Construction of a mini-RNA replicon in *Escherichia coli*

Akiko Kashiwagi [1,*] and Tetsuya Yomo[2,*]

[1]Faculty of Agriculture and Life Science, Hirosaki University, Hirosaki, Aomori, Japan
[2]School of Life Sciences, East China Normal University, Shanghai, People's Republic of China
*Corresponding author: E-mails: kashi_a1@hirosaki-u.ac.jp and tetsuyayomo@gmail.com

Abstract

How the ribonucleic acid (RNA) world transited to the deoxyribonucleic acid (DNA) world has remained controversial in evolutionary biology. At a certain time point in the transition from the RNA world to the DNA world, 'RNA replicons', in which RNAs produce proteins to replicate their coding RNA, and 'DNA replicons', in which DNAs produce RNA to synthesize proteins that replicate their coding DNA, can be assumed to coexist. The coexistent state of RNA replicons and DNA replicons is desired for experimental approaches to determine how the DNA world overtook the RNA world. We constructed a mini-RNA replicon in *Escherichia coli*. This mini-RNA replicon encoded the β subunit, one of the subunits of the Qβ replicase derived from the positive-sense single-stranded Qβ RNA phage and is replicated by the replicase in *E. coli*. To maintain the mini-RNA replicon persistently in *E. coli* cells, we employed a system of α complementation of LacZ that was dependent on the Qβ replicase, allowing the cells carrying the RNA replicon to grow in the lactose minimal medium selectively. The coexistent state of the mini-RNA replicon and DNA replicon (*E. coli* genome) was successively synthesized. The coexistent state can be used as a starting system to experimentally demonstrate the transition from the RNA–protein world to the DNA world, which will contribute to progress in the research field of the origin of life.

Key words: RNA world; mini-RNA replicon; *Escherichia coli*; positive-sense single-stranded RNA bacteriophage Qβ

Graphical Abstract
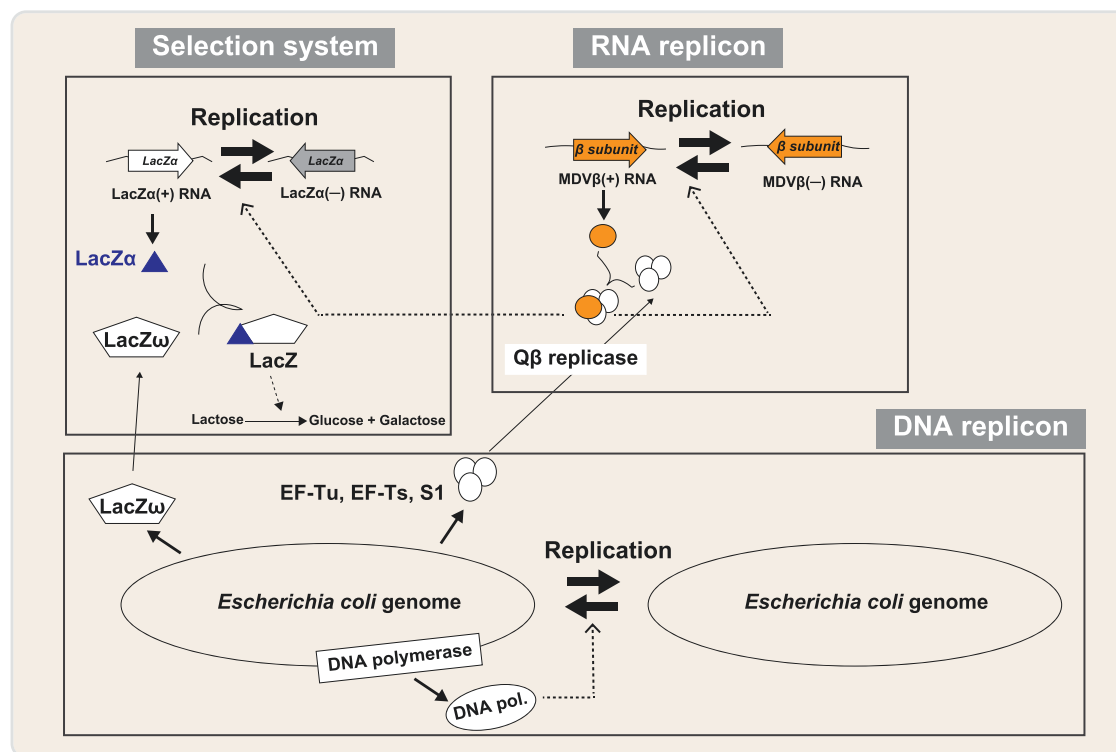
1. Introduction

The ribonucleic acid (RNA) world hypothesis is widely recognized as one of the 'origin of life' (1–5). In the RNA world, RNA had two functions: information storage and catalytic activities. Woese, Crick, and Orgel showed theoretically that RNA might have had catalytic activity (6–8). Subsequently, it was experimentally demonstrated that RNA has catalytic activity (9–12), and Gilbert proposed the concept of the 'RNA world' (1).

It has remained controversial how the RNA world transited to the deoxyribonucleic acid (DNA) world (4, 13–15). It is clear that the storage of information and catalytic functions were passed from RNA to DNA and proteins, respectively, although some catalytic activities remain in RNA as ribozymes (11, 12, 16). The RNA–protein (RNP) world has been proposed to claim that proteins appeared in the RNA world (3). Recently, it has been suggested that various worlds, such as a lipid world, metabolism world and protein world, coexisted and interacted with each other (17). It was shown that ribonucleotides and deoxyribonucleotides could appear at the same time (18). Furthermore, an RNA peptide was experimentally suggested to appear after the RNA world (19). Because DNA replication requires some complex multimer proteins such as polymerase and helicase, it can be assumed that protein synthesis machineries might already have existed before the appearance of the DNA world (14). When we consider the transition from the RNA world to the DNA world with catalytic functions of proteins, 'RNA replicons', in which RNA produces proteins to replicate their coding RNA, and 'DNA replicons', in which DNA produces RNA to synthesize proteins which replicate their coding DNA, can be assumed to have coexisted at a certain evolutionary time point (Figure 1). Probably, DNA replicons mostly overtook RNA replicons thereafter.

It is crucial to synthesize a coexistent state of an RNA replicon and DNA replicon to obtain insights into the transition from the RNA world to the DNA world and then to observe the process in which the DNA replicon overtakes the RNA replicon. The DNA genome in present-day organisms is a DNA replicon because it encodes DNA polymerase to replicate its own DNA. The RNA genome of RNA viruses is a replicon, as it encodes its own RNA replicase. Most RNA viruses proliferate explosively to kill the host cells. In other words, the RNA replicons overcome DNA replicons. However, some persistent viruses proliferate in cells but do not kill their hosts (20–25). The vesicular stomatitis virus (VSV) and the Sendai virus for animal cells have been studied well (22, 26, 27). VSV infected BHK21 cells and was maintained in the cell for over 5 years (28). In plants, the *Endornaviridae* viruses, which are stably founded as double-stranded RNAs, are vertically transmitted through seeds (29, 30). In fungi, the double-stranded RNA of mycovirus does not exhibit its extracellular transmission route; instead, it is distributed to daughter cells through host cell division (29, 31–33). In bacteria, *Escherichia coli* infected with the single–stranded RNA (ssRNA) phage f2 and *Pseudomonas phaseolicola* infected with the double-stranded RNA (dsRNA) phage φ6 turned out to be a carrier state (34–36). These persistent viruses are RNA replicons that parasitize their hosts to coexist with the host DNA genomes.

