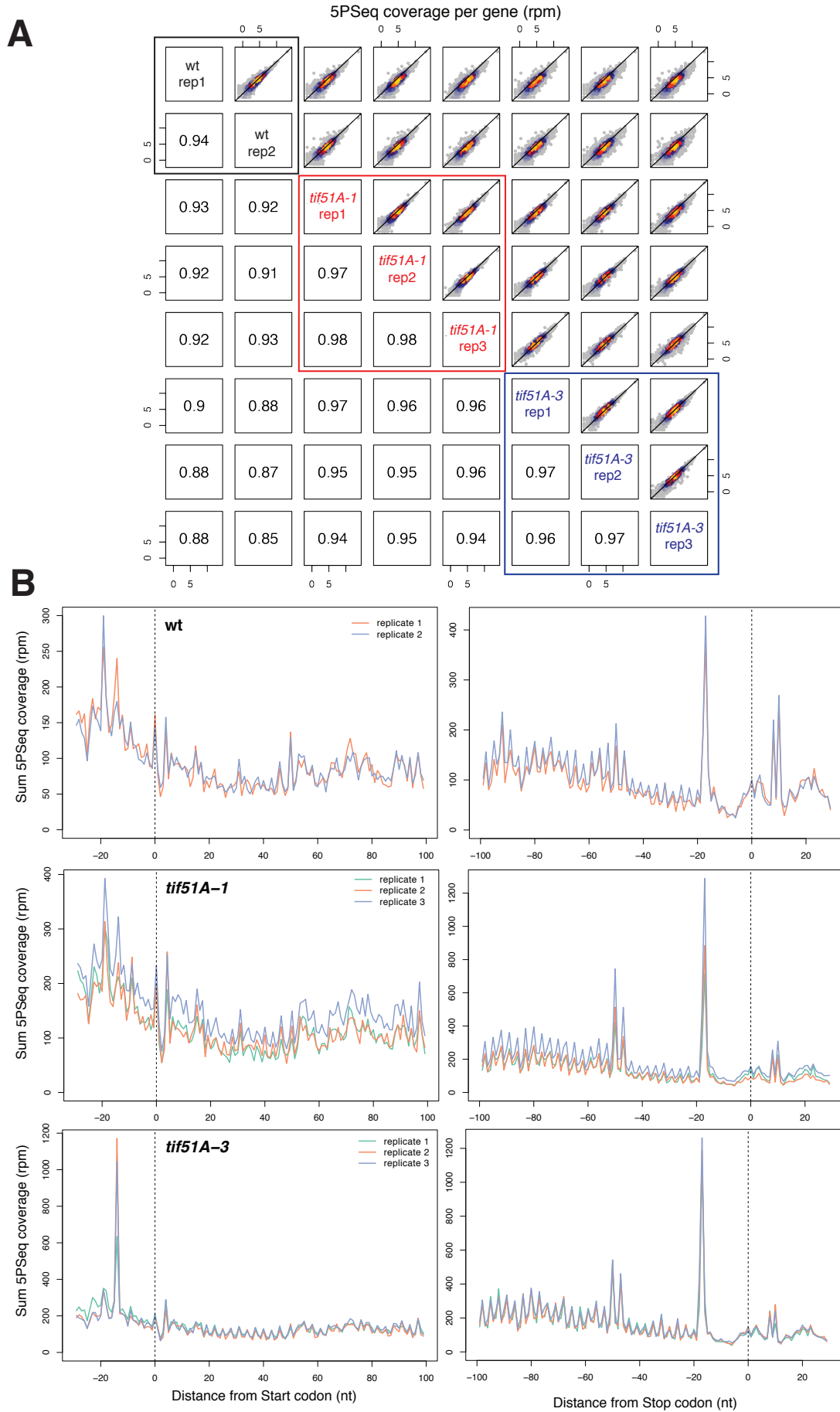


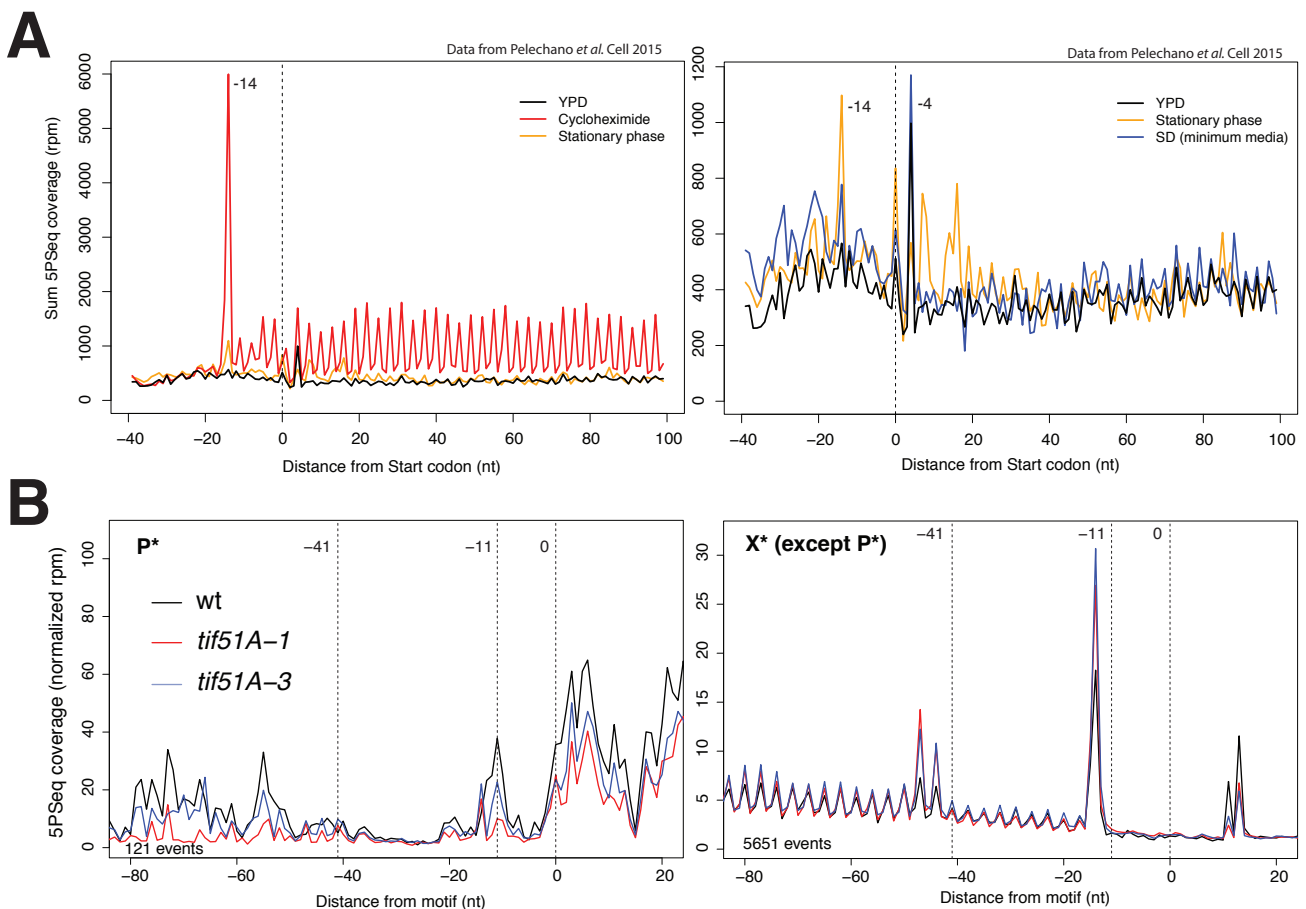
SUPPLEMENTARY DATA

Figure S1



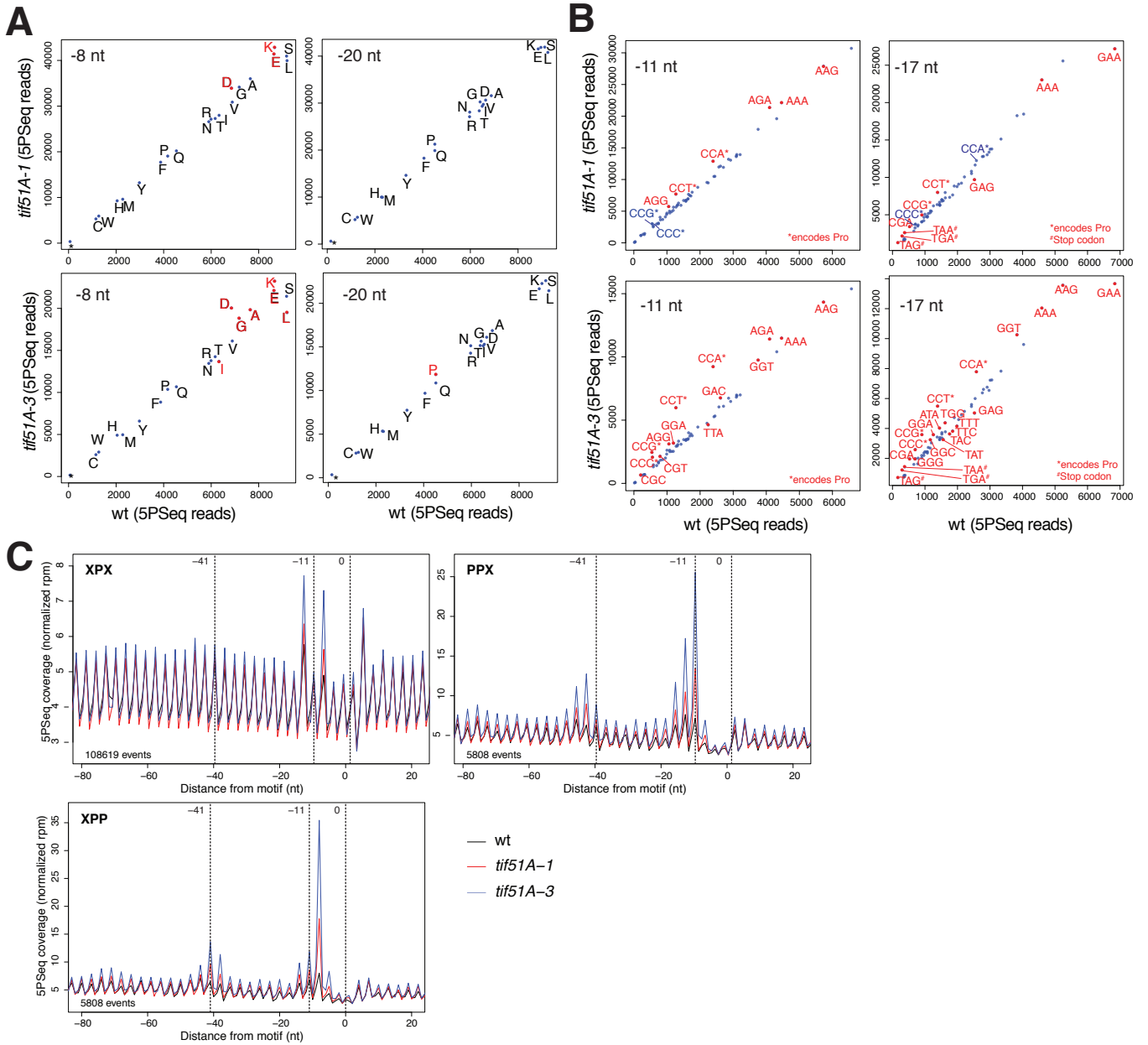
Supplementary Figure S1. Analysis of the reproducibility of the 5PSeq experiments. (A) Scatter plots showing the spearman correlation of reads (log₂-transformed) for each biological replicate of the 5PSeq experiments with the wild-type strain (BY4741) and the mutant *tif51A-1* and *tif51A-3* strains. (B) The metagenome analysis displaying the abundance of 5'P intermediates in reads per million (rpm) in relation to the ORF start (left panel) and stop (right panel) codons.

