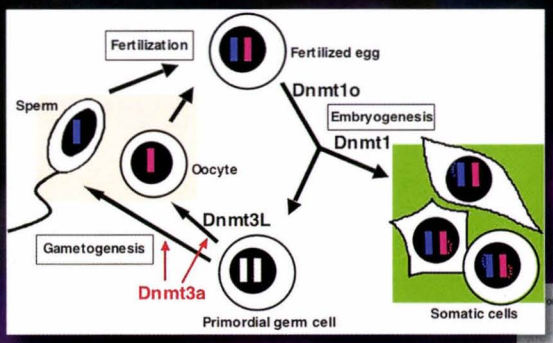
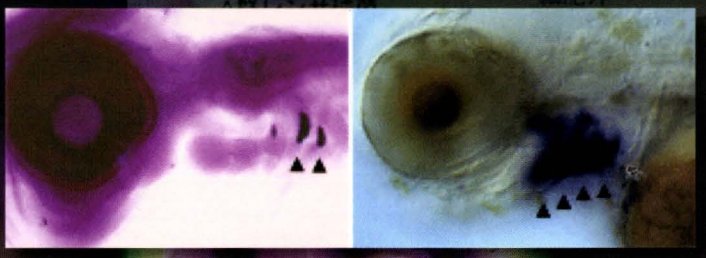
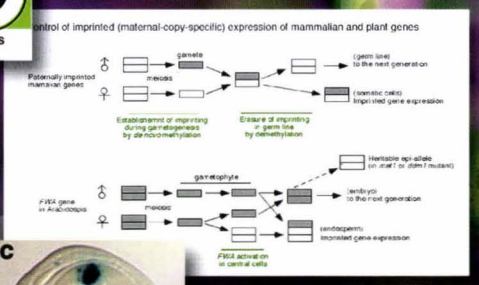
ANNUAL REPORT

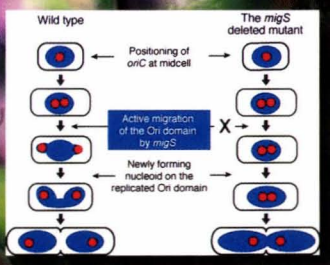
No.55
2004

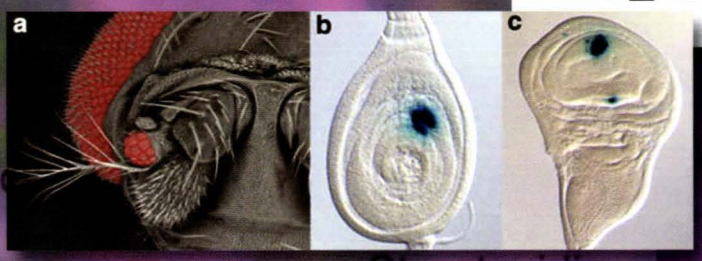
ISSN 0077-4995

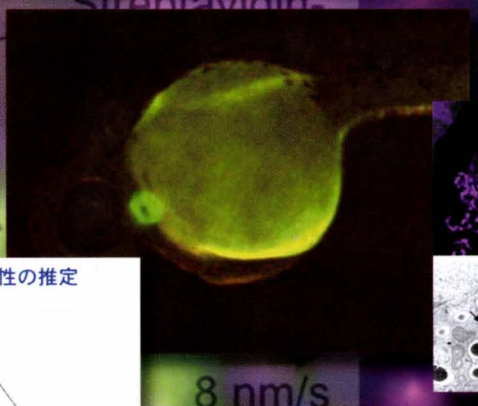


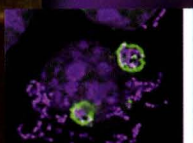


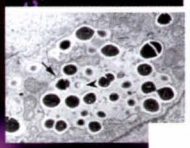


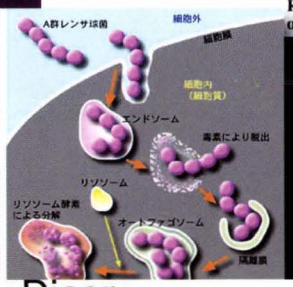


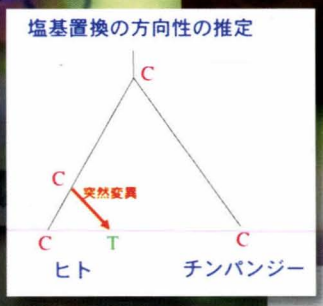


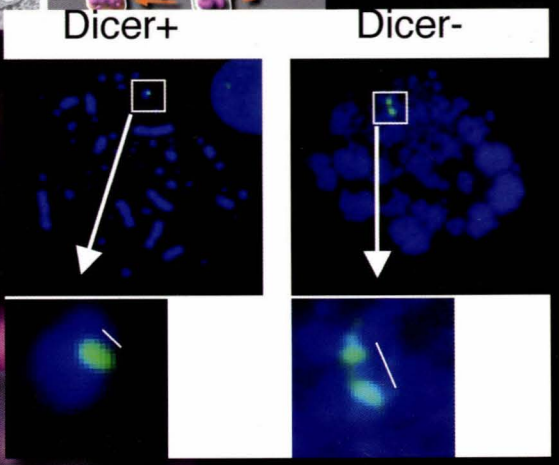


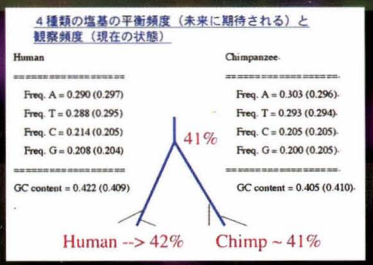


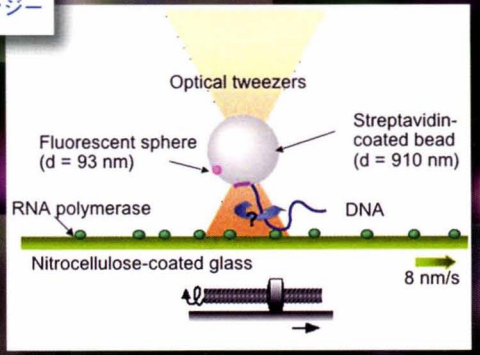






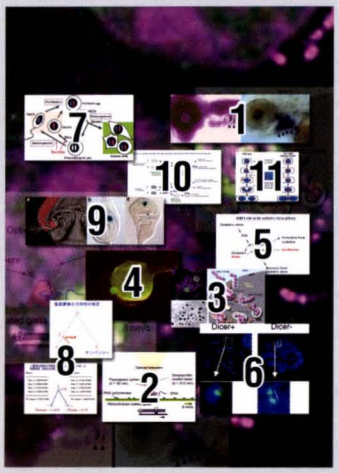






Research Organization of Information and Systems
NATIONAL INSTITUTE OF GENETICS
http://www.nig.ac.jp

Highlights of 2004 (http://www.nig.ac.jp/hot/index.html)



1. The origin of the parathyroid gland; persistence of gills in tetrapods
   Okabe M, Graham A.
   PNAS, *Vol.101 No.51*, 17716-17719, 2004

2. A DNA-binding protein can slide along a helical groove of DNA
   Kumiko Sakata-Sogawa, Nobuo Shimamoto.
   Proc. Nat. Acad. Sci. U.S.A, *101*, 14731-35, 2004

3. Autophagy defends cells against invading group A Streptococcus
   Nakagawa I, Amano A, Mizushima N, Yamamoto A, Yamaguchi H, Kamimoto T, Nara A, Funao J, Nakata M, Tsuda K, Hamada S, Yoshimori T
   Science, *306*, 1037-1040, 2004

4. A transposon-mediated gene trap approach identifies developmentally regulated genes in zebrafish.
   Kawakami, K., Takeda, H., Kawakami, N., Kobayashi, M., Matsuda, N., Mishina, M.
   Developmental Cell, *7*, 133-144, 2004.

5. MBF1 role in the oxidative stress defense
   Jindara M, Gaziova I, Uhlirova M, Okabe M, Hiromi Y, Hirose S.
   EMBO Journal, *23*, 3538-3547, 2004

6. Chromosome segregation and RNAi in vertebrate cells
   Fukagawa T, Nogami M, Yoshikawa M, Ikeno M, Okazaki T, Takami Y, Nakayama T, Oshimura M.
   Nature Cell Biology, *6*, 784-791, 2004

7. Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting.
   Masahiro Kaneda, Masaki Okano, Kenichiro Hata, Takashi Sado, Naomi Tsujimoto, En Li & Hiroyuki Sasaki
   Nature, *429*, 900-903, 2004

8. DNA sequence and comparative analysis of chimpanzee chromosome 22
   H. Watanabe, A. Fujiyama, M. Hattori, T. D. Taylor, A. Toyoda, Y. Kuroki, H. Nuchi, A. BenKahla, H. Lehrach, R. Sudbrak, M. Kube, S. Taenzer, P. Galgoczy, M. Platzer, M. Scharfe, G. Nordsiek, H. BloÅNcker, I. Hellmann, P. Khaitovich, S. PaÅNaÅNbo, R. Reinhardt, H.-J. Zheng9, X.-L. Zhang9, G.-F. Zhu9, B.-F. Wang9, G. Fu9, S.-X. Ren9, G.-P. Zhao, Z. Chen, Y.-S. Lee, J.-E. Cheong, S.-H. Choi, K.-M. Wu, T.-T. Liu, K.-J. Hsiao, S.-F. Tsai, C.-G. Kim, S. OOta, T. Kitano, Y. Kohara, N. Saitou, H.-S. Park, S.-Y. Wang, M.-L. Yaspo, and Y. Sakak
   Nature , *429*, 382-388, 2004

9. Diverse sensory organs arose via modification of a proto-sensory organ.
   Niwa N, Hiromi Y, Okabe M.
   Nature Genetics, *36* (3), 293-297, 2004

10. Imprinting of plants over generations
    Kinoshita T, Miura A, Choi Y, Kinoshita Y, Cao X, Jacobsen SE, Fischer RL, and Kakutani T.
    Science , *303*, 521-523, 2004

11. migS, a cis acting site for bipolar positioning of oriC on the Escherichia coli chromosome
    Yamaichi, Y. and Niki, H.
    EMBO Journal, *23*, 221-233, 2004

CONTENTS

G. CENTER FOR GENETIC RESOURCE INFORMATION

H. STRUCTURAL BIOLOGY CENTER

Introduction

The National Institute of Genetics (NIG) was established in 1949 as the central institute to study various aspects of genetics. It was reorganized in 1984 as an interuniversity research institute to promote collaborations with researchers at universities. Since 1988, NIG has been participating in graduate education as the Department of Genetics of the Graduate University for Advanced Studies (SOKENDAI). NIG also serves as a center for various genetic resources such as mutant strains, clones and vectors, and houses DDBJ, the DNA Data Bank of Japan, and a DNA sequencing center.

The history of NIG overlaps the period of a revolution in the field of life science. Genetics is no longer a discipline to study the rules and mechanisms of heredity, but has become the basis for all fields of life science. Molecular techniques now allow us not only to decipher the entire genome sequence of organisms including humans, but also to understand the details of higher biological phenomena: cell differentiation, morphogenesis, brain function, and evolution — the history of life itself. Currently, 36 research groups are actively performing pioneering and cutting-edge researches in these fields at NIG.

Recent generation of massive information on biological systems and their environment calls for new directions in life sciences, such as bioinformatics, system-level analysis, and theoretical approaches to extract knowledge from databases. To this end NIG and three other national institutes, the National Institute of Informatics, The Institute of Statistical Mathematics and the National Institute of Polar Research have formed a new organization, the Research Organization of Information and Systems (ROIS) in April 2004, as a part of the reform of national universities and research institutes in Japan. Interinstitutional collaborations within the new organization are in progress.

We welcome your comments and suggestions on our research activities and endeavors.

Yuji Kohara, Director-General

STAFF (as of December 31, 2004)

Director-General
KOHARA, Yuji, D. Sc.

Vice-Director
HIROSE, Susumu, D. Sc.

Member

1. Department of Molecular Genetics

YAMAO, Fumiaki, D. Sc., Head of the Department

Division of Molecular Genetics

FUKAGAWA, Tatsuo, D. Sc., Associate Professor
OKADA, Masahiro, D. Sc.

Division of Mutagenesis

YAMAO, Fumiaki, D. Sc., Professor
TSUTSUI, Yasuhiro, D.Med.

Molecular Mechanism Laboratory

MITSUZAWA, Hiroshi, D. Sc.
SEINO, Hiroaki, D. Sc.

Division of Nucleic Acid Chemistry

AIMOTO, Saburo, D. Sc., Adjunct Professor
KATAYAMA, Tsutomu, D. Sc., Adjunct Professor

2. Department of Cell Genetics

ARAKI, Hiroyuki, D. Sc., Head of the Department

Division of Cytogenetics

YOSHIMORI, Tamotsu, D. Med., Professor
UMEBAYASHI, Kyohei, D. Ag.

Division of Microbial Genetics

ARAKI, Hiroyuki, D. Sc., Professor
YASUDA, Seiichi, D. Sc., Associate Professor
KAMIMURA, Yoichiro, Dc. Med.
TANAKA, Seiji, D. Sc.

Division of Cytoplasmic Genetics

KOBAYASHI, Kazuto, Ph. D., D. Ag., Adjunct Professor
YAMAMOTO, Masayuki, D. Sc., Adjunct Professor

3. Department of Developmental Genetics

HIROMI, Yasushi, D. Sc., Head of the Department

Division of Developmental Genetics

HIROMI, Yasushi, D. Sc., Professor
FUJISAWA, Toshitaka, Ph. D., Associate Professor
SHIMIZU, Hiroshi, D. Eng.
OKABE, Masataka, M. D., Ph. D.
ASAOKA, Miho, D. Sc.

Division of Gene Expression

HIROSE, Susumu, D. Sc., Professor

NISHIOKA, Kenichi, D. Med.
FUSE, Naoyuki, D. Sc.

Division of Molecular and Developmental Biology
KAWAKAMI, Koichi, D. Sc., Associate Professor
KISHIMOTO, Yasuyuki, D. Sc.

Division of Physiological Genetics
ISHINO, Fumitoshi, D. Sc., Adjunct Professor
GOTOH, Yukiko, D. Sc., Adjunct Associate Professor

4. Department of Population Genetics
SAITOU, Naruya, Ph. D., Head of the Department

Division of Population Genetics
SAITOU, Naruya, Ph. D., Professor
TAKANO, Toshiyuki, D. Sc., Associate Professor
SUMIYAMA, Kenta, D. Sc.
TAKAHASHI, Aya, D. Ag.

Division of Evolutionary Genetics
IKEMURA, Toshimichi, D. Sc., Professor

Division of Theoretical Genetics
FUJIYAMA, Asao, D. Sc., Adjunct Professor
SUGANO, Sumio, D. Med., Adjunct Professor

5. Department of Integrated Genetics
SASAKI, Hiroyuki, D. Med., Head of the Department

Division of Human Genetics
SASAKI, Hiroyuki, D. Med., Professor
SADO, Takashi, D. Sc.
HATA, Kenichiro, D. Med.

Division of Agricultural Genetics
KAKUTANI, Tetsuji, D. Sc., Associate Professor
SHIBAHARA, Kei-ichi, M. D., Ph. D., Associate Professor
KINOSHITA, Tetsu, D. Sc.
OGAWA, Yuya, D. Sc.

Division of Brain Fanction
HIRATA, Tatsumi, D. Med., Associate Professor
KAWASAKI, Takahiko, D. Sc.

Division of Applied Genetics
SIOTA, Kunio, D. Ag., Adjunct Professor
ARAKI, Takashi, D. Sc., Adjunct Associate Professor

6. Genetics Strains Research Center
SHIROISHI, Toshihiko, D. Sc., Head of the Center

Mammalian Genetics Laboratory
SHIROISHI, Toshihiko, D. Sc., Professor
TAMURA, Masaru, D. Sc.

Mammalian Development Laboratory
SAGA, Yumiko, D. Sc., Professor
KOKUBO, Hiroki, D. Sc.
MITSUI, Kaoru, D. Med.

Mouse Genomics Resource Laboratory
KOIDE, Tsuyoshi, D. Med., Associate Professor
Model Fish Genomics Resource Laboratory
SAKAI, Noriyoshi, Ph. D., Associate Professor
SHINYA, Minori, D. Sc.
Plant Genetics Laboratory
KURATA, Nori, D. Ag., Professor
ITO, Yukihiro, D. Ag.
Microbial Genetics Laboratory
NISHIMURA, Akiko, D. Ag., Professor
Invertebrate Genetics Laboratory
UEDA, Ryu, D. Sc., Professor
TAKAHASHI, Kuniaki, D. Sc.
Labolatory for Frontier Research
ISSHIKI, Takako, D. Sc., Associate Professor
KUSANO, Ayumi, D. Sc.

7. Center for Genetic Resource Information
NISHIMURA, Akiko, D. Ag., Head of the Center
Genetic Informatics Laboratory
YAMAZAKI, Yukiko, D. Sc., Associate Professor
Genomu Biology Laboratory
KOHARA, Yuji, D. Sc., Professor
ANDACHI, Yoshiki, D. Sc.
Comparative Genomics Laboratory
Publicity and Intellectual Property Unit
TOMIKAWA, Munehiro, D. Pharm, Professor

8. Structral Biology Center
KATURA, Isao, D. Sc., Head of the Center
Biological Macromolecules Laboratory
TOKUNAGA, Makio, D. Sc., Professor
SHIINA, Nobuyuki, D. Sc.
Molecular Biomechanism Laboratory
SHIMAMOTO, Nobuo, D. Sc., Professor
NAKAYAMA, Hideki, D. Eng.
SUSA, Motoki, D. Sc.
Multicelluar Organization Laboratory
KATURA, Isao, D. Sc., Professor
KIMURA, Koutarou, D. Ag.
Biomolecular Structure Laboratory
SHIRAKIHARA, Yasuo, D. Sc., Associate Professor
ITO, Hiroshi, D. Sc.
Gene Network Laboratory
SUZUKI, Emiko, D. Med., Associate Professor

9. Center for Information Biology and DNA Data Bank of Japan
GOJOBORI, Takashi, D. Sc., Head of the Center
Laboratory for DNA Data Analysis

GOJOBORI, Takashi, D. Sc., Professor
IKEO, Kazuho, D. Sc., Associate Professor
SUZUKI, Yoshiyuki, M. D., Ph. D.

Laboratory for Gene-Product Informatics

NISHIKAWA, Ken, D. Sc., Professor
FUKUCHI, Satoshi, D. Sc.
KINJO, Akira, D. Sc.

Laboratory for Gene Function Research

TATENO, Yoshio, Ph. D., D. Sc., Professor
BARRERO, Robert A., D. Sc.

Laboratory for the Research and Development of Biological Databases

SUGAWARA, Hideaki, D. Eng., Professor
ABE, Takashi, D. Sc.

Gene-Expression Analysis

OKUBO, Kousaku, M. D., Ph. D., Professor
ITO, Koichi, M. D., Ph. D.

10. Radioisotope Center

NIKI, Hironori, D. Med., Associate Professor Head of the Center
OGATA, Yasuyuki, D. Pharm. Sci.

11. Experimental Farm

KURATA, Nori, D. Ag., Head of the Farm
NONOMURA, Ken-ichi, D. Ag.

12. Technical Section

ISHII, Yuriko, Chief of the Section

13. Department of Administration

IKINAGA, Tadatoshi, Head of the Department
ISHIDA, Yuzo, Chief of the General Affairs Section
KAWAGUCHI, Kenji, Chief of the Finance Section

ADVISORY COMMITTEE (as of December 31, 2004)

Chairman

HIROSE, Susumu; Professor, Department of Developmental Genetics, National Institute of Genetics

Vice-chairman

SEKIGUCHI, Mutsuo; Adjunct Professor, Fukuoka Dental College

Outside Members (Alphabetical order)

NAKAMURA, Haruki; Professor, Institute for Protein Research, Osaka University

NISHIDA, Eisuke; Professor, Graduate school of Biostudies, Kyoto University

OGAWA, Tomoko; Vice-Director, Iwate College of Nursing

OKADA, Norihiro; Professor, Department of Bioscience and Biotechnology, Tokyo Institute of Technology

OSUMI, Noriko; Professor, Graduate School of Medicine, Tohoku University

SHINOZAKI, Kazuo; Chief-Scientist, RIKEN

TACHIDA, Hidenori; Professor, Faculty of Sciences, Kyusyu University

TAKAGI, Toshihisa; Professor, Graduate School of Frontier Sciences, The University of Tokyo

TSUJI, Shoji; Professor, The University of Tokyo Hospital

Inside Members (Alphabetical order)

ARAKI, Hiroyuki; Professor, Department of Cell Genetics

GOJOBORI, Takashi; Professor, Center for Information Biology and DNA Data Bank of Japan

HIROMI, Yasushi; Professor, Department of Developmental Genetics

KATSURA, Isao; Professor, Structural Biology Center

NISHIMURA, Akiko; Professor, Genetic Strains Research Center

SAITOU, Naruya; Professor, Department of Population Genetics

SASAKI, Hiroyuki; Professor, Department of Integrated Genetics

SHIMAMOTO, Nobuo; Professor, Structural Biology Center

SHIROISHI, Toshihiko; Professor, Genetic Strains Research Center

YAMAO, Fumiaki; Professor, Department of Molecular Genetics

ADVISORY BOARD (as of December 31, 2004)

Members (Alphabetical order)

GEHRING, Walter J.; Professor, Biozentrum, University of Basel

GO, Michiko; Dean & Professor, Nagahama Institute of Bio-Science and Technology

HONJO, Tasuku; Dean & Professor, Graduate School of Medicine, Kyoto University

HUNT, Tim; Principal Scientist, Cancer Research UK London Research Institute

IWATSUKI, Kunio; Professor, University of the Air

OKAZAKI, Tsuneko; Director, JSPS Stockholm Office

SAKAKI, Yoshiyuki; Director, Genomic Sciences Center, RIKEN

SUGIMURA, Takashi; President Emeritus, National Cancer Center

SULSTON, John; Former Director-General, Wellcome Trust Sanger Institute

TAKEICHI, Masatoshi; Director, Center for Developmental Biology, RIKEN

A. DEPARTMENT OF MOLECULAR GENETICS

A-a. Division of Molecular Genetics Tatsuo Fukagawa Group

RESEARCH ACTIVITIES

(1) Identification of new centromere proteins in higher vertebrate cells

Masahiro Okada, Tetsuya Hori, Mi-Sun Kwon, Mayumi Takahashi and Tatsuo Fukagawa

The centromere plays a fundamental role in accurate chromosome segregation during mitosis and meiosis in eukaryotes. Its functions include sister chromatid adhesion and separation, microtubule attachment, chromosome movement and mitotic checkpoint control. Although chromosome segregation errors cause genetic diseases including some cancers, the mechanism by which centromeres interact with microtubules of the spindle apparatus during cell division is not fully understood.

To understand the function of the centromere, we were led to develop a genetic analysis method that utilizes the hyper-recombinogenic chicken B lymphocyte cell line DT40. The high level of homologous recombination in DT40 cells permit efficient targeted disruption of genes of interest. We have improved this system and have created several cell lines with conditional knockouts of several centromere proteins to investigate the molecular mechanism of centromere assembly and function. We have proposed a model for kinetochore assembly in vertebrate cells (Fukagawa et al., EMBO J., 2001; Nishihashi et al., Dev. Cell, 2002). In this model we explained that many unidentified proteins could be involved in kinetochore assembly. Therefore, we tried to identify new kinetogore proteins using the proteomics approach. We prepared cell lines in which expression of CENP-H was replaced by expression of CENP-H-Flag or CENP-H-GFP. We also prepared cell lines in which expression of CENP-I was replaced by expression of CENP-I-Flag or CENP-I-GFP. We extracted chromosome fraction from these cell lines and performed immunoprecipitation with anti-Flag or anti-GFP antibodies. Immunoprecipitates were then separated by SDS-PAGE and we identified several common bands from all cell lines. We analyzed these bands by LC-MS/MS and identified amino acids sequence of all bands by a peptide-mass-fingerprinting method. We newly identified five proteins and cloned cDNA of these proteins. We then made expression constructs for these cDNAs fused by GFP and investigated localization of these proteins. These proteins perfectly co-localized with CENP-H throughout the cell cycle. We conclude that these five proteins are constitutive centromere proteins (manuscript in preparation). We are creating conditional knockout cells for these proteins to understand mechanism of kinetochore assembly in vertebrate cells.

(2) Molecular analysis of the Nuf2-Hec1 complex that transiently localizes to centromere during mitosis

Tetsuya Hori, Yoshikazu Mikami, Kazuko Suzuki and Tatsuo Fukagawa

Nuf2 and Hec1 are evolutionarily conserved centromere proteins. To clarify the functions of these proteins in vertebrate cells, we characterized them in chicken DT40 cells (Hori et al., J. Cell Sci., 2003). We generated GFP fusion constructs of Nuf2 and Hec1 to examine in detail localization of these proteins during the cell cycle. We found that Nuf2 is associated with Hec1 throughout the cell cycle and that this complex is localized to the centrosomes during G1 and S phases and then moves through the nuclear membrane to the centromere in G2 phase. During mitosis, this complex is localized to the centromere. We also created conditional loss-of-function mutants of Nuf2 and Hec1. In both mutants, the cell cycle arrested at prometaphase, suggesting that the Nuf2-Hec1 complex is essential for mitotic progression. The inner centromere proteins CENP-A, -C, and -H and checkpoint protein BubR1 were localized to chromosomes in the mutant cells arrested at prometaphase, but Mad2 localization was abolished. Furthermore, photobleaching experiments revealed that the Nuf2-Hec1 complex is associated stably with

the centromere and that interaction of this complex with the centrosome is dynamic.

We also observed that CENP-H, which is a constitutive centromere component that localizes to the centromere throughout the cell cycle, interacts with the Nuf2 complex by yeast two-hybrid analysis. Co-immunoprecipitation experiments revealed that CENP-H interacts with the Nuf2 complex during mitosis in chicken DT40 cells. Photobleaching experiments showed that both Hec1 and CENP-H form stable associations with the centromeres during mitosis, suggesting that Hec1 acts as a structural component of centromeres during mitosis. On the basis of these results and published data, we propose that the Nuf2 complex functions as a connector between the inner and outer kinetochores.

We started the proteomics approach to identify other proteins that interact with the Nuf2-Hec1 complex. We identified chicken homolog of Spc24 and Spc25. We also identified several components that localizes centrosome and centromere and are now characterizing these components.

(3) Functional roles of the RNAi machinery in verterbrate centromeres

Tomoko Motohashi, Masahiro Nogami, Atsushi Fukushima and Tatsuo Fukagawa

RNAi-mediated silencing of gene expression occurs when double-stranded RNAs (dsRNAs) are cleaved by Dicer into 21- to 23-nt small interfering RNAs (siRNAs). These siRNAs guide a multicomponent nuclease, RNA-induced silencing complex (RISC), to degrade specific mRNAs. Although the Dicer-mediated gene-silencing is evolutionarily conserved system, the biological functions of the RNAi machinery are not fully understood. Genetic strategies have been used to examine the biological functions of the RNAi machinery in *C. elegans, Arabidopsis, Drosophila* and fungi. There are reports that the RNAi machinery is related to chromosome segregation in fission yeast. However, it is unclear whether the RNAi machinery is associated with chromosome segregation in vertebrate cells. To examine the biological function of the RNAi-related pathway in vertebrate cells, we generated a conditional loss-of-function mutant of Dicer in a chicken-human hybrid DT40 cell line that contains human chromosome 21. Loss of Dicer leads to cell death with accumulation of abnormal mitotic cells that show premature sister chromatid separation. Aberrant accumulation of transcripts from α-satellite sequences, which consist of human centromeric repeat DNAs, was detected in Dicer-deficient cells. Immunocytochemical analysis revealed abnormalities in localization of heterochromatin proteins, Rad21 cohesin protein, and BubR1 checkpoint protein, but core kinetochore proteins such as CENP-A and -C were normal. We conclude that Dicer-related RNAi machinery is involved in formation of the heterochromatin structure in higher vertebrate cells. We also examined expression profile of non-coding region using high-density DNA-microarray, when expression of Dicer was lost. We could identify several non-coding RNAs from this method. We are characterizing biological function these RNAs. We also started to create conditional knockout cells for other components involved in RNAi machinery such as *Ago*-family proteins. We would like to comprehensive understand relationship of RNAi machinery with centromere function.

PUBLICATIONS

Papers

1. Fukagawa, T., Nogami, M., Yoshikawa, M., Ikeno, M., Okazaki, T., Takami, Y., Nakayama, T. and Oshimura M. (2004). Dicer is essential for formetion of the heterochromatin structure in vertebrate cells. Nature Cell Biol. *6,* 784-791.

2. Fukagawa, T. (2004). Centromere DNA, proteins and kinetochore assembly in vertebrate cells. Chromosome Res. *12,* 557-567.

3. Fukagawa, T. (2004). Assembly of kinetochores in vertebrate cells. Exp. Cell Res. *296,* 21-27.

EDUCATION

1. Dr. Fukagawa gave several lectures at The Graduate University for Advanced Studies, Hayama (in Japanese).

2. Dr. Fukagawa was invited a seminar on "Kinetochore Assembly" at Kyoto University, Kyoto, August, 2004 (in Japanese).

3. Dr. Fukagawa gave a public lecture of National Institute of Genetics at National Museum of Emerging Science and Innovation (MIRAIKAN), Tokyo, October, 2004.

4. Dr. Fukagawa gave a special lecture at Tokyo Institute of Technology, Yokohama, December, 2004 (in Japanese).
5. Dr. Fukagawa was invited a seminar on "Kinetochore Assembly" at Tokyo Institute of Technology, Yokohama, December, 2004 (in Japanese).

A-b. Division of Mutagenesis Fumiaki Yamao Group

RESEARCH ACTIVITIES

(1) Multifunctional role of Mcl1/Slr3 in DNA replication, repair and chromatin structure

Yasuhiro Tsutsui, Toyoaki Natsume and Fumiaki Yamao

Slr3 in fission yeast is a member of eukaryotic family of WD repeat proteins and important for genome stability. *slr3* gene was originally identified as mutants showing synthetic lethality with Rad2 defect. Fission yeast Rad2 is a FEN-1 family nuclease processing Okazaki fragment during lagging strand DNA synthesis. Since recombination activity is essential for survival of cells defective in Rad2 function, genes that show, when mutated, synthetic lethality with Rad2-defect was expected to be involved in recombination or recombination repair of damaged DNA. Actually *slr3* mutants are sensitive to DNA damage reagents such as MMS, UV, bleomycin and hydroxyurea. Null mutant cell is temperature sensitive for growth, and shows severe growth retardation even at the permissive temperature. The mutant cells arrest at G2 phase with 2C DNA content at the restricted temperature, showing that the bulk DNA synthesis occurred. However, DNA double-strand breaks were found to accumulate in the arrested cells, which was a similar phenotype caused by defect of Okazaki fragment processing in such as *dna2* or *cdc17* mutants. It was further shown that Slr3 interacts not only genetically but also physically with DNA polymerase α. These results strongly suggest that Slr3 functions in lagging strand synthesis and Okazaki fragment processing in addition to DNA repair. On the other hand, *slr3* gene was found allelic to *mcl1* that is required for proper chromosomal segregation. Concomitantly *slr3* mutant cells were sensitive to TBZ. Consequently, highly frequent chromosome loss observed in *mcl1* mutant cells was attributed to defective cohesion of sister chromatids in centromere region. Recently, we also found that Slr3 affects heterochromatin structure, which is now under an extensive investigation (Natsume et. al). Thus, Slr3/Mcl1 protein plays multifunctional role to maintain chromosomal stability by connecting DNA replication, recombination, damage repair, cohesion and heterochromatin structure. This work was collaborated with H. Shinagawa's and H. Iwasaki's group in Osaka University and Yokohama-City University, respectively.

EDUCATION

1. Dr. Yamao was invited to give a seminar on a title of "Ubiquitin as a Multi-Player in Cell" at Keihanna Interaction Plaza, Kyoto, March (in Japanese).
2. Dr. Yamao gave a lecture at Department of Veterinary Medicine, Yamaguchi University, December (in Japanese).

A-c. Molecular Mechanism Laboratory Hiroshi Mitsuzawa Group

RESEARCH ACTIVITIES

(1) Identification of glyceraldehyde-3-phosphate dehydrogenase and actin as proteins that interact with the Rpb7 subunit of RNA polymerase II

Hiroshi Mitsuzawa

Eukaryotic RNA polymerase II (pol II), the enzyme that synthesizes all mRNAs and some non-coding RNAs, plays a central role in transcription. pol II consists of evolutionarily conserved 12 subunits, designated Rpb1 to Rpb12. The catalytic activity resides in a 10-subunit core, whereas the initiation of transcription requires two additional subunits, Rpb4 and Rpb7, which form a heterodimer to associate with the core.

To gain insight into the role of the Rpb7 subunit, we performed a two-hybrid screen by using the fission yeast *Schizosaccharomyces pombe* Rpb7 protein as bait. The screen yielded glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and actin. GAPDH and actin

were also found to be associated with *S. pombe* pol II complexes. Moreover, GAPDH and actin were affinity-purified from *S. pombe* extract with an Rpb4/Rpb7-coupled column. Although GAPDH is a well-characterized glycolytic enzyme abundant in the cytoplasm, nuclear GAPDH in human cells has been identified as a key component of a coactivator complex for histone H2B transcription. In addition, it has recently been reported that actin is required for formation of the preinitiation complex. Our results are consistent with these observations and further suggest that GAPDH and actin play roles in pol II transcription through the interaction with Rpb7.

(2) Identification of the general transcription factor TFIID in *S. pombe*

Hiroshi Mitsuzawa

TFIID, one of the general transcription factors required for transcription initiation by pol II, comprises the TATA-binding protein (TBP) and a set of TBP-associated factors (TAFs). Unlike those in human, *Drosophila,* and *Saccharomyces cerevisiae,* TAFs in *S. pombe* have not been characterized until recently. We have identified five *S. pombe* TAFs biochemically: spTAF111/130 (spTAF1), spTAF72 (spTAF5), spTAF73 (spTAF5b), spTAF50 (spTAF6), and Ptr6 (spTAF7).[1,2] (Given in parentheses are new unified names proposed recently.) Of particular interest are spTAF72 and spTAF73, which are likely to regulate genes involved in progression through M phase of the cell cycle. spTAF50 was shown to interact with the WD40-repeat domain of spTAF72. These interacting TAFs were found to be shared by the TFIID and SAGA complexes. Furthermore, we carried out BLAST searches for putative TAFs, revealing that the *S. pombe* genome contains nearly all TAF genes identified in other organisms.[1,2]

PUBLICATIONS

Reviews

1. Mitsuzawa, H. and Ishihama, A. (2004). RNA polymerase II transcription apparatus in *Schizosaccharomyces pombe.* Curr. Genet. *44*, 287-294.

Books

2. Kimura, M., Mitsuzawa, H. and Ishihama, A. (2004). RNA polymerases and accessory factors. In The Molecular Biology of *Schizosaccharomyces pombe.* R. Egel, ed. (Heidelberg, Germany: Springer-Verlag), pp.329-342.

A-c. Molecular Mechanism Laboratory Hiroaki Seino Group

RESEARCH ACTIVITIES

(1) An *in vitro* ubiquitination assay of mitotic cyclin

Hiroaki Seino

Cell cycle events are regulated by sequential activation and inactivation of Cdk kinases. Mitotic exit is accomplished by the inactivation of mitotic Cdk kinase, which is mainly achieved by degradation of cyclins by a ubiquitin-proteasome system.

Previously we reported that two ubiquitin-conjugating enzymes, UbcP1/Ubc4 and UbcP4/Ubc11, were responsible for degradation of mitotic cyclin Cdc13 in fission yeast. Each of these two ubiquitin-conjugating enzymes is essential for cell viability and responsible for degradation of Cdc13. These results suggest that the functions of these two ubiquitin-conjugating enzymes are not redundant and they have distinct functions for ubiquitination of Cdc13. Furthermore, we found that ubiquitin chains of Cdc13 were totally reduced in ubcP4/ubc11 mutant cells, whereas ubiquitin chains were short and not reduced in ubcP1/ubc4 mutant cells. Thus, we proposed a hypothesis that UbcP4/Ubc11 might be involved in initiation of ubiquitination, and UbcP1/Ubc4 might be involved in elongation of ubiquitin chains of Cdc13. However, this hypothesis has not been elucidated yet.

To clarify the functional differences between UbcP1/Ubc4 and UbcP4/Ubc11 for degradation of Cdc13, development of an *in vitro* assay system for ubiquitination for Cdc13 by using fission yeast components is required. Currently, I am attempting to develop this assay system for Cdc13. A ubiquitin-activating enzyme, these two ubiquitin-conjugating enzymes and substrate Cdc13 were expressed as recombinant proteins in bacterial cells and purified. One component of Cdc13-specific ubiquitin ligase anaphase promoting complex/cyclosome (APC/C) was tagged and expressed in

fission yeast cells, and APC/C was purified from fission yeast cells. Now I am examining the conditions for reconstitution of a ubiquitination reaction of Cdc13.

(2) Ubiquitin-conjugating enzyme(s) responsible for degradation of SCF substrates

Hiroaki Seino

One of the important ubiquitin pathways responsible for cell cycle regulation is the pathway involving APC ubiquitin ligase. Another important pathway for cell cycle regulation is the pathway involving SCF (Skp1-Cullin-F-box) ubiquitin ligase. The SCF-pathway mainly regulates the progression of G1- and S-phases of the cell cycle. Furthermore, several putative SCF components have been reported to be involved in many cell regulations, e.g., a DNA damage checkpoint mechanism, maintenance of chromosomes and other regulations in fission yeast and other organisms.

However, ubiquitin-conjugating enzyme(s) responsible for degradation of substrates of SCF ubiquitin ligase in fission yeast have not been elucidated yet. We found that polyubiquitinated proteins were significantly reduced in whole cell extract from ubcP1/ubc4 mutant. It is possible that the substrates of SCF are also ubiquitinated by the UbcP1/Ubc4-pathway. Thus, I am examining the stability of the substrates of SCF ubiquitin ligase Cdc18 in ubcP1/ubc4 mutant cells, and the results are now being obtained.

A-d. Division of Nucleic Acid Chemistry Saburo Aimoto Group

RESEARCH ACTIVITIES

(1) Development of a method for chemical synthesis of protein by using two different ligation methods

Saburo Aimoto (Institute for Protein Research, Osaka University)

A thioester method and the native chemical ligation method are useful for protein synthesis. Though both methods use peptide thioesters as building blocks, no route was known to use the thioester method after the peptide bond formation reaction by the native chemical ligation method. In the thioester method, the SH groups of the cysteine residues have to be protected by a group that is stable in the presence of silver ions as an activator of thioester groups. On the other hand, the native chemical ligation method produces an SH group at the condensation site for chemoselective ligation is carried out at -X-Cys- site using the cysteine residue at the N-terminal of the C-terminal building block. In order to overcome this problem, we searched protecting groups for the thiol groups that are easily introduced to the free SH group in polypeptide under mild conditions and stable in the presence of silver ions. We examined several thiol protecting groups. Among them, thiosulfonate group was the most promising protecting group. This group, however, was revealed to partially decompose under alkaline conditions. Thus new protecting groups are still under investigation.

(2) Preparation of peptide thioesters by the Fmoc-solid phase method via an on-resin N-S acyl shift

Saburo Aimoto (Institute for Protein Research, Osaka University)

Peptide thioesters are common intermediates in contemporary methods for protein synthesis. The peptide thioesters can be directly prepared either by Boc solid-phase peptide synthesis (SPPS) or by Fmoc SPPS using Fmoc(2-F)-amino acid derivatives. Peptide thioesters can be prepared by Fmoc SPPS indirectly via a safety catch resin, too. However each preparation method based on the Fmoc solid phase method has intrinsic difficulties such as low yields of products and recemization of the C-terminal amino acid residue. Therefore an innovative method was urgently requested. A new method under development is based on an on-resin N-S acyl rearrangement. Previously, we introduced the 2-mercapto-4, 5-dimethoxybenzyl (Dmmb) group as an auxiliary for extended chemical ligation. This auxiliary group can be removed by acid treatment following the ligation reaction. Unexpectedly, it was observed that N-S acyl rearrangement occurred, in part, during the acid treatment. Base on this finding, we searched for conditions to use this mechanism for preparation of peptide thioesters. For polypeptide synthesis, the prevention of a reverse acyl rearrangement is of prime importance in this context. A

protected peptide was assembled on a resin using the Dmmb-group as a linker followed by Fmoc-based SPPS. After peptide assembly, the resin was treated with a reagent containing trifluoroacetic acid to result in the removal of the protecting groups. At the same time, the N-S acyl rearrangement occurred, resulting in resin-bound peptide thioester. Finally, free peptide thioester was obtained following treatment with 2-mercaptoethanesulfonic acid in the presence of a base and subsequent resin wash. This method permits us to prepare peptide thioesters by a Fmoc solid-phase method without racemization.

(3) Synthesis of post-translationally modified histone H3

Saburo Aimoto (Institute for Protein Research, Osaka University)

Histones can undergo posttranslational modifications such as methylation, acetylation or/and phosphorylation. These modifications can play a role in gene regulation in an epigenetic manner. Focused on the N-terminal region of histone H3 we searched a strategy that would provide a synthetic procedure for post-translationally modified histones. First, synthesis of the N-terminal region of histone H3 peptides that contained Lys(Me_3) was examined by Boc and Fmoc SPPS. A desired product was obtained by the both methods though the Boc method gave better results than the Fmoc SPPS. Lys(Me_3)-containing peptide thioesters that were building blocks for histone H3 synthesis were prepared by Boc SPPS. Furthermore, two chemically synthesized building blocks covering [Lys(Me_3)^{4}]-histone H3(1-12) and [Lys(Me_3)^{36}]-histone H3(13-43) were condensed by the thioester method to give [Lys(Me_3)$^{4, 36}$]-histone H3(1-43). Along with this totally chemical synthetic method, we are developing a method that uses a biologically expressed histone H3 segment for the total synthesis of modified histone H3. In the near future, a variety of modified histone H3 with a full sequence will be synthesized by condensing chemically synthesized modified-amino-acid-containing peptide thioesters and biologically expressed histone H3 segments.

(4) Synthetic studies of G protein-coupled receptor, opioid receptor like-1 (ORL-1)

Saburo Aimoto (Institute for Protein Research, Osaka University)

About 30% of the human genome encodes membrane proteins. Among them G protein-coupled receptor family is the largest family of membrane receptor protein and is the target of most pharmaceuticals. Much of the information concerning the structure and function of these membrane proteins, however, remains to be uncovered because of the difficulties associated with biochemical sample preparation. As an alternative approach to obtaining membrane proteins, chemical synthesis represents a viable candidate. ORL-1 is a receptor of an opioid peptide, nociceptin. The strategy that is employed for the synthesis of the C-terminal region of ORL-1 is the combination of the native chemical ligation method and the thioester method. The first coupling between ORL-1(288-328) and ORL-1(329-370) was carried out in SDS solution by the native chemical ligation method. The coupling reaction conditions were thoroughly inspected in terms of pH of the solution, chemical composition of the reaction media. According to the defined conditions, ORL-1(288-370) that contained one transmembrane region and the C-terminal cytosolic tail region was obtained in the yield of 47%.[4] We are continuously searching the route to accomplish the total synthesis of ORL-1.

(5) Design of the inhibitors to human T-cell leukemia virus type-1 protease

Saburo Aimoto (Institute for Protein Research, Osaka University)

Human T-cell leukemia virus type 1 (HTLV-1), a retrovirus associated with a number of human diseases, was the first human retrovirus isolated from patients with adult T-cell leukemia/ lymphoma by Gallo et al. An HTLV-1 gene codes an asparatic protease (PR), which processes its own polyproteins, which are transcripted owing to three reading frames. As the result of a series of cis processing, a set of proteins is produced, which are necessary for viral replication. Thus, HTLV-1 PR plays a key role in the duplication of HTLV-1 in a manner analogous to the

human immunodeficiency virus type-1 protease in acquired immunodeficiency syndrome. In the design of potent protease inhibitors for this virus, the knowledge of the characteristics of HTLV-1 PR itself and its substrate specificities is critical. Then, we synthesized HTLV-1 PR and examined its substrate specificities. Based on the obtained data we designed HTLV-1 protease inhibitors of an olefin-containing cyclic peptide.[1)]

PUBLICATIONS

Papers

1. Bang, J.K., Hasegawa, K., Kawakami, T., Aimoto, S. and Akaji, K. (2004). Synthesis of an olefin-containing cyclic peptide using the solid-phase Honer-Emmons reaction. Tetrahedron Lett., *45,* 99-102.
2. Sekine, S., Kataoka, K., Tanaka, M., Nagata, H., Kawakami, T., Akaji, K., Aimoto, S. and Shizukuishi, S. (2004). Active domains of salivary statherin on apatitic surfaces for binding to Fusobacterium nucleatum cells. Microbiology. *150,* 2373-9.
3. Yamada, H., Sasaki, T., Niwa, S., Oishi, T., Murata, M., Kawakami, T. and Aimoto, S. (2004). Intact Glycation End Products Containing Carboxylmethyl-Lysine and Glyoxal Lysine Dimer Obtained from Synthetic Collagen Model Peptide, Bioorg. Med. Chem. Lett., *14,* 5677-80.
4. Sato, T., Saito, Y. and Aimoto, S. (2004). Synthesis of the C-terminal Region of Opioid Receptor Like 1 in an SDS Micelle by the Native Chemical Ligation: Effect of Thiol Additive and SDS Concentration on Ligation Efficiency. J. Peptide Sci., Epub. Dec. 20.
5. Onoda, A., Yamamoto, H., Yamada, Y., Lee, K., Adachi, S., Okamura, T. A., Yoshizawa-Kumagaye, K., Nakajima, K., Kawakami, T., Aimoto, S. and Ueyama, N. (2005). Switching of turn conformation in an aspartate anion peptide fragment by NH . . . O(-) hydrogen bonds. Biopolymers, Epub., Jan 4.

EDUCATION

1. Dr. S. Aimoto was invited to give a lecture at Kyoto University, January, 2005.

SOCIAL CONTRIBUTIONS AND OTHERS

1. Dr. S. Aimoto organized a joint symposium entitled "Protein Researches in Bioscience and Bioengineering", between Institute for Protein Research, Osaka University and Seoul National University, Osaka, June, 2004.
2. 相本三郎，日本ペプチド学会評議員.

A-d. Division of Nucleic Acid Chemistry Tsutomu Katayama Group

RESEARCH ACTIVITIES

(1) Molecular mechanism of DNA replication-coupled inactivation of the initiator protein in *Escherichia coli:* Interaction of DnaA with the sliding clamp-loaded DNA and the sliding clamp-Hda complex

Masayuki Su'etsugu, Makoto Takata[1], Toshio Kubota[2], Yusaku Matsuda and Tsutomu Katayama (Present addresses: [1]Sumitomo Pharmaceuticals Co., Osaka, Japan; [2]Department of Pharmacy, Kyushu University Hospital, Fukuoka, Japan)

In *Escherichia coli,* the ATP-DnaA protein initiates chromosomal replication. After the DNA polymerase III holoenzyme is loaded onto DNA, DnaA-bound ATP is hydrolyzed in a manner depending on Hda protein and the DNA-loaded form of the DNA polymerase III sliding clamp subunit, which yields ADP-DnaA, an inactivated form for initiation. This regulatory DnaA-inactivation represses extra initiation events. In this study, *in vitro* replication intermediates and structured DNA mimicking replicational intermediates were first used to identify structural prerequisites in the process of DnaA-ATP hydrolysis. Unlike duplex DNA loaded with sliding clamps, primer RNA-DNA heteroduplexes loaded with clamps were not associated with DnaA-ATP hydrolysis, and duplex DNA provided *in trans* did not rescue this defect. At least 40-bp duplex DNA is competent for the DnaA-ATP hydrolysis when a single clamp was loaded. The DnaA-ATP hydrolysis was inhibited when ATP-DnaA was tightly bound to a DnaA box-bearing oligonucleotide. These results imply that the DnaA-ATP hydrolysis involves the direct interaction of ATP-DnaA with duplex DNA flanking the sliding clamp. Furthermore, Hda protein formed a stable complex with the sliding clamp. Based on these, we suggest a mechanical basis in the DnaA-inactivation that ATP-DnaA interacts with the Hda-clamp complex with the aid of DNA binding.[4)]

(2) DiaA, a novel DnaA-binding protein, ensures the timely initiation of *E.coli* chromosome replication

Takuma Ishida, Nobuyoshi Akimitsu[1], Tamami Kashioka[2], Masakazu Hatano[3], Toshio Kubota[4], Yasuyuki Ogata[5], Kazuhisa Sekimizu[1] and Tsutomu Katayama (Present addresses: [1]Graduate School of Pharmaceutical Sciences, University of Tokyo, Japan; [2]Institute for Chinese Medicine, Nakamura City Hospital, Kochi, Japan; [3]Santen Pharmaceutical Co., Nara, Japan; [4]Department of Pharmacy, Kyushu University Hospital, Fukuoka, Japan; [5]Radioisotope Center, National Institute of Genetics, Mishima, Shizuoka, Japan)

The DnaA protein is the initiator of *E. coli* chromosomal replication. In this study, we identify a novel DnaA-associating protein, DiaA, which is required for the timely initiation of replication during the cell cycle. DiaA promotes the growth of specific temperature-sensitive *dnaA* mutants and ensures stable minichromosome maintenance, while DiaA does not decrease the cellular DnaA content. A *diaA*::Tn*5* mutation suppresses the cold-sensitive growth of an overinitiation-type *dnaA* mutant independently of SeqA, a negative modulator of initiation. Flow cytometry analyses reveal that the timing of replication initiation is disrupted in the *diaA* mutant cells as well as wild-type cells with pBR322 expressing the *diaA* gene. Gel-filtration and chemical cross-linking experiments show that purified DiaA forms a stable homodimer. Immunoblotting analysis indicates that a single cell contains about 280 DiaA dimers. DiaA stimulates minichromosome replication in an *in vitro* system especially when the level of DnaA included is limited. Moreover, specific and direct binding between DnaA and DiaA is observed, which requires a DnaA N-terminal region. DiaA binds to both ATP- and ADP-bound forms of DnaA with a similar affinity. Thus, we conclude that DiaA is a novel DnaA-associating factor that is crucial to ensure the timely initiation of chromosomal replication.[2]

(3) Reactivation of DnaA by DNA sequence-specific nucleotide exchange *in vitro*

Kazuyuki Fujimitsu and Tsutomu Katayama

In *Escherichia coli,* ATP-bound DnaA protein can initiate chromosomal replication. After initiation, DnaA-ATP is hydrolyzed by interactions with a complex containing a replicase subunit to yield the inactive ADP-DnaA. However, the mechanisms which regenerate ATP-DnaA from ADP-DnaA are not well understood. We report here that a 70-bp DNA segment promotes exchange of the DnaA-bound nucleotide in a sequence specific manner, thus reactivating the initiation function of DnaA *in vitro.* This segment contains a typical DnaA-binding 9-mer motif, the DnaA box, and two DnaA box-like sequences. The presence and precise composition of these three motifs are required for the DnaA-reactivating activity, which suggests that a highly ordered complex which includes multimeric DnaA molecules is formed for isomerization of DnaA. We named this DNA segment DARS, for DnaA-reactivating sequence. The role of DARS in regulation of DnaA function *in vivo* is discussed.[1]

(4) Novel heat shock protein HspQ stimulates the degradation of mutant DnaA protein in *Escherichia coli*

Toh-ru Shimuta, Kiyotaka Nakano[1], Yoko Yamaguchi, Shogo Ozaki, Kazuyuki Fujimitsu, Chika Matsunaga, Kenji Noguchi[2], Akiko Emoto[3] and Tsutomu Katayama (Present addresses: [1]Chugai Pharmaceutical Co., Tokyo, Japan; [2]Kaketsuken Institute, Kumamoto, Japan; [3]Department of Pharmacy, Saga University Hospital, Saga, Japan)

Escherichia coli DnaA protein initiates chromosomal replication and is an important regulatory target during the replication cycle. In this study, a suppressor mutation isolated by transposon mutagenesis was found to allow growth of the temperature-sensitive *dnaA508* and *dnaA167* mutants at 40°C. The suppressor consists of a transposon insertion in a previously annotated ORF, here termed *hspQ,* a novel heat shock gene whose promoter is recognized by the major heat shock sigma factor σ^{32}. Expression of *hspQ* on a pBR322 derivative inhibits growth of the *dnaA508* and *dnaA167* mutants at 30°C, whereas growth of *dnaA46* and other *dnaA* mutants is insensitive to changes in the level of *hspQ.* Cellular DnaA508 protein is degraded rapidly at elevated temperature, but *hspQ* disruption impedes this process. In contrast, DnaA46 protein is rapidly degraded in an *hspQ*-independent manner. Gel-filtration and chemical cross-linking experiments suggest that HspQ forms a

stable homodimer in solution and can form homomultimers consisting of about four monomers. Heat shock-induced proteases such as Clp contain homomultimers of subunit. We propose that HspQ is a new factor involved in the quality control of proteins and that it functions by excluding denatured proteins.[3)]

PUBLICATIONS

Papers

1. Fujimitsu, K. and Katayama, T. (2004). Reactivation of DnaA by DNA sequence-specific nucleotide exchange *in vitro*. Biochem. Biophys. Res. Commun., *322,* 411-419.

2. Ishida, T., Akimitsu, N., Kashioka, T., Hatano, M., Kubota, T., Ogata, Y., Sekimizu, K. and Katayama, T. (2004). DiaA, a novel DnaA-binding protein, ensures the initiation timing of *Escherichia coli* chromosome replication. J. Biol. Chem., *279,* 45546-45555.

3. Shimuta, T., Nakano, K., Yamaguchi, Y., Ozaki, S., Fujimitsu, K., Matsunaga, C., Noguchi, K., Emoto, A. and Katayama, T. (2004). Novel heat-shock protein HspQ stimulates the degradation of mutant DnaA protein in *Escherichia coli.* Genes Cells, *9,* 1151-1166.

4. Su'etsugu, M., Takata, M., Kubota, T., Matsuda, Y. and Katayama, T. (2004). Molecular mechanism of DNA replication-coupled inactivation of the initiator protein in *Escherichia coli:* Interaction of DnaA with the sliding clamp-loaded DNA and the sliding clamp-Hda complex. Genes Cells, *9,* 509-522.