The experimental setup suitable for getting insights into the transition requires an RNA and DNA replicon coexistence-dependent growth of host cells. The persistent RNA viruses mentioned previously may load a burden that reduces the growth rate of their host cells, and they will be removed if uninfected host cells take over the population of their host cells. Therefore, we need an experimental setup in which an RNA replicon gives the growth advantage to its host cells over uninfected cells. So far, an RNA replicon with genes of interest has been synthesized as RNA vectors by adding selection maker genes to persistent RNA virus genomes (37–43). The single-stranded (+) RNA genome of the Sindbis virus, which includes a *lacZ*, and the persistent infectious variants of the (–) stranded RNA Sendai virus genome, with an *egfp*, were passaged for several generations (44, 45). When the (+) stranded RNA genome of the brome mosaic virus or tobacco mosaic virus with an acetyltransferase gene (*cat*) was introduced into the protoplast, protein production and RNA genome replication were observed (42, 43). For *Pseudomonas syringae* pv. *phaseolicola*, the double-stranded segmented

RNA world   **Coexistence of RNA and DNA replicons**   **DNA world**
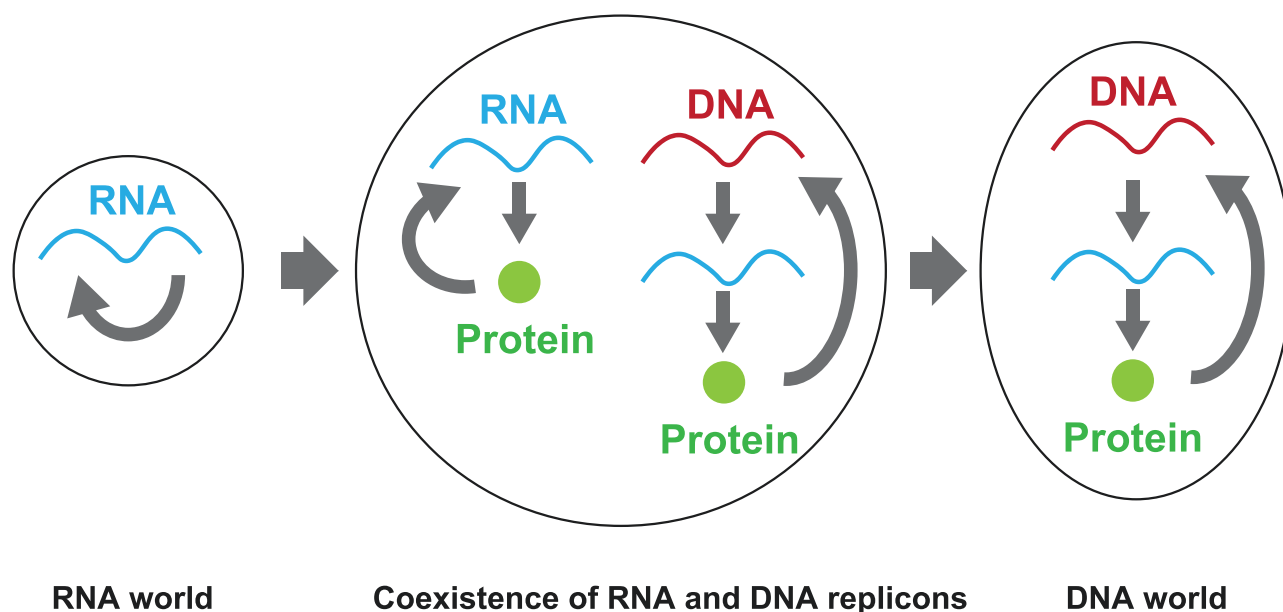
Figure 1. The transition of the replicable information carrier. From the RNA world to the DNA world, an RNA replicon and a DNA replicon coexisted at some intermediate time point.

RNA genome of φ6 was engineered with kanamycin resistance, β-galactosidase and GFP genes (39, 40, 46, 47). The host cells with these RNA vectors selectively grow to allow the stable coexistence of RNA and DNA replicons.

It is desirable to generate a simple RNA replicon that can be stably maintained with a selection maker in *E. coli* since the bacterium is the most-studied model organism and is more suitable for studying the 'origin of life' because of its simplicity than animal and plant cells are. The dsRNA genome of phage φ6, which has been successfully engineered in *P. syringae* pv. *phaseolicola*, as mentioned previously, has a complex segmented structure. A simple RNA replicon with a minimal size is needed for an experimental approach to the 'origin of life'. Although it is not an RNA replicon, Mills developed an RNA molecule that could be maintained in *E. coli* (48). The RNA molecule was replicated by Qβ RNA replicase supplied from a plasmid to produce its complementary strand, which encoded the λN protein, a selection maker for *E. coli*, thus allowing the cells with the RNA molecule to grow selectively.

In this study, we produced an *E. coli* strain in which a 'mini-RNA replicon' and a DNA replicon coexisted. The mini-RNA replicon in our system was an RNA molecule that encoded exclusively the β subunit of the Qβ replicase derived from the RNA bacteriophage Qβ. The β subunit gene was inserted into the middle of MDV-1, a short ssRNA sequence replicable using Qβ replicase (49, 50). The translated β subunit constitutes Qβ replicase with elongation factor (EF)-Tu, EF-Ts and S1 from *E. coli*. The matured enzyme replicated the original RNA, which was a mini-RNA replicon. By replacing the λN protein on the RNA molecule prepared by Mills (as described previously) with the α fragment of LacZ, we synthesized another RNA molecule to allow the *E. coli* cells with the two RNA molecules to grow selectively in a lactose minimal medium. One problem in Mills' system is that chloramphenicol resistance, the selective advantage brought about by the λN protein, is lost for a few passages with spontaneous mutations.

The synthesized coexistent state of the mini-RNA replicon and DNA replicon (*E. coli* genome) was confirmed to be persistent in *E. coli*. As our system involves present well-evolved macromolecules, it may not be very similar to the one that might have existed before the transition from the RNA world to the DNA world. However, our system and the ones present before the transition involve common competition between the two replicons for resources, translation machinery and others. Therefore, our system is an experimental setup toward obtaining experimental insights into how the RNA world transited to the DNA world.

2. Materials and methods
2.1 Strain and media
The *E. coli* DH5α strain and MDS™42*recA*⁻ Blue (Scrab Genomics, WI, USA) were used for plasmid construction and mini-RNA replicon passage, respectively. Lysogeny Broth (LB) medium (10 g/l tryptone, 5 g/l yeast extract and 10 g/l NaCl) and lactose minimal medium (62 mM K_2HPO_4, 39 mM KH_2PO_4, 15 mM ammonium sulfate, 1.8 mM $FeSO_4$, 15 mM thiamine hydrochloride, 0.2 mM $MgSO_4$ and 2.8 mM lactose) were used, together with 25 μM isopropyl-β-D-thiogalactopyranoside (IPTG), 100 μg/ml ampicillin sodium salt (Amp) and 50 μg/ml kanamycin sulfate (Km). The 10 or 20 nM doxycycline hydrochloride (Dox) was used to induce I-SceI.

