Expressible molecular colonies

Timur R. Samatov, Helena V. Chetverina and Alexander B. Chetverin*

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

Received July 20, 2005; Revised and Accepted September 19, 2005

ABSTRACT

Carrying out polymerase chain reaction in a gel layer generates a 2-D pattern of DNA colonies comprising pure genetic clones. Here we demonstrate that transcription, translation and protein folding can be performed in the same gel. The resulting nucleo-protein colonies mimic living cells by serving as compartments in which the synthesized RNAs and proteins co-localize with their templates. Yet, due to the absence of penetration barriers, such a molecular colony display allows cloned genes to be directly tested for the encoded functions. Now, the results imply that virtually any manipulations with genes and their expression products can be accomplished in vitro.

INTRODUCTION

Recent developments in genomics, proteomics and molecular engineering highlighted the importance and advantages of in vitro approaches, including PCR (1), microarray technology (2,3), cell-free protein synthesis (4–7) and in vitro display techniques (8–11) for systematic studies on gene expression and regulation, protein structure and function and selection of proteins and peptides that possess desired properties. In vitro methods provide for greater variation and tighter control of experimental conditions by a scientist than their in vivo counterparts; they are faster and more amenable to automation; are free from constraints imposed by living cells or cloning vectors; are less susceptible to natural selection; and allow larger DNA, RNA or protein libraries to be handled. Moreover, they permit modified nucleotides or amino acids, and even their unnatural analogs to be incorporated into nucleic acids and proteins, in order to specifically label them or to further expand their structural repertoire (12,13). By now, in vitro format has been implemented for almost every process involved, with the only, but important exception: isolation of individual molecular clones from DNA or RNA pools obtained by in vitro manipulations, as well as expressing and screening the clones in situ, is still performed using living cells, thus restricting the power of in vitro methods.

Here, we describe an approach that eliminates this restriction. It is based on the molecular colony technique (MCT) capable of generating a 2-D pattern of colonies of nucleic acids by amplifying them in a gel layer, each colony comprising many copies (a clone) of one starting RNA or DNA molecule (14,15). Earlier, MCT was employed for studies on chemical reactions between single RNA molecules (16), single nucleotide polymorphism genotyping and gene expression analysis (17,18), massively parallel sequencing of DNA fragments (19), studies on alternative pre-mRNA splicing (20), and extremely sensitive and reliable diagnostics (21,22). In this paper, we demonstrate that DNA clones can be transcribed and translated within their home colonies and screened according to properties of the expression products.

MATERIALS AND METHODS

Growing DNA colonies

PCR was performed essentially as described previously (21) in 0.4 mm-thick, 14 mm-diameter polyacrylamide gels of specified concentrations, containing 50 mM Tris–HCl (pH 8.6 at 25°C), 1 μg/μl BSA (fraction V, Amersham Biosciences), 2.5 mM MgCl₂, 0.2 mM each of dNTP, 3.6 ng/μl of Taq DNA A polymerase (from Thermus aquaticus), 0.02 ng/μl Pwo DNA polymerase (from Pyrococcus woesei) modified by a His₆ tag at the N-terminus (23), as well as a template and appropriate primers described in Supplementary Data. The DNA polymers used in this work were isolated as described (21). The gel was subjected to 40 cycles of PCR [melting at 94°C for 20 s, annealing at 55°C for 20 s and extension at 72°C for 90 s (obelin cDNA) or 150 s (luciferase and green fluorescent protein (GFP) cDNAs)], followed by incubation at 72°C for 5 min. DNA colonies were detected by blotting the gel with a Hybond™ N⁺ membrane (Amersham Biosciences), hybridizing it with a ³²P-labeled transcript synthesized from a corresponding plasmid with T7 RNA polymerase (16), and either autoradiographing or scanning the membrane with the Cyclone™ storage phosphor system (Packard Instrument).

Synthesis of luciferase cDNA

Total RNA from dried lanterns of firefly Luciola mingrelia (24), generously provided by Dr N.N. Ugarova (Moscow State Institute of Protein Research, Russian Academy of Sciences, Pushchino, Moscow Region 142290, Russia

*To whom correspondence should be addressed. Tel/Fax: +7 095 632 7871; Email: alexch@vega.protres.ru

© The Author 2005. Published by Oxford University Press. All rights reserved.

