Variable absorption of mutational trends by prion-forming domains during *Saccharomycetes* evolution

Paul M Harrison Corresp. 1

1 Dept. of Biology, McGill University, Montréal, Quebec, Canada

Corresponding Author: Paul M Harrison
Email address: paul.harrison@mcgill.ca

Prions are self-propagating alternative states of protein domains. They are linked to both diseases and functional protein roles in eukaryotes. Prion-forming domains in *Saccharomyces cerevisiae* are typically domains with high intrinsic protein disorder (*i.e.,* that remain unfolded in the cell most of the time), that are converted to self-replicating amyloid forms. It is still unclear what principles might govern the molecular evolution of prion-forming domains, and intrinsically disordered domains generally. Here, it is discovered that in a set of such prion-forming domains some evolve in the fungal class *Saccharomycetes* in such a way as to absorb general mutation biases across millions of years, whereas others do not, indicating a spectrum of selection pressures on composition and sequence. Thus, if the bias-absorbing prion formers are conserving a prion-forming capability, then this capability is not interfered with by the absorption of bias changes over vast evolutionary epochs. These results suggest methodology for assessing selection pressures on the composition of intrinsically disordered regions.

Variable absorption of mutational trends by prion-forming domains during *Saccharomycetes* evolution

Department of Biology,
McGill University,
Montreal, QC,
Canada.

Corresponding author:
Paul Harrison
Email: paul.harrison@mcgill.ca

Abstract:

Prions are self-propagating alternative states of protein domains. They are linked to both diseases and functional protein roles in eukaryotes. Prion-forming domains in *Saccharomyces cerevisiae* are typically domains with high intrinsic protein disorder (*i.e.,* that remain unfolded in the cell most of the time), that are converted to self-replicating amyloid forms. It is still unclear what principles might govern the molecular evolution of prion-forming domains, and intrinsically disordered domains generally. Here, it is discovered that in a set of such prion-forming domains some evolve in the fungal class *Saccharomycetes* in such a way as to absorb general mutation biases across millions of years, whereas others do not, indicating a spectrum of selection pressures on composition and sequence. Thus, if the bias-absorbing prion formers are conserving a prion-forming capability, then this capability is not interfered with by the absorption of bias changes over vast evolutionary epochs. These results suggest methodology for assessing selection pressures on the composition of intrinsically disordered regions.

Introduction

Prion formation and propagation have been studied extensively in the budding yeast *Saccharomyces cerevisiae*. The yeast *S. cerevisiae* has >200 prion-like proteins that have N/Q-rich domains of the sort observed in ≥8 known prion-formers (An et al. 2016). Such yeast prions have been linked to diverse phenomena including evolutionary capacitance, disease-like states, and large-scale genetic control. The first well-characterized yeast prions, that underlie the [PSI+] and [URE3] prions, are propagating amyloids of the proteins Sup35p and Ure2p respectively. The protein Sup35p is part of the translation termination complex. [PSI+] prion formation reduces translation termination efficiency and increases nonsense-codon read-through levels (Cox 1965;

Shorter & Lindquist 2005). This read-through has been shown to have a potential role in uncovering cryptic genetic variation (True et al. 2004; True & Lindquist 2000). [URE3] causes upregulation of poor nitrogen source usage, even when rich sources are available (Lacroute 1971; Wickner 1994; Wickner et al. 2004). Prion variants sometimes behave as budding-yeast diseases (McGlinchey et al. 2011; Nakayashiki et al. 2005). The [MOT3+] prion has been shown to have a possible role in control of transitions to multicellularity (Holmes et al. 2013). The stress-inducible cytoskeleton-linked budding-yeast protein Lsb2 (also known as Pin3) can form a metastable prion in response to high temperatures (Chernova et al. 2017a; Chernova et al. 2017b). There are now several known amyloid-based prions of *S. cerevisiae* (Harbi & Harrison 2014a; Harbi et al. 2012). Prion-forming proteins have also been discovered in the fungus *Podospora anserina* and the fission yeast *Schizosaccharomyces pombe* (Saupe 2011; Sideri et al. 2017). Amyloid-based budding yeast prion-forming regions tend to have high intrinsic disorder and a bias for asparagine (N) and/or glutamine (Q) residues (Harbi & Harrison 2014b). Several algorithms have been developed that annotate protein regions with high potential prion-forming propensity (Espinosa Angarica et al. 2013; Lancaster et al. 2014; Ross et al. 2013; Zambrano et al. 2015). Prion-like proteins in yeast and other organisms have more recently been linked to other processes, such as the formation of stress granules and other membraneless biomolecular condensates (Franzmann et al. 2018; Jain et al. 2016).

The original mammalian PrP domain is not biased for Ns and Qs, and is deeply conserved since a PrP ancestral gene emerged in early chordate evolution, likely through retrotransposition (Ehsani et al. 2011; Harrison et al. 2010; Westaway et al. 2011). The [PSI+] prion has an N/Q bias that is conserved across *Ascomycota* and *Basidiomycota,* which diverged >1 billion years ago (Harrison et al. 2007). A large population of yeast-prion-like proteins emerged *en masse* early in

Saccharomycetes evolution, as a result of mutational trends to form more polyasparagine runs, thus providing an evolutionary 'test set' from which several prion-forming domains seem to have developed (An et al. 2016). Prion-forming domains from *S. cerevisiae* tend to evolve more quickly as sequences than other prion-like domains but maintain their prion-like composition (Su & Harrison 2019). Eukaryotes often bear large numbers of these prion-like domains in their proteins. The slime mold *Dictyostelium* has >20% prion-like proteins (An & Harrison 2016; Malinovska et al. 2015), and there is evidence it has evolved a system to subvert prion formation (Malinovska & Alberti 2015; Malinovska et al. 2015). Other organisms such as *Drosophila melanogaster*, *Plasmodium falciparum* and the leech *Helobdella robusta* have high percentages of prion-like proteins in their proteomes (An & Harrison 2016; Pallares et al. 2018). In humans, several other yeast-prion-like proteins have links to neurodegeneration (Kim et al. 2013; Pokrishevsky et al. 2016; Sun et al. 2011). In *Aplysia* and *Drosophila*, such proteins have been linked to long-term memory formation (Khan et al. 2015; Si et al. 2010). Predicted prions can be observed in all the domains of life (Espinosa Angarica et al. 2013), including thousands in viruses and phages (Tetz & Tetz 2017; Tetz & Tetz 2018), and tens of thousands in bacteria (Harrison 2019; Iglesias et al. 2015). Bacterial prion-forming proteins have been detected experimentally (Molina-Garcia et al. 2018; Shahnawaz et al. 2017; Yuan et al. 2014; Yuan & Hochschild 2017). Bacterial prion-like proteins have a characteristic pattern of evolutionarily ancient, multi-phylum distribution coupled to sparse, intermittent conservation across their evolutionary range of species (Harrison 2019). About 5% of compositionally-biased dark matter (*i.e.*, regions that cannot be assigned as either structured or intrinsically disordered) in the known protein universe are predicted to be prion-like domains (Harrison 2018).

Here, the evolution of the sequences of prion-forming domains in *Saccharomycetes* is re-visited, but from the point of view of mutation biases. It is discovered that these protein regions have a spectrum of behaviour, variably absorbing mutation biases that are observable in the proteome as a whole, evidenced in the numbers of prion-like proteins, the % guanidine and cytidine (GC%) in the DNA, and the proportions of poly-asparagine and poly-glutamine.

