

# Reduction in the structural instability of cloned eukaryotic tandem-repeat DNA by low-temperature culturing of host bacteria

WATCHARAPORN THAPANA<sup>1,2</sup>, PENPORN SUJIWATTANARAT<sup>1,2</sup>,  
KORNSORN SRIKULNATH<sup>2</sup>, HIROHISA HIRAI<sup>1</sup> AND AKIHIKO KOGA<sup>1\*</sup>

<sup>1</sup>Primate Research Institute, Kyoto University, Inuyama City 484-8506, Japan

<sup>2</sup>Faculty of Science, Kasetsart University, Bangkok 10900, Thailand

(Received 20 July 2014; revised 24 September 2014)

## Summary

For accurate analyses of eukaryotic tandem-repeat DNA, it is often required to clone a genomic DNA fragment into a bacterial plasmid. It is, however, a serious problem that tandem-repeat DNA is frequently subjected to structural changes during maintenance or amplification in the host bacteria. Here, we show an example of a clear difference in the instability of tandem-repeat DNA between different culturing temperatures. A fragment of monkey centromeric DNA carried by pUC19 was considerably degraded by culturing bacteria at 37 °C, but the damage was reduced at 25 °C. Thus, culturing temperature is a significant factor for avoiding degradation, in addition to the genotype of the host bacteria.

Genomes of higher eukaryotes generally contain large amounts of tandem-repeat DNA, such as centromeric, telomeric and ribosomal DNA sequences. For accurate and fine structural analyses of such DNA, sequence data obtained from shotgun sequencing (even in the case of a hierarchical shotgun strategy) are often inadequate because of the difficulty of assembling contig sequences from sequence reads. This is considered a serious problem, especially for repeat DNA with high sequence homogeneity among repeat units. As a reflection of this problem, great portions of large-scale tandem-repeat regions are often put aside from the genome browsers built by genome sequencing projects.

When cloning of genomic DNA fragments (in a vector such as a plasmid, fosmid, bacterial artificial chromosome or yeast artificial chromosome) is included in the analysis of repetitive DNA, another factor makes it difficult to obtain accurate nucleotide sequences, that is, the low reliability of sequence reads due to the possible rearrangements occurring during the maintenance or amplification in cells of host bacteria or yeast. Degradation of cloned tandem-repeat DNA is a widely known phenomenon with a

number of illustrations in the literature (Brutlag *et al.*, 1977; Neil *et al.*, 1990; Song *et al.*, 2001). In the present study, we examined the effect of the culturing temperature of the host bacteria on the degree of degradation of a repetitive DNA fragment cloned into a bacterial plasmid.

Alpha satellite DNA is a major DNA component of primate centromeres (Maio, 1971; Willard, 1991). In our previous study (Prakhongcheep *et al.*, 2013 *b*), we found that the centromeres of Azara's owl monkey (*Aotus asazre*) carry two types of alpha satellite DNA that we designated as OwlAlp1 and OwlAlp2. These are tandem-repeat sequences consisting of 185-bp and 344-bp repeat units, respectively. Comparison of their sequences showed that OwlAlp2 is an alpha satellite DNA of the original type and that OwlAlp1 was derived from OwlAlp2 in the owl monkey lineage and subsequently expanded in the genome. For further study of their evolutionary history as well as their functions in the centromere, we initiated experiments to accurately determine their nucleotide sequences from cloned DNA fragments of >10 kb in length. We used a previously described method (Prakhongcheep *et al.*, 2013 *a*; Koga *et al.*, 2014). Briefly, the method included transfer of the genomic DNA fragment from pCC1FOS (a fosmid vector used for library construction) to pUC19 (a plasmid vector that carries the restriction enzyme recognition sites necessary for the next step), preparation

\* Corresponding author: Tel: +81 568 63 0567. Fax: +81 568 63 0085. E-mail: koga.akihiko.5n@kyoto-u.ac.jp

of a series of deletion clones by using an exonuclease, sequencing of the deletion clones by using a universal primer and conversion of the sequence reads into a contig sequence. These processes ran smoothly with OwlAlp1, but we faced a problem in the initial transfer step for OwlAlp2. The problem was that virtually all recombinant products carried fragments shorter than the original genomic DNA fragment. In these treatments, the bacterial culture was incubated at 37 °C, a widely used incubation temperature.

