DIVISION OF GENE EXPRESSION AND REGULATION II

Professor:	*Takashi Horiuchi*
Research Associate:	*Masumi Hidaka*
	Takehiko Kobayashi
	Ken-ichi Kodama
	**Katsuki Johzuka*
Technical Staff:	*Kohji Hayashi*
	***Yasushi Takeuchi*
Graduate Student:	****Hiroko Urawa*
	Naomi Serizawa

* *to Department of Biological Chemistry, University of California-Irvine, Irvine, California (Feb. 1. 1999-)*
** *to Division of Molecular Neurobiology in NIBB (May. 1 2000-)*
*** *She obtained her doctorate Sept. 2000.*

Homologous recombination, which may occur in all organisms, involves genetic exchange between two parent-derived homologous chromosomes in meiosis in addition to the repair of DNA damage induced by physical and chemical agents. As a result of our analyses of recombinational hotspots of *E. coli* and *S. cerevisiae*, in particular the activity related to DNA replication fork-blocking events, the physiological function of homologous recombination (especially in normally growing cells) is better understood. In 2000, work on the following three subjects has advanced our knowledge of the dynamics and function of the genome.

I. Identification of *cis*-essential sequences for amplification of the ribosomal RNA gene repeats in yeast.

We previously found that some recombinational hotspots in *E. coli* present in regions where progress of a DNA replication fork is severely impeded. Detailed analysis of this phenomenon revealed that when the fork is blocked, recombinational enzymes construct a new replication fork after breakdown of the arrested fork by introducing a double-stranded break at either of the sister chromatids. Consequently, this recombination reaction for renewal of the replication fork is a primary source of recombinational hotspots. Next, we showed that this phenomenon occurred not only in prokaryotes but also in eucaryotes. In most eucaryotes, the DNA replication fork blocking site, called *RFB* (replication fork barrier), is located in each unit of ribosomal multigenes (rDNA). It has been reported that in budding yeast (*S. cerevisiae*) when a non-transcriptional region (NTS) of the rDNA, in which the *RFB* site is located, is inserted into a non-rDNA region of chromosomes, recombination of a region close to the insertion site is enhanced. This enhancement is called *HOTl* activity and the DNA fragment required for activation is named *HOTI* DNA. We isolated a number of mutants, which are defective in *HOTl* activity, and found that some of them had a concomitant defect in fork blocking activity at the *RFB* site. Both of these defective phenotypes were restored to wild type by introducing a single gene, named *FOB1*, thereby suggesting that fork blockage acts as a trigger for recombination in yeast, as well as in bacteria. In addition to this finding, the detection of *RFB* activity within the rDNA cluster in various higher eucaryotic cells suggests that the fork renewal cycle occurring after the blockage is more general than expected.

The copy number of rDNA repeats is unstable. Although variation in the number of rDNA copies has been observed in various eucaryotes, the underlying mechanism remains unknown. Because fork arrest at the *RFB* site was a trigger for recombination, as mentioned above, we suspected that a fork-blocking event was involved in changing the copy number of rDNA repeats. Neither an increase nor a decrease in rDNA repeats occurred in *fob1* mutants, suggesting that the fork blocking event is required for changing the copy number, probably through a recombination process. As the fork blocking system has only been understood to prevent replication from "colliding" with rDNA transcription so far, this study indicates that the fork block system may have an entirely new function. If this is the case, *RFB* deleted rDNAs should not amplify.

This year, we tested this prediction and also attempted to identify *cis*-sequences, required for amplification of the copy number of rDNA. However, this kind of analysis was not feasible for a repeated gene with high copy number. Therefore, we constructed a strain with a rDNA copy number reduced to two copies of the original genomic locus by using a multi-copy plasmid, carrying a hygromycin B resistant copy of rDNA. Then, each sub-region in a single NTS1 region among the two rDNAs on the genome were replaced with a *URA3* marker gene in order to determine which sequences would be required for rDNA amplification, especially around the *RFB* site. Mutational analysis using this system showed that not only the *RFB* site but also the adjacent ~440 bp region in NTS1 (together called the EXP region) are required for the *FOB1*-dependent repeat expansion. This ~400 bp DNA element is not required for the fork arrest activity or the *HOT1* activity and therefore defines a function unique to rDNA repeat expansion (and presumably contraction) separate from *HOT1* and RFB activity. Our model, which indicates a mechanism by which the copy number of rDNA increases from two to 150, is shown in the accompanying Figure.