Methods

Data

The UniProt (Boeckmann et al. 2003) set of reference fungal proteomes for *Saccharomycetes* (73 organisms) was downloaded from www.uniprot.org in June 2017. Sets of proteins with prion-forming domains (Data S1) and their orthologs across *Saccharomycetes* were collated as previously described (Su & Harrison 2019).

Prion-like composition

Prion-like composition in orthologs was calculated in two ways, firstly using the PLAAC prion-like domain annotation program (Lancaster et al. 2014), and secondly using the fLPS program for annotation of compositional biases (Harrison 2017). These were both run using default parameters, except that for fLPS the expected frequency for glutamine and asparagine residues was set equal to 0.05. For PLAAC, both the PRD score and the LLR score were analysed; the former is an indicator of the overall amount of prion-like composition in an annotated bounded prion-like region, while the latter indicates the prion-like sequence composition of the best sequence window (Lancaster et al. 2014). Any PLAAC score values that are negative or ‘N/A’ in the output from PLAAC are set equal to 0.0 for the purposes of this analysis.

Measures of proteome bias

Several measures of compositional bias across proteomes/genomes were examined:

(i) %N (asparagine) in the proteome;

(ii) %Q (glutamine) in the proteome;

(iii) % poly-N in the proteome (with a minimum tract length of 3);

(iv) % poly-Q in the proteome (with a minimum tract length of 3);

(v) % poly-Q + poly-N in the proteome (with a minimum tract length of 3);

(vi) %GC in the DNA;

(vii) The fraction of N/Q-rich proteins in the proteome according to a specific fLPS bias P-value threshold (either 1e-08, 1e-10 or 1e-12);

(viii) The fraction of proteins in the proteome with prion-like composition according to the program PLAAC (with PRD score >0.0, ≥15.0 or ≥30.0, or similarly for LLR score).

Correlations

Both weighted and unweighted Pearson correlation coefficients were calculated to assess the correlations of individual prion-like composition with the general trends in the proteome. Weightings for plot points were calculated according to their closest similarity with another protein, calculated as (1-%**I**/100), where %**I** is the percentage sequence identity in the most significant BLASTP sequence alignment (Altschul et al. 1997). These weightings were summed appropriately, as described in previous analyses (Harrison 2019; Su & Harrison 2019). Results indicate that the overall outcomes for specific proteins are not affected by non-usage of such weightings (see below).

131

132 Results

133 *Ure2 protein*

134 As an initial example, the evolutionary behaviour of compositional biases in the prion-
135 forming domain of Ure2p was examined (Figures 1-2). The current data indicate that an ancestor
136 of the Ure2p prion-forming domain with a strong N/Q-rich prion-like composition originated early
137 in *Saccharomycetes* evolution (at least in the last common ancestor of the diverse families
138 *Debaryomycetaceae* and *Saccharomycetaceae*), in agreement with results in previous publications
139 (An et al. 2016; Harrison et al. 2007) (Figure 3). In general, there is a strong correlation between
140 the degree of bias in the N/Q-rich region of Ure2p and the degree of compositional bias in the
141 whole proteome/genome by several indicators (%polyasparagine or
142 %[polyasparagine+polyglutamine] or DNA GC% or fraction of N/Q-rich prion-like proteins with
143 fLPS P-value $<10^{-10}$) (Figure 1). The correlations with PLAAC prion-like composition score are
144 lower, but both measures have strong correlations with %GC in DNA (Figure 2). Thus, during the
145 surge in formation of prion-like regions during *Saccharomycetes* evolution (An et al. 2016), the
146 degree of N-bias in the individual prion-former Ure2p also increased in correlation with the general
147 trend as it panned out across various sub-clades.

148

149 *Other prion-forming proteins*

150 Of the known amyloid-based prions—as well as Ure2p—Swi1p, Cyc8p and Sup35p have
151 domains of prion-like composition or N/Q bias that are widespread across *Saccharomycetes* (in
152 84% of orthologs for Cyc8p, 98% for Swi1p, and 90% for Sup35p; Table S1), with such domains
153 of these latter three also arising in other *Ascomycota* clades (An et al. 2016; Harrison et al. 2007).

Furthermore, Pin3 protein also has a widespread prion-like domain across *Saccharomycetes*, there being 52/55 (95%) *Saccharomycetes* Pin3 orthologs having PLAAC LLR scores >15.0. However, the degree of conservation of N/Q-rich bias *per se* is lower for this protein with 38/55 (75%) having a fLPS compositional bias P-value ≤1e–10. The metastable prion domain of Pin3 is the only known amyloid-based prion in *S. cerevisiae* to demonstrate very little correlation for its prion-like compositional biases, indicating some selection pressure for composition of a different sort, that nonetheless may preserve prion-forming ability.

The other three cases (Mot3p, Rnq1p and Nu100p) have either more recent ancestry as novel prion-like domains within *Saccharomycetes* (in the case of Mot3p and Rnq1p), or they arise sporadically in fungal species (Nu100p) (An et al. 2016; Su & Harrison 2019). These three are thus not expected to demonstrate many significant correlations with measures of compositional bias, but nonetheless we see a mild negative correlation for Rnq1p and Mot3p with %Q in the proteome, which is not typical of the other prion-forming proteins, suggesting selection pressures against Q bias in these evolutionarily recently emergent proteins.

In general, there are strong correlations for Ure2p, Swip and Cyc8p with %N, %poly-N, %GC in DNA and with the numbers of proteins with prion-like composition (Tables 4-5). Within these general trends, these four demonstrate a spectrum of responses to the overall proteome-wide mutational trends, with Ure2p being the strongest correlator. Sup35p stands out as an exception; it shows on the whole weaker correlations generally with %N and %poly-N, and stronger correlations with %poly-Q than the other three. This may be because there is selection pressure to maintain a specific proportion of Qs in specific local patterns or ratios (MacLea et al. 2015).

There is one species that is often a far outlier on the plots, *Ascoidea rubescens*, an uncharacterized species that is the sole member of the family *Ascoideaceae*, which is

177 geographically widely distributed and typically grows in beetle galleries in dead wood. It has a
178 very high proportion of poly-N-rich proteins (Tables 1-2). Removal of this outlier species from
179 the correlation analysis causes a substantial increase in correlations with %N and %poly-N, but
180 not for %GC in DNA.

181 Thus, the three *S. cerevisiae* prion-forming proteomes Ure2p, Cyc8p and Swi1p appear to
182 absorb the general mutational trends linked to the surge in formation of prion-like domains, that
183 was observed previously (An et al. 2016). This trend is linked to a general decrease in %GC in the
184 DNA (Tables 1-2).

185 Two other separately studied prion-forming domains are from New1p and Pub1p (Li et al.
186 2014; Osherovich & Weissman 2001). These are both strongly correlated proteome-bias absorbers,
187 with Pub1p (which is an interaction hub for other prion-like proteins (Harbi & Harrison 2014b))
188 uniquely amongst all of the prion-forming domains displaying a strong correlation for both poly-
189 N and poly-Q (Tables 1-2). Pub1p is strongly correlated despite having a low number of
190 orthologous prion domains that have high bias for N and Q residues (53% with fLPS P-value ≤1e–
191 10; Table S1) indicating that there is still correlated behavior for the weaker N/Q biases for this
192 protein. Other prion-forming domains observed in the analysis of Alberti, *et al.* (Alberti et al.
193 2009), also display a similar spectrum of bias absorption across *Saccharomycetes* evolution (Table
194 S2). Highly-correlated bias absorbers from this data whose prion-like domains are widespread in
195 *Saccharomycetes* include Lsm4p and Gln3p, whereas other widespread prion-like domains show
196 little or no correlation, such as Ngr1p (Tables S1, S2).