To solve this issue, we established a method to detect which part of the entire process was causing the problem. The OwlAlp2 fragment was transferred to pUC19 (Fig. 1), and variation in the size of recombinant plasmids was examined at three different points (Fig. 2). Because the plasmid copy number per bacterial cell was considered a significant factor, we included two different temperatures (37 °C and 25 °C) in this test. pUC19 carries a mutant ColE1 replicon (Yanisch-Perron *et al.*, 1985) leading to several hundred copies per cell when incubated at 37 °C and smaller copy numbers when incubated at <30 °C (Twigg & Sherratt, 1980). After the introduction of recombinant plasmids into the bacteria, the culture was incubated at 25 °C until colonies were formed on ampicillin-containing agar plates (48 hours). A single colony was picked up and suspended in liquid medium, and the suspension was spread on agar plates. One set of plates was incubated at 25 °C, and another was incubated at 37 °C. After colony formation (48 hours at 25 °C, 14 hours at 37 °C), single colonies were picked up and plasmid sizes were examined (B01–24 and C01–24 for 25 °C and 37 °C, respectively). The bacterial host used was DH5 $\alpha$ . This is a commonly used strain of the *recA1* genotype that is effective in suppressing recombination (Taylor *et al.*, 1993).

Figure 3 shows that there was no detectable size variation among the colonies A01–12. Colonies B01–24 originated from colony A01, and their plasmid size was uniform (and the same as that of A01; 11 kb) except for a case of colony B19. B19 exhibited an extra faint band of a smaller size (7 kb), indicating that the final culture contained a mixture of two plasmids of different sizes. The 7-kb-carrying plasmid is thought to have been derived from an 11-kb-carrying plasmid during incubation at 25 °C. C01–24 showed a higher size variation. Different sizes were observed among many colonies, and there were colonies exhibiting three or more different sizes (C03, 14, 17, 20, 22, 23 and 24). Although fragments identical to that of A01 (11 kb) were observed in eight colonies (C01, 03, 09, 16, 17, 20, 22 and 23), the band coexisted with bands of other sizes, mostly as a minor band.

In all images of gel electrophoresis, the sizes of variant fragments were smaller than that of the original fragment. This does not necessarily imply that a plasmid of a larger size was not generated. The size

