Systematic analysis of AU-rich element expression in cancer reveals common functional clusters regulated by key RNA-binding proteins

Edward Hitti1*, Tala Bakheet1*, Norah Al-Souhibani1*, Walid Moghrabi1, Suhad Al-Yahya1, Maha Al-Ghamdi1, Maher Al-Saif1, Mohamed M. Shoukri2, András Lánczk3, Renaud Grépin5, Balázs Győrffy3,4, Gilles Pagès6, and Khalid S. A. Khabar1,#

1 Molecular BioMedicine Program and 2Department of Cell Biology, King Faisal Specialist Hospital & Research Centre, Riyadh, Saudi Arabia

3 MTA TTK Lendület Cancer Biomarker Research Group, Budapest, Hungary; 4 Dept. of Pediatrics, Semmelweis University, Budapest, Hungary

5 Centre Scientifique de Monaco Biomedical Department, Monaco, Principality of Monaco, 6 University of Nice, Institute for research on cancer and aging of Nice (IRCAN), Nice, France

* Joint first authors

# Corresponding Author:

Khalid S. A. Khabar, Ph.D.
Deputy Executive Director, Research Centre
Director, Molecular BioMedicine Program
Professor, Alfaisal University
King Faisal Specialist Hospital and Research Centre
(JCIA-Accredited Academic Medical Center)
P3354, MBC-03, Takhasusi Road
Riyadh 11211, Saudi Arabia
Phone: 966-11-442-7876, FAX: 966-11-442-4182
Email: khabar@kfshrc.edu.sa

Running Title: AU-rich elements-protein interactions in cancer

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

Acknowledgements
The research was supported by intramural funding to KSA (RAC-2150004) from the King Faisal Specialist Hospital & Research Centre.
ABSTRACT

Defects in AU-rich elements (ARE)-mediated post-transcriptional control can lead to several abnormal processes that underlie carcinogenesis. Here, we performed a systematic analysis of ARE-mRNA expression across multiple cancer types. First, the ARE database (ARED) was intersected with TCGA databases and others. A large set of ARE-mRNAs was over-represented in cancer and, unlike non-ARE-mRNAs, correlated with the reversed balance in the expression of the RNA-binding proteins: tristetraprolin (TTP, ZFP36) and HuR (ELAVL1). Serial statistical and functional enrichment clustering identified a cluster of 11 over-expressed ARE-mRNAs (CDC6, KIF11, PRC1, NEK2, NCAPG, CENPA, NUF2, KIF18A, CENPE, PBK, TOP2A) that negatively correlated with TTP/HuR mRNA ratios and involved in the mitotic cell cycle. This cluster was upregulated in a number of solid cancers. Experimentally, we demonstrated that the ARE-mRNA cluster is upregulated in a number of tumor breast cell lines when compared to non-invasive and normal-like breast cancer cells. RNA-IP demonstrated the association of the ARE-mRNAs with TTP and HuR. Experimental modulation of TTP or HuR expression led to changes in the mitosis ARE-mRNAs. Post-transcriptional reporter assays confirmed the functionality of AREs. Moreover, TTP augmented mitotic cell cycle arrest as demonstrated by flow cytometry and histone H3 phosphorylation. We found that poor breast-cancer patient survival was significantly associated with low TTP/HuR mRNA ratios and correlated with high levels of the mitotic ARE-mRNA signature. These results significantly broaden the role of AREs and their binding proteins in cancer, and demonstrate that TTP induces an anti-mitotic pathway that is diminished in cancer.

Keywords:
AU-rich elements/Breast cancer/ Cancer/Post-transcriptional control/ RNA-binding proteins
INTRODUCTION

AU-rich elements (AREs) act as key cis-acting factors in the post-transcriptional control of gene expression; abnormalities involving this pathway can occur in several diseases, including cancer (1). The AREs constitute instability determinant sequences located in the mRNA 3’UTR, and are fundamental to the transient expression of genes that regulate critical cellular functions, such as cell proliferation, apoptosis, cytokine response, and cellular motility (2,3). A growing body of evidence has linked aberrantly elevated and prolonged expression of ARE-encoding mRNAs to cancer, including those participating in angiogenesis, chemotaxis, and invasion. Examples of such genes include epidermal growth factor, estrogen receptor (ER), cyclooxygenase-2, vascular endothelial growth factor (VEGF), matrix metalloproteinase (MMP) 1, the chemokine receptor CXCR4, urokinase-type plasminogen activator (uPA) and its receptor (4-9). The ARE-mRNAs are regulated by trans-acting RNA-binding proteins notably tristetraprolin (TTP/ZFP36) and human antigen R (HuR/ELAVL1). TTP and HuR have antagonistic activities with regard to ARE-mRNA regulation; TTP destabilizes the mRNA and represses protein translation, whereas HuR is an mRNA-stability and translation-promoting factor (10-12).

The dysregulation of the expression and function of TTP and HuR has been of particular interest and significance as they affect several cancer processes including cellular growth (13,14). The expression of TTP and HuR is regulated at the levels of transcription, mRNA stability, and translation; and their activity can be regulated at the post-translational level by phosphorylation, cellular localization, and protein stability (10,13,15,16). The mRNAs of TTP and HuR contain AU-rich elements, and their proteins are capable of regulating their own expression in an ARE-mediated manner (17-20). Additionally, we identified a TTP-HuR axis in which TTP reduces HuR mRNA and translation (18,21). Furthermore, TTP mRNA levels can be reduced by the high cellular levels of the microRNA 29a in invasive breast cancer (22). The inhibition of miR-29a using a cell-permeable anti-miR-29a peptide-nucleic acid can normalize the TTP-HuR axis, which leads to a reduction in the levels of ARE-mRNAs in breast cancer cells (21).

