nature structural & molecular biology

Staufen1 senses overall transcript secondary structure to regulate translation

Emiliano P Ricci[1–3], Alper Kucukural[1–3], Can Cenik[1–4], Blandine C Mercier[1–3], Guramrit Singh[1–3], Erin E Heyer[1–3], Ami Ashar-Patel[1–3], Lingtao Peng[1–3] & Melissa J Moore[1–3]

Human Staufen1 (Stau1) is a double-stranded RNA (dsRNA)-binding protein implicated in multiple post-transcriptional gene-regulatory processes. Here we combined RNA immunoprecipitation in tandem (RIPiT) with RNase footprinting, formaldehyde cross-linking, sonication-mediated RNA fragmentation and deep sequencing to map Staufen1-binding sites transcriptome wide. We find that Stau1 binds complex secondary structures containing multiple short helices, many of which are formed by inverted Alu elements in annotated 3′ untranslated regions (UTRs) or in 'strongly distal' 3′ UTRs. Stau1 also interacts with actively translating ribosomes and with mRNA coding sequences (CDSs) and 3′ UTRs in proportion to their GC content and propensity to form internal secondary structure. On mRNAs with high CDS GC content, higher Stau1 levels lead to greater ribosome densities, thus suggesting a general role for Stau1 in modulating translation elongation through structured CDS regions. Our results also indicate that Stau1 regulates translation of transcription-regulatory proteins.

© 2013 Nature America, Inc. All rights reserved.

Staufen proteins are highly conserved dsRNA-binding proteins (dsRBPs) found in most bilateral animals[1]. Mammals contain two Staufen paralogs encoded by different loci. Stau1, expressed in most tissues, has a microtubule-binding domain, a dimerization domain and four conserved dsRNA-binding domains (dsRBDs), only two of which (dsRBDs 3 and 4) are necessary for dsRNA binding[2]. Within cells, Stau1 can make direct interactions both with itself and with Stau2, the more tissue-specific paralog[3]. Functionally, Staufen proteins are involved in multiple post-transcriptional regulatory processes. In flies, 3′ UTR–bound Staufen is required for proper localization and translational control of bicoid and prospero mRNAs during oogenesis[4,5]. In mammals, Stau1 has been implicated in mRNA transport to neuronal dendrites[6], regulation of translation via physical interaction with the ribosome[7], a form of translation-dependent mRNA degradation known as Staufen-mediated decay (SMD)[8–11], regulation of stress-granule homeostasis[12], alternative splicing, nuclear export and translation of a gene containing 3′-UTR CUG-repeat expansions[13]. Although Stau1 is not essential for mammalian development, neurons lacking Stau1 have dendritic spine-morphogenesis defects *in vitro*, and knockout mice have locomotor-activity deficits[14].

Crucial for the understanding of how Stau1 regulates gene expression is comprehensive knowledge of its intracellular RNA-binding sites. Although mammalian Stau1- and *Drosophila* Staufen-associated mRNAs were identified by microarray analysis after native RNA immunoprecipitation (RIP)[15–18], those studies were unable to directly map any individual Stau1-binding site, and subsequent bioinformatics analysis yielded no clear consensus for identified mammalian targets[16]. Thus, with the exception of a few well-characterized binding sites validated by mutagenesis[19,20], the exact target sites and RNA structures recognized by mammalian Stau1 remain to be determined. To address this, we here undertook a tandem affinity purification strategy (RIPiT[21]) to map Stau1-binding sites transcriptome wide in human tissue-cultured cells. We also knocked down and overexpressed Stau1 to measure functional consequences on target-mRNA levels and translation efficiency. Our results revealed a new role for Stau1 in regulating translation of GC-rich mRNAs by 'sensing' overall transcript secondary structure.

RESULTS

Transcriptome-wide mapping of Stau1-binding sites

Using the Flp-In system and a tetracycline promoter, we generated HEK293 cells that inducibly expressed a single Flag-tagged copy of either the Stau1 65-kDa spliced isoform (Stau1-WT) or a mutant version (Stau1-mut) containing point mutations in dsRBDs 3 and 4 known to disrupt binding to dsRNA[2] (**Fig. 1a**). Consistently with its propensity to bind dsRNA through the sugar-phosphate backbone[22] and with a previous report suggesting poor UV-cross-linking ability[23], we found that Stau1 cross-linked with very poor efficiency to poly(A)$^+$ RNA upon shortwave UV irradiation of living cells (**Supplementary Fig. 1a**). Therefore we used a RIPiT approach wherein initial immunoprecipitation (IP) with anti-Flag antibody was followed by affinity elution with Flag peptide and then a second IP with a polyclonal anti-Stau1 antibody. RIPiT was performed under two different regimens: (i) To finely-map stable Stau1 footprints, we extensively digested samples

[1]Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts, USA. [2]RNA Therapeutics Institute, University of Massachusetts Medical School, Worcester, Massachusetts, USA. [3]Howard Hughes Medical Institute, University of Massachusetts Medical School, Worcester, Massachusetts, USA. [4]Present address: Department of Genetics, Stanford University School of Medicine, Stanford, California, USA. Correspondence should be addressed to M.J.M. (melissa.moore@umassmed.edu).

Received 8 August; accepted 19 November; published online 15 December 2013; doi:10.1038/nsmb.2739

with RNase I in between native anti-Flag and native anti-Stau1 IPs, generating 30- to 50-nt Stau1-bound RNA fragments (FOOT libraries; **Fig. 1b** and **Supplementary Fig. 1c**). However, many of these short reads derived from Alu repeat elements (described below) and so were not uniquely mappable. Further, under native conditions, Stau1 can make new dsRNA associations after cell lysis (**Supplementary Fig. 1b**). (ii) Therefore, we also subjected cells to formaldehyde cross-linking before lysis, extensively sonicated the lysates to shear long RNAs into 200- to 300-nt fragments (thereby increasing their ability to be mapped) and performed a denaturing anti-Flag IP and then a native anti-Stau1 IP (CROSS libraries; **Fig. 1b** and **Supplementary Fig. 1d**). Cross-linking and subsequent

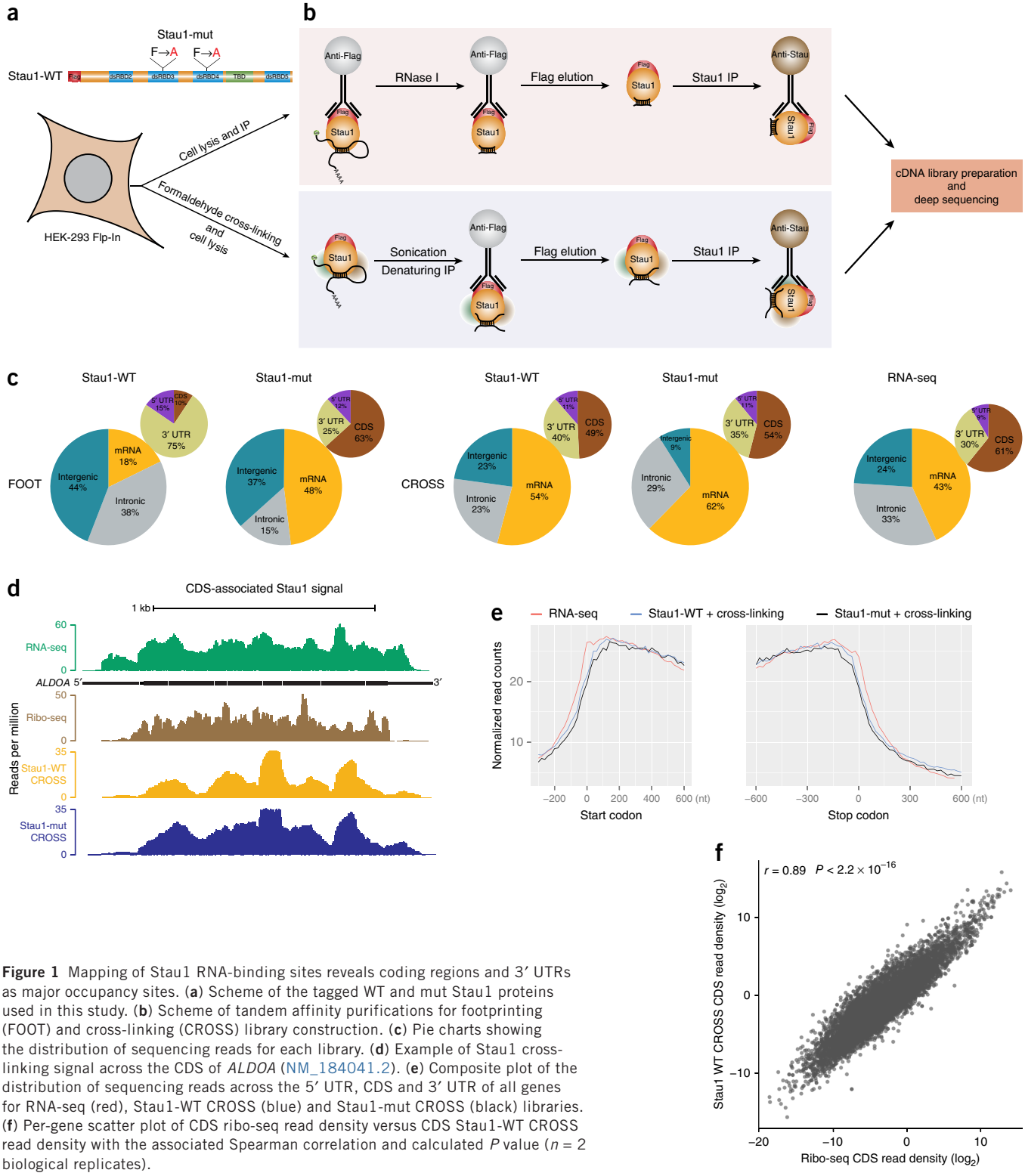

Figure 1 Mapping of Stau1 RNA-binding sites reveals coding regions and 3′ UTRs as major occupancy sites. (**a**) Scheme of the tagged WT and mut Stau1 proteins used in this study. (**b**) Scheme of tandem affinity purifications for footprinting (FOOT) and cross-linking (CROSS) library construction. (**c**) Pie charts showing the distribution of sequencing reads for each library. (**d**) Example of Stau1 cross-linking signal across the CDS of *ALDOA* (NM_184041.2). (**e**) Composite plot of the distribution of sequencing reads across the 5′ UTR, CDS and 3′ UTR of all genes for RNA-seq (red), Stau1-WT CROSS (blue) and Stau1-mut CROSS (black) libraries. (**f**) Per-gene scatter plot of CDS ribo-seq read density versus CDS Stau1-WT CROSS read density with the associated Spearman correlation and calculated *P* value ($n = 2$ biological replicates).

© 2013 Nature America, Inc. All rights reserved.

