Conserved mRNA-binding proteomes in eukaryotic organisms

Ana M Matia-González, Emma E Laing & André P Gerber

RNA-binding proteins (RBPs) are essential for post-transcriptional regulation of gene expression. Recent high-throughput screens have dramatically increased the number of experimentally identified RBPs; however, comprehensive identification of RBPs within living organisms is elusive. Here we describe the repertoire of 765 and 594 proteins that reproducibly interact with polyadenylated mRNAs in *Saccharomyces cerevisiae* and *Caenorhabditis elegans*, respectively. Furthermore, we report the differential association of mRNA-binding proteins (mRBPomes) upon induction of apoptosis in *C. elegans* L4-stage larvae. Strikingly, most proteins composing mRBPomes, including components of early metabolic pathways and the proteasome, are evolutionarily conserved between yeast and *C. elegans*. We speculate, on the basis of our evidence that glycolytic enzymes bind distinct glycolytic mRNAs, that enzyme-mRNA interactions relate to an ancient mechanism for post-transcriptional coordination of metabolic pathways that perhaps was established during the transition from the early 'RNA world' to the 'protein world'.

RBPs are key players in post-transcriptional gene regulation, performing essential functions that maintain cellular fitness^{1,2}. In particular, mRBPs mediate the processing of mRNA precursors in the nucleus, the export and localization of mRNAs to distinct subcellular regions in the cytoplasm and the translation and eventual decay of mRNAs¹. RBPs often bear characteristic RNA-binding domains (RBDs), such as the RNA-recognition motif (RRM), the hnRNP K homology (KH) domain or zinc-finger (ZnF) domains; many of these domains are believed to have originated at early stages of evolution^{1,3}. On the basis of these RBDs, hundreds of RBPs have been predicted to exist in eukaryotic organisms *in silico*³. However, recent systematic approaches to experimentally map mRBPs in yeast and human cultured cells, by use of protein microarrays^{4,5} or capturing of *in vivo* cross-linked mRNA-protein complexes and subsequent MS^{6–9}, have suggested that many proteins lacking canonical RBDs bind RNAs; these proteins include those with other well-established cellular functions, such as metabolic enzymes. Thus, to gain a full system-wide understanding of post-transcriptional regulation, it is critical that mRBPs be characterized *in vivo*. Although unicellular eukaryotes such as yeast exhibit many of the regulatory pathways seen in higher eukaryotes, identification of mRBPs in a living animal is vital to revealing additional key regulatory factors that mediate the post-transcriptional control essential for development and physiology.

Thus, to provide the first comprehensive catalog of mRBPs in the unicellular yeast *S. cerevisiae* and the multicellular nematode *C. elegans*, we developed a technique for the *in vivo* capture of mRNA-protein complexes. We observed dynamic changes of the mRNA-binding proteome (mRBPome) upon apoptotic stress in *C. elegans*. We also found remarkable conservation of mRBPomes between organisms. This conservation included many, if not all, components

of early metabolic pathways and protein complexes, thus suggesting an ancient origin of the mRBPome. Finally, we demonstrated that yeast glycolytic proteins can bind to their own and other glycolytic mRNAs, thus potentially establishing an early RNA regulon, which, we hypothesize, could create an additional layer for coordinated control of pathway activity to achieve metabolic control.

RESULTS

Defining the *S. cerevisiae* mRBPome

To identify the sets of proteins interacting with mRNAs within living unicellular (yeast) and multicellular (*C. elegans*) eukaryotic organisms, we adapted a protocol recently applied to cultured human and starved yeast cells^{6–8}. We grew *S. cerevisiae* cells in rich medium to mid-log phase and cross-linked proteins *in vivo* to nucleic acids by UV irradiation at 254 nm under conditions that preserved the integrity of RNA (Supplementary Fig. 1a); we subsequently purified poly(A)-containing RNAs from cell lysates via oligo(dT)₂₅ beads under stringent washing conditions (Online Methods). Because UV exposure also cross-links proteins to DNA¹⁰, we controlled for RNA-dependent binding of proteins by treating extracts with RNase ONE and tested for selective binding to poly(A) RNAs by adding excess competitor polyadenylic acids to extracts before mRNA isolation. A silver-stained gel showed selective enrichment of proteins in UV-cross-linked poly(A) mRNA eluates compared to the input extract, but no proteins or RNA were detectable in our control samples (Fig. 1a). Likewise, immunoblot analysis showed that Scp160p, a well-known yeast RBP¹¹, was present in the eluates; however, actin (Act1), a protein not thought to bind RNA, was not detectable (Fig. 1b). Neither protein was evident in any of the equally sized negative-control eluates. Together, these results showed that we were able to selectively capture mRBPs *in vivo*.

Department of Microbial and Cellular Sciences, School of Biosciences and Medicine, Faculty of Health and Medical Sciences, University of Surrey, Guildford, UK. Correspondence should be addressed to A.P.G. (a.gerber@surrey.ac.uk).

Received 6 July; accepted 23 October; published online 23 November 2015; doi:10.1038/nsmb.3128

Figure 1 Identification of mRBPs in *S. cerevisiae*. **(a)** Silver-stained polyacrylamide gel. Lanes 2–4, input extracts and poly(A) and RNase-treated control samples; lanes 6–8, eluates from poly(A) mRNA isolation. Lane 1, marker (M) with molecular weights (MW) indicated. **(b)** Immunoblot analysis with Scp160p- and Act1p-specific antibodies. Original images are shown in **Supplementary Data Set 1**. **(c)** Heat map representation of the abundance of 765 proteins composing the yeast mRBPome. Columns refer to three independent experiments and respective controls; rows represent individual proteins. For visualization purposes, the white-red color bar represents $\log_{10}$ -transformed raw (non-normalized) MS peak areas of respective proteins. The Venn diagram represents the overlap of number of identified proteins (n) across the three experiments. The P value (hypergeometric test) relates to the significance of overlap (Online Methods). **(d)** Selective samples of significantly shared GO terms from SGD ($P < 0.01$, FDR $< 5\%$) among proteins of the yeast mRBPome. Bars indicate the fraction of proteins annotated with the indicated GO term in the yeast mRBPome (765 proteins; black bars) and all GO-annotated proteins in the SGD (6,607 proteins; gray bars). **(e)** Selection of domains enriched in the mRBPome. Bars indicate the fraction of annotated proteins bearing at least one of the indicated domains (InterPro) in the yeast mRBPome (765 proteins; black bars) and the reference proteome (6,621 proteins in UniProt; gray bars). The number of proteins within each fraction is shown to the right. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$ at FDR $< 5\%$ (hypergeometric test).

To comprehensively identify the proteins bound to poly(A) mRNAs, defining the yeast mRBPome, we subjected samples from three independent experiments to MS analysis. At the same time, we analyzed respective control samples to demarcate nonspecific binders. We identified 765 proteins in at least two out of three replicate experiments with a false discovery rate (FDR) of less than 1% and represented by at least two different peptides (**Fig. 1c**; details in Online Methods; raw data in **Supplementary Table 1**).

Our yeast mRBPome contained 86 (72%) of the 120 RBPs that were previously identified in starved yeast cells⁸ ($P < 8 \times 10^{-55}$, hypergeometric test), and it significantly overlapped with the repertoire of yeast RBPs identified in screens for RBPs using protein microarrays or poly(A) purification from non-cross-linked cell extracts^{4,5} (overlap of data in **Supplementary Fig. 2a**). In agreement with previous observations¹² based on a list of 561 known and predicted proteins with RNA-related functions¹³ (205 of which were present in our mRBPome, $P < 2 \times 10^{-60}$, hypergeometric test), the mRNAs encoding proteins in our mRBPome, in comparison to proteins not present in our mRBPome, were significantly less stable but were highly expressed, with a tendency toward higher mRNA copy numbers and ribosome occupancies (**Supplementary Fig. 3a**). Moreover, the proteins composing the mRBPome tended to be more stable and abundant. In contrast, protein-expression noise, which is a measure for the extent of the variation of protein expression from cell to cell, was significantly lower for the mRBPome compared to all other proteins, thus suggesting that mRBPs are uniformly expressed across a homogenous population of cells¹². Finally, our experimentally determined mRBPome

consisted of proteins that were localized in all cellular compartments (**Supplementary Fig. 3b**). The observed slight overrepresentation of cytoplasmic proteins is consistent with our approach, in which we specifically captured mRNAs that are mainly associated with translation in the cytoplasm in rapidly grown yeast cells¹⁴.

