

# Site to site detection and analysis of C-to-U RNA editing in rice mitochondria-encoded ORFs

**Peng Zheng**

Zhejiang University

**Yuqing Huang**

Zhejiang University

**Hao Chen**

Zhejiang University

**Hao Du**

Zhejiang University

**Jumin Tu** (✉ [jtu@zju.edu.cn](mailto:jtu@zju.edu.cn))

Institute of Crop Science, College of Agriculture and Biotechnology, Zhejiang University

<https://orcid.org/0000-0002-5971-6527>

---

## Original article

**Keywords:** C-to-U RNA editing, Mitochondria-encoded ORFs, Edited codon, Rice

**Posted Date:** March 18th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-17600/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background** Cytidine to uridine (C-to-U) RNA editing is an important type of substitutional RNA editing. In plants, C-to-U modification is almost omnipresent in chloroplasts and mitochondria, where it mainly serves to restore evolutionarily conserved amino acids by changing the codon information. The important roles that C-to-U RNA editing plays in organelle biogenesis, adaptation to environmental changes, and signal transduction have been confirmed in different plant species. In rice mitochondria, 491 C-to-U editing sites have been identified previously, and case studies have elucidated the function of several C-to-U editing sites in rice, but the functional consequence of most C-to-U alteration need to be investigated further.

**Results** Here, we totally identified 539 C-to-U editing sites in rice mitochondria-encoded ORFs, 87.2% of these editing sites were observed on the first or the second base of a codon, resulting in the alteration of encoded amino acid. Moreover, we found some novel editing sites and several inaccurately annotated sites which may be functionally important, with the support of the highly conserved amino acids encoded by these edited codons. Finally, we annotated all 539 C-to-U RNA editing sites in their biological context.

**Conclusions** We detected 539 C-to-U editing sites in rice mitochondria-encoded ORFs and annotated them with more precise information, which will facilitate our investigation on the function of C-to-U editing events in rice.

# Background

RNA editing is an essential posttranscriptional or co-posttranscriptional biological process that specifically changes the nucleotide sequence of a primary transcript, making the genetic information in RNA different from that of the DNA template. The term RNA editing was introduced for the first time in 1986 to describe the addition and deletion of uridine nucleotides to and from mRNAs in the kinetoplast, the specialized mitochondrion of trypanosomes (Benne et al., 1986). Since its initial discovery, RNA editing has been found in a wide range of organisms, including basal eukaryotes, land plants, vertebrates, fungi, and viruses as well (Bowe and dePamphilis, 1996). RNA modifications due to RNA editing comprise nucleotide insertions/deletion or substitutions that can occur in the nucleus, in the cytoplasm as well as in organelles (plastids and mitochondria) (Gott and Emeson, 2000).

RNA editing in plants organelles occurs mainly in the form of cytidine to uridine (C-to-U) conversion, albeit other types of RNA editing including the opposite U-to-C alteration has also been seen especially in chloroplasts RNAs of some species (Takenaka et al., 2013), its frequency is much less than that of C-to-U editing. In flowering plants, several hundred C-to-U substitutions are observed in the two energy-producing organelles, the majority locating in mitochondria. Plant C-to-U editing was found to be associated with nucleus-encoded pentatricopeptide repeat (PPR) proteins (Takenaka et al., 2019), which are dominantly characterized by 30–40 amino acid repeated motifs that can recognize and bind single-strand RNAs (Yin et al., 2013). Several studies have demonstrated that other types of non-PPR proteins, including multiple

organellar RNA editing factor (MORF) (Bentolila et al., 2012) and organelle RNA recognition motif-containing (ORRM) proteins (Shi et al., 2015; Shi et al., 2016), are required to form the editosome machinery in flowering plants, although its molecular assembly mechanism remains largely unknown (Sun et al., 2016).

Most of plant organellar C-to-U RNA editing occurs in protein coding regions (usually at the first or second position of codons), typically leading to amino acid changes that appear to be evolutionary conserved (Gray, 2003). Although C-to-U alteration most frequently modified internal codons, in rare case it also create translation initiation (ACG-to-AUG) or termination codons (CAA-to-UAA). Therefore, this RNA modification is believed to act as a proofreading mechanism to correct DNA mutation at RNA level and to generate functional proteins (Edera et al., 2018). C-to-U RNA editing occasionally occurs in untranslated regions (5' UTR or 3' UTR), introns, rRNA and tRNA molecules, modulating splicing, transcript stability, and translation efficiency. Indeed, several RNA editing events in introns have proved to improve the stability of functionally relevant secondary structure motifs (Carrillo et al., 2001).

Several studies have demonstrated that C-to-U RNA editing is essential for normal plant growth and development. The alteration of C-to-U RNA editing pattern can lead to a series of plant developmental defects, such as impaired chloroplast and mitochondrial biogenesis (Sosso et al., 2012), retarded seedling growth (Lin et al., 2015), reduced embryo and endosperm development (Liu et al., 2013), and hypersensitivity to various abiotic stresses (Zhu et al., 2014). Moreover, loss of C-to-U RNA editing in plant mitochondria can result in male sterility, also termed as the cytoplasmic male sterile (CMS) (Kempken et al., 1998). These findings suggest that C-to-U RNA editing plays important roles in various plant developmental processes, including organelle biogenesis, adaptation to environmental changes, and signal transduction.