197 The above analysis uses the PLAAC PRDscore, to define the amount of prion-like
198 composition in a bounded region, and so reflecting more absorption of biases in a way analogous
199 to the working of the fLPS algorithm (Harrison 2017; Lancaster et al. 2014). The PLAAC log-

likelihood ratio (LLR) score has been used in the literature to pick out the most likely prion-forming sequence window within proteins (Alberti et al. 2009; An et al. 2016; Sideri et al. 2017; Tetz & Tetz 2018). Despite the restriction of a window of fixed size (41 amino-acid residues), these LLR scores also demonstrate a similar spectrum of bias absorption, with both strong and weak absorbers evident, albeit generally with less significance (Table S3).

It was checked whether the N/Q-rich regions are also rich in lysine, which is encoded by AT%-rich codons, like N (asparagine). Lysine has low prion formation propensity and charged residues are disruptive to prion formation and have low prion formation propensity (Lancaster et al. 2014; Osherovich & Weissman 2001). Lysine is a disorder-promoting residue (Oldfield & Dunker 2014) and some intrinsically disordered regions have high positive charge (Hatos et al. 2020; Necci et al. 2018). However, the N/Q-rich regions consistently in general have lower lysine content that the whole *Saccharomycetes* proteomes (Figure 4). Thus, these regions are not simply absorbing higher levels of AT% in their DNA through the embedding within them of amino-acid residues encoded by codons with high AT%.

Discussion

These results indicate that compositional aspects of many individual prion-formers behaved in a correlated way in relation to general trends as they panned out over millions of years across various sub-clades. Also, this surge in prion-like region formation is directly linked to a general trend for GC% decrease across the *Saccharomycetes* clade. However, some prion-forming domains resist the absorption of such mutational trends, such as the meta-stable prion-former Lsb2/Pin3 (Chernova et al. 2017b), despite it being as widely conserved as a protein as those that more easily absorb biases, such as Cyc8p and Swi1p. This suggests some greater selection pressure

on amino-acid composition. The Sup35p prion-forming domain also shows some special behavior: demonstrating a stronger correlation between overall proteome poly-Q levels and its own N or Q compositional bias as determined by the program fLPS. The Sup35 prion-forming domain has a subdomain with specific local patterns involving Q residues that is required for chaperone-dependent prion maintenance, that is separate from the N-terminal N/Q-rich region that is necessary for prion nucleation and fibre growth (MacLea et al. 2015). Also, the Sup35 prion-like domain has a more ancient origin before the last common ancestor of *Saccharomycetes*, and outside this clade it tends to have a predominant Q-bias that has been maintained within *Saccharomycetes*, resisting the trend for greater N-bias (An et al. 2016). However, this is also the behaviour of Cyc8p and Swi1p outside of *Saccharomycetes* (An et al. 2016), so this result is demonstrating an evolutionary behavior peculiar to Sup35p.

The Pub1p prion-forming domain shows strong correlations for both Q and N bias indicators. It is possible that proteins such as Pub1p that interact a lot with other prion-like proteins (Harbi & Harrison 2014b) 'need' to absorb more general compositional trends so that they can promiscuously bind with a large list of partners.

The results here provide a case study of mutational trend absorption by disordered regions generally. The results suggest some methodology for analyzing selection pressures on individual intrinsically disordered regions within the context of the behaviour of other sequences within the same proteome.

Conclusions

Thus, many prion-forming domains, and intrinsically disordered regions generally, are continually absorbing overall mutational trends in their proteomes, but this is modulated by

specific selection pressures. A spectrum of bias absorption is observed from Lsb2/Pin3---which shows little or no correlation---to Pub1, which shows very strong correlation to both asparagine- and glutamine-based biases.

Supplementary Materials

Data S1: FASTA-format file of the protein sequences of the proteins with prion-forming domains.

Table S1: Fraction of orthologs that have prion-like composition.

Table S2: Table of results for the other prion-forming proteins from ref. (Alberti et al. 2009).

Table S3: Correlations using PLAAC LLR score instead of PRD score.

Figure Legends

Figure 1: Correlation of various measures of mutational bias across proteomes versus the individual compositional bias in the Ure2p prion-forming domain, as judged by the fLPS program.

(a) Percentage of poly-N residues in the proteome.

(b) Percentage of (poly-N + poly-Q) residues in the proteome.

(c) DNA GC%.

(d) Fraction of N/Q-rich prion-like proteins with fLPS P-value <1e–10.

(e) Table of correlations and significances for plots (a) to (d).

Figure 2: As in Figure 1, except versus the individual PLAAC PRDscore in the Ure2p prion-forming domain.

(a) Percentage of poly-N residues in the proteome.

(b) Percentage of (poly-N + poly-Q) residues in the proteome.

(c) DNA GC%.

(d) Fraction of proteome with PLAAC score ≥15.0.

(e) Table of correlations and significances for plots (a) to (d).

Figure 3: Schematic evolutionary tree showing the distribution of orthologs with prion-like composition in different evolutionary families in the Uniprot reference set of fungal proteomes (Boeckmann et al. 2003). The organismal branching pattern from recent fungal phylogenies was used (Kurtzman & Robnett 2013; Shen et al. 2016). The number of species in

each family is given in brackets. The numbers of orthologs that are have fLPS P-value ≤1e-10 and PLAAC score ≥15.0 are listed in columns (Harrison 2017; Lancaster et al. 2014).

Figure 4: Correlations of %K (lysine residues) within the N/Q-rich regions of prion-forming proteins plotted versus the overall %K trend in proteomes. Blue points are for the set of known amyloid-based prions in Figure 4, and orange points for the total list of prion-forming domains including those listed in Table S2.

Table Legends

Table 1: Coloured table for a set of known prion-forming domains of the correlations (weighted and un-weighted) between the compositional bias (–log[fLPS P-value]), and a variety of parameters. Weighted correlations are the upper value in each cell, unweighted the lower value. Where removal of the common far outlier species *Ascoidea rubescens* causes increased significance for any correlation, the third and fourth rows in a cell display the correlation coefficients (in italics). For proteins which do not have an ortholog from *Ascoidea rubescens*, the name is labelled with '††'. If its removal causes no improvement in correlations, it is labelled with '†'. Correlations significant at ≤0.0005 are labelled *** and coloured green, significant at >0.0005 and ≤0.0016 labelled ** and coloured orange, and >0.0016, and ≤0.05 are labelled *). The threshold 0.0016 comes from a Bonferroni correction to allow for the fact that 31 sequences are being tested for a correlation against any specific proteome-wide property. In column one, the name is colour-coded according to the most significant correlation, with underlining if it is a 1-* correlation.