Figure S2



Supplementary Figure S2. (A) The metagenome analysis displaying the abundance of 5'P intermediates in reads per million (rpm) in relation to the ORF start codon in the cells grown in rich media (YPD, in black), in the stationary phase (in yellow), after cycloheximide treatment (CHX, in red) and in minimum media (SD, in blue). Data from Pelechano *et al.* (38). (B) The metagenome analysis displaying the abundance of 5'P intermediates in relation to the ORF stop codon. The identified pause is not dependent on the presence of an upstream proline residue. X refers to all amino acids, except proline.

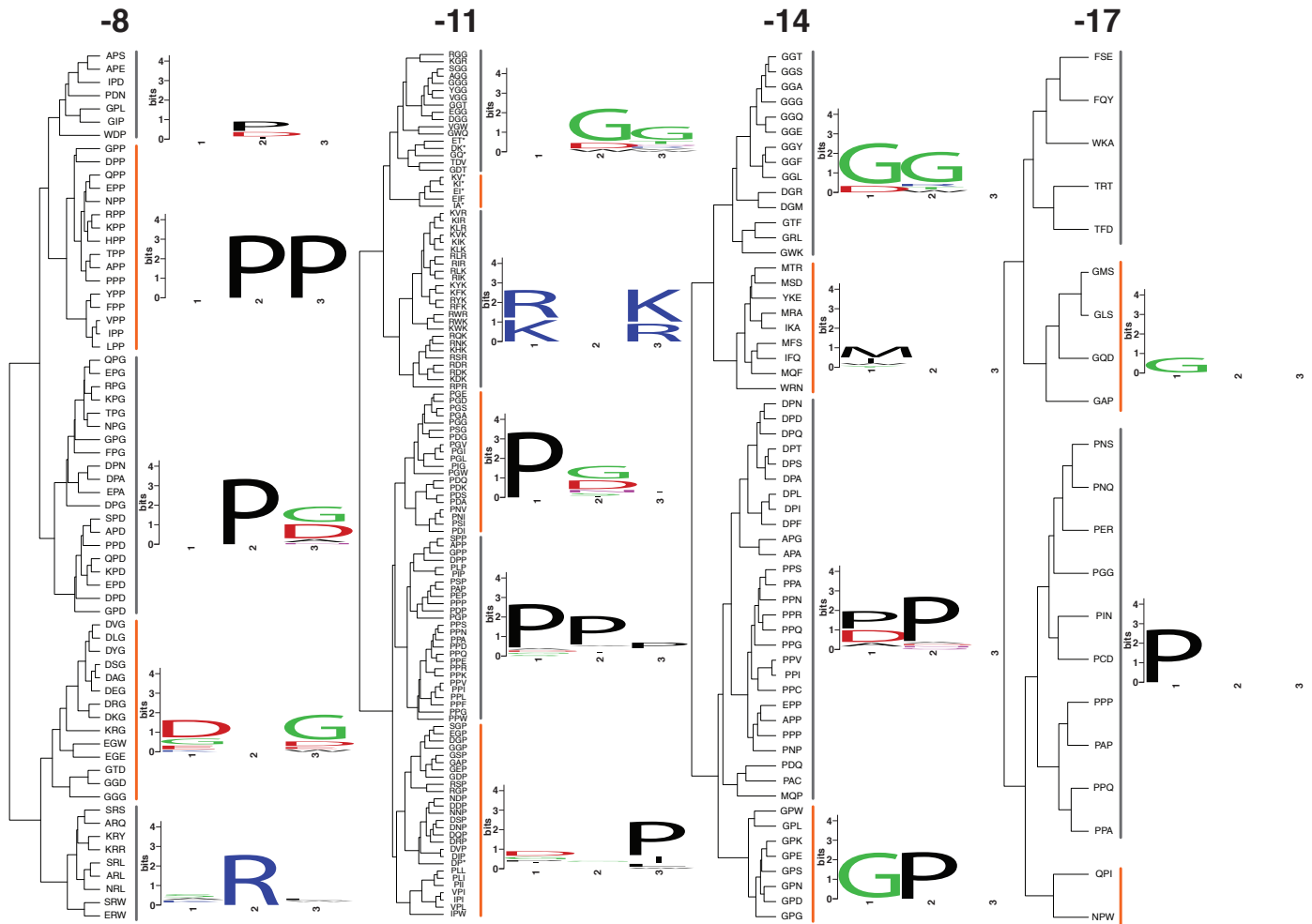
Figure S3



Supplementary Figure S3. (A) Differential ribosome pausing for all amino acids and (B) codons. Significantly regulated differences (adjusted p-val <0.01) in red. (C) Differential pausing depending on presence of proline codons and their context. The metagene represents the 5PSeq intermediates for the motifs, where the total number of reads for each metagene (-80 to +20 window) was normalized.