Rice is not only a major staple food crop for more than half of the population worldwide, but also an important monocot model for basic molecular and genetic studies. Detection of C-to-U RNA editing sites in rice mitochondria will not only facilitates the functional investigation of the genes essential for mitochondria biogenesis and plant development, but also provides a valuable clue to generate CMS lines to improve hybrid rice production. A previous study have identified a total of 491 C-to-U RNA editing sites in 33 open reading frames and one pseudogene of rice mitochondrial genome (Notsu et al., 2002). These editing sites are not always correctly or completely annotated in primary database such as GenBank, where they are often indicated as misc\_feature or as simple exception notes. To overcome these limitations, researchers have developed a series of specialized databases, among which REDIdb was the first specialized database to collect plant RNA editing events and annotate them in their biological context (Giudice et al., 2018). The latest version of REDIdb contains 26618 RNA editing sites distributed among 281 organisms, including those C-to-U RNA editing sites in rice mitochondria. Benefiting from the availability of well-annotated mitochondrial editing sites through specialized databases, the biological functions of several C-to-U RNA editing events in rice mitochondria have been identified successfully (Kim et al., 2009; Toda et al., 2012; Xiao et al., 2018). Nevertheless, some editing sites that are not or inaccurately annotated have been observed by independent research groups. These problems, as a

consequence, will hamper and even preclude the functional identification of rice genes involved in C-to-U RNA editing events.

With the aim to find novel C-to-U RNA editing sites and correct potential annotation errors, direct sequencing of cDNAs amplified by RT-PCR was used to test ORFs in rice mitochondria genome. We totally identified 539 C-to-U RNA editing in rice mitochondria, 48 more than the number reported previously. 87.2% of these editing sites were observed on the first or the second base of a codon, resulting in the alteration of encoded amino acid. We also found some novel editing sites and several inaccurately annotated sites. Finally, we annotated all 539 C-to-U RNA editing sites in their biological context. More precise information about C-to-U RNA editing sites in this study will facilitate our investigation on the function of C-to-U editing events in rice, at the same time provides a valid benchmark with which to compare rice mitochondria C-to-U modification.

## Results

### **Total 539 C-to-U RNA editing sites were detected in rice mitochondria-encoded ORFs**

Given the fact that most editing sites locate in the protein-coding regions, we only detected the C-to-U modifications in rice mitochondria-encoded ORFs. By analyzing gene organization in the rice mitochondria genome, Notus et al. (2002) identified 35 genes for annotated proteins, two pseudogenes for ribosomal protein, and several genes for tRNA and rRNA. Comparison between the mitochondrial cDNA sequences with the mitochondrial DNA sequence acquired a total of 491 C-to-U editing sites in the 33 protein-coding genes and one pseudogene in rice mitochondria (Table 1). In this study, in addition to the 35 ORFs for known proteins and two pseudogenes, we also investigated the 10 ORFs (orf152a, orf152b, orf161, orf153, orf162, orf165, orf176, orf187, orf224, and orf288), which is capable of encoding more than 150 amino acids and have been proved to be transcribed, even though no RNA editing site was found in any of these 10 ORFs in previous study (Notsu et al., 2002).

Table 1  
 Number of editing sites in rice mitochondria-encoded ORFs  
 reported by different studies.

Gene	Number of C-to-U editing sites		
	REDIdb	Notsu et al., 2002	In this study
atp1	5	5	5
atp6	16	17	17
atp9	8	8	8
ccmB (ccb2)	35	35	37
ccmC (ccb3)	35	36	36
ccmFC (ccb6c)	27	27	27
ccmFN (ccb6n)	31	31	38
cob (cytb)	19	19	19
cox1	4	4	5
cox2	19	19	20
cox3	1	1	12
matR	0	0	14
nad1	23	23	27
nad2	30	30	32
nad3	15	15	18
nad4	20	20	20
nad4L	10	10	11
nad5	11	11	11
nad6	18	18	4
nad7	32	32	35
nad9	12	12	12
orf25 (atp4)	9	9	9
orfB (atp8)	4	4	7
orfX	33	33	34

Gene	Number of C-to-U editing sites		
	REDIdb	Notsu et al., 2002	In this study
rpl2	1	1	1
rpl5	1	1	1
rpl16	12	12	13
rps1	3	3	2
rps2	10	10	10
rps3	10	10	12
rps4	15	15	18
rps7	2	2	2
pseudo-rps11	4	4	4
rps12	0	0	5
rps13	8	8	7
rps19	6	6	6
Total	489	491	539

In total we identified 539 significant C-to-U editing sites in rice mitochondria genome with a defined threshold of editing level above 0.10, and these RNA modifications are distributed in 36 different genes (Table 1). Compared with data from Notsu et al. (2002), although the number of editing sites in most genes has no or slight different, huge distinction was observed in several genes. For example, 1 editing sites in *cox3* and 18 in *nad6* have been identified by Notsu et al. (2002), while in this study 12 and 4 editing sites are detected in *cox3* and *nad6*, respectively. Interestingly, we also identified 14 C-to-U editing sites in *matR* and 5 sites in *rps12* (Table 1), but no C-to-U substitution has been reported in these two genes before. The numbers of C-to-U editing sites in each transcript are documented in Table 1. For reference, we have also compiled in Table 1 the numbers of C-to-U editing sites annotated in REDIdb (Giudice et al., 2018).

## Correction of annotation errors in REDIdb

REDIdb is a unique database focusing on collecting and annotating plant organellar RNA editing sites (Giudice et al., 2018). With its third release, all annotations have been manually check to correct potential errors. This does not mean, however, that there is no annotation error still existing in the bioinformatics resource. For example, in accession AB076666, the annotation for a C-to-U editing in rice mitochondrial gene *atp1* is reported at nucleotide 1291, which results in the conversion of amino acid from Pro to Ser (Fig. 1a). On the basis of sequencing results from three biological replicates, we found that the real

editing site is located one nucleotide downstream of the reported editing site at 1292, where a modification at the second codon position converts a genetically encoded amino acid Pro into Leu (Fig. 1b). This is also well-confirmed by our multiple alignment result that the Leu at that position is evolutionarily conserved in other plants (Fig. 1c). Coincidentally, a C-to-U substitution in *nad7* transcript is annotated to located at nucleotide 446, where the residue Pro encoded by CCG codon is converted to Leu encoded by TCG (Fig. 1d). However, our sequencing result indicates that the bona fide C-to-U editing site occurred at nucleotide 445 (Fig. 1e), creating a CTG codon to encode a conserved Ser as revealed in the annotation (Fig. 1d). To improve the annotation quality and facilitate the study of C-to-U editing sites in rice mitochondria, we have corrected all the annotation errors in REDIdb and compiled right annotations in Additional file 1: Table S1.

