Global regulation of alternative splicing by adenosine deaminase acting on RNA (ADAR)

O. Solomon, S. Oren, M. Safran, N. Deshet-Unger, P. Akiva, J. Jacob-Hirsch, K. Cesarkas, R. Kabesa, N. Amariglio, R. Unger, G. Rechavi, and E. Eyal

RNA **19** (2013) 1–14.

Speaker: Joseph Chuang-Chieh Lin

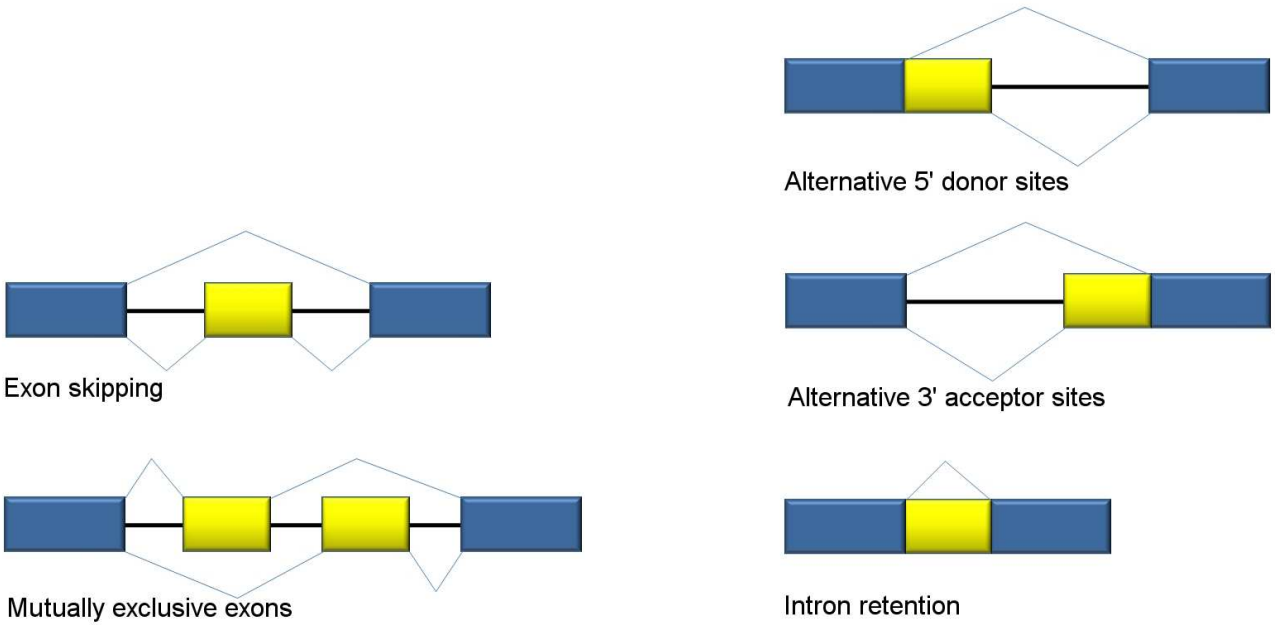
The Comparative & Evolutionary Genomics/Transcriptomics Lab.
Genomics Research Center, Academia Sinica
Taiwan

9 October 2013 (3 July 2013)

Outline

1. Introduction
2. Materials
3. Results
   - Collection of editing sites
   - A-to-I editing sites rarely fall within the canonical splicing motifs
   - *In silico* assay suggesting A-to-I RNA editing modifies SREs
   - ADAR has prominent influence on global splicing pattern
   - Splicing pattern changes by ADAR KD $\Leftrightarrow$ enrichment of A-to-I editing?
   - ADAR KD shows significant changes in splicing and RNA processing genes
4. Summary

Alternative splicing



- A major mechanism for gene regulation and transcriptome diversity.
- Yet the extent of the phenomenon, the regulation and specificity of the splicing machinery are only partially understood.

ADAR enzymes

- ADAR: **A**denosine **D**eaminase **A**cting on **R**NA.
- Adenosine-to-inosine (A-to-I) RNA editing of pre-mRNA by ADAR enzymes, which bind double strand RNAs.
  - ADAR (ADAR1): two distinct prevalent isoforms: p110 & p150.
  - ADARB1 (ADAR2).
  - ADARB2 (ADAR3).
- Crucial for development [Higuchi *et al.* *Nature* 2000 & Paz-Yaacov *et al.* *Proc. Natl. Acad. Sci.* 2010].
  - ADAR1 knockout: mice die in the embryonic stage.
  - ADAR2 knockout: mice suffer from seizures and die at an early age.
- Cytosine is base-paired by reverse transcriptase with inosine during cDNA synthesis.

Editing & pre-mRNA splicing

- A-to-I RNA **editing** and mRNA **splicing** are indeed coordinated in specific genes.
- This coordination may be governed by the RNA Pol II carboxy-terminal -domain (CTD) [Laurencikiene *et al*. *EMBO report* 2006; Ryman *et al*., *RNA* 2007].
  - CTD helps ensure that editing precedes splicing of the GluR-B transcript.
- Efficient exonic RNA editing often depends on intronic editing complementary sequences (ECS) for duplex formation.
- Examples for interrelations between editing and splicing are evident early in evolution [Jin *et al*., *BMC Evol. Biol.* 2007; Agrawal & Stormo, *RNA* 2005].

Editing & pre-mRNA splicing

- A-to-I RNA **editing** and mRNA **splicing** are indeed coordinated in specific genes.
- This coordination may be governed by the RNA Pol II carboxy-terminal -domain (CTD) [Laurencikiene *et al.* *EMBO report* 2006; Ryman *et al.*, *RNA* 2007].
  - CTD helps ensure that editing precedes splicing of the GluR-B transcript.
- Efficient exonic RNA editing often depends on intronic editing complementary sequences (ECS) for duplex formation.
- Examples for interrelations between editing and splicing are evident early in evolution [Jin *et al.*, *BMC Evol. Biol.* 2007; Agrawal & Stormo, *RNA* 2005].

Editing & pre-mRNA splicing (contd.)

- An editing site strengthens an enhancer signal within the exon [Agrawal & Stormo, *RNA* 2005].
- Exonization of intronic sequences:
  - RNA editing in the NARF gene in human might lead to exonization of an intronic sequence and the birth of the eighth exon of this gene [Lev-Maor *et al.*, *Genome Biol.* 2007].
- Editing at the branch site of PTPN6 gene in AML patients was found to impair splicing of the intron, with a probable role in leukemogenesis [Beghini *et al.*, *Hum. Mol. Genet.* 2000].

Contribution of this paper

To analyze how ADAR **globally** affects alternative mRNA splicing.

- A systematic approach;
- High-throughput expression analysis (exon-specific microarray) & sequencing of transcript data sets (ESTs/mRNA);
- Massively parallel sequencing (MPS; NGS).

Brief summary of the results

- A-to-I RNA editing rarely targets canonical splicing motifs.
  - Yet it was found to affect splicing regulatory elements (SREs) within exons.
- Cassette exons were found to be significantly enriched with A-to-I editing sites compared with constitutive exons.

Brief summary of the results (contd.)

- RNA-seq & exon-specific microarray revealed that ADAR knockdown in HepG2 & K562 cell lines leads to global changes in gene expression.
  - Hundreds of genes change their splicing patterns.
  - This **cannot** be explained by putative editing sites alone.
- Genes showing significant changes in their splicing pattern are frequently involved in RNA processing & splicing activity.
- Direct A-to-I RNA editing is NOT likely to be the primary mechanism for ADAR-mediated regulation of AS.
  - The regulation is suggested to be achieved by modulating *trans*-acting factors.

Brief summary of the results (contd.)

- RNA-seq & exon-specific microarray revealed that ADAR knockdown in HepG2 & K562 cell lines leads to global changes in gene expression.
  - Hundreds of genes change their splicing patterns.
  - This **cannot** be explained by putative editing sites alone.
- Genes showing significant changes in their splicing pattern are frequently involved in RNA processing & splicing activity.
- Direct A-to-I RNA editing is NOT likely to be the primary mechanism for ADAR-mediated regulation of AS.
  - The regulation is suggested to be achieved by modulating *trans*-acting factors.

