Transcriptome wide analysis of regulatory interactions of the RNA-binding protein HuR

Svetlana Lebedeva[1,*], Marvin Jens[1,*], Kathrin Theil[1], Björn Schwanhäusser[2], Matthias Selbach[2], Markus Landthaler[3], Nikolaus Rajewsky[1,†]

[1,2,3]Max Delbrück Center for Molecular Medicine, Robert-Rössle Strasse 10, 13125 Berlin, Germany (MDC)

[1]Laboratory of Systems Biology of Gene Regulatory Elements, MDC

[2]Laboratory of Cell Signalling and Mass Spectrometry, MDC

[3]Laboratory of RNA Biology and Posttranscriptional Regulation, MDC

*These authors contributed equally to this work.

[†]Correspondence rajewsky@mdc-berlin.de

Summary

Post-transcriptional gene regulation relies on hundreds of RNA binding proteins (RBPs) but the function of most RBPs is unknown. The human RBP HuR/ELAVL1 is a conserved mRNA stability regulator. We used PAR-CLIP, a recently developed method based on RNA-protein crosslinking, to identify transcriptome wide ~26,000 HuR binding sites. These sites were on average highly conserved, enriched for HuR binding motifs and mainly located in 3' untranslated regions. Surprisingly, many sites were intronic, implicating HuR in mRNA processing. Upon HuR knock down, mRNA levels and protein synthesis of thousands of target genes were down regulated, validating functionality. HuR and miRNA binding sites tended to reside nearby but generally did not overlap. Additionally, HuR knock down triggered strong and specific up regulation of miR-7. In summary, we identified thousands of direct and functional HuR targets, found a human miRNA controlled by

HuR, and propose a role for HuR in splicing.

Introduction

In eukaryotic cells, mRNA levels do not directly translate into protein levels because mRNA processing, transport, stability and translation are regulated post-transcriptionally. These fundamental processes are controlled and carried out by RNA-binding proteins (RBPs) and small RNAs which predominantly bind to specific elements located in the untranslated regions (UTRs) of target mRNAs.

There are hundreds of RBPs in the human genome, mostly with unknown function. However, it is believed that RBPs and small RNAs target mRNAs in an orchestrated way to regulate mRNA localization, half life time, and finally the amount of protein synthesized. The joint effect of these factors on mRNAs is hypothesized to constitute a "post-transcriptional regulatory code" (Keene, 2007). Deciphering this code requires the comprehensive identification of RBP binding sites. Since a RBP can have thousands of functional target sites (Licatalosi et al., 2008; Hafner et al., 2010), experiments need to reliably detect RBP-mRNA interactions on a transcriptome wide scale and resolve sites at nucleotide resolution. Previously, RIP-chip (RBP ImmunoPrecipitation and microarray analysis) (Keene et al., 2006) has identified numerous functional mRNA targets of RBPs, including HuR (Lal et al., 2004; de Silanes et al., 2004; Mukherjee et al., 2009). While the RIP-chip assay may also falsely detect indirect interactions or, depending on lysis conditions, even artefactual complexes formed in cell lysate (Mili & Steitz, 2004), it, more importantly, can neither identify the precise location of binding sites on mRNAs, nor detect potential binding to introns. Finally, the functional importance of the identified HuR interactions has only been tested in a few cases.

Current methods employ crosslinking of mRNA-RBP complexes by UV light to identify direct RBP binding sites (Ule et al., 2003; Licatalosi et al., 2008; Hafner et al., 2010; König et al., 2010), with iCLIP and PAR-CLIP providing nucleotide resolution. For our study of HuR we used PAR-CLIP (Photoactivatable ribonucleoside enhanced crosslinking and immunoprecipitation) (Hafner et al.,

2010) to comprehensively identify HuR binding sites in the transcriptome of human cells (HeLa). In this method, photoactivatable nucleosides such as 4-thiouridine (4SU) are incorporated into nascent RNAs providing a strongly enhanced crosslinking efficacy at a relatively short (~1 min) and low energy pulse of UV. Furthermore, crosslinked nucleosides leave a signature in sequencing reads that allows to identify binding sites at nucleotide resolution. PAR-CLIP has been used to identify transcriptome wide binding sites of several RBPs (Hafner et al., 2010).

The Hu/ELAV family of RNA-binding proteins is conserved across metazoans and has diverse functions in mRNA metabolism (Hinman & Lou, 2008). HuR is expressed broadly across tissues (Lu & Schneider, 2004) and a knock out is lethal in mice (Katsanou et al., 2009). The other human Hu proteins, HuB, HuC and HuD, are specific to neurons and required for nervous system development (Akamatsu et al., 1999). They regulate alternative splicing by binding U-rich elements in introns (Zhu et al., 2006). In contrast, only a single alternatively spliced exon has been reported to be skipped upon HuR binding to the exon (Izquierdo 2008).

The HuR protein offers a number of sites for post-translational modifications which allow for a shift from the normal, predominantly nuclear to a more cytoplasmic localization (Kim et al., 2008), especially under stress conditions (Bhattacharyya et al., 2006). In the cytoplasm, Hu proteins can stabilize mRNAs by binding to AU-rich elements (AREs) within 3'UTRs (Brennan & Steitz, 2001), consistent with *in vitro* studies of HuR (Hinman & Lou, 2008, and references within). However, the different classes of AU-rich elements recognized by HuR can also be bound by ~20 other ARE-binding proteins (Brennan & Steitz, 2001). The high resolution of our PAR-CLIP data allowed us to also explore the motif contents of binding sites in an unprecedented way, elucidating the specificity of HuR binding.

Since HuR is known to stabilize bound mRNAs, we used mRNA next generation sequencing to record changes in mRNA levels upon siRNA knock down, thereby testing the functional relevance of our identified binding sites. HuR was also proposed to promote translation, especially under stress conditions (Mazan-Mamczarz et al., 2003; Lal et al., 2005; Kawai et al., 2006). We therefore

employed state of the art mass spectrometry-based proteomics combined with stable isotope labeling (Ong et al., 2002, Selbach et al., 2008), quantifying changes in protein synthesis for thousands of proteins after HuR knock down.

Finally, it has been proposed that HuR can interact with microRNAs (miRNAs). miRNAs are single stranded ~22nt long non-coding RNAs which originate from hairpin precursors encoded as parts of introns or independent transcriptional units. miRNAs recognize their targets by base-pairing complementarity primarily within the 5'-most 6 to 8 nucleotides of the miRNA, the so called seed (Bartel, 2009). It has been shown that the >800 known human miRNAs regulate between 30-75% of all human genes (Krek et al., 2005; Lewis et al., 2005; Xie et al., 2005; Bartel, 2009). miRNAs act in complex with Argonaute (Ago) proteins as part of the RNA-induced silencing complex (RISC) to induce mRNA degradation or repress translation (Baek et al., 2008; Selbach et al., 2008; Filipowicz et al., 2008). HuR has been reported to recruit the miRNA let-7 to repress c-myc expression (Kim et al., 2009), whereas binding of HuR to the mRNA of CAT-1 has been reported to remove repression by miR-122 (Bhattacharyya et al., 2006). It has also been proposed that miRNA targets are enriched among the mRNAs bound by HuR (Mukherjee et al., 2009). We used miRNA target predictions (Lewis et al., 2005; Krek et al., 2005; Lall et al., 2006) and experimentally defined RISC binding data (Hafner et al., 2010) to investigate the spatial arrangement of HuR and miRNA binding sites. To detect miRNA expression changes upon HuR knock down we used next generation sequencing of small RNAs.