2.2 Plasmid DNA construction
The flow chart of plasmid construction is shown in Figure S1. Throughout the plasmid construction process, the Phusion High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA)

was employed for polymerase chain reaction (PCR), and the *E. coli* DH5α strain was used. The primers used in this study are shown in Table S1, with the number of primer sets. pDM1002 was kindly gifted by Dr D. R. Mills (48). All the constructs were confirmed by the dideoxynucleotide chain termination sequencing method.

pLacZα(–). The pASK-LacZα had 180 bases of the *lacZ* from the 5′ end in the pASK-IBA3 plus (IBA Lifesciences, Göttingen, Germany). The *lacZα* at the 5′ end was amplified by PCR using primers 1 and pASK-LacZα as a template, and the obtained PCR products were digested with XhoI. The digested fragment was inserted into the XhoI site of pDM1002.

pLacZα(–)_I-SceI_I-SceI site. An I-SceI recognition site was introduced into pLacZα(–) by PCR. PCR was conducted with primers 2 and 3, and the products were circularized using the In-Fusion® HD cloning kit (Takara Bio Inc., Japan). The derived plasmid was designated as pLacZα(–)_I-SceI site. The *tetR* from pASK-IBA3plus was inserted into pLacZα(–)_I-SceI site via BspHI digestion and ligation by T4 DNA ligase. The obtained plasmid was designated as pLacZα(–)_TetR_I-SceI site. The *I-SceI* derived from pCBASceI (#26477, Addgene, MA, USA) was amplified by PCR using primers 6. Two PCR fragments amplified by PCR using primers 4 and 5 and a PCR fragment containing the *I-SceI* were circularized using the In-Fusion® HD cloning kit. A duplicated sequence was removed as follows. Two PCR fragments amplified by PCR using primers 7 and 8 were circularized by the In-Fusion® HD cloning kit. The obtained plasmid was designated as pLacZα(–)_I-SceI_I-SceI site.

pLacZα(–)_I-SceI. The I-SceI site was removed from pLacZα(–)_I-SceI_I-SceI site. Two PCR fragments amplified by PCR using primers 9 and 10 were circularized using the In-Fusion® HD cloning kit.

Pβ. The his-tag sequence was removed from ptrcHisA (Thermo Fisher Scientific, MA, USA), and the ampicillin-resistance (beta-lactamase) gene was replaced by a kanamycin resistance gene. The PCR fragment containing the β subunit gene was amplified using primers 11 and pBAD33βH (51) as a template, and the obtained PCR fragment was introduced into the StyI and EcoRI sites by the In-Fusion® HD cloning kit. This plasmid was designated as ptrcβ_pBR322ori. The replication origin was replaced by pSC101ori as follows. Two PCR fragments obtained by PCR using primers 12 and 13 and ptrcβ_pBR322ori as a template and one PCR fragment obtained by PCR using primers 14 and pCL1920 (accession number AB236930, National BioResource Project, Japan) as a template were circularized by the In-Fusion® HD cloning kit. The obtained plasmid was designated as pβ.

Pβ_I-SceI site. The I-SceI recognition site was introduced as follows. Two PCR fragments amplified using primers 15 and 16 and pβ as a template were circularized by the In-Fusion® HD cloning kit.

pMDVβ. The replication origin of pN96, which carries a β subunit gene inside the MDV-1 sequence (kindly provided by Dr Ichihashi of the University of Tokyo), was replaced by pSC101ori as follows. One PCR fragment amplified using primers 17 and pN96

as a template, and two PCR fragments amplified using primers 18 and 19 and pβ as a template were circularized by the In-Fusion® HD cloning kit. The obtained plasmid was designated as pMDVβ.

pMDVβ_I-I-SceI site. The I-SceI recognition site was introduced into pMDVβ as follows. Two PCR fragments amplified using primers 15 and 16 with an I-SceI recognition site in the primer and pMDVβ as a template were circularized by the In-Fusion® HD cloning kit. The obtained plasmid was designated as the pMDVβ_I-I-SceI site.

2.3 Culture conditions and curing of plasmid DNA

The *E. coli* MDS42™ *recA*⁻ Blue strain was cultured in a lactose minimum medium with shaking at 160 rpm at 37°C. For curing plasmid DNA, we passaged *E. coli* in a lactose medium containing 25 μM IPTG, 100 μg/ml Amp and 50 μg/ml Km. We transferred a portion of the culture into a lactose medium containing 25 μM IPTG, followed by culture for 1.5 h at 37°C. We induced I-SceI by 10 nM Dox and cultured the cells for another 4 h. Then, the cells were diluted with LB and plated on LB agar medium or LB + Amp agar medium. For curing the pβ_I-SceI site, we picked 42 colonies from the LB + Amp agar medium and tested the loss of the pβ_I-SceI site by checking that they grew on LB + Amp agar medium and that they did not grow on LB + Amp + Km agar medium. For curing both plasmids, i.e. pLacZα(–)_I-SceI_I-SceI site and pMDVβ_ I-SceI site, we picked 47 colonies from the LB agar medium and tested the loss of both plasmids by checking that they did not grow on LB + Amp + Km, LB + Amp or LB + Km agar medium and that they grew on LB medium.

2.4 Genome sequence

The genome sequences of the *E. coli* MDS42™*recA*⁻Blue strain and the Adapted MDS42Blue strain were determined. The genome was extracted using Dneasy Blood & Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's instruction. The genomic DNAs of the MDS42™recA⁻Blue strain and Adapted MDS42Blue strain were sequenced with Illumina NovaSeq 6000 (Illumina Illumina, San Diego, CA) using 150-bp paired-end format by Rhelixa (Tokyo, Japan). The mean depth of MDS42™*recA*⁻Blue and Adapted MDS42Blue was 383.5 and 462.7, respectively. Low-quality bases and adapter sequences were removed using Cutadapt ver 4.1 (52). The reads that passed the quality check were aligned to reference sequence MDS42™ (GenBank accession number: AP012306.1) using BWA version 0.7.17-r1188 (53). Duplicate reads were identified using Picard version 2.27.4-SNAPSHOT (http://broadinstitute.github.io/picard/); subsequently, the single nucleotide polymorphisms and indels were detected using GATK version 4.2.6.1 (54). The effect of each variant was predicted using snpEFF version 5.1d (55).

2.5 Reverse transcription-quantitative polymerase chain reaction

The *in vitro* transcription (IVT) MDV-β RNA was used as the standard to calculate RNA molecules per cell. The IVT MDV-β RNA was synthesized using the MEGAscript™ T7 Transcription kit (Thermo Fisher Scientific Inc., MA, USA) and purified using the MEGAclear™ Transcription Clean-Up kit (Thermo Fisher Scientific Inc.). All the procedures for IVT synthesis were performed according to the manufacturer's instructions. The concentration of IVT MDV-β RNA was measured using Qubit™ RNA Broad

Range Assay Kit (Thermo Fisher Scientific Inc.). The molecular weight of IVT MDV-β was 655,439.2 calculated using the formula: An (457) × 329.2 + Un (559) × 306.2 + Cn (504) × 305.2 + Gn (521) × 345.2 + 159 (Thermo Fisher Scientific Inc.).