## Characterization of editing sites in rice mitochondria

All 539 C-to-U RNA editing sites in rice mitochondrial genome are unevenly distributed among different genes, with the percentage of editing events per gene C content ranging from 0.29% (*rpl2*) to 24.5% (*ccmB*) (Additional file 2: Table S2). The *ccmFN* gene, encoding N-terminal maturation subunit F of cytochromes *c*, is the most edited gene (38), whereas the *rpl2* and *rpl5*, both encoding ribosomal proteins, are the least edited genes (1 for each) (Table 1). The number of editing site for individual gene of a given mitochondrial complex is also quite variable. Only 4 C-to-U RNA editing events occurred in *nad6* gene, while 35 editing sites are identified in *nad7* (Table 1), although no significant correlation was observed between edited Cs and total Cs or gene length (Additional file 2: Table S2). Our data also confirm a level of species specificity of C-to-U editing. For instance, five edited site was detected in the rice *cox1* transcript, whereas the homologs from *Arabidopsis* and *V. vinifera* are edited at 0 and 22 positions, respectively (Picardi et al., 2010; Giudice et al., 2018).

Among these 539 C-to-U RNA editing sites, 33.95% (183) and 53.25% (287) occurred at the first and second base of a codon, respectively, and only 12.80% (69) occurred at the third base of the codon (Fig. 2a). This frequency is similar to that observed in mitochondria of *V. vinifera*. All the C-to-U modifications resulted in 512 codon alterations, of them 71 were synonymous (AA is unchanged) and 441 were nonsynonymous (AA is changed) (Fig. 2b). The number of editing sites in a codon is not limited to one. Actually, 485 out of the 512 codons harbor only a single editing site, and the rest 27 condons possess two editing sites for each. No codon that contains three editing sites was detected (Fig. 2c). The three most nonsynonymous changes induced by C-to-U editing were Ser-to-Leu (111), Pro-to-Leu (87) and Ser-to-Phe (67) (Fig. 2d). These three changes from hydrophilic to hydrophobic lead to the increase of content of hydrophobic amino acids. Indeed, 83.22% nonsynonymous changes eventually produced hydrophobic amino acids in rice mitochondrial, suggesting a biochemical function of C-to-U editing in protein modifications. In plants, these conversions mostly restore evolutionarily conserved amino acid. Moreover, Ser-to-Leu and Ser-to-Phe conversions potentially increase the hydrophobic residues in protein-protein interface while Pro-to-Leu substitutions can contribute to protien function by stabilizing 3D structures (Yura and Go, 2008).

As mentioned above, C-to-U editing sites at 512 codons in rice mitochondria are characterized by non-random distribution with regard to codon positions. Additionally, the bias of editing events towards specific codons were also observed. In particular, the top three edited codons were TCA (71) followed by CCA (49) and TCG (40), accounting for 31.25% of all edited codons (Fig. 2e). GCC and AGC are the only two C-containing codons that were never affected by C-to-U alteration. RNA editing in rice mitochondria creates the site of translation initiation in *nad1* and *nad4L* transcripts, and induces five stop codons in *rpl11*, *rpl16*, *atp6*, *atp9*, and *ccmFC* genes (Additional file 1: Table S1).

## **Novel C-to-U editing sites may be functionally important**

Besides the discovery of incorrectly annotated C-to-U events, we detected more than 40 novel editing sites that were not reported before. Some of these sites with low editing levels induced synonymous editing, whereas some sites highly edited generally resulted in nonsynonymous modification. Since the important biological function of several highly edited C-to-U editing sites located in the transcripts has been elucidated, it is possible that these highly edited sites are also functionally important.

For instance, *ccmFN* encodes a subunit of heme lyase complex required for the cytochrome c maturation. In maize, RNA editing of *ccmFN* transcript at position 1553 results in amino acid conversion from Ser to Phe. Abolishment of C-to-U editing at this site leads to the loss of CcmFN protein and a marked reduction of c-type cytochrome, and eventually compromises the development of maize embryo and endosperm (Sun et al., 2015). Actually, C-to-U editing-induced Phe at this site is highly conserved across different species, indicating that nonsynonymous C-to-U modification at some positions, at least at position 1553, is essential for the function of CcmFN protein. In this study, we observed 4 new highly edited sites in *ccmFN* transcript (Fig. 3a). Editing at positions 121, 263, 356 and 365 induced Pro-to-Ser, Pro-to-Leu, Ser-to-Phe, and Ser-to-Leu nonsynonymous substitutions, respectively (Fig. 3a). Multiple alignment of CcmFN proteins from various plant species indicated that these four amino acids encoded by edited codon are in high conservation (Fig. 3b). Compared with synonymous editing sites, nonsynonymous editing sites commonly showed higher conservation levels as well as editing levels (Giudice et al., 2018; Chu and Wei, 2019), and nonsynonymous sites are more essential and functional, suggesting that detection of novel C-to-U editing sites in this study will provide valuable source for rice biological study.

## **Discussion**

In plants, RNA editing was first identified as C-to-U conversions in mitochondrial mRNA for *coxII* (Covello and Gray, 1989), followed by its identification in chloroplasts mRNA for *rpl2* (Hoch et al., 1991). Since then, the occurrence of C-to-U editing has been widely discovered in plant organelles. With the advent of next-generation sequencing (NGS) technique, numerous novel RNA editing events have been uncovered based on the releasing of many complete plant organellar genomes and RNA-seq data. Databases are also developed to collect and annotate the greatly increased C-to-U RNA editing sites (Yura et al., 2009; Giudice et al., 2018; Li et al., 2019). However, the existence of annotation errors and the lack of novel editing sites in the database will, to some extent, limit the study explaining the biological functions of particular C-to-U editing events.