Here, we performed a systematic analysis of ARE-mRNA expression in invasive breast cancer and across different types of cancer. We found an ARE-mRNA cluster that is: enriched in a mitotic cell cycle, highly co-regulated by key RNA-binding proteins, and correlated with patients’ clinical outcome. We experimentally validated the expression patterns, functional responses to TTP and HuR, and the physical association of these mRNAs with their cognate proteins in cancer cells. Focusing efforts on pathways that significantly impact cancer processes, such as mitosis, may lead to new therapeutic approaches.
Materials and methods

Cancer patient data
The Oncomine web-based data mining platform (www.oncomine.com) was used to mine data from The Cancer Genome Atlas (TCGA) and METABRIC databases: http://cancergenome.nih.gov (23). Gene expression levels for 389 and 1556 invasive ductal breast cancer samples were downloaded for TCGA and METABRIC datasets, respectively, along with the 61 and 144 corresponding matched normal samples. Cancer upregulated genes with 1.7-fold increase in expression and Q < 0.0001 were retrieved. Intersecting 2396 upregulated genes with 3658 genes from the ARE database (ARED) (http://brp.kfshrc.edu.sa/ARED/) allowed the identification of the over-expressed AU-rich element-containing genes. Log2 median-centered intensity ratios were used. Other cancer datasets were used and extracted through Oncomine and Nextbio portals (Supplemental Table 1).

Regression and volcano modeling for correlation studies
See supplemental methods.

Clustering and functional annotation
Normal mixture was used for clustering and obtaining ARE-mRNAs with the most significant negative correlation with the TTP/HuR ratio. The input to the model was the 414 over-expressed ARE-mRNA data from TCGA dataset, specifically: the correlation R-values of TTP mRNA with ARE-mRNAs, R-values of HuR mRNA with ARE-mRNAs, and the fold change of tumor to normal tissues (T/N). The Akaike Information Criterion (AIC) value was used as an indicator of goodness of fit (smallest AIC). The over-expressed ARE-mRNAs were analyzed using the AIC value and Johnson's transformation, which resulted in a best-fit normal mixture. Supervised hierarchical clustering and heat map visualization of gene expression data were performed using Gene-E (Broad Institute). Functional enrichment was performed using several web-based resources, including GENeSet analysis http://bioinfo.vanderbilt.edu/webgestalt/analysis.php, Database for Annotation, Visualization and Integrated Discovery (DAVID), and STRING protein–protein interaction resource http://string-db.org/.

Cell culture and proliferation
Cell lines were obtained from ATCC (2014 and 2015 batches). The normal-like mammary cell lines MCF10A and MCF12A were cultured in DMEM: F12 medium (Thermofisher) fortified with HuMEC supplement. The tumor breast lines MDA-MB-231, MCF-7, and HEK293 cells were grown in DMEM. The tumor breast cell lines BT-474 and SKBR were cultured in DMEM: F12 and McCoy’s; respectively. All media were supplemented with 10% FBS and antibiotics. Tet-On Advanced HEK293 cells were obtained from ClonTech (2013) and were cultured in DMEM supplemented with 10% Tet-System Approved FBS (ClonTech), 100 ug/mL G418 (Sigma), and antibiotics. Stocks were made at the original date of obtaining the cells, and were usually passaged for no more than four months. Cellular proliferation was assessed by using real-time label-free electric impedance system (ACEA, San Diego) as previously described (24).
ARE reporter constructs and transfection
The RPS30 promoter-linked reporter expression vectors (25) containing the AREs (Supplementary Table 4), were constructed by ligating two annealed synthetic complementary oligonucleotides with BamHI and XbaI overhangs. The annealed DNA were cloned in the 3’UTR the RPS30-SGFP vector. The use of RPS30 promoter along with SGFP, an EGFP generated from UW-reduced coding region (26) allows selective post-transcriptional effects. Co-transfection with RPS23-linked RFP was performed in HEK293 using lipofectamine 2000 (Invitorgen). The induced expression of HA-tagged TTP and myc-tagged HuR were used as previously outlined (4,5,21).

Tet-on TTP induction and effect on ARE-SGFP reporters
Tet-On TTP expressing constructs were prepared by PCR as previously described (27). The HA-TTP construct was originally obtained from Dr. Perry Blackshear, NIH. The HEK-Tet-On advanced cells to study the TTP effect were seeded in six replicates in 96-well clear-bottom black plates and co-transfected overnight using lipofectamine 2000 with 50 ng of RPS30-SGFP-ARE constructs along with 10 ng of the Tet-On-TTP PCR product. Fluorescence was measured 24 hr after treating the cells with 0.2 μg/mL of doxycycline using the BD Pathway 435 imager (BD Biosciences) as reported previously (27).

TTP and HuR over-expression
Transient transfection of TTP in the TTP-deficient MDA-MB-231 cells was performed using lipofectaminetne 2000. For HuR overexpression experiments, MCF10A cells were seeded in 6-well plates and transfected with 1 μg of pcDNA3.1-HuR plasmid or vector alone using Lipofectamine 2000. Total RNA and protein was extracted 24 hours after transfection for qPCR of the 11 identified mRNAs and immunoblotting.

RNA Interference
Two different small interfering RNA (siRNA) duplexes were synthesized by Metabion (Germany): HuR siRNA-1 (sense: 5’UGUGAAGAUAUCGCGAC3’ and antisense: 5’GUCGCGGAUCACUUUCACA3’) and HuR siRNA-2 (sense: 5’GCCUUGUCAGCAGCAUUGG3’ and antisense 5’CCAAUGCUGCUAACCAGGC3’) while the scrambled siRNA sequence is as follows (sense: 5’GCAUGUAACCGCGUUC3’ and antisense: 5’GAACCGGUACACAUUGGCC3’). MDA-MB-231 or HEK293 cells were transfected with 50 nM of the siRNA for 48 hr, RNA was extracted for RT-QPCR or fluorescence was quantitated, respectively. HuR silencing efficiency was verified by immunoblotting.