Brief summary of the results (contd.)

- RNA-seq & exon-specific microarray revealed that ADAR knockdown in HepG2 & K562 cell lines leads to global changes in gene expression.
  - Hundreds of genes change their splicing patterns.
  - This **cannot** be explained by putative editing sites alone.
- Genes showing significant changes in their splicing pattern are frequently involved in RNA processing & splicing activity.
- Direct A-to-I RNA editing is NOT likely to be the primary mechanism for ADAR-mediated regulation of AS.
  - The regulation is suggested to be achieved by modulating *trans*-acting factors.

Brief summary of the results (contd.)

- RNA-seq & exon-specific microarray revealed that ADAR knockdown in HepG2 & K562 cell lines leads to global changes in gene expression.
  - Hundreds of genes change their splicing patterns.
  - This **cannot** be explained by putative editing sites alone.
- Genes showing significant changes in their splicing pattern are frequently involved in RNA processing & splicing activity.
- Direct A-to-I RNA editing is NOT likely to be the primary mechanism for ADAR-mediated regulation of AS.
  - The regulation is suggested to be achieved by modulating *trans*-acting factors.

Brief summary of the results (contd.)

- RNA-seq & exon-specific microarray revealed that ADAR knockdown in HepG2 & K562 cell lines leads to global changes in gene expression.
  - Hundreds of genes change their splicing patterns.
  - This **cannot** be explained by putative editing sites alone.
- Genes showing significant changes in their splicing pattern are frequently involved in RNA processing & splicing activity.
- Direct A-to-I RNA editing is NOT likely to be the primary mechanism for ADAR-mediated regulation of AS.
  - The regulation is suggested to be achieved by modulating *trans*-acting factors.

Genomic data

- hg19/GRCh37 (UCSC ftp site): DNA reference for detection of editing sites.
- RefSeq:
  - To define genes and exon–intron boundaries.
  - To determine 5'-SS and 3'-ss consensus motifs.
- dbSNP 132 (from UCSC table browser): SNP sites.
- Repeat data: taken from RepeatMasker.

Transcript data

- EST & mRNA data and their alignments to hg19:
  - EST: 249,717 reads; mRNA: 7,510,566 reads.
  - Taken from UCSC table browser.
- 454 Life Science RNA-seq data: SRA003647/SRP000614.
  - Including:
    - HBRR (Human Brain Reference RNA);
    - UHRR: Universal Human Reference RNA.
  - Alignment against hg19 using **BLAT**.
- RNA-Seq data (short reads):
  - U87MG cell line:
    - 2 x control + 2 x ADAR KD.
    - Downloaded (Bahn *et al*. Genome Res. 2012).
  - HepG2 & K562 cell line
    - Illumina GAIIX (76bp; paired-end; alignment: using **TopHat**).

Microarray experiment setting

- Using an exon-specific microarray (Affymetrix exon 1.0 st).
  - A probe set for each annotated exon.
  - ⋆ Feasible to compare expression levels of individual exons and thus to evaluate AS.
- For each cell line, control samples and ADAR KD samples were tested.

A-to-I RNA editing sites detection and collection

- Comparison of mRNA/EST/RNA-seq reads and the reference genome (preliminary SNV sites).
- Then process the following filters:
  - Quality assurance:
    1. The SNV site must be distal (20 bases) from sequence ends.
    2. Exclude known SNP sites.
    3. Enough support:
       $\geq$ 5% of all aligned reads that cover the site;
       $\geq$ 2 ESTs or $\geq$ 1 mRNA.
  - Others specific to ADAR enzymes:
    1. Clustering: [... A-to-G ... A-to-G ... A-to-G ...] $\Leftarrow$ 32 bases.
    2. Double strand: $\xleftarrow{12\text{ bp}}$ A-to-G $\xrightarrow{12\text{ bp}}$ 80% complementary to $\xleftarrow{1000\text{ bp}}$ A-to-G $\xrightarrow{1000\text{ bp}}$.
- Including previous published (putative) editing sites (e.g., DARNED).

A-to-I RNA editing sites detection and collection

- Comparison of mRNA/EST/RNA-seq reads and the reference genome (preliminary SNV sites).
- Then process the following filters:
  - Quality assurance:
    1. The SNV site must be distal (20 bases) from sequence ends.
    2. Exclude known SNP sites.
    3. Enough support:
       $\geq$ 5% of all aligned reads that cover the site;
       $\geq$ 2 ESTs or $\geq$ 1 mRNA.
  - Others specific to ADAR enzymes:
    1. Clustering: [. . . A-to-G . . . A-to-G . . . A-to-G . . .] $\Leftarrow$ 32 bases.
    2. Double strand: $\xleftarrow{12\text{ bp}}$ A-to-G $\xrightarrow{12\text{ bp}}$ 80% complementary to $\xleftarrow{1000\text{ bp}}$ A-to-G $\xrightarrow{1000\text{ bp}}$.
- Including previous published (putative) editing sites (e.g., DARNED).

A-to-I RNA editing sites detection and collection

- Comparison of mRNA/EST/RNA-seq reads and the reference genome (preliminary SNV sites).
- Then process the following filters:
  - Quality assurance:
    1. The SNV site must be distal (20 bases) from sequence ends.
    2. Exclude known SNP sites.
    3. Enough support:
       $\geq$ 5% of all aligned reads that cover the site;
       $\geq$ 2 ESTs or $\geq$ 1 mRNA.
  - Others specific to ADAR enzymes:
    1. Clustering: [. . . A-to-G . . . A-to-G . . . A-to-G . . .] $\Leftarrow$ 32 bases.
    2. Double strand: $\xleftarrow{12\text{ bp}}$ A-to-G $\xrightarrow{12\text{ bp}}$ 80% complementary to $\xleftarrow{1000\text{ bp}}$ A-to-G $\xrightarrow{1000\text{ bp}}$.
- Including previous published (putative) editing sites (e.g., DARNED).

A-to-I RNA editing sites detection and collection

- Comparison of mRNA/EST/RNA-seq reads and the reference genome (preliminary SNV sites).
- Then process the following filters:
  - Quality assurance:
    1. The SNV site must be distal (20 bases) from sequence ends.
    2. Exclude known SNP sites.
    3. Enough support:
       $\geq$ 5% of all aligned reads that cover the site;
       $\geq$ 2 ESTs or $\geq$ 1 mRNA.
  - Others specific to ADAR enzymes:
    1. Clustering: [. . . A-to-G . . . . . . A-to-G . . . A-to-G . . .] $\Leftarrow$ 32 bases.
    2. Double strand: $\xleftarrow{12\text{ bp}}$ A-to-G $\xrightarrow{12\text{ bp}}$ 80% complementary to $\xleftarrow{1000\text{ bp}}$ A-to-G $\xrightarrow{1000\text{ bp}}$.
- Including previous published (putative) editing sites (e.g., DARNED).

A-to-I RNA editing sites detection and collection

- Comparison of mRNA/EST/RNA-seq reads and the reference genome (preliminary SNV sites).
- Then process the following filters:
  - Quality assurance:
    1. The SNV site must be distal (20 bases) from sequence ends.
    2. Exclude known SNP sites.
    3. Enough support:
       $\geq$ 5% of all aligned reads that cover the site;
       $\geq$ 2 ESTs or $\geq$ 1 mRNA.
  - Others specific to ADAR enzymes:
    1. Clustering: [. . . A-to-G . . . ~~A-to-G~~ . . . A-to-G . . . A-to-G . . .] $\Leftarrow$ 32 bases.
    2. Double strand: $\xleftarrow{12\text{ bp}}$ A-to-G $\xrightarrow{12\text{ bp}}$ 80% complementary to $\xleftarrow{1000\text{ bp}}$ A-to-G $\xrightarrow{1000\text{ bp}}$.
- Including previous published (putative) editing sites (e.g., DARNED).