Reviews

5. 片山　勉(2004)「大腸菌染色体複製の開始とその制御に関わるAAA+スーパーファミリータンパク質」生化学 *76,* pp.1440-1443.

Books

6. 片山　勉，末次正幸，川上広宣(2004)「大腸菌ゲノムDNA複製系」，ゲノミクスとプロテオミクスの新展開～生物情報の解析と応用～(エヌ・ティー・エス，今中忠行監修) pp.57-64.

SOCIAL ACTIVITIES AND OTHERS

1. Dr. Katayama was awarded a Junior Research Leader of Kyushu University.

B. DEPARTMENT OF CELL GENETICS

B-a. Division of Cytogenetics Tamotsu Yoshimori Group

RESEARCH ACTIVITIES

Tamotsu Yoshimori

Most of membrane-bound organelles inside eukaryotic cells are linked each other by dynamic membrane trafficking regulated by a set of specific proteins. Membrane traffic is essential not only to survival of each cell but also to various functions organizing the multi-cellular system, e.g., formation of cell polarity and intercellular communication. We aim to unravel molecular mechanisms of membrane traffic and their roles in physiological functions and diseases in animals, which must produce knowledge contributing clinical medicine. We are now focusing on two trafficking routes; autophagy and the endosomal system. Autophagy is membrane traffic delivering cytoplasmic components to lysosomes for bulk degradation. The process is mediated by the formation of the double membrane-bound autophagosomes. Endosomes receive macromolecules taken up by endocytosis from outside. The cargo is then either sorted to lysosomes or recycled back to the plasma membrane.

(1) Three mammalian homologues of yeast Atg8 localize to autophagosomal membrane depending on post-translational processing

Yukiko Kabeya[1], Noboru Mizushima[2], Akitsugu Yamamoto[3], Satsuki Oshitani-Okamoto[1], Yoshinori Ohsumi[1] and Tamotsu Yoshimori ([1]National Institute for Basic Biology, [2]Tokyo Metropolitan Institute of Medical Science, [3]Nagahama Institute of Bio-Science and Technology)

We previously reported that rat LC3, a homologue of yeast Atg8, localizes to autophagosomal membranes after post-translational modifications. The C-terminal fragment of LC3 is cleaved immediately following synthesis to yield a cytosolic form, LC3-I. A subpopulation of LC3-I is further converted to an autophagosome-associating form, LC3-II by the ubiquitylation-like conjugation system. In this study, we showed that [^{14}C]-ethanolamine was preferentially incorporated into LC3-II, suggesting that LC3-II is a phosphatidylethanolamine (PE)-conjugated form like a membrane-binding form of yeast Atg8[1]. LC3-II can be a substrate of mammalian Atg4B, a homologue of yeast Atg8-PE deconjugase, supporting the idea that LC3-II is LC3-PE.

There are at least two other homologues of yeast Atg8, γ-aminobutyric-acid-type-A-receptor-associated protein (GABARAP) and Golgi-associated ATPase enhancer of 16kDa (GATE16) in mammals. We found that they also generate form II and its generation correlated with autophagosomal association of GABARAP and GATE16[1]. The results suggest that all mammalian Atg8 homologues receive a common modification to associate with autophagosomal membrane as the form II. GABARAP and GATE16 have been suggested to be involved in the $GABA_A$-receptor transport and an intra-Golgi transport, respectively. However, it remains a possibility that they participate in autophagy in addition to or instead of their functions originally described.

(2) Autophagy plays a pivotal role in survival during the early neonatal starvation period

Akiko Kuma[1], Masahiko Hatano[2], Makoto Matsui[1], Akitsugu Yamamoto[3], Haruaki Nakaya[2], Tamotsu Yoshimori, Yoshinori Ohsumi[4], Takeshi Tokuhisa[2] and Noboru Mizushima[1] ([1]Tokyo Metropolitan Institute of Medical Science, [2]Chiba University, [3]Nagahama Institute of Bio-Science and Technology, [4]National Institute for Basic Biology)

At birth the trans-placental nutrient supply is suddenly interrupted, and neonates face severe starvation until supply can be restored through milk nutrients. We showed that neonates adapt to this adverse circumstance by inducing autophagy to degrade cytoplasmic constituents within lysosomes[6]. The level of autophagy in mice, which was monitored

by GFP-LC3 expressed in the transgenic mice[1], remains low during embryogenesis; however, autophagy is immediately upregulated in various tissues after birth and is maintained at high levels for 3-12 h before returning to basal levels within 1-2 days. Mice deficient for Atg5, which is essential for autophagosome formation, appear almost normal at birth but die within 1 day of delivery. The survival time of starved Atg5-deficient neonates (~12 h) is much shorter than that of wild-type mice (~21 h) but can be prolonged by forced milk feeding. Atg5-deficient neonates exhibit reduced amino acid concentrations in plasma and tissues, and display signs of energy depletion. These results suggest that the production of amino acids by autophagic degradation of 'self' proteins, which allows for the maintenance of energy homeostasis, is important for survival during neonatal starvation.

(3) Autophagic machinery provides a cellular defense system against invasion by group A *Streptococcus*

Ichiro Nakagawa[1], Takahiro Kamimoto, Akitsugu Yamamoto[2], Noboru Mizushima[3], Kayoko Tsuda, Atsuki Nara, Junko Funao[1], Masanobu Nakata[1], Atsuo Amano[1], Shigeyuki Hamada[1] and Tamotsu Yoshimori ([1]Osaka University, [2]Nagahama Institute of Bio-Science and Technology, [2]Tokyo Metropolitan Institute of Medical Science)

A Gram-positive bacteria, group A *Streptococcus* (GAS) is the etiological agent for a diverse collection of human diseases. GAS invades non-phagocytic cells, but the destination of GAS after internalization is not well understood. We found that the autophagic machinery could effectively eliminate GAS within non-phagocytic cells[7]. GAS that had invaded HeLa cells or MEF was first trapped by endosomes and then successfully escaped to the cytoplasm using bacterial cytolysin SLO. Then, the cytoplasmic GAS was selectively sequestered by large, unique compartments bearing LC3, the formation of which was specifically induced by bacterial infection and required the Atg5 gene. The numbers of living GAS decreased to less than 10% at 4 h of post-infection and this reduction was completely suppressed in autophagy deficient $Atg5^{-/-}$ cells. In these cells, GAS survived, multiplied, and were released from the cells. The compartments trapping GAS eventually fused with lysosomes, in which enzymes degraded and killed the GAS. The autophagic machinery could only recognize GAS that had escaped from endosomes to the cytoplasm. Based on these results, we concluded that the autophagic machinery can act as an innate defence system against invading pathogens.

(4) Intracellular *Shigella* escapes from Autophagy

Michinaga Ogawa[1], Tamotsu Yoshimori, Toshihiko Suzuki[1], Hiroshi Sagara[1], Noboru Mizushima[2] and Chihiro Sasakawa[1] ([1]University of Tokyo, [2]Tokyo Metropolitan Institute of Medical Science)

Shigella is a group of Gram-negative bacteria causing Shigellosis. The invasion of Shigella into the colonic epithelium initiates the disease. IcsB, one of the *Shigella flexneri* effectors, is secreted via the type III secretion system of cytoplasmic bacteria and located on the bacterial surface. The *icsB* mutant is fully invasive, but defective in spreading within host cells. To clarify the role of IcsB in promoting infection, we investigated the intracellular behaviors of the *icsB* mutant[8]. As a result, we found that the mutant bacteria were trapped by autophagy in the cytoplasm. IcsB did not directly inhibit autophagy. Rather, *Shigella* VirG, a protein required for intracellular actin-based motility, induced autophagy by binding to Atg5. In non-mutant *Shigella,* this binding is competitively inhibited by IcsB binding to VirG. Thus, in contrast to GAS, *Shigella* in the cytoplasm is able to escape from attack of autophagy by secreting IcsB which act as an inhibitor of recognition of the bacteria by Atg5.

(5) The intracellular inclusions containing mutant α_1-antitrypsin Z are propagated in the absence of autophagic activity

Takahiro Kamimoto, Shisako Shoji, Noboru Mizushima[1], Kyohei Umebayashi, Tunda Hidvegi[2], David H. Perlmutter[2] and Tamotsu Yoshimori ([1]Tokyo Metropolitan Institute of Medical Science, [2]University of Pittsburgh)

Formation of intracellular inclusions comprised of terminally misfolded proteins is the most obvious hallmark of diseases that are collectively termed conformational disease. Mutant α_1-Antitrypsin Z (α_1-ATZ) protein accumulates within the endoplasmic

reticulum (ER) of the liver cells as an aggregated polymer and is associated with the development of chronic liver injury and hepatocellular carcinoma in hereditary α_1-antitrypsin (α_1-AT) deficiency. Previous studies have suggested that efficient intracellular degradation of α_1-ATZ is correlated with protection from liver disease in α_1-AT deficiency, and that the ubiquitin proteasome system accounts for a major, but not sole, route of α_1-ATZ disposal. Yet autophagy has also been implicated in the pathophysiology of α_1-AT deficiency. In this study, to provide genetic evidence that autophagy can mediate disposal of α_1-ATZ, we used autophagy deficient Atg5$^{-/-}$ cells. The results showed that in the absence of autophagy, degradation of α_1-ATZ was retarded and that α_1-ATZ accumulated in characteristic cellular inclusions colocalized with calnexin and ubiquitin. These data provide definitive evidence that autophagy can participate in the quality control/degradative pathway for α_1-ATZ and suggest that autophagic degradation plays a fundamental role in preventing the formation of characteristic cytoplasmic inclusions.

(6) Analysis of invasion of *Porphyromonas gingivalis* into cells by using beads coated with the bacterial vesicles

Kayoko Tsuda, Ichiro Nakagawa[1], Kyohei Umebayashi, Atsuo Amano[1] and Tamotsu Yoshimori ([1]Osaka University)

Porphyromonas gingivalis, Gram-negative anaerobic bacterium, is considered to be a bona fide pathogen of adult periodontitis, which is the most prevalent chronic disease among human over the world. *P. gingivalis* can internalize into primary gingival epithelial cells and other cell types. To clarify the mechanisms underlying the invasion and itinerary of the pathogen within host cells, we have used fluorescent beads coated with vesicles secreted by *P. gingivalis.* The beads efficiently entered into cells through membrane trafficking, while the internalization did not occur at all when beads were coated with BSA or the heat-inactivated vesicles. We have been investigating cellular components involved in the beads internalization and fate of the internalized beads.

(7) Reconstitution of initiation of autophagosome formation by using semi-intact cells

Shunsuke Kimura, Atsuki Nara, Yoshitaka Nagai[1] and Tamotsu Yoshimori ([1]Osaka University)

In vitro reconstitution of membrane trafficking by using the semi-intact cell systems is a powerful tool to resolve its molecular machinery. To elucidate mechanisms underlying autophagic membrane dynamics, we established semi-intact cells by treatment of cultured cells with a bacterial toxin, Streptolysin O, which forms micro pores on the plasma membranes and allow us to access the cytoplasm directly. We succeeded to reconstitute formation of the small vesicle decorated with GFP-Atg5, which is a precursor of autophagosome in the semi-intact cells by adding the cytosol fraction and ATP-regenerating system. The cytosol isolated from starved cells was more effective than that from cells cultured in nutrient-rich condition. Since autophagy is known to be induced by starvation, the result indicates that the starved cell cytosol contains factor(s) triggering autophagosome formation. Moreover, we found that adding of the recombinant protein containing expanded polyglutamine fragment, which causes a class of inherited neurodegenerative diseases, so-called polyglutamine diseases, including Huntington's disease, induced autophagy in this system. We are screening the cellular molecules recognizing the starvation signal or the expanded polyglutamine fragment to induce autopagy.

(8) Ubiquitin-dependent sorting of the epidermal growth factor receptor in the endocytic pathway

Kyohei Umebayashi and Tamotsu Yoshimori

When plasma membrane proteins are ubiquitinated, they follow the endocytic pathway to lysosomes. In the case of the epidermal growth factor receptor (EGFR), the Cbl E3 ligase is responsible for ligand-dependent ubiquitination. It has been considered that the ubiquitination of the receptor occurs in the plasma membrane, however, we found that Cbl is localized to early endosomes when cells are stimulated with EGF. Early endosomes are composed of distinct membrane domains, which may represent various transport directions from this organelle. Hrs sorts ubiquitinated cargoes to lysosomes, and is localized

to specific domains of early endosomal membrane. The localization patterns of Cbl and Hrs were overlapped very well, suggesting that Cbl is localized to early endosomal subdomains where cargoes destined for lysosomes are concentrated. It is known that expression of a dominant negative form Cbl(C381A) abolishes the ubiquitination of EGFR. Strikingly, we found that Cbl(C381A) does not inhibit the internalization of EGF from the plasma membrane. EGF could reach endosomes where it was colocalized with Cbl(C381A). Thus, ubiquitination of EGFR is not obligatory for the internalization process. Both the ubiquitination and the ubiquitin-dependent sorting of EGFR may occur in early endosomal subdomains. Tracking the localization after EGF stimulation suggested that Cbl binds EGFR in the plasma membrane, remains associated in early endosomes, and then becomes separated from the receptor. We have obtained an implication that an AAA ATPase SKD1 regulates the ubiquitinated status of EGFR, and will investigate the molecular mechanisms in detail.

PUBLICATIONS

Papers

1. Mizushima, N., Yamamoto, A., Matsui, M., Yoshimori, T. and Ohsumi, Y. (2004). *In vivo* analysis of autophagy in response to nutrient starvation using transgenic mice expressing a fluorescent autophagosome marker. Mol. Biol. Cell *15,* 1101-1111.

2. Prentice, E.W., Jerome, G., Yoshimori, T., Mizushima, N. and Denison, M.R. (2004). Coronavirus Replication Complex Formation Utilizes Components of Cellular Autophagy. J. Biol. Chem. *279,* 10136-10141.

3. Kabeya, Y., Mizushima, N., Oshitani-Okamoto, S., Ohsumi, Y. and Yoshimori, T. (2004). LC3, GABARAP and GATE16 localize to autophagosomal membrane depending on form-II formation. J. Cell Sci. *117,* 2805-2812.

4. Fujita, H., Umezuki, Y., Imamura, K., Ishikawa, D., Uchimura, S., Nara, A., Yoshimori, T., Hayashizaki, Y., Kawai, J., Ishidoh, K., Tanaka, Y. and Himeno, M. (2004). Mammalian class E Vps proteins, SBP1 and mVps2/CHMP2A, interact with and regulate the function of an AAA-ATPase SKD1/Vps4B. J. Cell Sci. *117,* 2997-3009.

5. Birkeland, H.C.G., Simonsen, A., Gillooly, D.J., Mizushima, N., Kuma, A., Yoshimori, T., Slagsvold, T., Brech, A. and Stenmark, H. (2004). Alfy, a novel FYVE domain-containing protein associated with protein granules and autophagic membranes. J. Cell. Sci. *117,* 4239-4251.

6. Kuma, A., Hatano, M., Matsui, M., Yamamoto, A., Nakaya, H., Yoshimori, T., Ohsumi, Y., Tokuhisa, T. and Mizushima, N. (2004). Role of autophagy during the early neonatal starvation period. Nature *432,* 1032-1036.

7. Nakagawa, I., Amano, A., Mizushima, N., Yamamoto, A., Yamaguchi, H., Kamimoto, T., Nara, A., Funao, J., Nakata, M., Tsuda, K., Hamada, S. and Yoshimori, T. (2004). Autophagy defenses cells against invading group A *Streptococcus.* Science *306,* 1037-1040.

8. Ogawa, M., Yoshimori, T., Suzuki, T., Sagara, H., Mizushima, N. and Sasakawa, C. (2004). Escape of Intracellular *Shigella* from Autophagy. Science published online 2 *December* [DOI: 10.1126/science.1106036].

9. Nakatsukasa, K., Okada, S., Umebayashi, K., Fukuda, R., Nishikawa, S. and Endo, T. (2004). Roles of *O*-mannosylation of aberrant proteins in reduction of the load for endoplasmic reticulum chaperones in yeast. J. Biol. Chem. *279,* 49762-49772.

Reviews

10. Yoshimori, T. (2004). Autophagy: a regulated bulk degradation process inside cells Biochem. Biophys. Res. Commun. *313,* 453-458.

11. 吉森　保(2004)蛋白質大規模分解システムとしてのオートファジー：明らかになってきた多彩な生理機能. 蛋白質核酸酵素増刊号「細胞における蛋白質の一生」*49,* 1029-1032.

12. 吉森　保(2004)オートファジーと疾患. 医学のあゆみ*211,* 147-151.

13. 梅林恭平(2004)ステロールとユビキチン—ポストゴルジでのタンパク質選別輸送における役割—日本農芸化学会誌*78,* 39-41.

EDUCATION

1. Dr. T. Yoshimori gave a lecture at Division of Pharmaceutical Science, Graduate School of Natural Science and Technology, Kanazawa University, January, 2004 (in Japanese).

2. Dr. T. Yoshimori gave a lecture at Kakegawa-nishi High School, Shizuoka-ken, March, 2003 (in Japanese).

3. Dr. T. Yoshimori gave a lecture at Graduate School of Integrated Science, Yokohama City University, March, 2004 (in Japanese).

4. Dr. T. Yoshimori was invited to give a seminar at Institute for Frontier Medical Science, Kyoto University, March, 2004 (in Japanese).
5. Dr. T. Yoshimori gave a lecture at Graduate School of Pharmaceutical Science, The University of Tokyo, July, 2004 (in Japanese).
6. Dr. T. Yoshimori was invited to give a seminar at Geneva University Sciences II, Geneva, September, 2004.
7. Dr. T. Yoshimori was invited to give a seminar at The Institute of Medical Science, The University of Tokyo, September, 2004 (in Japanese).
8. Dr. K. Umebayashi was invited to give a seminar at the Norwegian Radium Hospital, Oslo, September, 2004.

SOCIAL CONTRIBUTIONS AND OTHERS

1. Dr. T. Yoshimori was appointed for a councillor of the Japan Society for Cell Biology.
2. Dr. T. Yoshimori was appointed for an associate editor of "Cell Structure and Function".
3. Dr. T. Yoshimori was appointed for a member of editorial board of "Autophagy".
4. Dr. T. Yoshimori organized a symposium "Sorting and Selection in Membrane Traffiking" at The 57th Annual Meeting of Japan Society for Cell Biology, Osaka, May, 2004.

B-b. Division of Microbial Genetics Hiroyuki Araki Group

RESEARCH ACTIVITIES

We have been studying on eukaryotic chromosomal DNA replication and its regulation by the cell cycle. For this purpose, we have employed budding yeast, *Saccharomyces cerevisiae,* as a model system of eukaryotic cells. Using strong genetics of budding yeast, we have identified novel factors involving in chromosomal DNA replication and revealed their functions in chromosomal DNA replication. We have been also interested in chromosome instability caused by replication defect.

(1) The interaction between replication proteins for the initiation of DNA replication

Sachiko Sakamoto, Kazuyuki Hirai, Yoichiro Kamimura and Hiroyuki Araki

At replication origins in eukaryotes, the pre-RC (pre-Replicative Complex) forms from late M to G1 phases and other replication proteins including DNA polymerases assemble when CDK activity increases from G1/S boundary. We have reported three yeast complexes, Sld2-Dpb11, Sld3-Cdc45 and GINS, all of which associate with origins in a mutually dependent and the pre-RC dependent manner. Dpb11 has four BRCT repeats and form a complex with the Sld2 protein phosphorylated by CDK. The Sld3-Cdc45 complex is observed throughout the cell cycle. The GINS complex consisting of Sld5, Psf1, Psf2 and Psf3 subunits first associates with origins and then moves with replication forks.

To elucidate how the proteins assemble, we investigated the complex formation between GINS and other proteins by co-immunoprecipitation assay. GINS coprecipitated with Pol ε throughout the cell cycle and with Mcm, a component of the pre-RC, in the S phase. Moreover, when we treated the cells with the cross-linking agent before precipitation, Sld3, Sld2 and Dpb11 also coprecipitated with GINS. Even when S-CDK activity increased without the pre-RC in Cdc6-depleted cells, Sld2-Dpb11 and Pol ε coprecipitated with GINS, suggesting that they form a complex without origin-association. We therefore propose the pre-Landing Complex (pre-LC) containing Sld2, Dpb11, Pol ε and GINS, which is formed before associating with origins. The pre-LC seems to associate with the pre-RC via Sld3 since the C-terminal portion of Sld3 interacts with GINS whereas the N-terminal portion interacts with Cdc45. The pre-LC complex is observed only when CDK activity increases. Since the complex formation between Sld2 and Dpb11 depends on CDK-phosphorylation of Sld2 (see below), CDK regulates the pre-LC formation through the Sld2-Dpb11 complex formation.

In parallel with the in vivo analysis, we have tried to purify the proteins related to the pre-LC to know their function. First, we expressed all the subunits of GINS in insect cells and purified GINS to be a near homogeneity. We also expressed Dpb11 and Sld3 in *Escherichia coli* and partially purified them. The

purified GINS complex binds to the purified Dpb11 and Sld3, consistent with two hybrid assay and in vivo co-immunoprecipitation. We will further extend this analysis to all the components of the pre-LC.

(2) Regulatory mechanism of the complex formation between Sld2 and Dpb11 by cyclin-dependent kinase (CDK)

Yon-Soo Tak, Yoshimi Tanaka, Yoichiro Kamimura and Hiroyuki Araki

The initiation of chromosomal DNA replication in eukaryotes requires the CDK activity. However, how CDK regulates the initiation of DNA replication is not well elucidated since the substrates of CDK in DNA replication have not been identified except the Sld2 protein. Sld2 has a cluster of eleven CDK-dependent phosphorylation sites (S/T-P), six of which are preferred CDK phosphorylation sites. Phosphorylation of Sld2 enhanced its binding to Dpb11, which appears to be necessary for onset of DNA replication. We have studied how the complex formation between Dpb11 and Sld2 is regulated by CDK activity. Dpb11 has four copies of the BRCT domain, which is recently reported as a phospho-peptide binding module. A C-terminal pair of BRCT domains of Dpb11 binds to a short stretch of Sld2.

We purified the truncated Sld2 containing the Dpb11-binding stretch and 11-CDK phosphorylation sites and the C-terminal pair of BRCT domains of Dpb11. Using these purified proteins, we demonstrated that CDK-phosphorylation enhances the complex formation between Sld2 and Dpb11. To this in vitro binding reaction, we challenged various synthesized peptides with or without phosphorylation. It revealed that the 20-aa stretch with phosphorylation at threonine followed by proline (CDK-catalyzed phosphorylation site) competes with the complex formation between phosphorylated Sld2 and Dpb11. Thus, the 20-aa stretch with the phosphorylation functions as a binding stretch to Dpb11. Consistent with these observations, an alanine substitution of this threonine reduced the binding activity to Dpb11 in in vitro binding assay as well as two-hybrid assay. Furthermore, this alanine-substitution confers defect of cell growth.

We previously reported that the simultaneous alanine substitutions of all the canonical (preferred) CDK-phosphorylation sites of Sld2 confers defect of cell growth as well as reduced affinity to Dpb11 (Masumoto et al., Nature 415, 651, 2002). The 20-aa stretch does not contain these canonical phosphorylation sites. Moreover, the phosphomimetic replacement of this threonine by asparagine overrides the defect of the simultaneous substitutions of canonical phosphorylation-sites in a two-hybrid assay. These observations suggest that phosphorylation of canonical CDK-sites affects the phosphorylation of the threonine in the 20-aa stretch.

CDK is consisted of a catalytic subunit and a cyclin. Budding yeast has one catalytic subunit, Cdc28 and nine different cyclins, and thus 9 species of CDK. We expressed these CDK subunits in *E. coli* and purified typical CDK species. Using these purified CDKs, we found that Cdc28-Clb5, S-phase CDK most efficiently phosphorylates the Sld2 protein. We also identified efficient phosphorylation sites by CDK in both in vivo and in vitro using various Sld2 protein with mutations in CDK-phosphorylation sites. According to this assay, the threonine in the 20-aa stretch is an inefficient phosphorylation site. Moreover, the simultaneous mutations in canonical phosphorylation sites, some of which are actually efficient phosphorylation sites, reduced severely the phosphorylation of the threonine in the 20-aa stretch. We therefore propose that the threonine in the 20-aa stretch of Sld2 is phosphorylated cooperatively with other efficient phosphorylation sites and then Sld2 binds to Dpb11.

(3) CDK targets in the initiation of DNA replication

Seiji Tanaka and Hiroyuki Araki

The initiation of eukaryotic DNA replication is triggered by two essential kinases, Cdc7 and CDK. Requirement of Cdc7 in initiation can be bypassed with a *bob1-1* mutation, which is allelic to the one of the subunit of proposed replicative helicase, *MCM5*. Mutations that can bypass requirement of CDK, however, are not known so far. We have recently identified Sld2 as an essential substrate of CDK in the initiation of DNA replication. Phosphorylation of Sld2 by CDK enhances the interaction between Sld2 and its binding partner, Dpb11, which is likely to be important for association of the Sld2-Dpb11 complex to replication origins (Masumoto et al. Nature 415, 651-655, 2002).

To understand the individual processes in the initiation of DNA replication, we have tested whether Sld2 phosphorylation by CDK is sufficient for initiation. We have constructed a Sld2 mutant which might mimic a phosphorylated status of wild type Sld2. Although this mutant could support cell growth, CDK activity was still required for initiation. This suggests that CDK regulates multiple targets in the initiation of DNA replication. In order to identify these CDK target(s) other than Sld2, we have set a screening of mutants that show synthetic lethality with phosphomimetic sld2. The rationale of the screening is that cells might to be a lethal because of untimely DNA replication when all of the CDK targets in the initiation process including Sld2 are freed from strict regulation of CDK. Although we have not been successful to isolate that kind of mutants so far, we will continue the screening.

(4) Mechanisms for generating genomic instability

Seiji Tanaka and Hiroyuki Araki

Although genomic instability is a hallmark of human cancer cells, the mechanisms by which genomic instability is generated and selected for during oncogenesis remain obscure. In most human cancers, the pathway leading to the activation of the G1 cyclins is deregulated. Previously we hypothesized that G1 cyclin deregulation could cause genomic instability because the lack of a proper 'low Cdk' period in G1 may reduce numbers of functional pre-RCs that confer replication competence to cells. Using budding yeast as a model, we have shown that overexpression of the G1 cyclin, Cln2, inhibits the assembly of pre-RCs and induces gross chromosome rearrangements (GCRs) such as translocation, deletion of a chromosome arm, interstitial deletions and inversions. These results suggest that deregulation of G1 cyclins, selected for in oncogenesis because it confers clonal growth advantage, may also provide an important mechanism for generating genomic instability by inhibiting replication licensing. However, we still do not know how reducing origin licensing/firing contributes to genomic instability. We have focused on this question.

Reduced numbers of replication forks and slow DNA replication as a result may cause broken chromosomes in M phase or stalled replication forks that persist for longer periods. We have asked if these are good substrates for recombination and contribute to generating genomic instability. To test this, we have introduced a genomic sequence that is known to block the progression of replication fork and found that this induce higher rate of GCR. We are now proceeding the detailed analysis on this.

PUBLICATIONS

Papers

1. Iida, T. and Araki, H. (2004). Non-competitive counteractions of DNA polymerase ε and ISW2/yCHRAC for epigenetic inheritance of telomere-position effect in *Saccharomyces cerevisiae.* Mol. Cell. Biol. *24,* 217-227.

2. Mimura, S., Seki, T., Tanaka, S. and Diffley, J.F.X. (2004). Phophorylation-dependent binding of mitotic cyclins to Cdc6 contributes to DNA replication control. Nature *431,* 1118-1123.

Books

3. 荒木弘之(2004)真核生物の複製開始反応, DNA複製・修復がわかる(花岡文雄編集), 羊土社.

4. 荒木弘之(2004)DNAの複製, バイオテクノロジーのための基礎分子生物学(大嶋泰治・北本勝ひこ・原島俊・宮川都吉編), 化学同人.

EDUCATION

1. Dr. H. Araki gave a lecture at Medical School, University of Tokyo, February, 2004 (in Japanese).
2. Dr. H. Araki gave a seminar at Paterson Institute at Manchester, UK, April, 2004.
3. Dr. H. Araki gave a seminar at Clare Hall Laboratories, Cancer Research UK at South Mims, UK, April, 2004.
4. Dr. H. Araki gave a seminar at National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA, July, 2004.
5. Dr. H. Araki gave a lecture at Shizuoka University, October, 2004 (in Japanese).
6. Dr. H. Araki gave a lecture and a seminar at Tokyo Institute of Technology, November, 2004 (in Japanese).

B-b. Division of Microbial Genetics
Seiichi Yasuda Group

RESEARCH ACTIVITIES

(1) Mechanism of DnaA interaction with phospholipids

Seiichi Yasuda

The replication of chromosome in *Escherichia coli* is initiated at a unique chromosomal locus, *oriC*. A protein called DnaA catalyzes the first step in the initiation by binding and opening double-stranded DNA at an AT-rich region in *oriC*. The opened single-stranded region serves as the site of assembly of other proteins to form a replication complex. Comparative studies on the *dnaA* genes of many bacterial strains have established that DnaA is made of four functional domains. Among them, domain 3 is responsible for ATP-binding, and domain 4 for sequence-specific binding to *oriC* DNA. It has been known that acidic phospholipids such as cardiolipin release bound ATP and ADP from DnaA. This phenomenon was suggested to be involved in rejuvenation of used DnaA. It has been postulated that DnaA interacts with phospholipids at a site near the C-terminal portion of domain 3 because this site has an amino acid sequence that can form an alpha helix having a hydrophobic surface.

Previous studies with mutant DnaA proteins in this laboratory showed that the interaction of DnaA with phospholipids depends on the presence of the DNA-binding domain of DnaA, but not on that of the putative mem-brane-interaction site in domain III. This was further confirmed by quantitative measurements of DNA-binding and phospholipid-induced ATP-release using mutant DnaA proteins. DnaA mutants used were DnaADelM that has a deletion of the membrane-interaction site, DnaADelV that has a deletion of C-terminal 14 amino acids of the DNA-binding domain, and DnaADelS that has a deletion of almost all of the DNA-binding domain but that retains the membrane-interaction site. ATP binding of all these mutant proteins was essentially the same as that of wild type DnaA. Dissociation constants of DNA-binding were measured and found to be 5.75, 16.5, 115, and 825 nM for wild type DnaA, DnaADelM, DnaADelV and DnaADelS, respectively. The rate of ATP release versus cardiolipin concentration was determined and it was found that the order of cardiolipin sensitivity of these DnaA proteins was exactly reverse to that of the DNA binding constants. This suggests strongly that cardiolipin interacts with DnaA at the DNA-binding domain leading to the release of ATP from the protein.

During the course of this work it was found that cardiolipin inactivates the ATP, and DNA-binding activities of DnaA, but not the replication activity as measured in the presence of *oriC* and a crude protein fraction. The crude protein fraction was found to reactivate DnaA that had been inactivated by cardiolipin. Characterization of the reactivating factor is under way.

B-c. Division of Cytoplasmic Genetics
Kazuto Kobayashi Group

RESEARCH ACTIVITIES

(1) Mechanisms of Neural Circuitry Underlying Control of Behaviors

Kazuto Kobayashi

Control of behaviors is based on the mechanisms that mediate processing and regulation of information through the complex neural circuitry in the brain. Dysfunction in the neural circuitry is involved in the etiology and pathogenesis of some neural diseases. Our research group is interested in understanding the role of specific neurons or neural pathways that mediate motor control, memory, and emotion as well as the function of the signaling molecules that are involved in these behaviors. In the present year, we developed a Cre-*loxP* recombination system to study the role of signaling pathways in the function and development of the neural circuitry[2]. Using this system, we identified a critical role of Rho/Rho-kinase signaling pathway in the survival of developing motor neurons[1, 8]. In addition, we reported behavioral and physiological roles of the neural pathway originating from the striatum by using immunotoxin-mediated cell targeting[6, 9]. Immunotoxin-mediated cell targeting was also used to generate an animal model for autonomic diseases[10]. Furthermore, we identified subunit composition of $GABA_A$ receptors in midbrain

dopamine neurons by using the transgenic mouse strain that expresses GFP in these neurons[3]. This transgenic strain was also utilized for model experiments of cell transplantation for Parkinson's disease therapy[4, 5, 7].

PUBLICATIONS

Papers

1. Kobayashi, K., Takahashi, M., Matsushita, N., Miyazaki, J.-I., Yaginuma, H., Osumi, N., Kaibuchi, K. and Kobayashi, K. (2004). Survival of developing motor neurons by Rho GTPase signaling pathway through Rho-kinase. J. Neurosci. *24,* 3480-3488.
2. Matsushita, N., Kobayashi, K., Miyazaki, J. and Kobayashi, K. (2004). Fate of transient catecholaminergic cell types revealed by site-specific recombination in transgenic mice. J. Neurosci. Res. *78,* 7-15.
3. Okada, H., Matsushita, N., Kobayashi, K. and Kobayashi, K. (2004). Identification of $GABA_A$ receptor subunit variants in midbrain dopaminergic neurons. J. Neurochem. *89,* 7-14.
4. Saino-Saito, S., Sasaki, H., Volpe, B.T., Kobayashi, K., Berlin, R. and Baker, H. (2004). Differentiation of the dopaminergic phenotype in the olfactory system of neonatal and adult mice. J. Comp. Neurol. *479,* 389-398.
5. Yoshizaki, T., Inaji, M., Kouike, H., Shimazaki, T., Sawamoto, K., Ando, K., Date, I., Kobayashi, K., Suhara, T., Uchiyama, Y. and Okano, H. (2004). Isolation and transplantation of dopaminergic neurons generated from mouse embryonic stem cells. Neurosci. Lett. *363,* 33-37.

Reviews

6. 小林和人(2004)イムノトキシン細胞標的法：神経回路メカニズムの解析法. 脳21 *7,* 195-200.
7. 小林和人(2004)神経系の機能再生へのアプローチ. 理学療法学*31,* 448-450.
8. 小林憲太，小林和人(2004)Cre-*loxP*システム：コンディショナルな遺伝子改変技術. 脳21 *7,* 312-315.
9. 佐野裕美，小林和人(2004)線条体投射路を介する新しい神経回路メカニズム. Clinical Neuroscience *22,* 732.

Books

10. Kobayashi, K. and Nagatsu, T. (2004). Transgenic strategies in autonomic research. In The Primer on the Autonomic Nervous System, 2nd edition, Robertson, D. ed. (London: Elsevier Academic Press), pp.435-438.

EDUCATION

他大学／研究機関での講義やセミナー

1. Dr. K. Kobayashi gave a lecture at Kobe University School of Medicine, June, 2004 (In Japanese).
2. Dr. K. Kobayashi was invited to give a seminar on "Functional dissection of the neural circuitry that mediates dopamine-dependent behavior" at Wallenberg Neuroscience Center of Lund University, Lund, Sweden, September, 2004.
3. Dr. K. Kobayashi was invited to give a seminar on "Functional dissection of the neural circuitry that mediates dopamine-induced behavior" at Rutger University, New Jersey, USA, October, 2004.

市民講座，大学以外の学校での講演など

1. 小林和人「神経系の機能再生へのアプローチ」第39回日本理学療法学術大会イブニングセミナー，仙台，2004年5月.

客員教授等

大学共同利用機関法人　情報・システム研究機構　国立遺伝学研究所　客員教授

B-c. Division of Cytoplasmic Genetics Masayuki Yamamoto Group

RESEARCH ACTIVITIES

(1) Fission yeast Mes1p ensures the onset of meiosis II by blocking degradation of cyclin Cdc13p[2]

Masayuki Yamamoto

Meiosis is a special form of nuclear division to generate eggs, sperm, and spores in eukaryotes. Meiosis consists of the first (MI) and the second (MII) meiotic divisions, which occur consecutively. MI is reductional, in which homologous chromosomes derived from parents segregate. MII is equational, in which replicated sister-chromatids separate as in mitosis. MII is generally considered to mimic mitosis in mechanism. However, fission yeast Mes1p is essential for MII but dispensable for mitosis. We have demonstrated that Mes1p is a factor that suppresses destruction of cyclin Cdc13p at anaphase I. Mes1p inhibits the activity of APC/C (anaphase promoting complex/cyclosome) to degrade Cdc13p, thereby

playing a key role in saving a sufficient level of MPF (M-phase promoting factor) activity required for the execution of MII[2].

(2) The p150-Glued Ssm4p regulates microtubular dynamics and nuclear movement in fission yeast[4]

Masayuki Yamamoto

During vegetative growth of the fission yeast, microtubules nucleate from multiple MTOCs in the vicinity of the nucleus, polymerizing until they reach the end of the cell and then shrinking back to the cell middle. In response to mating pheromone, fission yeast undergoes a morphological switch from a vegetative to a shmooing growth pattern. The switch in growth mode is paralleled by a switch in microtubular dynamics. Microtubules nucleate mostly from a single MTOC and pull on the ends of the cell to move the nucleus back and forth. This movement continues after cellular and nuclear fusion in the zygote and is important to ensure correct chromosome pairing, recombination and segregation during meiosis. We have demonstrated that Ssm4p, a p150-glued protein, is induced specifically in response to pheromone and is required for this nuclear movement. Ssm4p is associated with the cytoplasmic dynein complex and regulates dynein heavy chain localization. We have also shown that Ssm4p functions in establishing the shmooing microtubular array[4].

(3) *C. elegans* DAZ-1 directs proper nuclear organization and cytoplasmic core formation during oogenesis[3]

Masayuki Yamamoto

The *Deleted in Azoospermia (DAZ)* family genes encode potential RNA-binding proteins that are expressed exclusively in germ cells in a wide range of metazoans. Mutations in *daz-1,* the only *DAZ* family gene in *Caenorhabditis elegans,* cause pachytene stage arrest of female germ cells, but do not affect spermatogenesis. We have demonstrated that DAZ-1 protein is most abundantly expressed in proliferating female germ cells. DAZ-1 is dispensable in males, but it is expressed also in male mitotic germ cells. By detailed phenotypic analyses we have found that loss of *daz-1* function causes multiple abnormalities as early as the onset of meiotic prophase, which include aberrant chromatin structure, small nucleoli, absence of the cytoplasmic core, and precocious cellularization[3]. It appears that DAZ-1 in *C. elegans* plays essential roles in female pre-meiotic and early meiotic germ cells, probably via regulating the translational activity of specific target genes required for the progression of oogenesis.

PUBLICATIONS

Papers

1. Inoue, T., Sugimoto, A., Suzuki, Y., Yamamoto, M., Tsujimoto, M., Inoue, K., Aoki, J. and Arai, H. (2004). Type II platelet-activating factor-acetylhydrolase is essential for epithelial morphogenesis in *Caenorhabditis elegans. Proc. Natl. Acad. Sci. USA. 101,* 13233-13238.
2. Izawa, D., Goto, M., Yamashita, A., Yamano, H. and Yamamoto, M. (2005). Fission yeast Mes1p ensures the onset of meiosis II by blocking degradation of cyclin Cdc13p. *Nature 434,* 529-533.
3. Maruyama, R., Endo, S., Sugimoto, A. and Yamamoto, M. (2005). *C. elegans* DAZ-1 is expressed in proliferating germ cells and directs proper nuclear organization and cytoplasmic core formation during oogenesis. *Dev. Biol. 277,* 142-154.
4. Niccoli, T., Yamashita, A., Nurse, P. and Yamamoto, M. (2004). The p150-Glued Ssm4p regulates microtubular dynamics and nuclear movement in fission yeast. *J. Cell Sci. 117,* 5543-5556.
5. Yamashita, A., Sato, M., Fujita, A., Yamamoto, M. and Toda, T. (2005). The roles of fission yeast Ase1 in mitotic cell division, meiotic nuclear oscillation and cytokinesis checkpoint signaling. *Mol. Biol. Cell 16,* 1378-1395.

SOCIAL CONTRIBUTIONS AND OTHERS

President: The Molecular Biology Society of Japan
Editor: Genes to Cells
Editorial Board: Current Genetics, YEAST

C. DEPARTMENT OF DEVELOPMENTAL GENETICS

C-a. Division of Developmental Genetics Yasushi Hiromi Group

RESEARCH ACTIVITIES

(1) Multi-dimensional axon network formation

Masaki Hiramoto and Yasushi Hiromi

A key feature of the CNS that allows efficient information processing is its repetitive modular structure. Such structure is found in the invertebrate ventral nerve cord, the vertebrate rhombomere or the columnar structures of the cortex. An individual module contains the complement of cell types and guidance cues required to form a neural circuit within the module. To construct a neural network, however, multiple modules must be interconnected by axons exiting the module border and then re-entering the next module. Such "boundary crossing" requires axons to adopt navigation strategies that are different from the rules that shaped their trajectories within the module unit. In bilateralians a boundary crossing problem occurs at the midline, the axis of symmetry. In animals with segmentally repeated body plan, an additional boundary crossing problem resides at the border of the repeat unit, segment boundary. In contrast to the well-studied midline crossing event, segmetal boundary crossing problem has never been subjected to an analysis, either in the vertebrate or the invertebrate CNS. We showed that two evolutionary conserved axon guidance molecules, secreted ligand Netrin and a guidance receptor ROBO, which are known to play important role in axon guidance across the midline, are also used as a molecular cassette to license longitudinal axons to cross the segmental boundary, thereby connecting modular structures in the CNS. This finding demonstrates that the dimension of the multi-modular structure can be increased by using the same molecular cassette multiple times.

(2) Intrinsic sub-axonal patterning in *Drosophila* neurons

Takeo Katsuki, Masaki Hiramoto and Yasushi Hiromi

During the development of the nervous system, neurons extend their axons over a long distance to their targets with the assistance of guidance cues and guidance receptors. Although a number of molecules that play instructive role in axonal navigation have been discovered, little is known about how the distributions of such guidance molecules are regulated in space and time. Immunohistochemical studies have revealed that the spatial distribution of guidance receptors *in vivo* is often restricted to specific segments of axons. Is such sub-axonal localization of guidance receptors generated by the intrinsic properties of neurons or by extrinsic signals? And how is the localization of transmembrane receptors maintained in the axonal membrane despite the fact that the membrane is continuous and fluid? Using a primary cell culture system of *Drosophila*, we demonstrated that isolated neurons possess an intrinsic property to generate sub-axonal localization of guidance molecules. FRAP analysis provided evidence that an underlying mechanism for the sub-axonal localization is compartmentalization of the axonal membrane by a diffusion barrier. We propose that the regulated expression of guidance receptors *in vivo* is based on the intrinsic sub-axonal patterning property of neurons.

(3) Identification of BP102 antigen which shows specific sub-axonal localization

Tohru Umemura, Takeo Katsuki and Yasushi Hiromi

A striking feature of the neuron is its regionalized cellular organization. Cellular compartments such as axon and dendrite have distinct molecular signatures that enable specific physiological functions. Axons themselves have regional specifications, such as the growth cone and the distal end and the initial segment at the base, which differ in cytoskeletal organizations that influence the dynamics of the axonal membrane and the movement of molecules within. Although the axonal shaft is seemingly uniform, a number of molecules are localized to specific sub-segments of

axons *in vivo*. Such sub-axonal localization of molecules is likely based on intrinsic patterning ability of the neuron that regionalizes the axon to sub-axonal "compartments" (see above). One of the molecules that exhibit intrinsic sub-axonal localization is the antigen of the monoclonal antibody BP102, which stains many axons in the *Drosophila* CNS highlighting the ladder-like axonal scaffold. In primary culture this antibody stains the middle or proximal segment of the axon, suggesting the existence of distinct sub-axonal compartments. We are trying to identify the BP102 antigen to investigate how it organizes itself to such sub-axonal compartments. We found that the antigen is a membrane protein of about 100kD, and requires glycosilation to be recognized by the antibody.

(4) Lamina-specific connection in mouse hippocampus

Fumikazu Suto, Hajime Fujisawa[1] and Yasushi Hiromi ([1]Nagoya University)

In many part of central nervous system, distinct populations of axons confine their terminal arbors and synapses to different subsets of laminae. These lamina-specificities can be divided into two categories, cellular and subcellular. Several molecules that are involved in lamina-specific projection at the cellular-specific level have recently been identified in the *Drosophila* visual system. However, little is known about the molecular mechanisms underlying the connection to a specific subcellular region, a process that requires recognition of a subcellular compartment. To understand the molecular mechanisms for the subcellular-specific connection, we are using mouse hippocampus as a model system. In the hippocampus, dentate granule cells (DGCs) project their axons (mossy fibers; MFs) mainly to the proximal-most part of apical dendrites of CA3 pyramidal cells and partly to the proximal basal dendrites in subcellular-specific manner. We found that axon guidance molecule Sema6A was localized on the dendritic field of CA3 pyramidal cells, and had repulsive activity for MFs, suggesting that Sema6A might be the signal that specify subcellular specificity of the MF projection. We thus analyzed the roles of Sema6A receptors, Plexin-A2 and Plexin-A4, on MF projection using mutant and normal explants. *plexin-A4*-deficient DGCs projected MFs abnormally to all parts of the apical and basal dendrites of genotypically normal CA3 pyramidal cells. In contrast, CA3 pyramidal cells that are mutant for *plexin-A2* failed to receive projections from genotypically normal MFs in the apical dendritic field, and received only in the proximal basal dendrite. These results suggest that during subcellular-specific MFs projection, Plexin-A4 mediates Sema6A repulsive activity in MFs, and Plexin-A2 regulates Sema6A activity and/or distribution in the CA3 field.

(5) *seven-up* controls switching of transcription factors that specify temporal identities of *Drosophila* neuroblasts

Makoto Kanai, Masataka Okabe and Yasushi Hiromi

Drosophila neuronal stem cell neuroblasts constantly change character upon division to produce a different type of progeny at the next division. Transcription factors Hunchback (HB), Krüppel (KR), Pdm (PDM), etc. are expressed sequentially in the developmental history of each neuroblast and act as determinants of birth-order identity. How neuroblast switches its expression profile from one transcription factor to the next is poorly understood. We showed that the HB-to-KR switch is directed by the nuclear receptor Seven-up. Seven-up expression is confined to a temporally restricted subsection within the neuroblast's lineage. Loss of *seven-up* function causes an increase in the number of HB-positive cells within several neuroblast lineages, whereas misexpression of *seven-up* leads to the loss of these early-born neurons. Lineage analysis provided evidence that *seven-up* is required to switch off HB at the proper time. Thus, *seven-up* modifies the self-renewal stem cell program to allow chronological change of cell fates, thereby generating neuronal diversity.

(6) A search for Seven-up-interacting molecules

Takayuki Hondoh and Yasushi Hiromi

Seven-up is a transcription factor containing a DNA binding domain and an evolutionarily-conserved domain that has homology to the ligand binding domains of nuclear receptors. The spatio-temporal expression pattern of Seven-up is tightly linked to the cell fate; in the developing *Drosophila* eye, *seven-up* is expressed in only four of the eight photoreceptor neurons, which are transformed to another neuronal

fate in *seven-up* mutant. In the embryonic CNS *seven-up* expression is temporally restricted to a subsection of the neuroblast lineage. The loss-of function mutation causes a "temporal transformation" of the fate of the neuroblast and its progeny (see above). Although Seven-up plays instructive roles in many cell fate decisions, how this transcription factor accomplishes such cell fate choices is not understood. We have prepared antibodies against Seven-up that would enable us to probe the endogenous Seven-up molecule. Using IP and ChIP methods, we are investigating the molecular environment around Seven-up.

(7) Glial-expression of Prospero integrates inputs from multiple signaling

Yoshihiro Yuasa and Yasushi Hiromi

Glial cells in the nervous system are morphologically and functionally diverse, performing various functions such as providing a scaffold for axonal migration, insulation of axons, and supporting neuronal survival. In *Drosophila,* interface glia such as longitudinal glia enwraps the longitudinal axon and, in some ways, resembles oligodendrocytes in vertebrates. Longitudinal glia also plays a key role in the formation of longitudinal axon pathway. Ten longitudinal glial cells that reside in each hemisegment are generated from a single precursor that divides while migrating towards the midline. These longitudinal glial cells likely comprise multiple glial sub-types, because some of them have specific association with distinct axon pathways. One molecule that is expressed in a subset of the longitudinal glia is a homeodomain transcription factor Prospero (PROS). PROS is expressed in six out of ten longitudinal glial cells, and is required for the generation of the longitudinal axonal scaffold and neuronal survival. In order to understand the gene regulatory cascade leading to the specification of longitudinal glial cells, we analyzed how PROS expression is regulated in the longitudinal glia. We found that the combinational expression of three transcriptional factors and Notch signaling is essential for the PROS expression in the longitudinal glia.

(8) A screening for genes involving germline stem cell establishment and niche formation in *Drosophila*

Miho Asaoka, Shuji Shigenobu[1], Satoru Kobayashi[1] and Yasushi Hiromi ([1]NIBB)

Stem cells play a central role in generating and maintaining most adult tissues in higher organisms. They are defined by their ability to self-renew and to produce numerous differentiated daughter cells. Stem cells in the adult tissue reside in a special microenvironment called "niche", which are thought to produce signals that support and maintain stem cell specification and function. However, how stem cell fate and niche are initially established during tissue development is currently unknown. In *Drosophila,* germline stem cells are derived from a subset of primordial germ cells (PGCs). We have previously shown that only the PGCs contacting somatic cells in the anterior half of embryonic gonad will become germline stem cells[3]. We also found that the somatic cells in the anterior half of the embryonic gonad already differ from the posterior somatic cells and likely constitute the niche in the adult. These results suggest that somatic cells in the anterior half of the embryonic gonad play a key role in the germline stem cell fate establishment and niche formation. As the first step for understanding molecular mechanisms of germline stem cell establishment and niche formation, we are focusing on genes that are expressed in the somatic cells in the anterior half of the embryonic gonad. Through EST and microarray analyses we identified 187 genes which are expressed only in the somatic cells in the embryonic gonad. *In situ* hybridization analyses show that some of these genes are expressed only in the anterior somatic cells in the embryonic gonad. We plan to test the function of these candidate genes in germline stem cell establishment and niche formation, using RNAi methods.

(9) The origin of the parathyroid gland

Masataka Okabe and Anthony Graham[1] ([1]MRC, Centre for Developmental Neurobiology, King's College London)

It has long been held that the parathyroid glands and parathyroid hormone evolved with the emergence of the tetrapods, reflecting a need for new controls on

calcium homeostasis in terrestrial, rather than aquatic, environments. Developmentally, the parathyroid gland is derived from the pharyngeal pouch endoderm, and studies in mice have shown that its formation is under the control of a key regulatory gene, Gcm-2. We have used a phylogenetic analysis of Gcm-2 to probe the evolutionary origins of the parathyroid gland. We show that in chicks, as in mice, Gcm-2 is expressed in the pharyngeal pouches and the forming parathyroid gland. We find that Gcm-2 is present not only in tetrapods but also in teleosts and chondrichthyans, and that in these species, Gcm-2 is expressed within the pharyngeal pouches and internal gill buds that derive from them in zebrafish (*Danio rerio*), a teleost, and dogfish (*Scyliorhinus canicula*), a chondrichthyan. We further demonstrate that Gcm-2 is required for the formation of the internal gill buds in zebrafish. We also have identified parathyroid hormone 1/2-encoding genes in fish and show that these genes are expressed by the gills. We further show that the gills express the calcium-sensing receptor, which is used in tetrapods to monitor serum calcium levels. These results indicate that the tetrapod parathyroid gland and the gills of fish are evolutionarily related structures, and that the parathyroid likely came into being as a result of the transformation of the gills during tetrapod evolution[4].

PUBLICATIONS

Papers

1. Niwa, N., Hiromi, Y. and Okabe, M. (2004). A conserved developmental program for sensory organ formation in *Drosophila melanogaster*. Nat. Genet. *36,* 293-297.
2. Jindra, M., Gaziova, I., Uhlirova, M., Okabe, M., Hiromi, Y. and Hirose, S. (2004). Coactivator MBF1 preserves the redox-dependent AP-1 activity during oxidative stress in *Drosophila*. EMBO J. *23,* 3538-3547.
3. Asaoka, M. and Lin, H. (2004). Germline stem cells in the *Drosophila* ovary descend from pole cells in the anterior region of the embryonic gonad. Development *131,* 5079-5089.
4. Okabe, M. and Graham, A. (2004). The origin of the parathyroid gland. Proc. Nat. Acad. Sci. USA *101,* 17716-17719.
5. 岡部正隆(2004)色覚バリアフリー. 東京都眼科医会報*187,* 3-9.

EDUCATION

1. Dr. Y. Hiromi gave a lecture course at Tokyo University of Agriculture and Technology. December, 2004 (in Japanese).

SOCIAL CONTRIBUTIONS AND OTHERS

1. Dr. Y. Hiromi served as an editor for Development, Growth & Differentiation.
2. Dr. Y. Hiromi served as a member of the council of The Genetics Society of Japan.
3. Dr. M. Okabe developed "Barrier-free presentation that is friendly to colorblind people", and enhanced its public awareness through web pages:
   ①岡部正隆, 伊藤啓 http://www.nig.ac.jp/color/
   ②Okabe, M. and Ito, K. http://jfly.iam.u-tokyo.ac.jp/color/
4. Dr. M. Asaoka contributed in making a DVD/Video「幹細胞とニッチ一見えてきた造血幹細胞のすみか一」(制作：桜映画社, 企画：中外製薬㈱, 監修：平嶋邦猛, 浅野茂隆, 中畑龍俊, 中内啓光).

C-a. Division of Developmental Genetics Toshitaka Fujisawa Group

RESEARCH ACTIVITIES

(1) Systematic identification of peptide signaling molecules in *Hydra*

Toshitaka Fujisawa, Chiemi Fujisawa, Eisuke Hayakawa, Yukihiko Noro, Akemi Hayashiuchi, Hiroshi Shimizu and Toshio Takahashi[1] ([1]Suntory Institute for Bioorganic Research, Osaka, Japan)

Systematic purification and structural determination has been carried out by using HPLC, peptide sequencing and mass spectrometry. About 50 new peptides were purified and subjected to structural determination. The peptide sequences were blasted against cDNA database we have created and identified 7 new peptide genes and 3 novel neuropeptides. Out of 3 neuropeptides, the first one belonged to the already known GLWamide family, the second belonged to the Hym-176 family and the last (Hym-65) to the hitherto unknown family. The Hym-65 gene presumably encode another peptide with a FRamide C-terminal motif.