Quantitative real-time PCR
Total RNA was extracted (TRI reagent, Sigma) and reverse transcribed as described previously (5). QPCR was performed in multiplex using FAM-labeled Taqman primer and probe sets (Applied Biosystems, Foster City, CA, USA) as outlined in Supplemental Methods, and normalized to VIC-labeled human GAPDH as the endogenous control. Samples were amplified in triplicate in a CFX96 cycler (Bio-Rad), and quantification of relative expression was performed using the ΔΔCt method.
Immunoprecipitation of RNP complexes
HEK-293 cells (10 cc plates) were transfected with lipofectamine 2000 overnight with 2 µg pcDNA3.1 vectors expressing either HA-tagged-TTP or myc-tagged HuR along with either 100 ng of RPS30-SGFP or RPS30-SGFP-TNF-ARE as controls. TTP and HuR-associated RNA species were obtained using the detailed RNA-IP protocol outlined in Supplemental method.

Western Blotting:
Western blotting was performed as previously described (5). Specific antibodies to TTP, HuR, and serine phosphorylated H3 histone were obtained from Santa Cruz. Anti-HA and anti-β actin was obtained from Roche and Cell Signaling, respectively.

Mitosis assessment: flow cytometry and histone phosphorylation
MDA-MB-231 Tet-Off cells (28) were seeded with or without Tetracycline (1 µg/ml). After 24 hr Nocodazole was added to the cells to enrich for G2/M phase for 16 hr. Generally, we used 30 µg/ml of the drug as higher concentrations (e.g., 100 µg/ml) led to over-arrested G2/M and toxicity. Cells were trypsinized, treated with DNase (2U/ul, Ambion), washed with PBS, fixed with 70% ethanol. Then, cells were washed with PBS, treated with RNase (100 µg/ml) and simultaneously stained with propidium iodide (25 µg/ml) for 30 min at 37°C for cell cycle analysis. Proteins extracted from the cells immunoblotting and probed with rabbit anti-phosphorylated Histone H3 (Ser 10, Santa Cruz).

Survival analysis
The Kaplan–Meier survival analyses were performed using the Kaplan–Meier plotter portal (http://kmplot.com/analysis/), a comprehensive dataset for survival analysis that covers cross-normalized expression data of 54,675 genes in 4,142 breast cancer patients (29). The database was built from the gene expression and survival data extracted from the European Genome-Phenome Archive (EGA) and the Gene Expression Omnibus (GEO) repositories. Overall survival, recurrence-free survival,, and distant metastasis free survival were determined using gene cluster stratification. In addition, survival analysis in the METABRIC and TCGA RNA-seq databases was performed separately. Associations between gene expression and patient survival were assessed by the Kaplan–Meier method (log-rank test, Graphpad 6.0 (30). The percentile threshold algorithm (25) was used to determine the optimal cut-off of the ARE-cluster members. The JETSET best probe set was selected in case multiple probe sets measured the same gene to ensure the optimal probe set for each gene. Hazard ratios and p-values were determined by Cox proportional hazards regression.
Results

ARE-gene expression and TTP deficiency are over-represented in cancer

We performed a comprehensive analysis of ARE-mRNA expression across several cancer types (Fig. 1A). We intersected the AU-rich element mRNA database (ARED) with over-expressed genes (tumor/normal: \( >1.7, Q< 0.0001; \) two-sample t-test) extracted from different large cancer datasets (Supplementary Table 1). We found that the cancer ARE-mRNAs were significantly over-represented (13–23%) when compared to the overall cancer-over-expressed mRNAs (~7.2%), which makes an enrichment of 2- to 3-fold, \( p<0.0001; \) One sample one-sided t-test) (Fig. 1B).

The TTP mRNA is generally found deficient in different cancers, with a 3- to 6-fold reduction compared to normal tissues (Fig. 1C). It was most deficient in breast cancer; density plots clearly illustrate the dramatic differences in TTP levels in normal versus cancer tissues in two large datasets (Fig. 1D). Due to this particular reduction in TTP gene expression and the large size of data available, further systemic analyses of ARE-mRNA expression patterns and their relation to TTP and HuR levels were conducted initially in breast cancer. All over-expressed genes with at least 1.7-fold increase \( (p < 0.0001, \) two-sample t-test) were retrieved from TCGA-IDBC data and crossed with the ARED (2). A total of 414 over-expressed ARE-mRNAs were derived (Supplemental Table 2).

Correlation modeling of TTP/HuR axis and alternative polyadenylation of HuR

We looked at the relationship between HuR and the expressed levels of the cancer ARE-mRNAs. HuR has three polyadenylation variants as previously reported by us and others (18-20). Therefore, we examined the probe sets (Affymetrix) in the TCGA breast cancer dataset, and found there were two different sets for HuR (ELAVL1) mRNA, one that targets the most abundant variant (2.7 kb) and another that targets the long 6 kb variant (Fig. 1E). The probe set that detects abundant HuR variants exhibits over-expression of HuR in breast cancer. In contrast, the probe set that specifically targets the distal (6 kb mRNA) 3'UTR appears to be under-expressed compared to normal tissues (Fig. 1F). In METABRIC dataset (23), there is only one probe (Illumina), which is directed to the distal 3’UTR and thus the 6 kb variant, and, indeed, gave under-expressed levels of HuR in tumor tissues compared to normal tissues (Fig. 1F). This observation is also found with other types of cancer; particularly, brain and lung cancer (Supplementary Fig. 1). Thus, we focused on the probe set that targets the most abundant form of HuR transcript variants, which also is over-expressed in cancers.