Table 2: Coloured table for a set of known prion-forming domains of the correlations (both weighted and un-weighted) between the prion-like composition (PLAAC PRDscore) and a variety of parameters. Weighted correlations are the upper value in each cell, unweighted the lower value. Where removal of the common far outlier species *Ascoidea rubescens* causes increased significance for any correlation, the third and fourth rows in a cell display the correlation coefficients (in italics). For proteins which do not have an ortholog from *Ascoidea rubescens*, the name is labelled with '††'. If its removal causes no improvement in correlations, it is labelled with '†'. Correlations significant at ≤0.0005 are labelled *** and coloured green, significant at >0.0005 and ≤0.0016 labelled ** and coloured orange, and >0.0016, and ≤0.05 are labelled *). The threshold 0.0016 comes from a Bonferroni correction to allow for the fact that 31 sequences are being tested for a correlation against any specific proteome-wide property. In column one, the name is colour-coded according to the most significant correlation, with underlining if it is a 1-* correlation.

References

Alberti S, Halfmann R, King O, Kapila A, and Lindquist S. 2009. A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. *Cell* 137:146-158.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, and Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389-3402.

An L, Fitzpatrick D, and Harrison PM. 2016. Emergence and evolution of yeast prion and prion-like proteins. *BMC Evol Biol* 16:24. 10.1186/s12862-016-0594-3

An L, and Harrison PM. 2016. The evolutionary scope and neurological disease linkage of yeast-prion-like proteins in humans. *Biol Direct* 11:32. 10.1186/s13062-016-0134-5

Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, Phan I, Pilbout S, and Schneider M. 2003. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res* 31:365-370.

Chernova TA, Chernoff YO, and Wilkinson KD. 2017a. Prion-based memory of heat stress in yeast. *Prion* 11:151-161. 10.1080/19336896.2017.1328342

Chernova TA, Kiktev DA, Romanyuk AV, Shanks JR, Laur O, Ali M, Ghosh A, Kim D, Yang Z, Mang M, Chernoff YO, and Wilkinson KD. 2017b. Yeast Short-Lived Actin-Associated Protein Forms a Metastable Prion in Response to Thermal Stress. *Cell Rep* 18:751-761. 10.1016/j.celrep.2016.12.082

Cox B. 1965. [PSI], a cytoplasmic suppressor of super-suppression in yeast. *Heredity* 20:505-521.

Ehsani S, Tao R, Pocanschi CL, Ren H, Harrison PM, and Schmitt-Ulms G. 2011. Evidence for retrogene origins of the prion gene family. *PLoS One* 6:e26800. 10.1371/journal.pone.0026800

Espinosa Angarica V, Ventura S, and Sancho J. 2013. Discovering putative prion sequences in complete proteomes using probabilistic representations of Q/N-rich domains. *BMC Genomics* 14:316. 10.1186/1471-2164-14-316

Franzmann TM, Jahnel M, Pozniakovsky A, Mahamid J, Holehouse AS, Nuske E, Richter D, Baumeister W, Grill SW, Pappu RV, Hyman AA, and Alberti S. 2018. Phase separation of a yeast prion protein promotes cellular fitness. *Science* 359. 10.1126/science.aao5654

353 Harbi D, and Harrison PM. 2014a. Classifying prion and prion-like phenomena. *Prion* 8.
354 Harbi D, and Harrison PM. 2014b. Interaction networks of prion, prionogenic and prion-like
355 proteins in budding yeast, and their role in gene regulation. *PLoS One* 9:e100615.
356 10.1371/journal.pone.0100615
357 Harbi D, Parthiban M, Gendoo DM, Ehsani S, Kumar M, Schmitt-Ulms G, Sowdhamini R, and
358 Harrison PM. 2012. PrionHome: a database of prions and other sequences relevant to
359 prion phenomena. *PLoS One* 7:e31785.
360 Harrison LB, Yu Z, Stajich JE, Dietrich FS, and Harrison PM. 2007. Evolution of budding yeast
361 prion-determinant sequences across diverse fungi. *J Mol Biol* 368:273-282.
362 Harrison PM. 2017. fLPS: Fast discovery of compositional biases for the protein universe. *BMC*
363 *Bioinformatics* 18:476. 10.1186/s12859-017-1906-3
364 Harrison PM. 2018. Compositionally biased dark matter in the protein universe.
365 *Proteomics*:e1800069. 10.1002/pmic.201800069
366 Harrison PM. 2019. Evolutionary behaviour of bacterial prion-like proteins. *PLoS One*
367 14:e0213030. 10.1371/journal.pone.0213030
368 Harrison PM, Khachane A, and Kumar M. 2010. Genomic assessment of the evolution of the
369 prion protein gene family in vertebrates. *Genomics* 95:268-277.
370 10.1016/j.ygeno.2010.02.008
371 Hatos A, Hajdu-Soltesz B, Monzon AM, Palopoli N, Alvarez L, Aykac-Fas B, Bassot C, Benitez
372 GI, Bevilacqua M, Chasapi A, Chemes L, Davey NE, Davidovic R, Dunker AK, Elofsson
373 A, Gobeill J, Foutel NSG, Sudha G, Guharoy M, Horvath T, Iglesias V, Kajava AV,
374 Kovacs OP, Lamb J, Lambrughi M, Lazar T, Leclercq JY, Leonardi E, Macedo-Ribeiro
375 S, Macossay-Castillo M, Maiani E, Manso JA, Marino-Buslje C, Martinez-Perez E,

376 Meszaros B, Micetic I, Minervini G, Murvai N, Necci M, Ouzounis CA, Pajkos M,
377 Paladin L, Pancsa R, Papaleo E, Parisi G, Pasche E, Barbosa Pereira PJ, Promponas VJ,
378 Pujols J, Quaglia F, Ruch P, Salvatore M, Schad E, Szabo B, Szaniszlo T, Tamana S,
379 Tantos A, Veljkovic N, Ventura S, Vranken W, Dosztanyi Z, Tompa P, Tosatto SCE, and
380 Piovesan D. 2020. DisProt: intrinsic protein disorder annotation in 2020. *Nucleic Acids*
381 *Res* 48:D269-D276. 10.1093/nar/gkz975
382 Holmes DL, Lancaster AK, Lindquist S, and Halfmann R. 2013. Heritable remodeling of yeast
383 multicellularity by an environmentally responsive prion. *Cell* 153:153-165.
384 Iglesias V, de Groot NS, and Ventura S. 2015. Computational analysis of candidate prion-like
385 proteins in bacteria and their role. *Front Microbiol* 6:1123. 10.3389/fmicb.2015.01123
386 Jain S, Wheeler JR, Walters RW, Agrawal A, Barsic A, and Parker R. 2016. ATPase-Modulated
387 Stress Granules Contain a Diverse Proteome and Substructure. *Cell* 164:487-498.
388 10.1016/j.cell.2015.12.038
389 Khan MR, Li L, Perez-Sanchez C, Saraf A, Florens L, Slaughter BD, Unruh JR, and Si K. 2015.
390 Amyloidogenic Oligomerization Transforms Drosophila Orb2 from a Translation
391 Repressor to an Activator. *Cell* 163:1468-1483. 10.1016/j.cell.2015.11.020
392 Kim HJ, Kim NC, Wang YD, Scarborough EA, Moore J, Diaz Z, MacLea KS, Freibaum B, Li S,
393 Molliex A, Kanagaraj AP, Carter R, Boylan KB, Wojtas AM, Rademakers R, Pinkus JL,
394 Greenberg SA, Trojanowski JQ, Traynor BJ, Smith BN, Topp S, Gkazi AS, Miller J,
395 Shaw CE, Kottlors M, Kirschner J, Pestronk A, Li YR, Ford AF, Gitler AD, Benatar M,
396 King OD, Kimonis VE, Ross ED, Weihl CC, Shorter J, and Taylor JP. 2013. Mutations in
397 prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and
398 ALS. *Nature* 495:467-473. 10.1038/nature11922