II. Enhanced homologous recombination caused by a non-transcribed spacer of the ribosomal RNA gene in *Arabidopsis*.

In yeast, a single rDNA unit structure consists of two rRNA genes, 35S pre-rRNA and 5S rRNA genes, and two non-transcribed spacer (NTS) regions, NTS1 and NTS2. NTS1 is flanked by the 3'-ends of the two rRNA genes and NTS2 is flanked by the 5'-ends of the two rRNA genes. A region containing NTS1-5SrRNA gene-NTS2, named *HOT1* DNA, has activity, that

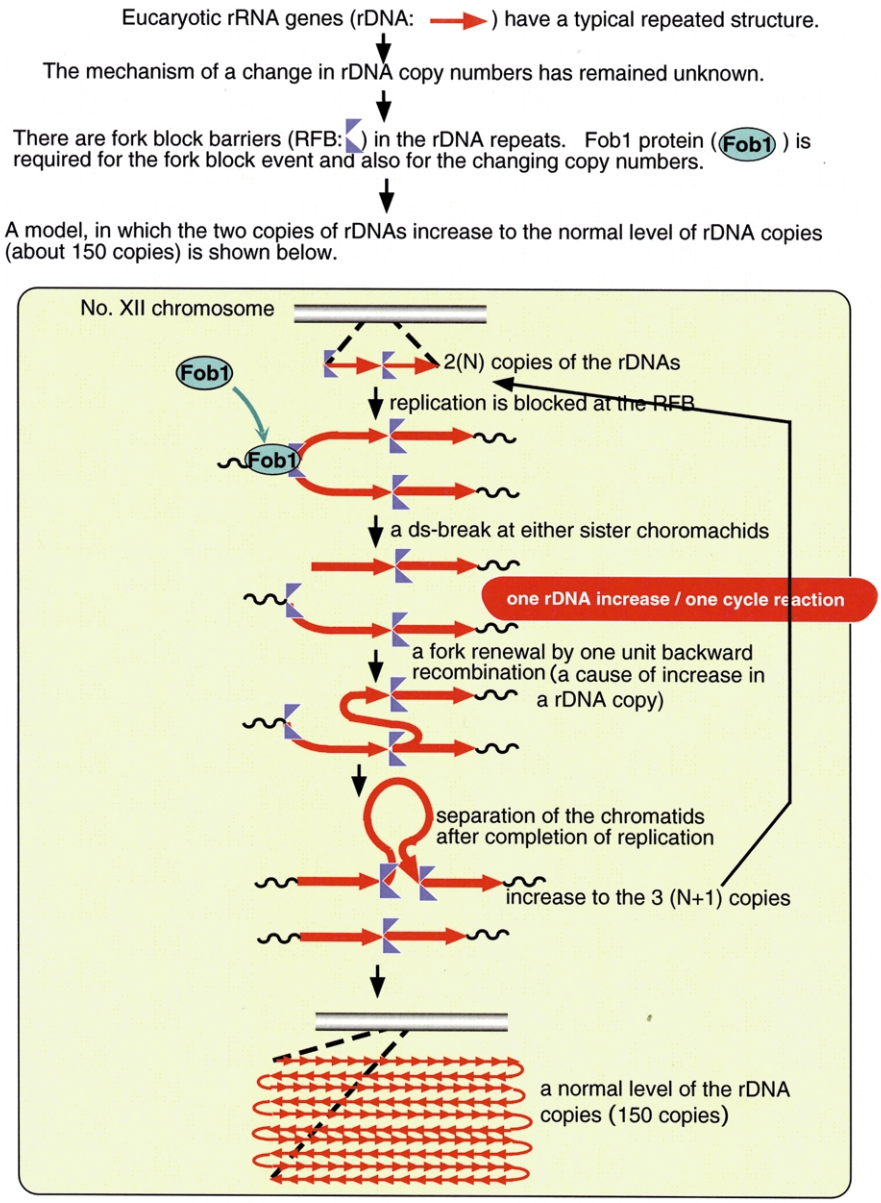

Figure. An increase mechanism for rDNA copy number. The figure shows that one reaction cycle increases the rDNA by one copy. Therefore, 150 reaction cycles can increase the copy number of the rDNA from two to 150.