It is conceivable that RNA editing shows tissue-specific and partial patterns in a given species (Picardi et al., 2010; Tseng et al., 2013). The tissue-specific sites could be resulted from the tissue-specific expressed genes. In this study, only leaf was used as sample to detect C-to-U editing sites, which consequently discarded the sites that specifically occurred in other rice tissues. Moreover, due to limitations of Sanger sequencing, the sites with editing level less than 0.10 were not considered as C-to-U editing sites. This setting condition will inevitably preclude some genuine editing sites which are extremely lowly edited in rice leaf, given the fact that editing levels of some sites are very dynamic during plant development and growth (Grosskopf and Mulligan, 1996; Peeters and Hanson, 2002; Tseng et al., 2013). Hence, it is believable that there should be more editing sites in rice mitochondrial ORFs.

Although the exact functions of most mitochondrial RNA events are unclear, some studies of rice PPR mutants that abolish C-to-U conversions at particular sites may prove the biological significance of these editing events. For example, the complete loss of seven RNA modifications on five mitochondrial transcripts in *org1* mutant results in delayed seed germination, retarded growth and sterility (Kim et al., 2009). In *pps1* RNAi plants, significant reduction of editing efficiency at five consecutive editing sites in *nad3* decreases the activity of several complexes in mitochondrial electron transport chain, resulted in delayed development and partial pollen sterility (Xiao et al., 2018). The loss of C-to-U RNA editing in these mutants and the associated development defect suggest that RNA editing to restore evolutionarily conserved amino acids in rice mitochondrial transcripts may be important for the function of encoded proteins. In this study, we discovered many novel editing sites and several incorrectly annotated sites via RT-PCR and Sanger sequencing. Some of them are highly edited and results in the restoration of conserved amino acids, implying that C-to-U RNA editing at these sites may have important biological functions. Our work provided a more perfect and precise information about C-to-U RNA editing sites in rice mitochondria, which will not only offer an excellent opportunity to investigate the effect of RNA editing on mitochondrial function, but also facilitate the functional identification of rice genes involved in C-to-U RNA editing events.

## Materials And Methods

### RNA isolation and cDNA synthesis

The complete mitochondrial genome sequence and information on the RNA editing sites in rice mitochondria are all determined in japonica cultivar Nipponbare (Notsu et al., 2002). To keep the genetic background the same, different tissues used for detection of C-to-U RNA editing sites were harvested from Nipponbare.

Editing sites were verified by reverse-transcription polymerase chain reaction (RT-PCR) experiments. For RNA isolation, rice tissues were separately frozen in liquid nitrogen, ground in a mortar with a pestle and extracted with RNAiso Plus reagent (TaKaRa). Quality and concentration of total RNA was analyzed using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Waltham). The first strand cDNA was synthesized with PrimeScript RT reagent Kit with gDNA eraser (TaKaRa) following the manufacturer's

protocol. Gene-specific primers that covered the full-length ORFs were used for cDNA amplification. 2  $\mu$ L of a 10-fold dilution of the cDNA solution were used as a template for RT-PCR. PCR products were purified with an AxyPrep™ PCR Cleanup kit (AXYGEN) and subjected to sequencing using ABI PRISM 3700. All primers for RT-PCR are listed in Additional file 3: Table S3.

## RNA editing analysis

RNA editing sites were identified in two ways: 1, cDNA sequences were aligned with their corresponding genomic sequences in BLASTN to obtain the mismatched sites. If the unpaired bases were T in cDNA but C in the genomic DNA, this site was considered as a C-to-U editing site; 2, Sequencing chromatograms were directly read to find the double-peaked sites. In such cases, if the two peaks indicate C and T, while a C is observed at the same position in the reference genome, the site was regarded as a editing site or not based on the editing level ( $T/C + T$ ), which calculated with the relative height of C and T peaks in the sequence chromatogram. To avoid the false-positive variants caused by technical limitations, the double-peaked sites with editing levels above 0.10 and appear in at least 2 samples were regarded as RNA editing sites (extremely low level is possibly caused by sequencing errors). Known C-to-U substitutions in rice mitochondria were downloaded from REDIdb database and served as reference to identify new editing sites.

## Abbreviations

ORF  
Open reading frame; C-to-U:Cytidine to uridine;PPR:Pentatricopeptide repeat; RT-PCR:Reverse transcription- polymerase chain reaction; CMS:cytoplasmic male sterile;

## Declarations

### Acknowledgments

The authors would like to thank Jan mehmood for the English writing.

### Author contributions

JM Tu, H Du, and P Zheng designed the experiments; P Zheng performed the most of experiments and analysed the data; P Zheng, YQ Huang, and H Chen wrote the manuscript; H Du and JM Tu revised the manuscript. All authors read and approved the final manuscript.

### Funding

This work was financially supported by the National Natural Science Foundation of China (Project no. 31371592 and 31501375).

### **Availability of Data and Materials**

The datasets supporting the conclusions of this article are included within the article (and its additional files).

### **Ethics Approval and Consent to Participate**

Not applicable.

### **Consent for Publication**

Not applicable.

### **Competing interests**

The authors declare that they have no competing interests.

