

主論文の要旨

**Cellular APOBEC3A deaminase drives mutations in
the SARS-CoV-2 genome**

〔 APOBEC3AがSARS-CoV-2ゲノムへ変異を導入する 〕

名古屋大学大学院医学系研究科 総合医学専攻
免疫不全統御学講座 免疫不全統御学分野

(指導：岩谷 靖雅 教授)

中田 佳宏

【緒言】

新型コロナウイルス感染症 (COVID-19: coronavirus diseases 2019) の原因である SARS-CoV-2 (Severe acute respiratory syndrome coronavirus 2) は 2019 年末に中国武漢市で初めて感染者が見つかり、その後急速に全世界へ感染が拡大した。SARS-CoV-2 はウイルスがコードする nsp14 (non-structural protein 14) がもつ exoribonuclease 活性によって他の RNA ウイルスと比較してゲノムの変異率を低く保っている。しかし、感染伝播の過程で徐々にゲノム配列に変異が蓄積し、新たなバリエーションが出現している。

一方、SARS-CoV-2 の感染拡大初期 (2020 年の前半) のウイルスゲノム配列を用いた分子疫学的研究により、シトシン (C) からウラシル (U) への変異 (C>U 変異) が他の変異と比べ多く生じていることが報告された。C>U 変異箇所の配列の特徴から、ウイルスゲノムへの変異導入にシチジン脱アミノ化酵素群である APOBEC ファミリータンパク質 (以下 APOBEC) が関与しているのではないかと予想された。しかし、直接的なウイルス学的検証はいまだなされておらず、APOBEC が SARS-CoV-2 ゲノム RNA へ変異導入できるか否かについては不明のままであった。そこで本研究では、SARS-CoV-2 ゲノムの変異に対する各種 APOBEC の影響を明らかにすることを目的に実験的検証を行った。

【方法】

2022 年 7 月 24 日に GISAID (Global Initiative on Sharing All Influenza Data) EpiCoV database から取得した SARS-CoV-2 のゲノム全長配列 (>29,000 塩基 [nt]) から、同じ遺伝子型が 2 本以上登録され、伝播していることが確からしい 2,051,393 配列を抽出した。加えて EpiCoV より、それら配列の検体採取日等のメタデータもあわせて取得した。これら配列を Wuhan 株の参照配列と比較し、変異の種類と座位数を解析した。

試験管内での APOBEC の影響を解析するために、ヒト Angiotensin-converting enzyme 2 (ACE2) 及び Transmembrane protease, serine 2 (TMPRSS2) が恒常的に発現したヒト胎児腎細胞由来 293T (293AT) 細胞に、7 種の APOBEC3 (A3) および、Activation-Induced Deaminase (AID)、APOBEC1 (A1) をそれぞれトランスフェクション法により導入した。過剰発現させた後、SARS-CoV-2 を感染させた。24 時間後にウイルスを回収し、ウイルスゲノム RNA 配列を Illumina MiSeq (Next generation sequencing: NGS) により解析した。誘導されたサンプル内存在頻度の変異を検出した。既報の情報から、1% の頻度を閾値とした。次に、SARS-CoV-2 の感染の標的となる気道、鼻腔ならびに肺由来の上皮系培養細胞および初代培養細胞に対して各種サイトカインによって刺激した際の A3 の mRNA 発現量を reverse transcription-droplet digital PCR (RT-ddPCR) 法により定量した。内在性の A3 によるウイルスのゲノム配列への影響については、ウイルス感染前及び感染後に A3 を誘導し、産生されたウイルスのゲノム配列を NGS により解析した。さらに、CRISPR/Cas9 によるゲノム編集で樹立した A3 Knock Out (KO) 細胞についても親株と同様の解析を行い、両者を比較した。

【結果】

2019年12月末から2022年6月末までの2.5年間に採取された検体のSARS-CoV-2ゲノムRNA配列をEpiCoVより取得し、それらを対象にバイオインフォマティクスによる変異解析を実施した。検体採取日と変異数の回帰線から、変異は総じて26.9 nt per year (nt/yr)で蓄積してきたことが示された。変異の中でも、APOBECの関与が想定されるC>U変異は10.2 nt/yrであり、他の変異と比較して、3.3倍以上多くかつ断続的に蓄積していたことが確認された(Figure 1)。このことから、APOBECのいずれかが繰り返しSARS-CoV-2ゲノムRNAに変異を導入してきた可能性が示唆された。

