**Title:**
In vitro transcription/translation-coupled DNA replication through the regeneration of 20 aminoacyl-tRNA synthetases

**Authors:**
Katsumi Hagino[1], Norikazu Ichihashi[1,2,3*]

**Affiliations:**
[1]Department of Life Science, Graduate School of Arts and Science, The University of Tokyo, Meguro, Tokyo 153-8902, Japan
[2]Komaba Institute for Science, The University of Tokyo, Meguro, Tokyo 153-8902, Japan
[3]Universal Biology Institute, The University of Tokyo, Meguro, Tokyo 153-8902, Japan

*Corresponding authors: ichihashi@bio.c.u-tokyo.ac.jp

**Abstract:**
The in vitro reconstruction of life-like self-reproducing systems is a major challenge in *in vitro* synthetic biology. Self-reproduction requires regeneration of all molecules involved in DNA replication, transcription, and translation. This study demonstrated the DNA replication and regeneration of major translation factors, 20 aminoacyl-tRNA synthetases (aaRS), in a reconstituted transcription/translation system (PURE system). First, we replicated each DNA that encode one of the 20 aaRSs through aaRS expression from the DNA (i.e., regeneration) by serial transfer experiments. Thereafter, we successively increased the number of aaRS genes and achieved simultaneous DNA replication and regeneration of all 20 aaRSs, which comprised approximately half the number of protein factors in the PURE system, except for ribosomes, by employing dialyzed reaction and sequence optimization. This study provides a step-by-step methodology for increasing the number of self-regenerative genes in self-reproducing artificial systems.

**Keywords:**
Self-reproduction; regeneration; artificial cell; aminoacyl-tRNA synthetase; in vitro synthetic biology.

## Introduction

Self-reproduction is a unique ability in living organisms. To date, various abilities of living organisms have been reconstituted in vitro to understand the design principles of living systems and the development of new biotechnologies[1–14]. However, the self-reproduction ability has not been implemented in an artificial system and remains a large challenge[6,15–19]. In living organisms, self-reproduction is achieved by synthesizing all macromolecules, such as DNA, RNA, and proteins, from low-molecular-weight compounds through DNA replication, transcription, and translation.

To construct an autonomous self-reproducing artificial system, the reconstituted translation system of Escherichia coli (PURE system) is a reasonable starting point because it is composed of the minimum elements for transcription and translation[20]. PURE systems consist of ribosomes, tRNAs, 36 translation proteins (TPs), including 20 aminoacyl-tRNA synthetases (aaRSs), and T7 RNA polymerase, as macromolecules. If all these RNA and proteins are regenerated from DNA that self-replicates in a PURE system, the whole system can self-reproduce.

Several studies have reported the synthesis of TPs in PURE systems. Awai et al. expressed 20 aaRSs in a PURE system and detected activities other than PheRS[21]. Wei and Endy expressed 36 TPs in a PURE system and detected the activities of 19 of 23 testable TFs after purification. Li et al. coexpressed most of the ribosomal proteins in one pot and detected the reconstitution of the 30S subunit[22]. Libicher et al. and Doerr et al. performed co-expression of multiple TPs in a PURE system from three large plasmids that encode most TPs[23]; however, the expression amounts were not sufficient to regenerate the original amount of TPs, and the activities of the expressed TFs were not verified[24,25]. Doerr et al. also reported that a substantial fraction of the co-expressed proteins is truncated due to inefficient ribosome processibity[25]. These studies did not directly demonstrate regeneration because the expressed TFs were not utilized for further translation. As studies that demonstrate the regeneration of TFs during transcription/translation reaction, Libicher performed regeneration of T7 RNA polymerase and adenylate kinase, and 12 aaRSs and RF1 up to second generations by a serial dilution experiment with the target proteins-depleted PURE system[26]. Lavickova et al. demonstrated the self-regeneration of T7 RNA polymerase or up to seven aaRSs using microfluidic reactors, which allows continuous expression of GFP for more than 24 h[27]. However, in these studies, the number of regenerative proteins was still limited, and DNA was continuously supplied by researchers (i.e., not replicated). The next important challenge is to increase the number of regenerative TFs and coupling with DNA replication.

This study constructed an in vitro system in which artificial genomic DNAs that encodes all 20 aaRSs replicates continuously for at least several generations in all aaRSs-depleted PURE systems by regenerating aaRSs from the DNA. First, we verified that each aaRS was expressed in the functional form in a PURE system, and the activity was sufficient for the replication of the DNA encoding itself. Thereafter, we performed continuous DNA replication through the expression of each aaRS using the serial dilution method in each aaRS-depleted PURE system. We further increased the regenerative aaRS numbers by adding circular DNAs that encode each aaRS and removing the corresponding aaRSs from the PURE system. We succeeded in DNA replication-coupled regeneration of up to five aaRSs under conventional conditions and all 20 aaRSs by employing dialyzed conditions after sequence and copy number optimization.

## Results

**The activity of aaRSs synthesized in a PURE system**

First, we examined the activity of 20 aaRSs expressed in a PURE system. The assay scheme is illustrated in Fig. 1a. First, DNA encoding aaRS (aaRS-DNA) was incubated in a PURE system to express aaRS. Thereafter, an aliquot of each reaction solution was diluted at a certain rate, as shown in Table S2, and added to the second PURE system, which lacks the aaRS of interest and contains a circular DNA that encodes phi29 DNA polymerase. If the synthesized aaRS in the first reaction is sufficiently active, it rescues the lack of aaRS in the second PURE system to express phi29 DNA polymerase, which catalyzes rolling-circle DNA replication using circular DNA as a template. DNA replication with aaRS DNA should be higher than that without aaRS-DNA. We observed significantly higher DNA replication with each aaRS-DNA than without aaRS-DNA for all 20 aaRSs (Fig. 1b). These results indicate that 20 aaRSs synthesized in the PURE system were sufficiently active to induce DNA replication.

In this experiment, the first reaction contained all aaRSs; thus, the first reaction was sufficiently diluted to decrease DNA replication without aaRS-DNA in the second reaction. The dilution rates were expected to reflect the requirements of each aaRS for DNA replication. For example, HisRS needs to be diluted by as much as $10^5$-fold because only a small amount of activity is sufficient to support DNA replication, whereas AlaRS needs to be diluted 10-fold because much activity is required. This information (Table S2) was used to classify 20 aaRSs based on their requirements in a subsequent experiment.

**DNA replication coupled with one of the 20 aaRS regeneration**

Furthermore, we tested whether DNA encoding one of the 20 aaRSs continuously replicated through the expression of aaRS from the DNA. The experimental scheme is illustrated in Fig. 2a. In the first reaction (round 1), a circular plasmid DNA encoding each of the 20 aaRSs (pET-aaRS) and the linear DNA encoding phi29 DNA polymerase (DNAP-DNA) were incubated in the customized PURE system, the composition of which was optimized for DNA replication (Table S1), containing all 20 aaRSs. During the incubation, phi29 DNA polymerase (DNAP) and each aaRS were expressed, and the DNAP catalyze rolling-circle replication of the circular DNA produced long DNA that encoded repetitive aaRS genes. Thereafter, the reaction solution was diluted 5-fold with another customized PURE system (PURE ΔaaRS) that lacks each target aaRS and contains DNAP-DNA, and incubated again. In this second reaction (round 2), both DNAP and the target aaRSs are expressed if the lacking aaRS protein is supplied (i.e., rescued) from the previous round of reaction. The expressed DNAP further replicates the long repetitive DNA through a repetitive DNA replication scheme that allows continuous DNA replication[28]. This serial dilution process was repeated for 8-10 rounds and aaRS gene concentration was measured by quantitative PCR (qPCR). If each target aaRS is continuously regenerated from replicating DNA, DNA replication should be sustained in later rounds. For the control, we conducted the same serial dilution experiment with a plasmid encoding GFP instead of aaRS (pET-GFP).

We conducted this serial dilution experiment with each of the 20 aaRSs. We observed that DNA was continuously replicated, and the concentrations were higher than those with pET-GFP control in the later rounds for all aaRSs except for GluRS, HisRS, and TrpRS (Fig. 2b). These results indicate that the DNA that encodes each of these 17 aaRSs can continuously replicate by regenerating self-encoded aaRS to complement the lack of aaRS activity. It should be noted that for GlnRS, we needed to reduce the initial pET-GlnRS concentration to 0.1 nM, one-tenth of the other pET-aaRSs, possibly because overexpression of GlnRS is harmful for translation, consistent with a previous study[29].