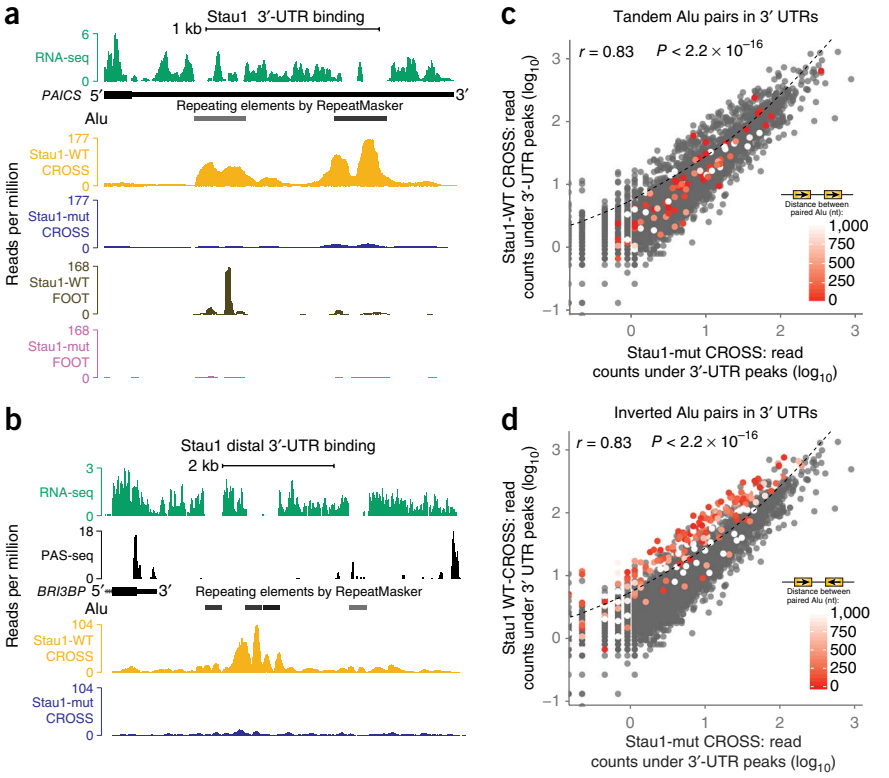

Figure 2 Inverted Alu pairs are an important class of Stau1-binding sites. (**a**,**b**) Distribution of sequencing reads obtained from RNA-seq (green), Stau1-WT CROSS (yellow), Stau1-mut CROSS (blue), Stau1-WT FOOT (brown), Stau1-mut FOOT (violet) and PAS-seq libraries (black) for the 3′ UTR of *PAICS* (NM_001079524) (**a**) and the strongly distal 3′ UTR of *BRI3BP* (NM_080626.5) (**b**). (**c**,**d**) Per-gene scatter plots of Stau1-WT CROSS and Stau1-mut CROSS read counts under called 3′-UTR Stau1-WT CROSS peak positions with associated Spearman correlation and calculated *P* values ($n = 2$ biological replicates). Genes containing a 3′ UTR Alu pair are colored with respect to the distance between each tandem Alu pair (**c**) or inverted Alu pair (**d**). The dashed line corresponds to the 2.7 cutoff in the ratio of Stau1-WT over mut read counts.

denaturation should both preserve weak *in situ* interactions that might otherwise dissociate during sample workup and prevent formation of any new interactions after cell lysis.

We sequenced all libraries constructed by 3′-adaptor ligation to RNA fragments, reverse transcription and circularization on GAII or HiSeq 2000 Illumina platforms and then mapped them to HG18 by using RefSeq gene annotations. Biological replicates of WT and mut CROSS and WT FOOT libraries exhibited extremely high correlations (r >0.98), thus indicating the reproducibility of the approach (**Supplementary Fig. 1e**).

Stau1 associates with translating ribosomes

In contrast to our previous exon junction complex (EJC) RIPiT libraries[24], all Stau1 FOOT libraries (WT and mut) were dominated by rRNA-mapping reads (14–30% versus 74–83%, respectively; **Supplementary Fig. 2**). Further, despite attempts to specifically deplete rRNA fragments during CROSS-library preparation, WT and mut CROSS libraries also contained abundant rRNA-mapping reads (**Supplementary Fig. 2**). These findings are consistent with a previous report that Stau1 cosediments with 60S ribosomal subunits via interactions independent of the functionality of dsRBDs 3 and 4 (ref. 2).

To further investigate this ribosome association, we performed sucrose sedimentation in the presence of inhibitors that either block elongation (cycloheximide) or initiation (harringtonine) (**Supplementary Fig. 3a**). In the presence of cycloheximide, both endogenous Stau1 and Flag–Stau1-WT cosedimented with 60S subunits, 80S monosomes and polysomes, with very little Stau1 observable in ribosome-free fractions at the top of the gradient. However, when lysates were treated with RNase before sedimentation, ~60% of Stau1 sedimented at the top of the gradient, with the remainder cosedimenting with 60S and 80S ribosomes (**Supplementary Fig. 3a**). This suggests that dsRBP-independent interactions with the ribosome are not the sole factor driving Stau1 polysome association. Finally, when translation initiation was blocked with harringtonine and elongating ribosomes allowed to complete translation (i.e., run off the mRNAs) before cell lysis, Stau1 sedimentation mirrored that of RPL26, an integral 60S protein. Both Stau1 and RPL26 rapidly shifted from heavy polysomal to 80S ribosome fractions upon inhibition of translation initiation (**Supplementary Fig. 3b**), suggesting that Stau1 associates with actively translating ribosomes.

Consistently with our ribosome-association data, approximately half of mRNA-mapping WT and mut CROSS reads (49% and 54%, respectively) mapped to coding exons (CDS regions; **Fig. 1c–e**). To test whether these CDS-mapping reads were due to Stau1 association with translating ribosomes, we compared their density to the density of ribosome footprints (ribo-seq; **Fig. 1f**). For both Stau1-WT (**Fig. 1f**; Spearman correlation = 0.89) and Stau1-mut (data not shown), CROSS read density strongly correlated with ribosome density in CDS regions. This correlation held for the entire gene population, thus suggesting that Stau1 generally associates with elongating ribosomes.

In sum, our data indicate that Stau1 is generally associated with the 60S ribosomal subunit, both on and off mRNA. Further, this ribosome association does not require dsRBD functionality but is partially dependent on RNA integrity. Last, Stau1 appears to associate with actively translating, not stalled, ribosomes.

Stau1 binds paired Alu elements in 3′ UTRs

Whereas WT and mut libraries were quite similar in their rRNA content, they were quite different with regard to Alu repeat–mapping reads. Alu repeats are ~300-nt primate-specific mobile elements in the short interspersed nuclear element family; the human genome contains ~1 million Alu elements, primarily in intergenic regions, introns and 3′ UTRs. Reads mapping to Alu repeats constituted 42% and 28% of non-rRNA–mapping reads in WT FOOT and CROSS libraries, respectively, but only 19% and 14% in the corresponding mut libraries (**Supplementary Fig. 2**). Greater Alu enrichment in WT libraries suggested that their interaction depended on Stau1's ability to bind dsRNA. Consistently with this, WT CROSS reads were often highly enriched over and adjacent to closely spaced Alu pairs likely to form dsRNA secondary structures. We detected such Alu-pair Stau1-binding sites on only two large

© 2013 Nature America, Inc. All rights reserved.

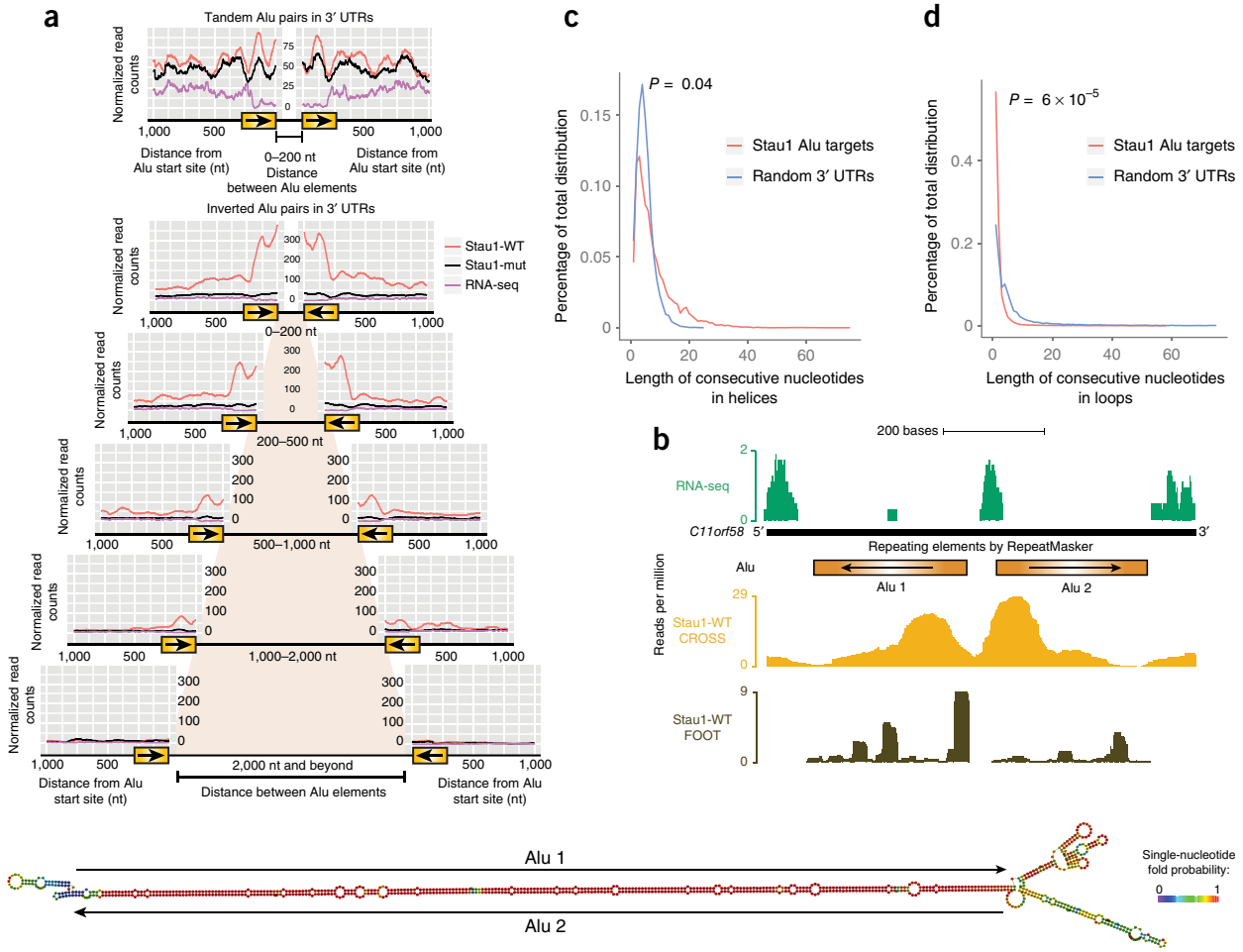

Figure 3 Characterization of the structural features of Stau1 Alu-binding sites. (**a**) Composite plot of Stau1-WT CROSS, Stau1-mut CROSS and RNA-seq read counts around tandem or inverted Alu pairs. Read counts normalized to host gene reads per kilobase per million mapped reads (RPKM) were determined for a region spanning 1 kb up- and downstream of the paired Alu elements separated by the indicated distance ($n = 2$ biological replicates). (**b**) Example of a 3′-UTR inverted Alu-binding site in *C11orf58* (NM_014267.5) showing read counts per million mapped (pmm) reads for RNA-seq, Stau1-WT CROSS and Stau1-mut CROSS libraries. The centroid secondary structure for this Alu pair predicted with the Vienna folding package[47] is shown below center. (**c**,**d**) Length distribution of predicted helices (**c**) and loop (**d**) regions within secondary structures of 3′-UTR inverted Alu pairs or 3′-UTR sequences of identical size randomly picked from nontarget genes. *P* values corresponding to the comparison of helix and loop length distributions between Stau1 Alu targets and random 3′ UTRs were calculated with the Wilcoxon rank-sum test ($n = 2$ biological replicates).

intergenic noncoding RNAs (NR_026757 and NR_026999) and minimally in introns (**Supplementary Fig. 4**). Conversely, they were highly enriched in 3′ UTRs (**Fig. 2a** and **Supplementary Fig. 4**) and in select 'intergenic' regions immediately 3′ to annotated 3′ UTRs (**Fig. 2b** and **Supplementary Fig. 4**). Polyadenylation-site sequencing (PAS-seq) revealed the intergenic regions to represent strongly distal 3′ UTRs[25] (**Fig. 2b**). Overall, we detected 515 strongly distal 3′ UTRs enriched for Stau1-WT CROSS reads (**Supplementary Table 1**), most of which contained multiple Alu pairs.

To identify those 3′ UTRs most enriched for dsRBD-dependent Stau1 binding, we called peaks in the WT CROSS libraries by using ASPeak (an expression-sensitive peak-calling algorithm[26]). We then compared, for each gene, the cumulative read counts under peak positions in the WT and mut CROSS libraries (**Fig. 2c**,**d**). Overall, the data sets were highly correlated ($r = 0.83$). Nonetheless, an outlier population ($n = 574$; **Supplementary Table 2**) exhibited much higher cross-linking (by a factor of 2.7) in WT than in mut (**Fig. 2c**,**d**); these outliers constitute a set of high-confidence 3′ UTRs displaying dsRNA-dependent Stau1 binding.

