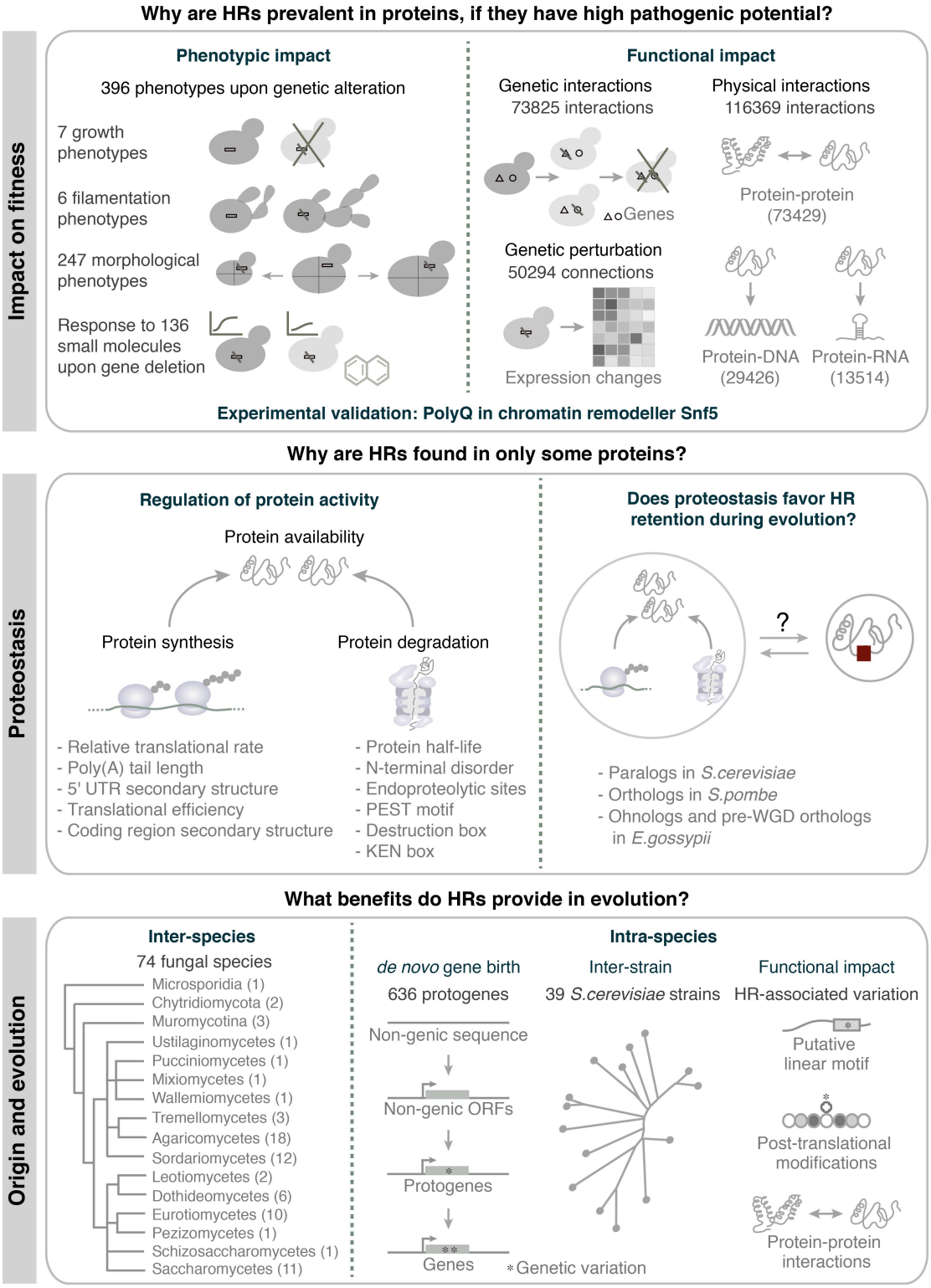
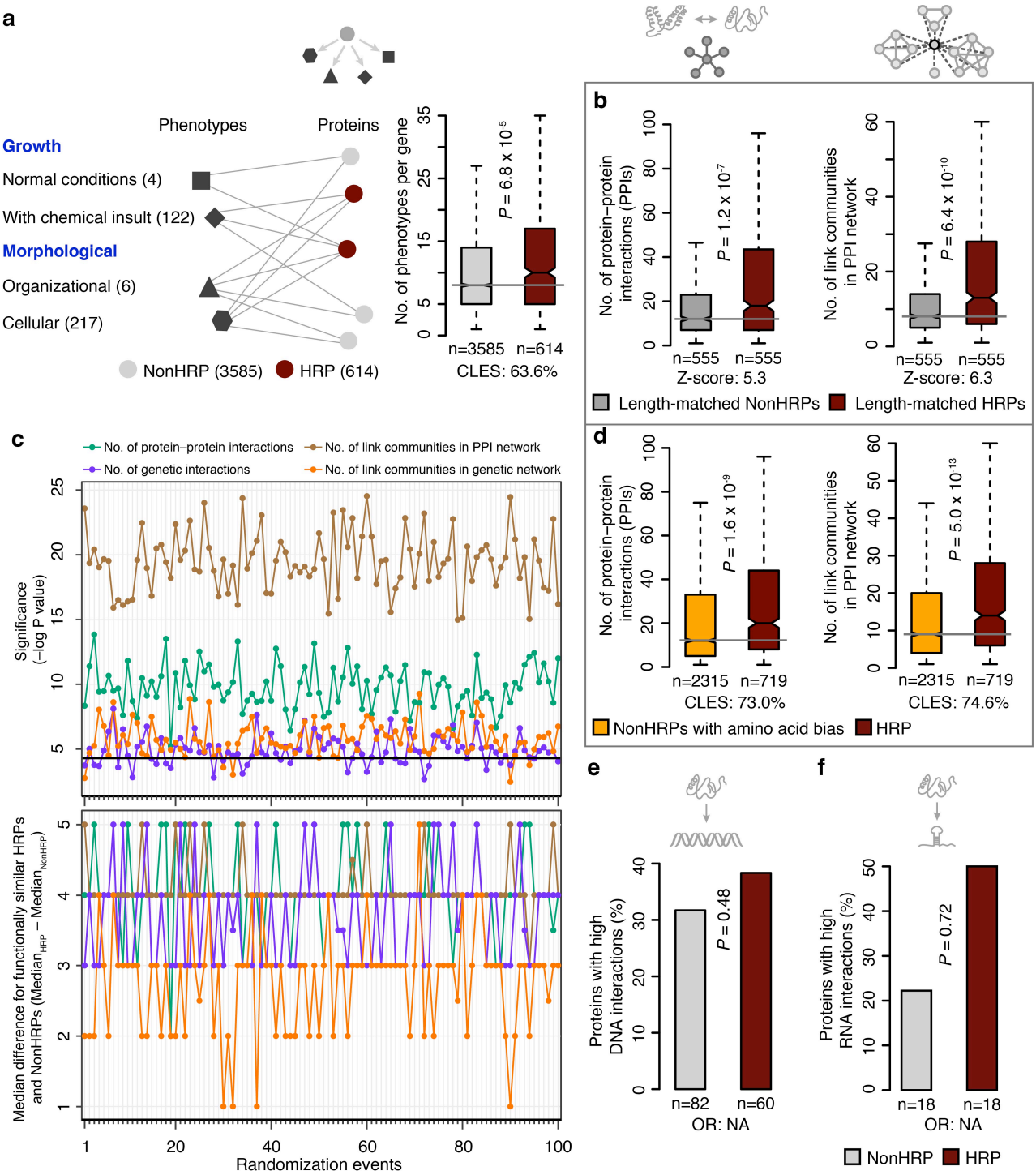


Supplementary Figure 1

Study outline, questions investigated and datasets analyzed.



