

Title	Stepwise evolution during ribosomal RNA gene cluster formation in Archaea
Sub Title	
Author	小林, 朝紀(Kobayashi, Asaki)
Publisher	慶應義塾大学湘南藤沢学会
Publication year	2014
Jtitle	生命と情報 No.21 (2014. ) ,p.133- 136
JaLC DOI	
Abstract	<p>Functionally related genes are often located closely to each other in the genome. Such characteristics could be derived through evolutionary processes. As a model case, I focused on the evolution of ribosomal RNA (rRNA) gene clusters in Archaea. Since a comprehensive analysis of the evolution of rRNA gene cluster formations has not been performed, the knowledge of them is still limited. Here, to understand the evolutionary changes of rRNA gene clusters in Archaea, a computational analysis of their patterns, localisation, and organisation on the complete archaeal genomes was performed.</p> <p>First, rRNA gene clusters were comprehensively analysed by characterising the distribution of rRNA genes and the distance between their adjacent genes in the archaeal genomes. The results showed that rRNA gene clusters mainly consist of 16S-23S in Crenarchaeota, whereas they mainly consist of 16S-23S-5S in Euryarchaeota. Notably, several rRNA gene clusters contained tRNA genes (e.g., 16S-tRNA-23S-5S) in Euryarchaeota. Moreover, the results suggest that, after the acquisition of the 16S-tRNA-23S-5S rRNA gene clusters in Euryarchaeota, either the whole cluster or only the 5S rRNA in the cluster was duplicated. Finally, a possible model explaining the evolutionary pathway for rRNA gene cluster formation in Archaea is proposed. This model provides valuable insights on the evolutionary process shaping archaeal genome organisation during early stages of life.</p>
Notes	慶應義塾大学湘南藤沢キャンパス先端生命科学研究会 2014年度学生論文集 卒業論文ダイジェスト
Genre	Technical Report
URL	<a href="https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO92001004-00000021-0133">https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO92001004-00000021-0133</a>

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the Keio Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

# **Stepwise evolution during ribosomal RNA gene cluster formation in Archaea**

Faculty of Environment and Information Studies

B4 Asaki Kobayashi

## **Abstract**

Functionally related genes are often located closely to each other in the genome. Such characteristics could be derived through evolutionary processes. As a model case, I focused on the evolution of ribosomal RNA (rRNA) gene clusters in Archaea. Since a comprehensive analysis of the evolution of rRNA gene cluster formations has not been performed, the knowledge of them is still limited. Here, to understand the evolutionary changes of rRNA gene clusters in Archaea, a computational analysis of their patterns, localisation, and organisation on the complete archaeal genomes was performed.

First, rRNA gene clusters were comprehensively analysed by characterising the distribution of rRNA genes and the distance between their adjacent genes in the archaeal genomes. The results showed that rRNA gene clusters mainly consist of 16S-23S in Crenarchaeota, whereas they mainly consist of 16S-23S-5S in Euryarchaeota. Notably, several rRNA gene clusters contained tRNA genes (e.g., 16S-tRNA-23S-5S) in Euryarchaeota. Moreover, the results suggest that, after the acquisition of the 16S-tRNA-23S-5S rRNA gene clusters in Euryarchaeota, either the whole cluster or only the 5S rRNA in the cluster was duplicated. Finally, a possible model explaining the evolutionary pathway for rRNA gene cluster formation in Archaea is proposed. This model provides valuable insights on the evolutionary process shaping archaeal genome organisation during early stages of life.

**Keywords:** Archaea, ribosomal RNA, gene cluster, evolution

## Introduction

Life is classified into three domains: Eukaryota, Bacteria and Archaea. Among the three phyla, Archaea is known to be formed of unicellular microorganisms, whose sizes and shapes are morphologically similar to prokaryotes (Bacteria). However, Archaea is distinguished from Bacteria due to their ether-membrane that differs from the bacterial ester-membrane. The three domains of life, represented on a universal tree, were proposed by Carl Woese *et al.*, based on the analysis of ribosomal RNA sequences and core protein translational genes (Woese *et al.*, 1977) (Figure 1). Since its discovery, research on Archaea has rapidly advanced, and is now considered as an ideal model organism to understand the evolution and origin of life as many archaeal species survive and thrive in extreme environments, such as hyperthermia and anoxia, which are similar to the environment of the primordial Earth.

