

**A Disrupted RNA Editing Balance Mediated by ADARs in Human Hepatocellular
Carcinoma**

Tim Hon Man Chan, Chi Ho Lin, Lihua Qi, Jing Fei, Yan Li, Kol Jia Yong, Ming Liu,
Yangyang Song, Raymond Kwok Kei Chow, Vanessa Hui En Ng, Yuan-Fei Yuan, Daniel G.
Tenen, Xin-Yuan Guan, Leilei Chen

Supplementary Figures 1-4, Tables 1-8

MATERIALS AND METHODS

RNA extraction and Illumina mRNA library preparation

Tumor and their adjacent non-tumorous (NT) tissues of 3 HCC patients (Case No. 448, 473 and 510) in GZ cohort were selected for RNA-Seq. All 3 patients are HBV-positive and HCV-negative. Total RNA was isolated using the mirVanaTM miRNA isolation kit (Ambion, Austin, TX, USA), and the total RNA was treated with the DNA-free kit (Ambion) for the removal of contaminated genomic DNA. PolyA⁺ RNA was purified using Dynabeads mRNA purification kit (Invitrogen, Carlsbad, CA) following the manufacturer's instructions. Approximately 100 ng of mRNA was fragmented by incubation for 5 min at 94°C in 5 × Array Fragmentation Buffer (Ambion). Double stranded cDNA was synthesized using the SuperScript Double-Stranded cDNA Synthesis kit (Invitrogen) using random hexamers. The reaction was purified using a QiaQuick PCR column (Qiagen, Valencia, CA).

Double-stranded cDNA fragments were repaired using the DNA Terminator End Repair Kit (Lucigen, Middleton, WI) and purified using a QiaQuick PCR column. The Klenow 3' to 5' exo-polymerase (NEB, Ipswich, MA) was used to add a single 'A' base to the 3' end of blunt phosphorylated DNA fragments. Following purification, the Illumina PE Adapter (Illumina, San Diego, CA) was ligated to the end of the DNA fragments using the Quick LigationTM Kit (NEB, Ipswich, MA). DNA fragments ranging from approximately 280 to 300 bp were excised from a 2% low-melting agarose gel. The fragments were enriched by 10 thermocycles using AccuPrimeTM Pfx DNA Polymerase (Invitrogen). The PCR product

was run on a Novex 8% TBE polyacrylamide gel (Invitrogen) and stained with SYBR Gold (Invitrogen). Gel slice containing the 340- to 360-bp fragments were excised and purified using the QiaQuick Gel Extraction Kit (Qiagen). The concentration of the gel-purified DNA fragments was measured using a ND-1000 UV/Vis spectrophotometer (NanoDrop Technologies, Wilmington, DE).

Solexa sequencing and read mapping

Cluster generation and sequencing were conducted using the Standard Cluster Generation kit v4, and 36-Cycle Sequencing kit v3 on the Illumina Cluster Station and GAIIx following the manufacturer's instructions. cDNA libraries from 3 paired HCC tumors and the corresponding NT counterparts (HCC448N/T, HCC473N/T, and HCC510N/T) were sequenced with 58-base single-reads. Raw data from the GAIIx were analyzed using the Illumina Real Time Analysis (RTA) v1.6 software. A phi-X 174 control lane was included in each Solexa run for matrix, phasing, and error rate estimations as recommended by the manufacturer. The error rate of the Phi-X control error rate was < 0.28% for all of the sequencing runs.

Ribosomal RNA sequences were first removed from GA reads by aligning them to 28S (NCBI RefSeq accession NR_003287.2), 18S (NCBI RefSeq accession NR_003286.2), human ribosomal DNA complete repeating unit (HSU13369) and mitochondrial ribosomal RNA (Ensembl transcript ID ENST00000387347 and ENST00000389680) using Bowtie¹ with default parameters. The high-quality reads were then aligned against the human genome assembly (NCBI Build 37.1/hg19) using TopHat v1.0.14² with the RefSeq refGene annotation,

which was downloaded from the UCSC Genome Browser.³ Finally, mapping results were processed with custom scripts and visualized on the UCSC Genome Browser as a custom track.

Plasmid construction

We amplified the full-length cDNA encoding ADAR1 p110 isoform (ADAR1 p110) by PCR with primers (F: 5'-CACCGAAAGAGGCAGGAACACCC-3'; R: 5'-CTATACTGGGCAGAGATAAAAAGTTC-3'). The full-length cDNA encoding ADAR2 was amplified by PCR with primers (F: 5'-CACCATGGATATAGAAGATGAAGAAAAC-3'; R: 5'-TCAGGGCGTGAGTGAGAACTGG-3'). Subsequently, the purified *ADAR1 p110* or *ADAR2* PCR products were ligated to pLenti6/V5-TOPO®vector (*pLenti6-ADAR1 p110* or *pLenti6-ADAR2*) (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions.