次に、各種APOBECを過剰発現させた293AT細胞に、SARS-CoV-2を感染させ、ウイルスゲノムRNAに特徴的なC>U変異が上昇するのを実験的に検証した。解析したAPOBECの中で、A3Aを発現した293ATから産生されたSARS-CoV-2のゲノム配列のみにAPOBEC3 signatureであるC>U変異が有意に多く認められた(Figure 2A)。一方、酵素活性欠失型であるA3AE72QおよびAID、A1ならびに他のA3ファミリーを発現した293ATから産生されたウイルスではC>U変異の蓄積は認められなかった。これらの結果から、A3AによりSARS-CoV-2ゲノムRNAに変異が導入されることが示唆された。また、293AT細胞にA3Aを過剰発現させた際に認めたウイルスゲノムRNA中のC>U変異のホットスポットがRNAの二次構造上に特徴ある分布をしているのか解析した。既報でselective 2'-hydroxyl acylation analysed by primer extension (SHAPE)法によって決定されたSARS-CoV-2のRNAの二次構造情報にマップした結果、hairpin loopあるいはbulgeといった一本鎖領域のウラシル・シトシン(UC) dinucleotide配列にホットスポットがあることがわかった(Figure 2B)。これらホットスポットの特徴は、A3Aが好んでDNAに変異導入するsignatureと同じであり、A3AがRNAでもDNAと同様のsignatureを好んで変異導入することが示唆された。

ここで、インターフェロン誘導遺伝子(Interferon [IFN] stimulated genes; ISGs)のひとつであるA3Aが呼吸器系上皮細胞で発現しているかを調べた。Calu-3をはじめとする各種呼吸器系の上皮細胞株や初代培養細胞に対して、I型IFNであるIFN- β と腫瘍壊死因子(Tumor Necrosis Factor- α : TNF- α)によって同時刺激するとA3AのmRNA発現レベルが約30~1,120倍上昇し、低酸素環境下ではさらに相加的な発現レベルの上昇効果が認められた(Figure 3)。このことから、呼吸器系の細胞において、ウイルス感染時などのIFN誘導時にA3Aが存在しうることが示唆された。さらに、Calu-3細胞の内在性のA3AによってC>U変異が誘導されるかを調べた。IFN- β +TNF- α で刺激したCalu-3細胞にSARS-CoV-2を感染させ、産生されたウイルスのゲノム配列を観察した。その結果、C>U変異の蓄積が認められたが、ウイルス感染後にIFN- β +TNF- α で刺激した場合にはC>U変異の蓄積は認められなかった(Figure 4A)。さらに、A3A KO Calu-3細胞ではウイルス感染前にIFN- β +TNF- α で刺激しても、親株で認められたC>U変異の蓄積は観察されなかった(Figure 4B)。これらのことから、呼吸器系組織においてA3Aが誘導され、感染したSARS-CoV-2のゲノムRNAにC>U変異が蓄積しうることが示唆された。

【考察】

SARS-CoV-2 の感染拡大初期の分子疫学的配列解析から C>U という特徴的な変異の蓄積が認められるという報告により APOBEC ファミリータンパク質の関与が予想されていた。しかし、直接的なウイルス学的検証はされていなかった。本研究では、(1)2022 年に至っても、断続的な C>U 変異の蓄積がウイルスゲノム RNA に生じていること、(2)APOBEC の中でも、A3A が SARS-CoV-2 ゲノム RNA に対して特に強い C>U 変異導入能があり、hairpin loop 上の UC>UU が好んで生じること、(3)I 型 IFN などの炎症系の刺激により、呼吸器系上皮細胞にて A3A が誘導され、誘導後に感染した SARS-CoV-2 のゲノム RNA に C>U 変異を導入できること、を明らかにした。

A3A を過剰発現させた 293AT 細胞において認めた C>U 変異のパターンと、SARS-CoV-2 の分子疫学的配列情報で観察された各 C>U 変異の頻度は完全に一致するものではなかった。この不一致に関しては、宿主における免疫や伝播の過程の中での選択が、おそらく深く関わっているものと考えられる。

