Dissertation

Inosine chemical labeling and molecular purification of edited RNA & DNA using maleimide

（マレイミドによる塩基編集を持つ核酸におけるイノシンの化学標識と分子精製技術の開発）

March 2024

YANG YUXI

Table of Contents

Overview

Background Adenosine-to-inosine (A-to-I) RNA editing, a base modification process catalyzed by adenosine deaminase acting on double-stranded RNA (ADAR), converts adenosine to inosine. Typically, adenosine pairs with uridine, but following editing, inosine preferentially pairs with cytidine, altering the genetic code and affecting the function of various transcripts. Recent findings indicate that ADAR can edit not only the canonical double-stranded RNA substrates but also RNA:DNA hybrid strands. Our current research focuses on determining whether A-to-I DNA editing, potentially the first endogenous RNA-dependent DNA adenosine deamination mechanism, occurs in human cells. However, because of the low abundance of inosine sites in the cells, identifying inosine in genomic DNA seemed to require higher sensitivity.

Purpose Therefore, we aim to develop a new method to accurately capture very subtle inosines in RNA and DNA molecules.

Results We investigated the potential of maleimide as a label to react with inosine, thereby establishing for the first time the chemical interaction between maleimide and inosine. The numerous available maleimide derivatives enable adaptable and innovative inosine labeling. Using fluorescein–maleimide derivatives, we confirmed the reactivity of maleimide with inosine and ascertained the efficiency and specificity of maleimide in labeling inosine within nucleic acids. Consequently, biotin–maleimide derivatives were further used to enrich RNA and DNA containing inosine. We termed this novel biochemical approach as Inosine Chemical Labeling and Affinity Molecular Purification (ICLAMP). This method could offer groundbreaking advantages over existing methods and provides a versatile platform with the potential to identify novel A-to-I editing sites in RNA and DNA.

Introduction

1. A-to-I editing

RNA post-transcriptional processing events, including alternative splicing, polyadenylation, capping, and RNA modification, determine the complexity of the transcriptome (Blencowe, 2006; Di Giammartino, Nishida and Manley, 2011). All four RNA bases and ribose sugars can be targeted for modification (Barbieri & Kouzarides, 2020). Interactions between different modifications may produce functional consequences (Rengaraj *et al.*, 2021). RNA editing is a ubiquitous and crucial post-transcriptional modification that alters the chemical structure of the base and ribose and can alter the base-pairing pattern, resulting in changes in genetic information, RNA structure, and functions, leading to cell fate regulation. Adenosine (A) to inosine (I) (A-to-I) RNA editing is a type of RNA editing that was discovered in *Xenopus laevis* oocytes and embryos in 1987 (Bass and Weintraub, 1987; Rebagliati and Melton, 1987), and was soon found in mammals (Wagner and Nishikura, 1988), which is an enzymatic activity causing the unwinding of double-stranded RNA (dsRNA). It was confirmed that A-to-I RNA editing was carried out by double-stranded RNA-specific adenosine deaminase (ADAR) enzymes (Figure 1) (Bass and Weintraub, 1987; Wagner and Nishikura, 1988; Melcher, Maas, Herb, Sprengel, Higuchi, *et al.*, 1996; Melcher, Maas, Herb, Sprengel, Seeburg, *et al.*, 1996). ADAR-mediated A-to-I RNA editing is dynamically regulated by various mechanisms to ensure cellular and organismic homeostasis (Vesely and Jantsch, 2021). The A-to-I RNA editing sites initially identified are located in the protein-coding region of the mRNA (Sommer *et al.*, 1991). Because inosine has a chemical structure similar to that of guanosine (G), inosine in the RNA transcripts acts as guanosine and pairs with cytidine during mRNA splicing and translational processes. With the development of next-generation sequencing (NGS) technologies and bioinformatics, the systematic exploration of new editing sites has been one of the canonical experiments in the A-to-I RNA editing research field (Li *et al.*, 2011). In addition to the discovery of editing sites in protein-coding sequences, widespread editing sites have also been found in non-coding RNAs, such as microRNAs (miRNAs) (Yang *et al.*, 2006; Kawahara *et al.*, 2008) and repetitive elements Alu (SINE), located in introns and untranslated regions (Athanasiadis, Rich and Maas, 2004; Sakurai *et al.*, 2010a; Bazak *et al.*, 2014). Abnormalities in coding and non-coding editing sites have been linked to brain development, viral infections, and human diseases (Erdmann *et al.*, 2020; Yang, Okada and Sakurai, 2021).

2. ADARs – the vital catalytic enzyme for A-to-I RNA editing

Adenosine-to-inosine (A-to-I) editing is catalyzed by adenosine deaminase acting on RNA (ADAR) enzymes. Three members of the ADAR family are encoded in the mammalian genome: ADAR1 (ADAR or DRADA) (Kim *et al.*, 1994), ADAR2 (ADARB1) (Melcher, Maas, Herb, Sprengel, Seeburg, *et al.*, 1996), and ADAR3 (ADARB2) (Melcher, Maas, Herb, Sprengel, Higuchi, *et al.*, 1996; Chen *et al.*, 2000) (Figure 2). ADAR1 and ADAR2 are ubiquitously expressed (Kim *et al.*, 1994; Melcher, Maas, Herb, Sprengel, Seeburg, *et al.*, 1996); however, ADAR3 expression is restricted to the brain (Melcher, Maas, Herb, Sprengel, Higuchi, *et al.*, 1996; Chen *et al.*, 2000). It has been demonstrated that ADAR1 and ADAR2 exhibit enzymatic activities (Kim *et al.*, 1994; Melcher, Maas, Herb, Sprengel, Seeburg, *et al.*, 1996). ADAR1 is responsible for most editing activities. No editing activity was observed for ADAR3 (Melcher, Maas, Herb, Sprengel, Higuchi, *et al.*, 1996; Chen *et al.*, 2000).

The domain structures of the three human ADARs are shown in Figure 2. Each protein had a catalytic deaminase domain in its C-terminal region. ADAR2 and ADAR3 contain two double-stranded RNA-binding domains (dsRDB) (Stefl *et al.*, 2006), whereas ADAR1 contains three domains. ADAR3 also has an Arg-rich, single-stranded RNA-binding domain in its N region (Chen *et al.*, 2000). In ADAR1, there is a nuclear localization sequence (NLS) located in the third dsRBD, and this sequence is located at the N-terminus of ADAR2 and ADAR3 (Strehblow, Hallegger and Jantsch, 2002; Desterro *et al.*, 2003). Furthermore, ADAR1 has two distinct isoforms: a shorter isoform, ADAR1p110, and a longer isoform, ADAR1p150. ADAR1p150 has a Z-DNA-binding domain (Zβ and Zα), whereas ADAR1p110 lacks a Zα domain at its N-terminus (Herbert *et al.*, 1997). A nuclear export sequence (NES) in the Zα domain (Poulsen *et al.*, 2001). ADAR2 and ADAR1p110 were mainly located in the nucleus. ADAR1p150 is located primarily in the cytoplasm but can shuttle between the nucleus and cytoplasm (Desterro *et al.*, 2003). In contrast to mammals, *Drosophila melanogaster* has only one *Adar* gene, which encodes a protein homologous to mammalian ADAR2 (Palladino *et al.*, 2000). In *Caenorhabditis elegans*, two ADAR proteins, ADR-1 and ADR-2, encoded by *adr-1* and *adr-2* genes, share common functional domains (Jin, Zhang and Li, 2009) (Figure 2).

One significant effect of ADAR1 in human diseases is its role in the innate immune system. Endogenous dsRNAs resemble virus-formed structures and activate the innate immune system by

activating melanoma differentiation-associated protein 5 (MDA5). Upon activation, MDA5 interacts with mitochondrial antiviral signaling proteins (MAVS), resulting in a severely harmful interferon (IFN) response. Embryonic lethality was discovered in mice homozygous for the ADAR1 null mutation. These mice also appeared to have massive apoptosis and an aberrant innate immune response (Wang *et al.*, 2004). The subsequent research supports this finding, suggesting that concurrent deletion of either the MDA5 receptor or the MAVS adaptor can rescue embryos from lethality (Mannion *et al.*, 2014; Liddicoat *et al.*, 2015). *Adar2*$^{-/-}$ mice are prone to seizures and death at a young age (within postnatal day 20) owing to the excessive influx of calcium ions through glutamate receptors (Higuchi *et al.*, 2000). Moreover, a study using *Adar2*-deficient mice proposed that Adar2-dependent A-to-I RNA editing is associated with circadian rhythms (Terajima *et al.*, 2017). However, ADAR3 lacks editing activity, and its insufficient catalytic ability has not yet been determined. A recent study introduced five-point mutations (A389V, V485I, E527Q, Q549R, and Q733D) in the ADAR3 deaminase domain that recovered ADAR3 editing activity, resulting in unique editing sites in transfected cells (Wang *et al.*, 2019). The analysis also assured ADAR3 binds to dsRNAs in living cells. This supports the observation that ADAR3 has a negative regulatory effect on RNA editing by competing with ADAR1 or ADAR2 on dsRNA substrates (Oakes *et al.*, 2017; Wang *et al.*, 2019). Mice lacking exon 3 of *Adar3* have increased anxiety levels and impaired hippocampal-dependent memory formation, suggesting that ADAR3 is involved in cognitive processes in mammals (Mladenova *et al.*, 2018).

3. A-to-I RNA editing sites in protein-coding sequences

A small number of gene transcripts are edited within their coding regions, creating new protein variants that contribute to the diversity of the proteome; this is classified as recoding-type editing. The total editing sites are more than 10 million in homo sapiens, while recoding sites are less than 0.15% (Duan, Tang and Lu, 2022). The effect of A-to-I RNA editing on the coding sequence depends on the position and mRNA species. The first and most studied recoding sites are located in the glutamate receptor (Sommer *et al.*, 1991). Mammalian glutamate receptor 2 (GRIA2) is a subunit of α-amino-3-hydroxy-5-methyl-4-isoxazole-propionic acid (AMPA) (Raymond *et al.*, 1999). The neurotransmitter glutamate released by the presynaptic density activates AMPA receptors, allowing the influx of Na^{+} or Ca^{2+} ions. ADAR2-mediated RNA editing changes glutamine codon (Q) to arginine codon (R) on GluA2 mRNA, resulting in the formation of a positively charged amino acid

residue that leads to a significant reduction in Ca^{2+} ion permeability (Figure 3) (Raymond *et al.*, 1999). Another critical role of GluA2 Q/R site editing is the regulation of endoplasmic reticulum (ER) transport of the AMPA subunit (Greger, Khatri and Ziff, 2002). The Q/R editing defect causes postnatal death in mice (Brusa *et al.*, 1995), and ADAR2-deficient mice show a similar phenotype (Higuchi *et al.*, 2000). This lethal phenotype was rescued when the mice had a point mutation of the glutamate codon to an arginine codon at the Q/R sites (Higuchi *et al.*, 2000). The second identified recording site was the serotonin 5-hydroxytryptamine$_{2C}$ receptor (5-HT_{2C}R) (Burns *et al.*, 1997). Five editing sites (A, B, C, D, and E) were identified in 5-HT_{2C}R (Burns et al., 1997). Editing of 5-HT_{2C}R isoforms resulted in reduced G-protein coupling efficiency and agonist binding (Figure 4) (Burns et al., 1997; Niswender et al., 1999; Price et al., 2001). In addition, the fully edited 5-HT_{2C}R showed sufficient expression on the cell surface (Marion et al., 2004). Kawahara et al. highlight the importance of 5-HT_{2C}R editing in energy metabolism by in vivo experiment (Kawahara, Grimberg, et al., 2008).

Other recoding sites in the transcriptome, such as *KCNA1* and *CACNA1D*, also demonstrated the function of A-to-I editing in ion channels. Voltage-gated delayed potassium channels (Kv1.1), encoded by the *KCNA1* gene, play an essential role in neuronal membrane excitability. The transcript encoding Kv1.1 converts the genomic codon ATT to ITT by A-to-I RNA editing, achieving I to V editing (isoleucine to valine) (Bhalla *et al.*, 2004). I/V editing specifically targeted the rapid inactivation of the K_V1.1 channel cavity, and recoding significantly accelerated the recovery of negative potential inactivation, which may be caused by the reduced affinity for the binding of the inactivating particle of the β-subunits (Figure 5) (Bhalla *et al.*, 2004). KCNA1 mRNAs editing also reduced drug and unsaturated fatty acid-blocking of the Kv1.1 channel (Decher *et al.*, 2010). A recent study compared the functional differences between edited (I400V) and non-edited isoforms of Kv1.1. The I400V isoform showed higher competitiveness in membrane translocation and significantly accelerated recovery from rapid inactivation, reflecting the functional significance of Kv1.1 RNA editing in neurons (Zhang *et al.*, 2021). Voltage-dependent calcium channels (Cav1.3) encoded by the *CACNA1D* gene mediate and participate in various calcium-dependent processes (Striessnig & Koschak, 2008). A-to-I RNA editing resulted in changes in three amino acids (I/M, Q/R, and Y/C) in the isoleucine-glutamine (IQ) domain of Cav1.3. The edited Cav1.3 channel significantly reduced Ca^{2+}feedback, reduced Ca^{2+}-dependent inactivation, and increased intracellular Ca^{2+} levels (Huang *et al.*, 2012).

Several well-studied recoding targets have been reported to be functionally important, such as the RNA editing sites in *AZIN1*, *CDK13,* and *FLNA.* Antizyme inhibitor 1 (AZIN1) encoded by *AZIN1* inhibits cell growth by promoting the degradation of cyclin. A-to-I RNA editing of AZIN1 is regulated by ADAR1, resulting in serine-to-glycine substitution (S/G) at AZIN1 residue 367 at β-chain 15 (β15) and predicted conformational change (Chen *et al.*, 2013). Edited AZIN1 has a stronger affinity for the antizyme, promoting hepatocellular carcinoma (HCC) cell proliferation by neutralizing the degradation of antizyme-mediated ornithine decarboxylase (ODC) and cyclin D1 (CCND1) (Chen *et al.*, 2013). Cyclin-dependent kinase 13 (CDK13) is a cyclin-dependent serine/threonine protein kinase. A high editing frequency was observed at Q/R (glutamate codon to arginine codon) in exon 1 of CDK13 (Maas *et al.*, 2011; Sakurai *et al.*, 2014a). RNA-seq results of tissue samples from patients with HCC showed overediting at two sites (Q103R and K96R). The editing level was positively correlated with ADAR1 (Dong *et al.*, 2018), suggesting that ADAR1-mediated CDK13 RNA editing may play a promoting role in the occurrence of HCC. Overediting at the Q103R site has also been identified in thyroid cancer tissues (Ramírez-Moya *et al.*, 2021), providing data support for A-to-I editing as an important cancer progression pathway. Filamin A (FLNA) is an actin-crosslinking protein. Loss of FLNA in mice results in vascular abnormalities and decreased vascular tension (Feng et al., 2006; Hart et al., 2006; Retailleau et al., 2016). A-to-I editing in exon 42 of FLNA leads to glutamine (Q) to arginine (R) amino acid exchange in the 22nd immunoglobulin (Ig)-like domain of the protein (Levanon et al., 2005; Stossel et al., 2001; Zhou et al., 2010). ADAR2 mediated RNA editing of FLNA pre-mRNA regulates vascular contraction and diastolic blood pressure (Jain et al., 2018). Recently, the same group showed that A-to-I RNA editing by FLNA regulates cell adhesion, migration, and mechanical properties (Jain et al., 2022).

4. A-to-I RNA editing in non-coding RNAs

Most A-to-I RNA editing occurs in non-coding regions of the transcriptome, such as microRNAs, long non-coding RNAs (lncRNAs), untranslated regions (UTRs), SINE repeats (primate: Alu), and circRNAs. Although non-coding editing cannot directly change the protein-coding ability of mRNA, it has many important functional implications.

MicroRNAs (miRNAs) are non-coding RNAs that are first transcribed into hairpin structures, known as primary miRNAs (pri-miRNAs). During maturation, pri-miRNA is cleaved by the

Drosha/DGCR8 complex into a precursor miRNA (pre-miRNA), which is further cleaved by the Dicer-TRBP complex in the cytoplasm and loaded into the RNA-induced silencing complex (RISC). Hairpins create dsRNA structures in pri-miRNAs and pre-miRNAs to enable miRNA editing (Yang *et al.*, 2006). RNA editing inhibits miRNA maturation by blocking the cleavage of pri-miRNA and pre-miRNA by drosha-DGCr8 and Dicer-TRBP complexes, respectively (Yang *et al.*, 2006; Kawahara, Zinshteyn, Chendrimada, *et al.*, 2007; Kawahara *et al.*, 2008; Iizasa *et al.*, 2010). A-to-I editing of pri-miRNAs can also lead to functional changes in the mature miRNAs. In addition, the role of miRNAs in translation inhibition is conferred by sequence complementarity between the miRNA seed region and target mRNA. The ability of mature miRNAs to redirect a new set of targets is conferred when the editing site is located in the seed region (Kawahara, Zinshteyn, Sethupathy, *et al.*, 2007; Vesely *et al.*, 2012; Pinto *et al.*, 2017; Wang *et al.*, 2017). Some studies have suggested that abnormal regulation of RNA editing by miRNAs is related to malignant changes and glioblastoma progression (Choudhury *et al.*, 2012; Tomaselli *et al.*, 2015; Paul *et al.*, 2017). Overediting the tumor suppressor miR-200b impairs its ability to inhibit epithelial-mesenchymal transformation (EMT) (Ramírez-Moya *et al.*, 2020). In addition to cancer, RNA editing of miRNAs has also been implicated in brain development (Mingardi *et al.*, 2018; Yang, Okada and Sakurai, 2021). The RNA editing level of some miRNAs gradually increases as the brain develops; for example, the miR-376 cluster, miR379-410 cluster, and miR-381 (Ekdahl *et al.*, 2012).

lncRNAs are defined as RNA transcripts that are over 200 nucleotides in length. A comprehensive analysis of RNA-seq data has revealed extensive RNA editing in the human transcriptome, including lncRNAs (Peng *et al.*, 2012). A-to-I editing of lncRNAs affects their secondary structure (Gong *et al.*, 2017), which can determine their functions in guiding, scaffolding, and isolating regulatory RNAs or proteins (Bhartiya and Scaria, 2016). lncRNA editing may generate new miRNA-targeting sites or remove miRNA-targeting sites on lncRNAs, thus regulating mRNA expression or inhibiting its function as miRNA decoys (Gong *et al.*, 2017).

In mRNA 3' UTRs, A-to-I RNA editing plays a role in changing gene expression through nuclear retention, degradation, and translation regulation. 3' UTR editing of mRNA transcripts has been shown to reduce the rate of mRNA translation and act as negative regulators of mRNA expression (Yang *et al.*, 2017). 3' UTR-associated A-to-I editing affects mRNA degradation by regulating the secondary structure of RNA and the accessibility of miRNA target genes (Brümmer *et al.*, 2017). In addition, editing of the 3' UTR can regulate miRNA binding, affecting post-transcriptional regulation

(Zhang *et al.*, 2016). A-to-I RNA editing has also been found in cathepsin S mRNA (CTSS) 3' UTR, which induces recruitment of the RNA-binding protein human antigen R (HuR) into the 3' UTR of the CTSS transcript, thereby regulating the stability and expression of CTSS mRNA (Stellos *et al.*, 2016).

Alu repeats are one of the most abundant types of retrotransposon insertions in primates and are frequently in the form of tandem repeats. Owing to the abundance of repeated sequences, it is easy to have two Alus in opposite directions to each other, which is the substrate of the ADAR enzyme. Most of the millions of editing sites detected in the human genome are located in primate-specific Alu sequences (Athanasiadis, Rich and Maas, 2004; Levanon *et al.*, 2004; Bazak *et al.*, 2014). One of the functions known is that editing of intron Alu may lead to exonization of Alu (Lev-Maor *et al.*, 2007; Sakurai *et al.*, 2010a), suggesting that A-to-I RNA editing may create new exons by constructing a functional splicing site and eliminating premature termination codons. The study of editing in dsRNA has implied the function of Alu in interferon response using multiple I-U mismatched synthetic dsRNAs (IU-dsRNAs) (Vitali and Scadden, 2010). IU-dsRNA, which resembles the widely edited Alu dsRNA, inhibits interferon transduction by specifically inhibiting retinoic acid-inducible gene 1 (RIG1) or MDA5 (Vitali and Scadden, 2010). Research using mouse models has revealed that ADAR1, particularly the ADAR1 p150 isoform, plays a vital role in dsRNA sensing and regulation of innate immunity (Wang *et al.*, 2004; Mannion *et al.*, 2014; Liddicoat *et al.*, 2015). In ADAR1-deficient mice, IU-dsRNA inhibits abnormally activated interferon pathways in embryonic fibroblasts (Mannion *et al.*, 2014). These results suggest that ADAR1 may mediate RNA editing in Alu to regulate the innate immune response. Furthermore, A-to-I RNA editing events have also been found to occur in intronic Alu bracketing circRNAs (single-stranded RNA, which forms a continuous closed loop). Hyperediting antagonizes circRNA formation by disrupting annealing between introns (Ivanov *et al.*, 2015).

5. A-to-I RNA editing in virus

Exogenous viral RNAs can be recognized and trigger a type I interferon response via the RIG-I or MDA5-MAVS pathway. There is increasing evidence that ADAR1-dependent A-to-I RNA editing plays a role in promoting or inhibiting the effects of some viruses. For a protective antiviral function, ADAR1 can edit dsRNA produced by viruses to influence viral infection (Samuel, 2011). The pro-

viral function of ADAR1 has been shown to enhance viral replication and proliferation by either direct editing of viral RNA or inhibition of RNA-activated protein kinase (PKR) (Gelinas *et al.*, 2011).

Measles virus (MV) is a negative-strand RNA virus that is usually transmitted through the respiratory tract and causes a fatal degenerative neurological disease called subacute sclerosing panencephalitis (SSPE) (Moss and Griffin, 2006). In SSPE, the viral genome contains high levels of A-to-I editing, with nearly half of the adenosine being mutated to guanosine. Editing sites were primarily found on matrix genes (Cattaneo and Billeter, 1992). Editing inhibited matrix protein production to reduce viral assembly and release and increased the persistence of MV infection in the central nervous system (Ward *et al.*, 2011). Moreover, deletion of the ADAR1p150 isoform in mouse embryonic fibroblasts (MEF) leads to increased susceptibility to measles virus, exhibiting antiviral properties of ADAR1p150 against measles virus infection (Ward *et al.*, 2011). In addition, another study showed that ADAR1 enhances viral growth by inhibiting the activation of PKR and interferon-regulated transcription factor 3 (IRF3), suggesting that ADAR1 plays a proviral role in measles virus infection (Toth *et al.*, 2009). It is not clear what determines the balance between the antiviral and proviral effects of ADAR during MV infections. Similar to measles virus, influenza A virus single-stranded RNA (ssRNA) is the target of A-to-I editing (Sarvestani *et al.*, 2013). ADAR1p150 prevents persistent RIG-I signaling during influenza virus A infection and promotes influenza virus replication by modulating the cellular antiviral defense pathways (Vogel *et al.*, 2020). Hepatitis C virus (HCV) infection is chronic in most infected individuals and is a significant cause of liver disease and hepatocellular carcinoma. In hepatocellular carcinoma cell lines, IFN-α treatment upregulated both ADAR1 and A-to-I editing (Taylor *et al.*, 2005). ADAR1 inhibits viral replication by modulating innate immune signaling (Pujantell *et al.*, 2018).

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which emerged at the end of 2019, is a positive single-stranded RNA ((+)ssRNA) virus. Recent evidence suggests that RNA editing may occur in the SARS-CoV-2 genome or transcriptome. In the initial RNA sequencing report, multiple RNA modification sites were detected in viral transcripts (Kim *et al.*, 2020). Although this report excludes the existence of A-to-I editing events, multiple RNA editing sites have been detected in the late SARS-CoV-2 transcriptome and genome (Giorgio *et al.*, 2020). Later, it was proposed that the A-to-G transformations found in SARS-CoV-2 are A-to-I RNA editing events mediated by ADAR1 (Picardi, Mansi and Pesole, 2021). Some studies have shown that after SARS-CoV-2 infection, viral intermediates activate the IFN response via MDA-5-mediated sensory activation (Rebendenne *et al.*,

2021; Sampaio *et al.*, 2021; Yin *et al.*, 2021). A recent study investigated RNA-seq data from samples infected with different types of SARS-CoV-2, suggesting that A-to-I editing is dynamically regulated, and that A-to-I editing may accelerate human SARS-CoV-2 evolution (Song *et al.*, 2021). In conclusion, although there is no direct method to identify the A-to-I editing site in SARS-CoV-2, the role of A-to-I RNA editing in the SARS-CoV-2 pandemic is worthy of continuous exploration. This will provide essential clues for studying viral replication, evolution, and treatment.

6. A-to-I RNA editing in RNA:DNA hybrids

The newly transcribed RNA separates from its template DNA strand immediately after transcription; however, occasionally, a stable RNA:DNA hybrid is formed, thus keeping the sense DNA in a single-stranded form (Figure 6). This three-stranded structure is called an R-loop. RNA:DNA hybrids play critical roles in various nuclear processes, particularly transcription and replication. A study has shown that ADARs can edit DNA strands that form RNA:DNA hybrids (Zheng, Lorenzo and Beal, 2017).The formation of the R-loop can cause abortive transcription and genome instability (Niehrs and Luke, 2020). Recent studies have shown that ADAR1p110 plays an essential role in the regulation of telomere R-loop formation and genomic stability (Shiromoto *et al.*, 2021a). A-to-I editing in RNA:DNA hybrids is necessary for DNA repair mediated by ADAR2 (Jimeno *et al.*, 2021). Recent studies reveal that ADAR1 edits both the RNA and the DNA strands of telomeric repeat RNA: DNA hybrids, regulating R-loop formation and genome stability at telomeres (Shiromoto *et al.*, 2021b). The discoveries that A-to-I editing can occur in RNA:DNA hybrids have broadened our understanding of the range of substrates that can be edited by ADAR enzymes and have important implications for our understanding of gene expression regulation and genome stability. Indeed, the presence of hypoxanthine as a base for deoxyinosine has long been detected as a product of spontaneous deamination of DNA (Nguyen *et al.*, 1992; Lindahl, 1993; Alseth, Dalhus and Bjørås, 2014). However, few attempts have been made to identify inosine sites on DNA until now, let alone the sparse A-to-I editing sites on RNA:DNA hybrids. This creates a pressing need for more precise and sensitive methodologies focused on identifying inosine sites within DNA.

