Division of Biosciences

Department of Computational Biology and Medical Sciences

Laboratory	Faculty	Introduction of research activities and laboratory	Key words	Projects or activities summer program students can participate
Frith Laboratory	Prof. Martin FRITH	Our ultimate aim is to decipher the functional and historical information in genome sequences. We do this using statistical models (such as hidden Markov models) and computational methods (such as enhanced suffix arrays and dynamic programming). A major approach is to compare and align related sequences to each other, to see how they have evolved. One recent focus is characterization of genome rearrangements in evolution and disease. Another long-term interest is promoter sequences and DNA motifs that regulate gene expression. Further interests are everything “weird”: malaria genomes (80% A+T), frameshifts (especially in microbial metagenomes), unexplained evolutionary conservation, trans-splicing, etc. Our physical location is partly in Kashiwa, and partly in Odaiba in central Tokyo.	1) Genome 2) Evolution 3) Orthology 4) Probability-based	Students are encouraged to pursue their own ideas on analyzing genetic sequences. There are broadly two types of project: biological investigation, and method development. Examples of biological investigation: survey the evolution of gene structure by gain or loss of splice sites, frameshifting, gene fusion or fission, etc; compare the evolution of mitochondrial versus plastid genomes; compare genome evolution to major body-form evolution (e.g. snakes, whales). Examples of method development: make a sensitive probabilistic model for finding distantly-related DNA sequences; devise a beautiful way to visualize complex sequence rearrangements; develop a way to extract specific rearrangement events from pair-wise alignments of long sequences (e.g. long DNA reads or whole genomes).
Morishita Laboratory	Prof. Shinichi MORISHITA	We have been attempting to develop efficient and accurate algorithms for uncovering “dark matters” in genomes that are hard to observe using traditional second generation DNA sequencers such as Illumina HiSeq and Ion torrent. Typical examples of dark matters are genomic sequences of long repetitive elements (LINE and LTR), centromeres, telomeres, homologous chromosomes, and microbiome. Towards this end, we exploit full potential of our third generation sequencers (PacBio Sequel, Oxford nanopore, 10X	1) Single-molecule sequencing 2) Single-cell sequencing 3) Centromeres 4) Microbiome 5) DNA methylation	We introduce basic ideas and algorithms for handling single-molecule sequencing data as well as how to operate third generation sequencers. We also provide a couple of open problems in this research fields. Afterwards, summer students are expected to propose and develop new ideas, applications, or algorithms through brainstorming with our graduate and undergraduate students. They can use highly

		Chromium) that realize single-molecule sequencing and are capable of sequencing very long DNA fragments of >10,000 base pairs. Extending single-molecule sequencing, we also have been devising efficient algorithms for observing DNA methylation states of dark matters, for example, CpG methylation of centromeric repeats and highly repetitive transposons, and 6mA in gut microbiome so as to understand their biological functions.		parallel computers with thousands of CPU cores if they are interested.