We further categorized the 765 proteins according to Gene Ontology (GO) annotation retrieved from the *Saccharomyces* Genome Database¹⁵ (SGD) (**Fig. 1d**; complete analysis of significantly enriched GO terms from the SGD at FDR $< 5\%$ in **Supplementary Table 2a–c**). We identified approximately half (47%) of the previously annotated RBPs (175 out of 370 proteins annotated as ‘RNA binding’, $P < 3 \times 10^{-70}$, hypergeometric test from SGD), most of the proteins (59%) annotated as ‘mRNA-poly(A) RNA binding’ (102 out of 172 annotated proteins, $P < 8 \times 10^{-55}$, hypergeometric test from SGD) and all six proteins with known poly(A)-binding function (Nab2, Npl3, Mtr4, Pab1, Sgn1 and Tpa1). Nevertheless, most of the proteins in the mRBPome (73%) have not been previously linked to RNA-binding functions and thus are likely to represent new RBPs or proteins with dual functions; this group included 325 enzymes, i.e., proteins with an annotated enzyme classification number in UniProt (42% of the mRBPome, $P < 0.001$ at FDR $< 5\%$, hypergeometric test; Online Methods). Most of these were enzymes functioning in metabolic processes, e.g., hydrolases, oxidoreductases and ligases.

We analyzed the yeast mRBPome for the occurrence of protein domains as annotated in the UniProt database¹⁶. The mRBPome was significantly enriched in prominent classical RBDs such as the hnRNPK homology domain (KH dom), the RRM (RRM dom) and

Figure 2 Identification of mRBPs in *C. elegans*. **(a)** Silver-stained polyacrylamide gel of UV-cross-linked samples from mixed-stage nematodes. **(b)** Immunoblot analysis with GLD-1- and CYC-1-specific antibodies. Original images of blots are shown in **Supplementary Data Set 1**. **(c)** Cluster heat map representing the abundance of 594 proteins composing the mRBPome in mixed-stage, L4-stage and L4+ENU stage animals treated with 5 mM ENU and in the corresponding negative-control samples. The white-blue scale shows raw (non-normalized) peak areas (log₁₀ transformed), and the number of proteins within groups is indicated to the left. **(d)** Relative changes of 96 proteins in the mRBPomes of synchronized L4-stage animals upon apoptosis (FDR < 5%). Rows indicate three pairwise comparisons within matched samples. In **a–c**, columns refer to proteins. Fold changes are indicated with the blue-yellow color bar. **(e)** Significantly shared GO terms among proteins of the *C. elegans* mRBPome (FWER-adjusted $P < 0.01$). Bar diagrams indicate the fraction of proteins in the indicated GO term among the *C. elegans* mRBPome (594 proteins; black bars) and the 22,817 GO-annotated proteins in the UniProt reference proteome (gray bars). **(f)** Domains enriched in the mRBPome. Bars relate to the fraction of proteins bearing at least one of the domains (InterPro) in the *C. elegans* mRBPome (594 proteins; black) and 26,165 proteins composing the UniProt reference proteome (gray). Numbers of proteins are shown to the right. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ at FDR < 5% (hypergeometric test). EEF-1B.1/2 refers to EEF-1B.1 and EEF-1B.2.

the Sm domain (Ribonucl LSM) (Fig. 1e). Less frequent domains, such as the S1 domain (four of the seven annotated proteins) and the La domain (all three annotated proteins), were also enriched in the mRBPome but did not reach statistical significance by the chosen statistical test (complete domain analysis in **Supplementary Table 2d**). Interestingly, this analysis also revealed significant enrichment of additional domains that have not been previously linked to RNA binding or that occur in proteins devoid of classical RBDs. This includes the P-loop NTPase domain, a common motif in ATP- and GTP-binding proteins; the heat-shock protein 70 domain and the chaperonin Cpn60-TCP-1 domain (Fig. 1e). Finally, we searched the yeast mRBPome for the overrepresentation of short amino acid sequence stretches: 707 tetramer and 64 pentamer amino acid sequences were significantly enriched in the yeast mRBPome compared to the reference proteome ($P < 0.05$ at FDR < 5%, hypergeometric test), excluding statistical bias from repeated motifs in the same protein (complete motif analysis in **Supplementary Table 2e,f**). Most of these motifs contain arginine, lysine, tyrosine and glycine residues (47 of 64 pentamers, i.e., GGRGG, RGGRG and GTGKT), which are also present in known RBPs and could relate to disordered regions that interact with RNA⁶.

Identification of *C. elegans* mRNA-interacting proteins

We next sought to decipher the mRBPome in a living animal, the nematode *C. elegans*. We thus applied UV-cross-linking and poly(A) purification to capture the *in vivo* mRBPome of mixed-stage animals representing all stages of the *C. elegans* life cycle and in synchronized animals at the fourth (final) larval stage (L4). At the L4 stage, important cellular processes such as spermatogenesis and apoptosis (programmed cell death) take place and are substantially controlled by post-transcriptional events¹⁷. Hence, we also investigated the mRBPome upon induction of germline apoptosis in synchronized L4-stage larvae with 5 mM *N*-nitroso-*N*-ethylurea (ENU), a potent alkylating mutagen¹⁸. The successful induction of apoptosis was monitored by the expression of *egl-1* mRNA (**Supplementary Fig. 4a**), a marker of germline apoptosis¹⁹.

In the same manner as for the yeast, we exposed the animals to UV under conditions that preserved the integrity of the RNA (**Supplementary Fig. 1b**) and then isolated poly(A) mRNAs (Online Methods). We found that the composition of proteins in the eluates of poly(A)-selected mRNAs was dissimilar to that of cell lysates, thus indicating the selective enrichment of mRBPs (Fig. 2a). To further confirm this, we used GLD-1, a KH-domain protein that

Figure 3 Conservation of the mRBPome across species. **(a)** Conservation between *S. cerevisiae* and *C. elegans* proteins²⁷. Yellow columns indicate the fraction of conserved *S. cerevisiae* proteins in *C. elegans* relative to the reference proteome (6,621 proteins) and mRBPome (765 proteins). Blue columns indicate the fraction of conserved *C. elegans* proteins in *S. cerevisiae* relative to the worm reference proteome (26,165 proteins) and mRBPome (594 proteins). Stars indicate significant differences between fractions ($P < 0.001$, hypergeometric test). **(b)** Venn diagram showing overlap of orthologous proteins of the mRBPome in *S. cerevisiae* (476 proteins) and *C. elegans* (330 proteins). The P value indicates the significance of overlap (hypergeometric test). **(c)** Schematic view of the glycolytic pathway. P, phosphate. Proteins highlighted in yellow are present exclusively in the yeast mRBPome; proteins in blue are found exclusively in the *C. elegans* mRBPome, and proteins shared in both mRBPomes are in green. DHAP, dihydroxyacetone phosphate; PEP, phosphoenolpyruvate; GA3P, glyceraldehyde 3-phosphate. **(d)** Schematic view of the proteasome composing the core particle (20S CP) that can bind to one or two regulatory particles (19S RP).

is required for key developmental processes in the *C. elegans* germ line²⁰, as a positive control and cytochrome C (CYC-1), a highly expressed mitochondrial protein not thought to bind RNA, as a negative control. As expected, GLD-1 was enriched in our mRNA isolates but not in the control samples, and CYC-1 was not present in eluates (Fig. 2b).