According to the manufacturer's manual, the total RNA was isolated using SV Total RNA Isolation System (Promega, WI, USA). We added Adapted MDS42Blue without plasmid DNA as a coprecipitant to harvest the small amount of cells. The amount of MDV-β was measured using THUNDERBIRD® Probe One-step reverse transcription-quantitative polymerase chain reaction (RT-qPCR) Kit (Toyobo Inc., Osaka, Japan). The primer for complementary DNA (cDNA) synthesis was cDNA_MDV_r (5′–TAAGCGAAT GTTGCGAGCACTGATCAGCAAAATCGCAACCC–3′), and primers for qPCR were real_MDV_f (5′–GACCATACAAGGGACCAAGAG–3′), real time_r1 (5′–TAAGCGAATGTTGCGAGCAC–3′) and a dual-labeled probe (5′–/56-FAM/CGGAAAGTA/ZEN/ACGATGGGTTGC CTCT/3IABkFQ/–3′) (Integrated DNA Technologies, IA, USA). We calculated the cell number using optical density at 600 nm $(OD_{600}) = 1$, calculated as 1×10^8 cells/ml. This conversion factor was determined using a known concentration of fluorescent latex beads (2 μm diameter, Polyscineces, Inc.) and *E. coli* cells using flow cytometry. The RNA molecules per cell without the pMDVβ_I-SceI site were calculated as follows. First, we calculated MDV–β RNA molecules per cell (α) with pMDVβ_I-SceI site cultured in lactose minimum medium with Km. Second, we calculated MDV–β RNA molecules per cell (β) cultured in lactose medium with 20 nM Dox and without Km. After obtaining the fraction of cells with a pMDVβ_I-SceI site that remained in the culture with 20 nM Dox and without Km (γ) by plate assay, the MDV–β RNA molecules per cell without pMDVβ_I-SceI site were obtained as β – γ ×α.

3. Results

3.1 *E. coli* growth system dependent on Qβ replicase

To maintain the mini-RNA replicon encoding the β subunit of Qβ replicase stably in *E. coli* cells, we constructed another RNA molecule that allowed the cells with Qβ replicase to grow selectively. As described in Section 1, we replaced the λN protein in the system prepared by Mills with the α fragment of LacZ for α complementation (56). The complementary strand of the sequence encoding the α fragment of LacZ was sandwiched with two recognition sites for Qβ replicase and placed under the *trp* promoter (pLacZα(–)) so that the complementary strand of the coding sequence of the *lacZα* (LacZα(–)RNA) was transcribed (Figure 2a). The transcribed LacZα(–)RNA was replicated by Qβ replicase into its complementary strand (LacZα(+)RNA) to synthesize α fragment of LacZ. The translated fragment was associated with the ω fragment from the *E. coli* genome to allow the cells to grow in a lactose minimal medium. Qβ replicase was constituted by the β subunit supplied from the other plasmid, pβ, together with EF-Tu, EF-Ts and S1 from the *E. coli* genome.

E. coli cells with pLacZα(–) and pβ grew in the lactose minimal medium (Figure 2b). As a control experiment, we cultured the cells with pLacZα(–) and the pΔβ plasmid; however, they did not grow in the lactose minimal medium. Furthermore, we introduced a vector that did not have the *lacZα* (empty vector). The cells with the empty vector and pβ or empty vector and pΔβ did not grow in the lactose minimal medium. This result indicated that pLacZα(–) afforded Qβ-replicase-dependent growth to *E. coli*.
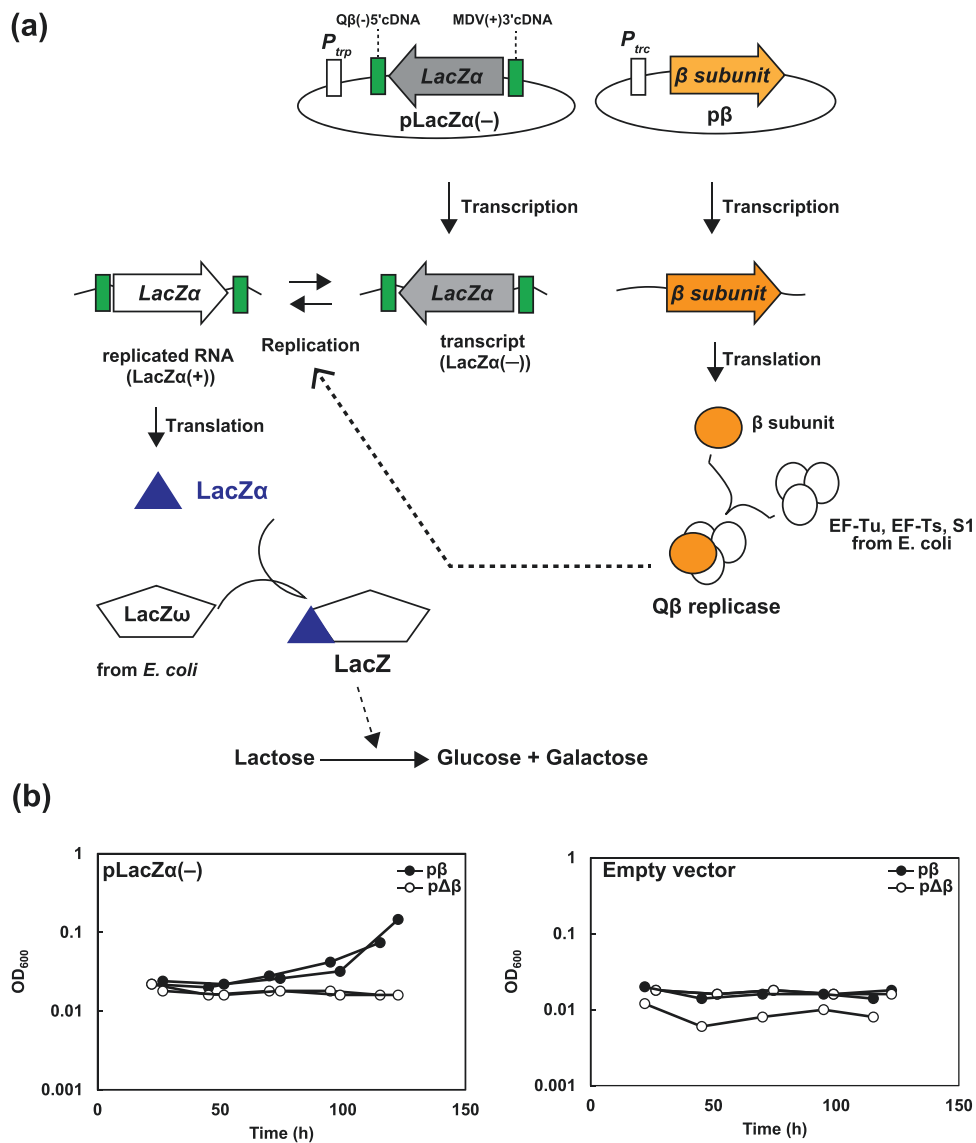
Figure 2. The *E. coli* growth dependent on Qβ replicase. (a) The operation principle of *E. coli* growth in the lactose minimal medium by Qβ replicase. *E. coli* was transformed with double plasmids, pLacZα(–) and pβ. The complementary strand of the LacZα-encoding gene with Qβ replicase recognition sites at both ends (LacZα(–)RNA) was transcribed by RNA polymerase of *E. coli*. The β subunit was produced from pβ and constituted Qβ replicase by assembling with three subunits from *E. coli*. The LacZα-encoding RNA strand (LacZα(+)RNA) was obtained by replication by Qβ replicase using the LacZα(–)RNA as a template. The LacZα protein was translated from the LacZα(+)RNA and assembled with LacZω from *E. coli* to produce LacZ (α complementation). Then, *E. coli* could grow in the lactose minimal medium. (b) Left: The growth of *E. coli* with pLacZα(–) and pβ or pLacZα(–) and pΔβ that did not have β subunit gene. Right: The growth of *E. coli* with empty vector (without LacZα(–)) and pβ or empty vector and pΔβ.

