Phase separation-mediated condensation of N protein-deaminase complex promotes SARS-CoV-2 mutagenesis

Xingxu Huang (huangxx@shanghaitech.edu.cn)
Zhejiang University School of Medicine https://orcid.org/0000-0001-8934-1247
Zhean Li
Zhejiang University School of Medicine
Lingling Luo
ShanghaiTech University
Xiaohui Ju
Tsinghua University
Shisheng Huang
ShanghaiTech University
Jia Liu
ShanghaiTech University https://orcid.org/0000-0001-9787-465X
Pumin Zhang
Zhejiang University School of Medicine
Tian Chi
ShanghaiTech University
Peixiang Ma
Shanghai Jiao Tong University School of Medicine
Cheng Huang
Shanghai University of Traditional Chinese Medicine
Qiang Ding
Tsinghua University
Yu Zhang
ShanghaiTech University

Article

Keywords: SARS-CoV-2, Nucleocapsid protein, RNA-protein condensates, Mutagenesis, Deaminases

Posted Date: April 27th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1566705/v1

Phase separation-mediated condensation of N protein-deaminase complex promotes SARS-CoV-2 mutagenesis

Zhean Li[1,6,9], Lingling Luo[2,9], Xiaohui Ju[3,9], Shisheng Huang[4,9], Jia Liu[5], Pumin Zhang[1], Tian Chi[4], Peixiang Ma[7,8], Cheng Huang[3], Xingxu Huang[1,4], Qiang Ding[3], Yu Zhang[1,4]

[1]Zhejiang Provincial Key Laboratory of Pancreatic Disease, The First Affiliated Hospital, and Institute of Translational Medicine, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

[2]School of Pharmacy, Shanghai University of Traditional Chinese Medicine, Shanghai, China.

[3]Center for Infectious Disease Research, School of Medicine, Tsinghua University, Beijing, China

[4]School of Life Science and Technology, ShanghaiTech University, Shanghai, China

[5]Shanghai Institute for Advanced Immunochemical Studies, ShanghaiTech University, Shanghai, China.

[6]Department of Urology & Andrology, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China.

[7]Shanghai Key Laboratory of Orthopedic Implants, Department of Orthopedic Surgery, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China;

[8]Guangzhou Laboratory, Guangzhou International Bio Island, Guangzhou, Guangdong, China;

[9]These authors contributed equally

[*]Correspondence: zhangy@shanghaitech.edu.cn; qding@tsinghua.edu.cn; huangxx@shanghaitech.edu.cn; chuang@shutcm.edu.cn.

Abstract

The emergence of SARS-CoV-2 variants poses enormous challenges to the prevention and control of COVID-19 with alterations in antigenicity, transmissibility and pathogenicity. The rapid evolution of RNA viruses could be caused by high mutation frequencies during replication, arising by replication errors, intergenomic recombination or even host deaminases. We sought to understand whether host deaminases are involved in SARS-CoV-2 mutation, and how they orchestrate host deaminases to trigger this process. Herein, we provided the experimental evidence that APOBEC and ADAR deaminases act as the driving forces for SARS-CoV-2 mutagenesis. Mechanistically, SARS-CoV-2 nucleocapsid (N) protein, which is responsible for packaging viral genomic RNA, complexes with host deaminases to facilitate viral RNA mutation. Moreover, N protein employs deaminases-involved condensates to further promote viral RNA mutation. Mutant N protein with F17A substitution, defective in entry of deaminases-involved RNA granules, leads to the decreased mutation of viral RNA, confirming the function of N protein-deaminase condensates on RNA editing. Our study sheds light on the novel mechanism of SARS-CoV-2 mutation during host-virus arms race.

Keywords: SARS-CoV-2, Nucleocapsid protein, RNA-protein condensates, Mutagenesis, Deaminases

Introduction

The continuous emergence of SARS-CoV-2 variants of concerns (VOC) with alterations in transmissibility, antigenicity and pathogenicity in human has posed an ongoing threat to global health[1-4]. Currently, the World Health Organization (WHO) has defined five variants as VOC, Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2) and Omicron (B.1.1.529), as well as two further variants of interest (VOI), Lambda (C.37) and Mu (B.1.621) and a new recombinant variant, XE[5, 6]. The VOCs and VOIs have exhibited abilities to evade neutralization by monoclonal antibodies or vaccine induced immunity, which presents urgent challenges for curbing COVID-19[2, 7-10].

SARS-CoV-2, belonging to subfamily Coronavirinae in the family Coronaviridae of the order Nidovirales, is a positive-sense, single-strand (ss) RNA virus and possesses a moderate genetic variability due to their RNA-proofreading machinery which utilizes nsp14 exonuclease to correct errors[11, 12]. In general, viral RNA mutations are emerged from two main sources which contain spontaneous random errors in replication made by RNA dependent RNA polymerase (RdRP) that are not corrected by the existing proofreading mechanism, and host-driven viral genome mutation, such as deaminase-mediated RNA editing and reactive oxygen species (ROS)[13-15].

The human deaminases, which consist of the adenosine deaminases acting on RNA (ADARs) and the apolipoprotein-B (*ApoB*) mRNA editing enzyme, catalytic polypeptide-like proteins (APOBECs)[16], play a critical role in the innate antiviral defense. The ADARs are able to deaminate adenosine in dsRNAs and convert it into inosine (A-to-I). Three types of ADAR genes (ADAR1, ADAR2 and ADAR3) have been identified in human genome. ADAR1 and ADAR2 are expressed in almost all human tissues and catalytically active for adenosine deamination, while ADAR3 is absent outside of the brain and has no reported deamination activity[17, 18]. ADAR1/2 possess antiviral effects by destabilizing dsRNA through introducing multiple A-to-I (G) substitutions, which were described for several RNA viruses, including influenza virus, lymphocytic choriomeningitis virus, rift valley fever virus, hepatitis C virus and

Zika virus[19, 20]. The APOBECs are a family of cytosine deaminases, which catalyze cytosine deamination to uracil (C-to-U) in single-stranded DNA or RNA. The APOBECs family contains eleven APOBEC members: APOBEC1, APOBEC2, APOBEC3 (A, B, C, D, F, G and H), APOBEC4 and activation-induced cytidine deaminase (AID)[21]. Similar to the ADARs, APOBEC proteins restrict viral replication due to their feature of hyper-editing, such as HIV and HBV[22]. APOBEC1, APOBEC3A and APOBEC3G are also reported to recognize and deaminate ssRNA to cause C-to-U mutation[23-26]. The ROS, also as an antiviral strategy, could oxidize nucleic acids to induce virus inactivation and mutagenesis, which is proposed to be related to the G-to U and C-to-A transversions[27, 28]. In the host-virus interaction, the host cells utilize the innate antiviral response to induce virus mutagenesis and inactivation, and thereby combat virus infection. On the contrary, the viruses develop multiple mechanisms to employ host-driven forces to achieve viral mutagenesis and evolution, and thereby enhances virus adaption and fitness.

The analysis of existing SARS-CoV-2 mutations and SARS-CoV-2 RNA-seq data from COVID-19 patients displayed predominant mutational patterns with specific signatures of A-to-I (G) and C-to-U (T) mutation, indicating that ADARs and APOBECs-mediated RNA editing likely plays a crucial role in SARS-CoV-2 mutagenesis[13, 14]. However, whether host deaminases are involved in SARS-CoV-2 mutation and how SARS-CoV-2 hijacks host deaminases to achieve viral mutation and evolution are still unknown. In our study, we provided the experimental evidence that host deaminases, ADARs and APOBECs, indeed act as the driving forces for RNA mutation of SARS-CoV-2. The potential mechanism is that host deaminases specifically interacted with N protein (a viral RNA binding protein in viral life cycle), which may link deaminases to target viral RNA for editing. Furthermore, N protein partitions into deaminases-involved RNA granule via phase separation, which concentrates RNA and proteins to enhance deaminases interaction with N protein and viral RNA, and thereby promotes RNA editing efficiency. Furthermore, N^{F17A} protein loses the capability to entry into stress granules (SGs), which impairs the interaction level with deaminase and decreases the

SARS-CoV-2 RNA mutagenesis, further supporting the function of deaminases-involved condensates on RNA editing. Our study may shed light on a novel therapeutic strategy against SARS-CoV-2 variants.

Results

N protein interacts with host deaminases and then guides host deaminases to edit viral RNA

The predominant mutational patterns in SARS-CoV-2 genome with the signatures of A-to-I (G) and C-to-U (T) mutation implies that APOBEC and ADAR-mediated RNA editing may contribute to shape the viral RNA mutational outcome and promote SARS-CoV-2 evolution[13, 16]. To confirm the role of the deaminases in viral mutation, we utilized an N protein-based trans-complementation SARS-CoV-2 cell culture model in Caco-2 cells[29], and performed RNA-seq to analyze mutational patterns of SARS-CoV-2 genome. With the reference sequence of SARS-CoV-2 genome (NC_045512), we analyzed the proportion of 12 single nucleotides variant (SNV) types based on four nucleotides (A: A>C, A>G, A>T; C: C>A, C>G, C>T; G: G>A, G>C, G>T; T: T>A, T>C, T>G). Consistent with the reported mutational pattern[13, 14], APOBEC-mediated C-to-U (C>T/G>A) and ADAR-mediated A-to-I (A>G/T>C) account for the predominant proportion in mutational ratio (Fig. 1a; Supplementary Fig. S1a), suggesting that the deaminase-induced RNA editing potentially occurred in SARS-CoV-2 RNA genome. However, SARS-CoV-2 infection did not alter the mRNA expression of APOBECs and ADARs (Supplementary Fig. S1b), suggesting that viral mutation is not induced by the change of deaminase level. The G-to-U (G>T) and C-to-A (C>A) transversions are the second main group for nucleotides C and G mutation (Fig. 1a; Supplementary Fig. S1a), which was likely contributed by reactive oxygen species (ROS) activity[27]. Since APOBECs or ADARs-catalyzed deamination of C-to-U or A-to-U is regulated by specific sequence contexts, we performed the analysis of the nucleotide contexts in the SARS-CoV-2 genome. We observed that the A<u>A</u> and G<u>A</u> are the most and the least preferable motifs for A-to-I mutation on the SARS-CoV-2 genome, respectively (Fig. 1b), which is similar to the context preference of hADARs[30,

[31]. On the other hand, C-to-U transitions on the SARS-CoV-2 genome preferentially occur at the downstream of adenosines or uridines, and the upstream of adenosines or uridines (Fig. 1b), which are in line with the reported motif preference for RNA editing by APOBECs[24, 32]. Notably, each nucleotides variant is evenly represented with respect to SNV frequency suggesting the strength of driving forces to induce SARS-CoV-2 mutation presents less difference (Fig. 1c).

To gain further insight into the role of host deaminase in SARS-CoV-2 mutation, we performed a series of virus passages on Caco-2 cells and collected the SARS-CoV-2 RNA after one and six passages to perform RNA-sequencing. At the first passage, the extensive C-to-U and A-to-I substitutions were displayed by the SARS-CoV-2 genome suggested that host deaminases induced viral genome hypermutation, the purpose of which is to mutate the viral RNA, and then restrict the viral RNA replication in the early stages of SARS-CoV-2 infection (Fig. 1d). Strikingly, the count of SNVs number dramatically decreased at the sixth passage of SARS-CoV-2 (Fig. 1d). One possible reason is that most of the mutations lead to the deficiency of viral RNA replication, and only mutations, i.e. T2578C, C5219T, A19736G and C27970T, which do not impair viral fitness could be maintained in the following passages (Supplementary Fig. S1c), further supporting that host deaminase indeed function as host restriction factors to induce viral genome hypermutation and inactivation upon SARS-CoV-2 infection. Conversely, the viruses could employ cellular deaminase to edit SARS-CoV-2 viral genome and achieve viral evolution, and thereby enhance virus fitness.