Altogether we identified endogenous HuR binding sites, transcriptome wide, using the high resolution PAR-CLIP assay and traced the effects of HuR depletion on a) the transcriptome, including changes in splicing, b) changes in protein synthesis and c) microRNA expression.

Results

PAR-CLIP identifies sites of endogenous HuR binding

To identify binding sites of HuR we used PAR-CLIP (Hafner et al., 2010, Methods) (Figure 1A,

left) in unstressed (Suppl. Figure S1) HeLa cells, performing IP of the endogenous HuR protein and using, independently, 4-thiouridine (4SU) and 6-thioguanosine (6SG) to assess a possible nucleotide bias. As our proteomics measurements required labeling of cells in a special medium we also performed PAR-CLIP on cells grown in SILAC medium. Altogether we performed three PAR-CLIP experiments (Figure 1F).

Efficient crosslinking lead to specific nucleotide conversion events during reverse transcription and next generation sequencing of RNA from each experiment: crosslinked 4SU and 6SG residues were converted into C and A, respectively (Figure 1B,C). These conversions mark the RBP binding site on the target RNA (Hafner et al., 2010). All PAR-CLIP sequencing data (Suppl. Table ST1) were analyzed with our computational pipeline to determine HuR binding sites at an estimated 5% false positive rate from filtered clusters of aligned reads (Methods, Suppl. Methods). Figure 1D shows an example of aligned reads and conversion events for the HuR target *Wnt5A* (Leandersson et al., 2006). A mean length of 27-39 nt for filtered clusters demonstrates the high resolution of PAR-CLIP (Figure 1E).

PAR-CLIP reproducibly identifies thousands of transcripts directly bound by HuR

In each PAR-CLIP experiment we identified ~15-20,000 HuR binding sites which could be assigned to genes, counting clusters in 5'UTR, CDS, 3'UTR and introns based on the RefSeq annotation (Figure 2A,B).

As anticipated, most sites were located in 3'UTRs, but a large fraction (~30-35%) fell into introns, consistent with the predominantly nuclear HuR localization. We combined the data from all PAR-CLIP experiments to derive a set of consensus binding sites supported by reads from at least two out of three experiments (Suppl. Methods). The binding sites are publicly available via our institute's database at http://dorina.mdc-berlin.de. The distribution of consensus sites is comparable to the results obtained from individual experiments (Figure 2C) and is largely independent of transcript expression (Suppl. Figure S2A).

We identify ~2,000 to ~3,500 HuR target genes independently in each of the experiments. 74% of the 6SG PAR-CLIP targets were reproduced using 4SU (Figure 2D) and the 4SU PAR-CLIP in normal medium confirmed the corresponding experiment in SILAC medium (Figure 2E). The larger number of consensus sites leads to a set of 4128 genes bound in exons and 746 genes with exclusive evidence for intronic binding of HuR. In total, 4874 genes were observed to interact with HuR via intronic or exonic sequence elements in at least two out of three PAR-CLIP experiments ("consensus set", Suppl. Table ST2). Target genes were subjected to a GO term enrichment analysis (Suppl. Table ST3). The enriched categories suggest that HuR preferentially targets other regulators of gene expression which act at the post-transcriptional or the transcriptional and post-translational level.

We compiled a list of 68 human HuR target genes with functional evidence described in the literature and expressed in HeLa cells (Suppl. Methods). We considered 9064 genes with an mRNA expression ≥ 5 FPKM (mRNA sequencing read pairs per kilobase of exon per million read pairs (Trapnell et al., 2009)) as expressed. Under the conservative assumption that all of the literature targets are bound by HuR in HeLa cells, we found the consensus set to be highly significantly enriched in known targets, recovering 55 out of 68 genes (81%, *P*-value < 2.7E-6 hypergeometric test).

We independently performed HuR IP in unlabeled cells and validated five HuR targets out of five tested with RT-PCR, including the effector of the RISC complex AGO2 and the non-coding RNA MALAT1 (Figure 2F). To compare RIP and PAR-CLIP more systematically, we quantified a larger number of transcripts in RIP by the NanoString nCounter Assay. NanoString is a RNA expression profiling technology based on counting individual mRNA molecules (Geiss et al., 2008). The assay quantified 236 genes in parallel, of which 97 were flagged as targets by PAR-CLIP (consensus set). The set of genes was not customized by us but simply picked from the available probesets. We considered 86 genes with more than 2-fold enrichment in RIP as "RIP-targets". 65 of the RIP-targets (76%) were also PAR-CLIP targets, validating 67% of our PAR-CLIP targets by this assay (Figure 2G). Comparison of the remaining 21 "RIP-only" genes and the 65 that were called by both RIP and PAR-CLIP (Figure 2H) revealed that PAR-CLIP targets were significantly more enriched in the IP

(*P*-value < 0.004 Mann-Whitney-U), indicating that RIP-only targets are less likely to be true positives.

Intronic HuR binding sites are highly conserved

In each cluster the position with the largest number of nucleotide conversion events indicates the preferred position of crosslinking. It provides a point of reference common to all binding sites, hereafter referred to as “anchor”. By aligning a large number of clusters on their anchors, we obtained high resolution maps revealing common features of HuR binding sites.

We first analyzed evolutionary conservation across 44 vertebrate species, spanning human to lamprey, using the PhyloP nucleotide conservation score (Pollard et al., 2010). We averaged the PhyloP score at a given distance from the anchor over all binding sites. Figure 3A shows the resulting conservation profile of intronic HuR binding sites. Randomly selected positions inside introns which harbor HuR binding sites served as a control. We found a highly conserved core of ~6nt around the anchor residing in a larger context of highly elevated conservation, indicating the proximity of larger functional elements.

Intronic HuR binding sites are associated with splice sites

Asking about the general distribution of HuR binding within introns, we averaged the presence of PAR-CLIP reads (hereafter referred to as “signal”) along all human introns (Suppl. Methods, Figure 3B). HuR binds almost uniformly within introns with a preference towards splice sites. Strikingly, there is a sharp peak close to the 3’ end and a broader preferred binding region towards the 5’ end. Figure 3C shows the signal around 5' and 3' splice sites at nucleotide resolution. A sharp peak is situated ~20nt upstream of the 3' splice site, potentially overlapping with the polypyrimidine tract in many introns. Together, our data indicate that HuR binding in introns preferentially occurs close to exons with a strong bias to bind just upstream of the 3’ splice site.

HuR binding sites in 3'UTRs are found nearby miRNA seeds but rarely overlap

Similar to intronic sites, 3'UTR binding sites of HuR (Figure 3D) show ~6nt around the anchor with high conservation. However, in 3'UTRs broad shoulders of high conservation appear on both sides, approaching background level with increasing distance. As 3'UTRs are hubs of post-transcriptional regulation we hypothesized the observed pattern may reflect the proximity of functional elements such as miRNA seeds. As little is known about the joint arrangement of RBP and miRNA binding, we computed the HuR signal along 3'UTRs and contrasted it with the publicly available PAR-CLIP data for Argonaute proteins in HEK 293 cells (Hafner et al., 2010) (Figure 3E). Ago proteins preferentially bind towards the boundaries of 3'UTRs, in accordance with previous reports (Grimson et al., 2007; Nielsen et al., 2007). In contrast, HuR binding is almost uniform along 3'UTRs but declines towards the stop codon and the polyadenylation site, on average avoiding the areas of Ago binding.