A-to-I RNA editing sites detection and collection

- Comparison of mRNA/EST/RNA-seq reads and the reference genome (preliminary SNV sites).
- Then process the following filters:
  - Quality assurance:
    1. The SNV site must be distal (20 bases) from sequence ends.
    2. Exclude known SNP sites.
    3. Enough support:
       $\geq$ 5% of all aligned reads that cover the site;
       $\geq$ 2 ESTs or $\geq$ 1 mRNA.
  - Others specific to ADAR enzymes:
    1. Clustering: [. . . A-to-G . . . ~~A-to-G~~ . . . A-to-G . . . A-to-G . . .] $\Leftarrow$ 32 bases.
    2. Double strand: $\xleftarrow{12\text{ bp}}$ A-to-G $\xrightarrow{12\text{ bp}}$ 80% complementary to $\xleftarrow{1000\text{ bp}}$ A-to-G $\xrightarrow{1000\text{ bp}}$.
- Including previous published (putative) editing sites (e.g., DARNED).

A-to-I RNA editing sites detection and collection

- Comparison of mRNA/EST/RNA-seq reads and the reference genome (preliminary SNV sites).
- Then process the following filters:
  - Quality assurance:
    1. The SNV site must be distal (20 bases) from sequence ends.
    2. Exclude known SNP sites.
    3. Enough support:
       $\geq$ 5% of all aligned reads that cover the site;
       $\geq$ 2 ESTs or $\geq$ 1 mRNA.
  - Others specific to ADAR enzymes:
    1. Clustering: [. . . A-to-G . . . ~~A-to-G~~ . . . A-to-G . . . A-to-G . . .] $\Leftarrow$ 32 bases.
    2. Double strand: $\xleftarrow{12\text{ bp}}$ A-to-G $\xrightarrow{12\text{ bp}}$ 80% complementary to $\xleftarrow{1000\text{ bp}}$ A-to-G $\xrightarrow{1000\text{ bp}}$.
- Including previous published (putative) editing sites (e.g., DARNED).

Verification of editing sites from RNA-seq analysis

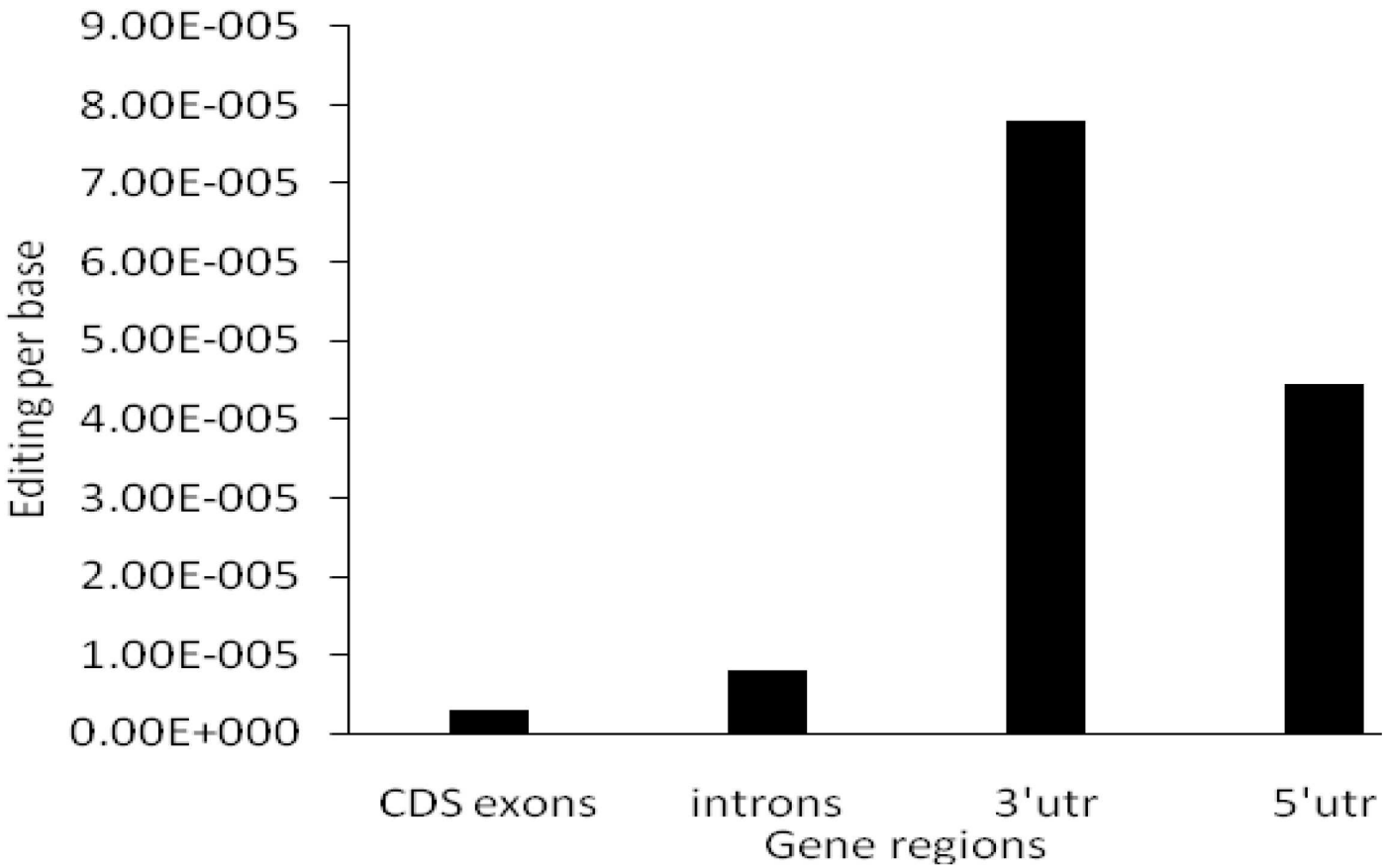
- To verify editing sites predicted based on mRNA/EST, using:
  - SAMtools pileup and varScan [Koboldt *et al.* *Bioinformatics* 2009] on the RNA-seq alignment and scanning for A-to-G or T-to-C.
  - Only sites in both replicates were considered.

Category of editing sites predicted based on mRNA/EST

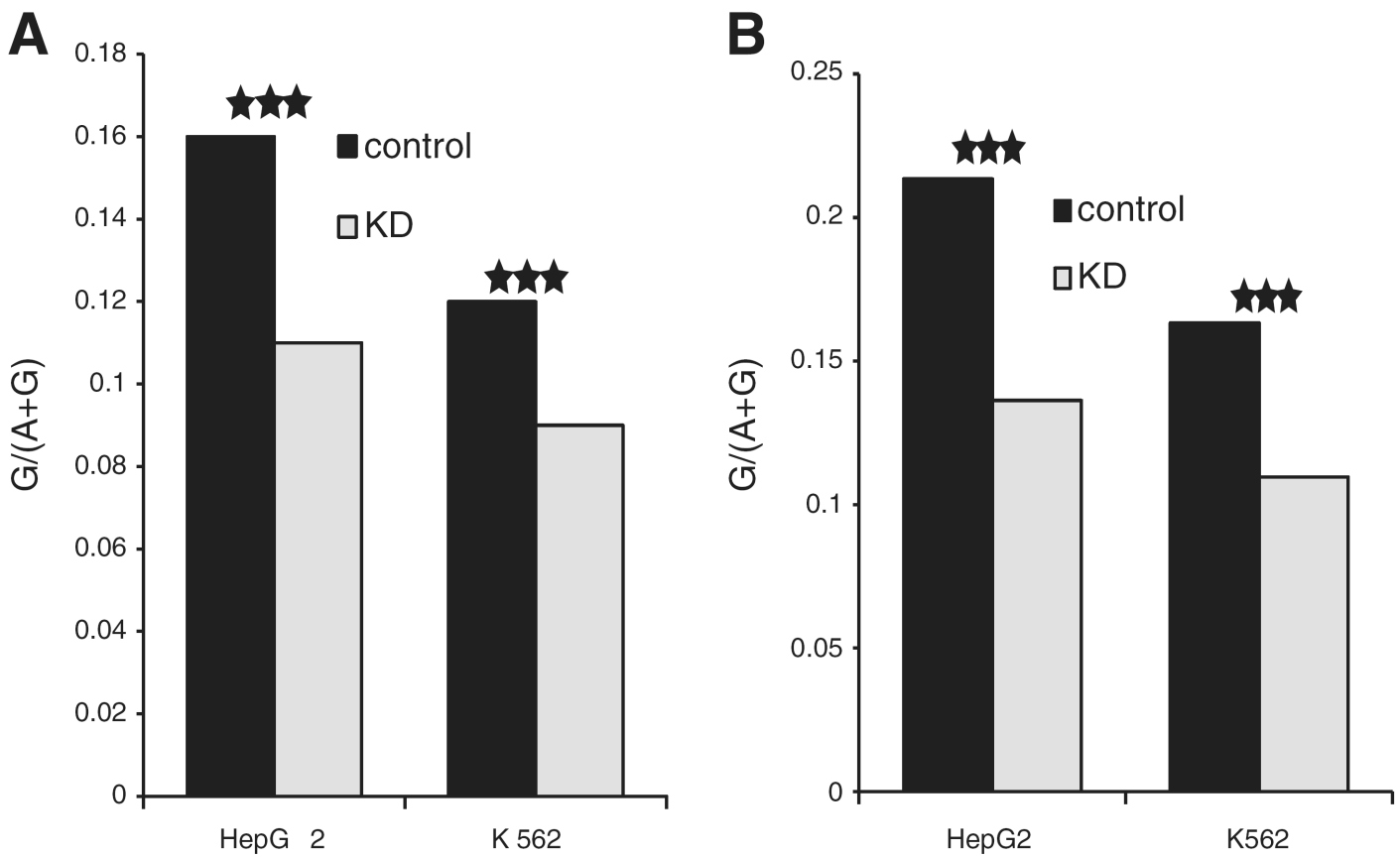
Sites passed all filters:

- 33,687 A-to-G sites (42%; total: 80,127 sites).
- 20,283 A-to-G sites in 3,630 genes (the rest: poorly annotated regions?)
  - 156 in CDS;
  - 14,958 in intronic regions;
  - 513 in non-coding exonic regions;
  - 1,115 in non-coding introns;
  - 3,452 in 3'UTR;
  - 89 in 5'UTR;

Category of editing sites predicted based on mRNA/EST (contd.)

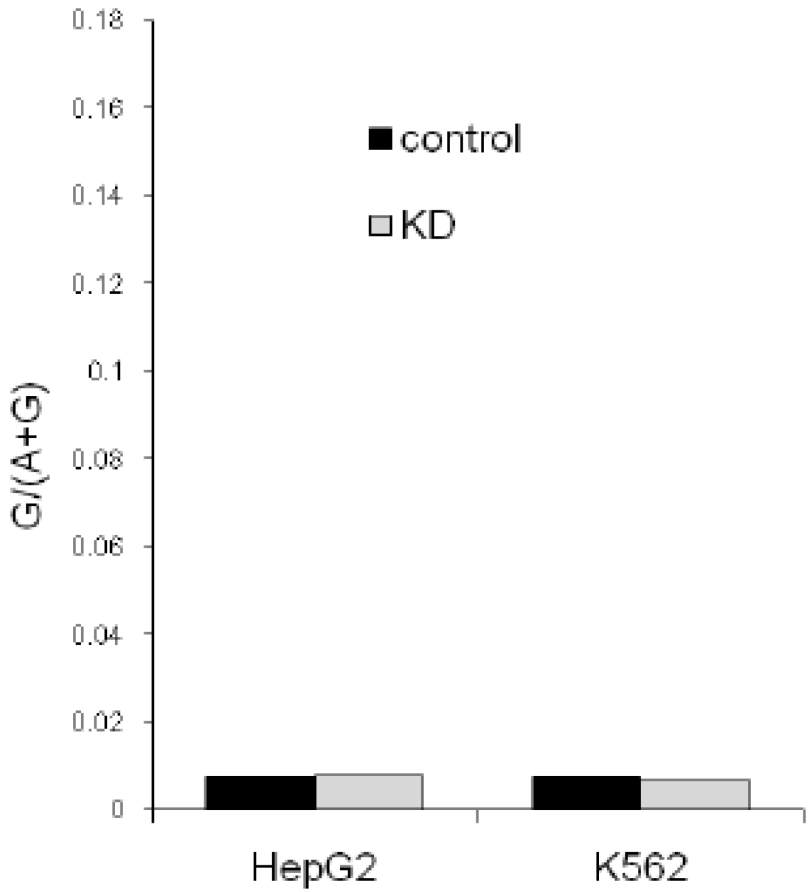


Reduction in editing level following ADAR KD based on RNA-seq



Reduction in editing level following ADAR KD (contd.)

- Adenosines not known to be edited (in edited genes).



Constitutive exons vs. cassette exons

- Categorize the editing sites in exons $\rightarrow$
  1. constitutive;
  2. cassette;
  3. other (e.g., retained introns, alternative 5' ss and 3' ss).

Constitutive exons vs. cassette exons (contd.)

- Enriched in cassette exons over constitutive exons.

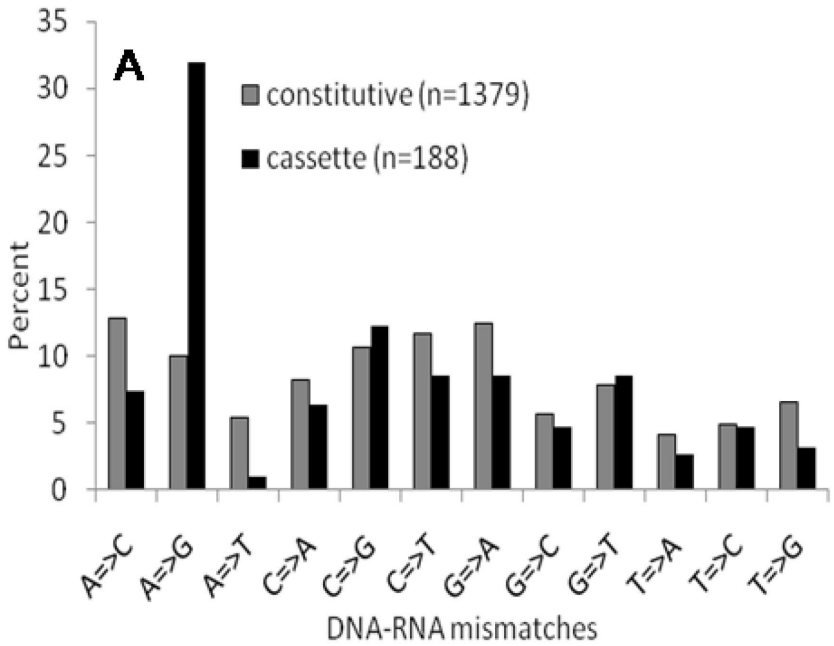
TABLE 1. A-to-I editing in constitutive and cassette exons

Exon type	No. of A-to-I sites[a]	Exon lengths (bases)	A-to-I per base
Constitutive	363	22,119,390	1.64×10^{-5}
Cassette	114	2,380,193	4.8×10^{-5}
Other	56	1,171,293	4.8×10^{-5}

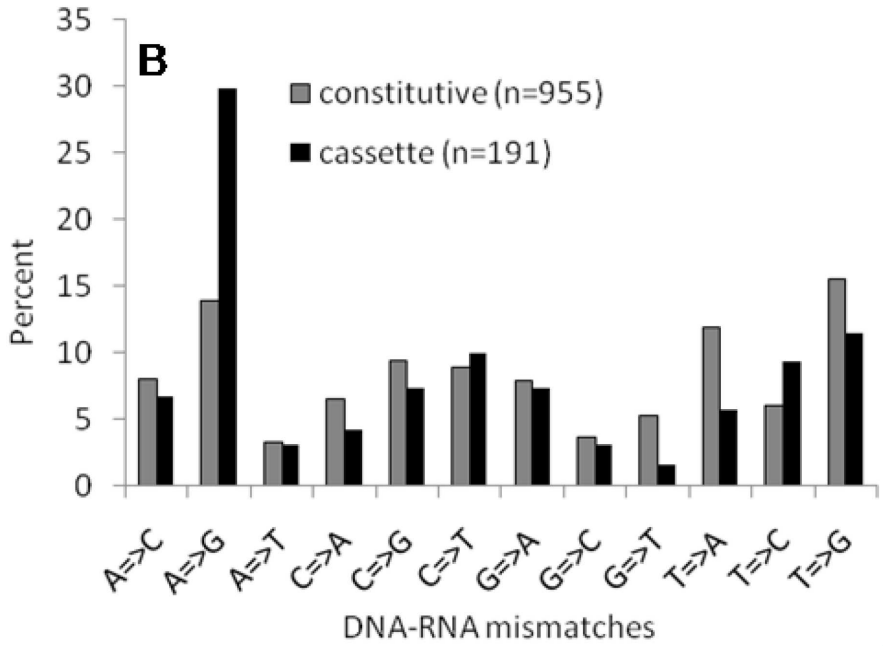
Overall, editing sites are enriched in cassette exons when normalized to the combined exon lengths (χ^2 P-value $< 10^{-10}$). Similar results were found using different editing sites sets (see Supplement 1, Supplemental Table S4).
[a]Editing sites were taken from DARNED (Kiran and Baranov 2010).