The function of host deaminases in driving SARS-CoV-2 viral genome mutations have been well characterized, but it is not yet known how SARS-CoV-2 virus hijacks host deaminases to deaminate viral RNA in the viral life cycle. SARS-CoV-2 nucleocapsid (N) protein, which is responsible for assembling viral RNA genome to form ribonucleoprotein (RNP) complex could interact host RNA-binding proteins and regulated SG formation to overcome host antiviral response[33, 34]. The deaminases (ADARs and APOBECs) also function as host RNA-binding proteins and favor RNA to occur deamination[17, 23]. In light of their function, we hypothesize that N protein act

as a host-virus intermediary to direct host-encoded cytidine/adenine deaminases to edit the viral RNA. To test this hypothesis, we first investigated whether N protein interacted with host deaminases. We systematically co-expressed GFP-tagged N protein and flag-tagged APOBEC family members (APOBEC1, activation-induced cytidine deaminase (AID), APOBEC2, APOBEC3 (A to F), and APOBEC4) or ADAR family members (ADAR1 and ADAR2) in HeLa cells, and then performed Co-IP experiment to detect the interaction of GFP-tagged N protein with each host deaminase. The results showed that N protein interacted with ADARs (ADAR1 and ADAR2) and some members of APOBECs, such as APOBEC1, AID, APOBEC3D, APOBEC3F, APOBEC3G and APOBEC3H, but not APOBEC2, APOBEC3A, APOBEC3C and APOBEC4 (Fig. 1e). Meanwhile, we transfected N gene in HeLa cells and then performed RNA immunoprecipitation (RIP) assay using anti-ADAR2 antibody to detect the interaction between ADAR2 and viral RNA. The result showed that ADAR2 bound the RNA of N protein, suggesting N protein directly bridges the host deaminases with viral RNA (Fig. 1f). Next, we validated whether host deaminases could induce mutations on the RNA which encodes viral proteins. We transfected viral N gene into HeLa cells and performed RNA-seq to analyze mutational signature of N gene RNA. Similar to mutational pattern of SARS-CoV-2 genome, the mutation bias toward transitions, including A>G/T>C and C>T/G>A, further confirming the host deaminase function (Fig. 1g). Moreover, A<u>A</u>G and G<u>C</u>A are the most preferable motifs for A-to-I mutation and C-to-U mutation, respectively (Supplementary Fig. S2a), which are consistent with the sequence context of N gene in the SARS-CoV-2 genome (Supplementary Fig. S2b). The similar results were observed on the RNA of spike gene in HeLa cells when spike gene and N gene were co-transfected into HeLa cells (Supplementary Fig. S2c). The SARS-CoV-2 genome encodes some non-structural proteins (nsp) to assist the viral RNA processing and replication, including nsp7, nsp8, nsp9, nsp12 and nsp16. To check the possibility of other viral factors interaction with host deaminases, we chose the aforementioned nsp and four structural proteins (S, N, E, M) to perform Co-IP assay with host deaminases. The results showed that APOBEC3G (A3G) or ADAR2 failed to interact with those nsp and structural proteins. However, N protein showed a strong

interaction with A3G or ADAR2 (Fig. 1h; Supplementary Fig. S2d). Thus, specific interaction between N protein with host deaminases is a potential mechanism for host-directed RNA editing of the SARS-CoV-2 genome.

The condensate formation of N-deaminase complex efficiently enhances the viral RNA mutagenesis

Our previously study have reported that N protein interacted with host RNA-binding proteins G3BP (the core proteins in SGs)[33], which provoked us to explore whether N protein interacts with G3BP and host deaminases to form multi-proteins complex. To this end, we co-expressed GFP-tagged N protein and HA-tagged A3G or ADAR2 in HeLa cells and then the cell lysate was immunoprecipitated by the antibody against GFP, followed by the detection of endogenous G3BP1 and A3G or ADAR2 using the antibody against G3BP1 and HA. N protein is able to interact with host G3BP1 and A3G or ADAR2 (Fig. 2a), suggesting that G3BP1 and host deaminases act as host binding partners of viral N protein and may be jointly involved in stress granules (SGs). SGs are the membrane-less dynamic structures containing large cytoplasmic mRNA-protein aggregates[35, 36]. Viral infection induces host SG formation which functions as a barrier to insulate the virus via inhibiting translation of viral mRNA[37-40]. SARS-CoV-2 RNA is able to induce SG formation[41], and the N protein co-localized in SGs[33]. We thus verified whether host deaminases co-localized with N protein in SGs. The cells were transfected with the plasmids encoding mCherry-tagged N protein and BFP-tagged deaminases AID, A3G, APOBEC3H or ADAR2, and then immunostained by antibodies against endogenous G3BP1 (SG marker). Under the stress condition, N protein specifically co-localized with G3BP1 and AID, APOBEC3G, APOBEC3H or ADAR2 in SGs, instead of other SARS-CoV-2 related proteins (Fig. 2b; Supplementary Fig. S3a-S3c). ADAR2 and APOBEC3H mainly localized to the nucleus in the resting conditions, however, they partially translocated to the cytoplasm and localized with N protein in SGs under the stress condition (Fig. 2b; Supplementary Fig. S3b). SGs, one type of RNA granule which is formed by phase separation, is able to concentrate RNA and RNA-binging proteins into condensed compartments[42, 43]. To examine whether

such compact organization can promote the interaction between deaminases and N protein or RNA. We transfected cDNA encoding N protein in HeLa cells and performed Co-IP and RIP assay. As our excepted, when N protein were recruited into SGs, the formation of RNA-protein condensates enhanced the interaction level between host deaminase and N protein or RNA (Fig. 2c and 2d). Phase-separated condensates is able to accelerate their related reaction by concentrating moleculars[44]. Thus, we next investigated whether the condensed RNA-deaminases interactions could boost the RNA-editing activity of host deaminases. For this purpose, we co-expressed N, spike or Orf 8 in HeLa cells and then treated cells with stress stimulator to induce SG formation. APOBEC-regulated C-to-U (C>T/G>A) and ADAR-mediated A-to-I (A>G/T>C) mutations still accounted for the predominant proration of mutational signature under stress condition (Supplementary Fig. S4), suggesting SG formation is unable to change deaminases function. Interestingly, the formation of N protein-deaminases condensates remarkably increased the number of SNVs in the transcripts of N, spike and Orf 8 (Fig. 2e-2g). Moreover, the frequency of SNVs was greatly upregulated (Fig. 2h and 2i), suggesting the condensation enhanced the RNA-deaminases interaction, and thereby promoted RNA mutagenesis. Surprisingly, the frequency of two VOC (variant of concern) mutants, A23403G (D614G in spike protein) and C27972T (Q27* in Orf 8) were significantly increased due to the condensate formation of N protein-deaminases (Fig. 2j and 2k), strongly suggesting that C-to-U (T) and A-to-I (G) mutations by APOBECs and ADARs may contribute to the increased fitness for SARS-CoV-2. These results demonstrated that N protein penetrated into deaminase-involved RNA granule to further promote SARS-CoV-2 RNA mutagenesis and evolution.

G3BP1/2 function as the core proteins to assemble SGs and the cells lacking G3BPs fail to form SGs in response to arsenite treatment[36]. To further dissect the role of RNA-deaminases condensates in host deaminase activity, we used CRISPR/Cas9 to generate HeLa cells harboring the double knockout of G3BP1 and G3BP2 (referred as G3BP1/2 dKO, Supplementary Fig. S5a) and transfected N gene to assess the RNA mutagenesis

of N gene. Consistent with the previous study[36], HeLa cells with G3BP1/2 dKO failed to form SGs after arsenite treatment (Supplementary Fig. S5b and S5c), and thereby destroying the condensate formation of N protein-deaminases (Fig.3a). Notably, the disruption of N protein-deaminases condensates impaired the APOBECs and ADARs-mediated RNA-editing efficiency (Fig. 3b and 3c). Thus, these results further revealed that N protein penetrated into deaminases-involved RNA granules to form N protein-deaminases condensates, which enhanced deaminase activity to facilitate RNA mutagenesis and evolution of SARS-CoV-2.

In addition, we similarly used CRISPR/Cas9 to generate HeLa cells with the double knockout of ADAR1 and ADAR2 (referred as ADAR1/2 dKO; Supplementary Fig. S5d) and transfected N gene to investigate the RNA mutagenesis of N gene. Importantly, ADAR1/2 dKO indeed decreased the SNVs number of A>G/T>C transitions (Fig. 3d), supporting our hypothesis that ADARs is involved in SARS-CoV-2 mutagenesis. To exclude the possibility that RNA mutation with A>G/T>C was induced by stress stimulation, we treated ADAR1/2 dKO HeLa cells with stress stimulator. ADAR1/2 dKO is unable to abolish SG formation, and N protein still possess the ability to localize in SGs under the stress condition (Supplementary Fig. S5e). However, the condensates formation failed to recover the A>G/T>C mutation of SNVs on the RNA of N gene, even under the stress condition (Fig. 3e), suggesting that condensates act as a contributor instead of an inducer on ADAR-induced A>G/T>C mutation.

RNA binding domain is required for ADAR2 to form RNA-protein condensates with N protein

ADAR1/2 consist two or more double-stranded RNA binding domains (dsRBDs) and a C-terminal adenosine-deaminase domain[17]. ADAR1 is ubiquitously distributed across several tissues, and ADAR2 is abundantly expressed in the lung, urinary bladder and cerebellum[45]. To identify the contribution of individual domains of ADARs to interaction with N protein and condensate formation, we chose ADAR2 and generated various truncated mutant constructs, including dsRBD1 (RNA binding domain 1, aa78-142), dsRBD2 (RNA binding domain 2, aa230-296), deaminase domain (aa317-710),

294 ΔdsRBD1 (lack of dsRBD1), ΔdsRBD2 (lack of dsRBD2) and ΔDea (lack of
295 deaminase domain) (Fig. 4a). We found dsRBD1 alone showed a high level of
296 interaction with N protein (Fig. 4a). On the contrary, dsRBD2 or deaminase domain
297 alone failed to form protein complex with N protein. Furthermore, deletion of dsRBD1
298 completely abolished the ADAR2 interaction with N protein (Fig. 4a). The mutants of
299 ΔdsRBD2 and ΔDea, which contain dsRBD1, possessed the ability of ADAR2
300 interaction with N protein (Fig. 4a), implying that dsRBD1 is the domain of ADAR2,
301 responsible for forming protein complex with N protein. Next, we assessed the domains
302 of ADAR2 involved in the co-localization with N protein in SGs. Consistent with the
303 result of interaction assay, dsRBD1 co-localized with N protein in SGs and ADAR2
304 lacking dsRBD1 failed to form proteins condensates with N protein (Fig. 4b;
305 Supplementary Fig. S6), illustrating that dsRBD1 is essential for ADAR2 to form the
306 condensates with N protein. SG assembly is regulated by phase separation and N
307 protein phase separated with G3BP1 into SGs[33, 36]. ADAR2-N protein complex
308 involved in SGs inspired us to investigate whether ADAR2 exhibits features of liquid-
309 like condensates. Due to the specific function of ADAR2-dsRBD1 on interaction and
310 SG localization, recombinant GFP-dsRBD1 fusion protein was purified from E. coli.
311 We found, ADAR2-dsRBD1 alone was unable to induce phase separation (Fig. 4c and
312 4d). To test whether RNA binding induces phase separation of ADAR2-dsRBD1 in vitro,
313 similar to other RNA binding protein, we incubated GFP-dsRBD1 protein with RNA
314 extracted from HeLa cells. Upon mixing, ADAR2-dsRBD1 rapidly formed
315 micrometer-sized liquid droplets, accompanied by increased fluorescence intensity and
316 larger droplet area (Fig. 4c and 4d). Fluorescence recovery after photobleaching (FRAP)
317 analysis revealed that GFP-dsRBD1 diffused rapidly within droplets and exhibited
318 liquid-like property (Fig. 4e). Furthermore, N protein partitioned into dsRBD1-RNA
319 droplets and enhanced phase separation (Fig. 4f and 4g). Taken together, these results
320 revealed that N protein-ADAR2 underwent an RNA-dependent phase separation to
321 form RNA-protein condensates.

322 **The disruption of N protein-localized in SGs attenuates the viral RNA mutagenesis**

Our previous study revealed that N protein interaction with G3BPs is prerequisite for N protein-localized in SGs, and the first intrinsically disordered region (IDR1) is required for N protein to interact with G3BP1 and localize in SGs[33]. Furthermore, the F17A mutation in IDR1 of N protein almost abolished the N-G3BP1interaction[46], which implies N protein with F17A mutation (termed as N^{F17A} protein) may fail to form RNA-protein condensates and localize in SGs. To further establish the effect of RNA-deaminase condensates on the viral RNA mutagenesis, we first assessed the ability of N^{F17A} protein localization in SGs. We constructed GFP-tagged N^{F17A} protein and performed Co-IP assays in HeLa cells (Fig. 5a). Consistent with the reported results of GST pull-down assay[46], the F17A mutation significantly impaired the N protein-G3BP1 interaction (Supplementary Fig. S7a), which then significantly weakened the ability of N^{F17A} protein to phase separate with G3BP1 and resulted in mislocalization of N^{F17A} protein in SGs (Fig. 5a-5c; Supplementary Fig. S7b andS7c). We next determined whether the F17A mutation affect the interaction between N protein and host deaminases. Intriguingly, the F17A mutation did not change N protein-ADAR2/A3G interaction (Fig. 5d), suggesting that ITFG motif is a specific region of N protein to interacted with G3BP1, instead of deaminases. However, the lack of N protein in RNA-granules resulted in the failure of N protein to form RNA-protein condensates with ADAR2 (Fig. 5e), which further weakened the interaction between ADAR2 and the RNA of N gene. Next, we examined whether blocking N protein to entry into deaminase-involved RNA granules affects deaminases-mediated the RNA editing activity of N gene. To this end, we transfected N WT or N^{F17A} gene into HeLa cells and then induced SG formation. Remarkably, N protein-ADAR2 complex unable to form condensates severely attenuated the interaction level between host deaminase and the RNA of N gene, and then impaired host deaminase -mediated RNA-editing ability (Fig. 5f and 5g).