7. Existing method for detecting A-to-I RNA editing sites

Canonical method

Since the discovery of the importance of A-to-I RNA editing in transcriptome and biological processes, several techniques have been developed for mapping inosine modifications. At present, the identification of editing sites depends on sequencing technology. The first human A-to-I RNA editing site was discovered by Sanger sequencing by comparing nucleotide differences between DNA and cDNA sequences (Figure 7) (Sanger, Nicklen and Coulson, 1977). The first example was the identification of the role of RNA editing in glutamate receptor ion channels in 1991 (Sommer *et al.*, 1991). Thereafter, a study identified five RNA editing sites in serotonin receptor 5-HT_{2C}R transcripts in 1997 (Burns *et al.*, 1997). Although Sanger sequencing was very accurate, the discovery of editing sites was very slow owing to technical defects and a cumbersome screening process. In the early 2000s, following the success of the Human Genome Project (Venter *et al.*, 2001), the first study to use comparative genomics to search for RNA editing sites identified 16 new editing sites in *Drosophila* (Hoopengardner *et al.*, 2003). Methods for aligning expressed sequence tag (ESTs) libraries to genome sequences have emerged in the same era. Four studies aligned millions of human EST or full-length cDNA transcripts to genome sequences and assessed A-to-G mismatches. More than 10,000 RNA editing sites have been identified, most of which are concentrated in the Alu sequences of non-coding regions of genes (Athanasiadis, Rich and Maas, 2004; Blow *et al.*, 2004; Kim *et al.*, 2004; Levanon *et al.*, 2004). The emergence of next-generation sequencing technology has promoted the further development of RNA editing site detection and has become a potential tool for using ESTs to predict editing sites. Compared with previous reports, the number of detected sites increased exponentially (Li *et al.*, 2009). Next-generation sequencing enables the de novo identification of editing sites without prior information. The first study, published in 2011, identified more than 10,000 exonic sites in which the RNA sequence did not match the DNA sequence (Li *et al.*, 2011). However, many false positives were found in later practice, as most RNA and DNA sequence differences led to these mismatches in sequencing, making it challenging to identify the actual editing sites (Kleinman and Majewski, 2012; Lin *et al.*, 2012; Pickrell, Gilad and Pritchard, 2012). Recently, Oxford Nanopore direct-RNA sequencing has been shown to be sensitive to RNA modifications. Inosine modifications in human poly(A) RNA have been detected by Nanopore (Workman *et al.*, 2019; Leger *et al.*, 2021).

Biochemical method

In addition to nucleotide sequencing, biochemical methods have been used to detect inosine sites in messenger RNAs. The first biochemical method developed in 1997 utilized RNase T1 to cleave inosine-containing sites. Typically, RNase T1 cleaves guanosine and inosine. They used glyoxal/borate to protect guanosines from cleavage and formed a compound (G*) (Figure 8) (Morse and Bass, 1997, 1999; Morse, 2004). Following the removal of the G* adduct from the RNA post-cleavage, an anchor sequence is ligated to the fragment containing inosine. This modified fragment serves as the template for reverse transcription and subsequent PCR amplification, a process that precedes the analysis of the sequence. However, precisely modulating glyoxal/boric acid treatments to make guanosine RNase T1 resistant remains challenging, as guanosines make up about 25% of RNA bases and are difficult to modify entirely. This incomplete modification muddles the detection of sparsely present inosine in long mRNA strands. Despite these challenges, the method's extension to next-generation sequencing library preparation demonstrates its significance in genomic research (Cattenoz *et al.*, 2013).

Another representative biochemical method, inosine chemical erasing (ICE) method, was developed and can address the sequencing limitations. This method uses acrylonitrile to chemically react with inosine, which is cyanoethylated to form 1-cyanoethylinosine by Michael addition, thereby inhibiting reverse transcriptase elongation at inosine nucleotides. Total RNA treated with acrylonitrile was amplified by PCR and sequenced directly. A large number of new A-to-I editing sites were predicted by comparing the untreated cDNA sequencing chromatograms (A and G mixed signals) with a cyanoethylated cDNA sequencing chromatogram (signal absence) (Figure 9) (Sakurai *et al.*, 2010a). The ICE method does not require genomic DNA as reference. It can distinguish inosine from signals generated by sequencing errors and false positives; however, the limitation of this method is that it can only be applied to sequences of interest. To identify new sites, they developed a new strategy using inosine chemical erasing in combination with the deep sequencing method ICE-seq, which detected approximately 20,000 new editing sites (Sakurai *et al.*, 2014a). However, ICE is primarily constrained by the chemical stability of acrylonitrile, which prevents the addition of functional groups for inosine labeling and enrichment (Figure 11).

Subsequent study synthesized acrylonitrile derivatives to achieve the conjugation of functional groups (Li, 2019). While acrylonitrile serves as a promising scaffold for chemical labeling of inosine, the synthesis of these reagents requires several synthetic and purification steps. Recent advancements

have introduced acrylamide derivatives such as acrylamidofluorescein and N-(4-ethynylphenyl)acrylamide to facilitate functional group labeling (Knutson, Ayele and Heemstra, 2018; Steve D Knutson *et al.*, 2020). However, these derivatives also complicate inosine detection due to the additional synthesis steps, resulting in inadequate efficiency and specificity.

Endonuclease-based methods

Endonuclease V, isolated from Escherichia coli (eEndoV), is characterized by its ability to identify and incise deoxyinosine-containing DNA, both in double-stranded and single-stranded forms. This enzyme specifically cleaves the second phosphodiester bonds 3' to a mismatched deoxyinosine, resulting in a nick with 3' hydroxyl and 5' phosphate groups (Moe *et al.*, 2003; Wu *et al.*, 2019). Kuraoka and colleagues have elucidated that human EndoV (hEndoV) exhibits a preference for RNA substrates over DNA (Morita *et al.*, 2013). hEndoV preferentially binds to RNA, and efficiently hydrolyzes the second phosphodiester bond 3' to inosine within regions of dsRNA that contain unpaired inosine-bearing single-stranded RNA (Vik *et al.*, 2013).

Endonuclease-based methods have been developed to enhance inosine site enrichment. An approach named EndoVIPER-seq (endonuclease V inosine precipitation enrichment sequencing) was demonstrated. They utilized the inosine-cleaving enzyme, endonuclease V, in conjunction with Ca^{2+} to promote inosine binding without cleavage and isolated transcripts containing inosine by affinity capture (Figure 10) (Steve D. Knutson *et al.*, 2020). Another method is EndoV-seq, the distinctive characteristics of hEndoV are leveraged to incise RNA at two bases downstream of inosine sites. Subsequently, a poly(U) sequence is appended to the 3' end of the cleaved RNA fragments. These fragments are selectively amplified and then utilized to construct a library for inosine detection via next-generation sequencing (Chen *et al.*, 2022). Moreover, a technique to detect inosine within DNA was developed by utilizing eEndoV (Zheng *et al.*, 2022). The eEndoV introduces nicks immediately after the inosine residues on the deoxyinosine-containing strands of fragmented double-stranded DNA. Following this incision, the DNA undergoes processing with E. coli DNA polymerase I and DNA ligase, which facilitates the incorporation of biotin-labeled dATP at the sites of inosine. These biotin-tagged locations are then selectively purified through biotin affinity, and the DNA fragments obtained are sequenced. The identification of inosine is deduced at positions where the typical guanine signal is supplanted by an adenine signal due to the presence of the incorporated biotin-dATP. However, these methods are incapable of simple inosine visualization, limited by enzymatic activity,

and are challenging in terms of enrichment efficiency. Despite these methodological advancements, the accurate identification of A-to-I editing sites remains complicated and requires further refinement.

8. Maleimide chemistry

Maleimide is a heterocyclic organic compound characterized by its five-membered ring, which includes one nitrogen atom and a carbon-carbon double bond (Figure 11). This double bond is conjugated with the carbonyl groups adjacent to the nitrogen, enhancing the reactivity of the ring. The electron-withdrawing nature of the carbonyl groups makes the double bond electrophilic, which readily reacts with nucleophiles such as the sulfhydryl groups (-SH) of cysteine residues in proteins (Ravasco *et al.*, 2018). Maleimides are notably reactive, with a propensity for undergoing conjugate additions through Michael reactions or cycloadditions via Diels-Alder reactions across their double bond (Renault *et al.*, 2018).

Maleimides play a pivotal role in protein modification due to their selective reactivity with cysteine residues, where they form stable thioether bonds. This specificity is exploited in bioconjugation techniques, particularly in the synthesis of antibody-drug conjugates for targeted therapies (Joubert *et al.*, 2017). Additionally, maleimide derivatives are employed in crosslinking studies to elucidate protein-protein interactions, contributing to the advancement of proteomics (Ravasco *et al.*, 2018). In the field of diagnostics and molecular imaging, maleimide-functionalized probes are instrumental for precise protein labeling, facilitating enhanced detection and analytical methods (Renault *et al.*, 2018). The reliability and selectivity of maleimide chemistry have established it as an essential strategy in protein engineering.

The market offers an array of maleimide derivatives, each tailored for specific applications in scientific research and industrial processes. These include N-Ethylmaleimide for protecting sulfhydryl groups, biotinylated maleimides for attaching biotin to proteins, PEGylated versions for enhancing solubility and biocompatibility, fluorescent maleimides for visual tracking in biological systems, crosslinking maleimides designed to form intermolecular connections, and photoactivatable maleimides that offer controlled conjugation through light exposure. This diversity in maleimide chemistry facilitates its integration into various fields, from drug development to materials engineering.

Research Purpose

A-to-I editing is crucial for transcriptome diversification and gene expression regulation. In addition to the editing sites in protein-coding sequences, widespread editing sites have been found in non-coding RNAs. The dysregulation of A-to-I editing is linked to several human diseases, including neurological disorders, cancer, and viral infections. The mechanisms and functions of A-to-I editing must be understood to decipher the complexities of gene regulation and dysregulation in human diseases. Thus, reliable and accurate methods to identify A-to-I editing sites in humans are essential for advancing our understanding of A-to-I editing and its role in human biology and disease.

Although most existing research on A-to-I editing focuses on double-stranded RNA, emerging evidence also indicates its occurrence in RNA:DNA hybrids. While A-to-I editing sites are well-documented in RNA (Ramaswami and Li, 2014; Picardi *et al.*, 2017), they are scarce in total cellular RNA, rendering their detection and study challenging. This issue is compounded by the largely unexplored realm of A-to-I editing within DNA. Moreover, the expression levels of A-to-I editing sites in non-coding RNAs are extremely low compared to those in canonical mRNAs, further complicating the detection of these sites.

Current inosine detection technologies, as summarized in the Introduction, face a range of limitations. Traditional methods that solely rely on NGS tend to produce false positives, compromising accuracy. Biochemical approaches using acrylonitrile have offered improvements; however, the chemical stability of acrylonitrile renders the visualization, detection, and specific enrichment of inosine unfeasible. Despite advancements with acrylamide and acrylonitrile derivatives, the complexity of their synthesis, along with suboptimal reaction efficiencies and specificities, has hindered the broader application of these methods.

In navigating the complexities of inosine detection, we have pinpointed commercially available maleimide as a prime candidate. Our intent is to employ maleimide derivatives for visualizing inosine and confirming its reactivity. We are also aiming to use maleimide derivatives to accomplish the enrichment of inosine-containing RNA and DNA, thereby constructing a rapid and accurate new method for inosine detection.

Schematic diagram of the ICLAMP method

The utilization of maleimide chemistry as a bioconjugation methodology stands out for its synthetic ease, notable reactivity, and practical applications (Ravasco *et al.*, 2018). This approach has been extensively recognized, especially in the selective modification of proteins using maleimide (Renault *et al.*, 2018). Despite its ubiquity, our study is the first to report its reactivity with inosine via Michael addition, as illustrated in Figures 11 and Figure 12. In this study, we employed a variety of commercially sourced maleimide derivatives, including those attached to functional groups like fluorescent markers or biotin, thereby broadening the method's applicability. Specifically, fluorescein-linked maleimide serves as an effective fluorescent marker for inosine, allowing for its visual detection (Figure 13). This process involves reacting nucleic acids with fluorescein-conjugated maleimide under meticulously controlled in vitro conditions. Post-labeling, any residual fluorescein-maleimide is eliminated through a detailed purification process. The fluorescence levels of the resulting product are then quantitatively evaluated using denaturing PAGE. Additionally, the incorporation of biotin into inosine nucleic acid bases is facilitated by biotin-maleimide. This enables the selective isolation of A-to-I-edited RNA or DNA through affinity purification, utilizing streptavidin-bonded materials (Figure 13). The effectiveness and reliability of each step in the ICLAMP methodology have been experimentally corroborated, as elaborated in subsequent sections:

Part 1. Inosine-specific fluorescent labeling and its associated chemical properties;

Part 2. Determination of nonspecific reactions of maleimide with other bases;

Part 3. Comparative analysis of the reactivity of maleimide, acrylonitrile, and acrylamide with inosine;

Part 4. Confirmation of reactivity between maleimide and inosine;

Part 5. Enrichment of inosine-containing DNA and RNA oligonucleotides using ICLAMP;

Part 6. Enrichment of inosine-containing total RNA using ICLAMP.

Comparative advantages and innovations of ICLAMP method

Right now, one of the main challenges in fully identifying RNA editing sites is the difficulty in confirming that the detected sites are genuine A-to-I editing sites. As introduced before, the traditional matched tissue method, which relies on detecting A-to-G mismatches by comparing RNA to genomic DNA, often produces false positives and has an accuracy of 30-60% (Figure 7 and 14). This is due to sequencing errors, genomic contamination, and SNPs that can be mistaken for editing sites. To improve detection accuracy, the ICE method was developed, using acrylonitrile to chemically modify inosine and eliminate G signals during sequencing, increasing accuracy to 97% (Figure 9 and 14). However, the ICE method does not cover low-abundance RNA or DNA. Aiming to expand the method's scope, we looked to integrate chemical labeling with biotin enrichment. Unfortunately, acrylonitrile's chemical stability prevents the addition of functional groups necessary for such enrichment.

To address this, we propose ICLAMP, an improved method that adopts maleimide to block reverse transcription, similar to ICE, while allowing for additional labeling (Figure 13 and 14). Maleimide can be tagged with various labels, such as fluorophores for visual detection. It also enables avidin-biotin interaction and antibody immunoprecipitation, allowing for the enrichment of inosine and the inclusion of low-abundance RNA and DNA in the analysis. ICLAMP not only matches the high accuracy of the ICE method but also significantly expands its application range, providing a more comprehensive tool for detecting and studying RNA modifications.

Part 1. Inosine-specific fluorescent labeling and its associated chemical properties

Materials and Method

Chemical compound preparation

Fluorescein–maleimide was purchased from Tokyo Chemical Industry. Fluorescein–maleimide stock solutions were prepared by dissolving powder in dimethyl sulfoxide (DMSO) to a concentration of 250 mM.

RNA and DNA Oligonucleotides

All RNA and DNA oligonucleotides were synthesized by Hokkaido System Science Co., Ltd. The edited sequences and nonedited controls were prepared with or without Cy5 at their 5′ termini. Oligonucleotides containing fluorescein amidite (FAM) and fluorescein isothiocyanate (FITC) labels were also prepared. The sequences used are shown in Table 1.

RNA or DNA oligonucleotide labeling and purification

Each RNA or DNA oligonucleotide (50 pmol) was dissolved in 38 μL triethylammonium acetate (TEAA) buffer (50% (v/v) ethanol, 3 M triethylammonium, pH adjusted to 8.6 with acetic acid) and vortexed vigorously. Fluorescein–maleimide was then added to the reaction solution to final concentrations of 0, 2, 10, or 50 mM. The reaction solution was replenished with DMSO to a final volume of 50 μL. The final oligonucleotide concentration was 1 pmol/μL. The reaction mixture was incubated at 70°C in the dark. At the indicated time points, the reaction was quenched by adding 10 volumes of cold (4°C) Buffer PNI (QIAquick Nucleotide Removal Kit). Excess maleimide was removed using the following purification steps, with all centrifugation steps performed as per the QIAquick nucleotide removal kit protocol: The quenched reaction mixture was loaded onto a QIAquick spin column and centrifuged. The column was washed five times with 750 μL Buffer PE (QIAGEN) to remove unreacted maleimide. The treated oligonucleotide was then eluted in 200 μL water and concentrated to 20–70 μL using an Eppendorf concentrator. An additional purification step involved mixing one volume of the concentrated sample with ten volumes of Buffer PNI. This mixture was then added to a MinElute column (QIAGEN MinElute PCR Purification kit), as per the QIAGEN MinElute PCR Purification kit protocol. Following two washes, all subsequent procedures followed the same QIAGEN MinElute PCR Purification kit protocol.

Polyacrylamide gel electrophoresis (PAGE) analysis of labeled oligonucleotides

The labeled and purified oligonucleotides were all resuspended in 15 μL water. 7 pmol of each purified sample was removed and an equal volume of 2x LS (80% formamide, 0.025% BPB, 0.025%

xylene cyanol, and 20% glycerol) was added, heated at 70°C for 10 min. The solutions were then loaded onto a denaturing PAGE gel containing 15% acrylamide and 7 M urea. For oligonucleotides without Cy5 tags, the gels were stained with toluidine blue or GelRed for RNA and DNA, respectively, and imaged using an iBright imaging system (Invitrogen). The band densities were quantified using ImageJ (1.52a). Each DNA sample was normalized by comparing the band intensity with that of a known amount of FITC-labeled DNA oligonucleotide (FITC-DNA-A or FITC-DNA-I). The percentage conversion was calculated as the molar ratio of fluorescein to Gel Red in each well. For Cy5-tagged oligonucleotides, the gels were imaged directly using iBright. Each sample was normalized by comparing the band intensity with those of known amounts of Cy5- and FITC-labeled DNA oligonucleotides (Cy5-DNA-A-FITC and Cy5-DNA-I-FITC). The percentage conversion was calculated as the molar ratio of fluorescein to Cy5 per well.

Results

First, we optimized the reaction conditions for inosine labeling using fluorescein–maleimide by designing and synthesizing two RNA oligonucleotides, RNA-6A and RNA-6I, containing six adenosines or inosines, respectively, at specific positions. The most favorable conditions were determined by varying reaction conditions such as the pH, temperature, and buffer systems. The optimal reaction conditions were determined using the reaction specificity index (I/A ratio), which was calculated as the ratio of the fluorescence intensity of RNA-6I to that of RNA-6A. The pKa value of each base (Figure 15) (Luyten *et al.*, 1998; Thaplyal and Bevilacqua, 2014; Gillingham, Geigle and Lilienfeld, 2016) was used to investigate the labeling efficiencies at pH 7.8–9.2, where the reaction efficiency to inosine is maximum relative to those of other bases. We identified pH 8.6 to be optimal for the reaction (Figure 16). The reaction temperatures were also compared in 10°C increments, revealing that the highest reaction efficiency and specificity could be achieved at 70°C (Figure 17). We then examined the effects of the reaction solution composition. Both 3 M (50% (v/v) ethanol and 3 M triethylammonium) and 1.1 M TEAA buffers (50% (v/v) ethanol and 1.1 M triethylammonium) enhanced the reaction efficiency (Figure 18). Based on these results, the optimal reaction conditions for inosine-specific maleimide labeling were TEAA buffer (3 M TEAA for oligonucleotides and 1.1 M TEAA for total RNA and genomic DNA), pH 8.6, and temperature 70°C.

We then applied the optimized inosine-labeling conditions to DNA oligonucleotides containing six deoxyadenosines (DNA-6dA) and six deoxyinosines (DNA-6dI). Maleimide is highly soluble in DMSO and facilitates the flexible adjustment of the maleimide reaction concentration. We therefore tested the difference in the reaction of RNA and DNA with two concentrations of maleimide: 2 mM (mild treatment) and 10 mM (medium treatment). RNA and DNA reacted rapidly with fluorescein–maleimide and produced substantial signals at 15 min with both the mild and medium treatments. The labeling efficiency for both RNA and DNA improved over time, with a similar trend observed. The fluorescence signals in the medium treatment group were substantially stronger than those of the mild treatment group (Figure 19).

The labeling efficiencies of maleimide on individual inosine bases were measured by synthesized a 21-base DNA strand containing only one deoxyinosine (DNA-dI). For comparison, a DNA strand was synthesized by replacing the deoxyinosine with a deoxyadenosine (DNA-dA). Notably, inosine was the only molecular difference between both DNA strands. We used the medium concentration of fluorescein–maleimide and performed time course experiments spanning 0–8 h for each DNA sample to investigate the correlation between reaction time and degree of labeling. As shown in Figure 20, the fluorescein labeling of DNA-dI increased over time and peaked at 1 h.

To enhance the specificity and quantification capabilities of our detection, we synthesized two single stranded DNAs (ssDNAs), Cy5-DNA-dA and Cy5-DNA-dI, which used the same sequences as the previous short DNA but included 5′-Cy5 fluorescent labels. We performed time course experiments that included the time points at 0, 15, 30, and 60 min, with all measurements made in triplicate (n=3). As shown in Figure 21, the fluorescent signal of Cy5-DNA-dI was detected as early as 15 min, whereas only a faint signal was observed at the same time point for Cy5-DNA-dA. With increasing reaction times, the Cy5 fluorescence bands migrated upward in the gel. This shift was attributed to the addition of a maleimide label to the DNA molecule, which increased its molecular weight. The fluorescence intensities of the DNA bands in the gels were analyzed in terms of their densities and normalized in terms of standard amounts of fluorescein and Cy5-labeled control oligonucleotides. Normalized data were used to calculate the labeling yield (% labeling) as a function of the reaction time. The I/A ratio was also calculated as the ratio of the labeling yields of Cy5-DNA-dI to Cy5-DNA-dA. Based on the data, the reaction efficiency was approximately 40% at 1 h. The labeling reaction occurred rapidly from 0 to 15 min, with the optimal I/A ratio occurred at 15 min corresponding to the least nonspecific reaction for A.

To ascertain the potential benefits of higher fluorescein–maleimide concentrations on labeling efficiency, the Cy5-labeled DNAs (Cy5-DNA-dI and Cy5-DNA-dA) were subjected to 50 mM fluorescein–maleimide labeling, and time course experiments were performed, with n=3 (Figure 22). The reaction consistently produced a replicable reaction efficiency of approximately 80% after 1 h reaction. The specific reactivity was observed at 15 min with 50 mM maleimide. Based on the results, increasing the maleimide concentration can achieve 100% inosine labeling efficiency. However, simultaneously it can cause nonspecific labeling of bases other than inosine, leading to lower I/A ratios. Thus, in the subsequent experiments, we can choose appropriate maleimide concentration and reaction time tailored to the specific aims of each experiment. We also utilized a DNA oligonucleotide Cy5-DNA-dG to detect the reactivity of 10 mM maleimide with guanosine, a structure similar to inosine, and performed time course experiments (Figure 23). The I/G ratio indicated that the optimal reaction conditions remained a reaction time of 15 min.

Part 2. Determination of nonspecific reactions of maleimide with other bases

Materials and Method

Chemical compound preparation

Fluorescein–maleimide was purchased from Tokyo Chemical Industry. Fluorescein–maleimide stock solutions were prepared by dissolving the powder in dimethyl sulfoxide (DMSO) to a concentration of 250 mM.

DNA Oligonucleotides

DNA oligonucleotides were synthesized by Hokkaido System Science Co., Ltd. The edited sequences and nonedited controls were prepared with Cy5 at their 5′ termini. The sequences used are shown in Table 1.

DNA oligonucleotide labeling and purification

For Fluorescein–maleimide labeling, DNA oligonucleotide (50 pmol) was dissolved in 38 µL triethylammonium acetate (TEAA) buffer (50% (v/v) ethanol, 3 M triethylammonium, pH adjusted to 8.6 with acetic acid) and vortexed vigorously. Fluorescein–maleimide or biotin–maleimide was then added to the reaction solution to final concentrations of 0, 2, 10, or 50 mM. The reaction solution was replenished with DMSO to a final volume of 50 µL. The final oligonucleotide concentration was 1 pmol/µL. The reaction mixture was incubated at 70°C in the dark. At the indicated time points, the reaction was quenched by adding 10 volumes of cold (4°C) Buffer PNI (QIAquick Nucleotide Removal Kit). Excess maleimide was removed using the following purification steps, with all centrifugation steps performed as per the QIAquick nucleotide removal kit protocol: The quenched reaction mixture was loaded onto a QIAquick spin column and centrifuged. The column was washed five times with 750 µL Buffer PE (QIAGEN) to remove unreacted maleimide. The treated oligonucleotide was then eluted in 200 µL water and concentrated to 20–70 µL using an Eppendorf concentrator. An additional purification step involved mixing one volume of the concentrated sample with ten volumes of Buffer PNI. This mixture was then added to a MinElute column (QIAGEN MinElute PCR Purification kit), as per the QIAGEN MinElute PCR Purification kit protocol. Following two washes, all subsequent procedures followed the same QIAGEN MinElute PCR Purification kit protocol.

Polyacrylamide gel electrophoresis (PAGE) analysis of labeled oligonucleotides

The labeled and purified oligonucleotides were all resuspended in 15 µL water. 7 pmol of each purified sample was removed and an equal volume of 2x LS (80% formamide, 0.025% BPB, 0.025% xylene cyanol, and 20% glycerol) was added, heated at 70°C for 10 min. The solutions were then loaded onto a denaturing PAGE gel containing 15% acrylamide and 7 M urea. The gels were imaged directly using iBright. Each sample was normalized by comparing the band intensity with those of known amounts of Cy5- and FITC-labeled DNA oligonucleotides (Cy5-DNA-A-FITC and Cy5-DNA-I-FITC). The percentage conversion was calculated as the molar ratio of fluorescein to Cy5 per well.

Results

After determining the optimal reaction conditions for maleimide labeling of inosine, we utilized five sets of short 13-nt oligonucleotides, labeled with Cy5 and organized into IA, IG, IC, IT, and IU groups (Figure 24). Each set of these oligonucleotides features identical sequences with the exception of the nucleotide being assessed, specifically A, G, C, T, and U, to evaluate the reactivity of maleimide at pH 8.6 (Figure 24). Each oligonucleotide was incubated with 50 mM fluorescein–maleimide for 60 min. When the reaction peaked at one hour, analysis of the ratio between the specific reactions to I and the nonspecific reactions to other bases showed that A and T exhibited the lowest nonspecific reactions. The nonspecific reaction of maleimide with U showed similar levels to those with G and C. Maleimide exhibited higher specificity in its reactivity toward inosine.

