Minireview

Replicable and recombinogenic RNAs

Alexander B. Chetverin*

Institute of Protein Research, Russian Academy of Sciences, 142290 Pushchino, Moscow, Russia

Received 3 March 2004; accepted 7 March 2004

Available online 7 April 2004

Edited by Horst Feldmann

Abstract This paper summarizes results of the 40-year studies on replication and recombination of RNA molecules in the cell-free amplification system of bacteriophage Q. Special attention is paid to the molecular colony technique that has provided for the discovery of the nature of "spontaneous" RNA synthesis by Q replicase and of the ability of RNA molecules to spontaneously rearrange their sequences under physiological conditions. Also discussed is the impact of these data on the concept of RNA World and on the development of new in vitro cloning and diagnostic tools.

Keywords: RNA replication; RNA recombination; Qβ replicase; Molecular colonies; RNA World; Molecular diagnostics

1. Introduction

The 40th anniversary of FEBS is a good occasion to commemorate two remarkable events that also have occurred four decades ago: the first report on the exchange of genetic information at the RNA level and the first cell-free exponential amplification of a nucleic acid.

In the beginning of the 1960s, Hirst [1] and Ledinko [2] reported on the exchange of genetic markers between related strains of poliovirus whose genome is composed of RNA. Although it was not precisely known at that time if the entire poliovirus genome consists of one RNA molecule and if the involvement of DNA intermediates can be absolutely excluded, these reports are commonly cited as marking the discovery of intermolecular RNA recombination, i.e., an exchange between RNA molecules with their segments. Anyway, those reports have made RNA recombination a matter of experimental research. At about the same time, the first RNA "replicase" (an enzyme capable of synthesizing RNA on an RNA template) was discovered in the lysates of *E. coli* cells infected with an RNA bacteriophage [3] and, slightly later, the exponential amplification of RNA was demonstrated in a cell-free system [4]. That system employed Qβ replicase, the RNA-dependent RNA polymerase of bacteriophage Qβ, and up to date it remains the only cell-free system capable of exponentially amplifying RNA. The two discoveries occurred so synchronously that it looked as if the providence delivered to the researchers a hint to use the cell-free amplification system for studies on RNA recombination. Ironically, the two fields remained in mutual ignorance for a quarter of century. A reason for this fact was probably the failure to detect recombination between RNA phages in experiments analogous to those performed with poliovirus, resulting in a belief that no RNA recombination is possible in a prokaryotic system [5]. This view persisted until 1988, when a natural Qβ replicase-amplifiable RNA was found whose recombinant origin was unequivocally demonstrated by sequencing [6].

This paper gives an updated review of in vitro studies on RNA replication and recombination in the Qβ amplification system and shows how their results help to imagine a likely scenario of the evolution in the RNA World, to unveil the role of RNA recombination in the contemporary DNA World, and to develop new genetic and diagnostic tools. A detailed review of earlier data can be found elsewhere [7].

2. Qβ replicase templates

In addition to the 4217 nt-long genomic Qβ RNA, Qβ replicase amplifies a number of RQ RNAs (termed so for being *R*eplicable by Qβ replicase), which are usually $\leqslant 250$ nt in length. The natural source of RQ RNAs is Qβ phage-infected *E. coli* cells [8,9] or Qβ phage itself [10]. Recently, many new RQ RNAs have been selected from random [11] or artificially designed [12] sequences, or produced by in vitro RNA recombination [13,14]. Approximately, 10^4 copies of a single genomic RNA molecule are produced in a Qβ phage infected *E. coli* cell in less than 1 h [15]. Amplification of small RQ RNAs is much faster: up to 10^{10} copies are produced at room temperature within 10 min in a cell-free system composed of purified Qβ replicase and all four rNTPs [7], and this is the absolute record of the rate of nucleic acid amplification. Since its discovery, there had been numerous attempts to utilize this extremely powerful cell-free amplification system for the amplification of desired RNAs, such as mRNAs or ribozymes, but that proved not an easy task, primarily because the mechanism Qβ replicase uses to recognize its templates is not known. Like DNA amplification in the polymerase chain reaction (PCR), RNA is amplified by Qβ replicase exponentially. This means that in each round of replication, the number of

* Fax: +7-095-924-0493.
E-mail address: alexch@vega.protres.ru (A.B. Chetverin).

Abbreviations: RQ RNA, a non-genomic RNA capable of exponential amplification by Qβ replicase; nt, nucleotide(s); PCR, polymerase chain reaction; MCT, molecular colony technique

0014-5793/$22.00
doi:10.1016/j.febslet.2004.03.066

RNA molecules increases by roughly a factor of 2, as long as replicase is in molar excess over the RNA amplified. Unlike PCR, no oligonucleotide primers and no temperature cycling are required. This is because Qβ replicase synthesizes RNA in a primer-independent manner, the product strand and its template are rendered unannealed throughout the replication round, and both the strands are available to being used as templates in the next round that commences immediately upon termination. Furthermore, unlike common DNA-directed RNA polymerases, Qβ replicase does not utilize transcriptional promoters [7]. Yet, it manifests a very high degree of selectivity in choosing which RNA to amplify: only a few of 10^{12} unique RNA sequences of 50–77 nt in length flanked by the correct 5′- and 3′-terminal clusters are recognized as templates [11]. This provides a rationale for the fact that Qβ replicase does not amplify most natural RNAs, including any tested cellular RNAs or genomic RNAs of other viruses [7].