We next investigated the structural features of these targets. To identify those containing Alu pairs, we wrote an algorithm to identify, transcriptome wide, pairs of full-length Alu elements in the same (tandem) or opposite (inverted) orientation. Overlaying the inter-Alu distance for tandem Alu pairs on the WT versus mut CROSS scatter plot (**Fig. 2c**) revealed no specific relationship between tandem pairs and Stau1 cross-linking. However, the inverted Alu-pair overlay revealed a striking coincidence with the above outlier population (**Fig. 2d**). Further, inverted Alu elements separated by the least distance were the most outlying (**Fig. 2d**).

We confirmed the inverse relationship between dsRBD-dependent Stau1 cross-linking efficiency and inverted-pair inter-Alu distance in composite plots (**Fig. 3a** and **Supplementary Fig. 5**). We found similar, but less striking, results for inverted pairs containing partial Alu elements and for inverted pairs in introns and strongly distal 3′ UTRs (**Supplementary Fig. 5a**). As expected, we observed no specific mapping of WT reads on tandem Alu pairs or mapping of mut reads on Alu pairs in either orientation (**Fig. 3a**). The inverse correlation between Stau1-WT cross-linking efficiency and

© 2013 Nature America, Inc. All rights reserved.

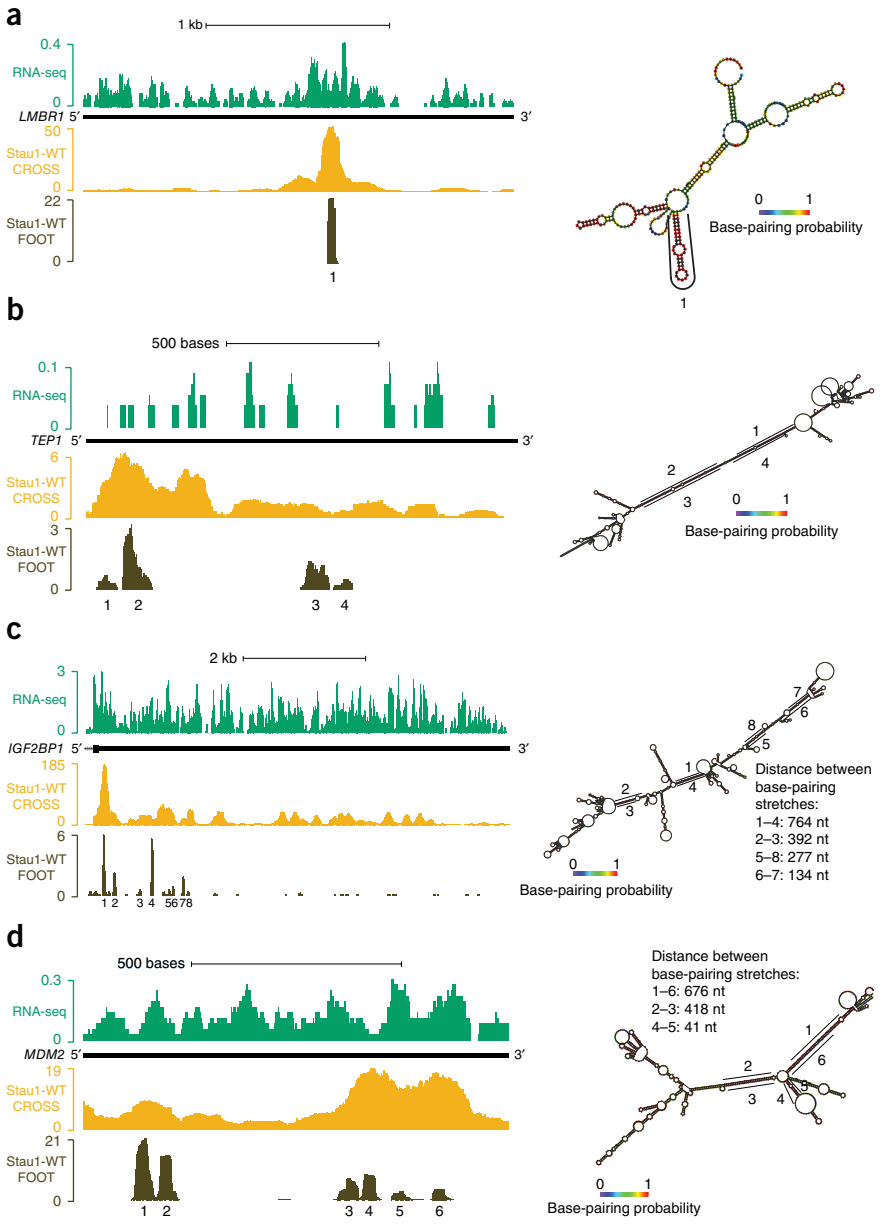

Figure 4 Examples of 3′-UTR non-Alu Staufen-binding sites. (**a**–**d**) Read distributions for RNA-seq (green), Stau1-WT CROSS (yellow) and Stau1-WT FOOT (brown) libraries on the 3′ UTRs of *LMBR1* (NM_022458.3) (**a**), *TEP1* (NM_007110.4) (**b**), *IGF2BP1* (NM_006546.3) (**c**) and *MDM2* (NM_002392) (**d**) (left) together with the corresponding centroid secondary structure colored for base-pairing probability as predicted by the Vienna folding package[47] (right). Numbers below the Stau1 WT FOOT track and in the predicted secondary structure correspond to Stau1-binding sites.

inverted-pair inter-Alu distance is consistent with the expectation that secondary-structure formation should inversely correlate with pairing-partner distance.

Because intact Alu elements are ~300 nt, inverted pairs containing full-length Alu elements could potentially form very long helices. However, individual elements in pairs exhibiting the highest WT cross-linking signal were often from different Alu families unlikely to be fully complementary. Consistently with this, *in silico* folding of an inverted Alu pair exhibiting one of the strongest Stau1-WT occupancies suggests the presence of many short helices interrupted by small loops (**Fig. 3b**). To assess the generalizability of this, we folded *in silico* all-full-length, 3′-UTR inverted Alu pairs highly enriched for Stau1-WT cross-linking and compared them to 3′-UTR sequences of similar length randomly chosen from nontarget genes. Histograms of predicted helix and loop lengths (**Fig. 3c**,**d**) revealed that Stau1-interacting Alu pairs tend to form structures with multiple helices containing <30 interrupted base pairs, spaced by 2- to 10-nt loops. Conversely, nontarget 3′ UTRs were predicted to have significantly shorter paired stretches (Wilcoxon rank-sum test, $P = 0.04$) interrupted by longer loops (Wilcoxon rank-sum test, $P = 6 \times 10^{-5}$).

Non-Alu 3′-UTR targets

Among the outlier population in **Figure 2d**, 201 of 574 contained clearly identifiable inverted Alu pairs (Alu targets), and 373 of 574 did not (non-Alu targets; **Supplementary Table 2**). Many of these non-Alu targets had clearly defined WT footprints in regions with high base-pairing probability (**Fig. 4**). A few contained a single strong footprint corresponding to a short stem-loop structure (**Fig. 4a**); others resembled inverted Alu pairs with many consecutive helices separated by short loops (**Fig. 4b**). The largest set, however, consisted of complex structures covering a few hundred nucleotides within which Stau1 footprints could be observed on multiple 7- to 40-bp helices (**Fig. 4c**,**d**). WT footprints were also present on the Arf1 3′ UTR, for which the precise Stau1-binding site was previously mapped by mutagenesis (**Supplementary Fig. 6a**)[20].

Comparison of FOOT and CROSS reads mapping to individual 3′ UTRs revealed that CROSS reads generally extended over much more of the 3′ UTR than did FOOT reads (for example, **Fig. 4c**). We could even observe extensive CROSS read coverage for many 3′ UTRs having no detectable footprints (**Supplementary Fig. 6b**). Greater abundance of such CROSS reads in WT libraries than in mut libraries indicated that they depended on Stau1's ability to bind dsRNA. This suggested that the kinetically stable Stau1-binding sites revealed by native footprinting represent only a small subset of RNA-interaction sites occurring within cells. Supporting the notion of many low-affinity Stau1-interaction sites *in vivo*, we observed a strong correlation ($r = 0.63$, $P < 2.2 \times 10^{-16}$) over all expressed genes between average per-nucleotide predicted secondary-structure strength (ΔG of the minimum free-energy structure/3′-UTR length) and the ratio of total WT/mut CROSS reads per 3′ UTR (**Fig. 5a**). We observed a similarly strong correlation ($r = 0.55$, $P < 2.2 \times 10^{-16}$) between this ratio and 3′-UTR GC content in all expressed genes (**Fig. 5b**).

We conclude that some Stau1-binding sites in 3′ UTRs consist of highly defined structures containing multiple short helices

© 2013 Nature America, Inc. All rights reserved.

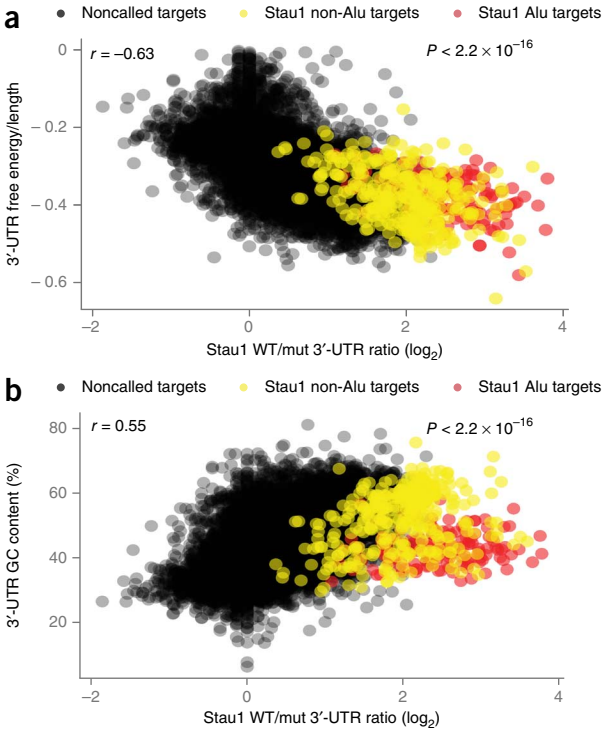

Figure 5 Stau1 binding to 3′ UTRs correlates with GC content and predicted secondary-structure free energy. (**a**) Per-gene scatter plot of 3′-UTR predicted secondary-structure free energy normalized by the length of the 3′ UTR (kcal/mol/nucleotide) against Stau1 WT/mut 3′-UTR ratio (log_2). (**b**) Per-gene scatter plot of 3′-UTR GC content (%) against Stau1 WT/mut 3′-UTR ratio (log_2) with associated Spearman correlation and calculated *P* values. Red and yellow dots correspond to called Alu and non-Alu binding sites, respectively ($n = 2$ biological replicates).

to which Stau1 binding is kinetically stable. Other binding sites, however, are more kinetically labile, and the extent of Stau1 occupancy on these sites is a function of overall 3′-UTR secondary structure–forming propensity, often driven by high GC content.

dsRBD-dependent binding of Stau1 to CDS regions

As previously discussed, both WT and mut CROSS reads mirrored ribosome density across CDS regions transcriptome wide (**Fig. 1f**). WT and mut reads were also similarly distributed relative to start and stop codons (**Fig. 1e**). In contrast to the general population, however, our 373 non-Alu 3′-UTR target genes had significantly greater CROSS reads in CDS regions for WT than for mut (**Fig. 6a**). This strong relationship between 3′ UTR and CDS WT/mut cross-linking initially suggested to us that dsRBD-dependent Stau1 binding within the 3′ UTR increases its association with CDS-bound ribosomes. Consistently with this, the correlation between preferential WT cross-linking in 3′-UTR and CDS regions held true for the entire mRNA population (**Fig. 6b**; $r = 0.61$, $P < 2.2 \times 10^{-16}$), with our identified 3′-UTR target genes simply being strongly skewed toward the higher end of both ratios. However, we also found that predicted per-nucleotide secondary structure–forming propensity and GC content were strongly correlated ($r = 0.55$ and 0.73, respectively, $P < 2.2 \times 10^{-16}$) between the 3′-UTR and CDS regions of individual genes (**Fig. 6c**,**d**); that is, the genes with high 3′-UTR secondary structure–forming propensity and GC content also tend to have high CDS secondary

