BMC Genomics

RESEARCH ARTICLE   **Open Access**

Factors affecting the rapid changes of protein under short-term heat stress

Bingjin Wu[1], Jianwen Qiao[2], Xiaoming Wang[1], Manshuang Liu[1], Shengbao Xu[1*] and Daojie Sun[1*]

Abstract

Background: Protein content determines the state of cells. The variation in protein abundance is crucial when organisms are in the early stages of heat stress, but the reasons affecting their changes are largely unknown.

Results: We quantified 47,535 mRNAs and 3742 proteins in the filling grains of wheat in two different thermal environments. The impact of mRNA abundance and sequence features involved in protein translation and degradation on protein expression was evaluated by regression analysis. Transcription, codon usage and amino acid frequency were the main drivers of changes in protein expression under heat stress, and their combined contribution explains 58.2 and 66.4% of the protein variation at 30 and 40 °C (20 °C as control), respectively. Transcription contributes more to alterations in protein content at 40 °C (31%) than at 30 °C (6%). Furthermore, the usage of codon AAG may be closely related to the rapid alteration of proteins under heat stress. The contributions of AAG were 24 and 13% at 30 and 40 °C, respectively.

Conclusion: In this study, we analyzed the factors affecting the changes in protein expression in the early stage of heat stress and evaluated their influence.

Keywords: Heat stress, Posttranscriptional regulation, Codon usage, AAG frequency, Wheat

Background

The fluctuation of protein abundance determine the state of cells, and the elements that affect protein expression have been extensively studied in recent years [1–4]. The rapid development of high-throughput technology provides technical support for research. Transcriptome sequencing is accurate and efficient, but due to the bottleneck of proteomics, it is incapable of quantifying a more comprehensive protein landscape. This discrepancy makes people usually use transcription to speculate about protein expression. Studies have shown that transcription is a weak proxy for protein abundance (R^2, 0.2–0.4) in various species [5–10], especially in stressful environments ($R^2 < 0.09$) [11–13]. In addition to the fact that protein synthesis fails to keep up with the pace of transcription, the reason for the weaker correlation under stress is that the synthesis of some proteins, such as transcription factor [14], does not depend on transcription and increases even before transcription changes [15–17]. Thus, it has been indicated that there are other regulatory mechanisms to help protein expression under stress conditions.

Although transcriptional regulation is important for protein expression, it is insufficient to represent protein variation. The genetic code, which is a template for protein synthesis, also contains information to regulate protein expression. The effects of amino acid [4], untranslated regions [18–21], length [21, 22], GC content [23–25], and mRNA secondary structure [26] have been confirmed in previous studies. Codon usage is regarded as one of the major factors in controlling elongation during translation, and the usage of preferred codons in the coding sequence can significantly increase the rate of protein synthesis [27–30]. In addition, microRNA-mediated

post-transcriptional gene silencing [31] and protein degradation [32] also play an important role in regulation of protein abundance.

Previous studies have focused on the correlation between overall transcription and protein. However, when an organism is subjected to stress, most of the protein expression is stagnated, and only a small portion of the proteins bypass this inhibition and are expressed rapidly in abundance to reduce damage. The most typical example is molecular chaperones, whose expression pattern is still a mystery.
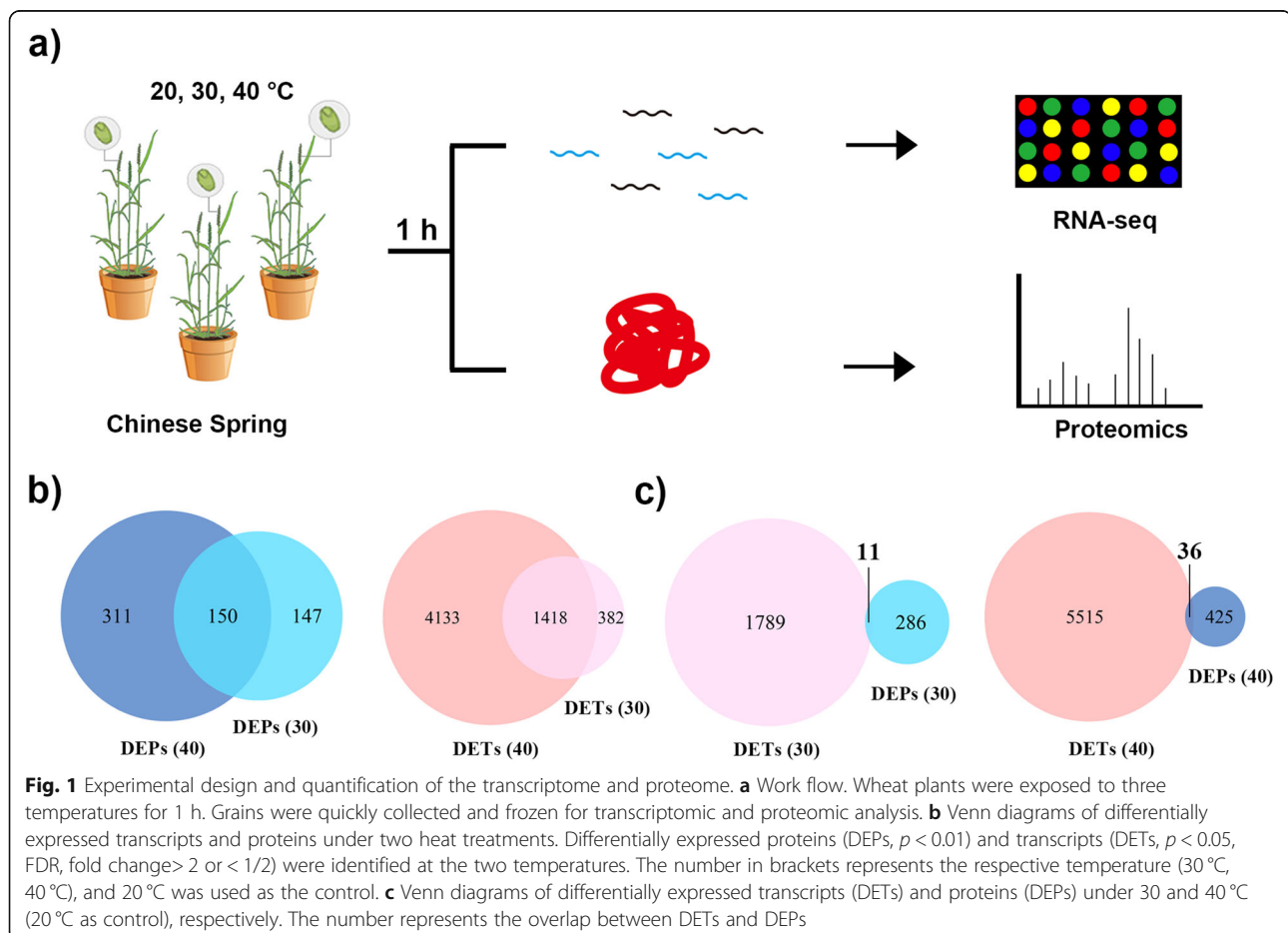
Wheat plants are sensitive to heat, especially in the middle and late stages of grain filling [33–36]. Organisms adapt to adversity by altering their protein abundance, and the factors that are related to fluctuations in protein abundance in the short term are still unclear. In this study, transcriptomic and proteomic analyses were performed on wheat grains under two types of short-term heat stress. The influencing factors of differentially expressed proteins were investigated in varying degrees of thermal environments. This study suggests that wheat grains have various response measures for different levels of heat stress

and that posttranscriptional regulation plays a crucial role in regulating protein expression to adapt to constantly changing temperatures.

