# Abstract

*Identification and analysis of the U-to-C RNA editing enzyme in plants for restoration of genetic disorders*

Ruchika,
Tsukahara Laboratory, JAIST

RNA editing is a post-transcriptional molecular process through which cells can be describe as site-specific changes in the nucleotide sequences within an RNA molecule that has been transcribed by RNA polymerase enzymes. Among various types of RNA editing, Adenine-to-Inosine, A-to-I(G), Cytidine-to-Uridine, C-to-U and Uridine-to-Cytidine, U-to-C are the most common. Studies have already revealed the mechanisms for A-to-I and C-to-U RNA editing which are generally caused by ADARs and APOBEC-AID deaminase family, respectively. However, the transaminase enzymes responsible for "Reverse" U-to-C editing is not yet discovered, although it is the abundant phenomenon in lower plant species like hornworts and rare in animals. Here I examined the U-to-C RNA editing in *Arabidopsis* tissues at different developmental stages of growth. In this study, the high-throughput RNA sequencing (RNA-seq) of 12-day-old and 20-day-old *Arabidopsis* seedlings was performed. The results showed that DEGs were expressed to higher levels in 12-day-old seedlings than in 20-day-old seedlings. Additionally, pentatricopeptide repeat (PPR) genes were also expressed at higher levels as indicated by the log2FC values. RNA-seq analysis of 12-day and 20-day-old Arabidopsis seedlings revealed candidates of U-to-C RNA editing events. Sanger sequencing of both gDNA and cDNA confirmed the seven U-to-C RNA editing sites.

Further, I investigated the U-to-C RNA editing-related genes in *Arabidopsis* tissues and the effects on mRNA stability, with a special focus on PPR proteins. Here, I have demonstrated the effects of the "reverse" RNA editing on the mRNA stability for all seven edited genes. In addition, the nuclear PPR gene (AT2G19280) in 12-day-old seedlings of *Arabidopsis thaliana* was also identified with U-to-C RNA editing. The U-to-C RNA editing sites were found in the untranslated region (3' UTR) of the mature mRNA and affect its secondary structure. Then, the correlation between U-to-C RNA editing-related genes and their mRNA abundance were examined. Furthermore, I investigated the effects of U-to-C RNA editing in Arabidopsis using the transcription inhibitor actinomycin D (Act D). The addition of Act D to the cell suspension culture of transgenic Arabidopsis generated by Agrobacterium-mediated transformation showed that single nucleotide base conversion adversely affected the mRNA secondary structure and stability.

Pentatricopeptide repeats (PPR) proteins are act as sequence-specific RNA-binding proteins within mitochondria and chloroplasts in almost all land plants. Genome-wide analysis of the hornworts, *Anthoceros agrestis*, revealed the PPR proteins in this species contain unique C-terminal DYW-like domains with specific signatures. In present work, I have explored the study on three different variants of C-terminal PPR proteins of hornworts, GRP-type, DRH- type and DYW-types. I investigated the RNA editing events by cloning the Hornworts PPR genes. A bacterial expression system was developed in which the Hornworts specific PPR protein variants were cloned with PPR56 (truncated DYW), *Physcomitrella patens* (moss) editing factor. We measured the gene expression levels of DYW variants. In addition, we demonstrated the functional homology of DYW domains with APOBEC1in human cells.

The MS2 system were initially developed with ADAR1 and APOBEC1 for A-to-I(G) and C to-U RNA editing, respectively. In this study, an expression system was designed with PPR56, *Physcomitrella patens* (moss) editing factor. We engineered an artificial RNA editing mechanism by combining the deaminase domain of plant derived PPR56 with a guideRNA (gRNA) which is complementary to target mRNA. As a target RNA, we used RNA encoding blue fluorescent protein (BFP) which was derived from the gene encoding GFP by a single T-to-C substitution. Upon transient expression of both components (PPR56 and gRNA), we confirmed the restoration of sequence of GFP revealing an editing efficiency of up to 85-100%, while previous developed system with APOBEC1 only showed about 20 % editing efficiency. Furthermore, we identified that the C-terminal E2-DYW domain of PPR56 is sufficient for C-to-U conversion in the MS2 system.

Lastly, I introduced a non-sequencing approach for the rapid detection of RNA editing using a portable micro-Temperature gradient Gel Electrophoresis (μTGGE). This is based on the principle of electrophoresis, which use temperature to denature the samples as it moves across the polyacrylamide gel. In this method, a fragment of double-

stranded DNA when heated, forming a gradient of double-stranded DNA to partially separated strands to completely separated single-stranded DNA. A sample of RNA editing with different nucleotide bases show the different melting profiles because of their different melting profile. Here, we demonstrated the difference between the melting profiles for edited and non-edited (wild type) RNAs. Reproducibility was evaluated from measuring the pattern similarity scores (PaSS) between the band patterns obtained with the edited and non-edited RNAs. This tool is providing a simple, and cost-effective method for detecting even a single base modification in RNAs. We expect that our rapid analyzing tool will foster further discoveries in this rapidly expanding field of molecular biology.

**Keywords:** U-to-C RNA editing, *Arabidopsis thaliana*, PPR56, DYW domains, hornworts, RNA-seq