Bioinformatics

1. Staffs and Students

Professor:	Hiroshi Tanaka	
Associate Professor:	Yoshihito Niimura	
Assistant Professor:	Soichi Ogishima	
Visiting Professor:	Hiroshi Mizushima	
Visiting Assistant Professor:	Yasen Mahmut	
Project Associate Professor:	Jun Nakaya,	Fengrong Ren,
	Takako Takai	
Project Lecturer:	Kazuro Shimokawa,	Kanae Oda
Project Assistant Professor:	Takeshi Hase (-Jan., Dec.-),	Naoki Hasegawa,
	Kaei Hiroi,	Keisuke Ido,
	Kaoru Mogushi,	Satoshi Shoji (-March),
	Masaki Morioka (-May)	
Technical Staff:	Masaya Itoda (-March),	Ken Miyaguchi,
	Shota Nemoto	
Graduate Students:	Isao Okada,	Takayuki Ohnishi,
	Emilio Campos,	Yuki Katayama,
	Yoshitomo Tanaka,	Shun-ya Takahashi,
	Hironobu Yamaguchi,	Yoshiyuki Kaneko,
	Ryosuke Ishiwata,	Hideaki Takata,
	Arihito Endo,	Satoshi Nagaie,
	Yasuko Takahashi,	Akiko Hatano,
	Taika Muto,	Yuki Tanaka,
	Hiromi Matsumae,	Izumi Yoshida,
	Hyeryung Kim,	Ken Miyaguchi,
	Kumiko Iijima,	Kyaw Tun,
	Satoru Suzuki,	Eiichi Ueno,
	Naoko Kasahara,	Tadashi Urashima,
	Sakiko Ohta,	Masataka Kikuchi,
	Yasuha Tanaka,	Hajime Sawai,
	Asami Suzuki,	Kasuo Iida,
	Taro Kishimoto,	Tadashi Ebara,
	Yasunori Ohto,	Chikako Shimizu,
	Afsaneh Eslami,	Shoko Nukaya,
	Hiroaki Hasegawa,	Wanping Aw,
	Kae Suzuki,	Reiko Yamaguchi,
	Satoshi Mizuno,	Norio Tanaka,
	Yukari Watanabe,	Tadaaki Katsuta,
	Moegi Maeda,	Kotaro Oda

2. Purpose of Education

Prof. Tanaka is charged with education of interdisciplinary medical informatics and bioinformatics. For undergraduate classes he educates “Clinical Informatics”, “Statistics for Health Science”, “Practice in Clinical Informatics II”, “Project Research”, and “Basics of Clinical Informatics”. For graduate classes he educates “Computational Biology”, “Bioinformatics Computation”, “Systems Pathology”, “Statistical Genetics and Medical Statistics”, “English Debate”, and “Practice in Global Linkage between University and Industry”, “Integrated Bioinformatics”, “Applied Biological Chemistry”, and “Integrated Translational Research”. He supervises 31 students of PhD course and 2 students in Master course in Graduate School of Medical and Dental Sciences and 6 students of PhD course and 9 students in Master course in Biomedical Science PhD Program. He is a principle investigator of “International Educational Program for Interdisciplinary Disease Science” granted by Program for Accelerating Internationalization of Higher (University) Education. In this year this program accepted and educated a graduate student of Glasgow University in our international joint-degree education for master

degree. A cooperative teaching course was provided to Japanese students in alliance with Harvard Medical School for the topics of translational research. He is also a co-principle investigator of "New Joint-education of Multi-disciplinary Students with Leading-edge Computer Technologies for OMICS Medicine" granted by JSPS Program for Enhancing Systemic Education in Graduate School. This is a joint program of Tokyo Institute of Technology (TITECH) and TMDU. This program provides a joint education system for students of the two universities to study different fields and to work together so as to develop advanced computational technologies in medicine. This program also implemented a super high-vision remote education system which connects two campuses of TITIECH and a campus of TMDU. He is also a principle investigator of a follow-up program of "Educational Program for Biomedical Omics Information Scientists", which had granted by Special Coordination Funds for Promoting Science and Technology. The follow-up program provided a series of seminars on advanced topics on regenerative medicine based in iPS technology, next sequencing technology and its medical application, and new trends in drug discovery. 48 people who had graduated this program participated in the follow-up program.

3. Research Subjects

In our laboratory, we conduct biological and medical researches from the viewpoint of Systems Biology.

Biological sciences: Recently, the whole genome sequences of diverse organisms have become available. Moreover, various "omix" information such as a proteome, transcriptome, and metabolome are currently accumulating. Our goal is to establish a grand-theory of biological sciences from the viewpoint of "evolving networks composed of biological molecules" by integrating omix information.

Medical sciences – Genomic and omix data are also utilized in the field of medicine. It has been revealed that most diseases are caused by the interaction among abnormalities of multiple genes, those at the tissue level, and environments. It is therefore possible to consider diseases as a system. From this standpoint, we try to establish the omix-based medicine.