Part 3. Comparative analysis of the reactivity of maleimide, acrylonitrile, and acrylamide with inosine

Materials and Method

Chemical compound preparation

Fluorescein–maleimide was purchased from Tokyo Chemical Industry. Fluorescein–maleimide stock solutions were separately prepared by dissolving in DMSO to a concentration of 250 mM. Acrylamidofluorescein was synthesized following the reported synthesis protocol (Knutson, Ayele and Heemstra, 2018).

DNA Oligonucleotides

DNA oligonucleotides were synthesized by Hokkaido System Science Co., Ltd. The edited sequences and nonedited controls were prepared with Cy5 at their 5′ termini. The sequences used are shown in Table 1.

DNA oligonucleotide labeling and purification

For Fluorescein–maleimide labeling, DNA oligonucleotide (50 pmol) was dissolved in 38 μL triethylammonium acetate (TEAA) buffer (50% (v/v) ethanol, 3 M triethylammonium, pH adjusted to 8.6 with acetic acid) and vortexed vigorously. Fluorescein–maleimide or biotin–maleimide was then added to the reaction solution to final concentrations of 0, 2, 10, or 50 mM. The reaction solution was replenished with DMSO to a final volume of 50 μL. The final oligonucleotide concentration was 1 pmol/μL. The reaction mixture was incubated at 70°C in the dark. At indicated time points, the reaction was quenched by adding 10 volumes of cold (4°C) Buffer PNI (QIAquick Nucleotide Removal Kit). All centrifugation steps performed as per the QIAquick nucleotide removal kit protocol: The quenched reaction mixture was loaded onto a QIAquick spin column and centrifuged. The column was washed five times with 750 μL Buffer PE (QIAGEN). The treated oligonucleotide was then eluted in 200 μL water and concentrated to 20–70 μL using an Eppendorf concentrator. An additional purification step involved mixing one volume of the concentrated sample with ten volumes of Buffer PNI. This mixture was then added to a MinElute column (QIAGEN MinElute PCR Purification kit), as per the QIAGEN MinElute PCR Purification kit protocol. Following two washes, all subsequent procedures followed the same QIAGEN MinElute PCR Purification kit protocol.

For acrylonitrile labeling, DNA oligonucleotides (38 pmol) were dissolved in water to make a solution totaling 4 μL. Then the DNA solution was incubated in 30 μL CE solution (50% (v/v) ethanol, 1.1 M triethylammonium, pH adjusted to 8.6 with acetic acid) with or without 4 μL 15.2 M acrylonitrile. The final oligonucleotide concentration was 1 pmol/μL. The reaction mixture was incubated at 70°C for 30 min. The treated DNA was purified using an RNeasy MinElute kit (QIAGEN) and ethanol precipitation.

For Acrylamidofluorescein labeling, DNA oligonucleotides (150 pmol) were added to a 15 μL solution of 250 mM acrylamidofluorescein in 50:50 EtOH:TEAA buffer and adjusted to pH 8.6. The reaction mixture was incubated at 70°C. At indicated time points, the reaction was quenched by adding 10 volumes of cold (4°C) Buffer PNI (QIAquick Nucleotide Removal Kit). All centrifugation steps performed as per the QIAquick nucleotide removal kit protocol: The quenched reaction mixture was loaded onto a QIAquick spin column and centrifuged. The column was washed five times with 750 μL Buffer PE (QIAGEN). The treated oligonucleotide was then eluted in 200 μL water and concentrated to 20–70 μL using an Eppendorf concentrator. An additional purification step involved mixing one volume of the concentrated sample with ten volumes of Buffer PNI. This mixture was then added to a MinElute column (QIAGEN MinElute PCR Purification kit), as per the QIAGEN

MinElute PCR Purification kit protocol. Following two washes, all subsequent procedures followed the same QIAGEN MinElute PCR Purification kit protocol.

Polyacrylamide gel electrophoresis (PAGE) analysis of labeled oligonucleotides

The labeled and purified oligonucleotides were all resuspended in 15 μL water. 7 pmol of each purified sample was removed and an equal volume of 2x LS (80% formamide, 0.025% BPB, 0.025% xylene cyanol, and 20% glycerol) was added, heated at 70°C for 10 min. The solutions were then loaded onto a denaturing PAGE gel containing 15% acrylamide and 7 M urea. The gels were imaged directly using iBright. Each sample was normalized by comparing the band intensity with those of known amounts of Cy5- and FITC-labeled DNA oligonucleotides (Cy5-DNA-A-FITC and Cy5-DNA-I-FITC). The percentage conversion was calculated as the molar ratio of fluorescein to Cy5 per well.

Results

To elucidate the maleimide reaction mechanism with inosine, we performed a comparative study using acrylonitrile (Figure 25), a known agent confirmed to result in the formation of 1-cyanoethylinosine (Yoshida and Ukita, 1968; Sakurai *et al.*, 2010b). Initially, acrylonitrile was used to convert inosine to 1-cyanoethylinosine. These cyanoethylated DNA oligonucleotides then underwent fluorescein–maleimide labeling (Figure 25). A significant decrease in labeling efficiency was observed, indicating that adduction at the N1 position of inosine inhibited the maleimide reaction. To compare the reactivity of maleimide with that of acrylonitrile, we designed competitive reactions incorporating both reagents. Under conditions optimized using both ICLAMP and ICE, the DNA oligonucleotides were exposed to a solution containing 1.6 M acrylonitrile and 10 mM fluorescein–maleimide in a 3 M TEAA buffer for 30 min (Figure 25). The minimal reduction in fluorescence intensity suggested that the reaction of maleimide with inosine was more rapid than that of acrylonitrile, highlighting the superior reactivity of maleimide toward inosine.

We then compared the labeling efficiency of our fluorescein–maleimide method with that of a previously reported acrylamidofluorescein technique.(Knutson, Ayele and Heemstra, 2018) Two types of ssDNA (Cy5-DNA-dA and Cy5-DNA-dI) were used instead of RNA oligonucleotides. A

time course experiment was performed over 24 h with 250 mM acrylamidofluorescein and 10 μM DNA oligonucleotides (Figure 26). We observed a weak fluorescence signal at 8 h and the labeling efficiency remained low even after 24 h (Figure 26, right). In contrast, a separate time course experiment using 10 mM fluorescein–maleimide and 1 μM DNA oligonucleotides resulted in specific labeling of inosine within 15 min and reached a peak efficiency at 1 h (Figure 26, left). Our method thus reduced the required labeling time by over 50-fold and increased the labeling efficiency by 3-fold while maintaining superior base specificity.

Part 4. Confirmation of reactivity between maleimide and inosine

Materials and Method

Chemical compound preparation

Maleimide, methylmaleimide, fluorescein–maleimide and biotin-maleimide were purchased from Tokyo Chemical Industry. The stock solutions were prepared by dissolving powder in dimethyl sulfoxide (DMSO).

RNA Oligonucleotides

All RNA oligonucleotides were synthesized by Hokkaido System Science Co., Ltd. The sequences used are shown in Table 1.

RNA oligonucleotide labeling and purification

RNA oligonucleotide (50 pmol) was dissolved in 38 μL triethylammonium acetate (TEAA) buffer (50% (v/v) ethanol, 3 M triethylammonium, pH adjusted to 8.6 with acetic acid) and vortexed vigorously. Maleimide was then added to the reaction solution to final concentrations of 0 or 50 mM. The reaction solution was replenished with DMSO to a final volume of 50 μL. The final oligonucleotide concentration was 1 pmol/μL. The reaction mixture was incubated at 70°C in the dark. At the indicated time points, 400 μL of cold water was added to the reaction mixture to stop the

reaction. Excess maleimide was removed as follows: The treated samples were purified via isopropanol precipitation and two Phenol-chloroform-isoamyl alcohol mixture (PCI, Sigma, pH 7.7-8.3) treatments, then treated with diethyl ether to remove the remaining phenol. Residual diethyl ether was removed using an Eppendorf concentrator. Finally, the treated RNA was purified via ethanol precipitation.

Liquid chromatography-mass spectrometry (LC/MS)

Maleimide-modified RNA oligonucleotide was digested with nuclease P1 and bacterial alkaline phosphatase and subjected to hydrophilic-interaction LC/MS analysis as described previously (Sakaguchi *et al.*, 2015). Data were analyzed using the Xcalibur Qual Browser (bundled with Thermo Xcalibur 4.3; Thermo Fisher Scientific).

Results

Following the comparative analysis of fluorescein–maleimide and acrylonitrile, we sought to further substantiate the specific reactivity between maleimide and inosine. To achieve this, we conducted experiments using native maleimide in reactions with inosine. An RNA oligo (RNA-6I) was treated with 50 mM maleimide for 60 min and digested into nucleosides, and then analyzed by LC/MS. Inosine and inosine-maleimide are clearly detected as protonated ions (Figure 27). Additionally, high-energy collisional dissociation spectra (Figure 27 and 28) revealed the segmentation of the base ion for the inosine-maleimide adducts, delineated by a dotted line.

Part 5. Enrichment of inosine-containing DNA and RNA oligonucleotides using ICLAMP

Materials and Method

Chemical compound preparation

Biotin–maleimide was purchased from Tokyo Chemical Industry. Biotin–maleimide stock solutions were prepared by dissolving each in dimethyl sulfoxide (DMSO) to a concentration of 250 mM.

RNA and DNA Oligonucleotides

All RNA and DNA oligonucleotides were synthesized by Hokkaido System Science Co., Ltd. The sequences used are shown in Table 1.

RNA or DNA oligonucleotide labeling and purification

Each RNA or DNA oligonucleotide (50 pmol) was dissolved in 38 μL triethylammonium acetate (TEAA) buffer (50% (v/v) ethanol, 3 M triethylammonium, pH adjusted to 8.6 with acetic acid) and vortexed vigorously. Fluorescein–maleimide was then added to the reaction solution to final concentrations of 0, 2, 10, or 50 mM. The reaction solution was replenished with DMSO to a final volume of 50 μL. The final oligonucleotide concentration was 1 pmol/μL. The reaction mixture was incubated at 70°C in the dark. At the indicated time points, the reaction was quenched by adding 10 volumes of cold (4°C) Buffer PNI (QIAquick Nucleotide Removal Kit). Excess maleimide was removed using the following purification steps, with all centrifugation steps performed as per the QIAquick nucleotide removal kit protocol: The quenched reaction mixture was loaded onto a QIAquick spin column and centrifuged. The column was washed five times with 750 μL Buffer PE (QIAGEN) to remove unreacted maleimide. The treated oligonucleotide was then eluted in 200 μL water and concentrated to 20–70 μL using an Eppendorf concentrator. An additional purification step involved mixing one volume of the concentrated sample with ten volumes of Buffer PNI. This mixture was then added to a MinElute column (QIAGEN MinElute PCR Purification kit), as per the QIAGEN MinElute PCR Purification kit protocol. Following two washes, all subsequent procedures followed the same QIAGEN MinElute PCR Purification kit protocol.

Affinity pull-down

Nucleic acids labeled with biotin–maleimide were pulled down using Dynabeads MyOne Streptavidin C1 magnetic beads (Thermo Fisher Scientific) as per the manufacturer's protocol. The

recommended buffers were modified by referring to the following detailed steps for the exact formulation. The tubes were pretreated to prevent the nonspecific adhesion of the beads and nucleic acids to the inner wall of the tubes during suspension. A 2 mL tube was filled with 1.5 mL of a blocking buffer (25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer at pH 7.5, 1 M NaCl, 12.5 mM ethylenediamine tetra-acetic acid (EDTA), 0.005% Tween 20, and 1% bovine serum albumin) and inverted for 1 h using an Intelli mixer (ELMI). The tube was then washed twice with 1 mL of 1x binding buffer (25 mM HEPES buffer at pH 7.5, 1 M NaCl, 12.5 mM EDTA, and 0.005% Tween 20). Appropriate quantities of the beads (10 μL for oligonucleotides and 120 μL for the total RNA) were removed and washed three times with 500 μL blocking buffer for 2 min; this step was repeated three times. The washed magnetic beads were resuspended in 2x binding buffer to a final concentration of 5 μg/μL (twice the original volume). An equal volume of biotinylated DNA or RNA in distilled water was added to the beads and incubated for 60 min at room temperature (25°C) with gentle rotation (Eppendorf ThermoMixer). The biotinylated DNA- or RNA-coated beads were separated using a magnet for 2–3 min. The supernatant was collected, and the coated beads were washed two or three times with 500 μL 1x washing buffer (5 mM HEPES buffer at pH 7.5, 200 mM NaCl, 2.5 mM EDTA, and 0.005% Tween 20). The nucleic acids were released by boiling the beads in 100% formamide at 70°C for 5 min. The tube was placed on a magnet and the supernatant solution was transferred to a new tube. The isolated nucleic acids were purified by ethanol precipitation.

Primer extension

To perform primer extension of the synthetic RNA oligonucleotides (RNA-1I), the pull-down RNA was incubated with 2 pmol primer (Cy5-PE-DNA) in a total reaction volume of 5.5 μL at 65°C for 5 min and placed on ice. Reverse transcription was performed at 50°C for 60 min in a 10 μL reaction mixture comprising 5.5 μL of the primer and RNA mixture, 5 mM DTT, 150 μM dNTPs (dATP, dTTP, dCTP, and dGTP mixture), 1x first-strand buffer (Invitrogen), and 200 units of SuperScript III reverse transcriptase (Invitrogen). Next, 10 μL of 2x LS (80% formamide, 0.025% BPB, 0.025% xylene cyanol, and 20% glycerol) was added to each solution and heated to 70°C for 10 min. The solutions were then loaded onto the PAGE gel. The cDNA was separated via 15% acrylamide–7 M urea PAGE and visualized using iBright.

Polyacrylamide gel electrophoresis (PAGE) analysis of labeled oligonucleotides

The labeled and purified oligonucleotides were all resuspended in 15 µL water. 7 pmol of each purified sample was removed and an equal volume of 2x LS (80% formamide, 0.025% BPB, 0.025% xylene cyanol, and 20% glycerol) was added, heated at 70°C for 10 min. The solutions were then loaded onto a denaturing PAGE gel containing 15% acrylamide and 7 M urea. The gels were imaged directly using iBright.

Determination of inosine enrichment Fold

Varying mixture of RNA-I-Cy5 and RNA-A-FAM were prepared as described (Knutson, Ayele and Heemstra, 2018). The RNA oligonucleotide mixtures were incubated with biotin-maleimide, and subsequently captured using magnetic beads coupled with streptavidin as mentioned above. The isolated RNA was eluted and transferred onto a membrane (Amersham Hybond-N+), then quantified the fluorescence intensity using a Typhoon scanner. To calculate fold enrichment, the final concentrations of RNA-A-FAM and RNA-I-Cy5 obtained after the pull-down were compared with their initial input ratios using the formula reported previously (Knutson, Ayele and Heemstra, 2018).

Results

After developing the inosine-specific base-labeling technique, we used another maleimide derivative, biotin–maleimide, for the affinity enrichment of inosine-containing molecules. To determine whether the reactivity of biotin-maleimide aligns with that of fluorescein-maleimide, we employed primer extension as a method to assess the reactivity of biotin-maleimide with inosine (Figure 29). Building on a previous study that identified inosine by primer extension (Sakurai *et al.*, 2010b), we investigated whether the adduction of maleimide to inosine could arrest cDNA extension. This approach was employed to confirm the presence and attachment position of maleimide. A short RNA (RNA-1I) template containing one inosine was subjected to 50 mM biotin–maleimide labeling at 0, 15, 30, and 60 min. A 5′-Cy5 labeled DNA primer (Cy5-PE) complementary to the 3′-end of the RNA was incubated with the labeled RNA. The RNA without maleimide labeling, primer extension stopped at the terminal base (Figure 29). The RNA with maleimide (MI) labeling resulted in a lower band ending at the inosine site. The result demonstrated that maleimide-inosine can inhibit base-

pairing during the cDNA extension process by reverse transcriptase. The reaction efficiency of biotin-maleimide is approximately 80% after 1 h reaction.

Subsequently, we utilized biotin-maleimide to pull down molecules containing inosine. This was achieved by binding the molecules to streptavidin-coupled magnetic beads (Figure 13 and 30). Four different oligonucleotides in length were used for simultaneous detection via electrophoresis (Figure 30a). DNA oligonucleotides containing one 5′-Cy5 fluorescent label and one deoxyinosine (Cy5-36nt-dI) were prepared as the primary targets for biotin labeling and affinity purification. FAM-30nt-dU and FAM-24nt-dA, each containing one 5′-FAM fluorescent label and one deoxyuridine or deoxyadenosine, respectively, were synthesized as negative controls to verify the nonspecific labeling and purification of bases other than inosine. Cy5-24nt-dI-Bio, containing a 5′-Cy5 fluorescent label and a 3′-biotin modification, was prepared as a positive control. Cy5-35nt-dI, FAM-30nt-dU, and FAM-24nt-dA were mixed in equal amounts with 50 mM biotin–maleimide and incubated for 60 min. Excess biotin–maleimide was removed via purification and an equal amount of the control oligonucleotide Cy5-24nt-dI-Bio was added (Figure 30a). Finally, affinity pull-down was performed with streptavidin-coupled magnetic beads. The results showed that the positive control oligonucleotides were pulled down in all experimental groups, those oligonucleotides that were labeled with biotin–maleimide at the inosine sites successfully isolated (Figure 30b).

To test the enrichment fold of our approach, we employed the same RNA sequences and experimental methodologies as previously described (Knutson, Ayele and Heemstra, 2018), with the modification of labeling the RNA sequences with Cy5 and FAM (RNA-I-Cy5 and RNA-A-FAM). These two RNA oligonucleotides were mixed in various proportions (Figure 31), followed by labeling with 50 mM biotin-maleimide for 60 min, and subsequently captured by affinity pull-down. As shown in Figure 31, biotin-maleimide labeling combined with streptavidin pull-down facilitates an enrichment of inosine-containing RNA oligonucleotides by upwards of 15-fold, and the enrichment factor increases as the proportion of inosine-containing RNA decreases.

Part 6. Enrichment of inosine-containing total RNA using ICLAMP

Materials and Method

Chemical compound preparation

Biotin–maleimide was purchased from Tokyo Chemical Industry. Biotin–maleimide stock solutions were prepared by dissolving each in dimethyl sulfoxide (DMSO) to a concentration of 250 mM.

Total RNA preparation

The total RNA was extracted from mouse brain tissues using an RNeasy Maxi Kit (QIAGEN). The mouse brain tissues were obtained from 8-week-old C57BL/6 mice in the Goitsuka Laboratory at the Tokyo University of Science. Upon collection, tissues were flash-frozen and stored at -80°C until use. All procedures were carried out in accordance with approved protocols of the Tokyo University of Science Animal Care and Use Committee. The total RNA (40 μg) was fragmented by incubating it in a 100 μL reaction solution containing magnesium ions (100 mM Tris-HCl and 2 mM $MgCl_2$) at 95°C for 8 min. Subsequently, the fragmented total RNA was purified by ethanol precipitation.

Total RNA labeling and purification

Fragmented total RNA (30 μg) was added to 38 μL of TEAA buffer (50% (v/v) ethanol, 1.1 M triethylammonium, pH adjusted to 8.6 with acetic acid) and vortexed vigorously. Biotin–maleimide was then added to the reaction solution to final concentrations of 0, 2, 10, and 50 mM. The reaction solution was then replenished with DMSO to a final volume of 50 μL and incubated at 70°C. At the indicated time points, 400 μL of cold water was added to the reaction mixture to stop the reaction. Excess maleimide was removed as follows: The treated samples were purified via isopropanol precipitation and two Phenol-chloroform-isoamyl alcohol mixture (PCI, Sigma, pH 7.7-8.3) treatments, then treated with diethyl ether to remove the remaining phenol. Residual diethyl ether was removed using an Eppendorf concentrator. Finally, the treated RNA was purified via ethanol precipitation.

Radioisotope labeling of DNA primers

The DNA primers located downstream of the glutamate receptor B (*Gria2*) mRNA Q/R site were labeled with radioisotopes. Briefly, 4 pmol of each DNA primer was mixed with 1 μL of 10x T4 polynucleotide kinase buffer, 1 μL (3.3 pmol) of γ-^{32}P ATP (3000 Ci/mmol, 10 μCi/μL), T4 polynucleotide kinase (NEB) 10 units, and water was added to bring the final volume to 10 μL. The solutions were reacted at 37°C for 30 min, immediately after which the enzyme was thermally inactivated via boiling at 95°C for 3 min, then placed on ice to cool. Next, 10 μL of 2x LS (80% formamide, 0.025% BPB, 0.025% xylene cyanol, and 20% glycerol) was added to each solution and heated to 95°C for 5 min. The solutions were then loaded onto a 10% acrylamide–7 M urea denaturing PAGE gel. The gel was imaged using an Amersham Typhoon scanner (Cytiva). The sections of the gel containing the ^{32}P-labeled primers were cut into 2 mm rectangles and transferred to 2 mL tubes. Next, 400 μL of an elution buffer (0.3 M sodium acetate at pH 5.2 and 0.1% sodium dodecyl sulfate) was added left to permeate for 90 min at room temperature (25°C) with continuous shaking. The eluted DNA supernatants were then transferred to 1.5 mL tubes and DNA pellets were collected via ethanol precipitation. The pellets were dissolved in 10 μL of water and the radioactivity of the labeled DNA primers was measured using a scintillation counter. The labeled DNA primers were reconstituted to 1,000,000–1,500,000 cpm/μL and used for primer extension.

Affinity pull-down

Nucleic acids labeled with biotin–maleimide were pulled down using Dynabeads MyOne Streptavidin C1 magnetic beads (Thermo Fisher Scientific) as per the manufacturer's protocol. The recommended buffers were modified by referring to the following detailed steps for the exact formulation. The tubes were pretreated to prevent the nonspecific adhesion of the beads and nucleic acids to the inner wall of the tubes during suspension. A 2 mL tube was filled with 1.5 mL of a blocking buffer (25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer at pH 7.5, 1 M NaCl, 12.5 mM ethylenediamine tetra-acetic acid (EDTA), 0.005% Tween 20, and 1% bovine serum albumin) and inverted for 1 h using an Intelli mixer (ELMI). The tube was then washed twice with 1 mL of 1x binding buffer (25 mM HEPES buffer at pH 7.5, 1 M NaCl, 12.5 mM EDTA, and 0.005% Tween 20). Appropriate quantities of the beads (10 μL for oligonucleotides and 120 μL for the total RNA) were removed and washed three times with 500 μL blocking buffer for 2 min; this

step was repeated three times. The washed magnetic beads were resuspended in 2x binding buffer to a final concentration of 5 µg/µL (twice the original volume). An equal volume of RNA in distilled water was added to the beads and incubated for 60 min at room temperature (25°C) with gentle rotation (Eppendorf ThermoMixer). The biotinylated RNA-coated beads were separated using a magnet for 2–3 min. The supernatant was collected, and the coated beads were washed two or three times with 500 µL 1x washing buffer (5 mM HEPES buffer at pH 7.5, 200 mM NaCl, 2.5 mM EDTA, and 0.005% Tween 20). The nucleic acids were released by boiling the beads in 100% formamide at 70°C for 5 min. The tube was placed on a magnet and the supernatant solution was transferred to a new tube. The isolated nucleic acids were purified by ethanol precipitation.

Primer extension

For primer extension of the mouse *Gria2* mRNA Q/R site, the purified total RNA was dissolved in 3 µL water, then mixed with 1 µL 5′-^{32}P-labeled DNA primer, 2 µL 5x glycerol RT buffer (250 mM Tris-HCl at pH 7.5, 375 mM KCl, 15 mM $MgCl_2$, and 50% glycerol), and 2 µL 5 M betaine (Merck-Sigma). The solution was incubated at 70°C for 2 min, then cooled to room temperature. Next, 0.5 µL 0.1 M DTT, 1 µL 1.5 mM dNTPs, and 100 units of SuperScript III reverse transcriptase (Invitrogen) were added to form a 10 µL reverse transcription reaction and placed at 50°C for 60 min. After the reaction, 10 µL of 2x loading solution (7 M urea, 0.05% bromophenol blue, 0.05% xylene cyanol, 1x TBE) and 20 µL of 100% formamide were added, heated at 95°C for 5 min. 10 µL of the solutions were then loaded onto a PAGE gel. The bands were separated via 15% acrylamide–7 M urea PAGE and visualized using an Amersham Typhoon scanner.

For primer extension of the mouse *Gria2* mRNA Q/R site with α-^{32}p dCTP, the purified total RNA was dissolved in 2 µL water and mixed with 0.5 pmol primer, 2 µL 5x glycerol RT buffer (250 mM Tris-HCl pH 7.5, 375 mM KCl, 15 mM $MgCl_2$, and 50% glycerol), and 2 µL 5 M betaine. The mixture was incubated at 70°C for 2 min and cooled to room temperature. Next, 0.5 µL 0.1 M DTT, 0.5 µL 3 mM dNTPs (dATP, dTTP, and dGTP mixture), 2 µL 3.3 µM α-^{32}p dCTP (3000 Ci/mmol, 10 µCi/µL), and 100 units of SuperScript III reverse transcriptase (Invitrogen) were added to form a 10 µL reverse transcription reaction mixture, which was then incubated at 50°C for 60 min. After the reaction, 10 µL of 2x loading solution (7 M urea, 0.05% bromophenol blue, 0.05% xylene cyanol, 1x TBE) and 20 µL of 100% formamide were added, heated at 95°C for 5 min. 10 µL of the solutions were then

loaded onto a PAGE gel. The bands were separated via 15% acrylamide–7 M urea PAGE and visualized using an Amersham Typhoon scanner.