Supplementary Figure 2

Control calculations for physiological importance and functional versatility of HRPs

(**a**) Scaled gene to phenotype (scaled-G2P) network of yeast non-essential genes, reconstructed by merging phenotypes that had more than 49% genetic overlap (left panel). Boxplot of distribution of the number of phenotypes in the scaled G2P network among HRPs and

NonHRPs (right panel). Statistical significance was assessed using Wilcoxon rank sum test, with effect sizes displayed as CLES.

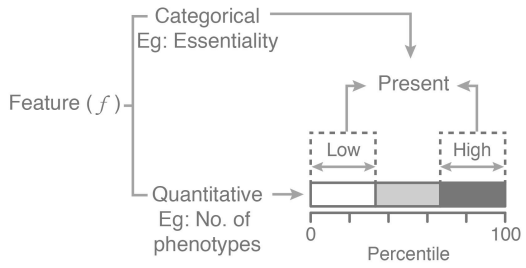
Different features such as (i) protein length (Ekman, D. *et al., Genome Biol* **7**, R45, 2006) and (ii) low complexity sequences, especially amino acid compositional bias (Coletta, A. *et al. BMC Syst Biol* **4**, 43, 2010), can influence the propensity of proteins to interact with multiple partners and hence its functionality. To investigate the effect of protein length on the functionality of HRPs, we obtained length matched controls and compared their distributions. To investigate whether the observed trends for functionality are similar between proteins with homorepeats and NonHRPs containing non-repeat amino acid bias we first identified all yeast NonHRPs that contain compositional bias. Boxplots of distribution of the number of protein-protein interactions and link communities in PPI network among (**b**) length-matched HRPs and NonHRPs and (**d**) HRPs and NonHRPs with amino acid bias. Statistical significance was assessed using Wilcoxon rank sum test and FDR corrected for multiple testing within each class, with effect sizes displayed as CLES. HRPs tend to have more interactions and are involved in more processes than their length-matched counterparts. HRPs have more protein interactions and participate in more biological processes than proteins with non-repetitive amino acid bias. Collectively, these results suggest that HRPs are more functionally versatile than proteins with similar lengths and those, which contain amino acid bias.

(**c**) Proteins with homorepeats are functionally more versatile than NonHRPs with similar biological functions. Significance estimates are provided as –log P values (upper panel) and effect size estimates are provided as median differences (lower panel) for each one of the 100 randomizations comparing distributions of HRPs and randomly selected but functionally similar NonHRPs for (i) No. of protein-protein interactions, (ii) No. of link communities in PPI network, (iii) No. of genetic interactions and (iv) No. of link communities in genetic network. Statistical significance was assessed using Wilcoxon rank sum test. The solid black line in the upper panel represents a p-value of 0.05.

Barplots showing percentages of HRPs and NonHRPs with a high number of (**e**) protein-DNA and (**f**) protein-RNA interactions. Since the number of data points was limited, we classified transcription factors and RNA binding proteins into tertiles representing proteins with low, medium and high number of protein-DNA and protein-RNA interactions, respectively. Statistical significance was assessed using Fisher's exact test, with effect size represented by OR.

a

Types of features analyzed

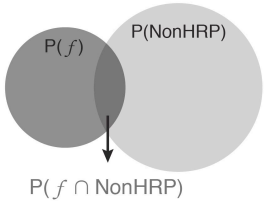


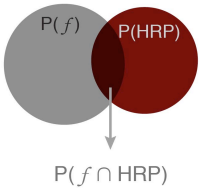
Likelihood Ratio (LR) Interpretation

LR > 1	Higher likelihood of finding a feature among HRPs
LR < 1	Higher likelihood of finding a feature among NonHRPs
LR ≈ 1	No bias in the likelihood

Conditional Probability and Likelihood Ratio estimation

$$\text{Probability(Event I Condition)} = \frac{\text{P (Event} \cap \text{Condition)}}{\text{P (Condition)}}$$





$$P(\text{ƒ I NonHRP}) = \frac{P(\text{ƒ} \cap \text{NonHRP})}{P(\text{NonHRP})} \qquad P(\text{ƒ I HRP}) = \frac{P(\text{ƒ} \cap \text{HRP})}{P(\text{HRP})}$$

