The efficacy and safety of SARS-CoV-2 vaccines mRNA1273 and BNT162b2 might be complicated by rampant C-to-U RNA editing

Zhongzheng Bian[1] · Ziqian Wu[1] · Nan Liu[1] · Xiao Jiang[2]

Abstract

The SARS-CoV-2 RNA vaccines are smartly designed to increase the synonymous codon usage by introducing multiple U-to-C mutations. This design would elevate the translation efficiency of vaccine RNAs. However, we found evidence to reason that the designed cytidines might be converted to uridines again by C-to-U RNA deamination in host cells. This C-to-U mechanism might be a main factor that affects the efficacy and safety of RNA vaccines.

Keywords SARS-CoV-2 · Vaccines · Synonymous codon usage · C-to-U RNA editing · Efficacy

Introduction

C-to-U RNA editing is a widespread mechanism in eukaryotes that modifies the cellular RNA by changing cytidines to uridines (Chu and Wei 2019b; Duan et al. 2023; Rajasekhar and Mulligan 1993). It is believed that the editing enzymes APOBECs are designed for restricting exogenous viruses (Harris and Dudley 2015), but to what extent the viral RNA is edited remains unclear.

Under this global COVID-19 pandemic, the emergence of SARS-CoV-2 vaccines contributed a lot to reducing the casualty and controlling the pandemic. One of the successful vaccines is the RNA vaccine. Generally, the design of SARS-CoV-2 RNA vaccines was to introduce multiple U-to-C mutations to the viral sequence in order to increase the synonymous codon usage and translation efficiency.

Synonymous codon usage is also a ubiquitous phenomenon existing in all living organisms (Chu and Wei 2021; Yu et al. 2021; Zhou et al. 2016). C/G-ending codons are generally favored by the hosts and have faster translation rates than the A/T(U)-ending codons (Chu and Wei 2019a; Li et al. 2021). Therefore, the U-to-C transition in vaccine sequences is a smart design that could elevate the translation efficiency of vaccine RNAs.

In this study, we would first acknowledge the success of the design of two RNA vaccines: mRNA1273 and BNT162b2. Then, we reason that the designed cytidines might be converted to uridines again by the cellular C-to-U RNA deamination system. This C-to-U mechanism sheds concerns on the efficacy and safety of RNA vaccines.

Results

The successful design of two SARS-CoV-2 RNA vaccines

The efficacy and safety of two SARS-CoV-2 RNA vaccines mRNA1273 (Baden et al. 2021) and BNT162b2 (Polack et al. 2020) has been previously reported. In this brief piece, we will first acknowledge the successful design of both RNA vaccines and then raise a concern that might complicate the efficacy and safety of RNA vaccines.

We obtained the two SARS-CoV-2 RNA vaccine sequences via https://github.com/NAalytics/Assemblies-of-putative-SARS-CoV2-spike-encoding-mRNA-sequences-for-vaccines-BNT-162b2-and-mRNA-1273/blob/main/Figure1Figure2_032321.fasta. Generally, when we aligned the sequence of the reference *S* gene with the two RNA vaccine sequences (Fig. 1A), we would obtain the

✉ Xiao Jiang
jx_kaishi@163.com

[1] Department of Emergency Medicine, The Affiliated Hospital of Qingdao University, Qingdao 266000 Shandong, China

[2] Department of Emergency Intensive Care Unit, The Affiliated Hospital of Qingdao University, Qingdao 266000 Shandong, China