Figure S4

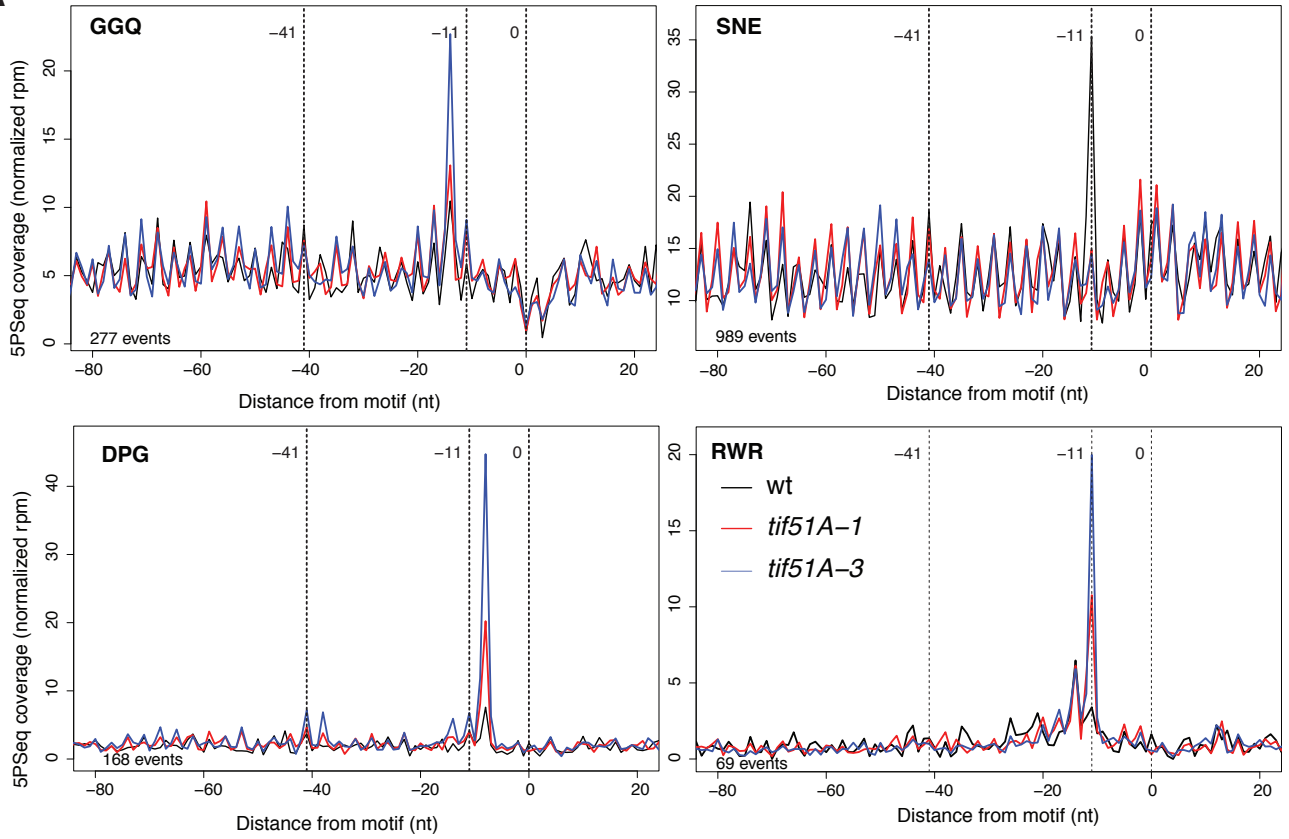
Pause-associated motifs



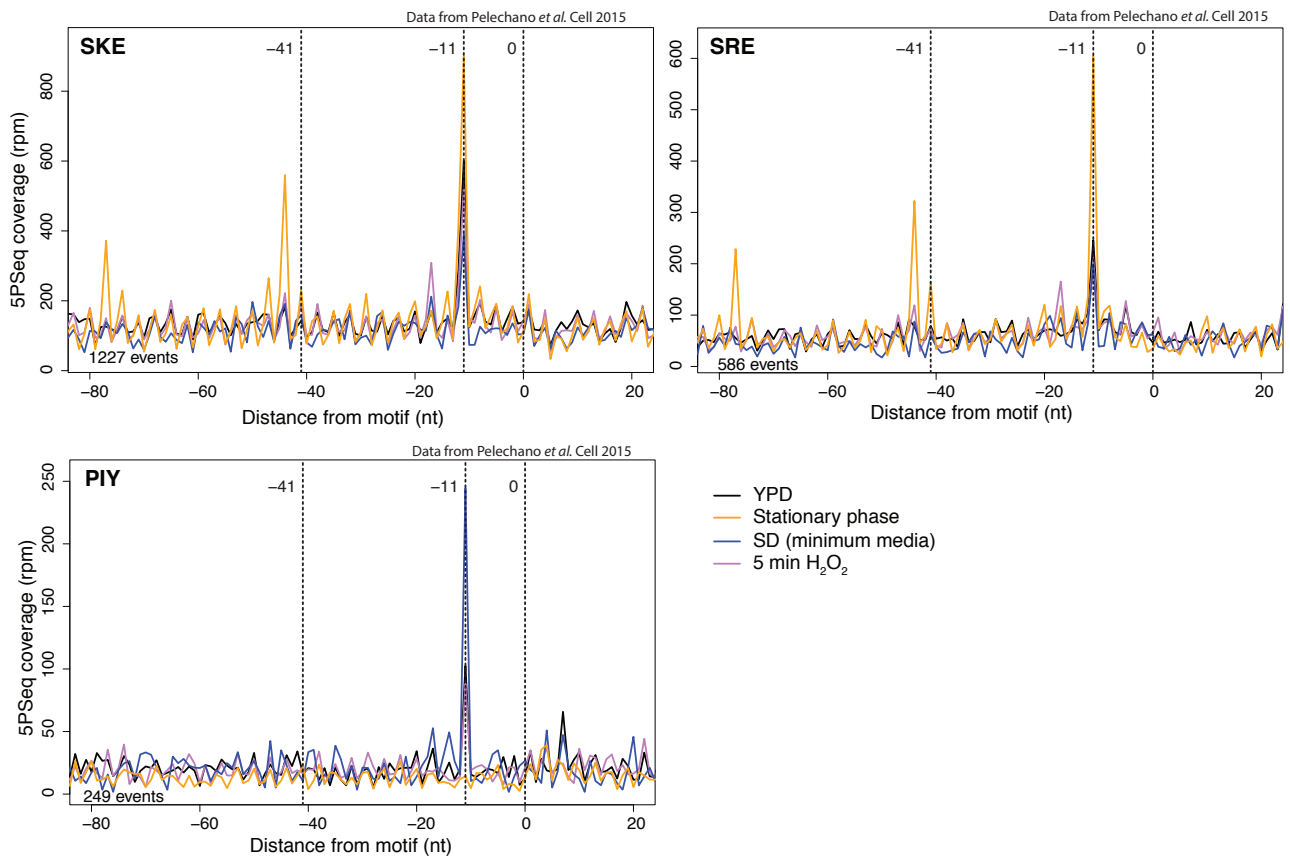
Supplementary Figure S4. The hierarchical cluster depicts the tripeptide motifs associated with eIF5A-dependent stall at different positions. The hierarchical cluster between motifs was computed using euclidean distances as defined by Biostrings (R package 2.40.2) and the complete linkage method.

