

Bioinformatics Center - Bioknowledge Systems -

<http://kanehisa.kuicr.kyoto-u.ac.jp/>



Prof
KANEHISA, Minoru
(D Sc)



Assoc Prof
GOTO, Susumu
(D Eng)



Assist Prof
HATTORI, Masahiro
(D Sc)



PD
KAWANO, Shin
(D Eng)



PD
SCHWARTZ, Jean-Marc
(D Eng)



PD
PERRET, Jean-Luc
(D Sc)



PD
JAUREGUI, Ruy
(D Sc)



PD
TANAKA, Nobuya
(D Eng)



PD
OKAMOTO, Shinobu
(D Sc)



PD
KARLSSON-
WHEELOCK, Åsa
(Ph D)



PD
WHEELOCK, Craig
(Ph D)

Research Associates (pt)

IGARASHI, Yoshinobu
YOSHIZAWA, Akiyasu
LIMVIPHUVADH, Vachiraneec
KOTERA, Masaaki

Researcher

HIRAKAWA, Mika

Students

MINOWA, Yosuke (D3)
ITOH, Masumi (D3)
HIZUKURI, Yoshiyuki (D3)
YAMANISHI, Yoshihiro (D3)
YAMADA, Takuji (D2)
OKUDA, Shujiro (D2)
MORIYA, Yuki (D2)
FUJITA, Masashi (D2)
SAKIYAMA, Tadahiko (D1)
TANAKA, Michiro (D1)

SAKAI, Hiroki (D1)
OH, Min-A (D1)
SASAKI, Naobumi (M3)
MIYAMA, Takashi (M2)
OTA, Koji (M2)
HONDA, Wataru (M1)
HASHIMOTO, Kosuke (M1)
MUTO, Ai (M1)
SAITO, Koichi (M1)

Scope of Research

Owing to continuous developments of high throughput experimental technologies, projects are going on not only to determine complete genome sequences of an increasing number of organisms, but also to analyze gene expression profiles both at the mRNA and protein levels and to catalog protein 3D structure families. Bioinformatics provides basic concepts as well as practical methods to go up from the molecular level to the cellular level, and eventually to still higher levels, to that of biological systems by analyzing complex interactions among building blocks and with dynamic environments. We have been developing such bioinformatics technologies and the KEGG system (<http://www.genome.jp/kegg/>), which is our attempt to uncover and utilize cellular functions through the reconstruction of protein interaction networks from genome information.

Research Activities (Year 2004)

Grants

Kanehisa M, Biological systems database (KEGG) and genome information science. Research for the Future Program, JSPS.

Kanehisa M, BRITE: Deductive Database of the Genome and the Biological System Based on Binary Rela-

tions, Bioinformatics Research and Development, JST.

Kanehisa M, Education and Research Organization for Genome Information Science, MEXT.

Kanehisa M, Knowledge Information Infrastructure for Genome Science. Kyoto University 21st Century COE Program, MEXT.

KEGG GLYCAN: Computational Approach to Comparative Glycomics

Glycans are an important class of biological macromolecules in addition to DNA and proteins. Under the KEGG project we released the GLYCAN database for carbohydrate structures and its associated information (Figure 1). We developed a novel method for conducting comprehensive analyses of comparative glycomics using glycan structures in KEGG GLYCAN. First, we developed a new similarity measure for comparing glycan structures taking into account the characteristic mechanisms of glycan biosynthesis, and we applied it to classify glycans of different species and tissues in the framework of support vector machines (SVMs). Next, we extracted characteristic functional units (motifs) of glycans (Figure 2).

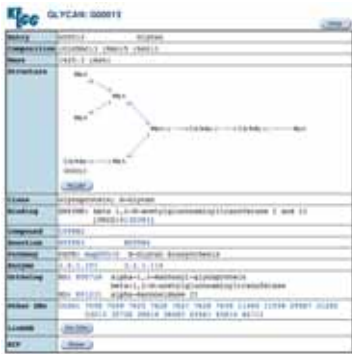


Figure 1. Structure based glycan database, KEGG GLYCAN.

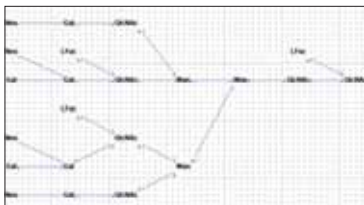


Figure 2. Extracted characteristic glycan structure of mouse.

KEGG Orthology: a Hierarchical Classification of All Functional Orthologs in the Gene Universe

KEGG Orthology (KO) is developed from “Ortholog IDs”, an extension of EC numbers in the KEGG PATHWAY database. It is one of the major projects of our

laboratory and is intended to identify and classify orthologous gene clusters among all species stored in the KEGG GENES database. Using the KEGG SSDB database, orthologs in the KEGG GENES database are classified and annotated both automatically and manually. Namely, orthologs are classified by a computational clustering method according to KEGG SSDB scores (Figure 3), and the function of each ortholog is further examined manually.

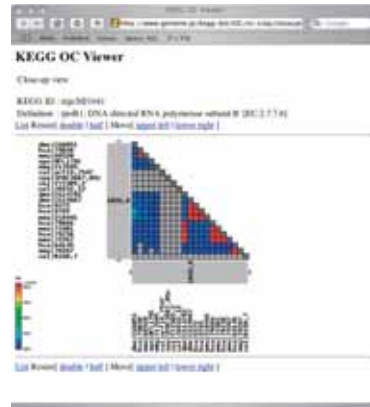


Figure 3. Ortholog Cluster Viewer on KEGG SSDB.

Many genome sequencing projects for eukaryotes are ongoing, and many draft genomes are available. Despite the delays in completing the genome project for eukaryotes, the draft sequences provide much important information. Therefore, the DGENES database was added to the KEGG database in order to store such draft genomes (Figure 4). We are currently developing automatic methods to annotate the genes in DGENES according to the KEGG Orthology.



Figure 4. KEGG DGENES database.

Kanehisa M, Development of the integrated database for bacterial genomes and their cellular function. Grant-in-Aid for Scientific Research on Priority Areas “Genome Biology”, MEXT.

Goto S, Construction and retrieval of highly integrated biological databases. Grant-in-Aid for Scientific Research on Priority Areas “Genome Information Science”, MEXT.