399 Kurtzman CP, and Robnett CJ. 2013. Relationships among genera of the Saccharomycotina
400 (Ascomycota) from multigene phylogenetic analysis of type species. *FEMS Yeast Res*
401 13:23-33. 10.1111/1567-1364.12006
402 Lacroute F. 1971. Non-Mendelian mutation allowing ureidosuccinic acid uptake in yeast. *J*
403 *Bacteriol* 106:519-522.
404 Lancaster AK, Nutter-Upham A, Lindquist S, and King OD. 2014. PLAAC: a web and
405 command-line application to identify proteins with prion-like amino acid composition.
406 *Bioinformatics* 30:2501-2502. 10.1093/bioinformatics/btu310
407 Li X, Rayman JB, Kandel ER, and Derkatch IL. 2014. Functional role of Tia1/Pub1 and Sup35
408 prion domains: directing protein synthesis machinery to the tubulin cytoskeleton. *Mol*
409 *Cell* 55:305-318. 10.1016/j.molcel.2014.05.027
410 MacLea KS, Paul KR, Ben-Musa Z, Waechter A, Shattuck JE, Gruca M, and Ross ED. 2015.
411 Distinct amino acid compositional requirements for formation and maintenance of the
412 [PSI(+)] prion in yeast. *Mol Cell Biol* 35:899-911. 10.1128/MCB.01020-14
413 Malinovska L, and Alberti S. 2015. Protein misfolding in Dictyostelium: Using a freak of nature
414 to gain insight into a universal problem. *Prion* 9:339-346.
415 10.1080/19336896.2015.1099799
416 Malinovska L, Palm S, Gibson K, Verbavatz JM, and Alberti S. 2015. Dictyostelium discoideum
417 has a highly Q/N-rich proteome and shows an unusual resilience to protein aggregation.
418 *Proc Natl Acad Sci U S A* 112:E2620-2629. 10.1073/pnas.1504459112
419 McGlinchey RP, Kryndushkin D, and Wickner RB. 2011. Suicidal [PSI+] is a lethal yeast prion.
420 *Proc Natl Acad Sci U S A* 108:5337-5341.

421 Molina-Garcia L, Gasset-Rosa F, Alamo MM, de la Espina SM, and Giraldo R. 2018.
422 Addressing Intracellular Amyloidosis in Bacteria with RepA-WH1, a Prion-Like Protein.
423 *Methods Mol Biol* 1779:289-312. 10.1007/978-1-4939-7816-8_18
424 Nakayashiki T, Kurtzman CP, Edskes HK, and Wickner RB. 2005. Yeast prions [URE3] and
425 [PSI+] are diseases. *Proc Natl Acad Sci U S A* 102:10575-10580.
426 Necci M, Piovesan D, Dosztanyi Z, Tompa P, and Tosatto SCE. 2018. A comprehensive
427 assessment of long intrinsic protein disorder from the DisProt database. *Bioinformatics*
428 34:445-452. 10.1093/bioinformatics/btx590
429 Oldfield CJ, and Dunker AK. 2014. Intrinsically disordered proteins and intrinsically disordered
430 protein regions. *Annu Rev Biochem* 83:553-584. 10.1146/annurev-biochem-072711-
431 164947
432 Osherovich LZ, and Weissman JS. 2001. Multiple Gln/Asn-rich prion domains confer
433 susceptibility to induction of the yeast [PSI(+)] prion. *Cell* 106:183-194. 10.1016/s0092-
434 8674(01)00440-8
435 Pallares I, de Groot NS, Iglesias V, Sant'Anna R, Biosca A, Fernandez-Busquets X, and Ventura
436 S. 2018. Discovering Putative Prion-Like Proteins in Plasmodium falciparum: A
437 Computational and Experimental Analysis. *Front Microbiol* 9:1737.
438 10.3389/fmicb.2018.01737
439 Pokrishevsky E, Grad LI, and Cashman NR. 2016. TDP-43 or FUS-induced misfolded human
440 wild-type SOD1 can propagate intercellularly in a prion-like fashion. *Sci Rep* 6:22155.
441 10.1038/srep22155

442 Ross ED, Maclea KS, Anderson C, and Ben-Hur A. 2013. A bioinformatics method for
443 identifying Q/N-rich prion-like domains in proteins. *Methods Mol Biol* 1017:219-228.
444 10.1007/978-1-62703-438-8_16
445 Saupe SJ. 2011. The [Het-s] prion of Podospora anserina and its role in heterokaryon
446 incompatibility. *Semin Cell Dev Biol* 22:460-468. 10.1016/j.semcdb.2011.02.019
447 Shahnawaz M, Park KW, Mukherjee A, Diaz-Espinoza R, and Soto C. 2017. Prion-like
448 characteristics of the bacterial protein Microcin E492. *Sci Rep* 7:45720.
449 10.1038/srep45720
450 Shen XX, Zhou X, Kominek J, Kurtzman CP, Hittinger CT, and Rokas A. 2016. Reconstructing
451 the Backbone of the Saccharomycotina Yeast Phylogeny Using Genome-Scale Data. *G3*
452 *(Bethesda)* 6:3927-3939. 10.1534/g3.116.034744
453 Shorter J, and Lindquist S. 2005. Prions as adaptive conduits of memory and inheritance. *Nat*
454 *Rev Genets* 6:435-450.
455 Si K, Choi YB, White-Grindley E, Majumdar A, and Kandel ER. 2010. Aplysia CPEB can form
456 prion-like multimers in sensory neurons that contribute to long-term facilitation. *Cell*
457 140:421-435. 10.1016/j.cell.2010.01.008
458 Sideri T, Yashiroda Y, Ellis DA, Rodriguez-Lopez M, Yoshida M, Tuite MF, and Bahler J.
459 2017. The copper transport-associated protein Ctr4 can form prion-like epigenetic
460 determinants in Schizosaccharomyces pombe. *Microb Cell* 4:16-28.
461 10.15698/mic2017.01.552
462 Su TY, and Harrison PM. 2019. Conservation of Prion-Like Composition and Sequence in Prion-
463 Formers and Prion-Like Proteins of Saccharomyces cerevisiae. *Front Mol Biosci* 6:54.
464 10.3389/fmolb.2019.00054

Sun Z, Diaz Z, Fang X, Hart MP, Chesi A, Shorter J, and Gitler AD. 2011. Molecular determinants and genetic modifiers of aggregation and toxicity for the ALS disease protein FUS/TLS. *PLoS Biol* 9:e1000614. 10.1371/journal.pbio.1000614

Tetz G, and Tetz V. 2017. Prion-Like Domains in Phagobiota. *Front Microbiol* 8:2239. 10.3389/fmicb.2017.02239

Tetz G, and Tetz V. 2018. Prion-like Domains in Eukaryotic Viruses. *Sci Rep* 8:8931. 10.1038/s41598-018-27256-w

True H, Berlin I, and Lindquist S. 2004. Epigenetic regulation of translation reveals hidden genetic variation to produce comlex traits. *Nature* 431:184-187.

True H, and Lindquist S. 2000. A yeast prion provides a mechanism for genetic variation and phenotypic diversity. *Nature* 407:477-483.