Results

To validate the ability of biotin–maleimide to label inosine-containing nucleic acids within cells and assess the effectiveness of affinity purification in removing interfering nucleic acids that lack inosine, we used a A-to-I editing site of *Gria2* mRNA in mouse brain where a glutamine CAG (Q) codon is edited to an arginine CIG (R) codon (Q-to-R protein editing), as the Q-to-R editing frequency is nearly 100% (Higuchi *et al.*, 1993; Seeburg, Higuchi and Sprengel, 1998). We designed a 5′-^{32}P-phosphorylated DNA primer located downstream of the Q/R site, following the previously reported ICE method (Sakurai *et al.*, 2010b), and fragmented the total RNA from mouse brains to approximately 800 nucleotides before performing biotin–maleimide labeling. The fragmented total RNA was treated with biotin–maleimide (MI+), followed by an affinity pull down (MI+PD+). The pull-down RNA was eluted from the magnetic beads as the first elution product (E1) and the flow-through solution (FT1) was collected. The RNA purified from FT1 was used in a secondary pull-down to obtain the second elution product (E2); the second flow-through solution (FT2) was also collected (Figure 32). The cyanoethylated total RNA (CE+, 1.6 M acrylonitrile for 60 min) was treated with acrylonitrile and used as a positive control (Sakurai *et al.*, 2010b). The products were used as templates for primer extension. Gradient experiments were first performed to determine the binding and enrichment of different biotin–maleimide concentrations with inosine in a 1.1 M TEAA buffer (Figure 33). As shown in Figure 33, higher concentrations of biotin–maleimide were associated with greater amounts of labeled inosine and greater amounts of isolated mRNA. The inosine-labeling yield by biotin–maleimide was similar to that of CE+ (1.6 M acrylonitrile for 60 min) under the strongest reaction conditions (50 mM maleimide for 60 min). Based on this observation, we further studied the pull-down efficiency specifically at the strongest reaction conditions. As shown in Figure 34, CE+ produced a band corresponding to the cDNA that was terminated immediately before the cyanoethylated inosine site in the Q-to-R site of the *Gria2* mRNA. When using untreated total RNA (MI-) as the template, cDNA synthesis continued from the inosine site to the end of the strand. A band was absent at the same position as that of the CE+ sample because the products were too long to be detected via PAGE. For MI+, we observed a band at the same position as that of CE+, which corresponded to the cDNA synthesis terminating before the maleimide-labeled inosine site. This

indicated the steric hinderance of base pairing between inosine and cytidine (C), as observed in cyanoethylated inosine. A band terminating at the inosine site was observed in both MI+PD+ samples. E1 exhibited a 55% yield (Figure 34). Although FT1 contained some inosine-labeled RNA, most were obtained in E2 after both pull downs (Figure 34). The inosine-labeled RNA yield in E1+E2 was 70%. Additionally, no bands corresponding to inosine were detected in FT2. These results indicated that the specific labeling and pull-down of nucleic acids containing inosine can be achieved within the total RNA fragments of actual tissues.

Considering that current chemical methods cannot completely overcome nonspecific labeling, even though maleimide has demonstrated better specificity and reactivity compared with other compounds, it still exhibits nonspecific labeling of other bases. This may introduce potential interference in subsequent sequencing analyses. Clearly, the critical factor is not the yield of affinity enrichment, but the purity of the product. Previous primer extension methods were unable to detect the purification status of RNA with nonspecific labeling. Therefore, we subjected the same products to primer extension in the presence of in the presence of α-^{32}P dCTP and unlabeled primers (Figure 35). The band signal strength was related to the number of cytidines in the cDNA. Non-specific labeling on bases interrupted primer extension, producing multiple cDNA bands terminated at these labeled sites. Clear bands were observed for CE+ and MI+, corresponding to the cDNAs terminated at the inosine site (Figure 35). However, multiple background bands, indicative of nonspecific labeling, were also detected. Notably, the application of affinity pull-down effectively removed the background noise in E1 and E2. This indicated the successful exclusion of non-specifically labeled molecules, underscoring the efficacy of our method in selectively purifying inosine-containing RNA.

Detecting novel inosine sites in total RNA and gDNA using ICLAMP-seq

Materials and Method

Chemical compound preparation

Biotin–maleimide was purchased from Tokyo Chemical Industry. Biotin–maleimide stock solutions were prepared by dissolving each in dimethyl sulfoxide (DMSO) to a concentration of 250 mM.

Total RNA preparation

The total RNA of cultured cells (HEK293T) was purified using a miRNeasy Mini kit (QIAGEN). The total RNA from mouse brain tissues was extracted using an RNeasy Maxi Kit (QIAGEN). The mouse brain tissues were obtained from 8-week-old C57BL/6 mice in the Goitsuka Laboratory at the Tokyo University of Science. Upon collection, tissues were flash-frozen and stored at -80°C until use. All procedures were carried out in accordance with approved protocols of the Tokyo University of Science Animal Care and Use Committee. The total RNA (40 μg) was fragmented by incubating it in a 100 μL reaction solution containing magnesium ions (100 mM Tris-HCl and 2 mM $MgCl_2$) at 95°C for 8 min. Subsequently, the fragmented total RNA was purified by ethanol precipitation.

Genomic DNA preparation

The genomic DNA (gDNA) of cultured cells was purified using a Blood & Cell Culture DNA Midi kit (QIAGEN). The gDNA was fragmented by sonication (SONICS).

Total RNA or gDNA labeling and purification

Fragmented total RNA or gDNA (30 μg) was added to 38 μL of TEAA buffer (50% (v/v) ethanol, 1.1 M triethylammonium, pH adjusted to 8.6 with acetic acid) and vortexed vigorously. Biotin–maleimide was then added to the reaction solution to final concentrations of 0, 2, 10, and 50 mM. The reaction solution was then replenished with DMSO to a final volume of 50 μL and incubated at 70°C. At the indicated time points, 400 μL of cold water was added to the reaction mixture to stop the reaction. Excess maleimide was removed as follows: The treated samples were purified via isopropanol precipitation and two Phenol-chloroform-isoamyl alcohol mixture (PCI, Sigma, pH 7.7-8.3) treatments, then treated with diethyl ether to remove the remaining phenol. Residual diethyl ether was removed using an Eppendorf concentrator. Finally, the treated RNA was purified via ethanol precipitation.

Affinity pull-down

Total RNA or gDNA labeled with biotin–maleimide were pulled down using Dynabeads MyOne Streptavidin C1 magnetic beads (Thermo Fisher Scientific) as per the manufacturer's protocol. The recommended buffers were modified by referring to the following detailed steps for the exact formulation. The tubes were pretreated to prevent the nonspecific adhesion of the beads and nucleic acids to the inner wall of the tubes during suspension. A 2 mL tube was filled with 1.5 mL of a blocking buffer (25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer at pH 7.5, 1 M NaCl, 12.5 mM ethylenediamine tetra-acetic acid (EDTA), 0.005% Tween 20, and 1% bovine serum albumin) and inverted for 1 h using an Intelli mixer (ELMI). The tube was then washed twice with 1 mL of 1x binding buffer (25 mM HEPES buffer at pH 7.5, 1 M NaCl, 12.5 mM EDTA, and 0.005% Tween 20). Appropriate quantities of the beads (10 μL for oligonucleotides and 120 μL for the total RNA) were removed and washed three times with 500 μL blocking buffer for 2 min; this step was repeated three times. The washed magnetic beads were resuspended in 2x binding buffer to a final concentration of 5 μg/μL (twice the original volume). An equal volume of biotinylated DNA or RNA in distilled water was added to the beads and incubated for 60 min at room temperature (25°C) with gentle rotation (Eppendorf ThermoMixer). The biotinylated DNA- or RNA-coated beads were separated using a magnet for 2–3 min. The supernatant was collected, and the coated beads were washed two or three times with 500 μL 1x washing buffer (5 mM HEPES buffer at pH 7.5, 200 mM NaCl, 2.5 mM EDTA, and 0.005% Tween 20). The nucleic acids were released by boiling the beads in 100% formamide at 70°C for 5 min. The tube was placed on a magnet and the supernatant solution was transferred to a new tube. The isolated nucleic acids were purified by ethanol precipitation.

Detecting editing sites in total RNA by ICLAMP-seq

In the ICLAMP method, selective enrichment and purification of nucleic acids containing inosine are achieved through the labeling of inosine with biotin-maleimide. This specific labeling technique allows for the targeted concentration of inosine-containing nucleic acids. To detect novel A-to-I editing sites in the human RNA and DNA, we have developed a new strategy using the ICLAMP method combined with next-generation sequencing (NGS), which we call ICLAMP-seq.

In cases where RNA is the target of analysis, the principles and procedures for ICLAMP-seq analysis are outlined in Figure 36. The first step of the procedure is the same as that of the ICLAMP method. For comparison, we utilized the ICE method as a positive control. The fragmented total RNA

is reacted with acrylonitrile (CE+) or biotin-maleimide (MI+) first. MI+ RNA will be further purified by affinity pull-down (MI+PD+), total RNA in flow through will be collected (FT). After the removal of ribosomal RNA (rRNA), the first-strand cDNA synthesis was carried out. In this step, the maleimide acts as a blocker of strand extension, causing premature termination and resulting in the production of shorter cDNA fragments. Following the synthesis of the second strand, both ends of the DNA strands undergo end repair and are subsequently ligated with adaptors. Then the gel electrophoresis was performed to separate targeted bands. The bands were arranged into two sets, one consisting of longer bands and the other of shorter bands. From the MI+PD+ group, the shorter bands were excised, while for the other groups, longer bands were excised. After the gel was purified, PCR amplification was conducted for use in Next Generation Sequencing (NGS). The sequencing reads were then mapped to the human transcriptome and genome. In the MI+PD+ condition, the site immediately following the termination point of the reads was identified as the editing site. Correspondingly, compared to the MI- condition, reads that were missing in both the MI+ and CE+ conditions were detected. The data analysis of RNA sequencing results is currently underway. The forthcoming objectives include the identification of new sites and the development and exploration of novel sequencing methodologies.

Detecting editing sites in gDNA by ICLAMP-seq

In cases where DNA is the target of analysis, the principles and procedures for ICLAMP-seq analysis are outlined in Figure 37. First, the fragmented gDNA samples including double-stranded DNA (dsDNA) and single-stranded DNA (ssDNA) are reacted with biotin-maleimide (MI+). MI+ DNA will be further purified by affinity pull-down (MI+PD+), gDNA in flow through will be collected (FT). Following the end repair and dA-tailing, the blunt-ended fragments with 5' phosphorylation were generated and a single A 3' overhang was created that enables ligation to adaptors with single T overhangs. However, the ssDNA will be removed during the adaptor ligation process with current NGS techniques. Then the ligated dsDNA is ready for PCR-free sequencing. In a comparative analysis with the MI- condition, the MI+ condition will reveal the detection of missing G signals. Furthermore, in the MI+PD+ condition, both missing G signals and absent A/T pairing information will be identified. Conversely, the FT samples will exhibit solely the presence of A/T pairing information. Investigations employing traditional DNA sequencing techniques to detect inosine sites are currently in progress. Given the limitations of conventional methods in detecting

ssDNA, our focus is shifting towards the development of novel DNA library preparation methodologies that can concurrently assess dsDNA and ssDNA.

Discussion

In this study, we reported that maleimide exhibited a high specificity and efficacy in RNA and DNA inosine labeling. Leveraging this capability, we developed ICLAMP. This approach used commercially available maleimide derivatives, namely fluorescein–maleimide for visualization and the quantification of labeling efficiency and biotin–maleimide for the targeted enrichment of inosine-containing nucleic acids. We have determined the optimal reaction conditions, efficiency, and specificity for maleimide to label inosine within nucleic acids. We have demonstrated that this method can selectively isolate and purify inosine-containing nucleic acids from complex nucleic acid mixtures. This approach is straightforward and rapid, and can serve as a nucleic acid preprocessing step to assist in elucidating the true scale and dynamics of A-to-I editing.

Chemically labeling modified bases is valuable for screening DNA and RNA modifications (Frommer *et al.*, 1992; Schaefer *et al.*, 2009). Chemical labeling coupled with affinity enrichment significantly improves the sensitivity and accuracy of modified base identification (Dominissini *et al.*, 2012; Li *et al.*, 2015, 2016; Wu, Amrutkar and Shao, 2018). Based on this concept, several methods for chemically labeling A-to-I editing sites have been developed. However, while some methods have labeling specificities, they often lack the flexibility for subsequent modifications (Sakurai *et al.*, 2010b, 2014b). Conversely, newly developed derivatives that address these limitations have poor labeling efficiencies and remain commercially inaccessible (Knutson, Ayele and Heemstra, 2018; Steve D Knutson *et al.*, 2020). The maleimide used in ICLAMP addressed these limitations, enabling efficient labeling of inosine and effective enrichment of nucleic acid molecules containing inosine. However, it is imperative to note that ICLAMP might also be potentially constrained by specificity issues. Notably, the observed reactivity between maleimide and uracil may seem problematic for the specific labeling and enrichment of inosine-containing nucleic acids. As the reaction time increases and the concentration of maleimide, the corresponding rate of nonspecific reactions can also escalate, leading to a reduced enrichment efficiency for actual inosine within RNA and DNA samples during the pull-down process. Employing lower concentrations of maleimide and shorter reaction durations might be

the most straightforward measures to enhance specificity. Moreover, the interference from uracil can be absent when analyzing DNA samples, significantly enhancing the precision of inosine labeling.

ICLAMP has potential applications beyond its core focus. Fluorescent inosine labeling can offer a rapid and efficient avenue to quantify inosines in nucleic acid solutions; however, further studies are needed to confirm this. While the affinity-based enrichment in ICLAMP could increase the detection sensitivity, its full utility in complementing NGS remains unexplored. This could serve as a pre-sequencing step for enriching inosine sites within nucleic acids, enhancing the data yields from limited input materials. The method could also be used to identify inosine in various nucleic acids, including those present in low quantities. As such, ICLAMP is a potentially valuable tool in both research and diagnostic applications. Moreover, the inherent flexibility of maleimide scaffolds enables the integration of various functional groups, ranging from Cy5 to streptavidin, thereby facilitating future studies into novel applications.

References

Alseth, I., Dalhus, B. and Bjørås, M. (2014) "Inosine in DNA and RNA," *Current Opinion in Genetics & Development*, 26, pp. 116–123. Available at: https://doi.org/10.1016/j.gde.2014.07.008.

Athanasiadis, A., Rich, A. and Maas, S. (2004) "Widespread A-to-I RNA Editing of Alu-Containing mRNAs in the Human Transcriptome," *PLoS Biology*, 2(12), p. e391. Available at: https://doi.org/10.1371/journal.pbio.0020391.

Bass, B.L. and Weintraub, H. (1987) "A developmentally regulated activity that unwinds RNA duplexes," *Cell*, 48(4), pp. 607–613. Available at: https://doi.org/10.1016/0092-8674(87)90239-x.

Bazak, L. *et al.* (2014) "A-to-I RNA editing occurs at over a hundred million genomic sites, located in a majority of human genes," *Genome Research*, 24(3), pp. 365–376. Available at: https://doi.org/10.1101/gr.164749.113.

Bhalla, T. *et al.* (2004) "Control of human potassium channel inactivation by editing of a small mRNA hairpin," *Nature Structural & Molecular Biology*, 11(10), pp. 950–956. Available at: https://doi.org/10.1038/nsmb825.

Bhartiya, D. and Scaria, V. (2016) "Genomic variations in non-coding RNAs: Structure, function and regulation," *Genomics*, 107(2–3), pp. 59–68. Available at: https://doi.org/10.1016/j.ygeno.2016.01.005.

Blencowe, B.J. (2006) "Alternative Splicing: New Insights from Global Analyses," *Cell*, 126(1), pp. 37–47. Available at: https://doi.org/10.1016/j.cell.2006.06.023.

Blow, M. *et al.* (2004) "A survey of RNA editing in human brain," *Genome Research*, 14(12), pp. 2379–2387. Available at: https://doi.org/10.1101/gr.2951204.

Brümmer, A. *et al.* (2017) "Structure-mediated modulation of mRNA abundance by A-to-I editing," *Nature Communications*, 8(1), p. 1255. Available at: https://doi.org/10.1038/s41467-017-01459-7.

Brusa, R. *et al.* (1995) "Early-Onset Epilepsy and Postnatal Lethality Associated with an Editing-Deficient GluR-B Allele in Mice," *Science*, 270(5242), pp. 1677–1680. Available at: https://doi.org/10.1126/science.270.5242.1677.

Burns, C.M. *et al.* (1997) "Regulation of serotonin-2C receptor G-protein coupling by RNA editing," *Nature*, 387(6630), pp. 303–308. Available at: https://doi.org/10.1038/387303a0.

Cattaneo, R. and Billeter, M.A. (1992) "Genetic Diversity of RNA Viruses," *Current Topics in Microbiology and Immunology*, pp. 63–74. Available at: https://doi.org/10.1007/978-3-642-77011-1_5.

Cattenoz, P. *et al.* (2013) "Transcriptome-wide identification of A > I RNA editing sites by inosine specific cleavage.," 19(2). Available at: https://doi.org/10.1261/rna.036202.112.

Chen, C.-X. *et al.* (2000) "A third member of the RNA-specific adenosine deaminase gene family, ADAR3, contains both single- and double-stranded RNA binding domains," *RNA*, 6(5), pp. 755–767. Available at: https://doi.org/10.1017/s1355838200000170.

Chen, J.-J. *et al.* (2022) "Single-Base Resolution Detection of Adenosine-to-Inosine RNA Editing by Endonuclease-Mediated Sequencing," *Analytical Chemistry*, 94(24), pp. 8740–8747. Available at: https://doi.org/10.1021/acs.analchem.2c01226.

Chen, L. *et al.* (2013) "Recoding RNA editing of AZIN1 predisposes to hepatocellular carcinoma," *Nature Medicine*, 19(2), pp. 209–216. Available at: https://doi.org/10.1038/nm.3043.

Choudhury, Y. *et al.* (2012) "Attenuated adenosine-to-inosine editing of microRNA-376a* promotes invasiveness of glioblastoma cells," *Journal of Clinical Investigation*, 122(11), pp. 4059–4076. Available at: https://doi.org/10.1172/jci62925.

Decher, N. *et al.* (2010) "RNA editing modulates the binding of drugs and highly unsaturated fatty acids to the open pore of Kv potassium channels," *The EMBO Journal*, 29(13), pp. 2101–2113. Available at: https://doi.org/10.1038/emboj.2010.88.

Desterro, J.M.P. *et al.* (2003) "Dynamic association of RNA-editing enzymes with the nucleolus," *Journal of Cell Science*, 116(9), pp. 1805–1818. Available at: https://doi.org/10.1242/jcs.00371.

Di Giammartino, D.C., Nishida, K. and Manley, J.L. (2011) "Mechanisms and Consequences of Alternative Polyadenylation," *Molecular Cell*, 43(6), pp. 853–866. Available at: https://doi.org/10.1016/j.molcel.2011.08.017.

Dominissini, D. *et al.* (2012) "Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq," *Nature*, 485(7397), pp. 201–206. Available at: https://doi.org/10.1038/nature11112.

Dong, X. *et al.* (2018) "CDK13 RNA Over-Editing Mediated by ADAR1 Associates with Poor Prognosis of Hepatocellular Carcinoma Patients," *Cellular Physiology and Biochemistry*, 47(6), pp. 2602–2612. Available at: https://doi.org/10.1159/000491656.

Duan, Y., Tang, X. and Lu, J. (2022) "Evolutionary driving forces of A-to-I editing in metazoans," *Wiley Interdisciplinary Reviews: RNA*, 13(1), p. e1666. Available at: https://doi.org/10.1002/wrna.1666.

Ekdahl, Y. *et al.* (2012) "A-to-I editing of microRNAs in the mammalian brain increases during development," *Genome Research*, 22(8), pp. 1477–1487. Available at: https://doi.org/10.1101/gr.131912.111.

Erdmann, E.A. *et al.* (2020) "To protect and modify double-stranded RNA – the critical roles of ADARs in development, immunity and oncogenesis," *Critical Reviews in Biochemistry and Molecular Biology*, 56(1), pp. 1–50. Available at: https://doi.org/10.1080/10409238.2020.1856768.

Frommer, M. *et al.* (1992) "A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands.," *Proceedings of the National Academy of Sciences of the United States of America*, 89(5), pp. 1827–1831. Available at: https://doi.org/10.1073/pnas.89.5.1827.

Gelinas, J.-F. *et al.* (2011) "Enhancement of Replication of RNA Viruses by ADAR1 via RNA Editing and Inhibition of RNA-Activated Protein Kinase," *Journal of Virology*, 85(17), pp. 8460–8466. Available at: https://doi.org/10.1128/jvi.00240-11.

Gillingham, D., Geigle, S. and Lilienfeld, O.A. von (2016) "Properties and reactivity of nucleic acids relevant to epigenomics, transcriptomics, and therapeutics." Available at: https://doi.org/10.1039/c5cs00271k.

Giorgio, S.D. *et al.* (2020) "Evidence for host-dependent RNA editing in the transcriptome of SARS-CoV-2," *Science Advances*, 6(25), p. eabb5813. Available at: https://doi.org/10.1126/sciadv.abb5813.

Gong, J. *et al.* (2017) "LNCediting: a database for functional effects of RNA editing in lncRNAs," *Nucleic Acids Research*, 45(Database issue), pp. D79–D84. Available at: https://doi.org/10.1093/nar/gkw835.

Greger, I.H., Khatri, L. and Ziff, E.B. (2002) "RNA Editing at Arg607 Controls AMPA Receptor Exit from the Endoplasmic Reticulum," *Neuron*, 34(5), pp. 759–772. Available at: https://doi.org/10.1016/s0896-6273(02)00693-1.

Herbert, A. *et al.* (1997) "A Z-DNA binding domain present in the human editing enzyme, double-stranded RNA adenosine deaminase," *Proceedings of the National Academy of Sciences*, 94(16), pp. 8421–8426. Available at: https://doi.org/10.1073/pnas.94.16.8421.

Higuchi, M. *et al.* (1993) "RNA editing of AMPA receptor subunit GluR-B: A base-paired intron-exon structure determines position and efficiency," *Cell*, 75(7), pp. 1361–1370. Available at: https://doi.org/10.1016/0092-8674(93)90622-w.

Higuchi, M. *et al.* (2000) "Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2," *Nature*, 406(6791), pp. 78–81. Available at: https://doi.org/10.1038/35017558.

Hoopengardner, B. *et al.* (2003) "Nervous System Targets of RNA Editing Identified by Comparative Genomics," *Science*, 301(5634), pp. 832–836. Available at: https://doi.org/10.1126/science.1086763.

Huang, H. *et al.* (2012) "RNA Editing of the IQ Domain in Cav1.3 Channels Modulates Their Ca2+-Dependent Inactivation," *Neuron*, 73(2), pp. 304–316. Available at: https://doi.org/10.1016/j.neuron.2011.11.022.

Iizasa, H. *et al.* (2010) "Editing of Epstein-Barr Virus-encoded BART6 MicroRNAs Controls Their Dicer Targeting and Consequently Affects Viral Latency*," *Journal of Biological Chemistry*, 285(43), pp. 33358–33370. Available at: https://doi.org/10.1074/jbc.m110.138362.

Ivanov, A. *et al.* (2015) "Analysis of Intron Sequences Reveals Hallmarks of Circular RNA Biogenesis in Animals," *Cell Reports*, 10(2), pp. 170–177. Available at: https://doi.org/10.1016/j.celrep.2014.12.019.

Jimeno, S. *et al.* (2021) "ADAR-mediated RNA editing of DNA:RNA hybrids is required for DNA double strand break repair," *Nature Communications*, 12(1), p. 5512. Available at: https://doi.org/10.1038/s41467-021-25790-2.

Jin, Y., Zhang, W. and Li, Q. (2009) "Origins and evolution of ADAR-mediated RNA editing," *IUBMB Life*, 61(6), pp. 572–578. Available at: https://doi.org/10.1002/iub.207.

Joubert, N. *et al.* (2017) "Towards antibody-drug conjugates and prodrug strategies with extracellular stimuli-responsive drug delivery in the tumor microenvironment for cancer therapy," *European Journal of Medicinal Chemistry*, 142, pp. 393–415. Available at: https://doi.org/10.1016/j.ejmech.2017.08.049.

Kawahara, Y., Zinshteyn, B., Sethupathy, P., *et al.* (2007) "Redirection of Silencing Targets by Adenosine-to-Inosine Editing of miRNAs," *Science*, 315(5815), pp. 1137–1140. Available at: https://doi.org/10.1126/science.1138050.

Kawahara, Y., Zinshteyn, B., Chendrimada, T.P., *et al.* (2007) "RNA editing of the microRNA-151 precursor blocks cleavage by the Dicer–TRBP complex," *EMBO reports*, 8(8), pp. 763–769. Available at: https://doi.org/10.1038/sj.embor.7401011.

Kawahara, Y. *et al.* (2008) "Frequency and fate of microRNA editing in human brain," *Nucleic Acids Research*, 36(16), pp. 5270–5280. Available at: https://doi.org/10.1093/nar/gkn479.

Kim, D. *et al.* (2020) "The Architecture of SARS-CoV-2 Transcriptome," *Cell*, 181(4), pp. 914-921.e10. Available at: https://doi.org/10.1016/j.cell.2020.04.011.

Kim, D.D.Y. *et al.* (2004) "Widespread RNA Editing of Embedded Alu Elements in the Human Transcriptome," *Genome Research*, 14(9), pp. 1719–1725. Available at: https://doi.org/10.1101/gr.2855504.

Kim, U. *et al.* (1994) "Molecular cloning of cDNA for double-stranded RNA adenosine deaminase, a candidate enzyme for nuclear RNA editing," *Proceedings of the National Academy of Sciences*, 91(24), pp. 11457–11461. Available at: https://doi.org/10.1073/pnas.91.24.11457.

Kleinman, C.L. and Majewski, J. (2012) "Comment on 'Widespread RNA and DNA Sequence Differences in the Human Transcriptome,'" *Science*, 335(6074), pp. 1302–1302. Available at: https://doi.org/10.1126/science.1209658.

Knutson, Steve D *et al.* (2020) "Chemical Profiling of A-to-I RNA Editing Using a Click-Compatible Phenylacrylamide," *Chemistry - A European Journal*, 26(44), pp. 9874–9878. Available at: https://doi.org/10.1002/chem.202001667.

Knutson, Steve D. *et al.* (2020) "Selective Enrichment of A-to-I Edited Transcripts from Cellular RNA Using Endonuclease V," *Journal of the American Chemical Society*, 142(11), pp. 5241–5251. Available at: https://doi.org/10.1021/jacs.9b13406.

Knutson, S.D., Ayele, T.M. and Heemstra, J.M. (2018) "Chemical Labeling and Affinity Capture of Inosine-Containing RNAs Using Acrylamidofluorescein." Available at: https://doi.org/10.1021/acs.bioconjchem.8b00541.

Leger, A. *et al.* (2021) "RNA modifications detection by comparative Nanopore direct RNA sequencing," *Nature Communications*, 12(1), p. 7198. Available at: https://doi.org/10.1038/s41467-021-27393-3.

Levanon, E.Y. *et al.* (2004) "Systematic identification of abundant A-to-I editing sites in the human transcriptome," *Nature Biotechnology*, 22(8), pp. 1001–1005. Available at: https://doi.org/10.1038/nbt996.