### **Author details**

Zhejiang Provincial Key Laboratory of Crop Germplasm Resources, Institute of Crop Science, Zhejiang University, No.866, Yu-Hang-Tang Road, Hangzhou 310058, China

## **References**

Benne R, Burg JVD, Brakenhoff JPJ, Sloof P, Boom JHV, Tromp MC (1986) Major transcript of the frameshifted *coxII* gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell* 46: 819-826

Bentolila S, Heller WP, Sun T, Babina AM, Friso G, van Wijk KJ, Hanson MR (2012) RIP1, a member of an *Arabidopsis* protein family, interacts with the protein RARE1 and broadly affects RNA editing. *Proc Natl Acad Sci U S A* 109: E1453-1461

- Bowe LM, dePamphilis CW (1996) Effects of RNA editing and gene processing on phylogenetic reconstruction. *Mol Biol Evol* 13: 1159-1166
- Carrillo C, Chapdelaine Y, Bonen L (2001) Variation in sequence and RNA editing within core domains of mitochondrial group II introns among plants. *Mol Gen Genet* 264: 595-603
- Chu D, Wei L (2019) The chloroplast and mitochondrial C-to-U RNA editing in *Arabidopsis thaliana* shows signals of adaptation. *Plant dir* 3: e00169
- Covello PS, Gray MW (1989) RNA editing in plant mitochondria. *Nature* 341: 662-666
- Edera AA, Gandini CL, Sanchez-Puerta MV (2018) Towards a comprehensive picture of C-to-U RNA editing sites in angiosperm mitochondria. *Plant Mol Biol* 97: 215-231
- Giudice CL, Pesole G, Picardi E (2018) REDIdb 3.0: A Comprehensive Collection of RNA Editing Events in Plant Organellar Genomes. *Front Plant Sci* 9: 482
- Gott JM, Emeson RB (2000) Functions and mechanisms of RNA editing. *Annu Rev Genet* 34: 499-531
- Gray MW (2003) Diversity and evolution of mitochondrial RNA editing systems. *IUBMB Life* 55: 227-233
- Grosskopf D, Mulligan RM (1996) Developmental- and tissue-specificity of RNA editing in mitochondria of suspension-cultured maize cells and seedlings. *Curr Genet* 29: 556-563
- Hoch B, Maier RM, Appel K, Igloi GL, Kössel H (1991) Editing of a chloroplast mRNA by creation of an initiation codon. *Nature* 353: 178-180
- Kempken F, Howad W, Pring DR (1998) Mutations at specific atp6 codons which cause human mitochondrial diseases also lead to male sterility in a plant. *FEBS Lett* 441: 159-160
- Kim SR, Yang JI, Moon S, Ryu CH, An K, Kim KM, Yim J, An G (2009) Rice *OGR1* encodes a pentatricopeptide repeat-DYW protein and is essential for RNA editing in mitochondria. *Plant J* 59: 738-749
- Li M, Xia L, Zhang Y, Niu G, Li M, Wang P, Zhang Y, Sang J, Zou D, Hu S, Hao L, Zhang Z (2019) Plant editosome database: a curated database of RNA editosome in plants. *Nucleic Acids Res* 47: D170-D174
- Lin DZ, Gong XD, Jiang Q, Zheng KL, Zhou H, Xu JL, Teng S, Dong YJ (2015) The rice *ALS3* encoding a novel pentatricopeptide repeat protein is required for chloroplast development and seedling growth. *Rice (N Y)* 8: 17. doi:10.1186/s12284-015-0050-9
- Liu YJ, Xiu ZH, Meeley R, Tan BC (2013) *Empty Pericarp5* Encodes a Pentatricopeptide Repeat Protein That Is Required for Mitochondrial RNA Editing and Seed Development in Maize. *Plant Cell* 25: 868-883

- Notsu Y, Masood S, Nishikawa T, Kubo N, Akiduki G, Nakazono M, Hirai A, Kadowaki K (2002) The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. *Mol Genet Genomics* 268: 434-445
- Peeters NM, Hanson MR (2002) Transcript abundance supercedes editing efficiency as a factor in developmental variation of chloroplast gene expression. *RNA (New York, N.Y.)* 8: 497-511
- Picardi E, Horner DS, Chiara M, Schiavon R, Valle G, Pesole G (2010) Large-scale detection and analysis of RNA editing in grape mtDNA by RNA deep-sequencing. *Nucleic Acids Res* 38: 4755-4767
- Shi X, Germain A, Hanson MR, Bentolila S (2016) RNA Recognition Motif-Containing Protein ORRM4 Broadly Affects Mitochondrial RNA Editing and Impacts Plant Development and Flowering. *Plant Physiol* 170: 294-309
- Shi X, Hanson MR, Bentolila S (2015) Two RNA recognition motif-containing proteins are plant mitochondrial editing factors. *Nucleic Acids Res* 43: 3814-3825
- Sosso D, Mbelo S, Vernoud V, Gendrot G, Dedieu A, Chambrier P, Dauzat M, Heurtevin L, Guyon V, Takenaka M, Rogowsky PM (2012) PPR2263, a DYW-Subgroup Pentatricopeptide repeat protein, is required for mitochondrial *nad5* and *cob* transcript editing, mitochondrion biogenesis, and maize growth. *Plant cell* 24: 676-691
- Sun F, Wang X, Bonnard G, Shen Y, Xiu Z, Li X, Gao D, Zhang Z, Tan BC (2015) *Empty pericarp7* encodes a mitochondrial E-subgroup pentatricopeptide repeat protein that is required for *ccmFN* editing, mitochondrial function and seed development in maize. *Plant J* 84: 283-295
- Sun T, Bentolila S, Hanson MR (2016) The Unexpected Diversity of Plant Organelle RNA Editosomes. *Trends Plant Sci* 21: 962-973
- Takenaka M, Jörg A, Burger M, Haag S (2019) RNA editing mutants as surrogates for mitochondrial SNP mutants. *Plant Physiol Bioch* 135: 310-321
- Takenaka M, Zehrmann A, Verbitskiy D, Hartel B, Brennicke A (2013) RNA editing in plants and its evolution. *Annu Rev Genet* 47: 335-352
- Toda T, Fujii S, Noguchi K, Kazama T, Toriyama K (2012) Rice *MPR25* encodes a pentatricopeptide repeat protein and is essential for RNA editing of *nad5* transcripts in mitochondria. *Plant J* 72: 450-460
- Tseng C-C, Lee C-J, Chung Y-T, Sung T-Y, Hsieh M-H (2013) Differential regulation of *Arabidopsis* plastid gene expression and RNA editing in non-photosynthetic tissues. *Plant Mol Biol* 82: 375-392
- Xiao H, Zhang Q, Qin X, Xu Y, Ni C, Huang J, Zhu L, Zhong F, Liu W, Yao G, Zhu Y, Hu J (2018) Rice *PPS1* encodes a DYW motif-containing pentatricopeptide repeat protein required for five consecutive RNA-editing sites of *nad3* in mitochondria. *New Phytol* 220: 878-892

