

Nascent-Seq Indicates Widespread Cotranscriptional RNA Editing in *Drosophila*

Joseph Rodriguez,¹ Jerome S. Menet,¹ and Michael Rosbash^{1,*}¹Howard Hughes Medical Institute, National Center for Behavioral Genomics, and Department of Biology, Brandeis University, Waltham, MA 02451, USA*Correspondence: rosbash@brandeis.edu

DOI 10.1016/j.molcel.2012.05.002

SUMMARY

The RNA editing enzyme ADAR chemically modifies adenosine (A) to inosine (I), which is interpreted by the ribosome as a guanosine. Here we assess cotranscriptional A-to-I editing in *Drosophila* by isolating nascent RNA from adult fly heads and subjecting samples to high throughput sequencing. There are a large number of edited sites within nascent exons. Nascent RNA from an ADAR-null strain was also sequenced, indicating that almost all A-to-I events require ADAR. Moreover, mRNA editing levels correlate with editing levels within the cognate nascent RNA sequence, indicating that the extent of editing is set cotranscriptionally. Surprisingly, the nascent data also identify an excess of intronic over exonic editing sites. These intronic sites occur preferentially within introns that are poorly spliced cotranscriptionally, suggesting a link between editing and splicing. We conclude that ADAR-mediated editing is more widespread than previously indicated and largely occurs cotranscriptionally.

INTRODUCTION

RNA editing increases the informational content of the genome. Moreover, independent experiments in different systems indicate that editing is essential for biological function. Although *Drosophila* null for ADAR are morphologically normal with no life span defect, they exhibit severe behavioral and locomotor defects (Palladino et al., 2000). They probably reflect the failure to edit transcripts encoding many key proteins involved in fast electrical and chemical neurotransmission, e.g., *shab* (Ryan et al., 2008) and *shaker* (Ingleby et al., 2009). The effects of messenger RNA (mRNA) editing in mammals are also striking. ADAR1 and ADAR2 are both essential in mice; ADAR1^{-/-} mice die as early embryos, and ADAR2^{-/-} mice die postnatally of juvenile onset epilepsy-like symptoms (Hartner et al., 2004; Higuchi et al., 2000; Wang et al., 2000, 2004). For example, the transcript encoding the GluR-B subunit of the AMPA receptor is an ADAR substrate. Even mice heterozygous for editing-incompetent GluR-B- engineered alleles (unedited/edited) do not survive (Brusa et al., 1995; Feldmeyer et al., 1999).

Editing has been shown to occur primarily within the pre-mRNA of target genes (Higuchi et al., 1993). Although the rules that define an ADAR substrate are still uncertain (Eggington et al., 2011), the introns of several well-studied examples contain functional elements termed editing site complementary sequences (ECSs). Intramolecular base pairing between the ECS and the target exon creates a recognition element for ADAR binding. Based on RNase sensitivity, there is evidence that ADAR is recruited to nascent RNA, at specific chromosomal loci (Das et al., 2007; Eckmann and Jantsch, 1999; Sallacz and Jantsch, 2005). However, the extent to which the covalent editing events occur on nascent RNA, i.e., prior to 3' end formation, has not been broadly investigated.

RESULTS

Cotranscriptional Editing of Pre-mRNA

To assess the extent of cotranscriptional editing in the *Drosophila* system, we isolated nascent RNA from adult fly heads and subjected 12 samples (two data sets each containing six samples; see the [Experimental Procedures](#)) to high-throughput sequencing using a standard Illumina single-end protocol. As expected, the data showed that nascent RNA contains significant levels of intron signal compared to cytoplasmic mRNA and to total mRNA (see [Figures S1A](#) and [S1B](#) available online for an example of Nascent-seq versus pA-seq). To aid in filtering out possible SNPS, we also sequenced DNA from the yellow white (yw) strain to 39X coverage and only considered sites that did not contain a single G in the DNA. After mapping the RNA sequences to the reference genome, we searched for the occurrence of one or more guanosines at each site in our 12 samples. Our most stringent criterion required a G in at least ten of the 12 samples (and five of six samples within each dataset) as well as zero Gs in the genomic sequence ([Figure 1A](#)). Using a log likelihood score, we classified sites into high or low ranking (see the [Experimental Procedures](#) for more details).

The nascent RNA contains 621 high ranking edited sites within RefSeq annotated exons indicating that cotranscriptional editing is widespread ([Figure 1A](#) and [Table S1](#)). Additionally, there are another 251 lower ranking sites that did not meet the most stringent thresholds. To address data reproducibility, we pooled the six samples from each data set and calculated the average percent editing level at every edited site within the individual sets; there was remarkably high reproducibility (R = 0.96; [Figure 1B](#)). Most sites also show a rather low percentage of nascent

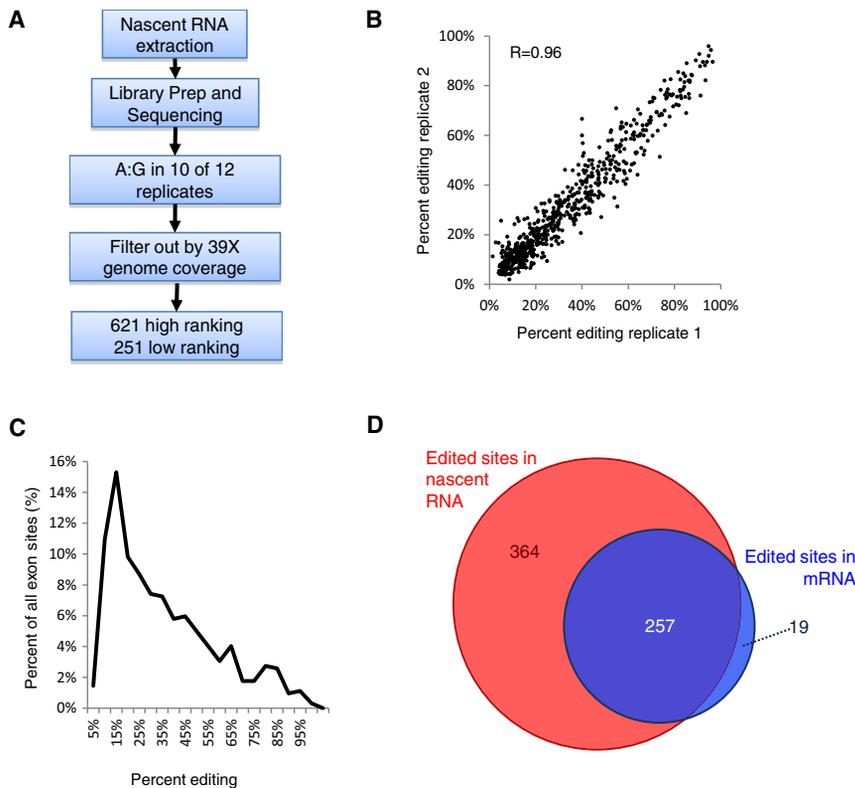


Figure 1. Editing of mRNA-Edited Sites Is Cotranscriptional

(A) Overview of workflow used to detect cotranscriptional editing. Putative edited sites occurring in ten of 12 replicates were filtered by 39X genomic sequence, resulting in 621 high-ranking cotranscriptionally edited sites.

(B) Comparison between the replicate sample editing levels show that the editing levels are highly correlated ($R = 0.96$). The percent editing levels per site are plotted.

(C) Distribution of editing levels is shown peaking at 15%.

(D) Overlap between the 621 high-ranking cotranscriptional exon editing sites and the 276 high-ranking mRNA exon editing sites. While a majority of mRNA sites (92%) were found to be cotranscriptionally edited, 41% of the nascent sites were found in the mRNA sites. Relaxing thresholds increases the mRNA overlap to 92% of the nascent sites.

See also Figure S1 and Table S1.

editing, peaking at 15% (Figure 1C). However, 22% of the sites are edited over 50%, i.e., $G > A$.

To compare nascent editing levels with more standard mRNA data, we sequenced pA RNA from two independent samples of *yw* fly heads (see the Experimental Procedures). These samples were from the same batch of fly heads used for the Nascent-seq data above. We also used the same analysis pipeline to identify editing sites in the mRNA data. However, since we only sequenced two mRNA samples, we required in addition that mRNA edited sites be observed in both samples.