Whole mount in situ hybridization indicated that *Hym-65* gene is expressed in a subpopulation of neurons throughout the body column. Hym-65 peptide induced body column contraction of epithelial polyp that is only made of epithelial cells and gland cells but has no other cells in the interstitial cell lineage including neurons. Thus, this peptide appears to act directly to the muscle cells presumably as a neurotransmitter at the neuro-muscular junctions. In contrast, the new Hym-176 family peptide strongly evoked contraction of ectodermal muscle of normal polyp but not of epithelial polyp suggesting that this peptide does not function as a neurotransmitter at the neuro-muscular junctions but acts via other neurons.

PUBLICATIONS

Papers

1. Horibata, Y., Sakaguchi, K., Okino, N., Iida, H., Inagaki, M., Fujisawa, T., Hama, Y. and Ito, M. (2004). Unique catabolic pathway of glycosphingolipids in a hydrozoan, *Hydra magnipapillata,* involving endoglycoceramidase. J. Biol. Chem. *279* (32), 33379-33389.
2. Shimizu, H., Koizumi, O., and Fujisawa, T. (2004). Three digestive movements in *Hydra* regulated by the diffuse nerve net in the body column. J. Comp. Physiol. A *190* (8), 623-630.
3. Hayakawa, E., Fujisawa, C. and Fujisawa, T. (2004). Involvement of *Hydra achaete-scute* gene, *CnASH* in the differentiation pathway of sensory neurons in the tentacles. Development, Genes and Evolution *214,* 486-492.
4. Alexopoulos, H., Boettger, A., Fisher, S., Levin, A., Wolf, A., Fujisawa, T., Hayakawa, S., Gojobori, T., Davies, J.A., David, C.N. and Bacon, J.P. (2004). Evolution of gap junctions: the missing link? Current Biology *14* (20), 879-880.
5. Takaku, Y., Hariyama, T. and Fujisawa, T. (2005). Motility of endodermal epithelial cells plays a majo role in reorganizing the two epithelial layers in *Hydra.* Mech. Dev. *122,* 109-122.

SOCIAL CONTRIBUTIONS AND OTHERS

各種受賞

The Zoological Society Prize for Toshitaka Fujisawa

C-b. Division of Gene Expression Susumu Hirose Group

RESEARCH ACTIVITIES

(1) Role of histone modifications and chromatin remodeling in epigenetic gene expression

Takahiro Nakayama, Tsukasa Shimojima, Kazuma Hanai, Kenichi Nishioka, Koji Akasaka[1] and Susumu Hirose ([1]Department of Biological Science, University of Tokyo, Tokyo, Japan)

In multicellular organisms some patterns of gene expression are remembered in the chromatin structure and maintained through many rounds of cell cycle. These are termed epigenetic gene expression. Two types of epigenetic gene expression are known in *Drosophila melanogaster:* (1) maintenance of *Hox* gene expression governed by *Polyomb (Pc)* and *trithorax (trx)* group genes, and (2) position effect variegation (PEV). *Trithorax-like (Trl)* encoding GAGA factor is involved in the both types of regulation. Thus *Trl* is a member of *trx* group and *Trl* mutation is an enhancer of PEV. We have shown that GAGA factor recruits FACT, a heterodimer of dSPT16 and dSSRP1, facilitates chromatin remodeling around its binding site and contributes to the maintenance of *Hox* gene expression.

When an actively transcribed *white (w)* gene is juxtaposed with heterochromatin by chromosome rearrangement such as w^{m4}, its expression is subject to variable but heritable silencing. This is PEV. We found that GAGA factor-FACT complex binds to a site just down stream of *w,* facilitates chromatin remodeling, and plays an important role in the maintenance of *w* expression against heterochromatin silencing.

RSF (remodeling and spacing factor) has been first purified from human cells as a heterodimer of RSF1 and SNF2H (a human counterpart of *Drosophila* ISWI), which can assemble regularly spaced nucleosome arrays *in vitro.* To investigate *in vivo* role of RSF, we analyzed its *Drosophila* counterpart, a heterodimer of dRSF1 and ISWI. Genetic studies implicate *Drosophila* RSF in PEV through facilitating the spreading of silent chromatin.

(2) Chromatin transcription

Mikage Nakajima, Kenichi Nishioka, Tadashi Wada[1], Hiroshi Handa[1] and Susumu Hirose ([1]Faculty of Bioscience and Biotechnology, Tokyo Institute of Technology, Nagatsuda, Yokohama, Japan)

SPT6 is a transcription factor that is conserved across species from yeast to human. While biochemical studies on human SPT6 have revealed it to be a transcription elongation factor, genetic studies on yeast SPT6 have suggested its role in the recovery of nucleosome structure from deformation due to the passage of RNA polymerase II through a nucleosome. To study *in vivo* role of SPT6 in mulicellular organism, we started genetic study on *Drosophila* SPT6. We also continued studies on SPT6 and FACT using chromatin transcription *in vitro* in collaboration with Drs. T. Wada and H. Handa.

(3) Role of DNA topology in the formation of active chromatin

Kuniharu Matsumoto, Hirofumi Furuhashi, Youhei Ogasawara and Susumu Hirose

Bulk DNA within the eukaryotic genome is torsionarily relaxed. However unconstrained negative supercoils of DNA have been detected in few local domains of the genome through preferential binding of psoralen. To make genome-wide survey for such domains, we introduced biotinylated psoralen into *Drosophila* salivary glands and visualized it on polytene chromosomes with fluorescent streptavidin. We observed bright psoralen signals on many transcriptionally active interbands and puffs. Upon heat shock, the signals appeared on heat-shock puffs. The signals were resistant to RNase treatment but disappeared or became faint by prior nicking of DNA or inhibition of transcription with α-amanitin. These data demonstrate that transcription-coupled, unconstrained negative supercoils of DNA exist in approximately 150 loci within the interphase genome.[4]

Supercoiling factor (SCF) is a protein capable of introducing negative supercoils into DNA in conjunction with topoisomerase II. Localization of SCF on puffs of polytene chromosomes has suggested its role in the formation of active chromatin. Knocking down of the SCF function *in vivo* by RNAi implicated SCF in twice activation of genes on male X chromosome. In good agreement with this, overexpression of SCF resulted in bloated appearance of the male X chromosome.

(4) Nucleosomal histone kinase-1

Hitoshi Aihara[1], Takeya Nakagawa[1], Kiyoshi Yasui[1], Tsutomu Ohta[2], Susumu Hirose, Naoshi Dhomae[3], Koji Takio[4], Mayumi Kaneko[5], Yukio Takeshima[5], Masami Muramatsu[6] and Takashi Ito[1] ([1]Department of Biochemistry, Nagasaki University School of Medicine, Nagasaki, Japan; [2]Medical Genomic Center, National Cancer Center Research Institute, Tokyo, Japan; [3]The Institute of Physical and Chemical Research, Wako-shi, Saitama, Japan; [4]RIKEN Harima Institute, Mikazuki-cho, Hyogo, Japan; [5]Second Department of Pathology, Hiroshima University School of Medicine, Hiroshima, Japan; [6]Saitama Medical School Research Center for Genomic Medicine, Hidaka, Saitama, Japan)

Posttranslational histone modifications are important for the regulation of many biological phenomena. Here, we show the purification and characterization of nucleosomal histone kinase-1 (NHK-1). NHK-1 has a high affinity for chromatin and phosphorylates a novel site, Thr 119, at the C terminus of H2A. Notably, NHK-1 specifically phosphorylates nucleosomal H2A, but not free H2A in solution. In *Drosophila* embryos, phosphorylated H2A Thr 119 is found in chromatin, but not in the soluble core histone pool. Immunostaining of NHK-1 revealed that it goes to chromatin during mitosis and is excluded from chromatin during S phase. Consistent with the shuttling of NHK-1 between chromatin and cytoplasm, H2A Thr 119 is phosphorylated during mitosis but not in S phase. These studies reveal that NHK-1-catalyzed phosphorylation of a conserved serine/threonine residue in H2A is a new component of the histone code that might be related to cell cycle progression.[1]

(5) Tips in analyzing antibodies directed against specific histone tail modifications

Kavitha Sarma[1], Kenichi Nishioka and Danny Reinberg[1] ([1]Howard Hughes Medical Institute, Division of Nucleic Acids Enzymology, Department of Biochemistry, University of Medicine and Dentistry of

New Jersay, Piscataway, NJ, USA)

Histone methylation has been known to exist for over 40 years but the enzymes that catalyze this reaction have remained elusive until the discovery that Suv39H1 methylates histone H3 specifically at lysine 9. This discovery was followed by a bevy of papers describing other methyltransferases specific for different residues and their apparent function *in vivo*. Histones are methylated as lysine as well as arginine residues. Lysines can be mono-, di-, or tri-methylated *in vivo*.

Studies on the effect of these modifications on gene regulation have been greatly facilitated by the production of antibodies "specific" for the modified state. The need to carefully characterize antibodies raised against methylated histone peptides stems from the observation by several laboratories that these antibodies can be promiscuous depending on several factors such as concentration, peptide context, substrate, etc. Due to this, several papers have been subject to scrutiny in recent months as the specificity of the antibodies used was questionable.

In this article, we present several parameters to be taken into consideration and some useful hints for systematic characterization of antibodies raised against methylated histone peptides. Although we have focused on antibodies against methylated residues: H3-K4, H3-K27, and H4-K20, the methods and procedures described herein are applicable for any antibody directed against the histone tail modification, including arginine methylation and lysine acetylation, among other modifications.[6]

(6) Functional analysis of transcriptional coactivator MBF1

Marek Jindra[1], Ivana Gaziova[1], Mirka Uhlirova[1], Masataka Okabe[2], Yasushi Hiromi[2], Kenichi Tsuda[3], Toshiro Tsuji[3], Ken-ichi Yamazaki[3] and Susumu Hirose ([1]Department of Molecular Biology, University of South Bohemia and Institute of Entomology ASCR, Ceske Budejovice, Czech Republic; [2]Division of Developmental Genetics, National Institute of Genetics; [3]Laobratory of Environmental Molecular Biology, Graduate School of Environmental Earch Science, Hokkaido University, Sapporo, Japan)

Basic leucine zipper proteins Jun and Fos form the dimeric transcription factor AP-1, essential for cell differentiation and immune and antioxidant defenses. AP-1 activity is controlled, in part, by the redox state of critical cysteine residues within the basic regions of Jun and Fos. The redox control is necessary since replacement of these cysteines contributes to oncogenic potential of Jun and Fos. How cells maintain the redox-dependent AP-1 activity at favorable levels is not known. We show that the conserved coactivator MBF1 is a positive modulator of AP-1. Via a direct interaction with the basic region of *Drosophila* Jun (D-Jun), MBF1 prevents an oxidative modification (S-cystenyl cystenylation) of the critical cysteine and stimulates AP-1 binding to DNA. Cytoplasmic MBF1 translocates to the nucleus together with a transfected D-Jun protein, suggesting that MBF1 protects nascent D-Jun also in *Drosophila* cells. Consistent with the role of AP-1 in antioxidant defense, *mbf1*-null mutants exhibit shorter life than $mbf1^+$ control in the presence of hydrogen peroxide (H_2O_2). An AP-1-dependent process of epithelial closure becomes sensitive to H_2O_2 in flies lacking *mbf1*. These results indicate that by preserving the redox-sensitive AP-1 activity, MBF1 provides an advantage during oxidative stress.[3]

Multiprotein bridging factor 1 (MBF1) is known to be a transcriptional co-activator that mediates transcriptional activation by bridging between an activator and a TATA-box binding protein (TBP). We demonstrated that expression of every three MBF1 from *Arabidopsis* partially rescues the yeast *mbf1* mutant phenotype, indicating that all of them function as co-activators for GCN4-dependent transcriptional activation. We also report that each of their subtypes shows distinct tissue-specific expression patterns and responses to phytohormones. These observations suggest that even though they share a similar biochemical function, each MBF1 has distinct roles in various tissues and conditions.[7]

(7) Differential functions of G protein and Baz-aPKC signaling pathways in Drosophila neuroblast asymmetric division

Yasushi Izumi[1], Nao Ohta[1], Asako Itoh-Furuya[1], Naoyuki Fuse and Fumio Matsuzaki[1] ([1]Laboratory for Cell Asymmetry, Center for Developmental Biology, Institute of Physical and Chemical Research, and CREST, Japan Science and Technology Corporation, Kobe, Japan)

Drosophila melanogaster neuroblasts (NBs) undergo asymmetric divisions during which cell-fate determinants localize asymmetrically, mitotic spindles orient along the apical-basal axis, and unequal-sized daughter cells appear. We identified here the first *Drosophila* mutant in the Gγ1 subunit of heterotrimeric G protein, which produces Gγ1 lacking its membrane anchor site and exhibits phenotypes identical to those of *Gβ 13F*, including abnormal spindle asymmetry and spindle orientation in NB divisions. This mutant fails to bind Gβ 13F to the membrane, indicating an essential role of cortical G1-Gβ13F signaling in asymmetric divisions. In *Gγ 1* and *Gβ 13F* mutant NBs, Pins-Gαi, which normally localize in the apical cortex, no longer distribute asymmetrically. However, the other apical components, Bazooka-atypical PKC-Par6-Inscuteable, still remain polarized and responsible for asymmetric Miranda localization, suggesting their dominant role in localizing cell-fate determinants. Further analysis of *Gβγ* and other mutants indicates a predominant role of Partner of Inscuteable-Gαi in spindle orientation. We thus suggest that the two apical signaling pathways have overlapping but different roles in asymmetric NB division.[2)]

(8) Introduction of $p16^{INK4a}$ inhibits telomerase activity through transcriptional suppression of human telomerase reverse transcriptase expression in human gliomas

Masahiro Saito[1], Kou Nakagawa[1], Katsuyuki Hamada[2], Susumu Hirose, Hironobu Harada[1], Shohei Kohno[1], Shigeyuki Nagato[1] and Takanori Ohnishi[1]. ([1]Department of Neurosurgery and [2]Obstetrics and Gynecology, Ehime University School of Medicine, Ehime Japan)

The p16 and p53 tumor suppressor proteins, which are frequently altered in malignant gliomas, have been noted as regulators of telomerase activity. However, the link between telomerase regulation and these suppressor proteins has not been adequately clarified. In the present study, we demonstrated that p16, as well as p53, suppress telomerase activity through transcriptional regulation of human telomerase reverse transcriptase (hTERT) in malignant glioma. To examine the effect of p16 and p53 on telomerase activity, we utilized wild-type p16 or p53 expression plasmid and three human glioma cell lines differing in their *p53* and *p16* status. Restoring *p16* significantly reduced the level of telomerase activity of glioma cells. Furthermore, cotransfection of the *p16* gene with 5'-deletion constructs of the *hTERT* promoter carrying Sp1 binding sites, repressed the transcriptional activity of *hTERT* promoter in *p16*-deleted cells. In addition, electrophoretic mobility shift assay revealed that p16 expression inhibited the binding of Sp1 to the consensus Sp1 responsive element, indicating that the recruitment of Sp1 to the *hTERT* proximal core promoter is inhibited by p16 protein. These results were similar to those from a *p53* transfection study in *p53*-mutated cells. These findings implicate *p16* in the transcriptional regulation of telomerase activity by inhibiting the function of Sp1 in human malignant gliomas.[5)]

PUBLICATIONS

Papers

1. Aihara, H., Nakagawa. T., Yasui, K., Ohta, T., Hirose, S., Dhomae, N., Takio, K., Kaneko, M., Takeshima, Y., Muramatsu, M. and Ito, T. (2004). Nucleosomal histone kinase-1 phosphorylates H2A Thr 119 during mitosis in the early *Drosophila* embryo. *Genes Dev., 18,* 877-888.

2. Izumi, Y., Ohta, N., Itoh-Furuya, A., Fuse, N. and Matsuzaki, F. (2004). Differential functions of G protein and Baz-aPKC signaling pathways in *Drosophila* neuroblast asymmetric division. *J. Cell Biol. 164,* 729-738.

3. Jindra, M., Gaziova, I., Uhlirova, M., Okabe, M., Hiromi, Y., Tsuda, K., Tsuji, T., Yamazaki, K. and Hirose, S. (2004). Coactivator MBF1 preserves the redox-dependent AP-1 activity during oxidative stress in *Drosophila. EMBO J., 23,* 3538-3547.

4. Matsumoto, K. and Hirose, S. (2004). Visualization of unconstrained negative supercoils of DNA on polytene chromosomes of *Drosophila. J. Cell Sci., 117,*

3797-3805.
5. Saito, M., Nakagawa, K., Hamada, K., Hirose, S., Harada, H., Kohno, S., Nagato, S. and Ohnishi, T. (2004). Introduction of *p16^{INK4a}* inhibits telomerase activity through transcriptional suppression of human telomerase reverse transcriptase expression in human gliomas. *Int. J. Oncol., 24,* 1213-1220.
6. Sarma, K., Nishioka, K. and Reinberg, D. (2004). Tips in analyzing antibodies directed against specific histone tail modifications. *Methods Enzymol. 376,* 255-269.
7. Tsuda, K., Tsuji, T., Hirose, S. and Yamazaki, K. (2004). Three *Arabidopsis* MBF1 homologs with distinct expression profiles play roles as transcriptional co-activators. *Plant Cell Physiol., 45,* 225-231.

Reviews

8. 兼崎琢磨，西岡憲一(2004)「*Hox* 遺伝子群におけるエピジェネティックス制御の分子機構」細胞工学，*23,* 1155-1161.
9. 西岡憲一(2004)「エピジェネティックス制御機構におけるメチル化ヒストンの役割」実験医学，*22,* 1361-1370.
10. 広瀬　進(2004)「ホメオティック遺伝子の発現制御」実験医学，*22,* 1371-1375.
11. 広瀬　進(2004)「MBF1 と寿命」医学のあゆみ，*211,* 769.

Books

12. 西岡憲一(2004)「メチル化ヒストンによるクロマチン構造のエピジェネティクス制御機構」"エピジェネティクス". 佐々木裕之編(シュプリンガー・フェアラーク東京)，pp31-50.
13. 広瀬　進(2004)「転写コファクター」"キーワードで理解する転写イラストマップ". 田村隆明編(羊土社)，pp52-58.

EDUCATION

1. Dr. S. Hirose was invited to give a seminar on "Chromatin remodeling as a barrier against the epigenetic silencing" at National Institute for Basic Biology, Okazaki, July, 2004 (in Japanese).
2. Dr. S. Hirose was invited to give a seminar on "GAGA factor-FACT complex as a boundary against the spreading of silent chromatin" at Tokyo Institute of Technology, Nagatsuda, August, 2004 (in Japanese).

SOCIAL CONTRIBUTION AND OTHERS

評価委員：生物資源研究所

特許

1. 出願番号：2004-315398，発明の名称：負の超らせんDNAの検出法，発明者：広瀬　進・松本国治，出願人：大学共同利用機関法人情報・システム研究機構

C-c. Division of Molecular and Developmental Biology
Koichi Kawakami Group

RESEARCH ACTIVITIES

(1) Gene trap and enhancer trap approaches in zebrafish

Koichi Kawakami, Yasuyuki Kishimoto, Kazuhide Asakawa, Tomoya Kotani, Saori Nagayoshi and Akihiro Urasaki

In order to understand the genetic basis for developmental processes in vertebrate, we have been using a small tropical fish, the zebrafish, as a model animal. Because it is practically possible to breed and maintain very large numbers of fish in the lab, and because zebrafish embryos develop in water, and are transparent, forward genetic approaches (i.e., collecting a large number of mutations affecting developmental processes and analyzing genes responsible for the mutant phenotypes) are feasible in the fish. The *Tol2* element is a transposable element identified from the genome of the medaka fish. Previously we found that the *Tol2* element encodes a fully functional transposase and showed that the *Tol2* element can transpose into the zebrafish genome in the germ lineage. To date, the *Tol2* element is the only natural transposon in vertebrate for which an autonomous element has been identified. We have been interested in developing novel genetic methodologies in zebrafish using the *Tol2* element. Using this *Tol2* transposon system, we have constructed various gene trap and enhancer trap vectors, that contain either a splice acceptor or a basal promoter and a promoterless fluorescent reporter gene (i.e., GFP, RFP, etc). The plasmid DNAs containing these transposon vectors were coinjected in zebrafish fertilized eggs with the transposase mRNA, and offspring from the injected fish were analyzed for expression of the reporter gene in specific tissues and

organs. The integration site of the transposon vector and the gene trapped by the vector insertion were cloned rapidly by PCR-based techniques such as inverse PCR and 5'RACE. These approaches should facilitate our understanding of vertebrate morphogenesis and organogenesis.

(2) Genetic analysis of zebrafish maternal-effect mutations affecting early embryogenesis

Yasuyuki Kishimoto, Sumito Koshida[1], Makoto Furutani-Seiki[2], Atsushi Kawakami[3], Hisato Kondoh[2,4] and Koichi Kawakami ([1]National Institutes of Natural Sciences, [2]Kondoh Differentiation Signaling Project, JST, [3]University of Tokyo, [4]Osaka University)

Maternal-effect genes play essential roles in early embryogenesis in many animals. We have carried out a genetic screen for mutations affecting such maternal-effect genes employing an F3 screen strategy, identifying six recessive mutations out of 60 mutagenized genomes (ref. 6). Among them, four mutations were kept and analyzed. Three of the mutations (*acytokinesis* mutations: ack^{kt5}, ack^{kt62} and ack^{kt119}) prohibited the cell cleavage in the embryos from homozygous females without affecting nuclear divisions. These embryos are defective in generating contractile rings; the ack^{kt62} mutation abolished the organization of cortical F-actin, while other mutations caused abortive contractile ring-like structures at ectopic sites. Defects in contractile ring formation leads to the absence of microtubule arrays at the prospective cleavage plane. Thus, these mutations reveal the sequence of events associated with cytokinesis. It is remarkable that in all acytokinetic embryos, daughter nuclei after mitosis are arranged in spatially normal positions, and maternal *vasa* mRNAs accumulate in the prospective planes of the first and the second cell cleavages despite complete loss of cytokinesis. This indicates that the basic cell architectures of early embryos are largely established by the autonomous activities of the mitotic apparatus, without any dependence on the cell cleavage machinery. The fourth mutation, *bobtail (btl),* caused a strong reduction of the tailbud outgrowth. The expression of *myoD* in somites and *eve1* in tailbud is reduced in the mutant embryos, whereas that of *ntl* in notochord and *pax2.1* in pronephros is comparable between wild-type and mutant embryos, suggesting that the *btl* gene product may regulate gene expression involved in myogenesis and posterior mesodermal development. We also found that the third ventricle in the tectal region is inhibited in the *btl* mutant embryo. In *btl* mutant embryos, expression of MHB genes such as *wnt1, fgf8* and *pax2.1* are once activated at the 8-somite stage, but eliminated from the MHB region afterward. Thus, the *btl* gene product is important for maintenance of the MHB gene expression. *btl* has been mapped on two BAC and PAC clones on chromosome 17. We are currently focusing on identification of the *btl* gene from these genomic clones.

(3) Transposition of the *Tol2* element in mouse embryonic stem cells

Koichi Kawakami and Tetsuo Noda[1] ([1]Tohoku University)

It has not been known whether *Tol2* can transpose in vertebrates other than fish. We investigated transposition of *Tol2* in mouse embryonic stem (ES) cells. We constructed a transposon donor plasmid containing a nonautonomous *Tol2* element with the neomycin resistance gene and a helper plasmid capable of expressing the transposase, and introduced the donor plasmid with various amounts of the helper plasmid by electroporation into mouse ES cells. The number of G418-resistant ES colonies increased as the amount of helper plasmid was increased, in a dose-dependent manner, indicating that the transposase activity elevated the integration efficiency. These G418-resistant ES colonies were cloned and the structure of the junction of the integrated *Tol2* element and the genomic DNA was analyzed by inverse PCR. In those clones, *Tol2* was surrounded by mouse genomic sequences and an 8-bp direct repeat was created adjacent to the both ends of *Tol2,* indicating that *Tol2* was integrated in the genome through transposition. The *Tol2* transposon system is thus active in mouse as well as in fish. We propose that it should be used as a genetic tool to develop novel gene transfer, transgenesis and mutagenesis methods in mammals (ref. 1).

(4) A transposon-mediated gene trap approach identifies developmentally regulated genes in zebrafish

Koichi Kawakami, Hisashi Takeda[1], Noriko Kawakami, Makoto Kobayashi[2], Naoto Matsuda[1] and Masayoshi Mishina[1] ([1]University of Tokyo, and [2]Tsukuba University)

We developed a novel gene trap method in zebrafish using the *Tol2* transposon system. First, we established a highly efficient transgenesis method in which a plasmid DNA containing the *Tol2* transposon vector and the transposase mRNA synthesized in vitro were coinjected into one-cell stage embryos. The transposon vector inserted in the genome could be transmitted to the F1 progeny at high frequencies, and regulated gene expression by a specific promoter could be recapitulated in transgenic fish. Then we constructed a transposon-based gene trap vector containing a splice acceptor and the GFP gene, performed a pilot screen for gene trapping, and obtained fish expressing GFP in temporally and spatially restricted patterns. We confirmed the endogenous transcripts were indeed trapped by the insertions, and the insertion could interfere with expression of the trapped gene. We propose our gene trap approach should facilitate studies of vertebrate development and organogenesis (ref. 2, 5).

(5) Excision of the *Tol2* transposable element of the medaka fish *Oryzias latipes* in *Xenopus laevis* and *Xenopus tropicalis*

Koichi Kawakami, Kosuke Imanaka[1], Mari Itoh[1] and Masanori Taira[1] ([1]University of Tokyo)

We demonstrated transposase-dependent excision of the *Tol2* element in *Xenopus laevis* and *Xenopus (Silurana) tropicalis* embryos. We coinjected a plasmid DNA containing a nonautonomous *Tol2* element and the transposase mRNA synthesized in vitro into two-cell-stage embryos, and analyzed DNA extracted from the injected embryos by polymerase chain reaction (PCR). We demonstrated that the *Tol2* element could be excised from the plasmid DNA in both *X. laevis* and *X. tropicalis* only when it was coinjected with the transposase mRNA. In most cases, a complete loss of the *Tol2* sequence was accompanied by addition of a short DNA sequence to the target sequence, indicating that transposase-dependent excision occurred. While these footprints were characteristic to those created upon excision of transposons of the hAT family, the additional bases found in *Xenopus* were longer and their structures were more complicated than those detected upon excision in zebrafish. This may reflect differences in the activities of host factors involved in either transposition, repair, or both between fish and frog. Our present study suggests that the *Tol2* transposon system should be used as a novel genetic tool to develop transgenesis and mutagenesis methods in *Xenopus* (ref. 3).

(6) The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC)

Gerhard D. S. and 113 others (The MGC project team).

The National Institutes of Health's Mammalian Gene Collection (MGC) project was designed to generate and sequence a publicly accessible cDNA resource containing a complete open reading frame (ORF) for every human and mouse gene. The project initially used a random strategy to select clones from a large number of cDNA libraries from diverse tissues. Candidate clones were chosen based on. 5'-EST sequences, and then fully sequenced to high accuracy and analyzed by algorithms developed for this project. Currently, more than 11,000 human and 10,000 mouse genes are represented in MGC by at least one clone with a full ORF. The random selection approach is now reaching a saturation point, and a transition to protocols targeted at the missing transcripts is now required to complete the mouse and human collections. Comparison of the sequence of the MGC clones to reference genome sequences reveals that most cDNA clones are of very high sequence quality, although it is likely that some cDNAs may carry missense variants as a consequence of experimental artifact, such as PCR, cloning, or reverse transcriptase errors. Recently, a rat cDNA component was added to the project, and ongoing frog (Xenopus) and zebrafish (Danio) cDNA projects were expanded to take advantage of the high-throughput MGC pipeline (ref. 4). We have made a contribution to this project by constructing the zebrafish full-length cDNA library.

PUBLICATIONS

Papers

1. Kawakami, K. and Noda, T. (2004). Transposition of the *Tol2* element, an *Ac*-like element from the Japanese medaka fish *Oryzias latipes,* in mouse embryonic stem cells. Genetics *166,* 895-899.

2. Kawakami, K., Takeda, H., Kawakami, N., Kobayashi, M., Matsuda, N. and Mishina, M. (2004). A transposon-mediated gene trap approach identifies developmentally regulated genes in zebrafish. Developmental Cell *7,* 133-144.

3. Kawakami, K., Imanaka, K., Itoh, M. and Taira, M. (2004). Excision of the *Tol2* transposable element of the medaka fish *Oryzias latipes* in *Xenopus laevis* and *Xenopus tropicalis.* Gene *338,* 93-98.

4. Gerhard, D. S. and 113 others (The MGC project team). (2004). The status, quality, and expansion of the NIH full-length cDNA project: The mammalian gene collection (MGC). Genome Research *14,* 2121-2127.

5. Kawakami, K. (2004). Transgenesis and gene trap methods in zebrafish by using the *Tol2* transposable element. Methods in Cell Biology *77,* 201-222.

6. Kishimoto, Y., Koshida, S., Furutani-Seiki, M. and Kondoh, H. (2004). Zebrafish maternal-effect mutations causing cytokinesis defect without affecting mitosis or equatorial vasa deposition. Mech Dev. *121,* 79-89.

Reviews

7. 川上浩一(2004)ゼブラフィッシュのファンクショナルゲノミクス.「ゲノミクス・プロテオミクスの新展開」今中忠行監修(エヌ・ティー・エス), pp362-372.

8. 小谷友也, 川上浩一(2004)ゼブラフィッシュの時空間特異的な遺伝子発現. 蛋白質核酸酵素*49,* 2111-2116.

SOCIAL CONTRIBUTIONS AND OTHERS

1. K. Kawakami served as a Chair of the 6th International Conference on Zebrafish Development and Genetics held at Madison, U.S.A. on July, 2004.

C-d. Division of Physiological Fumitoshi Ishino Group

RESEARCH ACTIVITIES

哺乳類の個体発生のエピジェネティック的理解/石野史敏
Epigenetic view of mammalian development

Fumitoshi Ishino

Genomic imprinting memories persist in somatic cells after fertilization, and need to be erased and re-established during germ cell development to reflect the sex of the individual. For the study on the re-establishment process of genomic imprinting memory in paternal germ cell development, we analyzed the patterns of each imprinted genes' expression and DNA methylation of DMRs in 3 paternally imprinted regions using embryonic day 10 fetuses each derived from several stage of gonocytes by nuclear transfer, and then could make a detailed profile on the procedure of paternal genomic imprinting memory establishing (in preparation).

Recently, establishment of germline stem (GS) cells and ES-like cells from postnatal mouse testis has been achieved by Kanatsu-Shinohara *et al.*[2]. Analysis of the paternally imprinted regions in these cell lines showed that the former had complete paternal imprinting status but the latter had intermediate status between paternal type and somatic cell type, indicating DNA demetylation occurred in the process from GS cells to ES-like cells[2].

PUBLICATIONS

Papers

1. Kai, M., Irie, M., Okutsu, T., Inoue, K., Ogonuki, N., Miki, H., Yokoyama, M., Suzuki, R., Muguruma, K., Fujita, H., Kohda, T., Ogura, A., Kaneko-Ishino, T. and Ishino, F. (2004). The novel dominant mutation *Dspd* leads to a severe spermatogenesis defect in mice. Biol. Reprod. *70* (5), 1213-1221.

2. Kanatsu-Shinohara, M., Inoue, K., Lee, J., Yoshimoto, M., Ogonuki, N,. Miki, H., Kato, T., Shiro, B., Kazuki, Y., Toyokuni, S., Toyoshima, M., Niwa, O., Oshimura, M., Heike, T., Nakahara, T., Ishino, F., Ogura, A. and Shinohara, T. (2004). Generation of pluripotent stem cells from neonatal mouse testis. Cell *119* (7), 1001-1012.

Reviews

3. 石野史敏, 金児ー石野知子(2004)ゲノムインプリンティングー片親性遺伝子発現の意味 特集 転写制御と疾患ー温故知新 医学のあゆみ *211* (2) 177-181.

4. 金児ー石野知子, 幸田尚, 石野史敏(2004)エピジェネティック制御によるゲノム機能の可塑性 特集 次世代の再生研究に向けた幹細胞の可塑性とそのメカニズム

実験医学　22 (1) 34-38.
5. 金児－石野知子，石野史敏(2004)哺乳類特異的ゲノム機能としての片親性遺伝子発現機構　特集　網羅的解析の先に見えてくるもの　ゲノム解析，プロテオミクスの最前線とその未来　細胞工学　23 (4), 433-438.

Books
6. 金児－石野知子，幸田尚，石野史敏(2004) Springer Reviewシリーズ「エピジェネティクス」3. ゲノムインプリンティング(佐々木裕之編) Springer-Verlarg-Tokyo pp.109-116.
7. 金児－石野知子，石野史敏(2004)エピジェネティクスの種間の相違　分かる実験医学シリーズ「注目のエピジェネティクスがわかる－ゲノムの修飾・構造変換と生命の多様性，疾患との関わり」(押村光雄編)羊土社 pp.97-102.

C-d. Division of Physiological Genetics Yukiko Gotoh Group

RESEARCH ACTIVITIES

(1) The Wnt-β-catenin pathway directs neuronal differentiation of cortical neural precursor cells

Yusuke Hirabayashi, Yasuhiro Itoh, Norihisa Masuyama and Yukiko Gotoh

Neural precursor cells (NPCs) have the ability to self-renew and to give rise to neuronal and glial lineages. The fate decision of NPCs between proliferation and differentiation determines the number of differentiated cells and the size of each region of the brain. However, the signals that regulate the timing of neuronal differentiation remain unclear. Here we show that the Wnt signaling inhibits self-renewal capacity of mouse cortical NPCs, and instructively promotes their neuronal differentiation. Overexpression of Wnt7a or of a stabilized form of β-catenin in mouse cortical NPC culture induced neuronal differentiation even in the presence of FGF2, a self-renewal-promoting factor in this system. Moreover, blockade of the Wnt signaling led to inhibition of neuronal differentiation of cortical NPCs in the developing mouse neocortex. Furthermore, the β-catenin/TCF complex appears to directly regulate the promoter of neurogenin1, a gene implicated in cortical neuronal differentiation. Importantly, stabilized β-catenin did not induce neuronal differentiation of cortical NPCs at earlier developmental stages, consistent with previous reports, suggesting stage-specific functions of the Wnt signaling. Collectively, these results reveal pivotal physiological roles for the Wnt signaling in neuronal differentiation.

(2) Hes binding to STAT3 mediates crosstalk between Notch and JAK-STAT signaling

Sachiko Kamakura, Koji Oishi, Takeshi Yoshimatsu, Norihisa Masuyama and Yukiko Gotoh

Although the Notch and JAK-STAT signaling pathways fulfill overlapping roles in growth and differentiation regulation, no coordination mechanism has been proposed to explain their relationship. Here we show that STAT3 is activated in the presence of active Notch as well as the Notch effectors Hes1 and Hes5. Hes proteins associate with JAK2 and STAT3, and facilitate complex formation between JAK2 and STAT3, thus promoting STAT3 phosphorylation and activation. Furthermore, suppression of endogenous Hes1 expression reduces growth factor induction of STAT3 phosphorylation. STAT3 appears to be essential for maintenance of radial glial cells and differentiation of astrocytes by Notch in the developing central nervous system. These results imply that direct protein-protein interactions coordinate cross-talk between the Notch-Hes and JAK-STAT pathways.

(3) Notch promotes survival of neural precursor cells via mechanisms distinct from those regulating neurogenesis

Koji Oishi, Sachiko Kamakura, Yuko Isazawa, Takeshi Yoshimatsu, Keisuke Kuida, Norihisa Masuyama and Yukiko Gotoh

During development of the mammalian brain, many neural precursor cells (NPCs) undergo apoptosis. The regulation of such cell death, however, is poorly understood. We now show that the survival of mouse embryonic NPCs in vitro was increased by culture at a high cell density and that this effect was attributable to activation of Notch signaling. Expression of an active form of Notch1 thus markedly promoted NPC survival. Hes proteins, key effectors of Notch signaling in inhibition of neurogenesis, were not sufficient for the survival-promoting effect of Notch1. This effect of

Notch1 required a region of the protein containing the RAM domain and was accompanied by up-regulation of the anti-apoptotic proteins Bcl-2 and Mcl-1. Moreover, knockdown of these proteins by RNA interference resulted in blockade of the Notch1-induced survival. These results reveal a new function of Notch, the promotion of NPC survival.

(4) JNK promotes Bax translocation to mitochondria through phosphorylation of 14-3-3 proteins

Fuminori Tsuruta, Jun Sunayama, Yasunori Mori, Norihisa Masuyama and Yukiko Gotoh

Targeted gene disruption studies have established that the c-Jun NH_2-terminal kinase (JNK) is required for the stress-induced release of mitochondrial cytochrome c and apoptosis, and that the Bax subfamily of Bcl-2-related proteins is essential for JNK-dependent apoptosis. However, the mechanism by which JNK regulates Bax has remained unsolved. Here we demonstrate that activated JNK promotes Bax translocation to mitochondria through phosphorylation of 14-3-3, a cytoplasmic anchor of Bax. Phosphorylation of 14-3-3 led to dissociation of Bax from this protein. Expression of phosphorylation-defective mutants of 14-3-3 blocked JNK-induced Bax translocation to mitochondria, cytochrome c release and apoptosis. Collectively, these results have revealed a key mechanism of Bax regulation in stress-induced apoptosis.

PUBLICATIONS

Papers

1. Hirabayashi, Y., Itoh, Y., Tabata, H., Nakajima, K., Akiyama, T., Masuyama, N. and Gotoh, Y. (2004). The Wnt-beta-catenin pathway directs neuronal differentiation of cortical neural precursor cells. Development, *131,* 2791-2801.
2. Kamakura, S., Oishi, K., Yoshimatsu, T., Nakafuku, M., Masuyama, N. and Gotoh, Y. (2004). Hes binding to STAT3 mediates crosstalk between Notch and JAK-STAT signaling. Nat. Cell Biol. *6,* 547-554.
3. Miyagi, S., Saito, T., Mizutani, K., Masuyama, N., Gotoh, Y., Iwama, A., Nakauchi, H., Masui, S., Niwa, H., Nishimoto, M., Muramatsu, M. and Okuda, A. (2004). The sox-2 regulatory regions display their activities in two distinct multipotent stem cells. Mol. Cell. Biol., *24,* 4207-4220.
4. Oishi, K., Kamakura, S., Isazawa, Y., Yoshimatsu, T., Kuida, K., Nakafuku, M., Masuyama, N. and Gotoh, Y. (2004). Notch promotes survival of neural precursor cells via mechanisms distinct from those regulating neurogenesis. Dev. Biol., *276,* 172-184.
5. Tsuruta, F., Sunayama, J., Mori, Y., Shimizu, S., Tsujimoto, Y., Yoshioka, K., Masuyama, N. and Gotoh, Y. (2004). JNK promotes Bax translocation to mitochondria through phosphorylation of 14-3-3 proteins. EMBO J. *23,* 1889-1899.
6. Sunayama, J., Ando, Y., Itoh, N., Tomiyama, A., Sakurada, K., Sugiyama, A., Kang, D., Tashiro, F., Gotoh, Y., Kuchino, Y. and Kitanaka, C. (2004). Physical and functional interaction between BH3-only protein Hrk and mitochondrial pore-forming protein p32. Cell Death Differ. *11,* 771-781.
7. Ogihara, T., Asano, T., Katagiri, H., Sakoda, H., Anai, M., Shojima, N., Ono, H., Fujishiro, M., Kushiyama, A., Fukushima, Y., Kikuchi, M., Noguchi, N., Aburatani, H., Gotoh, Y., Komuro, I. and Fujita, T. (2004). Oxidative stress induces insulin resistance by activating the nuclear factor-B pathway and disrupting normal subcellular distribution of phosphatidylinositol 3-kinase. Diabetologia *5,* 794-805.

Reviews

8. 大石康二・後藤由季子(2004)脳発生と細胞死，実験医学　増刊，*22* (11), pp1634-1639.
9. 大橋淳一郎，後藤由季子(2004)生存シグナル，生体の科学，*55* (5)，pp450-451.
10. 鎌倉幸子・吉松剛志，後藤由季子(2004)神経幹細胞の運命を制御するNotch-Hes経路とJAK-STAT3経路のクロストーク，実験医学，*22* (15), pp2167-2170.

SOCIAL CONTRIBUTIONS AND OTHERS

Oncogene (Editor)
Journal of Biochemistry (Associate Editor)
Cell Structure and Function (Associate Editor)

各賞受賞
日本癌学会奨励賞受賞
日本女性科学者の会奨励賞受賞

D. DEPARTMENT OF POPULATION GENETICS

D-a. Division of Population Genetics Naruya Saitou Group

RESEARCH ACTIVITIES

Saitou Naruya, Professor

(1) DNA sequence and comparative analysis of chimpanzee chromosome 22

The International Chimpanzee Chromosome 22 Consortium [Hidemi Watanabe[1,2], Choong-Gon Kim, Satoshi Oota, Takashi Kitano, Yuji Kohara, Naruya Saitou and Yoshiyuki Sakaki[2]] ([1]Nara Institute of Science, and Technology, [2]RIKEN,Genomic Sciences Center)

Human-chimpanzee comparative genome research is essential for narrowing down genetic changes involved in the acquisition of unique human features, such as highly developed cognitive functions, bipedalism or the use of complex language. Here, we report the high-quality DNA sequence of 33.3 megabases of chimpanzee chromosome 22. By comparing the whole sequence with the human counterpart, chromosome 21, we found that 1.44% of the chromosome consists of single-base substitutions in addition to nearly 68,000 insertions or deletions. These differences are sufficient to generate changes in most of the proteins. Indeed, 83% of the 231 coding sequences, including functionally important genes, show differences at the amino acid sequence level.

Furthermore, we demonstrate different expansion of particular subfamilies of retrotransposons between the lineages, suggesting different impacts of retrotranspositions on human and chimpanzee evolution. The genomic changes after speciation and their biological consequences seem more complex than originally hypothesized. For details, see ref. (1)

(2) Human specific amino acid changes found in 103 protein coding genes. Molecular Biology and Evolution

Takashi Kitano, Yuhua Liu[1], Shintaroh Ueda[2] and Naruya Saitou ([1]The Jackson Laoratory, USA, [2]Department of Biological Sciences, Graduate School of Science, University of Tokyo)

We humans have many characteristics that are different from those of the great apes. These human-specific characters must have arisen through mutations accumulated in the genome of our direct ancestor after the divergence of the last common ancestor with chimpanzee. Gene trees of human and great apes are necessary for extracting these human-specific genetic changes. We conducted a systematic analysis of 103 protein-coding genes for human, chimpanzee, gorilla, and orangutan. Nucleotide sequences for 18 genes were newly determined for this study, and those for the remaining genes were retrieved from the DDBJ/EMBL/GenBank database. The total number of amino acid changes in the human lineage was 147 for 26,199 codons (0.56%). The total number of amino acid changes in the human genome was, thus, estimated to be about 80,000. We applied the acceleration index test and Fisher's synonymous/nonsynonymous exact test for each gene tree to detect any human-specific enhancement of amino acid changes compared with ape branches. Six and two genes were shown to have significantly higher nonsynonymous changes at the human lineage from the acceleration index and exact tests, respectively. We also compared the distribution of the differences of the nonsynonymous substitutions on the human lineage and those on the great ape lineage. Two genes were more conserved in the ape lineage, whereas one gene was more conserved in the human lineage. These results suggest that a small proportion of protein-coding genes started to evolve differently in the human lineage after it diverged from the ape lineage. For details, see ref. (2)

(3) Genetic variation versus recombination rate in a structured population of mice

Aya Takahashi, Yuhua Liu[1] and Naruya Saitou ([1]The Jackson Laoratory, USA)

The correlation between genetic variation and

recombination rate was investigated in a structured mouse population. Nucleotide sequence data from 19 autosomal DNA loci from eight inbred strains of mouse (Mus musculus) sampled from three major subspecies were analyzed. The recombination rate was estimated from the comparison of genetic and physical map distances between markers flanking a 10-cM region of each locus. The strains were categorized into four groups (subpopulations) based on geography. By partitioning the genetic diversity into within-group and among-group variation, we detected a positive correlation between the recombination rate and nucleotide diversity within groups. The level of nucleotide differentiation among groups (G(ST)) showed a negative correlation with the rate of recombination. There was no significant correlation between recombination rate and nucleotide diversity when data from different subpopulations were pooled. No correlation was detected between recombination rate and nucleotide divergence of M. musculus and M. spicilegus. These patterns deviate from the strict neutral expectation under the constant nucleotide substitution rate, and they are likely to have been formed either by a hitchhiking effect of positively selected mutants or by background selection of deleterious mutants occurring in a subdivided population. Our series of comparisons show that because a real population always has some structure, incorporation of its information is important in detecting non-neutral evolution. For details, see ref. (3)

(4) Phylogenetic analysis of proteins associated in four major energy metabolism systems: photosynthesis, oxidative phosphorylation, nitrogen metabolism and sulfur metabolism

Takeshi Tomiki and Naruya Saitou

The four electron transfer energy metabolism systems, photosynthesis, aerobic respiration, denitrification, and sulfur respiration, are thought to be evolutionarily related because of the similarity of electron transfer patterns and the existence of some homologous proteins. How these systems have evolved is elusive. We therefore conducted a comprehensive homology search using PSI-BLAST, and phylogenetic analyses were conducted for the three homologous groups (groups 1-3) based on multiple alignments of domains defined in the Pfam database. There are five electron transfer types important for catalytic reaction in group 1, and many proteins bind molybdenum. Deletions of two domains led to loss of the function of binding molybdenum and ferredoxin, and these deletions seem to be critical for the electron transfer pattern changes in group 1. Two types of electron transfer were found in group 2, and all its member proteins bind siroheme and ferredoxin. Insertion of the pyridine nucleotide disulfide oxidoreductase domain seemed to be the critical point for the electron transfer pattern change in this group. The proteins belonging to group 3 are all flavin enzymes, and they bind flavin adenine dinucleotide (FAD) or flavin mononucleotide (FMN). Types of electron transfer in this group are divergent, but there are two common characteristics. NAD(P)H works as an electron donor or acceptor, and FAD or FMN transfers electrons from/to NAD(P)H. Electron transfer functions might be added to these common characteristics by the addition of functional domains through the evolution of group 3 proteins. Based on the phylogenetic analyses in this study and previous studies, we inferred the phylogeny of the energy metabolism systems as follows: photosynthesis (and possibly aerobic respiration) and the sulfur/nitrogen assimilation system first diverged, then the sulfur/nitrogen dissimilation system was produced from the latter system. For details, see ref. (4)

(5) Mitochondrial DNA Genealogy of Chimpanzees in Nimba Mountains and Bossou, West Africa

Makoto Shimada, Sachiko Hayakawa[1], Tatyana Humle[2], Shiho Fujita[3], Satoshi Hirata[4], Yukimaru Sugiyama[5] and Naruya Saitou ([1]Primate Research Institute, Kyoto University, [2]University of Stirling United Kingdom, [3]Faculty of Agriculture, Gifu University, [4]Hayashibara Biochemical Laboratories Inc., [5]Tokai-Gakuen University)

The chimpanzee populations of the Bossou and Nimba regions in West Africa were genetically surveyed to 1) reveal the genetic relationship between the Bossou and Nimba populations, and 2) elucidate the evolutionary relationship between the Bossou-Nimba and other West African populations. The chimpanzee group at Bossou is characterized by its small population size, no evidence of contact with neighboring populations, and no female immigration. It is believed that most females and adolescent males

emigrate from this population. To reveal the genetic signature of these characteristics, we examined the genetic diversity of Bossou and two neighboring populations (Seringbara and Yeale) in the Nimba Mountains by sequencing approximately 605 bp of the mitochondrial DNA (mtDNA) control region. A total of 20 distinct mtDNA variants were observed from 56 sequences of noninvasively collected, anonymous samples. Nucleotide diversity in the Nimba Mountain populations was 0.03-0.04, and did not differ significantly from that in the Bossou population. Very few mitochondrial variants are shared among the sites sampled, which suggests that there is little gene flow involving mtDNA. Nevertheless, no clear population structures were revealed in either population. A comparison with published sequences from West African chimpanzees (Pan troglodytes verus) indicates that the variants observed in the Bossou and Nimba regions are scattered throughout the subspecies, rather than clustered according to geographic region. This suggests that the Bossou-Nimba populations derived only recently from the common ancestral population of the West African chimpanzees, and did not pass through a bottleneck. For details, see ref. (5)

(6) Evolution of O alleles of the human ABO blood group gene

Francis Roubinet[1], Stephanie Despiau[1], Francesc Calafell[2], Fen Jin[3], Jaume Bertanpetit[2], Naruya Saitou and Antoine Blancher[1] ([1]Paul Sabatier University, Rangueil Hospital, [2]Universitat Pompeu Fabra, [3]The Institute of Genetics and Developmental Biology)

To date, at least 40 different alleles O have been characterized on the basis of exon 6 and exon 7 sequences but not always for intron 6. STUDY DESIGN AND METHODS: Among 415 individuals, from four continents (Africa, Europe, South America, and Asia), studied for exon 6 and exon 7 sequences, we selected 46 individuals (of respective phenotypes O [39], AB [3], B [3], or A [1]) for sequencing 1800-bp amplicons spanning exon 6, intron 6, and exon 7. The amplicons were characterized either by direct sequencing or after cloning when required. We defined 14 new intron 6 O allele sequences, including four recombinant alleles. Based on sequence comparison, a phylogenetic network was constructed. It confirmed recombinant allele origins and that most O alleles are derived by point mutations from the two worldwide distributed alleles O01 and O02. CONCLUSION: Allele O phylogenetic analysis suggests that the most frequent silencing mutation (deletion of a G in exon 6) appeared once in human evolution in the ancient O02 allele lineage and that allele O01 resulted from an interallele exchange between O02 and A101. Assuming constancy of evolutionary rate, diversification of the representative alleles of the three human ABO lineages (A101, B101, and O02) was estimated at 4.5 to 6 million years ago. For details, see ref. (6)

(7) Comparative genetics of functional trinucleotide tandem repeats in humans and apes

Aida Andres[1], Marta Soldevila[1], Oscar Lao[1], Victor Volpini[1], Nayuya Saitou, Howard Jacobs[1], Ikuo Hayasaka[2], Francesc Calafell[1] and Jaume Bertranpetit[1] ([1]Universitat Pompeu Fabra, [2]Kumamoto Primate Park, Sanwa Kagaku Kenkyusho)

Several human neurodegenerative disorders are caused by the expansion of polymorphic trinucleotide repeat regions. Many of these loci are functional short tandem repeats (STRs) located in brain-expressed genes, and their study is thus relevant from both a medical and an evolutionary point of view. The aims of our study are to infer the comparative pattern of variation and evolution of this set of loci in order to show species-specific features in this group of STRs and on their potential for expansion (therefore, an insight into evolutionary medicine) and to unravel whether any human-specific feature may be identified in brain-expressed genes involved in human disease. We analyzed the variability of the normal range of seven expanding STR CAG/CTG loci (SCA1, SCA2, SCA3-MJD, SCA6, SCA8, SCA12, and DRPLA) and two nonexpanding polymorphic CAG loci (KCNN3 and NCOA3) in humans, chimpanzees, gorillas, and orangutans. The study showed a general conservation of the repetitive tract and of the polymorphism in the four species and high heterogeneity among loci distributions. Humans present slightly larger alleles than the rest of species but a more relevant diference appears in variability levels: Humans are the species with the largest variance, although only for the expanding loci, suggesting a relationship between variability levels and expansion potential. The sequence analysis shows high levels of sequence

conservation among species, a lack of correspondence between interruption patterns and variability levels, and signs of conservative selective pressure for some of the STR loci. Only two loci (SCA1 and SCA8) show a human specific distribution, with larger alleles than the rest of species. This could account, at the same time, for a human-specific trait and a predisposition to disease through expansion. For details, see ref. (7)

(8) Polymorphisms in the Trace Amine Receptor 4 (TRAR4) Gene on Chromosome 6q23.2 Are Associated with Susceptibility to Schizophrenia

Jubao Duan[1], Alan Sanders[1], Cuiping Hou[1], Naruya Saitou, Takashi Kitano and Pablo Gejman[1] ([1]Department of Psychiatry and Behavioral Sciences, Northwestern University)

Several linkage studies across multiple population groups provide convergent support for a susceptibility locus for schizophrenia-and, more recently, for bipolar disorder-on chromosome 6q13-q26. We genotyped 192 European-ancestry and African American (AA) pedigrees with schizophrenia from samples that previously showed linkage evidence to 6q13-q26, focusing on the MOXD1-STX7-TRARs gene cluster at 6q23.2, which contains a number of prime candidate genes for schizophrenia. Thirty-one screening single-nucleotide polymorphisms (SNPs) were selected, providing a minimum coverage of at least 1 SNP/20 kb. The association observed with rs4305745 (P=.0014) within the TRAR4 (trace amine receptor 4) gene remained significant after correction for multiple testing. Evidence for association was proportionally stronger in the smaller AA sample. We performed database searches and sequenced genomic DNA in a 30-proband subsample to obtain a high-density map of 23 SNPs spanning 21.6 kb of this gene. Single-SNP analyses and also haplotype analyses revealed that rs4305745 and/or two other polymorphisms in perfect linkage disequilibrium (LD) with rs4305745 appear to be the most likely variants underlying the association of the TRAR4 region with schizophrenia. Comparative genomic analyses further revealed that rs4305745 and/or the associated polymorphisms in complete LD with rs4305745 could potentially affect gene expression. Moreover, RT-PCR studies of various human tissues, including brain, confirm that TRAR4 is preferentially expressed in those brain regions that have been implicated in the pathophysiology of schizophrenia. These data provide strong preliminary evidence that TRAR4 is a candidate gene for schizophrenia; replication is currently being attempted in additional clinical samples. For details, see ref. (8)

Kenta Sumiyama, Assistant Professor

(9) Evolutionary functional genomic analysis of cis-regulation in mammalian Dlx gene system

Kenta Sumiyama

Dlx genes are important regulators for mammalian early development. Mammalian Dlx genes has been duplicated through evolution and resulted in 6 genes in 3 clusters that have overlapping but also distinct functions. In branchial arches three Dlx gene clusters are expressed in nested pattern that is essential to establish positional information in jaw patterning. This nested expression pattern is found only in higher vertebrates so that studying genomic regulation mechanism through cis-regulatory elements would be of great interest in mammalian jaw evolution. Recent expansion of genome sequence information enables us to compare intergenic noncoding genomic sequences among different species. Although intergenic regions have been sometimes called "junk DNA" due to lack of knowledge about their functional importance, a large number of sequence conservations have been found in intergenic regions through comparative genomic sequence analysis. Such conservations are likely to be functional elements that are repressing, initiating or maintaining gene expressions. I focus on genomic analysis of cis-regulatory elements in Dlx3-7 gene clusters by using comparative genomics and further functional analysis by using transgenic mouse system.