Lev-Maor, G. *et al.* (2007) "RNA-editing-mediated exon evolution," *Genome Biology*, 8(2), p. R29. Available at: https://doi.org/10.1186/gb-2007-8-2-r29.

Li, J.B. *et al.* (2009) "Genome-Wide Identification of Human RNA Editing Sites by Parallel DNA Capturing and Sequencing," *Science*, 324(5931), pp. 1210–1213. Available at: https://doi.org/10.1126/science.1170995.

Li, M. *et al.* (2011) "Widespread RNA and DNA Sequence Differences in the Human Transcriptome," *Science*, 333(6038), pp. 53–58. Available at: https://doi.org/10.1126/science.1207018.

Li, M.G.K.K.C.D.V. and R.C.S.Y. (2019) "Identification of Adenosine-to-Inosine RNA Editing with Acrylonitrile Reagents," *Organic Letters*, 21(19), pp. 1–4. Available at: https://doi.org/10.1021/acs.orglett.9b02929.

Li, X. *et al.* (2015) "Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome," *Nature Chemical Biology*, 11(8), pp. 592–597. Available at: https://doi.org/10.1038/nchembio.1836.

Li, X. *et al.* (2016) "Transcriptome-wide mapping reveals reversible and dynamic N1-methyladenosine methylome," *Nature Publishing Group*, 12(5), pp. 311–316. Available at: https://doi.org/10.1038/nchembio.2040.

Liddicoat, B.J. *et al.* (2015) "RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself," *Science*, 349(6252), pp. 1115–1120. Available at: https://doi.org/10.1126/science.aac7049.

Lin, W. *et al.* (2012) "Comment on 'Widespread RNA and DNA Sequence Differences in the Human Transcriptome,'" *Science*, 335(6074), pp. 1302–1302. Available at: https://doi.org/10.1126/science.1210624.

Lindahl, T. (1993) "Instability and decay of the primary structure of DNA," *Nature*, 362(6422), pp. 709–715. Available at: https://doi.org/10.1038/362709a0.

Luyten, I. *et al.* (1998) "Determination of the Tautomeric Equlibrium of Ψ-Uridine in the Basic Solution," *The Journal of Organic Chemistry*, 63(4), pp. 1033–1040. Available at: https://doi.org/10.1021/jo971348o.

Maas, S. *et al.* (2011) "Genome-wide evaluation and discovery of vertebrate A-to-I RNA editing sites," *Biochemical and Biophysical Research Communications*, 412(3), pp. 407–412. Available at: https://doi.org/10.1016/j.bbrc.2011.07.075.

Mannion, N.M. *et al.* (2014) "The RNA-Editing Enzyme ADAR1 Controls Innate Immune Responses to RNA," *Cell Reports*, 9(4), pp. 1482–1494. Available at: https://doi.org/10.1016/j.celrep.2014.10.041.

Melcher, T., Maas, S., Herb, A., Sprengel, R., Seeburg, P.H., *et al.* (1996) "A mammalian RNA editing enzyme," *Nature*, 379(6564), pp. 460–464. Available at: https://doi.org/10.1038/379460a0.

Melcher, T., Maas, S., Herb, A., Sprengel, R., Higuchi, M., *et al.* (1996) "RED2, a Brain-specific Member of the RNA-specific Adenosine Deaminase Family*," *Journal of Biological Chemistry*, 271(50), pp. 31795–31798. Available at: https://doi.org/10.1074/jbc.271.50.31795.

Mingardi, J. *et al.* (2018) "miRNA Editing: New Insights into the Fast Control of Gene Expression in Health and Disease," *Molecular Neurobiology*, 55(10), pp. 7717–7727. Available at: https://doi.org/10.1007/s12035-018-0951-x.

Mladenova, D. *et al.* (2018) "Adar3 Is Involved in Learning and Memory in Mice," *Frontiers in Neuroscience*, 12, p. 243. Available at: https://doi.org/10.3389/fnins.2018.00243.

Moe, A. *et al.* (2003) "Incision at hypoxanthine residues in DNA by a mammalian homologue of the Escherichia coli antimutator enzyme endonuclease V.," *Nucleic acids research*, 31(14), pp. 3893–3900. Available at: https://doi.org/10.1093/nar/gkg472.

Morita, Y. *et al.* (2013) "Human endonuclease V is a ribonuclease specific for inosine-containing RNA," *Nature communications*, 4(1), p. 709. Available at: https://doi.org/10.1038/ncomms3273.

Morse, D.P. (2004) "Identification of substrates for adenosine deaminases that act on RNA.," *Methods in molecular biology (Clifton, N.J.)*, 265, pp. 199–218. Available at: https://doi.org/10.1385/1-59259-775-0:199.

Morse, D.P. and Bass, B.L. (1997) "Detection of Inosine in Messenger RNA by Inosine-Specific Cleavage †," *Biochemistry*, 36(28), pp. 8429–8434. Available at: https://doi.org/10.1021/bi9709607.

Morse, D.P. and Bass, B.L. (1999) "Long RNA hairpins that contain inosine are present in Caenorhabditis elegans poly(A)+ RNA," 96(11), pp. 6048–6053. Available at: https://doi.org/10.1073/pnas.96.11.6048.

Moss, W.J. and Griffin, D.E. (2006) "Global measles elimination," *Nature Reviews. Microbiology*, 4(12), pp. 900–908. Available at: https://doi.org/10.1038/nrmicro1550.

Nguyen, T. *et al.* (1992) "DNA damage and mutation in human cells exposed to nitric oxide in vitro.," *Proceedings of the National Academy of Sciences*, 89(7), pp. 3030–3034. Available at: https://doi.org/10.1073/pnas.89.7.3030.

Niehrs, C. and Luke, B. (2020) "Regulatory R-loops as facilitators of gene expression and genome stability," *Nature Reviews Molecular Cell Biology*, 21(3), pp. 167–178. Available at: https://doi.org/10.1038/s41580-019-0206-3.

Oakes, E. *et al.* (2017) "Adenosine Deaminase That Acts on RNA 3 (ADAR3) Binding to Glutamate Receptor Subunit B Pre-mRNA Inhibits RNA Editing in Glioblastoma*," *Journal of Biological Chemistry*, 292(10), pp. 4326–4335. Available at: https://doi.org/10.1074/jbc.m117.779868.

Palladino, M.J. *et al.* (2000) "A-to-I Pre-mRNA Editing in Drosophila Is Primarily Involved in Adult Nervous System Function and Integrity," *Cell*, 102(4), pp. 437–449. Available at: https://doi.org/10.1016/s0092-8674(00)00049-0.

Paul, D. *et al.* (2017) "A-to-I editing in human miRNAs is enriched in seed sequence, influenced by sequence contexts and significantly hypoedited in glioblastoma multiforme," *Scientific Reports*, 7(1), p. 2466. Available at: https://doi.org/10.1038/s41598-017-02397-6.

Peng, Z. *et al.* (2012) "Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome," *Nature Biotechnology*, 30(3), pp. 253–260. Available at: https://doi.org/10.1038/nbt.2122.

Picardi, E. *et al.* (2017) "REDIportal: a comprehensive database of A-to-I RNA editing events in humans," *Nucleic Acids Research*, 45(D1), pp. D750–D757. Available at: https://doi.org/10.1093/nar/gkw767.

Picardi, E., Mansi, L. and Pesole, G. (2021) "Detection of A-to-I RNA Editing in SARS-COV-2," *Genes*, 13(1), p. 41. Available at: https://doi.org/10.3390/genes13010041.

Pickrell, J.K., Gilad, Y. and Pritchard, J.K. (2012) "Comment on 'Widespread RNA and DNA Sequence Differences in the Human Transcriptome,'" *Science*, 335(6074), pp. 1302–1302. Available at: https://doi.org/10.1126/science.1210484.

Pinto, Y. *et al.* (2017) "Human cancer tissues exhibit reduced A-to-I editing of miRNAs coupled with elevated editing of their targets," *Nucleic Acids Research*, 46(1), pp. gkx1176-. Available at: https://doi.org/10.1093/nar/gkx1176.

Poulsen, H. *et al.* (2001) "CRM1 Mediates the Export of ADAR1 through a Nuclear Export Signal within the Z-DNA Binding Domain," *Molecular and Cellular Biology*, 21(22), pp. 7862–7871. Available at: https://doi.org/10.1128/mcb.21.22.7862-7871.2001.

Pujantell, M. *et al.* (2018) "ADAR1 affects HCV infection by modulating innate immune response," *Antiviral Research*, 156, pp. 116–127. Available at: https://doi.org/10.1016/j.antiviral.2018.05.012.

Ramaswami, G. and Li, J.B. (2014) "RADAR: a rigorously annotated database of A-to-I RNA editing," *Nucleic Acids Research*, 42(D1), pp. D109–D113. Available at: https://doi.org/10.1093/nar/gkt996.

Ramírez-Moya, J. *et al.* (2020) "ADAR1-mediated RNA editing is a novel oncogenic process in thyroid cancer and regulates miR-200 activity," *Oncogene*, 39(18), pp. 3738–3753. Available at: https://doi.org/10.1038/s41388-020-1248-x.

Ramírez-Moya, J. *et al.* (2021) "An ADAR1-dependent RNA editing event in the cyclin-dependent kinase CDK13 promotes thyroid cancer hallmarks," *Molecular Cancer*, 20(1), p. 115. Available at: https://doi.org/10.1186/s12943-021-01401-y.

Ravasco, J.M.J.M. *et al.* (2018) "Bioconjugation with Maleimides: A Useful Tool for Chemical Biology," *Chemistry - A European Journal*, 25(1), pp. 43–59. Available at: https://doi.org/10.1002/chem.201803174.

Raymond, D. *et al.* (1999) "The Glutamate Receptor Ion Channels," 51(1), p. 7. Available at: http://pharmrev.aspetjournals.org/content/51/1/7.abstract.

Rebagliati, M.R. and Melton, D.A. (1987) "Antisense RNA injections in fertilized frog eggs reveal an RNA duplex unwinding activity," *Cell*, 48(4), pp. 599–605. Available at: https://doi.org/10.1016/0092-8674(87)90238-8.

Rebendenne, A. *et al.* (2021) "SARS-CoV-2 Triggers an MDA-5-Dependent Interferon Response Which Is Unable To Control Replication in Lung Epithelial Cells," *Journal of Virology*, 95(8), pp. e02415-20. Available at: https://doi.org/10.1128/jvi.02415-20.

Renault, K. *et al.* (2018) "Covalent Modification of Biomolecules through Maleimide-Based Labeling Strategies," *Bioconjugate Chemistry*, 29(8), pp. 2497–2513. Available at: https://doi.org/10.1021/acs.bioconjchem.8b00252.

Rengaraj, P. *et al.* (2021) "Interplays of different types of epitranscriptomic mRNA modifications," *RNA Biology*, 18(Suppl 1), pp. 19–30. Available at: https://doi.org/10.1080/15476286.2021.1969113.

Sakaguchi, Y. *et al.* (2015) "Nucleoside Analysis by Hydrophilic Interaction Liquid Chromatography Coupled with Mass Spectrometry.," *Methods in enzymology*, 560, pp. 19–28. Available at: https://doi.org/10.1016/bs.mie.2015.03.015.

Sakurai, M. *et al.* (2010a) "Inosine cyanoethylation identifies A-to-I RNA editing sites in the human transcriptome," *Nature Chemical Biology*, 6(10), pp. 733–740. Available at: https://doi.org/10.1038/nchembio.434.

Sakurai, M. *et al.* (2010b) "Inosine cyanoethylation identifies A-to-I RNA editing sites in the human transcriptome," *Nature chemical biology*, 6(10), pp. 733–740. Available at: https://doi.org/10.1038/nchembio.434.

Sakurai, M. *et al.* (2014a) "A biochemical landscape of A-to-I RNA editing in the human brain transcriptome," *Genome Research*, 24(3), pp. 522–534. Available at: https://doi.org/10.1101/gr.162537.113.

Sakurai, M. *et al.* (2014b) "A biochemical landscape of A-to-I RNA editing in the human brain transcriptome," *Genome research*, 24(3), p. gr.162537.113-534. Available at: https://doi.org/10.1101/gr.162537.113.

Sampaio, N.G. *et al.* (2021) "The RNA sensor MDA5 detects SARS-CoV-2 infection," *Scientific Reports*, 11(1), p. 13638. Available at: https://doi.org/10.1038/s41598-021-92940-3.

Samuel, C.E. (2011) "Adenosine Deaminases Acting on RNA (ADARs) and A-to-I Editing," *Current Topics in Microbiology and Immunology*, 353, pp. 163–195. Available at: https://doi.org/10.1007/82_2011_148.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) "DNA sequencing with chain-terminating inhibitors," *Proceedings of the National Academy of Sciences*, 74(12), pp. 5463–5467. Available at: https://doi.org/10.1073/pnas.74.12.5463.

Sarvestani, S.T. *et al.* (2013) "Inosine-Mediated Modulation of RNA Sensing by Toll-Like Receptor 7 (TLR7) and TLR8," *Journal of Virology*, 88(2), pp. 799–810. Available at: https://doi.org/10.1128/jvi.01571-13.

Schaefer, M. *et al.* (2009) "RNA cytosine methylation analysis by bisulfite sequencing.," *Nucleic acids research*, 37(2), pp. e12–e12. Available at: https://doi.org/10.1093/nar/gkn954.

Seeburg, P.H., Higuchi, M. and Sprengel, R. (1998) "RNA editing of brain glutamate receptor channels: mechanism and physiology," *Brain Research Reviews*, 26(2–3), pp. 217–229. Available at: https://doi.org/10.1016/s0165-0173(97)00062-3.

Shiromoto, Y. *et al.* (2021a) "ADAR1 RNA editing enzyme regulates R-loop formation and genome stability at telomeres in cancer cells," *Nature Communications*, 12(1), p. 1654. Available at: https://doi.org/10.1038/s41467-021-21921-x.

Shiromoto, Y. *et al.* (2021b) "ADAR1 RNA editing enzyme regulates R-loop formation and genome stability at telomeres in cancer cells," *Nature Communications*, 12(1), p. 1654. Available at: https://doi.org/10.1038/s41467-021-21921-x.

Sommer, B. *et al.* (1991) "RNA editing in brain controls a determinant of ion flow in glutamate-gated channels," *Cell*, 67(1), pp. 11–19. Available at: https://doi.org/10.1016/0092-8674(91)90568-j.

Song, Y. *et al.* (2021) "ADAR mediated A-to-I RNA editing affects SARS-CoV-2 characteristics and fuels its evolution," *bioRxiv*, p. 2021.07.22.453345. Available at: https://doi.org/10.1101/2021.07.22.453345.

Stefl, R. *et al.* (2006) "Structure and Specific RNA Binding of ADAR2 Double-Stranded RNA Binding Motifs," *Structure*, 14(2), pp. 345–355. Available at: https://doi.org/10.1016/j.str.2005.11.013.

Stellos, K. *et al.* (2016) "Adenosine-to-inosine RNA editing controls cathepsin S expression in atherosclerosis by enabling HuR-mediated post-transcriptional regulation," *Nature Medicine*, 22(10), pp. 1140–1150. Available at: https://doi.org/10.1038/nm.4172.

Strehblow, A., Hallegger, M. and Jantsch, M.F. (2002) "Nucleocytoplasmic Distribution of Human RNA-editing Enzyme ADAR1 Is Modulated by Double-stranded RNA-binding Domains, a Leucine-rich Export Signal, and a Putative Dimerization Domain," *Molecular Biology of the Cell*, 13(11), pp. 3822–3835. Available at: https://doi.org/10.1091/mbc.e02-03-0161.

Taylor, D.R. *et al.* (2005) "New Antiviral Pathway That Mediates Hepatitis C Virus Replicon Interferon Sensitivity through ADAR1," *Journal of Virology*, 79(10), pp. 6291–6298. Available at: https://doi.org/10.1128/jvi.79.10.6291-6298.2005.

Terajima, H. *et al.* (2017) "ADARB1 catalyzes circadian A-to-I editing and regulates RNA rhythm," *Nature Genetics*, 49(1), pp. 146–151. Available at: https://doi.org/10.1038/ng.3731.

Thaplyal, P. and Bevilacqua, P.C. (2014) "Chapter Nine Experimental Approaches for Measuring pKa's in RNA and DNA," *Methods in Enzymology*, 549, pp. 189–219. Available at: https://doi.org/10.1016/b978-0-12-801122-5.00009-x.

Tomaselli, S. *et al.* (2015) "Modulation of microRNA editing, expression and processing by ADAR2 deaminase in glioblastoma," *Genome Biology*, 16(1), p. 5. Available at: https://doi.org/10.1186/s13059-014-0575-z.

Toth, A.M. *et al.* (2009) "RNA-specific Adenosine Deaminase ADAR1 Suppresses Measles Virus-induced Apoptosis and Activation of Protein Kinase PKR*," *Journal of Biological Chemistry*, 284(43), pp. 29350–29356. Available at: https://doi.org/10.1074/jbc.m109.045146.

Venter, J.C. *et al.* (2001) "The Sequence of the Human Genome," *Science*, 291(5507), pp. 1304–1351. Available at: https://doi.org/10.1126/science.1058040.

Vesely, C. *et al.* (2012) "Adenosine deaminases that act on RNA induce reproducible changes in abundance and sequence of embryonic miRNAs," *Genome Research*, 22(8), pp. 1468–1476. Available at: https://doi.org/10.1101/gr.133025.111.

Vesely, C. and Jantsch, M.F. (2021) "An I for an A: Dynamic Regulation of Adenosine Deamination-Mediated RNA Editing," *Genes*, 12(7), p. 1026. Available at: https://doi.org/10.3390/genes12071026.

Vik, E.S. *et al.* (2013) "Endonuclease V cleaves at inosines in RNA," *Nature communications*, 4(1), p. 5235. Available at: https://doi.org/10.1038/ncomms3271.

Vitali, P. and Scadden, A.D.J. (2010) "Double-stranded RNAs containing multiple IU pairs are sufficient to suppress interferon induction and apoptosis," *Nature Structural & Molecular Biology*, 17(9), pp. 1043–1050. Available at: https://doi.org/10.1038/nsmb.1864.

Vogel, O.A. *et al.* (2020) "The p150 Isoform of ADAR1 Blocks Sustained RLR signaling and Apoptosis during Influenza Virus Infection," *PLoS Pathogens*, 16(9), p. e1008842. Available at: https://doi.org/10.1371/journal.ppat.1008842.

Wagner, R.W. and Nishikura, K. (1988) "Cell cycle expression of RNA duplex unwindase activity in mammalian cells," *Molecular and Cellular Biology*, 8(2), pp. 770–777. Available at: https://doi.org/10.1128/mcb.8.2.770-777.1988.

Wang, Q. *et al.* (2004) "Stress-induced Apoptosis Associated with Null Mutation of ADAR1 RNA Editing Deaminase Gene*," *Journal of Biological Chemistry*, 279(6), pp. 4952–4961. Available at: https://doi.org/10.1074/jbc.m310162200.

Wang, Y. *et al.* (2017) "Systematic characterization of A-to-I RNA editing hotspots in microRNAs across human cancers," *Genome Research*, 27(7), pp. 1112–1125. Available at: https://doi.org/10.1101/gr.219741.116.

Wang, Y. *et al.* (2019) "RNA binding candidates for human ADAR3 from substrates of a gain of function mutant expressed in neuronal cells," *Nucleic Acids Research*, 47(20), pp. 10801–10814. Available at: https://doi.org/10.1093/nar/gkz815.

Ward, S.V. *et al.* (2011) "RNA editing enzyme adenosine deaminase is a restriction factor for controlling measles virus replication that also is required for embryogenesis," *Proceedings of the National Academy of Sciences*, 108(1), pp. 331–336. Available at: https://doi.org/10.1073/pnas.1017241108.

Workman, R.E. *et al.* (2019) "Nanopore native RNA sequencing of a human poly(A) transcriptome," *Nature Methods*, 16(12), pp. 1297–1305. Available at: https://doi.org/10.1038/s41592-019-0617-2.

Wu, J. *et al.* (2019) "Evolution of Inosine-Specific Endonuclease V from Bacterial DNase to Eukaryotic RNase," *Molecular Cell*, 76(1), pp. 44-56.e3. Available at: https://doi.org/10.1016/j.molcel.2019.06.046.

Wu, Q., Amrutkar, S.M. and Shao, F. (2018) "Sulfinate Based Selective Labeling of 5-Hydroxymethylcytosine: Application to Biotin Pull Down Assay," *Bioconjugate Chemistry*, 29(2), pp. 245–249. Available at: https://doi.org/10.1021/acs.bioconjchem.7b00826.

Yang, C.-C. *et al.* (2017) "ADAR1-mediated 3′ UTR editing and expression control of antiapoptosis genes fine-tunes cellular apoptosis response," *Cell Death & Disease*, 8(5), pp. e2833–e2833. Available at: https://doi.org/10.1038/cddis.2017.12.

Yang, W. *et al.* (2006) "Modulation of microRNA processing and expression through RNA editing by ADAR deaminases," *Nature Structural & Molecular Biology*, 13(1), pp. 13–21. Available at: https://doi.org/10.1038/nsmb1041.

Yang, Y., Okada, S. and Sakurai, M. (2021) "Adenosine-to-inosine RNA editing in neurological development and disease," *RNA Biology*, 18(7), pp. 1–15. Available at: https://doi.org/10.1080/15476286.2020.1867797.

Yin, X. *et al.* (2021) "MDA5 Governs the Innate Immune Response to SARS-CoV-2 in Lung Epithelial Cells," *Cell Reports*, 34(2), p. 108628. Available at: https://doi.org/10.1016/j.celrep.2020.108628.

Yoshida, M. and Ukita, T. (1968) "Modification of nucleosides and nucleotides VII. Selective cyanoethylation of inosine and pseudouridine in yeast transfer ribonucleic acid," *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis*, 157(3), pp. 455–465. Available at: https://doi.org/10.1016/0005-2787(68)90145-7.

Zhang, L. *et al.* (2016) "Altered RNA editing in 3′ UTR perturbs microRNA-mediated regulation of oncogenes and tumor-suppressors," *Scientific Reports*, 6(1), p. 23226. Available at: https://doi.org/10.1038/srep23226.

Zhang, L. *et al.* (2021) "Functional Differences Between Two Kv1.1 RNA Editing Isoforms: a Comparative Study on Neuronal Overexpression in Mouse Prefrontal Cortex," *Molecular Neurobiology*, 58(5), pp. 2046–2060. Available at: https://doi.org/10.1007/s12035-020-02229-1.

Zheng, X. *et al.* (2022) "Development of an Endonuclease V-Assisted Analytical Method for Sequencing Analysis of Deoxyinosine in DNA," *Analytical Chemistry*, 94(33), pp. 11627–11632. Available at: https://doi.org/10.1021/acs.analchem.2c02126.

Zheng, Y., Lorenzo, C. and Beal, P.A. (2017) “DNA editing in DNA/RNA hybrids by adenosine deaminases that act on RNA,” *Nucleic Acids Research*, 45(6), pp. 3369–3377. Available at: https://doi.org/10.1093/nar/gkx050.

Table. RNA and DNA Oligonucleotides used for this study

Names	Sequences
RNA-6A	5′-GGCGAGAGGCAAGAGGCGCGCAGUAGGGCGGCAGAAGCGGCGUAGCGGGCCGCGCGUCGGGC-3′
RNA-6I	5′-GGCGAGAGGCAIGAGGCGCGCIGUIGGGCGGCIGAIGCGGCGUIGCGGGCCGCGCGUCGGGC-3′
DNA-6dA	5′-GGCGAGAGGCAAGAGGCGCGCAGTAGGGCGGCAGAAGCGGCGTAGCGGGCCGCGCGTCGGGC-3′
DNA-6dI	5′-GGCGAGAGGCAIGAGGCGCGCIGTIGGGCGGCIGAIGCGGCGTIGCGGGCCGCGCGTCGGGC-3′
DNA-dA	5′-GACACACAAGCGACACAACGAG-3′
DNA-dI	5′-GACACACAAGCGICACAACGAG-3′
Cy5-DNA-dA	5′-[Cy5] GACACACAAGCGACACAACGAG-3′
Cy5-DNA-dI	5′-[Cy5] GACACACAAGCGICACAACGAG-3′
Cy5-DNA-dG	5′-[Cy5] GACACACAAGCGGCACAACGAG-3′
FITC-DNA-dA	5′-[FITC] GACACACAAGCGACACAACGAG-3′
FITC-DNA-dI	5′-[FITC] GACACACAAGCGICACAACGAG-3′
Cy5-DNA-dA-	5′-[Cy5] GACACACAAGCGACACAACGAG [FITC]-3′
Cy5-DNA-dI-	5′-[Cy5] GACACACAAGCGICACAACGAG[FITC]-3′
Cy5-IA-I	5′-[Cy5] GCTCGTIGTGCTC-3′
Cy5-IA-A	5′-[Cy5] GCTCGTAGTGCTC-3′
Cy5-IG-I	5′-[Cy5] ACTACTIACTACT-3′
Cy5-IG-G	5′-[Cy5] ACTACTGACTACT-3′
Cy5-IU-I	5′-[Cy5] GCACGAIGACGAC-3′
Cy5-IU-U	5′-[Cy5] GCACGAUGACGAC-3′
Cy5-IT-I	5′-[Cy5] GCACGAIGACGAC-3′
Cy5-IT-T	5′-[Cy5] GCACGATGACGAC-3′
Cy5-IC-I	5′-[Cy5] AGATGTIGTAGAT-3′
Cy5-IC-C	5′-[Cy5] AGATGTCGTAGAT-3′
Cy5-36nt-dI	5′-[Cy5] GACAGACAGCCICACAACAAGAGACCAGACACAGAG-3′
FAM-20nt-dA	5′-[FAM] GACAGACAGCCACACGACAG-3′
FAM-30nt-dU	5′-[FAM] GACAGACAGCCUCACAACAAGAGAAGAGAG-3′
Cy5-24nt-dI-Bio	5′-[Cy5] GACAGACAGCCICACAACAGACAG [Bio]-3′
RNA-1I	5′-UAUGCIGCUGCGCCCAGUUAUUUCGCCAAGGUUC-3′
RNA-I-Cy5	5′-[Cy5] GACACAUCCGCICAGCAACGAG-3′
RNA-A-FAM	5′-[FAM] GACACAUCCGCACAGCAACGAG--3′
Primers for primer extension	
Cy5-PE-DNA	5′-[Cy5] GAACCTTGGCGAAATAACTG-3′
Mouse Gria2	5′-GATCTTGGCGAAATATCGCATC-3′

Figures

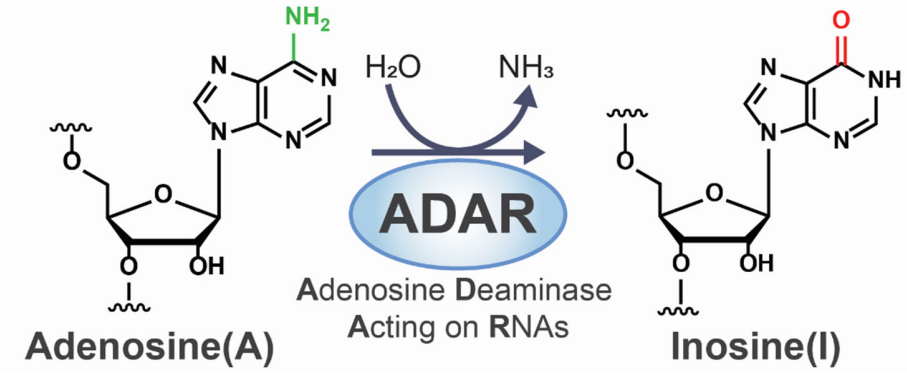


Figure 1. Hydrolytic deamination reaction catalyzed by adenosine deaminase acting on RNAs (ADAR). (a) ADARs catalyze adenosine to inosine (A-to-I) editing.