$$\text{Likelihood Ratio (LR)} = P(\text{ƒ I HRP}) / P(\text{ƒ I NonHRP})$$

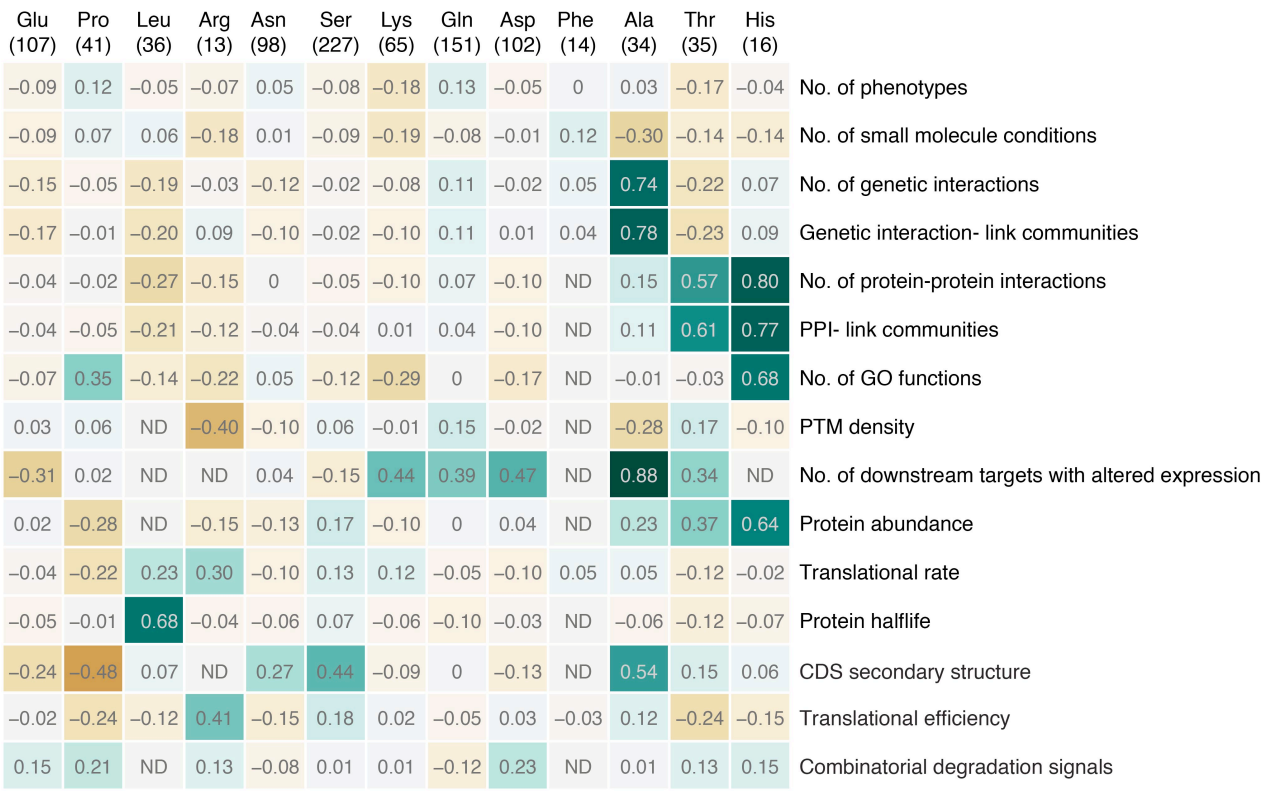
b

	HRP	Phe (14)	Leu (36)	Arg (13)	His (16)	Lys (65)	Ser (227)	Thr (35)	Ala (34)	Gln (151)	Asn (98)	Pro (41)	Glu (107)	Asp (102)	
Impact on fitness	1.23	0	0.68	0.94	0.38	2.36	0.73	1.05	1.98	1.46	1.25	1.05	1.95	1.56	Essentiality*
	1.30	0.68	1.24	0.24	0.60	0.69	0.99	0.64	1.31	1.20	1.40	1.09	0.92	1.06	No. of phenotypes (high)
	1.23	0.97	1.13	1.05	1.06	0.73	0.87	0.49	1.20	1.06	1.18	0.91	0.92	1.10	No. of small molecule conditions (high)
	1.22	0	0.71	0.49	1	0.88	0.74	0.64	1.13	0.97	1.14	0.93	0.83	1	No. of genetic interactions (high)
	1.18	0	0.89	0.50	1.01	0.89	0.70	0.55	1.23	0.96	1.08	0.94	0.81	0.92	Genetic interaction- link communities (high)
	1.45	0.24	0.85	0.78	0.63	1.51	0.85	1.16	1.89	1.73	1.28	2.23	1.90	1.56	No. of protein-protein interactions (high)
	1.59	0.25	0.68	0.54	0.88	1.30	1.08	1.21	1.86	1.93	1.62	2.49	1.97	1.62	PPI- link communities (high)
	1.31	0.24	1.13	1.05	0.43	1.20	1.20	0.78	1.50	1.31	1.46	1.74	1.24	1.20	No. of GO functions (high)
	3.22	0	1.18	3.26	0	0.98	0.75	0.61	0	0.42	0.87	1.03	0.40	0.83	Low protein solubility*
	2.13	0	0	0.36	1.47	1.38	0.89	1.61	1.52	1.12	1.44	1.15	1.10	1.29	PTM density (high)
	1.41	0	0.19	0.26	0.21	0.15	0.16	0.29	0.79	0.42	0.48	0.49	0.34	0.33	No. of downstream targets with altered expression (high)
Proteostasis	1.63	0.68	1.06	0.73	2.97	1.17	1.47	1.09	2.24	2.58	2.04	1.62	1.42	1.40	Over-expression lethality*
	1.07	0	0.16	0.45	0.73	0.36	0.47	0.42	0.35	0.60	0.69	0.50	0.41	0.46	Protein abundance (low)
	1.64	0.24	1.48	2.31	1.66	0.87	1.79	1.81	1.18	1.46	2	1.38	1.34	1.70	Translational rate (low)
	1.39	0	0.44	0.74	1.20	0.54	0.80	0.73	1.22	1.03	1.30	1.09	0.63	1	Protein halflife (low)
	2.62	0	0.27	0.50	0.40	0.79	0.51	0.74	0.57	0.24	0.63	0.87	1.18	0.98	Short poly(A) tail*
	1.25	0	0.52	0	0.59	0.39	0.44	0.45	0.74	0.54	0.32	0.53	0.32	0.28	5'UTR secondary structure (high)
	1.08	0	0.52	0	0.20	0.44	0.42	0.36	1.11	0.63	0.22	0.31	0.29	0.15	CDS secondary structure (high)
	1.25	0.83	1.69	1.12	1.27	1.20	1.12	0.50	0.85	1.09	1.39	0.85	1.22	1.48	Translational efficiency (low)
	2.05	0	0.44	1.84	1.99	0.86	1.23	1.14	1.40	1.90	2.11	1.75	1.34	1.33	Long N-terminal disorder*
	2.37	0	0.28	0.78	1.90	0.99	1.35	1.74	2.29	2.01	2.03	1.90	1.64	1.76	Endoproteolytic site*
	2.27	0	0.57	1.26	2.30	1.13	1.44	1.75	1.80	1.27	1.46	1.69	2.06	2.12	PEST motif*
	1.48	0	0.34	0.93	1.14	0.75	0.83	1.21	0.71	1.33	1.55	1.04	1.25	0.95	Destruction box*
	1.39	0	0.56	0	0.63	1.40	0.62	0.87	1.19	0.74	1.34	0.74	1.23	1.49	KEN box*
	1.95	0	0.33	0.92	1.86	0.82	1	1.36	1.58	1.26	1.82	1.31	1.56	1.40	Combinatorial degradation signals (high)

Supplementary Figure 3

Influence of the type of amino acid of homorepeat on different features analyzed in this study.

(**a**) Estimation of Likelihood Ratios (LRs) from conditional probabilities for different types of features (categorical or quantitative) analyzed in this study and its interpretation. Quantitative features were categorized into low or high classes based on tertile cut-offs. (**b**) The heatmap presents LRs for all HRPs and proteins with different amino acid repeat types, with the numbers in the parenthesis the corresponding to the number of proteins with a particular amino acid type. Categorical features are marked with asterisks. For quantitative features, the class analyzed (high or low) is presented in the parenthesis. LRs greater than 1 are marked in green, with the intensity of the color representing its magnitude.

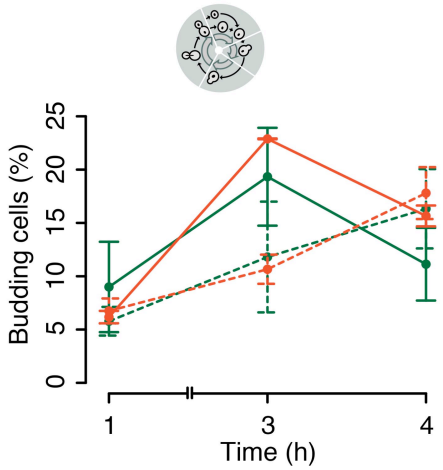


Supplementary Figure 4

Influence of the length of different amino acid repeat types on different features analyzed in this study.

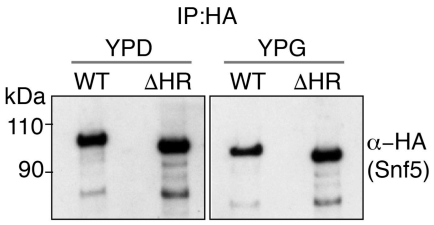
Each cell in the heatmap represents the Pearson correlation coefficient defining the extent and type of correlation (positive/negative) between the lengths of a specific amino acid repeat (given as column names) and the feature analyzed (row names). The numbers in the parentheses indicate the number of proteins with a particular amino acid repeat type. Correlation coefficients >0.70 were found to be statistically significant upon correcting for multiple testing.

a

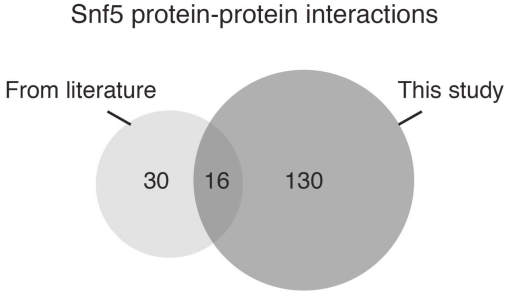


	Average budding cells (%)					
	YPD			YPG		
	1h	3h	4h	1h	3h	4h
WT	9.0	19.3	11.1	6.2	22.9	15.7
ΔHR	5.8	11.8	16.3	6.7	10.7	17.8
ANOVA *P*	3.2×10^{-2}			2.8×10^{-6}		