Figure S5

A



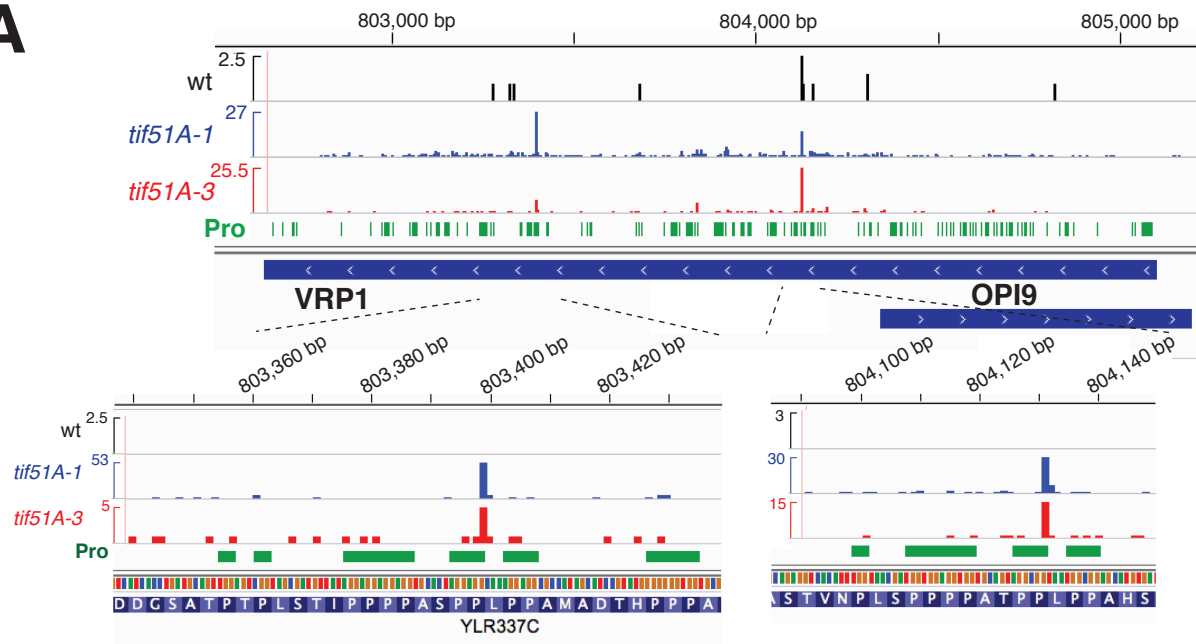
B



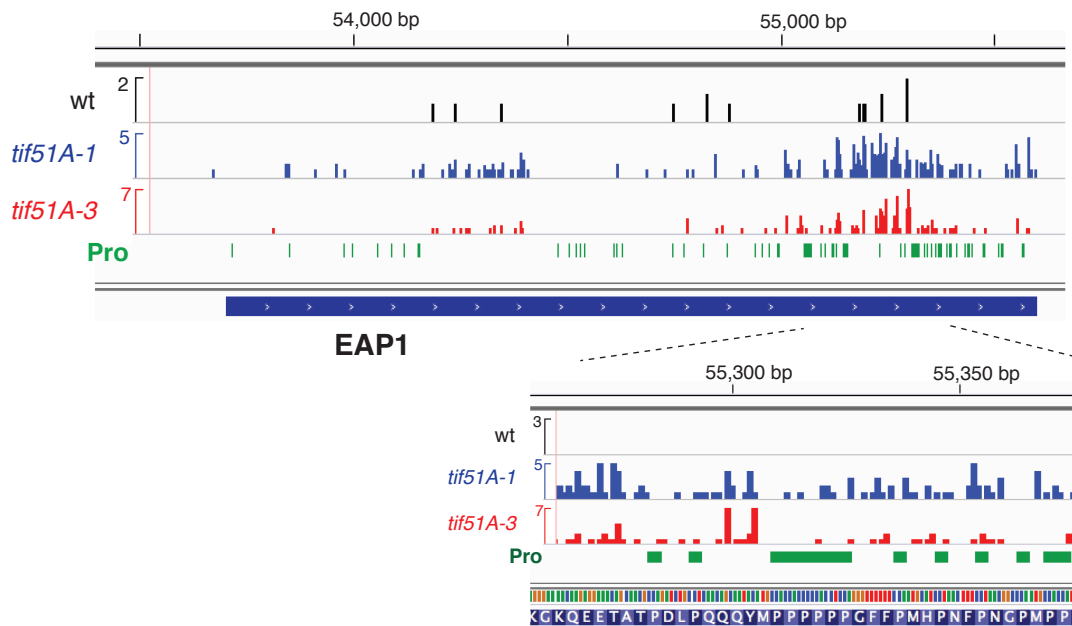
Supplementary Figure S5. The metagene representing the 5PSeq intermediates for the selected tripeptide motifs. The total number of reads for each metagene (-80 to +20 window) was normalized. (A) Selected tripeptide motifs that regulate the eIF5A-dependent pause. (B) Different physiological conditions