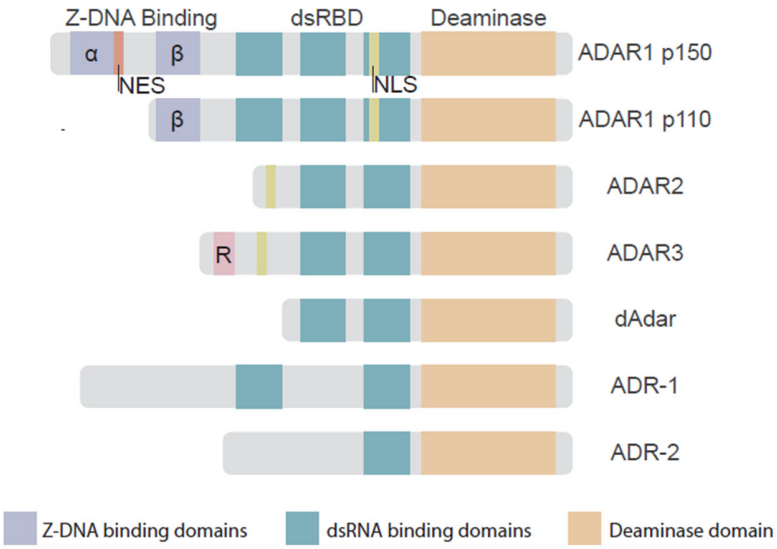


Figure 2. Domain architecture of three human ADAR family members (ADAR1, ADAR2, and ADAR3), two ADAR1 isoforms (p150 and p110), Drosophila melanogaster dAdar, and two Caenorhabditis elegans ADAR (ADR-1 and ADR-2). The domain structures shown are not strictly proportional. dsRBD, double-stranded RNA binding domains; R, Arg-rich single-stranded RNA binding domain; NES, nuclear export signal; NLS, nuclear localization signal.

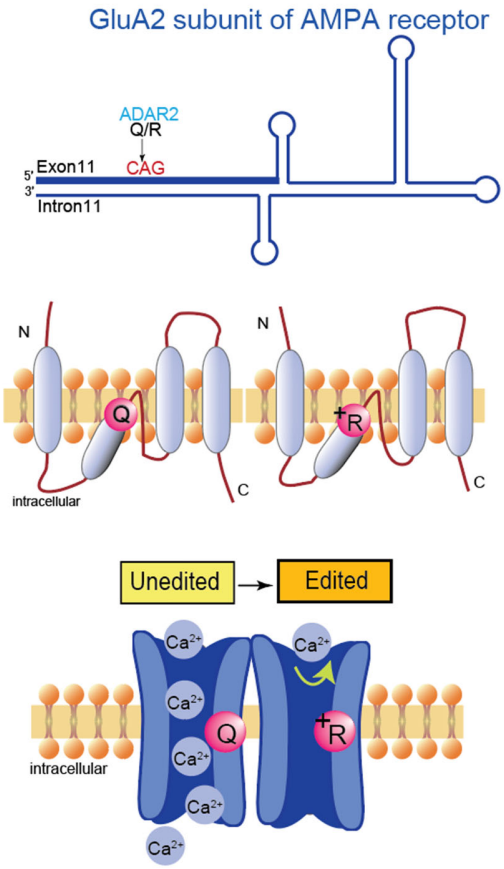


Figure 3. Top: The A-to-I editing site (Q/R) in the secondary structure of GluA2 pre-mRNA. Middle and Bottom: Schematic of the GluA2 subunit structure, contrasting edited and unedited forms. The RNA editing transition from a glutamate residue (Q) (middle left) to a positively charged arginine residue (R) (middle right), leading to reduced calcium permeability (bottom).

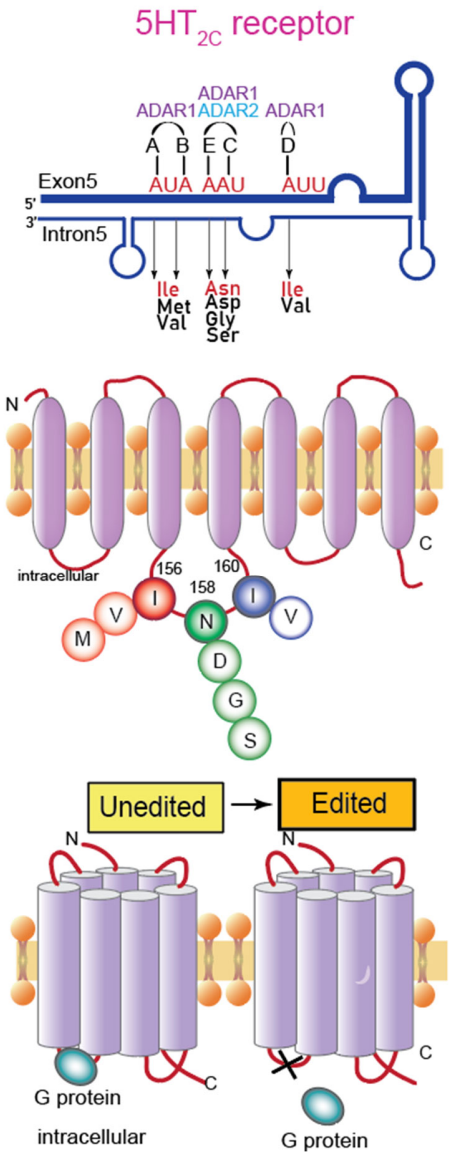


Figure 4. Top: The five editing sites (A-E) within the $5HT_{2C}R$ pre-mRNA's secondary structure, where exon 5 pairs with intron 5. Middle: The schematic of the 5-HT_{2C} receptor structure showing the differences between edited and unedited forms. RNA editing modifications in the $5HT_{2C}$ receptor, changing isoleucine (I) at position 156 to valine (V), or methionine (M) at position 158 to asparagine (N), aspartic acid (D), serine (S), or glycine (G), and isoleucine (I) at position 160 to valine (V). The edited $5HT_{2C}R$ isoforms demonstrate reduced G protein coupling efficiency (bottom left) compared to the unedited isoform (bottom right).

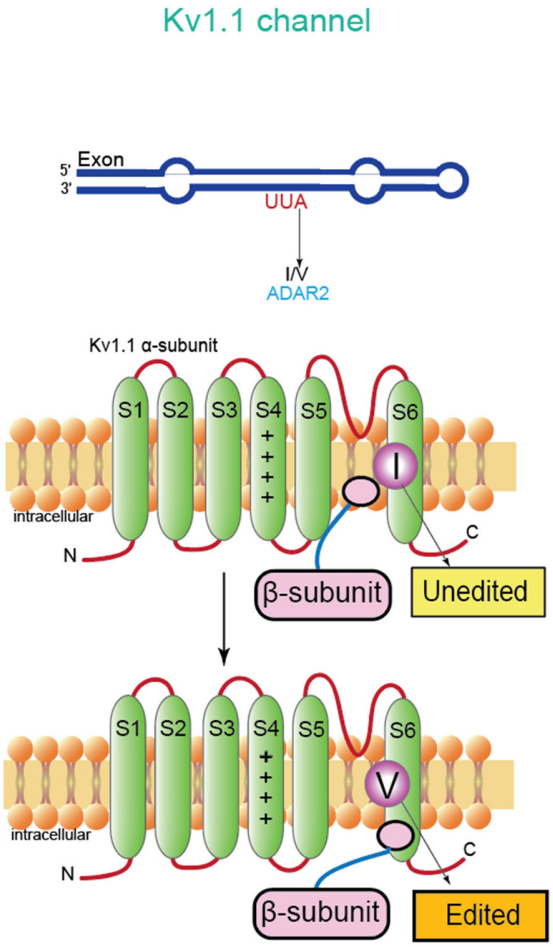


Figure 5. Top: The Kv1.1 mRNA's secondary structure contains the A-to-I editing site (I/V). Middle and Bottom: The structural schematic of the Kv1.1 channel delineates the functional variances in its edited and unedited states. The figures specifically point out the transmembrane voltage sensor (S4) and the K+ ion selectivity filter (S5–S6). The RNA editing in Kv1.1 transforms the isoleucine (I) (seen in the middle) into valine (V) (indicated at the bottom), diminishing Kv β1.1's binding affinity and improving the channel's recovery from inactivation.

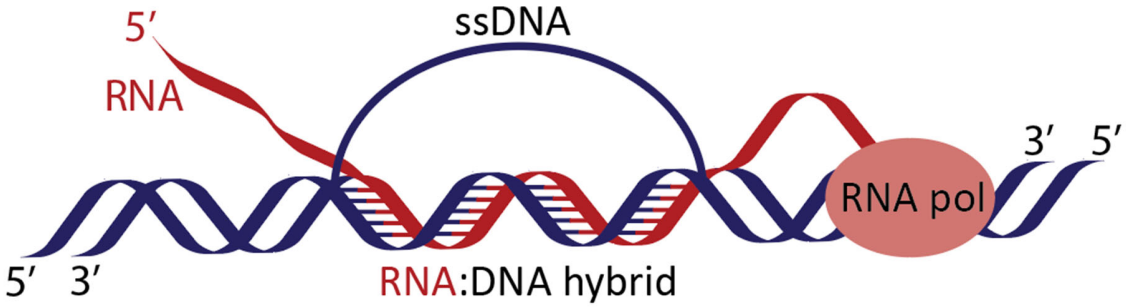


Figure 6. Schematic representation of an R-loop structure. In this configuration, the nascent RNA strand, depicted in red, is synthesized by the RNA polymerase (RNA pol) and hybridizes with the complementary DNA template strand. The non-template strand is exposed as single-stranded DNA (ssDNA).

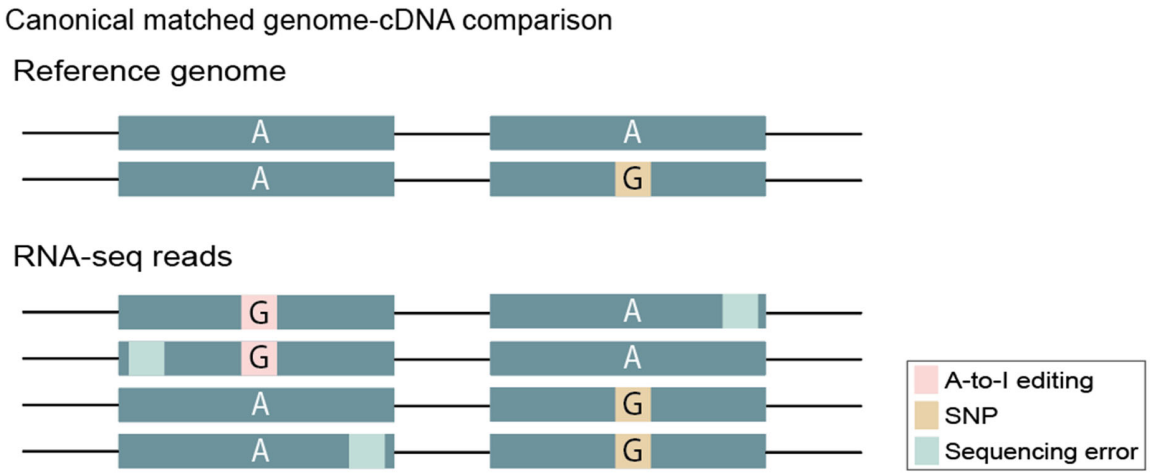


Figure 7. The canonical method detects editing sites by comparing nucleotide differences between DNA and cDNA sequences. Guanosine is present at the site of inosine in RNA after reverse transcription, and the equivalent site in the genomic DNA sequence is adenosine. Thus, A-to-I RNA editing sites can be detected as adenosine-guanosine (A to G) mismatches. In practice, there are many other sources of DNA-RNA sequence mismatch, including sequence polymorphism and sequencing errors, which make it challenging to identify the right editing site. SNP, single nucleotide polymorphism.

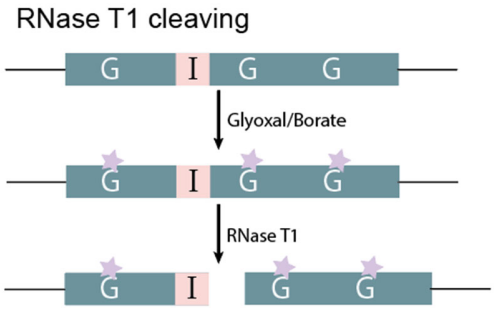


Figure 8. The cleavage of inosine by RNase T1. Stars indicate the protective effect of glyoxal/borate on guanosines.

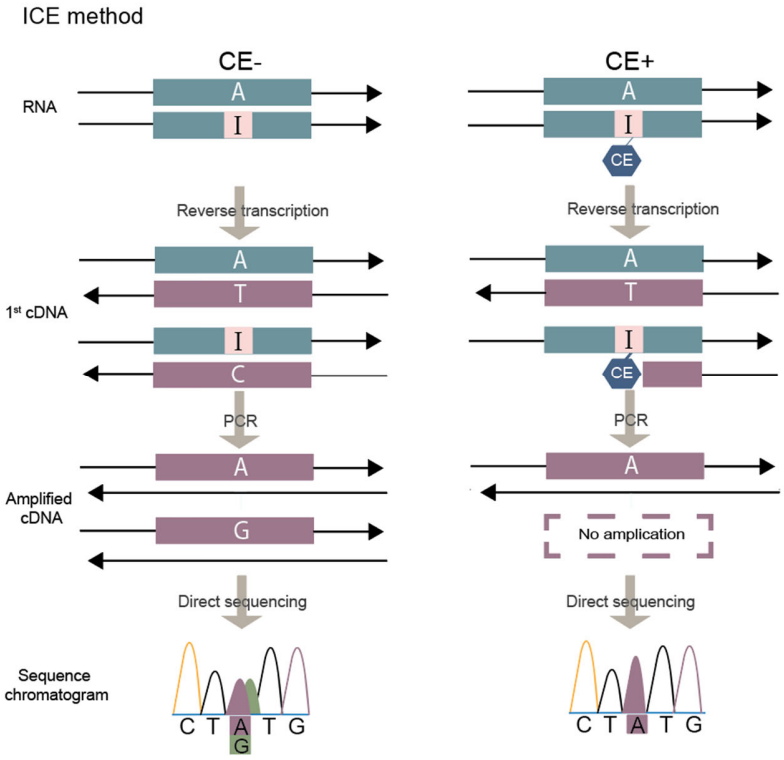


Figure 9. Schemes of ICE method. The non-cyanoethylated (CE-) RNA is on the left, and the cyanoethylated (CE+) RNA is on the right. RNA is shown in blue, cDNA in yellow. RNA reverse transcription produces first-strand cDNA. In CE- condition, RNA with A is converted into T and I to C. In the CE+ condition, the first strand extension of cDNA is blocked by CE at site I. After PCR amplification, the cDNA with editing sites T and C were amplified under CE- conditions. In the CE+ condition, the editing site is blocked by CE and cannot be amplified. After direct sequencing, mixed signals of A and G can be detected at the editing sites in the sequence chromatography under CE- conditions. In the CE+ condition, the G signal generated by A-to-I editing disappears and is replaced by A.

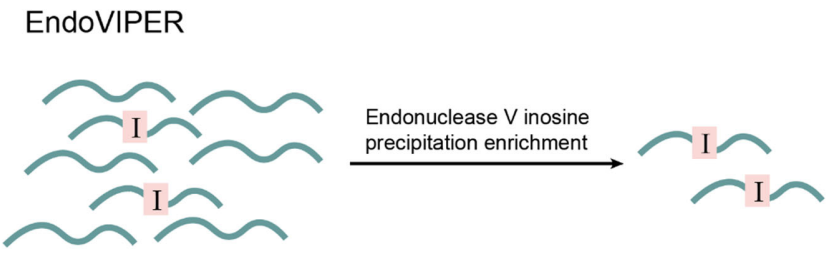


Figure 10. Enrichment of inosine from cellular RNA using Endonuclease V conjugated with Ca^{2+}.

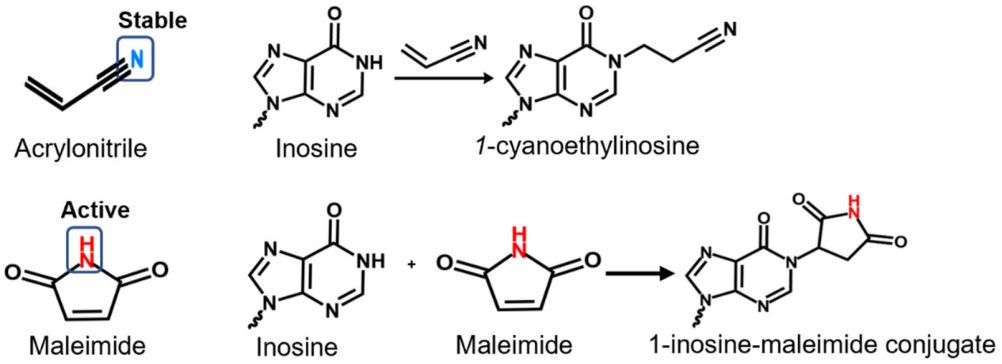


Figure 11. The structure of acrylonitrile and maleimide. And their chemical reaction with inosine.

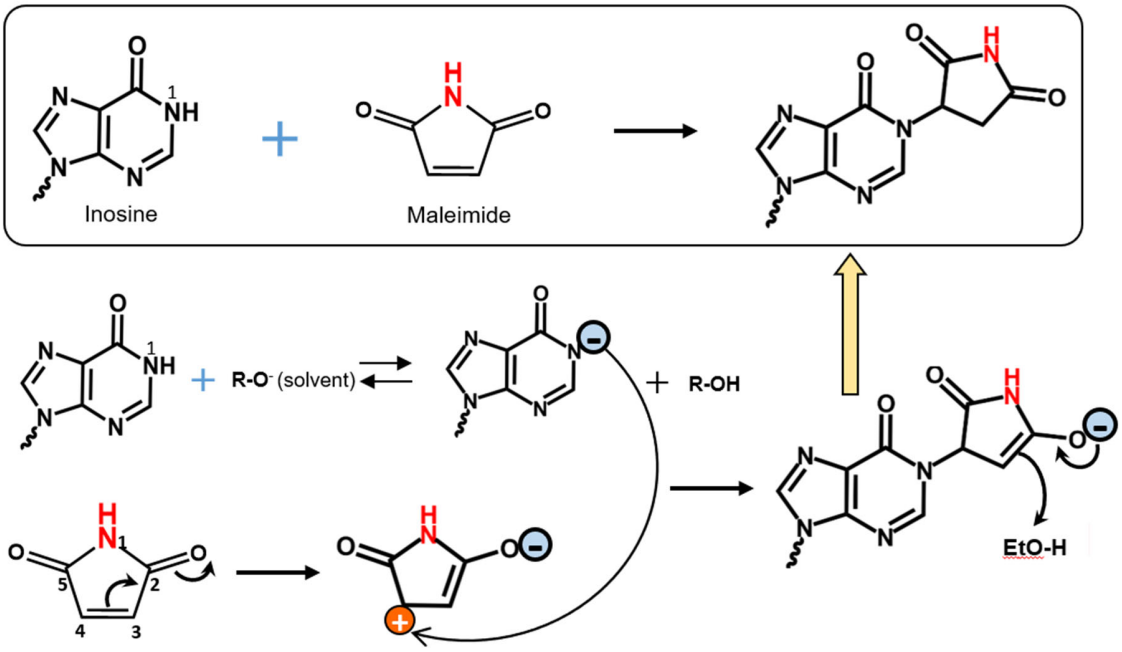


Figure 12. Mechanism of inosine-specific chemical modification by maleimide. In weakly alkaline solvents, the inosine base exhibits proton dissociation at the N1 position, resulting in an equilibrium between dissociated and non-dissociated forms, which imparts a negative charge to the N1 nitrogen. Maleimide's oxygen atom exerts an electron-withdrawing effect, generating a positive charge primarily on its 4th carbon. Given maleimide's symmetric structure, a positive charge can also manifest on the 3rd carbon. This charge interplay facilitates a Michael addition, leading to the labeling of the inosine base by maleimide.

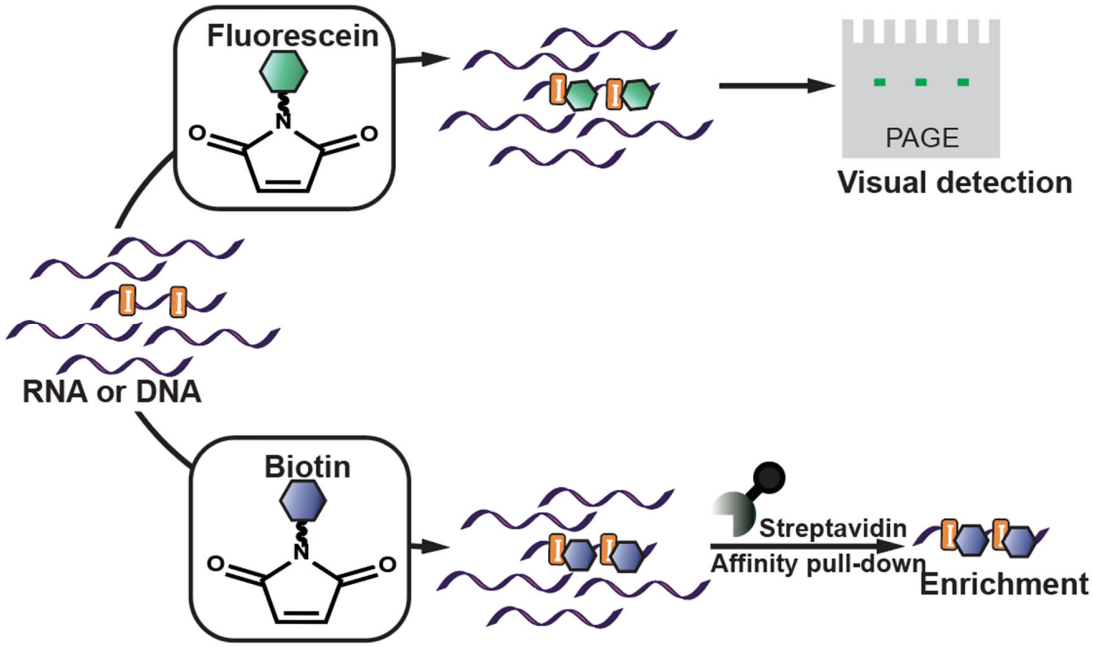


Figure 13. Inosine Chemical Labeling and Affinity Molecular Purification (ICLAMP). Fluorescein–maleimide added fluorescent labels to inosine-containing nucleic acids, allowing for visual detection. Biotin–maleimide enabled the chemical labeling of inosine and affinity purification of inosine-containing nucleic acids.

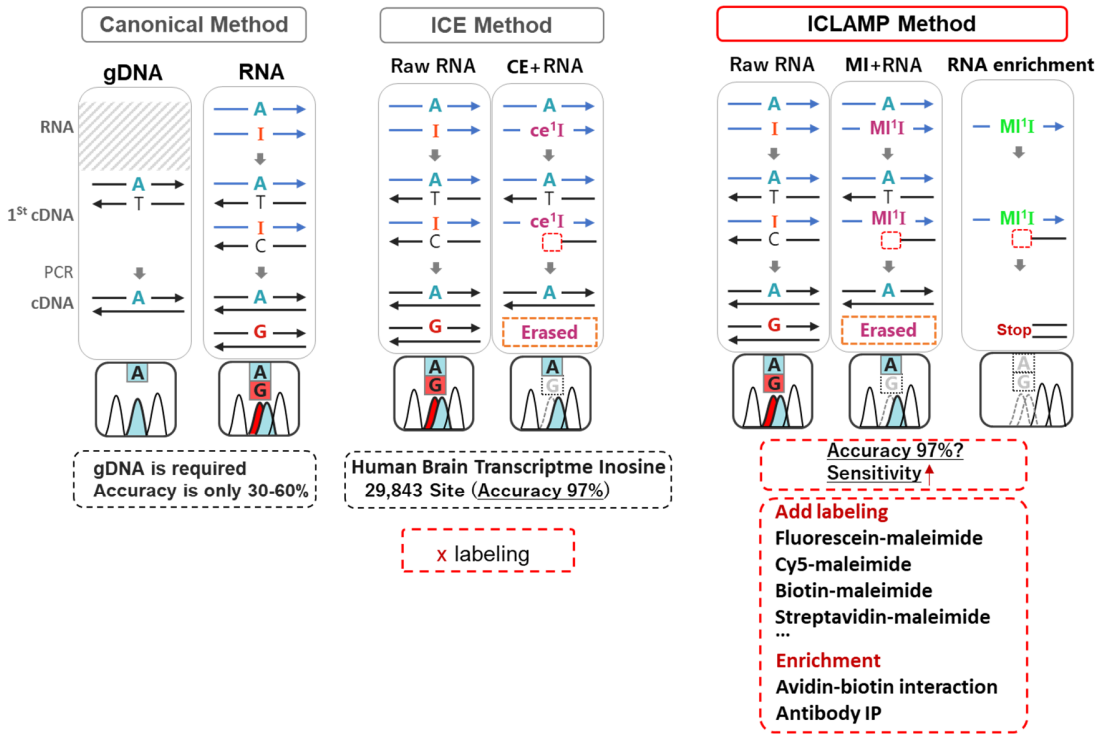


Figure 14. Comparative advantages and innovations of ICLAMP method.