3.2 Construction of the mini-RNA replicon

We placed the β subunit gene in the middle of a short sequence replicable with Qβ replicase (MDV-1) ([49](#), [50](#)) to synthesize the mini-RNA replicon, so that the RNA (MDVβ(+)RNA) transcribed from the plasmid, pMDVβ, produced the β subunit. This could constitute Qβ replicase with EF-Tu, EF-Ts and S1 to replicate MDVβ(+)RNA into MDVβ(–)RNA and vice versa, thus generating the mini-RNA replicon in our system.

We replaced pβ in the system depicted in [Figure 2a](#) with pMDVβ to confirm whether MDVβ(+)RNA worked as designed. Unfortunately, *E. coli* cells with pMDVβ did not grow in the lactose minimal medium. This unexpected failure could be explained by the insufficient expression of the β subunit gene sandwiched by MDV-1 or by a burden from the replication of MDVβ(–)RNA or MDV(+)RNA. However, the true cause of this problem remains to be investigated.

3.3 Preparation of the host cell strain for the mini-RNA replicon

We obtained an *E. coli* strain that exhibited high growth with MDVβ(+)RNA and LacZα(–)RNA through two passage experiments to circumvent this problem. For the first-passage experiment, we introduced the *I-SceI* downstream of the *tetA* promoter and an I-SceI recognition site into pLacZα(–), designated as pLacZα(–)_I-SceI_I-SceI site ([Figure 3a](#)). We also introduced an I-SceI recognition site into pβ, which was designated as the pβ_I-SceI site. The *E. coli* cells with the two plasmids were passaged four times to increase their growth rate in the lactose minimal medium. After the growth rate adaptation, followed by induction of I-SceI, we picked colonies that grew on LB + Amp agar medium but did not grow on LB +Km agar medium, to obtain the adapted strain only with pLacZα (–)_I-SceI_I-SceI site. Before the second-passage experiment, the adapted strain with pLacZα(–)_I-SceI_I-SceI site
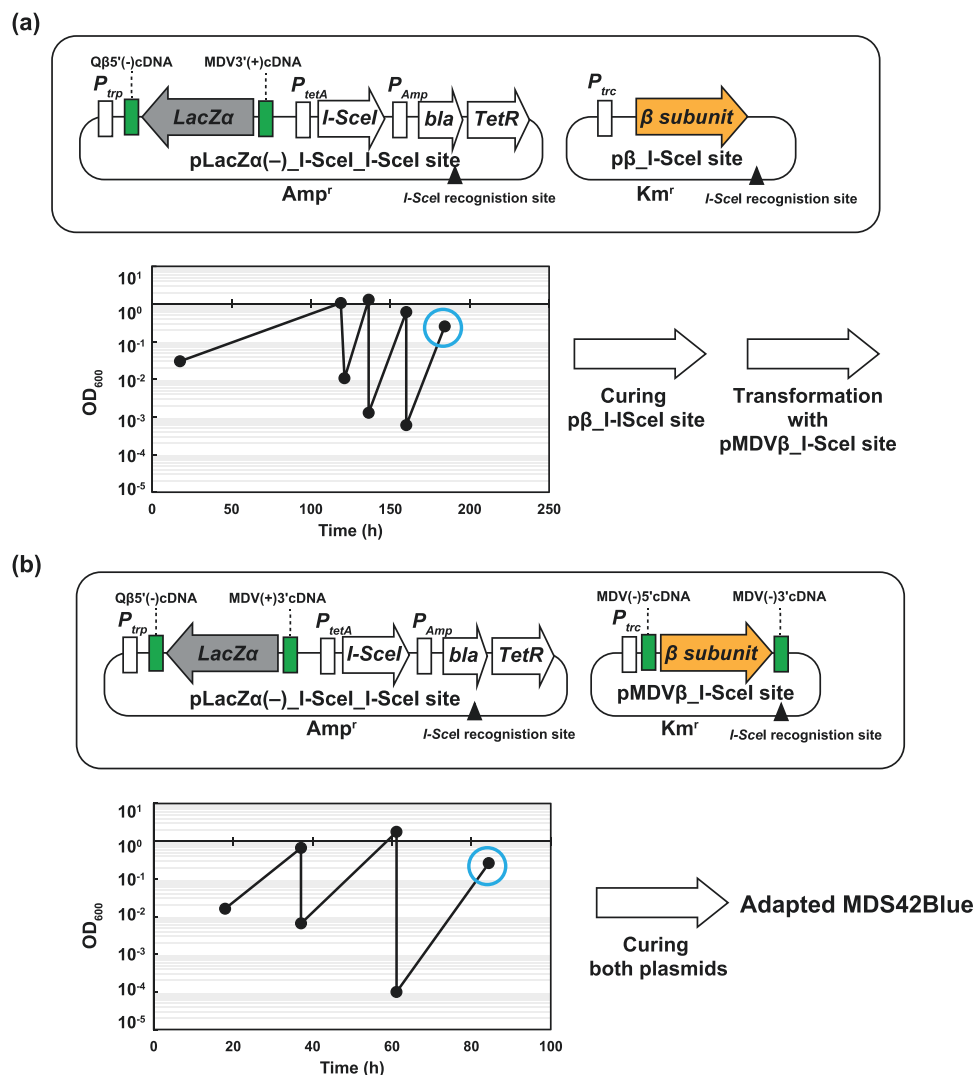
Figure 3. The adaptation of *E. coli* with pLacZα(−) and pMDVβ_I-SceI site to the lactose minimal medium. (a) *E. coli* with pLacZα(−)_I-SceI_I-SceI site and pβ_I-SceI site was passaged in lactose medium four times. The pβ_I-SceI site was cured by endonuclease I-SceI upon Dox induction at the time indicated by the circle. (b) The cells were transformed with pMDVβ_I-SceI site and passaged in the lactose medium three times. pLacZα(−)_I-SceI_I-SceI site and pMDVβ_I-SceI site were cured upon Dox induction at the time indicated by the circle.

was transformed with pMDVβ_I-SceI site. The pMDVβ_ISceI site was constructed by introducing an I-SceI recognition site into pMDVβ (Figure 3b). Throughout three passages in lactose minimal medium, the cells with the two plasmids showed a high growth rate, indicating the circumvention of the problem of the lack of growth of the original cells with MDVβ(+)RNA and LacZα(−)RNA. After the induction of I-SceI for curing both plasmids, we picked colonies that could not grow on LB + Amp agar medium or LB + Km agar medium but grew on LB agar medium. Thus, we obtained an adapted *E. coli* strain that could be used as a host cell for the mini-RNA replicon in our system (designated as Adapted MDS42Blue).

A whole-genome sequence was conducted using original and Adapted MDS42Blue to determine the mutation sites of Adapted MDS42Blue. Four point mutations were observed; two were nonsynonymous, and two were on intergenic regions (Table S2). One of the synonymous mutations was a substitution in 34 bases upstream of *lacZω*, and one of the nonsynonymous mutations was a substitution of chaperon *groL* that assists protein folding.