Establishment of ADAR1 or ADAR2 overexpression cells using a lentiviral expression system

Either the *pLenti6-ADAR1 p110* or *pLenti6-ADAR2* expression construct or the empty pLenti6/V5 vector was transfected into the 293FT cell line. Virus-containing supernatants were collected for subsequent transduction into SNU-423 cells. At 48 hours after virus transduction, half of cells were collected for the relevant assays. The remaining cells were cultured in complete medium containing blasticidin (3 µg/mL; Sigma-Aldrich) for

establishing stably transduced cell lines.

ADAR1 and ADAR2 knockdown/rescue experiments

The *ADAR1* and *ADAR2*-specific shRNA expression vectors (pRS-shADAR1 or pRS-shADAR2) and scrambled non-effective shRNA cassette in the pGFP-V-RS plasmid (pRS-scramble) were purchased from OriGene Technologies, Inc (Rockville, MD). The sequences of the shRNAs directed against *ADAR1* or *ADAR2* were as follows: shADAR1: CCTGTGGAATCCAGTGACATTGTGCCTAC and shADAR2: ACTCAAGTATGACTTCCTCTCCGAGAGCG. The pRS-shADAR1 or pRS-shADAR2 construct was transfected into SUN-423 cells stably expressing ADAR1 p110 or ADAR2 (423-AR1 or 423-AR2), respectively, using Lipofectamine 2000 (Invitrogen). The pRS-scramble construct was transfected into cells as a negative control. To generate an *ADAR1 p110* or *ADAR2* mutant construct preserving the native amino acid sequence, we introduced six mutations into the shADAR1 or shADAR2 targeting sequence (29nt), respectively. PCR-directed mutagenesis was performed using an inner forward or reverse primer containing six nucleotide alterations (AR1-rescue-forward: 5'-GAGAACGGAGAAGGCACAATCCCAGTAGAGTCAAGCGATATT -3'; AR1-rescue-reverse: 5'-AATATCGCTTGAGTCTACTGGGATTGTGCCTTCTCCGTTCTC-3'; AR2-rescue-forward: 5'-TTGAACGAACTGCGCCAGGACTGAAATACGATTTTCTGTCC-3';

AR2-rescue-reverse: 5'-

GGACAGAAAATCGTATTTTCAGTCCTGGGCGCAGTTCGTTCAA-3') with the corresponding external primers described above.

Twenty-four hours after transfection, the transfected cells were cultured for 3 days with 0.4 µg/mL puromycin (OriGene). Pooled populations of knockdown/rescue cells were subjected to *in vitro* experiments.

Complementary DNA (cDNA) synthesis and quantitative real-time PCR (QPCR)

Total RNA was isolated as described above. To quantify the *ADAR1* and *ADAR2* expression levels in clinical samples, equal amounts of cDNA were synthesized using the Advantage RT-for-PCR kit (Clontech, Mountain View, CA) and used for QPCR analysis. QPCR was performed using the SYBR Green PCR master mix (Applied Biosystems) and the following primers:

qADAR1-F(5'-CCCTTCAGCCACATCCTTC-3'), qADAR1-R(5'-GCCATCTGCTTTGCCACTT-3'), qADAR2-F(5'-CTGACACGCTCTTCAATGGTT-3') and qADAR2-R(5'-GGCGCAGTTCGTTCAAGAT-3'). 18S was amplified as an internal control using the following primers: q18S-F: 5'-CTCTTAGCTGAGTGTCCTCCGC-3'; q18S-R:

5'-CTGATCGTCTTCGAACCTCC-3'. PCR was performed using an ABI Prism 7900 System (Applied Biosystems), and data processing was performed using the ABI SDS v2.3 software (Applied Biosystems). For *ADAR1* and *ADAR2* expression in HCC and matched NT liver specimens, the relative target gene expression is indicated by $2^{-\Delta\Delta C_T}$ ($\Delta\Delta C_T = \Delta C_T$

$\Delta C_T^{\text{Tumor}} - \Delta C_T^{\text{Non-tumor}}$) and normalized to the average relative expression level in all of the NT tissues, which was defined as 1.0.

Gene Ontology Analysis

DAVID⁴ was used for functional enrichment analysis of the edited genes obtained from the RNA-Seq.

Focus Formation Assay

Briefly, 1×10^3 cells were seeded in a 6-well plate. After culture for 7 days, surviving colonies (>50 cells per colony) were counted and stained with crystal violet (Sigma-Aldrich). Triplicate independent experiments were performed and the data were expressed as the mean \pm SD of triplicate wells within the same experiment.

Cell migration assay

The transwell cell migration assay was performed using Bio-coat cell migration chambers (BD Biosciences) containing polyethylene terephthalate membranes (PET) of 8- μ m pore size according to the manufacturer's instructions. Briefly, 5×10^4 cells in FBS-free RPMI were added. RPMI supplemented with 20% FBS was added to the bottom chamber as a chemoattractant. After 24 hours, the number of cells that had migrated through the filter pores was counted in 10 fields under a 20 \times objective lens and imaged using SPOT imaging software (Nikon, Japan).