The online version of this article has been published under an open access model. Users are entitled to use, reproduce, disseminate, or display the open access version of this article for non-commercial purposes provided that: the original authorship is properly and fully attributed; the Journal and Oxford University Press are attributed as the original place of publication with the correct citation details given; if an article is subsequently reproduced or disseminated not in its entirety but only in part or as a derivative work this must be clearly indicated. For commercial re-use, please contact journals.permissions@oxfordjournals.org
University), was isolated according to a standard procedure (25). Because the resulting RNA preparation could not be completely dissolved in 0.5% SDS and 0.1 mM EDTA, it was additionally extracted with phenol (3 times), twice washed with chloroform and precipitated with ethanol. The resulting pellet was washed with 96% ethanol, dissolved in 0.1 mM EDTA, and used for reverse transcription with the SuperScript RNase H (Invitrogen) according to the manufacturer’s protocol, in 5 μl reactions containing 0.5 μg of the total firefly RNA and 0.25 μg oligo(dT)12-18.

**Transcription in molecular colonies**

After drying a PCR gel in vacuo, it was overlaid with 65 μl of a reaction mixture containing 100 mM Tris–HCl (pH 8.0 at 25°C), 20 mM MgCl2, 1 mM spermidine, 0.2 mM EDTA, 40 mM DTT, 4 mM each of rNTP and 50 ng/μl T7 RNA polymerase, and incubated at 4°C for 1 h, to allow all liquid to be entrapped. Transcription was carried out by incubating the gel at 37°C during 2 h, and monitored by blotting the gel with a Hybond™ N+ membrane and hybridizing the membrane with an appropriate 32P-labeled antisense mRNA sequence.

**Combined transcription-translation in molecular colonies**

Before drying a PCR gel in vacuo, it was twice extracted during 10 min with at least 20 volumes of a saline alcohol [a 45:55 (v/v) mixture of ethanol with 200 mM Na-citrate, 300 mM NaCl and 0.4 mM EDTA] (21) and then 3 times with a 50% ethanol. The dried gel was overlaid with 65 μl of a reaction mixture containing 30% (v/v) wheat germ lysate (from a Wheat Germ CECF Kit, Roche Diagnostics), 25 mM HEPES (pH 7.6), 2.5 mM MgCl2, 20 mM K-acetate, 1.6 mM DTT, 2 mM ATP, 1 mM GTP, 1 mM CTP, 1 mM UTP, 0.25 mM spermidine, 8 mM creatine phosphate, 60 μg/ml of creatine phosphokinase, 0.1 mM each of amino acid, 50 μg/ml yeast tRNA and 35 ng/μl T7 RNA polymerase. The reaction was carried out by incubating the gel at 25°C. Fluorescence of the synthesized GFP was monitored by scanning the gel at an indicated time intervals using a ScanArray™ Express microarray scanner (Perkin-Elmer) equipped with a 488 nm blue laser and a 508 nm emission filter.

**RESULTS**

For the purpose of the present study it is essential that molecular colonies mimic living cells by the ability to compartmentalize biochemical reactions (26). Hence, if one succeeded in carrying out in molecular colonies all reactions constituting the amplification and expression of entire genes, this would provide for both obtaining gene clones and screening them by the properties of co-localized expression products.

To achieve this goal, we used the PCR version of MCT (15,21), also termed ‘polony’ technology (27), whose template specificity is solely determined by pre-selected oligonucleotide primers matching the boundaries of DNA or RNA sequence(s) to be amplified. Since PCR involves repeated sample heating, thermostable media, such as polyacrylamide gel, must be used. Molecular colonies form because the gel matrix retards the motion of reaction products. This, in turn, may slow down gene amplification and expression reactions by obstructing the mobility of participating reagents and catalysts, especially of such giant biomolecules as ribosomes. Accordingly, the major concern regarding the feasibility of this approach was whether the yield of each of the contributing reactions (PCR, transcription and translation) would be high enough to enable the expression products in individual colonies to be tested. Therefore, we sought a polycyramylidase gel with the highest porosity, at which DNA colonies remain acceptably compact (see Supplementary Figure 1). In such a gel, almost every DNA molecule of up to 1.6 kb in length produced a colony of up to 10^8 of its copies (Figure 1). *Pwo* DNA polymerase (23) was included into the gel together with *Taq* DNA polymerase to improve the yield and fidelity of amplification (28).

