Published in final edited form as:

Microbiol Res. 2022 November ; 264: 127174. doi:10.1016/j.micres.2022.127174.

RNA editing in *Mycobacterium tuberculosis*

Yan Guo[a], **Yuri F. van der Heijden**[b,c,d], **Fernanda Maruri**[b,c], **Limin Jiang**[a], **Robert Morrison**[e], **Timothy R. Sterling**[b,c,*]

[a]Department of Internal Medicine, University of New Mexico, Albuquerque, NM, USA

[b]Division of Infectious Diseases, Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN, USA

[c]Vanderbilt Tuberculosis Center, Vanderbilt University School of Medicine, Nashville, TN, USA

[d]The Aurum Institute, Johannesburg, South Africa

[e]Seattle Children's Hospital, Seattle, WA, USA

Abstract

RNA editing, while studied thoroughly in humans, has been sporadically described in bacteria, and to our knowledge, has not been reported in Mycobacterium tuberculosis (Mtb). After thorough quality control and validation by repeated sequencing, by comparing sequences from RNA and DNA from high-throughput sequencing data, we report the first finding of three RNA editing events in two Mtb isolates.

Keywords

M. tuberculosis ; RNA editing

RNA editing, while studied thoroughly in humans, has been sporadically described in bacteria, and to our knowledge, has not been reported in *Mycobacterium tuberculosis* (Mtb). After thorough quality control and validation by repeated sequencing, by comparing sequences from RNA and DNA from high-throughput sequencing data, we report the first finding of three RNA editing events in two Mtb isolates.

RNA editing is one of the most studied post-transcriptional phenomena in eukaryotes. Especially in humans, millions of RNA editing sites have been described (Mansi et al., 2021). RNA editing events occur when the blueprints stored in DNA are altered at the RNA level, after transcription. Two canonical RNA editing processes have been identified in humans thus far: adenosine (A) to inosine (I) (A-to-I, aka A-to-G) by deamination, and cytosine to uracil (C-to-U aka C-to-T) by apolipoprotein B mRNA-editing enzyme, catalytic polypeptide (APOBEC) enzymes (Guo et al., 2019). Non-canonical RNA editing

*Corresponding author at: Division of Infectious Diseases, Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN, USA. timothy.sterling@vumc.org (T.R. Sterling).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.micres.2022.127174.

processes in humans have been described (Li et al., 2011; Guo et al., 2017), but their validity remains in debate (Piskol et al., 2013). Multiple studies have shown that RNA editing can affect cellular processes and impact diseases such as cancer by producing alternative protein sequences, altering drug sensitivity, and affecting cancer prognosis (Peng et al., 2018; Han et al., 2015; Wu et al., 2021). In humans, RNA editing events have been shown to cause important disruption in the binding sequences of transcription factors (TF), RNA binding proteins (RBP), and micro RNA seeds, leading to reprogrammed regulatory cascades (Jiang et al., 2020).

The study sample included 35 fluoroquinolone-resistant Mtb isolates from a population-based study of all newly diagnosed, culture-confirmed tuberculosis (TB) patients reported to the Tennessee Department of Health from January 2002 to December 2016 (Maruri et al., 2021). The study was approved by the Vanderbilt Institutional Review Board (IRB#050208). Mtb isolates were cultured with and without ofloxacin. Whole genome sequencing was performed on all 35 isolates after being grown in culture without ofloxacin on Illumina Hiseq 3000. RNA extracted from these 35 isolates cultured under both conditions was sequenced with 75 base pair single end sequencing on an Illumina NextSeq 500.

Mtb-specific quality control measures were conducted in two ways. First, lineage calls (Coll et al., 2014) and overall allele call profiles of all DNA and RNA datasets were compared to exclude isolates showing any sign of potential sample labelling issues. Second, allele counts at all variant sites per isolate were evaluated for evidence of multiplicity of infections (MOI). Polyclonal isolates typically have many variant call sites where the relative proportions of 2 + alleles vary in a consistent manner over many locations in the genome. Isolates with evidence of such multi-allelic sites were considered to be polyclonal, and excluded from analysis.