【結語】

本研究によって、呼吸器系上皮細胞では炎症系の刺激により A3A が誘導され、宿主由来の A3A による RNA editing により SARS-CoV-2 ゲノムの C>U 変異が増加し、SARS-CoV-2 の遺伝的多様性の拡大に寄与していることが明らかとなった。

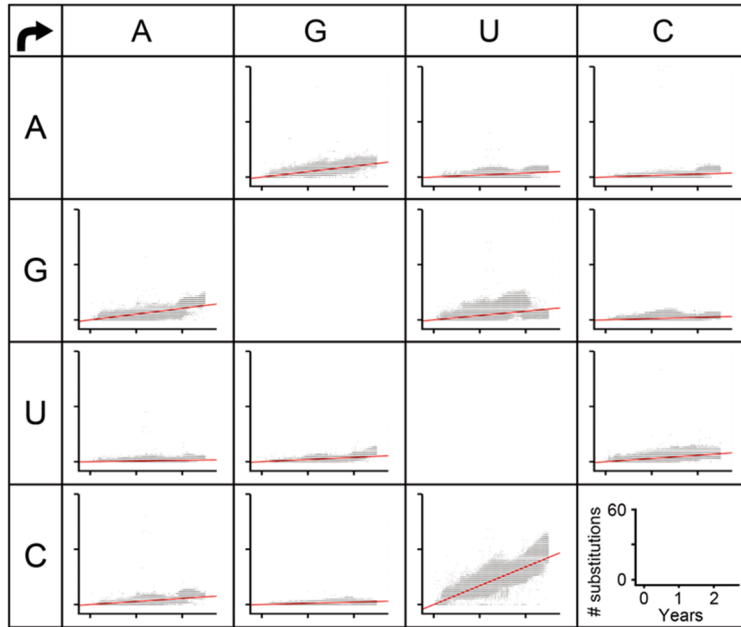


Figure 1. Chronological nucleotide changes observed in SARS-CoV-2 genome sequences in the GISAID database. Chronological changes in the numbers of mononucleotide substitutions per genome were analysed for each SARS-CoV-2 sequence ($n = 2,051,393$) relative to the Wuhan reference strain. The number was plotted against the sample collection date corresponding to the sequence in the GISAID database. The red lines represent linear regression lines. The years “0”, “1” and “2” on the x-axis represent 2020, 2021 and 2022, respectively.