3. What makes RNA replicable?

Qβ replicase copies its cognate (replicable) RNAs beginning at the 3′-terminal oligo(C) cluster. The cluster is usually 3–4, sometimes 2 nucleotides long. At the 5′ terminus, there is a matching oligo(G) cluster whose role is believed to code for the oligo(C) at the 3′ end of the complementary product strand [16]. Except for these terminal clusters, too short to provide for the observed selectivity, no other sequence common to all known replicable RNAs has been found. Other structural similarities were also noted, such as internal 8–15 nt-long pyrimidine-rich segments [11,17], the unpaired 3′ end, and a hairpin involving the 5′ end [12,18], but artificial RNAs designed to accommodate these features turned out to be non-replicable [12,19].

Recently [20] we found that if RQ135 RNA (a replicable species having 135 nt in length [21]) was cleaved into two fragments, neither of them could be amplified, which was expected because the integrity of the template was lost and its terminal clusters occurred in separate molecules. However, unexpectedly, each fragment appeared to be capable of directing the synthesis of its respective complementary copy, although the initiator oligo(C) cluster resided on the 3′ fragment only. To explore the unusual template activity of the 5′ fragment, we prepared an array of its variants with altered initiation regions and found that every variant can be copied, although with varying efficiency. However, template properties of the 5′ fragment variants, even those bearing oligo(C) clusters at the 3′ end, turned out to be entirely different from the properties of "legitimate" templates, the intact RQ135 RNA and its 3′ fragment. Most importantly, in contrast to the legitimate templates, none of the 5′ fragment variants was capable of the GTP-dependent formation of a stable replicative complex capable of elongation in the presence of aurintricarboxylic acid, a powerful inhibitor of RNA protein interactions. This was, in spite of the fact that each of the 5′ fragment variants inherited a pyrimidine-rich segment of the RQ135 RNA [17] and many of them possessed the secondary structure elements thought to be diagnostic features of replicable RNAs [12,18].

Thus, although Qβ replicase can initiate and elongate on a variety of RNAs, only some of them are recognized as legitimate templates. The results further suggest that the diagnostic feature discriminating between legitimate and illegitimate templates is a GTP-dependent step in initiation that induces a "closed" conformation of the replicative complex, which does not dissociate until the product strand is completed. A plausible role of the closed conformation is to render the complementary template and nascent strands non-paired, thereby providing for the exponential RNA amplification [20]. The structural features of legitimate templates, which determine the Qβ replicase commitment to entering the closed conformation and, ultimately, its template specificity, remain to be elucidated.

4. RQ RNAs as amplification vectors

In the absence of knowledge of the replication mechanism, there were attempts to amplify heterologous sequences by using natural replicating RNAs as vectors. To this end, a foreign sequence was embedded into an internal loop of an RQ RNA in such a manner as to minimally disturb the RNA tertiary structure. The first successful attempt of this sort made in 1983 [22] was given a great deal of enthusiasm by being declared to be the birth of recombinant RNA technology [23]. Later, however, it was understood that there are severe constraints on the embedded sequence: it should be highly structured and be no longer than about hundred nucleotides [24,25]. Common mRNAs could be amplified within RQ vectors only in vivo [26] or in the presence of a coupled cell-free translation system [27]. In these cases, amplification becomes possible because translating ribosomes coat the coding (sense) strand, thereby preventing its annealing with the antisense strand. Otherwise, the complementary strands would collapse into duplex, resulting in immediate cessation of RNA synthesis. However, inasmuch as ribosomes occupy sense strands and read them in opposite direction than Qβ replicase does, only antisense strands remain available as replicase templates. Therefore, amplification becomes asymmetric and rather linear than exponential [27].

Also, there were attempts to employ RQ RNA vectors, carrying short insert complementary to a target, as replicable probes for diagnostic purposes. The background RNA synthesis caused by non-specifically bound probes was eliminated by introducing compound binary probes consisting of two fragments capable of producing the full-sized replicable RNA when ligated upon hybridization next to one another on a target molecule if this is present in the analyzed sample [28]. However, this approach has not become a routine assay because even short inserts often inhibit replication of RQ RNA vectors, and because binary RNA probes can self-recombine to produce replicable RNAs in the absence of any target or ligase (see below).

5. RNA recombination can occur in the Qβ system

For a long time it was held that no RNA recombination could occur in bacterial cells, as far as all attempts to detect recombinant progeny upon co-infection of *E. coli* cells with RNA bacteriophages carrying distinct genetic markers were unsuccessful [5]. This contrasted the ease of observation of genetic recombinations in poliovirus and some other animal [29,30] or plant [31] viruses.

The first unequivocal evidence of the occurrence of RNA recombination in prokaryotic systems has been the sequence of RQ120 RNA, which contains, at the 5′ end, an 80-nt segment of the coat protein cistron of phage Qβ RNA and, at the 3′ end, the 3′-terminal 33-nt segment of *E. coli* $tRNA_1^{Asp}$ [6]. Remarkably, RQ120 RNA as well as other recombinant RQ RNAs discovered later, appeared to be a product of crossing-over between non-homologous sequences at non-homologous sites [32], in a sharp contrast to picornaviruses (including the poliovirus) and coronaviruses that manifest predominantly homologous recombination, in which parental and progeny RNAs are homologous to each other around the cross-over site [29,30]. Moreover, when homologous recombination was finally demonstrated for phage Qβ in specially designed experiments [33], it turned out to be a million times less frequent than in poliovirus, suggesting that different mechanisms may operate in these viruses [34].