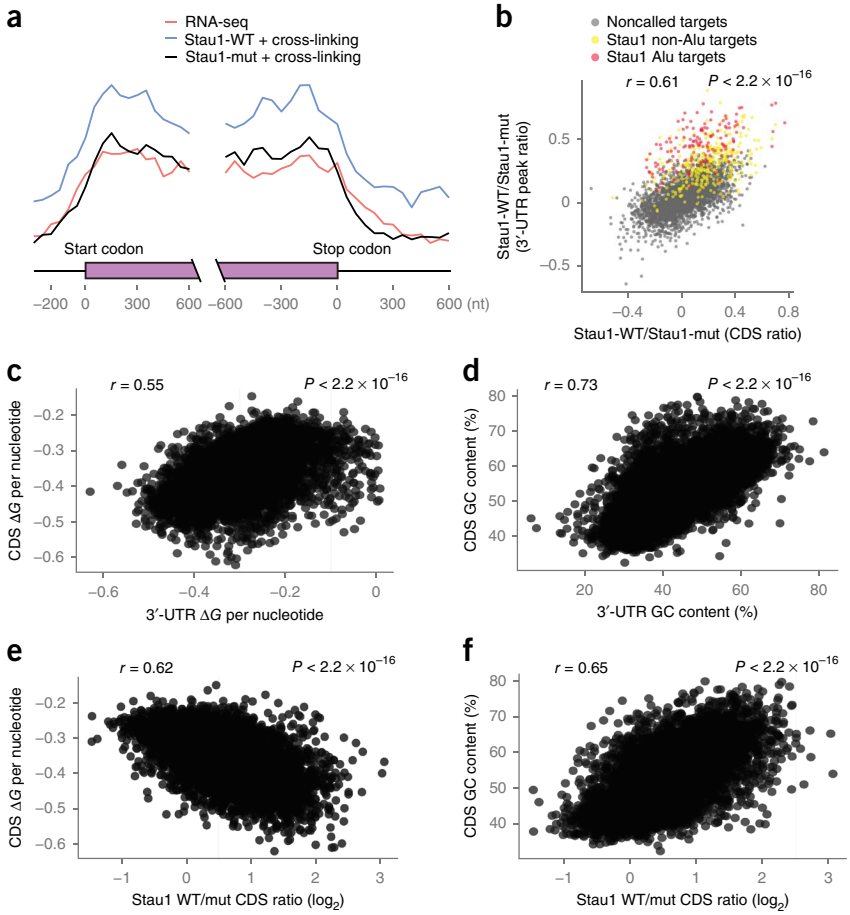

Figure 6 Stau1 occupancy on the CDS strongly correlates with GC content and predicted secondary-structure free energy. (**a**) Composite plot of the distribution of sequencing reads across the 5′ UTR, CDS and 3′ UTR of called Stau1-target genes for RNA-seq (red), Stau1-WT CROSS (blue) and Stau1-mut CROSS (black) libraries. (**b**) Per-gene scatter plot (log_{10}) of total CDS Stau1-WT CROSS read counts/total CDS Stau1-mut CROSS reads counts (CDS ratio) versus the ratio of Stau1-WT CROSS and Stau1-mut CROSS read counts under called 3′-UTR Stau1-WT CROSS peak positions (3′-UTR peak ratio). Red and yellow dots correspond to called Alu- and non-Alu–binding sites, respectively. (**c**) Per-gene scatter plot of 3′ UTR against CDS predicted secondary-structure free energy normalized by the length of the 3′ UTR (kcal/mol/nucleotide). (**d**) Per-gene scatter plot of 3′ UTR against CDS GC content (%). (**e**) Per-gene scatter plot of CDS predicted secondary-structure free energy normalized by the length of the CDS (kcal/mol/nucleotide) against Stau1 WT/mut 3′-UTR ratio (log_2). (**f**) Per-gene scatter plot of CDS GC content (%) against Stau1 WT/mut 3′-UTR ratio (log_2). All correlation coefficients and *P* values were calculated with the Spearman rank correlation ($n = 2$ biological replicates) throughout figure.

© 2013 Nature America, Inc. All rights reserved.

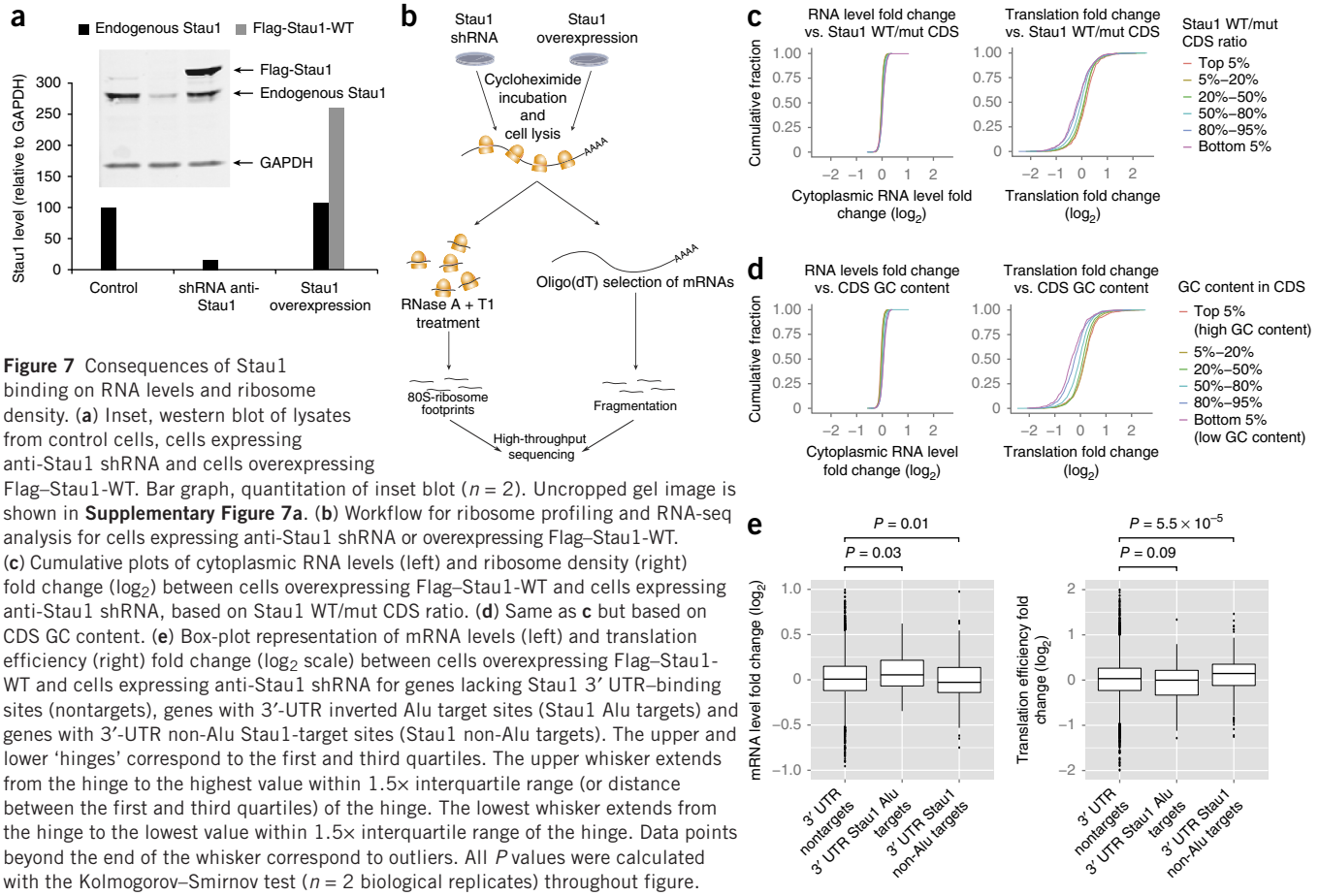

Figure 7 Consequences of Stau1 binding on RNA levels and ribosome density. (**a**) Inset, western blot of lysates from control cells, cells expressing anti-Stau1 shRNA and cells overexpressing Flag–Stau1-WT. Bar graph, quantitation of inset blot ($n = 2$). Uncropped gel image is shown in **Supplementary Figure 7a**. (**b**) Workflow for ribosome profiling and RNA-seq analysis for cells expressing anti-Stau1 shRNA or overexpressing Flag–Stau1-WT. (**c**) Cumulative plots of cytoplasmic RNA levels (left) and ribosome density (right) fold change ($\log_2$) between cells overexpressing Flag–Stau1-WT and cells expressing anti-Stau1 shRNA, based on Stau1 WT/mut CDS ratio. (**d**) Same as **c** but based on CDS GC content. (**e**) Box-plot representation of mRNA levels (left) and translation efficiency (right) fold change ($\log_2$ scale) between cells overexpressing Flag–Stau1-WT and cells expressing anti-Stau1 shRNA for genes lacking Stau1 3′ UTR–binding sites (nontargets), genes with 3′-UTR inverted Alu target sites (Stau1 Alu targets) and genes with 3′-UTR non-Alu Stau1-target sites (Stau1 non-Alu targets). The upper and lower 'hinges' correspond to the first and third quartiles. The upper whisker extends from the hinge to the highest value within 1.5× interquartile range (or distance between the first and third quartiles) of the hinge. The lowest whisker extends from the hinge to the lowest value within 1.5× interquartile range of the hinge. Data points beyond the end of the whisker correspond to outliers. All *P* values were calculated with the Kolmogorov–Smirnov test ($n = 2$ biological replicates) throughout figure.

structure–forming propensity and GC content. Consistently with this, preferential WT cross-linking in CDS regions strongly correlated with both predicted CDS secondary structure and GC content (**Fig. 6d** and **e**, $r = 0.62$ and 0.65, respectively, $P < 2.2 \times 10^{-16}$) and the 5′ UTR ($r = 0.2$; data not shown). These analyses suggest that enhanced Stau1-WT binding within CDS regions is primarily driven by GC content and secondary structure–forming propensity of the CDS itself, rather than by interactions of 3′ UTR–bound Stau1 with CDS-bound ribosomes.

From the above data, we conclude that Stau1 interacts to varying extents with the CDS and 3′-UTR regions of all cellular mRNAs in a manner dependent on their secondary structure–forming propensities. Further, the observed correlation between dsRBD-dependent Stau1 occupancy in CDS and 3′-UTR regions mainly reflects similar GC content between the CDS and 3′ UTR in individual genes rather than any direct effect of 3′-UTR binding on CDS binding. Instead, Stau1-WT occupancy on CDS regions appears to be driven by a combination of direct interactions with CDS secondary structures and its dsRBD-independent association with actively translating ribosomes.

Gene ontology analysis

To assess whether any particular gene classes were specifically enriched for dsRBD-dependent Stau1 binding, we performed gene ontology analysis using GeneCodis[27–29] (**Supplementary Table 3**). We obtained the most significant associations for the 469 genes having the highest WT/mut CDS cross-linking ratios (>1.9) and the 515 genes exhibiting high WT cross-linking to strongly distal 3′ UTRs. Both sets were highly enriched in transcription-regulatory proteins ($P = 7.1 \times 10^{-13}$ and $P = 1.1 \times 10^{-13}$, respectively). Among transcription-factor types, C2H2 zinc-finger proteins were the most enriched ($P = 4.5 \times 10^{-6}$), with homeobox and high-mobility group (HMG) proteins following close behind ($P = 1.3 \times 10^{-5}$ and 6.9×10^{-5}, respectively). Consistently with the strong correlation between Stau1 CDS and 3′-UTR occupancy, transcription-regulatory proteins were also highly enriched among our 373 non-Alu 3′-UTR targets ($P = 6.9 \times 10^{-7}$). Thus Stau1 may have a role in post-transcriptional regulation of transcription factors. Also enriched in the non-Alu and extended 3′-UTR targets ($P = 0.001$ and $P = 5.0 \times 10^{-5}$, respectively), but not in the 469 high CDS targets, were proteins involved in cell-cycle control.