Due to the conserved nature of portions of many bacterial genomes, and the possibility that a TB patient could have secondary bacterial infections, we surmised that coinfection might create alignment aritfacts similar to evidence of RNA editing. Therefore, we gathered five bacterial genomes (*Variovorax paradoxus*, *Bacillis subtilis*, *Escherichia coli*, *Staphylococcus aureus*, *Dermacoccus nishinomiyaensis*) common in humans, converted them to synthetic Illumina reads, and aligned them to the Mtb reference genome to identify genomic regions in Mtb that could receive spurious alignments from other bacteria. These included ribosomal RNA genes, molecular chaperones, and isocitrate dehydrogenase. Any possible RNA editing events identified in genes that received spurious bacterial alignments were discarded as being possible false positives. After thorough quality control, three of the 35 isolates were deemed to be clonal and of high quality for RNA editing detection.

RNA editing events were identified by comparing allele counts between DNA and RNA sequences with base quality PHRED score greater than 20 as the threshold. Three RNA editing events were identified (Table 1, Supplementary Table 1). The first RNA editing event occurred at position 841,326, at which whole genome sequencing showed that the DNA was dominated by allele T with 99% of the reads, and one 1% of the reads supported alternative allele C. In both RNAseq of RNA with and without ofloxacin, the percentage of the allele C increased from 1% in DNA to 66% in RNA with ofloxacin and 64% in RNA without

ofloxacin. For position 3471,001, the allele A increased from 5% in DNA to 70% in RNA with ofloxacin and 66% in RNA without ofloxacin. The last RNA editing event identified was at position 3,501,679 and was characterized by an increase of allele A from 0% in DNA to 64% in RNA with ofloxacin and 80% without ofloxacin.

To date, there have been no reports of RNA editing in Mtb, the bacterium that causes TB. According to the World Health Organization, 1.4 million people died of TB in 2020, and approximately 25% of the global population has latent Mtb infection (Houben and Dodd, 2016; WHO, 2021). Approximately 5–10% of persons with latent Mtb infection progress to active TB disease. It has been shown previously that single nucleotide variants in Mtb DNA gyrase can cause fluoroquinolone resistance in Mtb (Maruri et al., 2021). Fluoroquinolones are a commonly used drug class, and may be used to treat TB (Dorman et al., 2021). However, not all phenotypically fluoroquinolone-resistant Mtb strains have canonical resistance mutations identified, which suggests the presence of alternative resistance mechanisms, including efflux pumps (Maruri et al., 2021; Chien et al., 2017). RNA editing is a type of single nucleotide variant. As demonstrated previously in both humans and bacteria, RNA editing has potential functional effects, which may affect drug resistance (Han et al., 2015). Of note, none of the three genes affected by RNA editing in our study had evidence of canonical DNA gyrase mutations associated with fluoroquinolone resistance on whole genome sequencing. The one nonsynonymous RNA edit we identified was in Rv3135, which is in the proline-proline-glutamate (PPE) family, PPE50. Although PPE genes are highly polymorphic and therefore often not evaluated as a mechanism of drug resistance, a mutation in PPE51 has been associated with drug resistance in Mtb (Wang et al., 2020). Rv3102c is a component of an ABC transporter, which could be functionally relevant, via drug efflux (Kapopoulou et al., 2011). Although the functional correlate of the three RNA edits identified is unclear, a better understanding of RNA editing in Mtb may provide insights into mechanisms of drug resistance.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank David Sherman and Tige Rustad for their contributions to the gene expression work, including performance of RNA-seq on *M. tuberculosis* study isolates. We also wish to thank the Tennessee Public Health Laboratory for the clinical *M. tuberculosis* isolates, and the Tennessee Department of Health for their support of this collaboration. This work was supported by the following grants from the National Institute of Allergy and Infectious Diseases: R56 AI118361, R01 AI063200, K24 065298, and K08 AI106420. YG was supported by P30CA118100 from NIH/NCI of USA.

Data Availability

Data will be made available on request.