We performed MS analysis in three independent experiments (comprising ~120,000 animals each) of mixed-staged, L4 and matched L4 ENU-treated (L4+ENU) animals and analyzed respective control samples to demarcate nonspecific binding partners (Online Methods). Data were highly correlated between replicate experiments of mixed-stage samples (Pearson correlation, $r = 0.89$ – 0.97) and L4 and L4+ENU matched samples ($r = 0.95$ – 0.98), but the correlation was somewhat less preserved between replicate experiments of the L4 and L4+ENU animals ($r = 0.75$ – 0.92), possibly because of slight differences in the populations of L4-stage larvae (scatter plot comparing samples in Supplementary Fig. 4b). Applying the refinements that we used for yeast, we identified 594 proteins in at least two out of the nine samples with a FDR $< 1\%$ and at least two different peptides (Online Methods; raw data in Supplementary Table 3a). We considered this entire set of 594 proteins to be part of the *C. elegans* mRBPome, a significant proportion (136 proteins, 23%, $P < 1 \times 10^{-64}$, hypergeometric test) of which overlapped with a recent computationally predicted *C. elegans* RBPome²¹ (comparison in Supplementary Fig. 2b).

Under our stringent selection criteria, 93.5% of the experimentally captured *C. elegans* mRBPome was identifiable in mixed-stage animals (555 proteins), thus reflecting a comprehensive repertoire of mRBPs expressed during the entire nematode life cycle (Fig. 2c and Venn diagram showing occurrence of proteins across samples in Supplementary Fig. 4c). The remaining 39 proteins (6.5% of the mRBPome) were exclusive to L4-stage animals. These proteins included known RBPs with roles in germline development, such as

FBF-2, a PUF family RNA-binding protein that negatively regulates *fem-3* and *gld-1* expression; ALG-4, a member of the argonaute proteins involved in male germline fertility; and CPB-2, a cytoplasmic polyadenylation-dependent binding-protein homolog specifically expressed in the germ line²². Moreover, GST-5 has known roles in germline apoptosis²³ and has been recorded exclusively in apoptosis-induced animals. To further search for differences between L4 and L4+ENU animals, we performed relative quantification of MS data (Online Methods). 96 proteins were differentially enriched in these mRBPomes at an FDR $< 5\%$ (Fig. 2d; complete data set in Supplementary Table 3b). Of these, 41 proteins were more predominant in L4 and 55 proteins were more predominant in apoptosis-induced L4-stage animals. The former group was enriched in proteins involved in ‘cellular response to stress’ (Westfall and Young familywise error rate (FWER)-adjusted $P < 0.01$; Online Methods) and ‘eukaryotic translation elongation’ (FWER-adjusted $P < 0.05$), including EEF-1B.1 and EEF-1B.2. Conversely, the latter group contained proteins linked to apoptosis, germline development or stress response, such as EIF-3.K, a translation-initiation factor that positively regulates *ced-3* and thereby promotes programmed cell death²⁴; EIF-6, which is required for control of cell division in the germ line²⁵; and CTL-1 and PRDX-3, both of which are involved in the oxidative stress response²². Remarkably, six translation-initiation factors were prevalent in ENU-treated animals: EGL-45, EIF-3.I, EIF-3.K, EIF-6, IFE-2 and CLU-1 (FWER-adjusted $P < 0.01$), thus indicating the potential for selective remodeling of translation during apoptotic stress.

GO enrichment analysis revealed that, as expected, a significant fraction of the proteins composing the *C. elegans* mRBPome had roles in RNA metabolism, i.e., ‘mRNA metabolic process’ ($P < 0.001$; all P values in this paragraph are FWER adjusted unless otherwise noted), ‘RNA binding’ ($P < 0.001$), ‘ribonucleoprotein complex biogenesis’ ($P < 0.001$), and ‘translation’ ($P < 0.001$) (Fig. 2e; full list

Figure 4 Validation of mRNA-protein interactions by quantitative dual fluorescence-based mRNA detection assay. Normalized Alexa 594 fluorescence signal was used to monitor the presence of poly(A) mRNAs (left graph, black bars). As a control, the samples were treated with RNase ONE (gray bars). Data were normalized to untagged wild-type cells (ctrl). Error bars, s.e.m. ($n = 3$ independent IPs). * $P \leq 0.05$; ** $P \leq 0.01$ by paired two-tailed Student's t test. The right graph shows the normalized GFP fluorescence signal of immunopurified proteins. Data underlying the graphical representation are shown in the related Source Data.

of enriched GO terms in **Supplementary Table 4a–c**). Furthermore, many mRBPs were linked to the control of development, i.e., ‘developmental process’ ($P < 0.001$), ‘embryo development’ ($P < 0.05$) and ‘regulation of cell death’ ($P < 0.001$), results reminiscent of the increasingly recognized role of post-transcriptional control in these processes. As seen in yeast, many proteins of the *C. elegans* mRBPome were enzyme classification–annotated enzymes (227 proteins, 38% of the mRBPome, $P < 0.001$ at 5% FDR, hypergeometric test). They preferentially act in metabolic processes, for example, ‘primary metabolic process’, ‘carbohydrate derivate binding’, ‘glycolytic process’ and ‘tricarboxylic acid or Krebs cycle’ ($P < 0.001$), thus further supporting the idea that metabolic enzymes could have dual functions.

Protein domain analysis revealed significant overrepresentation of classical RNA-related protein domains, such as the DEAD-DEAH-box helicase or ZnF domains (for example, Znf CCHC) (**Fig. 2f**; full domain analysis in **Supplementary Table 4d**). However, certain RBDs were less well represented, such as KH and RRM domains. The reason for the absence of many proteins bearing these domains is not known, but the respective domains could target other types of RNA and/or have other functions, for example, acting as sites of protein-protein interaction²⁶. A search for short amino acid motifs with significant enrichment in the mRBPome identified 184 tetramer and 228 pentamer motifs ($P < 0.05$ at FDR < 5%, hypergeometric test) most of them bearing glycine, arginine, lysine or tyrosine residues (extended motif analysis in **Supplementary Table 4e,f**). Of note, 33 tetramer motifs and 7 pentamer motifs (RGGRG, GGRGG, DEAVA, TITND, GTGKT, LGGGT and QATKD) were also significantly enriched in the yeast mRBPome (**Supplementary Table 5a,b**).

Evolutionary conservation of the mRBPome

We next assessed whether our experimentally determined mRBPomes were evolutionarily conserved. Therefore, we retrieved protein homology information for *S. cerevisiae* and *C. elegans* from InParanoid²⁷, which revealed 1,841 orthologous protein pairs. We found that 7% of all *C. elegans* proteins in the reference proteome comprising 26,165 proteins matched a homologous yeast protein, and 28% of all *S. cerevisiae* proteins in the 6,621-protein reference proteome had an orthologous protein in *C. elegans* (**Fig. 3a**). Strikingly, more than half of the *C. elegans* mRBPome matched a *S. cerevisiae* ortholog (330 proteins, 56%), and almost two-thirds of the yeast mRBPome had a homolog in *C. elegans* (476 proteins, 62.3%). Furthermore, we found a highly significant overlap of orthologous pairs of proteins in both mRBPomes, thus suggesting high conservation of mRBPomes (**Fig. 3b**). For instance, the glycolytic pathway, which is significantly enriched in yeast ($P < 7 \times 10^{-27}$) and *C. elegans* ($P < 2 \times 10^{-7}$, P values from Wikipathway analysis; Online Methods) mRBPomes was contained in the intersection representing conserved orthologous mRBPs (**Fig. 3c** and complete analysis of orthologous proteins in mRBPomes in **Supplementary Table 5c–f**). Moreover, distinct proteins of large macromolecular complexes, such as the proteasome, are conserved and significantly enriched in the yeast ($P < 7 \times 10^{-11}$) and in the *C. elegans* ($P < 2 \times 10^{-19}$) mRBPomes (Wikipathway analysis). Interestingly, we found that the components of the 26S proteasome interact with mRNAs aligned along the surface and cover the cavity through which ubiquitinated proteins are degraded (**Fig. 3d**).