Fig. 1 The sequential features of the two RNA vaccines. (**A**) Alignment between reference *S* gene and two RNA vaccine sequences. (**B**) The Reference-Vaccine differences in CDS. Missense and synonymous sites were shown separately. Different colors represent the BNT162b2-specific substitutions (orange), mRNA1273-specific substitutions (blue), and shared substitutions (red). (**C**) Mutation types of the Reference-Vaccine differences on synonymous sites. (**D**) Alignment at codon level and comparison of their RSCU. (**E**) Comparison of RSCU between reference sequence and vaccine sequence. (**F**) For the shared substitution sites of two vaccines, the distribution of delta RSCU between vaccine codon and reference codon is shown
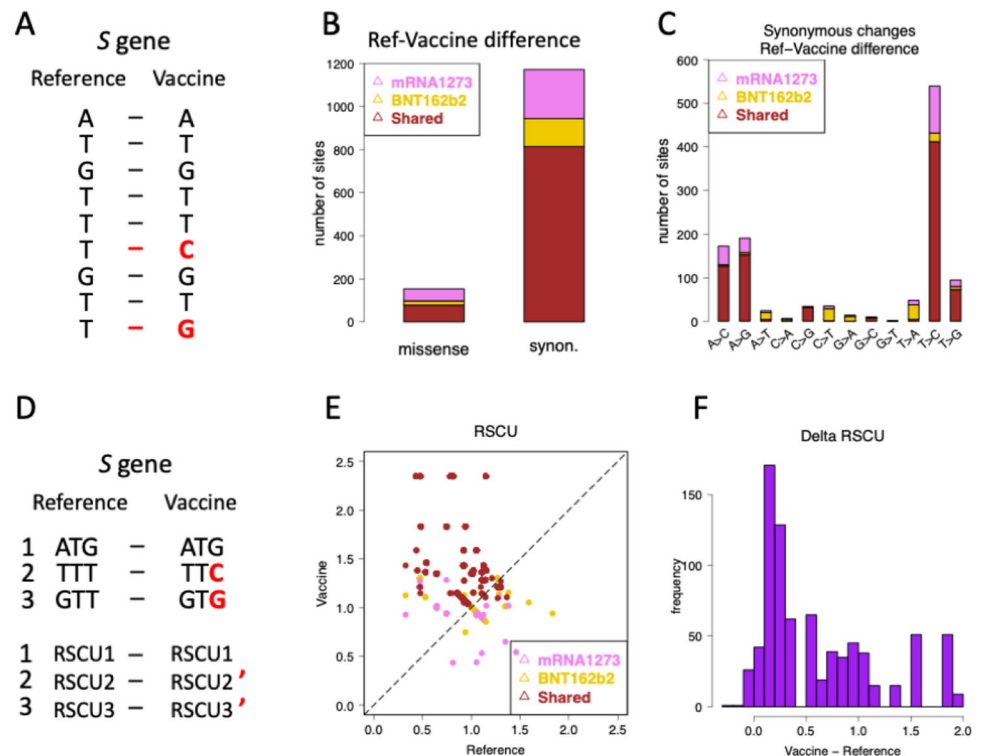
"Reference-Vaccine" differences: which are the nucleotides in vaccine sequences that are different from the reference *S* gene sequence. By looking at these Reference-Vaccine differences, we found that the design of the two vaccines has the following merits.

(1) Most (1171/1325 = 88.4%) of the Reference-Vaccine differences in CDS are synonymous substitutions (Fig. 1B) so that the amino acid sequence of spike protein is largely maintained. Only a few necessary amino acid changes are made

(2) Most (1115/1171 = 95.2%) of the synonymous substitutions increase or maintain the GC content (Fig. 1C). That is to say, when one calculates the relative synonymous codon usage (RSCU) of the reference codon in *S* gene and the codon in vaccines (Fig. 1D), one would find that the codons in vaccine sequences are optimized compared to the reference codons (Fig. 1E and F). This ensures the high translation efficiency and protein productivity of vaccine RNAs