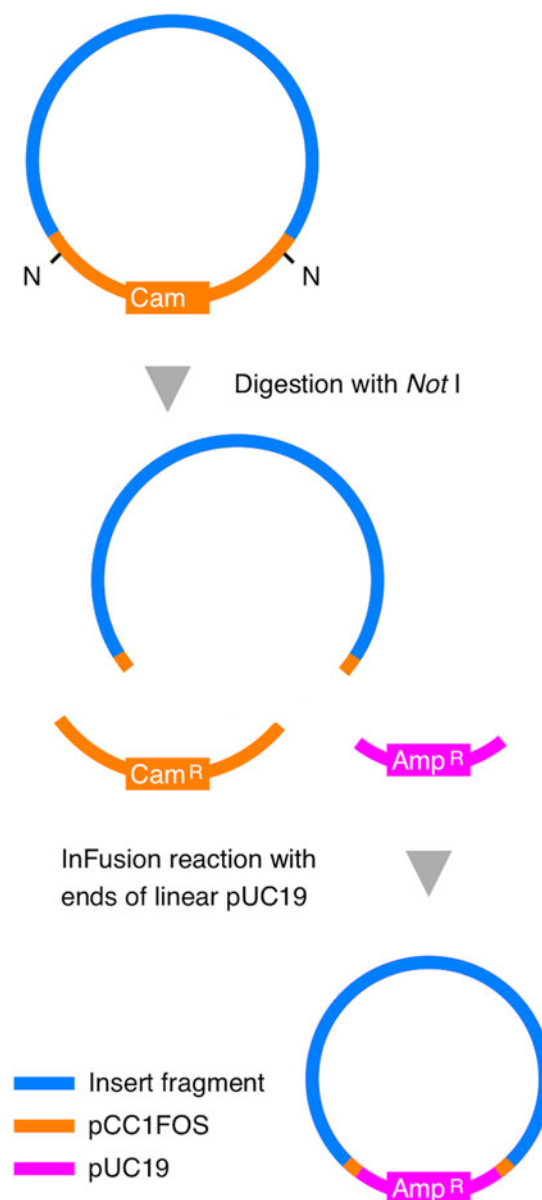


Fig. 1. Scheme for the transfer of an insert fragment from the pCC1FOS fosmid vector to the pUC19 plasmid vector. pCC1FOS has two *NotI* recognition sites located outside its cloning site. Fosmid DNA was digested with *NotI* and mixed with a linear fragment carrying the pUC19 sequence. The end regions of the fragment had been modified to carry 15-bp blocks corresponding to the end regions of the insert fragment. The DNA mixture was subjected to the InFusion reaction that caused recombination in the corresponding regions between pUC19 and the insert fragment. The reaction mixture was introduced into bacterial cells and the colonies formed on ampicillin-containing agar plates were selected randomly and used in subsequent experiments.

distribution observed here can be considered to have been affected by the efficiency of plasmid amplification; shorter plasmids would tend to have higher amplification efficiency.

As shown in the above results, the OwlAlp2 fragment was degraded during maintenance in host bacteria to a