References

Chien JY, Yu CJ, Hsueh PR, 2017. High incidence of fluoroquinolone resistance and effect of efflux pump inhibitors on moxifloxacin resistance among Mycobacterium tuberculosis isolates causing urinary tract infection in Taiwan. Int. J. Antimicrob. Agents 50 (3), 491–495. [PubMed: 28668684]

Coll F, et al. , 2014. A robust SNP barcode for typing Mycobacterium tuberculosis complex strains. Nat. Commun. 5, 4812. [PubMed: 25176035]

Dorman SE, et al. , 2021. Four-month rifapentine regimens with or without moxifloxacin for tuberculosis. New Engl. J. Med. 384 (18), 1705–1718. [PubMed: 33951360]

Guo Y, et al. , 2017. The discrepancy among single nucleotide variants detected by DNA and RNA high throughput sequencing data. BMC Genom. 18.

Guo Y, et al. , 2019. Single-nucleotide variants in human RNA: RNA editing and beyond. Brief. Funct. Genom. 18 (1), 30–39.

Han L, et al. , 2015. The genomic landscape and clinical relevance of A-to-I RNA editing in human cancers. Cancer Cell 28 (4), 515–528. [PubMed: 26439496]

Houben RMGJ, Dodd PJ, 2016. The global burden of latent tuberculosis infection: A Re-estimation using mathematical modelling. Plos Med. 13 (10).

Jiang L, et al. , 2020. SMDB: pivotal somatic sequence alterations reprogramming regulatory cascades. NAR Cancer 2 (4).

Kapopoulou A, Lew JM, Cole ST, 2011. The MycoBrowser portal: a comprehensive and manually annotated resource for mycobacterial genomes. Tuberculosis 91 (1), 8–13. [PubMed: 20980200]

Li M, et al. , 2011. Widespread RNA and DNA sequence differences in the human transcriptome. Science 333 (6038), 53–58. [PubMed: 21596952]

Mansi L, et al. , 2021. REDIportal: millions of novel A-to-I RNA editing events from thousands of RNAseq experiments. Nucleic Acids Res 49 (D1), D1012–D1019. [PubMed: 33104797]

Maruri F, et al. , 2021. Resistance-conferring mutations on whole-genome sequencing of fluoroquinolone-resistant and -susceptible mycobacterium tuberculosis isolates: a proposed threshold for identifying resistance. Clin. Infect. Dis. 72 (11), 1910–1918. [PubMed: 32348473]

Peng X, et al. , 2018. A-to-I RNA editing contributes to proteomic diversity in cancer. Cancer Cell 33 (5), 817–828 e7. [PubMed: 29706454]

Piskol R, et al. , 2013. Lack of evidence for existence of noncanonical RNA editing. Nat. Biotechnol. 31 (1), 19–20. [PubMed: 23302925]

Wang QL, et al. , 2020. PE/PPE proteins mediate nutrient transport across the outer membrane of Mycobacterium tuberculosis. Science 367 (6482), 1147 (-+). [PubMed: 32139546]

WHO. Tuberculosis. 2021; Available from: 〈https://www.who.int/en/news-room/fact-sheets/detail/tuberculosis〉 .

Wu YM, et al. , 2021. RNA editing affects cis-regulatory elements and predicts adverse cancer survival. Cancer Med. 10 (17), 6114–6127. [PubMed: 34319007]

Table 1

The three RNA editing events that were identified.

The number of filtered reads for each allele are provided.

				DNA				RNA with ofloxacin				RNA without ofloxacin			
Isolate	**Position**	**Gene**	**Editing type**	**A**	**C**	**T**	**G**	**A**	**C**	**T**	**G**	**A**	**C**	**T**	**G**
Isolate1	841,326	Rv0749	Synonymous	0	1	79	0	0	52	26	0	0	59	32	0
Isolate1	347,1001	Rv3102c	Synonymous	3	0	1	57	17	0	0	7	8	0	0	4
Isolate2	350,1679	Rv3135	Nonsynonymous	0	0	0	20	171	0	0	95	198	0	0	51