Next, we performed expression/correlation analysis by integrating both the levels of TTP and HuR mRNAs. The TTP/HuR mRNA ratios were computed since the balance of these RNA-binding proteins can influence the expression of certain ARE-mRNAs (21). We tested the correlation of the TTP/HuR mRNA ratio with the expression patterns of the 414 ARE-mRNAs using the fit model; the result was significant, \( R^2 = 0.98 \) \( (p < 0.001; \) Pearson’s correlation with Fisher’s test of significance Fig. 1G). Moreover, the volcano plot shows that a large proportion (47%) of ARE-mRNAs have a negative correlation with TTP/HuR mRNA level ratio (198 ARE-mRNAs with R-
values < −0.3 and p-values < 0.0001−(-Log_{10} > 4) compared to non-ARE mRNAs (only 2.8% of the 1979 non-ARE-mRNAs, Fig. 1H). Supplementary Table 3 shows the identities of the ARE genes, R-values and P-values that correlate with TTP, HuR, or TTP/HuR mRNAs. In general, TTP levels correlated negatively with ARE-mRNAs, while HuR positively correlated with the ARE-mRNAs when compared to non-ARE-mRNAs (Supplemental Fig. 2). Overall, this data clearly shows that it is much more likely for over-expressed ARE-genes to negatively correlate with low TTP/HuR ratio than over-expressed non-ARE genes.

**Step-wise functional clustering analyses for cancer ARE-mRNAs**

To derive further correlation and functional clustering of the cancer ARE-mRNAs that depend on TTP/HuR mRNA ratio, the Normal Mixture model was used to find the most significant cluster of genes (Fig. 2A). Although we have used stringent criteria on the expense of sensitivity, several known TTP/HuR ARE-mRNAs such as uPA (PLAU), PLAUR, and MMP13 were found (Supplementary Table 2). Five clusters were generated, and the total squared correlation among the three parameters was highest (R=0.81) in cluster five, which contained a 23-ARE-mRNA cluster (Fig. 2B). The expression levels of this cluster in the breast tumors and the normal tissues from the TCGA dataset are illustrated on a heat map demonstrating significant cancer upregulation for each cluster member (Fig. 2C).

Directed acyclic graphs (DAG) enriched in Gene Ontology (GO) categories, each as a node in a network, were built through GEneSeT analysis (Fig. 2D). The mitotic (M) cell cycle was the most significant node and comprises 11 genes, which are CDC6, CENPA, CENPE, KIF11, KIF18A, NCAPG, NEK2, NUF2, PBK, PRC1, and TOP2A. An enrichment (7-fold) of this category was observed when this ARE-mRNA cluster is compared to all TCGA over-expressed genes (p < 0.0001; Chi-square test). The protein interaction module, which is based on DAGs, reveals that the highest enrichment (p=6.70e−11) are in DNA replication-independent nucleosome assembly and microtubule motor activity; both are within the M-phase of cell cycle category (Fig. 2E). The use of STRING analysis confirmed the significant protein interactions among the gene nodes (Fig. 2E, large inset).

**Bioinformatics and gene expression of M cell cycle ARE-mRNAs**

The weighted means of the different ARE-clusters between tumor and normal tissue ratios were calculated (Fig. 3A). As a whole, the 414 over-expressed ARE-mRNAs were slightly, although statistically significantly (p < 0.001) higher than non-ARE over-expressed control set genes (n = 1979). However, the TTP/HuR-correlative 23-ARE-mRNA cluster weighted mean has significantly higher expression ratios (~5-fold) compared to both the total over-expressed ARE- and non-ARE-mRNAs (Fig. 3A). The 11 M cell cycle ARE-mRNAs were significantly higher than other gene sets (6-fold of the weighted mean, p < 0.001, two-sample t-test; Fig. 3A). The data of individual expression levels and correlations of the 11 ARE-mRNA cluster demonstrate significant upregulation and negative correlation with the TTP/HuR mRNA ratio (Supplementary Fig. 3 & 4).
Further, we looked at the expression patterns of the ARE-mRNA cluster across 10 different types of cancers. The expression of 11 ARE-mRNAs was upregulated in a number of different cancers as shown in the heat map (Fig. 3B). The up-regulation of the mean expression (non-weighted mean) of the M ARE-mRNA cluster is also obvious across the different cancers (Fig. 3C).

By RT-qPCR, we evaluated the expression levels of the M cell cycle ARE-mRNAs in a number of cell lines (Fig. 3D and E). Overall, the mRNA levels were higher in the triple-negative MDA-MB-231 cells compared to moderately-tumorigenic ER-positive, MCF-7, and normal-like breast cancer cells MCF10A and MCF-12A, respectively (Fig. 3D). Using different panel of cells, HER-positive/ER-negative status, we also found that all the ARE-mRNAs (except CENPA, which had very low levels) were highly expressed when compared to the normal-like cells (Fig. 3E). In particular, the mRNA levels in HER-positive ER-/PR-negative SKBR3 was the highest when compared to other lines in the panel (Fig. 3E). In general, these results are in agreement with the observations made in the patient tissues' datasets in which these ARE-mRNAs are over-expressed in invasive ductal breast cancer.

**Anti-mitotic effect of TTP**

Based on the computational approach, TTP could be functionally related to M-phase of the cell cycle. Thus, we have utilized TetO/TTP-regulated MDA-MB-231 cell line (28); with removal of tetracycline, the cells express TTP and exhibit reduced proliferation (Fig. 4A). When cells were treated with nocodazole, TTP behaves as an anti-mitotic protein and blocked cell cycle at G2/M phase as determined by flow cytometry (Fig. 4B). Furthermore, we showed that histone H3 phosphorylation, a marker of mitotic arrest, is higher when TTP is expressed in the nocodazole-treated cells (Fig. 4C). A confirmatory experiment was performed using transient expression of TTP in MDA-MB-231 cells (Fig. 4D).