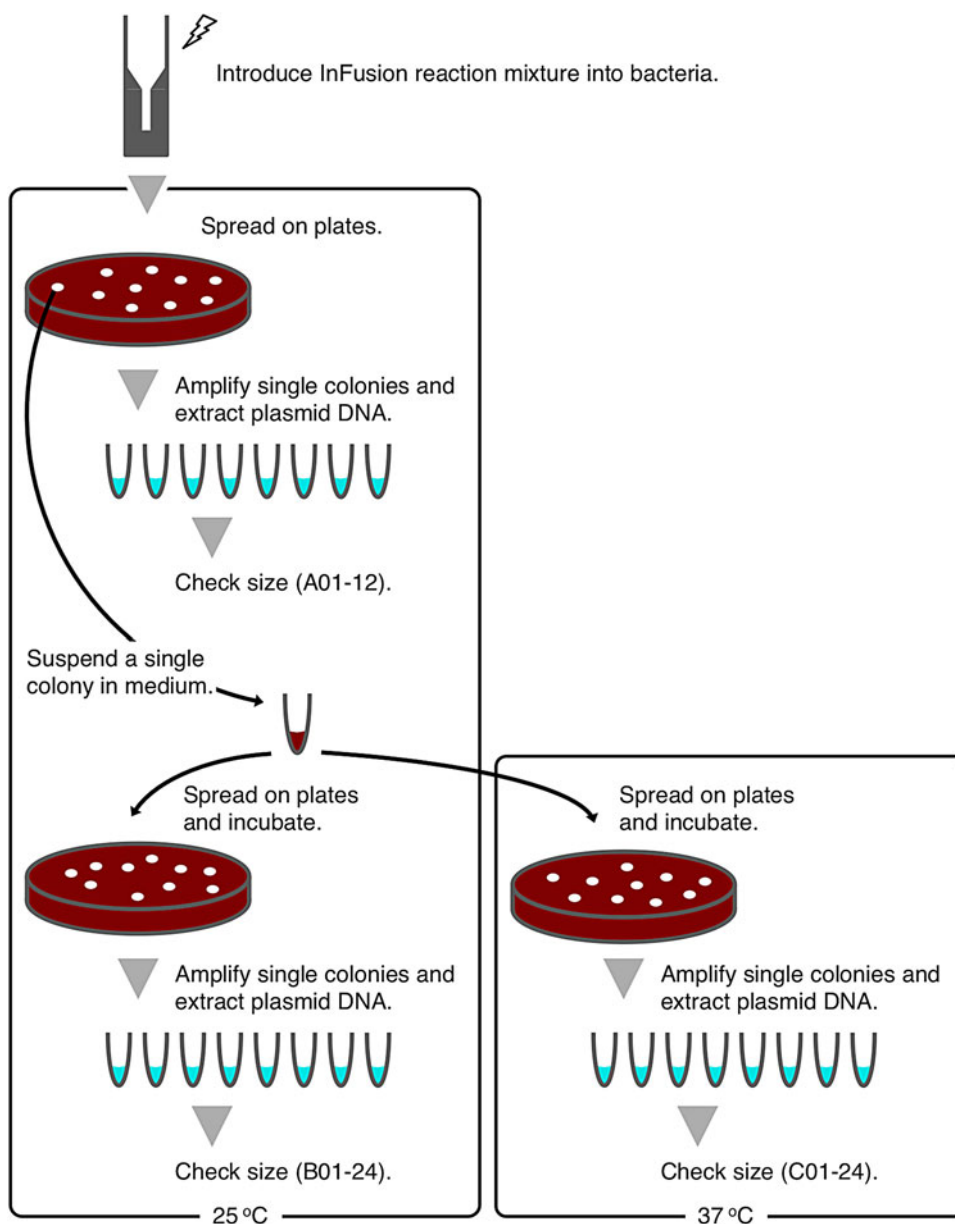


Fig. 2. Assay for the fragility of plasmid clones. Incubation of bacterial cells was conducted at 25 °C during the initial steps. Incubation at 25 °C or 37 °C was performed after picking up a single colony.

small extent at 25 °C and to a larger extent at 37 °C. The differences could be attributed to the copy number per cell; the higher chance of recombination between plasmid molecules at higher densities. It is, however, also possible that the difference was caused by other factors such as difference in the activity of enzymes involved in the recombination reaction. In addition to these assays using the DH5a strain, we conducted a rough assay corresponding to the C colonies, using the SURE strain whose genotype is *recB*, *recJ*, *uvrC*, *umuC* and *sbcC*. The combination of these mutations is known to be effective in suppressing a wide range of recombination and repair events (Doherty *et al.*, 1993). The results obtained were, however, essentially the same as those observed with the C colonies.

Bacterial artificial chromosomes and fosmids are used as cloning vectors for long DNA fragments (>40 kb), and plasmids are used mostly for short fragments (<15 kb). Even if the original genomic DNA fragment is cloned in a bacterial artificial chromosome or fosmid vector, the isolation of part of the insert fragment by subcloning it into a plasmid vector is often required. In such a case, a single-copy or low-copy-number plasmid, such as pBR322, pSC101 or pA15, is preferable. However, many plasmids that have been engineered to carry useful signals or cloning sites are high-copy-number plasmids. For example, pUC19, which we used in the present study, is one such highly developed, high-copy-number plasmid. Our results showed that culturing bacteria at a low

