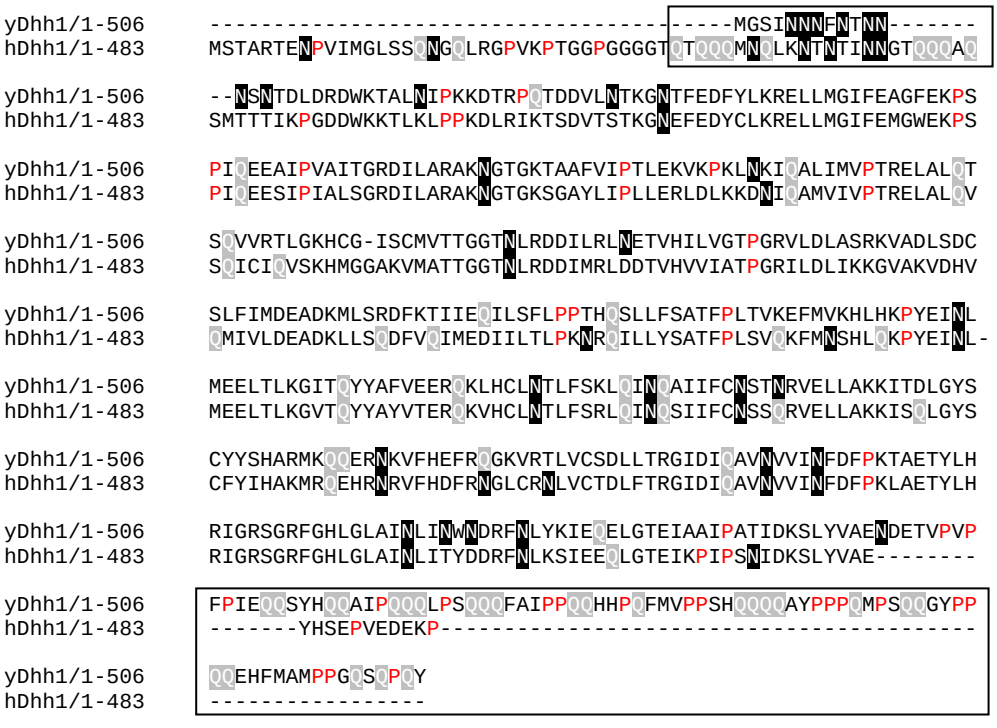


Table S1. Strains used in this study

Strain	Genotype	Reference
BY4741	*MAT***a** *his3Δ1 leu2Δ0 met15Δ ura3Δ0*	Brachmann et al., 1998
MCY4	*MAT***a** *ade1-101 his3-1 leu2-3,112 trp1-289 ura3-52 LEU2-GAL1-LSM4 GAL*$^+$	Cooper et al., 1995
MCY5	*MAT* **a**/*a ADE/ade1-101 HIS3/his3-1 TRP1/trp1-289 ura3-52/ura3-52 leu2-3,112/leu2-3,112 LSM4/lsm4Δ::LEU2, GAL*$^+$	Cooper et al., 1995
MRY71	*MAT***a** *ade1-101 his3-1 leu2-3,112 trp1-289 ura3-52 lsm4Δ::LEU2 GAL*$^+$ *[pUSS1]*	This work (derived from MCY5)
MRY73	*MAT***a** *ade1-101 his3-1 leu2-3,112 trp1-289 ura3-52 lsm4Δ::LEU2 GAL*$^+$ *[pMC12]*	This work (derived from MCY5)
MPS7	*MAT***a**, *ade2-1 his3-11,-15 leu2-3,112 trp1Δ ura3-1 lsm8Δ::TRP1 LSM7-13myc:HphMX6* [pAEM76 (pBM125, P_{GAL1}-*HA-LSM8*)]	Spiller et al., 2007a
Y10387	*MATa leu2Δ0 lys2Δ0 ura3Δ0 ccr4Δ::kanMX4*	Brachmann et al., 1998
Y14540	*MATa his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 xrn1Δ::kanMX4*	Brachmann et al., 1998