**Reporter assessment of ARE activity and response to TTP induction**

The putative AREs of the 11-gene cluster and their context sequences were cloned in the 3’ UTR of fluorescent reporter assays (Fig. 5A). A heat map was constructed to reflect the strength of the TTP target sequence UUAUUUAUU/UAUUUAUU (Fig. 5A), which shows that NEK2, CENPE, and NCAPG AREs have Class-II- like ARE while the rest are largely Class-I AREs; that is, AUUUA in a U-rich context (Fig. 5A and supplementary Table 4). We used HEK293 cell line, which is suitable for weak fluorescence-based reporter assays. Seven of the AREs caused a reduction in the reporter activity demonstrating functional and active AREs (Fig. 5B). Statistically significant reductions ranged from 25% (CDC6) to 50% (CENPE, NEK2), which share the strongest TTP target sequence score (Fig. 5B). The IL-8 ARE caused 52% reduction in reporter expression. The mRNAs for CENPA, PRC1, and TOP2A, although not statistically significant, they demonstrated a modest 25% reduction in reporter fluorescence compared to the non-ARE control (Fig. 5B). Also, most responded to TTP induction or HuR knockdown to varying degrees of reduction (Fig. 5C and D).
Physical association and response of the M ARE-mRNAs to TTP and HuR

We investigated the possible physical association of TTP with the 11 M cell cycle ARE-mRNAs by over-expressing an HA-tagged TTP (HA-TTP) construct in HEK-293 cells. The immunoprecipitation of TTP was confirmed by western blotting (Fig. 6A). The RNA-IP clearly shows that the SGFP-TNF-ARE reporter mRNA is highly enriched on the HA-coupled beads compared to the myc-coupled beads (> 30-fold enrichment). The ARE-containing known TTP target, IL-8 mRNA, is highly enriched in the TTP-HA IP sample (~23-fold) but not mRNAs of non-ARE housekeeping genes (Fig. 6A). Next, the association of 10 of the ARE-mRNAs was investigated (CENPA expression was below detection levels) by RT-QPCR and all were found to be enriched with TTP (Fig. 6A). The enrichment degree varied between 6-fold (CENPE) and 20-fold (PBK). As independent experiments, TTP-associated ARE-mRNAs were assessed in MDA-MB-231 using an alternative RNA-IP protocol (21) and further confirmed that TTP bound the ARE-mRNAs except CENPA which its levels were too low (Fig. 6B). Exogenous expression of TTP led to reduction of the majority of the 11 ARE-mRNAs (except NCAPG) (Fig. 6C).

The binding of HuR to the 11 ARE-mRNAs were assessed in a similar manner using transfected anti-myc tagged HuR. (Fig. 6D). The immunoblot shows the specificity of the antibody (Fig. 6D inset). Although the bindings to the different 11 ARE-mRNAs were modest compared to TTP, it is within the same degree as with the HuR mRNA target of IL-8 and is statistically significant (Fig. 6D). NEK2 mRNA has the strongest binding (3-fold) while TOP2A and PRC1 mRNAs were the weakest. These results demonstrate the physical association of the RNA-binding proteins TTP and HuR with the ARE-mRNA mitotic cluster.

We have also assessed the outcome of HuR knockdown in MDA-MB-231 (Fig. 6E), which has high levels of HuR (21), on the expression of the 11 ARE-mRNAs, and found that siRNA-mediated silencing caused significant reduction in the mRNA levels (Fig. 6F). Using the non-transformed MCF10A which generally has low levels of HuR (21), we found that over-expressed HuR was able to augment the M cell cycle ARE-mRNA levels (Fig. 6G).

Association of M cell cycle ARE cluster and cancer patients’ survival

Patient classification to low and high mean expression of the M cell cycle ARE-mRNA cluster (Fig. 7A) inversely corresponds as expected with the TTP/HuR mRNA ratio (Fig. 7B). Poor relapse-free survival (RFS, n = 4073 patient data) was observed with the high M cluster cohort (p <10e-10, log-Rank test; and hazard ratio (HR) = 2.4) (Fig. 7C). Similarly, poor outcome was observed in distant metastasis-free survival (DMFS, HR = 2.4) (Fig. 7D) and in overall survival (OS, HR = 2.13, Supplemental Fig. 5A). We repeated the survival analysis in the non-Affymetrix METABRIC database (n = 1959) and the RNA-seq data (n=781) and data demonstrate poor survival with higher mean expression of the mitotic ARE-mRNA cluster (Supplemental Figure 5 B and C).

Testing direct correlation of patients’ survival with TTP/HuR mRNA ratio stratification, we found that the cohort with the low TTP/HuR mRNA ratio had poor survival outcome (Fig. 7E). Poor survival (higher HR) in relation to the low TTP/HuR mRNA ratio, is more prominent with ER-positive cases and negative lymph node involvement (Fig. 7F). The 11 M ARE-mRNA cluster was
also more significantly associated with poor survival in case of ER-positive and HER-negative tissues (Fig. 7G). It appears that the TTP/HuR mRNA ratio and the ARE-mRNA cluster is more associated with survival in patients with lower grade tumors (Fig. 7F). Overall, high expression of the mitotic ARE cluster and their association with a low TTP/HuR mRNA ratio is indicative of poor survival in breast cancer patients.
Discussion

Alterations in ARE-mediated pathways are prominent post-transcriptional events that are increasingly studied in cancer. Only a small set of ARE-mRNAs have been previously shown to be over-expressed in cancer and regulated by the key RNA-binding proteins, TTP and HuR (For an updated list, see our review (1). In this study, we expanded cancer ARE-mRNAs to nearly 200 ARE-mRNAs that correlate negatively with TTP and positively with HuR expression. Further statistical and functional enrichment, and experimental validation led to the identification of 11 ARE-mRNAs, which code for proteins involved in the mitotic (M)-phase of the cell cycle. The 11 mitosis members perform critical roles in M phase of the cell cycle (31). Thus, TTP may promote G2/M arrest by the fine-tuning the temporal expression of these ARE-mRNAs.