To investigate local interactions we computed the density of conserved miRNA target sites (conserved ”seeds”) predicted by PicTar (Krek et al., 2005; Lall et al., 2006) or TargetScanS (Lewis et al., 2005) around HuR anchors (Figure 3F). We identified 740 HuR anchors directly overlapping miRNA seeds (Suppl. Table ST4). These sites are interesting candidates for HuR-miRNA interactions. However, randomly selecting positions in 3'UTRs results in comparable or even larger numbers. The observed profile suggests that generally, instead of overlapping with HuR binding sites, miRNA seeds occur in the immediate vicinity (~20nt) of HuR binding sites, consistent with the preference of miRNA target sites for an AU-rich sequence context (Grimson et al., 2007; Nielsen et al., 2007).

PAR-CLIP reproducibly recovers *in vitro* HuR motifs

To elucidate the sequence preference of HuR we counted 7-mer occurrences in 41nt windows centered on the anchors of binding sites. Figure 4A shows the log-frequencies of 7-mers with at least 10 counts for 4SU and 6SG experiments. The high-affinity motifs UUUUUUU, UUUAUUU and UUUGUUU from *in vitro* protein binding microarray experiments (Ray et al., 2009) are not only

most abundantly found, but their ranks also follow the described affinities. While G-containing 7-mers showed less enrichment in the 4SU experiment, overall there is a remarkable correlation (Spearman 0.88). On the transcript level, U and G content was the same for targets derived from 4SU and 6SG PAR-CLIP (Suppl. Figures S2C,D).

We used the miReduce algorithm (Sood et al., 2006) to search for words in mRNA sequences associated with changes in perturbation experiments, finding UAUUUAU occurrence in 3'UTRs to be highly significantly associated with mRNA reduction upon HuR knock down (Suppl. Table ST5). UAUUUAU constitutes the core of AREs and is a known high affinity HuR motif. AU-rich motifs are clearly enriched among HuR binding sites compared to all 3'UTR sequences while U and G content alone did not suffice for crosslinking or enrichment (Figure 4B). Consistent with the intronic binding of HuR, polypyrimidine rich sequences were observed as frequently as AREs. The motif analyses indicate that our PAR-CLIP experiments were able to quantitatively capture the *in vivo* binding preferences of HuR regardless of the used thionucleoside label.

HuR binds single stranded RNA with no preference for hairpins

RNA secondary structure can contribute to the specificity of RBP binding. HuR has been proposed to associate with hairpin loops (de Silanes et al., 2004) which later have been contested to contribute to specificity (Mukherjee et al., 2009). We computationally folded HuR binding sites in 3'UTRs (Suppl. Methods) and averaged the resulting base pairing probabilities, finding a substantially reduced pairing probability in the direct vicinity of HuR anchors but no indication of a hairpin structure (Figure 4C).

HuR knockdown induces specific down regulation of target gene expression

We used RNA interference to deplete HuR in HeLa cells and monitored changes on transcript levels with next generation paired-end mRNA sequencing (Methods). Transcript expression levels were estimated from the sequencing data (Suppl. Methods).

Results were highly reproducible between technical and biological replicates (Spearman 0.98 and 0.82, respectively, Suppl. Figure S3A). In addition, we validated mRNA expression levels by the NanoString nCounter assay (Spearman 0.82, Figure 5A).

To compare the effect of the HuR depletion on various groups of transcripts, we computed cumulative density fractions (Figure 5B-D). Here, for a given log fold change x, the fraction of genes with a change $\leq x$ is shown.

Consistent with the mRNA stabilizing function of HuR, targets identified by PAR-CLIP (purple) are significantly (P-value ≈ 0, t-test) more destabilized than non-targets (black), confirming the overall functionality of PAR-CLIP targets. The top 20% of targets with most binding sites showed strongest down regulation (pink line in Figure 5B). The down regulation was slightly weaker but also highly significant (P-value < 1E-06) for intronic targets (dashed line, Figure 5B).

To assess the effect of HuR on translation we performed pulsed SILAC ("pSILAC") proteomics measurements essentially as in (Selbach et al 2008) (Methods, Figure 5C) in an independent sample. In pSILAC, during a short time window (24h), newly synthesized proteins incorporate stable isotope labeled amino acids. The mass shift between HuR knock down (medium heavy label) and unperturbed conditions (heavy label) allowed to quantify thousands of changes in protein synthesis, independent of the unlabeled pool of pre-existing proteins (light). We could quantify changes in protein synthesis for ~4,300 proteins. Overall, the effects of HuR depletion on protein synthesis reflect the changes on mRNA levels with a specific and significant (P-value < 1E-04) reduction for HuR targets (Figure 5D). The protein synthesis of intronic targets is also significantly reduced (P-value < 0.01), with the impact of HuR depletion being relatively more pronounced than on mRNA level, consistent with a role of HuR in pre-mRNA processing (Discussion).

HuR and alternative splicing

Given the binding of HuR to introns, we screened our mRNA sequencing data for genes with HuR dependent alternative splicing. Quantifying changes in alternative exon inclusion upon HuR knock

down (Suppl. Methods), we found 51 candidate exons with either reduced (30 exons) or increased inclusion (21 exons) associated with HuR binding sites within 1kb into the flanking pre-mRNA or the exon itself (Suppl. Table ST7). Out of 6 tested candidate exons 4 showed significant changes in splicing in a PCR assay performed in two independent biological replicates (Figure 6). One of the genes showing HuR dependent alternative splicing is the splicing factor *PTBP2*. Upon HuR knock down the expression of exon 10, which is flanked by HuR binding sites, increases by 65% relative to the flanking exons 9 and 11. Skipping of exon 10 leads to nonsense mediated decay (NMD) (Spellman et al., 2007) and is employed in the crossregulation of the polypyrimidine-tract binding (PTB) proteins, suggesting an interplay of HuR with these splicing regulators.

HuR regulates miR-7 processing

Many miRNA precursors reside in introns, and their expression oftentimes correlates with the transcription of the host gene. However, miRNA expression can also be regulated by RBPs at the level of precursor processing (Guil & Cáceres, 2007; Rybak et al., 2008; Trabucchi et al., 2009).

We sequenced small RNAs in mock transfected and HuR knock down conditions and found miR-7 to be the only miRNA strongly regulated in the knock down (Figure 7A). The effect was dose-dependent: stronger knock down with siRNA1 caused ~20-fold up regulation, weaker knock down with siRNA2 ~3-5 fold (Figure 7B). In HeLa cells the *miR-7-1* locus in the last intron of *HNRNPK* is the source of mature miR-7 (Suppl. Discussion), which is barely detectable in wt HeLa cells (~100 sequencing reads). Therefore, as the housekeeping gene *HNRNPK* is highly expressed, the biogenesis of mature miR-7 must be strongly suppressed. Expression of *HNRNPK* does not change upon HuR knock down, indicating that the observed up regulation of mature miR-7 is due to a de-repression at the level of processing. The HuR binding sites in the intron and the surrounding exons (Figure 7C) suggest that HuR binding may directly influence the fate of the excised intron harboring the miR-7 precursor.