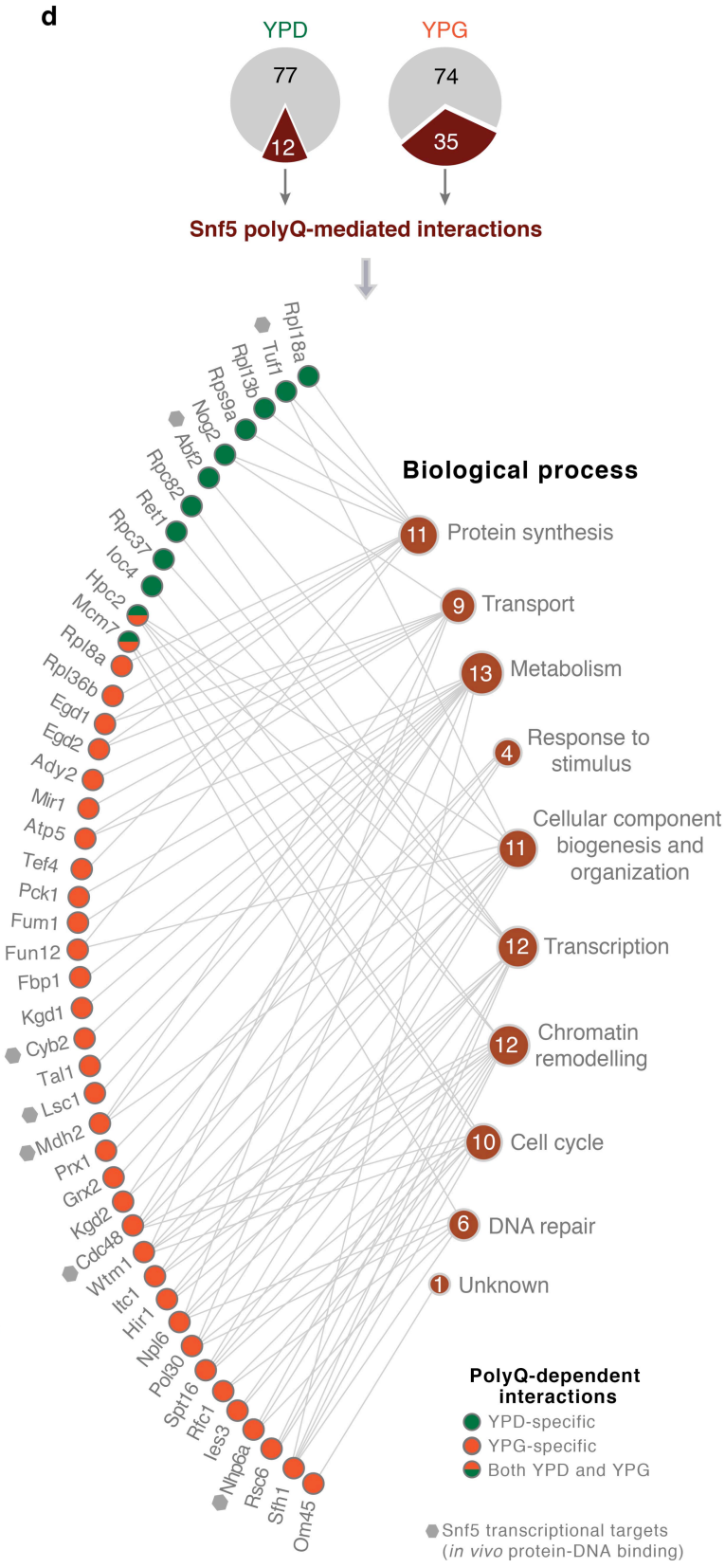
b



c



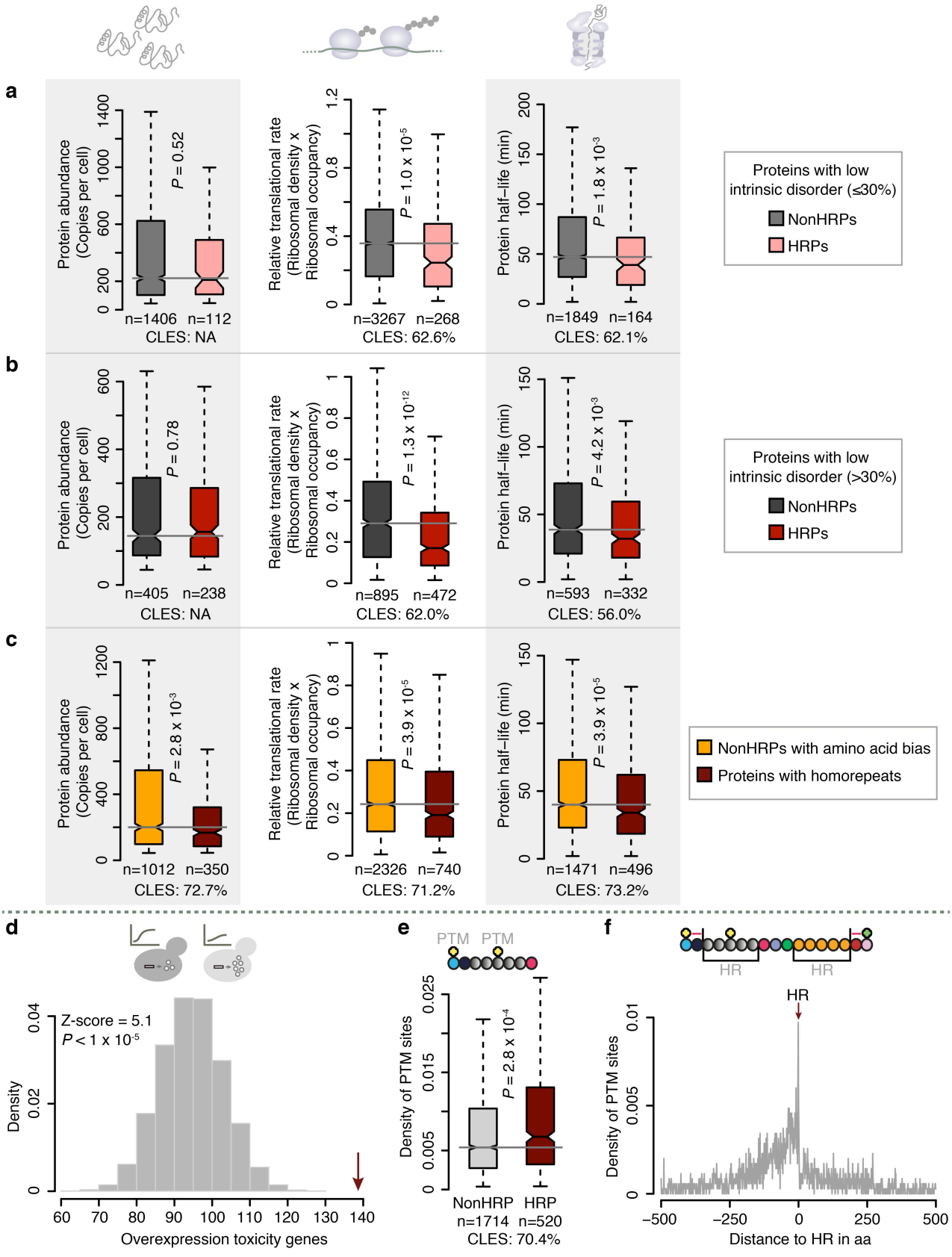
d



Supplementary Figure 5

PolyQ in Snf5 influences cell division and mediates protein-protein interactions.