The systemic study conducted here shows that ARE-mRNA expression is over-represented in multiple cancer types adding to the significance of the ARE-mediated post-transcriptional control in cancer. In these same types of tumors, TTP deficiency is observed and is linked to nearly half (198 genes) of the over-expressed ARE-mRNAs in breast cancer as assessed here. A deficiency of TTP can lead to the increased growth of cancer cells, because TTP targets those mRNAs involved in cellular proliferation, such as the previously identified cyclin D1 myc, E2F1 and IL-6 (32,33) and also the M cell cycle ARE-mRNAs newly identified here. Modulation of TTP in breast and other cancer cells by ectopic expression impacts cellular growth and attenuates tumor xenografts in mice (28,34,35) which in agreement with the patients’ data here.

While TTP is deficient in cancer, HuR, which stabilizes ARE-mRNAs is often over-expressed in tumors or its cytoplasmic localization is increased (36-41). HuR is more of a general RNA-stability promoting protein with preference towards AU/U-rich sequences (42). Here, we also expanded (more than 100 mRNAs) HuR-correlated ARE-mRNAs that are linked to cancer. HuR mRNA has three alternative polyadenylation forms of 1.5 kb, 6 kb long variant, and the most abundant form is 2.7 kb transcript (18). HuR autoregulates its own mRNA at the post-transcriptional level (18) (19,20). The lower expression of the long variant in contrast to the over-expression of the shorter (and abundant) mRNA variant in cancer can be explained by defects in the polyadenylation process, which can favor the abundant variant. This may be consistent with the cell line work of two groups (19,20). In general, HuR has been shown to block polyadenylation and cleavage at polyadenylation sites that are flanked by U-rich sequences (43). Thus, our data notably extend these laboratory HuR transcript expression patterns to patients’ tissues. Furthermore, this brings attention to the fact that different probe sets directed towards different regions of the same transcript, including polyadenylation variants, can have an impact on the interpretation of microarray expression results, including the expression direction.

We focused analysis on TTP/HuR mRNA aberrations and correlation with the cancer genes because a TTP-HuR axis exists in cells: TTP competes for HuR binding and can reduce HuR mRNA expression (18,21). Further mechanistic details showed that phosphorylation of TTP by MK2 alters that balance towards ARE-mRNA stabilization and translation (11,44), and more recent work confirmed that TTP targets the HuR mRNA (45). Another evidence for the TTP-HuR axis is
the ability of the miR-29a which targets TTP 3'UTR in breast cancer cells (22) to alter the normal TTP-HuR balance, while miR-29a inhibitor restores the reversed balance. This leads to reduction of the cellular levels of ARE-mRNAs and their proteins; for example, uPA and CXCR4 in cancer cells (5,21).

The M cell cycle ARE-mRNAs either as individuals or as a group (weighted mean) were not only over-expressed in the invasive breast tumor tissues but were over-expressed across a number of solid tumors, including liver, lung, bladder, brain, stomach, and kidney. This implies that there is a wide-spread effect of dysregulated ARE-mediated post-transcriptional processes in controlling M cell cycle in cancer. Other findings do not rule out other regulatory mechanisms that are abnormal in cancer such as transcription, since these genes are tightly co-expressed.

Another potential and important ramification of TTP deficiency and HuR over-expression is the impact on angiogenesis, tumor invasion, and metastasis as it targets ARE-mRNAs that mediate such processes, including IL-8, VEGF, IL-6, uPA, uPAR, CXCR4, and MMP13 (4-6,46,47). TTP has been shown to significantly reduce tumor cell migration and invasion while HuR increases these activities (4,5,21,48,49). The combined impact on mitosis and tumor invasion due to lack of TTP in breast cancer cells and patients may confirm the putative tumor suppressor role of TTP, including its anti-invasive role. The strikingly low number of patient samples with normal TTP and TTP/HuR levels allows the suggestion that the repression of TTP might be an "essential factor" at least in invasive breast cancer. Those who have levels that are not dramatically reduced had better prognostic outcomes. There was significant stratification of survival based on TTP/HuR mRNA ratios and the M-phase cell cycle ARE-mRNA clusters in breast cancer patients, particularly with low-grade, lymph-node negative and ER-positive status. Thus, targeting the TTP-HuR axis, and thereby reducing the M cell cycle and invasiveness of the cancer cells, either by increasing TTP or directly by reducing HuR, may have beneficial effects in patients. Our results can give more insights to the role of the post-transcriptional network in cancer maintenance with new therapeutic approaches in mind.
REFERENCES:


22. Gebeshuber CA, Zatloukal K, Martinez J. miR-29a suppresses tristetraprolin, which is a regulator of epithelial polarity and metastasis. EMBO reports 2009; 10:400-5.