Discussion

PAR-CLIP reproducibly identifies thousands of HuR target genes

Using PAR-CLIP we identified ~26,000 binding sites of endogenous HuR supported by two out of three independent experiments, discovering extensive binding to introns. The identified sites are enriched for known HuR motifs and show a distinct pattern of sequence conservation. Our motif analysis confirmed that HuR binds single stranded RNA with no further structural preferences and revealed substantial differences in the affinity to closely related sequences. For example, while some polypyrimidine motifs were abundantly bound, multiple consecutive cytidines rarely occurred. This indicates complex rules for HuR binding, beyond general affinity to U-rich, AU-rich or polypyrimidine sequences. According to our data HuR is an abundant protein (Suppl. Discussion) with versatile sequence recognition that interacts with up to 4874 genes, corresponding to ~50% of all HeLa cell transcripts, with 746 genes showing exclusively intronic binding. The overlaps between the PAR-CLIP consensus set and known HuR targets as well as our independently performed RIP assay are highly significant. Individual PAR-CLIP experiments also showed good reproducibility. The major cause for the remaining variance lies in the RNA digestion with RNAse T1 after pull-down, which determines the fragment size distribution available for sequencing and mapping. Moreover, the preference of RNAse T1 to cleave after G residues renders this influence sequence dependent. However, given the overall agreement between experiments we believe that we were able to control this step.

HuR knock down confirms functionality

Based on the mRNA stabilizing effect of HuR binding we validated the functional relevance of the reported interactions: siRNA knock down of HuR led to a highly significant destabilization of transcripts with HuR binding sites. This cannot be explained by off-target effects of the siRNA (Suppl. Discussion and Suppl. Figure S3B). The destabilization was most pronounced for

transcripts with many binding sites and unexpectedly extended to genes with intronic binding sites, indicating a function of HuR that is independent of inhibiting ARE-mediated decay. The pulsed SILAC experiment allowed us to trace the effects of HuR depletion further to changes in protein synthesis. The majority of HuR targets contain 3'UTR binding sites and show a reduction in protein synthesis that is highly correlated with the reduction of mRNA (Spearman 0.6, P-value < 1E-138, Suppl. Table ST6 lists all target genes with changes of mRNA and protein synthesis). In contrast, the significant reduction in protein output observed for intronic target genes (see also Suppl. Figure S3C) was stronger than and weakly correlated with the transcript-level changes (Spearman 0.43), indicating a functional role of HuR in pre-mRNA processing: Aberrantly spliced transcripts would be incapable of efficient translation and thus decouple total cellular mRNA levels from the rate of protein synthesis.

HuR interacts with introns and modulates splicing

HuR was reported to be associated with the spliceosome (Chen et al., 2007) and to affect a splicing reporter (Wang et al., 2010) but, unlike the neuronal Hu proteins, has not been reported to bind introns before. We find an association of HuR with splice sites that is reminiscent of the neuronal Hu proteins which bind U-rich sequences, promoting the skipping of proximal exons (Zhu et al., 2006).

Pointing in the same direction, HuR binds pyrimidine-rich sequences (Figure 4A,B) explaining in part the observed pattern of intronic binding with a peak ~20nt upstream of exons (Figure 3C). HuR also interacts directly with the human polypyrimidine tract binding (PTB) genes in multiple ways, apparently stabilizing the non-neuronal pattern with dominant *PTBP1* expression: HuR probably stabilizes *ROD1* and *PTBP1* by strong, ARE-containing 3'UTR binding sites. In contrast, the neuronal *PTBP2* shows only weak interaction with HuR via its 3'UTR but displays binding sites in the introns flanking exon 10, interestingly outside the polypyrimidine tract that can be bound by PTBP1. Skipping of this exon is known to cause nonsense mediated decay (NMD)(Spellman et al.,

2007). Consistent with a 4-fold increase in PTBP2 protein synthesis upon knock down of HuR but only modest reduction of PTBP1 we suspected that HuR acts together with PTBP1 to promote NMD of *PTBP2*. Experiments with a minigene confirmed the functionality of the HuR bound sequence element downstream of exon10. However, the observed effect of HuR binding was not conclusively resolved by our introduced mutations (not shown) and further experiments are required to determine the extent of regulation exerted directly by HuR on the splicing of *PTBP2* exon10.

HuR binding sites represent a preferred context for miRNA seeds

HuR targets tend to have long 3'UTRs (Suppl. Figure S2B) increasing the probability to be targeted by at least one miRNA. Yet, we found that direct overlap between HuR and miRNA binding sites occurs less often than expected by chance (Figures 3E, F). The known preference of miRNA seeds to reside in a context of AU-rich sequences (Grimson et al., 2007; Nielsen et al., 2007) provides an explanation for the more typical arrangement with miRNA seeds in the proximity of HuR binding sites. We report the 740 HuR binding sites that do overlap with conserved miRNA target sites (Suppl. Table ST4) as interesting candidates for direct HuR-miRNA interactions. We also cannot exclude the possibility of non-steric interactions between HuR and RISC for example by modulating RNA secondary structure.

Suppression of miR-7 biogenesis is relieved upon HuR knock down

An unexpected result was the selective, strong, and dose dependent up regulation of miR-7 upon HuR knock down. Remarkably, the ~20 fold induction of mature miR-7 was the most pronounced change in gene expression observed in any of our experiments (Figure 7A,B). Two recent studies reported post-transcriptional regulation of miR-7 biogenesis (Kefas et al., 2008; Wu et al., 2010). The identified HuR binding sites in the last intron of *HNRNPK* and the flanking exons suggest that miR-7 processing is directly controlled by HuR binding (Figure 7C), adding an evolutionarily old miRNA to the list of HuR targets. Given the extreme and specific response of miR-7 to HuR levels, it is intriguing to speculate that miR-7 might be one of the key targets of HuR. Since HuR and

HNRNPK seem to be expressed in the vast majority of human tissues, this regulatory relationship might be important in many biological contexts.

HuR as a hub for regulating RNA metabolism

A significant fraction of HuR target genes is involved in mRNA metabolism (Suppl. Table ST3), connecting HuR to a large network of post-transcriptional gene regulators. This includes central components of RISC: *AGO2* and *TNRC6B* bear multiple conserved 3'UTR binding sites of HuR. In light of only modest down regulation upon HuR depletion and the observed variety of HuR motifs we speculate that some cytoplasmic interactions with HuR may serve a different purpose than stabilization against ARE-mediated decay. HuR is known for its central role in mediating stress response and localizes to stress granules, potentially sequestering bound transcripts. For components of the miRNA pathway this may be a way to relieve miRNA repression upon stress. Of note, TNRC6B and AGO2 proteins meet in P-bodies with RNA degradation factors like the CCR4-NOT deadenylation complex (CNOT6,6L,7,8) and the RNA helicase Rck/p54 (DDX6), the transcripts of which are also high confidence HuR targets. We speculate that the function of HuR interactions with key components of different types of RNA granules, hot spots of RNA regulation, could also be relevant in the context of cellular stress response.