For aaRS, GluRS, HisRS, and TrpRS, DNA replication continued even in the control experiment with pET-GFP, possibly because a small amount of aaRS activity, which would remain even after omitting each aaRS, is sufficient for DNA replication. To test this possibility, we performed a single-round gene expression-coupled DNA replication assay in the PURE systems lacking each of the three aaRSs (Fig. 3a). The DNA replicated even without aaRSs in comparable amounts to that with all 20 aaRSs (Fig. 3b), confirming that the remaining activities of these aaRSs are sufficient for DNA replication. To circumvent this problem, we attempted to enhance the requirement of aaRS activity for DNA replication by adding unrelated DNAs as a translation load. For this translation load, linear DNAs encoding the other 19 aaRSs were used. This choice was reasonable because we eventually added similar DNAs in later experiments. In the presence of the unrelated DNAs (0.3 nM each), the DNA replication in the PURE system that omitted each aaRS (ΔGluRS, ΔHisRS, and ΔTrpRS) became significantly lower than that with all 20 aaRSs (Fig. 3c), indicating that the remaining activities are insufficient for DNA replication with the translation load. Thereafter, we performed a serial dilution experiment using pET-GluRS, pET-HisRS, and pET-TrpRS in the presence of the translational load and found that DNA replication with pET-aaRS continued and the concentration was higher than that with pET-GFP control in the later rounds (Fig. 3d). Taken together with the previous results in Fig. 2b, these results indicate that each of the 20 aaRS DNAs can continuously replicate by regenerating self-encoded aaRS to complement the lack of aaRS activity.

**DNA replication coupled with multiple aaRSs regeneration**

Thereafter, we simultaneously performed aaRS expression and DNA replication with multiple aaRSs. First, 20 types of aaRS were divided into four groups, from the most required Group 1 to the least required Group 4 (Fig. 4a), based on the dilution rates of the experiment in Fig. 1 (Table S2), which reflects the requirement of each aaRS for DNA replication. Thereafter, we conducted a serial dilution experiment for each group using the same serial dilution method, as shown in Fig. 2a. For Group 1, DNA replication was nearly stopped at round 3, and the DNA concentration of the five aaRS genes was similar to that of the pET-GFP control (Fig. 4b), indicating that the aaRSs in Group 1 were not sufficiently expressed to support their DNA replication. In contrast, for Groups 2 to 4, DNA replication was maintained until at least six rounds, and the DNA concentration of aaRS genes was higher than that of the pET-GFP control, indicating that these five aaRSs in Groups 2 – 4 were sufficiently expressed.

We performed simultaneous serial dilution experiments with 10 or 15 aaRSs using aaRS in Groups 2 to 4 (Fig. 4c). In all experiments, the DNA replication of pET-aaRS was maintained at round 4, where pET-GFP replication was almost stopped, whereas the difference in DNA concentration between pET-aaRS and pET-GFP was smaller than that shown in Fig. 4b. These results suggest that the expression levels were not sufficient to sustaining 10-15 aaRS regeneration.

**Dialyzed reaction**

To increase the expression of aaRSs, we employed a dialyzed reaction, in which low-molecular-weight compounds, such as amino acids, NTP, and dNTP, were supplied from the outer solution into the reaction mixture using a micro-dialyzer. First, we compared gene expression-coupled DNA replication and GFP expression in the dialyzed reaction with that in the nondialyzed reaction. Both DNA replication (Fig. 5a) and GFP expression (Fig. 5b) increased by approximately 3-folds in the dialyzed reaction. Thereafter, we employed the dialyzed reaction for the serial dilution experiment of the 15 aaRSs in Groups 2 to 4. We measured the DNA

concentration of one aaRS gene in each group during each round. We found that these aaRS DNA continued to replicate until round 4, and their concentrations were higher than those in the control experiment with pET-GFP (Fig. 5c). Next, we performed serial dilution experiments using all 20 pET-aaRSs. We found that DNA replication was almost stopped at round 3, and the DNA concentrations of four pET-aaRSs (AlaRS, SerRS, GlyRS, and TrpRS) were similar to those in the control experiment with pET-GFP. These results suggest that aaRS expression in the dialyzed reaction was still insufficient for all 20 aaRSs, including Group 1 aaRSs.

**Sequence modification to increase aaRS expression**

To investigate the reason for insufficient aaRS activity, particularly in Group 1, we measured the expression levels of 20 aaRS proteins. We performed the expression of each aaRS in the customized PURE system, which includes fluorescent-labeled lysin, and subjected it to SDS-PAGE analysis, followed by fluorescent imaging. The band intensities were significantly different among aaRSs, independent of their size (Fig. 6a). The quantified band intensities are presented in Fig. S6a. Thereafter, we evaluated the insufficiency of each aaRS expression for regeneration, which depends on the expression level and also the requirement of each aaRS for translation. To compare insufficiency, we divided the expression level of each aaRS (i.e., the band intensity) by each protein concentration in the PURE system, which represents the requirement for translation. The resultant values, named "satisfaction values," are shown in Fig. S6b. The aaRSs that represented lower satisfaction values were AlaRS, PheRS, IleRS, LeuRS, MetRS, ThrRS, and AsnRS, including all members in Group 1 and two aaRSs in Group 2.

To increase the expression of these aaRSs (AlaRS, PheRS, IleRS, LeuRS, MetRS, ThrRS, and AsnRS), we modified their sequences. Previously, a high GC ratio around the 1$^{st}$ codon has been reported to decrease translation efficiency[30]. Thereafter, we decreased the GC ratio of these seven aaRSs, which exceeded 50% in the 1$^{st}$ to 6$^{th}$ codons, without affecting the protein sequences. All modified aaRS-DNA (M) exhibited higher expression than the original aaRS DNA (O) (Fig. 6b), indicating that the GC ratio is one of the determining factors for expression. The quantified band intensities and satisfaction values are shown in Fig. S7a and S7b, respectively).

**DNA replication coupled with all 20 aaRSs regeneration**

Using the modified sequences, we performed a serial dilution experiment with 20 aaRSs in the dialysis reaction. In addition, we adjusted the initial concentrations (i.e., copy numbers) of each pET-aaRS with the range from 0.02 to 1.0 nM (Table. S4), based on the satisfaction values calculated from the expression data with a modified sequence (Fig. S7). The reaction scheme is shown in Fig. 7a, the same as in Fig. 2a, except that the reaction initiated with all 20 pET-aaRS and the PURE system used for dilution did not contain any aaRS proteins. With this experimental setup, the DNA replication of 20 aaRSs was maintained until round 5, and the DNA concentration of 20 aaRS genes was higher than that of the pET-GFP control (Fig. 7b). Similar results were obtained in a reproducible manner (Fig. S8b). These results demonstrate that 20 aaRSs are regenerated sufficiently to support continuous DNA replication in the PURE system, which lacks all 20 aaRSs. To directly examine aaRS regeneration, we performed the same dialyzed serial dilution experiment with fluorescent-labeled lysyl-tRNA to label newly synthesized proteins. After SDS-PAGE and fluorescence imaging, we continuously detected several bands corresponding to aaRSs until round 4 (Fig. S10), although it tended to diminish as the rounds proceeded.

We also evaluated translational activity during the serial dilution experiment. Aliquots of the reaction mixture at each round were diluted with the PURE system that lacked all aaRSs and contained linear DNA encoding GFP. The diluted mixture was then incubated for GFP expression. We measured the rate of increase in GFP fluorescence as a translation activity (Fig. 7c). For the control experiment, in which the serial dilution experiment was conducted with pET-GFP, GFP translation activity rapidly decreased to 25% at Round 2 and below the detection limit at Round 3-5, whereas for the serial dilution experiment with 20 pET-aaRSs, GFP translation activity decreased relatively slowly while maintaining 25% activity at Round 4, confirming that the translation activity was maintained to a certain extent even after dilution with the PURE system that lacked 20 aaRSs due to the regeneration of aaRSs.

**Discussion**

To realize self-reproductive artificial systems, the regeneration of translation proteins in an in vitro translation system is necessary. To date, two pioneering studies have tackled this challenge and successfully regenerate 7-13 aaRSs in their functional form[26,27]. In this study, we advanced this challenge toward the simultaneous regeneration of all 20 aaRSs, which comprises approximately half of the protein components in the PURE system, except for ribosomes, and succeeded in coupling the regeneration with the replication of DNAs that encode 20 aaRS by employing a dialysis reaction, modifying the DNA sequence, and adjusting the aaRS gene copy number. Although there is still room for improvement, these results provide a large step toward the development of self-reproducing artificial cells.