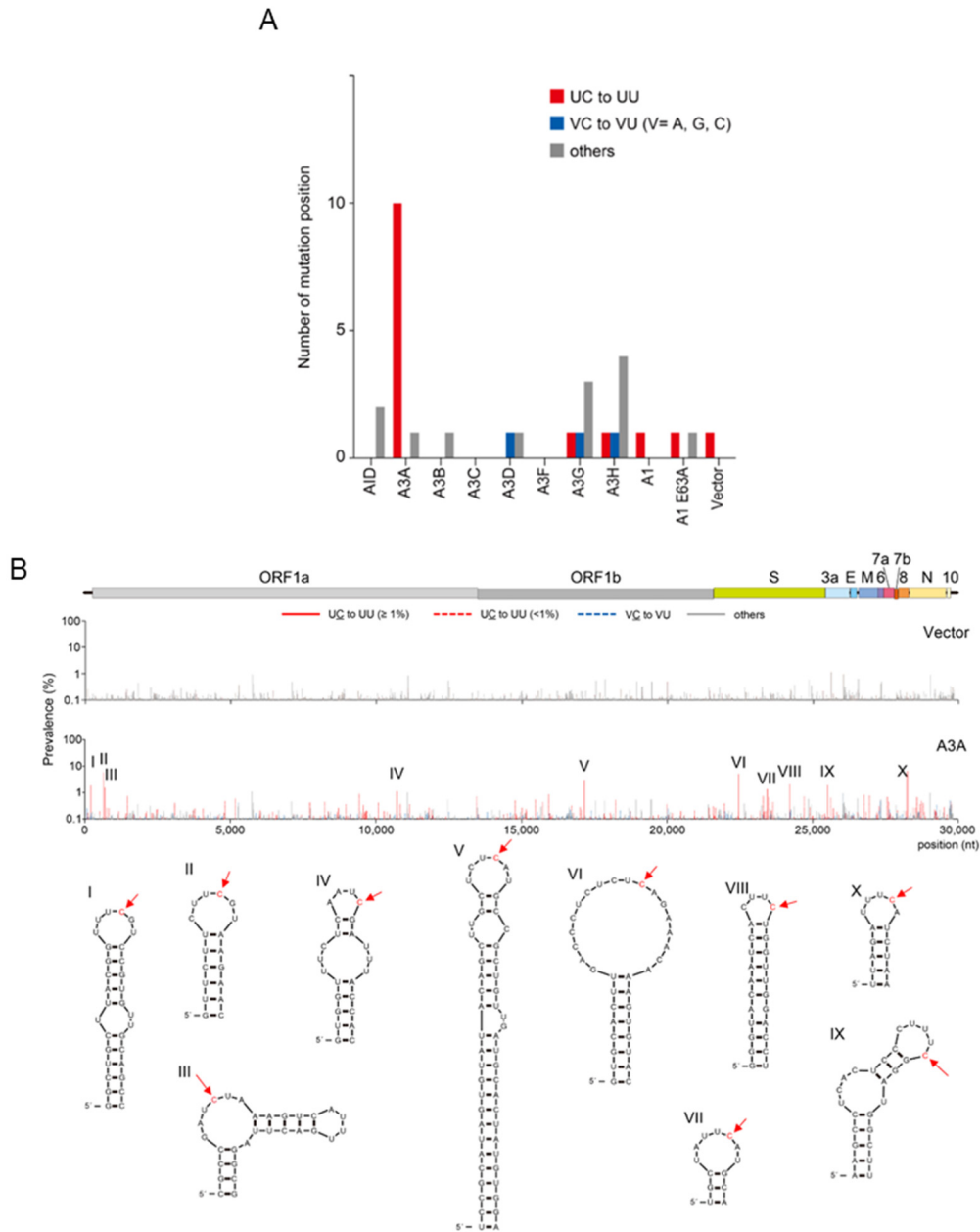


Figure 2. Exogenous A3A expression induces C-to-U mutations in the SARS-CoV-2 viral RNA (vRNA) genome, preferentially in single-stranded RNA (ssRNA) regions. (A) The number of positions with C-to-U substitutions in the viral genome was determined. AID, A3 (A-G, or H haplotype II [hapII]), A1 and A1 E63A proteins were transiently expressed in 293AT cells. Empty vector (Vector) was used as the negative control for the absence of deaminase expression. Thirty-six hours after transfection, the cells were infected with SARS-CoV-2 (B.1.1) (MOI = 0.5) and incubated for another 24 h. Viral genomes were sequenced on the Illumina MiSeq system. The graph shows the number of positions with C-to-U mutations in the dinucleotide context (UC-to-UU, VC-to-VU (V= not U), and others), with a prevalence of $\geq 1.0\%$ throughout the viral genome. The data are representative of three independent experiments. (B) The prevalence of mutations detected at each position in the viral genome is indicated in two bar graphs, along with a schematic diagram of the viral genomic structure. The upper and lower graphs represent the mutation prevalence (%) in the viruses produced from 293AT cells transfected with empty vector (Vector) or the A3A plasmid (A3A), respectively. The ten major positions of A3A-induced editing are labelled I-X. Secondary RNA structures around the edited positions (red arrows) were extracted from the results of a previous SHAPE study and are drawn in this figure.