Potential concerns caused by cellular RNA editing

However, despite the smart design of the vaccine sequences, there are potential concerns that might complicate the efficacy and safety of both vaccines. The concern comes from a ubiquitous molecular mechanism called RNA editing or RNA deamination that alters nucleotide sequence at RNA level. Although the accuracy of identifying RNA editing sites in SARS-CoV-2 remains a challenging task (Di Giorgio et al. 2020; Martignano et al. 2022; Zong et al. 2022), it is almost widely accepted that RNA editing indeed exists in the SARS-CoV-2 sequences (Liu et al. 2022; Picardi et al. 2021; Simmonds 2020; Song et al. 2022). There is even a notion that "nothing in SARS-CoV-2 makes sense except in the light of RNA deamination" (Zhao et al. 2022) because the RNA editing mechanism participated in every aspect of SARS-CoV-2 including sequence divergence and population genetics (Li et al. 2020a, b; Zhang et al. 2021). As a speculative support to the RNA editing in SARS-CoV-2, we reproduced a result that the majority of the mutations in millions of world-wide SARS-CoV-2 sequences (retrieved from literature (Zhu et al. 2022)) are C > T sites (Fig. 2A), representing the host-mediated C-to-U RNA editing. The C-to-U editing is inevitable in host cells so that the vaccine RNAs like BNT162b2 and mRNA1273 might also be subjected to this modification when transported into cells. Notably, as most of the substitutions in vaccine design are T > C sites (Fig. 1C); then, the rampant C-to-U editing might reverse the Cs to Us, making the design futile.

Although this undesired situation (the C-to-U editing on the designed Cs in vaccine RNAs) has not been directly observed, this possibility could be reflected from another angle. In the millions of world-wide SARS-CoV-2 sequences, if one looks at the allele frequency (AF) on synonymous sites, it turns out that the A > C AF is significantly
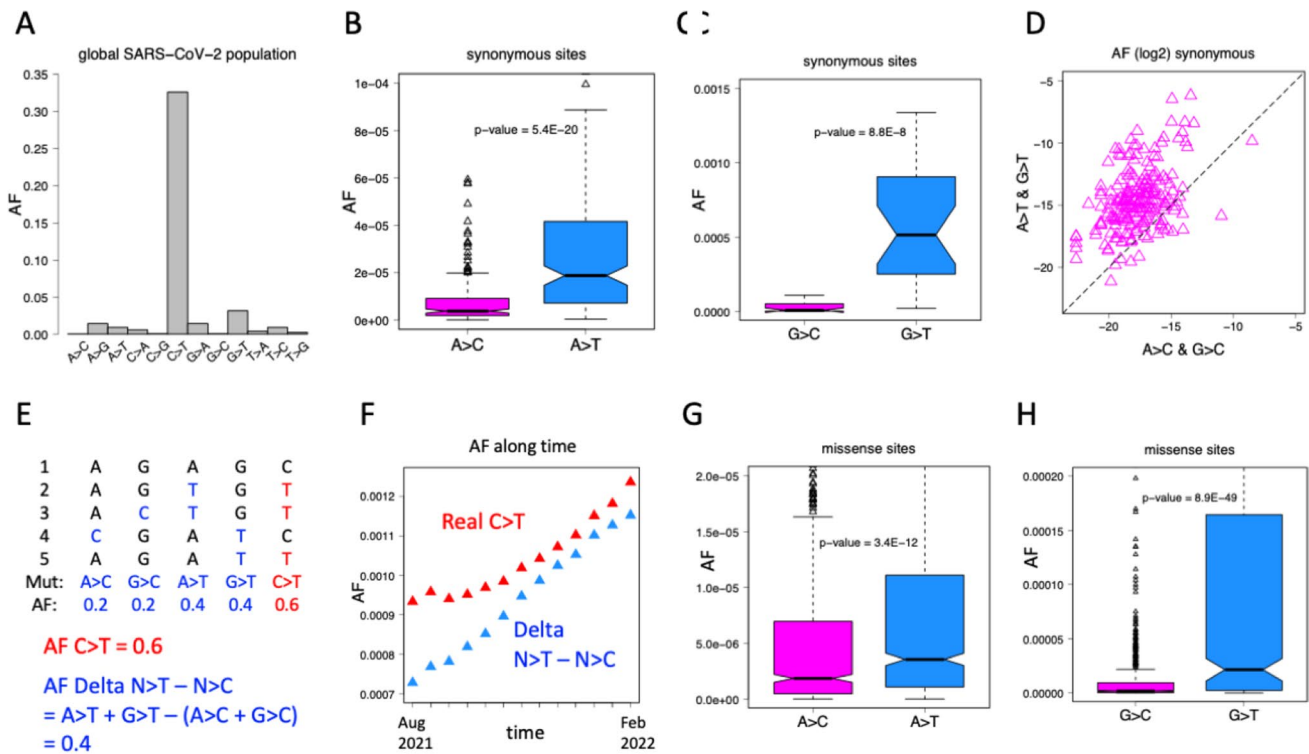
Fig. 2 Potential concerns on the two RNA vaccines. (**A**) Cumulated allele frequency (AF) of polymorphic sites in millions of worldwide SARS-CoV-2 strains. (**B**) Comparison between AF of A > C and A > T synonymous mutations. *P* value from Wilcoxon rank sum test. (**C**) Comparison between AF of G > C and G > T synonymous mutations. *P* value from Wilcoxon rank sum test. (**D**) Point-to-point demonstration of the AF of synonymous mutations in previous pan-els. (**E**) Illustrating the definition of "delta N > T – N > C" allele frequency. (**F**) The dynamic change of C > T allele frequency (red) and the "delta N > T – N > C" allele frequency (blue) from August 2021 to February 2022. (**G**) Comparison between AF of A > C and A > T missense mutations. *P* value from Wilcoxon rank sum test. (**H**) Comparison between AF of G > C and G > T missense mutations. *P* value from Wilcoxon rank sum test