6. The puzzle of spontaneous RNA synthesis

Once the occurrence of RNA recombination in Qβ phage-infected bacteria was established, demonstration of RNA recombination in the cell-free replication system would seem a technical matter. For example, one might mix the non-replicable 5′ and 3′ fragments of RQ RNA discussed above, incubate them with Qβ replicase and rNTPs and see if there is exponential RNA synthesis indicating that the two fragments have recombined to restore the replicable molecule. However, experiments of this type could not be carried out because of intense RNA synthesis that invariably took place in the Qβ replicase system irrespectively of the addition of these fragments or of any other RNA template [35]. The template-independent synthesis was claimed to be caused by RQ RNAs contaminating Qβ replicase preparations [36], but this was disputed, because the synthesis still occurred when the enzyme portion used might not contain even one replicable molecule [35,37].

The latter observations gave rise to a hypothesis that Qβ replicase can produce replicable RNAs de novo, without any template, by virtue of random condensation of nucleotides, with the fortuitous formation of replicable molecules and their subsequent evolution into rapidly amplifiable species [35,38]. If so, the Qβ replicase replication system could serve as an experimental model wherein replicable molecules are created within the span of the reaction time (one to few hours). However, there were facts that could hardly be reconciled with the de novo hypothesis. Thus, many RNAs isolated from the products of spontaneous synthesis turned out to be either identical to RQ RNAs isolated earlier or recombinants consisting of long pieces of other known RNAs (the above mentioned RQ120 RNA was one of them) [7]. This suggested that synthesis of these RNAs was in fact template-instructed, but how did those templates occur in the reaction mixture?

7. Detection of airborne RQ RNAs with the molecular colony technique

A similar problem was solved in 1860s by Louis Pasteur [39] who disproved the doctrine of spontaneous generation of life from non-living organic matter by demonstrating that no life could arise in a boiled meat broth unless solid particles heavier than air were allowed to enter. Those experiments also convincingly demonstrated that microorganisms are everywhere – even in the air.

Unfortunately, Pasteur's approach could not be used in this case, since boiling or otherwise killing of RNA would also kill Qβ replicase. Yet the source of RNA templates could be found out if it was possible to precisely monitor the number of RQ RNA molecules in the reaction mixture. Such a possibility was provided by the molecular colony technique (MCT), invented in this laboratory [10]. The idea of experiment was to carry out RNA amplification in a gel, rather than in solution. In this format, the copies of each replicable molecule would concentrate around the original template, giving rise to a molecular colony. By counting the number of RNA colonies one could determine how many replicable molecules had been entrapped in the gel. An ideologically similar approach was used by Robert Koch in 1881 [40], who demonstrated the formation of bacterial colonies in solidified culture media.

Fig. 1 shows Petri dishes in which the experiments were carried out. In each dish, two agarose layers were cast one on top of the other. The lower layer contained rNTP substrates, and the upper layer, prepared at the specified time, contained Qβ replicase. The RNA colonies grew on the interface between the layers and were made visible by staining with ethidium bromide. It is seen that, if the enzyme layer was cast immediately after solidification of the substrate layer (Fig. 1A), the number of RNA colonies was lower, than if it was cast 1 h later (Fig. 1B). Even fewer colonies grew if the layers were cast in a room where no experiments with RQ RNAs had been previously carried out (Fig. 1C). These experiments demonstrated that, like Pasteur's germs, RQ RNAs invade the reaction

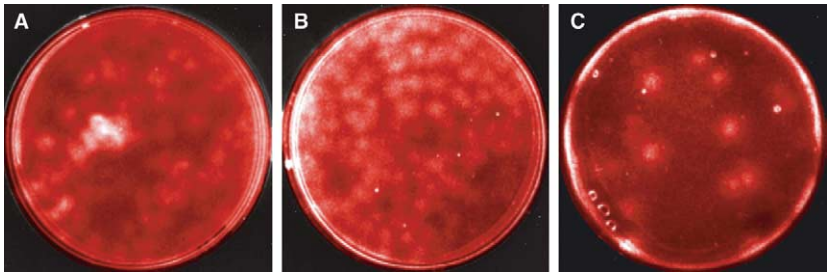

Fig. 1. Detection of airborne RQ RNAs using two-layer agarose sandwiches prepared in 35 mm Petri dishes [10]. The lower layer contained rNTPs, the upper layer contained Qβ replicase. The upper layer was cast without exposure (A, C) or after a one-hour exposure of the lower layer to air (B). The experiments were carried out in a room where RQ RNAs were often used (A, B) or in a remote room (C).

medium from air, and explained such puzzling features of the template-independent reactions as the failure to prevent RNA synthesis by exhaustive purification of Qβ replicase and the reproducible generation of the same RQ RNA species in independent experiments and in different laboratories [7].