The mRNA data indicate 276 high-ranking edited sites (Table S1). Significantly, 93% of these are present in the Nascent-seq, indicating that they are cotranscriptionally edited (Figure 1D); only 19 mRNA sites are not in these high-ranking nascent sites. In contrast, 41% of the high-ranking nascent sites were found in the mRNA data. We suspect that many of these sites are absent because of the rather low coverage, because only two samples were sequenced. Indeed, over 92% of the nascent sites are found in the mRNA data if the site is only required in one of the two mRNA preps (see below).

mRNA Editing Level Is Determined Cotranscriptionally

The 93% overlap between edited sites in mRNA and in nascent RNA allowed us to directly compare their editing levels. Although mRNA editing levels are slightly higher than nascent editing levels (8% on average; $p = 1 \times 10^{-21}$), the two are highly correlated ($R = 0.83$; Figure 2A). A similar high correlation of nascent editing levels is also observed with 12 pA-seq samples from

another strain, Canton-S (Cs) ($R = 0.90$; Figure 2B), indicating that the level of editing is set cotranscriptionally.

We then compared the rankings of exon base pair coverage, a measure of relative RNA abundance between the nascent and mRNA datasets (Figure 2C).

This was done by determining the log ratio of nascent read signal rankings divided by mRNA read signal rankings for edited genes compared to all genes. Edited genes have a lower relative ranking in mRNA compared to nascent RNA (Figure 2C; $p = 1 \times 10^{-81}$). The simplest interpretation is that edited mRNAs have a shorter half-life than nonedited mRNAs. Alternatively, editing affects transcription and causes a high nascent signal. Because longer mRNAs have lower mRNA/nascent RNA read ratios than shorter mRNAs (data not shown), we generated the same log ratios as a function of transcript length. The same lower relative mRNA ranking was observed even within restricted mRNA length distributions (Figures 2D, S2A, and S2B). We note that the distinction between edited and nonedited genes does not seem to be solely a characteristic of neuronal genes (data not shown).

Intron Editing Is Widespread

Editing within introns containing ECSs, has been previously described for the mouse *GluR-B* genes (Higuchi et al., 1993; Rueter et al., 1995). Because our nascent RNA data are highly enriched for intron signal (Figures S1A and S1B), we searched for genome-wide intron editing within RefSeq annotated introns. There were 729 high-ranking sites and an additional 171 low-ranking sites, indicating a surprisingly high level of nascent intron editing. This makes a total of 1,350 high ranking nascent intron plus exon sites, with only a small fraction of the latter located within 5' and 3' untranslated regions (UTRs) (Figure 3A and Table S1). The percentage of editing sites is slightly biased toward

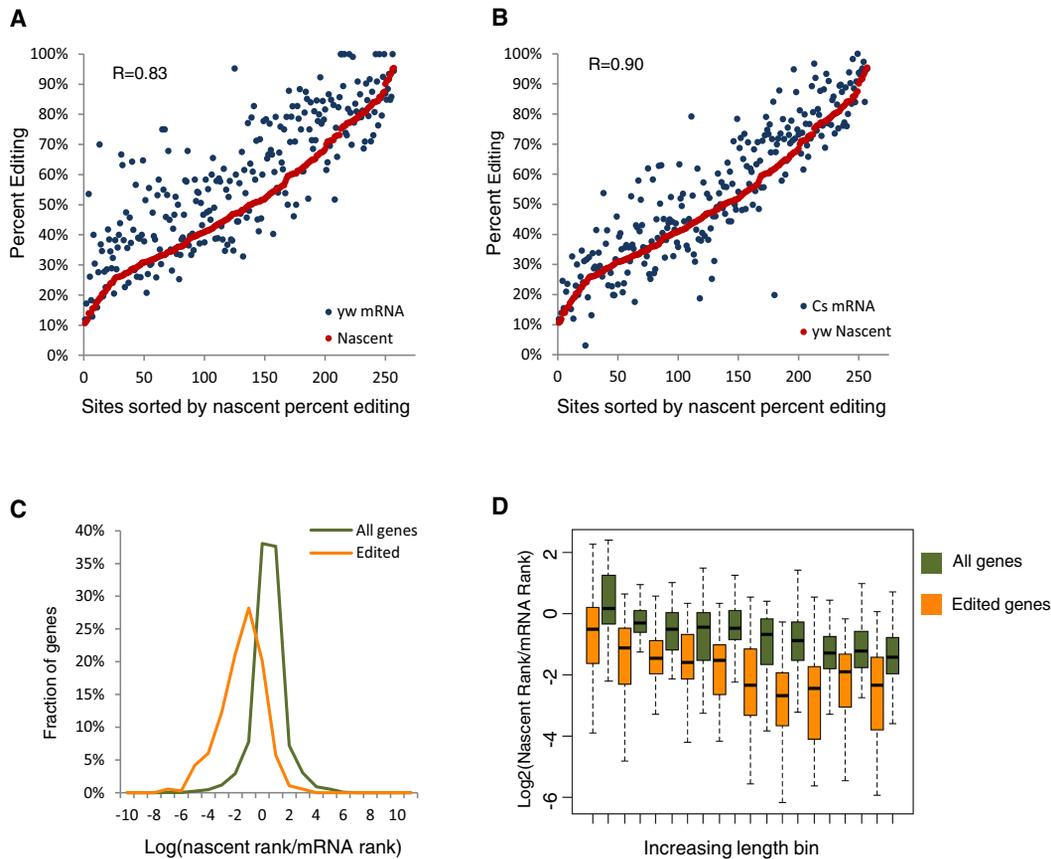


Figure 2. mRNA Editing Level Is Set Cotranscriptionally

(A) Comparison of the percent editing levels between nascent and mRNA data show that editing levels are set cotranscriptionally. The 257 exon sites found in both nascent and mRNA data sets were sorted by the nascent level and the pooled editing levels were plotted. Editing levels slightly rose in the mRNA data by an average of 8% (two-tailed paired t test, $p = 1 \times 10^{-21}$).

(B) mRNA editing levels in a different background (Cs) are also similar to the nascent editing levels.

(C) Distribution of nascent/mRNA rank is shown for both edited (orange) and all genes (green). These two distributions are different (two-tailed t test unequal variance, $p = 1 \times 10^{-81}$). Log₂ rank ratios are plotted. (See the main text for more details.)

(D) Box plot of the same data as in (C) but binned as a function of gene size. Edited genes are still different from all genes (two-tailed t test unequal variance, $p < 0.005$, Bonferroni corrected alpha).

See also Figure S2 and Table S1.

coding regions (36%; Figure 3A) relative to the percentage of coding nucleotides in the genome (27%); this suggests some positive selection for coding region editing or negative selection for editing elsewhere.

Specific or multiple editing, where one or only a few sites within a RNA are edited, is believed to occur on short imperfectly paired double-stranded RNA (dsRNA); promiscuous editing, i.e., many edited sites, occurs on long perfectly dsRNA regions (De-Cerbo and Carmichael, 2005b). We observed that intron editing and exon editing have similar distributions of specific, multiple, and promiscuous editing (Figure 3B). This suggests that intron editing follows the same rules as exon editing (see below).

Based on gene annotation data, we grouped nascent edited genes into three classes (Figure 3C). The first class, genes edited within introns as well as exons, contains 58 genes and represents 15% of all edited genes. This co-occurrence of intron editing and exon editing is much greater than expected for

independent events ($p = 1 \times 10^{-10}$), indicating that intron editing and exon editing are linked.

To determine whether exon and intron edited events occur within the same nascent transcripts in this first class of editing genes, we PCR amplified and cloned regions of three genes containing both intron and exon editing (see the Experimental Procedures). We then performed Sanger sequencing from both ends of individual clones. Exon and intron editing indeed occur within the same transcripts in all three genes (Figures 3D and S3A). Over 66% of these exon-edited transcripts contain at least one edited intron site, and over 79% of intron-edited transcripts contain at least one edited exon site. Furthermore, editing levels calculated from the Sanger sequencing data are remarkably well correlated with the levels determined from the Nascent-seq data ($R = 0.98$ exon sites, $R = 0.77$ intron sites; Figure S3B).