Validation of novel mRBPs

To validate novel and conserved mRBPs, we adapted a fluorescence-based mRNA-protein interaction assay to yeast²⁸ (Online Methods). We UV-cross-linked proteins to mRNAs *in vivo* and performed immunoprecipitation (IP) of endogenously GFP-tagged proteins from cell lysates under stringent conditions. We then assessed the abundance of proteins and bound mRNAs via GFP and hybridization with fluorescently labeled oligo(dT)₂₅, respectively (**Fig. 4**). As expected, we detected poly(A) mRNAs in the IP of three known mRBPs (Gis2, Khd1 and Pab1) but observed no RNA-dependent associations with untagged wild-type control cells or Ras1p, a GTPase not known to bind RNA. We further confirmed that mRNAs bound to Sbp1p and

Figure 5 Glycolytic enzymes selectively interact with glycolytic mRNAs. Agarose gel showing products from reverse-transcription (RT)-PCR reactions with gene-specific primers to detect glycolytic mRNAs (right) bound to immunopurified GFP or TAP-tagged proteins of indicated glycolytic proteins (top). Input, total RNA from cross-linked cells; ctrl, untagged control cells. Full-sized images of gels are shown in **Supplementary Data Set 1**.

YGR250p, two predicted mRBPs bearing classical RBDs; the glycolytic proteins Eno1 and Pfk2; the pentose phosphate-pathway enzymes Tal1p and Tkl1p; and the conserved proteasome components Pre10p and Rpt1p. Of note, we found no statistically significant correlation between the abundance of proteins (GFP) and the fluorescence signals inferred from the bound mRNAs (untreated, $r^2 = 0.11$ and $P = 0.29$; RNase treated, $r^2 = 0.05$ and $P = 0.48$, with r^2 from fitting of the linear model), showing that protein abundance is not a predictor of the amount of mRNAs bound. Intrigued by our finding that all yeast glycolytic enzymes were part of the mRBPome, and wondering what their specific mRNA targets could be, we speculated that they might bind to glycolytic mRNAs, thus providing potential for coordination of post-transcriptional pathways⁴. Indeed, we detected mRNAs encoding several glycolytic proteins in the IPs of GFP and/or tandem affinity purification (TAP)-tagged glycolytic enzymes (Pfk2, Gpm1 and Eno1), whereas unrelated mRNAs (Act1 and Ash1) were absent (Fig. 5). We detected no mRNAs in control isolates from untagged strains (mock). Furthermore, association of Pfk2p and Eno1p with glycolytic mRNAs was recapitulated in affinity-purified isolates obtained from non-cross-linked cells; the observed interactions were not ribosome dependent because of potential cotranslational assembly²⁹, as evidenced by the treatment of extracts with puromycin, a compound that disassembles ribosomes (Supplementary Fig. 5). In conclusion, these results strongly suggest that some glycolytic enzymes bind to their own mRNAs as well as to other select mRNAs for glycolytic proteins (although no physical or genetic interactions have been reported between Eno1, Gpm1 and Pfk2 (ref. 15)).

DISCUSSION

Our experimentally identified mRBPomes notably enhance and broaden the repertoire of known mRBPs in nature, particularly because we describe the first experimentally captured mRBPomes of living whole organisms and the deployment of different mRBPs in apoptosis. Although a significant proportion of our identified mRBPs were in agreement with previously reported data, we identified new mRBPs, functional characteristics (for example, enzymes), RNA binding-related protein domains and amino acid motifs, providing annotation for hundreds of new mRBPs (Figs. 1 and 2).

We also observed the differential arrangement of mRBPs upon induction of apoptosis in *C. elegans* L4-stage larvae, thus suggesting stress specificity within the mRBPomes of animals: 16% (96 proteins) of the 583 detectable mRBPs in L4-stage animals exhibited significantly altered abundance within the mRBPome upon induction of apoptotic stress (Fig. 2d). Interestingly, apoptotic stress decreased the association of translation-elongation factors within the mRBPome, whereas several translation-initiation factors became enriched, thus suggesting active remodeling of translation. Whether this relates to translation regulation of a subset of mRNAs by one or several initiation factors³⁰ or to changes of translation-elongation rates remains to be investigated. Monitoring of the dynamic responses of the mRBPome may thus establish a valuable new approach to identify potential post-transcriptional regulators in apoptosis or other physiological conditions, and in disease.

In addition to many of the annotated canonical RNA-binding domains, we found high enrichment of other motifs in both mRBPomes, some of which might act as RBDs. For example, the significant enrichment of the heat-shock protein 70 domain and the chaperonin Cnp60–TCP-1 domain in both mRBPomes is reminiscent of previous observations that certain heat-shock proteins act as RNA-binding entities *in vivo* and guide the folding of RNA substrates for subsequent degradation and translation^{31,32}. Domain analysis also revealed

limitations to our chosen UV-cross-linking approach. For instance, proteins bearing double-stranded RBDs (dsRBDs) were underrepresented in both mRBPomes: of the six annotated proteins containing dsRBDs in yeast, only Rps2p was identified; only 3 of the 18 annotated dsRBD-containing proteins were identified in *C. elegans* (RPS-2, MRPS-5 and RHA-1), all of them containing additional RBDs. These results may recapitulate the inefficiency of UV-cross-linking of proteins interacting with extended dsRNA in α -helical structures, and other methods may be used to capture dsRBPs with high efficiency³³. UV light cross-links proteins to RNA bases, preferentially uridine, if the protein and RNA are in proximity for a certain time. This could explain the capturing of RNA catalytic enzymes, such as *C. elegans* GLD-2, which is an enzyme that polyadenylates *gld-1* mRNA but has not been reported to directly bind to this mRNA, because it needs GLD-3 as a bridge³⁴. It is thus possible that we also identified proteins that, in current terms, ‘indirectly’ interact with mRNA, owing to the proximity of reactive centers to RNA.

We found significant conservation of experimentally determined mRBPs, thus suggesting elementary functions for the underlying protein-mRNA interactions in cell homeostasis (Fig. 3). Bioinformatics analysis of all 1,841 orthologous *S. cerevisiae* and *C. elegans* proteins showed that proteins associated with the UniProt keywords ‘RNA binding’, ‘mRNA processing’ or ‘proteasome’ were significantly overrepresented ($P < 0.001$ at FDR < 5%, hypergeometric test) whereas ‘DNA binding’ proteins were not particularly enriched (complete analysis of orthologs in Supplementary Table 5g,h). Our experimental data are thus in agreement with these *in silico* predictions, which probably reflect early establishment and sustainability of post-transcriptional regulators during evolution. In particular, we believe that conserved mRNA-binding properties of proteins composing ancient metabolic pathways, such as glycolysis, or protein complexes, such as the proteasome, could have emerged during the proposed transition from the RNA world to protein world, exerting activity that may have been lost in present cells.

The presence of proteasome subunits within both mRBPomes (Fig. 3d) is reminiscent of early findings that 19S prosomes, which were renamed ‘proteasomes’ after their role in proteolysis was discovered³⁵, may constitute an RNP complex involved in the negative control of mRNA translation³⁶. Moreover, it has been postulated that the 20S and 26S proteasomes exhibit RNA endoRNase activity³⁷ and that α subunits ($\alpha 1$, $\alpha 5$, $\alpha 6$ and $\alpha 7$), which are present in both of our mRBPomes, exhibit endoRNase activity³⁸. Inspired by the conceptual analogy between the proteasome and the RNA exosome³⁹, we thus speculate that the proteasome may still play a part in mRNA degradation, possibly via the catalytic activity of the 20S.