HuR is known to stabilize many transcripts encoding genes necessary for the immediate response to stress. This is reflected in our data, containing the known HuR target *HIF1A* (Galbán et al., 2008) and identified by us *TXNIP* (thioredoxin-interacting protein) and *PRKAA1*, a kinase that affects HuR localization upon stress (Wang et al., 2002), suggesting a feedback mechanism.

To conclude, we have identified and characterized an unprecedented number of functional HuR targets, promoting HuR to be a major hub for regulating RNA metabolism in the cell.

This regulation seems to involve known mechanisms recovered in this study (binding to target 3’ UTRs and stabilizing target mRNA levels) and mechanisms proposed in this study (regulation of splicing). Our data also implicate HuR in regulating members of the miRNA pathway and, specifi-

cally, miR-7. The methodology developed here can be used to study HuR in stress conditions, which will shed more light on the function of HuR and its many targets.

Methods

Transfection

Plasmids were transfected with Lipofectamine 2000 (Invitrogen), and siRNAs were transfected at a final concentration of 100 nM, using Lipofectamine RNAiMAX (Invitrogen). Controls (mock) contained only the transfection reagent.

Transcriptome sequencing

PolyA+ mRNA was isolated from 1µg of Trizol extracted total RNA using magnetic Oligo-dT_{25} beads (Invitrogen). NEBNext kit (NEB) and a customized protocol were used to prepare mRNA for sequencing (Suppl. Methods). The libraries were sequenced on a Illumina Genome Analyzer GAII or Illumina HiSeq for 2x76 cycles or 2x100 cycles (paired-end protocol).

Labeling of proteins, sample preparation and measurement by mass spectrometry

Cells were transferred to light SILAC medium 6h post transfection. Two days after transfection siRNA and mock-transfected cells were transferred to medium-heavy and heavy SILAC medium, respectively. After 24h of labeling cells were harvested and equal amounts of siRNA- and mock-transfected cells were combined. Proteins were extracted, separated by SDS-PAGE, trypsin-digested and analyzed by liquid chromatography tandem mass spectrometry (LC-MS/MS) on a high resolution instrument (LTQ-Orbitrap Velos, Thermo Fisher). Raw files were processed by MaxQuant (version 1.0.13.13) for peptide/protein identification at 1% FDR and quantification (Suppl. Methods).

was performed from 10 µg total RNA using the FlashPage Gel system (Ambion) and the standard Illumina small RNA library preparation protocol.

PAR-CLIP

The cells were labeled with 100 µM 4SU or 6SG. After labeling, procedure followed the PAR-CLIP protocol as described (Hafner et al., 2010). Briefly, UV irradiated cells were lysed in NP-40 lysis buffer. Immunoprecipitation was carried out with protein G magnetic beads (Invitrogen) coupled to HuR antibody (3A2, Santa Cruz, sc-5261) for 1h at 4 ºC. Beads were treated with CIP (NEB) and radioactively labeled. The crosslinked protein-RNA complexes were resolved on 4-12% NuPAGE gel (Invitrogen) and the 37kDa band corresponding to HuR was cut out. The RNA was isolated by electroelution followed by Proteinase K digestion and phenol-chloroform extraction and sequenced according to the standard small RNA protocol (Suppl. Methods for more details).

RIP-PCR

Immunoprecipitations were performed as described for PAR-CLIP. As negative control an anti-FLAG antibody (Sigma, F3165) was used. Typically, 5-10 15cm plates, 50-100µl Protein G beads and 10-20µg antibody were used per IP reaction. RNA was isolated from IP and analyzed by RT-PCR and agarose gel electrophoresis.

NanoString nCounter Assay

The NanoString nCounter Assay is available as a custom service by NanoString Technologies. Equal amounts (150ng) of RNA isolated from the IP with anti-HuR and anti-FLAG antibodies, as well as total RNA from mock and siRNA-transfected cells were analyzed in parallel using the nCounter Human Cancer Reference Kit (GXA-CR1-12).

RT-PCR

Trizol isolated RNA was treated with RQ1 DNase (Promega). cDNA synthesis was performed with

Superscript III (Invitrogen) with Oligo(dT) (T_{18}NN). PCR amplification was performed using 2x Green DreamTaq Master Mix (Fermentas), 0.5µM of each of the forward and reverse primers, and 1µl of cDNA for 30 cycles of 15 s at 94 ºC, 15 s at 60 ºC, and 20 s at 72ºC.

Quantification of splicing

After RT-PCR, the products were resolved by 8% TBE-PAGE. In parallel, PCR products were purified by phenol-chloroform extraction and analyzed by Agilent BioAnalyzer DNA 1000 Assay. PSI (Percent Spliced In) values were calculated as the molar ratio of the peak corresponding to the exon containing isoform and the sum of the peaks representing both isoforms.

PAR-CLIP computational pipeline

We developed a pipeline that performed all steps of the analysis from raw reads to cluster sets and target genes, in a largely automated and unbiased way. The emphasis was on stringent filtering and controlling the false-positive rate in the identification of binding sites (Suppl. Methods).

Briefly, reads were aligned to the human transcriptome (pre-mRNAs), allowing for up to one mismatch, insertion or deletion. Only uniquely mapping reads were retained.

We identified clusters of aligned PAR-CLIP reads continuously covering regions of pre-mRNA sequence. The number of T to C or G to A mismatches served as a crosslink score. We also assigned a quality score based on the number and positions of distinct reads contributing to the cluster (Suppl. Methods).

As the reads should originate from HuR-bound transcripts we regarded clusters aligning antisense to the annotated direction of transcription as false positives. We were thus able to select cutoffs on both scores such as to keep the estimated false positive rate below 5%. After filtering by these cutoffs we expect each remaining cluster to harbor at least one HuR binding site.

Accession numbers

The sequencing data have been deposited in the GEO database under the accesion number GSE29943.

Author contributions

SL designed the experiments. SL and KT conducted the experiments. MJ conducted and designed the computational analyses. BS performed the SILAC experiments, supervised by MS. ML provided guidance for the PAR-CLIP and knock down experiments. NR conceived and supervised the project. SL, MJ, and NR analyzed the data and wrote the paper.

Acknowledgements

We are grateful to Wu et al. (Department of Cell Biology, Duke University Medical Center) for providing their miR-7-1 minigene. We thank all members of the Rajewsky lab for discussion, Salah Ayoub and Lena von Oertzen for technical assistance, Claudia Langnick and Mirjam Feldkamp from the Wei Chen lab (MDC) for the sequencing runs, and Sebastian Memczak for help with miR-7 primers. We thank Kerstin Baethge from the Landthaler lab for technical help with PAR-CLIP. We acknowledge generous support by Markus Hafner & Tom Tuschl (The Rockefeller University) providing unpublished HuR PAR-CLIP data in HEK293 cells. MJ thanks the Deutsche Forschungsgemeinschaft for a fellowship in the International Research Training Group Genomics and Systems Biology of Molecular Networks (GRK 1360), SL thanks the international PhD program at the MDC.

The authors declare no conflict of interest.