One of the new aspects of the system developed here is the coupling of aaRS expression with DNA replication, a prerequisite for a self-reproductive artificial system. Here, we employed a simple DNA replication scheme, repetitive DNA replication, which we found in our previous study[28]. This scheme requires only one protein, phi29 DNA polymerase, for continuous replication, which is simpler than other DNA replication schemes used in the cells or viruses. In the repetitive DNA replication scheme, the polymerase synthesizes a long repetitive DNA from the initial circular DNA through rolling-circle replication and further replicates the long repetitive DNA while maintaining the DNA size within a certain range, possibly through repeating polymerization from random sites and hybridization of the produced single-stranded DNAs. An advantage of this replication scheme is the small number of required proteins, which saves the translational capacity of the PURE system and allows simultaneous expression of multiple aaRSs, as demonstrated in this study. Another advantage is the ease of increasing the number of genes for simultaneous expression and replication. In this scheme, multiple DNA fragments can be replicated simultaneously, and thus, genes can be increased by simply adding a circular DNA that encodes a new gene in the initial mixture. In addition, the copy number of each gene can be adjusted by simply changing the initial DNA concentration. Such flexible adjustment is difficult when using large polycistronic DNA. The use of this simple DNA replication scheme enhances the feasibility of realizing a self-regenerative artificial system.

One of the largest hurdles in achieving self-regenerative artificial systems is the insufficient translation activity of the reconstituted translation system. This study used PURE systems, reconstituted translation systems of *E. coli*, and found that up to 10 aaRSs expression was highest in the nondialyzed condition (Fig. 4c). Similar limitations were reported by Libicher et al. and Lavickova et al., who succeeded in regenerating up to 7 and 13 aaRSs[26,27]. To overcome this limitation, we employed two methods: dialysis and sequence modification. As reported previously[22,31] and also in this study (Fig. 5b), the dialyzed reaction enhanced translation, which means that the maximum translation is limited, at least partially, by the depletion of low-

molecular-weight compounds, such as amino acids and NTPs. Dialysis, which may mimic nutrient transportation in natural cells, is useful for achieving a self-regenerative system. Sequence modification is another effective method. As demonstrated in this study, the low GC ratio around the 1st codon enhanced gene expression for all aaRSs tested (Fig. 6b). We also found that the gene expression level varied significantly among aaRSs, consistent with previous expression studies[24,25]. Some aaRS expression was lower than the others even after sequence modification, implying that further improvement in the sequences is possible. However, it is not known how these sequences can be improved. To obtain this knowledge, we may need an evolutionary experiment to enhance the expression of TFs in the PURE system.

Several challenges remain in the development of self-regenerative artificial systems. Although we have demonstrated 20 aaRS regenerations, there remain other proteins and RNAs to be regenerated, such as translation initiation, elongation, releasing factors, tRNAs, and ribosomes. Recently, several studies have reported substantial progress in this direction, such as translation with in vitro-synthesized tRNAs[32–36] and other translation factors[37], and ribosome reconstitution of purified ribosomal proteins[22,38,39]; however, the translation factors in these studies were synthesized and purified in advance, except for our recent tRNA study[40], to disregard the expression level. To regenerate these RNA and proteins in the PURE system, further improvement of gene expression levels and activity of translation factors is required. This improvement may be possible using the evolutionary method as described above. In a previous study, we succeeded in the evolution of artificial genomic DNA that encodes phi29 DNA polymerase and a recombinase through a serial dilution cycle in a compartmentalized reaction[41]. The continuous DNA replication coupled with aaRS replication constructed in this study will be a basis for evolutionary experiments on translation factors.

## Methods

### DNA preparation

The circular DNA encoding the phi29 DNA polymerase used in the aaRS activity assay (Fig. 1) was prepared as follows. First, we PCR-amplified DNA fragments encoding phi29 DNA polymerase using primers 1 and 2 and the plasmid (pUC-phi29DNAPevo56_loxP_cre[42]) as a template. The PCR product was purified using the QIAquick PCR Purification Kit (QIAGEN), which was used for all DNA purification procedures in this study. The purified DNA fragment was digested with 0.6 U/μL SphI (TaKaRa, Japan) according to the manufacturer's instructions for 1 h at 37 °C and then self-ligated with 17.5 U/μL T4 DNA ligase (TaKaRa) at 16 °C overnight, followed by purification. The PCR product before SphI digestion was used as the linear DNA encoding phi29 DNA polymerase (DNAP-DNA).

The circular DNAs encoding each aaRS (pET-aaRS) were the expression plasmids that are used to prepare the PURE system[20,43] except for four aaRSs (AlaRS, AsnRS, PheRS, and ThrRS), which were originally encoded in the pQE vector. For these aaRSs, we moved the gene regions to the pET vector. Each aaRS gene region was PCR-amplified using pQE-aaRS plasmid and primers 3-7 and then ligated with vector fragment PCR-amplified using pET-GFP and primers 8 and 9 using the In-Fusion Cloning Kit (TaKaRa). The linear DNA encoding each aaRS used in Figs. 1 and 6 was amplified by PCR, using each pET-aaRS as a template and primers 10 and 11. The seven sequence-modified pET-aaRSs were prepared by PCR amplification followed by self-ligation using mutated primers 12-18 and the In-Fusion cloning Kit (TaKaRa).

The plasmid that encodes GFP under the T7 promoter (pET-GFP) was previously constructed as pETG5tag[44]. The linear DNA fragment encoding GFP used in Fig. 7b was prepared by PCR

using pET-GFP as a template and primers 10 and 11.

**Activity assay of the aaRSs synthesized in the PURE system (Fig. 1)**

First, the linear DNA encoding of each aaRS (2 nM) was incubated at 30 °C for 16 h using a commercially available PURE system (PURExpress 2.0, GeneFrontier) to express aaRS. Thereafter, an aliquot of the reaction solution was diluted with a dilution buffer (50 mM HEPES-KOH (pH7.6), 100 mM KCl, 10 mM $MgCl_2$, 7 mM 2-mercaptoethanol, and 30% glycerol) at different rates (Table S2) and added to the second customized PURE system for DNA replication. The second customized PURE system lacks each aaRS of interest and contains a circular DNA that encodes phi29 DNA polymerase and all other factors required for DNA replication, transcription, and translation, as described previously[41] except that the concentration of magnesium acetate was 10.8 mM. The composition was optimized for DNA replication (Table S1). After incubation at 30 °C for 16 h, the reaction solution was diluted 10000-fold with 1 mM EDTA (pH 8.0), and the DNA concentration was measured by qPCR using primer 19-20.

**Serial dilution experiment**

In round 1, a single or multiple circular plasmid DNA encoding each aaRS (pET-aaRS) and the linear DNA encoding phi29 DNA polymerase (DNAP-DNA) were incubated in the customized PURE system containing all 20 aaRSs (the complete composition is shown in Table S1) at 30 °C for 16 h. The initial pET-aaRS concentrations were 1.0, 0.2, and 0.1 nM for single (Fig. 2), five (Fig. 4b), and more (Figs. 4c and 5cd) aaRSs regeneration, respectively. The DNAP-DNA concentration was set to one-tenth of the total initial pET-aaRS concentration and was kept constant in the later rounds. For the single regeneration experiments with GlnRS, the initial pET-GlnRS concentration was decreased to 0.1 nM, whereas the DNAP-DNA concentration was the same (0.1 nM). For the experiment with the translation load (Fig. 3), the initial concentration of each pET-aaRS was 0.1 nM, the DNAP-DNA was 0.1 nM, and 19 linear DNAs encoding the other aaRSs were added at 0.3 nM each. For the 20 aaRS regeneration shown in Fig. 7, the initial total concentration of pET-aaRSs was 3 nM, and each aaRS concentration was varied from 0.02 to 1 nM (Table S4). In the subsequent rounds, the reaction solution after incubation in the previous round was diluted 5-fold with another customized PURE system (PURE ΔaaRS) that lacks aaRSs of interest and contains DNAP-DNA and incubated again at 30 °C for 16 h. This serial dilution process was repeated for the indicated rounds, and aaRS gene concentration was measured by qPCR. For the control experiments, pET-GFP was used instead of pET-aaRSs.