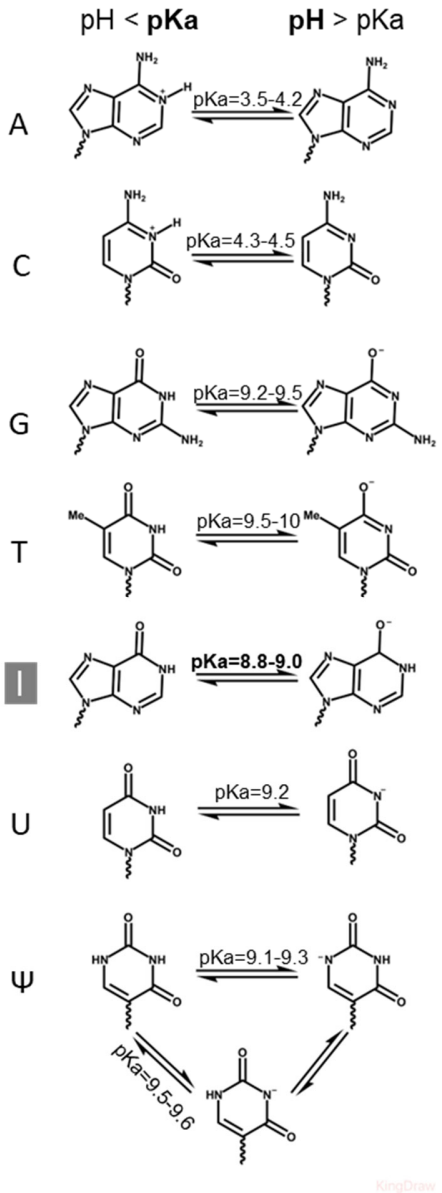


Figure 15. Comparison of acid dissociation constant pKa values for each base.

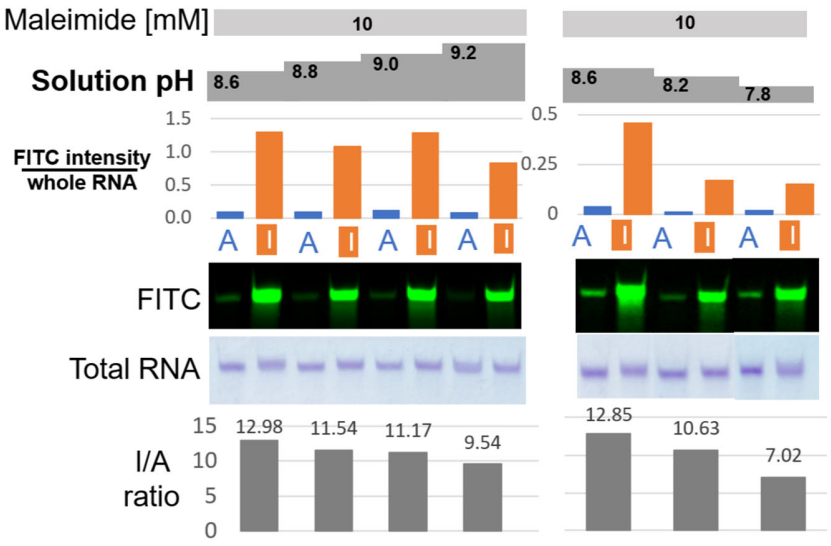


Figure 16. pH optimization of inosine-specific addition reaction conditions using fluorescent labels.

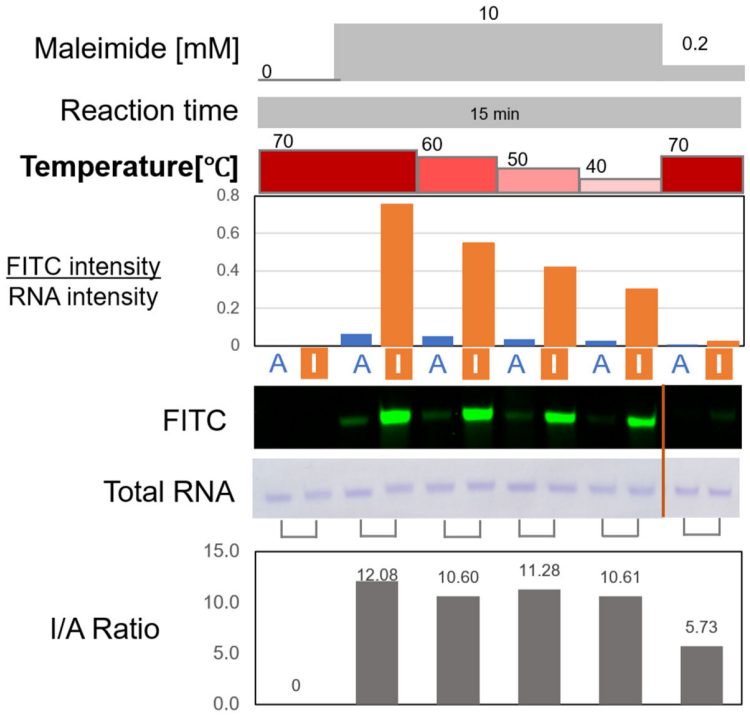
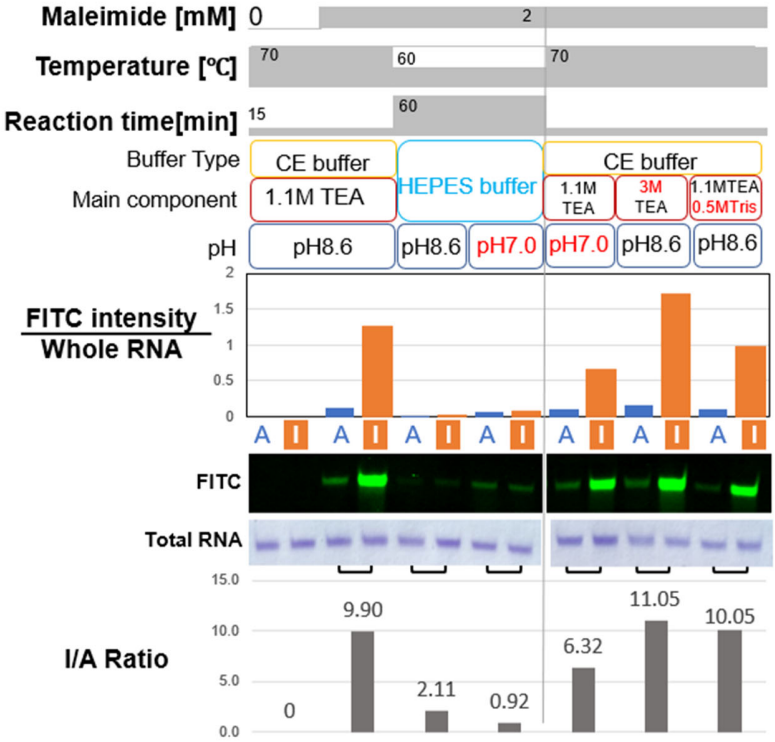


Figure 17. Temperature optimization of inosine-specific addition reaction conditions using fluorescent labels.



CE Buffer variant (TEA/EtOH base)

[Default]

1.1 M TEA-AcOH (pH8.6) 50% EtOH/water

1.1 M TEA-AcOH (pH7.0) 50% EtOH/water

3.0 M TEA-AcOH (pH8.6) 50% EtOH/water

1.1M TEA-AcOH & 0.5M Tris-AcOH(pH8.6) 50% EtOH/water

TEA= Triethylamine
Amine group may affect reaction specificity/efficiency

HEPES Buffer (Water base)

10MM HEPES-KOH 1mM EDTA DMSO10% pH8.6

10MM HEPES-KOH 1mM EDTA DMSO10% pH7.0

Figure 18. Optimization of RNA inosine-specific fluorescent label conditions.

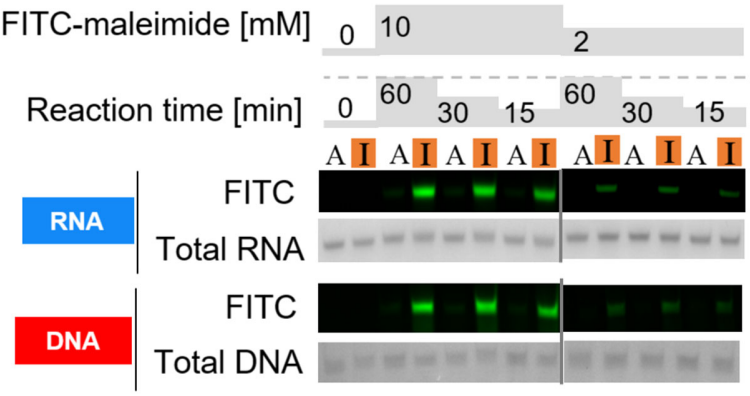


Figure 19. Comparison of the reactivities of fluorescein-maleimide with DNA and RNA oligonucleotides.

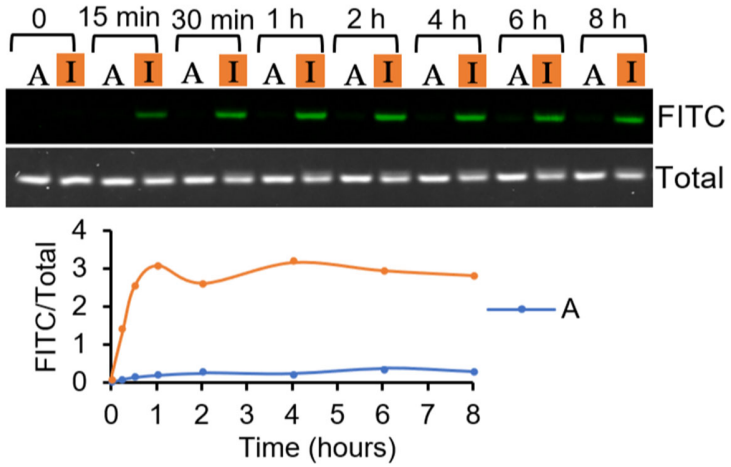


Figure 20. Reaction of DNA with fluorescein-maleimide over 8 h.

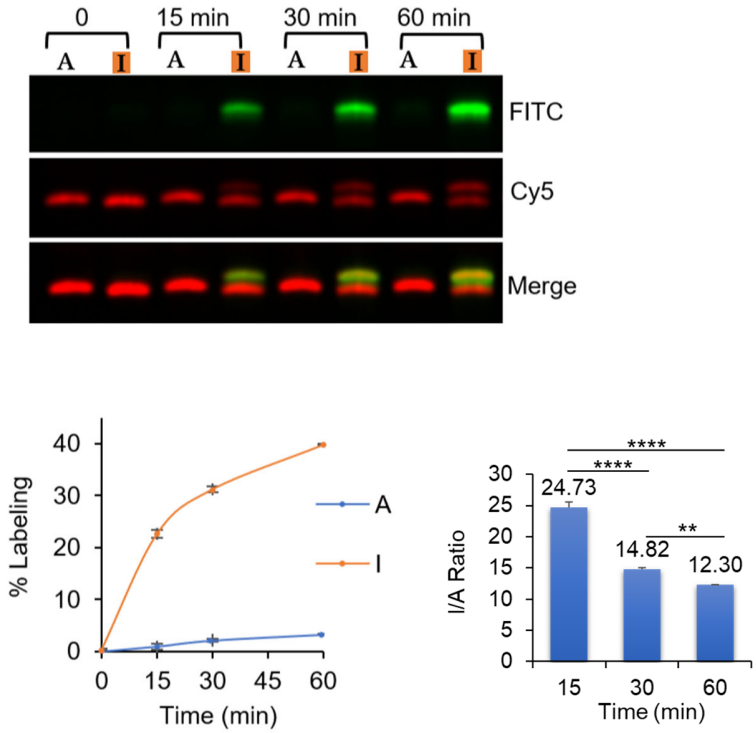


Figure 21. Reaction efficiency of DNA with 10 mM fluorescein-maleimide within 1 h. Densitometric quantification of oligonucleotide labeling and calculation of the I/A ratios. Values represent means ± standard deviation (SD) of results obtained from three independent experiments (n=3). Statistical significance was determined using one-way ANOVA followed by Tukey's multiple comparison test (** denotes $p < 0.01$, **** denotes $p < 0.0001$).

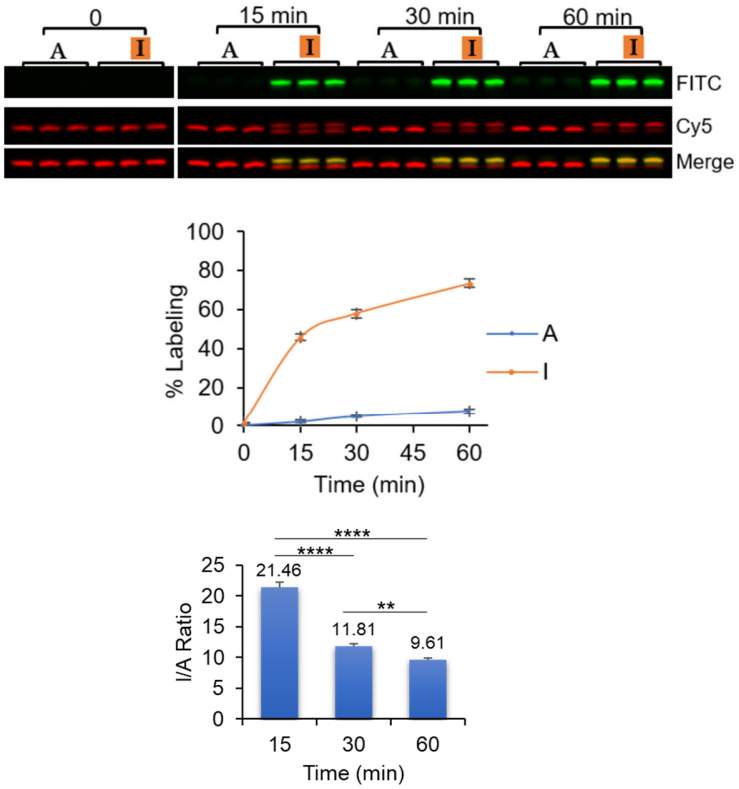


Figure 22. Reaction efficiency of DNA with 50 mM fluorescein-maleimide within 1 h. Densitometric quantification of oligonucleotide labeling and calculation of I/A ratios. Values represent means ± SD (n=3). Statistical significance was determined using one-way ANOVA followed by Tukey's multiple comparison test (** denotes $p < 0.01$, **** denotes $p < 0.0001$).

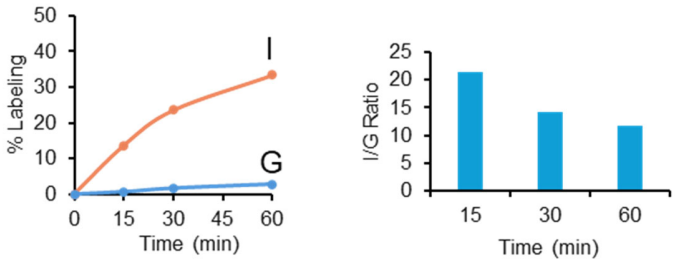


Figure 23. Reaction efficiency of DNA with 10 mM fluorescein-maleimide within 1 h and related I/G ratio.

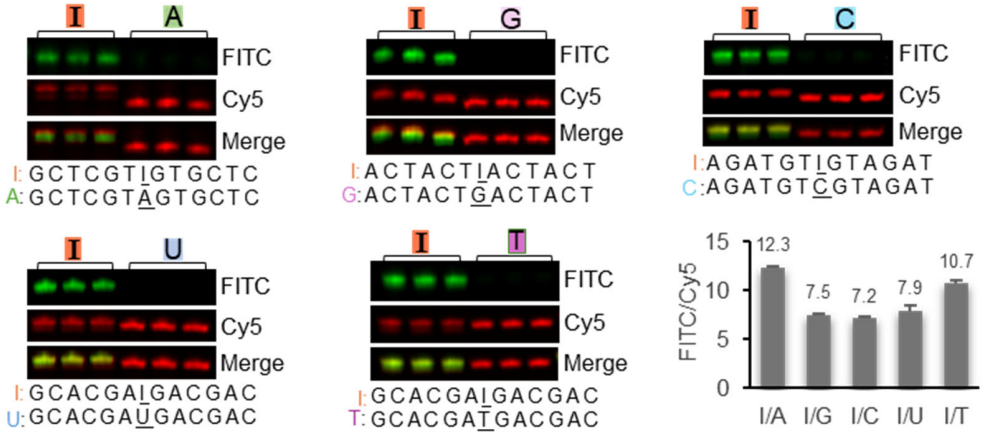


Figure 24. PAGE analysis of nonspecific reactions of maleimide with other bases and densitometric quantification. Data represents means ± SD (n=3).

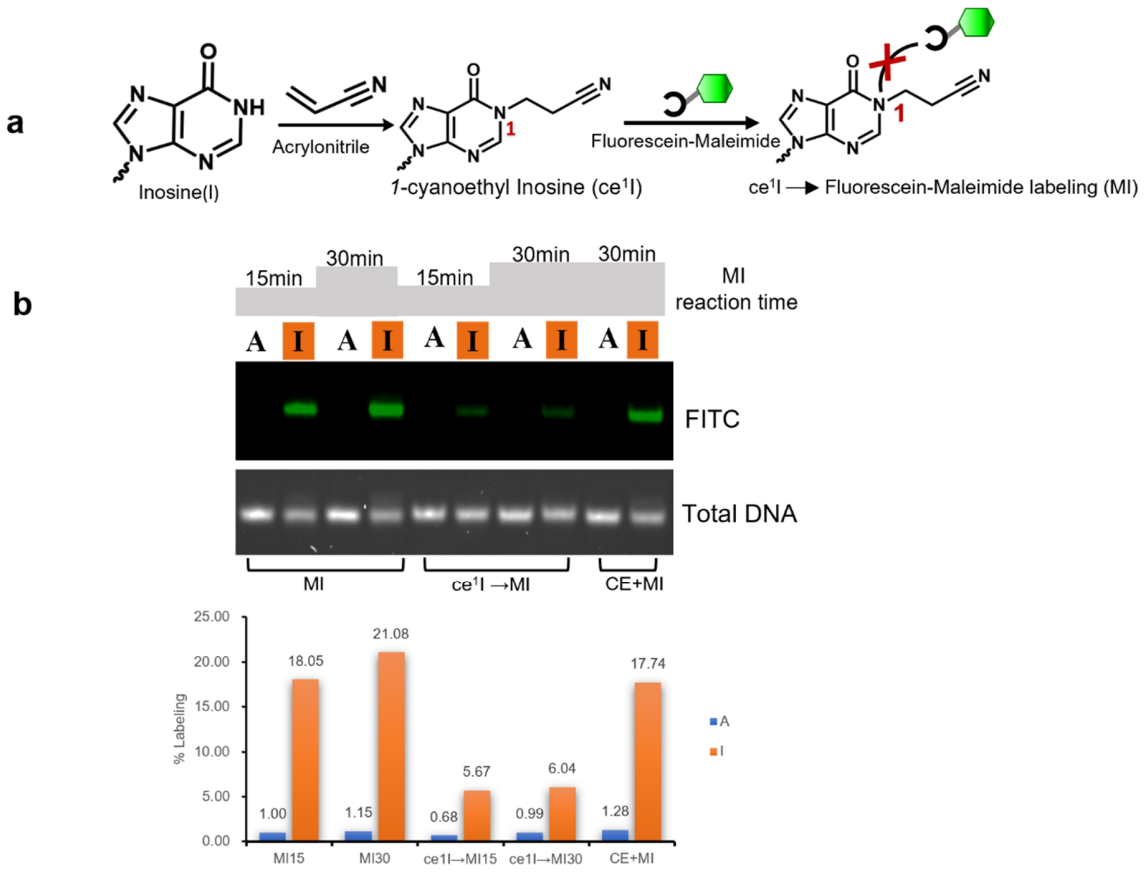


Figure 25. Confirmation of the maleimide–inosine reaction site. (a) Workflow based on the reactivity of acrylonitrile with inosine. Inosine first reacted with acrylonitrile, after which the reaction products were incubated with 10 mM fluorescein-maleimide for fluorescence intensity detection. (b) Reaction products were analyzed by PAGE and quantified by density. Ce^{1}I → MI represents the initial reaction between inosine and acrylonitrile. The resulting product subsequently reacted with maleimide. CE+MI represents the simultaneous reaction of inosine with a mixture of acrylonitrile and maleimide.

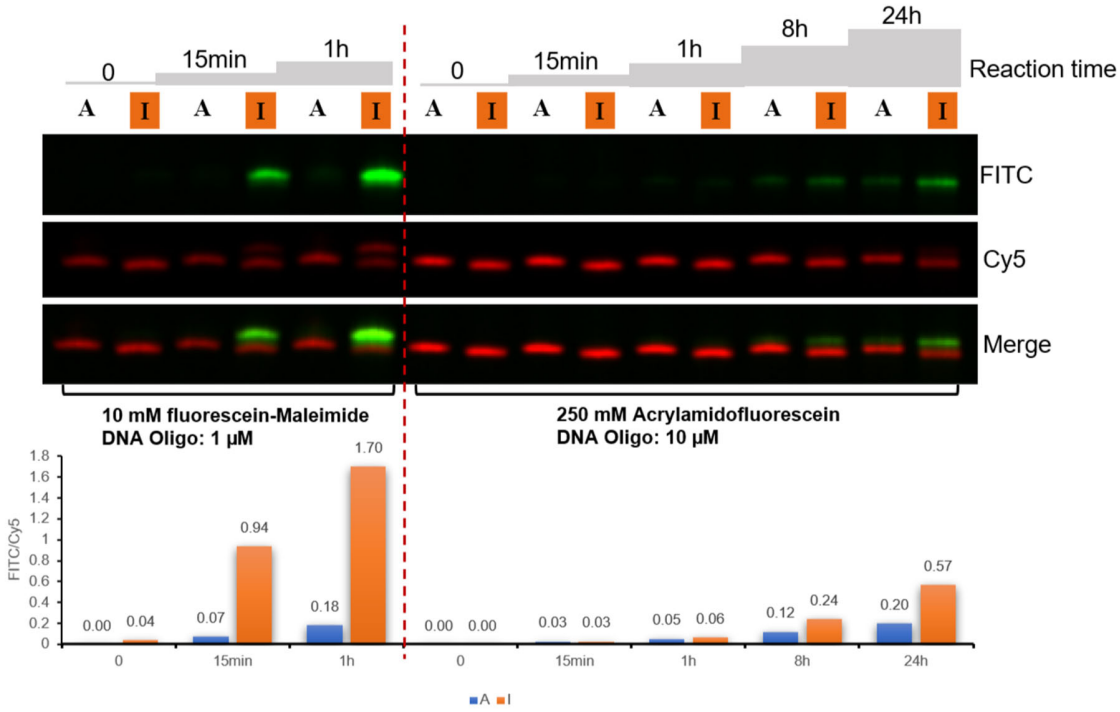


Figure 26. Comparison of the reactivities of fluorescein-maleimide and acrylamidofluorescein with inosine. PAGE analysis and densitometric quantification of the fluorescein-maleimide and acrylamidofluorescein reactivities; 1 μM DNA reacted with 10 mM fluorescein-maleimide for 1 h, while 250 mM acrylamidofluorescein reacted with 10 μM DNA for 24 h.

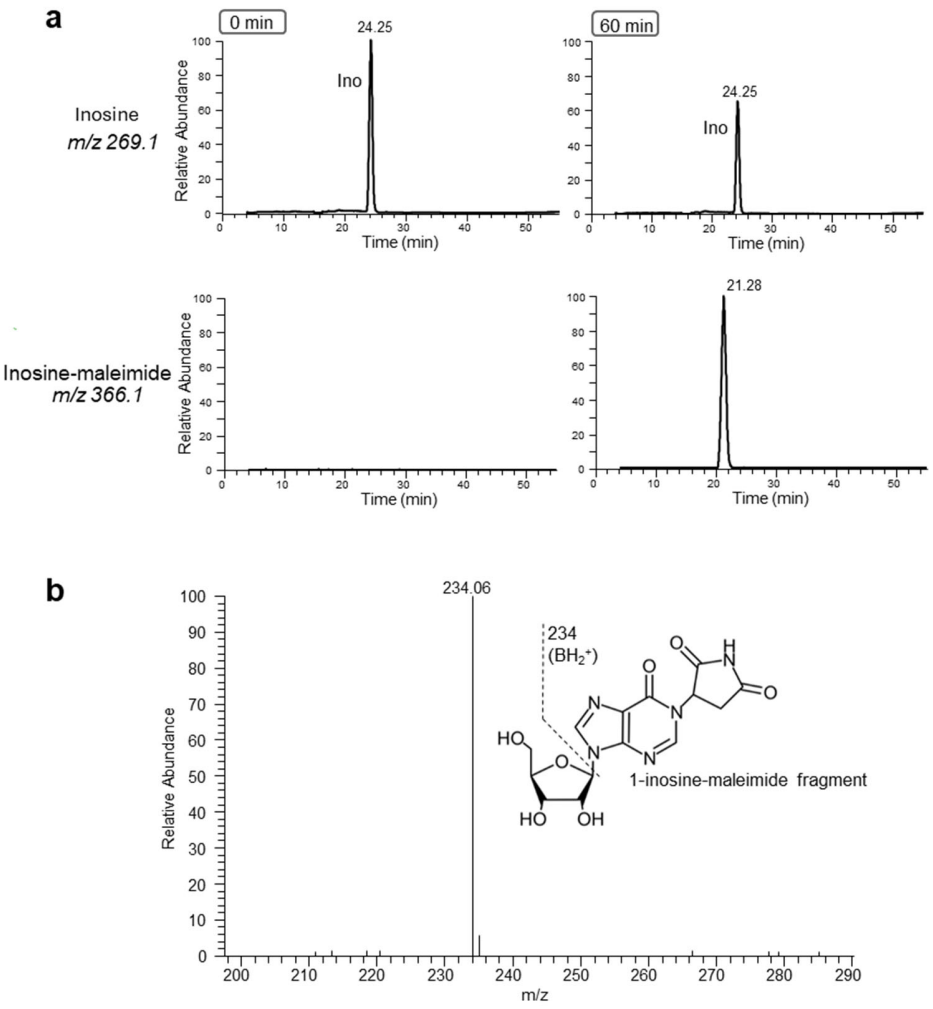


Figure 27. Extracted ion chromatogram of inosine and inosine-maleimide. (a) Reduced inosine peak intensity was normalized by adenosine in UV trace, and then quantified. (b) Higher-energy collisional dissociation spectra: base ion segmented by a dotted line was clearly detected.