3.4 Persistent coexistence of the mini-RNA replicon and the DNA replicon in *E. coli*

We investigated whether the mini-RNA replicon was replicated in the adapted *E. coli* strain, Adapted MDS42Blue. We introduced pMDVβ_I-SceI site and pLacZα(−)_I-SceI into Adapted MDS42Blue (Figure 4a). pLacZα(−)_I-SceI was constructed by eliminating the I-SceI site from pLacZα(−)_I-SceI_I-SceI site (Figure S1). First, we passaged the cells with the two plasmids in the lactose minimal medium without the induction of I-SceI (Figure 4b, black dots). We divided the culture into two tubes and added Dox (for the induction of I-SceI) into one of the two tubes. After 24 h of cultivation, we confirmed that only 0.07–0.08% of the cells held pMDVβ_I-SceI site by taking the ratio of colony forming unit (CFU) on LB + Amp + Km agar medium over CFU on LB + Amp agar medium (Figure 4c at 119 h). Although ∼99.9% of the cells lost pMDVβ_I-SceI site, they showed a high growth rate of $0.16 \pm 0.01 \, h^{-1}$ ($n = 7$), comparable to the growth rate of the cells without the induction of I-SceI $0.36 \pm 0.02 \, h^{-1}$ ($n = 5$), implying that MDVβ(+)RNA was kept without its continuous transcription from the pMDVβ_I-SceI site. We measured the MDVβ(+)RNA

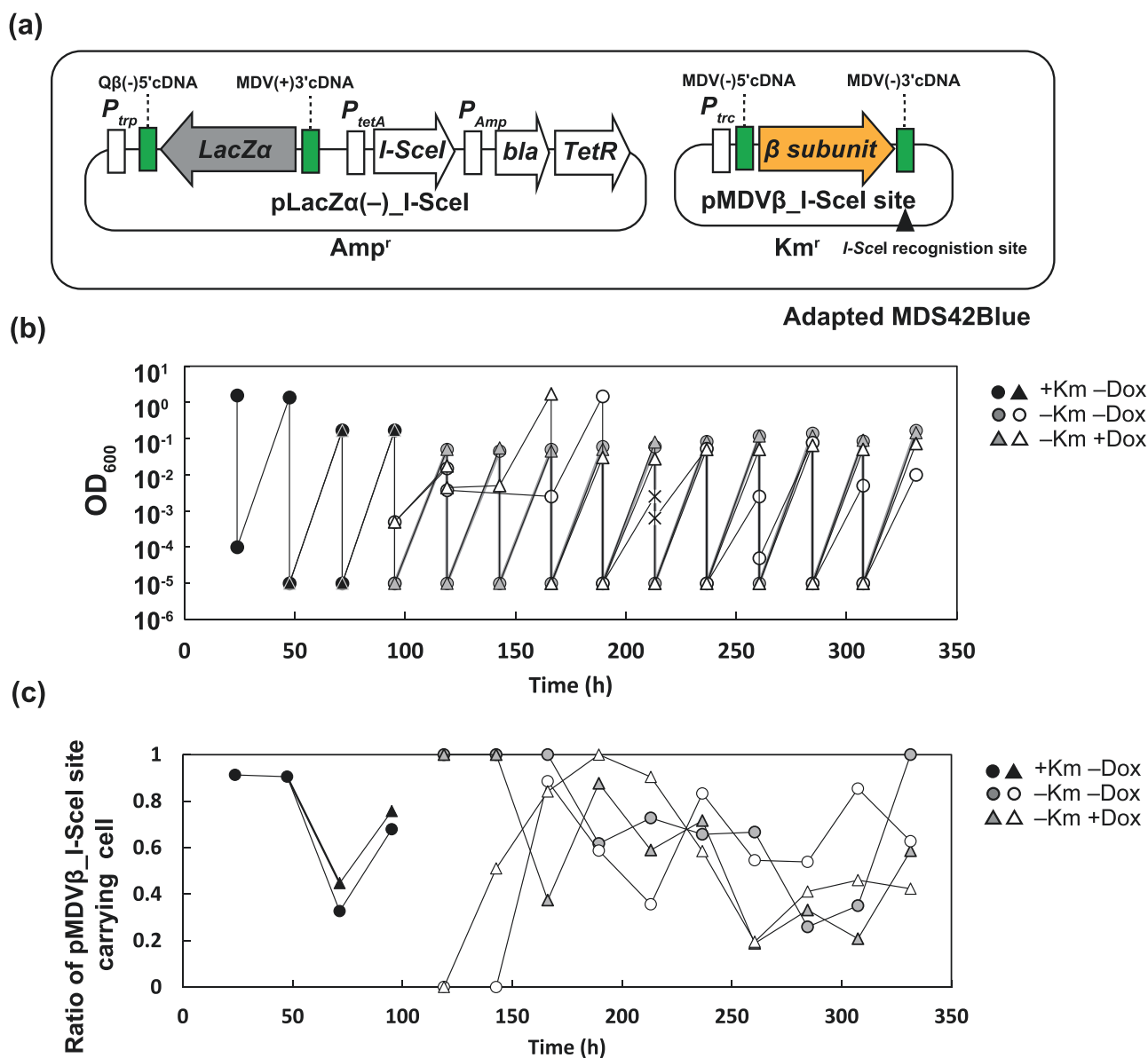**Figure 4.** The growth of *E. coli* with the mini-RNA replicon in the lactose minimal medium. (a) Adapted MDS42Blue was transformed with pLacZα(–)_I-SceI and pMDVβ_I-SceI site. (b) Cells were passaged in with Km in two lines and divided into two, respectively. One was treated with 20 nM Dox (white and gray triangles) and the other was not (white and gray circles). The (×) shows the detection limit (OD$_{600}$ = 0.0025) and the one-fourth dilution by passage. (c) The ratio of pMDVβ_I-SceI site-carrying cells. The ratio of CFU grown on LB with Km to CFU grown on LB with Amp and Km is shown.

molecules per cell by RT-qPCR. The RNA molecules of cells with pMDVβ_I-SceI site were 980 ± 310 molecules/cell ($n = 3$), and the RNA molecules of cells without pMDVβ_I-SceI site were 23 ± 24 molecules/cell ($n = 3$). The coexistent state of the mini-RNA replicon and DNA replicon in *E. coli* was successfully synthesized.

We confirmed the persistence of the mini-RNA replicon in *E. coli*. We performed an additional eight to nine passages with and without Dox induction of I-SceI (Figure 4b and c). Regardless of the induction, the fraction of cells without the pMDVβ_I-SceI site reached 60% at the end of the passage experiments. This may have stemmed from imperfect plasmid curing by I-SceI and the burden of the presence of the plasmid. The convergence of the ratio indicated that the cells without the pMDVβ_I-SceI site showed the same specific growth rate as that observed with the pMDVβ_I-SceI site, implying that MDVβ(–)RNA without continued transcription

from pMDVβ_I-SceI was kept during at least nine passages at the same concentration as in the cells with transcription from the plasmid. The coexistence of the mini-RNA replicon with the DNA replicon in *E. coli* cells was persistent.

4. Discussion

In this study, we constructed a persistent coexistence state of a mini-RNA replicon and DNA replicon in *E. coli*. The mini-RNA replicon encoding the β subunit of Qβ replicase was replicated by Qβ replicase in *E. coli*. To maintain the mini-RNA replicon persistently in *E. coli* cells, we constructed a system of α complementation that was dependent on Qβ replicase, thus allowing the *E. coli* cells with the mini-RNA replicon to grow in the lactose minimal medium selectively.