Shaker, a well-studied potassium channel, is a good example of this first class. It has six previously identified exon sites

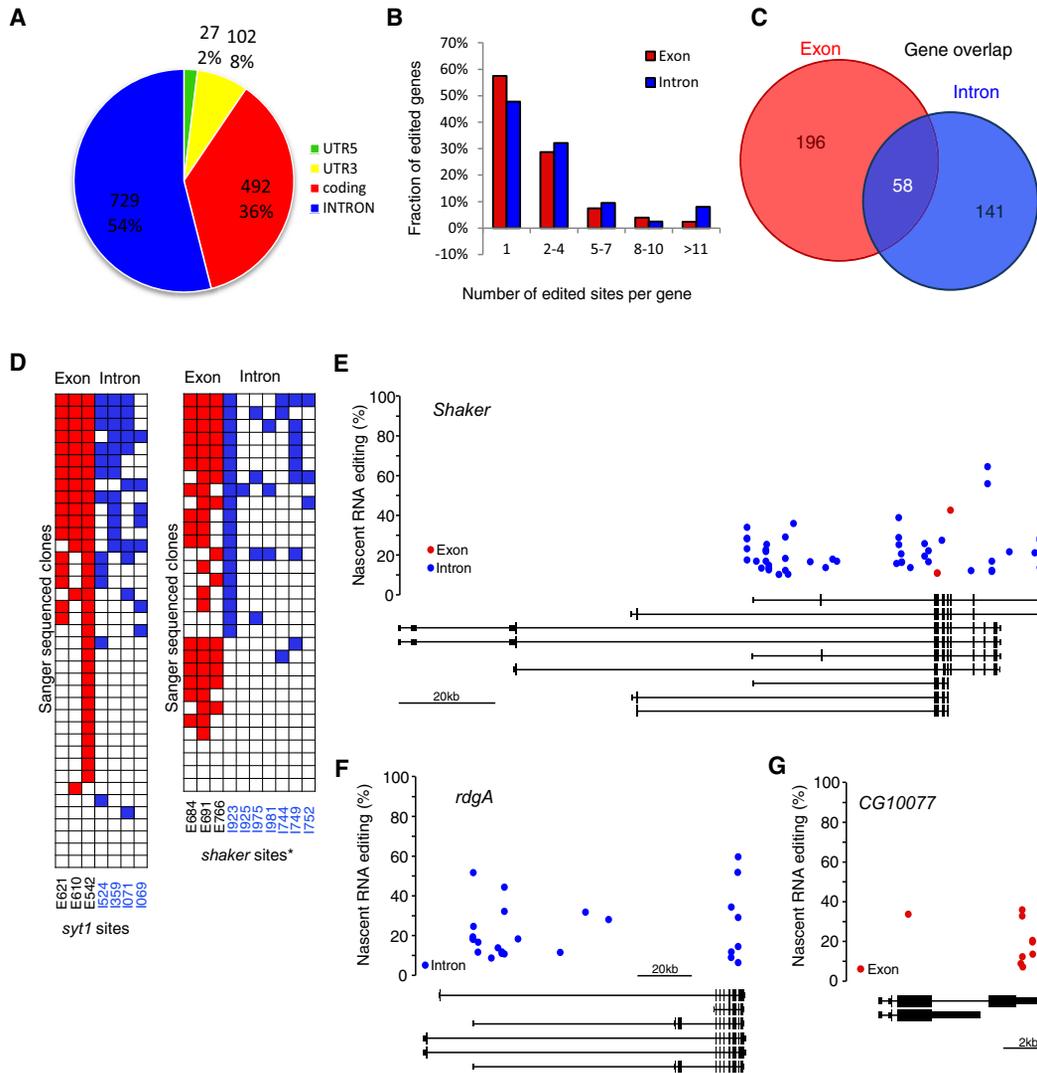


Figure 3. A Majority of Edited Sites Occurs within Introns

(A) Introns were scanned for editing using the same pipeline resulting in 729 high-ranking sites. Intron editing comprises 54% of all editing within the nascent data, in contrast to 46% within exons.

(B) Frequency of intron edited sites per gene follow a similar distribution as exon editing. Most genes contain one to four intron edited sites, while a few (<10%) contain more than 11.

(C) Overlap of exon edited genes and intron edited genes identify three classes of edited genes. Fifteen percent of edited genes are edited in both exon and intron. Thirty-six percent of edited genes are edited only within introns. Lastly, 49% of all edited genes are edited only in the exons.

(D) Intron and exon edited events in *syt1* and *shaker* genes occur on the same nascent transcripts. PCR amplicons were cloned into pGEM-T and individual clones were Sanger sequenced from the T7 and SP6 ends. The presence of a G in the chromatogram at the specified position is illustrated in red for exons and blue for introns. White represents the absence of editing, or an A in the Sanger sequencing.

(E) *shaker* is an example of a gene edited in both exon and intron. While six of six well-known exon sites (red) are found cotranscriptionally edited, we observe 52 sites (blue) within introns. Editing sites are plotted by position and level at the *shaker* locus. Asterisk denotes exon and intron sites found to occur on the same transcripts in Figure 3D.

(F) *rdgA* is an example of a gene edited only within the introns in which 26 cotranscriptionally edited sites are observed.

(G) *CG10077* is edited only within exons. We identified nine cotranscriptionally edited sites, five of which were previously identified in mRNA data from the modENCODE project.

See also Figure S3 and Tables S2 and S3.

(Hoopengardner et al., 2003; Ingleby et al., 2009), and the nascent data contain all six sites and no others. However, the nascent data also identify 52 intron-edited sites within this gene (Figure 3E). They are all distributed throughout the 3' half of the gene locus.

The second class, editing within introns only, accounts for 36% of all edited genes (Figure 3C). *rdgA*, a gene involved in vision, is an example and has 26 intron-edited sites (Figure 3F). They appear uniformly distributed across the transcription unit.

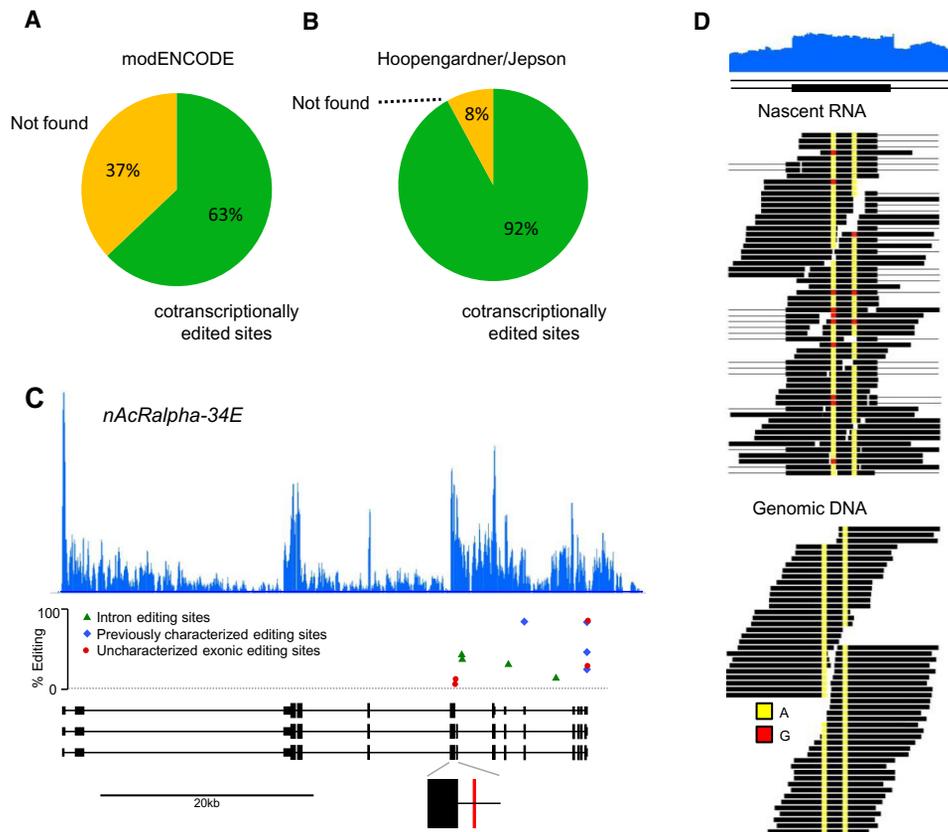


Figure 4. Comparison of Previously Identified mRNA Sites with Cotranscriptionally Edited Sites

(A) Comparison between the cotranscriptionally edited sites and the modENCODE mRNA edited sites shows that a majority (63%) of the (665) mRNA-edited sites are cotranscriptionally edited.

(B) Comparison between the cotranscriptionally edited sites and a pooled set of experimentally verified sites also shows that a majority (92%) of the (75) edited sites are cotranscriptionally edited.