Results

Quantification of the wheat grain transcriptome and proteome under short-term heat stress

To understand how filling grain quickly adapts to temperature variations, wheat (*Triticum aestivum* cv. Chinese Spring) plants were subjected to 20, 30, or 40 °C for 1 h. The most suitable growth temperature for wheat is 12–22 °C [37], while 30 and 40 °C are different levels of heat stress [38]. We evaluated both transcriptomic and proteomic profiles of wheat grains in the three environments (Fig. 1a). A total of 47,535 transcripts were acquired (FPKM > 1 in any circumstances) via the transcriptome. Of these, 1800 and 5551 were identified as differentially expressed transcripts (DETs; fold change > 2, $p < 0.05$, FDR; Table S1) under two types of heat environments (30 and 40 °C; 20 °C served as the control). Through TMT proteomic analysis, 3742 proteins were quantified, including 297 and 461 differentially expressed proteins (DEPs; T-test, $p < 0.01$; Table S2).

Fig. 1 Experimental design and quantification of the transcriptome and proteome. **a** Work flow. Wheat plants were exposed to three temperatures for 1 h. Grains were quickly collected and frozen for transcriptomic and proteomic analysis. **b** Venn diagrams of differentially expressed transcripts and proteins under two heat treatments. Differentially expressed proteins (DEPs, $p < 0.01$) and transcripts (DETs, $p < 0.05$, FDR, fold change> 2 or < 1/2) were identified at the two temperatures. The number in brackets represents the respective temperature (30 °C, 40 °C), and 20 °C was used as the control. **c** Venn diagrams of differentially expressed transcripts (DETs) and proteins (DEPs) under 30 and 40 °C (20 °C as control), respectively. The number represents the overlap between DETs and DEPs

Wu *et al. BMC Genomics*   (2021) 22:263

Page 3 of 11

More DETs and DEPs were identified under severe stress (40 °C) than under mild heat stress (30 °C), indicating that filling grains express more transcripts and proteins to respond to severe stress. In addition to 150 proteins and 1418 transcripts expressed under the two thermal treatments, wheat grains expressed specific genes and proteins in response to a certain degree of heat stress (Fig. 1b). These results revealed that grains have corresponding measures in response to various levels of heat stress.

What is striking is that there is little overlap between DETs and DEPs under short-term heat stress, whether it be mild (30 °C) or severe (40 °C) (Fig. 1c). That is, after a short-term heat stress, a small portion of the corresponding DETs and DEPs changed simultaneously, and the transcripts corresponding to most of the DEPs did not differ significantly.

Transcription and protein levels in response to heat stress

To further study the function of transcripts and proteins that respond to heat stress in a short time, GO and KEGG enrichment analysis was performed on differentially expressed transcripts and proteins under the two types of thermal environments.

Protein folding, one of the main ways in response to heat [39], was enriched according to GO enrichment analysis across DETs and DEPs in both treatments (Fig. 2a). Protein expression was triggered from the 1-h exposure to the thermal environment, regardless of whether it involved in mild or severe stress; the reaction speed of proteins to a disturbed environment may exceed our expectations.

More substances were enriched under severe heat stress than under mild stress at both the transcriptional and protein levels, not only increasing the number but also broadening the capability (Fig. 2). Consistent with few overlaps between DETs and DEPs (Fig. 1c), the enriched processes and pathways were also significantly different between the transcription and protein levels (Fig. 2a, b). In addition to "protein folding" and "protein processing in the endoplasmic reticulum" (Fig. 2), transcription and proteins play regulatory roles in different fields during short-term heat stress. These results imply that transcription may not be the main way to regulate the expression of proteins involved in rapid thermal responses.

As protein synthesis machine, the ribosome, was significantly enriched in the GO and KEGG analysis (Fig. 2). However, the expression of ribosomal proteins changed dramatically only at the protein level, rather than at the transcriptional level. That is to say, ribosome expression under short-term heat stress is free from transcriptional regulation or mRNA returns to undisturbed levels rapidly after translation.

Pattern of ribosomes in response to short-term heat stress

To clarify the expression patterns of ribosomal proteins under short-term heat stress, 141 ribosomal proteins identified via proteomics under two heat stresses were
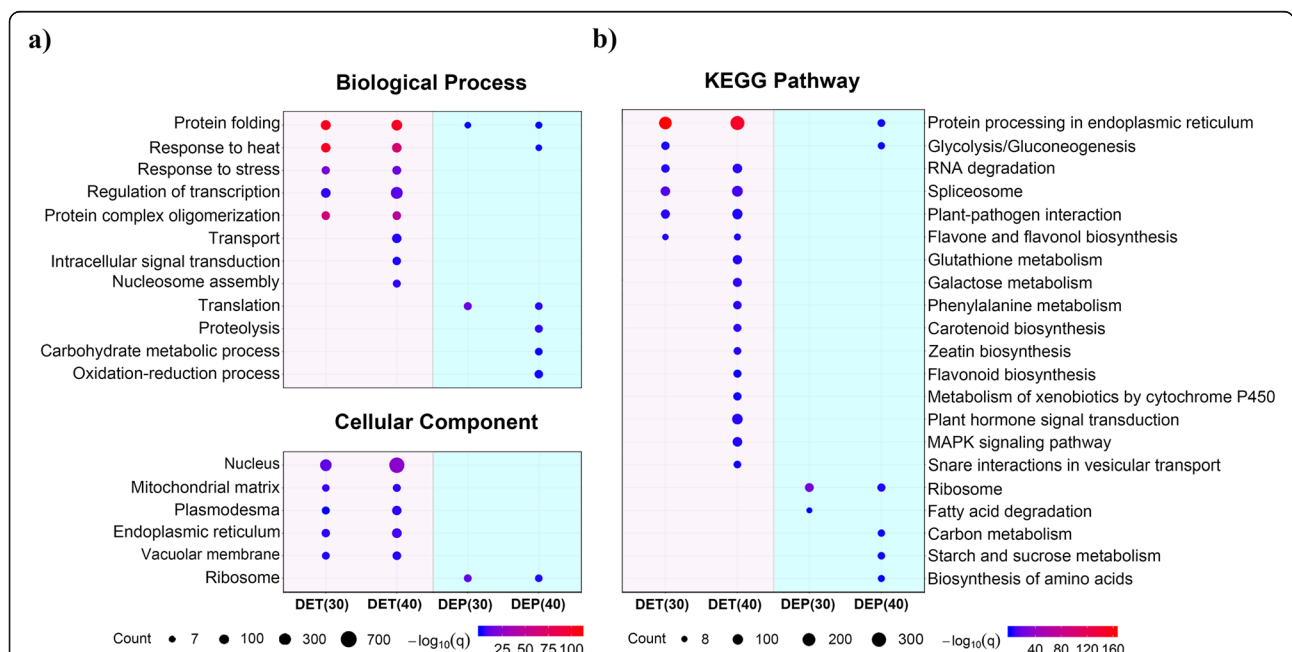
Fig. 2 Enrichment analysis of DETs and DEPs under two types of heat stress. **a** GO enrichment. Functional enrichment analysis of DETs and DEPs associated with biological processes and cellular components under 30 and 40 °C heat treatments ($N \geq 7$, q < 0.05). **b** KEGG enrichment. Enriched pathways of DETs and DEPs in the two types of heat environments ($N > 7$, q < 0.05)

used for expression pattern analysis. The expression of twenty-eight and twenty-two ribosomes changed significantly at 30 and 40 °C, respectively. In addition to one or two ribosomes whose expression changed at the transcriptional level, the expression of the others increased only at the protein level (Fig. 3). Detailed information is shown in the supplementary figure (Fig. S1).