The potential of MCT was explored by cloning a 1700 nt-long luciferase mRNA sequence from the total RNA prepared from dried lanterns of firefly *L.mingrelica* (24). Oligo(dT) served as a primer for reverse transcription of all the poly(A)-containing mRNAs, and the resulting cDNA preparation was used, in combination with sequence-specific oligonucleotide primers, for in-gel amplification of luciferase cDNA. A number of colonies hybridizable with a specific probe were produced by as low as 10 pg of total RNA (Figure 2a) which approximates the RNA content of an animal cell (29). This suggests that for MCT cloning, one can use the genetic material of a single cell without preliminary amplification, which is not achievable with in vivo cloning techniques because of a low sequence recovery, between 0.01 and 0.0001% of the input population (30). Figure 2b shows that material picked from colonies can be further amplified by the solution PCR to produce full-size luciferase cDNA capable of hybridization with a sequence-specific probe.

For practical use of MCT cloning it is important that cloned genes can be expressed and tested according to the encoded functions. However, proteins and nucleoprotein complexes responsible for transcription and translation would irreversibly denature at the high temperatures employed in PCR. Hence, gene amplification and expression steps must be separated, which could be done by a variety of means. For example, amplified genes could be expressed in another gel, to which
the contents of DNA colonies are partially transferred either through a direct contact with the amplification gel, or by using a blotting membrane. However, because of a low efficiency of transfer (see Supplementary Figure 2), the highest yield of expression can be achieved when genes are expressed in situ, in the same gel in which PCR had been carried out. We approached this goal as follows: after completion of PCR, the gel was dried and then reconstituted by soaking in a solution containing the components of a cell-free expression system.

This approach allowed us to perform in situ transcription, the first step of gene expression, with the colony pattern being perfectly preserved (Figure 3a). Judging by the increase of hybridization signal, at least 10 RNA copies of 1.7 kb in length were synthesized by phage T7 RNA polymerase per each DNA template.

However, using the same approach, we were unable to detect any protein synthesis in molecular colonies. In order to investigate reasons for this failure, we carried out a series of experiments on cell-free translation and combined transcription-translation, both in solution and in polyacrylamide gel. In the latter case, PCR colonies were mimicked by spotting the dried gel with miniature aliquots of a serially diluted mRNA or plasmid from which the fragment to be-expressed was excised with restriction exonucleases (Figure 3b). The results (see Supplementary Figure 3) show that translation occurs in polyacrylamide gel almost as efficiently as in solution. At the same time, every constituent of the PCR cocktail inhibits translation to a certain extent, with the buffer component being the most powerful inhibitor because of mutually exclusive pH requirements of PCR and translation. We found the following effective and simple solution of this problem: before drying the PCR gel, it is extracted with a saline alcohol and then desalted by washing with ethanol. This treatment eliminates the inhibitory action of PCR reagents (Figure 3b), in agreement with our earlier observations that saline alcohol is capable of extracting a variety of low molecular weight substances and proteins from nucleic acid pellets (21, 22). Also, this fixes DNA molecules within their home colonies.

We tested performance of the modified procedure by expressing the colonies generated by in-gel amplification of the cloned DNA was compared with ethidium bromide-stained DNA fragments from a BstEII digest of phage λ DNA (lane ‘M’) and the product of amplification of luciferase cDNA by solution PCR (lane ‘L’).
the number of template molecules introduced into the gel. Furthermore, gel images obtained by monitoring the fluorescence (Figure 3c) and by sequence-specific probe hybridization (Figure 3d) are nearly identical, indicating that fluorescence is emitted by the colonies of GFP-encoding DNA fragments and, hence, that in situ expression of those fragments resulted in the synthesis and folding of a functionally active protein.

Time course of the fluorescence emission by colonies follows the kinetics of maturation of this GFP variant (31) corrected for the time needed for GFP synthesis in solution (see Supplementary Figure 4), suggesting that gel matrix does not significantly affect the rates of transcription and translation. The time-dependent dissipation of the fluorescence pattern (Figure 3c and Supplementary Figure 4) is due to diffusion of the protein (GFP), rather than DNA moiety of colonies; the latter remains essentially compact (Figure 3d). Furthermore, protein diffusion is significant only after the first hour of the reaction, when translation levels are off (Supplementary Figure 4) and GFP fluorescence becomes detectable (Figure 3c) despite a delay caused by the slow reaction of protein fluorophore formation.