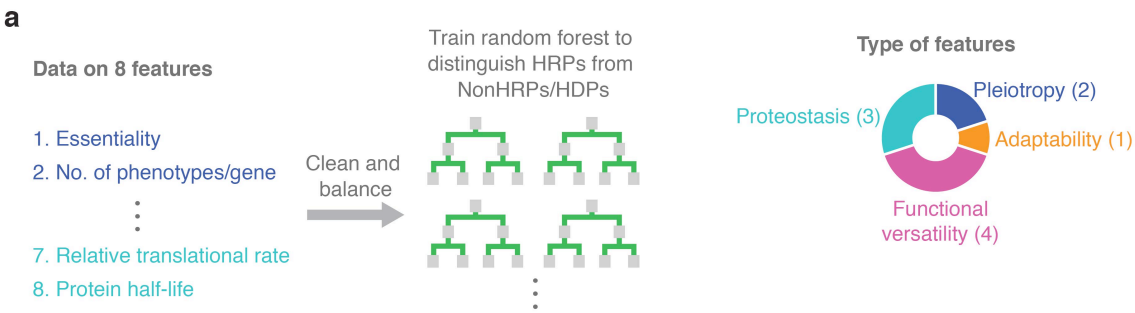
(**a**) Budding index of cells after re-inoculation of stationary phase cultures from WT or ΔHR in YPD or YPG after 1, 3 and 4h. Statistical significance was assessed using ANOVA. (**b**) Immunoblot analysis for HA-tagged WT and ΔHR Snf5 in the immunoprecipitate. (**c**) Number of Snf5 protein-protein interactions previously known, as retrieved from BioGRID (Chatr-Aryamontri, A. *et al.*, *Nucleic acids research* **41**, D816–23, 2013) and identified in this study. (**d**) Proteins interacting with Snf5 in a polyQ dependent manner (observed in WT but absent in ΔHR) in YPD and YPG connected to the diverse biological processes. The interactors are identified with a minimum of 3 unique peptide hits in at least 2 of the 3 experiments and classified according to the associated biological processes. The pie chart on top shows the number of polyQ-mediated interactions in each condition.


a
Protein abundance (Copies per cell)
P = 0.52
n=1406 n=112
CLES: NA
Relative translational rate (Ribosomal density x Ribosomal occupancy)
P = 1.0 x 10-5
n=3267 n=268
CLES: 62.6%
Protein half–life (min)
P = 1.8 x 10-3
n=1849 n=164
CLES: 62.1%
Proteins with low intrinsic disorder (≤30%)
NonHRPs
HRPs
b
Protein abundance (Copies per cell)
P = 0.78
n=405 n=238
CLES: NA
Relative translational rate (Ribosomal density x Ribosomal occupancy)
P = 1.3 x 10-12
n=895 n=472
CLES: 62.0%
Protein half–life (min)
P = 4.2 x 10-3
n=593 n=332
CLES: 56.0%
Proteins with low intrinsic disorder (>30%)
NonHRPs
HRPs
c
Protein abundance (Copies per cell)
P = 2.8 x 10-3
n=1012 n=350
CLES: 72.7%
Relative translational rate (Ribosomal density x Ribosomal occupancy)
P = 3.9 x 10-5
n=2326 n=740
CLES: 71.2%
Protein half–life (min)
P = 3.9 x 10-5
n=1471 n=496
CLES: 73.2%
NonHRPs with amino acid bias
Proteins with homorepeats
d
Z-score = 5.1
P < 1 x 10-5
Density
Overexpression toxicity genes
e
PTM PTM
Density of PTM sites
P = 2.8 x 10-4
NonHRP HRP
n=1714 n=520
CLES: 70.4%
f
HR HR
HR
Density of PTM sites
Distance to HR in aa

Supplementary Figure 6

Stringent regulation of HRP activity by coarse- and fine-tuning.

Previously we showed that intrinsically disordered proteins are tightly regulated (Gsponer, J., *et al. Science* **322**, 1365–8, 2008). This raises the question of whether the stringent regulation of HRPs is confounded by protein disorder. To address this, we compared the regulatory features between HRPs and NonHRPs with low and high intrinsic disorder content. Boxplot of distributions of protein abundance, translational rate and protein half-life among HRPs and NonHRPs with (**a**) low and (**b**) high disorder content. Irrespective of the disorder content, HRPs tend to have lower translational rates and low protein half-life compared to NonHRPs, though the protein abundance is comparable. Thus, stringent protein regulation of HRPs is more pronounced compared to NonHRPs with similar disorder content. To investigate if HRPs and proteins with non-repetitive amino acid bias are regulated with similar stringency, we compared the regulatory features between HRPs and NonHRPs with amino acid bias. (**c**) Boxplots showing the distributions of protein abundance, translational rate and protein half-life among HRPs and NonHRPs with amino acid bias. HRPs tend to be less abundant with lower translational rates and protein half-lives. This suggests that HRPs are more stringently controlled than NonHRPs without amino acid bias. Statistical significance was determined using Wilcoxon rank sum test corrected for multiple testing within each class of comparisons. Effect sizes are provided as CLES. (**d**) Distribution (grey histogram) showing the random expectation of the over-expression toxicity HRPs, with the red arrow showing the observed number of over-expression toxicity HRPs. Enrichment of HRPs in over-expression toxicity genes was tested using permutation test, by performing 10,000 iterations. The Z-score indicates the distance of the actual observation to the mean of random expectation in terms of number of standard deviation. P-values were estimated as the ratio of randomly observed proteins greater than or equal to the number of actually observed HRPs to the total number of random samples (10,000).Over-representation of HRPs among over-expression toxicty genes implies that they are stringently regulated by coarse-tuning, i.e., by regulating their abundance. (**e**) Box plot showing distribution of density of PTMs among NonHRPs and HRPs. Density of PTMs for each protein was estimated by obtaining a ratio of the total number of PTMs identified in a protein to the total number of amino acids in that protein. The P-value was estimated using Wilcoxon rank sum test and effect size displayed as CLES. More PTMs in HRPs suggest a stringent fine-tuning of their activity. (**f**) Distribution of the density of PTMs within (shown as 0 in the X-axis) and 500 amino acids on either side of the HR.

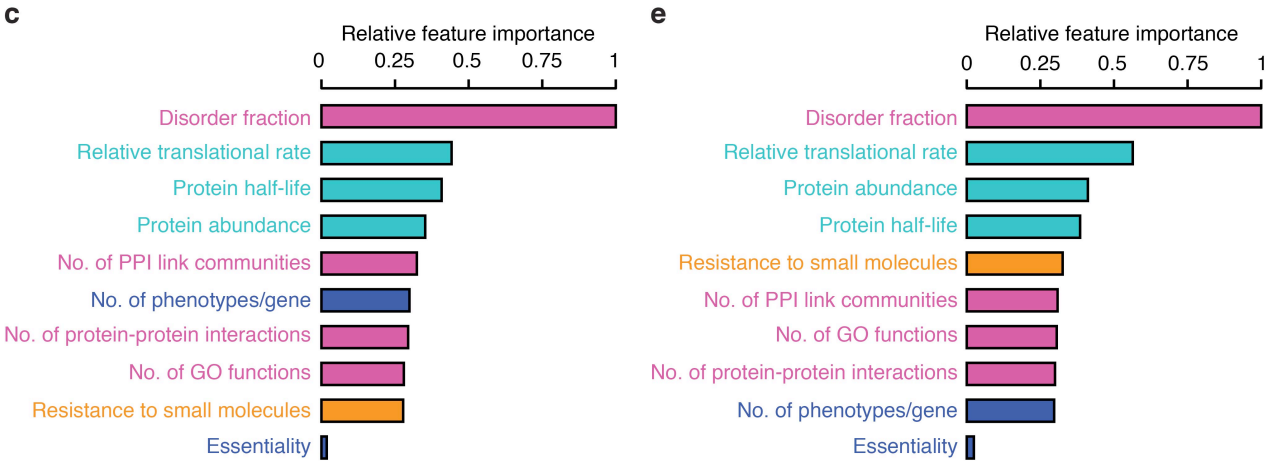


b

Class	Precision	Recall	F1	Support
NonHRP	0.74	0.82	0.78	206
HRP	0.78	0.68	0.73	188
Average/Total	0.76	0.75	0.75	394

d

Class	Precision	Recall	F1	Support
HDP	0.67	0.72	0.70	198
HRP	0.67	0.62	0.64	181
Average/Total	0.67	0.67	0.67	379