Westaway D, Daude N, Wohlgemuth S, and Harrison P. 2011. The PrP-like proteins Shadoo and Doppel. *Top Curr Chem* 305:225-256. 10.1007/128_2011_190

Wickner R. 1994. [URE3] as an altered URE2 protein: evidence for a prion analog in Saccharomyces cerevisiae. *Science* 264:528-530.

Wickner R, Edskes H, Roberts B, Baxa U, Pierce M, Ross E, and Brachmann A. 2004. Prions: proteins as genes and infectious entities. *Genes Dev* 18:470-485.

Yuan AH, Garrity SJ, Nako E, and Hochschild A. 2014. Prion propagation can occur in a prokaryote and requires the ClpB chaperone. *Elife* 3:e02949. 10.7554/eLife.02949

Yuan AH, and Hochschild A. 2017. A bacterial global regulator forms a prion. *Science* 355:198-201. 10.1126/science.aai7776

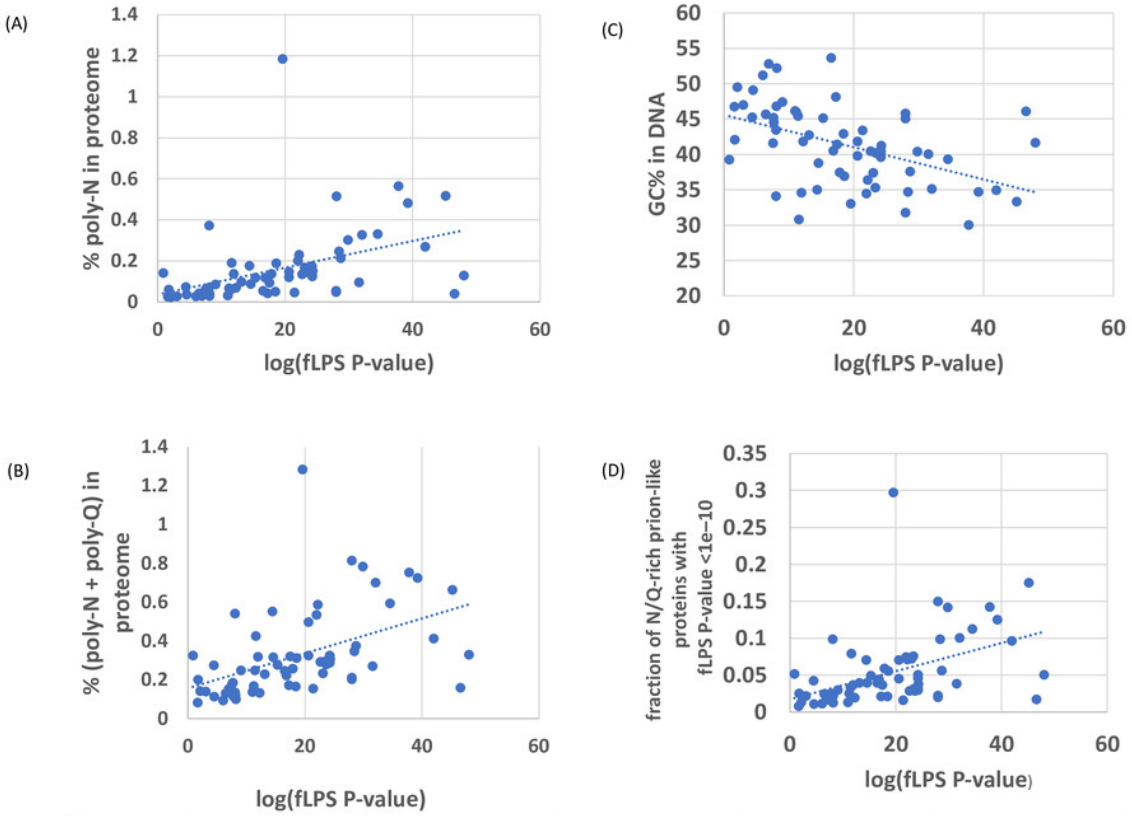
Zambrano R, Conchillo-Sole O, Iglesias V, Illa R, Rousseau F, Schymkowitz J, Sabate R, Daura X, and Ventura S. 2015. PrionW: a server to identify proteins containing

488 glutamine/asparagine rich prion-like domains and their amyloid cores. *Nucleic Acids Res*
489 43:W331-337. 10.1093/nar/gkv490
490

Figure 1

Correlation of various measures of mutational bias across proteomes versus the individual compositional bias in the Ure2p prion-forming domain, as judged by the fLPS program.

(a) Percentage of poly-N residues in the proteome. (b) Percentage of (poly-N + poly-Q) residues in the proteome. (c) DNA GC%. (d) Fraction of N/Q-rich prion-like proteins with fLPS P-value <1e–10. (e) Table of correlations and significances for plots (a) to (d).



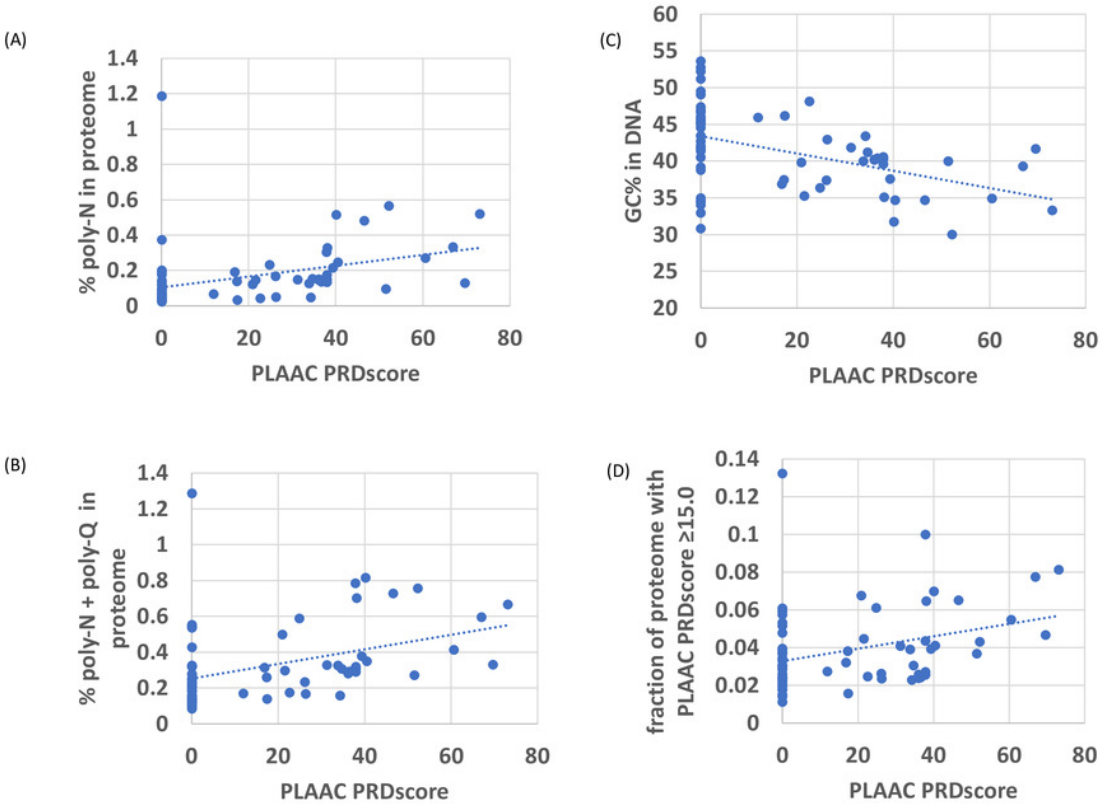
(E)

	Percentage of polyN in the proteome	Percentage of (polyN+polyQ) in the proteome	Fraction of prion-like domains fLPS P-value ≤10^{-10}	GC% in DNA
Log(fLPS P-value)	0.468 (P=0.00074)	0.527 (P<0.00001)	0.535 (P<0.00001)	−0.570 (P<0.00001)

Figure 2

As in Figure 1, except versus the individual PLAAC PRDscore in the Ure2p prion-forming domain.