PUBLICATIONS

Papers

1. The International Chimpanzee Chromosome 22 Consortium [Watanabe, H., Fujiyama, A., Hattori, M., Taylor, T.D., Toyoda, A., Kuroki, Y., Noguchi, H., Ben, Kahla, A., Lehrach, H., Sudbrak, R., Kube, M., Taenzer, S., Galgoczy, P., Platzer, M., Scharfe, M., Nordsiek, G., Blocker, H., Hellmann, I., Khaitovich, P., Paabo, S., Reinhardt, R., Zheng, H.-J., Zhang, X.-L., Zhu, G.-F., Wang, B.-F., Fu, G., Ren9, S.-X., Zhao, G.-P., Chen, Z.,

Lee, Y.-S., Cheong, J.-E., Choi, S.-H., Wu, K.-M., Liu, T.-T., Hsiao, K.-J., Tsai, S.-F., Kim, C.-G., Oota, S., Kitano, T., Kohara, Y., Saitou, N., Park, H.-S., Wang, S.-Y., Yaspo, M.-L. and Sakaki, Y.] (2004). DNA sequence and comparative analysis of chimpanzee chromosome 22. Nature, *429,* 382-388.

2. Kitano, T., Liu, Y.-H., Ueda, S. and Saitou, N. (2004). Human specific amino acid changes found in 103 protein coding genes. Molecular Biology and Evolution. *21,* 936-944.

3. Takahashi, A., Liu, Y.-H. and Saitou, N. (2004). Genetic variation versus recombination rate in a structured population of mice. Molecular Biology and Evolution, *21,* 404-409.

4. Tomiki, T. and Saitou, N. (2004). Phylogenetic analysis of proteins associated in four major energy metabolism systems: photosynthesis, oxidative phosphorylation, nitrogen metabolism and sulfur metabolism. Journal of Molecular Evolution, *59,* 158-176.

5. Shimada, M., Hayakawa, S., Hamle, T., Fujita, S., Hirata, S., Sugiyama, Y. and Saitou, N. (2004). Mitochondrial DNA Genealogy of Chimpanzees in Nimba Mountains and Bossou, West Africa. American Journal of Primatology, *64,* 261-275.

6. Roubinet, F., Despiau, S.-P., Calafell, F., Jin, F., Bertanpetit, J., Saitou, N. and Blancher, A. (2004). Evolution of O alleles of the human ABO blood group gene. Transfusion, *44,* 707-715.

7. Andres, A.M., Soldevila, M., Lao, O., Volpini, V., Saitou, N., Jacobs, H.T., Hayasaka, I., Calafell, F. and Bertranpetit, J. (2004). Comparative genetics of functional trinucleotide tandem repeats in humans and apes. Journal of Molecular Evolution, *59,* 329-339.

8. Duan, J., Martinez, M., Sanders, A.R., Hou, C., Saitou, N., Kitano, T., Mowry, B.J., Crowe, R.R., Silverman, J.M., Levinson, D.F. and Gejman, P.V. (2004). Polymorphisms in the Trace Amine Receptor 4 (TRAR4) Gene on Chromosome 6q23.2 Are Associated with Susceptibility to Schizophrenia. American Journal of Human Genetics, *75,* 624-638.

9. Abe, K., Noguchi, H., Tagawa, K., Yuzuriha, M., Toyoda, A., Kojima, T., Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriwaki, K. and Shiroishi, T. (2004). Contribution of Asian mouse subspecies Mus musculus molossinus to genomic constitution of strain C57BL/6J, as defined by BAC end sequence-SNP analysis. Genome Research, *14,* 2239-2247.

10. Imanishi, T. and many other authors including Saitou, N. and Sugano, S. (2004). Integrative annotation of 21,037 human genes validated by full-length cDNA clones. PLoS Biology, *2,* 1-20.

11. Nakajima, T., Wooding, S., Sakagami, T., Emi, M., Tokunaga, K., Tamiya, G., Ishigami, T., Umemura, S., Munkhbat, B., Jin, F., Guan-Jun, J., Hayasaka, I., Ishida, T., Saitou, N., Pavelka, K., Lalouel, J.M., Jorde, L.B. and Inoue, I. (2004). Natural selection and population history in the human angiotensinogen gene (AGT): 736 complete AGT sequences in chromosomes from around the world. American Journal of Human Genetics, *74,* 898-916.

12. 斎藤成也(2004)近隣結合法をはじめとする距離行列法. 連載「ゲノム進化学の展開」第4回. 数理科学, 42巻12号. 60-67頁.

13. 斎藤成也(2004)人類進化学の将来像. 科学, 74巻10号, 1246-1249頁.

14. 斎藤成也(2004)系統樹とは(その2). 連載「ゲノム進化学の展開」第3回. 数理科学, 42巻10号, 72-78頁.

15. 斎藤成也(2004)『ゲノムと進化ーゲノムから立ち昇る生命ー』. ワードマップシリーズ, 新曜社(9月10日発行).

16. 斎藤成也(2004)言語能力の遺伝的基礎. 大航海, 52号, 114-121頁.

17. 斎藤成也(2004)系統樹とは(その1). 連載「ゲノム進化学の展開」第2回. 数理科学, 42巻9号, 68-74頁.

18. 瀬名秀明・斎藤成也(2004)類人猿のゲノムで探る人間らしさの起源(対談). 瀬名秀明編『科学の最前線で研究者は何を見ているのか』, 52-70頁. (日経サイエンス, 2003年6月号, 94-99頁所収より)

19. 斎藤成也・隅山健太(2004)Hox遺伝子クラスターの進化ー脊椎動物の誕生からヒトまでー. 実験医学, Vol.22, No.12(8月号), 1677-1683頁.

20. 斎藤成也(2004)ネアンデルタールシンポジウムに参加して. GSJコミュニケーションズ, 79巻2号, 7-9頁.

21. 北野誉・斎藤成也(2004)類人猿とヒトのゲノム進化研究. 生体の科学, [特集]分子進化学の現在, 55巻3号, 252-256頁.

22. 斎藤成也(2004)はじめに. 連載「ゲノム進化学の展開」第1回. 数理科学, 42巻7号, 64-70頁.

23. 隅山健太・斎藤成也(2004)HOX遺伝子. 榊佳之・笹月健彦・油谷浩幸編『ヒトゲノムー生命システムの理解と医学への展開ー』, 91-93頁. 中山書店.

24. Lubert Stryer著, 入村達郎ら共訳(2004)『ストライヤー生化学(第5版)』. 2章・7章の翻訳. 東京化学同人.

EDUCATION

1. 斎藤成也：福井大学医学部法医学非常勤講義「中立進

化とゲノムの進化」，3月
2. 斎藤成也：総合研究大学院大学先導科学研究科生命体科学専攻集中講義「系統樹解析の理論と問題点について」・「ヒトと類人猿の比較ゲノム解析」，5月
3. 斎藤成也：東京大学理学部講義「分子進化学」5月，6月
4. 斎藤成也：熊本大学理学部集中講義「遺伝子とゲノムから見た生物の進化」，11月
5. 斎藤成也：埼玉大学理学部集中講義「遺伝子の進化と生物の進化」，12月

SOCIAL CONTRIBUTIONS AND OTHERS

1. 斎藤成也：Molecular Biology and Evolution, Associate Editor
2. 斎藤成也：GENE-Evolutionary Genetics, Editor
3. 斎藤成也：シリーズ進化学全7巻(岩波書店)　編集委員
4. 斎藤成也：日本遺伝学会　会計幹事
5. 斎藤成也：日本DNA多型学会　評議員
6. 斎藤成也：日本人類学会　評議員
7. 斎藤成也：財団法人遺伝学普及会　監事
8. 斎藤成也：公益信託進化学振興木村資生基金　運営委員
9. 斎藤成也：木原記念横浜生命科学振興財団より　第12回木原記念財団学術賞　受賞

D-a. Division of Population Genetics
Toshiyuki Takano Group

RESEARCH ACTIVITIES

(1) Inter-locus nonrandom association of polymorphisms in *Drosophila* chemoreceptor genes

Toshiyuki Takano-Shimizu, Yuriko Ishii, Akira Kawabe[1], Masanobu Itoh[2], Nobuyuki Inomata[3], Rumi Kondo[4] and Yutaka Inoue[5] ([1]University of Edinburgh, [2]Kyoto Institute of Technology, [3]Kyushu University, [4]Ochanomizu University, [5]Osaka University of Foreign Studies)

Some forms of multilocus selection with epistasis, such as truncation selection, can effectively reduce the mutation load. Many quantitative characters, including complex genetic diseases, are likely to be subject to these types of selection. However, direct measurement of selection in natural populations is difficult and the effect of epistasis on within-species variations remains unclear. Epistatic interaction in the fitness effect can generate linkage disequilibrium (LD). Therefore, we may detect the action of natural selection from its amount and pattern. We found a large number of inter-locus nonrandom associations between polymorphisms in 98 *Drosophila* chemoreceptor genes. LD was examined in two fly samples collected at the same location, but in different seasons. The amount of LD was much larger in the spring sample than in the autumn one. This between-sample difference was much more striking for the replacement polymorphisms than for the silent polymorphisms. The difference between the replacement and silent polymorphisms could not be attributed to differences in the mean marker distances. We also found a significant excess of associations between one frequent and one less common allele for the replacement polymorphisms, but not for the silent polymorphisms. It is unlikely that a simple seasonal bottleneck could explain all these differences in the scale of LD between the samples and between the replacement and silent polymorphisms. Natural selection is suggested to play a significant role in shaping the pattern of LD observed in this study. For details, see Paper 4.

(2) A high frequency null mutant of an odorant-binding protein, *Obp57e,* in *Drosophila melanogaster*

Aya Takahashi and Toshiyuki Takano-Shimizu

We have found a null mutant of an odorant-binding protein, *Obp57e,* in *Drosophila melanogaster.* This frameshift mutation, which is a 10 bp deletion in the coding region, is at a high frequency in the Kyoto population, and is also present in Taiwan and Africa. We have sequenced an about 1.5kb region including the tandemly duplicated gene, *Obp57d,* from 16 inbred lines sampled in Kyoto, Japan. The analyses showed a peak of nucleotide diversity and strong linkage disequilibrium around this mutation. This pattern suggests an elevated mutation rate or an influence of balancing selection in this region. The level of nucleotide divergence between *D. melanogaster* and *D. simulans* does not support the former possibility. Thus, this presence/absence polymorphism may be due to balancing selection, which takes advantage of the relatively weak functional constraint in members of a large gene family. In addition, the *Obp57d* gene

region showed an excess of high frequency derived mutants that is consistent with a pattern predicted under positive natural selection.

(3) Analyses of the mechanism and genetics of mate choice by *Drosophila melanogaster* males

Aya Takahashi and Toshiyuki Takano-Shimizu

We surveyed multiple strains of *D. melanogaster* and *D. simulans* for differences in heterospecific copulation frequencies, and identified two *D. melanogaster* inbred lines between which mating frequencies of males against *D. simulans* females differ by a large extent. One line was originated in West Africa and the other was collected in Taiwan. We also performed an intraspecific mating experiment and found that males from these two lines have different mating preference against conspecific females of different lines. A cuticular hydrocarbon transfer experiment revealed that this preference is likely to be affected by the differences in cuticular hydrocarbon components of the females. These two lines have been subjected to over 20 generations of sib-mating, and are suitable materials for genetic mapping. Our genetic analyses showed that the third chromosome has a major effect on these behavioral differences of males.

(4) Null alleles at Drosophila chemoreceptor genes

Akira Kawabe[1], Yuriko Ishii and Toshiyuki Takano-Shimizu ([1]University of Edinburgh)

We determined a total of about 85-kb sequence of 110 chemoreceptor genes of seven *D. melanogaster* strains and found 16 putative null alleles at 13 loci: 4 nonsense mutations, 7 small insertions/deletions causing frame shifts, and 5 long deletions encompassing parts of coding sequence. The seven strains analyzed in the present study had at least one null allele and, in total, 27 null alleles. Because we determined only one-half of the coding sequences, there might be eight null alleles per haploid genome of *D. melanogaster*. Importantly, the ratio (π_R/π_S) of the average nucleotide diversity (π) at the replacement sites to that at the synonymous sites suggested that the null alleles did not necessarily appear only in very low constraint loci, although the *Or85e* seems to be a pseudogene. On the other hand, the frequencies of most of these putative null alleles in a natural population are not very low: 36% of a deletion in *Or98b*, 30% of an insertion in *Gr2a*, 1% of a nonsense mutation in *Gr22b*, 10% of a nonsense mutation in *Gr36c*, 39% of a deletion in *Gr39aA*, 8% of a nonsense mutation in *Gr47a*, 12% of a nonsense mutation in *Gr58a* and 4% of a deletion in *Gr85a*. Then, the frequencies of null homozygotes are not expected to be very low. In sum, the results suggested that selection is weak but not completely relaxed in the *Or* and *Gr* genes. This condition may allow rapid evolution of the chemoreceptor genes. We found a single putative null mutation in 54 loci of a single *D. simulans* strain, which did not include *Or85e*. The expected number of null alleles at the chemoreceptor genes in *D. simulans* was comparable to that in *D. melanogaster*.

(5) Genetic and molecular basis of sex-comb-tooth-number variation in *Drosophila*

Haruki Tatsuta[1], Yuriko Ishii and Toshiyuki Takano-Shimizu ([1]National Institute of Environmental Studies)

The sex comb on the forelegs of *Drosophila* males is a secondary sexual trait, and the number of teeth on these combs varies greatly within and between species. To understand the relationship between the intra- and interspecific variations, we did quantitative trait loci (QTL) analyses of intra-specific sex-comb-tooth-number variations in five mapping populations with two inbred *Drosophila simulans* strains of large and small numbers of sex-comb teeth (whole-genome F_2, two second-chromosome substitution strains, and two third-chromosome congenic strains). From these mapping studies, we identified three major QTLs on the third chromosome and one of them likely matched the QTL responsible for the *simulans-mauritiana* differences. We focused our analyses on this region. Linkage disequilibrium survey for association study is now under way.

PUBLICATIONS

Papers

1. Kobayashi, S., Noro, Y., Nagano, H., Yoshida, K., Takano-Shimizu, T., Kishima, Y. and Sano, Y. Evidence for an evolutionary force that prevents epigenetic silencing between tail-to-tail rice genes with

a short spacer. Gene (in press).
2. Takahashi, A., Liu, Y.-H. and Saitou, N. (2004). Genetic variation versus recombination rate in a structured population of mice. Mol. Biol. Evol. *21,* 404-409.
3. Takahashi, A. and Ting, C.-T. (2004). Genetic basis of sexual isolation in *Drosophila melanogaster.* Genetica *120,* 273-284.
4. Takano-Shimizu, T., Kawabe, A., Inomata, N., Nanba, N., Kondo, R., Inoue, Y. and Itoh, M. (2004). Inter-locus nonrandom association of polymorphisms in *Drosophila* chemoreceptor genes. Proc. Natl. Acad. Sci. USA *101,* 14156-14161.

EDUCATION

1. Dr. T. Takano gave a lecture at The Graduate University for Advanced Studies, April, 2004 (in Japanese).
2. Dr. T. Takano was invited to give a lecture in "Shonan Lecture: Diversity and Evolution of Populations" at The Graduate University for Advanced Studies, August, 2004.
3. Dr. T. Takano organized a Collaborative Research Meeting at the National Institute of Genetics, Mishima, June, 2004.
4. Dr. T. Takano gave a two-day training course for Mishima Kita Junior High School students, October, 2004.

SOCIAL CONTRIBUTIONS AND OTHERS

1. Dr. T. Takano served as a member of the steering committee of National Bio-Resource Project "DROSOPHILA".
2. Drs. T. Takano and A. Takahashi served as members of the election administration committee of the Genetics Society of Japan.
3. Dr. T. Takano served as a member of the examination specialist committee for National Personnel Authority.

D-b. Division of Evolutionary Genetics (Until May) Toshimichi Ikemura Group

RESEARCH ACTIVITIES

(1) Phylogenetic classification of environmental DNA sequences including those from the Sargasso Sea and of rDNA sequences with Self-Organization Map (SOM)

Takashi Abe, Jian-Ping Song and Toshimichi Ikemura

The self-organizing map (SOM) is a powerful tool for clustering and visualizing high-dimensional complex data on a two-dimensional plane. Oligonucleotide frequency is an example of high-dimensional data, and we developed an SOM as a novel bioinformatics strategy to capture and visualize phylogenetic diversity of a wide variety of microbial genomic sequences obtained from an environmental sample. First we constructed SOMs for tetranucleotide frequencies in approximately 200,000 5-kb sequence fragments obtained from 1500 prokaryotes for which at least 10 kb of genomic sequence has been deposited in DDBJ/EMBL/GenBank (a total of 1.05 Gb). The sequences could be classified primarily according to 25 phylogenetic groups without information regarding the species. The classification was possible without orthologous sequence sets, and therefore, is especially useful for phylogenetic classification of novel sequences from poorly characterized species in environmental and clinical samples. We used the SOM method to classify 810,000 sequences recently reported by Venter *et al.* for pooled DNA samples from the Sargasso Sea near Bermuda. Phylogenetic diversity and novelty of the Sargasso Sea sequences was visualized on a single map, and sequences that were derived from a single genome but cloned independently could be reconstructed *in silico.*

Phylogenetic classification of genomic sequence fragments from a groundwater metagenome library and a human gut library was conducted using the tetranucleotide SOM. (For details, see *Nature Biotechnology,* 2005: *Canadian Journal of Microbiology,* 2005)

Because 16S rDNA sequences were highly conserved during evolution, their sequences have been used for detailed phylogenetic classification of prokaryotic species, including uncultured prokaryotes.

Approximately 20,000 16S rDNA sequences longer than 1 kb from 6,100 known prokaryotic species have been compiled in DDBJ/EMBL/GenBank. We constructed a tetranucleotide SOM with these 16S rDNA sequences after normalization for the sequence lengths. Clear clustering according to phylogenetic group was observed; 97% of sequences were classified into the correct group territory on the tetranucleotide SOM. The finding that the hit level of 16S rDNA classification into the correct group territory was higher than that of genomic sequences may indicate that the occurrence of horizontal transfer of rDNAs, if present, is lower than that of other genome portions. Combination of SOMs for genomic and 16S rDNA sequences will provide a tool for detailed phylogenetic studies of genomic sequence fragments from environmental uncultured prokaryotes.

(2) SOM classification of mammalian genomic and cDNA sequences according to function

Takashi Abe, Yoko Kosaka and Toshimichi Ikemura

In addition to protein-coding sequences (CDSs), 5' and 3' untranslated regions (UTRs) and transcription regulatory regions of eukaryotic genes have attracted wide attention because of their crucial roles in transcriptional and post-transcriptional regulation of gene activity. We constructed SOMs of tri- and tetranucleotide frequencies in all 1-kb sequences derived from human or mouse genome. When sequences of 5' and 3' UTRs, CDSs, and introns, as well as 1-kb upstream regions from transcriptional start sites, were mapped on these SOMs, a major portion of the sequences were clustered primarily according to the functional categories. This showed that SOM could detect sequence characteristics specific to the distinct functional categories. Importantly, the territory of each functional category was divided into multiple zones. Furthermore, when we constructed tetra- and pentanucleotide SOMs for human and mouse cDNA sequences, protein-coding and -noncoding cDNA sequences tended to be separated from each other. Because no information other than oligonucleotide frequencies is required for the map generation, SOM is a novel *in silico* method useful for identifying characteristic and diagnostic sequences for individual functional categories. Function-unknown sequences colocalized in a zone where sequences of known functions are clustered can be assumed presumably to have similar functions.

(3) Compilation of "GenomeWordDictionary"

Takashi Abe, Yoko Kosaka, Kiyomi Kita and Toshimichi Ikemura

In order to know the biological meaning of characteristic oligonucleotide sequences specified by SOM, it is important to systematically refer to literatures focusing on experimental studies of the oligonucleotide sequences. With regard to genomes on which experimental studies have advanced, many signal and motif sequences with functional activity are known. Referring to the SOM data, the signal and motif sequences that are not experimentally identified may be presumed newly. Furthermore, based on the SOM data, exploration of signal and motif candidates on the genomes of which experimental studies have not advanced may be possible. It should also be noted that a large amount of exploration information itself which is accumulated in the process of the above literature exploration should become significant and valuable data sets. We have compiled a new database called "GenomeWordDictionary" in the form of collection of the exploration results of papers which describe experimental facts regarding each oligonucleotide sequence. In the dictionary, oligonucleotides being composed of 4 letters (A, T, G, and C) are arranged in the alphabetical order, and the abstract of the paper reporting the experimental results are compiled. Since we have created the database using Oracle and Postgres relational-database systems, we can extract a dictionary for a particular phylogenetic group; e.g. Rice GenomeWordDictionary and Fly GenomeWordDictionary.

PUBLICATIONS

Papers

1. Uchiyama, T., Abe, T., Ikemura, T. and Watanabe, K. (2005). Substrate-induced gene-expression screening of environmental metagenome libraries for isolation of catabolic genes, *Nature Biotechnology, 23,* 88-93.

2. Abe, T., Kanaya, S., Kinouchi, M. and Ikemura, T. (2004). Genome informatics for unveiling hidden genome signatures. *Proceedings of the Institute of Statistical Mathematics, 52,* 207-215.

3. Abe, T., Ikemura, T., Kanaya, S., Kinouchi, M. and

Sugawara, H. (2004). Novel genome informatics for unveiling hidden signatures in genome sequences: self-organizing map (SOM) of oligonucleotide frequencies, *Proceedings of Information-Based Induction Sciences,* 94-99.

4. Hayashi, H., Abe, T., Sakamoto, M., Ohara, H., Ikemura, T., Sakka, K. and Benno, Y. (2005). Direct cloning of genes encoding novel xylanases from human gut, *Canadian Journal of Microbiology* (in press).

Journal Editor

5. GENE (Elsevier)
6. DNA Sequence (Harwood Academic Publisher)

D-c. Division of Theoretical Genetics Sumio Sugano Group

RESEARCH ACTIVITIES

(1) Functional Genomics Based on the full-length cDNA collection

Sumio Sugano

We have been constructing full-length cDNA libraries from various human tissues and cells using oligo-capping method. From these libraries, we have been isolating a number of clones and determining their 5' end sequences. Using these 5' end sequence data, we have selected about 30,000 clones and determined entire sequences. Part of the result of this activity and initial analysis of the fully sequenced clones were published this year[11]. We also gathered an international consortium, H-invitational, for the annotation of human full-length cDNAs with Takashi Gojobori and Nomura Nobuo and made H-invitational database[6].

The 5' end sequence data of full-length cDNAs are indispensable for identifying transcription start sites (tss). We made and are maintaining database of tss, DBTSS, with Kenta Nakai[15]. Using this database, we compared human and mouse tss and devised and evaluated promoter prediction programs[13] [14] [2]. Furthermore, we collaborated with Shin-Ichi Hashimoto and developed a method to isolate a 5' tags of mRNAs[4].

We also collaborated with various groups to make full-length cDNA libraries and to collect full-length cDNA clones. This includes, Zebrafish[3], red alga[7], and malaria[16]. We also helped isolating various full-length cDNA clones[1] [5] [8] [9] [10] [17] [18].

Finally, we made very interesting finding that some of the small open reading frames (ORF) within 5' un-translated region (UTR) of mRNA do make small proteins. About half of the mRNAs have small ORF in their 5' UTR. Thus, there is a possibility that unexpectedly high number of small proteins is translated from 5' UTR.

PUBLICATIONS

Papers

1. Adachi, S., Jigami, T., Yasui, T., Nakano, T., Ohwada, S., Omori, Y., Sugano, S., Ohkawara, B., Shibuya, H., Nakamura, T. and Akiyama, T., (2004). Role of a BCL9-related beta-catenin-binding protein, B9L, in tumorigenesis induced by aberrant activation of Wnt signaling. Cancer Res. *64,* 8496-8501.

2. Bajic, VB., Tan, SL., Suzuki, Y. and Sugano, S. (2004). Promoter prediction analysis on the whole human genome. Nat Biotechnol. *22,* 1467-1473.

3. Gerhard, DS., Wagner, L., Feingold, EA., Shenmen, CM., Grouse, LH., Schuler, G., Klein, SL., Old, S., Rasooly, R., Good, P., Guyer, M., Peck, AM., Derge, JG., Lipman, D., Collins, FS., Jang, W., Sherry, S., Feolo, M., Misquitta, L., Lee, E., Rotmistrovsky, K., Greenhut, SF., Schaefer, CF., Buetow, K., Bonner, TI., Haussler, D., Kent, J., Kiekhaus, M., Furey, T., Brent, M., Prange, C., Schreiber, K., Shapiro, N., Bhat, NK., Hopkins, RF., Hsie, F., Driscoll, T., Soares, MB., Casavant, TL., Scheetz, TE., Brown-stein, MJ., Usdin, TB., Toshiyuki, S., Carninci, P., Piao, Y., Dudekula, DB., Ko, MS., Kawakami, K., Suzuki, Y., Sugano, S., Gruber, CE., Smith, MR., Simmons, B., Moore, T., Waterman, R., Johnson, SL., Ruan, Y., Wei, CL., Mathavan, S., Gunaratne, PH., Wu, J., Garcia, AM., Hulyk, SW., Fuh, E., Yuan, Y., Sneed, A., Kowis, C., Hodgson, A., Muzny, DM., McPherson, J., Gibbs, RA., Fahey, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madari, A., Young, AC., Wetherby, KD., Granite, SJ., Kwong, PN., Brinkley, CP., Pearson, RL., Bouffard, GG., Blakesly, RW., Green, ED., Dickson, MC., Rodriguez, AC., Grimwood, J., Schmutz, J., Myers, RM., Butterfield, YS., Griffith, M., Griffith, OL., Krzywinski, MI., Liao, N., Morrin, R., Palmquist, D., Petrescu, AS., Skalska, U., Smailus, DE., Stott, JM., Schnerch, A., Schein, JE., Jones, SJ., Holt, RA., Baross,

A., Marra, MA., Clifton, S., Makowski, KA., Bosak, S. and Malek, J. (2004). MGC Project Team. The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). Genome Res. *14,* 2121-2127.

4. Hashimoto, SI., Suzuki, Y., Kasai, Y., Morohoshi, K., Yamada, T., Sese, J., Morishita, S., Sugano, S. and Matsushima, K. (2004). 5'-end SAGE for the analysis of transcriptional start sites. Nat Biotechnol. *22,* 1146-1149, Epub Aug 8.

5. Huang, DY., Kuo, YY., Lai, JS., Suzuki, Y., Sugano, S. and Chang, ZF. (2004). GATA-1 and NF-Y cooperate to mediate erythroid-specific transcription of Gfi-1B gene. Nucleic Acids Res. *32,* 3935-3946.

6. Imanishi, T., Itoh, T., Suzuki, Y., O'Donovan, C., Fukuchi, S., Koyanagi, KO., Barrero, RA., Tamura, T., Yamaguchi-Kabata, Y., Tanino, M., Yura, K., Miyazaki, S., Ikeo, K., Homma, K., Kasprzyk, A., Nishikawa, T., Hirakawa, M., Thierry-Mieg, J., Thierry-Mieg, D., Ashurst, J., Jia, L., Nakao, M., Thomas, MA., Mulder, N., Karavidopoulou, Y., Jin, L., Kim, S., Yasuda, T., Lenhard, B., Eveno, E., Suzuki, Y., Yamasaki, C., Takeda, JI., Gough, C., Hilton, P., Fujii, Y., Sakai, H., Tanaka, S., Amid, C., Bellgard, M., Bonaldo, Md, M., Bono, H., Bromberg, SK., Brookes, AJ., Bruford, E., Carninci, P., Chelala, C., Couillault, C., Souza, SJ., Debily, MA., Devignes, MD., Dubchak, I., Endo, T., Estreicher, A., Eyras, E., Fukami-Kobayashi, K., R, Gopinath, G., Graudens, E., Hahn, Y., Han, M., Han, ZG., Hanada, K., Hanaoka, H., Harada, E., Hashimoto, K., Hinz, U., Hirai, M., Hishiki, T., Hopkinson, I., Imbeaud, S., Inoko, H., Kanapin, A., Kaneko, Y., Kasukawa, T., Kelso, J., Kersey, P., Kikuno, R., Kimura, K., Korn, B., Kuryshev, V., Makalowska, I., Makino, T., Mano, S., Mariage-Samson, R., Mashima, J., Matsuda, H., Mewes, HW., Minoshima, S., Nagai, K., Nagasaki, H., Nagata, N., Nigam, R., Ogasawara, O., Ohara, O., Ohtsubo, M., Okada, N., Okido, T., Oota, S., Ota, M., Ota, T., Otsuki, T., Piatier-Tonneau, D., Poustka, A., Ren, SX., Saitou, N., Sakai, K., Sakamoto, S., Sakate, R., Schupp, I., Servant, F., Sherry, S., Shiba, R., Shimizu, N., Shimoyama, M., Simpson, AJ., Soares, B., Steward, C., Suwa, M., Suzuki, M., Takahashi, A., Tamiya, G., Tanaka, H., Taylor, T., Terwilliger, JD., Unneberg, P., Veeramachaneni, V., Watanabe, S., Wilming, L., Yasuda, N., Yoo, HS., Stodolsky, M., Makalowski, W., Go, M., Nakai, K., Takagi, T., Kanehisa, M., Sakaki, Y., Quackenbush, J., Okazaki, Y., Hayashizaki, Y., Hide, W., Chakraborty, R., Nishikawa, K., Sugawara, H., Tateno, Y., Chen, Z., Oishi, M., Tonellato, P., Apweiler, R., Okubo, K., Wagner, L., Wiemann, S., Strausberg, RL., Isogai, T., Auffray, C., Nomura, N., Gojobori, T. and Sugano, S. (2004). Integrative Annotation of 21,037 Human Genes Validated by Full-Length cDNA Clones. PLoS Biol. Apr 20 [Epub ahead of print].

7. Matsuzaki, M., Misumi, O., Shin-I, T., Maruyama, S., Takahara, M., Miyagishima, SY., Mori, T., Nishida, K., Yagisawa, F., Nishida, K., Yoshida, Y., Nishimura, Y., Nakao, S., Kobayashi, T., Momoyama, Y., Higashiyama, T., Minoda, A., Sano, M., Nomoto, H., Oishi, K., Hayashi, H., Ohta, F., Nishizaka, S., Haga, S., Miura, S., Morishita, T., Kabeya, Y., Terasawa, K., Suzuki, Y., Ishii, Y., Asakawa, S., Takano, H., Ohta, N., Kuroiwa, H., Tanaka, K., Shimizu, N., Sugano, S., Sato, N., Nozaki, H., Ogasawara, N., Kohara, Y. and Kuroiwa, T. (2004). Genome sequence of the ultrasmall unicellular red alga Cyanidioschyzon merolae 10D. Nature. *428,* 653-657.

8. Niida, A., Hiroko T., Kasai, M., Furukawa, Y., Nakamura, Y., Suzuki, Y., Sugano, S., Akiyama, T. (2004). DKK1, a negative regulator of Wnt signaling, is a target of the beta-catenin/TCF pathway. Oncogene. *23,* 8520-8526.

9. Noma, T., Yoshizawa, I., Sugano, S., Ogawa, N. and Kawano, Y. (2004). Somatropin promotes dermatophagoides farinae-specific IgE generation independently of IL-4 and IL-10. Asian Pac J Allergy Immunol. *22,* 133-142.

10. Nonaka, H., Sugano, S. and Miyajima, A. (2004). Serial analysis of gene expression in sinusoidal endothelial cells from normal and injured mouse liver. Biochem Biophys Res Commun. *324,* 15-24.

11. Ota ,T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Hotuta, T., Kusano, J.,

Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, TO., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S. (2004). Complete sequencing and characterization of 21,243 full-length human cDNAs. Nat Genet. *36,* 40-45.

12. Oyama, M., Itagaki, C., Hata, H., Suzuki, Y., Izumi, T., Natsume, T., Isobe, T. and Sugano, S. (2004). Analysis of small human proteins reveals the translation of upstream open reading frames of mRNAs. Genome Res. *10B,* 2048-2052.

13. Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J., Mizushima-Sugano, J., Kel, AE., Arakawa, T., Carninci, P., Kawai, J., Hayashizaki, Y., Takagi, T., Nakai, K. and Sugano, S. (2004). Large-scale collection and characterization of promoters of human and mouse genes. In Silico Biol. *4* (3), 0036 [Epub ahead of print].

14. Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J., Mizushima-Sugano, J., Nakai, K. and Sugano, S. (2004). Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions. Genome Res. *14,* 1711-1718.

15. Suzuki, Y., Yamashita, R., Sugano, S. and Nakai, K. (2004). DBTSS, DataBase of Transcriptional Start Sites: progress report 2004. Nucleic Acids Res. *32,* D78-81.

16. Watanabe, J., Sasaki, M., Suzuki, Y. and Sugano, S. (2004). Full-malaria 2004: an enlarged database for comparative studies of full-length cDNAs of malaria parasites, Plasmodium species. Nucleic Acids Res. *32,* D334-338.

17. Yamada, S., Ohira, M., Horie, H., Ando, K., Takayasu, H., Suzuki, Y., Sugano, S., Hirata, T., Goto, T., Matsunaga, T., Hiyama, E., Hayashi, Y., Ando, H., Suita, S., Kaneko, M., Sasaki, F., Hashizume, K., Ohnuma, N. and Nakagawara, A. (2004). Expression profiling and differential screening between hepatoblastomas and the corresponding normal livers: identification of high expression of the PLK1 oncogene as a poor-prognostic indicator of hepatoblastomas. Oncogene. *23,* 5901-5911.

18. Yu, L., Frith, MC., Suzuki, Y., Peterfreund, RA., Gearan, T., Sugano, S., Schwarzschild, MA., Weng, Z., Fink, JS. and Chen, JF. (2004). Characterization of genomic organization of the adenosine A2A receptor gene by molecular and bioinformatics analyses. Brain Res. *1000,* 156-173.

Databases

19. http://dbtss.hgc.jp/

20. https://www.jbirc.aist.go.jp/hinv/index.jsp

SOCIAL CONTRIBUTIONS AND OTHERS

学会活動

1. HUGO: Council
2. HUGO Pacific: Co-Chair
3. 遺伝子治療学会：評議員

E. DEPARTMENT OF INTEGRATED GENETICS

E-a. Division of Human Genetics Hiroyuki Sasaki Group

RESEARCH ACTIVITIES

(1) Establishment and maintenance of genomic imprinting in the germline and in early embryos

Hiroyuki Sasaki, Masahiro Kaneda, Ryutaro Hirasawa, Kenichiro Hata, Maki Kusumi, Takashi Sado, Kenji Kumaki, Hiroyasu Furuumi, Masaki Okano[1], Naomi Tsujimoto[2], En Li[2], Tomohiro Suzuki[3], Shigeharu Wakana[3] and Toshihiko Shiroishi[3] ([1]CDB, RIKEN; [2]Harvard Med. Sch.; [3]GSC, RIKEN)

DNA methylation serves as an important gene-marking mechanism for discrimination of the parental alleles of imprinted genes. Although *de novo* DNA methyltransferases of the Dnmt3 family are implicated in maternal imprinting, the lethality of conventional *Dnmt3a* and *Dnmt3b* knockout mice precluded further studies. We have disrupted *Dnmt3a* and *Dnmt3b* in male and female germ cells, leaving them intact in somatic cells, by conditional gene knockout technology[2), 3)]. Offspring from the *Dnmt3a* conditional mutant females died *in utero* and lacked methylation and allele-specific expression at maternally imprinted loci. The *Dnmt3a* conditional mutant males showed impaired spermatogenesis and a lack of methylation at paternally imprinted loci in spermatogonia. Although these defects closely resembled those of *Dnmt3L* knockout mice, exact contribution of Dnmt3a and Dnmt3L to paternal imprinting varied from locus to locus. By contrast, the *Dnmt3b* conditional mutants and their offspring showed no phenotype. These results indicate that Dnmt3a is the critical enzyme responsible for both paternal and maternal imprinting[2), 3)]. We also study how the primary imprints are maintained in preimplantation mouse embryos, and set out to screen ENU-treated mutant mouse stocks for mutations that affect the establishment of germline imprints in collaboration with a group in GSC, RIKEN.

(2) Comparative analyses of the distal imprinted domain on mouse chromnosome 7 and the orthologous domain on chicken chromosome 5

Hiroyuki Sasaki, Takaaki Yokomine, Wahyu Purbowasito[1], Hisao Shirohzu, Chikako Suda, Takashi Sado, Hisakazu Iwama, Kazuho Ikeo, Tetsuya Hori, Masaoki Tsuzuki[2], Shigeki Mizuno[3], Yo-ichi Matsuda[4], Chiyoko Sato[5], Katsuzumi Okumura[5], Tsunehiro Mukai[6], Mohamad Zubair[7], Ken Tsutsui[8], Reiko Kato[9], Atsushi Toyoda[9], Masahira Hattori[9] and Yoshiyuki Sakaki[9] ([1]Kyushu Univ.; [2]Hiroshima Univ.; [3]Nihon Univ.; [4]Hokkaido Univ.; [5]Mie Univ.; [6]Saga Univ.; [7]NIBB; [8]Okayama Univ.; [9]GSC, RIKEN)

Genomic imprinting, an epigenetic gene-marking phenomenon in the germline, causes parent-of-origin-specific monoallelic expression of a subset of mammalian genes. Imprinted genes tend to form clusters in the genome (imprinted domains), which may be related to the mechanism of imprinting or to the evolution of imprinting. As a step to understand the structural and functional characteristics of the imprinted domains, we have cloned and sequenced a 1-Mb imprinted domain in mouse chromosome 7F4/F5 and its orthologous domain in chicken chromosome 5 (0.5 Mb). Using the mouse YAC clone that we have isolated, Cerrato et al. found that this domain is in fact composed of two autonomous subdomains[1)]. We then found that the genes of the chicken domain are not imprinted and, furthermore, that the chicken domain lacks the unique tandem repeat cluster of 0.2 Mb, the *H19* gene, and the imprinting control elements, all of which are present in the mouse. The results indicate that the mammalian imprinted genes were already clustered in the common ancestors of mammals and birds and that the imprinting mechanism, which can affect multiple genes in the cluster, came in later during mammalian evolution[8), 9)]. We also mapped a total of 52 nuclear matrix attachment regions (MARs) in the imprinted mouse domain. We found that the MARs are unevenly distributed in the domain and that there is a large MAR cluster in the boundary region of two imprinted subdomains[6)].

(3) Imprinting mechanisms of the mouse Igf2/H19 sub-domain

Hiroyuki Sasaki, Yuzuru Kato, Ko Ishihara[1], Melanie Ehrlich[2], Walter Reith[3] and Mitsuyoshi Nakao[1] ([1]Kumamoto Univ.; [2]Tulane Univ.; [3]Univ. Geneva)

The imprinted mouse 7F4/F5 domain contains two linked imprinted genes *Igf2* and *H19* near its centromeric boundary: *Igf2* is paternally expressed and *H19* maternally expressed. It is known that the paternal-specific methylation of the differentially methylated region (DMR) upstream of *H19* is the primary signal for the *Igf2/H19* imprinting. We found that a winged-helix type DNA-binding protein called RFX1 or MDBP binds to the conserved sequences within the DMR. Interestingly, this protein binds to the target sequence preferentially when they are methylated at CpG sites. We are currently looking at RFX knockout mice to see whether these proteins play a role in *Igf2/H19* imprinting. We also wrote a review on the interactions between the DMRs by chromatin looping[10].

(4) Computer-assisted search for sequence features common to imprinted DMRs

Hiroyuki Sasaki, Hisato Kobayashi, Takashi Abe and Toshimichi Ikemura[1] ([1]SOKENDAI)

Although the imprinted DMRs, which show differential methylation depending on parental origin, often play crucial roles in imprinting, features common to the DMRs have not been identified. We therefore set out to look for sequence features common to the DMRs by computer-assisted programs. We first located the DMR sequences on self-organizing maps (SOMs) produced from the mouse genome sequences for di-, tri- and tetra-nucleotides. We found that most DMRs are located in the periphery of the SOMs: they are more CpG-rich than most of the genome but less CpG-rich than the CpG islands. More detailed studies on the DMR sequences are ongoing.

(5) Molecular pathology of ICF syndrome

Hiroyuki Sasaki, Hiroyasu Furuumi, Tadashi Kajii, Tomohiro Kamoda[1], Nobuaki Iwasaki[1], Naomi Tobita[1], Nobuko Fujiwara[1], Akira matsui[1], Yu-ichi Goto[2] and Takeo Kubota[2] ([1]Univ. Tsukuba; [2]Yamanashi Univ.)

We previously studied two Japanese families with ICF (immunodeficiency, centromeric instability, facial anomalies) syndrome, an autosomal recessive disorder with hypomethylation of satellite DNA, and found that they have mutations in the *de novo* DNA methyltransferase gene *DNMT3B*. We have now studied a new Japanese ICF case and found that this patient does not have any mutation in *DNMT3B*. This suggests that the molecular pathology of ICF syndrome is heterogeneous and that a gene other than *DNMT3B* can affect the methylation of pericentromeric satellite DNA[4].

(6) Development of a universal DNA chip system applicable to any organism

Hiroyuki Sasaki, Shin-ichi Mizuno[1], Tadafumi Iino[1], Hidetoshi Ozawa[1], Kosuke Tashiro[1] and Takashi Gojohbori ([1]Kyushu Univ.)

We carried out a collaborative research project with groups in Kyushu University to develop a universal DNA chip system that can be used to study expression of any gene in any organism. We established the basic chip design and protocols for this innovative chip system (patent application no. 2004-278122). Development of the universal DNA chip system for practical use is underway.

(7) Antisense regulation at the *Xist* locus

Takashi Sado, Tatsuya Ohhata[1], Yuko Hoki[1]and Hiroyuki Sasaki ([1]PRESTO, JST)

Xist (X-inactive specific transcript), which does not encode a protein, is essential for X-inactivation to occur in *cis*. Expression of *Xist* is negatively regulated in *cis* by its antisense gene *Tsix*. Disruption of *Tsix*, therefore, induces upregulation of *Xist* in *cis*. We have been studying the molecular mechanism of how *Tsix* regulates *Xist* expression in a *cis*-limited manner by gene targeting technology. We examined chromatin at the *Xist* locus on the *Tsix* deficient X chromosome in embryos. We found that in the absence of *Tsix*, methylation levels of CpG sites were reduced and chromatin structure became open at the *Xist* locus. In

addition, histones of this region were modified to be characteristic of active chromatin. These results suggest that *Tsix* plays a role in establishment of repressive chromatin at the *Xist* and silence the *Xist* gene.

(8) X-inactivation in mouse embryos deficient for histone methyltransferase G9a

Tatsuya Ohhata, Makoto Tachibana[1], Hiroyuki Sasaki, Yoichi Shinkai[1] and Takashi Sado ([1]Kyoto Univ.)

Accumulating evidence suggests that methylation of histone H3 at lysine 9 (K9) and 27 (K27) is implicated in X-inactivation. Histone methyltranferase G9a is the enzyme that catalyzes methylation of K9 and perhaps K27 in euchromatic region. We studied X-inactivation in mouse embryos deficient for G9a, which die around the early somite stage. RNA-FISH revealed that *Xist* was appropriately regulated in both males and females. Taking advantage of X-linked GFP transgenes, effects of functional loss of G9a on the maintenance of X-inactivation were analyzed. We did not observe reactivation of the hitherto inactivated GFP transgenes in both the embryonic and extraembryonic tissues, suggesting that X-inactivation is stably maintained in G9a-null embryos. The same result was obtained using X-linked LacZ transgenes located more distal to the X-linked GFP transgenes, indicating that inactive state of these transgenes was stably maintained regardless the distance from the X-inactivation center, from which X-inactivation initiates and spreads. The results suggest that the X-inactivation process is properly regulated in the absence of G9a[5]. It is likely that methylation of histone H3 at K9 and K27 on the inactive X chromosome is mediated by an enzyme(s) other than G9a.

(9) Role of Dnmt3L in spermatogenesis and genomic imprinting during oogenesis

Kenichiro Hata, Maki Kusumi, En Li[1] and Hiroyuki Sasaki ([1]Harvard Med. Sch.)

Dnmt3L (DNA cytosine-5-methyltransferase 3-Like) encodes a protein of 421 amino acid residues and harbors a putative zing finger domain that shares a high degree of homology with the PHD-like domain of DNA methyltransferases Dnmt3a and Dnmt3b. The C-terminal part of Dnmt3L is related to DNA cytosine-5-methyltransferase, but it does not possess critical motifs for methyltransferase activity. We have generated *Dnmt3L*-deficient mice by gene targeting. While *Dnmt3L*$^{-/-}$ female mice grew normally, all embryos from pregnant *Dnmt3L* $^{-/-}$ mothers died around E10.5. The maternally methylated imprinted genes, e.g. *Igf2r* and *Peg1*, were hypomethylated in embryos derived from *Dnmt3L* $^{-/-}$ females x *Dnmt3L* $^{+/+}$ males, but paternally methylated imprinted genes were unaffected. Also, *Dnmt3L*$^{-/-}$ male mice showed severe defects in spermatogenesis, which is similar to, but severer than, the phenotype displayed by *Dnmt3a*$^{-/-}$ mice. We speculate that Dnmt3L functions via interactions with Dnmt3a and/or Dnmt3b to control DNA methylation in developing germ cells. Dnmt3L may be involved in not only the establishment of genomic imprinting but also DNA methylation of other regions.

PUBLICATIONS

Papers

1. Cerrato, F., Sparago, A., Zou, X., Dean, W., Bruggemann, M., Sasaki, H., Reik, W. and Riccio, A. (2005) The two domain hypothesis in Beckwith-Wiedemann syndrome: autonomous imprinting of the telomeric domain of the distal chromosome 7 cluster. **Hum. Mol. Genet.** (in press).

2. Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E. and Sasaki, H. (2004). Essential role for *de novo* DNA methyltransferase Dnmt3a in paternal and maternal imprinting. **Nature** *429*, 900-903.

3. Kaneda, M., Sado, T., Okano, M., Hata, K., Tsujimoto, N., Li, E. and Sasaki, H. (2004). Role of *de novo* DNA methyltransferases in initiation of genomic imprinting and X-chromosome inactivation. **Cold Spring Harbor Symp. Quant. Biol.** (in press).

4. Kubota, T., Furuumi, H., Kamoda, T., Iwasaki, N., Tobita, N., Fujiwara, N., Goto, Y., Matsui, A., Sasaki, H. and Kajii, T. (2004). .ICF syndrome in a girl with DNA hypomethylation but without detectable *DNMT3B* mutation. **Am. J. Med. Genet.** *129*A, 290-293.

5. Ohhata, T., Tachibana, M., Tada, M., Tada, T., Sasaki, H., Shinkai, Y. and Sado, T. (2004). X-inactivation is stably maintained in mouse embryos deficient for histone methyltransferase *G9a*. **Genesis** *40*, 151-156.

6. Purbowasito, W., Suda, C., Yokomine, T., Zubair, M., Sado, T., Tsutsui, K. and Sasaki, H. (2004). Large-scale identification and mapping of nuclear matrix-attachment regions in the distal imprinted domain of mouse chromosome 7. **DNA Res.** *11*, 391-407.

7. Sado, T., Okano, M., Li, E. and Sasaki, H. (2004). De novo methylation is dispensable for the initiation and propagation of X chromosome inactivation. **Development** *131*, 975-982.

8. Shirohzu, H., Yokomine, T., Sato, C., Kato, R., Toyoda, A., Purbowasito, W., Suda, C., Mukai, T., Hattori, M., Okumura, K., Sakaki, Y. and Sasaki, H. (2004). A 210-kb segment of tandem repeats and retroelements located between imprinted subdomains of mouse distal chromosome 7. **DNA Res.** *11*, 325-334.

9. Yokomine, T., Shirohzu, H., Purbowasito, W., Toyoda, A., Iwama, H., Ikeo, K., Hori, T., Mizuno, S., Tsudzuki, M., Matsuda, Y., Hattori, M., Sakaki, Y. and Sasaki, H. (2005). Structural and functional analysis of a 0.5-Mb chicken region orthologous to the imprinted mammalian *Ascl2/Mash2-Igf2-H19* region. **Genome Res.** (in press).

Reviews

10. Kato, Y. and Sasaki, H. (2005). Imprinting and looping: epigenetic marks control interactions between regulatory elements. **BioEssays** (in press).

11. 石原宏，佐々木裕之，中尾光善(2004)クロマチンインスレーターの構造と機能．わかる実験医学シリーズ注目のエピジェネティクスがわかる63-70.

12. 金田正弘，佐渡敬，佐々木裕之(2004)遺伝子に刷り込まれた記憶—細胞分化とDNAのメチル化・ゲノムインプリンティング．生物の科学　遺伝*58*, 69-74.

13. 金田正弘，佐々木裕之(2004)HOT PRESS：DNAメチル化酵素Dnmt3aが生殖系列でゲノム刷り込み(インプリンティング)を行う．細胞工学*23*, 934-935.

14. 熊木健治，佐々木裕之．幹細胞とゲノムインプリンティング．バイオインダストリー(印刷中).

15. 佐々木裕之(2004)エピジェネティクスと疾患. Molecular Medicine 41(増刊号)ヒトゲノム—ヒトの理解と疾病の克服332-338.

16. 佐々木裕之(2005). ゲノムインプリンティング，新しい潮流をもたらしたスーパーモデル(Overview). Molecular Medicine(印刷中).

17. 佐渡敬(2004) X染色体遺伝子量補償. ゲノム医学*4*, 51-57.

18. 佐渡敬，大畑樹也(2005). マウス胚体外組織におけるインプリント型X染色体不活性化. Molecular Medicine (印刷中).

19. 秦健一郎，佐々木裕之(2005)：生殖系列におけるインプリント成立機構．Molecular Medicine(印刷中).

20. 古海弘康，佐々木裕之(2004)シリーズ最新医学講座・転写因子12．エピジェネティクス制御とその異常．臨床検査*48*, 1673-1679.

Books

21. Sasaki, H. (2005). DNA methylation in epigenetics, development and imprinting. In: **Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics** (eds. Dunn, M., Jorde, L., Little, P. & Subramaniam, S.), John Wiley & Sons, Chichester (in press).

22. 佐々木裕之編(2004)Springer Reviews シリーズ—エピジェネティクス，シュプリンガー・フェアラーク東京.

23. 佐渡敬(2004)性染色体の遺伝子量補償．Springer Reviewsシリーズ—エピジェネティクス(佐々木裕之編)，pp.117-127，シュプリンガー・フェアラーク東京.

24. 秦健一郎(2004)生殖，発生の異常とエピジェネティクス. Springer Reviewsシリーズ—エピジェネティクス(佐々木裕之編)，pp.183-190，シュプリンガー・フェアラーク東京.

25. 平澤竜太郎，佐々木裕之．DNAメチル化．再生医療教科書シリーズ，第3巻「再生医療のための分子生物学(仲野徹・赤池敏宏監修)」(印刷中).

26. 古海弘康，佐々木裕之(2004)哺乳類のメンデル遺伝するエピジェネティクス. Springer Reviewsシリーズ—エピジェネティクス(佐々木裕之編)，pp.129-134，シュプリンガー・フェアラーク東京.

27. 佐渡敬(2004)（日本語訳)Hall, I.M., Grewal, S.I.S. Structure and Function of Heterochromatin: Imprication fro Epigenetic Gene Silencing and Genome Organization. In "RNAi: A guide to gene silencing" (Ed. Hannon, G.J.)(日本語版監修　中村義一)，pp.199-223，メディカル・サイエンス・インターナショナル.

EDUCATION

1. Prof. H. Sasaki was invited to give a seminar "Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting" at the Hospital for Sick Children, Toronto, Canada, June, 2004.

2. 佐々木裕之：山梨大学医学部講義，玉穂町，2004年6月.

3. 秦健一郎：NHK高校講座「生物」，2004年10月放送.

4. 佐々木裕之：北海道大学医学部講義，札幌，2004年10月.

5. 佐々木裕之：北里大学理学部講義，相模原，2004年10月.

SOCIAL CONTRIBUTIONS AND OTHERS

1. 特許出願番号：2004-278122，発明の名称：核酸マイクロアレイ及び核酸プローブの設計方法並びに遺伝子検出方法，発明者：佐々木裕之他４名，出願人：九州大学，情報・システム研究機構.
2. 佐々木裕之：(財)遺伝学普及会評議員.
3. 秦健一郎，佐々木裕之：韮山高校生インターンシップ２名受け入れ，2004年７月.
4. 佐渡敬，佐々木裕之：夏休み体験入学プログラム１名受け入れ，2004年８月.
5. 佐々木裕之：静岡県立静岡がんセンター研究所遺伝子組換え実験安全委員会委員.

E-b. Division of Agricultural Genetics Tetsuji Kakutani Group

RESEARCH ACTIVITIES

(1) Developmental abnormalities induced by DNA methylation mutation of Arabidopsis

Tetsuji Kakutani, Tetsu Kinoshita, Asuka Miura, Yuki Kinoshita, Masaomi Kato and Hodetoshi Saze

Arabidopsis provides a genetically tractable system to learn role of DNA methylation, since viable mutants with reduced genomic DNA methylation are available. Arabidopsis MET1 (METHYLTRANSFERASE 1, ortholog of mammalian DNA methyltransferase Dnmt1) is necessary for maintaining genomic cytosine methylation at 5'-CG-3' sites. Arabidopsis additionally methylates non-CG sites using CHROMOMETHYLASE3 (CMT3). A third Arabidopsis gene necessary for DNA methylation is *DDM1 (DECREASE IN DNA METHYLATION)*, which encodes a chromatin remodeling factor (Jeddeloh et al., 1999, Nat. Genet 22, 94-). The *ddm1* mutation affects both CG and non-CG methylation. A striking feature of *ddm1* mutation is that it induces a variety of developmental abnormalities by causing heritable change in other loci (Kakutani et al., 1996, PNAS 93, 12406-). The molecular basis has been clarified in two of the loci directly causing the developmental abnormalities (see below). In addition, we are starting genetic characterization of other types of developmental abnormalities, such as *bonsai* (Kakutani et al., 2005).