The last argument of the advocates of the de novo hypothesis was that replicable RNA species arose even in sealed capillaries after periods of time much longer than would be needed for a single RQ RNA molecule to produce detectable progeny [18,41]. However, the sequence of at least some of the produced RNAs [18] appeared to be a mosaic of pieces of RNAs studied in that laboratory [7], suggesting that the replicable species had been generated by recombination from non-replicating RNA fragments. Thus, we can repeat after Pasteur that "there is no known circumstance in which it can be confirmed that microscopic beings come into the world without germs, without parents similar to themselves" (as cited in [42]).

8. Use of the molecular colony technique to monitoring RNA recombination

Detection of airborne RQ RNAs has been the first application of MCT. These experiments demonstrated diagnostic potential of MCT by revealing its ability to detect even solitary "infectious" molecules. The advantages of MCT as a diagnostic tool are discussed later. In this section, I will consider the utility of MCT to monitoring chemical reactions between single molecules; in particular, to studying RNA recombination. Earlier, RNA recombination was exclusively studied in vivo, using viruses and their defective genomes amplified in living cells. However, because of limitations of the in vivo systems, those studies could not answer even such basic questions as whether the recombination is performed by viral replicase, by host cell proteins, or by RNA molecules themselves. Moreover, the in vivo studies did not definitely rule out the involvement of DNA intermediates and, therefore, the possibility that recombination occurs at the DNA, rather than at the RNA, level.

Due to its ability to detect even single replicable RNA molecules, MCT allows RNA recombination experiments to be carried out in vitro as easily as in vivo, but without restraints inherent to the in vivo systems. To observe RNA recombination in vitro, the above mentioned non-replicable 5′ and 3′ fragments of RQ RNA are mixed and seeded on a Qβ replicase containing agarose layer, which is then covered with a nylon membrane impregnated with rNTPs [13]. The role of the nylon membrane is to retard diffusion of RNA molecules by reversibly binding them, thereby reducing the size of RNA colonies and hence increasing the resolving power of MCT

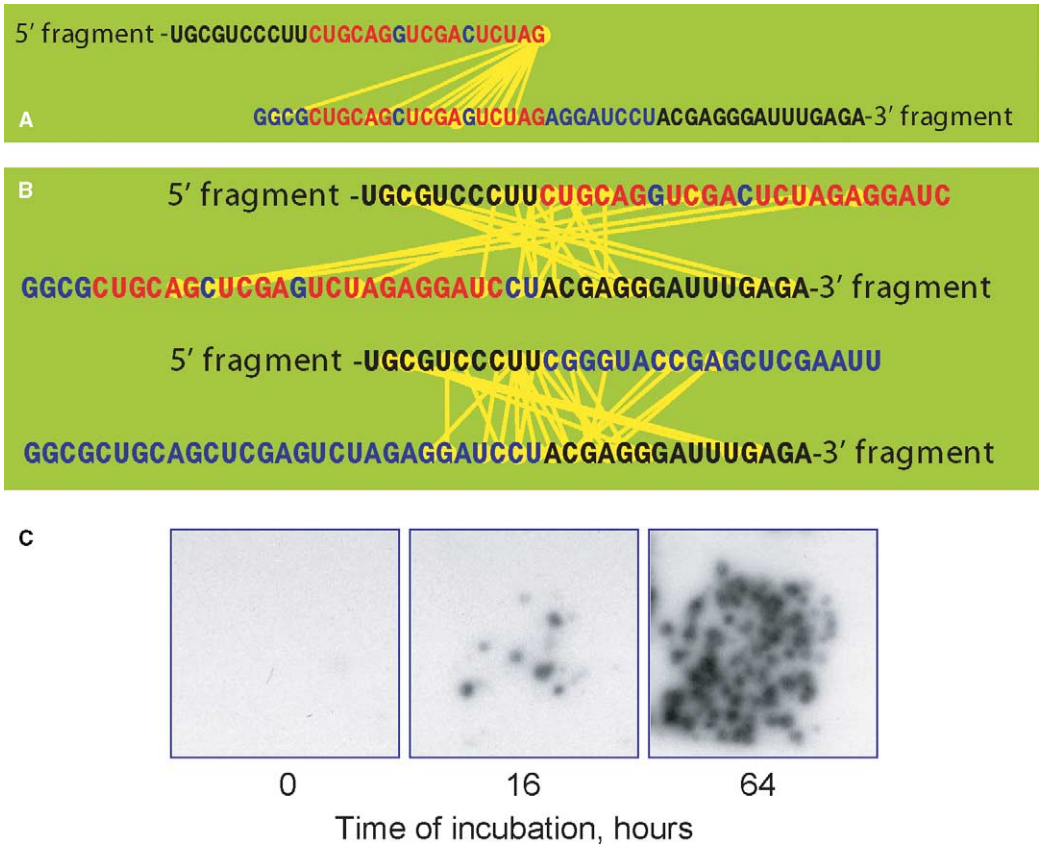

Fig. 2. Recombination between supplementing 5′ and 3′ fragments of an RQ RNA. A: Nucleotides that become joined (connected with yellow lines) in the recombinants generated in the presence of Qβ replicase (Samatov, T.R. and Chetverin, A.B., unpublished data). B: Same in the absence of any protein [14]. Black letters indicate sequences derived from RQ135 RNA [21], colored letters indicate artificial extensions of the fragments, red letters indicate homologous segments. C: Time course of the generation of replicable RNAs by self-recombination from the 5′ and 3′ fragments at 37 °C in the presence of 10 mM $MgCl_2$ and buffer (10 mM Tris–HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA) [14]. RNA colonies are detected by hybridization with ^{32}P-labeled probes.