The significant proportion of metabolic enzymes within both mRBPomes indicates the possibility of cross-talk between metabolism and RNA biology. Several human enzymes are known to selectively bind mRNAs; these enzymes include the glycolytic proteins glyceraldehyde-3-phosphate dehydrogenase (GAPDH; termed Tdh1, Tdh2 and Tdh3 in yeast), enolase (ENOL-1 and Eno1,2), aldolase (ALDO and Fba1) and phosphoglycerate kinase (PGK)^{40,41}. However, we were particularly intrigued by our finding that all 16 enzymes that mediate the 10 steps of glycolysis bound mRNAs in yeast, as did some of the orthologous proteins in *C. elegans* (Fig. 3c). Moreover, because we found that glycolytic proteins can bind both their own mRNAs and other glycolytic mRNAs, we postulate that glycolysis may be additionally coordinated and/or controlled by specific enzyme-mRNA interactions within the pathway. In one scenario, mRNA-protein interactions could activate or modify enzymatic activities; however, this may appear less likely in light of the tremendous excess of glycolytic enzymes (~100,000 copies) to mRNAs (~10 copies) in cells.

In another scenario, metabolic enzymes may directly control glycolytic mRNAs in the cytoplasm, possibly affecting mRNA localization, translation or decay. If so, we hypothesize that this function could relate to an ancient mechanism for the coordination of pathway activity to achieve metabolic control—perhaps establishing an early RNA regulon, a structure defined by mRBPs, that coordinates the expression of mRNAs encoding functionally related proteins⁴².

METHODS

Methods and any associated references are available in the [online version of the paper](#).

Note: Any Supplementary Information and Source Data files are available in the [online version of the paper](#).

Accession codes. Proteomics data have been deposited in the ProteomeXchange Consortium database under accession codes [PXD002293](#) and [PXD002226](#), corresponding to the *S. cerevisiae* and *C. elegans* mRBPomes, respectively.

ACKNOWLEDGMENTS

We are grateful to K. Heesom (Proteomics Facility, University of Bristol) for performing the MS analysis; M. Hengartner and D. Subasic (Institute of Molecular Life Sciences, University of Zurich) for support and *C. elegans* strains; S. Leidel (Max-Planck-Institute for Molecular Biomedicine, Münster) for GFP and TAP-tagged *S. cerevisiae* strains; M. Seedorf (Center for Molecular Biology, University of Heidelberg) and R. Ciosk (Friedrich Miescher Institute for Biomedical Research) for anti-Scp160 and anti-GLD-1 antibodies, respectively; D. Pérez-Mendoza and D. Subasic for reading of the manuscript; and members of the Gerber laboratory and the Sinergia project for discussions. This study was funded by a 'Sinergia' grant (CSRII3-141942 (A.P.G.)) from the Swiss National Science Foundation and (in part) by the Biotechnology and Biological Sciences Research Council (BB/K009303/1 (A.P.G.)).

AUTHOR CONTRIBUTIONS

A.M.M.-G. and A.P.G. conceived and designed the experiments. A.M.M.-G. performed laboratory experiments. E.E.L. performed bioinformatics analyses. All authors analyzed data, discussed the results, and wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Glisovic, T., Bachorik, J.L., Yong, J. & Dreyfuss, G. RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett.* **582**, 1977–1986 (2008).
- Lukong, K.E., Chang, K.W., Khandjian, E.W. & Richard, S. RNA-binding proteins in human genetic disease. *Trends Genet.* **24**, 416–425 (2008).
- Gerstberger, S., Hafner, M. & Tuschl, T. A census of human RNA-binding proteins. *Nat. Rev. Genet.* **15**, 829–845 (2014).
- Scherrer, T., Mittal, N., Janga, S.C. & Gerber, A.P. A screen for RNA-binding proteins in yeast indicates dual functions for many enzymes. *PLoS ONE* **5**, e15499 (2010).
- Tsvetanova, N.G., Klass, D.M., Salzman, J. & Brown, P.O. Proteome-wide search reveals unexpected RNA-binding proteins in *Saccharomyces cerevisiae*. *PLoS ONE* **5**, e12671 (2010).
- Castello, A. *et al.* Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* **149**, 1393–1406 (2012).
- Baltz, A.G. *et al.* The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol. Cell* **46**, 674–690 (2012).
- Mitchell, S.F., Jain, S., She, M. & Parker, R. Global analysis of yeast mRNPs. *Nat. Struct. Mol. Biol.* **20**, 127–133 (2013).
- Kwon, S.C. *et al.* The RNA-binding protein repertoire of embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1122–1130 (2013).
- Zhang, L., Zhang, K., Prandl, R. & Schoffl, F. Detecting DNA-binding of proteins *in vivo* by UV-crosslinking and immunoprecipitation. *Biochem. Biophys. Res. Commun.* **322**, 705–711 (2004).