lower than the A > T AF (Fig. 2B), and similarly, the G > C AF is significantly lower than the G > T AF (Fig. 2C). The natural selection acting on synonymous mutations is much weaker than the selection force acting on missense mutations so that the AF of synonymous sites would more faithfully mirror the mutation rate. A > C and A > T are both transversions and should have similar occurrence rates. A plausible explanation for the lower AF in A > C than A > T is that the Cs is converted to Ts via C-to-U RNA editing. The same goes for the difference between G > C and G > T (Fig. 2D).

Moreover, the mutation spectrum is dynamic. At a particular time-point, one could calculate the "delta N > T – N > C" allele frequency to represent the discrepant AF between A > T + G > T mutations and A > C + G > C mutations (Fig. 2E). If C-to-U RNA editing converts Cs to Us at these A > C or G > C sites (like secondary hit), then the "delta N > T – N > C" value should be correlated with the observed AF of C > T mutations (which represents the strength of C-to-U editing). Strikingly, from August 2021 to February 2022, the "delta N > T – N > C" is highly correlated with the AF of C > T mutations (Fig. 2F). This suggests

that C-to-U editing might have converted Cs to Us on the already-mutated A > C and G > C sites.

Potential C-to-U editing on missense mutation sites raises concerns on safety problems

Notably, we try to distinguish the two terminologies "efficacy" and "safety". "Efficacy" is commonly more related to the gene expression ability or RNA translation efficiency (and therefore related to codon usage). However, "safety" should be connected to qualitative changes caused by missense mutations. Thus, we also looked at missense mutations in SARS-CoV-2. AF of A > C sites was significantly lower than AF of A > T sites (Fig. 2G); AF of G > C sites was significantly lower than AF of G > T sites (Fig. 2H). These patterns were similar to what we observed for synonymous sites (although the AF of missense sites was globally lower than AF of synonymous sites). Again, it is compelling evidence suggesting that the Cs in the SARS-CoV-2 sequences have been extensively converted to Ts (Us), presumably by APOBEC-mediated C-to-U RNA editing.

Discussion

We have listed two major facts: (1) the main design of BNT162b2 and mRNA1273 RNA vaccine is to change many Ts (Us) to Cs at synonymous sites to facilitate RNA translation, (2) on the other hand, C-to-U RNA editing by APOBECs is rampant and inevitable for the SARS-CoV-2 RNAs in host cells. Based on these two facts, it is intuitive to concern that the designed T > C changes in vaccine RNAs might be reversed by C-to-U RNA editing.