(C) *nAcRalpha-34E* (*dalp5*) contains ten cotranscriptional edited sites within exons 6 were identical to the eight previously identified in the pooled data set. Two sites are found located within an alternatively spliced exon (red) that also have significant intron signal. We also identify four intronic sites for a total of 14 sites. The locations and editing levels are plotted for exon and intron sites. Plotted in blue are the sequenced nascent reads per base pair at the *nAcRalpha-34E* locus. A negative slope and intron signal are observed, indicative of nascent RNA.

(D) Comparison of nascent and genomic reads at the alternatively spliced exon locus of *nAcRalpha-34E*, containing two sites. Both sites are present in close proximity to each other, with the first one being slightly higher edited. G nucleotides are in red, A nucleotides in yellow. Editing levels for each site are plotted for uncharacterized exon- (red circle), previously characterized exon- (blue diamond), and intron- (green triangle) edited sites.

The third class contains genes edited in exons only and accounts for almost half of all edited genes (49%; Figure 3C). An example is *CG10077* with nine sites distributed between the 5' and 3' exons (Figure 3G). Five of these sites were recently identified by the modENCODE project (Graveley et al., 2011). Because this consortium generated a deep mRNA dataset by sequencing mRNA from a diverse set of developmental tissues including whole adults (Graveley et al., 2011), we were surprised that the nascent data identified all five of their *CG10077* sites as well as four more.

Most Identified mRNA Sites Are Cotranscriptionally Edited

To facilitate a more comprehensive comparison with the modENCODE data, we first removed modENCODE sites with one or more Gs and lower than 10× coverage in our genomic

sequencing data. Of the 665 remaining modENCODE sites, our high-ranking nascent head data identified 47% of them, 51% including lower-ranking sites, and up to 63% when only requiring the presence of a site in six of 12 nascent samples (Figure 4A). Conversely, 49% of the nascent sites are absent from the extensive modENCODE data set. Because a much larger fraction of the nascent sites are present in our yw head mRNA samples despite lower coverage, one simple explanation is differences in tissue sources. If much of *Drosophila* editing is from the nervous system, then the choice of diverse developmental stages and tissues by the modENCODE project might have underemphasized a key source of editing.

In addition to the modENCODE data set (Graveley et al., 2011), there is another experimentally verified list of edited mRNA sites (Hoopengardner et al., 2003; Jepson et al., 2011). The nascent data contain 81% of them, 92% when considering sites edited

in six of 12 samples (Figure 4B). For example, the nascent data identify ten edited sites within exons of the acetylcholine receptor subunit *nAcRalpha-34E* (*alpha5*) (Figure 4C). Of these ten, six were identical to the eight edited sites previously identified within this gene (Grauso et al., 2002; Hoopengardner et al., 2003; Jepson et al., 2011). We conclude that at least 75% of these previously identified sites are cotranscriptionally edited, consistent with the rest of the data indicating that most editing occurs cotranscriptionally.

The two extra sites present only in the nascent data lie within an exon that contributes to the ligand-binding domain of *nAcRalpha-34E*; it is also an alternatively spliced exon (Figures 4C and 4D) (Grauso et al., 2002). Interestingly, the nascent sequencing data also suggest that the intron surrounding this exon has a lower cotranscriptional splicing efficiency, i.e., a higher level of intron retention in nascent RNA than average introns in the nascent data (Figure 4D). Although these two sites could be absent from and not relevant to mature mRNA and therefore to coding potential, alternative splicing might make some edited sites more difficult to detect in mRNA than in nascent RNA.

Most Exon as well as Intron Sites Are Not Edited in ADAR Mutant Flies

To validate the identified nascent exon and intron edited sites and also to control for background noise, we sequenced nascent RNA from an ADAR mutant fly strain, ADAR0. It has severely compromised ADAR activity (see the Experimental Procedures). Given that we obtained very few viable ADAR0 flies, we modified our standard nascent RNA protocol to accommodate fewer fly heads and also sequenced the same number of yw and FM7A control fly heads. We prepared libraries from two replicates of the three strains and examined the 1,350 exon and intron sites. Only sites that occurred in both replicates of the yw and FM7 controls were considered, which resulted in 609 editing events: 374 exon sites and 235 intron sites.

The editing levels of almost all exon and intron sites were dramatically reduced in ADAR0 flies (Figures 5A and 5B). We used these data to calculate false positive rates for exon and intron editing, which were 4.5% and 5.1%, respectively. We conclude that the vast majority of newly identified sites are real and due to ADAR activity.

Intron Editing Occurs within the *syt1* ECS and Is Evolutionarily Conserved

Some ADAR substrates are formed when the target exon forms a double stranded structure with a complementary sequence (ECS) in the intron (Egebjerg et al., 1994; Hanrahan et al., 2000; Herb et al., 1996; Higuchi et al., 1993; Reenan, 2005). Moreover an intron containing the ECS in the transcript encoding GluR-B is edited (Higuchi et al., 1993; Rueter et al., 1995). To gain further insight into intron editing, we focused on two *syt1* ECS regions that are exceptionally well characterized for *syt1* exon editing (Reenan, 2005). These two regions are located within the subsequent intron and are necessary for formation of the ADAR substrate and for exon editing (Figure 6A). Mismatches introduced into these regions abolish editing within the exon and can be rescued with compensatory mutations that restore

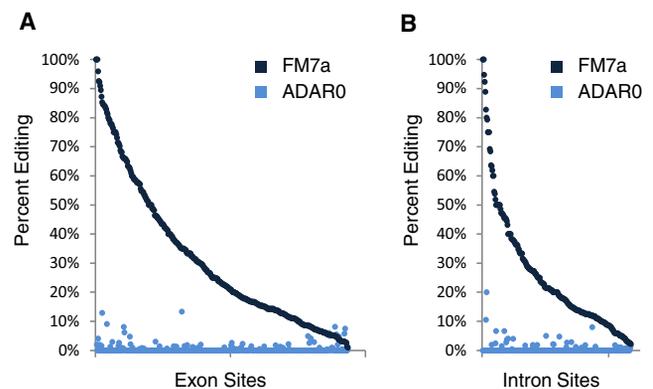


Figure 5. Most Exon and Intron Sites Are Poorly Edited in ADAR0 Flies

Comparison of editing levels in Nascent-seq data from ADAR0 and FM7a flies reveals that most exon and intron sites have a reduced editing level in the ADAR0 flies. Editing levels of the 1350 editing intron and exon sites were determined in ADAR0 and control groups. Only edited sites observed in both replicates of control groups were considered resulting in 609 total sites (e.g., FM7A and yw flies; see the Experimental Procedures). A false-positive rate of 4.5% and 5.1% was observed for exon and intron sites, respectively.

(A) Exon editing levels are shown for each of the 374 sites in ADAR0 and FM7a samples.

(B) Intron editing levels are shown for each of the 235 sites in ADAR0 and FM7a samples.

See also Table S1.

pairing (Reenan, 2005). We searched for editing within this intron and found four intronic sites. All four are located within the two ECS regions and directly opposite each other in the dsRNA structure (Figure 6A).

We next asked whether all intron-edited sites including the sites in *syt1* were evolutionary conserved across different *Drosophila* species. We first addressed the issue computationally and did not observe significant conservation of edited intron adenosines compared to all intron nucleotides sequences (data not shown). We then addressed the issue experimentally, by selecting two *syt1* intronic sites (I069 and I071 sites; Figure 6A) and assessing intron editing by Sanger sequencing PCR-amplified complementary DNA (cDNA) and genomic DNA fragments across four *Drosophila* species (including the distantly related *D. mojavensis*). Editing of both sites was apparent in all four species (Figure 6B), indicating they have been evolutionary conserved over 63 million years (Tamura et al., 2004).

Relationship between Editing and Splicing

Given that a 5' splice site is contained within the *syt1* pseudoknot structure (Reenan, 2005), we considered that splicing may compete with editing. To test this hypothesis, we compared the Nascent-seq signal within this intron between ADAR0 and controls. Significantly higher levels of intron signal were present in the ADAR0 strain compared to both yw and background genotype controls (Figure 6C), suggesting that ADAR activity increases *syt1* cotranscriptional splicing. This increased intron retention (1.55-fold increase, Figure 6D) was confirmed by qRT-PCR of total RNA (2.76-fold increase, $p < 0.05$; Figure 6E).