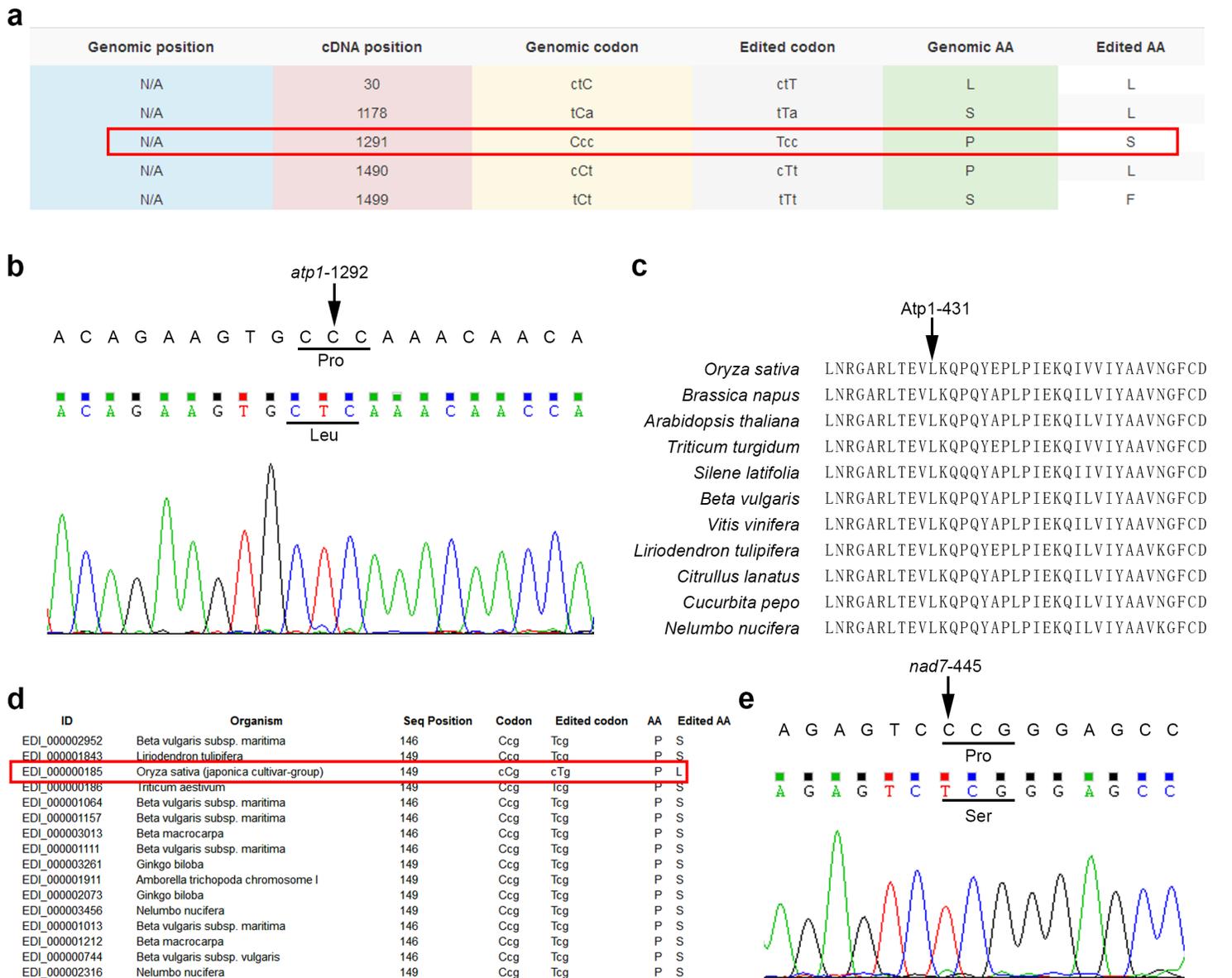
Yin P, Li QX, Yan CY, Liu Y, Liu JJ, Yu F, Wang Z, Long JF, He JH, Wang HW, Wang JW, Zhu JK, Shi YG, Yan NE (2013) Structural basis for the modular recognition of single-stranded RNA by PPR proteins. *Nature* 504: 168-173

Yura K, Go M (2008) Correlation between amino acid residues converted by RNA editing and functional residues in protein three-dimensional structures in plant organelles. *BMC Plant Biol* 8: 79

Yura K, Sulaiman S, Hatta Y, Shionyu M, Go M (2009) RESOPS: a database for analyzing the correspondence of RNA editing sites to protein three-dimensional structures. *Plant Cell Physiol* 50: 1865-1873

Zhu Q, Dugardeyn J, Zhang C, Muhlenbock P, Eastmond PJ, Valcke R, De Coninck B, Oden S, Karampelias M, Cammue BP, Prinsen E, Van Der Straeten D (2014) The *Arabidopsis thaliana* RNA editing factor SLO2, which affects the mitochondrial electron transport chain, participates in multiple stress and hormone responses. *Mol Plant* 7: 290-310