**Dialyzed reaction**

For the dialyzed reaction, 20 μL of the reaction solution of the serial dilution experiment, which additionally contained 1/20 volume of DnaK mix and 1/40 volume of GroEL mix (Gene Frontier), was placed in a microdialyzer (Scienova) and dialyzed against 200 μL of dialysis buffer (Table S3) with shaking at 500 rpm at 30 °C overnight. The dialysis buffer contained all components except for all proteins and tRNAs in the customized PURE system, and the concentration of magnesium acetate and dNTPs were changed to 6.36 mM and 0.06 mM, respectively. The dNTP and magnesium acetate concentrations in the customized PURE system used for dilution were also decreased to 0.06 mM and 8.39 mM, respectively. The dialysis buffer also contained 100 μg/mL fosfomycin (Wako Chemical) and 50 μ μg/mL ampicillin to prevent bacterial growth. In the serial dilution experiment (Fig. 7), the dialysis started at round 2.

**Analysis of aaRS expression amount**

Linear DNA (5 nM) encoding each aaRS was incubated at 30 °C for 8 h in the customized

PURE system containing fluorescent-labeled lysyl-tRNA (FluoroTect GreenLys, Promega). After expression, an aliquot (5 μL) was treated with 0.5 μL of 5 mg/ml RNase A (QIAGEN, Hilden, Germany) at 37 °C for 30 min, incubated at 95 °C for 5 min in SDS sample buffer (50 mM Tris-HCl (pH 7.4), 2% SDS, 0.86 M 2-mercaptoethanol, and 10% glycerol) and subjected to 10% SDS-PAGE. The synthesized fluorescent-labeled proteins were detected using FUSION-SL4 (Vilber-Lourmat) and band intensities were analyzed.

**Translation activity assay in Fig. 7c**

Aliquots of the reaction mixtures during the serial transfer experiment were diluted 5-fold with the customized PURE system that lacks all 20 aaRSs and contains linear DNA (10 nM) that encodes GFP. The mixture was incubated at 30 °C for 12 h. GFP fluorescence was measured every 30 min (Mx3005P, Agilent Technologies) (Fig. S9). From the time-course data, the slope of the five measured points, where the fluorescence increased linearly, was defined as the translation activity. Translation activity was set to zero when the slope became negative.

**Acknowledgements**

This work was supported by JST, CREST Grant Number JPMJCR20S1, Japan.

**Author contributions**

K. H. and N.I. designed the project and wrote the manuscript. K. H. performed all experiments and analysis.

**Competing interests**

The authors declare no competing interests.

**Materials & Correspondence**

Correspondence and requests for materials should be addressed to N.I.

**Reference**

1. Szostak, J. W., Bartel, D. P. & Luisi, P. L. Synthesizing life. *Nature* **409**, 387–390 (2001).
2. Forster, A. C. & Church, G. M. Towards synthesis of a minimal cell. *Mol. Syst. Biol.* **2**, 45 (2006).
3. Damiano, L. & Stano, P. On the "Life-Likeness" of Synthetic Cells. *Front. Bioeng. Biotechnol.* **8**, 1–6 (2020).
4. Cho, E. & Lu, Y. Compartmentalizing cell-free systems: Toward creating life-like artificial cells and beyond. *ACS Synth. Biol.* **9**, 2881–2901 (2020).
5. Ivanov, I. *et al.* Bottom-Up Synthesis of Artificial Cells: Recent Highlights and Future Challenges. *Annu. Rev. Chem. Biomol. Eng.* **12**, 287–308 (2021).
6. Wang, C., Yang, J. & Lu, Y. Modularize and Unite: Toward Creating a Functional Artificial Cell. *Front. Mol. Biosci.* **8**, 1–14 (2021).
7. Laohakunakorn, N. *et al.* Bottom-Up Construction of Complex Biomolecular Systems With Cell-Free Synthetic Biology. *Front. Bioeng. Biotechnol.* **8**, 213 (2020).
8. Olivi, L. *et al.* Towards a synthetic cell cycle. *Nat. Commun.* **12**, (2021).
9. Ichihashi, N. What can we learn from the construction of in vitro replication systems? *Ann. N. Y. Acad. Sci.* **1447**, 144–156 (2019).
10. Forlin, M., Lentini, R. & Mansy, S. S. Cellular imitations. *Curr. Opin. Chem. Biol.* **16**, 586–592 (2012).

11. Buddingh', B. C. & van Hest, J. C. M. Artificial Cells: Synthetic Compartments with Life-like Functionality and Adaptivity. *Acc. Chem. Res.* **50**, 769–777 (2017).
12. Yewdall, N. A., Mason, A. F. & Hest, J. C. M. Van. The hallmarks of living systems: Towards creating artificial cells. *Interface Focus* **8**, 20180023 (2018).
13. Silverman, A. D., Karim, A. S. & Jewett, M. C. Cell-free gene expression: an expanded repertoire of applications. *Nat. Rev. Genet.* **21**, 151–170 (2020).
14. Noireaux, V. & Liu, A. P. The New Age of Cell-Free Biology. *Annu. Rev. Biomed. Eng.* **22**, 51–77 (2020).
15. Blain, J. C. & Szostak, J. W. Progress Toward Synthetic Cells. *http://dx.doi.org/10.1146/annurev-biochem-080411-124036* **83**, 615–640 (2014).
16. Gaut, N. J. & Adamala, K. P. Reconstituting Natural Cell Elements in Synthetic Cells. *Adv. Biol.* **5**, 2000188 (2021).
17. Jia, T. Z., Wang, P. H., Niwa, T. & Mamajanov, I. Connecting primitive phase separation to biotechnology, synthetic biology, and engineering. *J. Biosci. 2021 463* **46**, 1–28 (2021).
18. Stano, P. Gene Expression Inside Liposomes: From Early Studies to Current Protocols. *Chem. – A Eur. J.* **25**, 7798–7814 (2019).
19. Bergquist, P. L., Siddiqui, S. & Sunna, A. Cell-Free Biocatalysis for the Production of Platform Chemicals. *Front. Energy Res.* **8**, 193 (2020).
20. Shimizu, Y. *et al.* Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* **19**, 751–755 (2001).
21. Awai, T., Ichihashi, N. & Yomo, T. Activities of 20 aminoacyl-tRNA synthetases expressed in a reconstituted translation system in Escherichia coli. *Biochem. Biophys. Reports* **3**, 140–143 (2015).
22. Li, J. *et al.* Cogenerating Synthetic Parts toward a Self-Replicating System. *ACS Synth. Biol.* **6**, 1327–1336 (2017).
23. Shepherd, T. R. *et al.* De novodesign and synthesis of a 30-cistron translation-factor module. *Nucleic Acids Res.* **45**, 10895–10905 (2017).
24. Libicher, K., Hornberger, R., Heymann, M. & Mutschler, H. In vitro self-replication and multicistronic expression of large synthetic genomes. *Nat. Commun.* **11**, (2020).
25. Doerr, A., Foschepoth, D., Forster, A. C. & Danelon, C. In vitro synthesis of 32 translation-factor proteins from a single template reveals impaired ribosomal processivity. *Sci. Rep.* **11**, 1–12 (2021).
26. Libicher, K. & Mutschler, H. Probing self-regeneration of essential protein factors required for: In vitro translation activity by serial transfer. *Chem. Commun.* **56**, 15426–15429 (2020).
27. Lavickova, B., Laohakunakorn, N. & Maerkl, S. J. A partially self-regenerating synthetic cell. *Nat. Commun.* **11**, (2020).
28. Okauchi, H., Sakatani, Y., Otsuka, K. & Ichihashi, N. Minimization of Elements for Isothermal DNA Replication by an Evolutionary Approach. *ACS Synth. Biol.* **9**, 1771–1780 (2020).
29. Sherman, J. M., Rogers, M. J. & Söll, D. Competition of aminoacyl-tRNA synthetases for tRNA ensures the accuracy of aminoacylation. *Nucleic Acids Res.* **20**, 2847–2852 (1992).
30. Gene Frontier web site, 6 Tips about the template DNA Design. https://purefrex.genefrontier.com/resources/tech_notes/template-dna-design.html.
31. Kazuta, Y., Matsuura, T., Ichihashi, N. & Yomo, T. Synthesis of milligram quantities of proteins using a reconstituted in vitro protein synthesis system. *J. Biosci. Bioeng.* **118**, (2014).
32. Hibi, K. *et al.* Reconstituted cell-free protein synthesis using in vitro transcribed