The connection between RNA editing and splicing of *sytl* inspired us to address this relationship genome wide. To this end, we used several metrics that reflect splicing, including intron retention; it measures the sequencing signal within each intron normalized to the exon signal across each gene (Khodor et al., 2011). Edited introns have significantly higher intron retention compared to all introns (median of 0.468 compared to 0.268, respectively; $p < 1 \times 10^{-15}$; Figures 6F, S4A, and S4B). Moreover, the introns surrounding exon-edited only genes exhibit significantly less intron signal (median upstream intron: 0.21, median downstream intron: 0.185; $p < 1 \times 10^{-5}$), when compared to the population of introns (median of 0.268). We interpret these results to indicate that intron editing occurs preferentially on introns that are poorly cotranscriptionally spliced. Because there is no global change in intron retention of intron-edited genes in ADAR0 flies (Figure S4C), either another editing component inhibits cotranscriptional splicing or the poor cotranscriptional splicing is upstream of editing, perhaps to facilitate ADAR recognition or function.

DISCUSSION

The data shown here indicate that Nascent-seq is an efficient and comprehensive strategy to identify editing sites. Although a minor contribution by post-transcriptional editing could account for the slight increase (8%) in mRNA editing levels compared to nascent RNA (Figure 2A), most adenosine-to-inosine (A-to-I) editing occurs cotranscriptionally. Moreover, it requires ADAR, as indicated by sequencing of the ADAR-null strain. In addition, this genome-wide study of A-to-I conversion revealed a surprisingly large number of intronic sites. A paper published during revision of this manuscript also reported extensive intron editing in the human system (Peng et al., 2012). The nascent RNA approach used here also suggests mechanistic links between editing and cotranscriptional splicing, as well as between intron editing and exon editing.

For example, the statistically significant overlap of intron and exon editing on the same genes suggests that they occur on the same transcripts, which is confirmed by the Sanger sequencing results (Figure 3D). Taken together with the dramatic reduction of intron as well as exon editing in the ADAR0 strain, we suggest that they are mechanistically similar and in some cases linked, e.g., the same ADAR-containing complex could catalyze intron as well as exon editing within one transcript. Given the possibility of unannotated alternate splicing (Figure 4D), intron editing may in some cases function similarly to exon editing and expand coding potential.

Because introns and their ECS regions have been shown to play an integral role in defining most well-studied ADAR substrates, one can imagine that intron editing facilitates or inhibits exon editing. A positive or regulatory role of intron editing on exon editing is possible in the case of the 15% of edited genes with editing events within introns as well as exons. Regulatory interactions might occur only indirectly, i.e., without ECS editing. However, a direct effect of ECS editing on exon editing is plausible in the case of *sytl* with four intron-edited sites within the known ECS (Figure 6A).

In contrast to the genes with both intron and exon editing, almost half of all edited genes only had exon editing, indicating

that it is biased toward exons and likely under positive selection. Most *Drosophila* introns are underrepresented in nascent RNA (Figure 6F), indicating that they are efficiently spliced cotranscriptionally (Khodor et al., 2011). This might cause an underestimate of intron editing frequency, which could also be related to the slight but significantly lower intron retention observed in exon-only edited genes (Figure 6F). However, more than 75% of “exon-only edited genes” are still missing intron editing when the thresholds for identifying editing are reduced from ten of 12 samples to six of 12 samples (data not shown). We cannot exclude the possibility that deeper sequencing would alter some of these conclusions, quantitatively or perhaps even qualitatively.

The introns of exon-only edited genes may still contain functional ECS regions, or these exons might form ADAR substrates autonomously. This is the case for an editing site that is recognized by ADAR without an intron (Keegan et al., 2005) and for the transcript of the intronless gene *K_v1.1* (Bhalla et al., 2004). These examples suggest that even mature transcripts have the potentially to be edited. Another alternative is that some exon-only edited genes use *trans*-acting RNAs to form ds RNA regions.

There are two extreme models for the relationship between editing and splicing: splicing occurs before editing, or editing occurs before splicing. If splicing is first, intron editing could be negatively impacted because potential intron substrate regions might be removed from nascent RNA before intron editing can occur. This raises an apparent paradox: might splicing also inhibit exon editing by removing a required ECS from pre-mRNA prior to editing? If editing is first, might it inhibit splicing (Bratt and Ohman, 2003; Laurencikiene et al., 2006)? This possibility is supported by the finding that the introns surrounding the *nAcRalpha-34E* alternative exon containing the two extra edited sites are inefficiently spliced (Figure 4D). Also consistent with this notion is the fact that edited introns on average have significantly more intron retention than all introns (Figure 6F). Alternatively, this might just reflect inefficient cotranscriptional splicing, which is mechanistically unrelated to editing. However, the inefficient splicing might still help promote editing.

Exon editing might even promote splicing. In the case of exon-only edited genes, this might help remove introns before less efficient intron editing events have a chance to occur. It is interesting in this context that numerous ADAR substrate pre-mRNAs sequester splicing signals within editing-relevant dsRNA structures (Burns et al., 1997; Hanrahan et al., 2000; Higuchi et al., 1993; Reenan, 2005). For example, the *sytl* intron forms a pseudoknot structure that contains the 5' splice site (Reenan, 2005). Moreover, there is a significant increase in intron retention of the *sytl* intron in ADAR0 nascent RNA compared to controls (Figures 6C–6E). Editing within the *sytl* DI and DII dsRNA may therefore destabilize these structures (Bass and Weintraub, 1988; Serra et al., 2004) and allow splicing to proceed.

However, there is no general effect on the retention of edited introns in ADAR0 nascent RNA compared to controls, e.g., the higher intron retention of edited introns may be upstream of editing. The *sytl*-edited intron therefore appears to be an exception. There are even a small number of edited introns that respond in the opposite way, namely, a decrease in intron retention in the

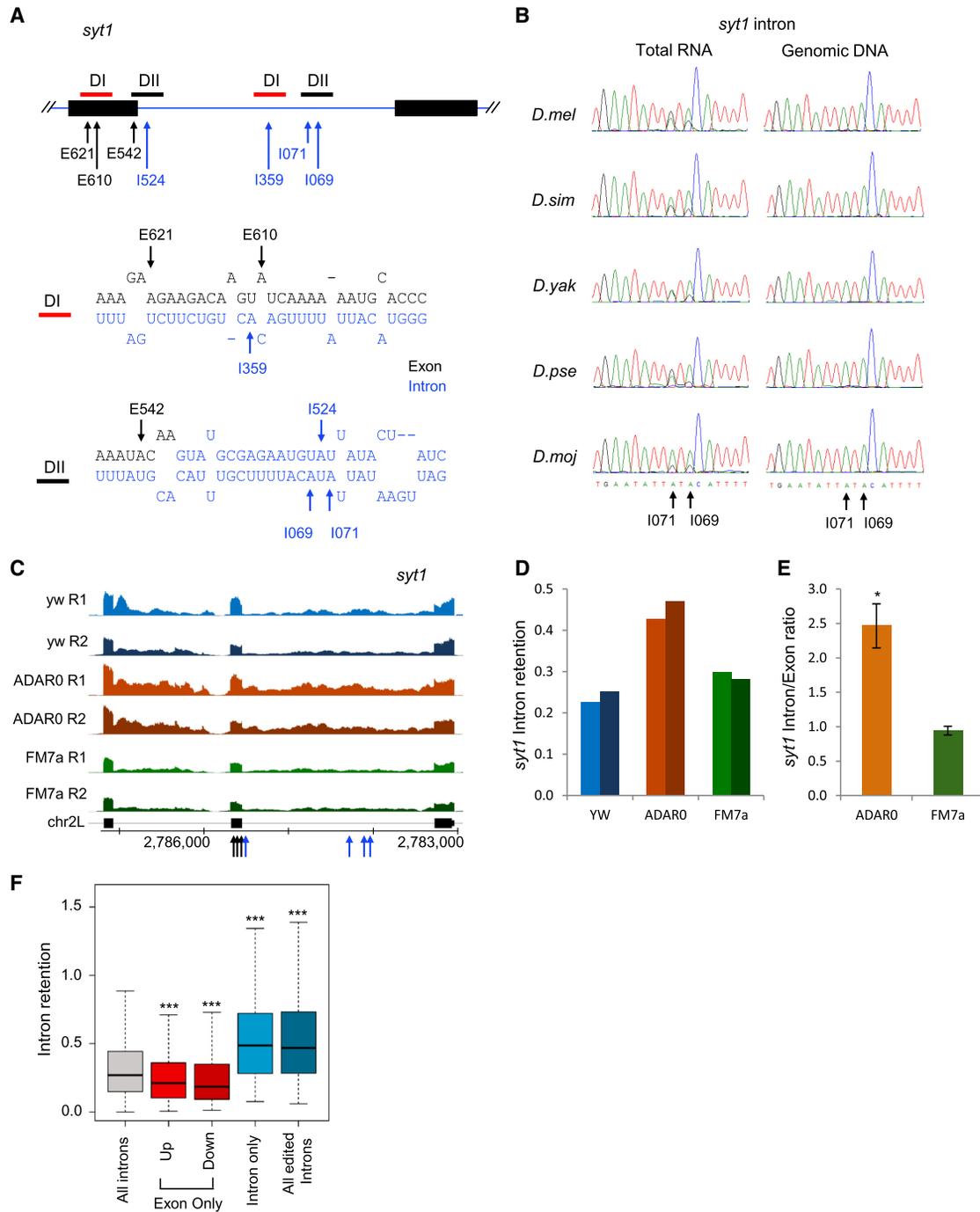


Figure 6. *syt1* Intron Editing Is Conserved and a Link between Editing and Splicing