[43]. This experimental scheme provides for positive selection of recombination molecules in which the two fragments are arranged in the manner as they are in the original RQ RNA. Such molecules are replicable and produce RNA colonies whose number reflects the recombination frequency.

In accordance with the in vivo data, recombination in this cell-free system appeared to be non-homologous and occurred at a frequency characteristic of non-homologous recombination in RNA viruses, $\approx 10^{-5}$/nt [34]. Inasmuch as the system contained only pure Qβ replicase and rNTPs, one could conclude that no cellular proteins or DNA intermediates were involved. Recombination was virtually prevented by periodate oxidation of the 5′ fragment eliminating its 3′-terminal hydroxyl group. This fact, as well as the recombinant sequences (Fig. 2A), was consistent with the hypothesis that recombination is brought about by a transesterification reaction in which the free 3′ hydroxyl of the 5′ fragment attacks phosphodiester bonds or the 5′-terminal triphosphate group of the 3′ fragment [13].

9. RNA molecules are intrinsically recombinogenic

From the above results it was not clear whether the reaction between RQ RNA fragments was performed by Qβ replicase or by RNA molecules themselves. To explore this alternative, the experimental scheme was modified to separate the recombination and replication events [14]. To this end, a mixture of the RNA fragments was incubated under chosen conditions and, prior to seeding on the Qβ replicase-containing agarose, the fragments were oxidized to suppress further recombination by the above mechanism. Hence, RNA colonies would only grow if recombination had occurred before the oxidation step. In the absence of Qβ replicase and rNTPs, recombination between the fragments turned out to be several orders of magnitude slower, suggesting that it is somehow catalyzed by the replicase. The residual recombination was not due to incomplete elimination of the 3′ hydroxyls, as it might be expected: repeated oxidation of the fragments either before or after incubation did not change its rate. Furthermore, its mechanism appeared to be entirely different from the mechanism of the replicase-catalyzed reaction: the fragments react by internal segments not involving the 3′ hydroxyls (Fig. 2B). Similar reactions can also occur *in cis*, resulting in deletion of internal RNA segments [14].

Cross-over sites are randomly distributed (Fig. 2B), indicating that cryptic ribozyme structures are not involved. It follows that RNA is intrinsically recombinogenic and, inasmuch as self-recombination requires nothing but RNA itself and Mg^{2+}, it must be ubiquitous in nature. Most probably, self-recombination occurs via a Mg^{2+}-catalyzed RNA cleavage generating fragments with 2′,3′ cyclic phosphate and 5′ hydroxyl termini, which are then cross-ligated. An alternative mechanism involving intermediate formation of a branched structure [44] was apparently excluded by the fact that yeast debranching ribonuclease, which selectively cleaves the 2′–5′ internucleotide bonds [45], did not reduce the reaction yield (Chetverina, H.V. and Chetverin, A.B., unpublished).

The rate of RNA self-recombination is low (Fig. 2C), $\approx 10^{-9}$ h^{-1} per internucleotide bond at 37 °C [14]. Yet, spontaneous rearrangements in RNA sequences might play an important role in the evolution of both RNA and DNA genomes. Even if not increased by the action of cellular proteins, this rate provides for the generation of a new recombinant RNA in a human cell every minute, yielding up to 10^{20} recombinant molecules during the life span of the human body [14]. Reverse transcription and integration of even a minute fraction of them would provide for a significant change of the human genome, and must be considered among factors affecting the genetic variability and the probability of spontaneous oncogenic transformation.

10. Implications for the RNA World

As noted above, there is no evidence that RNA creation/evolution can be so fast as to generate replicable RNAs from mononucleotides within hours, but it seems quite likely that replicable RNAs and their replicases could have arisen on a much longer timescale on Earth or on some other planet, giving rise to the RNA World [46].

The studies on the replicable RNAs and their self-recombination give some clues to what might be likely features of the RNA World.

(1) RNA colonies [10] might be a primitive, pre-cellular form of compartmentation. It has been understood that natural selection works with ensembles of molecules, rather than with individual molecules. This means that there must be some kind of compartmentation linking replicase to its products in order that natural selection be able to identify a gene that makes a better product [47,48]. Mixed RNA colonies comprised of more than one RNA species and growing in moist clays or other porous solids could perform the function of such a compartment [48,49]. This form of compartmentation is not confronted with the problem of transportation through hydrophobic barriers that inevitably arise if lipid membranes are employed [48].
(2) Dissemination of RNA molecules through atmosphere [10] might be a mechanism for the lateral transfer of genetic information between RNA colonies, an equivalent of sex, which is required for a high rate of evolution [50].
(3) Spontaneous recombinations and rearrangements of polyribonucleotides [14] might be a mechanism for creation of complex structures and for producing new combinations of genetic elements, generating the variability needed for natural selection [46]. These reactions might be antecedent to ribozyme functions, and could drive the creation of long ribozyme molecules from short oligonucleotides that had been generated by nucleotide condensation [49].