tRNAs. *Commun. Biol.* **3**, (2020).
33. Fujino, T., Tozaki, M. & Murakami, H. An Amino Acid-Swapped Genetic Code. *ACS Synth. Biol.* **9**, 2703–2713 (2020).
34. Iwane, Y. *et al.* Expanding the amino acid repertoire of ribosomal polypeptide synthesis via the artificial division of codon boxes. *Nat. Chem.* **8**, 317–325 (2016).
35. Chen, J., Chen, M. & Zhu, T. F. Translating protein enzymes without aminoacyl-tRNA synthetases. *Chem* **7**, (2021).
36. Calles, J., Justice, I., Brinkley, D., Garcia, A. & Endy, D. Fail-safe genetic codes designed to intrinsically contain engineered organisms. *Nucleic Acids Res.* **47**, 10439–10451 (2019).
37. Endy, D. Experimental tests of functional molecular regeneration via a standard framework for coordinating synthetic cell building. *bioRxiv* 2021.03.03.433818 (2021).
38. Tamaru, D., Amikura, K., Shimizu, Y., Nierhaus, K. H. & Ueda, T. Reconstitution of 30S ribosomal subunits in vitro using ribosome biogenesis factors. *RNA* **24**, 1512–1519 (2018).
39. Shimojo, M. *et al.* In vitro reconstitution of functional small ribosomal subunit assembly for comprehensive analysis of ribosomal elements in E. coli. *Commun. Biol.* **3**, 142–142 (2020).
40. Miyachi, R., Shimizu, Y., Ichihashi, N. *Jxiv* 10.51094/jxiv.35 (2022) doi:DOI: https://doi.org/10.51094/jxiv.35.
41. Okauchi, H. & Ichihashi, N. Continuous Cell-Free Replication and Evolution of Artificial Genomic DNA in a Compartmentalized Gene Expression System. *ACS Synth. Biol.* **10**, 3507–3517 (2021).
42. Sakatani, Y., Yomo, T. & Ichihashi, N. Self-replication of circular DNA by a self-encoded DNA polymerase through rolling-circle replication and recombination. *Sci. Rep.* **8**, 13089 (2018).
43. Shimizu, Y., Kanamori, T. & Ueda, T. Protein synthesis by pure translation systems. *Methods* **36**, 299–304 (2005).
44. Ito, Y., Kawama, T., Urabe, I. & Yomo, T. Evolution of an Arbitrary Sequence in Solubility. *J. Mol. Evol.* **58**, 196–202 (2004).

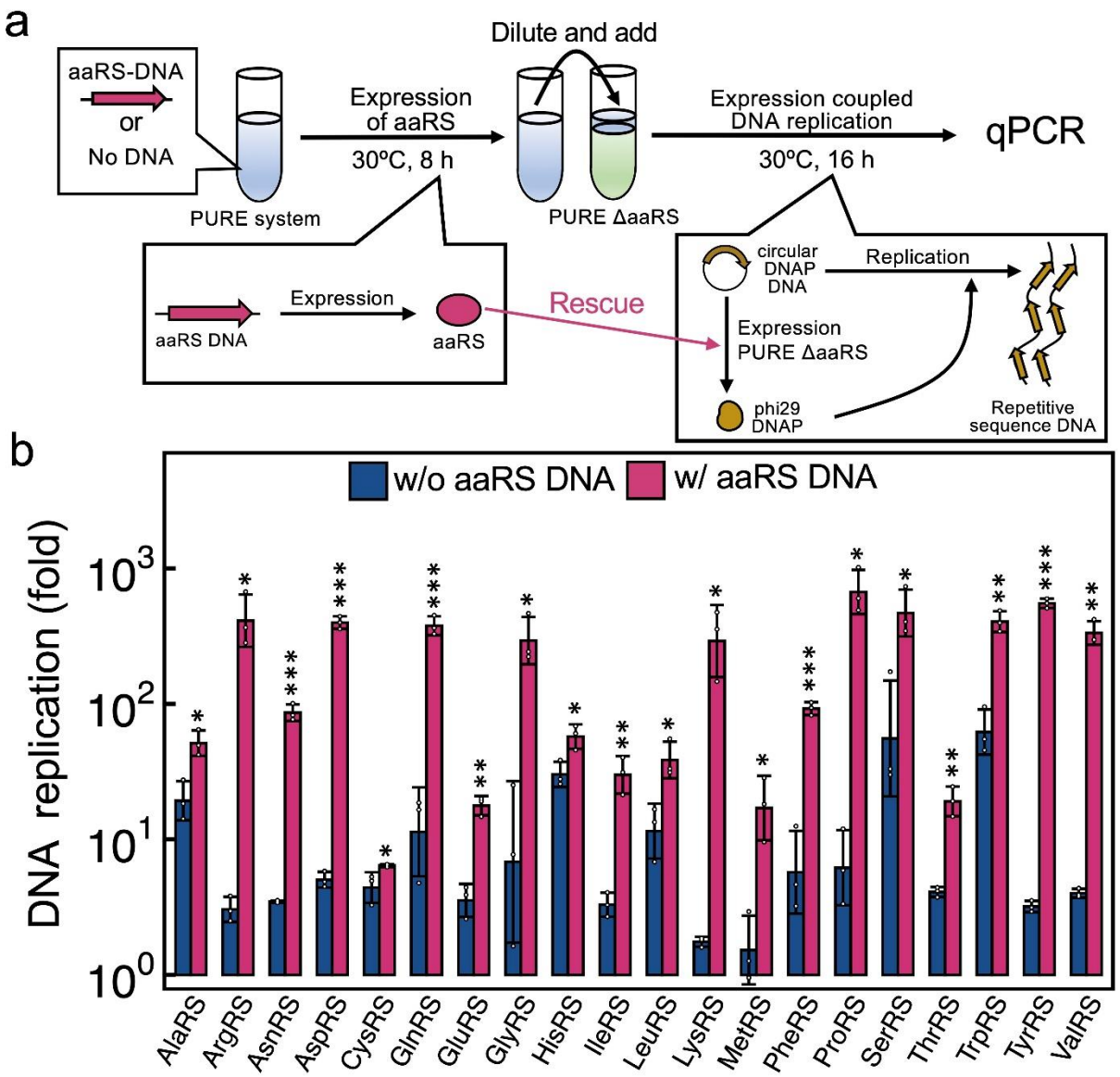


**Figure 1. The activity assay of each aaRS expressed in a PURE system.**
a) Scheme of the assay method. In the first reaction, each aaRS was expressed in a PURE system (PURExpress 2.0) at 30˚C for 8 h. An aliquot of each reaction solution was diluted at a difference rate shown in Table S2 and added to the customized PURE system that lacks the aaRS of interest and contains a circular DNA that encodes phi29 DNA polymerase. The second reaction mixture was incubated at 30˚C for 16 h, during which phi29 DNA polymerase is expressed depending on the aaRS expressed in the first reaction and performs rolling-circle DNA replication using the circular DNA as a template. The DNA concentration was measured by quantitative PCR (qPCR), and the DNA replication rate was calculated by dividing the concentration at 16 h by that at 0 h. b) Replication rate of the circular DNAP-DNA with or without aaRS-DNA. Error bars represent standard deviations of three independent experiments. Significance test was performed by student's t-test (*: $p$-value < 0.05, **: $p$-value <0.01, ***: $p$-value < 0.001).