Functional consequences of varying Stau1 protein levels

To directly test the functional consequences of Stau1 binding, we next varied intracellular Stau1 concentration (**Fig. 7a**,**b**). Transduction of HEK293 FLP-in cells with a lentivirus expressing an anti-Stau1 short hairpin RNA (shRNA) stably reduced endogenous Stau1 to ~20% normal levels (UNDER, **Fig. 7a**). Incubation of our stably integrated Flag–Stau1-WT cells overnight (16 h) with a high level of doxycycline induced transgene overexpression by 300–400% relative to endogenous Stau1 (OVER). We then assessed effects of Stau1 depletion or overexpression by preparing cytoplasmic poly(A)$^+$ RNA-seq and ribo-seq libraries.

RNA-seq and ribo-seq read counts on individual genes were highly correlated both between biological replicates ($r \geq 0.98$; E.P.R., unpublished data) and between UNDER and OVER samples ($r \geq 0.98$; **Supplementary Fig. 7**). Other than *STAU1* itself, there were no clear outlier genes between UNDER and OVER conditions for either

© 2013 Nature America, Inc. All rights reserved.

RNA-seq or ribo-seq. Further, no significant changes in alternative-splicing patterns could be detected (data not shown); thus, at least in HEK cells, binding of Stau1 in introns is of little apparent consequence for pre-mRNA splicing. However, small negative correlations between RNA levels and Stau1 levels could be detected when we ordered transcripts by CDS GC content or preferential Stau1-WT CDS cross-linking (**Fig. 7c**,**d**). That is, transcripts with high CDS GC content (which drives greater Stau1-WT CDS binding) exhibited slightly lower cytoplasmic mRNA abundances when Stau1 was overexpressed than underexpressed.

The strongest observable effect of varying Stau1 concentration was on ribosome occupancy. Cumulative histograms revealed positive relationships between ribosome occupancy and both Stau1-WT CDS cross-linking and CDS GC content across the entire transcriptome (**Fig. 7c**,**d**; Spearman correlation $r = 0.21$ and $r = 0.34$, respectively, $P < 2.2 \times 10^{-16}$). That is, genes with higher Stau1 CDS occupancy and higher CDS GC content exhibited increased ribosome occupancy upon Stau1 overexpression compared to Stau1 knockdown; conversely, genes with lower Stau1 CDS occupancy and CDS GC content exhibited decreased ribosome occupancy upon Stau1 overexpression compared to Stau1 knockdown. This suggests that higher Stau1 protein levels increase ribosome occupancy on high-GC-content transcripts at the expense of low-GC-content transcripts. Ontology analysis of the 400 genes exhibiting the greatest increase in ribosome occupancy between UNDER and OVER conditions revealed significant enrichments for transcription-regulatory proteins ($P = 0.004$) and zinc-binding proteins ($P = 1.1 \times 10^{-6}$; **Supplementary Table 3**), the same terms obtained above for genes exhibiting the highest CDS and extended 3′-UTR Stau1 occupancies.

Although we observed the strongest effects of varying Stau1 protein levels for genes with high Stau1 CDS occupancy, we also examined the effects of Stau1 over- and underexpression on our 3′-UTR non-Alu and Alu target sets. Ribosome occupancy increased slightly on non-Alu 3′-UTR targets (10% change from UNDER to OVER; $P = 0.00005$) when compared to the total population, whereas their mRNA levels decreased slightly (−2% change from UNDER to OVER; $P = 0.01$). Thus, non-Alu targets behaved like high-GC-content mRNAs. Conversely, Alu targets exhibited no significant change in ribosome occupancy, but their cytoplasmic mRNA levels increased upon Stau1 upregulation (+8% change from UNDER to OVER; $P = 0.03$; **Fig. 7e**). Therefore, mRNAs containing 3′-UTR inverted Alu pairs behave differently from other cellular mRNAs in response to Stau1 abundance. For the strongly distal 3′-UTR Stau1-binding sites, we detected no significant effect of Stau1 expression on either mRNA levels or ribosome occupancy (E.P.R., unpublished data), possibly because such isoforms represent only a minor fraction of transcripts from individual loci.

To confirm that changes in Stau1 levels are of little consequence for levels of mRNAs with 3′ UTR–binding sites, we performed quantitative reverse-transcription PCR (qRT-PCR) on several Alu and non-Alu 3′-UTR target mRNAs including Arf1, a previously identified SMD target (**Supplementary Fig. 8**). Consistently with our RNA-seq results, neither downregulation nor overexpression of Stau1 had a significant impact on the abundance of tested targets (**Supplementary Fig. 8a**). Experiments performed in two other cell lines (Huh7 and SK-Hep1) in which either Stau1 or Stau2 or both were downregulated yielded similar results (**Supplementary Fig. 8b**,**c**).

Taken together, our results indicate that Stau1 binding to the CDS results in increased ribosome occupancy and in decreased mRNA levels proportionate to both the amount of bound Stau1 and the GC content of the target mRNA. Further, at least in the cell lines we tested, Stau1 binding within the 3′ UTR appears to be of little or no consequence for translation efficiency or steady-state mRNA levels.

DISCUSSION

Like many RNA-binding factors, the *Drosophila* and mammalian Staufen proteins have been implicated in multiple post-transcriptional processes including alternative splicing[13], RNA localization[4,6,30–32], translational activation[7] and translation-dependent mRNA decay[8–11,14,20,33,34]. Which activity is observed depends on the cellular context, the identity of the bound RNA and the location of the binding site on the target RNA. Many of Staufen's previously documented activities parallel those of the EJC[7–11,19,20,33,35–38]. To better understand EJC function, we recently determined the complete EJC RNA-binding landscape in HEK293 cells[24]. Here we undertook the same analysis for Stau1.

Up to now, confirmed Staufen-binding sites were limited to a few well-characterized structures[19,20]. Broader identification of Staufen-associated mRNAs has been attempted in various organisms by combination of native RIP protocols with microarray analyses[15–18]. Unfortunately, however, such methodologies have yielded no consensus as to general features of Staufen targets. One recent study of Staufen-associated mRNAs from *Drosophila* oocytes reported enrichment of three different secondary-structural motifs that might explain Staufen binding specificity in flies[16]. However, the authors were unable to identify similar structural motifs among human Staufen-associated mRNAs from available native mammalian Stau1 and Stau2 RIP microarray data[15]. We show here that human Stau1 generally associates with actively translating ribosomes; therefore, it is impossible to discriminate between sites of direct Stau1-mRNA interaction via dsRNA binding and sites of indirect Stau1-mRNA association via elongating ribosomes without some sort of footprinting approach. Further, because of (i) Stau1's strong ribosome association, (ii) the prevalence of kinetically labile Stau1-binding sites *in vivo* and (iii) Stau1's ability to form new interactions with dsRNA after cell lysis, native RIP experiments are likely to be biased toward both highly translated mRNAs and RNAs containing the most stable sites of direct Stau1-dsRNA interaction. Our experimental design, which combined formaldehyde cross-linking and fragmentation of Stau1-associated RNAs, using both WT and mut proteins, allowed us to both avoid binding-site reassortment after cell lysis and discriminate between binding modes that do or do not require Stau1 dsRBD functionality.

The majority of non-rRNA reads in our cross-linked libraries mapped sense to 3′ UTRs and CDS regions. Within 3′ UTRs, we identified numerous high-occupancy Stau1-binding sites composed of either inverted Alu pairs (Alu targets) or sequences with extremely high secondary structure–forming propensity (non-Alu targets). Observable native Stau1 footprints showed that these structures often consist of several closely spaced helices separated by short loops. Bioinformatics analysis of the footprints, however, failed to identify any particular enriched motif (A.K. and E.P.R., unpublished data), results consistent with the idea that Staufen recognizes dsRNA in a sequence-independent manner[39–41].

Unexpectedly, in addition to detecting strong binding to large RNA secondary structures, we also detected extensive dsRBD-dependent Stau1 cross-linking extending throughout the entire length of 3′ UTRs and CDS regions. This cross-linking strongly correlated with both GC content and per-nucleotide predicted secondary-structure strength. Because GC content in CDS and 3′-UTR regions also correlate, mRNAs exhibiting preferential Stau1-WT 3′-UTR cross-linking also tend to exhibit preferential Stau1-WT CDS cross-linking. Inverted Alu pairs, the 3′ UTRs containing them and their associated CDS regions,

© 2013 Nature America, Inc. All rights reserved.

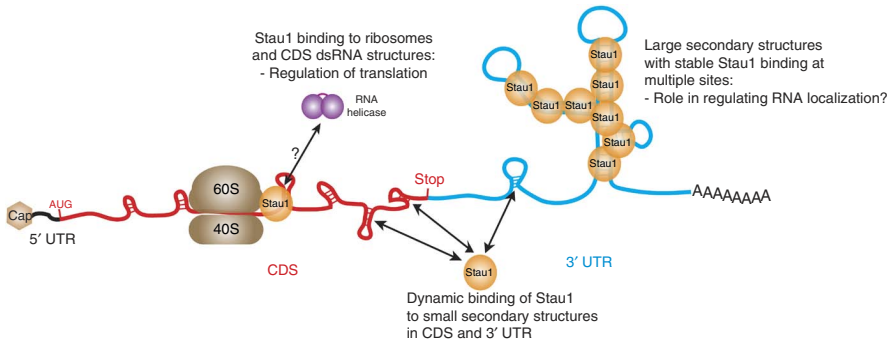

Figure 8 Model of Stau1 RNA binding and its functional role in translation. Stau1 interacts with both actively translating ribosomes and secondary structures in CDS and 3′-UTR regions. Some 3′ UTRs contain highly complex secondary structures (for example, inverted Alu pairs) that serve as kinetically stable Stau1-binding sites. However, Stau1 also makes transient interactions with smaller secondary structures throughout CDS and 3′-UTR regions. Formation of these structures is a function of overall CDS and 3′-UTR GC content. Whereas interaction of Stau1 with 3′-UTR Alu pairs has a small positive effect on cytoplasmic mRNA levels, high Stau1 CDS occupancy both increases ribosome density and slightly decreases cytoplasmic mRNA levels.

© 2013 Nature America, Inc. All rights reserved.

however, exhibit average GC content. Despite high Stau1 occupancy on such 3′ UTRs, their CDS occupancies are close to levels that would be expected from their GC content alone. We therefore conclude that the strongest feature driving dsRBD-dependent Stau1 binding within CDS regions is the secondary structure–forming propensity of the CDS itself. Thus dsRBD-dependent Stau1 binding to 3′ UTRs appears to be functionally uncoupled from dsRBD-dependent Stau1 binding to CDS regions, with the correlation between 3′ UTR and CDS cross-linking driven primarily by GC-content similarity.

Our results suggest that endogenous Stau1 RNA targets can be divided into two broad classes dependent on their structural topology. One class corresponds to stable RNA secondary structures such as inverted Alu pairs and other sequences with extremely high secondary structure–forming propensity. Such elements are capable of simultaneously binding multiple Staufen molecules whose association may be further stabilized by multimerization. Close association of multiple Staufen-binding sites would assure continuous Staufen occupancy even though individual protein molecules might come and go. It is of note that we generally detected such binding sites in annotated 3′ UTRs and extended 3′ UTRs, the latter being particularly rich in inverted Alu pairs. Recently, extended 3′ UTRs were shown to be especially prevalent in the brain. Because Stau1 is known to have a role in dendritic mRNA targeting, these stable RNA secondary structures with their long-lived Stau1 associations could well be the functional binding sites through which Stau1 promotes proper subcellular mRNA localization in neurons.

