Title of Thesis:
Computational approaches for comparing biological data in the fields of genomics and RNA-protein interactions

Summary of Thesis:
To clarify the role played by biological components such as nucleic acids and proteins in the system called life, it is extremely important to develop a systematic method to compare and examine genetic information and the interactions between these components. In addition, to observe and understand the system of life, this biological information should ideally be stored and analyzed in an exactly comparable format. This thesis addresses this problem via two main approaches. First, I explain the development of Restauro-G, a highly efficient automatic genome re-annotation software for bacterial genomes. Restauro-G enabled comparative genome analysis of genomic sequence information, which is increasing at an explosive rate. Second, I developed a method to extract features related to the binding sites of tRNA and aminoacyl synthetase complex in a format that can be adequately compared based on the principles of information science. This comparative analysis revealed the commonality and specificity of the aminoacyl–tRNA synthetase interaction region. I succeeded in handling the genomic information that forms the basis of information flow in Central Dogma and the information on interactions between nucleic acid (RNA) and protein, which are responsible for major cellular functions. This result is expected to improve our understanding of biological phenomena in systems biology.

Keyword: Bioinformatics, Transfer RNA, Genome, Translation, Aminoacyl tRNA synthetase