- Frey, S., Pool, M. & Seedorf, M. Scp160p, an RNA-binding, polysome-associated protein, localizes to the endoplasmic reticulum of *Saccharomyces cerevisiae* in a microtubule-dependent manner. *J. Biol. Chem.* **276**, 15905–15912 (2001).
- Mittal, N., Roy, N., Babu, M.M. & Janga, S.C. Dissecting the expression dynamics of RNA-binding proteins in posttranscriptional regulatory networks. *Proc. Natl. Acad. Sci. USA* **106**, 20300–20305 (2009).
- Hogan, D.J., Riordan, D.P., Gerber, A.P., Herschlag, D. & Brown, P.O. Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. *PLoS Biol.* **6**, e255 (2008).
- Arava, Y. *et al.* Genome-wide analysis of mRNA translation profiles in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **100**, 3889–3894 (2003).
- Costanzo, M.C. *et al.* *Saccharomyces* genome database provides new regulation data. *Nucleic Acids Res.* **42**, D717–D725 (2014).
- UniProt Consortium. Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* **42**, D191–D198 (2014).
- Thomas, M.P. & Lieberman, J. Live or let die: posttranscriptional gene regulation in cell stress and cell death. *Immunol. Rev.* **253**, 237–252 (2013).
- Gartner, A., Milstein, S., Ahmed, S., Hodgkin, J. & Hengartner, M.O. A conserved checkpoint pathway mediates DNA damage-induced apoptosis and cell cycle arrest in *C. elegans*. *Mol. Cell* **5**, 435–443 (2000).
- Conradt, B. & Horvitz, H.R. The *C. elegans* protein EGL-1 is required for programmed cell death and interacts with the Bcl-2-like protein CED-9. *Cell* **93**, 519–529 (1998).
- Francis, R., Maine, E. & Schedl, T. Analysis of the multiple roles of *gld-1* in germline development: interactions with the sex determination cascade and the *gld-1* signaling pathway. *Genetics* **139**, 607–630 (1995).
- Tamburino, A.M., Ryder, S.P. & Walhout, A.J. A compendium of *Caenorhabditis elegans* RNA binding proteins predicts extensive regulation at multiple levels. *G3 (Bethesda)* **3**, 297–304 (2013).
- Harris, T.W. *et al.* WormBase 2014: new views of curated biology. *Nucleic Acids Res.* **42**, D789–D793 (2014).
- Lettre, G. *et al.* Genome-wide RNAi identifies p53-dependent and -independent regulators of germ cell apoptosis in *C. elegans*. *Cell Death Differ.* **11**, 1198–1203 (2004).
- Huang, C.Y. *et al.* *C. elegans* EIF-3.K promotes programmed cell death through CED-3 caspase. *PLoS ONE* **7**, e36584 (2012).
- Voutev, R., Killian, D.J., Ahn, J.H. & Hubbard, E.J. Alterations in ribosome biogenesis cause specific defects in *C. elegans* hermaphrodite gonadogenesis. *Dev. Biol.* **298**, 45–58 (2006).
- Cléry, A., Blatter, M. & Allain, F.H. RNA recognition motifs: boring? Not quite. *Curr. Opin. Struct. Biol.* **18**, 290–298 (2008).
- Ostlund, G. *et al.* InParanoid 7: new algorithms and tools for eukaryotic orthology analysis. *Nucleic Acids Res.* **38**, D196–D203 (2010).
- Strein, C., Alleaume, A.M., Rothbauer, U., Hentze, M.W. & Castello, A. A versatile assay for RNA-binding proteins in living cells. *RNA* **20**, 721–731 (2014).
- Halbach, A. *et al.* Cotranslational assembly of the yeast SET1C histone methyltransferase complex. *EMBO J.* **28**, 2959–2970 (2009).
- Lee, A.S., Kranzusch, P.J. & Cate, J.H. eIF3 targets cell-proliferation messenger RNAs for translational activation or repression. *Nature* **522**, 111–114 (2015).
- Henics, T. *et al.* Mammalian Hsp70 and Hsp110 proteins bind to RNA motifs involved in mRNA stability. *J. Biol. Chem.* **274**, 17318–17324 (1999).
- Zimmer, C., von Gabain, A. & Henics, T. Analysis of sequence-specific binding of RNA to Hsp70 and its various homologs indicates the involvement of N- and C-terminal interactions. *RNA* **7**, 1628–1637 (2001).
- Liu, Z.R., Wilkie, A.M., Clemens, M.J. & Smith, C.W. Detection of double-stranded RNA-protein interactions by methylene blue-mediated photo-crosslinking. *RNA* **2**, 611–621 (1996).
- Suh, N., Jedamzik, B., Eckmann, C.R., Wickens, M. & Kimble, J. The GLD-2 poly(A) polymerase activates *gld-1* mRNA in the *Caenorhabditis elegans* germ line. *Proc. Natl. Acad. Sci. USA* **103**, 15108–15112 (2006).
- Baumeister, W., Walz, J., Zuhl, F. & Seemuller, E. The proteasome: paradigm of a self-compartmentalizing protease. *Cell* **92**, 367–380 (1998).
- Schmid, H.P. *et al.* The prosome: an ubiquitous morphologically distinct RNP particle associated with repressed mRNPs and containing specific ScRNA and a characteristic set of proteins. *EMBO J.* **3**, 29–34 (1984).
- Kulichkova, V.A. *et al.* 26S proteasome exhibits endoribonuclease activity controlled by extra-cellular stimuli. *Cell Cycle* **9**, 840–849 (2010).
- Mittenberg, A. *et al.* Mass-spectrometric analysis of proteasome subunits exhibiting endoribonuclease activity. *Cell Tissue Biol.* **8**, 423–440 (2014).
- Makino, D.L., Halbach, F. & Conti, E. The RNA exosome and proteasome: common principles of degradation control. *Nat. Rev. Mol. Cell Biol.* **14**, 654–660 (2013).
- Ciesla, J. Metabolic enzymes that bind RNA: yet another level of cellular regulatory network? *Acta Biochim. Pol.* **53**, 11–32 (2006).
- Hentze, M.W. & Preiss, T. The REM phase of gene regulation. *Trends Biochem. Sci.* **35**, 423–426 (2010).
- Keene, J.D. RNA regulons: coordination of post-transcriptional events. *Nat. Rev. Genet.* **8**, 533–543 (2007).

ONLINE METHODS

In vivo capture of mRBPs in *S. cerevisiae*. Strain BY4741 (MATa *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*) was grown in 500 ml YPD medium (1% yeast extract, 2% peptone, and 2% D-glucose) at 30 °C with constant shaking at 220 r.p.m. Cells were collected at mid-log phase (OD₆₀₀ of ~0.6) by centrifugation and washed three times with phosphate-buffered saline (PBS). For UV-cross-linking, the cells were resuspended in 25 ml of PBS and exposed to 1,200 mJ/cm² of 254-nm UV light in a Stratallinker 1800 (Stratagene) with two 2-min breaks on ice and gentle mixing. Cells were resuspended in 4 ml lysis buffer (100 mM Tris-HCl, pH 7.5, 500 mM LiCl, 10 mM EDTA, 1% Triton X-100, 5 mM DTT, 20 U ml⁻¹ DNase I (Promega, M6101), 100 U ml⁻¹ RNasin (Promega, N2611), and complete EDTA-free protease-inhibitor cocktail (Roche, 11836170001)) and mechanically broken with glass beads in a Tissue Lyser (RETSCH MM200, Qiagen) for 10 min at 30 Hz at 4 °C. The cooled lysate was cleared by three sequential centrifugations at 4 °C at 3,000g for 3 min, and 5,000g and 10,000g for 5 min each. Two negative controls were introduced at this stage: the extract was supplemented with 100 U RNase ONE (Promega, M4265) for 2 h at 37 °C to digest RNA, or 20 mg of poly(A) (Sigma, P9403) was added to the extract for competition experiments. To control for the integrity of RNA, total RNA was isolated from 50 µl of extracts with the ZR RNA MiniPrep kit (Zymo Research, R1065). RNA was quantified with a NanoDrop ND-2000 device and visualized on a 1% agarose gel stained with RedSafe (iNTRON Biotechnology, 21141). The presence of mRNAs in poly(A) eluates was controlled by reverse-transcription (RT)-PCR (data not shown). To capture polyadenylated RNAs for mRBPome analysis, 1 mg of oligo(dT)₂₅ Dynabeads (Life Technologies, 61011) was equilibrated in lysis buffer, mixed with the extracts (~5 mg) and incubated on a shaker for 10 min at room temperature (RT). The beads were collected with a magnet, and the supernatant was recovered for repeat incubations (described below). The beads were washed once with 500 µl of wash buffer A (10 mM Tris-HCl, pH 7.5, 600 mM LiCl, 1 mM EDTA, and 0.1% Triton X-100) and twice with 500 µl wash buffer B (10 mM Tris-HCl, pH 7.5, 600 mM LiCl, and 1 mM EDTA). The poly(A) RNA was eluted from beads in 60 µl of 10 mM Tris-HCl, pH 7.5, at 80 °C for 2 min and collected. The entire procedure was repeated twice by reapplying the supernatant to the oligo(dT)₂₅ beads, which were recovered after elution and washed three times with lysis buffer before reapplication. The three sequential eluates were combined and concentrated to 70 µl in a 0.5 ml Microcon-10kDa centrifugal filter unit with an Ultracel-10 membrane (Millipore, MRCPT010). In total, we subjected six samples for MS analysis: three independent experiments of the mRBPome, two RNase ONE-treated samples and one poly(A) competition experiment.

***C. elegans* cultures and apoptosis induction.** Bristol N2 worms were cultured at 20 °C on NGM plates (0.3% NaCl, 1.7% agar, 0.25% peptone, 1 mM CaCl₂, 5 µg/ml cholesterol, 1 mM MgSO₄, and 25 mM KPO₄ buffer, pH 6.0) seeded with OP50 strain *Escherichia coli* according to standard procedures (<http://www.wormbook.org/>). Synchronization was reached by bleaching gravid adults. To induce germ-cell death, ~120,000 worms were synchronized in late L4 and treated with 5 mM of *N*-nitroso-*N*-ethylurea (ENU, Sigma, N3385) in 4 ml M9 buffer (0.3% KH₂PO₄, 0.6% Na₂HPO₄, 0.5% NaCl, and 1 mM MgSO₄) in 15-ml Falcon tubes for 4 h with constant rotation¹⁸.