Next, we further utilized an N^{F17A}-based transcomplementation SARS-CoV-2 cell culture model to characterize the role of N-deaminase complex-involved RNA condensates in deamination activity on SARS-CoV-2 genome. In agreement with the

mutational results of SARS-CoV-2 based on WT N protein, F17A mutation did not change the mutational signature that APOBEC-mediated C-to-U (C>T/G>A) and ADAR-mediated A-to-I (A>G/T>C) occupy the predominant proportion in mutational ratio (Supplementary Fig. S7d), since N^{F17A} protein possessed the capability to interact with host deaminases. Importantly, F17A mutation in N protein remarkably decreased the ratio of T>C and C>T/G>A changes, and weakened the function of deaminase on viral mutagenesis, which abolished the enrichment ratio of SNVs (A19736G and C27970T) induced by deamination activity (Fig. 5h and 5i). Collectively, these results demonstrated that the condensate formation of RNA-N-deaminase complex is critical for the SARS-CoV-2 mutagenesis.

Discussion

In host-virus interaction, APOBEC and ADAR-mediated RNA editing often function as the host immune response against viral infection [13, 20, 47]. Conversely, viruses exploit host deaminases to drive genetic diversity and achieve viral evolution[48, 49]. Here, we utilized an N protein based transcomplementation SARS-CoV-2 cell culture model to fully characterize the RNA mutational signature of SARS-CoV-2, providing the experimental evidences that host deaminases-mediated RNA editing is able to enhance viral mutation and drive virus evolution. Mechanistically, we demonstrated that N protein acts as a bridge to guide host deaminases to target viral RNA for editing. Furthermore, N protein penetrates into deaminase-involved RNA granule via RNA-dependent phase separation to form N protein-deaminase condensates, an organizational strategy that not only concentrates RNAs and proteins but also promotes deaminases-induced RNA editing efficiency to viral genome. More importantly, N $protein^{F17A}$ fails to entry into deaminases-involved RNA granule, which greatly impairs host deaminases activity to viral RNA, further establishing the role of N protein-deaminases condensates in viral RNA mutation (Fig. 6). Our findings thus highlight a mechanism involving phase separation-regulated N protein-deaminases condensation that enables deaminases to drive viral mutation and evolution.

RNA virus mutations are often arisen from spontaneous random errors during viral

replication and host-driven antiviral defense systems. Coronaviruses have proof-reading machinery in viral replication achieved by nsp14 exonuclease, thus possessing a relatively high fidelity in viral transcription and replication processes[34]. However, the persistent accumulation of new mutations in SARS-CoV-2 genome implies host-driven forces involving in viral genome mutation. Here, we examined SARS-CoV-2 RNA sequence in cell culture and found the A-to-I (G) and C-to-U (T) mutation account for the predominant mutational patterns, which is consistent with database analysis of SARS-CoV-2 genomic variations and RNA sequence analysis from COVID-19 patients[13, 14], suggesting host antiviral response via APOBEC and ADAR-mediated RNA editing is the one of drivers for SARS-CoV-2 genome mutation. Moreover, the number of APOBEC or ADAR-induced SNVs on SARS-CoV-2 genome was significantly decreased after several passages, demonstrating that the deaminases function as host antiviral response to induce the extensive mutation of SARS-CoV-2 RNA in the early viral infection, which may lead to viral RNA degradation and inhibit viral RNA replication. However, SARS-CoV-2 is likely to express functional protein to combat host restriction factors and weaken the deaminases activity, which lead to the reduced the number of SNVs in SARS-CoV-2 genome. The similar process has been previous reported for APOBEC3G, a powerful arm of the host defence against HIV and other retroviruses via bringing “typographical errors” during viral replication[50-52]. On the contrary, HIV-encoded virion infectivity factor (Vif) induces APOBEC3G degradation and protects the virus from APOBCE3G-induced inactivation[52]. However, the frequency of the SNVs in undegraded viral RNA was greatly promoted, suggesting SARS-CoV-2 in turn utilizes host deaminases-mediated RNA editing to fuels its evolution. Another interesting feature is that the transversions of G-to-U (G>T) and C-to-A (C>A) are the second main group for nucleotides C and G mutation, which was possibly associated with the mutagenic activity of ROS[16, 28]. Furthermore, the transversions of A-to-U (A>T) and U-to-A (T>A) on SARS-CoV-2 genomes appear in the second main group for nucleotides A and T change, but the mechanistic basis for this process is currently unknown and still need to be further studied. Thus, the mutational outcome of SARS-CoV-2 genome are controlled by multi-driving forces

411 from host and virus.

412 For RNA-protein interaction, the SARS-CoV-2 N protein, which binds the viral RNA,
413 was reported to be a most significantly enriched viral protein in SARS-CoV-2 RNA
414 purifications[53]. ADARs were identified as the SARS-CoV-2 RNA binder in host cell
415 proteins[53]. For protein-protein interaction, APOBEC3G associated with coronavirus N
416 protein to package into SARS-CoV virus-like particles[54]. A SARS-CoV-2 protein
417 interaction map revealed N protein interacted with several RNA binding proteins[34].
418 Although several clues were shown to be implied in deaminases targeting viral RNA,
419 the precise mechanism remains elusive. In this study, we demonstrated that N protein,
420 but not other viral SARS-CoV-2 RNA binders, interacted with host deaminases
421 containing APOBEC1, AID, APOBEC3D, APOBEC3F, APOBEC3G, APOBEC3H,
422 ADAR1 and ADAR2, suggesting that N protein acts as a specific protein binder for
423 host deaminases and viral RNA, and then link them together. For APOBECs family,
424 only APOBEC1, APOBEC3A and APOBEC3G are shown to possess RNA editing
425 activities[24, 26, 55, 56]. The further study need to be perform to identify the member of host
426 deaminases which indeed target SARS-CoV-2 RNA and edit viral RNA.

427 Membraneless organelles, serve as an alternative approach to concentrate specific
428 proteins and nucleic acids and increase the local concentration, thereby significantly
429 strengthening binding affinity[44, 57-59]. Consistently, we observed that the condensate
430 formation of N protein-deaminases complex enable ADAR2 to bind more RNA of N
431 protein. The biological function of specific molecular-enriched condensates is to amply
432 signals and accelerate the reaction rate[58]. Notably, we showed that N protein penetrates
433 into deaminases-involved RNA granule to form N protein-deaminases condensates,
434 which greatly promoted the RNA mutational rate of N protein. Conversely, N^{F17A} failed
435 to be recruited into deaminases-involved RNA granule, which dramatically reduced the
436 RNA editing efficiency of deaminases. The potential mechanism is that RNA granule,
437 as a condense organization, increases the local concentration of viral RNA and
438 deaminases to enhance deamination activity. These results suggest that phase separation
439 may serve as a regulatory mechanism to control the deamination activity, and provide

a novel organizational strategy to promote viral RNA editing efficiency via concentrating RNAs and deaminases. However, phase separation-mediated condensation often acts as a double edged sword for regulating bimolecular activity. Phase separation forms a selective environment that limit catalytic function of the cytosolic exonuclease TREX1[60]. RNA polymerization are inhibited in the polyamine-containing condensation[61]. On the contrary, polyethylene glycol (PEG)/dextran aqueous two-phase system (ATPS)-mediated phase separation greatly increased RNA catalytic reaction[62]. Thus, understanding the physicochemical features of micro-condensates containing the size, chemical makeup and the structure stability is necessary for predictions of deamination activity.

In summary, we here provide the experimental evidences that host deaminases function as the driving force to induce mutation of SARS-CoV-2 genome. During this process, N protein acts as a bridge to guide deaminases to target viral RNA. Furthermore, virus exploits host RNA granule to promote deaminases-mediated RNA editing activity and fuel its evolution. These insights help to predict viral mutation trend, and also provide new insights into SARS-CoV-2 therapeutic development.

Methods

Plasmid construction. DNA fragments encoding the APOBEC family proteins, ADAR family proteins and SARS-CoV-2 proteins annotated in this manuscript were synthesized by Genscript and amplified by PCR using Phanta® Max Super-Fidelity DNA Polymerase (Vazyme, P505-d1). While the coding region of ADAR2 deletions were amplified by PCR from a plasmid containing full-length ADAR2 with appropriate sets of primers. Exnase (Vazyme, C214-02-AF) was used to insert these sequences into one of the following backbone: CMV-mCherry-Flag, CMV-mCherry-HA, CMV-BFP-Flag, CMV-BFP-HA, CMV-GFP, CMV-Flag. All plasmid inserts were confirmed by BioSune Sanger sequencing.

Cell culture, transfection, and harvest. HeLa cells were cultured in DMEM medium supplemented with 10% bovine growth serum and 1% penicillin/streptomycin at 37 °C in a 5% CO2 humidified incubator. For plasmid transfection, cells were seeded on 6-

well plates and transfected at approximately 70% confluence using Lipofectamin 3000 reagent (Invitrogen, L3000008). 48 h after transfection, cells were collected from Fluorescence Activating Cell Sorter (FACS) according to the fluorescence driven by the transfected plasmid.

Generation of G3BP1/2 or ADAR1/2 knock-out cell lines. HeLa cells were seeded in a 24-well plate and transfected with 300 ng sgRNA and 700 ng spCas9 plasmid using Lipofectamine 3000 transfection kit (Invitrogen, L3000015). 48 h after transfection, cells were sorted in 96-well plates using fluorescence-activated cell sorting (FACS). The genomic DNA of edited cells was extracted using QuickExtract™ DNA Extraction Solution (Lucigen, QE09050), the sequences of interest was amplified by PCR and analyzed by BioSune Sanger sequencing. While individual clones were lysed in RIPA buffer (Beyotime, P0013B) contained with 1 mmol/L phenylmethanesulfonyl fluoride (PMSF) and 1% (v/v) protease inhibitor cocktail (Thermo Fisher scientific, 78443) on ice for 15 min. The lysate was centrifuged at 12,000 rpm for 15 min at 4 °C. The supernatant was harvested and used for SDS-PAGE analyzed with G3BP1 antibody (proteintech, 66486-1-Ig), G3BP2 antibody (proteintech, 16276-1-AP), ADAR1 antibody (proteintech, 14330-1-AP) and ADARB1 antibody (proteintech, 22248-1-AP), respectively. The sgRNAs used are listed in Supplementary Table 1.

Co-immunoprecipitation. Plasmids were transfected into HeLa cells using Lipofectamin 3000 reagent (Invitrogen, L3000008). 48 h after transfection, cells were harvested and lysed with RIPA buffer (Beyotime, P0013B) contained with 1 mmol/L PMSF and 1% (v/v) protease inhibitor cocktail (Thermo Fisher scientific, 78443). The cell lysates were incubated on ice for 15 min and centrifuged at 16,000 × g at 4 °C for 20 min, the supernatant was incubated with pre-equilibrated anti-flag beads (Bimake, B26101) or anti-GFP beads (KT Health, KTSM1333) for 4 h at 4 °C. Then, the beads were washed three times with PBS supplemented with 0.05% (V/V) Tween-20 to remove nonspecific binding proteins. 1 × loading buffer containing SDS were added into each sample and heated at 98 °C for 10 min to elute bound proteins. The supernatant was harvested and used for SDS-PAGE analysis.

RNA immunoprecipitation (RIP) assays. Plasmids encoding the DNA fragments of

N wild type or N^{F17A} mutant were transfected into HeLa cells, the experiments were performed using RIP assay kit (Geneseed Biotech, P0102) according to the manufacturer's instructions with the following antibodies: rabbit anti-ADARB1 (Proteintech, 22248-1-AP) and normal rabbit IgG (Abcam, ab190475). The immunoprecipitated RNA was purified and analyzed by RT-qPCR, sequences for primers are listed in Supplementary Table 2.