(a) Percentage of poly-N residues in the proteome. (b) Percentage of (poly-N + poly-Q) residues in the proteome. (c) DNA GC%. (d) Fraction of proteome with PLAAC score ≥15.0. (e) Table of correlations and significances for plots (a) to (d).



(E)

	Fraction of polyN in proteome	Fraction of (polyN+polyQ)	Fraction of prion-like domains fPLAAC PRD score ≥15.0	GC% in DNA
Log(fLPS P-value)	0.388 (P=0.0013)	0.441 (P=0.00021)	0.424 (P=0.00039)	−0.539 (P<0.00001)

Figure 3

Schematic evolutionary tree showing the distribution of orthologs with prion-like composition in different evolutionary families in the Uniprot reference set of fungal proteomes.

The organismal branching pattern from recent fungal phylogenies was used (Kurtzman & Robnett 2013; Shen et al. 2016) . The number of species in each family is given in brackets. The numbers of orthologs that are have fLPS P-value ≤1e-10 and PLAAC score ≥15.0 are listed in columns (Harrison 2017; Lancaster et al. 2014) .

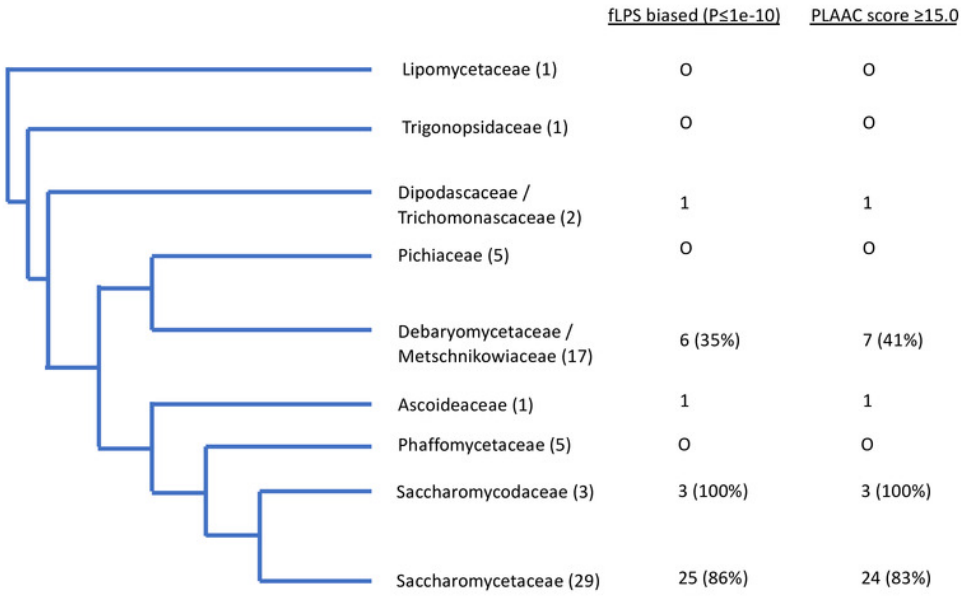


Table 1(on next page)

Coloured table for a set of known prion-forming domains of the correlations (weighted and un-weighted) between the compositional bias (–log[fLPS P-value]), and a variety of parameters.

Weighted correlations are the upper value in each cell, unweighted the lower value. Where removal of the common far outlier species *Ascoidea rubescens* causes increased significance for any correlation, the third and fourth rows in a cell display the correlation coefficients (in italics). For proteins which do not have an ortholog from *Ascoidea rubescens*, the name is labelled with '††'. If its removal causes no improvement in correlations, it is labelled with '†'. Correlations significant at ≤0.0005 are labelled *** and coloured green, significant at >0.0005 and ≤0.0016 labelled ** and coloured orange, and >0.0016, and ≤0.05 are labelled *). The threshold 0.0016 comes from a Bonferroni correction to allow for the fact that 31 sequences are being tested for a correlation against any specific proteome-wide property. In column one, the name is colour-coded according to the most significant correlation, with underlining if it is a 1-* correlation.

Protein (Number of orthologs in brackets)	%N in proteome	%Q in proteome	%poly-N	%poly-Q	%poly-Q+%poly-N	DNA GC%	Fraction of N/Q-rich proteins in the proteome by fLPS bias threshold		
							Threshold 1e-08	Threshold 1e-10	Threshold 1e-12
Known amyloid-based prions in *S. cerevisiae*									
Sup35 **P05453 (62)**	0.237 0.136 *0.350 ** *0.316 **	0.042 0.060 *0.019* *0.021*	0.132 0.035 *0.315 ** *0.307 **	0.415 ** 0.380 ** *0.411 *** *0.370 **	0.278 ** 0.180 *0.409 *** *0.385 **	–0.351 ** –0.263 ** *–0.388 ** *–0.323 **	0.218 0.142 *0.348 ** *0.348 **	0.202 0.119 *0.353 ** *0.356 **	0.187 0.095 *0.366 ** *0.368 **
Swi1 †† **P09547 (56)**	0.661 *** 0.628 ***	–0.149 –0.184	0.603 *** 0.570 ***	0.074 0.016	0.544 *** 0.473 ***	–0.498 *** –0.510 ***	0.643 *** 0.600 ***	0.627 *** 0.586 ***	0.607 *** 0.568 ***
Cyc8 **P14922 (61)**	0.387 ** 0.251 *0.522 **** *0.368 **	0.292 ** 0.320 ** *0.278 ** *0.307 **	0.320 ** 0.165 *0.577 **** *0.350 **	0.361 ** 0.305 ** *0.354 ** *0.297 **	0.409 ** 0.254 ** *0.567 **** *0.382 **	–0.472 *** –0.305 ** *–0.507 **** *–0.334 **	0.398 ** 0.240 *0.563 **** *0.374 **	0.385 ** 0.234 *0.581 **** *0.394 **	0.364 ** 0.225 *0.595 **** *0.418 **
Ure2 **P23202 (66)**	0.571 *** 0.485 *** *0.682 **** *0.566 ****	0.241 0.253 ** *0.246 ** *0.259 **	0.468 *** 0.420 *** *0.676 **** *0.590 ****	0.357 ** 0.330 ** *0.361 ** *0.332 **	0.527 *** 0.476 *** *0.651 **** *0.563 ****	–0.570 *** –0.478 *** *–0.584 **** *–0.484 ****	0.556 *** 0.470 *** *0.687 **** *0.568 ****	0.535 *** 0.453 *** *0.696 **** *0.576 ****	0.495 *** 0.423 *** *0.690 **** *0.572 ****
Rnq1 † **P25367 (26)**	0.139 0.230	–0.381 –0.431 *	0.096 0.159	–0.193 –0.197	0.037 0.090	–0.080 –0.159	0.081 0.070	0.053 0.040	0.010 0.001
Mot3 † **P54785 (25)**	0.460 ** 0.393	–0.420 ** –0.507 **	0.395 0.264	0.371 0.268	0.439 ** 0.299	–0.468 ** –0.409 **	0.385 0.140	0.386 0.129	0.399 ** 0.146
Nu100 † **Q02629 (11)**	0.154 0.224	0.107 0.148	–0.110 –0.058	–0.105 0.013	–0.518 –0.499	–0.008 –0.090	–0.012 –0.017	–0.027 –0.030	–0.047 –0.042
Pin3 † **Q06449 (55)**	0.198 0.179	0.022 –0.030	0.230 0.200	0.000 –0.014	–0.121 –0.046	–0.183 –0.169	0.209 0.165	0.198 0.153	0.179 0.137
Other prion-forming domains discussed in the text									
New1 †† **Q08972 (63)**	0.566 *** 0.521 ***	0.269 ** 0.261 **	0.476 *** 0.442 ***	0.191 0.188	0.482 *** 0.439 ***	–0.482 *** –0.449 ***	0.513 *** 0.468 ***	0.501 *** 0.458 ***	0.486 *** 0.446 ***
Pub1 **P32588 (62)**	0.469 *** 0.457 *** *0.466 **** *0.450 ****	0.365 ** 0.243 *0.401 *** *0.278 **	0.484 *** 0.426 *** *0.551 **** *0.459 ****	0.707 *** 0.620 *** *0.734 **** *0.646 ****	0.686 *** 0.597 *** *0.728 **** *0.622 ****	–0.547 *** –0.532 *** *–0.534 **** *–0.518 ****	0.545 *** 0.449 *** *0.567 **** *0.447 ****	0.545 *** 0.448 *** *0.584 **** *0.459 ****	0.533 *** 0.442 *** *0.594 **** *0.471 ****