We modified Mills' Qβ replicase-dependent *E. coli* growth system to allow the cells with our mini-RNA replicon selectively grow through α complementation. Mills constructed a system in which *E. coli* grew depending on Qβ replicase (48). He inserted the complementary RNA strand of the N protein gene between 145 bases of the complementary strand of the 3′ end of the Qβ genome sequence and 147 bases of the MDV sequence. The system encompassed another two plasmids; one encoded the *β subunit* of Qβ replicase, and the other carried a *cat* located downstream of a transcription terminator. After the replication of the complementary RNA in the presence of Qβ replicase made of a β subunit supplied from the first plasmid together with EF-Tu, EF-Ts and S1 from the *E. coli* genome, the N protein was translated. With the translated N protein, the mRNA of *cat* was transcribed from the second plasmid to produce CAT, thus allowing Qβ replicase-dependent cell growth in the presence of chloramphenicol.

It should be noted that Mills' RNA is not an RNA replicon because it does not encode any protein for its replication. One problem in Mills' system is that the *E. coli* cells quickly become chloramphenicol resistant after a few passages. To circumvent this problem, we replaced the N protein with the α fragment of LacZ to allow the cells with our mini-RNA replicon to grow stably. The cells with our mini-RNA replicon were stably passaged for >10 days in the lactose minimal medium. We confirmed that no mutations that damage our Qβ replicase-dependent cell growth system appeared on the *E. coli* genome during the passages (data not shown).

We synthesized a mini-RNA replicon encoding only the β subunit of Qβ replicase. Some RNA replicon-type RNA vectors have been developed by modifying persistently infecting RNA virus and phage genomes (37–43). The constructed coexistent state of the mini-RNA replicon and the DNA genome in *E. coli* cells is suitable for research on the 'Origin of life' because of its simplicity compared with the previously developed RNA replicon-type RNA vectors. We can observe the evolutionary dynamic of the coexistent state with passages of many generations and whole-genome sequencing of both the RNA replicon and the DNA replicon (*E. coli* genome).

The next challenge in our coexistence state is the expansion of the size of the RNA replicon. Two routes of expansion are possible. One consists of inserting several genes from the *E. coli* genome into the mini-RNA replicon obtained in this study. The other way consists of modifying persistent or pseudolysogenized Qβ phage mutants in the same way as for developing RNA replicon-type RNA vectors. It is known that Qβ phage enters a pseudolysogeny state if it loses the function of the A2 protein for cell lysis (57). After a substantial expansion of the size of the RNA replicon, we will be able to obtain insights regarding what happened on the long RNA replicon in the ancient RNP world that coexisted with the DNA world.

Using the coexistence state of the mini-RNA replicon and DNA replicon, it is possible to obtain insight into the transition process from the RNA world to the DNA world. Questions related to the transition process, such as how did some of the functional information on the RNA replicon move to the DNA replicon and to what extent were the suitable mutation rates of the RNA replicon and DNA replicon dedicated to enhancing the transition process or keeping the coexistence state, can be addressed using the approaches of experimental evolution and synthetic biology. Addressing these questions experimentally will contribute to progress in the research field of the origin of life.

Supplementary data

Data availability

Material availability

Funding

Acknowledgments

References

1. Gilbert,W. (1986) The RNA world. *Nature*, **319**, 618.
2. Joyce,G.F. and Orgel,L.E. (1999) Prospects for understanding the origin of the RNA world. In: Gesteland FR, Cech TR, Atkins JF (eds). *The RNA World: The Nature of Modern RNA Suggests a Prebiotic RNA*. Cold Spring Harbor Laboratory Press, New York, pp. 49–77.
3. Jeffares,D.C., Poole,A.M. and Penny,D. (1998) Relics from the RNA world. *J. Mol. Evol.*, **46**, 18–36.
4. Poole,A.M., Jeffares,D.C. and Penny,D. (1998) The path from the RNA world. *J. Mol. Evol.*, **46**, 1–17.
5. Kun,A., Szilagyi,A., Konnyu,B., Boza,G., Zachar,I. and Szathmary,E. (2015) The dynamics of the RNA world: insights and challenges. *Ann. N Y Acad. Sci.*, **1341**, 75–95.
6. Woese,C. (1967) The genetic code: the molecular basis for genetic expression. Harper & Rwo Pub., New York, pp. 179–195.
7. Crick,F.H. (1968) The origin of the genetic code. *J. Mol. Biol.*, **38**, 367–379.
8. Orgel,L.E. (1968) Evolution of the genetic apparatus. *J. Mol. Biol.*, **38**, 381–393.
9. Kruger,K., Grabowski,P.J., Zaug,A.J., Sands,J., Gottschling,D.E. and Cech,T.R. (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell*, **31**, 147–157.
10. Guerrier-Takada,C., Gardiner,K., Marsh,T., Pace,N. and Altman,S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, **35**, 849–857.
11. Prody,G.A., Bakos,J.T., Buzayan,J.M., Schneider,I.R. and Bruening,G. (1986) Autolytic processing of dimeric plant virus satellite RNA. *Science*, **231**, 1577–1580.

12. Hutchins,C.J., Rathjen,P.D., Forster,A.C. and Symons,R.H. (1986) Self-cleavage of plus and minus RNA transcripts of avocado sunblotch viroid. *Nucleic Acids Res.*, **14**, 3627–3640.

13. Orgel,L.E. (2004) Prebiotic chemistry and the origin of the RNA world. *Crit. Rev. Biochem. Mol. Biol.*, **39**, 99–123.

14. Forterre,P. (2005) The two ages of the RNA world, and the transition to the DNA world: a story of viruses and cells. *Biochimie*, **87**, 793–803.

15. Robertson,M.P. and Joyce,G.F. (2012) The origins of the RNA world. *Cold Spring Harb. Perspect. Biol.*, **4**, a003608.

16. Doherty,E.A. and Doudna,J.A. (2000) Ribozyme structures and mechanisms. *Annu. Rev. Biochem.*, **69**, 597–615.

17. Krishnamurthy,R. (2017) Giving rise to life: transition from prebiotic chemistry to protobiology. *Acc. Chem. Res.*, **50**, 455–459.

18. Xu,J., Chmela,V., Green,N.J., Russell,D.A., Janicki,M.J., Gora,R.W., Szabla,R., Bond,A., Sutherland,J. *et al.* (2020) Selective prebiotic formation of RNA pyrimidine and DNA purine nucleosides. *Nature*, **582**, 60–66.

19. Muller,F., Escobar,L., Xu,F., Wegrzyn,E., Nainyte,M., Amatov,T., Chan,C.-Y., Pichler,A., Carell,T. *et al.* (2022) A prebiotically plausible scenario of an RNA-peptide world. *Nature*, **605**, 279–284.

20. Walker,D.L. (1964) The viral carrier state in animal cell cultures. *Prog. Med. Virol.*, **6**, 111–148.

21. Rustigian,R. (1966) Persistent infection of cells in culture by measles virus. I. Development and characteristics of HeLa sublines persistently infected with complete virus. *J. Bacteriol.*, **92**, 1792–1804.

22. Holland,J.J. and Villarreal,L.P. (1974) Persistent noncytocidal vesicular stomatitis virus infections mediated by defective T particles that suppress virion transcriptase. *Proc. Natl. Acad. Sci. U S A*, **71**, 2956–2960.