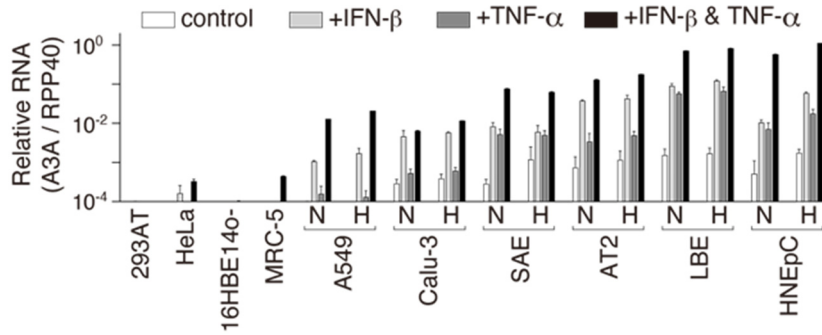


Figure 3. A3A mRNA expression in human airway and lung cells. The A3A mRNA level in each cell line was quantified by RT-ddPCR 18 h after treatment without (control) or with IFN- β and/or TNF- α under normoxic conditions (N). A549, Calu-3, SAE, AT2, and LBE cells and HNEpCs were also cultured under hypoxic conditions (H). The mean mRNA copy numbers of A3A relative to those of the housekeeping gene RPP40 are shown (n=3).

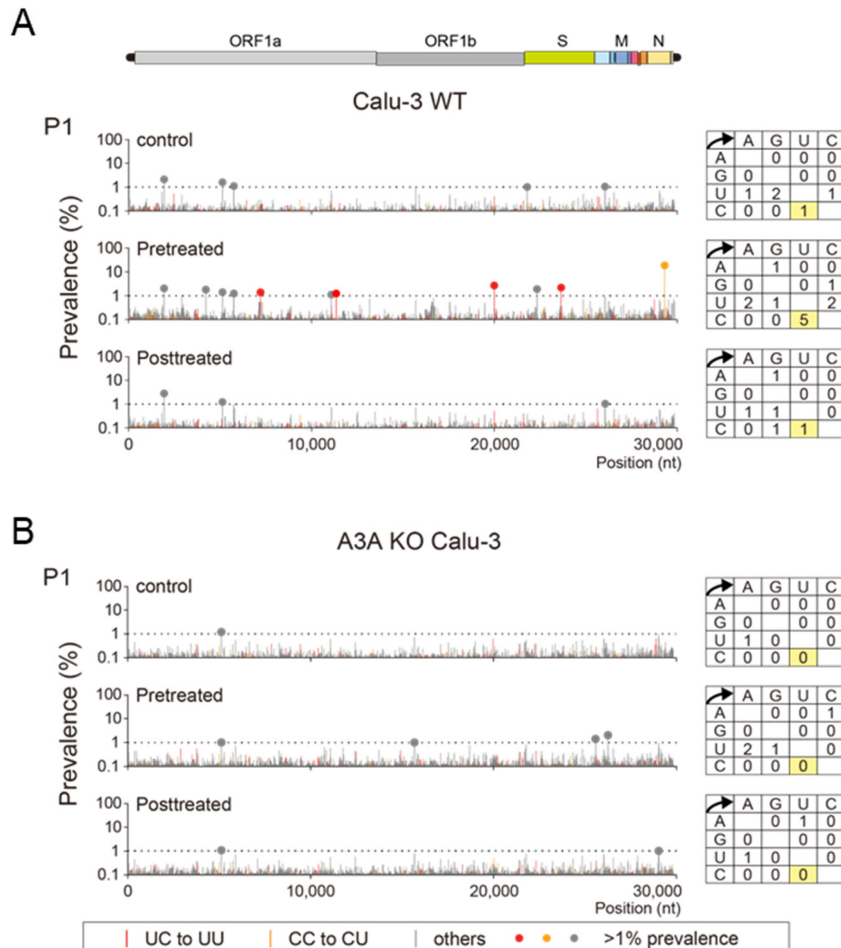


Figure 4. A3A expression induced by IFN- β and TNF- α in Calu-3 cells increases UC-to-UU mutations in the SARS-CoV-2 genome. (A and B) Effects of IFN- β and TNF- α treatment on C-to-U mutations in the SARS-CoV-2 genome were analysed using Calu-3 cells and A3A-KO Calu-3 cells (clone #15). The cells were left untreated (control) or pretreated (Pretreated) with IFN- β and TNF- α . Seventy-two hours post-infection with SARS-CoV-2 (B.1.1), the culture supernatants were harvested. The infected cells were treated with IFN- β and TNF- α (Posttreated), and the supernatants were collected. The position and prevalence of the mutations in the genome at passage (P1) were analysed and are shown in bar graphs. The closed circles indicate mutations with a prevalence of $\geq 1\%$.