As shown in Fig. 3, ribosomal protein was expressed in large quantities, and transcription had not changed or returned to a resting state in a short time, regardless of mild (30 °C) or severe (40 °C) heat stress. This class of proteins may adopt an unknown strategy to achieve a large amount even if protein expression is suppressed.

Factors affecting the rapid alteration of proteins under short-term heat stress

To investigate the factors affecting the variation in DEP expression under two types of short-term heat conditions, 107 factors and protein changes under two kinds of heat stress were used for regression analysis respectively. These candidate items were found to be involved in regulating protein expression in previous studies, including transcriptional alteration, codon usage, amino acid frequency, length, and a base frequency of coding sequences or untranslated regions.

We used fold change (fold change = treatment/control) as a measure to evaluate the changes in transcription and protein abundance. Through correlation analysis, in addition to the transcription, more than 20 sequence characteristics were identified as being significantly related to the alteration in protein expression (Table 1). Although transcription significantly correlated with protein variation under both treatments (0.24 at 30 °C, $p < 0.01$; 0.55 at 40 °C, $p < 0.01$), the correlation was stronger under severe heat stress. These results suggested that transcription provided greater support for altering protein expression under severe heat stress.

Interestingly, the frequency of lysine was very strongly correlated with protein variation, even more so than transcription was (Table 1). Lysine is encoded by two codons, AAG and AAA. AAG was also associated with protein changes (0.48 at 30 °C, $p < 0.01$; 0.40 at 40 °C,
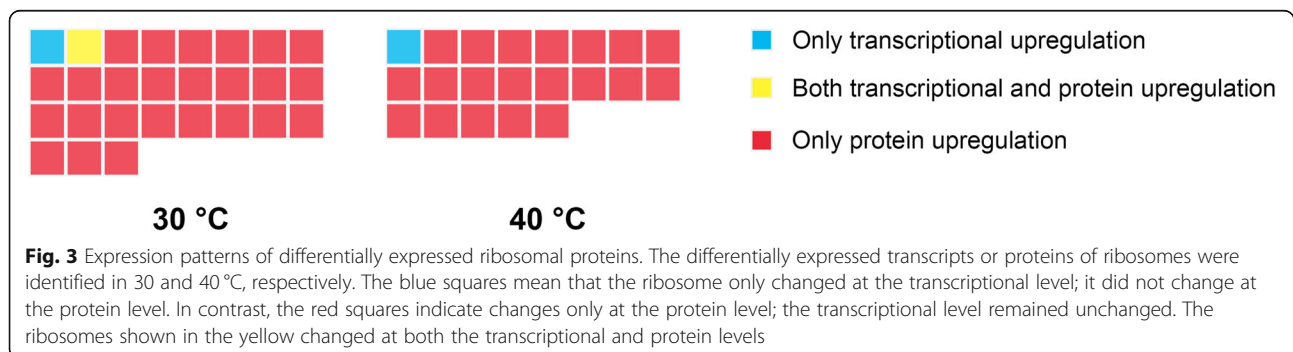
$p < 0.01$), while AAA was not. It is easy to determine that the correlation coefficient between AAG frequency and protein alteration was nearly the same as that of lysine. Namely, AAG usage, rather than lysine frequency, is one of the major factors regulating changes in protein abundance. In addition, several codons, amino acids, and base frequency also showed a strong relationship with protein variation in different situations. Sequences containing more factors which positively associated with protein expression and less negatively associated factors are more likely to respond to short-term heat stress.

Contribution of factors to protein expression under heat stress

To explore the underlying mechanism of the adjustment of heat-responsive proteins in two types of heat stress, a regression analysis of the 107 factors was performed to evaluate their contribution. We applied multivariate adaptive regression spline (MARS) to fit the model and calculate the importance of each variable under two thermal environments, and linear regression and elastic net were used for verification. The equations fitted by the factors explained 58.2 and 66.4% of the changes in protein under mild (30 °C) and severe (40 °C) heat stress, respectively (Fig. 4, Table S3). Codon usage, transcription, and amino acid frequency had the most significant impact on protein expression.

Limited contribution (6%) of transcription to protein changes occurred under mild heat stress (30 °C); however, transcription had a dramatic effect (31%) on protein alteration when plants were under extreme heat stress (40 °C) (Fig. 4). These findings showed that transcriptional regulation may have a marginal impact under mild stress, but have a significant role in the acute thermal response.

Interestingly, codon usage largely affected protein expression (37% at 30 °C; 25% at 40 °C), whether under mild or severe heat stress, indicating that codon preference strongly supports the rapid variation in proteins under heat stress. In line with the correlation results (Table 1), the codon AAG may play an important role in the rapid alteration of proteins under heat stress. The

Fig. 3 Expression patterns of differentially expressed ribosomal proteins. The differentially expressed transcripts or proteins of ribosomes were identified in 30 and 40 °C, respectively. The blue squares mean that the ribosome only changed at the transcriptional level; it did not change at the protein level. In contrast, the red squares indicate changes only at the protein level; the transcriptional level remained unchanged. The ribosomes shown in the yellow changed at both the transcriptional and protein levels

Table 1 Correlation analysis

Characteristic	P (30)	P (40)
Transcript		
Transcript fold change	0.24**	0.55**
Amino Acid Frequency		
A	−0.22**	0.09
C	−0.11	−0.24**
D	−0.30**	−0.16**
E	0.03	0.26**
K	0.48**	0.39**
P	−0.27**	−0.06
R	0.34**	0.21**
S	−0.20**	−0.22**
Y	0.10	−0.22**
Codon Usage		
TAC (Y)	0.06	−0.22**
GAC (D)	−0.27**	−0.1
AAG (K)	0.48**	0.40**
AAA (K)	0.02	0.05
GAG (E)	0.07	0.30**
CCG (P)	−0.27**	−0.10
AGC (S)	−0.30**	−0.21**
TGC (C)	−0.09	−0.22**
CGC (R)	0.18**	0.20**
AGG (R)	0.29**	0.11*
CGT (R)	0.39**	0.12*
Base Frequency		
A (CDS)	0.20**	0.15**
T (3′UTR)	0.25**	0.23**
GC (5′UTR)	0.27**	0.15**

Factors related to protein expression at two heat treatment. P (30) and P (40) represent fold changes at 30 °C and 40 °C, respectively. ** The correlation is significant at the 0.01 level (2-tailed). * The correlation is significant at the 0.05 level (2-tailed)

contributions of AAG were 24 and 13% at 30 and 40 °C, respectively (Fig. 4). Although the lysine content was also significantly correlated with variation of protein expression, the frequency of lysine (K) was removed from the equation after MARS analysis. That is, the critical function for rapid protein expression is AAG instead of lysine. Similar results were also obtained by linear regression and elastic net (Table S4), supporting the credibility of the equation.

When an organism responds to environmental disturbances, only transcriptional regulation may be too slow. The mechanism of regulating protein expression is written in the sequence, which may be the fastest and most effective response mode.

The relationship between protein expression and the AAG occurrence frequency under short-term heat stress

To verify whether the up-regulated proteins under short-term heat stress are rich in AAG codon, two previously published proteomic datasets of yeast subjected to heat stress were analyzed [40, 41]. The AAG codon occurrence frequencies were calculated for the transcripts corresponding to the whole-genome proteins (all the proteins annotated in yeast genome), the proteins identified by proteome analysis and the differentially expressed proteins under heat stress (Fig. 5a, b).

