Usage Report for Fiscal Year 2020

Project Title:

Predicting 3D structure of the human genome

Name:
○Saumya Agrawal (1), Michiel JL de Hoon (1)
Laboratory at RIKEN:
(1) Center for Integrative Medical Sciences, Laboratory for Applied Computational Genomics

1． Background and purpose of the project, relationship of the project with other projects

Transcription of the human genome yields mostly long non-coding RNAs (lncRNAs). Systematic functional annotation of lncRNAs is challenging due to their low expression level, cell type-specific occurrence, poor sequence and conservation between orthologs. Currently, 95% of human lncRNAs have no functional characterization. Using Hi-C and Cap Analysis of Gene Expression (CAGE) data in 18 human cell types, we located genomic regions in 3D proximity to lncRNA genes and identified functional clusters of interacting mRNA genes, lncRNAs and enhancers.

2． Specific usage status of the system and calculation method

The Hi-C data was mapped to the genome and the mapped reads were processed using different Hi-C related tools for the analysis. Identifying interacting genomic regions is a computationally intensive process and the supercomputer facility was used extensively to determine these structures.

3． Result and Conclusion

We provide a cell type-specific functional annotation for 40 7,688 out of 14,198 (54.14%) lncRNAs. The cell type-specific lncRNA functional annotations are provided through an interactive visualization web portal at https://fantom.gsc.riken.jp/zenbu/reports/#FANTOM6_HiC

4． Schedule and prospect for the future

We will continue the analysis as we are integrating different data sets types to further interpret the results.

5． If no job was executed, specify the reason.

Not applicable

Fiscal Year 2020 List of Publications Resulting from the Use of the supercomputer

[Paper accepted by a journal]

None

[Conference Proceedings]

None

[Oral presentation]

1. Agrawal Saumya, Alam Tanvir, Koido Masaru, Kulakovskiy Ivan, Severin Jessica, Abugessaisa Imad, Buyan Andrey, Dostie Josee, Ito Masayoshi, Kondo Naoto, Li Yunjing, Mendez Mickael , Ramilowski Jordan, Yagi Ken, Yasuzawa Kayoko, Yip Chi Wai, Okazaki Yasushi, Hoffman Michael M, Strug Lisa, Hon Chung Chau, Terao Chikashi, Kasukawa Takeya , Makeev Vsevolod, Shin Jay, Carninci Piero, De Hoon Michiel. The biological role of long noncoding RNAs in human cells inferred from chromatin conformation data, The ENCODE Research Applications and Users Meeting, Online, Barcelona, Spain September 30, (2020)

2. Agrawal Saumya, Alam Tanvir, Koido Masaru, Kulakovskiy Ivan, Severin Jessica, Abugessaisa Imad, Buyan Andrey, Dostie Josee, Ito Masayoshi, Kondo Naoto, Li Yunjing, Mendez Mickael , Ramilowski Jordan, Yagi Ken, Yasuzawa Kayoko, Yip Chi Wai, Okazaki Yasushi, Hoffman Michael M, Strug Lisa, Hon Chung Chau, Terao Chikashi, Kasukawa Takeya , Makeev Vsevolod, Shin Jay, Carninci Piero, De Hoon Michiel. The biological role of long noncoding RNAs in iPSC and differentiated human cells inferred from chromatin conformation data, IMS-Stanford symposium, Online, RIKEN, Japan November 5, (2020)

[Poster presentation]

None

[Others (Book, Press release, etc.)]

None