stimulates recombination at a region close to a non-rDNA locus where the *HOT1* DNA is inserted. The stimulated recombination activity is called *HOT1* activity. Further analysis revealed that only two small non-contiguous regions, called E and I, are required for *HOT1* activity. E and I are located in NTS1 and NTS2, respectively and correspond to enhancer and initiator of RNA polymerase I dependent 35S rRNA transcription. The rDNA repeated structure is essentially the same in various eucaryotes as in yeast, except in other eucaryotes 5S rRNA genes are located at a different locus from the 35S rRNA gene cluster. In addition, fork-blocking activity has been detected in the NTS regions of the rDNA array in all eucaryotes tested so far. Here, we investigated whether the non-transcribed spacer region in the rDNA cluster of higher eucaryotes has hotspot activity similar to that observed in yeast. We chose *Arabidopsis* as an experimental model, because a reliable recombination assay system had been developed and recombination event, if it occurs anywhere-whole body of the plant, is easily visible. To analyze the effect of the NTS on mitotic homologous recombination in the plant, we constructed transgenic lines of *Arabidopsis* containing the NTS region and a recombination substrate, in which two 3'- and 5'- deleted *uidA* genes, with partially overlapping sequences are separated by a Hyg^r gene. Histochemical GUS-staining monitored reconstitution of functional *uidA* genes by homologous recombination. We found that recombination occurred more frequently in all organs tested in F

(Fork Block) lines transgenic for NTS than in C (Control) lines without NTS. The average number of GUS^+ spots on leaves in F lines was more than nine-fold higher than in C lines. Furthermore, by genomic Southern blotting analysis, post-recombinational molecules were detected in a transgenic line, F43, which had an extremely high number of GUS^+ blue spots. These results strongly suggests that NTS-dependent enhancement of homologous recombination could be a common feature in higher plant as well as yeast.

III. *E. coli* genomic structure and function

Analysis of whole genomic sequence of *E. coli* had been completed in 1997. While a strain Japan team sequenced is W3110, the other US team did is MG1655. These strains were separated from a common ancestor during or right after World War II. In order to understand the micro-evolution of these strains, we are sequencing a previously undetermined region of the W3110 genome and will finished this spring. Our goal is to compare this genome with the MG1655 genomes at the nucleotide level.

In order to identify a minimal set of genes required for the duplication of a single cell and to elucidate the function of those genes, we proposed a new project and a established a new collaboration team, headed by Mori (NAIST). Since last year our project has been supported by CREST from JST and is being carried out by the following four groups: (1) resources, (2) informatics, (3) database, and (4) functional analysis. At present, the resources group (1) has cloned all ORF candidates, constructed *E. coli* micro-arrays using the cloned ORFs and started to analyze gene expression under various conditions, obtaining results that will be published. Other plans, such as the construction of a set of mutant strains, in which each gene is disrupted, are in progress. These strains will be distributed to workers world wide, stimulating functional analysis of the *E. coli* genome, and eventually contributing to defining an integrated functional network of genes in *E. coli.*

Publication list:

Mori, H., Isono, K., Horiuchi, T., and Miki, T. (2000) Functional genomics of *Escherichia coli* in Japan. *Res. Microbiol.* 151, 121-128.

Kobayashi, T., Nomura, M., and Horiuchi, T. (2001) Identification of DNA cis-elements essential for expansion of ribosomal DNA repeats in *Saccharomyces cerevisiae. Mol. Cell. Biol.* 21, 136-147.