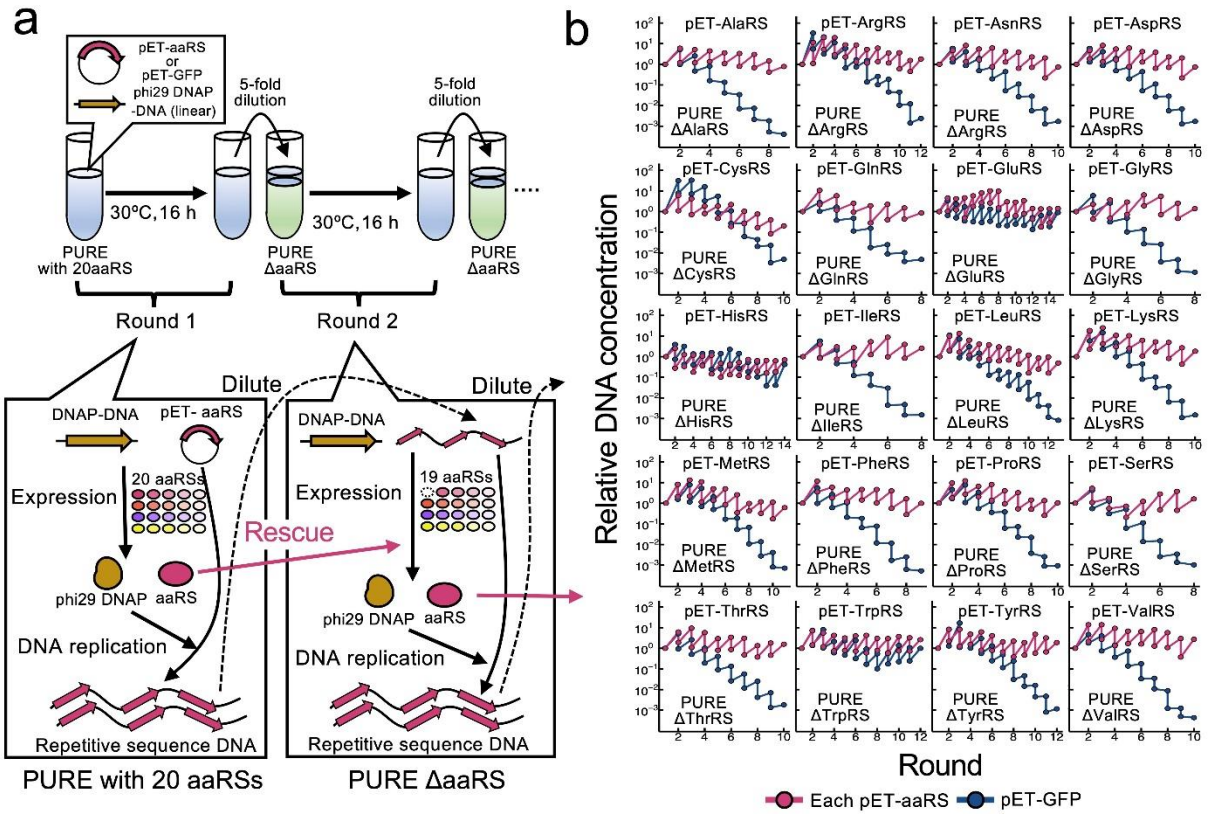


**Figure 2. DNA replication coupled with one of the 20 aaRS regeneration.**
a) Scheme of the serial dilution experiment. The first reaction mixture contains 1 nM pET-aaRS (0.1 nM for GlnRS) or 1 nM pET-GFP and 0.1 nM linear DNA encoding phi29 DNA polymerase (DNAP-DNA) in the customized PURE system that contains all 20 aaRSs. In the first reaction (Round 1), each aaRS and phi29 DNA polymerase are expressed, and pET-aaRS (or pET-GFP) are replicated through rolling-circle replication to synthesize long repetitive sequence DNA at 30˚C for 16 h. In Round 2, the reaction solution of the previous round is diluted 5-fold with the customized PURE system that lacks each one of 20 aaRSs (PURE ΔaaRS) and contains DNAP-DNA. If the aaRS is sufficiently synthesized in the previous round, the lack of each of 20 aaRSs is rescued, and thus DNAP and aaRS are expressed again. The DNAP further replicates the long repetitive sequence DNA. Thereafter, the reaction mixture is diluted again with PURE ΔaaRS for the next round of reaction, where DNA replication continues if a sufficient amount of active aaRS is expressed from the repetitive sequence DNA.
b) Trajectories of DNA concentration in the serial dilution experiments. DNA concentrations of each aaRS gene or gfp gene before and after the dilution step was measured by qPCR and normalized based on the concentration after Round 1. The raw data before the normalization are shown in Fig. S1.

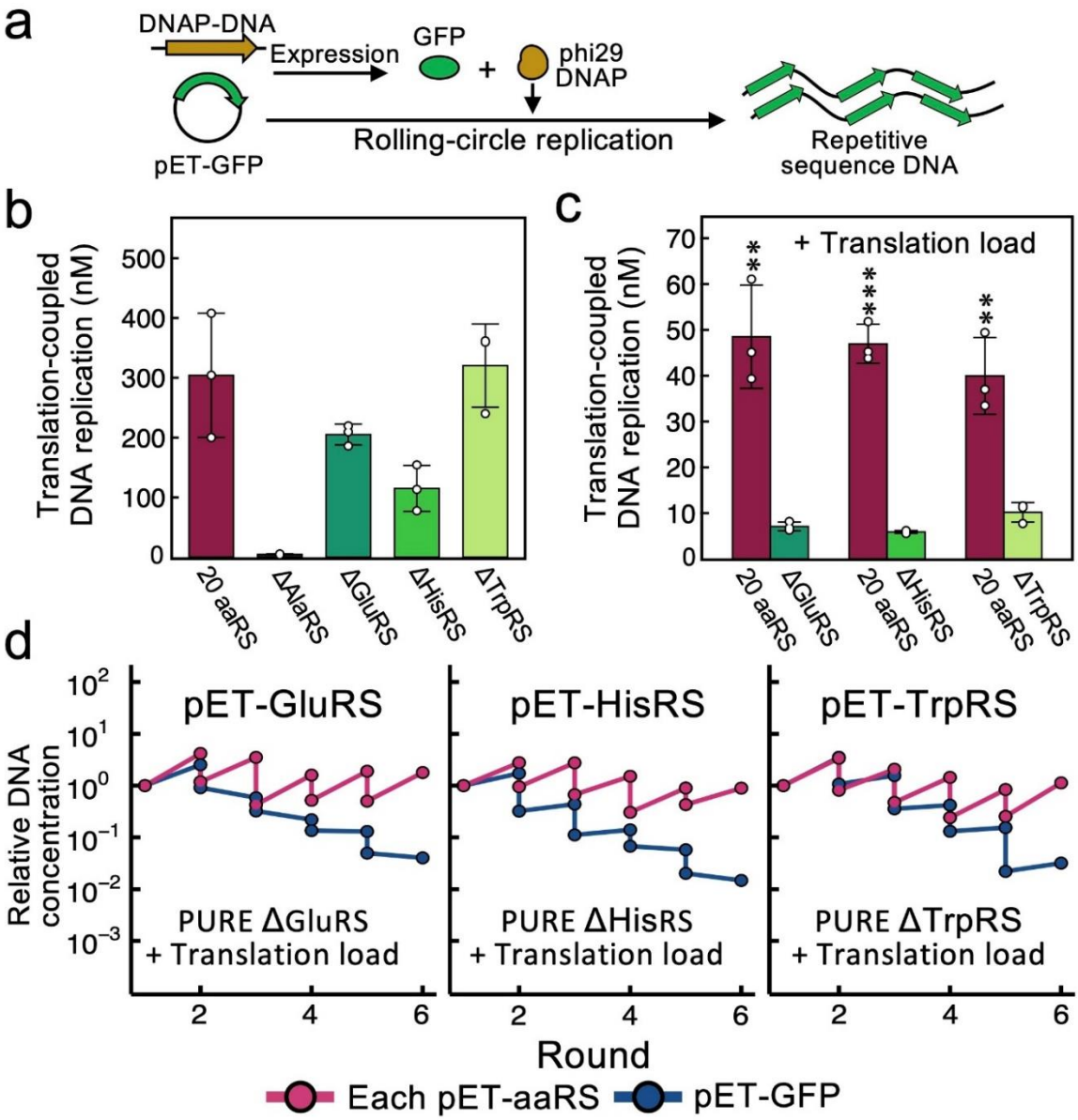


**Figure 3. Gene expression-coupled DNA replication assay in the PURE systems that lacks AlaRS, GluRS, HisRS and TrpRS.**

a) Scheme of the single-round DNA replication coupled with the expression of phi29 DNA polymerase in the PURE systems, one of the three aaRSs is omitted. If the omitted aaRS is remained in other components, pET-GFP is replicated through the expression of phi29 DNAP. b) Replication of pET-GFP DNA (initially 1 nM) in the PURE systems that lack each aaRS (AlaRS, GluRS, HisRS, and TrpRS). The DNA concentration was measured by qPCR. Error bars represents standard deviations of three independent experiments. The result of AlaRS is shown for comparison. c) Replication of pET-GFP DNA (initially 1 nM) in the PURE systems, which lacks each aaRS (GluRS, HisRS, or TrpRS) and contains the other 19 aaRS DNA (0.3 nM each) as a translation load. Error bars represent standard deviations of three independent experiments. Significance test was performed by student's t-test (**: *p*-value <0.01, ***: *p*-value < 0.001). d) Trajectory of aaRS or GFP gene concentration in the serial dilution experiment as shown in Fig. 2a for GluRS, HisRS, and TrpRS in the presence of the translation load. DNA concentrations of each aaRS gene before the dilution step was measured by qPCR and normalized based on the concentration after Round 1. The raw data before the normalization are shown in Fig. S2.