There are grounds to believe that studies on the Qβ system will continue to feed the ideas on how the RNA World might operate. For example, elucidation of the mechanism by which Qβ replicase prevents the template and its complementary copy from being hydrogen-bonded along the entire length [7] might help to approach the yet unsolved issue of how this function could be fulfilled by ribozyme replicases [47].

11. Applications of the molecular colony technique to in vitro cloning and diagnostics

Of course, the potentials of MCT are much greater than just the demonstration of airborne replicable RNAs and the studies on RNA recombination. The unique feature of MCT distinguishing it from other methods for nucleic acid amplification is

that amplified molecules are spatially separated. This results in weakening or (given that the sample is properly diluted) complete elimination of a competition between the molecular species; allows the individual amplifiable molecules to be monitored, counted and analyzed; provides for a direct high-throughput screening of a large number of molecules; and makes the isolation of homogeneous molecule populations (cloning) possible. Therefore, the most obvious applications of MCT would seem to be cell-free molecular cloning and molecular diagnostics. Unfortunately, the structural restraints Qβ replicase imposes upon its templates and the ability of RNAs to self-recombine (see above) limit the utility of the Qβ replicase version of MCT for these purposes. However, MCT can utilize any enzymatic reaction that provides for the exponential amplification of nucleic acids [51], e.g., PCR or isothermal amplification reactions, such as 3SR (self-sustained sequence replication [52]) or SDA (strand displacement amplification [53]). The most promising results were obtained using the PCR version of MCT (PCR-MCT) [51], also termed "polony" (polymerase colony) technology [54]. Since PCR involves repeated sample heating, thermostable media, such as a polyacrylamide gel, are used in place of the agarose gel.

One of the most promising applications of PCR-MCT is molecular diagnostics. MCT format digitalizes the assay; it makes single target molecules visible and their quantitation straightforward, by simply counting the number of molecular colonies. MCT format has been shown to be capable of eliminating any competition between different targets, even if their ratio varies more than million-fold. It also eliminates the interference from a non-specific synthesis occurring due to mishybridization of primers with non-target nucleic acids that are often present in clinical samples in a trillion-fold excess over the assayed target (Fig. 3). These features greatly increase the reliability and sensitivity of molecular diagnostics as compared to those achievable by the conventional solution assays [55]. The latest experiments closely mimicking real clinical diagnostics demonstrate that PCR-MCT detects, in 100-μL human blood aliquots, 100% molecules of DNA targets and 50% molecules of RNA targets. This corresponds to the sensitivity of 1 and 2 molecules, respectively (Chetverina, H.V., Falaleeva, M.V. and Chetverin, A.B., submitted), which is the highest diagnostic sensitivity ever achieved.

In addition, PCR-MCT has been used for establishing the physical linkage between distant genetic markers according to their ability to produce mixed colonies and for precisely quantifying the relative expression of different alleles of the same gene by counting the number of their respective colonies [56], single nucleotide polymorphism genotyping and expression profiling of heterogeneous cell populations [57], characterizing various cancer-associated genomic abnormalities [58], single molecule profiling of alternative pre-mRNA splicing [59], and massively parallel in situ sequencing of DNA fragments amplified in the form of molecular colonies [60].

Future developments of MCT will include cell-free cloning of entire genes and expressing (transcribing and translating) them directly within molecular colonies [51]. This will provide for