Immunofluorescence. HeLa cells were seeded to reach 60% confluence and transfected with plasmids in 24-well glass bottom plates (Cellvis, Mountain View, USA). To induce SGs, the cells were treated with 200 μmol/L sodium arsenite for 45 min. The cells were fixed in 4% paraformaldehyde for 15 min at room temperature and washed three times with PBS. After incubated with blocking buffer (5% (v/v) normal goat serum (Bioss Antibodies, C-0005), 0.3% (v/v) Triton X-100, 94.7% (v/v) PBS) for 2 h at room temperature, cells were incubated with primary antibodies against G3BP1 (Proteintech, 66486-1-Ig; 1:200 dilution), TIA1 (Proteintech, 12133-2-AP; 1:200 dilution) or Flag (Cell Signaling Technology, 14793s; 1:1000 dilution) overnight at 4 °C. The cells were washed three times in PBS and incubated with Alexa Fluor 488/647 conjugated secondary antibodies for 1 h. Cell nuclei were stained with 4,6-diamidino-2-phenylindole (DAPI; Yeasen, 40728ES03). Images were acquired with a Nikon spinning disk microscope.

Western blotting. The same amount of protein samples were separated by SDS-PAGE and transferred to methanol-activated polyvinylidene fluoride (PVDF) membranes (Millipore, IPVH00010). The membranes were blocked for 1 h at room temperature in Tris-buffered saline and 0.1% Tween-20 (TBST) containing 5% (wt/vol) nonfat milk. After that, the membranes were incubated with primary antibodies with the dilution ratio according to manufacturer's instruction at 4 °C overnight. The membranes were washed three times in TBST and the protein bands were detected with horse radish peroxidase-conjugated secondary antibodies and Immobilon Western enhanced chemiluminescent solution (Millipore, WBKLS0100). The protein levels were normalized by probing the same blots with GAPDH antibody (Absin, abs132004; 1:5000 dilution).

SARS-CoV-2 GFP/ΔN trVLP infection. Caco-2 cells expressing WT N protein are infected with SARS-CoV-2 GFP/ΔN P1 trVLP[29] at MOI of 0.1. 24 hours post infection, cells are washed with Dulbecco's Phosphate-Buffered Saline (DPBS) and collected with TRIzol reagent (Invitrogen, 15596018). Uninfected Caco-2 cells expressing WT N protein are mock control.

SARS-CoV-2 GFP/ΔN trVLP passage. SARS-CoV-2 GFP/ΔN P0 trVLPs[29] are used to infect Caco-2 cells expressing WT N protein or F17A mutant N protein at a multiplicity of infection (MOI) of 0.1 for 24 hours, supernatants (P1) are collected to further infect Caco-2 cells expressing WT N protein or F17A mutant N protein. Cell cultures collected from each subsequent passage on the Caco-2-N WT or F17A mutant cells were defined as P1 to P6, respectively. Cells are washed with DPBS and collected with Trizols (Invitrogen, 15596018).

SARS-CoV-2 mutation analysis. Total viral RNA in the supernatant were extracted using TRIzol reagent (Invitrogen, 15596018). Complementary DNA (cDNA) was synthesized from RNA using the HiScript III 1st Strand cDNA synthesis Kit (Vazyme, R312-01). Libraries were generated using the TruePrepTM DNA Library Prep Kit V2 for Illumina (Vazyme, TD503-01) and sequenced by Illumina Hiseq X Ten platform. The RNA-seq data were aligned to the SARS-CoV-2 reference genome (NC_045512) using STAR software (Version 2.5.1).

Plasmids encoding the DNA fragment of SARS-CoV-2 genes were transfected into HeLa cells. 48 h after transfection, cells were treated with 120 μmol/L sodium arsenite to generate stress granules for 10 h. Then, cells were harvested and the RNA was extracted by TRIzol reagent (Invitrogen, 15596018). 1 μg RNA was reversely transcribed into cDNA using Reverse Transcriptase (Vazyme, R223-01-AB). The target sites were amplified with Phanta Max Super-Fidelity DNA Polymerase (Vazyme, P505-d1) and subjected to high-throughput sequencing with the Illumina Hiseq X Ten platform. Sequences for primers are listed in Supplementary Table 3. The preprocessed data were then aligned to the plasmids sequence using bwa software (Version 0.7.17).

All variants from RNA-seq and deep-seq were identified and calculated using bam-readcount with parameters −q 20 −b 30. The SNVs filter is based on a minimum

threshold of 20 reads.

Protein purification. DNA fragments encoding the proteins of interest were cloned into pET-28a vector. Proteins were expressed in E. coli BL21 (DE3) (Trans, CD601). E. coli were induced with 1.0 mmol/L IPTG at 16 °C for 14 h. Cells were ultrasonic broken down centrifuged, the supernatant was collected and added 0.4 μg/mL DNase and RNAase to remove the possible bound DNA and RNA to the protein. The interest protein was purified by Ni resin (TaKaRa, 635660) and AEKTA purifier. The purified protein was stored in 20 mM Tris (pH 8.4), 500 mM NaCl, 10% (v/v) glycerol at -80 °C before further analysis.

Fluorescence recovery after pohotobleaching (FRAP) analysis. FRAP analysis was performed using the Nikon Spinning Disk microscope equipped with two laser systems. 1-2 images were acquired before the center of the indicated protein droplets was bleached with laser at 488 nm. Images were collected every 10 s for 2 min after photobleaching.

Statistical analysis. All statistical analyses were done using OriginPro (2019b, OriginLab, Northampton, USA) or Microsoft Excel (Professional 2019, Microsoft Corporation, Redmond, USA). The outcomes of all statistical tests including number of samples and *P* values were revealed in the corresponding figure legends. Results were presented as mean ± SEM. The significance of *P* values is represented as follows: ns > 0.05, *$P < 0.05$; **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$, ns, no significant.

Acknowledgements

We thank all members of Huang lab for helpful discussions and Dr. Haitao Yang from ShanghaiTech University for insightful discussion; Molecular Imaging Core Facility and Molecular and Cellular Biology Core Facility for support with imaging and cellular biology experiments. This work is supported by the National Key R&D Program of China (2021YFA0804702), the leading talents of Guangdong province program (2016LJ06S386), National Natural Science Foundation of China (22177073), Emergency Key Program of Guangzhou Laboratory (EKPG21-18), National Natural Science Foundation of China (32070153), Tsinghua University Spring Breeze Fund (2021Z99CFY030) and Beijing Municipal Natural Science Foundation (M21001).

Author contributions

Y.Z., X.H., C.H. and Q.D. conceived and designed the entire project. Y.Z., Z.L. and L.L. supervised and performed the bulk of the experiments and analyzed nearly all experimental data. L.L. generated constructs and cell lines. S.H. analyzed next-generation sequencing data. Z.L. and L.L. assisted with imaging experiments. X.J. and Q.D. prepared all viral RNA sample of the SARS-CoV-2 for next-generation sequencing. Y.Z., Z.L. and L.L. wrote the manuscript. All authors contributed idea and reviewed the manuscript.

Data availability

Sequence data were submitted to NCBI and are accessible with BioProject ID: PRJNA824251. The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Tegally, H. *et al.* Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature* **592**, 438-+ (2021).
2. Schmidt, F. *et al.* Plasma Neutralization of the SARS-CoV-2 Omicron Variant. *New Engl J Med* **386**, 599-601 (2022).
3. Karim, S.S.A. & Karim, Q.A. Omicron SARS-CoV-2 variant: a new chapter in the COVID-19 pandemic. *Lancet* **398**, 2126-+ (2021).
4. Muik, A. *et al.* Neutralization of SARS-CoV-2 lineage B.1.1.7 pseudovirus by BNT162b2 vaccine-elicited human sera. *Science* **371**, 1152-+ (2021).
5. Maxmen, A. Omicron Blind Spots: Why It's Hard to Track Coronavirus Variants. *Nature* **600**, 579-579 (2021).
6. WHO. Coronavirus (COVID-19) Dashboard. https://covid19.who.int/. (2022).
7. Dejnirattisai, W. *et al.* Reduced neutralisation of SARS-CoV-2 omicron B.1.1.529 variant by post-immunisation serum. *Lancet* **399**, 234-236 (2022).
8. Collie, S., Champion, J., Moultrie, H., Bekker, L.G. & Gray, G. Effectiveness of BNT162b2 Vaccine against Omicron Variant in South Africa. *New Engl J Med* **386**, 494-496 (2022).
9. Liu, L.H. *et al.* Striking antibody evasion manifested by the Omicron variant of SARS-CoV-2. *Nature* (2021).
10. Hu, J. *et al.* Increased immune escape of the new SARS-CoV-2 variant of concern Omicron. *Cell Mol Immunol* **19**, 293-295 (2022).
11. Yin, W.C. *et al.* Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science* **368**, 1499-+ (2020).

12. Maio, N. *et al.* Fe-S cofactors in the SARS-CoV-2 RNA-dependent RNA polymerase are potential antiviral targets. *Science* **373**, 236-+ (2021).
13. Wang, R., Hozumi, Y., Zheng, Y.H., Yin, C.C. & Wei, G.W. Host Immune Response Driving SARS-CoV-2 Evolution. *Viruses-Basel* **12** (2020).
14. Di Giorgio, S., Martignano, F., Torcia, M.G., Mattiuz, G. & Conticello, S.G. Evidence for host-dependent RNA editing in the transcriptome of SARS-CoV-2. *Science Advances* **6** (2020).
15. Zhang, Y.P. *et al.* Fast evolution of SARS-CoV-2 driven by deamination systems in hosts. *Future Virol* **16**, 587-590 (2021).
16. Mourier, T. *et al.* Host-directed editing of the SARS-CoV-2 genome. *Biochem Bioph Res Co* **538**, 35-39 (2021).
17. Rusk, N. RNA editing with endogenous ADAR. *Nat Methods* **16**, 285-285 (2019).
18. Roth, S.H., Levanon, E.Y. & Eisenberg, E. Genome-wide quantification of ADAR adenosine-to-inosine RNA editing activity. *Nat Methods* **16**, 1131-+ (2019).
19. Piontkivska, H., Frederick, M., Miyamoto, M.M. & Wayne, M.L. RNA editing by the host ADAR system affects the molecular evolution of the Zika virus. *Ecol Evol* **7**, 4475-4485 (2017).
20. Samuel, C.E. Adenosine deaminases acting on RNA (ADARs) are both antiviral and proviral. *Virology* **411**, 180-193 (2011).
21. Salter, J.D., Bennett, R.P. & Smith, H.C. The APOBEC Protein Family: United by Structure, Divergent in Function. *Trends Biochem Sci* **41**, 578-594 (2016).
22. Stavrou, S. & Ross, S.R. APOBEC3 Proteins in Viral Immunity. *J Immunol* **195**, 4565-4570 (2015).
23. Bishop, K.N., Holmes, R.K., Sheehy, A.M. & Malim, M.H. APOBEC-mediated editing of viral RNA. *Science* **305**, 645-645 (2004).
24. Sharma, S. *et al.* APOBEC3A cytidine deaminase induces RNA editing in monocytes and macrophages. *Nature Communications* **6** (2015).
25. Dance, G.S.C., Sowden, M.P., Yang, Y. & Smith, H.C. APOBEC-1 dependent cytidine to uridine editing of apolipoprotein B RNA in yeast. *Nucleic Acids Research* **28**, 424-429 (2000).
26. Sharma, S., Patnaik, S.K., Taggart, R.T. & Baysal, B.E. The double-domain cytidine deaminase APOBEC3G is a cellular site-specific RNA editing enzyme. *Sci Rep-Uk* **6** (2016).
27. Valyi-Nagy, T. & Dermody, T.S. Role of oxidative damage in the pathogenesis of viral infections of the nervous system. *Histol Histopathol* **20**, 957-967 (2005).
28. Smith, E.C. The not-so-infinite malleability of RNA viruses: Viral and cellular determinants of RNA virus mutation rates. *Plos Pathogens* **13** (2017).
29. Ju, X.H. *et al.* A novel cell culture system modeling the SARS-CoV-2 life cycle. *Plos Pathogens* **17** (2021).
30. Eggington, J.M., Greene, T. & Bass, B.L. Predicting sites of ADAR editing in double-stranded RNA. *Nature Communications* **2** (2011).
31. Bass, B.L. RNA editing by adenosine deaminases that act on RNA. *Annu Rev Biochem* **71**, 817-846 (2002).
32. Rosenberg, B.R., Hamilton, C.E., Mwangi, M.M., Dewell, S. & Papavasiliou, F.N. Transcriptome-wide sequencing reveals numerous APOBEC1 mRNA-editing targets in transcript 3 ' UTRs. *Nature Structural & Molecular Biology* **18**, 230-U300 (2011).