(2) Epigenetic behavior of *CACTA* transposon

Asuka Miura, Masaomi Kato, Kazuya Takashima, Yuki Kinoshita and Tetsuji Kakutani

Through genetic characterization of one of the *ddm1*-induced developmental abnormalities, we identified a novel endogenous Arabidopsis transposon, named *CACTA1*. This transposon transposes and increases in the copy number in DNA hypomethylation background in *ddm1* mutant (Miura et al., 2001, Nature 411, 212-). Loss of DNA methylation seems to be sufficient for mobilization of *CACTA1*, because it was mobilized in mutants of DNA methyltransferase genes, *MET1* and *CMT3*. High frequency transposition of *CACTA* elements was detected in *cmt3-met1* double mutants. Single mutants in either *met1* or *cmt3* were much less effective in mobilization. Thus CG and non-CG methylation systems redundantly function for immobilization of transposons (Kato et al., 2003 Curr Biol 13, 421-426). *CMT3* gene and non-CG methylation in plants may have evolved as an additional epigenetic tag dedicated to transposon control.

CACTA1 activated by the *ddm1* mutation remained mobile in the presence of the wild type *DDM1* gene, suggesting that de novo silencing is not efficient for the defense of the genome against *CACTA* movement. The defense depends on maintenance of transposon silencing over generations (Kato et al., 2004).

Differentiation of gene-rich and transposon-rich (gene-poor) regions is a common feature in plant genomes. That may be due to preferential integration of transposons to gene-poor regions or may be due to purifying selection against transposon insertion to gene-rich regions in natural populations. To evaluate the possible contribution of natural selection to the formation of transposon distribution pattern, we examined the distribution of the *CACTA* transposons in genomes of 19 natural variants (ecotypes) and compared that to integration induced in laboratory in the *ddm1* mutants. Sequences similar to mobile *CACTA1* copy distribute among the ecotypes and show high polymorphism in genomic localization. Despite the high polymorphism, the copy number was low in all the examined ecotypes and they localized preferentially in pericentromeric and transposon-rich regions (Miura et al, 2004). This contrasts to transposition induced in laboratory, in which the integration sites are less biased and the copy number

frequently increases. In addtion, transposition induced in DDM1 wild type background also showed unbiased integration sprcificity (Kato et al., 2004). The differences in the integration sites may be due to natural selection against deleterious insertion into chromosomal arm regions.

(3) Inheritance of epigenetic developmental abnormality

Yuki Kinoshita, Asuka Miura, Tetsu Kinoshita and Tetsuji Kakutani

Another developmental abnormality, late flowering trait, was induced by ectopic expression of *FWA* gene associated with hypomethylation of tandem repeat upstream of the coding region (Soppe et al., 2000, Mol Cell 6, 791-). Interesting thing is that change in nucleotide sequence was also not observed in *fwa-1* and *fwa-2* alleles isolated by conventional mutagenesis. In both cases, over-expression associated with the hypomethylation resulted in the phenotypes. Combining *ddm1* mutation and linkage analysis is useful for identifying epigenetically regulated genes affecting plant development (Kakutani, 1997, Plant J., 12, 1447-). The *ddm1*-induced late flowering trait as well as the hypomethylation and ectopic expression of the *FWA* gene were stably inherited over generations even in the presence of the wild type *DDM1* copy (manuscript in preparation).

(4) Epigenetic control of *FWA* gene expression in endosperm

Tetsu Kinoshita, Asuka Miura, Yuki Kinoshita and Tetsuji Kakutani

Although *FWA* is ectopically expressed in the epigenetic alleles stated above, the role of FWA gene product in normal development remained unknown. To understand why the *FWA* gene is epigenetically controlled, we further examined expression of this gene during normal development in wild type. Results of GFP reporter system and direct detection of the transcript both suggest that *FWA* is expressed specifically in the endosperm. Endosperm is a plant tissue analogous to mammalian placenta; it serves as nutritional support to the embryo. Furthermore, the *FWA* gene was expressed in parent-of-origin-specific manner; only maternal gene is expressed. The *FWA* imprint depends on the maintenance DNA methyltransferase MET1, as is the case in mammals. Unlike mammals, however, the *FWA* imprint is not established by allele-specific *de novo* methylation. It is established by maternal gametophyte-specific gene activation, which depends on a DNA glycosylase gene, *DEMETER*. Since endosperm does not contribute to the next generation, the activated *FWA* gene need not be silenced again. Double fertilization enables plants to use such 'one-way' control of imprinting and DNA methylation in endosperm (Kinoshita et al, 2004).

PUBLICATIONS

Papers

1. Kinoshita, T., Miura, A., Choi, Y., Kinoshita, Y., Cao, X., Jacobsen, SE., Fischer, RL. and Kakutani, T. (2004). One-way control of *FWA* imprinting in *Arabidopsis* endosperm by DNA methylation. Science *303*, 521-523.
2. Miura, A., Kato, M., Watanabe, K., Kawabe, A., Kotani, H. and Kakutani, T. (2004). Genomic localization of endogenous mobile CACTA family transposons in natural variants of *Arabidopsis thaliana*. MGG *270*, 524-532.
3. Kato, M., Takashima, K. and Kakutani, T. (2004). Epigenetic control of CACTA transposon mobility in Arabidopsis thaliana. Genetics *168*, 961-969.
4. Kakutani, T., Kato, M., Kinoshita, T. and Miura, A. (2005). Control of Development and Transposon Movement by DNA Methylation in Arabidopsis thaliana. Cold Spring Harbor Symposia on Quantitative Biology. *69* in press.
5. Bachmair, A., Garber, K., Takeda, S., Sugimoto, K., Kakutani, T. and Hirochika, H. (2004). Biochemical analysis of long terminal repeat retrotransposons. Methods in Molecular Biology *260*, 73-82.

EDUCATION

1. 角谷徹仁　東京大学医科学研究所で講義, 2004年5月
2. 角谷徹仁　京都大学農学部で集中講義, 2004年7月
3. 角谷徹仁　東京大学農学部で集中講義, 2004年11月

E-b. Division of Agricultural Genetics Keiichi Shibahara Group

RESEARCH ACTIVITIES

(1) Mechanism of nucleosome assembly during DNA replication

Yasunari Takami[1], Tatsuya Ono, Tatsuo Nakayama[1] and Kei-ichi Shibahara ([1]Department of Biochemistry, Miyazaki Medical College, University of Miyazaki)

A newly replicated DNA is assembled into nucleosome soon after the passage of replication fork. This rapid formation of nucleosome is functionally linked to DNA replication machineries and plays a critical role for the maintenance of genome integrity and epigenetic states of chromatin in proliferating cells. Two histone binding proteins, CAF-1 (Chromatin Assembly Factor-1) and ASF1 (Anti-Silencing Function1), are involved in some process of this nucleosome assembly reaction.

In collaboration with Takami and Nakayama in University of Miyazaki, using CAF-1p150-or p60-deficint chicken DT40 cells, we recently showed that without CAF-1 function, S-phase progression was delayed and a rapid nucleosome assembly during DNA replication was disturbed *in vivo*. This is the first direct evidence for the involvement of CAF-1 in a rapid nucleosome assembly during DNA replication *in vivo*. In addition, CAF-1 was suggested in involving in Chk1-dependent checkpoint pathway after the treatment with HU.

(2) Physiological implications of CAF-1 and CAF-1-dependent nucleosome assembly in higher eukaryotes

Tatsuya Ono, Hidetaka Kaya[1], Shin Takeda[2], Jerzty Paszkowski[2], Takashi Araki[1] and Kei-ichi Shibahara ([1]Department of Botany, Graduate School of Science, Kyoto University, [2]Department of Plant Biology, University of Geneva)

The genetic approach of *Arabidopsis* is a powerful tool to see physiological implications of CAF-1 and CAF-1dependent nucleosome assembly in higher eukaryotes. We analyzed loss-of-function mutants of *caf-1 (fasciata)* in *Arabidopsis* and showed that those *caf-1* mutants displayed severely disturbed cellular and functional organization of both meristems (*Cell 104*, 131-142. 2001). We recently showed that transcriptional gene silencing (TGS) of endogenous *CACTA* transposons was released infrequently in a stochastic manner in *fas*, without decreasing DNA methylation. Other endogenous silent genes at different chromosomal sites were also transcriptionally activated non-concomitantly with each other. Furthermore, TGS of the silent transgene *β-glucuronidase (GUS)* was also de-repressed randomly in *fas* mutant plants, irrespectively of developmental abnormalities, and the activated state of *GUS* was maintained during growth to produce clusters of cells expressing *GUS*. We suggest that CAF-1 contributes to the stable inheritance of epigenetic states through multiple rounds of cell divisions and that defects in CAF-1 functions explain the stochastic occurrence of pleiotropic phenotypes in *fas* mutants.

In collaboration with J. Paszkowski in University of Geneva, we analyzed *brul* mutants and found striking phenotypic similarities to *fas* mutants. Epistatic analysis of *brul* and *fas* mutants suggests that BRU1 and CAF-1 have overlapping functions. BRU1 encodes a novel nuclear protein containing two protein-protein interaction domains. However, the molecular function of BRU1 and the molecular link between BRU1 and CAF-1 is still unknown. The future analysis of BRU1 would provide yet another clue to understanding the mechanisms of epigenetic maintenance.

(3) Role of histon macroH2A in gene repression

Yuya Ogawa, Tatsuya Ono and Kei-ichi Shibahara

The unusual histone variant macroH2A (mH2A) is associated with transcriptional repression as it appears to be enriched in the inactive X chromosome by immunostaining. *In vitro* analysis showed that non-histone region of mH2A in carboxy-terminal prevents transcription factor from interacting with chromatin reconstituted with mH2A, but, the molecular function of mH2A in modulating chromatin structures remains unknown.

To reveal the roles of mH2A, we purified mH2A-containing nucleosome by affinity purification with anti-epitope tag antibodies, and eventually, we found mono-ubiquitinated form of mH2A.

Mono-ubiquitination of H2A and H2B has been shown to be important in regulating chromatin structure and transcription. We are currently trying to identify the ubiquitinated site of mH2A and reveal its implications in modulation of chromatin structures and transcriptional regulations.

(4) Differential methylation of *Xite* and CTCF sites in *Tsix* in mouse X-inactivation

Rebecca Maxfield Boumil[1,2,3], Yuya Ogawa, Bryan K. Sun[1,2,3] and Jeannie T. Lee[1,2,3] ([1]Massachusetts General Hospital, [2]Harvard Medical School and [3]Howard Hughes Medical Institute)

In X-chromosome inactivation (XCI) in mammalian female cells, one of two X-chromosomes is inactivated in the imprinted or stochastic manner. Although some loci influencing the choice decision have been genetically identified, the primary marks for imprinting and random choice remain undefined. We examined the role of DNA methylation, a mechanism known to regulate imprinting in autosomal loci, in XCI. We found differentially methylated domains (DMDs) in *Tsix* and *Xite*, two genes previously shown to influence choice. The DMDs occured in gametes and disappeared in ES cells with two active X chromosomes. Because the DNA methylation pattern reflects events of XCI, we propose that DMDs in *Tsix* and *Xite* are candidates for the primary marks of epigenetic regulation in XCI.

PUBLICATIONS

Papers

1. Takeda, S., Tadele, Z., Hofmann, I., Probst, A.V., Angelis, K.J., Kaya, H., Araki, T., Mengiste, T., Mittelsten Scheid, O., **Shibahara, K.,** Scheel, D. and Paszkowski, J. (2004). BRU1, a novel link between responses to DNA damage and epigenetic gene silencing in Arabidopsis. ***Genes Dev** 18*, 782-93.

2. Chromatin Assembly Factor-1 Dependent Nucleosome Assembly Coupled with DNA Replication, is Essential for Cell Proliferation in Vertebrate Cells. Takami, Y., Ono, T., Fukagawa, T., **Shibahara, K.** and Nakayama. T. Under revision in ***Mol Cell Biol.***

3. Mutations in CAF1 of *Arabidopsis* release transcriptional gene silencing in a stochastic manner. Ono, T., Kaya, H., Takeda, S., Abe, M., Ogawa, Y., Kato, M., Kakutani, T., Scheid, O.M., Araki, T. and **Shibahara, K.** Submitted

Reviews

4. 久保知大, 柴原慶一(2004)DNA複製及びヌクレオチド除去修復に伴うヌクレオソーム形成, わかる実験医学「DNA複製とDNA修復」.

特許

特願2005-65442, 遺伝子導入方法, ジーンターゲッティング方法及びトランスジェニック植物の製造方法, 柴原慶一

E-c. Division of Brain Function Tatsumi Hirata Group

RESEARCH ACTIVITIES

(1) Theoretical Consideration of Olfactory Axon Targeting with an Activity-Dependent Neural Network Model

Hirokazu Tozaki and Tatsumi Hirata

Olfactory sensory neurons that express a given odorant receptor target their axons onto a few specific glomeruli in the olfactory bulb. Although the odorant receptor plays an indispensable role in olfactory axon targeting, the mechanisms underlying this guidance are largely unknown. In particular, there is much controversy regarding the involvement of activity-dependent mechanism in the targeting process. In this study, we developed an activity-dependent self-organization model of the glomerular layer in the olfactory bulb and simulated the targeting of olfactory axons onto the layer[3]. Our model successfully constructed discrete glomeruli that received olfactory axons expressing a common odorant receptor through odorant-evoked neural activities. Furthermore, our model explained the perplexing experimental results that have been reported in olfactory axon targeting. For example, the segregated projection of the axons that express the same odorant receptor transcribed from the different alleles was successfully simulated if the genetically modified allele was assumed to express a smaller amount of the receptor protein. The activity-dependent model even explained the inconsistent effects of disruption of the activity-evoking ion channel on axons expressing different odorant receptors, although some of these results were regarded as

evidence for activity-independency of the olfactory targeting. Taken together, the activity-dependent targeting of olfactory axons seems to be a simple probable mechanism that can provide a unified explanation of glomerular formation.

(2) Systematic Screening and Identification of the Antigens Recognized by Monoclonal Antibodies Raised against the Developing Lateral Olfactory Tract

Takahiko Kawasaki, Yoshiko Takagi, Hitoshi Yamatani and Tatsumi Hirata

During development, olfactory bulb axons navigate the complex microenvironment composed of myriads of molecules in the telencephalon to construct a bundle called the lateral olfactory tract. The axons, themselves, also express thousands of different molecules during the navigation. In this study, we produced and characterized six monoclonal antibodies that label the lateral olfactory tract and the surroundings in a unique pattern[2]. The labeling profiles suggested that the antigen molecules recognized by each antibody are heterogeneously distributed around the developing lateral olfactory tract. We developed an efficient screening method to identify the antigen molecules by combining expression of a cDNA library in COS-7 cells and the subsequent immunohistochemical staining of the cells. The systematic screening successfully identified the specific cDNA clones for all of the monoclonal antibodies, which highly probably code for the antigen molecules, and therefore unveiled the molecular nature of local components that embrace the developing lateral olfactory tract.

(3) Chronotopic Organization of Olfactory Bulb Axons in the Lateral Olfactory Tract

Hitoshi Yamatani, Yasufumi Sato and Tatsumi Hirata

The arrangement of axons in a tract can have a specific consequence on the organization of functional neuronal circuits. We found that olfactory bulb axons are chronologically arranged in the lateral olfactory tract[4]. Newly-differentiated projection neurons over the whole olfactory bulb are similarly marked with transient expression of c-kit protein. Their axons are assembled together and project into the ventral superficial part of the tract, displacing the older axons. This special assembly of the axons explains the non-topographical relationships between the olfactory bulb and the lateral olfactory tract axons that have been described in previous studies, and could possibly influence the subsequent selection of the olfactory target areas by these axons.

PUBLICATIONS

Papers

1. Aizawa, H., Sato, Y., Maekawa, M., Fujisawa, H., Hirata, T. and Yuasa, S. (2004). Development of the amygdalohypothalamic projection in the mouse embryonic forebrain. Anat. Embryol. *208*, 249-264.
2. Kawasaki, T., Takagi, Y., Yamatani, H. and Hirata, T. (2004). Systematic screening and identification of the antigens recognized by monoclonal antibodies raised against the developing lateral olfactory tract. J. Neurobiol. *62*, 330-340.
3. Tozaki, H., Tanaka, S. and Hirata, T. (2004). Theoretical consideration of olfactory axon projection with an activity-dependent neural network model. Mol. Cell. Neurosci. *26*, 503-517.
4. Yamatani, H., Sato, Y., Fujisawa, H. and Hirata, T. (2004). Chronotopic organization of olfactory bulb axons in the lateral olfactory tract. J. Comp. Neurol. *475*, 247-260.

E-d. Division of Applied Genetics Takashi Araki Group

RESEARCH ACTIVITIES

(1) Physiological roles of chromatin assembly factor-1 in epigenetic control of meristem activity during post-embryonic development in plants

Takashi Araki[1, 2], Mitsutomo Abe[1, 3], Hidetaka Kaya[4] and Kei-ichi Shibahara ([1]Graduate School of Science, Kyoto University, [2]Japan Science and Technology Agency, [3]PROBRAIN, [4]Genome & Drug Research Center, Tokyo University of Science)

Loss-of-function mutants of chromatin assembly factor-1 (CAF-1) in *Arabidopsis, fasciata (fas)*, show a wide variety of morphological abnormalities and unique defects in the expression of *WUSCHEL* and *SCARECROW* genes in meristems (Cell 104, 131-142.

2001). We found that transcriptional gene silencing (TGS) of endogenous *CACTA* transposons was released in a stochastic manner in *fas*, without decreasing global levels of DNA methylation. Other endogenous silent genes at different chromosomal sites, a DNA transposon *AtMu1* and a hypothetical gene *T5L23.26* at a heterochromatin knob, were also transcriptionally activated. Activation of the three different silent loci occurred non-concomitantly to each other. Furthermore, TGS of a silent β-glucuronidase (GUS) transgene was also de-repressed randomly in *fas* plants, without apparent correlation with developmental abnormalities. Activated state of *GUS* was maintained during growth to produce clusters of cells expressing GUS. Based on these observations, we propose a model that CAF-1 contributes to the stable inheritance of epigenetic states through multiple rounds of cell divisions. Defect of CAF-1 function explains the stochastic occurrence of pleiotropic morphological phenotypes observed in *fas* mutants (Ogawa, Kaya, and others, submitted). We also investigated other genes involved in epigenetic control of meristem activity, such as *BRUSHY1 (BRU1)* and *ANTI-SILENCING FUNCTION1 (ASF1)* (in preparation).

(2) Genes involved in the integration of various genetic pathways for regulation of the floral transition

Takashi Araki[1,2], Mitsutomo Abe[1,3], Sumiko Yamamoto[1,2], Yasufumi Daimon[1], Ayako Yamaguchi[1], Yoko Ikeda[1], Michitaka Notaguchi[1] and Masaki Kobayashi[1] ([1]Graduate School of Science, Kyoto University, [2]Japan Science and Technology Agency, [3]PROBRAIN)

Floral transition in *Arabidopsis* is regulated by several pathways which converge on the transcriptional regulation of floral pathway integrators including *FT*. *FT* is a direct target of CO and encodes a protein with similarity to mammalian proteins (PEBP/RKIP) involved in cellular signaling. *FT* transcription is immediately induced in cotyledon and leaf vascular tissues upon transfer from short-day to inductive long-day photoperiods. Promotion of flowering by *FT* requires the activity of another flowering-time gene *FD* which encodes a bZIP transcription factor preferentially expressed in the shoot apex. FD is involved in transcriptional activation of the floral meristem identity genes *AP1* and *CAL* redundantly with LFY. *ft; lfy* and *fd; lfy* double mutants are very similar in severe reduction of *AP1* mRNA levels and strong defects in floral specification. Loss of *FT* function suppresses ectopic up-regulation of *AP1* in seedlings by *FD* over-expression. Mutant forms of FD which lack a C-terminal potential phosphorylation site cannot interact with FT in yeast cells and fail to complement *fd* late-flowering phenotype even by over-expression. These and other evidences suggest that *FT* and *FD* are inter-dependent in promotion of floral transition and activation of *AP1* expression (in preparation). Since the activity of FD, which is preferentially expressed in shoot apex, seems to require protein/protein interaction with FT, shoot apex is likely the site of action of FT protein. Consistent with this, restoration of *FT* function in whole region or outermost layer (L1) of the shoot apex can rescue *ft* late-flowering. These raise an interesting possibility that the FT protein may represent a long-distance signal generated in photoperiodically-induced leaves (mainly in vascular tissues) and act at the shoot apex to initiate floral development. *Arabidopsis* genome has a homolog of *FT*, called *TWIN SISTER OF FT (TSF)*. We found that *TSF* acts redundantly with *FT* as a floral pathway integrator (Yamaguchi et al., submitted).

PUBLICATIONS

Papers

1. Takeda, S., Tadele, Z., Hofmann, I., Probst, A.V., Angelis, K.J., Kaya, H., Araki, T., Mengiste, T., Scheid, O.M., Shibahara, K., Scheel, D. and Paszkowski, J. (2004). *BRU1*, a novel link between responses to DNA damage and epigenetic gene silencing in *Arabidopsis*. Genes Dev *18*, 782-93.

Reviews

2. 荒木崇(2004)“花成を制御する遺伝因子と環境因子の相互作用”,『植物の環境応答と形態形成のクロストーク』, 岩渕雅樹・篠崎一雄(編), シュプリンガー・フェアラーク東京, pp.57-63.

Books

3. 荒木崇(2004)“第24章　花成の調節”,『Taiz & Zeiger 植物生理学　第三版』, 島崎研一郎・西谷和彦(監訳), 培風館, pp.568-601.

EDUCATION

1. 荒木崇　京都大学大学院理学研究科・理学部(兼担)助教授
2. 荒木崇　近畿大学農学部集中講義，2004年4－5月
3. 荒木崇　筑波大学遺伝子実験センターセミナー，2004年11月

SOCIAL CONTRIBUTIONS AND OTHERS

1. 荒木崇　2004年(第11回)日本植物生理学会論文賞(共同受賞)
2. 荒木崇　Plant and Cell Physiology (editorial board)
3. 荒木崇　日本植物生理学会(評議員)
4. 荒木崇　岡山県生物科学総合研究所所外研究協力員
5. 荒木崇　特定非営利活動法人『シニア自然大学』の地球環境大学講座(代表：河野昭一京都大学名誉教授)講師，2004年4－10月

E-d. Division of Applied Genetics
Kunio Shiota Group

RESEARCH ACTIVITIES

(1) The Sall3 locus is an epigenetic hotspot of aberrant DNA methylation associated with placentomegaly of cloned mice

Jun Ohagane[1], Teruhiko Wakayama[2, 3], Sho Senda[1], Yukiko Yamazaki[3], Kimiko Inoue[2], Atsuo Ogura[2], Joel Marh[3], Satoshi Tanaka[1], Ryuzo Yanagimashi[3] and Kunio Shiota[1] ([1]Univ. Tokyo; [2]RIKEN; [3]Hawaii Univ.)

Cloned offspring develop a variety of abnormal phenotypes such as increased body weight (large fetus syndrome), pulmonary hypertension, placental overgrowth, respiratory problems and early death. We had identified several aberrantly methylated loci in the tissues of full-term cloned fetuses. Interestingly, each cloned animal has a different DNA methylation pattern and the extent of hyper- or hypo-methylation varies among the individuals (Ohgane *et al.*, Genesis 30, 45-50, 2001). In contrast, overgrowth of the placenta is one of the commonly observed symptoms in all cloned mice regardless of the sex and strain of animal and the type of donor cell. Thus, there may be genomic loci associated with the abnormal placental development in cloned mice and the genomic loci frequently associated with the epigenetic error have been explored in the cloned animals. We identified the Spalt-like gene 3 (*Sall3*) locus as a hypermethylated region in the placental genome of cloned mice. The *Sall3* locus has a CpG island containing a T-DMR. The T-DMR sequence is conserved in the human genome at the SALL3 locus of chromosome 18q23, which has been suggested to be involved in the 18q deletion syndrome. Intriguingly, larger placentas were more heavily methylated at the *Sall3* locus in cloned mice. This epigenetic error was found in all cloned mice examined regardless of sex, mouse strain and the type of donor cells. In contrast, the placentas of in vitro fertilized (IVF) and intracytoplasmic sperm injected (ICSI) mice did not show such hypermethylation, suggesting that aberrant hypermethylation at the *Sall3* locus is associated with abnormal placental development caused by nuclear transfer of somatic cells. Thus the *Sall3* locus is the area with frequent epigenetic errors in cloned mice. These data suggest that there exists at least genetic locus that is highly susceptible to epigenetic error caused by nuclear transfer.

(2) Epigenetic control of mouse *Oct-4* gene expression in embryonic stem cells and trophoblast stem cells

Naoko Hattori[1], Koichiro Nishino[1, 2], Yeoung-Gyu Ko[1], Naka Hattori[1, 2], Jun Ohgane[1], Satoshi Tanaka[1] and Kunio Shiota[1] ([1]Univ. Tokyo; [2]Bio-oriented Technology Research Advancement Institution)

The first cell differentiation event in mammalian embryogenesis segregates inner cell mass lineage from the trophectoderm at the blastocyst stage. *Oct-4*, a member of the POU family of transcription factors, is necessary for the pluripotency of the inner cell mass lineage. Embryonic stem (ES) cells, which contribute to all of embryonic lineages, express the *Oct-4* gene. Trophoblast stem (TS) cells, which have the ability to differentiate into trophoblast lineage in vitro, never contribute to embryonic proper tissues in chimeras and differentiate only into trophoblastic cells in the placenta. Expression of the *Oct-4* gene was undetectable and severely repressed in trophoblastic lineage, including the stem cells. We found that the culture of TS cells with 5-aza-2'-deoxycytidine or trichostatin A caused the activation of the *Oct-4* gene. Analysis of the DNA methylation status of mouse *Oct-4*

gene upstream region revealed that *Oct-4* enhancer/promoter region was hypomethylated in ES cells but hypermethylated in TS cells. Furthermore, in vitro methylation suppressed *Oct-4* enhancer/promoter activity in reporter assay. In the placenta of *Dnmt1*(n/n) mutant mice, most of the CpGs in the enhancer/promoter region were unmethylated, and *Oct-4* gene expression was aberrantly detected. Chromatin immunoprecipitation assay revealed that *Oct-4* enhancer/promoter region was hyperacetylated in ES cells compared with TS cells, thus demonstrating that DNA methylation status is closely linked to the chromatin structure of the *Oct-4* gene. Here we propose that the epigenetic mechanism, consisting of DNA methylation and chromatin remodeling, underlies the developmental stage- and cell type-specific mechanism of *Oct-4* gene expression.

(3) DNA methylation-mediated control of *Sry* gene expression in mouse gonadal development

Koichiro Nishino[1,2], Naoko Hattori[1], Satoshi Tanaka[1] and Kunio Shiota[1] ([1]Univ. Tokyo; [2]Bio-oriented Technology Research Advancement Institution)

DNA methylation at CpG sequences is involved in tissue-specific and developmentally regulated gene expression. The *Sry* (sex determining region on the Y chromosome) gene encodes a master protein for initiating testis differentiation in mammals, and its expression is restricted to gonadal somatic cells at 10.5-12.5 days post coitum (dpc) in the mouse. We found that *in vitro* methylation of the 5'-flanking region of the *Sry* gene caused suppression of reporter activity, implying that *Sry* gene expression could be regulated by DNA methylation-mediated gene silencing. Bisulfite restriction mapping and sodium bisulfite sequencing revealed that the 5'-flanking region of the *Sry* gene was hypermethylated in the 8.5 dpc embryos in which the *Sry* gene was not expressed. Importantly, this region was specifically hypomethylated in the gonad at 11.5 dpc, while the hypermethylated status was maintained in tissues that do not express the *Sry* gene. We concluded that expression of the *Sry* gene is under the control of an epigenetic mechanism mediated by DNA methylation.

(4) Skewed X-inactivation in cloned mice

Sho Senda[1], Teruhiko Wakayama[2], Yukiko Yamazaki[2], Jun Ohgane[1], Naka Hattori[1], Satoshi Tanaka[1], Ryuzo Yanagimachi[2] and Kunio Shiota[1] ([1]Univ. Tokyo; [2]Hawaii Univ.)

In female mammals, dosage compensation for X-linked genes is accomplished by inactivation of one of two X chromosomes. The X-inactivation ratio (a percentage of the cells with inactivated maternal X chromosomes in the whole cells) is skewed as a consequence of various genetic mutations, and has been observed in a number of X-linked disorders. We previously reported that phenotypically normal full-term cloned mouse fetuses had loci with inappropriate DNA methylation. Thus, cloned mice are excellent models to study abnormal epigenetic events in mammalian development. In the present study, we analyzed X-inactivation ratios in adult female cloned mice (B6C3F1). Kidneys of eight naturally produced controls and 11 cloned mice were analyzed. Although variations in X-inactivation ratio among the mice were observed in both groups, the distributions were significantly different (Ansary-Bradley test, $P<0.01$). In particular, 2 of 11 cloned mice showed skewed X-inactivation ratios (19.2% and 86.8%). Similarly, in intestine, 1 of 10 cloned mice had a skewed ratio (75.7%). Skewed X-inactivation was observed to various degrees in different tissues of different individuals, suggesting that skewed X-inactivation in cloned mice is the result of secondary cell selection in combination with stochastic distortion of primary choice. The present study is the first demonstration that skewed X-inactivation occurs in cloned animals. This finding is important for understanding both nuclear transfer technology and etiology of X-linked disorders.

(5) Non-coding RNA directed DNA demethylation of Sphk1 CpG island

Takuya Imamura[1], Soshi Yamamoto[1], Jun Ohgane[1], Naka Hattori[1], Satoshi Tanaka[1] and Kunio Shiota[1] ([1]Univ. Tokyo)

The formation of DNA methylation patterns is one of the epigenetic events that underlie mammalian development. The Sphk1 CpG island is a target for tissue-dependent DNA methylation as well as a

template for generating multiple subtypes. The number of mammalian non-coding RNA genes is rapidly expanding. In this study, we found endogenous antisense transcripts, Khps1 subtypes with different sizes (600-20,000nt). A subtype, Khps1a, was a 1290-bp, non-coding, 5'-capped and 3'-polyadenylated RNA that originated from the CpG island and overlapped with a tissue-dependent differentially methylated region (T-DMR) of Sphk1. Intriguingly, overexpression of two fragments of Khps1 caused demethylation of CG sites in the T-DMR. Furthermore, this RNA-directed demethylation was associated with DNA methylation at three CC(A/T)GG sites in the T-DMR. The link between the RNA-directed CG demethylation and non-CG methylation provides a novel mechanism of epigenetic regulation and potential tool for epigenetic manipulation of mammalian cells.

(6) Preference of DNA methyltransferases for CpG islands in normal cells

Naka Hattori[1], Tetsuya Abe[1], Masako Suzuki[1], Tomooki Matsuyama[2], Shigeo Yoshida[2], En Li[3] and Kunio Shiota[1] ([1]Univ. Tokyo; [2]RIKEN; [3]Massachusets General Hospital)

In vitro studies indicated that DNA methyltransferase1 (*Dnmt1*) prefers hemimethylated DNA compared to unmethylated DNA, while *Dnmt3a*/3b methylate CpG dinucleotides without preference for hemimethylated or unmethylated DNA. Collectively, the fact that *Dnmt1* is localized to DNA replication foci and associated with the methyl-CpG binding protein, MeCP2, which directs DNA methyltransferase activity to hemimethylated DNA[2], implies that *Dnmt1* is involved in maintenance methylation *in vivo* to preserve methylation patterns in genomic DNA, and that *Dnmt3a*/3b function as *de novo* methyltransferases. However, a direct association between DNMTs was reported, and it was demonstrated that *DNMT1* and *DNMT3B* function cooperatively for maintenance methylation in the human cancer cell line. Therefore, categorizing Dnmts into maintenance and *de novo* DNA methylation may not be appropriate when attempting to characterize *in vivo* mechanisms involved in determining or establishing DNA methylation profiles in the genome. To address the question how the T-DMR of CpG islands are regulated by Dnmts *in vivo*, we analyzed the genome-wide DNA methylation pattern focussing on CpG islands and found that each Dnmt has target preferences depending the genomic component (in preparation).

(7) Stage-by-stage change in DNA methylation status of DNA methyltransferase 1 *(Dnmt1)* locus during mouse early development

Yeoung-Gyu Ko[1], Koichiro Nishino[1], Naoko Hattori[1], Yoshikazu Arai[1], Satoshi Tanaka[1] and Kunio Shiota[1] ([1]Univ. Tokyo)

Methylation of DNA is involved in tissue-specific gene control, and establishment of DNA methylation pattern in the genome is thought to be essential for embryonic development. Three isoforms of DNA methyltransferase 1 *(Dnmt1)* transcripts, *Dnmt1s*, *Dnmt1o* and *Dnmt1p*, are produced by alternative usage of multiple first exons. *Dnmt1s* is expressed in somatic cells. *Dnmt1p* is found only in pachytene spermatocytes, whereas *Dnmt1o* is specific to oocytes and preimplantation embryos. Here we determined that there is a tissue-dependent differentially methylated region (T-DMR) in the 5' region of *Dnmt1o* but not in that of the *Dnmt1s/1p*. The methylation status of the *Dnmt1o* T-DMR was distinctively different in the oocyte from that in the sperm and adult somatic tissues, and changed at each stage from fertilization to blastocyst stage, suggesting that active methylation and demethylation occur during preimplantation development. The T-DMR was highly methylated in somatic cells and ES cells. Analysis using Dnmt-deficient ES cell lines revealed that each of *Dnmt1, Dnmt3a* and *Dnmt3b* is partially responsible for maintenance of methylation of *Dnmt1o* T-DMR. In particular, there are compensatory and cooperative roles between *Dnmt3a* and *Dnmt3b*. Thus, the regulatory region of *Dnmt1o*, but not of *Dnmt1s/1p*, appeared to be a target of DNA methylation. The present study also suggested that the DNA methylation status of the gene region dynamically changes during embryogenesis independently of the change in the bulk DNA methylation status.

PUBLICATIONS

Papers

1. Ohgane, J., Wakayama, T., Senda, S., Yamazaki, Y., Inoue, K., Ogura, A., Marh, J., Tanaka, S.,

Yanagimachi, R. and Shiota, K. (2004). The *Sall3* locus is an epigenetic hotspot of aberrant DNA methylation associated with placentomegaly of cloned mice. Genes to Cells, *9*, 253-260.

2. Hattori, N., Nishino, K., Ko, Y.G., Ohgane, J., Tanaka, S. and Shiota, K. (2004). Epigenetic control of mouse *Oct-4* gene expression in embryonic stem cells and trophoblast stem cells. J. Biol. Chem., *279*, 17063-17069.

3. Nishino, K., Hattori, N., Tanaka, S. and Shiota, K. (2004). DNA methylation-mediated control of Sry gene expression in mouse gonadal development. J. Biol. Chem., *279*, 22306-22313.

4. Senda, S., Wakayama, T., Yamazaki, Y., Ohgane, J., Hattori, N., Tanaka, S., Yanagimachi, R. and Shiota, K. (2004). Skewed X-inactivation in cloned mice. Biochem. Biophys. Res. Commun., *321*, 38-44.

5. Imamura, T., Yamamoto, S., Ohgane, J., Hattori, N., Tanaka, S. and Shiota, K. (2004). Non-coding RNA directed DNA demethylation of Sphk1 CpG island. Biochem. Biophys. Res. Commun., *322*, 593-600.

6. Hattori, N., Abe, T., Suzuki, M., Matsuyama, T., Yoshida, S., Li, E. and Shiota, K. (2004). Preference of DNA methyltransferases for CpG islands in mouse embryonic stem cells. Genome Res., *14*, 1733-1740.

7. Ko, Y. G., Nishino, K., Hattori, N., Arai, Y., Tanaka, S. and Shiota, K. (2005). Stage-by-stage change in DNA methylation status of DNA methyltransferase 1 *(Dnmt1)* locus during mouse early development. J. Biol. Chem., in press.

Reviews

8. Shiota, K. (2004). DNA methyaltion profiles of CpG islands for Cellular differentiation and development in mammals. Cytogenetic and Genome Res., *105*, 325-334.

9. Ohgane, J., Hattori, N. and Shiota, K. (2004). Analysis of tissue-specific DNA methylation during development. Methods Mol. Biol., *289*, 371-382.

10. 坂本英樹，塩田邦郎(2004)発生プログラムと組織・細胞特異的DNAメチル化プロファイルの形成，わかる実験医学シリーズ「注目のエピジェネティクスがわかる」(羊土社)，90-95.

11. 鈴木雅子，塩田邦郎(2004)発生と治療用クローニングのエピジェネティックス．Medical Science Digest (ニュー・サイエンス社)，*30*, 26-30.

EDUCATION

1. Dr. K. Shiota was invited to give a seminar on "Cell- and tissue-specific DNA methylation profiles in mammalian genome: Epigenetics of Embryo Development", at Canadian Workshop on "Human Reproduction & Reproductive Biology", Ottawa, Canada, May 3-5, 2004.

2. Dr. K. Shiota was invited to give a lecture on "DNA methylation profiles of CpG islands during cellular differentiation and development in mammals." At McGill University, Reproductive and Developmental Biology Seminar, Montreal, Canada, May 6, 2004.

3. Dr. K. Shiota was invited to give a lecture on "DNA methylation profiles of CpG islands for cellular-differentiation and development in mammals", at 13th Conference of the International Society of Differentiation, Honolulu, USA, September 5-9, 2004.

4. Dr. K. Shiota was invited to give a seminar on "Epigenetic marks by DNA methylation specific to cell types", at IX International Congress of Reproductive Immunology, Hakone, Japan, October 11-15, 2004.

5. Dr. K. Shiota was invited to give a seminar on "DNA methylation profiles for evaluation of epigenetic risk", at International Symposium on Environmental Endocrine Disrupters 2004, Nagoya, Japan, December 15-17, 2004.

SOCIAL CONTRIBUTIONS AND OTHERS

1. 東京大学大学院農学生命科学研究科　教授
2. 東京大学大学院農学生命科学研究科応用動物科学専攻
3. 日本繁殖生物学会　理事
4. 生殖免疫学会　評議員
5. 日本獣医学会生理・生化学分科会　副会長
6. 日本再生医療学会　評議員

F. GENETIC STRAINS RESEARCH CENTER

F-a. Mammalian Genetics Laboratory Toshihiko Shiroishi Group

RESEARCH ACTIVITIES

(1) Elimination of a long-range *cis*-regulatory module causes complete loss of limb-specific *Shh* expression and truncation of the mouse limb

Tomoko Sagai, Masaki Hosoya, Youichi Mizushina, Masaru Tamura and Toshihiko Shiroishi

Point mutations in a conserved non-coding region in intron 5 of the *Lmbr1* locus, which is 1 Mb away from the sonic hedgehog *(Shh)* coding sequence, are responsible for mouse and human preaxial polydactyly with mirror-image digit duplications. In the mouse mutants, ectopic *Shh* expression is observed in the anterior mesenchyme of limb buds. Previously we showed that the conserved intronic sequence functions as a *cis*-acting regulator for limb specific expression of *Shh*. The phylogenetic studies have also shown that this sequence is highly conserved among tetrapods, and even in teleost fishes. Paired fins of teleost fishes and tetrapod limbs have evolved from common ancestral appendages, and polarized *Shh* expression is commonly observed in fins and limbs. Recently we found that this conserved sequence motif is also physically linked to the *Shh* coding sequence in teleost fish, medaka, by homology search of a newly available genomic sequence database. On the other hand, this sequence has been lost in certain limbless species of reptiles and amphibians, such as snakes and a limbless newt.

For the precise functional analysis of the intronic sequence, we eliminated it from mouse genome by means of ES targeting. The targeted mutant mice showed complete loss of *Shh* expression in the limb buds and truncation of skeletal elements distal to the stylopod/zeugopod junction. The result revealed that the sequence contains a major limb-specific *Shh* enhancer necessary for distal limb development. All results suggested that the conserved intronic sequence evolved in a common ancestor of fishes and tetrapods to control polarized expression of *Shh* in fins and limbs. It is also possible that loss of the conserved intronic sequence represents one way by which limblessness may have evolved in vertebrate species.

In addition, in order to find other putative *cis*-regulators in the intervening sequence between the *Shh* and *Lmbr1* genes, we conducted comparative genomics to search for conserved sequence in the genomes of medaka fish and mammals. We found three clustering blocks of conserved non-coding sequences between medaka fish and mammals, and the medaka sequences are located in the same scaffold as the *Shh* coding sequence. The ordering of the three blocks is conserved between medaka fish and mammals, with the most distant block found in intron 5 of *Lmbr1*. The function analysis for these conserved non-coding sequences is also underway.

(2) Genetic incompatibility between X-linked loci and two autosomal regions causes hybrid breakdown between two mouse subspecies

Ayako Oka, Nobuo Takagi[1], Kiyotaka Toshimori[2], Toshiyuki Takano[3], Akihiko Mita, Yoichi Mizushina, Noriko Yamatani, Hiromi Yamamoto, Kazuo Moriwaki[4] and Toshihiko Shiroishi ([1]Department of economics, Hokusei Gakuen University, [2]Department of Anatomy and Developmental Biology (G1) Graduate School of Medicine, Chiba University, [3]Division of Population Genetics, NIG, [4]RIKEN, BRC)

Hybrid breakdown is a type of reproductive failure that appears after F_2 generation of crosses between different species or subspecies. It is caused by incompatibility between alleles of interacting genes. We have studied the hybrid breakdown using a consomic strain, C57BL/6J-X^{MSM}, in which the X chromosome of C57BL/6J (derived mostly from *Mus musculus domesticus*) is substituted by the X chromosome of the MSM/Ms strain *(M. m. molossinus)*. Males of this consomic strain are sterile, whereas F_1 hybrids between C57BL/6J and MSM/Ms are completely fertile. This finding implies that incompatibility between X-linked gene(s) and other chromosomal gene(s) causes

the hybrid breakdown observed in C57BL/6J-X^{MSM}. The previous our study revealed that at least three X-linked QTLs are responsible for the sperm head abnormality and male sterility of C57BL/6J-X^{MSM}.

In this study, we conducted the whole-genome QTL analysis to detect loci that interact with the X-linked genes. The result suggested that one locus in Chr 1 and two loci in Chr 11 interact with the X-linked genes for proper male reproduction in parental strains, and disruption of this interaction causes the hybrid breakdown. Phenotype characterization showed that sperms of C57BL/6J-X^{MSM} fail to penetrate the zona pellucida of egg, and is responsible for disability to fertilize. This result implies that the causative genes located in X-chr, Chr 1 and Chr 11 orchestrate sperm function to penetrate zona pellucida.

(3) Comparison of BAC-end genome sequences of Japanese wild mice-derived MSM/Ms strain with the whole genome sequence of standard laboratory strain C57BL/6J detected vast amount of SNPs

Kuniya Abe[1], Hideki Noguchi[2], Keiko Tagawa[3], Misako Yuzuriha[1], Atsushi Toyoda[2], Toshio Kojima[2], Kiyoshi Ezawa[4], Naruya Saitou[4], Masahira Hattori[2], Yoshiyuki Sakaki[2], Kazuo Moriwaki[1] and Toshihiko Shiroishi ([1]RIKEN BRC, [2]RIKEN GSC, [3]Inst. Mol. Embryology, Kumamoto Univ., [4]Div. of Population Genet., NIG)

MSM/Ms is an inbred strain established from the Japanese wild mouse, *Mus musculus molossinus* in National Institute of Genetics, Mishima. Inbreeding generation numbers of this strain reached to F73 as of the end of 2004. It is believed that subspecies *molossinus* has substantially contributed to the genome constitution of common laboratory strains of mice, although the majority of their genome is derived from the west European *M. m. domesticus*. Information on the *molossinus* genome is thus essential not only for genetic studies involving *molossinus* but also for characterization of common laboratory strains. We constructed an arrayed BAC library from male MSM/Ms genomic DNA, covering ~11× genome equivalent. Both ends of 176,256 BAC clone inserts were sequenced, and 62,988 BAC end-sequence (BES) pairs were mapped onto the C57BL/6J genome (NCBI mouse Build 30), covering 2,228,164 kbp or 89% of the total genome. Taking advantage of the BES map data, we established a computer-based clone screening system. Comparison of the MSM/Ms and C57BL/6J sequences revealed 489,200 candidate SNPs in 51,137,941 bp sequenced. The overall nucleotide substitution rate was as high as 0.0096. The distribution of SNPs along the C57BL/6J genome was not uniform: the majority of the genome showed a high SNP rate, and only 5.2% of the genome showed an extremely low SNP rate (percentage identity=0.9997); these sequences are likely derived from the *molossinus* genome. This result indicated usefulness of the MSM/Ms strain in genetical analysis particularly in cross expreiments with domesticus derived standard inbred strains.

(4) Characterization of *M00745* mutants resembling *Rim3* mutants

Shigekazu Tanaka, Hiroshi Masuya[1], Shigeru Wakana[1], Masaru Tamura and Toshihiko Shiroishi ([1]RIKEN Genomic Sciences Center)

A dominant skin mutant *M00745* was generated by the ENU-mutagenesis project of RIKEN Genomic Sciences Center, and it exhibits scarring alopecia resembling the phenotype of *Rim3* and *Re*^{den}. First, we carried out linkage analysis based on totally 70 backcross progeny generated from a backcross of (*M00745*/+ x JF1) F$_1$ to JF1. This inter-subspecific backcross showed tight linkage of *M00745* to a marker *D11Mit124* and the *GsdmA* cluster. Sequencing of the *M00745* mutant revealed a point mutation in the C-terminus of *GsdmA-3*, which is the causative gene for *Rim3* and *Re*^{den}. This single base substitution leads to a nonsense mutation at the amino acid residue conserved among all members of the Gsdm family. It suggested that the conserved C-terminus of *GsdmA-3* is important for regulation of proliferation and differentiation of epidermal cells.

To elucidate phenotypic difference of two mutations, *Rim3* and *M00745*, we employed histological analysis of the two mutant mice. Degenerated hair follicles, atrophy of sebaceous glands, hyperproliferation of epidermis and hair follicles were commonly observed in the both mutants, but epidermis of *M00745*/+ mice was much thicker than that of the *Rim3*/+ mouse. Another difference observed between *Rim3*/+ and *M00745*/+ mice was late-onset corneal opacity. *Rim3*/+ mice exhibited corneal opacity after 3 months of age, whereas *M00745*/+ mice did not even after 10 months of age. We found that meibomian

glands of the *Rim3*/+ mouse were cystic and seemed to be dysfunctional, but *M00745*/+ mouse has cells containing meibum. It indicated that meibum lipid is depleted in *Rim3*/+, but not in *M00745*/+. The affected structure of the meibomian glands may explain the late-onset corneal opacity specifically observed in *Rim3*/+ mice. Difference of the phenotypes between *Rim3*/+ and *M00745*/+ mice is possibly attributable to difference of the mutation type in the C-terminus of GsdmA-3, although we can not exclude another possibility that different genetic background of the *Rim3*/+ and *M00745*/+ is responsible for the difference.

PUBLICATIONS

Papers

1. Abe, K., Noguchi, H., Tagawa, K., Yuzuriha, M., Toyoda, A., Kojima, T., Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriwaki, K. and Shiroishi, T. (2004). Contribution of Asian mouse subspecies *Mus musculus molossinus* to genomic constitution of strain *C57BL/6J*, as defined by BAC-end sequence-SNP analysis. *Genome Res. 14*, 2439-2447.
2. Oka, A., Mita, A., Sakurai-Yamatani, N., Yamamoto, H., Takagi, N., Takano-Shimizu, T., Toshimori, K., Moriwaki, K. and Shiroishi, T. (2004). Hybrid breakdown caused by substitution of the X chromosome between two mouse subspecies. *Genetics. 166*, 913-924.
3. Nemoto, M., Morita, Y., Mishima, Y., Takahashi, S., Nomura, T., Ushiki, T., Shiroishi, T., Kikkawa, Y., Yonekawa, H. and Kominami, R. (2004). *Ahl3*, a third locus on mouse chromosome 17 affecting age-related hearing loss. *Biochem Biophys Res Commun. 324*, 1283-1238.
4. Sagai, T., Masuya, H., Tamura, M., Shimizu, K., Yada, Y., Wakana, S., Gondo, Y., Noda, T. and Shiroishi, T. (2004). Phylogenetic conservation of a limb-specific, cis-acting regulator of Sonic hedgehog *(Shh). Mamm Genome. 15*, 23-34.
5. Sagai, T., Hosoya, M., Mizushina, Y., Tamura, M. and Shiroishi, T.(2005). Elimination of a long-range cis-regulatory module causes complete loss of limb-specific *Shh* expression and truncation of the mouse limb. *Development 132*, 797-803.
6. Sakai, T., Miura, I., Yamada-Ishibashi, S., Wakita, Y., Kohara, Y., Yamazaki, Y., Inoue, T., Kominami, R., Moriwaki, K., Shiroishi, T., Yonekawa, H. and Kikkawa, Y. (2004). Update of mouse microsatellite database of Japan (MMDBJ). *Exp Anim. 53*, 151-154.

Books

7. 城石俊彦(2004)「亜種マウスのゲノム比較ーどうして表現型が違うのか」ヒトゲノム・生命システムの理解と医学への展開, Molecular Medicine Vol.41臨時増刊.

EDUCATION

1. 城石俊彦「野生マウスを利用したゲノム機能解析」ゲノムひろば2004 in福岡, 福岡, 8月, 2004. (in Japanese).
2. Dr. T. Shiroishi organized a satellite symposium at the 36th Annual Meeting of The Japanese Society of Genetics, Osaka, May, 2004.
3. Dr. T. Shiroishi gave a lecture at the Tohoku University, Sendai, September, 2004 (in Japanese).
4. Dr. T. Shiroishi was invited to give a seminar on "Mouse Limb development" at the Tohoku University, Sendai, September, 2004 (in Japanese).
5. Dr. T. Shiroishi organized a satellite symposium at the 27th Annual Meeting of the Molecular Biology Society of Japan, Kobe, December, 2004.

SOCIAL CONTRIBUTION AND OTHERS

Dr. T. Shiroishi served as an editor for Mammalian Genome.
Dr. T. Shiroishi served as an editor for Genes and Genetic Systems.
Dr. T. Shiroishi served as a member of the council of The Genetics Society of Japan.
Dr. T. Shiroishi served as a member of the council of Japanese Association for Laboratory Animal Science.
Dr. T. Shiroishi is a project director of Mouse Functional Genomics Research Group, Genome Science Center, RIKEN.

F-b. Mammalian Development Laboratory Yumiko Saga Group

RESEARCH ACTIVITIES

(1) Molecular mechanism of somite segmentation

Mitsuru Morimoto, Yu Takahashi[1] and Yumiko Saga
([1]National Institute of Health Sciences)

The somite is the first morphologically distinct segmental unit formed in a vertebrate embryo and gives rise to metameric structures such as vertebrae, ribs and skeletal muscles. A 'clock and wavefront' model has been proposed to explain the underlying mechanism, in which the periodicity is generated by a 'clock' in the posterior PSM and this temporal periodicity is then translated into the segmental units in the 'wavefront'. The wavefront is thought to exist in the anterior PSM and progress backwards at a constant rate. The majority of the oscillating genes are related to Notch-signaling pathway. However, an important question is whether the level of Notch activity really oscillates and how such oscillation is translated into a segmental pattern in the anterior PSM. We have succeeded in visualizing the levels of Notch1-activity in mice, by using an antibody against an activated form of Notch1, and show that it oscillates in the posterior PSM but is arrested in the anterior PSM. Detailed analyses of the distribution of an activated form of Notch1 and Mesp2 protein in the anterior PSM demonstrate that somite boundaries are formed at the interface between Notch1-activated and -repressed domains and that Mesp2 protein localizes in the Notch1-respressed domain. Thus, Mesp2 plays a crucial role in translation of the temporal oscillation of Notch activity into the formation of regularly-spaced somites.

Somitogenesis is not only an attractive example of metameric pattern formation but is also a good model system for studies of morphogenesis, particularly epithelial-mesenchymal interconversion in vertebrate embryos. Mesp1 and Mesp2 are homologous bHLH transcription factors that are co-expressed in the anterior presomitic mesoderm (PSM) just prior to somite formation. Analysis of possible functional redundancy of Mesp1 and Mesp2 has been prevented by the early developmental arrest of Mesp1/Mesp2 double-null embryos. We performed chimera analysis using either Mesp2-null cells or Mesp1/Mesp2 double-null cells, to clarify (1) possible functional redundancy and the relative contributions of both Mesp1 and Mesp2 in somitogenesis and (2) the cell autonomy of Mesp functions in several aspects of somitogenesis. Both Mesp2-null and Mesp1/Mesp2 double-null cells fail to form initial segment borders and to acquire rostral properties, confirming that the contribution of Mesp1 is trivial in these aspects. In contrast, Mesp1/Mesp2 double-null cells contribute to neither epithelial somite nor dermomyotome formation while Mesp2-null cells partially contribute to incomplete somites and dermomyotome. This indicates that Mesp1 has a significant role in the epithelialization of somitic mesoderm. We have found that the roles of the Mesp genes in epithelialization and establishing rostral properties are cell autonomous. However, we also found that epithelial somite formation with normal rostro-caudal patterning by wild-type cells was severely disrupted by the presence of Mesp mutant cells, showing non-cell autonomous effects and supporting our previous hypothesis that Mesp2 is responsible for the rostro-caudal patterning process itself in the anterior PSM via cellular interaction.

(2) Regulation of *Mesp1* and *Mesp2* expression

Masayuki Oginuma, Yukuto Yasuhiko[1] and Yumiko Saga ([1]National Institute of Health Sciences)

Mesp1 and *Mesp2* are both expressed in the early mesoderm and presomitic mesoderm (PSM) just before segmentation. In order to analyze global transcriptional regulation of both *Mesp1* and *Mesp2*, we started to use BAC transgenic strategy. *Mesp1* and *Mesp2* are located in head to head orientation and separated only by 16 kb. Since the expression pattern is very similar, we have introduced *ALP* gene in the *Mesp1* locus and *LacZ* gene in the *Mesp2* locus of a BAC clone using homologous recombination in bacteria. By introducing this BAC transgene, we were able to reproduce both gene expressions in a single embryo. We have generated several BAC constructs that have deletions in the possible regulatory region. Using those transgenes, we will find regulatory regions required for faithful expression of *Mesp1* and *Mesp2* in vivo.

For Mesp2 specific enhancer, we have already determined the minimum sequence required for PSM expression. Further analyses revealed that the expression was regulated by both Tbx6 and Notch signaling. Although we have identified the direct binding of Tbx6, no RBP-Jk-binding is found, indicating that another mechanism is involved in the *Mesp2* regulation, which is mediated via Notch signaling.