Constitutive exons vs. cassette exons (contd.)

- Mismatches in exons (mRNA/EST).



- Mismatches at 150-base intronic regions flanking exons (mRNA/EST).



Constitutive exons vs. cassette exons (contd.)

- A possible explanation:
  - Cassette and their flanking introns contain more *Alu* repeats.
- Nevertheless, counting only non-*Alu* editing sites:
  - cassette vs. constitutive:
    1.52×10^{-5} sites/base vs. 1.29×10^{-5} sites/base. (not significant)

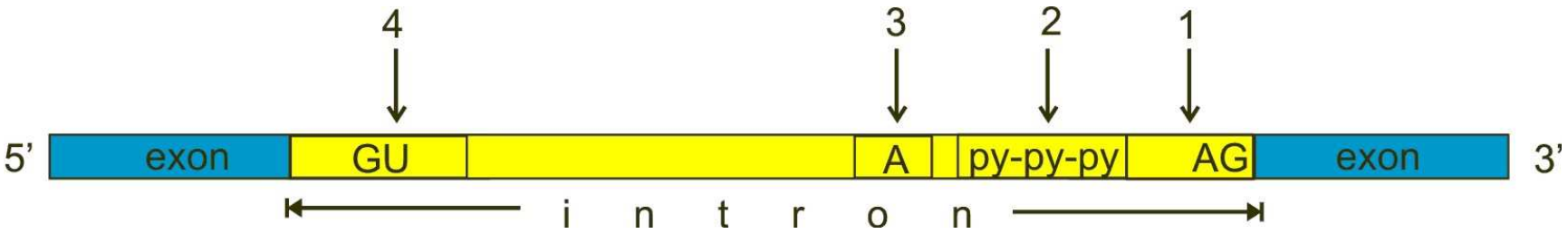
Constitutive exons vs. cassette exons (contd.)

- A possible explanation:
  - Cassette and their flanking introns contain more *Alu* repeats.
- Nevertheless, counting only non-*Alu* editing sites:
  - cassette vs. constitutive:
    1.52×10^{-5} sites/base vs. 1.29×10^{-5} sites/base. (not significant)

Constitutive exons vs. cassette exons (contd.)

- A possible explanation:
  - Cassette and their flanking introns contain more *Alu* repeats.
- Nevertheless, counting only non-*Alu* editing sites:
  - cassette vs. constitutive:
    1.52×10^{-5} sites/base vs. 1.29×10^{-5} sites/base. (not significant)

Fundamental sequence motifs required for canonical splicing



1: 3' splice site (5' ss);
2: poly-pyrimidine-tract (PPT);
3: branch point
   - branch site (BS) consensus motif: yUnAy);
4: 5' splice site (5' ss).

Editing sites rarely fall within the primary consensus sites of canonical splicing

- For editing sites within 5' ss or 3' ss consensus motifs:
  - Only 3 and 2 are within 5' ss and 3' ss resp (mRNA/EST).
  - Only 24 are within 5' ss or 3' ss for ALL data.
- For editing sites overlapping the branch-site consensus motif yUnAy:
  - None.

Splicing regulatory elements (Data preparation)

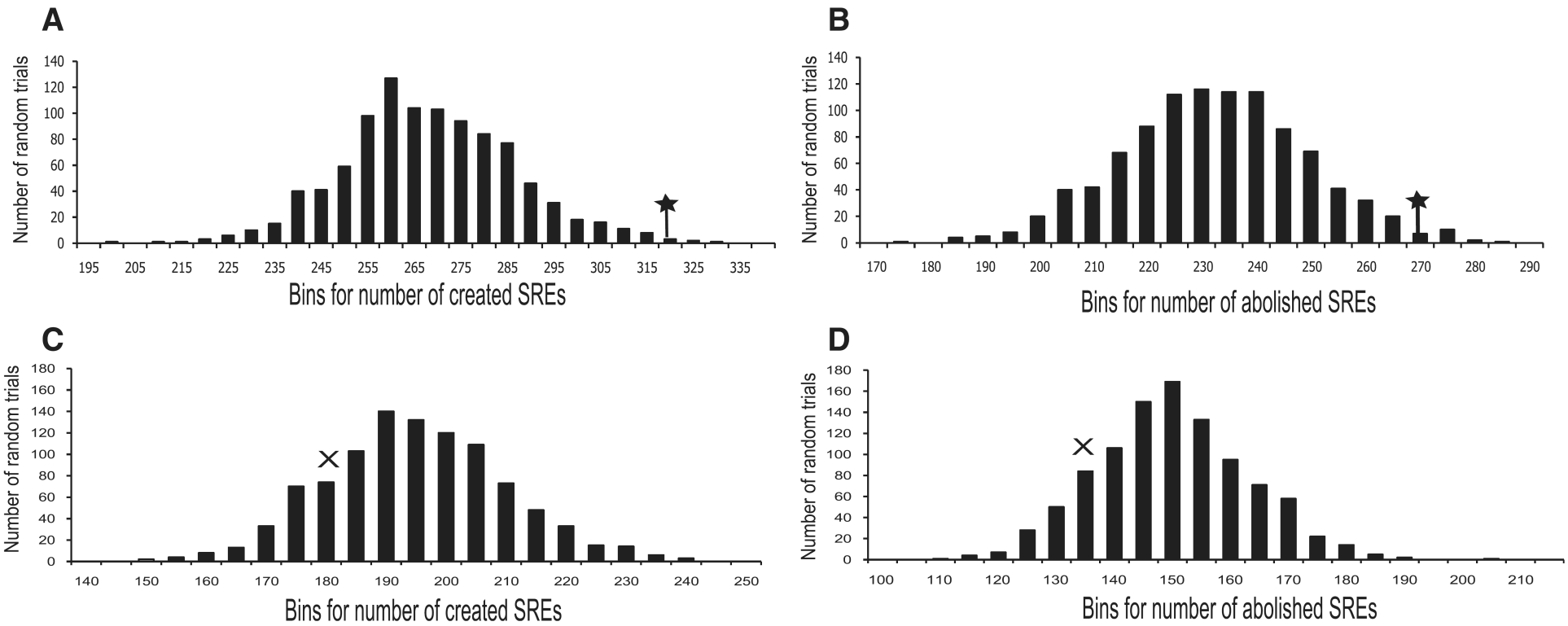
- Experimentally proved SREs:
  - Akerman *et al*. [*Genome Biol.* 2009].
  - Piva *et al*. [*Bioinformatics* 2009].
- ComputationI verification:
  - `Skippy` [Woolfe *et al*. *Genome Biol.* 2010].

Created/abolished SREs

- ⋆ Splicing factor proteins (e.g., hnRNPs and SR-proteins) bind to SREs in mRNA and regulate the type and the level of intron inclusion/exclusion.
- The approach: Searching for potential SREs in exons that have ≥ 1 putative editing site.
  - Control: randomly select adenosines in the exon and then change them to guanosines.
    - Count the cases where the changes create or abolish SREs.
  - Repeat the process for 1,000 times.
- The same procedure was applied for SREs located within the 150-base intronic regions flanking the exons.

Created/abolished SREs

- ⋆ Splicing factor proteins (e.g., hnRNPs and SR-proteins) bind to SREs in mRNA and regulate the type and the level of intron inclusion/exclusion.
- The approach: Searching for potential SREs in exons that have ≥ 1 putative editing site.
  - Control: randomly select adenosines in the exon and then change them to guanosines.
    - Count the cases where the changes create or abolish SREs.
  - Repeat the process for 1,000 times.
- The same procedure was applied for SREs located within the 150-base intronic regions flanking the exons.