23. Rima,R.K. and Martin,S.J. (1976) Persistent infection of tissue culture cells by RNA viruses. *Med. Microbiol. Immunol.*, **162**, 89–119.

24. Oldstone,M.B.A., Buchmeier,M.J. and Doyle,M.V. (1978) Immunobiology of persistent virus infection. In: Stevens JG, Todaro GJ, Fox CF (eds). *Persistent Viruses*. Academic Press, New York, pp. 43–56.

25. Friesen,P., Scotti,P., Longworth,J. and Rueckert,R. (1980) Black beetle virus: propagation in *Drosophila* line 1 cells and an infection-resistant subline carrying endogenous black beetle virus-related particles. *J. Virol.*, **35**, 741–747.

26. Roux,L. and Holland,J.J. (1979) Role of defective interfering particles of Sendai virus in persistent infections. *Virology*, **93**, 91–103.

27. Vignuzzi,M. and Lopez,C.B. (2019) Defective viral genomes are key drivers of the virus-host interaction. *Nat. Microbiol.*, **4**, 1075–1087.

28. Holland,J.J., Grabau,E.A., Jones,C.L. and Semler,B.L. (1979) Evolution of multiple genome mutations during long-term persistent infection by vesicular stomatitis virus. *Cell*, **16**, 495–504.

29. Brown,G.G. and Finnegan,P.M. (1989) RNA plasmids. *Int. Rev. Cytol.*, **117**, 1–56.

30. Wakarchuk,D.A. and Hamilton,R.I. (1990) Partial nucleotide sequence from enigmatic dsRNAs in *Phaseolus vulgaris*. *Plant Mol. Biol.*, **14**, 637–639.

31. Bruenn,J.A. (1980) Virus-like particles of yeast. *Annu. Rev. Microbiol.*, **34**, 49–68.

32. Ghabrial,S.A. and Suzuki,N. (2009) Viruses of plant pathogenic fungi. *Annu. Rev. Phytopathol.*, **47**, 353–384.

33. Ghabrial,S.A., Caston,J.R., Jiang,D., Nibert,M.L. and Suzuki,N. (2015) 50-plus years of fungal viruses. *Virology*, **479–480**, 356–368.

34. Davern,C.I. (1964) The isolation and characterization of an RNA bacteriophage. *Aust J. Biol. Sci.*, **17**, 719–725.

35. Zgaga,V. (1977) Lysogeny by f2 phage? *Nature*, **267**, 860–862.

36. Romantschuk,M. and Bamford,D.H. (1981) φ6-resistant phage-producing mutants of *Pseudomonas phaseolicola*. *J. Gen. Virol.*, **56**, 287–295.

37. Kim,D.Y., Atasheva,S., McAuley,A.J., Plante,J.A., Frolova,E.I., Beasley,D.W., Frolov,I. *et al.* (2014) Enhancement of protein expression by alphavirus replicons by designing self-replicating subgenomic RNAs. *Proc. Natl. Acad. Sci. U S A*, **111**, 10708–10713.

38. Khakhar,A. and Voytas,D.F. (2021) RNA viral vectors for accelerating plant synthetic biology. *Front. Plant Sci.*, **12**, 668580.

39. Onodera,S., Olkkonen,V.M., Gottlieb,P., Strassman,J., Qiao,X.Y., Bamford,D.H., Mindich,L. *et al.* (1992) Construction of a transducing virus from double-stranded RNA bacteriophage φ6: establishment of carrier states in host cells. *J. Virol.*, **66**, 190–196.

40. Mindich,L. (1999) Reverse genetics of dsRNA bacteriophage φ6. *Adv. Virus Res.*, **53**, 341–353.

41. Schott,J.W., Morgan,M., Galla,M. and Schambach,A. (2016) Viral and synthetic RNA vector technologies and applications. *Mol. Ther.*, **24**, 1513–1527.

42. French,R., Janda,M. and Ahlquist,P. (1986) Bacterial gene inserted in an engineered RNA virus: efficient expression in monocotyledonous plant cells. *Science*, **231**, 1294–1297.

43. Takamatsu,N., Ishikawa,M., Meshi,T. and Okada,Y. (1987) Expression of bacterial chloramphenicol acetyltransferase gene in tobacco plants mediated by TMV-RNA. *EMBO J.*, **6**, 307–311.

44. Agapov,E.V., Frolov,I., Lindenbach,B.D., Pragai,B.M., Schlesinger,S. and Rice,C.M. (1998) Noncytopathic Sindbis virus RNA vectors for heterologous gene expression. *Proc. Natl. Acad. Sci .U S A*, **95**, 12989–12994.

45. Nishimura,K., Segawa,H., Goto,T., Morishita,M., Masago,A., Takahashi,H., Ohmiya,Y., Sakaguchi,T., Asada,M., Imamura,T. *et al.* (2007) Persistent and stable gene expression by a cytoplasmic RNA replicon based on a noncytopathic variant Sendai virus. *J. Biol. Chem.*, **282**, 27383–27391.

46. Onodera,S., Qiao,X., Gottlieb,P., Strassman,J., Frilander,M. and Mindich,L. (1993) RNA structure and heterologous recombination in the double-stranded RNA bacteriophage φ6. *J. Virol.*, **67**, 4914–4922.

47. Sun,Y., Qiao,X. and Mindich,L. (2004) Construction of carrier state viruses with partial genomes of the segmented dsRNA bacteriophages. *Virology*, **319**, 274–279.

48. Mills,D.R. (1988) Engineered recombinant messenger RNA can be replicated and expressed inside bacterial cells by an RNA bacteriophage replicase. *J. Mol. Biol.*, **200**, 489–500.

49. Mills,D.R., Kramer,F.R. and Spiegelman,S. (1973) Complete nucleotide sequence of a replicating RNA molecule. *Science*, **180**, 916–927.

50. Nishihara,T., Mills,D.R. and Kramer,F.R. (1983) Localization of the Qβ replicase recognition site in MDV-1 RNA. *J. Biochem.*, **93**, 669–674.

51. Kita,H., Cho,J., Matsuura,T., Nakaishi,T., Taniguchi,I., Ichikawa,T., Shima,Y., Urabe,I., Yomo,T. *et al.* (2006) Functional Qβ replicase genetically fusing essential subunits EF-Ts and EF-Tu with beta-subunit. *J. Biosci. Bioeng.*, **101**, 421–426.

52. Martin,M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*, **17**, 10–12.

53. Li,H. and Durbin,R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.

54. McKenna,A., Hanna,M., Banks,E., Sivachenko,A., Cibulskis,K., Kernytsky,A., Garimella,K., Altshuler,D., Gabriel,S., Daly,M. *et al.* (2010) The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.*, **20**, 1297–1303.

55. Cingolani,P., Platts,A., Wang,L.L., Coon,M., Nguyen,T., Wang,L., Land,S.J., Lu,X., Ruden,D.M. *et al.* (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)*, **6**, 80–92.

56. Welply,J.K., Fowler,A.V. and Zabin,I. (1981) beta-Galactosidase alpha-complementation. Effect of single amino acid substitutions. *J. Biol. Chem.*, **256**, 6811–6816.

57. Watanabe,I., Sakurai,T., Furuse,K. and Ando,A. (1979) "Pseudolysogenization" by RNA phage Qβ. *Microbiol. Immunol.*, **23**, 1077–1083.