(A) *syt1* exon and ECS pairing contains two dsRNA domains DI and DII (Reenan, 2005). These domains consist of an exon and intron pairing and are important for substrate definition. Cotranscriptionally edited sites were identified in *syt1* exon and intron regions. Exon sites are illustrated by black arrows, intron sites illustrated by blue arrows. Edited sites are found on the opposite strand of paired dsRNA structures for both DI and DII. Structures were folded with mfold.

(B) *syt1* intron sites are edited in different *Drosophila* species. Sanger sequencing chromatograms are shown for total RNA and genomic DNA. Validation was performed for each species from two independent replicates of total RNA.

(C) *syt1* intron signal from high throughput sequencing is higher in ADAR0 flies compared to controls yw and FM7a. yw replicates in blue, ADAR0 replicates in orange, and FM7a replicates in green. IGB genome illustration shown with genome annotation in black. All samples normalized by the number of uniquely mapped reads. Location of exon sites are illustrated by black arrows. Intron sites are illustrated by blue arrows.

(D) Intron retention from the high throughput sequencing data is quantified for each replicate and sample. Intron retention is defined as the average base pair signal within the intron divided by the average base pair signal across all exons.

ADAR0 nascent RNA compared to control nascent RNA (data not shown).

Nonetheless, there is good evidence in the literature for a regulatory interface between splicing and editing (Rieder and Reenan, 2012). For example, an intron-editing site has been reported to convert a cryptic/benign splice site into a favored splicing acceptor; it induces an alternative splicing event within an ADAR2 transcript (Dawson et al., 2004). There is other evidence that editing promotes splicing: ADAR2^{-/-} mice not only have significantly less edited GluR-B pre-mRNA but also accumulate pre-mRNA; they also have 5-fold less mRNA than the wild-type (Higuchi et al., 2000). Based on the fact that *Drosophila* editing levels are set cotranscriptionally (Figures 2A and 2B) and that most splicing occurs cotranscriptionally (Khodor et al., 2011), we favor the notion that some intron-only editing impacts pre-mRNA splicing. Alternatively or in addition, cotranscriptional intron editing may impact other aspects of nuclear RNA processing as suggested by its effect on transcript retention within nuclei (DeCervo and Carmichael, 2005a; Zhang and Carmichael, 2001).

EXPERIMENTAL PROCEDURES

Fly Strains

ADAR0/FM7a flies were obtained from the Reenan lab and were generated as a result of previous work (Jepson et al., 2011). This strain was generated from homologous recombination events at the ADAR locus. And although the extent of the deletion is unknown, at least the catalytic domain is deleted. *D. simulans* (14021-0251.194), *D. yakuba* (14021-0261.01), *D. pseudoobscura* (14011-0121.94), *D. mojavensis* (15081-1352.22) fly strains were obtained from the Griffith Lab but originate from the *Drosophila* Species Stock Center at the University of California, San Diego. *D. melanogaster* yw and Cs lab stock flies were used as controls for wild-type editing.

Nascent RNA Extraction

yw flies were entrained to three days of 12 hr:12 hr light and dark cycles at 25°. Flies were collected at 4 hr intervals on the fourth day and frozen on dry ice. Fly heads were then isolated with brass sieves. Nascent RNA was extracted as described (Khodor et al., 2011). In brief, purified nuclei were isolated from 1 ml fly heads for each sample and resuspended in 1 ml nuclear lysis buffer. An equal volume of 2X NUN (NaCl, Urea, NP40) buffer was added to each sample and incubated on ice for 20 min. The nascent RNA containing complex was pelleted. Nascent RNA was extracted from the NUN pellet with Invitrogen TRIzol Reagent according to the manufacturer's protocol.

Total RNA Extraction

Total RNA was extracted with TRIzol Reagent (Invitrogen) according to the manufacturer's protocol.

Library Construction and Sequencing

Sequencing library construction was performed via the standard Illumina protocol as previously described (Khodor et al., 2011) (see the Supplemental Experimental Procedures).

(E) qRT-PCR validation from total RNA of the increased intron signal in *sytl* between ADAR0 and control FM7a. The bar graph plots the ratio of a PCR amplicon within the intron and a PCR amplicon within the neighboring exon (see the Experimental Procedures). We observe a significant difference in intron and exon ratio between the ADAR0 and FM7a ($n = 3$, two-tailed t test unequal variance, $p < 0.01$). Asterisk denotes $p < 0.05$. The error bars represent the standard deviation of the qRT-PCR measurements from three independent samples.

(F) Intron retention of introns neighboring exon only edited genes is significantly lower than all introns in transcribed genes (median upstream intron, 0.21; median downstream intron, 0.185; population median of 0.268; Kruskal-Wallis, $p < 1 \times 10^{-5}$). Intron retention of intron only edited genes and all intron-edited genes is significantly higher than all introns in transcribed genes (medians of 0.487 and 0.468, respectively; population median of 0.268; Kruskal-Wallis, $p < 1 \times 10^{-15}$). Three asterisks denote $p < 0.001$.

See also Figure S4 and Tables S1, S2, and S3.

Genomic DNA Extraction

Genomic DNA was extracted via standard protocol with minor modifications (see the Supplemental Experimental Procedures).

PCR and qRT-PCR

PCR and RT-PCR were performed with Platinum Taq (Invitrogen) and Syber Master Mix (QIAGEN), respectively, with standard protocols as described (Abruzzi et al., 2011). qRT-PCR from ADAR0 and FM7a cDNA samples were performed from three independent replicates each. t tests with unequal variances were used to test significance between the samples. The sequences of the primers used are supplied in Table S3.

PCR Purification and Sanger Sequencing

PCR amplicons were gel purified and Sanger sequenced by Genewiz. Two replicates of cDNA made from two independent samples of total RNA were used for the I069 and I071 intron validation across species. Thirty-one to 39 plasmids from the pGEM-T cloning were sequenced for each cloned amplicon, from both T7 and SP6 ends.

pGEM-T Cloning

PCR amplicons from cDNA made from an independent sample of Nascent RNA were gel purified and cloned into pGEM-T (Promega) according to the manufacturer's protocol. Cloned amplicons were transfected into competent cells and plated on LB 100/XGAL media. White colonies were selected and grown in liquid media, and plasmids were extracted with a miniprep kit (QIAGEN).

Sequencing Alignment

Nascent RNA, yw mRNA, and yw genomic libraries were aligned to the dm3 *Drosophila melanogaster* genome with Tophat (Trapnell et al., 2009) with the parameters "-m 1 -F 0-microexon-search-no-closure-search -G exon20110421.gtf-solexa1.3-quals -l 50000." Cs paired-end samples were aligned with Tophat with the parameters "-r 50 -m 1 -F 0 -l 50000 -g 1 -G exon20110421.gtf," and reads with overlapping ends were removed and re-mapped. Cs single-end samples were aligned with Tophat using the parameters "-m 1 -F 0 -g 1-microexon-search-no-closure-search -G exon20110421.gtf-solexa1.3-quals -l 50000." We used annotation from the University of California, Santa Cruz (UCSC) Genes and Gene Prediction group, flybaseGeneTrack/table Apr. 2006 assembly to aid Tophat in the alignment of splice junctions (exon20110421.gtf).