(3) Search for target genes of Mesp2 transcription factor

Kaoru Mitsui, Yoshiro Nakajima and Yumiko Saga

Mesp2 transcription factor is critically important for both segment border formation and establishing rostro-caudal patterning of somites. However, the direct target genes are not identified yet. In order to obtain information of the target sequences, we have employed SELEX method. Random nucleotide oligomers with PCR primers were mixed with Mesp2 protein and the possible binding sequences were identified. However, the sequence is different from so-called E-box or N-box sequences that are known to be consensus sequences for bHLH-type transcription factor. However, the similar sequence is found in the promoter region of Delta-like 1 gene that is known to be negatively regulated by Mesp2. We are currently examining the sequence by generating transgenic mice.

In addition, the analyses of Mesp2-knockout mouse and the genetic studies have identified several genes affected by the Mesp2 during somitogenesis. One of genes that are downregulated in the Mesp2-knockout mouse is *EphA4*. We are expecting that the *EphA4* might be a direct target of Mesp2 protein since the expression domain is very similar to that of *Mesp2* and the expression disappears in the Mesp2-null background. To address this question, we have searched the enhancer sequence of *EphA4*. Using transgenic mouse strategy, we have succeeded to identify the minimum sequence of *EphA4* enhancer. The enhancer contains multiple E-box sequences and the deletion of some of them results in the loss of enhancer activity, which strongly indicates that these E-boxes are responsible for the activation of *EphA4* by Mesp2.

(4) Cardiovascular development and Notch signaling

Hiroki Kokubo, Yusuke Watanabe, Yoshiaki Okamura, Wataru Saito and Yumiko Saga

Notch signaling is required for multiple aspects of cardiovascular development, including arterial-venous differentiation, septation and cushion formation. Despite recognition of the importance of the Notch pathway in normal cardiovascular development, the proximate downstream effectors are not yet known. Likely candidate effectors are members of the hesr (hairy and enhancer of split related) family of bHLH transcription factors. However, mutational analysis of individual *hesr* genes has so far failed to elucidate their role in all Notch-mediated cardiovascular signaling events. An example of this is evident for mutants of *gridlock*, the zebrafish counterpart of mouse *hesr2*, which have vascular defects, whereas mouse *hesr2* mutants have only cardiac defects. One possible explanation for these differences could be functional redundancy between *hesr* family members. Mice lacking the *hesr1* gene are viable and fertile, whereas knockout mouse of both *hesr1 and hesr2* is embryonic lethal at 11.5 days postcoitum (dpc) and recapitulates most of the known cardiovascular phenotypes of disrupted Notch pathway mutants including defects in arterial-venous specification, septation and cushion formation. Taken together, our results demonstrate a requirement for hesr1 and hesr2 in mediating Notch signaling in the developing cardiac and vascular systems. In addition, we have tried to find out arterial specific enhancer of *hesr1* gene. There are at least 5 RBP-Jk binding sites in the upstream region of *hesr1* gene and disruption of these sites resulted in the great reduction of *hesr1* expression. However, further detailed transgenic analyses revealed that another enhancer sequence might be involved in the regulation of *hesr1* in the artery.

We are also studying Notch function in heart morphogenesis using transgenic mouse, which has activated Notch1 after floxed CAT gene under the control of CAG promoter. We can achieve forced Notch activation by intercrossing the transgenic mouse with Cre-expressing mouse. Since *Mesp1* is expressed in the heart precursor cells, we can drive Notch activation only in *Mesp1*-lineage using *Mesp1*-cre mouse. The trans-heterozygous mouse exhibits heart abnormality, which is characterized by abnormal myocardial trabeculation and AVC formation. To know the downstream genes involved in the abnormal morphogenesis, we have conducted GeneChip analysis, by which ectopic induction of *Wnt2, BMP6, Ilet-1* was detected in addition to *hesr1*. Since hesr1 is known to be a direct target of Notch signaling, we have first asked whether these changes in gene expression is mediated by hesr1 or not by activating Notch1 in the absence of hesr1. Interestingly, the changes in gene expression were observed even in the absence of hesr1, indicating that these genes are up-regulated by a hesr1-independent Notch signaling pathway.

(5) Functional analysis of mouse *nanos* genes

Masayuki Tsuda, Atushi Suzuki, Hitomi Suzuki, Makoto Kiso and Yumiko Saga

Previously we have isolated three mouse *nanos* genes *(nanos1, nanos2* and *nanos3).* Among them, we focus on function of nanos2 and nanos3 since these are specifically expressed and play important roles on germ cell development.

We have shown that nanos2 is expressed in the germ cells in both embryonic and adult testes and disruption of nanos2 resulted in a complete loss of germ cells in the testis. To understand the molecular mechanism leading to the loss of germ cells, we have to know the direct targets of nanos2 since it is known that nanos protein works as a translational repressor in the *Drosophila* germ cells. To achieve this, we have first tried to generate good antibodies against nanos2 and naons3 to be used for immunoprecipitation. Using purified nanos2 and nanos3 protein expressed in E.coli, we have succeeded to generate antibodies for both proteins. The nanos2 antibody can be used for immunoprecipitation of nanos2 protein from embryonic testes. Therefore, this antibody would be useful to identify not only the target genes but also proteins interacting with nanos2 in future studies. We are also interested in the regulation of *nanos2* expression. The *nanos2* expression starts in the PGC after entering male gonad and the expression is testis-specific and not observed in any other tissues. Using transgenic mouse, we have identified a core enhancer region required for the testis specific expression of *nanos2.* The identification of the upstream signal would be a key to understand a mechanism of early male germ cell specification.

PUBLICATIONS

Papers

1. Haraguchi, S., Saga, Y., Naito, K., Inoue, H. and Seto, A. (2004). Specific gene silencing in the pre-implantation stage mouse embryo by an siRNA expression vector system. Mol Reprod Dev. *68*, 17-24.
2. Kokubo, H., Miyagawa-Tomita, S., Tomimatsu, H., Nakashima, Y., Nakazawa, M., Saga, Y. and Johnson, RL. (2004). Targeted disruption of hesr2 results in atrioventricular valve anomalies that lead to heart dysfunction. Circ Res. *95*, 540-7.
3. Okazaki, N., Kikuno, R., Ohara, R., Inamoto, S., Koseki, H., Hiraoka, S., Saga, Y., Seino, S., Nishimura, M., Kaisho, T., Hoshino, K., Kitamura, H., Nagase, T., Ohara, O. and Koga, H. (2004). Prediction of the coding sequences of mouse homologues of KIAA gene: IV. The complete nucleotide sequences of 500 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries. DNA Res. *11*, 205-18.
4. Okazaki, N., Kikuno, R., Ohara, R., Inamoto, S., Koseki, H., Hiraoka, S., Saga, Y., Kitamura, H., Nakagawa, T., Nagase, T., Ohara, O., Koga, H. (2004). Prediction of the coding sequences of mouse homologues of FLJ genes: the complete nucleotide sequences of 110 mouse FLJ-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries. DNA Res. *11*, 127-35.
5. Kii, I., Amizuka, N., Shimomura, J., Saga, Y., Kudo, A. (2004). Cell-cell interaction mediated by cadherin-11 directly regulates the differentiation of mesenchymal cells into the cells of the osteo-lineage and the chondro-lineage. J Bone Miner Res. *19*, 1840-9.
6. Ishikawa, A., Kitajima, S., Takahashi, Y., Kokubo, H., Kanno, J., Inoue, T. and Saga, Y. (2004). Mouse Nkd1, a Wnt antagonist, exhibits oscillatory gene expression in the PSM under the control of Notch signaling. Mech Dev. *121*, 1443-53.

Books

7. 相賀裕美子(2005)Notchシグナルの多様性, 実験医学(増刊), 「発生・分化・再生研究2005」*23*, 64-72.

EDUCATION

1. Dr. Y. Saga gave a lecture at the Fujita Health University, January, 2004 (in Japanese).
2. Dr. Y. Saga gave a lecture at the Keio University, Aprile, 2004 (in Japanese).
3. Dr. H. Kokubo gave a lecture at the Tokyo University of Science on "Functions of hesr1 and hesr2 in the heart morphogenesis", September, 2004 (in Japanese).

F-c. Mouse Genomics Resource Laboratory Tsuyoshi Koide Group

RESEARCH ACTIVITIES

(1) Identification of genetic factors responsible for difference in spontaneous home-cage activity between KJR and C57BL/6

Juzoh Umemori and Tsuyoshi Koide

It has been previously reported that mice move spontaneously in the habituated home-cage according to a light/dark rhythm, and most mice are active during the dark period. This kind of activity is called a spontaneous home-cage activity. In a variety of strains, mice of C57 were relatively hypoactive, in contrast to mice of KJR were especially hyperactive in habituated home-cage. We studied the spontaneous home-cage activity in terms of the ethological components, genetics factors and biological functions.

For the ethological and genetic studies, F2 progeny (BKF2) were made between C57BL/6J and KJR strains, and analyzed. Ethological studies showed that total home-cage activity (THA) could be divided to two independent components, active time (AT) that indicated temporal elements of spontaneous home-cage activity, and average activity (AA) that indicated quantitative elements. A little correlation between AT and AA suggested that AA and AT are independent measurement and share only a few common genetic factors.

In order to study the genetic bases responsible for the differences in spontaneous home-cage activity between KJR and C57BL/6J, we conducted quantitative trait loci (QTL) analyses using BKF2 progeny. QTL analyses showed three significant QTLs associated with the spontaneous home-cage activity. These QTLs were designated as *hyperlocomotive activity related QTL1 (Hylaq1), Hylaq2 and Hylaq3.* (All *Hylaq* loci controlled THA that was defined by summing up effects of the AT and the AA.) *Hylaq1* was located in a middle region on chr2 and mainly controlled activity time. *Hylaq2* was located on the distal side of *Hylaq1*on chr2 and controlled both of AT and AA. *Hyaq3* was located near the telomeric region on chr10, and controlled mainly AA.

It has been reported that dopamine (DA) system in basal ganglia in brain is associated with movement and spontaneous locomotive activity. To search for candidate genes in the above QTL region, and to understand the biological mechanism that is associated with the spontaneous home-cage activity, strain-comparative analyses on the DA system have been conducted with pharmacological and biochemical methods. The results showed that difference in home-cage activity might be associated with neither DA synthesis, release nor uptake in presynapse but with downstream pathway in DA system following activation of DA receptors in dopaminergic neuron. In particular, function of D1 DA pathway that has role of suppressing the spontaneous activity is possibly reduced in KJR.

(2) A systematic analysis of genetic factors associated with the diversity of mouse behavior

Akinori Nishi, Toshihiko Shiroishi[1] and Tsuyoshi Koide ([1]Mammalian Genetics Laboratory)

A great variation of behavioral pattern is observed in a variety of animals, such as among individuals, populations and species. A number of studies indicated that the behavioral diversity is greatly influenced by genetic factors. The aim of our study is to identify the genetic factors associated with the diversity of behavioral traits.

In order to study genetic mechanism associated with the behavioral diversity, we are using mouse strains, C57BL/6 and MSM. C57BL/6 is a widely used laboratory strain. The other strain, MSM, was established from Japanese wild mice after the 20 generations of brother-sister breeding. These two strains are genetically different to a great extent and known to exhibit a behavioral difference in a variety of behavioral tests.

Recently, consomic strains, in which one of the chromosomes of C57BL/6 was replaced with the corresponding chromosome from MSM, were established for a series of chromosomes. These strains are expected as powerful tools for investigating complex traits, which are influenced by multiple genetic factors. In this project, we applied a systematic test battery on the consomic strains to reveal genetic factors associated with behavioral traits. As a behavioral test battery, the home cage test, novel home-cage test and light/dark box test were applied.

In the tests, behavioral traits in a habituated condition were measured by the home cage test, the activity in the novel place and habituation for the novel situation were evaluated by the novel home cage test, and anxiety related place preference behavior was measured by using the light/dark box test.

As the result of these tests, we found that B6-6CMSM consomic strain, which carries MSM-derived chromosomal segment from D6Mit166 to D6Mit12 on chromosome six, exhibited significantly lower activity in the home cage test and the novel home cage test, and indicated higher anxiety related behavior in the light/dark box test comparing to C57BL/6.

(3) Descriptive and Temporal Analysis of the Open-field Behavior in Wild-derived Mouse Strain

Aki Takahashi, Toshihiko Shiroishi[1] and Tsuyoshi Koide ([1]Mammalian Genetics Laboratory)

The open-field is one of the widely used methods for measuring emotionality. In this test, an animal is simply placed into a novel, brightly lit, arena from which escape is barred. It is known that genetic factors make a large contribution to phenotypic variation in various measures of open-field behavior. Most study have used two major variables, ambulation and defecation, that totaled per test. However, mice actually show more various behaviors such as sniffing, rearing, and grooming in the field, and frequency of those behaviors change with time. In this point of view, detailed analysis of the open-field behavior is important to understand psychological status of mice. We therefore described detail of open-field behavior with 12 ethograms (sniffing, locomotion, stretching, leaning, rearing, grooming, face-washing, digging, gnawing, jumping, pausing, freezing) by counting presence of those behaviors in 5 seconds. We previously examined descriptive and temporal analysis of open-field behavior in genetically diverse mouse strains: ten wild-derived mouse strains (PGN2, BFM/2, HMI, CAST/Ei, NJL, BLG2, CHD, SWN, KJR, MSM), one fancy strain JF1, and one laboratory strain C57BL/6. It was revealed that those strains had a great diversity in the temporal changes of those open-field behaviors.

To elucidate genetic mechanisms underlying the open-field behavior, we have been performing this descriptive study on consomic mouse strains made by C57BL/6 and MSM. These parental strains exhibited great difference in the open-field behavior, C57BL/6 tended to show locomotion and leaning in the later session but MSM exhibited grooming and freezing. Each consomic mouse strain has same genetic background as C57BL/6 except for one chromosome which is introduced from MSM. By examining a series of strains, we will be able to reveal the chromosomes that are associated with the behavioral difference between C57BL/6 and MSM. A half of those strains have been so far analyzed, and some of the strains exhibited different behavior comparing to the parental strain C57BL/6. One strain carrying chromosome 17 from MSM, B6-17MSM, showed reduced locomotion and rearing, and extended risk assessment behavior, stretching. In order to find genetic locus responsible for this behavioral pattern, we will conduct further genetic analysis using B6-17MSM and C57BL/6 strains.

PUBLICATIONS

Reviews

1. 古瀬民生，小出剛(2004)痛覚感受性の個体差にかかわる遺伝学的解析. オピオイド研究の進歩と展望, 155-170. 鎮痛薬・オピオイド研究会編　ネオメディカル.
2. 小出　剛(2004)野生由来マウス系統を用いた自発運動性の行動遺伝学的解析. 動物心理学研究*54*, 59-66.

EDUCATION

1. Dr. T. Koide gave a seminar at Tokyo Institute of Psychiatry, February, 2004 (in Japanese).
2. Dr. T. Koide gave a lecture at Institute of Psychology, University of Tsukuba, October, 2004 (in Japanese).

F-d. Model Fish Genomics Resource
Noriyoshi Sakai Group

RESEARCH ACTIVITIES

(1) Gene targeting system with RNA interference (RNAi) in zebrafish

Minori Shinya, Kimiko Saka and Noriyoshi Sakai

Gene silencing via short interfering RNAs (siRNAs) has proved to be a useful tool in studying gene function

in plants, invertebrates and mammalian systems. To date, gene silencing effects of siRNAs have confirmed in the zebrafish, which is an emerging model for developmental and diseases analysis. However, the effects were temporal (only in early developmental stages) and sometimes mosaic in an embryo because of the method injecting siRNAs into the one-cell stage of embryos. We recently succeeded in the production of transgenic zebrafish from *in vitro*-cultured sperm[1]. The advantage of this technique is that the mosaicism inherent in other conventional transgenic methods is avoided. Our aim in the present study is the establishment of a rapid system with cultured sperm to generate transgenic zebrafish for gene silencing by siRNA. To achieve this, we first selected the targeted genes that we have already known the silenced phenotype and also the phenotype can be easily recognized. Then, we established the zebrafish cultured cells expressing the targeted genes, in order to determine the best sequence of siRNA for specific suppression of the gene by transfecting several possible designed siRNAs. After finding out the best siRNA, we will construct the retroviral vector of DNA vector-based siRNA constructs and then infect to the cultured spermatogonia with the virus. By *in vitro* insemination with the sperm derived from the infected spermatogonia, we can obtain the transgenic zebrafish producing siRNA in all the cells. Once the above system is developed, the inducible siRNA in the specific cells or organs will be easily conducted, and thus, the technique leads to a high-throughput system for genome-wide loss-of-function studies.

(2) Analysis of functionally distinctive testicular cell lines of zebrafish to support male germ cell development

Kayoko Kurita, Kataaki Okubo[1], Masaru Matsuda[1], Yoshitaka Nagahama[1] and Noriyoshi Sakai ([1]Laboratory of Reproductive Biology, National Institute for Basic Biology)

Sertoli cells are important to germ cells in everything from male sex-determination to spermatogenesis. In spermatogenesis, Sertoli cells interact directly with germ cells in the testis to induce the complex process required for the production of functional sperm. These cells mediate the production of many molecules as well as cell junctions and adhesion. The function of many of these molecules and the regulation of gene expression remain unclear. We recently established two testicular cell lines of zebrafish with distinct functions to support the development of male germ cells[2]. Twelve cell lines were established by single-colony isolation from tumor-like testis-derived ZtA6 cells. Multiple features characteristic of Sertoli cells such as phagocytic activity and transcription specific genes (*sox9a* and Wilms' tumor suppressor *WT1*) were analyzed in the lines. The lines, ZtA6-2 and ZtA6-12, showed almost the same characteristics as Sertoli cells, but exhibited distinctive features when male germ cells were co-cultured with each line as feeders. The *in vitro* fertilization by the culture of germ cells with ZtA6-12 produced more embryos than that with ZtA6-2. In contrast, ZtA6-2 gave rise to significantly larger clumps of germ cells after a 12-day culture compared to the ZtA6-12 cell line. Expression of *vas*, strongly expressed in spermatogonia and premeiotic spermatocytes, was prolonged in the culture using ZtA6-2 feeders, while it was reduced with the germ cells on the ZtA6-12 feeders. Compared with the previous results obtained on the original ZtA6 cells, these results suggested that the function of the ZtA6-2 cells was directed to stimulate the proliferation of spermatogonia, and ZtA6-12 to stimulate the differentiation into sperm.

These cell lines will facilitate investigation of Sertoli cell molecules that contribute to the proliferation and differentiation of spermatogonia. We are currently working on the molecular cloning of genes expressed specifically in each line. 205 clones and 256 clones were isolated by the subtraction experiments with ZtA6-2 cDNA minus ZtA6-12 cDNA and *vice versa*. Several have already been confirmed by screening with dot blot hybridization and virtual Northern hybridization to cDNA of each line. *In situ* hybridization for sections of a testis showed that some of the cDNAs were expressed in Sertoli cells. Analysis of other cDNAs is under investigation.

(3) Culture condition for zebrafish spermatogonial stem cells

Kenji Saito and Noriyoshi Sakai

Spermatogonial stem cells (SSCs) maintain spermatogenesis by continuous production of the daughter cells and cells that differentiate into spermatozoa. Sertoli cells are the only type of somatic cells that closely interact with SSCs to create a

favorable environment inter testis. However, the regulatory mechanisms are still unclear because of the difficulty of identifying and manipulating an individual SSC and the surroundings.

To establish *in vitro* system in which SSCs proliferate continuously, we performed the isolation procedure of A type spermatogonia and determined a culture condition to proliferate A type spermatogonia on a Sertoli cell feeder layer. When zebrafish were treated with busulfan reagent, we found the decline of differentiating germ cells, such as spermatocytes and spermatids, while A type spermatogonia increased after 4 days of treatment. Then, on day 7 germ cells of the testis were consisted of only A type spermatogonia and sperm. In teleosts, a single spermatogonium that enclosed within germinal cysts of Sertoli cells is defined as A type. Furthermore, we observed that the A type spermatogonia proliferate rapidly after the damage with busulfan reagent. The definition and the observation suggest that A type spermatogonia are SSCs of a teleost. When enzymatically dissociated testicular cells containing A type spermatogonia were co-cultured on ZtA6-6 Sertoli cell line, proliferation of A type spermatogonia was observed without differentiation by a BrdU incorporation experiment after 14 days of culture. These results indicate that A type spermatogonia self-renew in the culture condition that might represent testicular microenvironments to maintain SSCs.

(4) Analysis of the ability of cultured embryonic cells derived from different developmental stages to induce the anterior-posterior axis

Megumi Hashiguchi and Noriyoshi Sakai

One of the most fascinating phenomena in primary embryonic induction is the regional specificity of the neural structures that are produced. Primary embryonic induction can be divided into three major components; head specific, trunk specific and tail specific induction.

We established the condition to culture zebrafish embryonic cells continuously without any artificial immortalization treatment. When cultured cells were transplanted into a blastulae embryo, the cells from different developmental stages had different abilities to induce specific second axes. Cells derived from earlier stage embryos (gastrula stage) induce complete anterior structure. Cells from the pharyngula stage (embryos segmented) induce anterior structures either with cyclopia or without an eye. Cells from later stages (early larva) induce posterior structure with otic vesicles and a heart. Interestingly, cells from later stages had low induction efficiency, but the efficiency increased when proliferation of the cells was arrested with mitomycin C. Whole-mount *in situ* hybridization for *emx-1* (telencephalon marker), *krox 20* (rhombomere marker), and *shh* (notochord marker) indicated similar patterns of gene expression to the specific induction. Cultured cells derived from embryos at various developmental stages, therefore, change their properties to induce the second axis from an anterior to a posterior orientation according to their developmental stage. In addition, cells may secrete more inducer(s) when proliferation is stopped, particularly in later stages. These cultured cells can be used to find cellular factors involved in anterior-posterior specific induction.

PUBLICATIONS

Papers

1. Kurita, K., Burgess, S.M. and Sakai, N. (2004). Transgenic zebrafish produced by retroviral infection of *in vitro*-cultured sperm. Proc. Natl. Acad. Sci. USA *101*, 1263-1267.

2. Kurita, K. and Sakai, N. (2004). Functionally distinctive testicular cell lines of zebrafish to support male germ cell development. Mol. Reprod. Dev. *67*, 430-438.

3. Matsumoto, T., Yukawa, W., Nozaki, Y., Nakashige, R., Shinya, M., Makino, S., Yagura, M., Ikuta, T., Imanishi, T., Inoko, H., Tamiya, G. and Gojobori, T. (2004). Novel algorithm for automated genotyping of microsatellites. Nucleic Acids Res *32*, 6069-6077.

Reviews

4. Sakai, N. (2004). Genetically modified sperm in fish. ISB News Report *April*, 1-3.

EDUCATION

他大学／機関での講義やセミナー

1. Dr. N. Sakai gave a seminar on "Genetically modified sperm in fish" at Shizuoka University, March, 2004 (in Japanese).

2. Dr. N. Sakai gave a course of lectures at the Department of Marine Bioscience of Fukui Prefectural University, April-August, 2004 (in Japanese).

3. Dr. N. Sakai was invited to give a seminar on "Genetically modified sperm in fish" at National Institute of Child Health and Human Development of NIH, Bethesda, USA, July, 2004.

F-e. Plant Genetics Laboratory Nori Kurata Group

RESEARCH ACTIVITIES

(1) Molecular cytogenetics of plant meiosis analyzed with rice mutants

Ken-ichi Nonomura, Mutsuko Nakano, Mitsugu Eiguchi, Akio Miyao*, Hirohiko Hirochika* and Nori Kurata (*Natl. Inst. Agrobiol. Sci.)

This study aims to dissect genetic machinery controlling meiosis and sporogenesis of higher plants. By screening of insertional mutant lines tagged by the *Tos17*, the endogenous retrotransposon of rice, we could identify and isolate two meiotic genes of interest; *PAIR1 (HOMOLOGOUS PAIRING ABERRATION IN RICE MEIOSIS 1)* (Nonomura et al. 2004a) and *PAIR2* (Nonomura et al. 2004b). The *PAIR2* gene encodes a protein of 610 amino acids, homologous to the yeast *HOP1* and *Arabidopsis ASY1*. We raised a polyclonal antibody against recombinant PAIR2 protein and identified the PAIR2 expression in rice meiocytes by western blotting and immuno-fluorescent analyses. The PAIR2 protein began to accumulate in the nucleus of meiocytes following the onset of pre-meiotic DNA synthesis, and disappeared at metaphase I. In early meiosis I, the PAIR2 antibody stains the lateral elements, which are components of chromosomal axes and play important roles in mediating homologous chromosome pairing. The PAIR2 protein was removed from the chromosomes as soon as the homologous chromatins aligned and paired. This observation suggests the PAIR2 function in recruiting the central components, which are important to establish a rigid structure between lateral elements of homologous partners.

The immuno-cytological analysis at diakinesis, prior to metaphase I, revealed that the most PAIR2 proteins were removed from chromosomes, but not from the centromeric regions. The centromere-specific histone H3 variant always colocalized with the PAIR2 through early to late meiosis I. In addition, our study revealed that the rice centromeres paired preceding to usual homologous chromosome pairing. We are now underway to make it sure that the centromere pairs at early meiosis I prior to pairing of other chromosomal regions in rice meiocytes.

(2) Structural and functional analysis of rice *OsHAP3* genes

Yukihiro Ito, Thiruvengadam Thirumurugan, Kazumaru Miyoshi and Nori Kurata

We identified 11 genes *(OsHAP3A to OsHAP3K)* which encode a HAP3/NF-YB subunit of CCAAT-box binding complex in rice by cDNA screening and a database search. We showed that three genes, *OsHAP3A, OsHAP3B* and *OsHAP3C*, are involved in chloroplast biogenesis (Miyoshi et al. 2003). We next examined the function of *OsHAP3E*, which is most closely related to *LEC1* and *L1L*. The *LEC1* and *L1L* are *Arabidopsis* members of the HAP3/NF-YB subunit and are critical regulators of embryogenesis. The *OsHAP3E* was also expressed in developing embryo and young panicle. The antisense plants of *OsHAP3E* showed slightly increased plant height, and its overexpressing plants showed dwarfism, erected leaves and abnormal inflorescence with the increased number of spikelets. Microarray analysis showed that brassinosteroid-related genes and MADS-box genes were affected in the overexpressing plants. These results suggested that *OsHAP3E* may be involved in brassinosteroid-related response and panicle development. Yeast two-hybrid analysis showed that OsHAP3E interacted with several specific members of OsHAP2 and OsHAP5, which are other subunits of heterotrimeric CCAAT-box binding complex. These results suggested that several different combination of HAP complexes with OsHAP3E can be formed in rice cells. Considering that *LEC1* and *L1L* are critical regulators of embryogenesis in *Arabidopsis, OsHAP3A, B* and *C* are controlling factors of chloroplast biogenesis and *OsHAP3E* is thought to regulate hormonal and/or influorescence development pathways, plant *HAP3/NF-YB* genes are suggested to have diverse functions depending on each gene.

(3) Regulation of expression of *KNOX* family class 1 homeobox genes of rice

Yukihiro Ito and Nori Kurata

KNOX family class1 homeobox genes play a vital function for shoot apical meristem (SAM) formation and maintenance. We found that cytokinin, which is necessary for shoot regeneration, induced the expression of *OSH1*, a member of the *KNOX* family, in the callus. Expression of *OSH1* was observed an hour after cytokinin treatment, whereas expression of *ORR6*, a homologue of a cytokinin early response gene, was detected within 30 min. It was assumed that the cytokinin-induced gene expression is under control of a histidine phosphotransfer system which consists of a histidine kinase, a phosphotransmitter protein and a response regulator. To examine this possibility, we generated transgenic plants overexpressing *COS3* histidine kinase gene, *OHP2* phosphotransmitter gene or *ORR1* response regulator gene. These transgenic plants showed enhanced expression of *OSH1* upon treatment with cytokinin. This suggested that these genes are involved in the cytokinin-induced *OSH1* expression. Analysis of transgenic plants with dominant negative constructs of these genes is underway.

(4) Generation and screening of enhancer trap lines of rice

Yukihiro Ito and Nori Kurata

To isolate tissue-specific enhancers and valuable mutants defective in various steps of rice development and to clone the corresponding genes, we are generating enhancer trap lines of rice. We employed an enhancer trap system based on the *Ac/Ds* two-element system and the GUS reporter gene. We already have established the system by which *Ds-GUS* transposants are efficiently identified in rice (Ito et al. 2004).

This year we screened about 2,000 transposed lines and identified 146 lines with tissue-specific GUS activity. GUS expression was observed in various organs including embryo (7 lines), leaf (59 lines), root (22 lines), panicle (11 lines), flower (22 lines), seed (18 lines) and shoot apex (7 lines). We also determined the *Ds-GUS* insertion site of the 27 GUS-positive lines. We identified a trapped gene by a *Ds-GUS* transposition in a line P1772, which expressed the GUS on a leaf lesion after wounding, as an example. In this line the *Ds-GUS* was shown to be inserted near a full-length cDNA AK073621 on chromosome 12, and this gene was expressed transiently in the leaf after wounding. This indicated that the enhancer of this gene was trapped in P1772 and that our trap system can usefully work not only on enhancer selection but also on molecular identification of the trapped genes. Cloning of other trapped enhancers is underway. These data will be opened through Oryzabase (http://www.shigen.nig.ac.jp/rice/oryzabase/top/top.jsp).

(5) Positional cloning of a segregation distortion gene detected in a progeny of a cross between Japonica and Indica rice

Yoshiaki Harushima and Nori Kurata

The aim of this study is isolation of the most prominent barrier on chromosome 3 detected in the F2 of Nipponbare-Kasalath hybrid by positional cloning, and elucidation of the molecular nature of the individual reproductive barrier. We have clarified the pollen with Kasalath genotype at the gametophyte gene preferentially fertilize eggs by 94% probability in the maternal plant that is heterozygote or Kasalath homozygote at the interactive locus on chromosome 6.

For fine mapping the gametophyte gene, we have selected plants with recombination in the candidate region from 5691 F2 and 473 backcross plants and mapped it by their selfed progeny test. A prominent causal gene is an *Argonaute* gene. Expression of this *Argonaute* gene in both Nipponbare pollen and Kasalath pollen was confirmed. RT-PCR and micro-array analysis showed there was no difference in the expression level. However, high frequency of non-synonymous substitutions, 8 out of 9 single nucleotide changes in a coding region, was detected. This suggests rapid evolution that is common feature of reproductive barriers. There are two possibilities to explain Kasalath pollen preferential fertilization; one is Kasalath pollen gene accelerates fertilization, another is Nipponbare pollen gene retards it in pollen competition. For complementation test, two kinds of Nipponbare near isogenic lines, both the aimed gametophyte gene and the interactive gene regions are Kasalath homozygote

and only the interactive gene region is Kasalath homozygote, were transformed by the Nipponbare or Kasalath genomic fragment containing the *Argonaute* with *GFP*. These transformants are now growing and their selfed seeds will be tested for the transmission of introduced *Argonaute* and *GFP*.

(6) Identification of nuclear proteins showing various sub-nuclear distribution in rice

Tadzunu Suzuki, Kazuki Moriguchi and Nori Kurata

We screend rice nuclear proteins from three developmental stages; young panicle, flowering panicle and regenerating calli by using NTT (nuclear transportation trap) system and identified over 500 genes including many novel proteins and many transcription factors. A subset of these genes were examined for their nuclear localization by introducing GFP-fusion constructs into onion epidermal cells. A variety of protein localization in the cell; for instance, nuclear peripheral, foci-like, speckle-like, chromatin-associated and matrix-associated distribution, were observed (Moriguchi et al. 2005). Some of these genes were revealed to be expressed at specific stages of rice development by RT-PCR and *in situ* hybridization. One of these which we designated OsAHP1 was further characterized for its tissue specificity and sub-nuclear localization by detecting with an antibody produced. The results showed that this protein is expressed transiently in the floral tissue and co-localized with chromatins in the nucleus (Moriguchi et al. 2005).

(7) Comparative genomics among cultivated and wild rice species

Yukie Sano, Hiroyuki Kanamori*, Nobukazu Namiki*, Yukiko Yamazaki and Nori Kurata (*STAFF institute, Tsukuba)

We have started to analyze gene structure and expression differences between cultivated rice, *Oryza sativa* (AA genome species), and wild rice of *O. punctata* (BB genome) and *O. officinalis* (CC genome). Total of 4,000 cDNA clones randomly picked up from the libraries of SAM (shoot apical meristem) or very early influorescence were subjected to EST sequencing and blast search analyses. Some clones were further analyzed for full length clone sequences. Totally about 2,900 clones have been qualified in their sequences. Blast searches showed about 70% of the clones had highly identical sequences with rice cDNA or genomic sequences. Around 3% (from 7% to 0%) base substitution and in frame or out of frame insertion/deletion were detected between cultivated and wild rice in the clones which showed identities with rice sequences. Simultaneous examination for expression profiling using 22K microarrays revealed scores of genes which showed quite large differences in their expression between cultivated and wild rice. We are going to identify gene structure of clones having no similarity with rice and showing much higher expression in either cultivated or wild rice.

(8) Rice genetic resource project in NBRP and Oryzabase

Nori Kurata, Ken-ichi Nonomura, Yukihiro Ito and Yukiko Yamazaki

The National Bioresource Project (NBRP) was organized in 2002 to conserve and distribute biological resources for scientific communities. Rice is one of the organisms whose resources are required to amplify, conserve and distribute to the scientific community as genetic resources. In the NIG, we are dealing 2,000 wild rice accessions covering 10 genomes and 23 species collected all over the world for over 50 years. Four rice sub-centers in other institutions dealing with different kinds of genetic materials are joined to this project. Resource materials and their information can be accessed at; http://shigen.lab.nig.ac.jp/rice/oryzabase/nbrpStrains/aboutNbrp.jsp;jsessionid=9AD3EACD512A1DA03ABB65874CF6381D.tomcat4_6 in the Oryzabase.

The Oryzabase is a comprehensive rice biological database composed of 15 sections of information. Major data characteristic in the Oryzabase are; mutants and their trait genes collection, wild rice collection and development/anatomy of rice, together with genetic maps, physical maps, basic biological data and so on. The rice genetic resources committee centered in the NIG is responsible to collect and curate the contents of information and construction, and maintenance of those data is carried out by the genetic informatics laboratory. The DB is accessible at; http://www.shigen.nig.ac.jp/rice/oryzabase/top/top.jsp

Publications

Papers

1. Miyoshi, K., Ahn, B-O., Kawakatsu, T., Ito, Y., Itoh, J-I., Nagato, Y. and Kurata, N. (2004). *PLASTOCHRON1*, a timekeeper of leaf initiation in rice, encodes cytochrome P450. *Proc. Natl. Acad. Sci. USA 101*, 875-880.
2. Ito, Y., Chujo, A., Eiguchi, M. and Kurata, N. (2004). Radial axis differentiation in a globular embryo is marked by *HAZ*, a PHD-finger homeobox gene of rice. *Gene 331*, 9-15.
3. Nonomura, K.I., Nakano, M., Murata, K., Miyoshi, K., Eiguchi, M., Miyao, A., Hirochika, H. and Kurata, N. (2004). The insertional mutation of rice *PAIR2* gene, the ortholog of *Arabidopsis ASY1*, caused a defect in homologous chromosome pairing in meiosis. Mol. Genet. Genomics *271*, 121-129.
4. Nonomura, K.I., Nakano, M., Fukuda, T., Eiguchi, M., Miyao, A., Hirochika, H., Kurata, N. (2004). The novel gene *HOMOLOGOUS PAIRING ABERRATION IN RICE MEIOSIS 1* of rice encodes a putative coiled-coil protein required for homologous chromosome pairing in meiosis. Plant Cell *16*, 1008-1020.
5. Salina, E.A., Adonina, I., Vatolina, T. and Kurata, N. (2004). A comparative analysis of the composition and organization of two subtelomeric repeat families of *Aegilops speltoides* Tausch.and related species. Genetica *122*, 227-237.
6. Moriguchi, K., Suzuki, T., Ito, Y., Yamazaki, Y., Niwa, Y. and Kurata, N. (2005). Functional isolation of novel nuclear proteins showing a variety of sub-nuclear localizations. Plant Cell *17*, 389-403.

Reviews

7. Itoh, J.I., Nonomura, K.I., Ikeda, K., Yamaki, S., Inukai, Y., Yamagishi, H., Kitano, H. and Nagato, Y. (2005). Rice plant development: from zygote to spikelet. Plant Cell Physiol. *46*, 23-47.
8. Kurata, N., Miyoshi, K., Nonomura, K.I., Yamazaki, Y. and Ito, Y. (2005). Rice mutants and genes related to organ development, morphogenesis and physiological traits. Plant Cell Physiol. *46*, 48-62.

Book

1. 野々村賢一・土井一行(2004)「生殖研究」, イネゲノム配列解読で何ができるのか(農業生物資源研＋農文協, 矢野昌裕・松岡信編), 90-102.
2. 倉田のり(2005)パキテン染色体標本, クロモソーム：植物染色体研究法(養賢堂, 福井・向井・谷口編)印刷中.

EDUCATION

1. 倉田のり：北海道大学集中講義：札幌８月
2. 倉田のり：名古屋大学集中講義：名古屋11月
3. 倉田のり：東京大学集中講義：東京12月

SOCIAL CONTRIBUTIONS

1. 倉田のり：日本学術会議育種学連携委員
2. 倉田のり：日本育種学会幹事
3. 倉田のり：日本育種学会学会賞選考委員
4. 倉田のり：生物遺伝資源イネ小委員会委員長
5. 倉田のり：農林水産省評価専門委員
6. 倉田のり：Rice Genetics Newsletter Editor
7. 倉田のり：NSF project advisory committee member

F-f. Microbial Genetics Laboratory Akiko Nishimura Group

RESEARCH ACTIVITIES

(1) Dynamics of bacterial tubulin: coupling model between DNA replication and cell division

Ippei Inoue, Kenji Yasuda and Akiko Nishimura

The most approaches understanding cell division have asked whether cell division occurs coupling with DNA replication. To solve this problem, we examined the dynamics of FtsZ, an essential key protein of *Escherichia coli* cell division, in relation to the stage of DNA replication by creating a fusion protein containing FtsZ and a green fluorescent protein (GFP). FtsZ started to assemble at potential division site coincident with initiation of DNA replication, all FtsZ assembled in the ring structure coincident with termination of replication, and the ring constricted after nucleoids separation. These results suggest that cell division occurs in association with DNA replication.

(2) A complete set of *Escherichia coli* open reading frames in mobile plasmids and their successful application to the systematic identification of cell division mutant *(fts)* genes

Kimiko Saka, Maki Tadenuma, Shinsuke Nakade, Noriko Tanaka, Hideaki Sugawara, Ken Nishikawa, Nobuyuki Ichiyoshi, Masanari Kitagawa, Hirotada Mori, Naotake Ogasawara and Akiko Nishimura

To facilitate genetic studies of *Escherichia coli*, we constructed a complete set of mobile plasmid clones of intact open reading frames (ORFs). The vector for mobile plasmid clones carries the following genes and sequences; (i) *bom, rom,* and *mob* of ColE1, which enable the plasmid to transfer from F^+ to F^- by F-mediated conjugation; (ii) P_{tac}/*lacI*^{q}, which enables control of the expression of the cloned ORF by IPTG; (iii) the ColE1 replication origins, *ori* and *pri*, to maintain a low plasmid copy number in the cell under normal growth condition; (iv) the ampicillin resistant gene, *bla*, for selection; (v) the unique multi-cloning site of pJF118HE; (vi) universal primer recognition sites -21M13 and SP6 for confirmation of the cloned ORF by sequencing; and (vii) *rrnB* with a transcriptional terminator which prevents read-through. Since replication of the plasmid *in vivo* does not require protein synthesis, the plasmid can be enriched during inhibition of protein synthesis that leads to inhibition of chromosomal replication without blocking the plasmid replication. To clone an ORF with its native SD into pNT3, we amplified the predicted ORF along with 20 bp of the upstream region by PCR using primer sets designed to contain the appropriate restriction sites. With this procedure, native products should be produced even if fragments longer than necessary are cloned. The plasmids carrying each ORF were introduced into an F^+ *recA* strain and stored in 96-well microtiter plates. In this way, 96 clones can be transferred simultaneously to F^- bacteria using the conjugative system. In order to simplify the screening method, we investigated the possibility of searching for positive clones from a mixed pool of unrelated clones. We first made a test stock composed of subpools of 48, 24, 12, or 6 clones within single wells of a 96-well microtiter plate. A complementation test was carried out using *ftsA, ftsI, ftsW,* and *ftsZ* mutants. All four mutants were complemented by the subpools of 48, 24, 12, or 6 clones, or by a single clone. There were fewer and smaller colonies in the *ftsZ*-positive patches on the 1 mM IPTG plate than on the 0.1 mM IPTG or no IPTG plates. The results are supported by the facts that low-level expression of *ftsZ* (less than two times) can complement the *ftsZ* mutant, but high level expression of *ftsZ* (more than four times) causes inhibition of growth of the *ftsZ* mutant. The size of the *ftsA*-positive patches was the largest on the 1 mM IPTG-supplemented plate and very small on a plate without IPTG. The *ftsA* mutants might need high-level expression of FtsA for normal growth when it is provided from plasmid-born *ftsA*. The results for *ftsI* were similar to *ftsA*. Changes in IPTG concentration did not affect the size of the *ftsW*-positive patches. These results indicate that the 48-clone pool is effective for complementation studies if we use the three selection plates containing 0, 0.1, and 1 mM of IPTG. Having demonstrated that we can identify the positive ORF within a mixture of 48 clones, we created two types of stock: a single microtiter plate containing pools of 48 clones in each well, and a second stock composed of 45 microtiter plates containing the individual clones. In summary, to find the clone that complements the mutant, we first identify the positive mixed pool and then determine which clone in the mixture complements the mutant. This provides a convenient procedure for systematic identification of ORFs that suppress or complement mutations.

(3) Global regulation of cell division: Isolation of a whole set of cell division mutants

Kimiko Saka, Noriko Matsumoto and Akiko Nishimura

The entire nucleotide sequence of *E. coli* has been analyzed, and 4311 ORFs have been demonstrated, but the functions of more than half of these ORFs remain unknown. The greater part of these ORFs are considered to be involved in coordinating cell proliferation. To thoroughly analyze the hierarchy and network responses in expression of cell division genes, as one of the model cases for post-genome analysis, we have identified a whole set of cell division genes using above mobile plasmid clone sets by their ability to complement a filament-temperature-sensitive (*fts*) cell division mutants. A total of 339 *fts* strains from the Hirota thermosensitive (Ts) mutant bank, which form multi-cellular filaments at 41°C without immediate

arrest of DNA synthesis or an increase in cell mass were tested. We found that 278 *fts* mutants were complemented by 403 of ORF clones. Of these, 69% of the *fts* mutants were complemented by one ORF each. Sequence analysis of genomic DNA of 10 of these *fts* mutants showed that all had missense mutation in the corresponding ORF. However, 15% of the *fts* mutants were complemented by two ORFs, and 16% by three ORFs. These may contain the allele of the *fts* mutant gene and a high-dosage suppressor gene(s). Sequence analysis of five of these *fts* mutants showed that all five had missense mutation corresponding to one ORF and no mutation in the remaining ORF(s). The cog database (http://www.ncbi.nlm.nih.gov/COG/new/) was consulted for functional annotation of individual genes. From the functional annotation analysis, known cell division genes were found in only 6% of the *fts* mutants, while 30% were unknown. Of all genes identified, 24% were involved in a basic process, such as DNA replication or protein synthesis, despite the fact that the mutants involved in this category had presumably been excluded by the process of selection of the *fts* mutants. 19% were involved in cellular processes, such as ion transport and signal transduction; and 21% were involved in metabolism. These results suggest that various intracellular reactions are coordinated with cell division and that the *E. coli* cell cycle is the result of the coordination of multiple independent processes, so that each daughter cell receives a complete copy of cell components. We are currently planning to purify these mutations in cells with wild-type background, and thoroughly analyze the hierarchy and network responses in expression of cell division genes using DNA chip technologies along with these mutants, as one of the models for post-genome analysis for global cellular regulation in *E. coli*.

PUBLICATIONS

Papers

1. Saka, K., Tadenuma, N., Nakade, S., Tanaka, N., Sugawara, H., Nishikawa, K., Ichiyoshi, N., Kitagawa, M., Mori, H., Ogasawara, N. and Nishimura, A. (2004). A complete set of *Escherichia coli* open reading frames in mobile plasmids and their successful application to the systematic identification of cell division mutant (*fts*) genes. (DNA res., imprinting).

Database

2. http://www.shigen.nig.ac.jp/ecoli/strain/top.jsp

SOCIAL CONTRIBUTIONS AND OTHERS

特許

1. 出願番号未定，マルチウェルプレート，西村昭子・梶谷隆文，大学共同利用機関法人情報システム研究機構・有限会社カジックストレーディング
2. 特許2004-195247，超薄型マルチウェルプレートの製造法，西村昭子・富川宗博・相吉三宏，大学共同利用機関法人情報システム研究機構・静宏産業株式会社

学会活動

Dr. A. Nishimura was appointed for a member of editorial board of "Microbiology and Culture Collections".

F-g. Invertebrate Genetics Laboratory Ryu Ueda Group

RESEARCH ACTIVITIES

(1) RNAi mutant fly bank for comprehensive analyses of gene function in Drosophila

Ryu Ueda, Misako Taniguchi, Yukiko Sado[1], Kaoru Saigo[2] and Kuniaki Takahashi ([1]JST, [2]Graduate School of Science, University of Tokyo)

Genome sequencing projects have revealed the number of genes for several model organisms for genetics. The small worm *Caenolabditis elegans*, which is composed of only 959 cells, has 19,000 genes in its genome. On the other hand, *Drosophila melanogaster*, which has a long and sophisticated alimentary canal, a tubular heart that circulates hemolymph, and a large brain composed of over 10^4 cells, harbors only 13,800 protein-coding genes. Considering there is such a small number of fly genes, each one of them may have an essential function in fly development and behavior. In other words, it may be easy to detect and analyze gene function in the fly by reverse genetics because the abnormal phenotype will frequently appear when knocking down a target gene whose function is unknown. We are planning to investigate the function of fly genes comprehensively as a suitable model for studying the functional genomics of multicellular

organisms.

How does one investigate the function of all 13,800 genes in the fly? We use RNA interference (RNAi) to knock down the activity of the target gene. RNAi is one of the emerging technologies with which to investigate gene function in multicellular organisms. When introduced into the cell, double stranded RNA (dsRNA) works as a specific mutagen for each gene. That is, dsRNA recognizes host mRNA and digests it in a sequence-specific manner, and consequently brings a loss-of-function mutation phenotype to the host cell. The detailed mechanism of this RNAi phenomenon has not yet been elucidated, but it works efficiently in many multicellular organisms, including humans.

We coupled the RNAi with the GAL4-UAS gene expression system to induce a conditional loss-of-function mutation in the fly. The GAL4-UAS system is a binary system for inducing transgene expression, in which two fly lines are used. One is the GAL4 driver fly line, which expresses yeast transcription factor, GAL4, in a specific cell/tissue or at a specific developmental stage in favor of the GAL4 transgene. The other fly line harbors a transgene on the chromosome, in which an appropriate gene to be expressed is fused to the UAS promoter, the GAL4 target. When these two fly lines are crossed with each other, we can observe in the fly progeny that the GAL4 protein induces target transgene expression in a driver-specific conditional fashion. In this GAL4-UAS system, when we use a UAS-transgene having an inverted repeat (IR) sequence, the transcribed RNA may form a dsRNA in the cell and induce a loss-of-function mutation by the RNAi mechanism. Such inducible RNAi caused by the transcription of an IR sequence was first successfully adopted to gene function analysis in *C. elegans*. It was then also found to be effective in fly genetics. By making a UAS-transformation vector containing an IR sequence of the gene predicted by the fly genome project, and by introducing it into a fly line (IR fly), a mutant phenotype of the gene can be easily observed in any cell or at any developmental stage of the progeny, whenever the IR fly is crossed to an appropriate GAL4 driver fly.

We are expanding this inducible RNAi to the whole genome of the fly. This process involves two major procedures.

1) *in vitro* construction of transformation vectors containing an IR sequence from each of the 13,800 predicted genes.

2) Transformation of IR vectors by injecting them into fly eggs and establishment of IR fly lines by traditional genetic methods.

As of the end of 2004, over 6,900 transformation vectors had been constructed, 6,300 of which have been successfully introduced into the fly. We will continue this work in the next year, and may be able to add up more than 2,000 IR fly lines to our fly bank.

Along with the establishment of IR fly lines, basic characterization of the target genes is conducted using these fly lines. All of the IR fly lines are crossed to the Act5C-GAL4 fly. The Act5C-GAL4 induces the UAS-transgene in all cells at all developmental stages. Thus, if the gene targeted by RNAi has functions that are indispensable for fly development, the progeny of IR and GAL4 flies should die before the adult flies emerge. Among the 1954 genes tested, 51.1% of the fly lines showed lethality. This value is rather high compared to that obtained by classical genetics (25%), while the fact that many of the genes tested here were considered to have important functions in various aspects of fly development by our collaborators may bring about such a high score. Detailed analyses on known genes and greater accumulation of data are necessary. We are currently collaborating with 70 groups. The usefulness of inducible RNAi for investigating gene function in *Drosophila* is being revealed in many aspects. We published 5 papers using RNAi flies[3)~7)] and 2 papers on RNAi mechanism[1), 2)] in 2004.

This work was supported in part by financial assistance to Dr. Ueda from the Mitsubishi Kagaku Institute of Life Sciences (MITILS).

PUBLICATIONS

Papers

1. Ui-Tei, K., Naito, Y., Takahashi, F., Haraguchi, T., Ohki-Hamazaki, H., Juni, A., Ueda, R. and Saigo, K. (2004). Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference. Nucleic Acids Res., *32*, 936-948.

2. Ui-Tei, K., Ueda, R., Zenno, S., Takahashi, F., Doi, N., Naito, Y., Yamamoto, M., Hashimoto, N., Takahashi, K., Hamada, T., Tokunaga, T. and Saigo, K. (2004). RNA interference induced by transient or stable expression of hairpin structures of double stranded RNA in *Drosophila* and mammalian cells.

Molekulyarnaya biologiya, *38*, 228-238.
3. Pili-Floury, S., Leulier, F., Takahashi, K., Saigo, K., Samain, E., Ueda, R. and Lemaitre, B. (2004). *In vivo* RNA interference analysis reveals an unexpected role for GNBP1 in the defence against Gram-positive bacterial infection in *Drosophila* adults. J. Biol. Chem. *279*, 12848-12853.
4. Ichimiya, T., Manya, H., Ohmae, Y., Yoshida, H., Takahashi, K., Ueda, R., Endo, T. and Nishihara, S. (2004). The twisted-abdomen phenotype of *Drosophila* POMT1 and POMT2 mutants coincides with their heterophilic protein O-mannosyltransferase activity. J. Biol. Chem., *279*, 42638-42647.
5. Kamimura, K., Rhodes, J.M., Ueda, R., McNeely, M., Shukla, D., Kimata, K., Spear, P.G., Shworak, N.W. and Nakato, H. (2004). Regulation of Notch signaling by *Drosophila* heparan sulfate 3-O sulfotransferase. J. Cell Biol., *166*, 1069-1079.
6. Koganeya, A. K., Sasamura, T., Oshima, E., Suzuki, E., Nishihara, S., Ueda, R. and Hirabayashi, Y. (2004). *Drosophila* glucosylceramide synthase: A negative regulator of cell death mediated by proapoptotic factors. J. Biol. Chem., *279*, 35995-36002.
7. Ishimaru, S., Ueda, R., Hinohara, Y., Ohtani, M. and Hanafusa, H. (2004). PVR plays a critical role via JNK activation in thorax closure during Drosophila metamorphosis. The EMBO J., *23*, 3984-3994.

Reviews

8. Nishihara, S., Ueda, R., Goto, S., Toyoda, H., Ishida, H. and Nakamura, M. (2004). Approach for functional analysis of glycan using RNA interference. Glycoconj J., *21*, 63-68.

EDUCATION

1. R. Ueda was invited to give a seminar on "RNAi mutant fly bank" at Keio University, Tokyo, Mar., 2004 (in Japanese).
2. R. Ueda was invited to give a seminar on "RNAi technology" at the satellite meeting of the annual meeting of Japan society for Bioscience, Biotechnology, and Agrochemistry, Hiroshima, Mar., 2004 (in Japanese).
3. R. Ueda was appointed as an adjunct lecturer at Tsukuba University, Department of biological sciences to give a course on Genetics.

SOCIAL CONTRIBUTIONS AND OTHERS

1. R. Ueda joined advisory committee of National Institute of Agrobiological Sciences at Tsukuba.

F-h. Laboratory for Frontier Research Takako Isshiki Group

RESEARCH ACTIVITIES

(1) Analysis of temporal specification within late phases of Drosophila neuroblast lineage

Takako Isshiki and Ayumi Kusano

During development, neural progenitors often generate diverse cell types in an invariant order, changing their property over time. Although substantial progresses have been made in understanding the molecular mechanisms of how different cell fates are generated in order, many parts of the mechanisms still remain unknown, especially in vertebrate. We utilize the *Drosophila* central nervous system (CNS) as an excellent model system to study temporal specification of cell fates within a lineage. We previously showed that *Drosophila* neuroblasts sequentially express the transcription factors Hunchback, Krüppel, Pdm and Castor over time. However, most neuroblasts divide additional 10 times on average after they start expressing Castor. Thus, there must be subsequent mechanisms for temporal specification.

To identify yet unidentified factors involved in temporal specification within neuroblast lineage, we searched a public dataset of the expression profiles of the *Drosophila* gene transcripts. We found several transcription factors expressed later than Castor in most neuroblasts. Next, we have investigated and elucidated the precise order of their expression. Furthermore, we start investigating how neuroblasts change their property during larval stages. Recently, it has been reported that a couple of mouse homologs of these factors are also expressed in cerebral cortex in temporally regulated manner. This fact suggests that the mechanisms we are studying might be conserved across species.