modulate the relative pausing of the tripeptides negatively affected by presence of eIF5A. Data from Pelechano *et al.* (38).

Figure S6

A

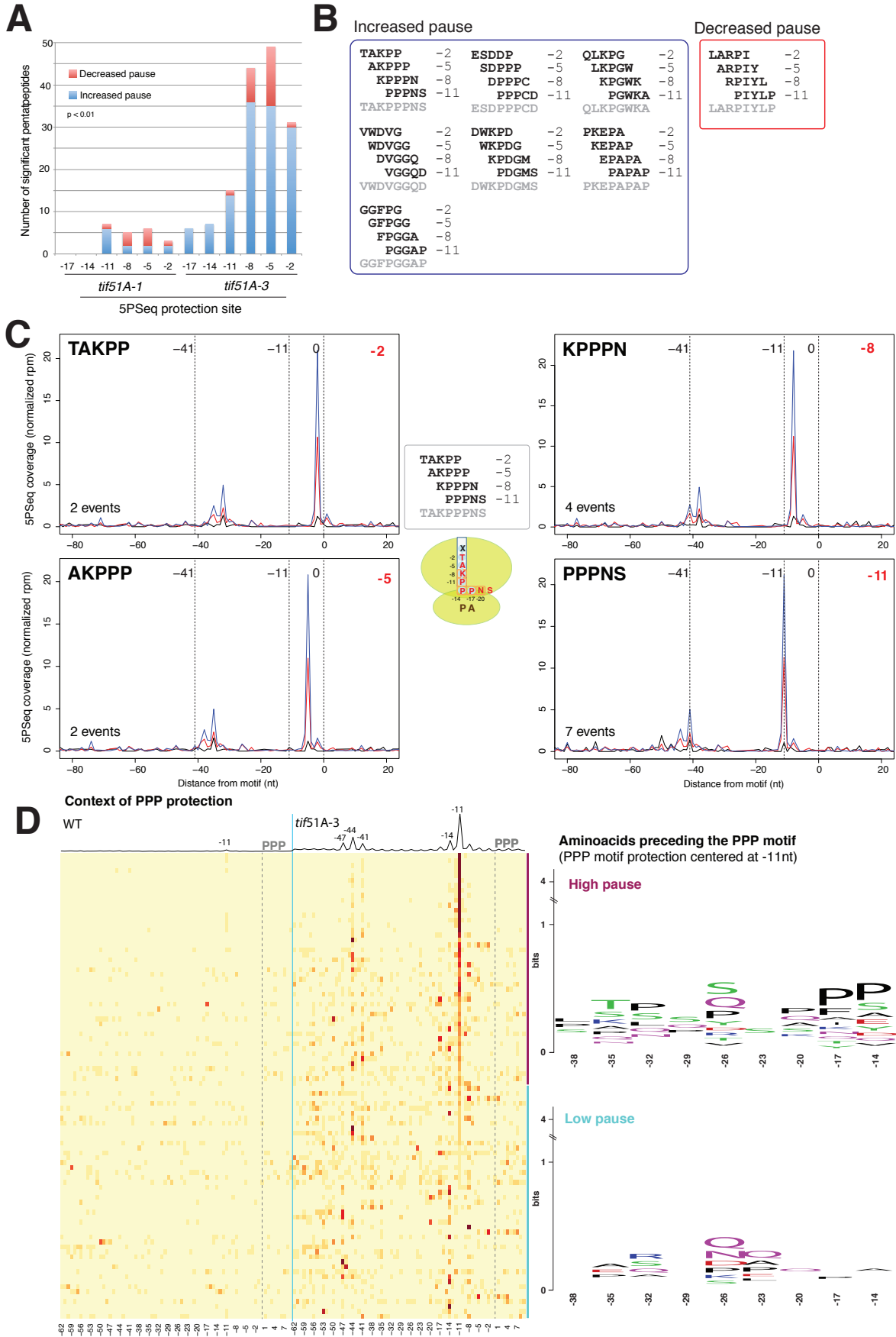


B



Supplementary Figure S6. Coverage of 5PSeq reads (collapsed to the single 5' nucleotide) in the wild type (black), *tif51A-1* (blue) and *tif51A-3* (red) for *VRP1* (A) and *EAP1* (B). Proline codons are depicted in green.