The second class consists of smaller and more labile secondary structures as might occur in GC-rich CDS regions. Here our data indicate that transient Stau1 binding, perhaps by Stau1 molecules simultaneously interacting with elongating ribosomes, has a role in regulating translation. We arrived at this conclusion by analyzing cytoplasmic poly(A)$^+$ RNA-seq and ribo-seq data from cells under- and overexpressing Stau1. This allowed us to assess the effects of varying intracellular Stau1 concentration on both cytoplasmic mRNA levels and ribosome occupancy. Observable changes in mRNA levels were extremely subtle. Consistently with recent data indicating that Stau1 binding to mRNAs containing inverted Alu elements enhances their nucleocytoplasmic export[42], we did observe a small positive effect of increasing Stau1 on cytoplasmic mRNA levels for our 3′-UTR Alu targets. Conversely, for all other sets of mRNAs exhibiting preferential Stau1-WT cross-linking, Stau1 levels negatively influenced cytoplasmic mRNA levels proportionately to CDS Stau1 occupancy but not to 3′-UTR occupancy. Thus we could find little evidence for SMD driven by 3′ UTR–bound Stau1, either over the entire mRNA population or for previously identified SMD targets. Instead, higher Stau1 levels led to a preferential increase in ribosome density on high-GC-content mRNAs.

We propose a model (**Fig. 8**) based on these findings, wherein ribosome-bound Stau1 molecules transiently interact with short dsRNA helices throughout the CDS and 3′ UTR. In the CDS, such interactions somehow serve to increase ribosome density. Because Stau1 interacts with actively translating ribosomes, the increase in ribosome density may reflect increased translation efficiency. One possibility is that Stau1 helps ribosomes elongate through otherwise inhibitory secondary structures by recruiting factors such as RNA helicase A (RHA or DHX9) to disrupt them. RHA is a positive regulator of translation on mRNAs containing 5′-UTR secondary structures[43] and is known to copurify with Stau1 (ref. 44 and E.P.R., unpublished data). Another abundant translational-regulatory protein that binds ribosomes and cross-links throughout CDS regions is the fragile X protein, FMRP[45]. FMRP, however, is a negative regulator of translation. Whereas deletion of either FMRP or Stau1 causes neurological defects, the phenotypes are opposite: absence of FMRP leads to dendritic spine overgrowth[46], whereas absence of Stau1 results in fewer spines[14]. Thus it is possible that FMRP and Stau1 have opposing roles in synaptic protein production, with FMRP inhibiting translation and Stau1 promoting it.

Finally, mRNAs encoding transcription-regulatory proteins were recently reported as being enriched in *Drosophila* Staufen RIP samples[16]. Consistently with this, we found that mRNAs encoding transcription factors of the C2H2 zinc-finger, HMG and homeobox families were highly enriched among mRNAs exhibiting the highest preferential 3′-UTR and CDS Stau1-WT occupancy. Transcription factors and zinc-binding proteins were also highly enriched among the mRNAs whose ribosome density was most positively affected by Stau1 protein levels. Thus Stau1 may have a previously unrecognized role in the translational regulation of transcription-regulatory proteins.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. High-throughput sequencing data corresponding to native and cross-linked Stau1 RIPiT experiments as well as PAS-seq, RNA-seq and ribo-seq have been deposited in the GEO database under accession number GSE52447.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We would like to acknowledge M. Garber, A. Bicknell, A. Noma and J. Braun for comments on the manuscript and the University of Massachusetts Medical School Deep Sequencing Core for technical advice. We thank P.S. Chen for technical help in preparing plasmids and cell lines. We also thank H. Ozadam for technical advice in using RNA structure–prediction tools. Finally, we thank M. Janas, R. Lakshmi

and D. Morrissey from Novartis for kindly sharing total RNA from Huh7 and HepG2 cells upon Stau1 and Stau2 knockdown. M.J.M. is supported as a Howard Hughes Medical Institute Investigator.

AUTHOR CONTRIBUTIONS

E.P.R. and M.J.M. conceived the study, designed the experiments and wrote the manuscript. E.P.R. performed the experiments. A.K. conducted most bioinformatics analyses. C.C. designed and implemented the ribo-seq analysis. E.P.R. and A.K. designed and performed GC-content and secondary structure–prediction analyses. B.C.M. and G.S. contributed with tandem-affinity purification of Stau1 complexes. E.E.H. contributed with cDNA library preparation for high-throughput sequencing. A.A.-P. prepared PAS-seq cDNA libraries. L.P. participated in quantitative PCR analysis.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

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

1. Kerner, P., Degnan, S.M., Marchand, L., Degnan, B.M. & Vervoort, M. Evolution of RNA-binding proteins in animals: insights from genome-wide analysis in the sponge *Amphimedon queenslandica*. *Mol. Biol. Evol.* **28**, 2289–2303 (2011).
2. Luo, M., Duchaîne, T.F. & Desgroseillers, L. Molecular mapping of the determinants involved in human Staufen-ribosome association. *Biochem. J.* **365**, 817–824 (2002).
3. Martel, C. *et al.* Multimerization of Staufen1 in live cells. *RNA* **16**, 585–597 (2010).
4. St Johnston, D., Beuchle, D. & Nüsslein-Volhard, C. Staufen, a gene required to localize maternal RNAs in the *Drosophila* egg. *Cell* **66**, 51–63 (1991).
5. Ferrandon, D., Koch, I., Westhof, E. & Nüsslein-Volhard, C. RNA-RNA interaction is required for the formation of specific bicoid mRNA 3′ UTR–STAUFEN ribonucleoprotein particles. *EMBO J.* **16**, 1751–1758 (1997).
6. Köhrmann, M. *et al.* Microtubule-dependent recruitment of Staufen-green fluorescent protein into large RNA-containing granules and subsequent dendritic transport in living hippocampal neurons. *Mol. Biol. Cell* **10**, 2945–2953 (1999).
7. Dugré-Brisson, S. *et al.* Interaction of Staufen1 with the 5′ end of mRNA facilitates translation of these RNAs. *Nucleic Acids Res.* **33**, 4797–4812 (2005).
8. Kim, Y.K., Furic, L., Desgroseillers, L. & Maquat, L.E. Mammalian Staufen1 recruits Upf1 to specific mRNA 3′UTRs so as to elicit mRNA decay. *Cell* **120**, 195–208 (2005).
9. Gong, C. & Maquat, L.E. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3′ UTRs via Alu elements. *Nature* **470**, 284–288 (2011).
10. Gleghorn, M.L., Gong, C., Kielkopf, C.L. & Maquat, L.E. Staufen1 dimerizes through a conserved motif and a degenerate dsRNA-binding domain to promote mRNA decay. *Nat. Struct. Mol. Biol.* **20**, 515–524 (2013).
11. Park, E., Gleghorn, M.L. & Maquat, L.E. Staufen2 functions in Staufen1-mediated mRNA decay by binding to itself and its paralog and promoting UPF1 helicase but not ATPase activity. *Proc. Natl. Acad. Sci. USA* **110**, 405–412 (2013).
12. Thomas, M.G., Martinez Tosar, L.J., Desbats, M.A., Leishman, C.C. & Boccaccio, G.L. Mammalian Staufen 1 is recruited to stress granules and impairs their assembly. *J. Cell Sci.* **122**, 563–573 (2009).
13. Ravel-Chapuis, A. *et al.* The RNA-binding protein Staufen1 is increased in DM1 skeletal muscle and promotes alternative pre-mRNA splicing. *J. Cell Biol.* **196**, 699–712 (2012).
14. Vessey, J.P. *et al.* A loss of function allele for murine Staufen1 leads to impairment of dendritic Staufen1-RNP delivery and dendritic spine morphogenesis. *Proc. Natl. Acad. Sci. USA* **105**, 16374–16379 (2008).
15. Furic, L., Maher-Laporte, M. & Desgroseillers, L. A genome-wide approach identifies distinct but overlapping subsets of cellular mRNAs associated with Staufen1- and Staufen2-containing ribonucleoprotein complexes. *RNA* **14**, 324–335 (2008).
16. Laver, J.D. *et al.* Genome-wide analysis of Staufen-associated mRNAs identifies secondary structures that confer target specificity. *Nucleic Acids Res.* **41**, 9438–9460 (2013).
17. Maher-Laporte, M. & Desgroseillers, L. Genome wide identification of Staufen2-bound mRNAs in embryonic rat brains. *BMB Rep.* **43**, 344–348 (2010).
18. Kusek, G. *et al.* Asymmetric segregation of the double-stranded RNA binding protein Staufen2 during mammalian neural stem cell divisions promotes lineage progression. *Cell Stem Cell* **11**, 505–516 (2012).
19. Ferrandon, D., Elphick, L., Nüsslein-Volhard, C. & St Johnston, D. Staufen protein associates with the 3′UTR of bicoid mRNA to form particles that move in a microtubule-dependent manner. *Cell* **79**, 1221–1232 (1994).
20. Kim, Y.K. *et al.* Staufen1 regulates diverse classes of mammalian transcripts. *EMBO J.* **26**, 2670–2681 (2007).
21. Singh, G., Ricci, E.P. & Moore, M.J. RIPiT-Seq: a high-throughput approach for footprinting RNA:protein complexes. *Methods* doi:10.1016/j.ymeth.2013.09.013 (2 October 2013).
22. Ramos, A. *et al.* RNA recognition by a Staufen double-stranded RNA-binding domain. *EMBO J.* **19**, 997–1009 (2000).
23. Liu, Z.R., Wilkie, A.M., Clemens, M.J. & Smith, C.W. Detection of double-stranded RNA-protein interactions by methylene blue-mediated photo-crosslinking. *RNA* **2**, 611–621 (1996).
24. Singh, G. *et al.* The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell* **151**, 750–764 (2012).
25. Miura, P., Shenker, S., Andreu-Agullo, C., Westholm, J.O. & Lai, E.C. Widespread and extensive lengthening of 3′ UTRs in the mammalian brain. *Genome Res.* **23**, 812–825 (2013).
26. Kucukural, A., Özadam, H., Singh, G., Moore, M.J. & Cenik, C. ASPeak: an abundance sensitive peak detection algorithm for RIP-Seq. *Bioinformatics* **29**, 2485–2486 (2013).
27. Carmona-Saez, P., Chagoyen, M., Tirado, F., Carazo, J.M. & Pascual-Montano, A. GENECODIS: a web-based tool for finding significant concurrent annotations in gene lists. *Genome Biol.* **8**, R3 (2007).
28. Nogales-Cadenas, R. *et al.* GeneCodis: interpreting gene lists through enrichment analysis and integration of diverse biological information. *Nucleic Acids Res.* **37**, W317–W322 (2009).
29. Tabas-Madrid, D., Nogales-Cadenas, R. & Pascual-Montano, A. GeneCodis3: a non-redundant and modular enrichment analysis tool for functional genomics. *Nucleic Acids Res.* **40**, W478–W483 (2012).
30. Micklem, D.R., Adams, J., Grünert, S. & St Johnston, D. Distinct roles of two conserved Staufen domains in *oskar* mRNA localization and translation. *EMBO J.* **19**, 1366–1377 (2000).
31. Macchi, P. *et al.* The brain-specific double-stranded RNA-binding protein Staufen2: nucleolar accumulation and isoform-specific exportin-5-dependent export. *J. Biol. Chem.* **279**, 31440–31444 (2004).
32. Kiebler, M.A. *et al.* The mammalian staufen protein localizes to the somatodendritic domain of cultured hippocampal neurons: implications for its involvement in mRNA transport. *J. Neurosci.* **19**, 288–297 (1999).
33. Cho, H. *et al.* Staufen1-mediated mRNA decay functions in adipogenesis. *Mol. Cell* **46**, 495–506 (2012).
34. Kretz, M. *et al.* Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature* **493**, 231–235 (2013).
35. Ghosh, S., Marchand, V., Gáspár, I. & Ephrussi, A. Control of RNP motility and localization by a splicing-dependent structure in *oskar* mRNA. *Nat. Struct. Mol. Biol.* **19**, 441–449 (2012).
36. Nott, A., Le Hir, H. & Moore, M.J. Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. *Genes Dev.* **18**, 210–222 (2004).
37. Wiegand, H.L., Lu, S. & Cullen, B.R. Exon junction complexes mediate the enhancing effect of splicing on mRNA expression. *Proc. Natl. Acad. Sci. USA* **100**, 11327–11332 (2003).
38. Ivanov, P.V., Gehring, N.H., Kunz, J.B., Hentze, M.W. & Kulozik, A.E. Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. *EMBO J.* **27**, 736–747 (2008).
39. Marión, R.M., Fortes, P., Beloso, A., Dotti, C. & Ortín, J. A human sequence homologue of Staufen is an RNA-binding protein that is associated with polysomes and localizes to the rough endoplasmic reticulum. *Mol. Cell Biol.* **19**, 2212–2219 (1999).
40. Wickham, L., Duchaîne, T., Luo, M., Nabi, I.R. & DesGroseillers, L. Mammalian staufen is a double-stranded-RNA- and tubulin-binding protein which localizes to the rough endoplasmic reticulum. *Mol. Cell Biol.* **19**, 2220–2230 (1999).
41. Monshausen, M. *et al.* Two rat brain staufen isoforms differentially bind RNA. *J. Neurochem.* **76**, 155–165 (2001).
42. Elbarbary, R.A., Li, W., Tian, B. & Maquat, L.E. STAU1 binding 3′ UTR IRAlus complements nuclear retention to protect cells from PKR-mediated translational shutdown. *Genes Dev.* **27**, 1495–1510 (2013).
43. Hartman, T.R. *et al.* RNA helicase A is necessary for translation of selected messenger RNAs. *Nat. Struct. Mol. Biol.* **13**, 509–516 (2006).
44. Villacé, P., Marión, R.M. & Ortín, J. The composition of Staufen-containing RNA granules from human cells indicates their role in the regulated transport and translation of messenger RNAs. *Nucleic Acids Res.* **32**, 2411–2420 (2004).
45. Darnell, J.C. *et al.* FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* **146**, 247–261 (2011).
46. Comery, T.A. *et al.* Abnormal dendritic spines in fragile X knockout mice: maturation and pruning deficits. *Proc. Natl. Acad. Sci. USA* **94**, 5401–5404 (1997).
47. Gruber, A.R., Lorenz, R., Bernhart, S.H., Neuböck, R. & Hofacker, I.L. The Vienna RNA websuite. *Nucleic Acids Res.* **36**, W70–W74 (2008).