References

Akamatsu, W.; Okano, H. J.; Osumi, N.; Inoue, T.; Nakamura, S.; Sakakibara, S.; Miura, M.; Matsuo, N.; Darnell, R. B. & Okano, H. (1999). Mammalian ELAV-like neuronal RNA-binding pro-

teins HuB and HuC promote neuronal development in both the central and the peripheral nervous systems. Proc Natl Acad Sci U S A *96*, 9885-9890.

Baek, D.; Villén, J.; Shin, C.; Camargo, F. D.; Gygi, S. P. & Bartel, D. P. (2008). The impact of microRNAs on protein output. Nature *455*, 64-71.

Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215-233.

Bhattacharyya, S. N.; Habermacher, R.; Martine, U.; Closs, E. I. & Filipowicz, W. (2006). Relief of microRNA-mediated translational repression in human cells subjected to stress. Cell *125*, 1111-1124.

Brennan, C. M. & Steitz, J. A. (2001). HuR and mRNA stability. Cell Mol Life Sci *58*, 266-277.

Chen, Y.-I. G.; Moore, R. E.; Ge, H. Y.; Young, M. K.; Lee, T. D. & Stevens, S. W. (2007). Proteomic analysis of in vivo-assembled pre-mRNA splicing complexes expands the catalog of participating factors. Nucleic Acids Res *35*, 3928-3944.

Filipowicz, W.; Bhattacharyya, S. N. & Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat Rev Genet *9*, 102-114.

Galbán, S.; Kuwano, Y.; Pullmann, R.; Martindale, J. L.; Kim, H. H.; Lal, A.; Abdelmohsen, K.; Yang, X.; Dang, Y.; Liu, J. O.; Lewis, S. M.; Holcik, M. & Gorospe, M. (2008). RNA-binding proteins HuR and PTB promote the translation of hypoxia-inducible factor 1alpha. Mol Cell Biol *28*, 93-107.

Geiss, G. K.; Bumgarner, R. E.; Birditt, B.; Dahl, T.; Dowidar, N.; Dunaway, D. L.; Fell, H. P.; Ferree, S.; George, R. D.; Grogan, T.; James, J. J.; Maysuria, M.; Mitton, J. D.; Oliveri, P.; Osborn, J. L.; Peng, T.; Ratcliffe, A. L.; Webster, P. J.; Davidson, E. H.; Hood, L. & Dimitrov, K. (2008). Direct multiplexed measurement of gene expression with color-coded probe pairs. Nat Biotechnol *26*, 317-325.

Grimson, A.; Farh, K. K.-H.; Johnston, W. K.; Garrett-Engele, P.; Lim, L. P. & Bartel, D. P. (2007). MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell *27*, 91-105.

Guil, S. & Cáceres, J. F. (2007). The multifunctional RNA-binding protein hnRNP A1 is required

for processing of miR-18a. Nat Struct Mol Biol *14*, 591-596.

Hafner, M.; Landthaler, M.; Burger, L.; Khorshid, M.; Hausser, J.; Berninger, P.; Rothballer, A.; Ascano, M.; Jungkamp, A.-C.; Munschauer, M.; Ulrich, A.; Wardle, G. S.; Dewell, S.; Zavolan, M. & Tuschl, T. (2010). Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell *141*, 129-141.

Hinman, M. N. & Lou, H. (2008). Diverse molecular functions of Hu proteins. Cell Mol Life Sci *65*, 3168-3181.

Izquierdo, J. M. (2008). Hu antigen R (HuR) functions as an alternative pre-mRNA splicing regulator of Fas apoptosis-promoting receptor on exon definition. J Biol Chem *283*, 19077-19084.

Katsanou, V.; Milatos, S.; Yiakouvaki, A.; Sgantzis, N.; Kotsoni, A.; Alexiou, M.; Harokopos, V.; Aidinis, V.; Hemberger, M. & Kontoyiannis, D. L. (2009). The RNA-binding protein Elavl1/HuR is essential for placental branching morphogenesis and embryonic development. Mol Cell Biol *29*, 2762-2776.

Kawai, T.; Lal, A.; Yang, X.; Galban, S.; Mazan-Mamczarz, K. & Gorospe, M. (2006). Translational control of cytochrome c by RNA-binding proteins TIA-1 and HuR. Mol Cell Biol *26*, 3295-3307.

Keene, J. D. (2007). RNA regulons: coordination of post-transcriptional events. Nat Rev Genet *8*, 533-543.

Keene, J. D.; Komisarow, J. M. & Friedersdorf, M. B. (2006). RIP-Chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. Nat Protoc *1*, 302-307.

Kefas, B.; Godlewski, J.; Comeau, L.; Li, Y.; Abounader, R.; Hawkinson, M.; Lee, J.; Fine, H.; Chiocca, E. A.; Lawler, S. & Purow, B. (2008). microRNA-7 inhibits the epidermal growth factor receptor and the Akt pathway and is down-regulated in glioblastoma. Cancer Res *68*, 3566-3572.

Kim, H. H.; Abdelmohsen, K.; Lal, A.; Pullmann, R.; Yang, X.; Galban, S.; Srikantan, S.; Martindale, J. L.; Blethrow, J.; Shokat, K. M. & Gorospe, M. (2008). Nuclear HuR accumulation through phosphorylation by Cdk1. Genes Dev *22*, 1804-1815.

Kim, H. H.; Kuwano, Y.; Srikantan, S.; Lee, E. K.; Martindale, J. L. & Gorospe, M. (2009). HuR

recruits let-7/RISC to repress c-Myc expression. Genes Dev *23*, 1743-1748.

Krek, A.; Grün, D.; Poy, M. N.; Wolf, R.; Rosenberg, L.; Epstein, E. J.; MacMenamin, P.; da Piedade, I.; Gunsalus, K. C.; Stoffel, M. & Rajewsky, N. (2005). Combinatorial microRNA target predictions. Nat Genet *37*, 495-500.

König, J.; Zarnack, K.; Rot, G.; Curk, T.; Kayikci, M.; Zupan, B.; Turner, D. J.; Luscombe, N. M. & Ule, J. (2010). iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol *17*, 909-915.

Lal, A.; Kawai, T.; Yang, X.; Mazan-Mamczarz, K. & Gorospe, M. (2005). Antiapoptotic function of RNA-binding protein HuR effected through prothymosin alpha. EMBO J *24*, 1852-1862.

Lal, A.; Mazan-Mamczarz, K.; Kawai, T.; Yang, X.; Martindale, J. L. & Gorospe, M. (2004). Concurrent versus individual binding of HuR and AUF1 to common labile target mRNAs. EMBO J *23*, 3092-3102.

Lall, S.; Grün, D.; Krek, A.; Chen, K.; Wang, Y.-L.; Dewey, C. N.; Sood, P.; Colombo, T.; Bray, N.; Macmenamin, P.; Kao, H.-L.; Gunsalus, K. C.; Pachter, L.; Piano, F. & Rajewsky, N. (2006). A genome-wide map of conserved microRNA targets in C. elegans. Curr Biol *16*, 460-471.

Leandersson, K.; Riesbeck, K. & Andersson, T. (2006). Wnt-5a mRNA translation is suppressed by the Elav-like protein HuR in human breast epithelial cells. Nucleic Acids Res *34*, 3988-3999.

Lewis, B. P.; Burge, C. B. & Bartel, D. P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell *120*, 15-20.