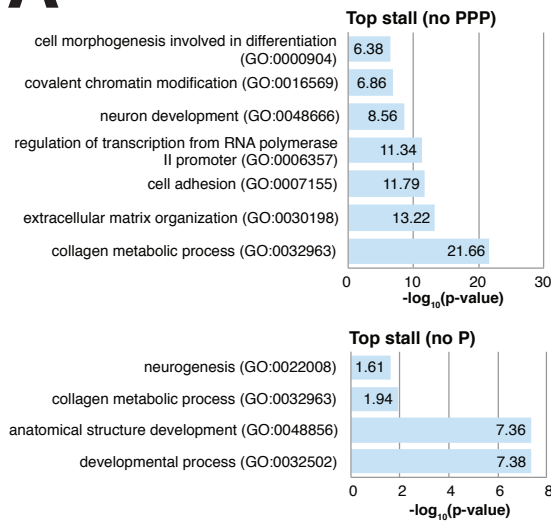
Figure S7



Supplementary Figure S7. (A) Number of significant pentapeptides in relation to the ribosome position. (B) Examples of extended motifs constructed by the concatenation of the identified pentapeptide motifs. (C) An example of the ribosome pausing caused by the four different pentapeptides that constitute the TAKPPPNS motif. (D) Heatmap of the 5PSeq motif specific protection patterns according to the upstream context of PPP motifs for the wild type and *tif51A-3* mutant strains. Motifs were sorted according to their relative protection at position -11 nt. Weblogo was performed for the upstream sequence of both motifs with high (red) and low (blue) ribosome stalls. Only the motif upstream regions with at least 10 reads were considered for the analysis.

Figure S8

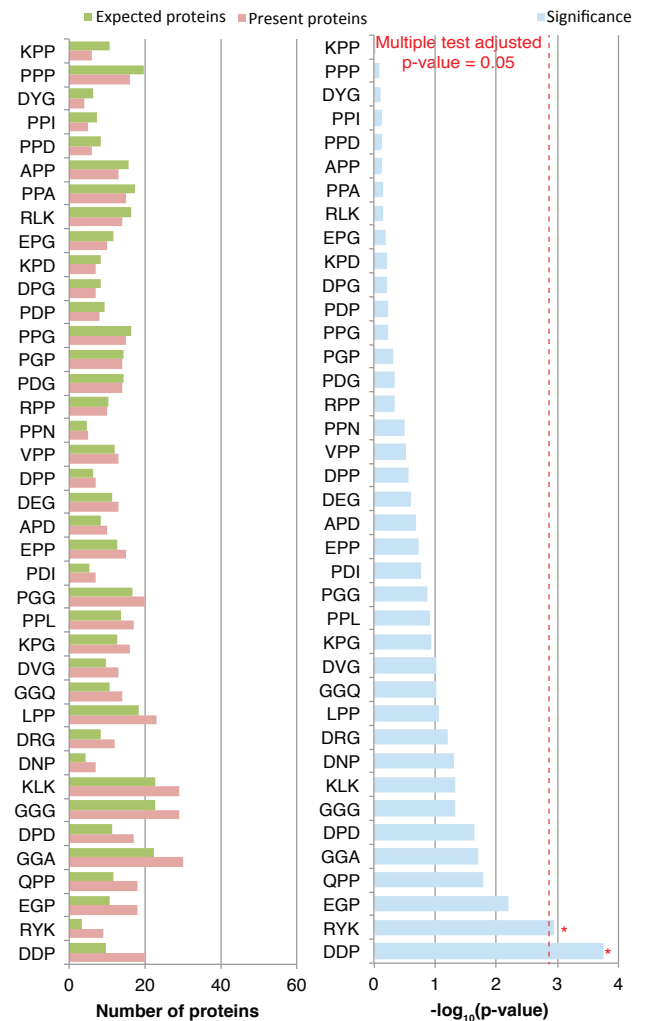
A



Data from Mandal *et al.* Sci Rep 2015
(30 proteins identified or more)

B

Human proteins downregulated after eIF5A depletion



Supplementary Figure S8. (A) The selected gene ontology terms significantly enriched in the human proteins with high content (>25 for the top-stall motifs, except PPP, and >10 motifs for the top-stall motifs, except those with P) of the eIF5A-dependent motifs. (B) The down-regulated proteins after eIF5A depletion identified by Mandal *et al.* (35) in relation to the identified tripeptide motifs. Randomly expected proteins (green), identified (pink) and significance (according to the hypergeometric distribution, blue). Significance level for Bonferroni-adjusted multiple testing is depicted in red and by *. Only the motifs present in at least 30 identified proteins were considered for the analysis.

Table S1. Main features of the 241 tripeptide eIF5A-regulated motifs.