Figure legends

**Figure 1: Systemic analysis of ARE-mRNAs and TTP expression in cancer.** A) Schematic representation illustrating the approach used for deriving a list of over-expressed ARE-mRNAs in cancer types. Gene expression levels were mined from 20 datasets representing 10 cancer types (Datasets, Supplemental Table 1). The threshold for over-expression was set as 1.7-fold of tumor to normal (T/N) and Q < 0.0001. Over-expressed ARE-mRNAs for each dataset were determined by intersecting total over-expressed genes with the ARED database. B) Percentage of ARE-mRNAs that are over-expressed in the different cancer types in proportion to the total expressed mRNA in each cancer. C) T/N level Fold of TTP mRNAs were obtained from each cancer dataset. D) Density frequency plot of TTP mRNA levels in normal (green) and tumor tissues (brown) using the indicated dataset. E) Schematic representation of the three different HuR polyadenylation mRNA variants. The relative probe target locations (arrows) are shown (Affymetrix probe starts with A_ and Illumina with ILMN). pA: poly A signal. Boxes (between vertical lines) denote exons. F) Tumor/Normal levels of HuR mRNA by probe ID. G) Scatter plot using fit model for the correlation between TTP/HuR mRNA ratios and the cancer 414 ARE-mRNAs. H) Volcano plots for TTP/HuR mRNA ratio comparative correlation with ARE- and non-ARE cancer mRNAs by significance (−log_{10} p-value) as described in Materials and Methods. Pie plots depict the percentages of ARE-mRNAs (upper) or non-ARE-mRNAs (lower) that according to the correlation. Vertical dotted lines denote: negative correlation (R<-0.3, p<0.001), or positive correlation (R>0.3) with TTP expression levels. *Pearson’s correlation with Fisher’s exact test for significance (p-value) was used.*

**Figure 2: Clustering analysis of cancer ARE-mRNAs with TTP and HuR expression.** A) Normal mixture clustering analysis based on the correlation (R) of the T/N expression ratios (Fold) of the 414 over-expressed breast cancer ARE-mRNAs with that of TTP and HuR. B) Heat map visualizing the most significant cluster (Cluster Five) containing 23 ARE-mRNA showing fold (T/N mRNA ratio) and ARE-mRNA correlations (R-values) with HuR and TTP mRNA levels. Legend shows the scale for each parameter. C) Heat map visualizing the over-expression of the resultant 23 ARE-mRNA cluster in the invasive ductal breast cancer. D) Functional annotation of the 23 genes using GeneSeT analysis: enriched GO categories are in red (multiple tests adjusted P-value) while those in black are their non-enriched parents. Pie charts showed the most enriched GO categories using DAVID. E) Modules of enriched protein interactions (in red lines) on directed acyclic graphs (DAGs). Green nodes indicate the mitotic proteins. The large inset shows the STRING analysis of known and predicted protein–protein interactions among the 11 ARE-mRNA gene products. The nodes indicate proteins, and the edges indicate interaction between these genes.

**Figure 3. Expression of the ARE-mRNA mitotic cluster in cancer.** A) The weighted means of T/N ratio for the indicated ARE-mRNA clusters that correlated with TTP/HuR. Level 2 raw data was obtained from TCGA invasive ductal breast cancer dataset. B) A heat map was generated to visualize the over-expression of the mitotic cell cycle ARE-mRNAs in 10 different cancer types. C) The mean expression (non-weighted mean) of the combined fold increase of the 11 M ARE-
mRNAs was derived from the Nextbio portal using several cancer types. No SEM data are given from the Nextbio output, but P-values are reported (numbers are shown). D) The normal MCF-10A and MCF-12A, as well as the breast tumor cell lines triple-negative MDA-MB-231 and ER-positive/Her-negative, MCF-7, were grown in six-well culture plates. Total RNA was extracted, and the mRNAs were quantitated by RT-QPCR using TaqMan specific primers and were normalized to the housekeeping gene, RPLPO. E) The mRNA levels of the 11 ARE-mRNA cluster in MCF10A and the HER2 positive cell lines BT474 and SK-BR3 using RT-QPCR as above. Data are from one experiment representative of at least two independent experiments. In both D and E, statistical significance was assessed by one-way ANOVA, with underlined values of *p< 0.05, **P<0.01, *** P<0.001. Post-Hoc tests, Dunnett’test for comparison with MCF10A or Tukey’s (with each normal-tumor pair) were used, *p < 0.01, **p < 0.0001.

Figure 4. The anti-mitotic effect of TTP
(A) Continuous monitoring of cell proliferation of MDA-MB-231 Tet-off cells in the presence or absence of tetracycline (1 μg/ml). (B) After 24 hr of tetracycline treatment, nocodazole (30 μg/ml) was added to the cells for 16 hr to enrich for G2/M-positive cells. Data are representative of three flow cytometry experiments. (C) MDA-MB-231 Tet-Off cells were treated with tetracycline and nocodazole as indicated in B and protein was extracted for immunoblotting with anti-p-Histone H3 or β-actin. The experiment is representative of two independent experiments. (D) Transient expression of TTP and histone H3 phosphorylation. MDA cells were transfected with TTP or vector, followed by nocadzole treatment (16 hr). Western blotting showed anti-phosphorylated H3 and actin control.

Figure 5. Reporter assessment of ARE activity and response to TTP induction. A) A scheme representing the AREs of the M cycle ARE-mRNAs along with the cluster membership according to the ARED. The color score represents strength of TTP sequence in which 1 = perfect match with UUAUUUAUUU. B) HEK293 cells were transfected with the RPS30-SGFP-ARE reporters along with control vector (non-ARE) and positive control (IL-8 ARE) as described in Materials and Methods. Fluorescence was measured after 24 hr and presented as % Mean of control ± SEM. C) TTP effect on ARE-reporter expression. A Tet-O inducible TTP expression cassette was co-transfected with each of the reporter constructs in HEK-Tet-ON cells. TTP was induced by 0.2 μg/mL doxycycline treatment, and the effect of TTP was assessed 24 hr later by quantifying the level of the reporter fluorescence. D) HuR knockdown and ARE reporter activity. HEK293 cells were treated with 50 nM control siRNA or HuR siRNA for 24hr and transfected as above. Fluorescence was quantified after another 24 hr. *p<0.05, **p < 0.005, ***p < 0.001, ****p < 0.0001 using Student’s t-test.