669 33. Luo, L.L. *et al.* SARS-CoV-2 nucleocapsid protein phase separates with G3BPs to
670 disassemble stress granules and facilitate viral production. *Sci Bull* **66**, 1194-1204 (2021).
671 34. Gordon, D.E. *et al.* A SARS-CoV-2 protein interaction map reveals targets for drug
672 repurposing. *Nature* (2020).
673 35. Buchan, J.R. & Parker, R. Eukaryotic stress granules: the ins and outs of translation. *Mol*
674 *Cell* **36**, 932-941 (2009).
675 36. Yang, P. *et al.* G3BP1 Is a Tunable Switch that Triggers Phase Separation to Assemble
676 Stress Granules. *Cell* **181**, 325-345 e328 (2020).
677 37. Tsai, W.C. & Lloyd, R.E. Cytoplasmic RNA Granules and Viral Infection. *Annu Rev Virol* **1**,
678 147-170 (2014).
679 38. Onomoto, K., Yoneyama, M., Fung, G., Kato, H. & Fujita, T. Antiviral innate immunity and
680 stress granule responses. *Trends Immunol* **35**, 420-428 (2014).
681 39. White, J.P. & Lloyd, R.E. Regulation of stress granules in virus systems. *Trends Microbiol*
682 **20**, 175-183 (2012).
683 40. Lloyd, R.E. Regulation of stress granules and P-bodies during RNA virus infection. *Wiley*
684 *Interdiscip Rev RNA* **4**, 317-331 (2013).
685 41. Zheng, Z.Q., Wang, S.Y., Xu, Z.S., Fu, Y.Z. & Wang, Y.Y. SARS-CoV-2 nucleocapsid protein
686 impairs stress granule formation to promote viral replication. *Cell Discov* **7**, 38 (2021).
687 42. Banani, S.F., Lee, H.O., Hyman, A.A. & Rosen, M.K. Biomolecular condensates: organizers
688 of cellular biochemistry. *Nat Rev Mol Cell Bio* **18**, 285-298 (2017).
689 43. Tsang, B. *et al.* Phosphoregulated FMRP phase separation models activity-dependent
690 translation through bidirectional control of mRNA granule formation. *P Natl Acad Sci USA*
691 **116**, 4218-4227 (2019).
692 44. Kent, S. *et al.* Phase-Separated Transcriptional Condensates Accelerate Target-Search
693 Process Revealed by Live-Cell Single-Molecule Imaging. *Cell Reports* **33** (2020).
694 45. Chen, G.H., Katrekar, D. & Mali, P. RNA-Guided Adenosine Deaminases: Advances and
695 Challenges for Therapeutic RNA Editing. *Biochemistry* **58**, 1947-1957 (2019).
696 46. Huang, W.J. *et al.* Molecular determinants for regulation of G3BP1/2 phase separation by
697 the SARS-CoV-2 nucleocapsid protein. *Cell Discov* **7** (2021).
698 47. Olson, M.E., Harris, R.S. & Harki, D.A. APOBEC Enzymes as Targets for Virus and Cancer
699 Therapy. *Cell Chem Biol* **25**, 36-49 (2018).
700 48. Wang, W.Q., Zhao, H.Y. & Han, G.Z. Host-Virus Arms Races Drive Elevated Adaptive
701 Evolution in Viral Receptors. *Journal of Virology* **94** (2020).
702 49. Dolja, V.V. Virus-host arms race as a shaping force of virus evolution. *Phytopathology* **99**,
703 S165-S165 (2009).
704 50. Mangeat, B. *et al.* Broad antiretroviral defence by human APOBEC3G through lethal
705 editing of nascent reverse transcripts. *Nature* **424**, 99-103 (2003).
706 51. Esnault, C. *et al.* APOBEC3G cytidine deaminase inhibits retrotransposition of endogenous
707 retroviruses. *Nature* **433**, 430-433 (2005).
708 52. Sheehy, A.M., Gaddis, N.C. & Malim, M.H. The antiretroviral enzyme APOBEC3G is
709 degraded by the proteasome in response to HIV-1 Vif. *Nature Medicine* **9**, 1404-1407
710 (2003).
711 53. Schmidt, N. *et al.* The SARS-CoV-2 RNA-protein interactome in infected human cells. *Nat*
712 *Microbiol* **6**, 339-+ (2021).

54. Wang, S.M. & Wang, C.T. APOBEC3G cytidine deaminase association with coronavirus nucleocapsid protein. *Virology* **388**, 112-120 (2009).
55. Teng, B.B., Burant, C.F. & Davidson, N.O. Molecular-Cloning of an Apolipoprotein-B Messenger-Rna Editing Protein. *Science* **260**, 1816-1819 (1993).
56. Wolfe, A.D., Arnold, D.B. & Chen, X.J.S. Comparison of RNA Editing Activity of APOBEC1-A1CF and APOBEC1-RBM47 Complexes Reconstituted in HEK293T Cells. *Journal of Molecular Biology* **431**, 1506-1517 (2019).
57. Alberti, S. Phase separation in biology. *Current Biology* **27**, R1097-R1102 (2017).
58. Antoniou, D. & Schwartz, S.D. Phase Space Bottlenecks in Enzymatic Reactions. *J Phys Chem B* **120**, 433-439 (2016).
59. Tauber, D., Tauber, G. & Parker, R. Mechanisms and Regulation of RNA Condensation in RNP Granule Formation. *Trends Biochem Sci* **45**, 764-778 (2020).
60. Zhou, W., Mohr, L., Maciejowski, J. & Kranzusch, P.J. cGAS phase separation inhibits TREX1-mediated DNA degradation and enhances cytosolic DNA sensing. *Molecular Cell* **81**, 739-+ (2021).
61. Drobot, B. *et al.* Compartmentalised RNA catalysis in membrane-free coacervate protocells. *Nature Communications* **9** (2018).
62. Strulson, C.A., Molden, R.C., Keating, C.D. & Bevilacqua, P.C. RNA catalysis through compartmentalization. *Nat Chem* **4**, 941-946 (2012).

Figure legends

Fig. 1 SARS-CoV-2 N protein interacts with host deaminases and links deaminases to edit viral RNA. a, Average proportion of nucleotide changes in the SARS-CoV-2 transcriptomes. n = 3 (allelic fraction ≥ 0.1%). The X-axis is the original nucleotide of SARS-CoV-2 genome, and indicative colour blocks represent nucleotides to which original nucleotide were mutated. **b,** Sequence contexts for A-to-I (G) and C-to-U (T) edited sites in viral transcriptome. **c**, Allelic fraction of SNVs identified in the SARS-CoV-2 transcriptomes using SARS-CoV-2 N protein-based genetic complementation system. n = 3 (allelic fraction ≥ 0.02%). **d,** The number of SNVs identified in the SARS-CoV-2 transcriptomes using SARS-CoV-2 N protein-based genetic complementation system after one and six passages (P1 and P6) (allelic fraction ≥ 0.1%, n = 3). Data are shown as mean ± SEM. **e,** N protein interacted with host deaminases. HeLa cells were co-transfected with plasmids encoding GFP-tagged N protein and Flag control or Flag-tagged host deaminases, including APOBEC1 (A1), activation-induced cytidine deaminase (AID), APOBEC2 (A2), APOBEC3 (3A-3D and 3F-3H), APOBEC4 (A4), ADAR1 (AD1) and ADAR2 (AD2). Cell lysates were immunoprecipitated with an anti-Flag antibody and the bound proteins were then analysed by western blotting. **f,** The enrichment of N protein mRNA was measured by RIP assay in HeLa cells (n = 3). HeLa cells were transfected with N gene with GFP tag and lysates were collected 48 h after transfection. Lysates were prepared and split for incubation with mouse anti-ADAR2 antibody. Co-precipitated RNA was analyzed by qRT-PCR (n = 3). Data are shown as mean ± SEM. **g,** Average proportion of nucleotide changes in the N gene in HeLa cells. Cells were transfected with the N gene and then were lysed in TRIzol reagent. The mRNA was reversely transcribed and cDNA were amplified. The amplicons were analysed by deep sequencing. Values and error bars were represented as the mean ± SEM of three independent biological replicates (allelic fraction ≥ 0.02%). The X-axis is the original nucleotide of N gene, and indicative colour blocks represent nucleotides to which original nucleotide were mutated. **h,** ADAR2 specifically interacted with SARS-CoV-2 N protein. HeLa cells were co-transfected with plasmids encoding HA-tagged ADAR2 and GFP control or GFP-tagged SARS-CoV-2 genes, including nsp7

(n7), nsp8 (n8), nsp9 (n9), nsp12 (n12), nsp16 (n16), spike (S), E, M and N. Cell lysates were immunoprecipitated with an anti-GFP antibody and the bound proteins were then analysed by western blotting. Statistical analysis was performed with two-tailed unpaired t-test. ns > 0.05, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. ns, no significant.

Fig. 2 The condensation of N protein-deaminases complex efficiently enhances viral RNA mutagenesis. a, N protein interacted with host G3BP1 and deaminase APOBEC3G or ADAR2. HeLa cells were co-transfected with plasmids encoding GFP-tagged N protein or GFP control and HA-tagged APOBEC3G (A3G) or ADAR2. Cell lysates were immunoprecipitated with an anti-GFP antibody and the bound proteins were then analysed by western blotting with antibodies against HA or G3BP1. **b,** APOBEC3G or ADAR2 co-localized with N protein in stress granules in HeLa cells. The HeLa cells transfected with the N gene and APOBEC3G or ADAR2 were treated with sodium arsenite (AS) for 45 min to induces SG formation, followed by immunostaining for the N protein (red), A3G or ADAR2 (blue) and endogenous G3BP1. Scale bar: 20 μm. **c,** The condensate formation of N protein-deaminases complex enhanced N protein interaction with A3G or ADAR2 protein. HeLa cells transfected with N gene and Flag control or Flag-tagged deaminases, including AID, A3G, A3H and ADAR2, were treated with or without AS. Cell lysates were immunoprecipitated with an anti-flag antibody and the bound proteins were then analysed by Western blotting. **d,** The enrichment of N protein mRNA was measured by RIP assay in HeLa cells treated with or without AS treatment. HeLa cells were transfected with N gene with GFP tag for 48 h and then treated with or without AS treatment for 45 min. Lysates were prepared and split for incubation with mouse anti-ADAR2 antibody. Co-precipitated RNA was analyzed by qRT-PCR (n = 3). Data are shown as mean ± SEM. **e,** The formation of RNP condensation promoted the account of SNVs identified in the N gene in HeLa cells. Cells were transfected with the N gene with or without sodium arsenite (AS) for 10 h and then lysed in TRIzol reagent. The mRNA was reversely transcribed and cDNA were amplified. The amplicons were analysed by deep sequencing. Values and error bars were represented as the mean ± SEM of three

independent biological replicates (allelic fraction ≥ 0.02%). **f,** The formation of RNP condensation promoted the account of SNVs identified in the spike gene in HeLa cells. Values and error bars were represented as the mean ± SEM of three independent biological replicates (allelic fraction ≥ 0.02%). **g,** The formation of RNP condensation promoted the account of SNVs identified in the Orf 8 in HeLa cells. Values and error bars were represented as the mean ± SEM of three independent biological replicates (allelic fraction ≥ 0.02%). **h,** Allelic fraction of SNVs for each nucleotide change in the N gene in HeLa cells treated with or without AS treatment (n = 3). Data are shown as mean ± SEM. **i,** Allelic fraction of SNVs for each nucleotide change in the spike gene in HeLa cells treated with or without AS treatment (n =3). Data are shown as mean ± SEM. **j,** Allelic fraction of SNVs for A23403G (spike protein: D614G) in HeLa cells (n =3). Data are shown as mean ± SEM. Data were normalized to 100% by the allelic fraction of HeLa cells without AS treatment. **k,** Allelic fraction of SNVs for C27972T (ORF8:Q27*) in HeLa cells (n =3). Data are shown as mean ± SEM. Data were normalized to 100% by the allelic fraction of HeLa cells without AS treatment. Statistical analysis was performed with two-tailed unpaired t-test. ns > 0.05, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. ns, no significant.

Fig. 3 The disruption of N protein/deaminases-involved condensates attenuates viral RNA mutagenesis. a, G3BP1/2 dKO abolished the formation of N protein/deaminases-involved condensates. **b,** The disruption of N protein/deaminases-involved condensates attenuated ADAR-mediated RNA editing. Values and error bars were represented as the mean ± SEM of three independent biological replicates (allelic fraction ≥ 0.03%). **c,** The disruption of N protein/deaminases-involved condensates attenuated APOBEC-mediated RNA editing. Values and error bars were represented as the mean ± SEM of three independent biological replicates (allelic fraction ≥0.03%). **d,** ADAR1/2 dKO reduced ADAR-mediated RNA editing. Values and error bars were represented as the mean ± SEM of three independent biological replicates (allelic fraction ≥ 0.03%). **e,** The RNP condensation failed to increase ADAR-mediated RNA mutagenesis due to ADAR1/2 dKO. Values and error bars were represented as the mean ± SEM of three independent biological replicates (allelic fraction ≥ 0.03%). Statistical

analysis was performed with two-tailed unpaired t-test. ns > 0.05, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. ns, no significant.