Compared with the whole-genome proteins, higher AAG usage was observed in the coding sequence of protein identified in the proteomic analysis (Fig. 5a, b). Furthermore, the transcripts encoding up-regulated proteins under short-term heat stress have significantly higher AAG frequency than that of the proteins identified in the proteome analysis, which indicated that codon AAG might play an important role in the rapid expression of proteins responsive to heat stress (Fig. 5a, b).

Discussion

Here, we performed transcriptome and proteome sequencing of filling grains subjected to different degrees of heat stress (30 °C, 40 °C) for 1 h. Transcription and elements related to protein expression were investigated for rapid protein variation under two types of short-term heat stress.

Factors affecting protein expression

In recent decades, the contribution of transcription to protein has been widely discussed in different kingdoms and environmental conditions. Yeast is one of the simplest eukaryotic organisms. In yeast, transcription determined about 60% of protein expression [1, 42]. In mice and humans, 40 and 27% of protein expression were controlled by transcription [6, 43, 44], respectively. It is indicated that the more complex the organism, the more limited the role of transcriptional regulation.

Previous studies mainly focused on the relationship between total transcription and protein. However, our research has concentrated on the factors that affect the expression of changed proteins under disturbed conditions. Therefore, the contribution of transcription to protein expression depends on the degree of stress. We evaluated the contribution of transcript abundance changes to protein expression under various degrees of heat stress. The transcriptional regulation of proteins under 40 °C (31%) was more intense than that under 30 °C (6%). The posttranscriptional regulation always played a stable role (36–52%). Posttranscriptional regulation played a more important role than did transcription under mild stress, while under severe heat stress, transcription and posttranscriptional regulation worked
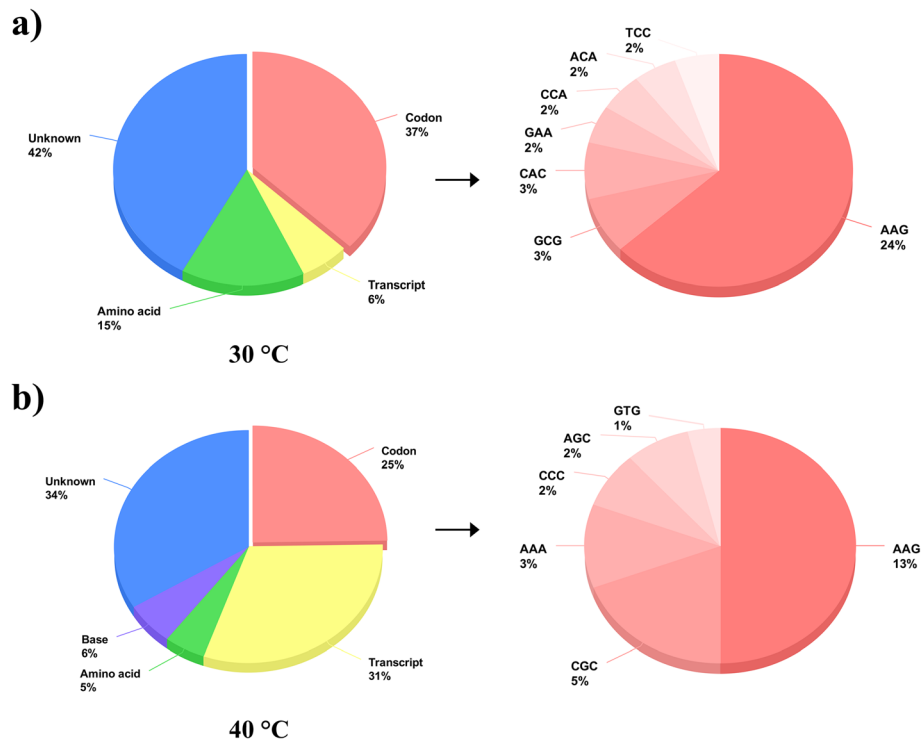
Fig. 4 Contributions of various factors to protein expression under two types heat stress. **a** and **b** Factors and their contributions to protein expression under 30 and 40 °C. By analyzing transcriptome data, proteome data and 107 sequence characteristics, the factors affecting protein expression and their contributions were shown on the pie chart under 30 and 40 °C, respectively. The pie on the right is an expanded view of the codon usage

synergistically to express the desired protein faster and stronger.

In previous studies, researchers often used the codon adaptation index (CAI) to characterize the impact of codons on translation, which explain up to 6% of protein expression [1]. We evaluated the effect of each codon on protein changes under heat stress, explaining 25–37% of protein changes. The impact of codons on protein expression may be beyond our knowledge. The usage of amino acids is rarely involved in previous studies. Our research indicated that the usage frequency of amino acids may also be related to protein expression.
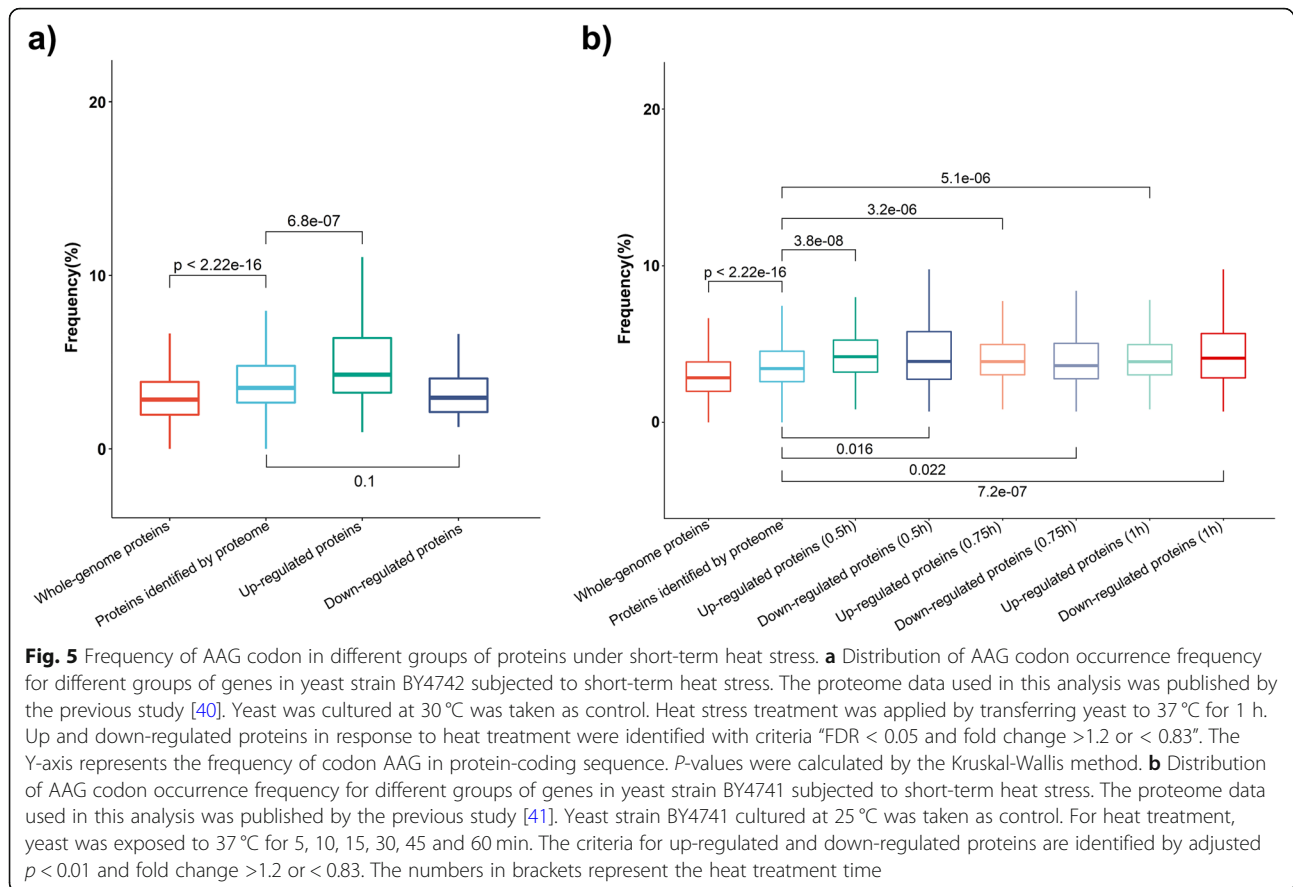
Ribosomal protein in response to heat stress

Protein synthesis is the most time- and energy-consuming process in organisms. If something goes wrong, the consequences can be catastrophic, especially under adverse conditions. Therefore, as protein synthesis machines, ribosomes need to adjust their production plan in time to quickly express the required proteins to reduce the damage caused by stress.