Archaeal species possess circular genomes, operon structure, ribosomal binding region, and other characteristic features and structures that are similar to Bacteria, but their gene organisation, e.g., the presence of introns in the rRNA genes (Yip *et al.*, 2013), is similar to Eukarya, as well as the fact they possess RNA polymerase II-like polymerase (Rawlands *et al.*, 1994, Langer *et al.*, 1995, Kosa *et al.*, 1997, Hirata *et al.*, 2009). A recent hypothesis (Eocyte-hypothesis) in contrast with the three-domain hypothesis, also proposed only two primary domains, as proposed by “archaeal-host hypothesis for eukaryotic origins” (Embley and Martin, 2006, Williams *et al.*, 2013). Therefore, Archaea is a key to reveal the molecular evolution of Eukaryotes.

Ribosomal RNAs (rRNA) catalyse the protein synthesis as they are major components of ribosomes, that are one of the most ancient and most complex molecular machineries (Moss *et al.*, 2007). The amount of rRNA molecules occupy over 80% of the total cellular RNA. In Archaea, there are four types of rRNA genes (23S, 5S, 5.8S and 7S rRNA genes) that are packaged with ribosomal proteins to form the large subunits of ribosomes, whereas 16S rRNA genes form the small subunits (Santoro, 2005). The rRNA genes are essential for biological activities, especially 16S rRNA genes, and are highly conserved in all living organisms, with almost no mutations. Nevertheless, 16S rRNA genes possess a sequence whose regions are relatively easy to mutate. Therefore analysing its sequence allows understanding the relationships between distant relative species. 23S and 5S rRNA gene sequences have also been used for this purpose, as well as the 16S-23S intron sequences in Bacteria (Maguire and Zimmermann, 2001).

Previous studies have brought light to many aspects of the rRNA transcriptional machineries in Archaea. Several rRNA transcription units have been reported in different archaeal species, for example in Methanogens (e.g., *Methanobacterium thermoautotrophicum*), extreme halophiles (e.g., *Halobacterium halobium*), and extreme thermophiles (e.g., *Sulfolobus solfataricus*). For instance, 16S-23S genes clusters have been observed in *Desulfurococcus mobilis* and reported as “crenarchaeal-specific clusters”, while 16S-23S-5S clusters and 16S-tRNA<sup>Ala</sup>-23S-5S clusters have been shown in *Methanobacterium vanniellii*, and *Methanobacterium thermoautotrophicum* and therefore reported as “euryarchaeal-specific clusters” for (Garrett *et al.*, 1991).

However, how those clusters have formed through evolutionary processes is still not understood. The purpose of this study is therefore to comprehensively determine rRNA gene clusters in archaeal genomes, and to reveal the evolutionary process through which rRNA gene clusters are formed. To do so, rRNA gene information was collected from GenBank database using RefSeq annotations (Table 1), and distances between rRNA genes in archaeal genomes were calculated. According to the distribution of rRNA genes, rRNA gene clusters were predicted on 144 archaeal genomes, and were then evolutionary analysed by mapping along the phylogenetic tree.

Finally, a possible model explaining the evolutionary pathway for rRNA gene cluster formation in Archaea is proposed.

Phylum	Species	Number of rRNA genes					Total
		16S	23S	5S	7S	5.8S	
Euryarchaeota	99	184	184	231	11	-	610
Crenarchaeota	37	37	37	38	3	1	116
Thaumarchaeota	6	6	6	5	-	1	18
Korarchaeota	1	1	1	1	-	-	3
Nanoarchaeota	1	1	1	1	-	-	3
Total	144	229	229	276	14	2	750

Table 1. Number of rRNA genes.

## Results and Discussion

論文投稿前のため非公開.

## Conclusions

rRNA genes clusters have been shown in archaeal species, however, the mechanisms through which these clusters organise through evolutionary process are still not yet understood. Here, through the bioinformatics analysis of 144 archaeal genomes collected from public GenBank databases, a comprehensive determination of rRNA gene clusters was performed. The results were mapped along the archaeal phylogenetic tree, providing evidence on the evolutionary processes that shaped the formation of rRNA genes clusters: 16S-23S clusters are the first to be formed. Next, 5S and tRNA genes are included in the clusters, especially forming 16S-tRNA-23S-5S clusters. However, early-diverged Euryarchaeota still show 16S-23S clusters with duplicated 5S rRNA genes, while several other several Euryarchaeota species showed both 16S-23S based and 16S-23S-5S based clusters, indicating two possible evolutionary pathways after the acquisition of the 16S-tRNA-23S-5S clusters: either the whole cluster containing 5S rRNA gene, or only the 5S rRNA in the cluster was duplicated. Interestingly, recently diverged Euryarchaeota showed rRNA genes that were scattered in the genomes, suggesting reversed evolutionary changes of rRNA gene cluster formation. Another interesting feature of rRNA genes clusters is that the inclusion of tRNA genes. In particular, Euryarchaeota showed significant enrichment of tRNA genes inclusion in rRNA gene clusters, e.g., between 16S and 23S rRNA genes for tRNA<sup>Ala</sup>, between 5S rRNA genes for tRNA<sup>Asp</sup> and tRNA<sup>Lys</sup>. In addition, the type of tRNA anticodon included in rRNA gene clusters was found to be unique, e.g., TGC for tRNA<sup>Ala</sup>, TTT for tRNA<sup>Lys</sup>, GCA for tRNA<sup>Cys</sup>, and GTC for tRNA<sup>Asp</sup>, indicating a specific evolutionary selection the tRNAs during cluster inclusion.