The "restorative effect" of C-to-U editing not only makes the vaccine design futile but also incurs a series of concerns on the safety and efficacy of RNA vaccines. When BNT162b2 or mRNA1273 were transported into host cells, the sites subjected to C-to-U editing are random and unpredictable. Stronger C-to-U editing restores more T > C sites, weakening the effect of the vaccines. If the protein produced from "intact vaccine RNA sequence" is fully secure and efficient, then the final safety and efficacy of vaccines depends on the APOBEC activity in an individual or a particular cell. This largely complicates our estimation on the effectiveness of vaccines. It is possible that the variable responses of different individuals vaccinated by mRNA1273 (Baden et al. 2021) and BNT162b2 (Polack et al. 2020) are partially due to the difference in APOBEC activity and the extent of C-to-U RNA editing.

Indeed, the key statement of our article remains speculative as we have not yet known the frequency of C-to-U editing in vaccinated cells. However, cellular (endogenous) mRNA editing by APOBECs, especially APOBEC3A, has been, so far, reported by several studies (Sharma and Baysal 2017; Sharma et al. 2017, 2015). The C-to-U editing levels in human cells mostly ranged between 5 and 40% (Sharma et al. 2015). If this editing percentage roughly reflects the editing status of vaccine RNAs, then the efficacy of vaccines might be largely reduced. Fortunately, not every C in an RNA is editable. For example, APOBEC3A typically edits a C in the CAUC* tetra-loop flanked by several palindromic nucleotides (Sharma et al. 2015). This fact might reduce the damage of RNA editing to the vaccine RNAs. Nevertheless, as long as in vivo C-to-U RNA editing exists, its potential effect on SARS-CoV-2 vaccines should not be ignored.

In summary, the efficacy and safety of SARS-CoV-2 vaccines mRNA1273 and BNT162b2 might be complicated by rampant C-to-U RNA editing. This naturally occurring RNA modification is inevitable and should not be ignored when evaluating the RNA vaccines.

Materials and methods

Data acquisition

The vaccine sequences were downloaded from the following link: https://github.com/NAalytics/Assemblies-of-putat ive-SARS-CoV2-spike-encoding-mRNA-sequences-for-vaccines-BNT-162b2-and-mRNA-1273/blob/main/Figur e1Figure2_032321.fastaThe time-course SARS-CoV-2 population data were retrieved from literature (Zhu et al. 2022).

Statistical analysis

All statistics and graphical works were done in the R environment (version 3.6.3).

Declarations

Ethics approval This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest The authors declare no competing interests.

References

Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, Diemert D, Spector SA, Rouphael N, Creech CB et al (2021) Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. N Engl J Med 384:403–416

Chu D, Wei L (2019a) Characterizing the heat response of Arabidopsis thaliana from the perspective of codon usage bias and translational regulation. J Plant Physiol 240:153012

Chu D, Wei L (2019b) The chloroplast and mitochondrial C-to-U RNA editing in Arabidopsis thaliana shows signals of adaptation. Plant Direct 3:e00169

Chu D, Wei L (2021) Direct in vivo observation of the effect of codon usage bias on gene expression in Arabidopsis hybrids. J Plant Physiol 265:153490

Di Giorgio S, Martignano F, Torcia MG, Mattiuz G, Conticello SG (2020) Evidence for host-dependent RNA editing in the transcriptome of SARS-CoV-2. Sci Adv 6:eabb5813

Duan Y, Cai W, Li H (2023) Chloroplast C-to-U RNA editing in vascular plants is adaptive due to its restorative effect: testing the restorative hypothesis. RNA 29:141–152

Harris RS, Dudley JP (2015) APOBECs and virus restriction. Virol 479–480:131–145

Li Y, Yang X, Wang N, Wang H, Yin B, Yang X, Jiang W (2020a) GC usage of SARS-CoV-2 genes might adapt to the environment of human lung expressed genes. Mol Genet Genomics 295:1537–1546

Li Y, Yang XN, Wang N, Wang HY, Yin B, Yang XP, Jiang WQ (2020b) The divergence between SARS-CoV-2 and RaTG13 might be overestimated due to the extensive RNA modification. Future Virol 15:341–347