A total of 141 ribosomal proteins were identified via proteomics, among which 28 and 22 were up-regulated under two types of short-term thermal stress (30 °C and 40 °C, respectively). Due to the limitation of proteomics technology, ribosomal proteins with altered expression

should be amplified proportionately. Approximately 14–20% of ribosomes are involved in the translation of heat-responsive proteins. This feature of ribosomes involves the preferential translation of a subset of functionally-related mRNAs and has been a popular research topic in recent years [45–48]. Heterogeneous ribosomes facilitate the synthesis of stress response proteins and help cells cope with environmental changes.

As shown in Fig. 3, a portion of the upregulated ribosomal protein expression may not be associated with transcriptional alteration. This pattern of regulation of ribosomes is often observed in cancer and stress research. In the study of rhabdomyosarcoma, it was found that the abundance of eL36 and eL42 (60S ribosomal protein L36 and L42) increases, while the corresponding mRNA decreases [49]. Several ribosomal proteins, such as CAC1787 (30S RPS2), CAC3105 (30S RPS4), CAC3147 (50S RPL1) and CAC3132 (50S RPL4), are also disconnected from transcription under butanol stress in *Clostridium acetobutylicum* [50]. The short-term adaptation of cells to new states usually requires the involvement of posttranscriptional mechanisms, because transcriptional regulation alone would be too slow. High-level translation of existing transcripts can help to synthesize needed proteins rapidly, while the targeted degradation of proteins can accelerate the removal of

Fig. 5 Frequency of AAG codon in different groups of proteins under short-term heat stress. **a** Distribution of AAG codon occurrence frequency for different groups of genes in yeast strain BY4742 subjected to short-term heat stress. The proteome data used in this analysis was published by the previous study [40]. Yeast was cultured at 30 °C was taken as control. Heat stress treatment was applied by transferring yeast to 37 °C for 1 h. Up and down-regulated proteins in response to heat treatment were identified with criteria "FDR < 0.05 and fold change >1.2 or < 0.83". The Y-axis represents the frequency of codon AAG in protein-coding sequence. P-values were calculated by the Kruskal-Wallis method. **b** Distribution of AAG codon occurrence frequency for different groups of genes in yeast strain BY4741 subjected to short-term heat stress. The proteome data used in this analysis was published by the previous study [41]. Yeast strain BY4741 cultured at 25 °C was taken as control. For heat treatment, yeast was exposed to 37 °C for 5, 10, 15, 30, 45 and 60 min. The criteria for up-regulated and down-regulated proteins are identified by adjusted *p* < 0.01 and fold change >1.2 or < 0.83. The numbers in brackets represent the heat treatment time

unnecessary proteins [16]. This adjustment by the organism can lead to quick adaptations in response to environmental disturbances.

Codon-based regulation of protein expression

From the regression results (Fig. 4), changes in codon use were more closely related to protein expression than transcription. Among codons, AAG (Lys) is the most remarkable, contributing 24 and 13% in the two thermal environments.

Involvement of the AAG codon has been shown in many studies. By the use of the model to predict the reaction of proteins after doubling the codon usage, AAG was shown to have the most significant effect on increased protein expression in human tissue [51]. All codons encoding lysine are replaced with AAA or AAG, and the protein expression before and after replacement is observed and compared under heat stress. It can be used to test whether AAG can be quickly translated under heat stress. The results of this experiment will be interesting.

The selection of AAG by heat-responsive proteins is also related to tRNA modification. m1A-modified tRNA has a higher affinity for the elongation factor EF1A (elongation factor 1-alpha), which delivers tRNA to the ribosome [52]. ALKBH1 (histone H2A dioxygenase) is

an RNA demethylase that mediates the removal of the methyl group from the N1-methyladenosine (m1A) in tRNA. tRNA$^{Lys}_{CUU}$, which complementarily pairs with AAG in the coding sequence, is one of the most important binders of ALKBH1 [53]. Therefore, tRNA$^{Lys}_{CUU}$ is widely modified by m1A to promote translation efficiency. Studies have shown that heat shock significantly increases m1A levels [54], so that when an organism is subjected to heat stress, a AAG-rich sequences are translated efficiently.

On the other hand, tRNA$^{Lys}_{UUU}$ needs to be modified with mcm^{5}s^{2}U^{34}, an important form of post-transcriptional modification of tRNA, to maintain efficient translation of AAA codons. Studies have indicated that this modification occurs at a low level under high temperature, leading to stagnant translation of AAA-rich sequences, while AAG does not require this modification [55–57]. In addition, the continuous codon AAA easily causes ribosome sliding and premature termination [58]. Therefore, the expression of AAA-rich sequences will slow down or stagnate at high temperatures and may not participate in the thermal response. Genes related to ALKBH1 and mcm^{5}s^{2}U^{34} are overexpressed and the expression of sequences is observed which rich in AAG and AAA. Thus, it can be judged whether tRNA

Wu *et al. BMC Genomics*   (2021) 22:263

Page 8 of 11

modification is involved in participating in determining which codon translates fastest under heat stress.

Above all, the codon AAG may be chosen as a means to quickly and easily identify whether mRNA is highly expressed in a thermal environment.

AAG-rich genes

We enriched the genes with a high frequency of AAG (frequency > 0.08) in the whole wheat genome (Fig. S2); nucleosomes, ribosomes, and molecular chaperones were significantly enriched.

Half of the ribosomes and the vast majority of nucleosomes have an abundance of AAG presence, while only 15% of molecular chaperones do. Ribosomes rich in AAG have been confirmed [59], and they usually have a shorter sequence and are in an active transcriptional state. This is conducive to the use of existing mRNA for efficient expression after environmental disturbance.

In addition to posttranscriptional regulation of molecular chaperone expression, transcriptional regulation also plays an important role. Unlike ribosomes, chaperones do not have an enormous presence and needs a large amount in a short time to help misfolded proteins. This explains the mystery that has been unresolved for a long time how heat shock proteins quickly respond to heat stress.

In conclusion, through a systematic analysis of the factors of changes in protein expression under heat stress, the related factors of protein expression and their influence have been described under short-term heat stress. High expression of housekeeping and heat-responsive genes may have solved evolutionarily the problem of rapid expression under heat stress by increasing the ratio of AAG.

Conclusions

By analyzing the transcriptome and proteome data under two kinds of heat stress, the factors were revealed which affect the rapid expression of proteins. Codon usage may play an important role in the rapid translation of proteins, especially for AAG. Moreover, the ability of transcriptional regulation changed according to the degree of heat stress. Transcription and post-transcriptional regulation worked synergistically to express the desired protein faster under heat stress. Our study revealed the main factors affecting the changes of protein expression in the short-term heat stress and explained the potential mechanism that heat-responsive protein expressed rapidly under heat stress.