In vivo capture of mRBPs in *C. elegans*. ~120,000 worms were pelleted and washed three times in M9 buffer; this was followed by 15-min incubation with 15 ml M9 buffer on a rotatory wheel. Worms were transferred to an NGM plate and exposed to UV light (254 nm) at 300 mJ/cm² in a Stratallinker 1800 (Stratagene) and collected in M9 buffer and resuspended in lysis buffer (100 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA, 0.75% IGEPAL, 1 mM DTT, 20 U ml⁻¹ DNase I (Promega), 100 U ml⁻¹ RNasin (Promega), and complete EDTA-free protease-inhibitor cocktail (Roche)). To prepare extracts, the worms or larvae were ground in a mortar filled with liquid nitrogen, and the lysate including the fat layer was subsequently clarified at 14,000g for 10 min and passed through a 0.45-µm filter (Millipore). For the controls, worm extracts were treated with 100 U RNase ONE (Promega), or 20 mg of poly(A) (Sigma) was added for competition experiments. The poly(A) RNA and cross-linked RBPs were isolated from 1.25 ml (~12 mg) of mixed-stage worm extracts and 1 ml (~10 mg) of synchronized L4 worm extracts in three sequential rounds, as described for yeast with minor modification: the volumes of beads and buffers were scaled up in relation to the amount of input extract; and 500 mM LiCl was used in wash buffers. In total, we subjected

18 samples to MS analysis: three independent experiments containing ~120,000 animals of mixed-stage worms as well as two RNase ONE-treated and one poly(A)-competitor control; three independent experiments of synchronized larvae, stage 4 (L4), and three matched experiments of synchronized L4 treated with 5 mM ENU, along with their corresponding three poly(A) competition control experiments.

LC-MS/MS and protein identification. MS analysis was performed at the Proteomics Facility, University of Bristol. 50 µl of the samples (~8–10 µg of protein) was run on a 4–12% NuPAGE Novex acrylamide gel (Life Sciences). The gel lane was cut into one slice and subjected to in-gel tryptic digestion with a ProGest automated digestion unit (Digilab UK). The resulting peptides were fractionated with a Dionex Ultimate 3000 nanoHPLC system in line with an LTQ-Orbitrap Velos mass spectrometer (Thermo Scientific). In brief, peptides in 1% (vol/vol) formic acid were injected onto an Acclaim PepMap C18 nanotrap column (Dionex). After washing with 0.5% (v/v) acetonitrile 0.1% (v/v) formic acid, peptides were resolved on a 250 mm × 75 µm Acclaim PepMap C18 reverse-phase analytical column (Dionex) over a 150-min organic gradient, with seven gradient segments (1–6% solvent B over 1 min, 6–15% B over 58 min, 15–32% B over 58 min, 32–40% B over 3 min, and 40–90% B over 1 min, held at 90% B for 6 min and then reduced to 1% B over 1 min) with a flow rate of 300 nl min⁻¹. Solvent A was 0.1% formic acid, and solvent B was aqueous 80% acetonitrile in 0.1% formic acid. Peptides were ionized by nano-electrospray ionization at 2.1 kV with a stainless-steel emitter with an internal diameter of 30 µm (Thermo Scientific) and a capillary temperature of 250 °C. Tandem mass spectra were acquired with an LTQ Orbitrap Velos mass spectrometer controlled by Xcalibur 2.1 software (Thermo Scientific) and operated in data-dependent acquisition mode. The Orbitrap was set to analyze the survey scans at 60,000 resolution (at m/z 400) in the mass range m/z 300 to 2,000 and the top 20 multiply charged ions in each duty cycle selected for MS/MS in the LTQ linear ion trap. Charge-state filtering, in which unassigned precursor ions were not selected for fragmentation and dynamic exclusion (repeat count, 1; repeat duration, 30 s; exclusion list size, 500), was used. Fragmentation conditions in the LTQ were as follows: normalized collision energy, 40%; activation q, 0.25; activation time 10 ms; and minimum ion selection intensity, 500 counts.

The raw data files were processed and quantified with Proteome Discoverer software v1.2 (Thermo Scientific) and searched against the SwissProt SPECIES database with the Mascot algorithm (version 2.4). Peptide precursor mass tolerance was set at 10 p.p.m., and MS/MS tolerance was set at 0.8 Da. Search criteria included carbamidomethylation of cysteine (+57.0214) as a fixed modification and oxidation of methionine (+15.9949) as a variable modification. Searches were performed with full tryptic digestion, and a maximum of one missed cleavage was allowed. The reverse database search option was enabled, and all peptide data were filtered to satisfy an FDR threshold of <1%.

To define a list of proteins referred to as the yeast mRBPome, we initially considered the 1,425 proteins that were enriched in at least one of the three independent experiments with an FDR of at least 1% at the protein identification and quantification level (raw data in **Supplementary Table 1**). Of these, we selected 973 proteins that were identified with at least two peptides and for which data were obtained in two out of three independent experiments, leaving 765 proteins. 116 of these proteins were also identified in at least one of our three control samples (63 and 111 proteins in two RNase ONE-treated controls and 16 proteins in poly(A) competition). However, the respective proteins were not excluded, because peptide numbers were at least two-fold higher in mRBP samples compared to the controls (except Pab1p, 1.6-fold).

To define the *C. elegans* mRBPome, we initially considered the 1,016 proteins that were enriched in at least one of the nine UV-cross-linked mRBPome samples with a FDR <1% (raw data in **Supplementary Table 3a**). We further selected those proteins that were represented by at least two peptides (691 proteins) in at least two of the nine samples, revealing 594 proteins. 43 of them were also seen in at least one of our nine negative-control samples but were not excluded from the mRBPome because the number of peptides for respective proteins was at least two-fold lower in the negative controls compared to mRBP samples.

***C. elegans* MS data processing and label-free quantification.** Only data for proteins identified as having at least two peptides across samples were subjected to processing and analysis (peak areas for proteins not meeting our criteria were set

to zero). For each protein within the experiment, the 'raw' MS peak area in the technical control sample was subtracted from the corresponding 'raw' MS peak area in the biological sample to produce 'background-subtracted' peak areas. Subsequently, to normalize each background-subtracted peak area, as a fraction of the total area measured by MS for a given sample, each background-subtracted peak area in a sample was divided by the total background-subtracted area for that sample (normalized data in **Supplementary Table 3b**). An arbitrary value of 0.000001 was then added to all normalized peak areas to avoid zero values and to enable the calculation of a fold change between conditions that included presence versus absence. To determine the relative changes in protein abundance between *C. elegans* L4+ENU and L4 mRBPomes, we calculated $\log_2$ fold changes for each of the three replicates by $\log_2$ (L4+ENU_normalized_area_replicate_x / L4_normalized_area_replicate_x), where x is the same replicate number, owing to the paired nature of the experimental design. We then determined the average $\log_2$ fold change between L4+ENU and L4 treated animals across all three replicates. To estimate an FDR for declaring differentially abundant proteins between conditions, we generated 10,000 data sets by randomly shuffling the experimentally obtained fold changes of replicates. From this randomized data set, we then computed average $\log_2$ fold changes to assess, for each possible $\log_2$ L4+ENU/L4 threshold of -10 to 10 in 0.5 increments, the fraction of the 10,000 randomly generated fold changes that were above that threshold. From these values, it was possible to linearly interpolate FDRs for all experimentally obtained (nonrandom) fold changes. Thus, a suitable $\log_2$ L4+ENU/L4 threshold to determine differential protein abundance with a false discovery rate of less than 5% can be defined as having less than 500 of the 10,000 randomly generated $\log_2$ L4+ENU/L4 values greater than that threshold (fold changes and FDRs in **Supplementary Table 3b**).