Motif	5'P signal (Position at ribosome) ¹	Effect in <i>eIF5A</i> ^{ts}	% Motifs ²	Top motifs	Fold Change ³	P _{adj} -value	Abundance in yeast proteome ⁴
PPX/XPP	Prolines at -14 (P-site) and -11	Neg.	14.6	KPP PPP PPN PPD	9.00 7.03 5.99 5.97	3.35E-62 1.52E-62 1.34E-21 1.68E-25	326 (317) 828(565) 337(317) 259(243)
XGG/GGX	Glycines at -17 (A-site) and -14 (P-site)	Neg.	8.0	PGG GGQ GGP GGA	4.67 3.41 3.01 2.83	6.77E-23 1.84E-12 6.97E-09 4.88E-14	297(278) 279(261) 276(251) 610(522)
(R/K)X(R/K)	Basic AAs at -17 (A-site) and -11	Neg.	12.4	RWR KYK RYK RWK	3.63 3.61 3.58 2.87	1.16E-07 3.07E-10 7.70E-12 7.82E-06	72 (72) 478 (446) 273 (265) 100 (99)
DX(P/G/D) or D(P/G)X	First aspartic at -14 (P-site) or -11	Neg.	13.3	DPG DPP DYG DDP	5.06 4.63 3.78 3.66	1.51E-23 2.10E-15 4.85E-12 1.28E-16	176 (173) 184 (181) 224 (219) 433 (406)
PGX/XPG ⁵	Glycine at -14 (P-site) and proline at -11	Neg.	8.4	PGW PGP KPG EPG	6.41 4.44 4.32 4.20	3.88E-21 3.92E-11 1.29E-22 2.62E-16	77 (76) 235 (352) 374 (290) 246 (236)
GPX/XGP ⁵	Proline at -17 (A-site) and glycine at -14 (P-site)	Neg.	5.8	EGP RGP GPD DGP	3.71 3.32 2.95 2.75	5.26E-13 2.20E-08 8.59E-07 3.53E-06	249 (239) 216 (204) 228 (223) 263 (253)
SXE	Glutamic at -17 (A-site) and serine at -11	Pos.	26.7	SRE SKE SPE SNE	0.36 0.37 0.41 0.43	6.18E-07 7.10E-13 1.07E-04 1.41E-05	609 (564) 1336 (1126) 768 (694) 1036 (886)

¹Most frequent positions, defined as in Figure 2A. AA indicates amino acids.

²% respect to the group of motifs causing same effect on translation (neg. or pos).

³Fold change of 5'P signal of *tif51A-1* or *tif51A-3* respect to the wild-type.

⁴It is indicated the number of motifs and, into brackets, the number of proteins.

⁵Excluding PP or GG containing motifs.

Table S2. Gene Ontology categories (GO) over-represented in the *S. cerevisiae* proteins containing the pentapeptide eIF5A-regulated motifs.

Motif*	# Prot.	GO_term*	P-value	Gene(s) annotated to the term
QLKPG	6	cupric ion binding mating pheromone activity	7.0E-06 4.2E-05	MF(ALPHA)2/YGL089C;MF(ALPHA)1/YPL187W MF(ALPHA)2/YGL089C;MF(ALPHA)1/YPL187W
PPPNS	7	cyanamide metabolic process	4.1E-05	DDI2/YFL061W;DDI3/YNL335W
DVGGQ	7	GTPase activity	6.5E-09	ARF2/YDL137W;ARF1/YDL192W;GPA2/YER020W;ARF3/YOR094W;ARL3/YPL051W
GGFPG	5	tRNA import into nucleus protein folding RNA import into nucleus protein targeting to ER proteolysis	2.1E-08 4.6E-08 4.1E-07 5.2E-07 1.4E-05	SSA2/YLL024C;SIS1/YNL007C;YDJ1/YNL064C SSA1/YAL005C;SSA3/YBL075C;SSA2/YLL024C;SIS1/YNL007C;YDJ1/YNL064C SSA2/YLL024C;SIS1/YNL007C;YDJ1/YNL064C SSA1/YAL005C;SSA3/YBL075C;SSA2/YLL024C;YDJ1/YNL064C SSA1/YAL005C;SSA3/YBL075C;SSA2/YLL024C;SIS1/YNL007C;YDJ1/YNL064C
PPPPP	41	actin cytoskeleton organization actin cortical patch mating projection tip	2.1E-07 2.3E-17 2.1E-05	AIM3/YBR108W;ABP1/YCR088W;BNR1/YIL159W;PAN1/YIR006C;VRP1/YLR337C;APP1/YNL094W;SRV2/YNL138W;BNI1/YNL271C;LAS17/YOR181W;SCD5/YOR329C AIM3/YBR108W;ABP1/YCR088W;LDB17/YDL146W;PAN1/YIR006C;MYO3/YKL129C;VRP1/YLR337C;MYO5/YMR109W;APP1/YNL094W;INP52/YNL106C;SRV2/YNL138W;INP53/YOR109W;LAS17/YOR181W;SCD5/YOR329C ABP1/YCR088W;PAN1/YIR006C;VRP1/YLR337C;MYO5/YMR109W;INP52/YNL106C;SRV2/YNL138W;BNI1/YNL271C;LAS17/YOR181W
DKGRL	5	SRP-dependent cotranslational protein targeting to membrane protein targeting to ER unfolded protein binding ATPase activity	2.5E-12 3.8E-10 1.1E-09 9.5E-07	SSA1/YAL005C;SSA3/YBL075C;SSA4/YER103W;KAR2/YJL034W;SSA2/YLL024C SSA1/YAL005C;SSA3/YBL075C;SSA4/YER103W;KAR2/YJL034W;SSA2/YLL024C SSA1/YAL005C;SSA3/YBL075C;SSA4/YER103W;KAR2/YJL034W;SSA2/YLL024C SSA1/YAL005C;SSA3/YBL075C;SSA4/YER103W;KAR2/YJL034W;SSA2/YLL024C

*GOs computed using GO Term Finder (v0.82): <http://www.yeastgenome.org/cgi-bin/GO/goTermFinder.pl>