Methods

Plant materials and growth conditions

Chinese Spring (*Triticum aestivum* L.) is thought to be a Sichuan variety. The wide application of this variety and its derived genetic stocks has greatly advanced wheat genetics, including the recent achievement of genome sequencing of wheat.

Wheat seeds (Chinese Spring) was grown in a greenhouse, and daily care was taken to avoid stress. The main stem of the plant was labeled when the first flower appeared on the spike. Twelve days after flowering, they were transferred to growth chambers with a temperature of 20 °C, and a 14/10 h day/night photoperiod for 3 days of adaptation; all the plants were in good condition. All the plants were divided into three groups (approximately 50 plants); the plants in two groups were quickly transferred to incubators that were preheated to 30 °C and 40 °C, and the plants in the other group remained at 20 °C as the control. The grains on the main stem were collected at the same time after exposure to three temperature environments for 1 h (in light), immediately frozen in liquid N_2, and stored at − 80 °C for transcriptomic and proteomic analysis.

RNA sequencing

Grains from three independent plants in each treatment were taken for mRNA sequencing. The RNAs of 9 samples were subjected to 150 bp paired-end sequencing using the Illumina HiSeq X Ten platform. The sequencing depth is 10X. Trimmomatic (version 0.36) [60] was used to remove the adapters and filter low-quality reads and bases from the next-generation sequencing (NGS) data. The RNA-seq reads were aligned to the wheat reference genome, IWGSC v1.0 [https://urgi.versailles.inra.fr/download/iwgsc/IWGSC_RefSeq_Assemblies/v1.0/]. FPKM (fragments per kilobase per million) were calculated using RSEM version 1.3.0 [61] and edgeR version 3.24.3 [62] were used to identify the differentially expressed transcripts. Scripts for processing NGS data have been uploaded to GitHub (https://github.com/shangguanhehe/NGS.git).

Isobaric tandem mass tag (TMT)-labeled quantitative proteomics

A total of 12 samples (four biological replicates each) under the three types of treatment were quantified via proteomics. After wheat grain proteins were extracted, they were digested in a solution with trypsin and labeled with a TMT isobaric mass tagging kit (Thermo Fisher Scientific). The method of protein extraction was based on Wang et al. [63]. The mixture was then physically separated by high-performance liquid chromatography (HPLC) and further analyzed for peptides using mass spectrometry (MS).

The resulting MS/MS spectra were processed using the MaxQuant search engine (version 1.5.2.8). Tandem mass spectra were searched against the wheat protein database (https://urgi.versailles.inra.fr/download/iwgsc/IWGSC_RefSeq_Annotations/v1.1/iwgsc_refseqv1.1_

genes_2017July06.zip) concatenated with the reverse decoy database. Trypsin/P was specified as a cleavage enzyme allowing up to 2 missings. The first search and main search range were set to 5 ppm, and 0.02 Da of fragment ions. Carbamidomethylation on Cys was specified as a fixed modification, and oxidation on Met, and oxidation on Met and acetylation on the protein N-terminus were specified as variable modifications. False discovery rate (FDR) thresholds for protein, peptide, and modification sites were specified at 1%.

Three or more identified proteins among the four biological replicates were considered as reliable quantitative data. The missing values were filled with averages of the other three and normalized.

Bioinformatic analysis
Hypergeometric distributions (phyper, R) were used for GO and KEGG enrichment analysis to test the significance of items; furthermore, the *p*-value was adjusted by the false discovery rate (FDR) to reduce the probability of false positives. The entries were chosen with a sufficiently large count ($N \geq 7$) and a suitable q-value (FDR < 0.05). Annotations of wheat genes were obtained from the URGI website (https://urgi.versailles.inra.fr/download/iwgsc/IWGSC_RefSeq_Annotations/v1.0/).

Factors affecting protein expression
Transcription and sequence characteristics were used to analyze the effects on protein expression under thermal conditions. Protein sequences, coding sequences (CDS), and annotation files were obtained from the website (https://urgi.versailles.inra.fr/download/iwgsc/IWGSC_RefSeq_Annotations/v1.0/), and untranslated region (UTR) were extracted from the genome annotation gff3 file. Sequence features such as the length of the sequences, base, codon, and amino acid usage frequency were obtained via python scripts (https://github.com/shangguanhehe/Systems-biology.git). Transcript and protein fold changes were obtained from transcriptome and proteome identification and the subsequent data analysis.

Multivariate adaptive regression splines (MARS)
Multivariate adaptive regression splines (MARS) was used to assess the individual and combined contribution of the selected features to protein fold changes. MARS is a statistical technique for modeling data and it is an extension of linear regression that captures nonlinearities and interactions between variables [64]. MARS analysis was implemented via the 'earth' package (version 5.1.2) on the R platform. To make the results more reliable, linear regression (IBM SPSS, version 22) and elastic net (R, glmnet 2.0–18) were used for verification.

Abbreviations
FPKM: Fragments per kilobase per million; DET: Differentially expressed transcript; DEP: Differentially expressed protein; MARS: Multivariate adaptive regression spline; TMT: Tandem mass tag; MS: Mass spectrometry; FDR: False discovery rate

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12864-021-07560-y.

Additional file 1: Figure S1. Changes in the transcription and protein levels of 141 ribosomes identified in two thermal environments. T30, T40, P30 and P40 represent transcriptional changes at 30 ℃, transcriptional changes at 40 ℃, protein changes at 30 ℃ and protein changes at 40 ℃, respectively.

Additional file 2: Figure S2. Function of codon-rich AAG genes. **a)** Among the whole-genome data of wheat, genes with an AAG frequency >0.08 in the coding sequence were subjected to enrichment analysis. **b)** GO enrichment of AAG-rich genes ($N > 60$, q < 1e-20).

Additional file 3: Table S1. Identification of DETs.

Additional file 4: Table S2. Identification of DEPs.

Additional file 5: Table S3. MARS regression results.

Additional file 6: Table S4. Linear and elastic net regression.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]State Key Laboratory of Crop Stress Biology for Arid Areas, College of Agronomy, Northwest A&F University, Yangling 712100, Shaanxi, China. [2]State Key Laboratory of Crop Stress Biology for Arid Areas, College of Plant Protection, Northwest A&F University, Yangling 712100, Shaanxi, China.

References
1. Lahtvee PJ, Sánchez BJ, Smialowska A, Kasvandik S, Elsemman IE, Gatto F, et al. Absolute quantification of protein and mRNA abundances demonstrate variability in gene-specific translation efficiency in yeast. Cell Systems. 2017;4(5):495–504. https://doi.org/10.1016/j.cels.2017.03.003.

2.   Chen W-H, van Noort V, Lluch-Senar M, Hennrich ML, H. Wodke JA, Yus E, Alibés A, Roma G, Mende DR, Pesavento C *et al*: Integration of multi-omics data of a genome-reduced bacterium: prevalence of post-transcriptional regulation and its correlation with protein abundances. Nucleic Acids Res 2016, 44(3):1192–1202, DOI: https://doi.org/10.1093/nar/gkw004.

3.   Budak H, Hussain B, Khan Z, Ozturk NZ, Ullah N. From genetics to functional genomics: improvement in drought signaling and tolerance in wheat. Front Plant Sci. 2015;6:1012.

4.   Vogel C, De Sousa AR, Ko D, Le SY, Shapiro BA, Burns SC, et al. Sequence signatures and mRNA concentration can explain two-thirds of protein abundance variation in a human cell line. Mol Syst Biol. 2010;6(400):1–9.