© 2013 Nature America, Inc. All rights reserved.

ONLINE METHODS

Plasmids and cell lines. pcDNA5-TetO-Flag was previously described[38]. A cDNA encoding Stau1 (HindIII-NotI) was inserted into the polylinker of pcDNA5-TetO-Flag. A cDNA encoding the Stau1 mutant lacking RNA-binding activity was created by PCR using primers carrying the mutations described in ref. 2, in which phenylalanines at position 216 in dsRBD3 and at position 319 in dsRBD4 were mutated into alanines.

Stable cell lines were generated as described in ref. 38. In these cells, the expression level of the stably integrated Flag-tagged protein was optimized by titration of doxycycline (Dox; 0–2,000 ng ml^{-1}) to determine a concentration at which exogenous protein expression levels were comparable to those of endogenous counterparts.

Generation of Stau1-knockdown cell line. HEK293T LentiX cells (Clontech) were transfected with pGIPZ encoding shRNAs directed against Stau1 (Open Biosystems, CloneID: V2LHS_42695), pPAX2 and pMD2.G at a 12:9:3 ratio with Lipofectamine 2000 (Invitrogen). 2 d after transfection, the supernatant of transfected cells was collected and passed through a 0.45-μm filter. To generate the Stau1-knockdown cell line, HEK293 TRex cells (Invitrogen) were transduced with 7 mL of the supernatant of lentiviral-producing cells in the presence of 10 μg ml^{-1} polybrene for 6 h. Transduced cells were then selected in the presence of puromycin (3 μg ml^{-1}) for 2 weeks.

Stau1 RIPiT. The procedure was performed essentially as described in (ref. 38). For each Staufen purification, TRex-HEK293 cells containing a stable copy of Flag-tagged Stau proteins (Stau1-WT and Stau1-mut) or control cells (expressing Flag tag only) were grown in four 15-cm plates. Expression of the Flag-tagged protein was induced with doxycycline for 16 h. 1 h before cell harvesting, cycloheximide (CHX) was added to 100 μg ml^{-1}. The monolayer was rinsed and harvested in phosphate-buffered saline (PBS) containing 100 μg ml^{-1} CHX. The cells were lysed in 3 ml hypotonic lysis buffer (20 mM Tris-HCl, pH 7.5, 15 mM NaCl, 10 mM EDTA, 0.5% NP-40, 0.1% Triton X-100, 1× EDTA-free protease inhibitor cocktail (ROCHE), and 100 μg ml^{-1} CHX) for 10 min on ice. The suspension was sonicated (Branson Digital Sonifier-250) at 40% amplitude with a Microtip for a total of 20 s (in 2-s bursts with 10-s intervals). NaCl was adjusted to 150 mM, and the lysate was cleared by centrifugation at 15,000*g* for 10 min at 4 °C. The lysate was incubated for 2 h at 4 °C with 420 μl of anti-Flag agarose beads (50% slurry, Sigma) prewashed twice with 10 ml isotonic wash buffer (IsoWB) (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, and 0.1% NP-40). The RNA–protein (RNP) complexes captured on beads were washed four times (4 × 10 ml) with 10 ml IsoWB. After the fourth wash, bound RNP complexes were incubated with one bed volume of IsoWB containing 1 U μl^{-1} of RNase I for 10 min at 37 °C with intermittent shaking. RNP complexes were again washed four times with 10 ml IsoWB. Flag epitope–containing complexes were affinity eluted from the beads in one bed volume of IsoWB containing 250 μg.ml^{-1} Flag peptide with gentle shaking at 4 °C for 2 h. To prepare the recovered elution for input into a second IP, its volume was adjusted to 400 μl and its composition adjusted to that of the lysis buffer above with NaCl at 150 mM. The suspension was incubated with 7 μg of anti-Stau1 antibody (ab105398, Abcam, validation of IP in **Supplementary Fig. 1c**,**d**) precoupled to 35 μl of Protein G Dynabeads (Invitrogen) according to the manufacturer's instructions. Immunoprecipitation was carried out at 4 °C for 2 h. Captured RNP complexes were washed six times with 1 ml of ice-cold IsoWB and eluted with 200 μl of clear sample buffer (100 mM Tris-HCl, pH 6.8, 4% SDS, 10 mM EDTA and 100 mM DTT) at 25 °C for 5 min and subsequently at 95 °C for 2 min.

For IPs under protein–cross-linking conditions, cells were collected and rinsed once in PBS + CHX and then resuspended in PBS + CHX. Formaldehyde was added to 0.1%, and the suspension was gently mixed at RT for 10 min. A one-tenth volume of quenching buffer (2.5 M glycine, and 25 mM Tris base) was added. Cells were pelleted and lysed in hypotonic lysis buffer supplemented with 0.1% SDS and 0.1% sodium deoxycholate. Sonication after cell lysis was performed at 40% amplitude with a Microtip for a total of 90 s (in 5-s bursts with 30-s intervals). After Flag IP as described above, IP samples were washed twice with IsoWB + 0.1% SDS and 0.1% sodium deoxycholate and then with IsoWB. All subsequent steps were as above with omission of RNase I treatment.

RIPiT RNA extraction. The volume of RIPiT elution was extracted as described in ref. 38. For Stau1 cross-linked RIPiT experiments, extracted RNAs were depleted of rRNA before cDNA library construction with the Ribozero rRNA-removal kit from Epicentre.

Preparation of samples for RNA-seq. 75 μg of total RNA were poly(A)-selected with the Dynabeads mRNA-purification kit (Invitrogen). After poly(A) selection, mRNAs were fragmented with RNA-fragmentation buffer (Ambion) for 4 min and 30 s at 70 °C to obtain fragments 100–125 nt long. After fragmentation, RNAs were precipitated in three volumes of 100% ethanol at −20 °C overnight. After a wash with 70% ethanol, RNA was resuspended in 5 μl of water and the 3′ ends dephosphorylated with PNK (New England BioLabs) for 1 h at 37 °C. After this, RNAs were subjected to cDNA library preparation.

Poly(A)-site sequencing (PAS-seq). 75 μg of total RNA were poly(A)-selected with the Dynabeads mRNA purification kit (Invitrogen). After poly(A) selection, mRNAs were fragmented with RNA fragmentation buffer (Ambion) for 5 min at 70 °C to obtain fragments 60–80 nt long. After fragmentation, RNAs were reverse transcribed with an anchored and barcoded oligo$(dT)_{21}$VN (where V corresponds to A, C or G residues) containing the sequences complementary to Illumina's paired-end primers (PE1.0 and PE2.0) separated by a polyethylene glycol (PEG) spacer. After reverse transcription (RT) with Superscript III (Invitrogen), cDNAs were size-selected and circularized with Circligase I (Epicentre) for 4 h at 60 °C. This was followed by inactivation at 80 °C for 10 min. After circularization, cDNAs were PCR-amplified with Illumina's PE1.0 and 2.0 primers for a total of 14 cycles. PCR products were size-selected on a nondenaturing polyacrylamide gel and sent for high-throughput sequencing on Illumina's HiSeq2000 platform.

Quantitative RT-PCR. Total RNA from HEK293, Huh7 and HepG2 cells transfected with control shRNA or a Stau1-targeting shRNA construct was reverse transcribed with the Vilo RT kit from Life Technologies. Obtained cDNAs were then used as input for quantitative PCR analysis as described in ref. 48 with the following primers: *GAPDH* forward, 5′ tccaccaccctgttgctgtag 3′ and reverse, 5′acccactcctccacctttgac3′; *Arf1* forward, 5′ atcttcgcctcccgactc 3′ and reverse, 5′atgcttgtggacaggtgga3′; *C11orf58* forward, 5′ cagacgacgatctgggatct 3′ and reverse, 5′ tgatctcctataacaagacgaccag 3′; *PAICS* forward, 5′ aaggaaaagctgcaatctcaa 3′ and reverse, 5′ ccccacattttctggtgaag 3′; *MDM2* forward, 5′ catgcctgcccactttaga 3′ and reverse, 5′ ggaggctcccaactgctt 3′; *MDM4* forward, 5′ agggatgaaatgcttcttgg 3′ and reverse, 5′ aaggttgctatgaggtctaccttg 3′.

Sucrose-gradient sedimentation of Flag–Stau1-WT cells. HEK293 cells were plated at 5 million in a 150-mm^2 plate and Flag–Stau1-WT expression induced overnight with doxycycline at 0.5 ng ml^{-1}. 16 h after induction, cells were either incubated with cycloheximide (100 μg ml^{-1}) for 10 min or with harringtonine (2 μg ml^{-1}) for either 3, 10 or 40 min. This was followed by incubation with cycloheximide (100 μg ml^{-1}) for 10 min. Cells were then washed in PBS + cycloheximide (100 μg ml^{-1}) or PBS + harringtonine (2 μg ml^{-1}) + cycloheximide (100 μg ml^{-1}) and scraped. Cells were then lysed in 1 ml of lysis buffer (10 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 100 mM KCl, 1% Triton X-100, 2 mM DTT, 100 μg/ml cycloheximide and 1× Protease-Inhibitor Cocktail EDTA-free (Roche). Lysate was homogenized by gentle pipetting up and down with a P1000 pipettor for a total of eight strokes and incubated at 4 °C for 10 min. The lysate was centrifuged at 1,300*g* for 10 min at 4 °C; the supernatant was recovered. After this, samples were loaded on top of a 10–50% (w/v) sucrose gradient (20 mM HEPES-KOH, pH 7.4, 5 mM $MgCl_2$, 100 mM KCl, 2 mM DTT, and 100 μg ml^{-1} cycloheximide) and centrifuged in a SW-40ti rotor at 35,000 r.p.m. for 2 h 40 min at 4 °C. Samples were fractionated into 14 individual samples that were subjected to SDS-PAGE to monitor Stau1 (ab105398, Abcam, 1:1,000 dilution) and Rpl26 (Bethyl, A300-685A, 1:1,000 dilution) levels by western-blotting. Experimental validation of antibodies used for western blots can be found at the manufacturers' websites.