Matrigel Invasion Assay

We performed invasion assays using 24-well BioCoat Matrigel Invasion Chambers (BD Biosciences) according to the manufacturer's instructions. Briefly, 2×10^5 cells FBS-free RPMI were added to the top chamber, and 20% FBS in RPMI was added to the bottom chamber as a chemoattractant. After 22 hours of incubation, cells that invaded the Matrigel were fixed and stained with crystal violet (Sigma-Aldrich). The number of cells was counted in 10 fields under a 20 \times objective lens and imaged using SPOT imaging software (Nikon, Japan).

***In Vivo* Tumorigenicity Assay**

We subcutaneously injected approximately 1×10^7 cells into the right flank of 4- to 5-week-old male severe combined immunodeficient (SCID) mice. We monitored tumor formation in the SCID mice over an 8-week period and calculated the tumor volume weekly by the formula V (volume) = $0.5 \times L$ (length) $\times W$ (width) $\times W$. All animal experiments were approved by and performed in accordance with the Institutional Animal Care and Use Committees of National University of Singapore.

Antibodies and Western blot analysis

Mouse anti-ADAR1 and anti- β -actin antibodies were purchased from Abcam (Cambridge,

MA). The mouse anti-ADAR2 and anti-GAPDH antibodies were purchased from Sigma-Aldrich (Missouri, USA) and Santa Cruz Biotechnology. Protein lysates were quantified and resolved on a SDS-PAGE gel, transferred onto a polyvinylidene difluoride (PVDF) membrane (Millipore, Billerica, MA, USA), and immunoblotted with a primary antibody, followed by incubation with a secondary antibody. The blots were visualized by enhanced chemiluminescence (GE Healthcare, Buckinghamshire, UK).

Immunohistochemical staining (IHC)

The tissue microarray (TMA) blocks were sectioned (5 mM thick) for IHC staining. Briefly, sections were deparaffinized and rehydrated. The endogenous peroxidase activity was blocked with 3% hydrogen peroxide (H₂O₂) for 10min. For antigen retrieval, the slides were immersed in 10 mM citrate buffer (pH 6.0) and boiled for 15 min in a microwave oven. Non-specific binding was blocked with 5% normal goat serum for 10min. The slides were incubated in a 1:100 dilution of anti-ADAR1 (Abcam) at 4°C overnight in a humidified chamber. The slides were then sequentially incubated with biotinylated goat anti-mouse IgG (1:100 dilution, Santa Cruz) for 30 min at room temperature, streptavidin-peroxidase conjugate for 30 min at room temperature. Isotope-matched human IgG was used in each case as a negative control. Finally, the 3, 5-diaminobenzidine (DAB) Substrate Kit (Dako Ltd., Carpinteria, CA) was used for color development followed by Mayer's hematoxylin counterstaining. Based on staining intensities, the ADAR1 immunoreactivity was scored as negative (0) (total absence of staining), weak expression (1) (faint staining in <50%, or

moderate staining in <25% of tumor cells), moderate expression (2) (moderate staining in $\geq 25\%$ to <75%, or strong staining in <25% of tumor cells), and strong expression (3) (moderate staining in $\geq 75\%$, or strong staining in $\geq 25\%$ of tumor cells).

REFERENCES

1. **Langmead B**, Trapnell C, Pop M, *et al.* Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009;**10**:R25.
2. **Trapnell C**, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics.* 2009;**25**:1105-11.
3. **Zweig AS**, Karolchik D, Kuhn RM, *et al.* UCSC genome browser tutorial. *Genomics.* 2008;**92**:75-84.
4. **Huang da W**, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* 2009;**4**:44-57.

FIGURE LEGENDS

Supplementary Figure 1. (A) Distribution of the editing levels of potential A-to-I (G) editing sites. The data are presented as dot plots and indicate the mean \pm SD of all 6 samples (3 pairs of primary HCC and matched NT liver specimens). (B) Validation of inferred editing sites from RNA-Seq by Sanger sequencing. The sequencing chromatograms of 5 representative gene loci are shown. The editing positions are indicated by arrows. The top trace is genomic DNA (gDNA), and the bottom trace is cDNA.

Supplementary Figure 2. The sequence chromatograms of the *AZINI*, *FLNB*, *COPA* and *UTP14C* gDNA sequences in the indicated tumor and matched NT liver samples. An arrow indicates the editing position.

Supplementary Figure 3. (A) The numbers of potential editing sites within *Alu* sequences in 3 pairs of primary HCC and matched NT liver specimens. (B) Validation of inferred editing sites from RNA-Seq by Sanger sequencing. The editing positions are indicated by arrows.

Supplementary Figure 4. (A) XTT proliferation assays showing growth rates of the indicated stable cell lines. Triplicate independent experiments were performed and the data

were expressed as the mean \pm SD of triplicate wells within the same experiment (Unpaired, two-tailed Student's *t* test). **(B, C)** Quantification of cells from the indicated cells that invaded through Matrigel-coated membrane **(B)** or migrated through the polyethylene terephthalate (PET)-membrane **(C)** (Unpaired, two-tailed Student's *t* test). Scale bar, 200 μ m.