Table S2. Plasmids used in this study

Plasmid	Description	Reference
pGFP-N-FUS	*CEN6*, *URA3* shuttle vector with MCS for N-terminal GFP tagging under control of *MET25* promoter	Niedenthal et al., 1996
pGFP-C-FUS	*CEN6*, *URA3* shuttle vector with MCS for C-terminal GFP tagging under control of *MET25* promoter	Niedenthal et al., 1996
pUSS1	1.7 kb *Hind*III genomic fragment with *LSM4* ORF, promoter and terminator sequences cloned into pFL39 (*CEN*6, *TRP1*)	Cooper et al., 1995
pMC12	Modified pFL39 (*CEN6*, *TRP1*) with a 1.36 kb *Hind*III/*Sal*I fragment containing truncated *LSM4* (aa 1-92) with *LSM4* promoter region	This work
pMC12 was cloned as follows: A 1.36 kb fragment containing the first 276 bp of the *LSM4* ORF and upstream sequences was generated by PCR using pBS-SDB23 (genomic *BamH*I restriction fragment cloned into pBluescript KS^+ cloned such that the *LSM4* gene faces in the same direction as the T7 promoter; described by Cooper *et al.* (1995)) as template and primers T7 (AATACGACTCACTATAG) and M0055 (GCAGTCGACTTGCTGCTTGACCTT). The resulting product was cut with *Hind*III and *Sal*I, and cloned between the *Hind*III and *Sal*I restriction sites of pFL39, upstream of a stop-codon linker (TCGATTAACTAACTAG and TCGACTAGTTAGTTAA) inserted into the *Sal*I restriction site of this vector. The resulting vector expresses the first 92 amino acids of Lsm4p from its native promoter, but lacks its normal 3' UTR.		
pMPSLsm4	pGFP-N-LSM4: *LSM4* coding region inserted in frame with *GFP* coding region in pGFP-N-FUS	This work
pMPSLsm4D1	pGFP-N-lsm4ΔC: *lsm4ΔC* coding region (aa 1-93) inserted in frame with *GFP* coding region	This work
pMPSLsm4D2	pGFP-N-lsm4C: *lsm4C* coding region (aa 92-187) inserted in frame with *GFP* coding region	This work
pGFP-N-LSM1	*LSM1* coding region in frame with N-terminal GFP tag	This work
pMPSLsm2	pGFP-N-LSM2: *LSM2* coding region in frame with N-	This work

	terminal GFP tag in pGFP-N-FUS	
pMPSLsm6	pGFP-N-LSM6: *LSM6* coding region in frame with N-terminal GFP tag in pGFP-N-FUS	This work
pMPSLsm7	pGFP-N-LSM7: *LSM7* coding region in frame with N-terminal GFP tag in pGFP-N-FUS	This work
pMPS2	pGFP-N-FUS, with *LHP1* coding region inserted in frame with *GFP* coding region	Spiller et al., 2007a
pMR83	pGFP-C-LSM8: *LSM8* coding region in frame with C-terminal GFP in pGFP-C-FUS	Spiller et al., 2007b
pRP1155	*DCP2* with C-terminal RFP tag in *CEN*, *LEU* shuttle vector (gift from R. Parker)	Teixeira et al., 2005
pMR171	Promoter-*DCP2-RFP-PGK1* terminator from pRP1155 cloned into pRS313 (*CEN6*, *HIS3*)	This work
pPM119	pUN50 (*CEN4*, *URA3*) with P_{GAL}-*PGKpGmini* reporter (gift from D. Tollervey)	Mitchell and Tollervey, 2003
pMR202	pGFP-C-CCR4(1-229): N-terminal 229 residues of Ccr4p in frame with GFP-tag in pGFP-C-FUS	This work
pMR203	pGFP-C-POP2(1-156): N-terminal 156 residues of Pop2p in frame with GFP-tag in pGFP-C-FUS	This work
pMR204	pGFP-N-DHH1(427-506): C-terminal 80 residues of Dhh1p in frame with GFP-tag in pGFP-N-FUS	This work
pMR210	pGFP-N-DHH1: *DHH1* coding sequence in frame with GFP-tag in pGFP-N-FUS	This work
pMR211	pGFP-N-DHH1ΔC: Dhh1p aa 1-427 in frame with GFP-tag in pGFP-N-FUS	This work
pMR212	pGFP-N-CCR4: *CCR4* coding sequence in frame with GFP-tag in pGFP-N-FUS	This work
pMR213	pGFP-N-CCR4ΔN2: Ccr4p aa 250-837 in frame with GFP-tag in pGFP-N-FUS	This work

pMR214	pGFP-N-POP2: *POP2* coding sequence in frame with GFP-tag in pGFP-N-FUS	This work
pMR215	pGFP-N-POP2ΔN: Pop2p aa 147-433 in frame with GFP-tag in pGFP-N-FUS	This work
pMR216	pGFP-C-POP2: *POP2* coding sequence in frame with GFP-tag in pGFP-C-FUS	This work
pMR217	pGFP-C-POP2ΔN: Pop2p aa 147-433 in frame with GFP-tag in pGFP-C-FUS	This work
pMR218	pGFP-C-CCR4ΔN: Ccr4p aa 148-837 in frame with GFP-tag in pGFP-C-FUS	This work