- **Note:** Only 244 editing sites within 193 exons of 188 genes, while AS analysis of RNA-seq and microarray data revealed thousands of AS regions.

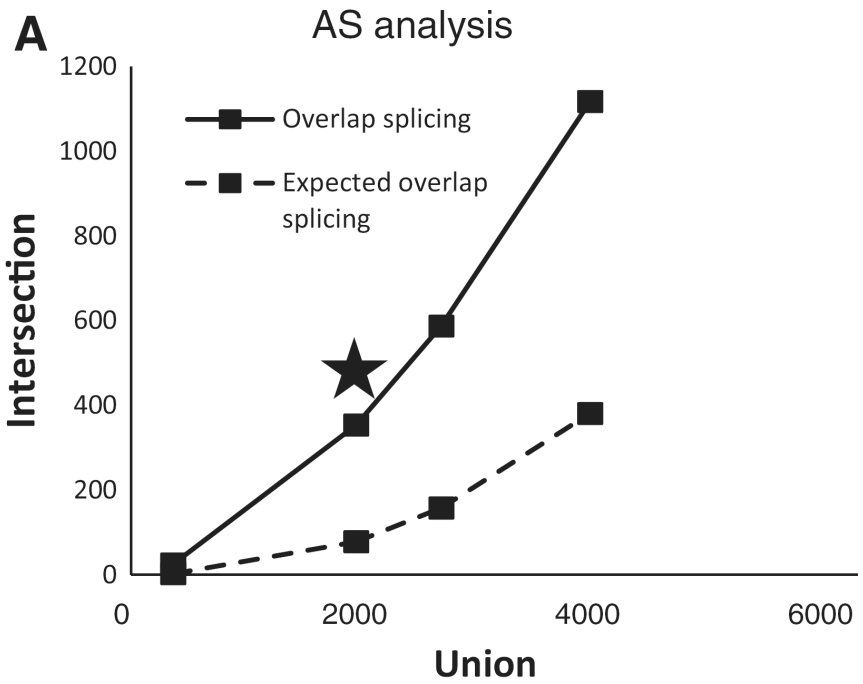
Microarray analysis

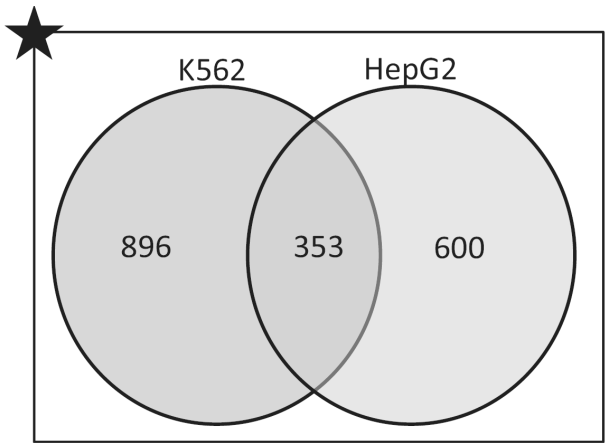
- Tools:
  - Partek genomic suite (for microarray data).
  - AS ANOVA (for splicing analysis) & express ANOVA (for total gene expression level).
- Only genes without a large change in expression (fold-change ≤ 1.5) were considered in the AS analysis.
- Lists of AS genes were extracted using four different significant thresholds (AS ANOVA P-values of 0.001, 0.01, 0.015, 0.025).
- Under the independence assumption:

$$N_{\text{exp}} = \left(\frac{N_{\text{HepG2}}}{N}\right) \times \left(\frac{N_{\text{K562}}}{N}\right) \times N.$$

  - N: the total number of genes represented on the array.
  - N_{HepG2}: the number of detected genes in HepG2.
  - N_{K562}: the number of detected genes in K562.

Microarray analysis (contd.)





Differential expression & alternative splicing analysis of RNA-seq

- Both are done using AltAnalyze.

  Emig *et al.* *Nucleic Acids Research* 2010 & Salomonis *et al.* *Proc. Natl. Acad. Sci.* 2010.

- Differential expression analysis:
  - Count the reads for each constitutive exons in Ensembl genes (V.62) and use AltAnalyze quintile normalization.
  - Use fold-change ratios of 1.4, 2, and 2.5.
  - # DE genes in both cell lines (as they did in the microarray analysis).

A simple exmaple of quintitle normalization

A	5	4	3
B	2	1	4
C	3	4	6
D	4	2	8

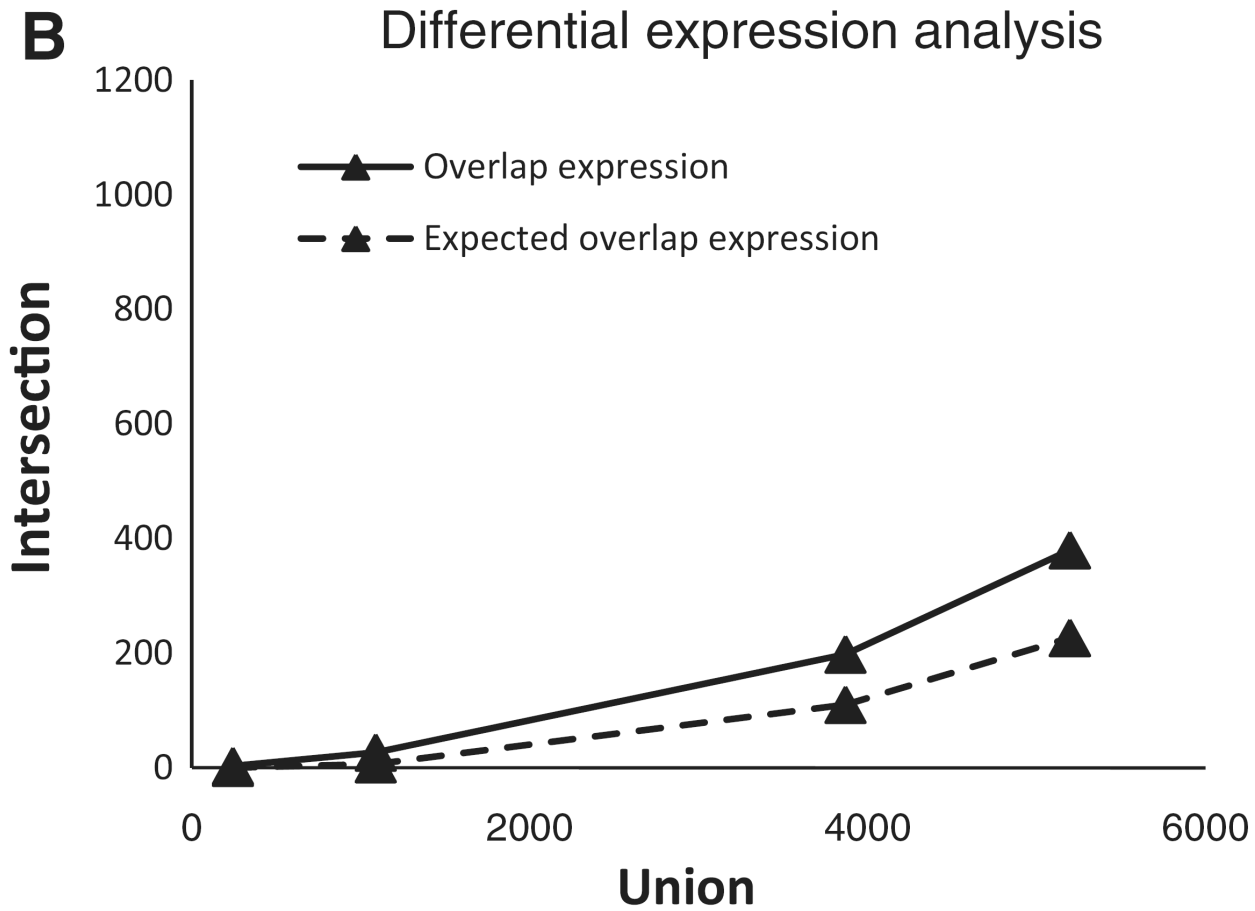
$\Longleftrightarrow$

A	iv	iii	i
B	i	i	ii
C	ii	iii	iii
D	iii	ii	iv

```
A    5    4    3    becomes A 2 1 3
B    2    1    4    becomes B 3 2 4
C    3    4    6    becomes C 4 4 6
D    4    2    8    becomes D 5 4 8
```