(2) Investigation of molecular functions of Krüppel type zinc finger proteins in Development of the CNS

Ayumi Kusano

A *C. elegans* heterochronic gene, *lin-29*, encodes a Krüppel type zinc finger protein. Last year, we reported that *Drosophila melanogaster* Lin-29 is expressed transiently in neuroblast at very late stages of embryogenesis and subsequently in their late-born progeny. We have been investigating molecular functions of DmLin-29, by conducting genetic analysis. Our analysis revealed that DmLin-29 and Krüppel have distinct functions, although they share high homology in their zinc fingers, that are thought to bind DNA directly. This fact suggests that not their zinc fingers but other parts of the proteins are mainly responsible for making the differences in their function. Aiming to elucidate the molecular mechanism that makes the differences, we have start searching proteins that can form a complex with either only one of Krüppel or DmLin-29.

(3) In situ Detection of microRNA

Katsutomo Okamura

Hundreds of microRNAs are found in each animal, so microRNAs are thought to have a fundamental role in regulating gene expression. Northern blot analyses of their expression strongly suggest that many microRNAs have their function in development. However, the lack of a conventional method in detecting spacio-temporal distribution of microRNA makes it very difficult to study their functions in development. Okamura has developed an efficient method for detecting mature form of microRNA *in situ* in *Drosophila* embryos. So far, he has identified a few microRNAs expressed in specific cells in the nervous system. Future analyses on neural microRNAs should provide us comprehensive understanding of gene networks that control neural cell fate specification, and may reveal novel roles for microRNAs.

PUBLICATIONS

Reviews

1. Isshiki, T. and Doe, CQ. (2004). Maintaining Youth in Drosophila Neural Progenitors. Cell Cycle. *3*, 296-299.
2. 一色孝子(2004)神経幹細胞システムによる神経細胞特異化の分子機構蛋白質核酸酵素(増刊) *49*, 228-233.

G. CENTER FOR GENETIC RESOURCE INFORMATION

G-a. Genetic Informatics Laboratory Yukiko Yamazaki Group

RESEARCH ACTIVITIES

1 SHIGEN project

1.1 E coli Databases-E. coli strain DB and PEC (Profiling of Escherichia coli Chromosome) database.

Takehiro Yamakawa, Junichi Kato, Akiko Nishimura and Yukiko Yamazaki

As a result of the National BioResource Project, a lot of new E. coli resources were accumulated. These resources include mutant strains established from research activities of individual researchers, Genomic DNA plasmid such as Cosmid collection and pLC plasmid, and deletion clones for each ORF (mobile plasmid clones and deletion mutants). E. coli strain database was rebuilt in order to compile these resources and allow users to request them through the internet directly from the list page after searching. Gene information for each ORF in the strain database was linked to the GenoBase (E. coli W3110 strain Genome Database) and to the PEC (Profiling of E. coli MG1655 chromosome). Since newly developed large scale chromosomal deletion mutants of E. coli that lack 2.4% to 29.7% of the parental chromosome (Ref.5) are available for use, "Essential Genes and Minimal Genome" section was added to PEC and a minimal genomic view and detailed information of each deletion mutants were also provided. By successfully reducing the genome size of E. coli, the number of genes which were classified as "unknown" decreased. Gene's essential information as well as huge amount of resource collections would become very useful to precede E. coli researches. PEC not only collects latest public information of MG1655 but also gives comparative genomics platform among microbe genomes to researchers.

PEC database can be accessed through the SHIGEN server (http://www.shigen.nig.ac.jp/ecoli/pec/).

1.2 CARD R-DB and IMSR (International Mouse Strain Resources)

Takehiro Yamakawa, Hideki Kato, Naomi Nakagata, Kenichi Yamamura and Yukiko Yamazaki

CARD R-BASE is a database of transgenic mouse strains established and deposited by individual researchers. All resources are stored as frozen embryos and available on request. The database contains genetic backgrounds, genes destroyed/transferred, and relevant human diseases for each strain. In this year, we introduced new links from CARD R-BASE to the gene details page of the Mouse Genome Informatics (MGI). CARD-R-BASE became a member of International Mouse Strain Resources (IMSR) since a new version was implemented in July 2004. IMSR is a searchable online database of mouse strains and stocks available worldwide and maintained by Jackson Laboratory. Before merging data set of CARD R-BASE into IMSR, manual annotation on nomenclature of strain type, mutant type and gene/allele name was required. In order to encourage researchers, who deposit mouse strains, to get used to the controlled vocabulary of strain name, a new tool named "GODFATHER" was developed. GODFATHER is a simple naming tool with web-interface and it is divided into three levels depending on the level of user's knowledge. By filling necessary information such as background mouse name, gene name and laboratory code, the user can obtain a candidate name for the mouse strain. An English version of GODFATHER is also planned to be developed. CARD R-BASE also made bilateral linking to the Exchangeable Gene Trap Clones (EGTC) which provides trapped gene sequence information and its homology search results.

The International Federation of Mouse Resource (IFMR) established in 2004 is a collaborating group of Mouse Repository and Resource Centers worldwide whose collective goal is to archive and provide strains of mice as cryopreserved embryos and gametes, ES cell lines, and live breeding stock to the research community. CARD R-BASE as well as RIKEN BRC (BioResource Center) are members of IFMR. The goals of the IFMR are (1) to coordinate repositories and

resource centers, (2) to establish consistent, highest quality animal health standards in all resource centers, (3) to provide genetic verification and quality control of genetic background and mutations, and (4) to provide resource training to enhance user ability to utilize cryopreserved resources. IMSR database would be a good start of IFMR database.

The CARD R-DB is accessible at http://cardb.cc.kumamoto-u.ac.jp/transgenic/.

1.3 Mouse polymorphism database

Takehiro Yamakawa, Eri Kibukawa, Toshihiko Shiroishi and Yukiko Yamazaki

Mouse microsatellite database (MMDBJ) provided SSLPs among different mouse strains and cSNPs between B6 and MSM mouses based on their full length cDNA sequence information. Since a huge amount of MSM BAC end sequences as well as B6/MSM gSNPs are now available for use, the name of the database was changed from MMDBJ to Mouse Polymorphism Database and all information was put together in a genomic viewer. We applied the Ensembl, which is freely provided by EBI, as a genome browser and added original data of MSM BAC ends, MSM/B6 cSNPs and microsatellite polymorphism as DAS (Distributed annotation source) and then customized the browser. Although keeping the Ensembl data up-to-date is not a simple task, adding original data as DAS sources to Ensembl instead of developing an original genomic browser from scratch was quicker to implement. All data are accessible through the internet now. Then we started developing our genomic browser (shigen genomic browser) to overcome several functional limitations of Ensembl.

Mouse Polymorphism Database is released at http://shigen.lab.nig.ac.jp/mouse/polymorphism/

1.4 Oryzabase and Plant Ontology

Takehiro Yamakawa, Nori Kurata and Yukiko Yamazaki

Oryzabase is a comprehensive rice database to collect as much as public knowledge information and provide them with a user friendly interface to support rice research activities. As a member of the National BioResource Project, large amount of new rice resources were collected and ready to be distributed to researchers through the Oryzabase. We developed an online resource order system which allows users to send e-mails to order from each resource center directly from the resource list page. The system selects an appropriated Material Transfer Agreement (MTA) form depending on which resource center maintains the requested resource, and a reply mail is automatically sent back to senders as an attached file.

Oryzabase started to provide genomic sequence information using shigen-genome-browse.

POC (Plant Ontology Consortium) is a consortium that aims to develop, curate and share controlled vocabulary that describe plant growth stages and structures (anatomy). Japanese rice researchers independently started to clarify organ specific developmental stage and collected stage specifically expressed genes and their mutants from journal articles. Oryzabase integrated these data into the already existing genes and mutant collections. We have been developing multiple ontology platform in order to integrate several different ontologies and vocabularies consistently. Three dimensional matrixes consisting of "time" (development/growth), "space" (location/structure/anatomy), and "features" (characteristics) was conceptually build at the early developmental stage. Recently Zebrafish group (ZFIN) started to establish the PATO (Phenotype And Trait Ontology) to describe mutant phenotypes. PATO is described with a combination of terms and their attribute and has a more general concept than our "features" which means that most species can share this concept. Therefore, Oryzabase started to apply PATO and is replacing "feature" with PATO type expression. As a result of collecting all available information as much as possible and as soon as possible, the current Oryzabase have a lot of inconsistencies on genes, mutants and phenotypes. Our next task will be to resolve these problems.

Oryzabase is available at http://www.shigen.nig.ac.jp/rice/oryzabase/.

1.5 KOMUGI

Takehiro Yamakawa, Takashi Endo, Yasunari Ogihara, Hitoshi Tsujimoto, Taihachi Kawahara, Tetsuro Sasakuma and Yukiko Yamazaki

KOMUGI is a wheat resource database and is

unique as a wheat specific database in the world. Bioresources of KOMUGI contains wild strains; Landraces; experimental strains such as chromosome lines, cytoplasm substitution lines, and mutants; and DNA resources such as EST clones and array. The KOMUGI is also a member of the National BioResource Project and the e-mail resource order system has been implemented. Array data was newly incorporated into KOMUGI database and is used internally at the moment. The task of adding most recent information to KOMUGI gene dictionary remained although the "Macgene" system was developed for this purpose last year. Since MacGene is a stand alone type system instead of a web-based platform, there are still problems for dictionary keepers to share a common reference number when they work individually at different locations. Making cross linking between sequence accessions and KOMUGI genes is also a task that will be carried out next.

KOMUGI database is available at http://www.shigen.nig.ac.jp/wheat/komugi/.

1.6 NBRP-databases

As an information center of the National BioResource Project, we intensively support the construction of resource databases of each resource center including the following species so that all resource centers could make their resources public.

1.6.1 NBRP-C. elegans

Miharu Ikizawa, Shohei Mitani, Hiroshi Kagoshima, and Yukiko Yamazaki

NBRP-C. elegans provides deletion mutants on request after isolating targeted gene among mutant pool. Users can see the status of mutant screening on the web page and resource curator also manage the screening processes using a web-interface. This system also gives results of statistic analysis. The database has bilateral cross-linkings to WormBase, a worldwide famous C. elegans database. Promoter information is also part of this database and it is maintained by several volunteers. The common search site for resource database and promoter database are now available.

1.6.2 NBRP-Silkworm

Miharu Ikizawa, Yutaka Banno, Hiroshi Fujii, and Yukiko Yamazaki

NBTP-Silkworm database consists of the following three parts, (1) gene dictionary, (2) references, and (3) strain phenotypes. Since the traditional classification of phenotypes and resources was not consistent enough to construct an electronic database, manual annotation was performed intensively. As there is no silkworm gene database in the world, this database is the first gene dictionary of silkworm and it will be very useful for scientific researches.

We are planning to incorporate genomic information when it is opened to the public.

1.6.3 NBRP-Legume (Lotus and Glycine) database

Takehiro Yamakawa, Shoko Isobe, Masatsugu Hashiguchi, Ryo Akashi, Satoshi Tabata, and Yukiko yamazaki

NBRP-LegumeBase is a resource ordering site shared by Lotus japonica resource database and Glycine max. resource database. Lotus japonica database provides wild accessions, root culture, Rebominant Inbred Lines as resources. Some strains were characterized in phenotype when growing in northern areas and southern areas. Although Glycine max database currently only contains wild species, Glycine soja, RIL as well as DNA resources such as EST, full length cDNA clones are expected to be added. Legume follows rice (monocot model plant) and arabidopsis (dicot model plant) as the third model plant.

1.6.4 NBRP-Medaka database

Takehiro Yamakawa, Yuko Wakamatsu, and Yukiko Yamazaki

The NBRP Medaka site web server was moved from Nagoya University to National Institute of Genetics this year. Medaka Genomic Information has been incorporated into the construction of the Medaka resource database.

The Medaka site also prepares the web-based protocols "Medaka Book" in collaboration with several Medaka researchers using PukiWiki software. The

Medaka web-atlas is also among the attractive contents available and the brain part is now under construction. Interactive interface and glossary will be integrated into the atlas images and illustrations.

Medaka group applied the Zebrafish PATO ontology to describe mutant phenotypes so that the database would be designed to handle the PATO data.

1.6.5 NBRP-Drosophila database

Takehiro Yamakawa, Masatoshi Tomaru, Masatoshi Yamamoto, Ryu Ueda, and Yukiko Yamazaki

NBRP Drosophila consists of four organizations and there are 4 individual databases. In order to make these databases more useful without changing the current system, we developed an one-stop-shop site "Flystock" from where user can search resources for four databases at once and make order directly from the resource list page.

We are planning to develop a maintenance system to manage each database and to make connections between Flystock and each database.

PUBLICATION

Papers

1. Mochida, K., Yamazaki, Y. and Ogihara, Y. (2004). Discrimination of homologous gene expression in hexaploid wheat by SNP analysis of contigs grouped from a large number of expressed sequence tags. *Mol. Gen. Genomics. 270*, 371-377.
2. Sakai, T., Miura, I., Yamada-Ishibashi, S., Akita, Y., Kohara, Y., Yamazaki, Y., Inoue, T., Kominami, R., Moriwaki, K., Shiroishi, T., Yonekawa, H. and Kikkawa, Y. (2004). Update of Mouse Microsatellite Database of Japan (MMDBJ). *Exp. Anim. 53 (2)*, 151-154.
3. Kawai, Y., Ishii, Y., Arakawa, K., Uemura, K., Saitoh, B., Nishimura, J., Kitazawa, H., Yamazaki, Y., Tateno, Y., Itoh, T. and Saito, T. (2004). Structural and Functional Differences in Two Cyclic Bacteriocins with the Same Sequences Produced by Lactobacilli, *Applied and Environmental Microbiology, 70 (5)*, 2906-2911.
4. Yamazaki, Y., Nagato, Y. and Kurata, N. (2004). Oryzabase (Integrated Rice Database) in 2004, *Rice Genetics Newsletter*, 20, 9-10.
5. Hashimoto, M., Ichimura, T., Mozoguchi, H., Tanaka, K., Fujimitsu, K., Keyamura, K., Ote, T., Yamakawa, T., Yamazaki, Y., Mori, H., Katayama, T. and Kato, J. (2005). Cell size and nucleoid organization of engineered Escherichia coli cells with a reduced genome. *Mol. Microbiol. 55 (1)*, 137-149.
6. Moriguchi, K., Suzuki, T., Ito, Y., Yamazaki, Y., Niwa, Y. and Kurata, N. (2005). Functional Isolation of Novel Nuclear Proteins Showing a Variety of Subnuclear Localizations. *The Plant Cell 17*, 389-403.

Reviews/Books

7. 荒　武，山崎由紀子(2004)「大腸菌ゲノムデータベース」ゲノミクス・プロテオミクスの新展開，エヌ・ティ・エヌ，15-21.
8. 山崎由紀子(2004)「バイオリソースセンター」蛋白質核酸酵素Vol.49 No.11, 1956-1963.
9. 山崎由紀子(2004)「モデル動物の作製と維持」別冊，1-17.

Database

10. JMSR http://www.shigen.nig.ac.jp/mouse/jmsr/
11. Mouse Polymorphism DB http://www.shigen.nig.ac.jp/mouse/mmdbj/
12. CARD R-BASE http://cardb.cc.kumamoto-u.ac.jp/transgenic/
13. FlyStock http://218.44.182.89/%7Eflystock/html/index_j.html
14. NigFly http://www.shigen.nig.ac.jp/fly/
15. Oryzabase http://www.shigen.nig.ac.jp/rice/oryzabase/
16. KOMUGI http://www.shigen.nig.ac.jp/wheat/komugi/
17. Barley Germplasm Database http://www.shigen.nig.ac.jp/barley/
18. PEC http://www.shigen.nig.ac.jp/ecoli/pec/
19. E.coli Strain Database http://www.shigen.nig.ac.jp/ecoli/strain/
20. GRW http://www.shigen.nig.ac.jp/shigen/grw/
21. WGR http://www.shigen.nig.ac.jp/shigen/wgr/
22. NBRP-ChrysanthemumDatabase http://www.shigen.nig.ac.jp/kiku/
23. NBRP-Algae DB http://www.shigen.nig.ac.jp/algae/
24. NBRP-Silkworm http://shigen.lab.nig.ac.jp/silkwormbase/
25. NBRP-Legume http://shigen.lab.nig.ac.jp/legume/legumebase/
26. NBRP-C.elegans http://www.shigen.nig.ac.jp/cele/
27. NBRP http://www.nbrp.jp/

EDUCATION

京大併任助教授
教育セミナーフォーラム３月
熊本大学客員教授

SOCIAL CONTRIBUTIONS AND OTHERS

学術会議遺伝資源研究連絡委員
運営委員：ミヤコグサ，ES細胞，メダカ
評価委員：生物資源研究所BRC
検討委員：理研BRC

G-b. Publicity and Intellectual Property Unit
Munehiro Tomikawa Group

RESEARCH ACTIVITIES

2004 Annual Publicity and Intellectual Property Unit Report

This unit makes various efforts to return the outcome of research generated by NIG to the industrial world and the public while respecting the scientists' priorities in terms of scientific ideas, partnership and publications.

To be more precise, what we have done is to establish scientific results as intellectual property, explain the property rights to the industrial world intelligibly, seek collaboration with commercial industries for actualizing those results and establish partnerships. This shows the importance of carrying out both public relations and the acquisition of intellectual property.

This unit was set up in March 2003. Since April 2003, a secretary, Ms. Nozomi Yokoyama and Ms. Kiyomi Kuwahara have devoted all their energy to this work from starting up, and had assisted by creating the MTA(Material Transfer Agreement) form, and getting it off the ground. They have also been supporting various works such as planning and promotion of publicity activities and procedures of application for intellectual property rights. The results of our work in 2004 are shown as follows.

(1) Acquirement and utilization of Intellectual Property Rights

We have continued to promote a sense of entitlement in the scientists belonging to NIG and Division of Life Science in the Graduate University for Advanced Studies (Sokendai) by giving them speeches on why entitlement is necessary to the outcome of basic research. Also, in visiting laboratories, monitoring the scientific progress and thinking of each scientist, we have explained how intellectual property rights can be gained from their results. We have also established an efficient patenting procedure. We have prepared an invention exposition format to minimize the scientists' burden in informing our unit of their scientific results. Our unit helps the scientists avoid the troublesome task of acquiring of patent rights to the practical applications of their inventions.

In terms of patent applications, we have tried to promote acquirement of the patent rights so the NIG scientist does not need to bear the costs necessary for the preparation of patent documents. At this time, eight domestic patents have been applied for and four patents are under preparation, of which 15 as a whole will be applied for as NIG patents (including nine joint applications and three international patents) in the Research Organization of Information and Systems (ROIS). One patent has already come into marketing as of November 2003. We got the royalty with a small amount from its first business at April 2004.

Secondly, we have been involved in making guidelines for the Material Transfer Agreement which stipulates ownership, ethical codes and compensation for biological resources based on research results. At present, three kinds of agreements, depending on the positioning of biological resources as intellectual property, have been created and are in use. Each has a simple version corresponding to specific materials in the research field such as yeast, worms and flies (Form 1) and an intermediate version without provision for sharing of industrial property rights (Form 2) and a complete version with provision for sharing of industrial property rights (Form 3). They have been authorized and, a total of 272 agreements (approximately 80% of them are international) have been signed by the end of this year.

In addition, we have participated in the Shizuoka-TLO-Yaramaika (STLO) campaign, which facilitates putting patents owned by the NIG into

practical use.

(2) Acceleration of public relations activities

Ideally, this kind of work should be developed over a broad spectrum, however, we have focused our attention on the following main points. We have given priority to our response to visitors, promotion at academic conferences, introduction of research results to the public, and the search for potential collaboration.

Firstly, we had 9 groups of visitors (total 225 persons) in 2004: technical school students, legal apprentices, high school students, schoolteacher interns, members of the judiciary, elderly community college students and others. An overview of this institute, genetic science and current laboratory experiments was given to the visitors.

Secondly, promotion at academic conferences has been carried out in order to obtain more excellent students and scientists. We introduced the current research results obtained from NIG scientists at the Japanese Cell Biology Meeting (in Osaka), the Japanese Society for Developmental Biology meeting (in Nagoya), and the Japanese Molecular-Biology meeting (in Kobe).

Thirdly, we presented fourteen specifically excellent research results in the Hot News column on the website as an introduction of research results and academic awards for life science to the public. We have also had fourteen research results inserted in main-stream newspapers through cooperation with the media. Furthermore, we introduced two scientists to companies which were interested in their work, and fortunately, this effort led to further business-academia collaboration.

(3) Collaboration Research with academia, industry and government, and Social Action Work

I have been conducting research of Fuji Pharma Valley Cancer Diagnostics Study Group consisting of academia, industry and government. The aim of this project is to increase Research and Development potential in the Mt. Fuji area and its cities of Numazu, Mishima, Fujinomia, Fuji and Nagaizumi and to study together with companies operating in the field of medicine as well as the pharmaceutical industry on the development of cancer diagnostics through genomics and proteomics approaches. Fortunately, we could apply one patent regarding cancer diagnostics based on higher order structure of chromatin DNA this October.

Recently, the participation of public facilities, such as universities and NIG, is increasing in the Government-Industry-University-Research Cooperation Conference. In June, We attended first Third Government-Industry-University Cooperation Conference as a panelist on behalf of NIG, which was held at Kyoto. It was quite meaningful to be able to develop a relationship with industrial people who were interested in our patent and products.

In conclusion, acquiring intellectual property and engaging in returning excellent technology from scientific results to the industrial world through public relations will be important factors in gaining wide public recognition of NIG. These efforts are expected to facilitate the financial management of NIG, after turning into independent administrative entities.

H. STRUCTURAL BIOLOGY CENTER

H-a. Biological Macromolecules Laboratory Makio Tokunaga Group

RESEARCH ACTIVITIES

(1) Single Molecule Imaging and Quantitative Analysis of Nuclear Transport in Cells using Highly Inclined and Laminated Optical sheet microscopy

Makio Tokunaga and Naoko Imamoto[1] ([1]Cellular Dynamics Laboratory, Riken)

What is a key to enter the nuclei through nuclear pores? This question has been answered by visualizing single molecules inside cells. Clear video images of single molecules translocating into the nuclei are obtained using novel fluorescence microscopy, Highly Inclined and Laminated Optical sheet (HILO) microscopy. Very little is known about the interactions between transport molecules and the assembled nuclear pore complex (NPC) because of its large supramolecular structure. Obtained single-molecule video images inside cells are very clear, therefore kinetic parameters of the molecular interactions in cells are obtained through quantitative analysis. We have discovered how many molecules interact, how strong the interactions are, and which molecular interaction is the open sesame to nuclear import. We have opened up a new way to quantify molecular dynamics and interactions at the single-cell level, which is 'open sesame' toward quantitative molecular and system biology.

(2) RNG105: A Novel Regulatory Protein in Neuronal RNA Granules Responsible for Stimulation-Dependent Local Translation

Nobuyuki Shiina, Kazumi Shinkura and Makio Tokunaga

mRNA translocation and subsequent local translation in neuronal dendrites are important bases for long-term synaptic plasticity, but responsible molecules have not been fully identified. We previously identified RNG105 (RNA granule protein 105) as a component of RNA granules, which play central roles in the transport of mRNAs to the dendrites. The RNG105-localizing RNA granules contain mRNAs, such as CaM kinase II alpha, CREB and BDNF mRNAs, whose translational products play key roles in synaptic plasticity.

In this year, we have found that RNG105 contains an RNA-binding motif, and shown that RNG105 is a regulatory protein for local translation in dendrites of hippocampal neurons. RNG105 has an ability to repress translation in vivo, consistent with the finding that the RNA granule is translationally arrested in the basal conditions. Dissociation of RNG105 from the RNA granules is induced by BDNF (brain-derived neurotrophic factor), a growth factor responsible for synaptic plasticity. The dissociation of RNG105 is significantly correlated with the induction of local translation of the mRNAs near the granules. These findings indicate that RNG105 is responsible for the local translational regulation in neuronal dendrites, and suggest its implication in the regulation of local synaptic plasticity in a stimulation-dependent manner.

(3) Single Hydrogen Bonds of DNA Base Pairs Detected in Unzipping Force by Intermolecular Force Microscopy

Michio Hiroshima[1] and Makio Tokunaga ([1]Single Molecule Immunoimaging, Research Center for Allergy and Immunology, RIKEN)

Single hydrogen bonds of DNA base pairs have been measured by unzipping double-stranded DNA oligomers with an intermolecular force microscope (IMF). To detect such ultrafine forces, high resolution of force as well as high accuracy in controlling the probe position is required. IFM has achieved the force resolution of subpico-newton using ultrasensitive cantilevers. The probe position is controlled with nanometer accuracy using a feedback system, which uses laser radiation pressure to reduce thermal fluctuation of the cantilever.

Force vs. extension curves showed repeated force peaks of 10-15pN. Auto- or cross-correlation analysis and averaging of force curves were carried out to

reduce noises in the force curve. The previous studies showed that the force for separating poly(G-C) DNA was 1.5 to 2 times stronger than that for poly(A-T). However, no difference was found in the force between individual G-C and A-T base pairs. The force curve of individual G-C and A-T pairs showed three and two peaks, respectively, which are assigned to single hydrogen bonds. The force is variable but the work is constant. The work to break single hydrogen bonds is 1.3 kB・T, in other words, about 1.3-fold of the thermal energy. This is the first report to detect the force of single hydrogen bonds in biological macromolecules.

(4) Mechanism for passive force generation of invertebrate connectin revealed by single molecule measurement

Michio Hiroshima[1], Atushi Fukuzawa[2], Koscak Maruyama[2], Sumiko Kimura[2] and Makio Tokunaga ([1]Single Molecule Immunoimaging, Research Center for Allergy and Immunology, RIKEN, [2]Department of Biology, Chiba University)

Connectin is an elastic protein in muscle and is thought to keep the thick filament at the center of sarcomere or protect a myofibril from damages by extraordinary loads. Invertebrate connectin (I-connectin) is a 1960 kDa elastic protein linking the Z line to the tip of the myosin filament in the giant sarcomere of crayfish claw closer muscle. There are several extensible regions in I-connectin: two long PEVK regions, one unique sequence region and Ser-, Glu- and Lys-rich 68 residue-repeats called SEK repeats.

The force measurements of the single recombinant polypeptide including SEK or PEVK regions was performed by intermolecular force microscopy (IFM). The force versus extension curves were well fit to the wormlike chain (WLC) model. The obtained persistence lengths were 0.38±0.1nm (n=63) of SEK peptide and 3.07±1.0nm (n=35) of PEVK peptide. The persistence lengths well explain the elastic behavior of SEK and PEVK regions in muscle fiber. The value of 0.38nm indicates that the SEK region is a random coil for full length. This is the first observation of an entropic elasticity of fully random coil region contributing to the physiological function of invertebrate connectin.

(5) Asymmetric nucleocytoplasmic transport revealed by a novel assay system using planner reconstituted nuclear envelope

Atsuhito Okonogi, Michio Hiroshima[1], Nobuyuki Shiina, Shingo Kose[2], Naoko Imamoto[2] and Makio Tokunaga ([1]Single Molecule Immunoimaging, RCAI, Riken, [2]Cellular Dynamics Laboratory, Riken)

We have developed a novel *in vitro* assay system of nucleocytoplasmic transport. We aim at application of the method to new single-molecule experiments, imaging and nano- or force-measurement. Nuclear envelope was formed on a planar surface of a small agarose plate. Agarose plates were modified with glutation, and were coated with GST-RanGDP fusion protein. Nuclear envelope was formed onto the RanGDP-coated surface using extracts from *Xenopus laevis* frog eggs. Formation of Nuclear Pore Complexes was confirmed by observing import of fluorescently labeled proteins. Import assays using fluorescently labeled proteins showed that the reconstituted NPCs retain the transport activity. The dependence of the import activity on RanGTP concentration showed cooperativity.

Import and export of importin a was examined in the absence of RanGTP and other soluble factors using the present cell-free in vitro assay system. Asymmetry between import and export was discovered in the interaction of importin a and the NPC in the absence of Ran-GTP and other soluble factors. In the apparent equilibrium, import of importin a showed a difference in the concentration between at the cytoplasmic side and at the nucleoplasmic side, whereas export of importin a showed no difference. The concentration difference of importin a is found to correlate with the amount of binding of importin a with the NPC. The asymmetry between import and export was also found in the transport of importin a in the presence of NLS cargo molecules. In the presence of NLS-protein, importin a showed no biding with the NPC. The present finding shows that the NPC has an asymmetric feature in the interactions with cargo molecules and transport factors between import and export.

(6) Molecular Imaging of translation initiation factors in neuronal dendrites

Hiraku Miyagi, Nobuyuki Shiina and Makio Tokunaga

Local protein synthesis in neuronal dendrites is required for synaptic plasticity, which is associated with long-term memory storage. This protein synthesis is reported to be induced at activated postsynaptic sites. In order to investigate when and where the translation is initiated locally in the dendrites, we visualized eukaryotic translation initiation factors (eIF) 4E and eIF4G in dendrites of rat hippocampal neurons.

Immunostaning of rat hippocampal primary neurons showed that most of eIF4E and eIF4G were spreaded in distinct granular structures, and few of them were colocalized in some granular structures in the dendrites. Immunostaning of rat hippocampal slices showed essentially the same results as in the case of the primary cultures. Stimulation by brain-derived neurotrophic factor (BDNF) increased colocalization of eIF4E and eIF4G in both rat hippocampal slices and primary neurons. As the association of eIF4E and eIF4G is known to trigger translation initiation, the increase of their colocalization suggests that translation is induced. Immunostaning of rat hippocampal slices and primary neurons for eIF4G and PSD-95, a marker protein for postsynapses, showed that colocalization of eIF4G and PSD-95 was increased by BDNF stimulation. Colocalization of eIF4E and PSD-95 is also increased by same stimulation. In GFP-expressed primary neurons, colocalization of eIF4E and eIF4G in spines is increased by BDNF stimulation. These results indicate that translation in the neuronal dendrites is inhibited by separating spatially eIF4E and eIF4G, and translation is induced by BDNF stimulation at postsynaptic areas.

PUBLICATIONS

Papers

1. Kitamura, K., Tokunaga, M., Esaki, S., Iwane, A.H. and Yanagida, T. (2004). Mechanism of muscle contraction based on stochastic mechanical features of single actomyosin motors observed in vitro. BIOPHYSICS, in press.

EDUCATION

他大学／研究機関での講義やセミナー

1. Dr. M. Tokunaga gave a lecture at Tokyo Institute of Technology, Department of Bioscience and Biotechnology, August, 2004 (in Japanese).
2. Dr. M. Tokunaga gave a Seminar at the University of Tokyo, Department of Applied Chemistry, School of Engineering, September, 2004 (in Japanese).

SOCIAL CONTRIBUTIONS AND OTHERS

各種委員

1. 徳永万喜洋　バイオテクノロジー開発技術研究組合「細胞内ネットワークのダイナミズム解析技術開発」研究開発委員
2. 徳永万喜洋　科学技術振興調整費研究評価部会「細胞・生体システム研究評価WG」科学技術・学術審議会専門委員

集会／シンポジウムの主宰

1. Dr. M. Tokunaga organized a symposium entitled "Frontier of Single Molecule, GFP and Bioimaging" supported by Grant-in-Aid for Scientific Research on Priority Areas "Molecular Imaging" from MEXT, Tokyo International Forum, March, 2004.

H-b. Molecular Biomechanism Laboratory Nobuo Shimamoto Group

RESEARCH ACTIBITIES

(1) The branched mechanism of transcription initiation in *E. coli*

Motoki Susa[1], Shouji Yagi[1] and Nobuo Shimamoto[1]
([1]Structural Biology Center, National Institute of Genetics and and Department of Genetics, School of Life Science, The Graduate University for Advanced Studies)

For several decades, the mechanism of transcription initiation has been assumed to be a sequence of three essential steps: formation of a complex between RNA polymerase and a promoter (closed complex), formation of another complex with partially melted DNA duplex to form phosphodiester bonds (open complex and chemical reaction), and escape of RNA polymerase from the promoter

associated with the progress of RNA elongation (promoter clearance). There is a process called "abortive initiation", an iterative synthesis and release of oligo-RNA molecules. This process is often excluded from the mainstream of the mechanism, because its role as well as its occurrence in vivo has been unknown. However, this process has been observed in vitro with all prokaryotic and eukaryotic RNA polymerases so far isolated. The products of the process, abortive transcripts, are typically 2 to 15 nucleotides in length. On the assumption that these short transcripts are "unsuccessful precursors" of the full-length transcript, abortive synthesis has been considered to precede promoter clearance.

We have been clarifying for several years that this sequential mechanism is not the case and the initiation follows branched pathways, one of which contains the moribund complex, being defined as a complex that produces only abortive and no full-length transcripts. Followings are its characteristics. 1. The moribund complex, as well as the productive complex that synthesizes full-length product, are formed from the same homogeneous fraction of enzyme molecules, and dissociation of the molecules from the promoter DNA cancels any difference between them. 2. Structural differences between these complexes have been demonstrated. 3. At some promoters, a moribund complex is converted into a dead-end complex that still retains a short transcript but has no elongation activity. Therefore, the initiation pathway is branched into the conventional productive pathway and the abortive pathway that can lead towards a dead end. 4. The fates of a moribund complex are either inactivation as a dead-end complex, dissociation from the DNA, or direct conversion into a productive complex, and the rates of these reactions vary with the promoter. 5. There are factors that affect the fate of the moribund complex in a manner that depends on the promoter.

To examine the existence and significance of the branched pathway *in vivo*, we selected GreA and GreB for clues. At the λP_RAL promoter these factors enhance conversion of the moribund complex into the productive one, in the presence of high concentrations of initiating nucleoside triphosphate *in vitro*. If the branched mechanism exists *in vivo*, absence of the Gre factors should result in reduction of productive transcription from promoters at which the moribund complex is susceptible to these factors. We constructed a double-disruptant of *E.coli*, Δ*greA*Δ*greB*, and then arbitrarily selected 10 genes from among those whose levels of transcripts in the mutant strain were found to be lower than those in the parental *greA*$^+$*greB*$^+$ strain. Finally, the promoter for three of these genes, *atpC* (*uncC*), *cspA*, and *rpsA*, passed a further conventional test which confirmed that they displayed a branched initiation pathway in a reconstituted transcription system composed of purified components. The results obtained prove that the branched initiation pathway exists *in vivo* and is utilized in regulation of transcription initiation from some promoters, through modulation of the fraction of polymerase-promoter complexes entering each branch of the pathway.

In this year, we examined various physiological conditions that affect on the levels of the Gre factors. The determination of the level of GreB was observed to be constitutive. The level of GreA remained the same through the growth phase, and did not respond much to the richness of the culture media. However, it decreased into half in aerobic conditions, indicating that some genes are regulated by GreA. Therefore, the regulatory circuit responding to the levels of proteins involving GreA, namely the branched pathway mechanism, is working in cells.

(2) Applicability of thermodynamics to equilibria in biology

Nobuo Shimamoto[1] and Jyun-ichi Tomizawa[1]
([1]Structural Biology Center, National Institute of Genetics and Department of Genetics, School of Life Science, The Graduate University for Advanced Studies)

Most DNA-binding proteins are biologically functional as a specific complex, one containing a special short DNA segment. Such a complex is usually assumed as a state tenable for thermodynamic analysis of binding equilibrium. Thus, forward and backward reactions should balance at equilibrium in every pathway, and the affinity should be independent of the length of DNA. However, we have found that the balance at equilibrium is broken for some proteins by their sliding along DNA during association but not dissociation and that their affinities for their specific sites dependent on the length of DNA harboring the sites. This seeming disagreement is explained by an indeliberate use of the state of specific complex in

thermodynamics. In the presence of sliding, the state does not satisfy the second law (the ergodic condition) and thus is disqualified for thermodynamic analysis. A general treatment of binding equilibrium, while maintaining the specific complex as a distinct state, is proposed on the base of the master equation or chemical kinetics.

(3) Systematic search for promoters encoded in the genomic DNA sequences of E. coli

Nobuo Shimamoto[1], Hideki Nakayama[1] and Hironori Aromaki[2] ([1]Structural Biology Center, National Institute of Genetics and Department of Genetics, School of Life Science, The Graduate University for Advanced Studies. [2]Daiichi College of Pharmaceutical Sciences)

Irrespective of a pile of compiled sequences identified as promoters for E. coli vegetative RNA polymerase holoenzyme, σ^{70} holoenzyme, there is no successful methods to predict strength of a promoter. To construct such prediction method, we started to design a functional SELEX to select promoter sequences. The complexity of random oligo-DNA available in a lab scale, $4^{12\sim14}$ is too small to cover the all-possible promoters. Therefore we limited the candidate sequences to those involved in the genomic sequence. In order to construct and select a library, we transferred and amplified parts of the genomic sequence with PCR with a single primer. As a drawback of the use of single primer, the constructed library contains DNA fragments generated from in vitro recombination. A theoretical method to diminish the effect of such recombinants has been developed.

PUBLICATIONS

Papers

1. Sakata-Sogawa, K. and Shimamoto, N. (2004). RNA polymerase can track a DNA groove during promoter search. ***Proc. Nat. Sci. U.S.A.*** **101**, 14731-14735.

Reviews

2. 嶋本伸雄(2004)レーザートラップのナノ操作標準技術としての確立. 医学の歩み, 医歯薬出版.

3. 中山秀喜(2004)第2章プロテインチップの応用技術「抗体チップとDNAチップの比較」Inバイオチップの最新技術と応用(松永是監修)シーエムシー出版.

Books

4. 嶋本伸雄(2004)バイオテクノロジーとナノテクノロジーの融合, 早稲田大学産業技術創成研究所.

EDUCATION

Dean of School of Life Science, Graduate School of Advansed Studies

Lectures in other universities and institutes

1. Shimamoto, N. "RNA polymerase and a secret of nano-biomachines", Center For DNA Fingerprinting and Diagnostics (CDFD), Hyderabad, India January, 2004.
2. Shimamoto, N. "The firm start-point of nanobiology and biological nanotechnology: Understanding the characteristics of biological nanomachines" (in Japanese,「ナノバイオの揺るぎない出発点生体のナノマシンの特色」), Graduate School of Engeneering, The university of Tokyo, June 2004.
3. Shimamoto, N. "Chemical and Biological consequences of one-dimensional diffusion of proteins along DNA School of Chemistry", Seoul National University, Seoul, Korea, September 2004.
4. Shimamoto, N. "RNA polymerase and a secret of nano-biomachines", Department of Life Sciences, Seoul National University, Seoul, Korea, September 2004.
5. Shimamoto, N. "Proteins as biologically functional molecules in Fundamental Course of Molecular and Cellular Biology (KAST)" (in Japanese基礎から学ぶKAST分子細胞生物学コース：機能分子としてのタンパク質) The Institute of Medical Science, The University of Tokyo October 2004.
6. Shimamoto, N. "Acey-deucy biological nanotechnology: Truths and ghosts in nano-biological science and technology (in Japanese,「玉か石か？：ナノバイオを通して見る科学と技術の虚実」), The Osaka Cabinet of Commerce and Industry (大阪商工会議所), November 2004.

SOCIAL CONTRIBUTIONS AND OTHERS

Organization of International Meeting

1. Shimamoto, N and Shivashankar, G. V. (Organizers),
Nakayama, H and Susa, M. (Organizing Members)
JSPS-DST Asia Academic Seminar:
NIG-NCBS International Workshop on Single-molecule Biophysics, Bangalore, India January, 2004. (25 International Invitees from US, France, Swiss,

Israel, Korea, Japan, India and 30 Asian Trainees from India, Japan, China, Korea, Malaysia, Singapore)

2. Member of international Committee of Asian Conference of Transcription (ACT), the delegate of Japan. (Organizing members of ACT8, Bangkok, November, 2004)

特許

1. 嶋本伸雄・中山秀喜・荒牧弘範(発明者), 国立遺伝学研究所長(出願者), 2004年3月3日, ゲノムライブラリー作成方法, および同方法により作成されたゲノムライブラリー, 特願2004-59900, 機構整理番号, U2003P394

2. 嶋本伸雄・中山秀喜・荒牧弘範(発明者), 国立遺伝学研究所長が代表する日本国(出願者), 2004年3月16日, ゲノムライブラリー作成方法, および同方法により作成されたゲノムライブラリー, PCT/JP2004/003507, 機構整理番号, U2003P103

3. 嶋本伸雄・須佐太樹・福島和久(発明者), 横河電機株式会社(出願者), 2004年10月6日, 生体高分子検出方法およびバイオチップ並びに抗体固定法及び抗体固定基板, 米国およびドイツに出願

4. 嶋本伸雄・須佐太樹・福島和久(発明者), 横河電機株式会社(出願者), 2004年10月6日, 生体高分子検出方法およびバイオチップ並びに抗体固定法及び抗体固定基板, 中国特許出願番号200410090010.5

学会活動等

1. Member of international Committee of Asian Conference of Transcription(ACT), the delegate of Japan.(Organizing members of ACT8, Bangkok, November, 2004)

2. 嶋本伸雄, 財団法人未踏科学技術協会「生命を測る」組織幹事

各種委員

1. 嶋本伸雄, 文部科学省科学技術・学術審議会専門委員

2. 嶋本伸雄, NEDO「ナノバイオテクノロジー産業化推進調査」委員, WG

H-c. Multicellular Organization Laboratory Isao Katsura Group

RESEARCH ACTIVITIES

(1) Analysis of synthetic dauer-constitutive mutants in the nematode *Caenorhabditis elegans*

Tomoko Yabe, Kotaro Kimura, Takeshi Ishihara and Isao Katsura

If newly hatched larvae of *C. elegans* are in a crowded state and given a limited food supply, they sense the environmental signal of pheromone and food with a head sensory organ called amphid, and deviate from the normal life cycle, ending in non-feeding larvae called dauer larvae. Since the assay of dauer larva formation is much less time-consuming than behavioral assays such as chemotaxis assays, we are analyzing the pathway of sensory signals in the amphid neural circuit by detecting dauer larva formation as the output. We found that mutations in more than 50 known genes show synthetic dauer-constitutive (synDaf) phenotypes, i.e., they induce dauer larva formation in certain mutant backgrounds, regardless of the environmental conditions. The synthetic nature of the phenotype, we think, is based on the pathway of sensory signals. Namely, the signals are transmitted through parallel routes, and therefore two mutations are required to block them. We are determining the combinations of mutations for the synDaf phenotype and the pattern of suppression of the synDaf phenotype by various suppressor mutations. In this way we hope to elucidate the pathway of sensory signals for dauer regulation.

Furthermore, to identify new genes required for the sensory signal transduction, we have isolated and mapped 44 mutations that show the synDaf phenotype in the *unc-31(e169)* background, where *unc-31* gene encodes CAPS protein, which acts in secretion from dense core vesicles. Eight of the mutations map in 4 known genes, but most of the remaining 36 mutations, which map at least in 13 genes, seem to be located in novel genes, which we named *sdf* genes. Of these genes, *sdf-9*, *sdf-13*, and *sdf-14* have been cloned. *sdf-9* gene encoded a protein tyrosine phosphatase-like molecule, was expressed in a pair of neuron-associated cells called XXXL/R, and regulated dauer larva formation

in the steroid hormone signaling pathway (Ohkura, K. et al.: Development *130*, 3237-3248, 2003). *sdf-13* encoded a homologue of the transcription factors Tbx2 andTbx3, was expressed in AWB, AWC, and ASJ sensory neurons as well as many pharyngeal neurons, and controlled olfactory adaptation in AWC and dauer larva formation in cells other than AWC (possibly ASJ)[1]. *sdf-14* gene was the same as *mrp-1* gene, which was formerly identified by its homology to multidrug resistance-associated protein genes in mammals. A functional *sdf-14*::GFP fusion gene was expressed in many tissues including neurons, pharyngeal muscles and intestinal cells. Experiments with extrinsic cell-specific promoters revealed that expression in two, but not one, of the three tissues rescue the mutant phenotype efficiently. Interestingly, human MRP1 could substitute for *C. elegans* MRP-1 in dauer larva regulation, and an inhibitor of the MRP1 export activity impaired this function, showing that the export activity is required for normal dauer larva regulation.

In the year 2004, we found by epistatis analysis that *sdf-14* acts neither in the cGMP nor TGF-β signaling pathway among the four signaling pathways involved in dauer larva regulation. The result is consistent with that of the cell-specific expression experiments, because genes in these two pathways act only in neurons. We found that *sdf-14* mutations strongly enhance the dauer-constitutive phenotype of the *daf-2(e1370)* mutation, which confers resistance to many environmental stresses. Furthermore, sodium arsenite, which is a substrate of human MRP1, as well as high temperature (27°C) enhanced the dauer larva formation of the *unc-31(e169)* mutant, which is known to block the *daf-2* insulin signaling pathway. Thus, *sdf-14* gene may be involved in heat and heavy metal stress responses. The proof of this possibility as well as its relation to dauer larva formation remains to be studied.

(2) *C. elegans* mutants in the "associative" learning with odorants and food

Ichiro Torayama, Hiroshi Ichijo, Kotaro Kimura, Takeshi Ishihara and Isao Katsura

The nematode *C. elegans* provides a good system for the study of learning with a combination of two stimuli. However, the mechanism of this learning looks different from that of classical conditioning, because (a) the unconditioned stimulus is usually limited to food or starvation, and because (b) the learning is efficient, if the conditioned stimulus is presented at the same time as but not before the unconditioned stimulus. To elucidate the molecular mechanism of such "associative" learning in *C. elegans*, we are isolating and characterizing mutants that show abnormality in the learning with butanone and food/starvation. It is known that butanone attracts wild-type animals without conditioning. Conditioning with butanone and starvation decreased the efficiency of chemotaxis to butanone, while conditioning with butanone and food increased it. We isolated mutants in these behaviors, some of which showed decrease in the efficiency of chemotaxis after conditioning with food and butanone, while others are attracted efficiently by butanone only after conditioning with food and butanone. Of those mutants, *ut305* and *ut306*, which belong to the former category, have been studied in detail. *ut305*, but not *ut306*, showed abnormality in adaptation to isoamyl alcohol and benzaldehyde, which are sensed by the same type of sensory neurons (AWC) as butanone. *ut305* showed inefficient chemotaxis to butanone at low concentration, although it showed normal chemotaxis at the concentration used in the learning assay. *ut305* gene encoded a novel protein containing predicted transmembrane domains and showing limited homology to the *Drosophila* Raw protein.

In the year 2004, we examined the rescue of *ut305* mutant phenotypes by expressing the wild-type *ut305* gene in various cells, using extrinsic promoters. The mutant phenotypes were rescued, if the wild-type *ut305* gene was expressed in neurons including AWC sensory neurons, although we formerly detected the fluorescence of a functional *ut305*::GFP fusion gene in AIA interneurons and many pharyngeal neurons, but not in AWC neurons. Furthermore, *str-2*::GFP fusion gene, which is normally expressed in only one of the AWCL/R neurons, was not expressed in either of them in the *ut305* mutant, while it was expressed in both of them when the wild-type *ut305* gene was overexpressed in AWC neurons. These results indicated that *ut305* gene acts in AWC neurons for the learning with butanone and food, for the adaptation to benzaldehyde and isoamyl alcohol and for the regulation of *str-2* expression. We are now making specific antibodies to the *ut305* protein to examine whether the *ut305* protein is present in AWC neurons. We are also determining the position of *ut305* gene in the cascade of the

regulation of *str-2* expression, in which Ca^{2+} signaling and the ASK1 MAP kinase cascade are involved. We also cloned *ut306* gene and analyzed the mutant phenotypes. The study will be continued in 2005.

(3) Molecular genetic studies on sensory integration and behavioral plasticity in *C. elegans*

Takeshi Ishihara, Yuichi Iino[1], Akiko Mohri[2], Ikue Mori[2], Keiko Gengyo-Ando[3], Shohei Mitani[3] and Isao Katsura ([1]Molecular Genetics Laboratory, University of Tokyo, [2]Division of Biological Science, Nagoya University, [3]Department of Physiology, Tokyo Women's Medical University School of Medicine)

Animals receive environmental cues, select and integrate necessary information, and make proper responses, while all these steps can be modified by experience or memory. In *C. elegans*, many behavioral mutants defective in chemotaxis and thermotaxis, for instance, have been isolated and analyzed, and the molecular mechanisms of sensation have been elucidated. On this basis and as a next step, we are analyzing mutants that show abnormality in the learning and selection (evaluation) of sensory signals, to elucidate novel mechanisms of higher order sensory signal processing.

C. elegans shows avoidance of copper ion and chemotaxis to odorants by receiving these stimuli with different sensory neurons in the head. We developed a behavioral assay for the interaction of two sensory signals: aversive copper ion and attractive odorant, diacetyl. Wild-type animals change their preference. between the responses, depending on the relative concentration of copper ion and odorants. On the basis of the *C. elegans* neural circuitry, the result suggests that the two sensory signals interact with each other in a neural circuit consisting of about 10 pairs of neurons. While well-fed animals are usually used for this assay, we found that animals starved for 5 hours tend to prefer chemotaxis to odorants. The change is due to the desensitization of copper ion avoidance by starvation, and can be suppressed by serotonin, which mimics the effect of food. This desensitization is advantageous in natural environment, because starved animals can search for food over a wider area.

To elucidate the mechanism of sensory integration in the neuronal circuit, we are isolating and analyzing mutants that show abnormality in this assay. The *hen-1* mutants showed much weaker tendency to cross the Cu^{2+} barrier when migrating toward attractive odorants than the wild type, although that the *hen-1* mutants had defects neither in the chemotaxis toward the attractive odorant nor in the avoidance of Cu^{2+} ion per se.

To elucidate molecular mechanisms for the sensory integration, we cloned the *hen-1* gene and found that it encodes a secretory protein with an LDL receptor ligand binding domain, LDLa. This domain in HEN-1 is most similar to that domain of *Drosophila* signaling molecule Jeb, which regulates migration and differentiation of visceral mesodermal precursor cells. Immunostaining by using antibody against recombinant HEN-1 protein revealed that the gene product is localized in the axon and cell body of each one pair of sensory and inter-neurons. The localization in the axon was abolished in *unc-104* (kinesin KIF1A homologue) mutants, which show defects in the transport of synaptic vesicles. Expression studies with various promoters showed that this gene acts non-cell-autonomously in the mature nervous system.

The *hen-1* mutants also show abnormality in learning by paired presentation of starvation and NaCl (collaboration with Dr. Iino, University of Tokyo) and by paired presentation of starvation and temperature (collaboration with Ms Mohri and Dr. Mori, Nagoya University). Wild-type animals show chemotaxis to NaCl under a well-fed condition, although they avoid NaCl after conditioned with starvation and NaCl. The *hen-1* mutants show a weaker behavioral change than the wild type after the conditioning, although they show normal chemotaxis to NaCl under a well-fed condition. Wild-type animals prefer the cultivation temperature under a well-fed condition, while they avoid that temperature after conditioned in the absence of food at the same cultivation temperature. Although the *hen-1* mutants show normal thermotaxis under a well-fed condition, they do not avoid the cultivation temperature after conditioned in the absence of food. Since starvation was used to induce plasticity in both learning assays, we analyzed whether *hen-1* animals can sense starvation, but we could not find any abnormality in the behavior after simple starvation. These results indicate that the *hen-1* show defects in the behavioral plasticity after paired presentation of starvation and NaCl or starvation and temperature, although it responds normally to each of these stimuli.

Molecular genetic analyses of HEN-1 suggest that

HEN-1 functions as a neuronal modulator for sensory integration and learning. To elucidate the molecular mechanisms of this neuromodulation, we started investigating the protein interacting with the HEN-1 protein. First, we developed a binding assay for identification of receptors for HEN-1. By using a HEN-1-alkaline phosphatase fusion protein as a ligand, which was expressed by HEK293 cells, we found that HEN-1 specifically binds a subpopulation of the primary culture cells in *C. elegans*, suggesting that receptors for HEN-1 exist in these cells.

Recently, Jeb in *Drosophila* was reported to regulate development of mesodermal cells through the receptor tyrosine kinase DAlk, which is a homologue of human proto-oncogene Alk. Since the LDLa domain in Jeb is similar to that in HEN-1, we started to analyze the *scd-2* gene, which encodes a receptor tyrosine kinase similar to DAlk. Expression analyses by using an *scd-2* promoter-GFP construct suggested that SCD-2 is expressed in several sensory neurons and interneurons. *scd-2* mutants showed the same behavioral defects as *hen-1* mutants in paradigms for sensory integration and learning. The expression of SCD-2 in *scd-2* mutants driven by neuron specific promoters and a heatshock promoter suggested that SCD-2 functions in the mature nervous system, like HEN-1. These results suggest that SCD-2 may be a receptor for HEN-1. Determination of the cells where SCD-2 functions for sensory integration and learning may reveal the center of informational processing in *C. elegans*.

(4) Genetic analysis of plasticity of avoidance behaviors in *C. elegans*

Kotaro Kimura and Isao Katsura

C. elegans avoid many toxic and/or hazardous signals, such as high osmotic strength, acidic pH and SDS, as well as several volatile and soluble small compounds. In general, preceded stimulation(s) of animal's sensory neuron lead to a reduction in the magnitude of its response (adaptation), and *C. elegans* has been shown to exhibit adaptation to attractive odors or to body touch after pre-exposure to the stimulus. However, plasticity of the avoidance behaviors of *C. elegans* to the signals mentioned above is poorly understood. We wondered whether the avoidance behaviors of *C. elegans* are modulated by preceded experience of the same stimulus.

We found that pre-exposure to 2-nonanone, one of the repellent odors, enhances the avoidance behavior of the animals to the odor. For example, the avoidance behavior to 10% (v/v) 2-nonanone of the wild-type animals was significantly enhanced after pre-exposure. The magnitude of the enhanced avoidance behavior was comparable to the behavior to 30% 2-nonanone of previously unexposed animals, suggesting that pre-exposure may enhance the sensitivity to 2-nonanone by about 3-fold.

Enhancement after pre-exposure was also observed for 1-octanol, another repellent odor, but not for osmolarity as the repellent stimulus, suggesting that only certain type(s) of repellent stimuli trigger this phenomenon, and that it requires specific molecular and/or neuronal mechanisms.

We are interested in the sensitization because of the following reasons: (1) Not all types of the repellent stimuli trigger the sensitization. (2) Sensitization to repellent odors and tastes has not been reported in *C. elegans* to our knowledge. (3) The molecular mechanisms of sensitization to odor, taste or even light in other animals are poorly understood. We believe that genetic analysis of the sensitization may reveal novel molecular mechanism of regulation of signal sensation.