Fig. 4 RNA binding domain is required for ADAR2 to form phase separation-mediated RNP condensates with N protein. a, Characterization of ADAR2 mutants required for interaction with N protein. Top, schematic domain structure of ADAR2. HeLa cells were co-transfected with plasmids encoding N protein and Flag control or Flag-tagged host ADAR2 truncations. Cell lysates were immunoprecipitated with an anti-Flag antibody and the bound proteins were then analysed by Western blotting. **b,** Characterization of ADAR2 mutants required for SG-localization. The HeLa cells co-transfected with N gene and BFP–tagged ADAR2 mutants were treated with sodium arsenite for 45 min to induce SGs, followed by immunostaining for N protein (red) and G3BP1 (green). **c,** Time-lapse imaging of RNA-induced RBD1 phase separation. Liquid droplets formed after mixing of ADAR2-RBD1 protein with 75 ng/μL RNA. **d,** Fluorescence intensity of ADAR2-RBD1 liquid droplets forming in the presence of 75 ng/μL RNA. Data are shown as mean ± SEM. **e,** FRAP recovery images and recovery curve of ADAR2-RBD1 liquid droplets forming in the presence of RNA. The dotted square displays the region of photobleaching. Data are shown as mean ± SEM (n = 30). Scale bar, 10 μm. **f,** N protein phase separated with ADAR2-RBD1 in an RNA-dependent manner. The concentration of each protein in buffer is 2 μM. Scale bar, 10 μm. **g**, Column scatter charts display the droplets area in reactions relative to Figure. 3f. Data are shown as mean ± SEM (n = 10). Statistical analysis was performed with two-tailed unpaired t-test. ns > 0.05, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. ns, no significant.

Fig. 5 N^{F17A} protein fails to penetrate into deaminases-involved RNP condensates, which impaires deaminases-mediated viral RNA editing activity. a, Schematic domain structure of NF17A protein. **b**, F17A mutation in N protein impaired the N protein phase separation with G3BP1. Scale bar, 10 μm. **c**, Column scatter charts display the droplets area in reactions related to panel **b**. **d,** N^{F17A} protein still possessed the interaction ability with ADAR2. HeLa cells were co-transfected with plasmids encoding GFP-tagged N wild type or N^{F17A} mutant and HA-tagged ADAR2. Cell

lysates were immunoprecipitated with an anti-GFP antibody and the bound proteins were then analysed by Western blotting. **e,** N^{F17A} protein failed to penetrate into deaminases-involved RNP condensates. The HeLa cells co-transfected with N wild type or N^{F17A} mutant were treated with sodium arsenite for 45 min to induce SGs, followed by immunostaining for N protein (green) and the endogenous ADAR2 (red). **f,** The enrichment of N protein mRNA was measured by RIP assay in HeLa cells treated with or without AS treatment. HeLa cells were transfected with N wild type or N^{F17A} mutant with GFP tag for 48 h and then treated with or without AS treatment for 45 min. Lysates were prepared and split for incubation with mouse anti-ADAR2 antibody. Co-precipitated RNA was analyzed by qRT-PCR (n = 3). Data are shown as mean ± SEM. **g,** N^{F17A} protein failed to penetrate into deaminases-involved RNP condensates, which impaired deaminases-mediated viral RNA editing activity. Values and error bars were represented as the mean ± SEM of three independent biological replicates (allelic fraction ≥0.03%). **h,** Average proportion of nucleotide changes in the SARS-CoV-2 transcriptomes using SARS-CoV-2 N protein WT or N^{F17A}-based genetic complementation system. Values and error bars were represented as the mean ± SEM of three independent biological replicates. **i,** Allelic fraction of SNVs for A19736G and C27970T in the SARS-CoV-2 transcriptomes using SARS-CoV-2 WT N protein or N^{F17A}-based genetic complementation system. Values and error bars were represented as the mean ± SEM of three independent biological replicates. Statistical analysis was performed with two-tailed unpaired t-test. ns > 0.05, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. ns, no significant.

Fig. 6 Schematic illustration depicting how SARS-CoV-2 hijacks host deaminases to achieve viral mutagenesis and evolution.

Figures

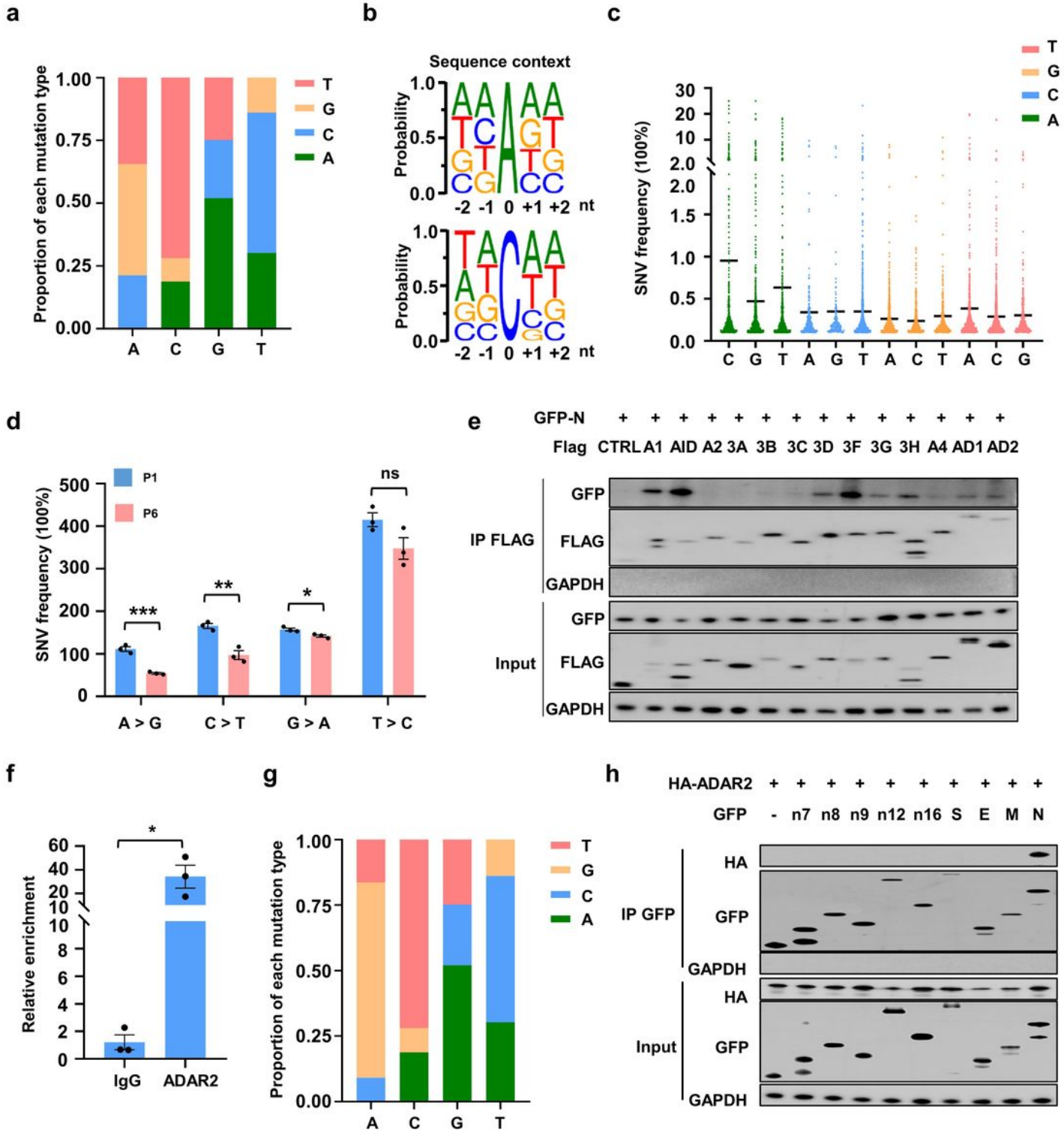

Figure 1

SARS-CoV-2 N protein interacts with host deaminases and links deaminases to edit viral RNA. a, Average proportion of nucleotide changes in the SARS-CoV-2 transcriptomes. n = 3 (allelic fraction ≥ 0.1%). The X-axis is the original nucleotide of SARS-CoV-2 genome, and indicative colour blocks represent nucleotides

to which original nucleotide were mutated. b, Sequence contexts for A-to-I (G) and C-to-U (T) edited sites in viral transcriptome. c, Allelic fraction of SNVs identified in the SARS CoV-2 transcriptomes using SARS-CoV-2 N protein-based genetic complementation system. n = 3 (allelic fraction ≥ 0.02%). d, The number of SNVs identified in the SARS CoV-2 transcriptomes using SARS-CoV-2 N protein-based genetic complementation system after one and six passages (P1 and P6) (allelic fraction ≥ 0.1%, n = 3). Data are shown as mean ± SEM. e, N protein interacted with host deaminases. HeLa cells were co-transfected with plasmids encoding GFP-tagged N protein and Flag control or Flag tagged host deaminases, including APOBEC1 (A1), activation-induced cytidine deaminase (AID), APOBEC2 (A2), APOBEC3 (3A-3D and 3F-3H), APOBEC4 (A4), ADAR1 (AD1) and ADAR2 (AD2). Cell lysates were immunoprecipitated with an anti Flag antibody and the bound proteins were then analysed by western blotting. f, The enrichment of N protein mRNA was measured by RIP assay in HeLa cells (n = 3). HeLa cells were transfected with N gene with GFP tag and lysates were collected 48 h after transfection. Lysates were prepared and split for incubation with mouse anti-ADAR2 antibody. Co-precipitated RNA was analyzed by qRT-PCR (n = 3). Data are shown as mean ± SEM. g, Average proportion of nucleotide changes in the N gene in HeLa cells. Cells were transfected with the N gene and then were lysed in TRIzol reagent. The mRNA was reversely transcribed and cDNA were amplified. The amplicons were analysed by deep sequencing. Values and error bars were represented as the mean ± SEM of three independent biological replicates (allelic fraction ≥ 0.02%). The X-axis is the original nucleotide of N gene, and indicative colour blocks represent nucleotides to which original nucleotide were mutated. h, ADAR2 specifically interacted with SARS-CoV-2 N protein. HeLa cells were co-transfected with plasmids encoding H tagged ADAR2 and GFP control or GFP-tagged SARS-CoV-2 genes, including nsp7 (n7), nsp8 (n8), nsp9 (n9), nsp12 (n12), nsp16 (n16), spike (S), E, M and N. Cell lysates were immunoprecipitated with an anti-GFP antibody and the bound proteins were then analysed by western blotting. Statistical analysis was performed with two-tailed unpaired t-test. ns > 0.05, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. ns, no significant.