Table 2(on next page)

Coloured table for a set of known prion-forming domains of the correlations (both weighted and un-weighted) between the prion-like composition (PLAAC PRDscore) and a variety of parameters.

Weighted correlations are the upper value in each cell, unweighted the lower value. Where removal of the common far outlier species *Ascoidea rubescens* causes increased significance for any correlation, the third and fourth rows in a cell display the correlation coefficients (in italics). For proteins which do not have an ortholog from *Ascoidea rubescens*, the name is labelled with '††'. If its removal causes no improvement in correlations, it is labelled with '†'. Correlations significant at ≤0.0005 are labelled *** and coloured green, significant at >0.0005 and ≤0.0016 labelled ** and coloured orange, and >0.0016, and ≤0.05 are labelled *). The threshold 0.0016 comes from a Bonferroni correction to allow for the fact that 31 sequences are being tested for a correlation against any specific proteome-wide property. In column one, the name is colour-coded according to the most significant correlation, with underlining if it is a 1-* correlation.

Protein (Number of orthologs in brackets)	%N in proteome	%Q in proteome	%poly-N	%poly-Q	%poly-Q+%poly-N	DNA GC%	Fraction of prion-like proteins in the proteome by PLAAC PRDscore		
							≥0.0	≥15.0	≥30.0
Known amyloid-based prions in *S. cerevisiae*									
Sup35 P05453 (62)	0.292 * 0.160 *0.457 **** *0.407 **	0.268 * 0.254 * *0.245* *0.215*	0.174 0.040 *0.437 **** *0.411 **	0.423 ** 0.372 * *0.421 *** *0.363 *	0.313 * 0.181 *0.497 **** *0.454 ****	–0.345 * –0.252 * *–0.401 *** *–0.336 *	0.429 *** 0.307 * *0.574 **** *0.528 ****	0.369 * 0.224 *0.560 **** *0.506 ****	0.273 * 0.108 *0.525 **** *0.461 ****
Swi1 †† P09547 (56)	0.475 *** 0.465 ***	–0.206 –0.200	0.451 *** 0.431 **	0.074 0.054	0.414 ** 0.375 *	–0.471 *** –0.470 ***	0.460 *** 0.443 **	0.464 *** 0.441 **	0.442 ** 0.411 *
Cyc8 P14922 (61)	0.353 * 0.244 *0.453 **** *0.328 *	0.250 0.285 * *0.242* *0.279 *	0.325 * 0.183 *0.535 **** *0.324 *	0.421 ** 0.356 * *0.417 *** *0.353 *	0.438 *** 0.288 * *0.569 **** *0.389 *	–0.458 *** –0.301 * *–0.482 **** *–0.319 *	0.563 *** 0.462 *** *0.645 **** *0.544 ****	0.486 *** 0.385 * *0.608 **** *0.503 ****	0.375 * 0.274 * *0.548 **** *0.429 **
Ure2 P23202 (66)	0.495 *** 0.448 *** *0.683 **** *0.594 ****	0.151 0.087 *0.130* *0.071*	0.388 ** 0.369 * *0.704 **** *0.631 ****	0.308 * 0.246 * *0.297 ** *0.239*	0.441 *** 0.401 ** *0.645 **** *0.548 ****	–0.539 *** –0.453 *** *–0.586 **** *–0.483 ****	0.494 *** 0.388 ** *0.627 **** *0.484 ****	0.424 *** 0.333 * *0.615 **** *0.465 ****	0.314 * 0.226 *0.567 **** *0.393 **
Rnq1 † P25367 (26)	–0.001 0.079	–0.264 –0.340	–0.046 0.005	–0.267 –0.242	–0.107 –0.058	0.035 –0.027	–0.128 –0.137	–0.128 –0.139	–0.133 –0.199
Mot3 † P54785 (25)	0.149 0.153	–0.166 –0.336	0.135 0.057	0.314 0.213	0.196 0.103	–0.159 –0.172	0.114 –0.107	0.212 –0.014	0.283 –0.046
Nu100 † Q02629 (11)	0.090 0.169	0.283 0.303	–0.185 –0.122	–0.030 0.075	–0.165 –0.083	–0.003 –0.084	0.304 0.290	0.142 0.160	0.021 –0.024
Pin3 † Q06449 (55)	0.000 –0.010	0.282 * 0.226	0.025 0.005	0.222 0.202	0.121 0.100	–0.081 –0.067	0.112 0.006	0.113 0.029	0.159 0.056
Other prion-forming domains discussed in the text									
New1 †† Q08972 (63)	0.368 * 0.339 *	0.236 0.250 *	0.301 * 0.288 *	0.183 0.226	0.326 * 0.327 *	–0.369 * –0.368 *	0.412 * 0.419 *	0.377 * 0.380 *	0.299 * 0.291 *
Pub1 P32588 (62)	0.226 0.241 *0.232* *0.247*	0.521 *** 0.424 ** *0.540 **** *0.443 ****	0.300 * 0.255 * *0.381 ** *0.290 *	0.756 *** 0.679 *** *0.771 **** *0.703 ****	0.559 *** 0.479 *** *0.628 **** *0.535 ****	–0.279 * –0.309 * *–0.274 ** *–0.303 *	0.597 *** 0.504 *** *0.631 **** *0.532 ****	0.605 *** 0.509 *** *0.674 **** *0.570 ****	0.570 *** 0.465 *** *0.695 **** *0.571 ****

Figure 4

Correlations of %K (lysine residues) within the N/Q-rich regions of prion-forming proteins plotted versus the overall %K trend in proteomes.

Blue points are for the set of known amyloid-based prions in Figure 4, and orange points for the total list of prion-forming domains including those listed in Table S2.