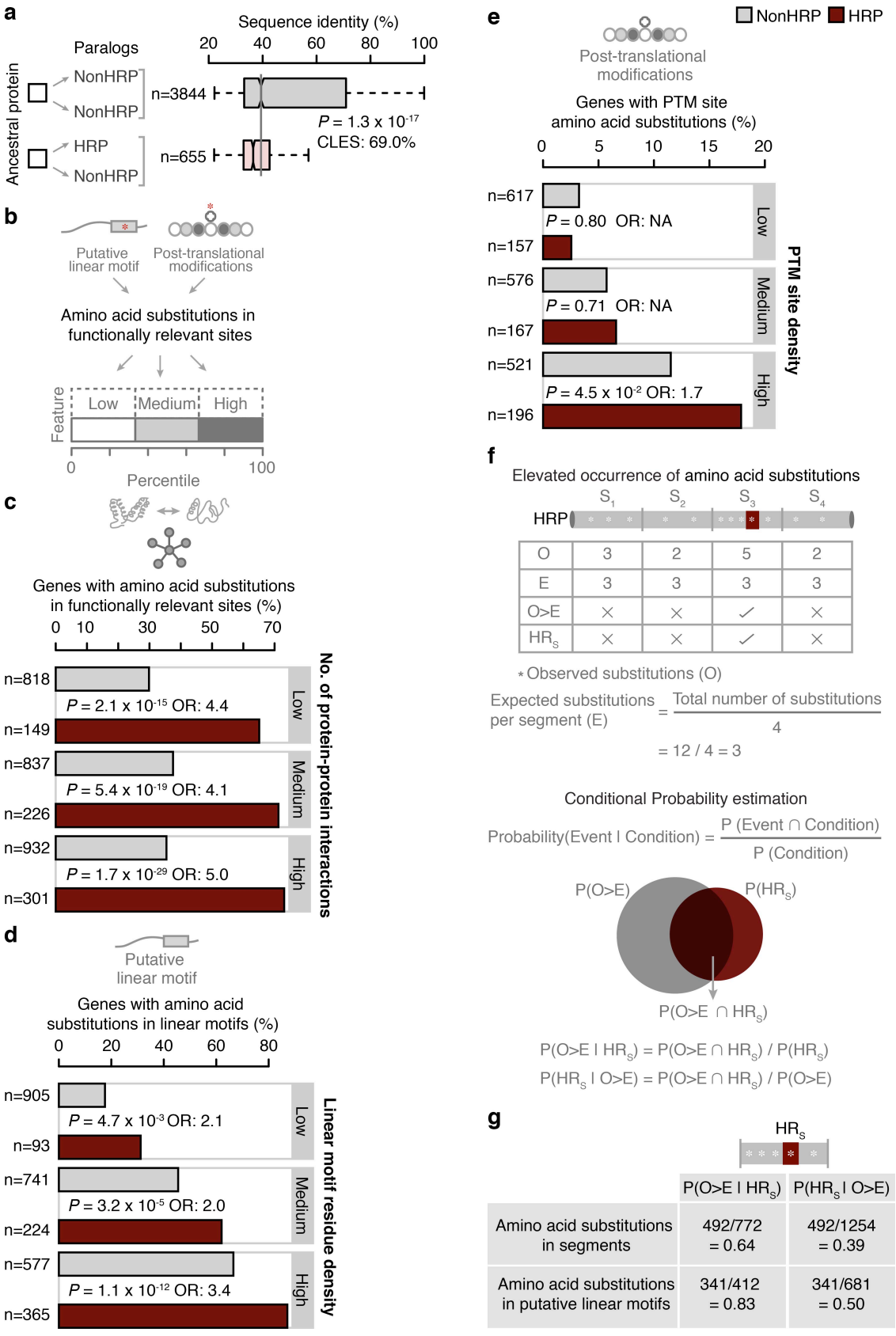


Supplementary Figure 7

Pleiotropic effects, adaptability, functional versatility and proteostasis collectively differentiate HRPs from NonHRPs and highly disordered NonHRPs.

(**a**) The left panel depicts the work flow of a random forest (RF) model that was trained on eight features to distinguish (i) proteins with (HRPs) and without HRs (NonHRPs) and (ii) HRPs from highly disordered NonHRPs (HDP; intrinsic disorder >30%). The eight features considered for this machine learning approach pertained to pleiotropy, adaptability, functional versatility and proteostasis (right panel). The number of features for each class is provided in the parenthesis. (**b**) The RF model trained on the eight features is able to distinguish between HRPs and Non HRPs with an overall accuracy of 0.75. Test set precision, recall, f1 scores, and sample sizes (support) are provided for each class. Genes with three or more missing values among the eight features were disregarded. This resulted in 3819 genes with values for at least five features. For some of these genes, the missing values were imputed using multiple imputation by chained equations, using the `mice` R package (arguments to mice method: method="pmm", m=5, maxit=5; van Buuren, S. *et al. J Stat Softw* **45**, doi:10.18637/jss.v045.i03, 2011). The outcome class was balanced (data before balancing: 16% HRP, n=3819; data after balancing: 48%, n=1261) by down-sampling the majority class size to approximately the size of the minority class (down sampled majority n=700). The cleaned and balanced feature set was split into 70% training and 30% testing data. The training data was used as the input to a random forest model composed of 50 classification trees, with hyperparameters (purity criterion, maximum features nodes consider for splitting, and max tree depth) optimized using 10-fold cross-validation and grid search (yielding optima: criterion="entropy", max_features="log2", max_depth=None). (**c**) Importance of different features for distinguishing HRPs from NonHRPs relative to the most important feature, disorder fraction. The most predictive features are related to disorder and proteostasis, with the remaining features being of approximately equal importance, with relatively low contribution from essentiality. Since disorder fraction was an important feature for distinguishing HRPs from NonHRPs, we investigated if HRPs could be distinguished from highly disordered NonHRPs (HDPs; with disorder fraction >30%). (**d**) The RF model is able to distinguish HRPs from HDPs with an overall

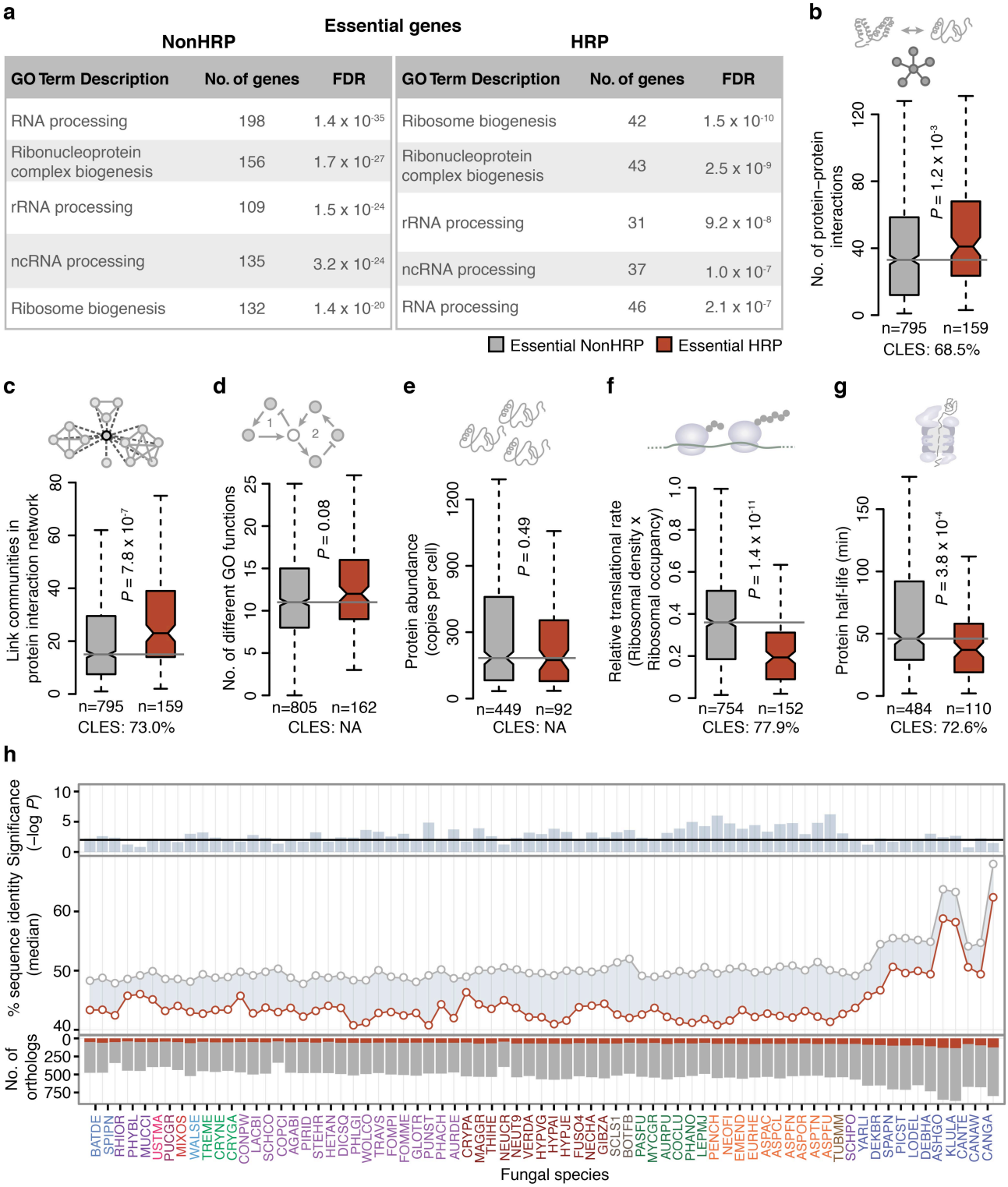
accuracy of 0.67. (**e**) Importance of different features for distinguishing HRPs from HDPs, relative to disorder fraction. Similar to our observations for NonHRPs, the most important features are related to proteostasis, with other features being of approximately equal importance and low contribution from essentiality.