RNA Editing Site Identification

Base frequencies were calculated within exons and introns of UCSC annotation. Genes with multiple isoforms were flattened, where overlapping exons generate one exon. Base positions with one or more Gs in the nascent data sets and zero Gs in the sequenced genomic DNA were identified. We required that editing sites occur in at least five of six samples within each set of replicate time points, for a total of ten of 12 independent occurrences for each site. To avoid potential mismapping of reads at splice junctions by Tophat, we required that edited sites occur in at least one of the two middle quadrants of at least one read. Intronic sites that occurred within ten bases of an annotated splice site were also discarded. We also ranked our data sets by using the following g test log likelihood metric:

$$a = \text{EditedBases} \quad b = \text{TotalBases} \quad c = \text{GenomicEditBases} \\ d = \text{GenomicTotalBases}$$

$$\log \text{likelihood metric}(\log l) = 2 * a * \ln\left(\frac{a}{\text{EditExp}}\right) + 2 * c * \ln\left(\frac{c}{\text{GExp}}\right)$$

$$\text{EditExp} = b * \frac{(a+c)}{(b+d)}$$

$$\text{GExp} = d * \frac{(a+c)}{(b+d)}$$

When considering only zero G bases in the genomic, the equation simplifies to

$$\log l = 2 * a * \log\left(\frac{a}{\text{EditExp}}\right).$$

This metric was calculated for each sample and summed over all 12 replicates with 12 degrees of freedom. A one sided chi squared p value was generated and used to separate editing sites into high ranking with a 1×10^{-6} cutoff and low ranking with a 0.05 cutoff. We applied a similar approach to the yw pA-seq data, with the exception that we required that the editing site occur in both samples (two of two).

Editing levels of our identified Nascent sites were also calculated in six paired-end 36 bp sequenced mRNA samples, as well as six single-end 72b p sequenced mRNA samples of the fly strain Cs. Editing levels were then calculated for the sites found in the nascent analysis.

Determination of Editing Level

For reproducibility of the nascent level, we pooled the editing counts for each replicate set of six time points together and calculated the percent editing level. The final percent editing level was determined by pooling the editing counts for all 12 samples and dividing by the total pooled counts of the 12 samples. Editing was similarly calculated for the yw pA-seq data and the small-scale Nascent-seq data by pooling of both replicate samples. The editing level for the Cs pA-seq data was calculated by pooling of all 12 samples. A two-tailed paired t test was used to test the significance of the observed editing levels between the nascent and yw mRNA data ($p = 1 \times 10^{-21}$).

Inferred Half-Life

All genes were ranked by the pooled average of nascent read coverage per base pair and the pooled average of the yw mRNA read coverage per base pair. A ranking of 1 indicates the gene contains the highest number of reads per base pair. We observed that edited genes were among the most highly expressed in the Nascent-seq data, even when controlling for transcript length (Figure S2C). The log base 2 ratio of the nascent rank/mRNA rank was calculated. We looked at a few functional categories of genes known to have long relative half-lives (Wang et al., 2002) and observed a good correlation with inferred half-lives (Figure S2D). We observed a significant difference between edited and all genes (two-tailed t test with unequal variances, $p = 1 \times 10^{-81}$). Larger genes also had longer inferred half-lives by this log ratio metric (data not shown), therefore genes were grouped into ten bins of 37 genes, leaving out the 14 longest genes and 11 ambiguous genes in which the editing site falls into two genes on the same strand. Unique genes of similar or larger size were selected into respective background distribution bins (Figure S2A). Only background genes with a minimum nascent average of four reads per base pair were considered. Because the variances of the log ratio metric were different between groups (Levene's test of Error Variances, $p \leq 1 \times 10^{-4}$), we did not use an ANOVA test. Alternatively, using t tests with unequal variance within the individual bins, we observed a significant difference between edited and background genes in 9 of 10 bins. ($p < 0.005$, two-tailed t test with unequal variances, Bonferroni corrected alpha; Figure S2B).

Edited Gene Overlap

Edited genes were classified into exon, intron, or exon and intron. A significant overlap was observed (two-tailed Fisher's exact test, $p = 1 \times 10^{-10}$) using all genes with a minimum NUN signal average of two reads per base pair, resulting in 7,703 genes total genes.

Annotation of Edited Sites

Sites found in an alternatively spliced exon or within a UTR were assigned preferentially to exons and not introns. UTRs and splice sites were generated from annotation downloaded from the UCSC Genes and Gene Prediction group, flybaseGene Track/table Apr. 2006 assembly.

modENCODE, Hoopengardner, and Jepson Data Preparation

We downloaded and filtered data sets by using our genomic DNA criterion, removing sites that contained one or more Gs in the gDNA seq or had lower than $10 \times$ coverage. We obtained 665 and 75 sites, respectively.

Determining False-Positive Rate

The exon and intron editing sites identified in the 12 Nascent-seq replicates were scanned in the small-scale ADAR0, FM7a, yw Nascent-seq data. We required editing to occur in both replicates of the FM7a and yw data, and sites with zero sequence coverage in any sample were not considered. This resulted in 609 total editing sites: 374 exon sites and 235 intron sites. Editing levels from each of the 12 replicates of the main yw Nascent-seq data were used to determine the fifth percentile. Both replicates for each small-scale (ADAR0, FM7a, yw) Nascent-seq data set were pooled to determine the final editing level. If the editing level was greater than the fifth percentile, it was considered a false positive. We then calculated the false-positive rate.

Splicing Metrics

Intron retention, 5' splice site (ss) and 3' ss ratio were calculated as described (Khodor et al., 2011). In brief, intron retention was calculated by dividing the average base pair coverage within the intron by the average base pair coverage within the exon. The 5' ss ratio was calculated as the average base pair coverage 25 bp into the intron from the 5' ss, divided by the average base pair coverage 25 bp into the exon from the 5' ss. The 3' ss was similarly calculated. Exon and intron UCSC annotation were handled as described above. Kruskal-Wallis nonparametric tests were used to test significance between the groups.

Additional methods are provided in Supplemental Experimental Procedures.

ACCESSION NUMBERS

All raw sequencing data are available for download from NCBI Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE37232.

SUPPLEMENTAL INFORMATION

Supplemental Information includes three tables, Supplemental Experimental Procedures, and four figures and can be found with this article online at doi:10.1016/j.molcel.2012.05.002.

ACKNOWLEDGMENTS

J.R., J.S.M., and M.R. conceived and designed the experiments; J.R. performed the experiments and bioinformatic analyses; J.R., J.S.M., and M.R. analyzed the data; and J.R., J.S.M., and M.R. wrote the manuscript. The authors declare no competing financial interests. We thank Katharine Abruzzi for helpful comments. We thank the Robert Reenan for kindly providing the ADAR0 fly strain and for helpful comments. We thank Peter Bronk, Chris Vecsey, and Leslie Griffith for fly strains. This research was supported by the Howard Hughes Medical Institute. J.R. was supported in part by a grant from the National Science Foundation (Integrative Graduate Education and Research Traineeship; DGE 0549390) and from the NIH (Genetics Training Grant; 5T32GM007122).