[1] Difference in gene duplicability may explain difference in overall structure of protein-protein interaction networks among eukaryotes

To uncover the evolutionary mechanisms of protein-protein interaction networks (PINs), we investigated PINs from several eukaryotes, i.e., yeast, worm, fly, human, and malaria parasite. Many researchers suggested that a disassortative network, in which interactions between high- and low-degree genes, are favored while hub-hub interactions are suppressed. It was postulated that a disassortative structure minimizes unfavorable cross-talks between different hub-centric functional modules and was positively selected in evolution. Thus, many researchers believe that a disassortative structure is a common feature among eukaryotes. However, we showed that the yeast, worm, fly, and human PINs are disassortative while the malaria parasite PIN is not. By conducting simulation studies on the basis of a duplication-divergence model, we showed that a preferential duplication of low- and high-degree genes can generate networks with disassortative structure and those without disassortative structure, respectively. From this observation, we hypothesized that the difference in degree (# of interactions) dependence on gene duplications accounts for the difference in assortativity of PINs among eukaryotes. By comparing 55 proteomes in eukaryotes, we revealed that genes with lower degrees showed higher gene duplicabilities in the yeast, worm, and fly, while high-degree genes tend to have high duplicabilities in the malaria parasite. The observation supports the above hypothesis.

Our results suggest that disassortative structures in PINs among eukaryotes are merely a byproduct of preferential duplications of low-degree genes, which might be caused by a living environment of an organism.

[2] Investigation of disease mechanism using omics-based analysis

Recent advances in analysis techniques in molecular biology have led to the investigation of genome-wide data such as genome, transcriptome and proteome. In order to reveal the underlying biological mechanisms from such a large amount of "omics" data, integration of biomedical knowledge with multivariate statistical analysis or machine learning methods is one of the most crucial tasks for bioinformatics research. We have been performing collaborative research with our university hospital and other institutes mainly based on transcriptome analysis using DNA microarray, including the following topics: 1) identification of diagnosis marker for early relapse in hepatocellular carcinoma patients, 2) development of predictive marker for metastatic relapse in colorectal cancer, and 3) analysis of the molecular mechanisms of epithelial-mesenchymal transition in breast cancer cells.

[3] Bioinformatics on disease Omics data

The i2b2 (Informatics for Integrating Biology and the Bedside) is a database system developed by Harvard Medical School to facilitate integration of clinical patients data collected in various forms and by various people such as university hospitals, clinics, and organizations of patients. The i2b2 is designed to enable integration of many different data by ontology-based object-oriented database technologies. We constructed the i2b2 database with Japanese clinical patients data from our university hospital. We developed a computational pipeline to extract disease names from doctor's comments in Japanese and translate them into English using Natural Language Processing techniques. Comprehensive information on proteins plays an important role in elucidating molecular progress of a disease. There had been no concise and systematic method to identify terminus of proteins in proteomics. We applied a unique enzyme that digested N- and C- terminus of a Lys residue and developed a new method to identify protein terminus in MASS data. We investigated human plasma by our method and detected un-known protein terminus that may have pathogenic links to a disease.

[4] iCOD: an integrated clinical omics database based on the systems-pathology view of disease

Variety of information relating between genome and the pathological findings in disease will yield a wealth of clues to discover new function, the role of genes and pathways, and future medicine. In addition to molecular information such as gene expression and genome copy number, detailed clinical information is essential for such systematic omics analysis. In order to provide a basic platform to realize a future medicine based on the integration of molecular and clinico-pathological information of disease, we have developed an integrated clinical omics database (iCOD) in which comprehensive disease information of the patients is collected, including not only molecular omics data such as CGH (Comparative Genomic Hybridization) and gene expression profiles but also comprehensive clinical information such as clinical manifestations, medical images (CT, X-ray, ultrasounds, etc), laboratory tests, drug histories, pathological findings and even life-style/environmental information. The iCOD is developed to combine the molecular and clinico-pathological information of the patients to provide the holistic understanding of the disease. Furthermore, we developed several kinds of integrated view maps of disease in the iCOD, which summarize the comprehensive patient data to provide the information for the interrelation between the molecular omics data and clinico-pathological findings as well as estimation for the disease pathways, such as three layer-linked disease map, disease pathway map, and pathome-genome map. With these utilities, our iCOD aims to contribute to provide the omics basis of the disease as well as to promote the pathway-directed disease view. The iCOD database is available online, containing 140 patient cases of hepatocellular carcinoma, with raw data of each case as supplemental data set to download. The iCOD and supplemental data can be accessed at http://omics.tmd.ac.jp/icod_pub_eng.