a
Sequence identity (%)
20 40 60 80 100
Paralogs
Ancestral protein
NonHRP
NonHRP
n=3844
P = 1.3 x 10-17
CLES: 69.0%
HRP
NonHRP
n=655
b
Putative linear motif
Post-translational modifications
Amino acid substitutions in functionally relevant sites
Feature
Low
Medium
High
0
100
Percentile
c
Genes with amino acid substitutions in functionally relevant sites (%)
0 10 30 50 70
n=818
P = 2.1 x 10-15 OR: 4.4
n=149
Low
n=837
P = 5.4 x 10-19 OR: 4.1
n=226
Medium
n=932
P = 1.7 x 10-29 OR: 5.0
n=301
High
No. of protein-protein interactions
d
Putative linear motif
Genes with amino acid substitutions in linear motifs (%)
0 20 40 60 80
n=905
P = 4.7 x 10-3 OR: 2.1
n=93
Low
n=741
P = 3.2 x 10-5 OR: 2.0
n=224
Medium
n=577
P = 1.1 x 10-12 OR: 3.4
n=365
High
Linear motif residue density
e
NonHRP
HRP
Post-translational modifications
Genes with PTM site amino acid substitutions (%)
0 5 10 15 20
n=617
P = 0.80 OR: NA
n=157
Low
n=576
P = 0.71 OR: NA
n=167
Medium
n=521
P = 4.5 x 10-2 OR: 1.7
n=196
High
PTM site density
f
Elevated occurrence of amino acid substitutions
S1 S2 S3 S4
HRP
O 3 2 5 2
E 3 3 3 3
O>E × × ✓ ×
HRS × × ✓ ×
* Observed substitutions (O)
Expected substitutions per segment (E) = Total number of substitutions / 4
= 12 / 4 = 3
Conditional Probability estimation
Probability(Event I Condition) = P (Event ∩ Condition) / P (Condition)
P(O>E)
P(HRS)
P(O>E ∩ HRS)
P(O>E I HRS) = P(O>E ∩ HRS) / P(HRS)
P(HRS I O>E) = P(O>E ∩ HRS) / P(O>E)
g
HRS
P(O>E I HRS)
P(HRS I O>E)
Amino acid substitutions in segments
492/772 = 0.64
492/1254 = 0.39
Amino acid substitutions in putative linear motifs
341/412 = 0.83
341/681 = 0.50

Supplementary Figure 8

Evolutionary benefits associated with HRs.

(**a**) Box plot showing the distribution of sequence identities among similar NonHRP:NonHRP) and divergent (HRP:NonHRP) pairs of yeast paralogs. Statistical signficance was estimated using Wilcoxon rank-sum test, with effect size displayed as CLES. To test if our observations related to HR-associated amino acid substitutions (Fig. 8) are confounded by high number of protein- protein interactions or higher density of linear motifs or PTM sites associated with HRPs, we classified yeast proteins into three bins of low, medium and high using tertile cut-offs for each of the features (panel **b**). Linear motif residue density was estimated by obtaining the ratio of residues predicted to form putative linear motifs over the entire length of the protein. In each bin, we tested for differences in the proportion of NonHRPs and HRPs with and without amino acid substitutions in the functionally relevant sites. Barplots showing proportion of NonHRPs and HRPs with amino acid substitutions affecting (**c**) functionally relevant sites (putative linear motifs and/or PTM sites) among low, medium and high bins of proteins classified based on the number of protein-protein interactions, (**d**) putative linear motifs among different bins of proteins classified based on the density of linear motif residues and (**e**) PTM sites among different bins of proteins classified based on the density of PTM sites. Statistical significance within each bin for each attribute was tested using Fisher's exact test and corrected for multiple testing (FDR). If a feature is a confounder, then there will not be any difference in the proportion of NonHRPs and HRPs with amino acid substitutions across different bins of that feature. Higher proportion of HRPs tend to harbor substitutions affecting functionally relevant sites compared to NonHRPs across all bins, with matched number of protein-protein interactions (panel **c**). Similarly, across all bins of linear motif residue density, higher proportion of HRPs harbor substitutions within putative linear motifs compared to the matched NonHRPs (panel **d**). HRPs show a significant difference for the proportion of genes with amino acid subsitutions affecting PTM sites especially in the 'high' bin of PTM site density (panel **e**). These findings suggest that the amino acid substitutions affecting functionally relevant sites in HRPs are independent of the number of protein interactions that a HRP participates in, or the density of linear motif residues or PTM sites. (**f**) Estimation of conditional probability for observing more amino acid substitutions than expected by chance in segments containing HRs and conditional probability for observing HRs in segments with more observed amino acid substitutions than expected by chance (denoted by P(O>E | HRs) and P(HRs | O>E), respectively). Each protein was divided into four equal segments and the expected number of amino acid substitutions per segment was calculated as shown. Subsequently, segments with more amino acid substitutions than expected and those that contained HRs were identified. Using a similar approach, segments with more amino acid substitutions in putative linear motifs were defined. Due to very few data points, substitutions at the PTM sites were not considered for estimating conditional probabilities. (**g**) Conditional probability values of finding more amino acid substitutions than expected by chance in HR segments and that of finding a HR given a segment contains more amino acid substitutions (first row). The second row provides the conditional probability values of finding more amino acid substitutions affecting putative linear motifs than expected by chance in HR segments and that of finding a HR given a segment contains more linear motif affecting amino acid substitutions.



Supplementary Figure 9

Essential HRPs constitute a part of the rapidly adaptable part of the proteome, facilitated by stringent proteostasis.