Received: September 28, 2011

Revised: March 8, 2012

Accepted: April 16, 2012

Published online: May 31, 2012

REFERENCES

- Abruzzi, K.C., Rodriguez, J., Menet, J.S., Desrochers, J., Zadina, A., Luo, W., Tkachev, S., and Rosbash, M. (2011). *Drosophila* CLOCK target gene characterization: implications for circadian tissue-specific gene expression. *Genes Dev.* 25, 2374–2386.
- Bass, B.L., and Weintraub, H. (1988). An unwinding activity that covalently modifies its double-stranded RNA substrate. *Cell* 55, 1089–1098.
- Bhalla, T., Rosenthal, J.J., Holmgren, M., and Reenan, R. (2004). Control of human potassium channel inactivation by editing of a small mRNA hairpin. *Nat. Struct. Mol. Biol.* 11, 950–956.
- Bratt, E., and Ohman, M. (2003). Coordination of editing and splicing of glutamate receptor pre-mRNA. *RNA* 9, 309–318.
- Brusa, R., Zimmermann, F., Koh, D.S., Feldmeyer, D., Gass, P., Seeburg, P.H., and Sprengel, R. (1995). Early-onset epilepsy and postnatal lethality associated with an editing-deficient GluR-B allele in mice. *Science* 270, 1677–1680.
- Burns, C.M., Chu, H., Rueter, S.M., Hutchinson, L.K., Canton, H., Sanders-Bush, E., and Emeson, R.B. (1997). Regulation of serotonin-2C receptor G-protein coupling by RNA editing. *Nature* 387, 303–308.
- Das, R., Yu, J., Zhang, Z., Gygi, M.P., Krainer, A.R., Gygi, S.P., and Reed, R. (2007). SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. *Mol. Cell* 26, 867–881.
- Dawson, T.R., Sansam, C.L., and Emeson, R.B. (2004). Structure and sequence determinants required for the RNA editing of ADAR2 substrates. *J. Biol. Chem.* 279, 4941–4951.
- DeCervo, J., and Carmichael, G.G. (2005a). Retention and repression: fates of hyperedited RNAs in the nucleus. *Curr. Opin. Cell Biol.* 17, 302–308.
- DeCervo, J., and Carmichael, G.G. (2005b). SINEs point to abundant editing in the human genome. *Genome Biol.* 6, 216.
- Eckmann, C.R., and Jantsch, M.F. (1999). The RNA-editing enzyme ADAR1 is localized to the nascent ribonucleoprotein matrix on *Xenopus* lampbrush chromosomes but specifically associates with an atypical loop. *J. Cell Biol.* 144, 603–615.
- Egebjerg, J., Kukekov, V., and Heinemann, S.F. (1994). Intron sequence directs RNA editing of the glutamate receptor subunit GluR2 coding sequence. *Proc. Natl. Acad. Sci. USA* 91, 10270–10274.
- Eggington, J.M., Greene, T., and Bass, B.L. (2011). Predicting sites of ADAR editing in double-stranded RNA. *Nat Commun* 2, 319.
- Feldmeyer, D., Kask, K., Brusa, R., Kornau, H.C., Kolhekar, R., Rozov, A., Burnashev, N., Jensen, V., Hvalby, O., Sprengel, R., and Seeburg, P.H. (1999). Neurological dysfunctions in mice expressing different levels of the Q/R site-unedited AMPAR subunit GluR-B. *Nat. Neurosci.* 2, 57–64.
- Grauso, M., Reenan, R.A., Culetto, E., and Sattelle, D.B. (2002). Novel putative nicotinic acetylcholine receptor subunit genes, *Dalpha5*, *Dalpha6* and *Dalpha7*, in *Drosophila melanogaster* identify a new and highly conserved target of adenosine deaminase acting on RNA-mediated A-to-I pre-mRNA editing. *Genetics* 160, 1519–1533.
- Graveley, B.R., Brooks, A.N., Carlson, J.W., Duff, M.O., Landolin, J.M., Yang, L., Artieri, C.G., van Baren, M.J., Boley, N., Booth, B.W., et al. (2011). The developmental transcriptome of *Drosophila melanogaster*. *Nature* 471, 473–479.
- Hanrahan, C.J., Palladino, M.J., Ganetzky, B., and Reenan, R.A. (2000). RNA editing of the *Drosophila* para Na(+) channel transcript. Evolutionary conservation and developmental regulation. *Genetics* 155, 1149–1160.
- Hartner, J.C., Schmittwolf, C., Kispert, A., Müller, A.M., Higuchi, M., and Seeburg, P.H. (2004). Liver disintegration in the mouse embryo caused by deficiency in the RNA-editing enzyme ADAR1. *J. Biol. Chem.* 279, 4894–4902.
- Herb, A., Higuchi, M., Sprengel, R., and Seeburg, P.H. (1996). Q/R site editing in kainate receptor GluR5 and GluR6 pre-mRNAs requires distant intronic sequences. *Proc. Natl. Acad. Sci. USA* 93, 1875–1880.
- Higuchi, M., Single, F.N., Köhler, M., Sommer, B., Sprengel, R., and Seeburg, P.H. (1993). RNA editing of AMPA receptor subunit GluR-B: a base-paired intron-exon structure determines position and efficiency. *Cell* 75, 1361–1370.
- Higuchi, M., Maas, S., Single, F.N., Hartner, J., Rozov, A., Burnashev, N., Feldmeyer, D., Sprengel, R., and Seeburg, P.H. (2000). Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. *Nature* 406, 78–81.
- Hoopengardner, B., Bhalla, T., Staber, C., and Reenan, R. (2003). Nervous system targets of RNA editing identified by comparative genomics. *Science* 301, 832–836.
- Ingleby, L., Maloney, R., Jepson, J., Horn, R., and Reenan, R. (2009). Regulated RNA editing and functional epistasis in Shaker potassium channels. *J. Gen. Physiol.* 133, 17–27.
- Jepson, J.E., Savva, Y.A., Yokose, C., Sugden, A.U., Sahin, A., and Reenan, R.A. (2011). Engineered alterations in RNA editing modulate complex behavior in *Drosophila*: regulatory diversity of adenosine deaminase acting on RNA (ADAR) targets. *J. Biol. Chem.* 286, 8325–8337.
- Keegan, L.P., Brindle, J., Gallo, A., Leroy, A., Reenan, R.A., and O'Connell, M.A. (2005). Tuning of RNA editing by ADAR is required in *Drosophila*. *EMBO J.* 24, 2183–2193.
- Khodor, Y.L., Rodriguez, J., Abruzzi, K.C., Tang, C.H., Marr, M.T., 2nd, and Rosbash, M. (2011). Nascent-seq indicates widespread cotranscriptional pre-mRNA splicing in *Drosophila*. *Genes Dev.* 25, 2502–2512.
- Laurencikienė, J., Källman, A.M., Fong, N., Bentley, D.L., and Ohman, M. (2006). RNA editing and alternative splicing: the importance of co-transcriptional coordination. *EMBO Rep.* 7, 303–307.
- Palladino, M.J., Keegan, L.P., O'Connell, M.A., and Reenan, R.A. (2000). A-to-I pre-mRNA editing in *Drosophila* is primarily involved in adult nervous system function and integrity. *Cell* 102, 437–449.
- Peng, Z., Cheng, Y., Tan, B.C., Kang, L., Tian, Z., Zhu, Y., Zhang, W., Liang, Y., Hu, X., Tan, X., et al. (2012). Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome. *Nat. Biotechnol.* 30, 253–260.
- Reenan, R.A. (2005). Molecular determinants and guided evolution of species-specific RNA editing. *Nature* 434, 409–413.
- Rieder, L.E., and Reenan, R.A. (2012). The intricate relationship between RNA structure, editing, and splicing. *Semin. Cell Dev. Biol.* 23, 281–288.
- Rueter, S.M., Burns, C.M., Coode, S.A., Mookherjee, P., and Emeson, R.B. (1995). Glutamate receptor RNA editing in vitro by enzymatic conversion of adenosine to inosine. *Science* 267, 1491–1494.
- Ryan, M.Y., Maloney, R., Reenan, R., and Horn, R. (2008). Characterization of five RNA editing sites in Shab potassium channels. *Channels (Austin)* 2, 202–209.
- Sallacz, N.B., and Jantsch, M.F. (2005). Chromosomal storage of the RNA-editing enzyme ADAR1 in *Xenopus* oocytes. *Mol. Biol. Cell* 16, 3377–3386.
- Serra, M.J., Smolter, P.E., and Westhof, E. (2004). Pronounced instability of tandem IU base pairs in RNA. *Nucleic Acids Res.* 32, 1824–1828.
- Tamura, K., Subramanian, S., and Kumar, S. (2004). Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks. *Mol. Biol. Evol.* 21, 36–44.
- Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105–1111.
- Wang, Q., Khillan, J., Gadue, P., and Nishikura, K. (2000). Requirement of the RNA editing deaminase ADAR1 gene for embryonic erythropoiesis. *Science* 290, 1765–1768.
- Wang, Y., Liu, C.L., Storey, J.D., Tibshirani, R.J., Herschlag, D., and Brown, P.O. (2002). Precision and functional specificity in mRNA decay. *Proc. Natl. Acad. Sci. USA* 99, 5860–5865.
- Wang, Q., Miyakoda, M., Yang, W., Khillan, J., Stachura, D.L., Weiss, M.J., and Nishikura, K. (2004). Stress-induced apoptosis associated with null mutation of ADAR1 RNA editing deaminase gene. *J. Biol. Chem.* 279, 4952–4961.
- Zhang, Z., and Carmichael, G.G. (2001). The fate of dsRNA in the nucleus: a p54(nrb)-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. *Cell* 106, 465–475.