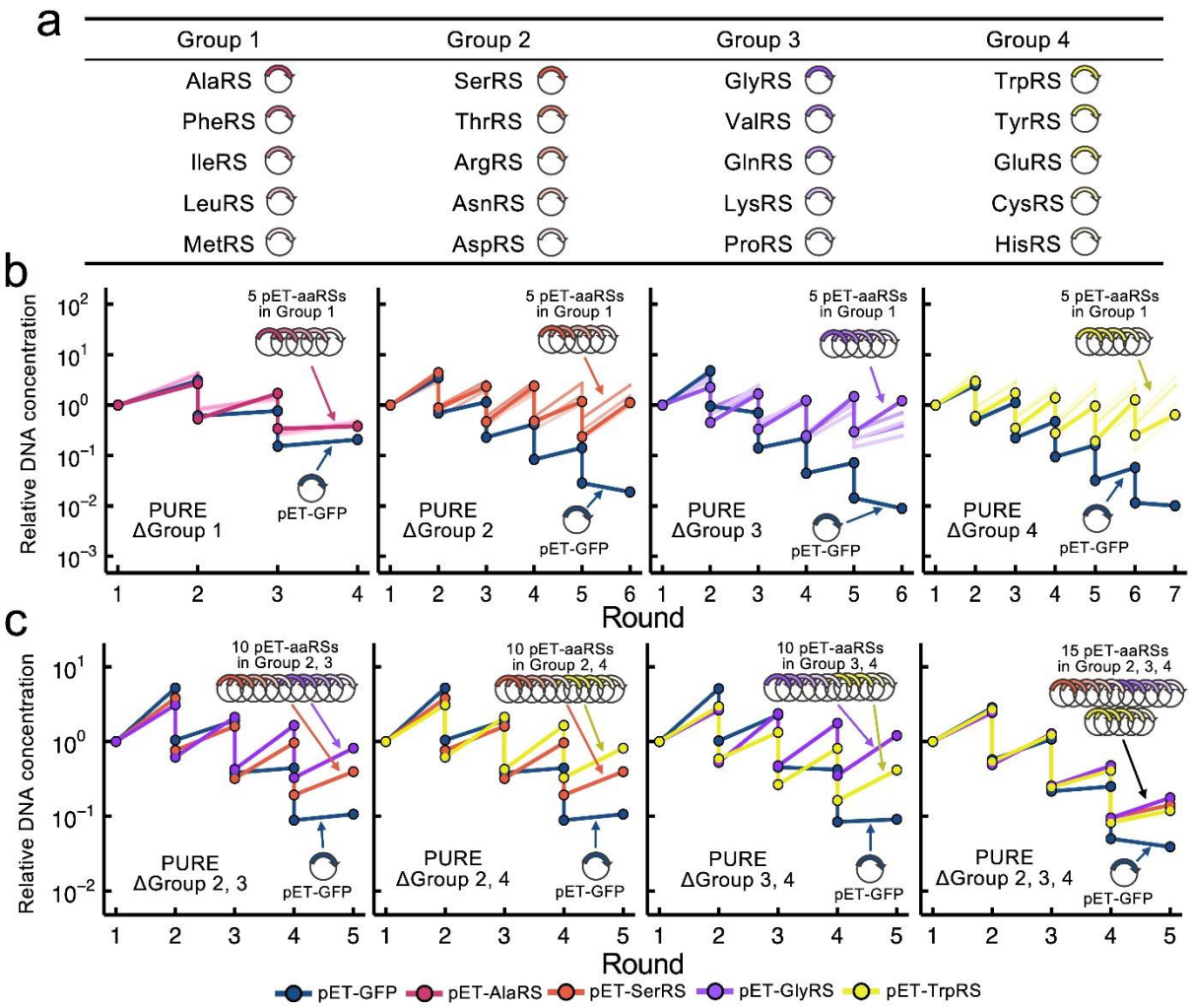


**Figure 4. DNA replication coupled with multiple aaRS regeneration.**
a) Twenty aaRSs are divided into four groups according to the requirement for DNA replication. b-c) Trajectories of DNA concentrations in the serial dilution experiment for the simultaneous regeneration of five (b) or 10-15 (c) aaRSs. Each customized PURE system that lacks multiple aaRSs of interest were used for dilution. DNA concentrations of each aaRS gene after the dilution step was measured by qPCR and normalized based on the concentration after Round 1. For 10-15 simultaneous experiments (c), only one aaRS gene in each group was measured. DNA concentrations of each aaRS gene before and after the dilution step was measured by qPCR and normalized based on the concentration after Round 1. The raw data before the normalization are shown in Figs. S3 and S4.

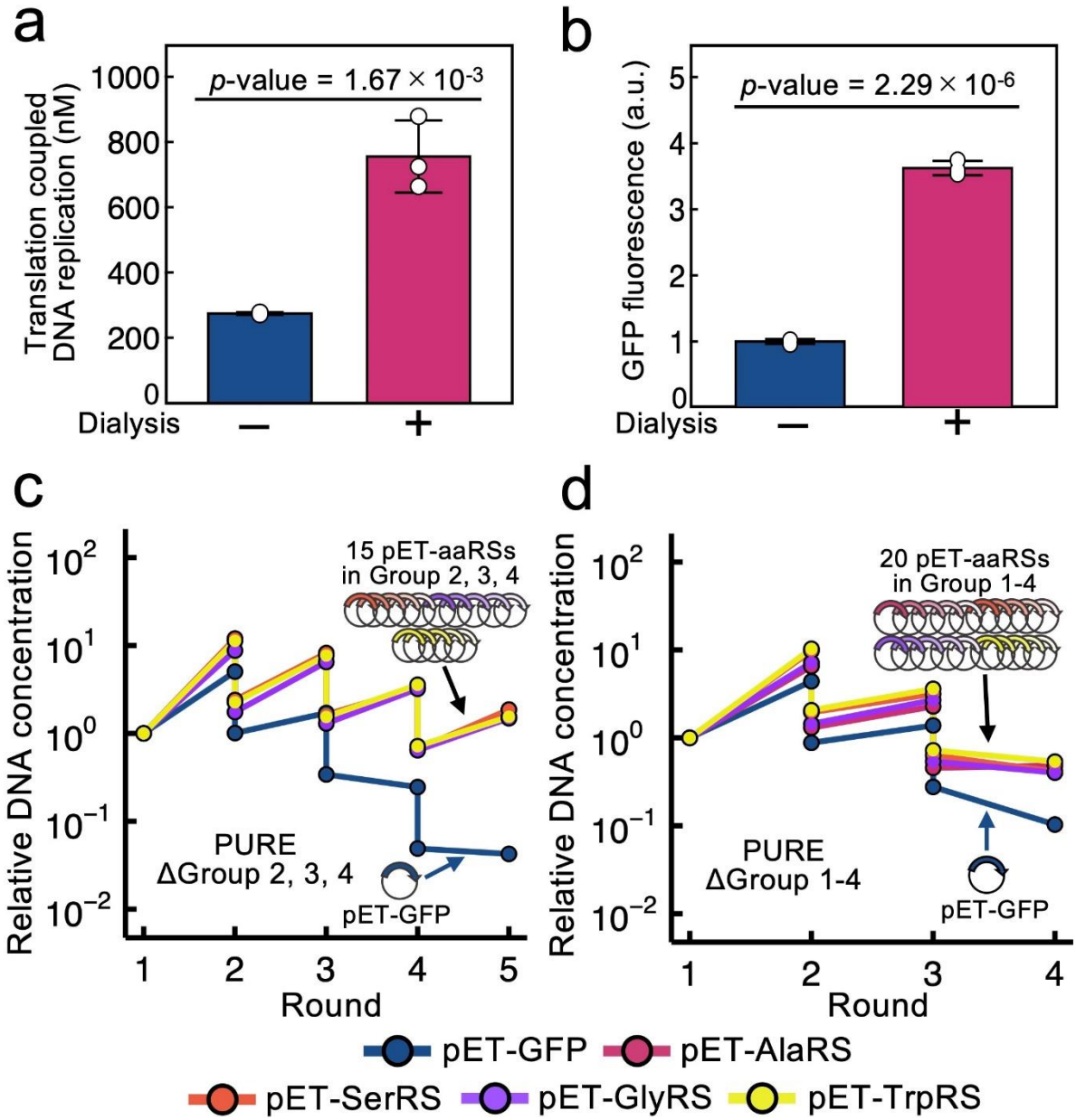


**Figure 5. Dialyzed reaction.**

a, b) The effect of dialysis on the single-round translation-coupled DNA replication assay. (a) The GFP gene concentration measured by qPCR. (b) GFP fluorescence. The assay was performed as shown in Fig. 3a. The reaction mixture contains DNAP-DNA (0.1 nM) and pET-GFP (1 nM) in the customized PURE system that contains all 20 aaRSs and was incubated at 30˚C for 16 h. Error bars represent standard deviations of three independent experiments. The significance test was performed by student's t-test. c, d) Trajectories of pET-aaRSs and pET-GFP in the dialyzed serial dilution experiment with 15 (c) and 20 (d) aaRSs. DNA concentrations of each aaRS gene before and after the dilution step was measured by qPCR and normalized based on the concentration after Round 1. The raw data before the normalization are shown in Fig. S5.