4. Publications

[Original Papers]

1. Hase T, Niimura Y, Tanaka H: Difference in gene duplicability may explain the difference in overall structure of protein-protein interaction networks among eukaryotes, BMC evolutionary biology, 10:358, 2010
2. Shimokawa K, Mogushi K, Shoji S, Hiraishi A, Mizushima H, Tanaka H: iCOD: an integrated clinical omics database based on the systems-pathology view of disease, BMC Genomics, 11:S19, 2010
3. Shimizu S, Iida S, Ishiguro M, Uetake H, Ishikawa T, Takagi Y, Kobayashi H, Higuchi T, Enomoto M, Mogushi K, Mizushima H, Tanaka H, Sugihara K: Methylated BNIP3 gene in colorectal cancer prognosis, Oncology Letters, 1:865-72, 2010
4. Takano S, Sogawa K, Yoshitomi H, Shida T, Mogushi K, Kimura F, Shimizu H, Yoshidome H, Ohtsuka M, Kato A, Ishihara T, Tanaka H, Yokosuka O, Nomura F, Miyazaki M: Increased circulating cell signalling phosphoproteins in sera are useful for the detection of pancreatic cancer, British journal of Cancer, 103:223-31, 2010
5. Nakahara I, Miyamoto M, Shibata T, Akashi-Tanaka S, Mogushi K, Oda K, Ueno M, Takakura N, Mizushima H, Tanaka H, Ohta T: Up-regulation of PSF1 Promotes the Growth of Breast Cancer Cells, Genes to Cells, 15:1015-24, 2010
6. Ohashi K, Ota S, Ohno-Machado L, Tanaka H: Smart medical environment at the point of care:Auto-tracking clinical interventions at the bed side using RFID technology, Computers in Biology and Medicine, 40:545-54, 2010
7. Tanaka H: Omics-based Medicine and Systems Pathology, Methods of Information in Medicine, 49:173-85, 2010
8. Yoshida T, Kobayashi T, Itoda M, Muto T, Miyaguchi K, Mogushi K, Shoji S, Shimokawa K, Iida S, Uetake H, Ishikawa T, Sugihara K, Mizushima H, Tanaka H: Clinical omics analysis of colorectal cancer incorporating copy number aberration and gene expression data, Cancer Informatics, 9:147-61, 2010
9. Matsuyama T, Ishikawa T, Mogushi K, Yoshida T, Iida S, Uetake H, Mizushima H, Tanaka H, Sugihara K: MUC12 mRNA Expression is an independent marker of prognosis in stage II and stage III colorectal cancer, International

Journal of Cancer, 127:2292-99, 2010
10. Tanaka S, Mogushi K, Yasen M, Noguchi N, Kubo A, Nakamura N, Ito K, Miki Y, Inazawa J, Tanaka H, Arii S: Gene expression phenotypes for vascular invasiveness of hepatocellular carcinomas, Surgery, 147:405-14, 2010
11. Nakaya J, Kimura M, Hiroi K, Ido K, Yang W, Tanaka H: Genomic sequence variation markup language (GSVML), International Journal of Medical Informatics, 79:130-42, 2010
12. Watanabe K, Kurihara Y, Nakamura T, Tanaka H: Design of a Low-Frequency Microphone for Mobile Phones and Its Application to Ubiquitous Medical and Healthcare Monitoring, IEEE SENSORS JOURNAL 1, 10:934-41, 2010
13. Nakagawa M, Sakamoto N, Ueyama M, Mogushi K, Nagaie S, Itsui Y, Azuma S, Kakinuma S, Tanaka H, Enomoto N, Watanabe M: Mutations in the interferon sensitivity determining region and virological response to combination therapy with Pegylated-interferon alpha 2b plus ribavirin in patients with chronic hepatitis C-1b infection, Journal of Gastroenterology, 45:656-65, 2010
14. Nishimura-Sakurai Y, Sakamoto N, Mogushi K, Nagaie S, Nakagawa M, Itsui Y, Sekine-Osajima Y, Tasaka-Fujita M, Onuki-Karakama Y, Suda G., Mishima K, Yamamoto M, Ueyama M, Funaoka Y, Watanabe T, Chen CH, Kakinuma S, Tsuchiya K, Tanaka H, Enomoto N, Watanabe M: Comparison of HCV-associated gene expression and cell signaling pathways in cells without HCV replicon and in replicon-cured cell, Journal of Gastroenterology, 45:523-36, 2010
15. Tsubota A, Mogushi K, Nariai K, Tanaka H, Tada N: IQGAP1 and vimentin may be key regulator genes in naturally occurring hepatotumorigenesis induced by oxidative stress, Carcinogenesis, 31:504-11, 2010
16. Matsui A, Go Y, Niimura Y: Degeneration of olfactory receptor gene repertories in primates: No direct link to full trichromatic vision, Mol. Biol. Evol, 27:1192-1200, 2010
17. Nakagawa S, Niimura Y, Miura K, Gojobori T: Dynamic evolution of translation initiation mechanisms in prokaryotes, Proc. Natl. Acad. Sci. USA, 107:6382–87, 2010
18. Obulhasim G, Fujii H, Matsumoto T, Yasen M, Abe M, Matsuoka S, Ohtsuji N, Hino O: Mesothelin gene expression and promoter methylation/hypomethylation in gynecological tumors, Eur J Gynaecol Oncol, 31:63-71, 2010
19. Aihara A, Tanaka S, Yasen M, Matsumura S, Mitsunori Y, Murakata A, Noguchi N, Kudo A, Nakamura N, Ito K, Arii S: The selective Aurora B kinase inhibitor AZD1152 as a novel treatment for hepatocellular carcinoma, J Hepatol, 52: 63-71, 2010