Immunoblot analysis. 0.1% and 0.05% of the yeast and worm input extract respectively (~5 µg of protein) and 10% of the eluates were resolved on 4–15% SDS polyacrylamide gels and transferred to polyvinylidene difluoride (PVDF) membranes (Thermo Scientific Pierce). Membranes were blocked in PBS with 0.1% Tween-20 and 5% low fat milk, probed with the indicated antibodies and horseradish peroxidase (HRP)-coupled secondary antibodies, and developed with the Immobilon Western Chemiluminescent HRP Substrate (Millipore). Blots were recorded with a FluorChem (Alpha Innotech). The following antibodies were used: rabbit anti-Scp160 (ref. 11) (1:10,000), mouse anti-Act1 (1:2,500; MP Biomedicals, 0869100), mouse anti-GLD-1 (ref. 43) (1:50), mouse anti-CYC-1 (1:1,000; Invitrogen, 456100), mouse anti-GFP (1:2,000; Roche, 11814460001), HRP-conjugated donkey anti-rabbit IgG (1:5,000; Amersham, NA9340V), and HRP-conjugated sheep anti-mouse IgG (1:5,000; Amersham, NXA931). The validation of all commercial primary antibodies is provided on the manufacturers' websites.

IP of proteins from UV-cross-linked samples and fluorescence mRNA-binding assay. The fluorescence RNA-binding assay is based on a protocol²⁸ that we adapted to yeast. Yeast strains expressing endogenously GFP-tagged proteins⁴⁴ were verified by immunoblot analysis (data not shown). Strain BY4741 was used for mock-control affinity isolations. Cells were grown in 75 ml of YPD at 30 °C with constant shaking at 220 r.p.m. and collected at mid-log phase by centrifugation and washed with PBS. The cells were resuspended in 4 ml of PBS and exposed to UV light for cross-linking as described above. Cells were collected and resuspended in 650 µl of lysis buffer (100 mM Tris-HCl, pH 7.5, 500 mM LiCl, 10 mM EDTA, 1% Triton X-100, 5 mM DTT, 20 U ml⁻¹ DNase I (Promega), 100 U ml⁻¹ RNasin (Promega), and complete EDTA-free protease-inhibitor cocktail (Roche)), and extracts were prepared as described above.

To IP GFP-tagged proteins, 500 µl cell extract (2.5 mg protein) was incubated with 20 µl of preequilibrated GFP-Trap_A agarose beads (Chromotek, gta-20) for 2 h under shaking at 1,000 r.p.m. in a ThermoThermoshaker (PeqLab) at 4 °C. Beads were washed twice with 750 µl of wash buffer C (10 mM Tris-HCl, pH 7.5, 500 mM LiCl, 1 mM EDTA, 0.1% Triton X-100, 0.05% SDS, 50 U ml⁻¹ RNasin (Promega), and complete EDTA-free protease-inhibitor cocktail (Roche)) and collected by centrifugation at 2,500g for 2 min at 4 °C. The beads were then resuspended in 800 µl of wash buffer C, and then half of the sample (400 µl) was incubated with 100 U ml⁻¹ of RNase ONE (Promega), and the other half was left

untreated. Samples were incubated for 30 min at 4 °C. Beads were washed three times with 750 µl of wash buffer C and incubated with 500 µl of blocking buffer (10 mM Tris-HCl, pH 7.5, 500 mM LiCl, 1 mM EDTA, 0.01% Triton X-100, 100 µg/ml *E. coli* tRNA, 100 µg/ml BSA, 50 U ml⁻¹ RNasin (Promega), and EDTA-free protease-inhibitor cocktail (Roche)) for 15 min in a shaker at 4 °C. After blocking, the samples were incubated with 500 µl of hybridization buffer (10 mM Tris-HCl, pH 7.5, 500 mM LiCl, 1 mM EDTA, 0.01% Triton X-100, 0.05% LiDS, 5 mM DTT, and 100 U ml⁻¹ RNasin (Promega)) supplemented with 40 nM of oligo(dT)₂₅ labeled with Alexa 594 (IDT) for 1 h in the dark in a shaker at 4 °C. Finally, samples were washed three times with 500 µl of wash buffer D (10 mM Tris-HCl, pH 7.5, 500 mM LiCl, 1 mM EDTA, 0.01% Triton X-100, 0.01% LiDS, 5 mM DTT, 50 U ml⁻¹ RNasin (Promega), and complete EDTA-free protease-inhibitor cocktail (Roche)). Beads were resuspended in 100 µl of wash buffer D and transferred to an opaque 96-well plate. Fluorescence measurements were performed in a FLUOstar Omega microplate reader (BMG LABTECH), with the gain set to the positive control (Pab1) and with the following filters: GFP, Ex485-12, Em520; Alexa594, Ex584, Em620-10. Background signal from the buffer was subtracted, and the data were normalized to the average of the mock-control sample (triplicates). Three independent experiments were performed for each tagged strain, and fluorescence for each experiment was measured in triplicate.

TAP-tagged proteins⁴⁵ were immunopurified in the same manner as for the GFP-tagged proteins, with the following modifications: 100 µl of Pan mouse IgG Dynabeads (Life Technologies, 11041) was used to capture tagged proteins, and the proteins were released from beads in 100 µl elution buffer (10 mM Tris-HCl, pH 7.5, 150 mM LiCl, 1 mM EDTA, and 1 mM DTT) containing 80 U ml⁻¹ of AcTEV protease (Life Technologies, 12575-023) for 2 h at 19 °C. IP was controlled by immunoblot analysis (data not shown).

Puromycin treatment and affinity isolation of tagged proteins. Isolation of TAP-tagged proteins from non-cross-linked cells and puromycin treatment of extracts is described in **Supplementary Note 1**.

RT-PCR. To assess induction of apoptosis in L4 worms, total RNA was isolated from 50 µl of extracts with a ZR RNA MiniPrep Kit (Zymo Research) with in-column DNA digestion. RT was performed with 500 ng of RNA combined with a mixture of oligo(dT)₁₈ and random hexamer primers and the Transcriptor High Fidelity cDNA Synthesis Kit, according to the manufacturers' instructions (Roche, 05091284001). PCR was performed with 1 µl of complementary DNA (cDNA) for 5 min at 94 °C, 35 cycles at 94 °C for 30 s, 57 °C for 30 s, 72 °C for 40 s, and 8 min at 72 °C with the primers listed in **Supplementary Table 6**. Mpk-1 was used as a housekeeping-gene control.

To identify mRNA targets of glycolytic enzymes, total RNA was isolated from 50 µl of extract (input) and 100 µl of GFP or TAP IPs. 500 ng of input total RNA or 9.4 µl of immunoprecipitated RNA (~28% of IP) were combined with a mixture of oligo(dT)₁₈ and random hexamer primers for RT with the Transcriptor High Fidelity cDNA Synthesis Kit (Roche). PCR was performed with 1 µl of cDNA reaction with gene specific primer pairs (**Supplementary Table 6**). PCR was performed for 5 min at 94 °C, 33 cycles at 94 °C for 30 s, 57 °C for 30 s, 72 °C for 40 s, and 8 min at 72 °C. For amplification of *ENO1* and *SET1*, 30 cycles were used.

Bioinformatics and statistical analysis. A detailed description of statistical tests and bioinformatics (GO, Wikipathway, domain and motif analysis) is given in **Supplementary Note 1**. The R code script for performing the hypergeometric test and Benjamini and Hochberg (BH) FDR correction with the p.adjust function is given in **Supplementary Note 2**.

43. Scheckel, C., Gaidatzis, D., Wright, J.E. & Ciosk, R. Genome-wide analysis of GLD-1-mediated mRNA regulation suggests a role in mRNA storage. *PLoS Genet.* **8**, e1002742 (2012).

44. Huh, W.K. *et al.* Global analysis of protein localization in budding yeast. *Nature* **425**, 686–691 (2003).

45. Ghaemmaghami, S. *et al.* Global analysis of protein expression in yeast. *Nature* **425**, 737–741 (2003).