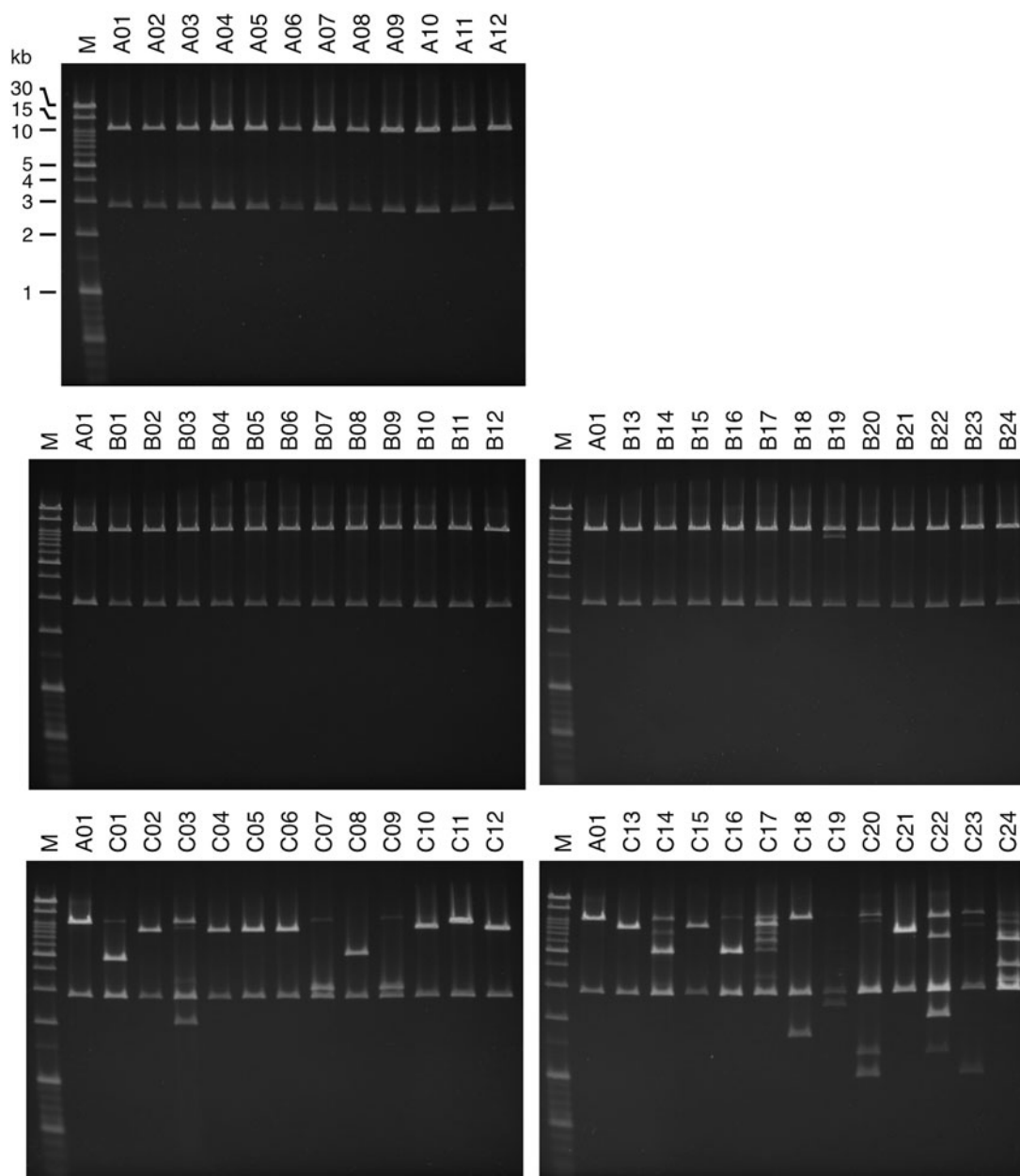


Fig. 3. Results of agarose gel electrophoreses for size variation among plasmid clones. pCC1FOS carried two *Hind*III recognition sites just adjacent to its cloning site, and the OwlAlp2 fragment did not have sites for this enzyme. The two sites were transferred to pUC19 together with the entire OwlAlp2 fragment. Because the pUC19 portion did not contain a *Hind*III site, digestion of pUC19 clones with *Hind*III produced the insert fragment (originally 11 kb) and a vector fragment (2.9 kb) that consisted of the pUC19 backbone and short fragments originating from pCC1FOS. “M” stands for size marker. The sizes of the marker fragments are shown along the left margin of the top picture.

temperature (25 °C) is effective in reducing the degradation frequency even if pUC19, and possibly its derivatives, is used as a vector.