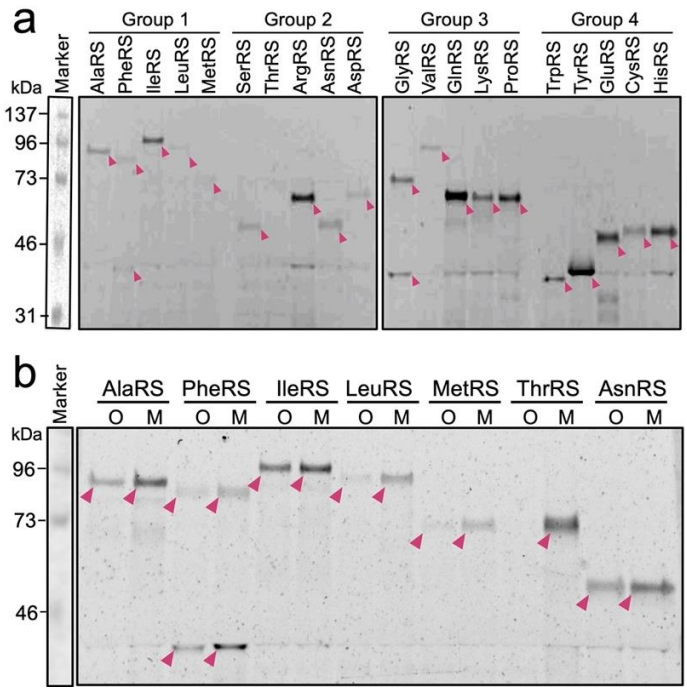


**Figure 6. Expression amounts of aaRS proteins.**

a) Expression amounts of the original 20 aaRSs. Each aaRS was expressed from each linear aaRS DNA (5 nM) at 30˚C for 8 h in the customized PURE system that contains all 20 aaRSs and a fluorescent labeled lysyl-tRNA. The mixture was subjected to SDS-PAGE and the fluorescence was detected. The expected bands of each aaRS are indicated by the red arrowheads. b) Effect of sequence modification of the seven aaRSs, which exhibit lower satisfactory values. The results of the original (O) and modified (M) are shown. The sequences before and after modifications are shown in Table S5.

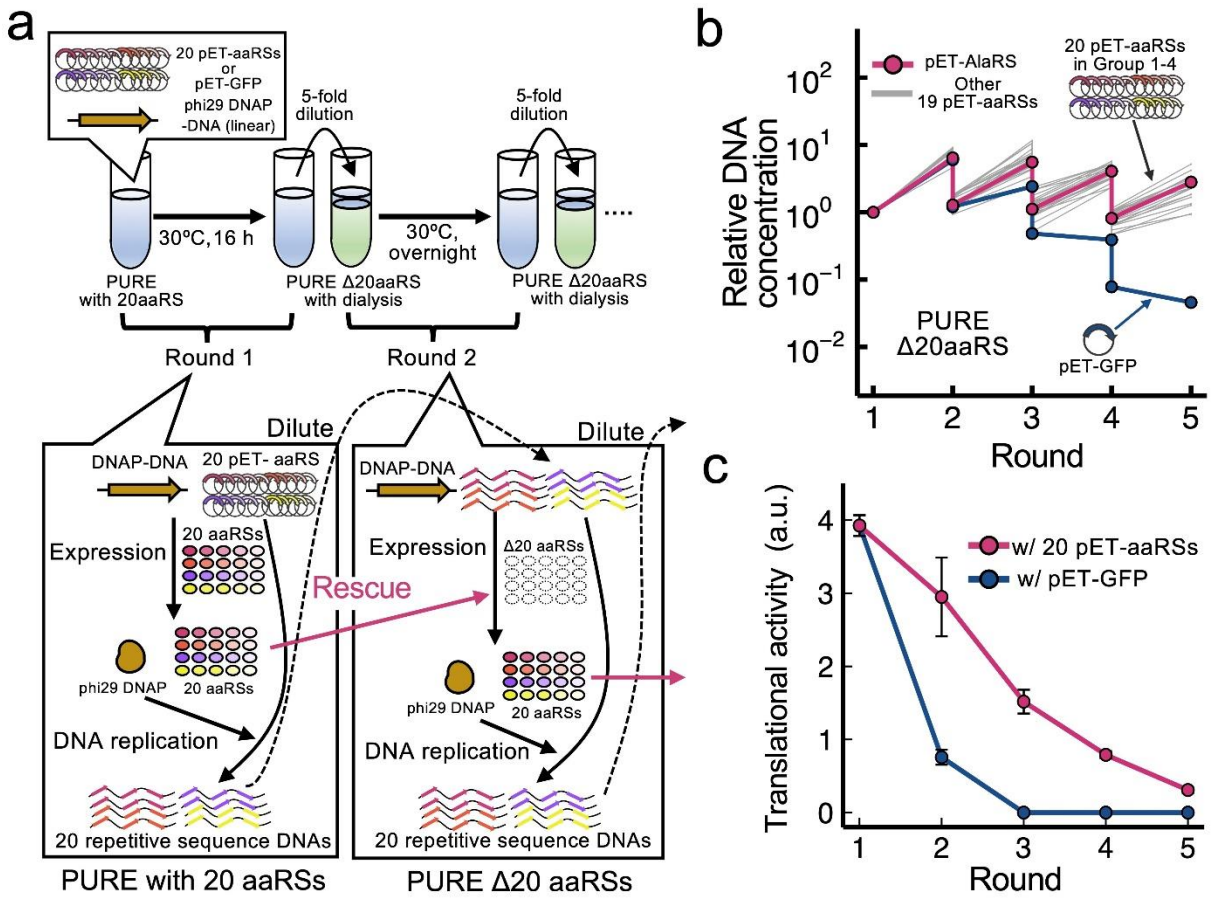


**Figure 7. DNA replication coupled with all 20 aaRS regeneration.**
a) Scheme of the serial dilution experiment. The first reaction mixture contains 20 pET-aaRSs (0.02-1.0 nM, the initial concentration of each pET-aaRS is shown in Table. S4) or pET-GFP (1 nM) as a control, and 0.1 nM linear DNA encoding phi29 DNA polymerase (DNAP-DNA) in the customized PURE system that contains all 20 aaRSs. In the first reaction (Round 1), each aaRS and phi29 DNA polymerase are expressed, and pET-aaRSs (or pET-GFP) are replicated through rolling-circle replication to synthesize long repetitive sequence DNA at 30˚C for 16 h. In Round 2, the reaction solution of the previous round is diluted 5-fold with the customized PURE system that lacks all 20 aaRSs and contains DNAP-DNA (0.1 nM). If the aaRSs are sufficiently synthesized in the previous round, the lack of 20 aaRSs is rescued, and thus DNAP and aaRSs are expressed again. The DNAP further replicates the long repetitive sequence DNA. The reaction mixture is then diluted again with PURE∆aaRSs for the next round of reaction, where DNA replication continues if sufficient amount of active aaRSs are expressed from the repetitive sequence DNA. The reaction mixture was dialyzed from round 2. b) Trajectories of DNA concentrations in the dialyzed serial dilution experiment. DNA concentrations of each aaRS gene before and after the dilution step were measured by qPCR and normalized based on the concentration after Round 1. The raw data before the normalization is shown in Fig. S8a. c) Translational activity of the reaction mixture during the serial dilution experiment conducted in b. Aliquots of the reaction mixture at each round was mixed with the PURE system that lacks all aaRSs and contains linear DNA that contains GFP gene (10 nM) and incubated at 30˚C for 12 h. The increase rate of GFP fluorescent was evaluated as translation activity. The time-course data are shown in Fig. S9. Error bars represents standard deviations of three independent experiments.