5.   Corbin RW, Paliy O, Yang F, Shabanowitz J, Platt M, Lyons CE, et al. Toward a protein profile of Escherichia coli: comparison to its transcription profile. Proc Natl Acad Sci U S A. 2003;100(16):9232–7. https://doi.org/10.1073/pnas.1533294100.

6.   Tian Q, Stepaniants SB, Mao M, Weng L, Feetham MC, Doyle MJ, et al. Integrated genomic and proteomic analyses of gene expression in mammalian cells. Mol Cell Proteomics. 2004;3(10):960–9. https://doi.org/10.1074/mcp.M400055-MCP200.

7.   Nie L, Wu G, Zhang W. Correlation between mRNA and protein abundance in Desulfovibrio vulgaris: a multiple regression to identify sources of variations. Biochem Biophys Res Commun. 2006;339(2):603–10. https://doi.org/10.1016/j.bbrc.2005.11.055.

8.   Baerenfaller K, Grossmann J, Grobei MA, Hull R, Hirsch-Hoffmann M, Yalovsky S, et al. Genome-scale proteomics reveals Arabidopsis thaliana gene models and proteome dynamics. Science. 2008;320(5878):938–41. https://doi.org/10.1126/science.1157956.

9.   Schrimpf SP, Weiss M, Reiter L, Ahrens CH, Jovanovic M, Malmström J, et al. Comparative functional analysis of the Caenorhabditis elegans and Drosophila melanogaster proteomes. PLoS Biol. 2009;7(3):e1000048. https://doi.org/10.1371/journal.pbio.1000048.

10. Casas-Vila N, Bluhm A, Sayols S, Dinges N, Dejung M, Altenhein T, et al. The developmental proteome of Drosophila melanogaster. Genome Res. 2017; 27(7):1273–85. https://doi.org/10.1101/gr.213694.116.

11. Lee MV, Topper SE, Hubler SL, Hose J, Wenger CD, Coon JJ, et al. A dynamic model of proteome changes reveals new roles for transcript alteration in yeast. Mol Syst Biol. 2011;7(514):1–12.

12. Lackner DH, Schmidt MW, Wu S, Wolf DA, Bähler J. Regulation of transcriptome, translation, and proteome in response to environmental stress in fission yeast. Genome Biol. 2012;13(4):R25. https://doi.org/10.1186/gb-2012-13-4-r25.

13. Mühlhofer M, Berchtold E, Stratil CG, Csaba G, Kunold E, Bach NC, et al. The heat shock response in yeast maintains protein homeostasis by chaperoning and replenishing proteins. Cell Rep. 2019;29(13):4593–607. https://doi.org/10.1016/j.celrep.2019.11.109.

14. Dhaliwal NK, Abatti LE, Mitchell JA. KLF4 protein stability regulated by interaction with pluripotency transcription factors overrides transcriptional control. Genes Dev. 2019;33(15–16):1069–82. https://doi.org/10.1101/gad.324319.119.

15. Jovanovic M, Rooney MS, Mertins P, Przybylski D, Chevrier N, Satija R, et al. Dynamic profiling of the protein life cycle in response to pathogens. Science. 2015;347(6226):1259038. https://doi.org/10.1126/science.1259038.

16. Liu Y, Beyer A, Aebersold R. On the dependency of cellular protein levels on mRNA abundance. Cell. 2016;165(3):535–50. https://doi.org/10.1016/j.cell.2016.03.014.

17. Crawford RA, Pavitt GD. Translational regulation in response to stress in Saccharomyces cerevisiae. Yeast. 2019;36(1):5–21. https://doi.org/10.1002/yea.3349.

18. Kozak M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell. 1986;44(2): 283–92. https://doi.org/10.1016/0092-8674(86)90762-2.

19. Lackner DH, Beilharz TH, Marguerat S, Mata J, Watt S, Schubert F, et al. A network of multiple regulatory layers shapes gene expression in fission yeast. Mol Cell. 2007;26(1):145–55. https://doi.org/10.1016/j.molcel.2007.03.002.

20. Calvo SE, Pagliarini DJ, Mootha VK. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proc Natl Acad Sci U S A. 2009;106(18):7507–12. https://doi.org/10.1073/pnas.0810916106.

21. Mayr C, Bartel DP. Widespread shortening of 3′UTRs by alternative cleavage and Polyadenylation activates oncogenes in Cancer cells. Cell. 2009;138(4): 673–84. https://doi.org/10.1016/j.cell.2009.06.016.

22. Coghlan A, Wolfe KH. Relationship of codon bias to mRNA concentration and protein length in Saccharomyces cerevisiae. Yeast. 2000;16(12):1131–45. https://doi.org/10.1002/1097-0061(20000915)16:12<1131::AID-YEA609>3.0.CO;2-F.

23. Courel M, Clément Y, Bossevain C, Foretek D, Vidal Cruchez O, Yi Z, et al. GC content shapes mRNA storage and decay in human cells. eLife. 2019;8: e49708. https://doi.org/10.7554/eLife.49708.

24. Litterman AJ, Kageyama R, Le Tonqueze O, Zhao W, Gagnon JD, Goodarzi H, et al. A massively parallel 3′ UTR reporter assay reveals relationships between nucleotide content, sequence conservation, and mRNA destabilization. Genome Res. 2019;29(6):896–906. https://doi.org/10.1101/gr.242552.118.

25. Mordstein C, Savisaar R, Young RS, Bazile J, Talmane L, Luft J, et al. Codon usage and splicing jointly influence mRNA localization. Cell Systems. 2020; 10(4):351–62. https://doi.org/10.1016/j.cels.2020.03.001.

26. Mauger DM, Cabral BJ, Presnyak V, Su SV, Reid DW, Goodman B, et al. mRNA structure regulates protein expression through changes in functional half-life. Proc Natl Acad Sci. 2019;116(48):24075–83. https://doi.org/10.1073/pnas.1908052116.

27. Horn D. Codon usage suggests that translational selection has a major impact on protein expression in trypanosomatids. BMC Genomics. 2008;9(1): 2–2. https://doi.org/10.1186/1471-2164-9-2.

28. Rudolph KLMM, Schmitt BM, Villar D, White RJ, Marioni JC, Kutter C, et al. Codon-driven translational efficiency is stable across diverse mammalian cell states. PLoS Genet. 2016;12(5):e1006024. https://doi.org/10.1371/journal.pgen.1006024.

29. Frumkin I, Lajoie MJ, Gregg CJ, Hornung G, Church GM, Pilpel Y. Codon usage of highly expressed genes affects proteome-wide translation efficiency. Proc Natl Acad Sci. 2018;115(21):E4940–9. https://doi.org/10.1073/pnas.1719375115.

30. Mittal P, Brindle J, Stephen J, Plotkin JB, Kudla G. Codon usage influences fitness through RNA toxicity. Proc Natl Acad Sci. 2018;115(34):8639–44. https://doi.org/10.1073/pnas.1810022115.

31. Ravichandran S, Ragupathy R, Edwards T, Domaratzki M, Cloutier S. MicroRNA-guided regulation of heat stress response in wheat. BMC Genomics. 2019;20(1):488. https://doi.org/10.1186/s12864-019-5799-6.

32. Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, et al. Global quantification of mammalian gene expression control. Nature. 2011; 473(7347):337–42. https://doi.org/10.1038/nature10098.

33. Semenov MA, Shewry PR. Modelling predicts that heat stress, not drought, will increase vulnerability of wheat in Europe. Sci Rep. 2011;1:1–5.