Figure 6. RNA/protein immunoprecipitation and modulation of TTP and HuR A) RNA/Protein co-immunoprecipitation of TTP with the putative M ARE-mRNA cluster. HA-TTP HEK293 transfected cells were lysed and TTP protein was immunoprecipitated with anti-HA coupled beads. The negative control (NC) IP was performed with anti-myc. TTP-IP was confirmed by western blotting (insets). TTP-associated mRNA levels were assessed by RT-QPCR and normalized to RPLPO as background control. The non-ARE mRNAs of GAPDH, GUS, and RPLP1
were used as internal negative controls while IL-8 ARE-mRNA as control. B) MDA-MB-231 cells were transfected with TTP expression plasmid for 24 hr. Cells were lysed and TTP protein was immunoprecipitated using anti-TTP antibody or normal IgG control antibody. Associated mRNA was quantified by RT-QPCR using gene-specific probes normalized to GAPDH. C) TTP-mediated downregulation of M ARE-mRNA in MDA-MB-231 cells. Cells were transfected with 10 ng TTP-pcDNA3.1 plasmid and 500 ng DNA with pcDNA3.1 empty vector or transfected with 500 ng empty vector alone. Total RNA was extracted after 24 hr and RT-QPCR of 11-genes plus positive and negative controls were performed as in (A). D) The binding of myc-HuR to the 11 M ARE-mRNAs was assessed in a similar manner to A. The myc-HuR IP was performed on anti-myc coupled beads, and the negative control IP with anti-HA coupled beads. The same positive and negative control mRNAs were used as in A). (E, F) HuR silencing and the M ARE-mRNA cluster in MDA-MB-231. The cells (0.5 x 10^6) were transfected with 50 nM HuR siRNA-1, HuR siRNA-2 or scrambled siRNA control for 48 hr. HuR silencing efficiency was determined by Western blotting. The mRNA levels were determined by RT-QPCR using HuR siRNA-2 samples. Data are from one representative experiment of at least two independent experiments. G) Effect of HuR over-expression on the M ARE-mRNA levels in MCF10A. Cells (0.5 X 10^6 cells) were transfected with 1 ug of pcDNA3.1-HuR plasmid or vector alone and RNA was extracted after 24 hr for RT-QPCR. The results are from one experiment representative of two independent experiments. For all data above: * P <0.05, ** P <0.005, *** P < 0.0001 (Student’ t-test) shown for each ARE mRNA in comparison to the control.

Figure 7. Clinical outcome of the M cell cycle TTP/HuR correlated ARE-mRNA expression
A) A heatmap for M ARE-mRNA expression cohorts. Data in Kaplan–Meier portal was used to classify the data into low and high expressing M ARE-mRNA cohorts in relation to the clinical outcome. The dotted line represents the best cutoff. B) TTP/HuR mRNA ratio levels of the two expression cohorts. The best cut-off was based on the expression levels of the ARE-mRNA cluster shown in (A). C) Kaplan–Meier survival curves generated in accordance with expression classifiers in (A) using RFS: relapse-free survival (C) or DFMS, distant metastasis free survival (D). E) Kaplan–Meier survival curve using TTP/HuR mRNA classifiers. Grade, lymph-node involvement, and receptor status in patients with TTP/HuR mRNA classification (F) and M cell cycle ARE-mRNA cluster (G). P values (numbers) and hazard ratio (HR) are indicated. Q square test was used to evaluate the statistical significance between each group member, *P<0.01, **<0.001, ***p<0.0001.
A

Cell index vs. Hours

-TTP

+TTP

B

-TTP + Nocodazole

+TTP + Nocodazole

Cell Count

0 200 400 600 800 1000

G0/G1 18.33%
S 21.21%
G2/M 58.49%

G0/G1 12.08%
S 11.06%
G2/M 75.87%

C

+ Nocodazole

38
17
12
38

Ratio 1 1.5

β-actin

D

-TTP + TTP -TTP + TTP

Cell Count

0 200 400 600 800 1000

G0/G1 65.77%
S 18.61%
G2/M 15.23%

G0/G1 66.75%
S 12.95%
G2/M 19.83%

Ratio 0.4 0.9

β-actin

TTP

phosph-3H
A

<table>
<thead>
<tr>
<th>Gene</th>
<th>ARE Target</th>
<th>TTP Cluster</th>
<th>Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDC6</td>
<td>ATATATATAT</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>CENPA</td>
<td>TAAATACATT</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>CENPE</td>
<td>TTATATTTTG</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>KIF11</td>
<td>TATAATATTAT</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>KIF15A</td>
<td>ATATATAAAAAT</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>NCAPG</td>
<td>ATATATTACT</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>NEK2</td>
<td>CATTATTATT</td>
<td>III</td>
<td></td>
</tr>
<tr>
<td>NUF2</td>
<td>AATATATATA</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>PBK</td>
<td>TAAATATATA</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>PRC1</td>
<td>TATATATTCT</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>TOP2A</td>
<td>AAATATTTAT</td>
<td>V</td>
<td></td>
</tr>
</tbody>
</table>

0.7 1

B

RPS30 Prom  SGFP  BGH 3’UTR

RPS30 Prom  SGFP  ARE BGH 3’UTR

C

RPS30 Prom  SGFP  ARE BGH 3’UTR

TeTO3-Min Prom  TTP  3’UTR

+ Doxycycline (TTP-On)

D

sRNA HuR/ctrl siRNA

Non-ARE  NEK2  CDC6  CENPA  KIF11  KIF18A  NUF2  PBK  PRC1  TOP2A  NCA2

Fluorescence levels, % of control