Licatalosi, D. D.; Mele, A.; Fak, J. J.; Ule, J.; Kayikci, M.; Chi, S. W.; Clark, T. A.; Schweitzer, A. C.; Blume, J. E.; Wang, X.; Darnell, J. C. & Darnell, R. B. (2008). HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature *456*, 464-469.

Lu, J.-Y. & Schneider, R. J. (2004). Tissue distribution of AU-rich mRNA-binding proteins involved in regulation of mRNA decay. J Biol Chem *279*, 12974-12979.

Mazan-Mamczarz, K.; Galbán, S.; de Silanes, I. L.; Martindale, J. L.; Atasoy, U.; Keene, J. D. & Gorospe, M. (2003). RNA-binding protein HuR enhances p53 translation in response to ultraviolet light irradiation. Proc Natl Acad Sci U S A *100*, 8354-8359.

Mili, S. & Steitz, J. A. (2004). Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. RNA *10*, 1692-1694.

Mukherjee, N.; Lager, P. J.; Friedersdorf, M. B.; Thompson, M. A. & Keene, J. D. (2009). Coordinated posttranscriptional mRNA population dynamics during T-cell activation. Mol Syst Biol *5*, 288.

Nielsen, C. B.; Shomron, N.; Sandberg, R.; Hornstein, E.; Kitzman, J. & Burge, C. B. (2007). Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. RNA *13*, 1894-1910.

Ong, S.-E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A. & Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics *1*, 376-386.

Pollard, K. S.; Hubisz, M. J.; Rosenbloom, K. R. & Siepel, A. (2010). Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res *20*, 110-121.

Ray, D.; Kazan, H.; Chan, E. T.; Castillo, L. P.; Chaudhry, S.; Talukder, S.; Blencowe, B. J.; Morris, Q. & Hughes, T. R. (2009). Rapid and systematic analysis of the RNA recognition specificities of RNA-binding proteins. Nat Biotechnol *27*, 667-670.

Rybak, A.; Fuchs, H.; Smirnova, L.; Brandt, C.; Pohl, E. E.; Nitsch, R. & Wulczyn, F. G. (2008). A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. Nat Cell Biol *10*, 987-993.

Selbach, M.; Schwanhäusser, B.; Thierfelder, N.; Fang, Z.; Khanin, R. & Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. Nature *455*, 58-63.

de Silanes, I. L.; Zhan, M.; Lal, A.; Yang, X. & Gorospe, M. (2004). Identification of a target RNA motif for RNA-binding protein HuR. Proc Natl Acad Sci U S A *101*, 2987-2992.

Sood, P.; Krek, A.; Zavolan, M.; Macino, G. & Rajewsky, N. (2006). Cell-type-specific signatures of microRNAs on target mRNA expression. Proc Natl Acad Sci U S A *103*, 2746-2751.

Spellman, R.; Llorian, M. & Smith, C. W. J. (2007). Crossregulation and functional redundancy between the splicing regulator PTB and its paralogs nPTB and ROD1. Mol Cell *27*, 420-434.

Trabucchi, M.; Briata, P.; Garcia-Mayoral, M.; Haase, A. D.; Filipowicz, W.; Ramos, A.; Gherzi, R. & Rosenfeld, M. G. (2009). The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. Nature *459*, 1010-1014.

Trapnell, C.; Pachter, L. & Salzberg, S. L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.

Ule, J.; Jensen, K. B.; Ruggiu, M.; Mele, A.; Ule, A. & Darnell, R. B. (2003). CLIP identifies Nova-regulated RNA networks in the brain. Science *302*, 1212-1215.

Wang, H.; Molfenter, J.; Zhu, H. & Lou, H. (2010). Promotion of exon 6 inclusion in HuD pre-mRNA by Hu protein family members. Nucleic Acids Res *38*, 3760-3770.

Wang, W.; Fan, J.; Yang, X.; Fürer-Galban, S.; de Silanes, I. L.; von Kobbe, C.; Guo, J.; Georas, S. N.; Foufelle, F.; Hardie, D. G.; Carling, D. & Gorospe, M. (2002). AMP-activated kinase regulates cytoplasmic HuR. Mol Cell Biol *22*, 3425-3436.

Wu, H.; Sun, S.; Tu, K.; Gao, Y.; Xie, B.; Krainer, A. R. & Zhu, J. (2010). A splicing-independent function of SF2/ASF in microRNA processing. Mol Cell *38*, 67-77.

Xie, X.; Lu, J.; Kulbokas, E. J.; Golub, T. R.; Mootha, V.; Lindblad-Toh, K.; Lander, E. S. & Kellis, M. (2005). Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. Nature *434*, 338-345.

Zhu, H.; Hasman, R. A.; Barron, V. A.; Luo, G. & Lou, H. (2006). A nuclear function of Hu proteins as neuron-specific alternative RNA processing regulators. Mol Biol Cell *17*, 5105-5114.

Figure legends

Figure 1. PAR-CLIP, HuR perturbation, mRNA and protein measurements.

(A) Outline of experiments. PAR-CLIP of endogenous HuR was performed with HeLa cells using RNA labeling with 4-thiouridine (4SU) or 6-thioguanosine (6SG). Genome wide impacts of HuR siRNA knock down on transcript levels were measured by mRNA sequencing ("RNA-seq") and on protein synthesis by pulsed SILAC shotgun proteomics ("pSILAC").

(B,C) Specific mismatches in aligned reads demonstrate efficient mRNA-RBP crosslinking. The frequency of nucleotide mismatches in PAR-CLIP reads aligned to mature mRNAs is shown for 4SU (red) and 6SG (blue). T to C and G to A mismatches are the signature of efficient crosslinking.

(D) Representative example of PAR-CLIP data. The coverage by aligned reads (gray) and nucleotide conversions (purple) are shown along the *WNT5A* gene, a known target of HuR. Binding sites inferred by our pipeline are indicated as purple boxes. Potentially spurious clusters overlapping repetitive elements are discarded. **Insert: Example of a PAR-CLIP consensus cluster.** The *WNT5A* mRNA sequence is shown in uppercase letters at the top. Aligned PAR-CLIP reads are shown in lowercase with mismatches highlighted (T to C in red for 4SU, G to A in blue for 6SG). xN denotes N counts for a read.

(E) PAR-CLIP clusters are typically small. Length histogram of clusters identified in 4SU (red) and 6SG (blue) PAR-CLIP. The average PAR-CLIP cluster size is 27nt to 39nt.

(F) Overview of the samples and experiments.

Figure 2. Identification, analysis and validation of HuR target sites.

(A-C) The distribution of binding sites along transcripts is reproducible and reveals prominent binding within introns.

Binding sites predominantly reside in 3'UTRs. A surprisingly large fraction is intronic. **(A)** Here, reads from 4SU PAR-CLIP in DMEM and SILAC medium were pooled **(B)** 6SG PAR-CLIP was performed in SILAC medium. **(C)** The distribution is not changed if consensus clusters are derived from all libraries.

(D,E) Identification of thousands of reproducible HuR target genes.

Venn-diagrams showing the overlap of target genes between PAR-CLIP experiments. **(D)** The majority of targets (65% to 74%) is detected by 6SG and 4SU labeling.