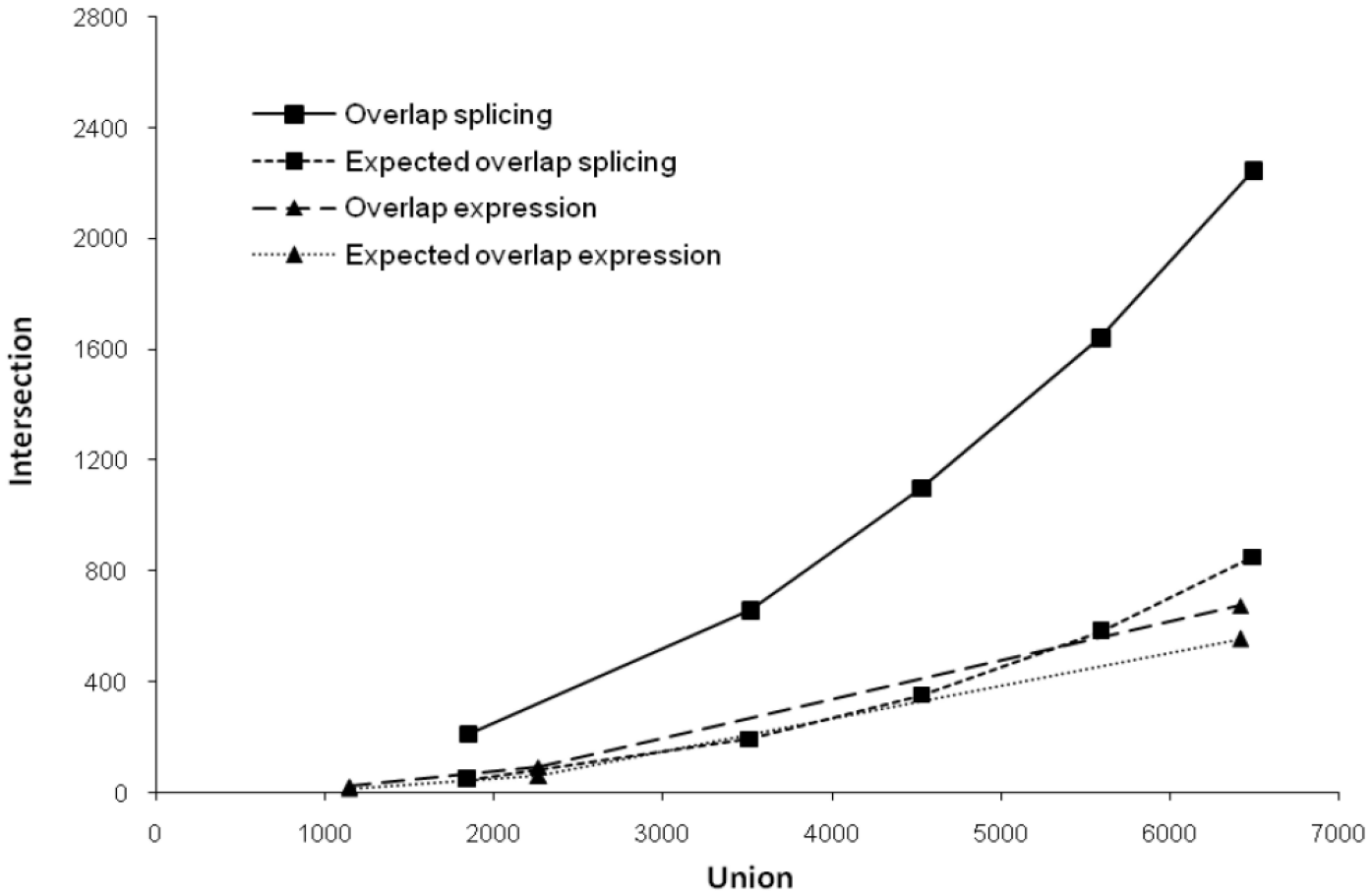
```
A (2 1 3)/3 = 2.00 = rank i
B (3 2 4)/3 = 3.00 = rank ii
C (4 4 6)/3 = 4.67 = rank iii
D (5 4 8)/3 = 5.67 = rank iv
```

A	5.67	4.67	2.00
B	2.00	2.00	3.00
C	3.00	4.67	4.67
D	4.67	3.00	5.67


B
Differential expression analysis
Overlap expression
Expected overlap expression
Intersection
1200
1000
800
600
400
200
0
0
2000
4000
6000
Union

Differential expression & alternative splicing analysis of RNA-seq (contd.)

- Alternative splicing analysis:
  - Base on alignment and junctions obtained by TopHat.
  - Using AltAnalyze
    - Splice-index (SI).
      - ⋆ Expression of a given exon relative to the total gene expression.
    - Analysis of splicing by isoform reciprocity (ASPIRE) [Ule *et al*. *Nature Genetics* 2005].
      - ⋆ Using reads mapped to junctions.
- Only exons detected as AS by both methods and whose genes were NOT found to be differentially expressed ($|\text{fold-change}| \leq 2$) were considered as AS exons.
  - Genes with ≥ 1 AS exon were considered as AS genes.


Overlap splicing
Expected overlap splicing
Overlap expression
Expected overlap expression
Intersection
Union
0
400
800
1200
1600
2000
2400
2800
0
1000
2000
3000
4000
5000
6000
7000

Are genes that significantly change their splicing pattern in ADAR KD cells enriched with A-to-I RNA editing sites?

TABLE 2. Editing in AS genes detected by RNA-seq

Cell line	Splicing pattern	Number of A-to-I sites[a]	Genomic size (bases)	A-to-I per base
HepG2	Changed	12,043	338,829,808	3.5×10^{-5}
HepG2	Unchanged	15,973	731,790,993	2.18×10^{-5}
K562	Changed	12,971	343,111,651	3.78×10^{-5}
K562	Unchanged	14,897	679,627,158	2.19×10^{-5}

Editing sites are enriched in AS genes (χ^2 P-value $< 10^{-10}$ for both HepG2 and K562).
[a]Editing sites were taken from DARNED (Kiran and Baranov 2010).

- It is still difficult to determine if editing mediates the AS regulation, or whether other ADAR-dependent events are responsible for this effect on AS.
  - Based on the RNA-seq experiments, only 17 verified editing sites in the vicinity (1,000 bases flanking) of the detected AS cassette exons are detected.

⋆ Direct editing is not the sole explanation for AS.

Significant changes in RNA processing and splicing machinery genes?

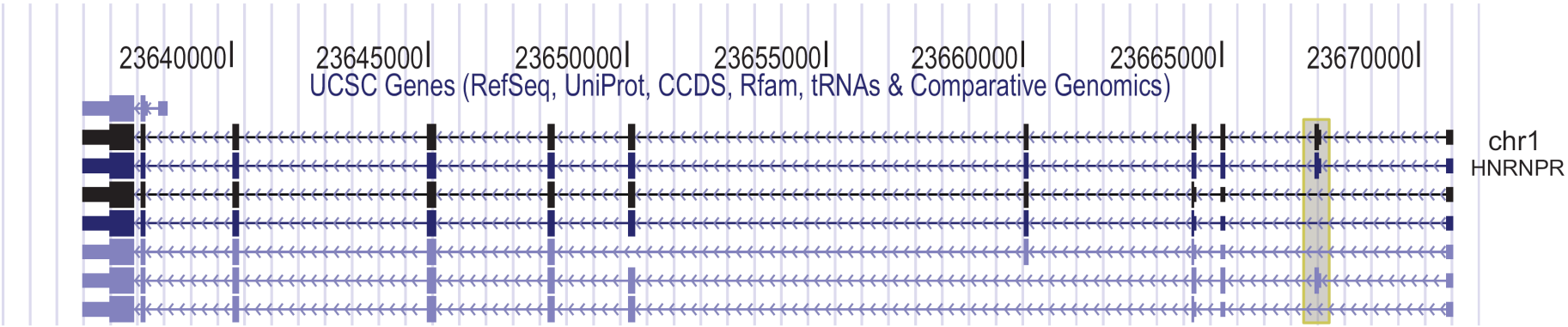
TABLE 3. Detected AS genes are enriched with RNA processing and splicing functions