By comparing the fluorescence intensities of colonies with those of known amounts of pre-synthesized mature GFP (see Supplementary Figure 4), we inferred that one colony contained on average 10^4 GFP molecules. This corresponds to a surface density of ~40 pg of protein per square millimeter, which is sufficient for the detection of synthesized proteins by routine assays (as low as 1 pg or even smaller amount of protein can be detected with commercially available kits, as reported by vendors; see, e.g. http://www.bio-rad.com, http://www.amershambiosciences.com, http://www.piercenet.com), and implies that in this system, ~10 protein molecules are synthesized per each DNA template.

**DISCUSSION**

In this study, we demonstrated that genes can be cloned, expressed and screened entirely in vitro and in situ, without any involvement of living cells. This has been made possible by carrying out gene amplification, transcription and translation reactions in a gel. In this format, copies of each amplifiable DNA molecule concentrate, together with their expression products, around the progenitor template in the form of a spherical colony (14,15). In certain sense, each molecular colony may be considered as a non-enveloped cell whose ‘genome’ is comprised of multiple copies of a single gene.

It should be noted that several methodologies based on the exhaustive dilution principle, such as digital PCR (32), sorting on oligonucleotide arrays (33), on microbeads (34), or on microbeads contained in water-in-oil emulsion compartments (35,36), could provide for obtaining individual clones, e.g. when the sorting compartments are rarely populated or when the DNA-to-beads ratio is low. However, in contrast to molecular colonies which, like bacterial colonies, inherently represent clones, the use of such technologies does not automatically lead to obtaining molecular clones, and clonal purity of the resulting preparations needs to be verified by direct methods. Therefore, in vivo cloning, such as in bacterial cells, is used at final steps to isolate and analyze individual clones from the obtained samples (37), and the use of terms ‘sorting’ and ‘enrichment’, rather than ‘cloning’, is more justified in these cases.

The technology reported here can be immediately used as a research tool in a number of areas. As far as each gene is physically linked to its expression product by being located in the same molecular colony, such a molecular colony display could aid to identification of genes performing certain functions, and could complement in vitro display methods (38) for rapid high-throughput screening of protein or peptide libraries, alone or at final steps of in vitro selection procedures, such as DNA (11), mRNA (10) or ribosome (8,9) display. The generated 2-D pattern of expressible molecular colonies can function as a sort of self-assembling DNA, RNA or protein array. Transcription in molecular colonies could be helpful for rapid selection of ribozymes and RNA aptamers (30). Further developments of this technology will likely include replacement of PCR with isothermal methods, such as 3SR (39), NASBA (40), strand-displacement (41) or rolling-circle amplification (42) and use of cell-free translation systems entirely composed of purified components (6).

In a number of respects, the reported technology is advantageous over the in vivo cloning and display methods. It is the first technology that provides for true molecular cloning, rather than cloning of cells or viruses harboring the gene of interest. Therefore, there is no need in cloning vectors, in the transformation of cells (which is always inefficient), or in isolating the cloned genes from cellular DNA. This allows up to 100% members of a genetic library to be cloned, expressed and tested compared from 0.0001 to 0.01% characteristic for methods relying on vectors and cell transformation (30). Unlike in vivo display methods, linking of a protein or peptide to its gene is achieved without fusing it to a tag sequence or to another protein; therefore, its native fold and properties are not disturbed. Many genes and their expression products can be simultaneously tested in molecular colonies directly, because there are no cell walls, membranes or emulsion compartments. Moreover, the expressed clones can be interrogated under conditions different from the transcription/translation and analytes can be soaked into the gels to detect activity of the macromolecules synthesized in colonies. Finally, genes can be amplified and expressed in the absence of natural selection and in the presence of unnatural nucleotides and amino acids.

In conclusion, we would like to note that demonstration of the capability of molecular colonies to synthesize proteins has an interesting implication for the pre-cellular RNA world namely, that RNA colonies possibly growing in moist clays or other porous substrates (26,43) might have created a translation apparatus and became selectable by functions of the encoded proteins even before they acquired a membrane envelope.

**SUPPLEMENTARY DATA**

Supplementary Data are available at NAR Online.

**ACKNOWLEDGEMENTS**

We thank Prof. A.S.Spirin for critical reading the manuscript; Drs E.S.Vysotski, V.N.Ksenzenko, N.N.Ugarova, J.Kur and S.Dabrowski for providing the genetic materials used.
in this work; Drs A.A.Minin, L.A.Shaloiko, V.A.Shirovok, V.I.Ugarov and V.A.Yashin for advice and help in carrying out some preliminary experiments; and A.G.Androsova, N.I.Androsova, L.V.Shutova, E.A.Uzlova and Z.V.Valina for technical assistance. This work was supported in part by program ‘Molecular and Cell Biology’ of the Russian Academy of Sciences and an International Research Scholar’s Award from the Howard Hughes Medical Institute to A.B.C. Funding to pay the Open Access publication charges for this article was provided by the Howard Hughes Medical Institute.