Li Q, Li J, Yu CP, Chang S, Xie LL, Wang S (2021) Synonymous mutations that regulate translation speed might play a non-negligible role in liver cancer development. BMC Cancer 21:388

Liu X, Liu X, Zhou J, Dong Y, Jiang W, Jiang W (2022) Rampant C-to-U deamination accounts for the intrinsically high mutation rate in SARS-CoV-2 spike gene. RNA 28:917–926

Martignano F, Di Giorgio S, Mattiuz G, Conticello SG (2022) Commentary on "Poor evidence for host-dependent regular RNA

editing in the transcriptome of SARS-CoV-2." J Appl Genet 63:423–428

Picardi E, Mansi L, Pesole G (2021) Detection of A-to-I RNA editing in SARS-COV-2. Genes (Basel) 13 no. 1: 41. https://doi.org/10.3390/genes13010041

Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, Perez JL, Perez Marc G, Moreira ED, Zerbini C et al (2020) Safety and efficacy of the BNT162b2 mRNA COVID-19 vaccine. N Engl J Med 383:2603–2615

Rajasekhar VK, Mulligan RM (1993) RNA editing in plant-mitochondria - alpha-phosphate is retained during C-to-U conversion in messenger-RNAs. Plant Cell 5:1843–1852

Sharma S, Baysal BE (2017) Stem-loop structure preference for site-specific RNA editing by APOBEC3A and APOBEC3G. PeerJ 5:e4136

Sharma S, Patnaik SK, Taggart RT, Kannisto ED, Enriquez SM, Gollnick P, Baysal BE (2015) APOBEC3A cytidine deaminase induces RNA editing in monocytes and macrophages. Nat Commun 6:6881

Sharma S, Patnaik SK, Kemer Z, Baysal BE (2017) Transient overexpression of exogenous APOBEC3A causes C-to-U RNA editing of thousands of genes. RNA Biol 14:603–610

Simmonds, P. (2020). Rampant C-- > U hypermutation in the genomes of SARS-CoV-2 and other coronaviruses: causes and consequences for their short- and long-term evolutionary trajectories. mSphere 5(3). https://doi.org/10.1128/mSphere.00408-20

Song Y, He X, Yang W, Wu Y, Cui J, Tang T, Zhang R (2022) Virus-specific editing identification approach reveals the landscape of A-to-I editing and its impacts on SARS-CoV-2 characteristics and evolution. Nucleic Acids Res 50:2509–2521

Yu YY, Li Y, Dong Y, Wang XK, Li CX, Jiang WQ (2021) Natural selection on synonymous mutations in SARS-CoV-2 and the impact on estimating divergence time. Future Virol 16:447–450

Zhang YP, Jiang W, Li Y, Jin XJ, Yang XP, Zhang PR, Jiang WQ, Yin B (2021) Fast evolution of SARS-CoV-2 driven by deamination systems in hosts. Future Virol 16:587–590

Zhao M, Li C, Dong Y, Wang X, Jiang W, Chen Y (2022) Nothing in SARS-CoV-2 makes sense except in the light of RNA modification? Future Virol 17(11). https://doi.org/10.2217/fvl-2022-0043

Zhou Z, Dang Y, Zhou M, Li L, Yu CH, Fu J, Chen S, Liu Y (2016) Codon usage is an important determinant of gene expression levels largely through its effects on transcription. Proc Natl Acad Sci U S A 113:E6117–E6125

Zhu L, Wang Q, Zhang W, Hu H, Xu K (2022) Evidence for selection on SARS-CoV-2 RNA translation revealed by the evolutionary dynamics of mutations in UTRs and CDSs. RNA Biol 19:866–876

Zong J, Zhang Y, Guo F, Wang C, Li H, Lin G, Jiang W, Song X, Zhang X, Huang F et al (2022) Poor evidence for host-dependent regular RNA editing in the transcriptome of SARS-CoV-2. J Appl Genet 63:413–421