(**a**) Enrichment of Gene Ontology (GO) biological processes among essential HRPs and NonHRPs in yeast. Both essential HRPs and NonHRPs are over-represented in similar biological processes. Boxplot of distributions of the number of (**b**) protein-protein interactions, (**c**) link communities in protein-protein interaction network, (**d**) different GO terms a gene is associated with, reflecting its functional dviersity, (**e**) protein abundance, (**f**) translational rate and (**g**) protein half-life among essential HRPs and NonHRPs. Statistical significance was assessed using Wilcoxon rank sum test and P-values were corrected for multiple testing (FDR). (**h**) Sequence divergence among yeast essential HRPs and NonHRPs with their one-to-one orthologs in 73 fungal species belonging to 16 distinct fungal classes. Species names are abbreviated. Taxonomic and phylogenetic details of the fungal species are provided in Supplementary Notes. Median divergences of essential HRPs (red circle) and NonHRPs (grey circle) with their orthologs in each species are shown in the middle panel. The upper panel provides the statistical significance estimated by comparing the distribution of divergence of yeast essential HRPs and NonHRPs with their respective orthologs in each species. The black line shows a P-value cut-off of 0.01, corrected for multiple testing. The bottom panel shows the number of orthologs of yeast essential HRP (red) and NonHRP (grey) in each species. The average median difference between essential NonHRPs and HRPs is 6.3%, which corresponds to ~19 missense variant positions for a 300 amino acid long protein. Though both HRP and NonHRP essential genes show enrichment for similar processes, essential proteins that contain a HR are more functionally versatile, stringently regulated and undergo accelerated sequence divergence and thereby may constitute the rapidly adaptable part of the genome.

a

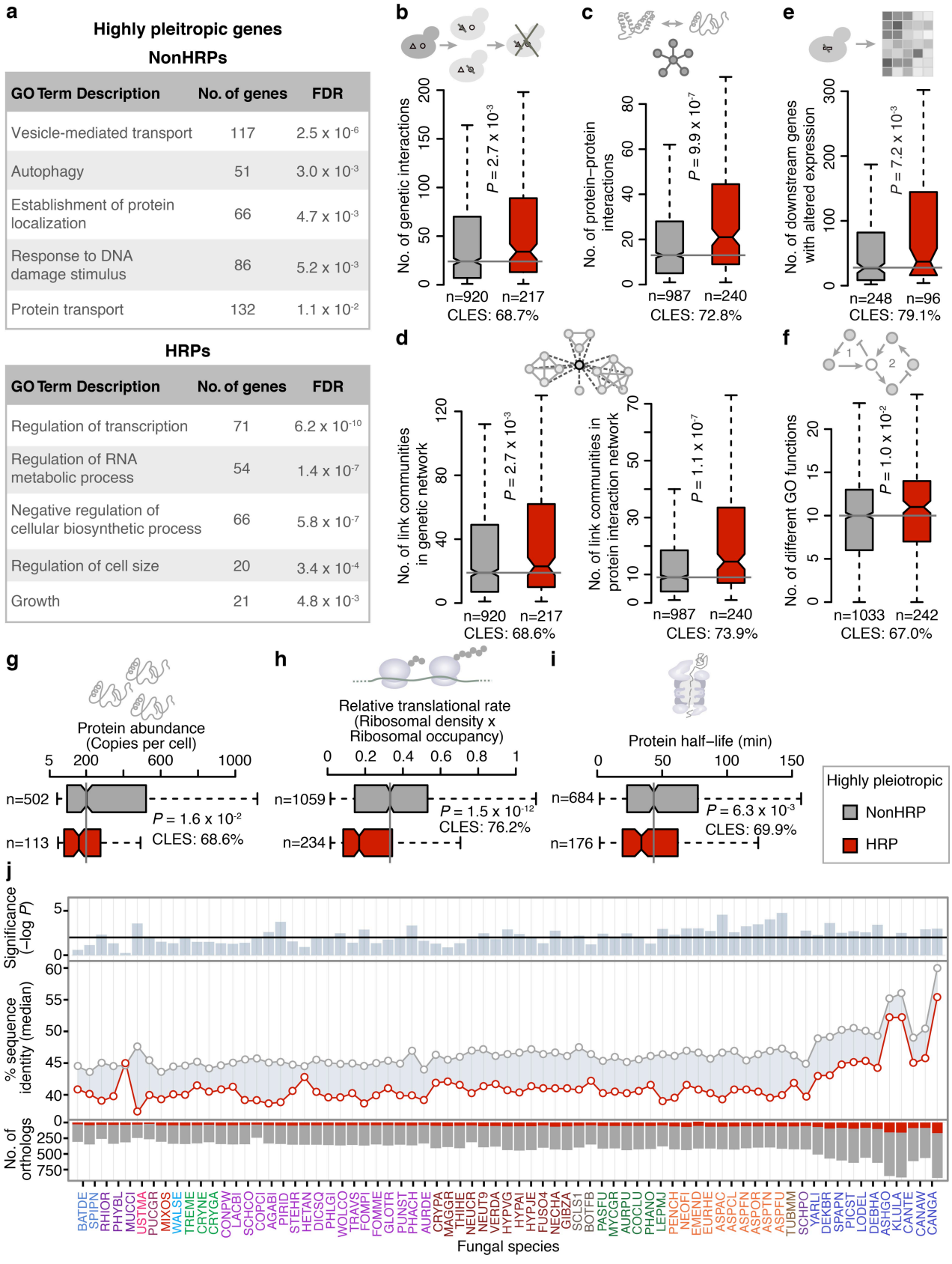
Highly pleitropic genes

NonHRPs

GO Term Description	No. of genes	FDR
Vesicle-mediated transport	117	2.5×10^{-6}
Autophagy	51	3.0×10^{-3}
Establishment of protein localization	66	4.7×10^{-3}
Response to DNA damage stimulus	86	5.2×10^{-3}
Protein transport	132	1.1×10^{-2}

HRPs

GO Term Description	No. of genes	FDR
Regulation of transcription	71	6.2×10^{-10}
Regulation of RNA metabolic process	54	1.4×10^{-7}
Negative regulation of cellular biosynthetic process	66	5.8×10^{-7}
Regulation of cell size	20	3.4×10^{-4}
Growth	21	4.8×10^{-3}



Supplementary Figure 10

Highly pleiotropic HRPs constitute a part of the rapidly adaptable part of the proteome, facilitated by stringent proteostasis.

We selected the top 33% of pleiotropic genes (defined by tertiles) in the G2P network and classified them into highly pleiotropic HRPs and NonHRPs. (**a**) Enrichment of Gene Ontology (GO) biological processes among yeast highly pleiotropic HRPs and NonHRPs. Boxplots of distributions of the number of (**b**) genetic interactions, (**c**) protein-protein interactions, (**d**) link communities in genetic and protein-protein interaction network, (**e**) different GO terms a gene is associated with, reflecting its functionality, (**f**) genes whose expression is altered upon deletion of regulators (from the gene-perturbation network), (**g**) protein abundance, (**h**) translational rate and (**i**) protein half-life among highly pleiotropic HRPs and NonHRPs. Statistical significance was assessed using Wilcoxon rank sum test, corrected for multiple testing. **(j**) Sequence divergence among yeast highly pleiotropic HRPs and NonHRPs with their one-to-one orthologs in 73 fungal species belonging to 16 distinct fungal classes. Species names are abbreviated. Taxonomic and phylogenetic details of the fungal species are provided in Supplementary Notes. Median divergences of highly pleiotropic HRPs (red circle) and NonHRPs (grey circle) with their orthologs in each species are shown in the middle panel. The upper panel provides the statistical significance estimated by comparing the distribution of divergence of yeast highly pleiotropic HRPs and NonHRPs with their respective orthologs in each species. The black line shows a P-value cut-off of 0.01, corrected for multiple testing. The bottom panel shows the number of orthologs of yeast highly pleiotropic HRP (red) and NonHRP (grey) in each species. The average median difference between highly pleiotropic NonHRPs and HRPs is 5.2% which corresponds to ~15 missense variant positions for a 300 amino acid long protein. These findings suggest that highly pleiotropic proteins that contain a HR affect diverse functions through multiple interactions, are stringently regulated, and undergo accelerated divergence and hence may constitute the rapidly adaptable part of the proteome. In contrast, hNonHRPs with housekeeping functions seem to constitute the relatively slowly evolving core of the proteome.