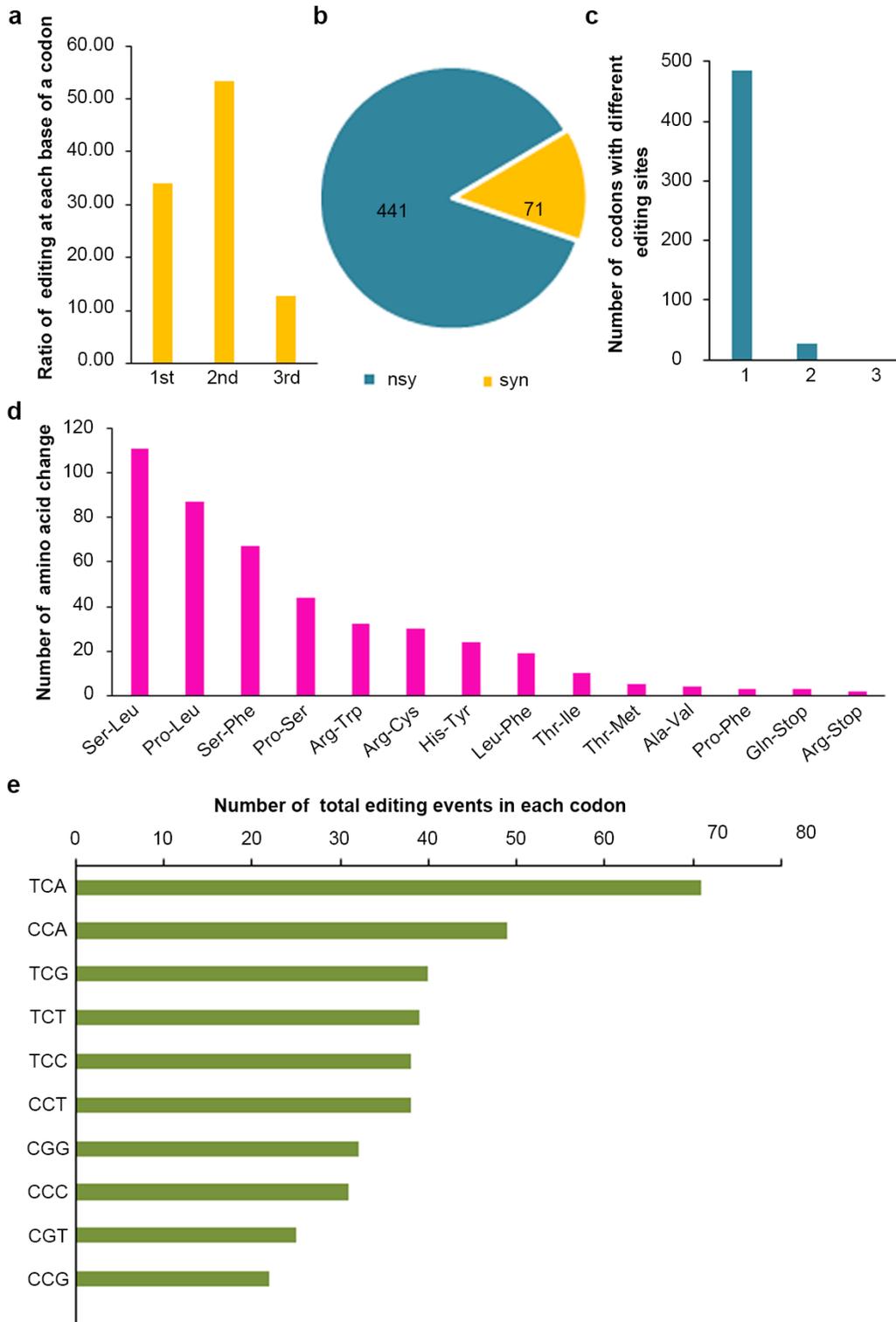
## Figures



**Figure 1**

Confirming and correcting the error annotations in REDIdb. A, Annotation of RNA editing sites in atp1 transcript in REDIdb. Red box indicates the incorrectly annotated RNA editing at atp1-1291. B, Sequence chromatogram of PCR-amplified atp1 cDNA. Arrow indicates the editing site atp1-1292. Amino acids before and after editing are shown below the codons indicated by black lines. C, Alignment of amino acid sequences of rice Atp1 and its orthologs from other plants species around the amino acid Atp1-431. The sequences were retrieved from REDIdb and aligned with Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>). Arrow indicates the amino acid for rice Atp1-431 encoded by edited codon. Note that amino acid sequences from REDIdb are translated from edited transcripts. D, Annotations retrieved from REDIdb about amino acid alteration at Nad7-149 in rice and its counterpart in other species. Red box indicates the alteration of amino acid 149 resulted from incorrectly annotated RNA editing site of rice nad7-446. E, Sequence chromatogram of PCR-amplified nad7 cDNA. Arrow indicates

