Simplification of the Genetic Code: Restricted Diversity of Genetically Encoded Amino Acids. K Amikura[1,2], A Kawahara-Kobayashi[2], and D. Kiga[1,2], [1]Department of computational intelligence and systems science, Interdisciplinary graduate school of science and engineering, Tokyo Institute of Technology, [2]Earth-Life Science Institute, Tokyo Institute of Technology. ffiliation (4259 Nagatsuta-cho, Midori-ku, YOKOHAMA, Japan, 226-8503. kiga@dis.titech.ac.jp).

At earlier stages in the evolution of the universal genetic code, fewer than 20 amino acids were considered to be used. Although this notion is supported by a wide range of data, the actual existence and function of the genetic codes with a limited set of canonical amino acids have not been addressed experimentally, in contrast to the successful development of the expanded codes. Recently, we constructed artificial genetic codes involving a reduced alphabet [1]. In one of the codes, a tRNA(Ala) variant with the Trp anticodon reassigns alanine to an unassigned UGG codon in the Escherichia coli S30 cell-free translation system lacking tryptophan. We confirmed that the efficiency and accuracy of protein synthesis by this Trp-lacking code were comparable to those by the universal genetic code, by an amino acid composition analysis, GFP fluorescence measurements and the crystal structure determination. We also showed that another code, in which UGU/UGC codons are assigned to Ser, synthesizes an active enzyme.

This method will provide not only new insights into primordial genetic codes, but also an essential protein engineering tool for the assessment of the early stages of protein evolution. To create a protein with improved activity relative to that of the wild-type, random mutagenesis by an error-prone polymerase chain reaction is widely used in a directed evolution process involving multiple rounds of mutagenesis and selection. For a simplified protein containing less than 20 amino acid species, however, efficient evolution with the random mutation strategy has been prevented by the reappearance of codons, generated by mutation, for the specific amino acids to be excluded. In this work, we showed that the simplified code completely excludes the specific amino acid from the genetic code. Therefore, even if the codon for the specific amino acid in the universal code appears in the sequence through a mutation, the amino acid to be excluded is not incorporated within the protein. As a result, the simplified codes will allow us to efficiently search the sequence space of simplified proteins.

In this presentation, we will show the generality of our method for the simplification, by constructing other types of further simplified codes including a 16-amino-acid code.

Reference: [1] Kawahara-Kobayashi A, et al. *Nucleic Acids Res.* 2012. 40(20):10576-84. (featured article, top 5% of papers)