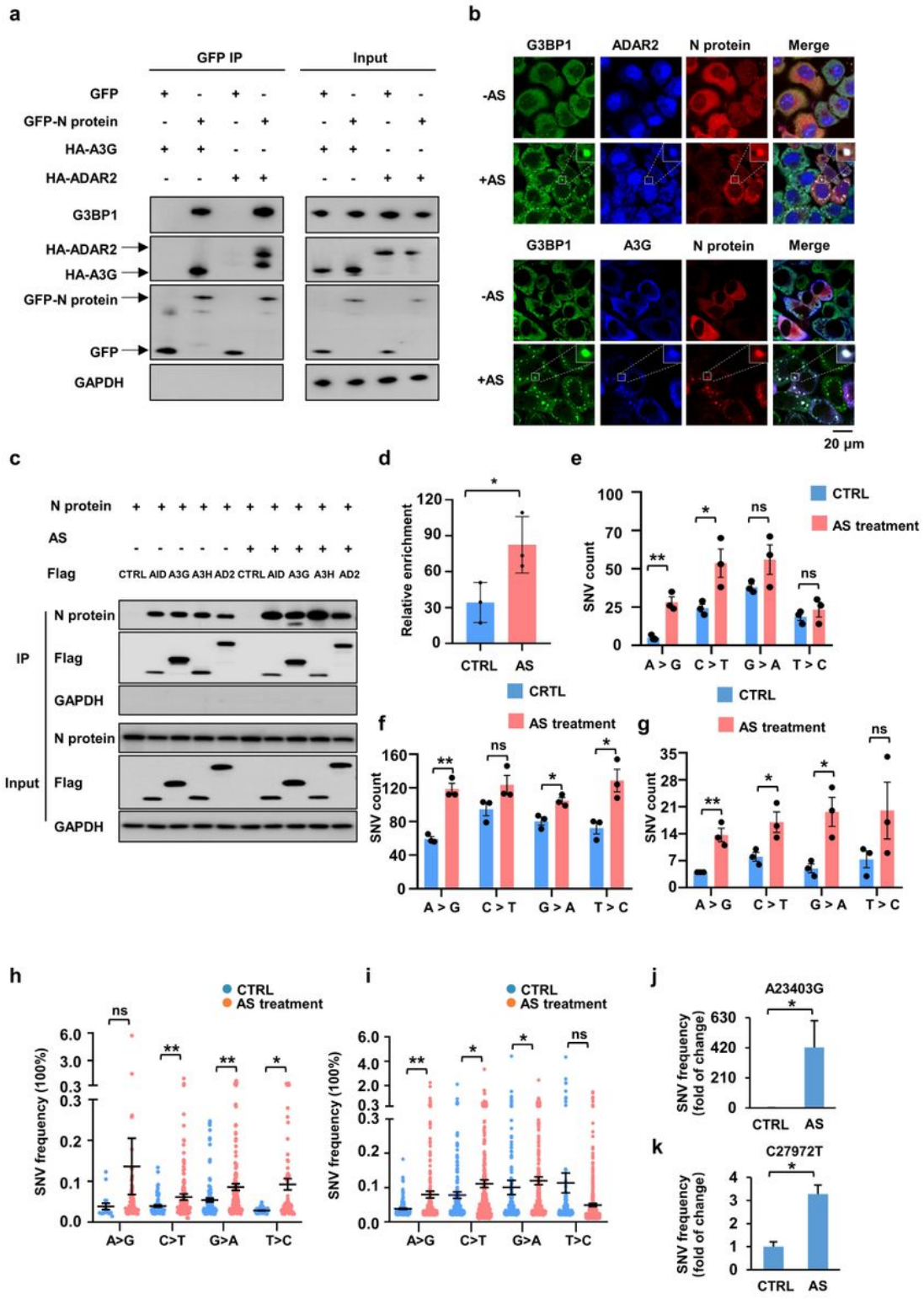

Figure 2

The condensation of N protein-deaminases complex efficiently enhances viral RNA mutagenesis. a, N protein interacted with host G3BP1 and deaminase APOBEC3G or ADAR2. HeLa cells were co-transfected with plasmids encoding GF tagged N protein or GFP control and HA-tagged APOBEC3G (A3G) or ADAR2. Cell lysates were immunoprecipitated with an anti-GFP antibody and the bound proteins were then analysed by western blotting with antibodies against HA or G3BP1. b, APOBEC3G or ADAR2 co-localized

with N protein in stress granules in HeLa cells. The HeLa cells transfected with the N gene and APOBEC3G or ADAR2 were treate with sodium arsenite (AS) for 45 min to induces SG formation, followed by immunostaining for the N protein (red), A3G or ADAR2 (blue) and endogenous G3BP1. Scale bar: 20 μm. c, The condensate formation of N protein-deaminases complex enhanced N protein interaction with A3G or ADAR2 protein. HeLa cells transfected with N gene and Flag control or Flag-tagged deaminases, including AID, A3G, A3H and ADAR2, were treated with or without AS. Cell lysates were immunoprecipitated with an anti-flag antibody and the bound proteins were then analysed by Western blotting. d, The enrichment of N protein mRNA was measured by RIP assay in HeLa cells treated with or without AS treatment. HeLa cells were transfected with N gene with GFP tag for 48 h and then treated with or without AS treatment for 45 min. Lysates were prepared and split for incubation with mouse anti-ADAR2 antibody. Co precipitated RNA was analyzed by qRT-PCR (n = 3). Data are shown as mean ± SEM. e, The formation of RNP condensation promoted the account of SNVs identified in the N gene in HeLa cells. Cells were transfected with the N gene with or without sodium arsenite (AS) for 10 h and then lysed in TRIzol reagent. The mRNA was reversely transcribed and cDNA were amplified. The amplicons were analysed by deep sequencing. Values and error bars were represented as the mean ± SEM of three independent biological replicates (allelic fraction ≥ 0.02%). f, The formation of RNP condensation promoted the account of SNVs identified in the spike gene in HeLa cells. Values and error bars were represented as the mean ± SEM of three independent biological replicates (allelic fraction ≥ 0.02%). g, The formation of RNP condensation promoted the account of SNVs identified in the Orf 8 in HeLa cells. Values and error bars were represented as the mean ± SEM of three independent biological replicates (allelic fraction ≥ 0.02%). h, Allelic fraction of SNVs for each nucleotide change in the N gene in HeLa cells treated with or without AS treatment (n = 3). Data are shown as mean ± SEM. i, Allelic fraction of SNVs for each nucleotide change in the spike gene in HeLa cells treated with or without AS treatment (n =3). Data are shown as mean ± SEM. j, Allelic fraction of SNVs for A23403G (spike protein: D614G) in HeLa cells (n =3). Data are shown as mean ± SEM. Data were normalized to 100% by the allelic fraction of HeLa cells without AS treatment. k, Allelic fraction of SNVs for C27972T (ORF8:Q27*) in HeLa cells (n =3). Data are shown as mean ± SEM. Data were normalized to 100% by the allelic fraction of HeLa cells without AS treatment.Statistical analysis was performed with two-tailed unpaired t-test. ns > 0.05, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. ns, no significant.

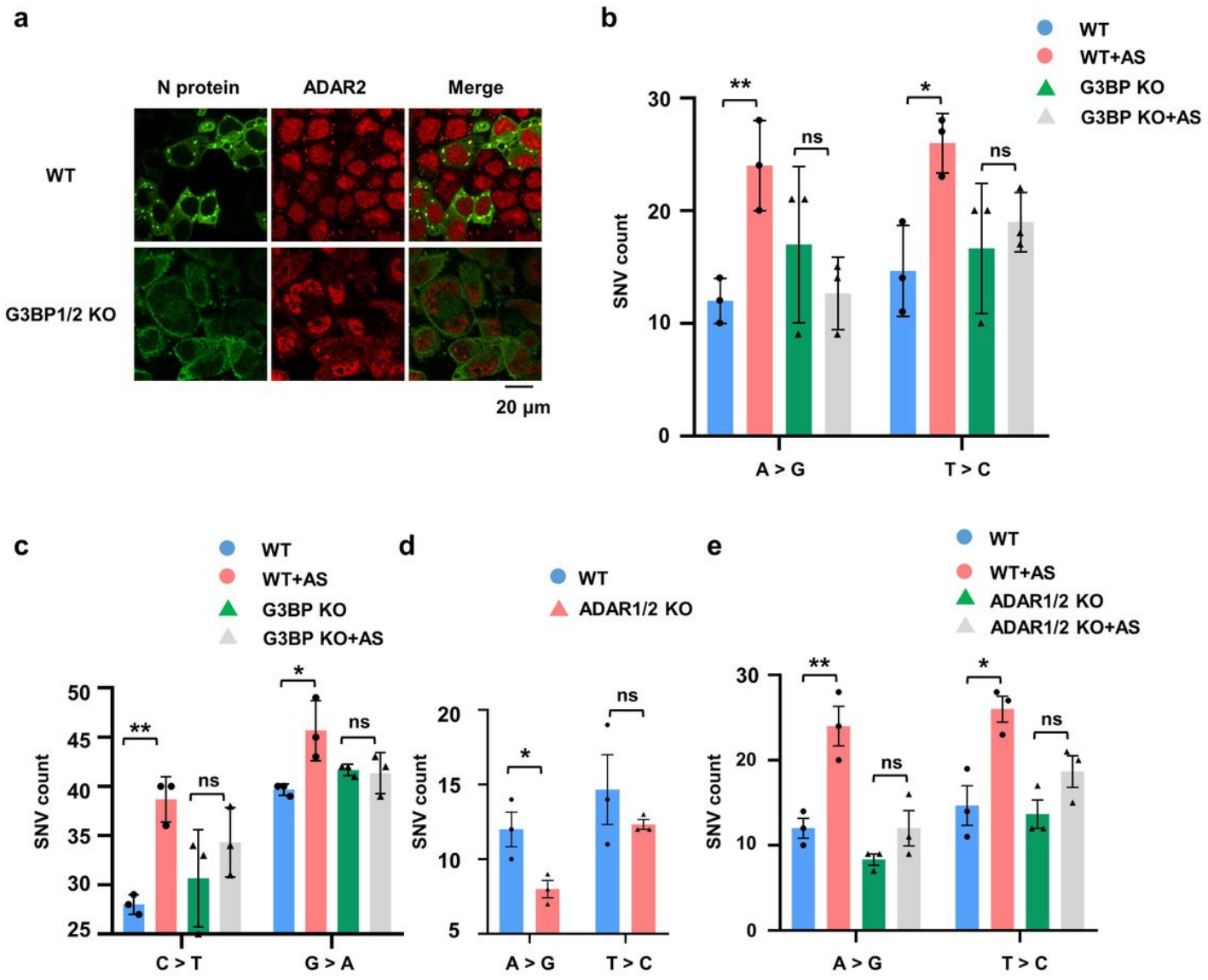

Figure 3

The disruption of N protein/deaminases-involved condensates attenuates viral RNA mutagenesis. a, G3BP1/2 dKO abolished the formation of N protein/deaminases-involved condensates. b, The disruption of N protein/deaminases involved condensates attenuated ADAR-mediated RNA editing. Values and error bars were represented as the mean ± SEM of three independent biological replicates (allelic fraction ≥ 0.03%). c, The disruption of N protein/deaminases-involved condensates attenuated APOBEC-mediated RNA editing. Values and error bars were represented as the mean ± SEM of three independent biological replicates (allelic fraction ≥0.03%). d, ADAR1/2 dKO reduced ADAR-mediated RNA editing. Values and error bars were represented as the mean ± SEM of three independent biological replicates (allelic fraction ≥ 0.03%). e, The RNP condensation failed to increase ADAR-mediated RNA mutagenesis due to ADAR1/2 dKO. Values and error bars were represented as the mean ± SEM of three independent biological replicates (allelic fraction ≥ 0.03%). Statistical analysis was performed with two-tailed unpaired t-test. ns > 0.05, *P < 0.05, **P < 840 0.01, ***P < 0.001, ****P < 0.0001. ns, no significant.

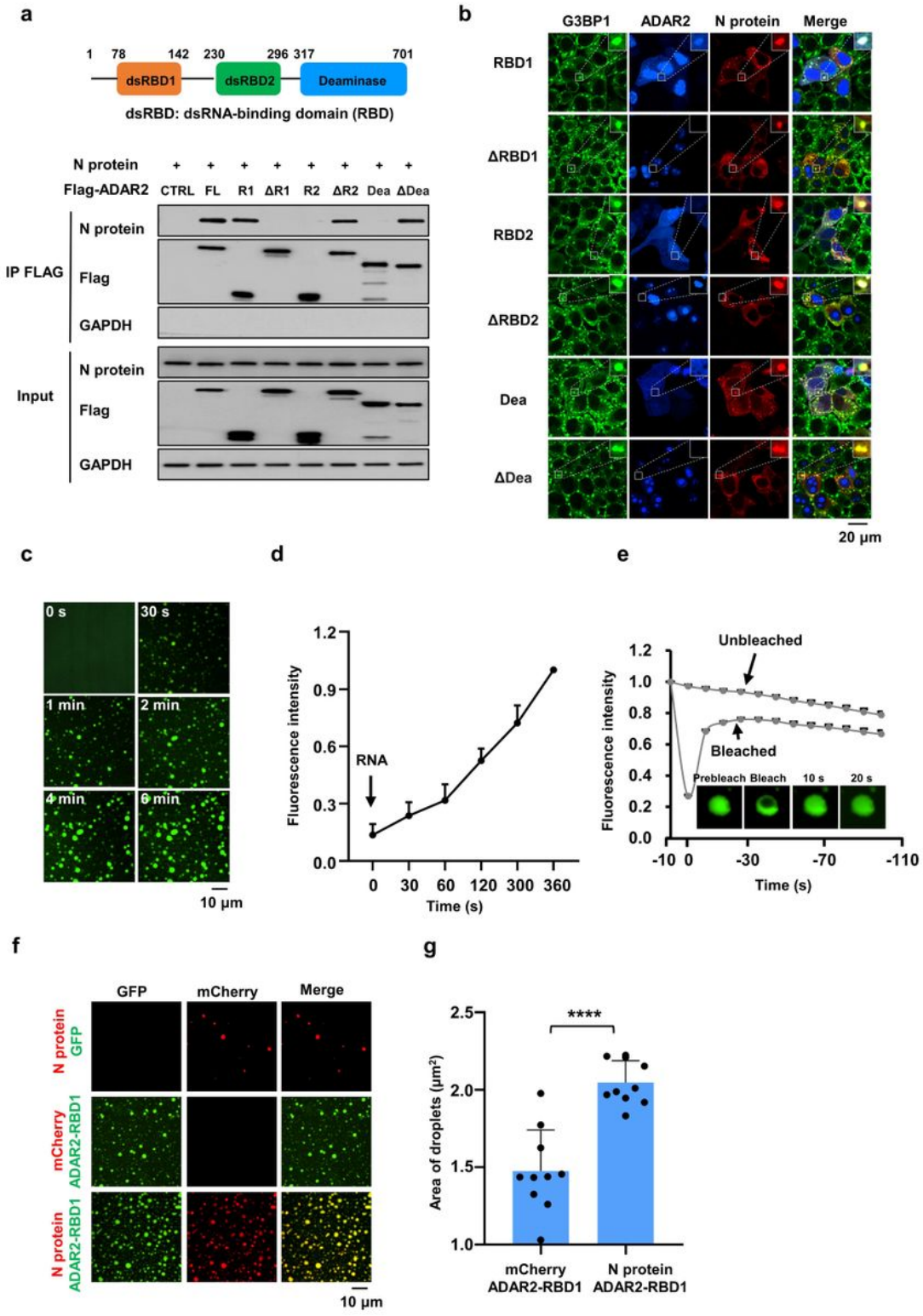

Figure 4

RNA binding domain is required for ADAR2 to form phase separation mediated RNP condensates with N protein. a, Characterization of ADAR2 mutants required for interaction with N protein. Top, schematic domain structure of ADAR2. HeLa cells were co-transfected with plasmids encoding N protein and Flag control or Flag-tagged host ADAR2 truncations. Cell lysates were immunoprecipitated with an anti-Flag antibody and the bound proteins were then analysed by Western blotting. b, Characterization of ADAR2