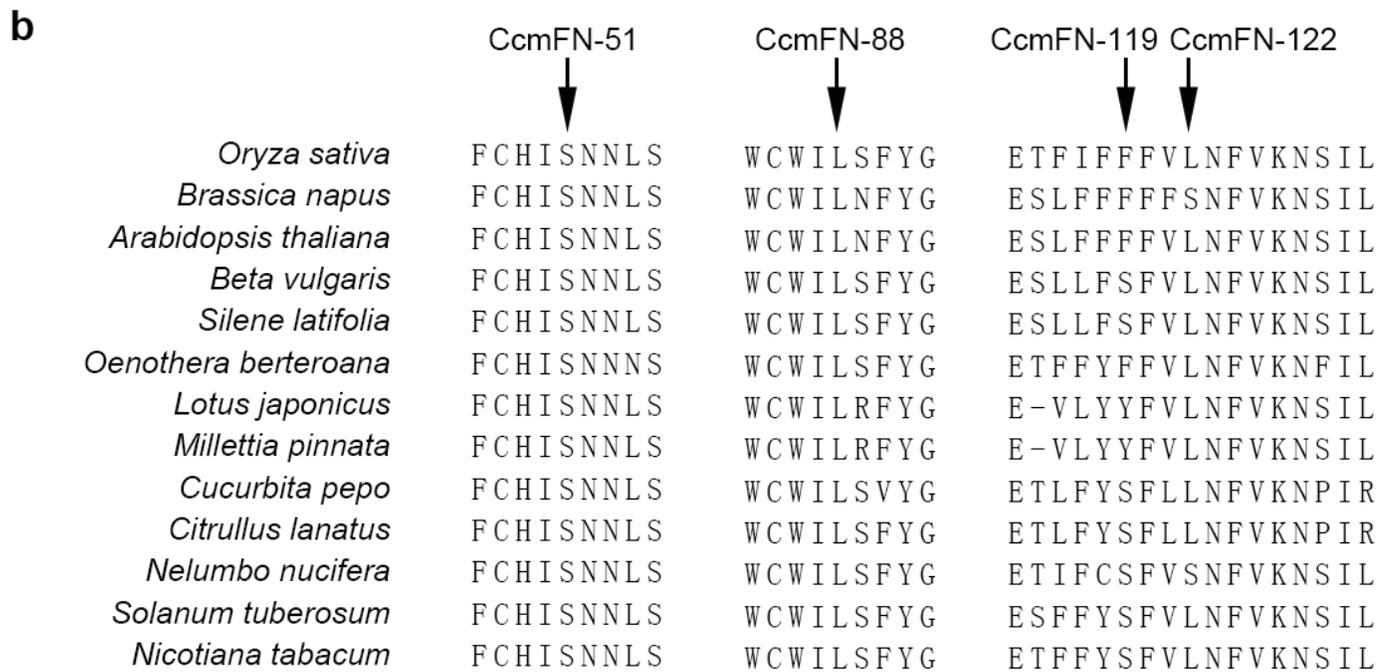
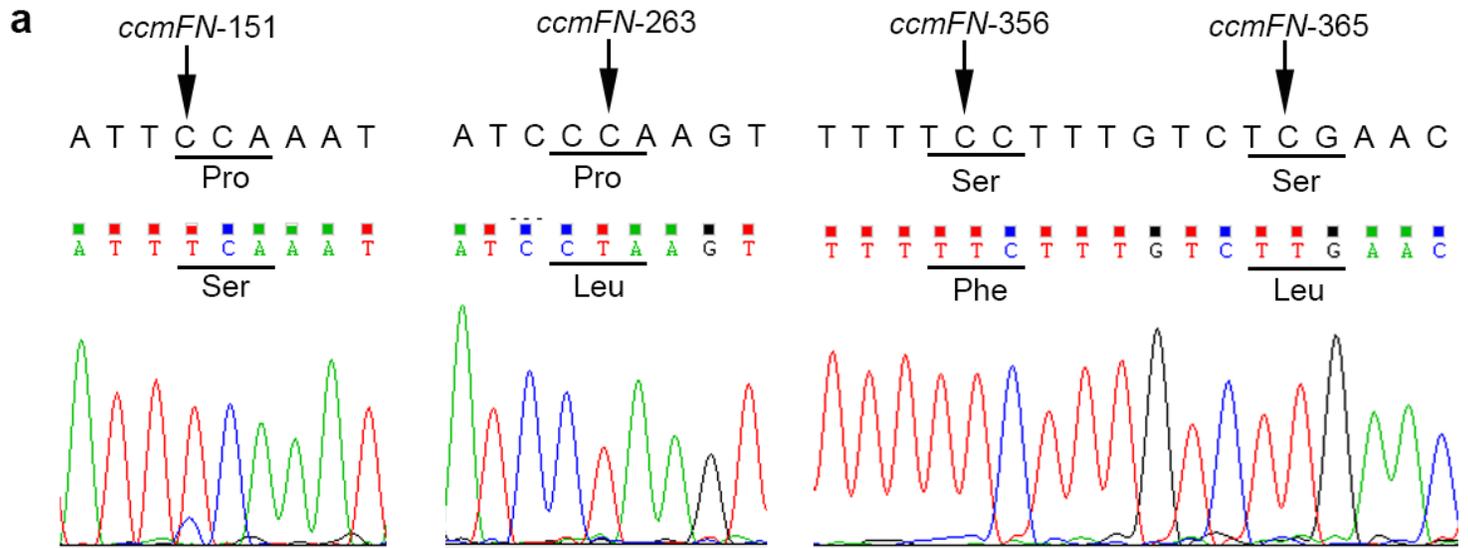
the editing site nad7-445. Amino acids before and after editing are shown below the codons indicated by black line.



**Figure 2**

Principal statistics of C-to-U RNA editing sites in ORFs of rice mitochondria. A, Distribution of editing sites at each codon position. B, Number of nonsynonymous (nsy) and synonymous (syn) amino acid

conversions by RNA editing. C, Number of codons with different number of editing sites. D, Number of amino acid changes induced by C-to-U RNA editing. E, Distribution of editing events in each codon.



**Figure 3**

Identification of novel RNA editing sites in rice mitochondria-encoded ORFs. A, Sequence chromatogram of PCR-amplified *ccmFN* cDNA. Arrow indicates the novel editing sites *ccmFN-151*, *ccmFN-263*, *ccmFN-356*, and *ccmFN-365*. Amino acids before and after editing are shown below the codons indicated by black lines. B, Alignment of amino acid sequences from rice *CcmFN* and its orthologs from other plant species. The sequences of *Oryza sativa*, *Citrullus lanatus*, *Lotus japonicus*, *Cucurbita pepo*, *Oenothera berteroa*, *Arabidopsis thaliana*, *Silene latifolia*, *Silene latifolia*, *Nelumbo nucifera* and *Millettia pinnata* were retrieved from REDIdb and sequences of *Nicotiana tabacum* (BAD83458.2), *Beta vulgaris* (NP\_063987.2) and *Solanum tuberosum* (QE76363.1) were downloaded from NCBI. All the sequences

were aligned with Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>). Arrows indicate the amino acids for rice CcmFN-51, CcmFN-88, CcmFN-119, and CcmFN-122 encoded by edited codons. Note that amino acid sequences from NCBI are translated from unedited transcripts.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1.xlsx](#)
- [TableS3.xls](#)
- [TableS2.docx](#)