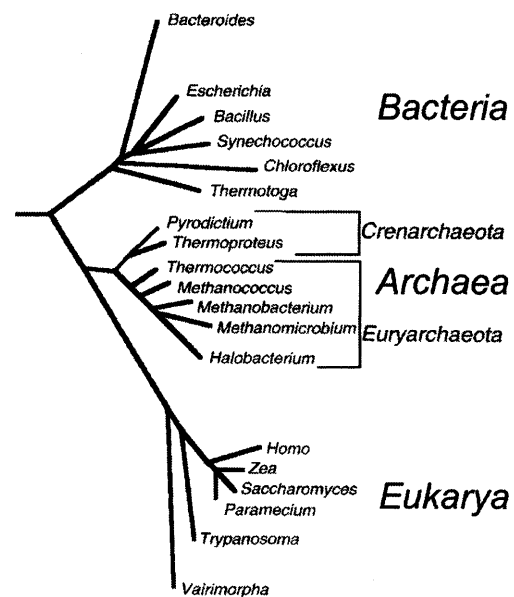


Figure 1. Universal Phylogenetic Tree. Adapted from Olsen and Woese, 1997

In summary, the model explaining the evolutionary pathway for rRNA gene cluster formation in Archaea proposed in this study may provides valuable knowledge on evolution of species during early stages of life. The analysis conducted here may also serve future researches to understand the acquisition of specific tRNA gene in rRNA gene clusters.

## Materials and Methods

論文投稿前のため非公開.

## Acknowledgements

I would like to thank Professor Akio Kanai for his great support of my research, and for giving me opportunities to present in conferences in Japan and Québec. I would like to thank all RNA analysis group members who have supported me during these years, especially my adviser Dr. Kiyofumi Hamashima for meaningful advice and kind support. I could make a lot of precious memories with Biocampers and my contemporaries. Also I would like to thank Dr. Vincent Piras and Dr. Cornelia Amariei for their kind support. Finally, I would like to thank Professor Masaru Tomita for providing a stimulating environment to do my research in IAB.

## References

- Embley TM, Martin W. Eukaryotic evolution, changes and challenges. *Nature* 440:623-30 (2006).
- Garrett RA, Dalgaard J, Larsen N, Kjems J, Mankin AS. Archaeal rRNA operons. *Trends Biochem Sci* 16:22-26 (1991).
- Hirata A, Murakami KS. Archaeal RNA polymerase *Curr Opin Struct Biol* 19:724-731 (2009).
- Kosa PF, Ghosh G, DeDecker BS, Sigler PB. The 2.1-Å crystal structure of an archaeal preinitiation complex: TATA-box-binding protein/transcriptionfactor (II)B core/TATA-box. *Proc Natl Acad Sci U S A* 94:6042-6047 (1997).
- Langer D, Hain J, Thuriaux P, Zillig W. Transcription in Archaea- Similarity to that in Eucarya. *Proc Natl Acad Sci U S A* 92:5768-5772 (1995).
- Maguire BA, Zimmermann RA. The ribosome in focus. *Cell*. 104:813-6 (2001).
- Moss T, Langlois F, Gagnon-Kugler T, Stefanovsky V. A housekeeper with power of attorney: the rRNA genes in ribosome biogenesis. *Cell Mol Life Sci* 64:29-49 (2007).
- Rowlands T, Baumann P, Jackson SP. The TATA-binding protein: a general transcription factor in eukaryotes and archaeobacteria. *Science* 264:1326-1329 (1994).
- Santoro R. The silence of the ribosomal RNA genes. *Cell Mol Life Sci* 62:2067-2079 (2005).
- Williams TA, Foster PG, Cox CJ, Embley TM. An archaeal origin of eukaryotes supports only two primary domains of life. *Nature* 504:231-236 (2013).
- Woese CR, Fox GE. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A* 74:5088-5090 (1977).
- Yip WS, Vincent NG, Baserga SJ. Ribonucleoproteins in archaeal pre-rRNA processing and modification. *Archaea* 2013:614735 (2013).