(E) Overlap of targets from 4SU PAR-CLIP in DMEM and SILAC medium. The identification of

target genes is largely independent of the culturing medium.

(F) Validation of PAR-CLIP targets by RIP and RT-PCR. RT-PCR on RNA from HuR IP validated five PAR-CLIP targets out of five tested. The known HuR target *EIF4E* served as a positive control, isotype-matched anti-FLAG antibody as unspecific control and the highly abundant *GAPDH* (not detected in PAR-CLIP) as a negative control.

(G) Validation of PAR-CLIP targets by RIP and NanoString nCounter gene expression system. We analyzed RNA from HuR-RIP and control using the NanoString system. Eighty-six out of 236 were more than two fold enriched in HuR vs. control IP ("RIP targets"). 97 PAR-CLIP targets were present on the NanoString chip. 65 of them were also at least two fold enriched in HuR-IP ("PAR-CLIP and RIP targets").

(H) Genes enriched in HuR-IP but not detected by PAR-CLIP had significantly weaker enrichment in the IP than PAR-CLIP targets.

See also Suppl. Fig. S2 and Suppl. Tables ST1, ST2.

Figure 3. Detailed analysis of HuR binding sites.

(A) Intronic binding sites of HuR are highly conserved. Average PhyloP nucleotide conservation score at a given distance from crosslink positions ("anchors"). Intronic binding sites display a conserved core of ~6nt and reside in a larger context of elevated conservation.

Random positions were drawn from the same introns to serve as control. The control is not a flat line as the window sometimes overlaps with neighboring exons. The gray envelope represents the standard error of the mean.

(B) Intronic binding sites preferentially locate close to splice sites. The presence of PAR-CLIP reads (HuR "signal") was averaged along all human introns. The signal is uniform within introns but peaks sharply near the 3' splice site and also increases towards the 5' splice site.

(C) Detailed view of splice site context. While direct overlap with splice sites is rare, the HuR signal in the 5' region of an intron is almost as high as in the adjacent exon. HuR binding peaks sharply within ~20nt upstream of the 3' splice sites.

(D) 3'UTR binding sites of HuR are highly conserved. Conservation profile analogous to 3A. 3'UTR binding sites display a core of ~6 conserved nucleotides and flanking regions of high conservation, possibly indicating other regulatory elements.

Random positions from the same 3'UTRs serve as control. The control is not a flat line as the windows sometimes overlap with the CDS on the left and intergenic regions on the right.

(E) HuR and AGO binding profiles on 3'UTRs are different. AGO PAR-CLIP read presence peaks in the beginning and in the end of 3'UTRs. HuR appears to avoid proximity to coding sequences and close proximity to the site of polyadenylation.

(F) miRNA seeds are proximal to but rarely overlap HuR sites. Density of predicted conserved miRNA seeds around anchors in 3'UTRs. HuR anchors and seeds display no tendency for direct overlap but the larger context (10-20nt) shows an elevated seed density.

See also Suppl. Table ST4.

Figure 4. Sequence motifs and secondary structure of HuR binding sites

(A) Binding motifs are reproduced by 4SU and 6SG PAR-CLIP.

Log_{10} frequencies of 7-mers occurring close to HuR anchors are correlated between 4SU and 6SG experiments (Spearman 0.86). Sequences with at least one G are more visible in the 6SG PAR-CLIP. The most abundant motifs UUUUUUU, UUUAUUU and UUUGUUU match known high affinity *in vitro* motifs. AU-rich elements (AREs) and polypyrimidine motifs are also frequent.

(B) Known HuR motifs and AREs are enriched in binding sites. Enrichment of selected 7-mers. Protein binding microarray motifs and AU-rich elements (but not GU-rich elements) are enriched in both 4SU and 6SG-derived clusters compared to all human 3'UTR sequences. Polypyrimidine rich motifs are enriched compared to their background frequency in 3'UTRs as well as introns.

(C) HuR binding sites have weak secondary structure. Sequences of length 201nt centered on HuR anchors in 3'UTRs were computationally folded. The average base pairing probability is strongly reduced close to the anchor, consistent with the low base pairing energy of AU and U-rich sequences (Figure 4A). The peak in the center is due to the guanosines contributed by the 6SG

experiment which can also form G:U wobble base pairs.

Figure 5: HuR perturbation experiments.

(A) mRNA quantification by RNA sequencing is validated by NanoString nCounter Assay. PolyA+ mRNA levels were inferred from RNA sequencing as FPKM (Fragments/read-pairs Per Kilobase of exon, per Million read-pairs). $\log_{10}$(FPKM) and $\log_{10}$(NanoString counts) for RNA (mock-transfection, DMEM) correlate well (Spearman 0.82, red line: best fit).

(B) HuR target mRNAs are destabilized upon knock down of HuR. Cumulative density fractions of mRNA $\log_2$ fold changes. HuR targets are destabilized upon knock down of HuR. Targets with most binding sites (pink) show the strongest effect. Genes with exclusively intronic binding of HuR (dashed line) are also highly significantly down regulated. **Insert:** Western blot validation of HuR knock down.

(C) pSILAC measures changes in protein synthesis. Cellular proteins incorporate heavy (mock) and medium-heavy (HuR knock down) amino acids on special medium for 24h. The mass-shift allows to measure the difference in newly synthesized protein between normal and HuR depleted cells, using LC-MS/MS.

(D) Protein synthesis of HuR targets is reduced upon HuR knock down. Cumulative density fractions of protein synthesis $\log_2$ fold changes. Exonic and intronic targets of HuR are significantly down regulated after knock down. **Insert**: Western blot validation of HuR knock down.

See also Suppl. Figure S3 and Suppl. Tables ST5, ST6.

Figure 6. Alternative exons with HuR dependent alternative splicing.

Alternatively spliced exons and flanking exons for *ZNF207*, *GANAB*, *PTBP2* and *DST* are shown together with HuR PAR-CLIP clusters (black boxes) and RNA sequencing depth-of-coverage profiles in mock- and siRNA transfected cells. PCR results in a biological replicate are shown in the center and percent spliced in (PSI) values computed from BioAnalyzer quantifications on the right. Error bars represent standard deviation for three technical replicates. While *ZNF207* exon9, *GANAB*

exon6 and *PTBP2* exon10 show an increase in PSI upon HuR depletion, *DST* shows a decrease. See also Suppl. Table ST7.

Figure 7: Suppression of miR-7 biogenesis correlates with HuR expression.

(A) Mature miR-7 is strongly up regulated upon knock down of HuR. Small RNAs were sequenced from mock- and siRNA1 transfected cells. The plot shows the $\log_{10}$ of read counts for all mature miRNAs expressed in HeLa cells.

(B) Validation of miR-7 up regulation by TaqMan quantitative RT-PCR. Endogenous control was U6 snRNA. Error bars indicate 95% confidence interval. The extent of miR-7 up regulation correlates with the strength of the HuR knock down (siRNA2 was less efficient). Expression of let-7b (control) does not change.

(C) HuR binds directly to the last intron of *HNRNPK* that contains *miR-7-1*. The last two exons of *HNRNPK* are separated by a short intron hosting the *miR-7-1* precursor (grey box). Black boxes indicate HuR binding sites. RNA sequencing shows high expression of *HNRNPK* independent of HuR knock down.