Conflict of interest statement
Dr Alexander B. Chetverin are authors of patents (US patents 5,616,478; 5,958,698; 6,001,568; Russian patents 2,048,522; 2,114,175; 2,114,915) that disclose the molecular colony technique and may be licensed for use.

REFERENCES

1. Saiki,R.K., Gelfand,D.H., Stoffel,S., Scharf,S.J., Higuchi,R., Horn,G.T., Mullis,K.B. and Erlich,H.A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science, 239, 487–491.
2. Schena,M., Shalon,D., Davis,R.W. and Brown,P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science, 270, 467–470.
3. MacBeath,G. and Scherer,S.L. (2000) Printing proteins as microarrays for high-throughput function determination. Science, 289, 1760–1763.
4. Spirin,A.S., Baranov,V.I., Ryabova,L.A., Ovodov,S.Y. and Alakhov,Y.B. (1988) A continuous cell-free translation system capable of producing polypeptides in high yield. Science, 242, 1162–1164.
5. Kigawa,T., Yabuki,T., Yoshida,Y., Tsutsui,M., Ito,Y., Shibata,T. and Shibatou,S. (1999) Cell-free production and stable-isotope labeling of milligram quantities of proteins. FEBS Lett., 442, 19–15.
6. Shimizu,Y., Inoue,A., Tomari,Y., Suzuki,T., Yokogawa,T., Nishikawa,K. and Ueda,T. (2001) Cell-free translation reconstituted with purified components. Nat. Biotechnol., 19, 751–755.
7. Sawasaki,T., Ogasawara,T., Morishita,R. and Endo,Y. (2002) A cell-free protein synthesis system for high-throughput proteomics. Proc. Natl Acad. Sci. USA, 99, 14652–14657.
8. Mathews,I.C., Bhattach,K.R. and Dower,W.J. (1994) An in vitro polysome display system for identifying ligands from very large peptide libraries. Proc. Natl Acad. Sci. USA, 91, 9022–9026.
9. Hanes,J. and Pluckthun,A. (1997) In vitro selection and evolution of functional proteins by using ribosome display. Proc. Natl Acad. Sci. USA, 94, 4937–4942.
10. Roberts,R.W. and Szostak,J.W. (1997) RNA-protein fusions for the in vitro selection of peptides and proteins. Proc. Natl Acad. Sci. USA, 94, 12297–12302.
11. Tawfik,D.S. and Griffiths,A.D. (1998) Man-made cell-like compartments for molecular evolution. Nat. Biotechnol., 16, 652–656.
12. Hiroa,I., Ohtsuki,T., Fujitwara,T., Mitsu,J., Yokogawa,T., Okum,I., Nakayama,H., Takio,K., Yabuki,T., Kigaw,T. et al. (2002) An unnatural base pair for incorporating amino acid analogs into proteins. Nat. Biotechnol., 20, 177–182.
13. Frankel,A., Li,S., Stark,S.R. and Roberts,R.W. (2003) Unnatural RNA display libraries. Curr. Opin. Struct. Biol., 13, 506–512.
14. Chetverin,A.B., Chetverina,H.V. and Munishkin,A.V. (1991) On the nature of spontaneous RNA synthesis by Q beta replicase. J. Mol. Biol., 222, 2–9.
15. Chetverin,A.B. and Chetverina,H.V. (1997) U.S. Patent 5,616,478.
16. Chetverin,A.B., Chetverina,H.V., Demidenko,A.A. and Ugarov,V.I. (1997) Nonhomologous RNA recombination in a cell-free system: evidence for a transsterification mechanism guided by secondary structure. Cell, 88, 503–513.
17. Mitra,R.D., Butty,V.L., Shendure,J., Williams,B.R., Housman,D.E. and Church,G.M. (2003) Digital genotyping and haplotyping with polymerase colonies. Proc. Natl Acad. Sci. USA, 100, 5926–5931.
18. Butz,J.A., Yan,H., Mikkilineni,V. and Edwards,J.S. (2004) Detection of allelic variations of human gene expression by polymerase colonies. BMC Genet., 5, 3.