In addition to the effect of the culturing temperature, our results show the possible involvement of higher order DNA structure in determining degradation frequency. In the photographs of C01–24 (Fig. 3), the size distribution among the clones is not likely to be random: eight clones (C02, 04, 05, 06, 10, 12, 13 and 15) exhibited bands of approximately 8 kb and four clones (C08, 14, 16

and 17) showed bands of approximately 5 kb. Bands of other sizes were unique to a single clone or common to two clones at most. The differences among those commonly appearing bands, including the size of their original fragment (11 kb; A01), are multiples of 3 kb. In our extension study, we obtained a 10.8-kb contig sequence of OwlAlp2 (GenBank accession number LC002884) taking care to avoid degradation, and found a higher order repeat structure in which nine basic repeat units composed a larger repeat unit (Koga

*et al.*, unpublished observations). The length of the larger repeat units was 3073 bp. We did not encounter any problems in subcloning the OwlAlp1 sequence, as described earlier in the present report, and we have not found a higher order repeat structure in the OwlAlp1 contig sequence obtained so far. These results suggest that the higher order repeat structure is involved in the occurrence of structural rearrangements and the provision of rearrangement hotspots.

We are grateful to Takehiko Kobayashi and Toshio Mouri for helpful discussions. This work was supported by Grants-in-Aid (25650152 to AK, 23470098 to AK and 24255009 to HH) from the Japan Society for the Promotion of Science (JSPS). PS was supported by the Royal Golden Jubilee PhD Program of the Thai Government.

#### Declaration of interest

None.

#### References

- Brutlag, D., Fry, K., Nelson, T. & Hung, P. (1977). Synthesis of hybrid bacterial plasmids containing highly repeated satellite DNA. *Cell* **10**, 509–519.
- Doherty, J. P., Lindeman, R., Trent, R. J., Graham, M. W. & Woodcock, D. M. (1993). *Escherichia coli* host strains SURE and SRB fail to preserve a palindrome cloned in lambda phage: improved alternate host strains. *Gene* **124**, 29–35.
- Koga, A., Hirai, Y., Terada, S., Jahan, I., Baicharoen, S., Arsaithamkul, V. & Hirai, H. (2014). Evolutionary origin of higher-order repeat structure in alpha-satellite DNA of primate centromeres. *DNA Research* **21**, 407–415.
- Maio, J. J. (1971). DNA strand reassociation and polyribonucleotide binding in the African green monkey, *Cercopithecus aethiops*. *Journal of Molecular Biology* **56**, 579–595.
- Neil, D. L., Villasante, A., Fisher, R. B., Vetric, D., Cox, B. & Tyler-Smith, C. (1990). Structural instability of human tandemly repeated DNA sequences cloned in yeast artificial chromosome vectors. *Nucleic Acids Research* **18**, 1421–1428.
- Prakhongcheep, O., Chaiprasertsri, N., Terada, S., Hirai, Y., Srikulnath, K., Hirai, H. & Koga, A. (2013 a). Heterochromatin blocks constituting the entire short arms of acrocentric chromosomes of Azara's owl monkey: formation processes inferred from chromosomal locations. *DNA Research* **20**, 461–470.
- Prakhongcheep, O., Hirai, Y., Hara, T., Srikulnath, K., Hirai, H. & Koga, A. (2013 b). Two types of alpha satellite DNA in distinct chromosomal locations in Azara's owl monkey. *DNA Research* **20**, 235–240.
- Song, J., Dong, F., Lilly, J. W., Stupar, R. M. & Jiang, J. (2001). Instability of bacterial artificial chromosome (BAC) clones containing tandemly repeated DNA sequences. *Genome* **44**, 463–469.
- Taylor, R. G., Walker, D. C. & McInnes, R. R. (1993). *E. coli* host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic Acids Research* **21**, 1677–1678.
- Twigg, A. J. & Sherratt, D. (1980). Trans-complementable copy-number mutants of plasmid ColE1. *Nature* **283**, 216–218.
- Willard, H. F. (1991). Evolution of alpha satellite. *Current Opinion in Genetics & Development* **1**, 509–514.
- Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119.