mutants required for SG-localization. The HeLa cells co transfected with N gene and BFP–tagged ADAR2 mutants were treated with sodium arsenite for 45 min to induce SGs, followed by immunostaining for N protein (red) and G3BP1 (green). c, Time-lapse imaging of RNA-induced RBD1 phase separation. Liquid droplets formed after mixing of ADAR2-RBD1 protein with 75 ng/μL RNA. d, Fluorescence intensity of ADAR2-RBD1 liquid droplets forming in the presence of 75 ng/μL RNA. Data are shown as mean ± SEM. e, FRAP recovery images and recovery curve of ADAR2-RBD1 liquid droplets forming in the presence of RNA. The dotted square displays the region of photobleaching. Data are shown as mean ± SEM (n = 30). Scale bar, 10 μm. f, N protein phase separated with ADAR2-RBD1 in an RNA dependent manner. The concentration of each protein in buffer is 2 μM. Scale bar, 10 μm. g, Column scatter charts display the droplets area in reactions relative to Figure. 3f. Data are shown as mean ± SEM (n = 10). Statistical analysis was performed with two tailed unpaired t-test. ns > 0.05, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. ns, no significant.

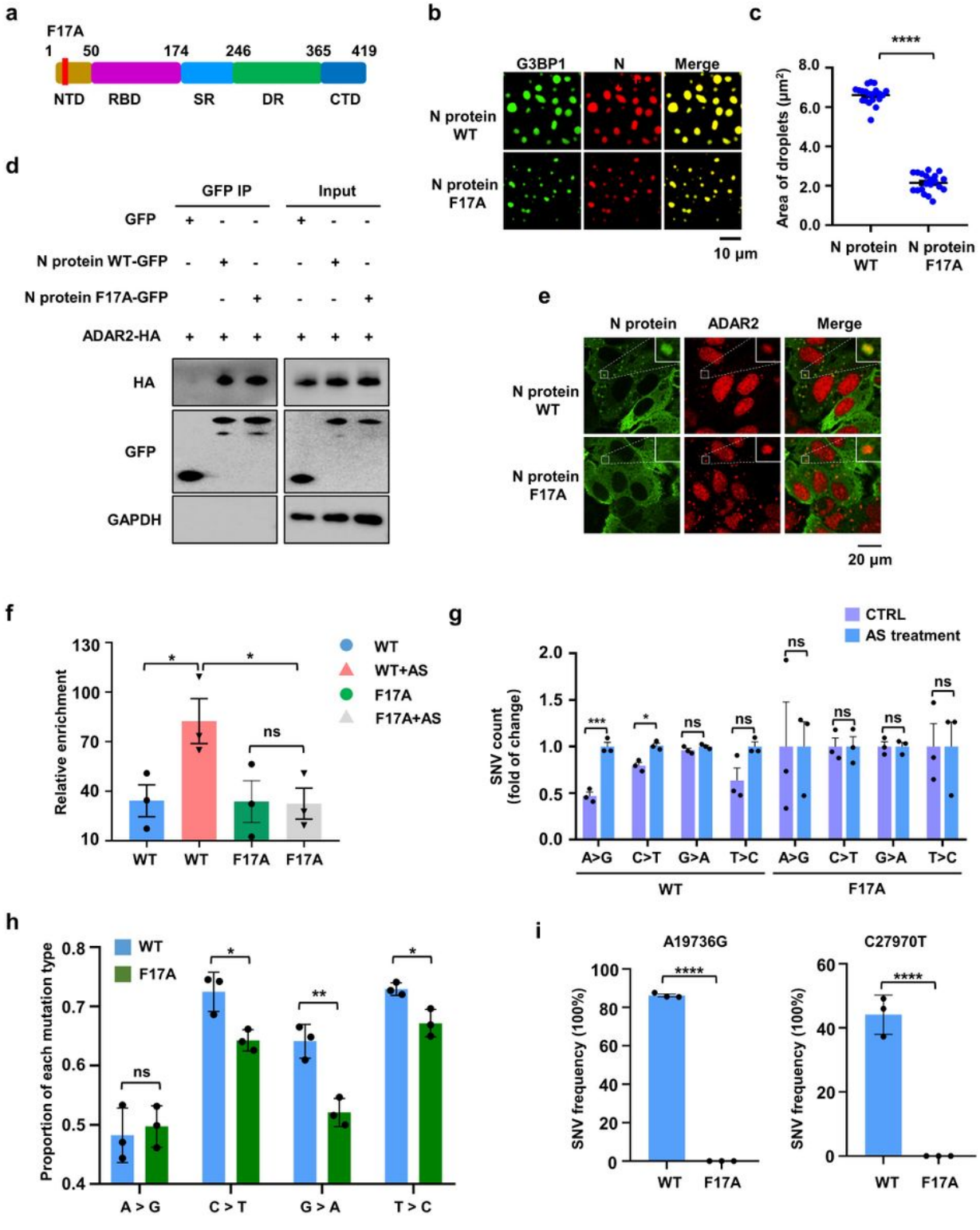

Figure 5

NF17A 862 protein fails to penetrate into deaminases-involved RNP condensates, which impaires deaminases-mediated viral RNA editing activity. a, Schematic domain structure of NF17A protein. b, F17A mutation in N protein impaired the N protein phase separation with G3BP1. Scale bar, 10 μm. c, Column scatter charts display the droplets area in reactions related to panel b. d, N F17A protein still possessed the interaction ability with ADAR2. HeLa cells were co-transfected with plasmids encoding GFP-tagged N

wild type or N F17A mutant and HA-tagged ADAR2. Cell lysates were immunoprecipitated with an anti-GFP antibody and the bound proteins were then analysed by Western blotting. e, N F17A protein failed to penetrate into deaminases-involved RNP condensates. The HeLa cells co-transfected with N wild type or N F17A mutant were treated with sodium arsenite for 45 min to induce SGs, followed by immunostaining for N protein (green) and the endogenous ADAR2 (red). f, The enrichment of N protein mRNA was measured by RIP assay in HeLa cells treated with or without AS treatment. HeLa cells were transfected with N wild type or N F17A mutant with GFP tag for 48 h and then treated with or without AS treatment for 45 min. Lysates were prepared and split for incubation with mouse anti-ADAR2 antibody. Co878 precipitated RNA was analyzed by qRT-PCR (n = 3). Data are shown as mean ± SEM. g, N F17A protein failed to penetrate into deaminases-involved RNP condensates, which impaired deaminases-mediated viral RNA editing activity. Values and error bars were represented as the mean ± SEM of three independent biological replicates (allelic fraction ≥0.03%). h, Average proportion of nucleotide changes in the SARS-CoV-2 transcriptomes using SARS-CoV-2 N protein WT or NF17A -based genetic complementation system. Values and error bars were represented as the mean ± SEM of three independent biological replicates. i, Allelic fraction of SNVs for A19736G and C27970T in the SARS-CoV-2 transcriptomes using SARS-CoV-2 WT N protein or N F17A -based genetic complementation system. Values and error bars were represented as the mean ± SEM of three independent biological replicates. Statistical analysis was performed with two-tailed unpaired t-test. ns > 0.05, *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. ns, no significant.

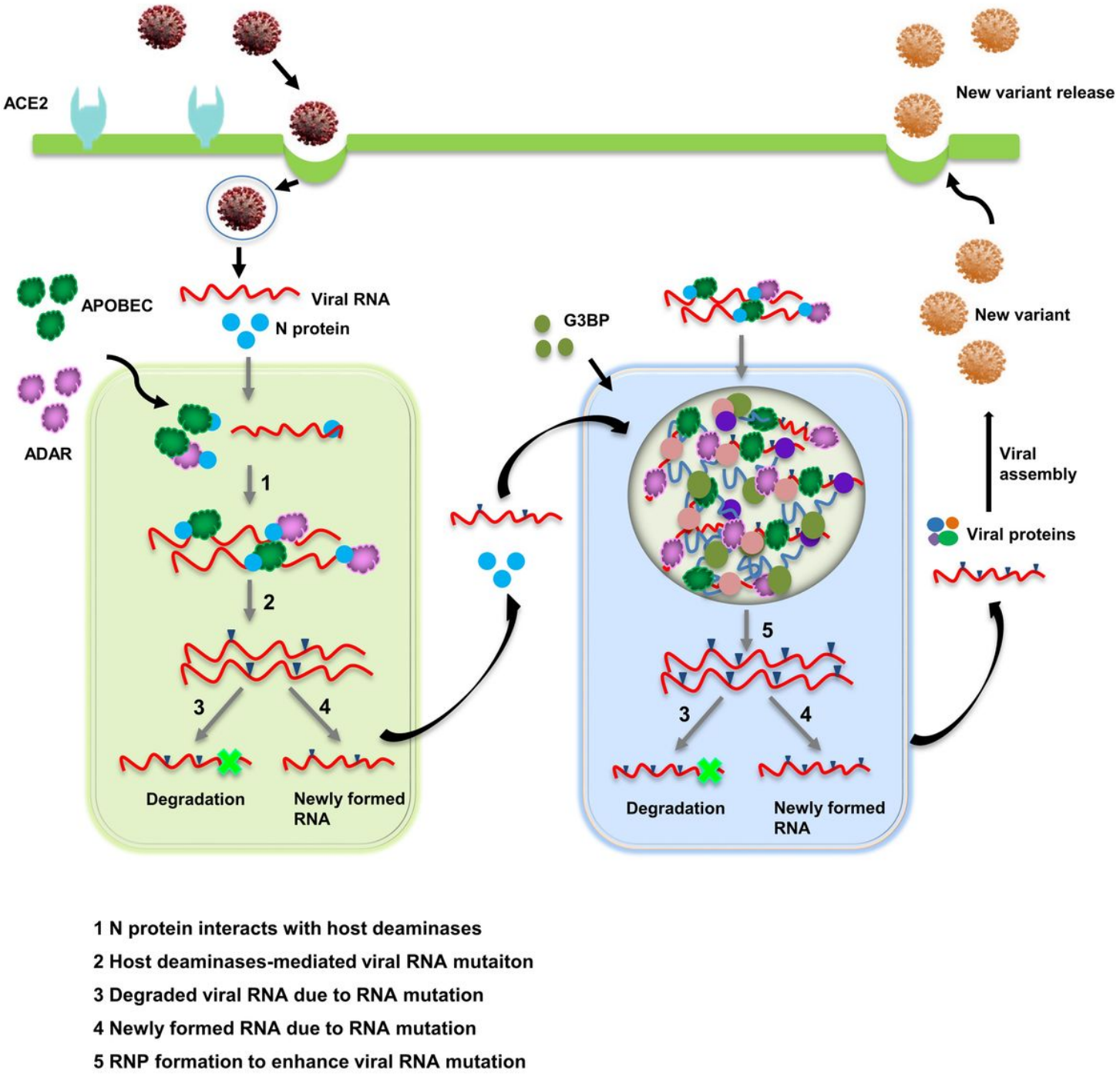

Figure 6

Schematic illustration depicting how SARS-CoV-2 hijacks host deaminases to achieve viral mutagenesis and evolution.

Supplementary Files

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

- Supplementalinformation.pdf