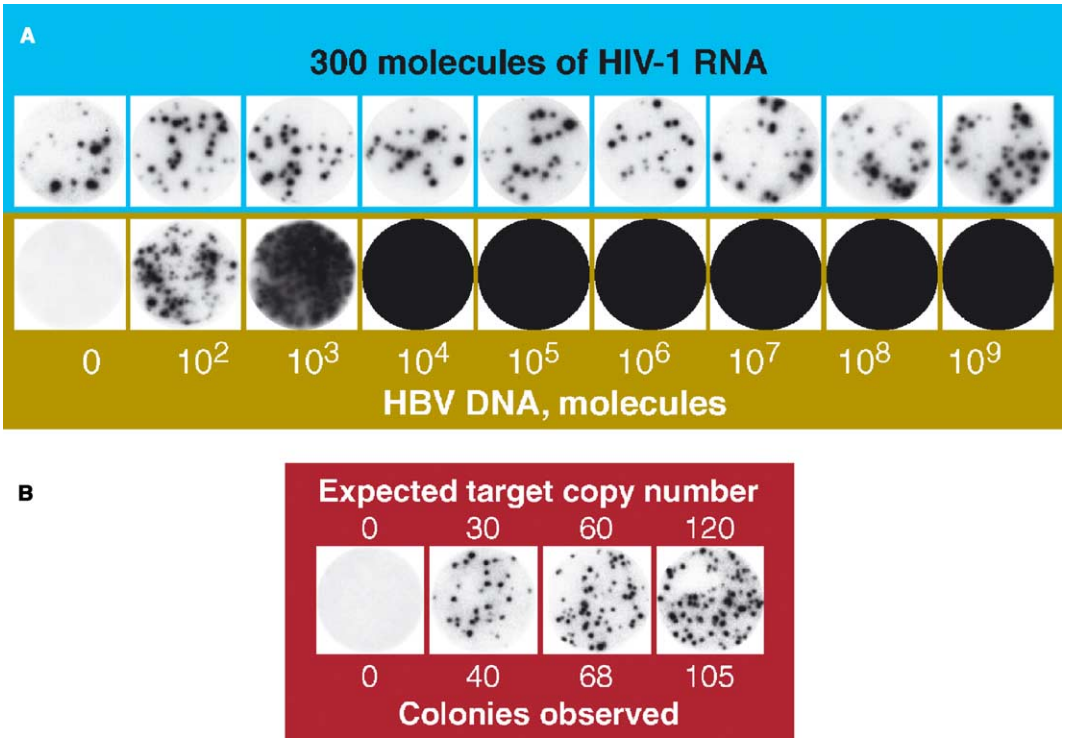

Fig. 3. Assay of RNA and DNA targets with the PCR version of the molecular colony technique. A: Lack of competition between targets in a multiplex assay. Colonies produced by 300 molecules of human immunodeficiency virus type-1 (HIV-1) RNA in the presence of the indicated number of concurrently amplifying molecules of human hepatitis B virus (HBV) DNA. Each of the nine gels was blotted with a nylon membrane that was first hybridized with an HIV-1-specific ^{32}P-labeled probe (top row) and then with an HBV-specific probe (bottom row). B: Detection of HBV DNA molecules in human blood. Total nucleic acids were isolated from 60-μl aliquots of the whole human blood to which diluted samples, expected to contain the indicated number of HBV DNA molecules, had been added. The blood aliquots contained nucleic acids equivalent by weight to 10^{13} molecules of the target. (Reprinted from [55] by permission of BioTechniques/Eaton Publishing.)

screening the colonies for the ability of the synthesized proteins to perform a particular enzymatic reaction or to bind a particular ligand (including antigens, antibodies, or nucleic acids). In combination with ribosome display strategies [61,62], this will provide for the in vitro selection of genes coding for proteins with desired functions, a cell-free alternative to the phage display. Compared to the traditional in vivo techniques, this is a true molecular cloning. In this case, there is no need in cloning vectors, in transformation of cells which is always very inefficient, or in isolation of the cloned nucleic acids – since each colony comprises a homogenous DNA preparation. Moreover, the colonies comprise naked RNA or DNA that can be directly analyzed in situ. In other words, MCT can do everything the traditional in vivo gene cloning can, and even more.

Acknowledgements: I thank Helena Chetverina for helpful discussions. This work was supported by the Program for Physico-Chemical Biology of the Russian Academy of Sciences, the Russian Foundation for Basic Research Grant No. 02-04-48320, INTAS (International Association for the promotion of co-operation with scientists from the New Independent States of the former Soviet Union) Grant No. 01-2012, and an International Research Scholar's award from the Howard Hughes Medical Institute to A.B.C.