Oligo(dT) pulldown of poly(A) RNAs after UV exposure of living cells. HEK293 Flp-In cells were exposed to 0, 0.2, 0.4 and 0.8 J cm^{-2} of 254-nm UV light. Cells were then scraped and lysed with binding buffer (0.5 M NaCl, 10 mM Tris-HCl, pH 7.5, 0.5% SDS, 0.1 mM EDTA and protease inhibitor cocktail). Cell lysates were passed through a 22-gauge syringe needle five times and spun at 15,000*g* for 10 min. Cleared cell lysates were added to oligo(dT) cellulose beads (Ambion, AM10020) previously washed in binding buffer. After 1 h of

© 2013 Nature America, Inc. All rights reserved.

incubation at room temperature, beads were washed three times with binding buffer and once with nondenaturing wash buffer (0.5 M NaCl, 10 mM Tris-HCl, pH 7.5, 0.1% NP-40, 0.1% Triton-X 100 and 0.2 mM EDTA). Finally, beads were resuspended in elution buffer (10 mM Tris-HCl, pH 7.5, and 1 mM EDTA) complemented with RNase A/T1 cocktail (Ambion AM2286) and incubated at 37 °C for 1 h. The supernatant was finally recovered and loaded on a 12% SDS-PAGE for western-blotting of hnRNPA1 and Stau1.

Stau1 RNA reassociation test. HEK293 Flp-In, Flag–Stau1-WT or Flag–Stau1-mut cells were harvested and lysed as described in the Stau1 RIPiT section (above). After cell lysis, 0.1 pmol of a radiolabeled *in vitro*–transcribed *Arf1* 3′-UTR sequence (labeled with [α-^{32}P]UTP) was added for each 150-mm^2 plate. The remaining procedure is identical to that described in the Stau1 RIPiT section until the Flag elution step. Flag eluates on each sample were monitored for radioactivity with liquid scintillation.

Ribosome profiling. HEK293 Flag–Stau1-WT cells incubated in the presence of 1 ng ml^{-1} of doxycycline (Stau1 overexpression) or not (control) as well as HEK293 cells expressing the shRNAs against Stau1 (Stau1 knockdown) were seeded at 5 million cells in a 150-mm dish. After 16 h of culture, cycloheximide was added to 100 μg ml^{-1} for 10 min. Cells were then washed two times in ice-cold PBS + cycloheximide (100 μg ml^{-1}) and scraped in 1 ml of PBS + cycloheximide (100 μg ml^{-1}). Cells were pelleted at 500*g* for 5 min at 4 °C and lysed in 1 ml of lysis buffer (10 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 100 mM KCl, 1% Triton X-100, 2 mM DTT, 100 μg/ml cycloheximide and 1× Protease-Inhibitor Cocktail EDTA-free (Roche)). Lysate was homogenized with a P1000 pipettor by gentle pipetting up and down for a total of eight strokes and incubated at 4 °C for 10 min. The lysate was centrifuged at 1,300*g* for 10 min at 4 °C, the supernatant recovered and the absorbance at 260 nm measured. For the footprinting, 5 A_{260} units of the cleared cell lysates were incubated with 300 units of RNase T1 (Fermentas) and 500 ng of RNase A (Ambion) for 30 min at RT. After this, samples were loaded on top of a 10–50% (w/v) linear sucrose gradient (20 mM HEPES-KOH, pH 7.4, 5 mM $MgCl_2$, 100 mM KCl, 2 mM DTT and 100 μg ml^{-1} of cycloheximide) and centrifuged in a SW-40ti rotor at 35,000 r.p.m. for 2 h 40 min at 4 °C.

Samples were then collected from the top of the gradient while absorbance was measured at 254 nm and the fraction corresponding to 80S monosomes recovered. The collected fraction was complemented with SDS to 1% final and proteinase K (200 μg ml^{-1}) and then incubated at 42 °C for 45 min. After proteinase K treatment, RNA was extracted with one volume of phenol (pH 4.5)/chloroform/isoamyl alcohol (25:24:1). The recovered aqueous phase was supplemented with 20 μg of glycogen, 300 mM sodium acetate, pH 5.2, and 10 mM $MgCl_2$. RNA was precipitated with three volumes of 100% ethanol at −20 °C overnight. After a wash with 70% ethanol, RNA was resuspended in 5 μl of water and the 3′ ends dephosphorylated with PNK (New England BioLabs) in MES buffer (100 mM MES-NaOH, pH 5.5, 10 mM $MgCl_2$, 10 mM β-mercaptoethanol and 300 mM NaCl) at 37 °C for 3 h. Dephosphorylated RNA footprints were then resolved on a 15% acrylamide (19:1), 8 M urea denaturing gel for 1 h 30 min at 35 W and fragments ranging from 26 nt to 32 nt size-selected from the gel. Size-selected RNAs were extracted from the gel slice by overnight nutation at RT in RNA elution buffer (300 mM NaCl, and 10 mM EDTA). The recovered aqueous phase was supplemented with 20 μg of glycogen, 300 mM sodium acetate, pH 5.2, and 10 mM $MgCl_2$. RNA was precipitated with three volumes of 100% ethanol at −20 °C overnight. After a wash with 70% ethanol, RNA was resuspended in 5 μl of water and subjected to cDNA library construction.

The remaining undigested cell lysates were extracted with an equal volume of phenol (pH 4.5)/chloroform/isoamyl alcohol (25:24:1). The recovered aqueous phase was supplemented with 20 μg of glycogen, 300 mM sodium acetate, pH 5.2, and 10 mM $MgCl_2$. RNA was precipitated with three volumes of 100% ethanol at −20 °C overnight. After a wash with 70% ethanol, RNA was resuspended in 9 μl of water and fragmented with RNA fragmentation buffer (Ambion) at 70 °C for 4 min 30 s in order to obtain RNA fragments of 100–150 nt. Fragmented RNAs were supplemented with 20 μg of glycogen, 300 mM sodium acetate, pH 5.2, and 10 mM $MgCl_2$ and precipitated with three volumes of 100% ethanol at −20 °C overnight. After a wash with 70% ethanol, RNA was resuspended in 5 μl of water and dephosphorylated as described above with PNK. After 3′-end dephosphorylation, RNA fragments were subjected to cDNA library construction.

Illumina cDNA library construction. cDNA libraries were prepared with a homemade kit (E.E.H., unpublished data). Briefly, RNA fragments with a 3′-OH were ligated to a preadenylated DNA adaptor. Following this, ligated RNAs were reverse transcribed with Superscript III (Invitrogen) with a barcoded reverse-transcription primer that anneals to the preadenylated adaptor. After reverse transcription, cDNAs were resolved in a denaturing gel (10% acrylamide and 8 M urea) for 1 h and 45 min at 35 W. Gel-purified cDNAs were then circularized with CircLigase I (Epicentre) and PCR-amplified with Illumina's paired-end primers 1.0 and 2.0 for 5 cycles (ribosome footprints), 12 cycles (Stau1 RIPiT) or 16 cycles (RNA-seq libraries). PCR amplicons were gel-purified and submitted for sequencing on the Illumina HiSeq 2000 platform.

Mapping of high-throughput sequencing reads. First, reads were split with respect to their 5′-barcode sequence. After this, 5′-barcode and 3′-adaptor sequences were removed from reads. Reads were then aligned to University of California, Santa Cruz (UCSC) human hg18 assembly with TopHat[49]. Unmapped reads from TopHat were then mapped with Bowtie[50] to a custom set of sequences including 18S, 28S, 45S, 5S and 5.8S rRNA, repeat elements, small-nuclear RNAs (snRNAs), tRNAs, microRNAs and pre-microRNAs.

Transcript-level quantification and normalization for all high-throughput sequencing libraries. Read counts from all high-throughput sequencing libraries were normalized to the total number of mapped reads. When a single read aligned across the boundary of two different regions (for example, CDS and 3′ UTR), the read was divided proportionally to the aligned length in the given region. To quantify gene expression, reads per kilobase per million of mapped reads (RPKM) were calculated for the most abundant isoform of each gene.

Transcriptome-wide pairing of inverted and tandem Alu elements. Genomic coordinates for all Alu elements were obtained from the repeat-masker track of the UCSC genome browser. With the BedTools[51] intersectBed function with –s(strand) option, we obtained coordinates of all Alu elements located in 3′ UTRs, distal 3′ UTRs and introns. After this step, Alu elements in Watson and Crick strands were separated. To pair inverted Alu elements, we used the ClosestBed function between Alu elements in the Watson and Crick strands. The same steps were also performed to detect same-strand pairs on Watson-Watson and Crick-Crick pairs. In this case, tandem pairs that had an inverted Alu pair less than 2,000 nt apart were excluded from the analysis.

Definition of distal 3′-UTR regions. To define distal 3′-UTR regions, we used PAS-seq mapped reads. For this, peaks were called from the PAS-seq library with ASPeak. With the BedIntersect function (BedTools) we found all genes that had a polyadenylation site within 10,000 nt of the canonical polyadenylation site, provided that they were upstream of the transcription start site of the downstream gene. If multiple peaks were called within that interval, the called peak most distal to the canonical polyadenylation site was selected.

Counting of sequencing reads for Alu pairs separated by different distances. With the genomic coordinates of tandem and inverted Alu pairs described above, we created a new file for each region (3′ UTR, distal 3′ UTR and introns) that contained the genomic coordinates of each Alu element within every pair in addition to 1,000 nt upstream of the 5′-most Alu element and 1,000 nt downstream of the 3′-most Alu element. After this, read counts from Stau1-WT CROSS, Stau1-mut CROSS and RNA-seq were obtained for each defined region within pairs and then normalized with RNA-seq RPKM value for the same interval. Normalized read counts for each interval were added for all Alu pairs located between 0 and 200 nt, 200 and 500 nt, 500 and 1,000 nt, 1,000 and 2,000 nt, and 2,000 nt and beyond.

Secondary-structure analysis of inverted Alu pairs. To analyze the secondary structure of inverted Alu pairs, we obtained the sequence of both Alu elements, including the region separating them. As a control, same-length sequences were randomly chosen from the list of genes devoid of Stau1-target sites. After the sequences were obtained, the Vienna Package RNAfold 2.1.1 (ref. 47) was used to predict secondary structures.

Analysis of ribosome-profiling reads. To calculate translational-efficiency changes upon knockdown or overexpression of Stau1, we used a generalized

© 2013 Nature America, Inc. All rights reserved.

doi:10.1038/nsmb.2739

linear model (GLM). We used the number of sequencing reads mapping to the annotated coding region (or ORF) for each RefSeq transcript. In the GLM, we used the cell type (overexpression of Stau1WT, Stau1-shRNA), sequence-library preparation batch and type of sequence data (RNA-seq or ribo-seq) as predictor variables of the number of mapped reads per transcript. We had two biological replicates for all conditions, which were used to estimate a biological variability in the number of counts. We used a trended dispersion estimation method, following ref. 52. To extract translation-efficiency changes upon a given treatment, we used the contrast between type of sequence data, ribo-seq versus RNA-seq, in each pair of conditions.

48. Ricci, E.P. *et al.* Translation of intronless RNAs is strongly stimulated by the Epstein-Barr virus mRNA export factor EB2. *Nucleic Acids Res.* **37**, 4932–4943 (2009).
49. Trapnell, C., Pachter, L. & Salzberg, S.L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
50. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
51. Quinlan, A.R. & Hall, I.M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
52. McCarthy, D.J., Chen, Y. & Smyth, G.K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012).

© 2013 Nature America, Inc. All rights reserved.