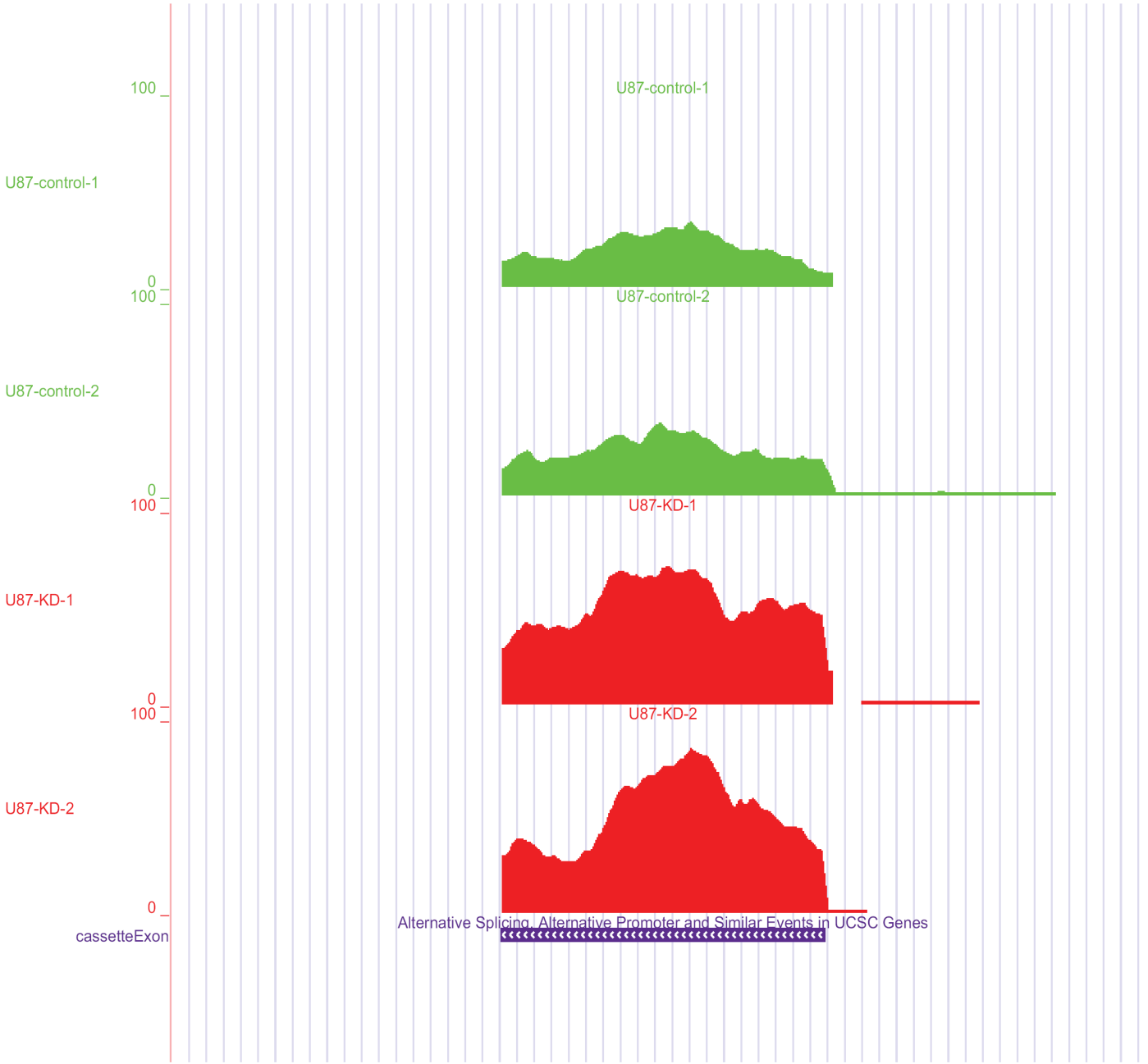
GO		HepG2		K562		U87MG	
ID	Term	No. of AS genes	FDR	No. of AS genes	FDR	No. of AS genes	FDR
6396	RNA processing	229	1.2×10^{-17}	246	2.8×10^{-19}	130	5.4×10^{-9}
16071	mRNA metabolic process	153	2.9×10^{-10}	168	1.0×10^{-12}	95	7.8×10^{-9}
3723	RNA binding	290	6.0×10^{-20}	305	2.1×10^{-20}	152	8.8×10^{-7}
8380	RNA splicing	123	2.0×10^{-9}	128	8.4×10^{-9}	80	2.5×10^{-8}
398	Nuclear mRNA splicing, via spliceosome	71	5.6×10^{-6}	71	1.8×10^{-4}	43	0.002

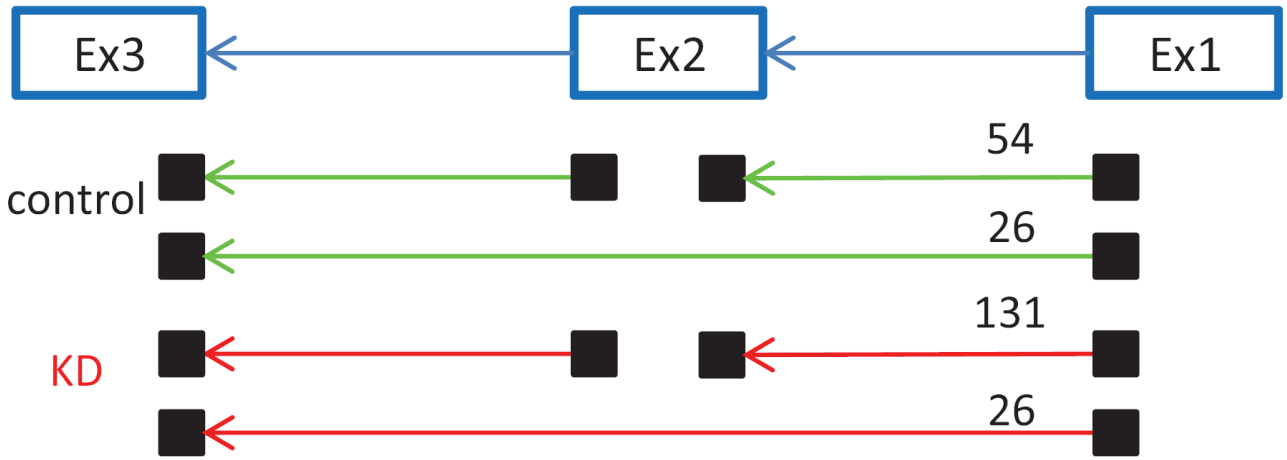
All three RNA-seq data sets (HepG2, K562, and U87MG) show significant enrichment for these functions.

Significant changes in RNA processing and splicing machinery genes?

- ADAR-dependent AS in *HNRNPR* based on U87MG RNA seq [Bahn et al. *Genome Res.* 2012].




100
U87-control-1
U87-control-1
0
100
U87-control-2
U87-control-2
0
100
U87-KD-1
U87-KD-1
0
100
U87-KD-2
U87-KD-2
0
cassetteExon
Alternative Splicing, Alternative Promoter and Similar Events in UCSC Genes



- The KD sample contains more reads supporting inclusion of the second exon than reads supporting its exclusion (P-value: 0.0075).

Summary of the results

- A-to-I RNA editing rarely targets canonical splicing motifs, yet it affects SREs within exons.
- Cassette exons were found to be significantly enriched with A-to-I editing sites compared with constitutive exons.
- ADAR knockdown in HepG2 & K562 cell lines leads to global changes in gene expression and hundreds of genes changing their splicing pattern, yet this cannot be explained by putative editing sites alone.
- Genes showing significant changes in their splicing pattern are frequently involved in RNA processing & splicing activity.
- The primary mechanism for ADAR-mediated regulation of AS is suggested to be achieved by modulating *trans*-acting factors.

Thank you.