34. Lobell DB, Tebaldi C. Getting caught with our plants down: the risks of a global crop yield slowdown from climate trends in the next two decades. Environ Res Lett. 2014;9(7):074003. https://doi.org/10.1088/1748-9326/9/7/074003.

35. Tack J, Barkley A, Nalley LL. Effect of warming temperatures on US wheat yields. Proc Natl Acad Sci U S A. 2015;112(22):6931–6. https://doi.org/10.1073/pnas.1415181112.

36. Tao F, Zhang Z, Zhang S, Rötter RP. Heat stress impacts on wheat growth and yield were reduced in the Huang-Huai-Hai plain of China in the past three decades. Eur J Agron. 2015;71:44–52. https://doi.org/10.1016/j.eja.2015.08.003.

37. Farooq M, Bramley H, Palta JA, Siddique KHM. Heat stress in wheat during reproductive and grain-filling phases. Crit Rev Plant Sci. 2011;30(6):491–507. https://doi.org/10.1080/07352689.2011.615687.

38. Stone PJ, Nicolas ME. A survey of the effects of high temperature during grain filling on yield and quality of 75 wheat cultivars. Aust J Agric Res. 1995;46(3):475–92. https://doi.org/10.1071/AR9950475.

39. Zhu JK. Abiotic stress signaling and responses in plants. Cell. 2016;167(2): 313–24. https://doi.org/10.1016/j.cell.2016.08.029.

40. Jarnuczak AF, Albornoz MG, Eyers CE, Grant CM, Hubbard SJ. A quantitative and temporal map of proteostasis during heat shock in Saccharomyces cerevisiae. Mol Omics. 2018;14(1):37–52. https://doi.org/10.1039/C7MO00050B.

41. Storey AJ, Hardman RE, Byrum SD, Mackintosh SG, Edmondson RD, Wahls WP, et al. Accurate and sensitive quantitation of the dynamic heat shock proteome using tandem mass tags. J Proteome Res. 2020;19(3):1183–95. https://doi.org/10.1021/acs.jproteome.9b00704.

42. Lawless C, Holman SW, Brownridge P, Lanthaler K, Harman VM, Watkins R, et al. Direct and absolute quantification of over 1800 yeast proteins via selected reaction monitoring. Mol Cell Proteomics. 2016;15(4):1309–22. https://doi.org/10.1074/mcp.M115.054288.

Wu *et al. BMC Genomics*   (2021) 22:263

Page 11 of 11

43. Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J, et al. Deep proteome and transcriptome mapping of a human cancer cell line. Mol Syst Biol. 2011;7(1):548. https://doi.org/10.1038/msb.2011.81.

44. Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, et al. Correction: corrigendum: global quantification of mammalian gene expression control. Nature. 2013;495(7439):126–7. https://doi.org/10.1038/nature11848.

45. Komili S, Farny NG, Roth FP, Silver PA. Functional specificity among ribosomal proteins regulates gene expression. Cell. 2007;131(3):557–71. https://doi.org/10.1016/j.cell.2007.08.037.

46. Kondrashov N, Pusic A, Stumpf CR, Shimizu K, Hsieh AC, Xue S, et al. Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning. Cell. 2011;145(3):383–97. https://doi.org/10.1016/j.cell.2011.03.028.

47. Shi Z, Fujii K, Kovary KM, Genuth NR, Röst HL, Teruel MN, et al. Heterogeneous ribosomes preferentially translate distinct subpools of mRNAs genome-wide. Mol Cell. 2017;67(1):71–83. https://doi.org/10.1016/j.molcel.2017.05.021.

48. Gerst JE. Pimp my ribosome: ribosomal protein Paralogs specify translational control. Trends Genet. 2018;34(11):832–45. https://doi.org/10.1016/j.tig.2018.08.004.

49. Shaikho S, Dobson CC, Naing T, Samanfar B, Hajikarimloo M, Golshani A, et al. Elevated levels of ribosomal proteins eL36 and eL42 control expression of Hsp90 in rhabdomyosarcoma. Translation. 2016;4(2):e1244395. https://doi.org/10.1080/21690731.2016.1244395.

50. Venkataramanan KP, Min L, Hou S, Jones SW, Ralston MT, Lee KH, et al. Complex and extensive post-transcriptional regulation revealed by integrative proteomic and transcriptomic analysis of metabolite stress response in Clostridium acetobutylicum. Biotechnology for Biofuels. 2015;8(1):1–29.

51. Eraslan B, Wang D, Gusic M, Prokisch H, Hallström BM, Uhlén M, et al. Quantification and discovery of sequence determinants of protein-per-mRNA amount in 29 human tissues. Mol Syst Biol. 2019;15(2):e8513. https://doi.org/10.15252/msb.20188513.

52. Pan T. Modifications and functional genomics of human transfer RNA. Cell Res. 2018;28(4):395–404. https://doi.org/10.1038/s41422-018-0013-y.

53. Liu F, Clark W, Luo G, Wang X, Fu Y, Wei J, et al. ALKBH1-mediated tRNA Demethylation regulates translation. Cell. 2016;167(3):816–28. https://doi.org/10.1016/j.cell.2016.09.038.

54. Dominissini D, Nachtergaele S, Moshitch-Moshkovitz S, Peer E, Kol N, Ben-Haim MS, et al. The dynamic N1-methyladenosine methylome in eukaryotic messenger RNA. Nature. 2016;530(7591):441–6. https://doi.org/10.1038/nature16998.

55. Fernández-Vázquez J, Vargas-Pérez I, Sansó M, Buhne K, Carmona M, Paulo E, et al. Modification of tRNALysUUU by Elongator is essential for efficient translation of stress mRNAs. PLoS Genet. 2013;9(7):e1003647. https://doi.org/10.1371/journal.pgen.1003647.

56. Zinshteyn B, Gilbert WV. Loss of a conserved tRNA anticodon modification perturbs cellular signaling. PLoS Genet. 2013;9(8):e1003675. https://doi.org/10.1371/journal.pgen.1003675.

57. Damon JR, Pincus D, Ploegh HL. tRNA thiolation links translation to stress responses in Saccharomyces cerevisiae. Mol Biol Cell. 2014;26(2):270–82. https://doi.org/10.1091/mbc.E14-06-1145.

58. Koutmou KS, Schuller AP, Brunelle JL, Radhakrishnan A, Djuranovic S, Green R. Ribosomes slide on lysine-encoding homopolymeric a stretches. eLife. 2015;4:e05534. https://doi.org/10.7554/eLife.05534.

59. Ishii K, Washio T, Uechi T, Yoshihama M, Kenmochi N, Tomita M. Characteristics and clustering of human ribosomal protein genes. BMC Genomics. 2006;7(1):37. https://doi.org/10.1186/1471-2164-7-37.

60. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20. https://doi.org/10.1093/bioinformatics/btu170.

61. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011;12(1):323. https://doi.org/10.1186/1471-2105-12-323.

62. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2009;26(1):139–40. https://doi.org/10.1093/bioinformatics/btp616.

63. Wang X, Hou L, Lu Y, Wu B, Gong X, Liu M, et al. Metabolic adaptation of wheat grain contributes to a stable filling rate under heat stress. J Exp Bot. 2018;69(22):5531–45. https://doi.org/10.1093/jxb/ery303.

64. Friedman JH. Multivariate adaptive regression splines. Ann Stat. 1991;19(1):1–67.

65. Vizcaíno JA, Csordas A, Del-Toro N, Dianes JA, Griss J, Lavidas I, et al. 2016 update of the PRIDE database and its related tools. Nucleic Acids Res. 2016;44(D1):D447–56. https://doi.org/10.1093/nar/gkv1145.

Publisher's Note