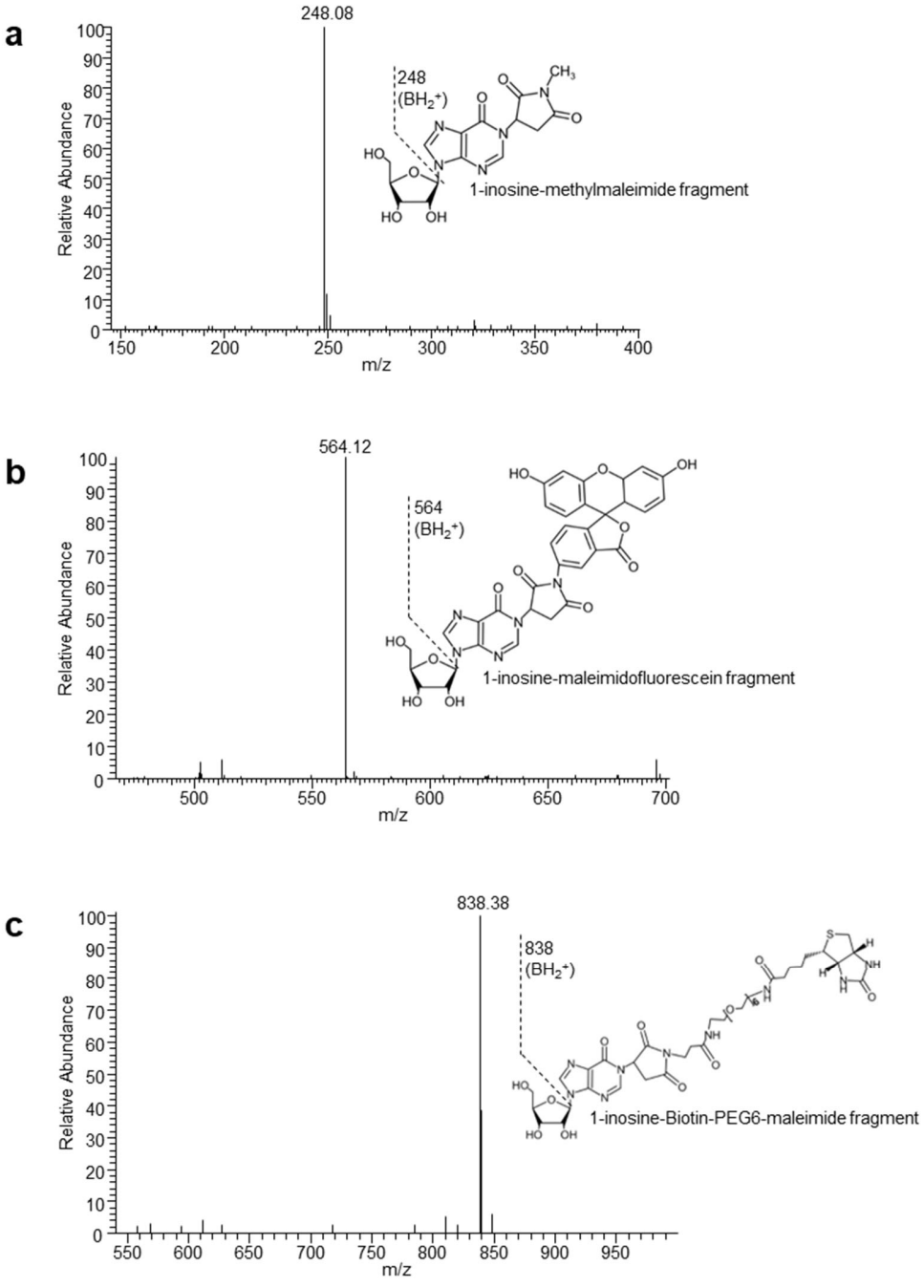


Figure 28. Mass spectrometric analyses of inosine and maleimide derivative reaction products. Higher-energy collisional dissociation mass spectra: Base ion segmented for (a) inosine-methylmaleimide, (b) inosine-fluorescein-maleimide, (c) inosine-biotin-PEG6-maleimide shown by dotted lines were detected.

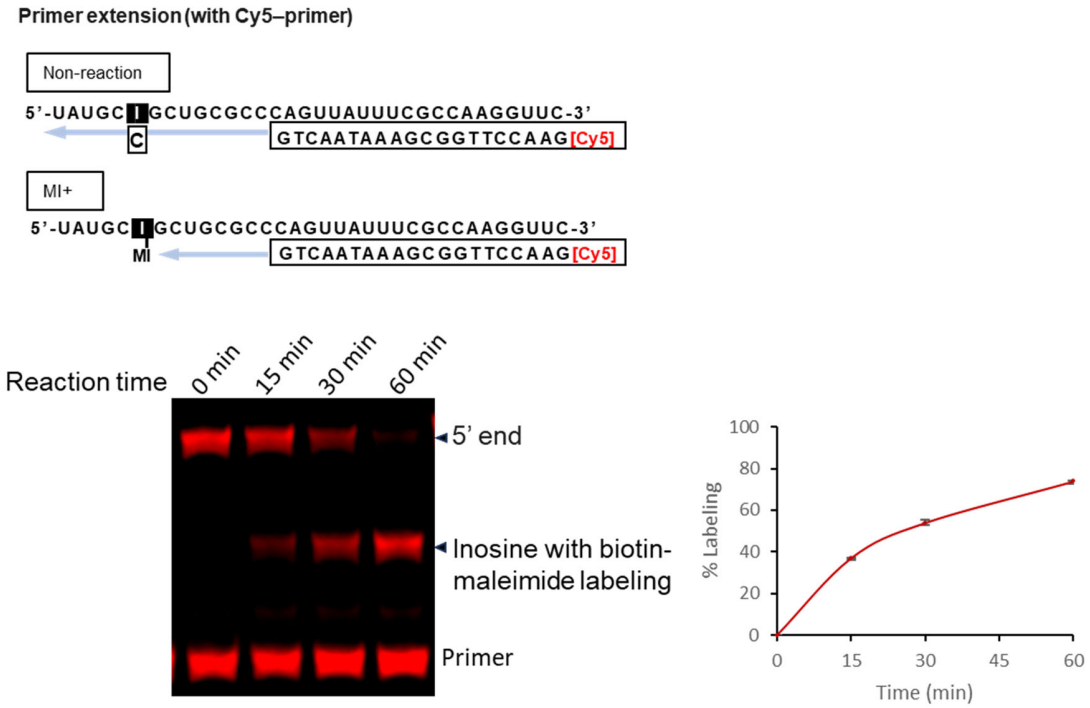


Figure 29. Reactivity assessment of inosine-specific biotin-maleimide labeling by primer extension. Values represent means ± standard deviation (SD) of results obtained from three independent experiments (n = 3).

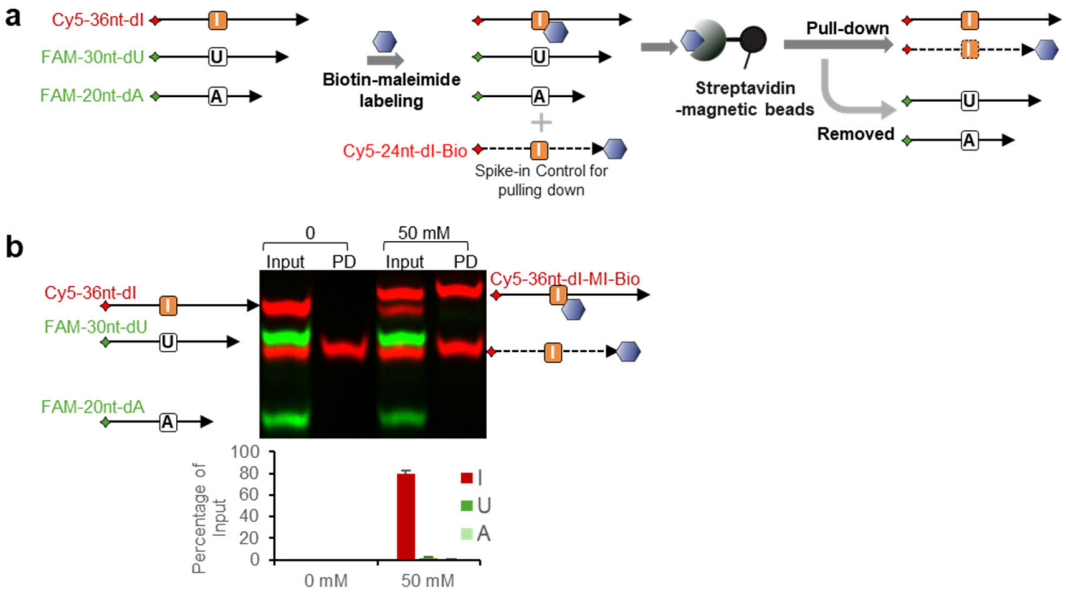


Figure 30. (a) Schematic of the pull-down workflow for DNA oligonucleotides. (b) PAGE analysis and densitometric quantification of the pulled down DNA. Biotin–maleimide labeling shifted the band distribution to a higher position in the gel.

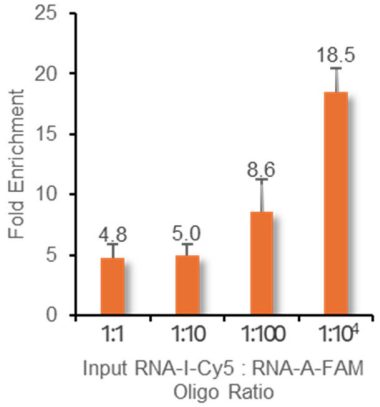


Figure 31. Fold enrichment of RNA oligonucleotides containing inosine isolated from a variable mixture. Values represent means ± SD (n=3).

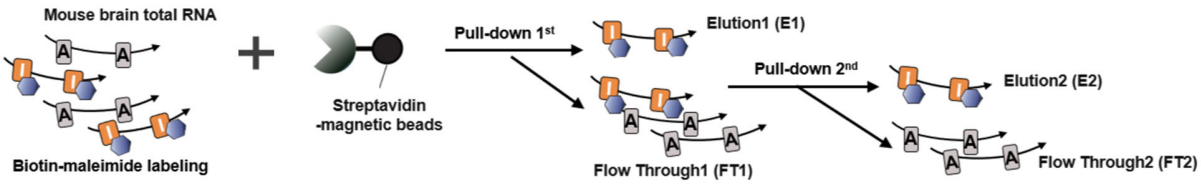


Figure 32. Schematic of the pull-down workflow.

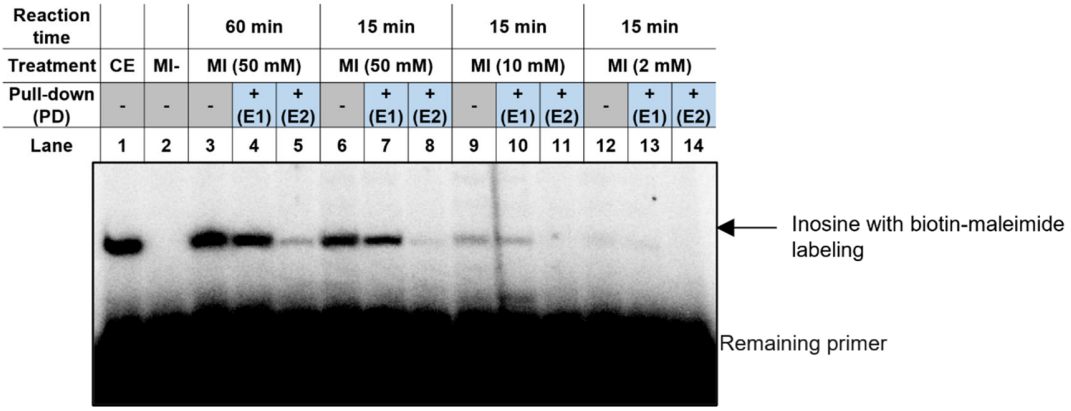


Figure 33. Detection of biotin–maleimide labeling and enrichment of inosine at the Q-to-R site of mouse *Gria2* mRNA using primer extension under gradient conditions.

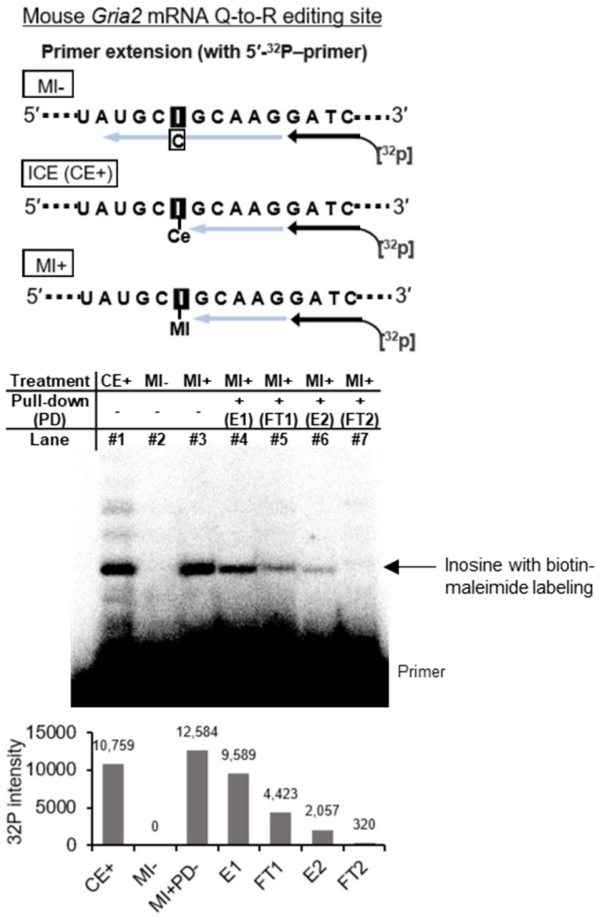


Figure 34. Primer extension to detect pulled down inosine at the Q-to-R site in mouse *Gria2* mRNA. The extension began a few bases downstream from the inosine site in mouse *Gria2* mRNA and was detected by autoradiography via a ^{32}P-labeled DNA primer at the 5′-end. Under MI- conditions, the primer (black arrow) underwent extension to synthesize a cDNA (blue arrow), crossing the inosine and terminating at the 5′ end of the RNA. Under CE+ and MI+ conditions, transcription terminated at the position immediately adjacent to the labeled inosine, as shown on the right.

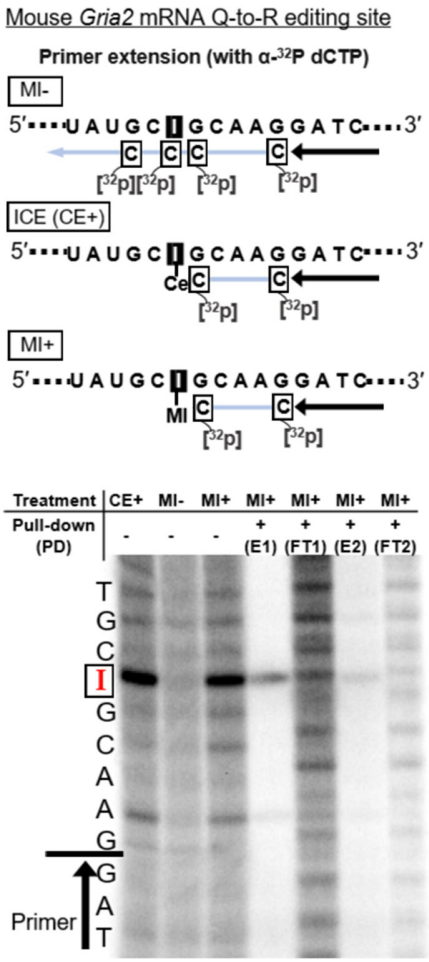


Figure 35. Primer extension to detect pulled down inosine at the Q-to-R site in mouse *Gria2* mRNA by inserting α-^{32}P dCTP. Under MI- conditions, the primer (black arrow) was extended to synthesize a cDNA spanning the inosines (blue arrow) and terminating at the 5′ end of the RNA. Under CE+ and MI+ conditions, transcription terminated at the position immediately adjacent to the labeled inosine, producing a cDNA containing two ^{32}P-labeled cytosines, as shown on the right.

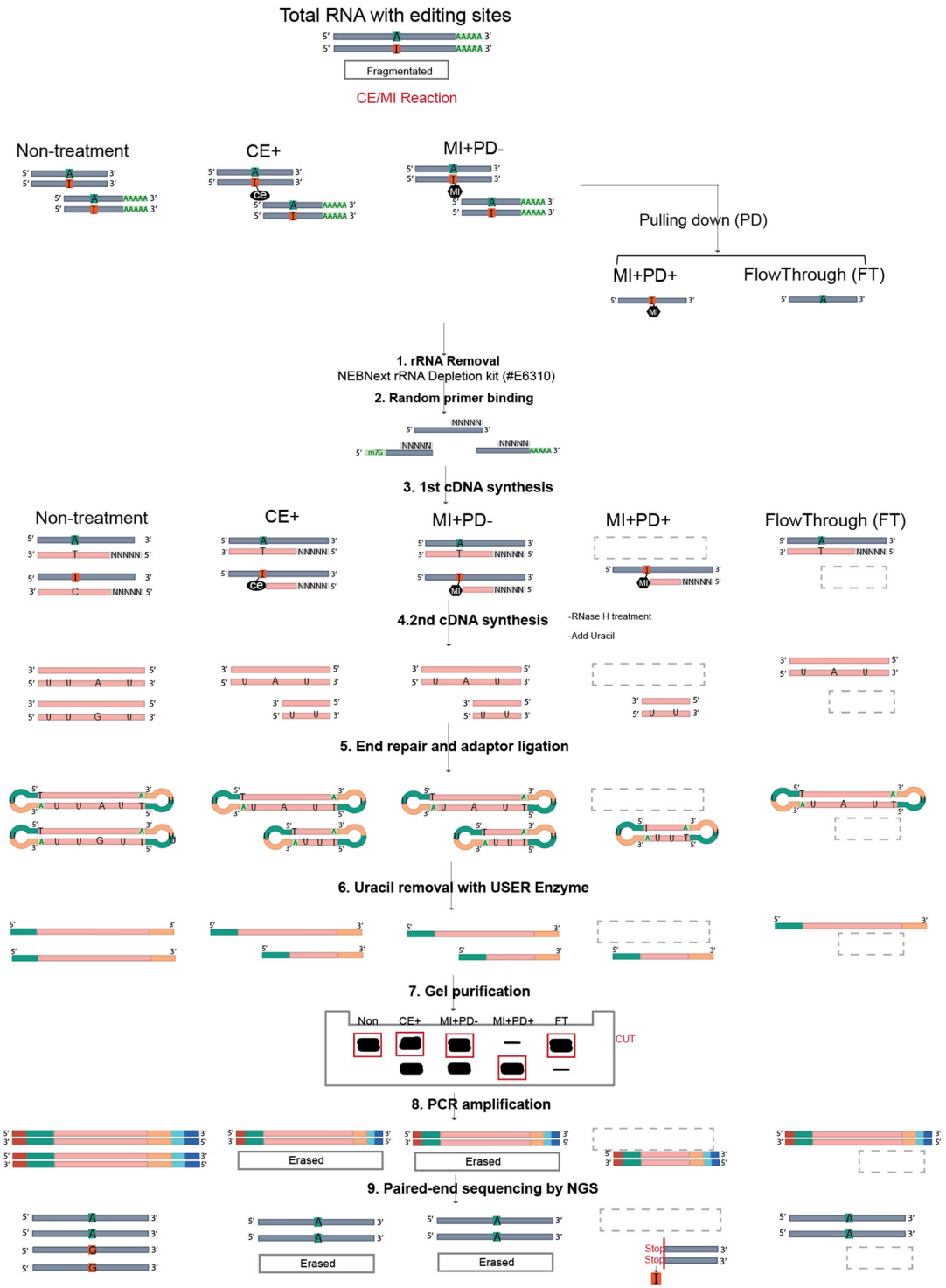


Figure 36. Flowchart for detecting new editing sites in RNA by ICLAMP-seq.

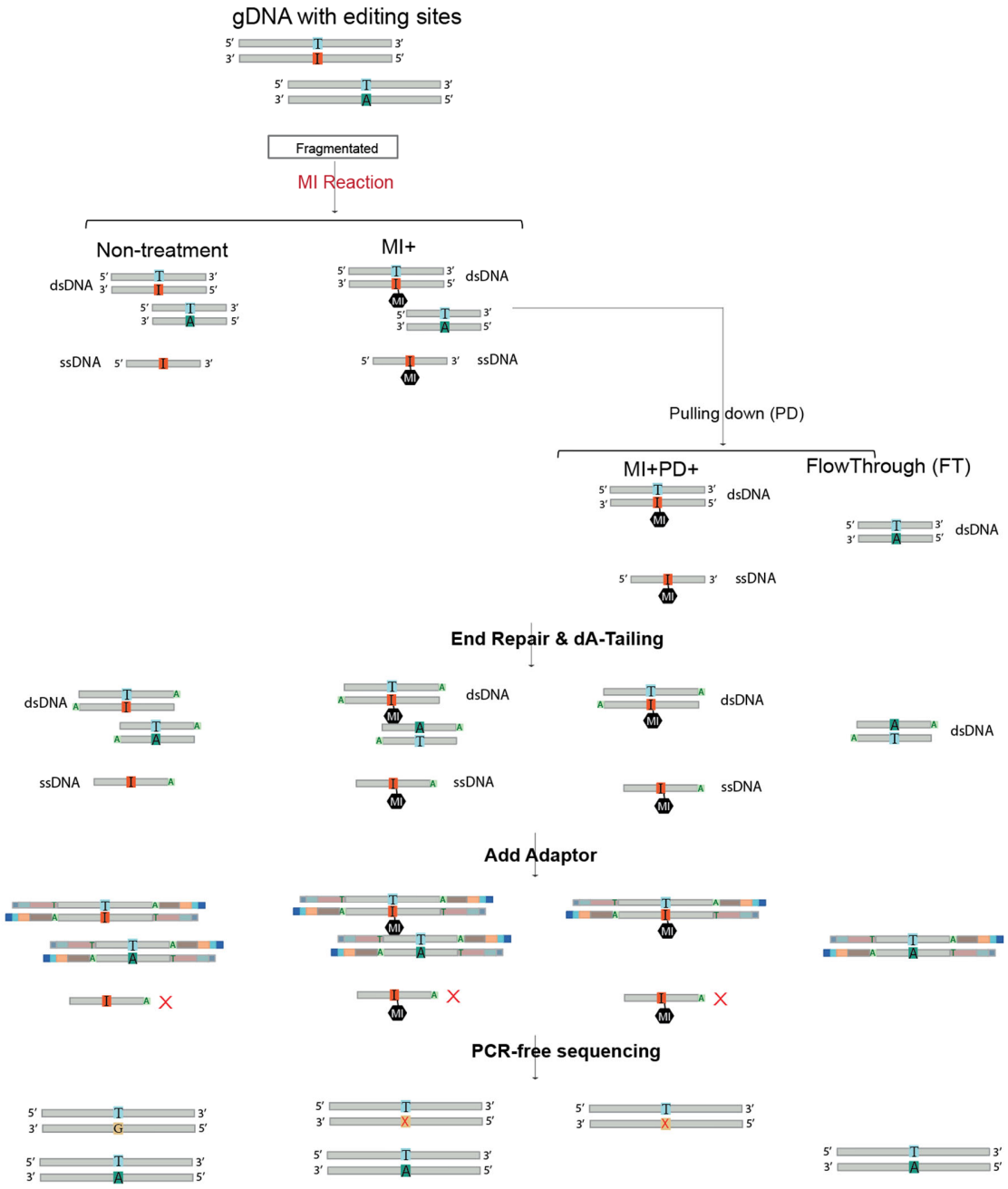


Figure 37. Flowchart for detecting new editing sites in DNA by ICLAMP-seq.

Acknowledgment

I would like to express my deepest gratitude to several individuals whose invaluable contributions and support have been fundamental to the completion of this study.

First and foremost, I extend my sincere thanks to my mentor, Dr. Masayuki Sakurai from the Institute of Biomedical Sciences, Tokyo University of Science, for his exemplary guidance throughout this journey. From the initial stages of experimental design to the intricate process of writing this paper, his expertise, patience, and insightful feedback have been instrumental in shaping my research and academic growth. His unwavering support and encouragement have not only enhanced my research skills but also inspired me to pursue excellence in my scholarly endeavors.

I am also deeply indebted to my supervisor, Dr. Ryo Goitsuka from the Institute of Biomedical Sciences, Tokyo University of Science, whose assistance was crucial in providing the experimental animals necessary for this study. His extensive support during various application processes and readiness to share his vast knowledge and resources have greatly contributed to the success of my research.

My heartfelt appreciation goes to two doctors from Faculty of Pharmaceutical Sciences in Tokyo University of Science, Dr. Kazuki Sato and Dr. Takeshi Wada, who played a pivotal role in synthesizing the acrylamidofluorescein essential for my experiments. Their technical expertise and willingness to collaborate have been invaluable to my project.

I extend my sincere gratitude to Dr. Yuriko Sakaguchi, Dr. Ayaka Murayama, and Dr. Tsutomu Suzuki from the Graduate School of Engineering, University of Tokyo. Their expertise and assistance were instrumental in the successful acquisition of the LC/MS results that significantly contributed to the advancement of my research.

I would also like to acknowledge the guidance and support provided by the former laboratory member, Dr. Shunpei Okada. Though he is no longer with us, his early contributions to my research and development have left a lasting impact on my academic journey.

Special thanks to Koki Nakayama, who graduated with a master's degree, for his collaboration in the preliminary stages of the research, which was crucial in laying the groundwork for the later phases of the study.

I am also grateful to the secretary, Ms. Miyoko Nakano, whose support extended beyond academic matters. Her assistance in administrative tasks, coupled with her caring advice on personal matters, has created a nurturing and conducive environment for my studies.

My appreciation also extends to all my peers in the laboratory. Their camaraderie, shared insights, and collaborative spirit have not only made this journey enjoyable but also immensely enriching.

Last but certainly not least, I owe a debt of gratitude to my parents. Their unconditional love, sacrifice, and belief in my abilities have been the cornerstone of my achievements. This accomplishment is as much theirs as it is mine.

In sum, the completion of this study would not have been possible without the collective support and guidance of all these individuals, to whom I am eternally grateful.