References

[1] Hirst, G.K. (1962) Cold Spring Harbor Symp. Quant. Biol. 27, 303–309.
[2] Ledinko, N. (1963) Virology 180, 107–119.
[3] Haruna, I., Nozu, K., Ohtaka, Y. and Spiegelman, S. (1963) Proc. Natl. Acad. Sci. USA 50, 905–911.
[4] Haruna, I. and Spiegelman, S. (1965) Science 150, 884–886.
[5] Horiuchi, K. (1975) in: RNA Phages (Zinder, N.D., Ed.), pp. 29–50, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[6] Munishkin, A.V., Voronin, L.A. and Chetverin, A.B. (1988) Nature 333, 473–475.
[7] Chetverin, A.B. and Spirin, A.S. (1995) Prog. Nucleic Acid Res. Mol. Biol. 51, 225–270.
[8] Banerjee, A.K., Rensing, U. and August, J.T. (1969) J. Mol. Biol. 45, 181–193.
[9] Avota, E., Berzins, V., Grens, E., Vishnevsky, Y., Luce, R. and Biebricher, C.K. (1998) J. Mol. Biol. 276, 7–17.
[10] Chetverin, A.B., Chetverina, H.V. and Munishkin, A.V. (1991) J. Mol. Biol. 222, 3–9.
[11] Brown, D. and Gold, L. (1995) Biochemistry 34, 14775–14782.
[12] Zamora, H., Luce, R. and Biebricher, C.K. (1995) Biochemistry 34, 1261–1266.
[13] Chetverin, A.B., Chetverina, H.V., Demidenko, A.A. and Ugarov, V.I. (1997) Cell 88, 503–513.
[14] Chetverina, H.V., Demidenko, A.A., Ugarov, V.I. and Chetverin, A.B. (1999) FEBS Lett. 450, 89–94.
[15] Weissmann, C. (1974) FEBS Lett. 40, S10–S18.
[16] Weissmann, C., Billeter, M.A., Goodman, H.M., Hindley, J. and Weber, H. (1973) Annu. Rev. Biochem. 42, 303–328.
[17] Brown, D. and Gold, L. (1995) Biochemistry 34, 14765–14774.
[18] Biebricher, C.K. and Luce, R. (1993) Biochemistry 32, 4848–4854.
[19] Brown, D. and Gold, L. (1996) Proc. Natl. Acad. Sci. USA 93, 11558–11562.
[20] Ugarov, V.I., Demidenko, A.A. and Chetverin, A.B. (2003) J. Biol. Chem. 278, 44139–44146.
[21] Munishkin, A.V., Voronin, L.A., Ugarov, V.I., Bondareva, L.A., Chetverina, H.V. and Chetverin, A.B. (1991) J. Mol. Biol. 221, 463–472.
[22] Miele, E.A., Mills, D.R. and Kramer, F.R. (1983) J. Mol. Biol. 171, 281–295.
[23] Lewin, R. (1983) Science 222, 1313–1315.
[24] Axelrod, V.D., Brown, E., Priano, C. and Mills, D.R. (1991) Virology 184, 595–608.
[25] Arora, R., Priano, C., Jacobson, A.B. and Mills, D.R. (1996) J. Mol. Biol. 258, 433–446.
[26] Mills, D.R. (1988) J. Mol. Biol. 200, 489–500.
[27] Morozov, I.Yu., Ugarov, V.I., Chetverin, A.B. and Spirin, A.S. (1993) Proc. Natl. Acad. Sci. USA 90, 9325–9329.
[28] Tyagi, S., Landegren, U., Tazi, M., Lizardi, P.M. and Kramer, F.R. (1996) Proc. Natl. Acad. Sci. USA 93, 5395–5400.
[29] Lai, M.M.C. (1992) Microbiol. Rev. 56, 61–79.
[30] Agol, V.I. (1997) Semin. Virol. 8, 77–84.
[31] Bujarski, J.J. and Kaesberg, P. (1986) Nature 321, 528–531.
[32] Chetverin, A.B. (1997) Semin. Virol. 8, 121–129.
[33] Palasingam, K. and Shaklee, P.N. (1992) J. Virol. 66, 2435–2442.
[34] Chetverin, A.B. (1999) FEBS Lett. 460, 1–5.
[35] Sumper, M. and Luce, R. (1975) Proc. Natl. Acad. Sci. USA 72, 162–166.
[36] Hill, D. and Blumenthal, T. (1983) Nature 301, 350–352.
[37] Biebricher, C.K., Eigen, M. and Luce, R. (1986) Nature 321, 89–91.
[38] Biebricher, C.K., Eigen, M. and Luce, R. (1981) J. Mol. Biol. 148, 369–390.
[39] Pasteur, L. (1860) C. R. Acad. Sci. 50, 303–307.
[40] Koch, R. (1881) Mittheilungen aus dem Kaiserlichen Gesundheitsamte 1, 1–48.
[41] McCaskill, J.S. and Bauer, G.J. (1993) Proc. Natl. Acad. Sci. USA 90, 4191–4195.
[42] Dubos, R.J. (1950) Louis Pasteur: Free Lance of Science. DaCapo Press, New York. p. 187.
[43] Chetverina, H.V. and Chetverin, A.B. (1993) Nucleic Acids Res. 21, 2349–2353.
[44] Chetverin, A.B. (1999) Mol. Biol. (Moscow) 33, 985–996.
[45] Nam, K., Hudson, R.H.E., Chapman, K.B., Ganeshan, K., Damha, M.J. and Boeke, J.D. (1994) J. Biol. Chem. 269, 20613–20621.
[46] Gilbert, W. (1986) Nature 319, 618.
[47] Gilbert, W. and de Souza, S.J. (1999) in: The RNA World, 2nd edn. (Gesteland, R.F., Cech, T.R. and Atkins, J.F., Eds.), pp. 221–231, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[48] Szostak, J.W. (1999) in: Size Limits of Very Small Microorganisms: Proceedings of a Workshop, National Academy Press, Washington, DC, pp. 120–125.
[49] Spirin, A.S. (2002) FEBS Lett. 530, 4–8.
[50] Woese, C. (1998) Proc. Natl. Acad. Sci. USA 95, 6854–6859.
[51] Chetverin, A.B., Chetverina, H.V. (1997), US Patent 5,616,478.
[52] Guatelli, J.C., Whitfield, K.M., Kwoh, D.Y., Barringer, K.J., Richman, D.D. and Gingeras, T.R. (1990) Proc. Natl. Acad. Sci. USA 87, 1874–1878.
[53] Walker, G.T., Little, M.C., Nadeau, J.G. and Shank, D.D. (1992) Proc. Natl. Acad. Sci. USA 89, 392–396.
[54] Mitra, R.D. and Church, G.M. (1999) Nucleic Acids Res. 27, e34.
[55] Chetverina, H.V., Samatov, T.R., Ugarov, V.I. and Chetverin, A.B. (2002) BioTechniques 33, 150–156.
[56] Mitra, R.D., Butty, V., Shendure, J., Williams, B.R., Housman, D.E. and Church, G.M. (2003) Proc. Natl. Acad. Sci. USA 100, 5926–5931.
[57] Merritt, J., DiTonno, J.R., Mitra, R.D., Church, G.M. and Edwards, J.S. (2003) Nucleic Acids Res. 31, e84.
[58] Butz, J., Wickstrom, E. and Edwards, J. (2003) BMC Biotechnol. 3, 11.
[59] Zhu, J., Shendure, J., Mitra, R.D. and Church, G.M. (2003) Science 301, 836–838.
[60] Mitra, R.D., Shendure, J., Olejnik, J., Olejnik, E.K. and Church, G.M. (2003) Anal. Biochem. 320, 55–65.
[61] Hanes, J. and Pluckthun, A. (1997) Proc. Natl. Acad. Sci. USA 94, 4937–4942.
[62] Roberts, R.W. and Szostak, J.W. (1997) Proc. Natl. Acad. Sci. USA 94, 12297–12302.