(5) Class 1 *flr* mutants of the nematode *Caenorhabditis elegans*

Yuri Kobayashi, Kotaro Kimura, Takeshi Ishihara and Isao Katsura

Class 1 *flr* mutants of *C. elegans*, which map in *flr-1*, *flr-3* and *flr-4*, were originally isolated by resistance to 0.4 mg/ml NaF (Katsura, I. *et al.*: Genetics *136*, 145-154, 1994). They also show many other phenotypes including slow growth, short defecation cycle periods, frequent skip of the expulsion step of defecation, synthetic abnormality in dauer larva formation, weak tendency to stay on food, and hypersensitivity to serotonin. The *flr-1* gene encodes an ion channel belonging to the DEG/ENaC (*C. elegans* degenerins and mammalian amiloride-sensitive epithelial sodium channels) superfamily, while *flr-4* and *flr-3* code for a novel Ser/Thr protein kinase and a kinase-like molecule, respectively, both having a hydrophobic domain on the carboxyl terminus. A functional *flr-1*::GFP fusion gene is expressed only in

the intestinal cells from the comma stage of embryos to the adult stage, while a functional *flr-4*::GFP is expressed in the intestinal cells from the 1.5-fold stage, in the isthmus of the pharynx from the 3-fold stage and in a pair of head neurons called AUA from L1 larvae to adults. Moreover, the expression of various *flr-3::lacZ* and *flr-3*::GFP fusion genes is detected only in the intestine. Temperature-shift experiments of a *flr-4(ts)* mutant showed that the activity of FLR-4 is required at the time of defecation assay (i.e., young adults) for normal defecation cycle periods. We therefore think that class 1 *flr* genes constitute a regulatory system that acts in the differentiated intestinal cells.

In 2004, we confirmed using an extrinsic promoter that expression of the wild-type *flr-4* gene in the intestine of *flr-4* mutants is sufficient for the wild-type phenotypes in growth, defecation cycle periods, expulsion, and dauer regulation[3]. This is consistent with our old experiments showing that killing of AUA neurons in wild-type animals and *flr-4* mutant animals does not change their defecation phenotypes. Thus, we could show that the three class 1 *flr* genes act in the intestinal cells to regulate various functions. We are studying how they interact with one another.

(6) Class 2 *flr* mutants of the nematode *Caenorhabditis elegans*

Akane Oishi, Kotaro Kimura, Takeshi Ishihara and Isao Katsura

Class 2 *flr* mutations were isolated mostly as suppressors of the slow growth or serotonin-hypersensitivity of class 1 *flr* mutations. Besides these phenotypes, they also suppress the dauer larva formation abnormality and weak tendency to stay on food, but not the defecation abnormalities or strong fluoride-resistance. By themselves, class 2 *flr* mutations show the phenotypes of weak resistance to NaF and short average longevity as compared with wild-type animals. The phenotypes suggest two possibilities on the relationship between class 1 and class 2 *flr* genes. (a) Class 2 *flr* genes may act downstream of the class 1 regulatory pathway. At the downstream, the regulatory pathway bifurcates into two branches, the growth/dauer branch and the defecation branch, while class 2 genes act in the former branch and not the latter. (b) Class 2 *flr* genes may act antagonistically to class 1 genes, while the threshold of the phenotypes is different between the growth/dauer phenotypes and the defecation phenotypes.

Class 2 mutations map in four genes, *flr-2, flr-5, flr-6* and *flr-7*, of which only *flr-2* has been cloned. *flr-2* encodes a secretory protein belonging to the gremlin/DAN/cerberus family. A functional *flr-2*::GFP fusion gene was expressed in some neurons in the head and the tail as well as many pharyngeal neurons. FLR-2::alkaline phosphatase fusion protein molecules bound specifically to the intracellular compartment of a limited number of *C. elegans* primary culture cells. We therefore screened for *C. elegans* cDNA clones whose expression in COS cells enabled the cells to bind to FLR-2::alkaline phosphatase. By screening 264 pools of cDNA, where one pool consists of about 1000 clones, we obtained a single positive clone (ZK20.1) encoding a secretory protein.

In 2004, we carried out cell-specific expression experiments, using extrinsic promoters. Although a functional *flr-2*::GFP fusion gene was expressed only in neurons, expression of the wild-type *flr-2* gene in the body wall muscle and intestine, respectively, rescued the *flr-2* mutant phenotype, as assayed by the recovery of the slow growth phenotype of *flr-2; flr-1* double mutants. Rescue with neuronal promoters was successful, only when we used low concentrations of the construct DNA for transformation. Thus, it seems that although FLR-2 is expressed and secreted by some neurons under natural conditions, body wall muscles and intestinal cells also have the ability to secrete it. Besides these experiments, we are working on the functional interaction between *flr-2* and ZK20.1, by isolating a deletion mutant in the ZK20.1 gene and looking for its phenotype.

PUBLICATIONS

Papers

1. Miyahara, K., Suzuki, N., Ishihara, T., Tsuchiya, E. and Katsura, I. (2004). TBX2/TBX3 transcriptional factor homologue controls olfactory adaptation in *Caenorhabditis elegans*. J. Neurobiol. *58*, 392-402.

2. Kimura, K.D., Miyawaki, A., Matsumoto, K. and Mori, I. (2004). The *C. elegans* thermosensory neuron AFD responds to warming. Curr. Biol. *14*, 1291-1295.

3. Take-uchi, M., Kobayashi, Y., Kimura, K., Ishihara, T. and Katsura, I. (2005). FLR-4, a novel Serine/Threonine protein kinase, regulates defecation rhythm in *Caenorhabditis elegans*. Mol. Biol. Cell, *16*, 1355-1365.

Reviews

4. 石原　健(2004)「線虫学習行動の分子遺伝学：連合学習制御遺伝子」蛋白質・核酸・酵素*49*, 450-455.

5. 桂　　勲(2005)「線虫*C. elegans*の発生・分化研究とゲノム情報」実験医学増刊*23 (1)*, 157-163.

EDUCATION

1. Dr. I. Katsura gave a lecture at Kyoto University, Graduate School of Science, June 2004 (in Japanese).

2. Dr. I. Katsura gave a lecture at Kyushu University Graduate School, Faculty of Sciences, June 2004 (in Japanese).

3. Dr. I. Katsura was invited to give a seminar on "Molecular biological analysis of the behavior of the nematode *C. elegans*" at Kyushu University Graduate School, Faculty of Sciences, June 2004 (in Japanese).

4. Dr. I. Katsura gave a lecture at the University of Tokyo, Graduate School of Arts and Sciences, July 2004 (in Japanese).

SOCIAL CONTRIBUTION AND OTHERS

1. 特願2004-314935「ヒトMRP1阻害剤のスクリーニング法」発明者：矢部智子，桂　　勲，石原　健，鈴木教郎，出願人：大学共同利用機関法人情報・システム研究機構.

2. Dr. I. Katsura served as one of the associate editors of the journal "Genes to Cells".

3. 毛利秀雄，勝見允行，木村武二，中西剋爾，守　隆夫，高橋正征，桂　　勲，他12名「高等学校　生物Ⅱ」三省堂2004.

H-d. Biomolecular Structure Laboratory Yasuo Shirakihara Group

RESEARCH ACTIVITIES

(1) Crystallographic Study of ATP synthase

Yasuo Shirakihara and Aya Shiratori

ATP synthase is responsible for ATP production in living cells, and is a membrane protein located in the energy conversion membrane. ATP synthase consists of a channel Fo portion (about 100,000 dalton, subunit composition of $ab_2c_{8\text{-}12}$) and a large soluble catalytic F1 portion (380,000 dalton, $\alpha_3\beta_3\gamma\delta\varepsilon$). The unique rotational catalysis mechanism of F1 includes rotation of the rod-like γ subunit which is thought to control the conformations of the three catalytic β-subunits in a cyclic manner by its rotation. Starting from elucidation of the $\alpha_3\beta_3$ sub-assembly structure of the thermophilic F1, we have been moving up to the higher sub-assembly. The $\alpha_3\beta_3\gamma$ sub-assembly was difficult to crystallize, but the $\alpha_3\beta_3\gamma\varepsilon$ sub-assembly gave crystals that allowed to see a novel conformation of F1. We are now dealing with the holo-enzyme, ATP synthase.

The membrane protein ATP synthase is still a challenging target for a structural study, in view of relatively few solved structures of the membrane proteins so far. After an initial trial for preparation and crystallization of the protein made last year, we have continued examination of both aspects extensively. Detergents such as decy-maltoside, undecyl-maltoside, dodecyl-maltoside, tridecyl-malltoside were good in preparation, but others like alkyl glucosides were not. Among various columns examined, useful ones were Q-sepharose high performance, Superdex prep grade 200 columns. An array of Q-sepharose high performance, Superdex prep grade 200 and the second Q-sepharose high performance columns yielded a preparation of crystallization grade, but a smaller array of Superdex prep grade 200 and Q-sepharose columns only was indistinguishable from the first array. Initial tiny crystals are now replaced by crystals with typical dimensions of 0.2mm, 0.2mm, 0.01mm after extensive and systematic crystallization condition search. Among the detergents mentioned above, only decy-maltoside and dodecyl-maltoside gave diffracting crystals, though all of the four detergents gave similar-looking crystals. ADP is the most favorite ligand, as others like AMPPNP, an ATP analogue, gave crystals that diffracted poorly. Other factors such as temperature in crystallization, kind and concentration of monovalent salts and those of divalent salts have been optimized. ATP synthase in our crystals contains all the 8 subunits, in contrast to yeast ATP synthase that lacks a and b subunits in crystals. Our crystals diffract to a resolution of about 7 A. Further refinement of the crystallization conditions to get better diffracting crystals is in progress.

The preparation and crystallization study was done in collaboration with Satoshi MURAKAMI at Institute of scientific and industrial research, Osaka University, Toshiharu SUZUKI and Masasuke YOSHIDA, at Research Laboratory of Resource

Utilization, Tokyo Institute of technology.

(2) Comprehensive Crystallographic Study of Transcription factors and Genome-partitioning Factors from *E.coli*

Yasuo Shirakihara and Aya Shiratori

In *E.coli*, more than 160 transcription factors control transcription of their target gene(s) by binding to both their specific DNA sequence and the transcription apparatus. Sixty-five such transcription factors had been purified in Ishihama laboratory. In Niki laboratory, a number of novel proteins, that are judged to play roles in the genome and plasmid partitioning, have been prepared. Setting these transcription and genome partitioning factors as targets for a structural study, we are doing comprehensive structure determination in the Protein 3000 project (the sub-field of 'transcription and translation'). In the previous first two years, we examined 65 proteins for their crystals but identified only three proteins that gave crystals of diffraction quality. This year, we have concentrated to refine the crystals obtained so far.

Mlc, an *E.coli* transcription factor, was found to diffract to about 3 A resolution. The pattern was recorded after reducing the aggregating tendency of the Mlc crystals. Phasing was hampered by either unavailability of appropriate heavy atom derivatives or inability of a Se-methionine substituted form of the protein to form crystals.

YmcB, a protein believed to be involved in the genome and plasmid partitioning, was found to form extremely good crystals, diffracting beyond 2 A even with the laboratory beam source, after a sustained crystallization condition search. We did some expression experiments on YmcB, as the delivered preparation ran out and the subsequent expression in our hands turned out to be difficult. By changing expression conditions from the previous one (BL21(pLys), 30° incubation, IPTG induction) to new ones (BL21, 25° or 20° incubation, no induction), we got two preparations: from 25° incubation, a preparation of slightly shorter fragment of the molecule: from 20° incubation, a preparation of the intact molecule. The initial delivered preparation was a mix of the two, and the species responsible for crystals was found to be the fragment. Collection of diffraction data and analysis of the data is in progress.

This work has been done in collaboration with Akira ISHIHAMA, Emi KANDA (Nippon Institute for Biological Science), Hironori NIKI, Rie INABA, Katsynori YATA, Yasushi OGATA (isotope center).

(3) Comprehensive structure-based functional studies on transcription factors

Hiroshi Itou and Yasuo Shirakihara

Genomic-DNA information on number of organisms is now available. Using the information, comprehensive structure analysis of transcription factors for their structure-based functional understanding is in progress. Our research targets are transcription factors (including putative ones) from hyperthermophilic archaeon *Pyrococcus horikoshii*, mesophilic bacteria *Corynebacterium glutamicum* and vertebrates *Homo sapience*. 162 genes were selected as target, and 146 of 162 were cloned, over-expressed (125 of 142) and purified (49 of 125). These purified proteins were tried crystallization for X-ray crystal structure analysis, and ten initial crystals were obtained. We had already succeeded in structure analysis of three of crystals, PH1161 and PH1932 from *P.horikoshii* and CGL2612 from *C.glutamicum*. PH1161 protein is a homologue of bacterial transcriptional activator TenA, and PH1932 and CGL2612 are homologue of transcriptional repressor protein ArsR and QacR, respectively. As a functional approach, target DNA sequences for these proteins were determined using SELEX (Systematic Evolution of Ligand by EXponential enrichment) method. This led to a tentative identification of the target genes for these transcription factors. This combination of tertiary structure analysis and binding DNA sequence determination by SELEX method is a powerful tool that provides us with key information on biological functions of transcription factors newly discovered by genomic analysis.

This work has been done in collaboration with Ui OKADA and Isao TANAKA at Division of Biological Sciences, Graduate School of Science, Hokkaido University.

(4) Crystallographic Study of Transcription factors from *Pseudomonas aeruginosa*

Yasuo Shirakihara and Aya Shiratori

Pseudomonas aeruginosa is the well-known

opportunistic bacterial pathogen, and a number of transcription factors responsible for the pathogenicity have been identified. Among those, we are doing structural study of PtxR, PtxS, PhzR, and PA3547.

Last year, we analyzed the MAD data from PhzR crystals and got an electron density map which showed clear helical densities in some part. However, inspired by a lesson from a failing expression experiment of a plant cell-death related protein SAG12 (molecular mass of about 30k) where SAG12 was not expressed but lactamase from a vector was instead expressed, C terminus amino acid sequence of PhzR (26K) expressed in our hands was examined. The sequence corresponded to that of C terminus residues of lactamase from Staphylococcus aureus, not to that of PhzR. There is a strong suspect that our PhzR crystals were not genuine. Diffraction data have been re-examined in this regard.

This work has been done in collaboration with Hironori ARAMAKI (Daiichi Pharmaceutical College).

(5) Crystallization of Kid, a chromosome mover

Yasuo Shirakihara and Aya Shiratori

Kid is involved in spindle formation and chromosome movements in mitosis/meiosis. Kid consists of three domains: an N-terminal kinesin-like motor domain (35-370), a C-terminal chromosome-binding domain (594-647) and a connecting domain (371-593) including a coiled-coil region. Our Previous trials to crystallize the N-terminal motor and the C-terminal chromosome-binding domains had failed, partly because both preparations contained a GST-tag that was attached for expression and easy preparation. We tried this year to crystallize a longer N-terminal fragment (1-462), and got small crystals from jeffamine or isopropanol. The crystals were formed only in the presence of MgADP. Further refinement of the crystallization conditions is in progress.

We also tried to express the full-length Kid using our favorite *E. coli* expression system described last year. B834(DE3), JM109(DE3) and Rosetta produced the protein, though each in slightly different conditions. Expression was checked by a Western blot analysis. Efforts are being made to get a pure preparation in 10mg scale.

The study was done in collaboration with Noriko TOKAI and Jun-ichiro INOUE at the Institute of medical science, the university of Tokyo.

PUBLICATIONS

Papers

1. Itou, H., Yao, M., Watanabe, N. and Tanaka, I. (2004). Structure analysis of PH1161 protein, a transcriptional activator TenA homologue from the hyperthermophilic archaeon *Pyrococcus horikoshii*. *Acta Cryst. D60*, 1094-1100.

H-e. Gene Network Laboratory
Emiko Suzuki Group

RESEARCH ACTIVITIES

(1) Molecular and fine morphological studies on the phototransduction in Drosophila

Emiko Suzuki

Phototransduction in Drosophila photoreceptor cells is a G-protein coupled phosphoinositide (PI)-signaling cascade. It is the most rapid process among G-protein coupled cascade known to date (receptor potential is generated within ~20 milliseconds after light stimulation), and the responsiveness is dynamically regulated. We are studying how such elaborated signaling cascade is accomplished, from the viewpoint of intracellular organization of the molecules involved in this system. Our recent studies have shown that the core components of this cascade, phospholipase C (NORPA), protein kinase C (INAC) and TRP channel protein, form a macromolecular complex "transducisome" in the photoreceptive microvilli, by binding to a scaffold protein, INAD (collaboration with Dr. Charles Zuker's group in U.C. San Diego). The rapidness of phototransduction is dependent on this topological coupling of signaling molecules. In addition to these core components, there are several regulatory proteins, and proteins involved in the supporting system of phototransduciton. This year, we studied some of these protein functions. As for the back-up system of PI metabolism, we have found that many of the enzymes for PI-regeneration are localized on subrhabdomeric cisternae (SRC), the specialized smooth endoplasmic reticulum adjacent to the photoreceptive microvilli.

Among these, we focused on the diacylglycerol kinase 2 (RDGA) that is essential not only for the regulation of the TRP channel activity but also for the structural maintenance of photoreceptors. We asked which domain of RDGA protein molecule is important for the intracellular targeting and/or its function. By the expression studies of modified RDGA proteins in photoreceptors, we found that the cysteine-rich region of RDGA molecule is essential for the protein localization and enzyme antivities. The western blotting analysis showed that the RDGA protein lacking the cysteine-rich region is unstable. This suggests that the proper targeting of RDGA protein is important for the protein stability.

Besides PI metabolism, rhodopsin metabolism is important for the regulation of phototransduction. This year we studied Sunglass (SUN) protein that appears to be involved in rhodopsin degradation. This project was done in collaboration with Dr. Craig Montell's Group at Johns Hopkins University School of Medicine. We found that SUN protein co-localizes with rhodopsin in the photoreceptor cell organelles involved in protein degradation, and found that *sun* mutant is resistant to the rhodopsin degradation induced by bright light[1].

(2) Fine morphological studies of the synaptic target recognition

Emiko Suzuki

The neuromuscular junction of Drosophila embryonic body wall musculature is one of the ideal models for studying development and function of synapses at the single cell level. Each hemisegment of an embryo/larva has only 30 muscle cells, innervated by about 40 motoneurons. The neuromuscular projection forms in the late embryos. During these stages, one can fillet dissect the embryos and observe or manipulate the development of neuromuscular networks under a light microscope. Our previous studies on this system using combination of single cell labeling and electron microscopy, have revealed the target-specific interaction of pre- and post-synaptic cells with filopodia. This year we studied this process by immuno-scanning electron microscopy we have developed recently. With this technique, we could observe the intimate interaction of clustered muscle filopodia (myopodia) and neuro-filopodia, that is not recognizable by fluorescence microscopy.

(3) Fine morphological analysis of the function of a Ser/Thr kinase, UNC51 in Drosophila nervous system

Hirofumi Toda*, Hiroaki Mochizuki*, Emiko Suzuki and Katsuo Furukubo-Tokunaga* (*University of Tsukuba)

UNC51 has been known as a Ser/Thr kinase necessary for the axonal formation in *C. elegans* and mice. We have identified Drosophila gene orthologue of *unc51* in an attempt to study the detailed molecular function of UNC51 in axonal formation by use of Drosophila molecular genetics. We localized the UNC51 protein to the vesicular structures in axons and synaptic terminals by immuno-electron microscopy. Further, we found that the *unc51* gene mutation affects axonal transport of vesicular cell organelles, by the immuofluorescence microscopy of synaptotagmin and electron microscopic analysis. Studies on genetic interactions showed that UNC51 regulates axonal elongation cooperatively with TRIO and Rac GTPases. These results suggest that UNC51 is involved in axonal development through interaction with the proteins that regulate axonal membrane trafficking.

PUBLICATIONS

Papers

1. Xu, H., Lee, S.-J., Suzuki, E., Dugan, K.D., Stoddard, A., Li, H.-S., Chodosh, L.A. and Montell, C. (2004). A lysosomal tetraspanin associated with retinal degeneration identified via a genome-wide screen. EMBO J. *23*, 811-822.

2. Kohyama-Koganeya, A., Sasamura, T., Oshima, E., Suzuki, E., Nishihara, S., Ueda, R. and Hirabayashi, Y. (2004). *Drosophila* glucosylceramide synthase. J. Biol. Chem. *279*, 35995-36002.

Review

3. 鈴木えみ子(2004)「標的認識分子から見たシナプス形成機構」細胞, *36*, 66-69.

I. CENTER FOR INFORMATION BIOLOGY AND DNA DATA BANK OF JAPAN

I-a. Laboratory for DNA Data Analysis Takashi Gojobori Group

RESEARCH ACTIVITIES

(1) Comparative analysis of gene expression for convergent evolution of camera eye between octopus and human

Atsushi Ogura, Kazuho Ikeo and Takashi Gojobori

Although the camera eye of the octopus is very similar to that of humans, phylogenetic and embryological analyses have suggested that their camera eyes have been acquired independently. It has been known as a typical example of convergent evolution. To study the molecular basis of convergent evolution of camera eyes, we conducted a comparative analysis of gene expression in octopus and human camera eyes. We sequenced 16,432 ESTs of the octopus eye, leading to 1052 nonredundant genes that have matches in the protein database. Comparing these 1052 genes with 13,303 already-known ESTs of the human eye, 729 (69.3%) genes were commonly expressed between the human and octopus eyes. On the contrary, when we compared octopus eye ESTs with human connective tissue ESTs, the expression similarity was quite low. To trace the evolutionary changes that are potentially responsible for camera eye formation, we also compared octopus-eye ESTs with the completed genome sequences of other organisms. We found that 1019 out of the 1052 genes had already existed at the common ancestor of bilateria, and 875 genes were conserved between humans and octopuses. It suggests that a larger number of conserved genes and their similar gene expression may be responsible for the convergent evolution of the camera eye.

(2) Molecular hierarchy in neurons differentiated from mouse ES cells containing a single human chromosome 21

Chi Chiu Wang, Mitsutaka Kadota, Ryuichi Nishigaki, Yasuhiro Kazuki, Yasuaki Shirayoshi, Michael Scott Rogers, Takashi Gojobori, Kazuho Ikeo and Mitsuo Oshimura

Defects in neurogenesis and neuronal differentiation in the fetal brain of Down syndrome (DS) patients lead to the apparent neuropathological abnormalities and contribute to the phenotypic characters of mental retardation, and premature development of Alzheimer's disease, those being the most common phenotype in DS. In order to understand the molecular mechanism underlying the cause of phenotypic abnormalities in the DS brain, we have utilized an *in vitro* model of TT2F mouse embryonic stem cells containing a single human chromosome 21 (hChr21) to study neuron development and neuronal differentiation by microarray containing 15K developmentally expressed cDNAs. Defective neuronal differentiation in the presence of extra hChr21 manifested primarily the post-transcriptional and translational modification, such as Mrpl10, SNAPC3, Srprb, and SF3a60 in the early neuronal stem cell stage, and Mrps18a, Eef1g, and Ubce8 in the late differentiated stage. Hierarchical clustering patterned specific expression of hChr21 gene dosage effects on neuron outgrowth, migration, and differentiation, such as Syngr2, Dncic2, Eif3sf, and Peg3.

(3) Strong positive selection and recombination drive the antigenic variation of the PilE protein of the human pathogen *Neisseria meningitidis*

Daniel T. Andrews and Takashi Gojobori

The PilE protein is the major component of the *Neisseria meningitidis* pilus, which is encoded by the pilE/pilS locus that includes an expressed gene and eight homologous silent fragments. The silent gene fragments have been shown to recombine through gene conversion with the expressed gene and thereby provide a means by which novel antigenic variants of the PilE protein can be generated. We have analyzed the evolutionary rate of the pilE gene using the nucleotide sequences of two complete pilE/pilS loci.

The very high rate of evolution displayed by the PilE protein appears driven by both recombination and positive selection. Within the semivariable region of the pilE and pilS genes, recombination appears to occur within multiple small sequence blocks that lie between conserved sequence elements. Within the hypervariable region, positive selection was identified from comparison of the silent and expressed genes. The unusual gene conversion mechanism that operates at the pilE/pilS locus is a strategy employed by *N. meningitidis* to enhance mutation of certain regions of the PilE protein. The silent copies of the gene effectively allow "parallelized" evolution of pilE, thus enabling the encoded protein to rapidly explore a large area of sequence space in an effort to find novel antigenic variants.

(4) Difference of organelles involved in membrane traffic

Hajime Ohyanagi and Takashi Gojobori

Contrary to prokaryotes, eukaryotes have a number of membranous organelles involved in membrane traffic in cells. Therefore, the gain of membrane systems could be one of the most epochal events in the evolution of eukaryotes. However, details about the evolution of a membrane system are still unclear. Each protein involved in a membrane system must have been evolved with its own characteristics, so that proteins localized in the same position in cells may show the same characteristics in the context of molecular evolution. With the aim of giving insight to the evolutionary studies of membrane systems, we are conducting the following analyses. First, we obtained subcellular localization data of yeast proteins that have already published (http://yeastgfp.ucsf.edu/). Second, in order to classify these yeast proteins into groups according to their evolutionary emergence times, we performed the BLAST reciprocal best hit analysis of yeast proteome against the proteomes of other species. From these analyses, we observed that protein sets in endoplasmic reticulum and Golgi apparatus, major components of the membrane traffic system, clearly showed evolutionary emergence times different from each other. Thus, it suggests that endoplasmic reticulum and Golgi apparatus appeared independently in the evolution of eukaryotes.

(5) Highly conserved upstream sequences for transcription factor genes and implications for the regulatory network

Hisakazu Iwama and Takashi Gojobori

Identifying evolutionarily conserved blocks in orthologous genomic sequences is an effective way to detect regulatory elements. In this study, with the aim of elucidating the architecture of the regulatory network, we systematically estimated the degree of conservation of the upstream sequences of 3,750 human-mouse orthologue pairs along 8-kb stretches. We found that the genes with high upstream conservation are predominantly transcription factor (TF) genes. In particular, developmental process-related TF genes showed significantly higher conservation of the upstream sequences than other TF genes. Such extreme upstream conservation of the developmental process-related TF genes suggests that the regulatory networks involved with developmental processes have been evolutionarily well conserved in both human and mouse lineages.

(6) Computational prediction of microRNA (miRNA) target gene in human and its experimental varidation

Hitomi Sakurai, Roberto Barrero, Shiho Hayakawa, Takuro Tamura, Yoshio Tateno, Ikeo Kazuho, Imanishi Tadashi and Takashi Gojobori

MicroRNAs (miRNAs) form a novel class of small RNA genes of 21-25 nucleotides derived from highly conserved hairpin precursors (pre-miRNAs) present from mollusc to mammals. miRNAs act as post-transcriptional repressors of target transcripts via mRNA degradation or translation inhibition. We present a novel large scale RNA-affinity screening tool combined with the evaluation of RNA secondary structures to identify putative microRNA recognition elements (MREs) on target mRNAs. We identified 1,476 MREs for 115 known mammalian miRNAs. Eleven of the top predicted MREs for *Let-7a, miR-20, miR-97* and *miR-182a* were analyzed using a reporter assay. As a result, we found that seven MREs mediated inhibition of reporter expression. It suggests that our target prediction algorithm, RNAFFY, identifies highly reliable miRNA targets.

(7) Detection of apoptosis during planarian regeneration by the expression of apoptosis-related genes and TUNEL assay

Jung Shan Hwang, Chiyoko Kobayashi, Kiyokazu Agata, Kazuho Ikeo and Takashi Gojobori

Apoptosis is a tightly organized cell death process that plays a crucial role in metazoan development, but it has not yet been revealed whether apoptotic events are involved in the process of regeneration. Here, we tried to detect apoptotic cells during planarian regeneration using the TdT-mediated dUTP nick-end labeling (TUNEL) assay as well as the expression of apoptosis-related genes. Three novel cDNAs were isolated from a planarian cDNA library and shown to be closely related to other metazoan caspases at the amino acid sequence level. One of these cDNAs, Caspase-like gene 3 (DjClg3), was expressed primarily in apoptotic cells by double detections with the TUNEL assay. Whole mount *in situ* studies indicated that DjClg3 was expressed in the cells of the mesenchymal space and also around the pharynx of the intact body. Its expression in the regenerating head piece was seen in the blastema and less significantly in the brain, while in the regenerating tail piece, DjClg3 expression was detected uniformly throughout the entire region. In parallel experiments, we performed *in situ* TUNEL assays to localize the regions where cell death occurred during regeneration and comparable results to the DjClg3 expression patterns were obtained. This is the first report to show that planarians have apoptosis-related genes, and the results suggest that the apoptotic mechanism probably takes place to a large extent in normal intact worms as well as during their regeneration. We hypothesize that the presence of apoptosis in planarians may have a role in controlling cell numbers, eliminating unnecessary tissues or cells and remodeling the old tissues of regenerating body parts.

(8) Analysis of sexual dimorphism of gene expression in mouse brain

Kazuya Yuge, Kazuho Ikeo and Takashi Gojobori

There are sexual differences in morphological features of the mammalian brain. The so-called sexual dimorphism of mammalian brains is thought to be determined by the gonadal hormones; this view is called "gonadal origin theory". However, recent studies suggest that differences in gene expression of sex-liked genes in neuronal cells directly generate sexual dimorphism in the brain. To examine this possibility we conducted microarray experiments to screen the genes that were expressed differentially in the brain between male and female mice before gonadal hormone secretion. In this study we have identified 57 female enhanced genes and 24 male enhanced genes in brains at embryonic day 10.5 (E 10.5). These results indicate that sexual differences in gene expression in neuronal cells before gonadal hormone secretion would play an important role in sexual dimorphism in the brain.

(9) Seeking for signes of aging in Hydra, a primitive metazoan

Kengo Yoshida, Hwang Jung Shang, Chimi Fujisawa, Toshitaka Fujisawa, Kazuho Ikeo and Takashi Gojobori

Aging occurs in a variety of organisms including yeast, nematode, fly, and mammals. However, hydra belonging to phylum Cnidaria has been considered to be immortal (Martinez 1998). This apparent immortality was observed when hydra were maintained so that they underwent asexual reproduction, budding. Do they remain immortal if they undergo sexual reproduction? A previous study using *Hydra Oligactis* showed that the depression occurred after sexual reproduction (Littlefield 1985). To seek for possible signs of aging in hydra, we studied the process of the depression in detail. As a result, reduction in the capacity of food capture, digestive movements and contractile movements was all observed. Moreover, exponential increase in mortality rate of population was also observed. From these observations, we conclude that aging-like phenomenon occurs in hydra. Aging research on hydra, one of the most primitive metazoans, has potential for providing further insight into the understanding of evolution of aging.

(10) A large variation in the rates of synonymous substitution for RNA viruses and its relationship to a diversity of viral infection and transmission modes

Kousuke Hanada, Yoshiyuki Suzuki and Takashi

Gojobori

RNA viruses successfully adapt to various environments by repeatedly producing new mutants, often through generating a number of nucleotide substitutions. To estimate the degree of variation in mutation rates of RNA viruses and to understand the source of such variation, we studied the synonymous substitution rate because synonymous substitution is exempt from functional constraints at the protein level, and its rate reflects the mutation rate to a great extent. We estimated the synonymous substitution rates for a total of 49 different species of RNA viruses, and we found that the rates had tremendous variation by 5 orders of magnitude (from 1.3×10^{-7} to 6.2×10^{-2} / synonymous site/year). Comparing the synonymous substitution rates with the replication frequencies and replication error rates for the RNA viruses, we found that the main source of the rate variation was differences in the replication frequency because the rates of replication error were roughly constant over different RNA viruses. Moreover, we examined a relationship between viral life strategies and synonymous substitution rates to understand which viral life strategies affect replication frequencies. The results show that the variation of synonymous substitution rates has been influenced most by either the difference in the infection modes or the differences in the transmission modes. In conclusion, the variation of mutation rates for RNA viruses is caused by different replication frequencies, which are affected strongly by the infection and transmission modes.

(11) Evolutionary processes of gene splicing and gene silencing

Lihua Jin, Yoshiyuki Suzuki, Kazuho Ikeo and Takashi Gojobori

Both gene splicing and gene silencing have become known to contribute significantly to functional diversification of genes and gene networks. For gene splicing, we conducted an evolutionary study of the relationship between gene duplication and alternative splicing. For gene silencing, we studied the evolutionary process of small RNA-guided pathways, focusing on the Rnase III family enzymes. For the first topic, we mentioned that gene duplication and alternative splicing (AS) were the two major evolutionary mechanisms that could bring the functional variation through the diversification of genes and gene products. The purpose of this research is to understand the evolutionary relationship between the two different mechanisms, utilizing the available data resources. The results of this study showed that the proportion of the AS loci in the singleton gene group was less than that in the duplicated gene group. Moreover, we found that the duplicated genes tended to have more AS isoforms than singleton genes. These results suggest that gene duplication would induce more alternative splicing events on duplicated copies than on singletons possibly by reducing the functional constraints on the duplicates. For the second topic, we pointed out that the RNA-mediated gene silencing pathways were evolutionarily conserved processes. It highlights a fundamental role of short RNAs in eukatyotic gene regulation and antiviral defense. Recently, three distinct small RNA-directed silencing pathways were observed, such as the destruction of mRNA *via* siRNA, inhibition of mRNA translation *via* miRNA, and epigenetic gene silencing *via* siRNA. It was also found that in these pathways, the members of ribonclease III family played important roles in diverse RNA maturation and decay. Here we investigated the evolution of Rnase III nucleases, Dicer as representative, to further figure out the evolutionary relationship among the three gene silencing pathways. With the advantage of using genomic sequences as the subjects of homology search in the annotated and un-annotated genomic regions, we were able to detect possible candidates for a gene of dicer and its two functional domains. Moreover, we found that representative prokaryotes including eubacteria and archbacteria lacked completely the PAZ domain of Dicer. These results clearly show the taxonomy-dependent evolution of the RNA-mediated gene silencing pathways. The results obtained in this study provide the information for the understanding of the evolutionary origin and relationships of the three pathways. The information may also be useful for the conducting of the relevant experiments. As a summary, these results suggested that changes in the genome and regulatory network were closely related evolutionary events, implying that gene multiplication and functional constraints were the two important sources for the change in gene function in the networks of transcription and gene regulation in evolution.

(12) The bioinformatics challenges in comparative analysis of cereal genomes-an overview

Matthew Bellgard, Jia Ye, Takashi Gojobori and Rudi Appels

Comparative genomic analysis is the cornerstone of *in silico*-based approaches to understanding biological systems and processes across cereal species, such as rice, wheat and barley, in order to identify genes of agronomic interest. The size of the genomic repositories is nearly doubling every year, and this has significant implications on the way bioinformatics analyses are carried out. The concepts and technology underpinning bioinformatics as applied to comparative genomic analysis are discusse as the overview.

(13) Various adaptations for the perpetual darkness in the diversification process from an eyed surface-dwelling form to an eyeless cave-dwelling form of Mexican tetra, *Astyanax mexicanus*, by cDNA microarrays

Nobuhiko Tanaka, Shozo Yokoyama, Kazuho Ikeo and Takashi Gojobori

Mexican tetra (*Astyanax mexicanus*), a single species with eyed surface and eyeless cave dwelling forms, gives a unique opportunity of the evolutionary process of gene diversification during a short period of time. In this study, we tried to find genes related to the evolution of eyeless cave form from the eyed surface form by cDNA microarrays, each of which has 3,070 non-redundant clones from the entire body of an eyed adult fish. Target genes for the microarrays were made from the entire body of each an eyed fish and an eyeless fish. On the basis of the results of comparative microarray analysis for 716 functionally known genes, relatively highly expressed genes in the surface fish and in the cavefish were listed up, respectively. In the surface fish, we for at that there were relatively highly expression of eye formation-related genes, digestive system-related genes, energy metabolism-related genes, lipocalin-type prostaglandin D synthase gene, elastase genes. The eye formation-related genes may be highly expressed because surface fish have well-developed eyes whereas cavefish have degenerated eyes. The digestive system-related and energy metabolism-related genes may be also highly expressed because the surface fish are more active than the cavefish. The lipocalin-type prostaglandin D synthase protein is related to the regulation of circadian sleep-wake cycles. The highly expression of the protein gene corresponds to the regressive circadian activity in the cavefish. The elastase digests elastin fibers in the skin, when the ultraviolet (UV) light reaches the dermis in the skin. It is suggested that the elastase genes is highly expressed because the surface fish has been exposed to sufficiently sunny conditions compared with the cavefish. On the other hand, in the cavefish, relatively highly expression of transferrin genes, lipid metabolism-related genes, sex differentiation-related genes were observed. Transferrin proteins are related to the oxygen transportation in the blood. To effectively transport oxygen in the blood of the cavefish living in the water with the lower concentration of oxygen at the dark cave environments, transferrin genes may be required to have relatively highly expression in the cavefish. The highly expression of lipid metabolism-related genes would be due to efficient acquirement of the energy under circumstances of at least temporary shortness in food. The highly expressions of the sex differentiation-related genes (diazepam binding inhibitor gene and granulin 1 gene) in the cavefish would be related to negative sex differentiation because of the lack of any kind of display/courtship in the cavefish. In conclusion, various adaptations for the perpetual darkness were evolutionarily raised in the diversification process from the surface fish to the cavefish. This is the first report that differences of gene expression between the surface fish and the cavefish are extensively examined.

(14) Evolutionary analysis of transcriptional coactivator MBF1

Qing-Xin Liu, Kazuho Ikeo, Susumu Hirose and Takashi Gojobori

Transcriptional coactivators play a crucial role in gene expression by communicating between regulatory factors and the basal transcription machinery. How a coactivator evolves was poorly understood. We have taken a phylogenetic approach to analyze an evolutionarily history of coactivator MBF1 and TATA element-binding protein TBP. We found the following points. (1) MBF1 and TBP are evolutionarily conserved from *Archaea* to human to achieve transcription

initiation. (2) Archaeal MBF1 carries a DNA-binding domain consisting of a Zn-ribbon which is absent from its eukaryotic counterparts. Thus, Eukaryotic MBF1 can interact with various regulatory factors while archaeal MBF1 has a single partner. (3) The interaction between MBF1 and TBP is also conserved from *Archaea* to human. In *Archaea*, MBF1 binds to TBP through Lysine or Asparagine to Glutamic acid interaction. In eukaryotes, MBF1 binds to TBP through Asparatic acid or Glutamic acid to Glutamine interaction. (4) The phylogenetic tree of MBF1 is very similar to that of TBP, indicating that MBF1 coevolved with TBP.

(15) Pigment cell lineage-specific expression activity of the ascidian tyrosinase-related gene

Reiko Toyoda, Akiko Kasai, Shigeru Sato, Shuichi Wada, Hidetoshi Saiga, Kazuho Ikeo, Takashi Gojobori, Takaharu Numakunai and Hiroaki Yamamoto

Solitary ascidian tadpole larvae develop two types of black pigment cells in the major sensory organs of the brain. Such pigment cells have been demonstrated to express the melanogenic genes, tyrosinase and Tyrp/TRP (tyrosinase-related protein). To understand the genetic and developmental mechanisms underlying the differentiation of chordate pigment cells, we examined the function of the promoter region of Tyrp/TRP gene, an ascidian (*Halocynthia roretzi*) tyrosinase family gene. The expression of the gene in pigment cell lineage starts at the early-mid gastrula stages. To identify the transcriptional regulatory region of the gene allowing cell-type-specific expression, a deletion series of the HrTyrp 5' flanking region fused to a lacZ reporter gene was constructed and microinjected into ascidian fertilized eggs. The region of 73 bp in HrTyrp was identified as sufficient for expression in pigment cell-precursors of tailbud stage embryos. It is noteworthy that there is no M-box element highly conserved in the promoters for vertebrate tyrosinase family genes such as tyrosinase, Tyrp1/TRP-1 and Tyrp2/TRP-2 (Dct). Although the regulatory system of ascidian pigment-cell development is likely to contain most factors critical to vertebrate pigment-cell development, there might be critical differences in the mode of regulation, such as the developmental timing of interactions of factors, proteins and genes, involved in pigment cell differentiation and pigmentation.

(16) DDBJ in the stream of various biological data

Satoru Miyazaki, Hideaki Sugawara, Kazuho Ikeo, Takashi Gojobori and Yoshio Tateno

In the past year we at DDBJ (http://www.ddbj.nig.ac.jp) have made a steady increase in the number of data submissions with a 50.6% increment in the number of bases or 46.5% increment in the number of entries. Among them the genome data of man, ascidian and rice hold the top three. Our activity has extended to providing a tool that enables sequence retrieval using regular expressions, and to launching our SOAP server and web services to facilitate the acquisition of proper data and tools from a huge number of biological data resources on websites worldwide. We have also opened our public gene expression database, CIBEX.

(17) Note on the maximum likelihood estimation of haplotype frequencies

Shuhei Mano, Norikazu Yasuda, Toru Katoh, Kenichi Tounai, Hidetoshi Inoko, Tadashi Imanishi, Gen Tamiya and Takashi Gojobori

The maximum likelihood estimation (MLE) is one of the most popular ways to estimate haplotype frequencies of a population with genotype data whose linkage phases are unknown. The MLE is commonly implemented in the use of the Expectation-Maximization (EM) algorithm. It is known that the EM algorithm carries the risk that an estimator may converge erroneously to one of the local maxima or saddle points of the likelihood surface, resulting in serious errors in the MLE of haplotype frequencies. In this note, by theoretical treatments we present the necessary and sufficient conditions that the local maxima or saddle points on the likelihood surface appear. As a rule of thumb, that the difference between the coupling and repulsive haplotype frequencies in phase known individuals is 3/2 times larger than the frequency of phase ambiguous individuals is the sufficient condition that the likelihood surface is unimodal. Moreover, we present the analytic solution to the biallelic two-locus problem, and construct a general algorithm to obtain the global maximum.

(18) The evolutionary rate of a protein influenced by features of the interacting partners

Takashi Makino and Takashi Gojobori

We focused upon how the evolutionary rates of proteins were influenced by the characteristic features of PPIs. Because the recent advancement of molecular technologies enables us to understand actual features of protein to protein interactions (PPIs), it becomes possible to make objective descriptions about the characteristic features of the proteins in the PPI networks. In this analysis, we defined a protein having a larger number of PPI partners of the same functional class as the SF (Same Function) protein, and a protein having a larger number of PPI partners of different functional classes as the DF (Different Function) protein. We also classified proteins in the PPI networks into respective proteins in dense and sparse parts of the PPI network, denoting these proteins as the DP (Dense part) and SP (Sparse Part) proteins, respectively. Because these two classifications were independent of each other, we classified the proteins in PPIs further into the four categories, the SF-DP, SF-SP, DF-DP and DF-SP proteins. Then, we compared the evolutionary rates between the SF and DF proteins, between the DP and SP proteins, and among the four categories. As a result, we found that the DF proteins evolved at a slower rate than the SF proteins. We also found that the SP proteins evolved at a slower rate than the DP proteins. In particular, we pointed out that the DF-SP proteins evolved at the slowest rate in the proteins examined. Because all these differences in the evolutionary rates are statistically significant, it is suggested that the proteins with their PPI partners belonging to different functional classes and occupying a sparse part of the PPI network are under strong functional constraints. It follows that those proteins are very important for the maintenance and survival of the PPI network.

(19) Integrative annotation of 21,037 human genes validated by full-length cDNA clones

Tadashi Imanishi, (154 authors), Takashi Gojobori and Sumio Sugano

The human genome sequence defines our inherent biological potential; the realization of the biology encoded therein requires knowledge of the function of each gene. Currently, our knowledge in this area is still limited. Several lines of investigation have been used to elucidate the structure and function of the genes in the human genome. Even so, gene prediction remains a difficult task, as the varieties of transcripts of a gene may vary to a great extent. We thus performed an exhaustive integrative characterization of 41,118 full-length cDNAs that capture the gene transcripts as complete functional cassettes, providing an unequivocal report of structural and functional diversity at the gene level. Our international collaboration has validated 21,037 human gene candidates by analysis of high-quality full-length cDNA clones through curation using unified criteria. This led to the identification of 5,155 new gene candidates. It also manifested the most reliable way to control the quality of the cDNA clones. We have developed a human gene database, called the H-Invitational Database (H-InvDB; http://www.h-invitational.jp/). It provides the following: integrative annotation of human genes, description of gene structures, details of novel alternative splicing isoforms, non-protein-coding RNAs, functional domains, subcellular localizations, metabolic pathways, predictions of protein three-dimensional structure, mapping of known single nucleotide polymorphisms (SNPs), identification of polymorphic microsatellite repeats within human genes, and comparative results with mouse full-length cDNAs. The H-InvDB analysis has shown that up to 4% of the human genome sequence (National Center for Biotechnology Information build 34 assembly) may contain misassembled or missing regions. We found that 6.5% of the human gene candidates (1,377 loci) did not have a good protein-coding open reading frame, of which 296 loci are strong candidates for non-protein-coding RNA genes. In addition, among 72,027 uniquely mapped SNPs and insertions/deletions localized within human genes, 13,215 nonsynonymous SNPs, 315 nonsense SNPs, and 452 indels occurred in coding regions. Together with 25 polymorphic microsatellite repeats present in coding regions, they may alter protein structure, causing phenotypic effects or resulting in disease. The H-InvDB platform represents a substantial contribution to resources needed for the exploration of human biology and pathology.

(20) Japanese domesticated chickens derived from Shamo traditional fighting cocks

Tomoyoshi Komiyama, Kazuho Ikeo, Yoshio Tateno and Takashi Gojobori

With the aim of elucidating the evolutionary origin of Japanese domesticated chickens, we examined 85 chicken mtDNA sequences. Thirty-four various ornamental chickens, 42 fighting cocks (*Shamo*), and nine long-crowing chickens (*Naganakidori*) were included in these samples. Of the *Shamo*, 18 were sampled from Okinawa, while the remaining 24 were collected in other islands around Japan. In addition, three Southeast Asian Junglefowls were used as a reference to determine the common ancestor of from Okinawa that clearly diverged from the other Japanese domesticated chickens studied. We found that all Japanese domesticated chickens, including the ornamental varieties and *Naganakidori*, were derived from the ancestors of the *Shamo* in Okinawa. To create novel varieties of ornamental chickens, intensive artificial selection is imposed on ancestral *Shamo* population, resulting in profoundly differentiation of Japanese domesticated chickens.

(21) The evolutionary origin of long-crowing chicken: its evolutionary relationship with fighting cocks disclosed by the mtDNA sequence analysis

Tomoyoshi Komiyama, Kazuho Ikeo, Yoshio Tateno and Takashi Gojobori

Chickens with exceptionally long crow are often favored all over the world, and connoisseur breeders have bred certain types of chicken exclusively for this trait. In Japan, three chicken varieties have been specifically bred to develop an exceptionally long crow of over 15 seconds. Although these three long-crowing chickens, Naganakidori, are honored as heritage varieties of Japan, the domestication process and genealogical origin of long-crowing chickens remain unclear. The purpose of this study is to clarify these issues using nucleotide sequences of the mitochondrial DNA D-loop region. Blood samples from a total of nine long-crowing chickens and 74 chickens from 11 Japanese native varieties were collected. DNA sequence data of two Junglefowl species were also collected from the International DNA database (DDBJ /EMBL/ GenBank) for use as the outgroup. A phylogenetic tree was then constructed revealing that all three Naganakidori varieties were monophyletic and originated from a fighting cock, a Shamo, for cockfighting. These results suggest that these three long-crowing chickens share a common origin in spite of their conspicuously different characters, and that human cultures favoring long-crowing chickens might have been preceded by a tradition of cockfighting. Moreover, these long-crowing varieties first separated from the fighting cocks of Okinawa, which is geographically closer to Southern China and Indochina than Mainland Japan (Honshu/Kyushu). This implies that Japanese long-crowing chickens were first brought to Mainland Japan as fighting cocks from the surrounding regions of Southern China or Indochina and through Okinawa.

(22) Novel algorithm for automated genotyping of microsatellites

Toshiko Matsumoto, Wataru Yukawa, Yasuyuki Nozaki, Ryo Nakashige, Minori Shinya, Satoshi Makino, Masaru Yagura, Tomoki Ikuta, Tadashi Imanishi, Hidetoshi Inoko, Gen Tamiya and Takashi Gojobori

Microsatellites or short tandem repeats (STRs) are abundant in the human genome with easily assayed polymorphisms, providing powerful genetic tools for mapping both Mendelian and complex traits. Microsatellite genotyping requires detection of the products of polymerase chain reaction (PCR) amplification by electrophoresis, and analysis of the peak data for discrimination of the true allele. A high-throughput genotyping approach requires computer-based automation at both the detection and analysis phases. In order to achieve this, complicated peak patterns from individual alleles must be interpreted in order to assign alleles. Previous methods consider limited types of noise peaks and cannot provide enough accuracy. By pattern recognition of various types of noise peaks, such as stutter peaks and additional peaks, we have achieved an overall average accuracy of 94% for allele calling in our actual data. Our algorithm is crucial for a high-throughput genotyping system for microsatellite markers by reducing manual editing and human errors.

(23) Evolution of vitamin b6 (pyridoxine) metabolism by gain and loss of genes

Tsuyoshi Tanaka, Yoshio Tateno and Takashi Gojobori

Vitamin B(6) (VB6) functions as a cofactor of many diverse enzymes in amino acid metabolism. Three metabolic pathways for pyridoxal 5'-phosphate (PLP; the active form of VB6) are known: the *de novo* pathway, the salvage pathway, and the fungal type pathway. Most unicellular organisms and plants biosynthesize VB6 using one or two of these three biosynthetic pathways. However, animals such as insects and mammals do not possess any of the pathways and, thus, need to intake VB6 in their diet to survive. It is conceivable that breakdowns of these pathways occurred in the evolutionary lineages of insects and mammals, and one of the major reasons for this would be the loss of pertinent genes. We studied the evolution of VB6 biosynthesis from the view of the gain and loss of 10 pertinent genes in 122 species whose genome sequences were completely determined. The results revealed that each gene in the pathways was lost more than once in the entire evolutionary lineages of the 122 species. We also found the following three points regarding the evolution of PLP biosynthesis: (1) the breakdown of the PLP biosynthetic pathways occurred independently at least three times in animal lineages, (2) the de novo pathway was formed by the generation of pdxB in gamma-proteobacteria, and (3) the order of the gene loss in VB6 metabolism was conserved among different evolutionary lineages. These results suggest that the evolution of VB6 metabolism was subject to gains and frequent losses of related genes in the 122 species examined. This dynamic nature of the evolutionary changes must have been responsible for the breakdowns of the pathways, resulting in profound differentiation of heterotrophy among the species.

(24) Biased biological functions of horizontally transferred genes in prokaryotic genomes

Yoji Nakamura, Takeshi Itoh, Hideo Matsuda and Takashi Gojobori

Horizontal gene transfer is one of the main mechanisms contributing to microbial genome diversification. To clarify the overall picture of interspecific gene flow among prokaryotes, we developed a new method for detecting horizontally transferred genes and their possible donors by Bayesian inference with training models for nucleotide composition. Our method gives the average posterior probability (horizontal transfer index) for each gene sequence, with a low horizontal transfer index indicating recent horizontal transfer. We found that 14% of open reading frames in 116 prokaryotic complete genomes were subjected to recent horizontal transfer. Based on this data set, we quantitatively determined that the biological functions of horizontally transferred genes, except mobile element genes, are biased to three categories: cell surface, DNA binding and pathogenicity-related functions. Thus, the transferability of genes seems to depend heavily on their functions.

(25) False positive selection identified by ML-based methods: examples from the *Sig1* gene of the diatom *Thalassiosira weissflogii* and the *tax* gene of a human T-cell lymphotropic virus

Yoshiyuki Suzuki and Masatoshi Nei

Sexually induced gene 1 (*Sig1*) in the centric diatom *Thalassiosira weissflogii* is considered to encode a gamete recognition protein. Sorhannus (2003) analyzed nucleotide sequences of *Sig1* using parsimony analysis and the maximum-likelihood (ML)-based Bayesian method for inferring positive selection at single amino acid sites and reported that positively selected sites were detected by the latter method but not by the former. He then concluded that for this type of study the ML-based method is more reliable than parsimony analysis. Here we show that his results apparently represent false-positive cases of the ML-based method and that there is no solid evidence that this gene contains positively selected sites. We further demonstrate that in the *tax* gene of human T-cell lymphotropic virus type I (HTLV-I) all codon sites, including invariable sites, can be inferred as positively selected sites by the ML-based method. These observations indicate that the ML-based method may produce many false-positive sites. One of the main reasons for the occurrence of false positives is that in the ML-based method codon sites are grouped into several categories with different nonsynonymous/

synonymous rate ratios (ω's) on a purely statistical basis and positive selection is inferred indirectly by examining whether the average ω for each category is greater than 1 or not. In parsimony analysis, however, the evolutionary change of nucleotides at each codon site is examined. For this reason, parsimony-based methods rarely produce false positives and are safer than ML-based methods for detecting positive selection at individual codon sites, though a large number of sequences are necessary.

(26) Negative selection on neutralization epitopes of poliovirus surface proteins: implications for prediction of candidate epitopes for immunization

Yoshiyuki Suzuki

For development of effective vaccines against viruses, it is of importance to choose appropriate epitopes as the target for immunization. These epitopes should eventually be determined experimentally, but it would be helpful if we could predict candidate epitopes computationally because it accelerates the entire process. To predict candidate epitopes for immunization, it is of great interest to characterize the target epitopes of poliovirus vaccine, which has empirically proven to be the most effective among all vaccines available. Here I show that almost all amino acid sites of poliovirus surface proteins VP1, VP2, and VP3 including neutralization epitopes are negatively selected and no site is under positive selection. These results, together with those obtained in previous studies, indicate that vaccines directed against epitopes which consist of negatively selected sites protect vaccinees more effectively than those directed against epitopes which contain positively selected sites. These observations suggest that candidate epitopes for immunization are predicted by the molecular evolutionary analysis of viral protein (and its coding nucleotide) sequences, as the epitopes which consist exclusively of negatively selected amino acid sites.

(27) New Methods for Detecting Positive Selection at Single Amino Acid Sites

Yoshiyuki Suzuki

Inferring positive selection at single amino acid sites is of particular importance for studying evolutionary mechanisms of a protein. For this purpose, Suzuki and Gojobori (1999) developed a method (SG method) for comparing the rates of synonymous and nonsynonymous substitutions at each codon site in a protein-coding nucleotide sequence, using ancestral codons at interior nodes of the phylogenetic tree as inferred by the maximum parsimony method. In the SG method, however, selective neutrality of nucleotide substitutions cannot be tested at codon sites, where only termination codons are inferred at any interior node or the number of equally parsimonious inferences of ancestral codons at all interior nodes exceeds 10,000. Here I present a modified SG method which is free from these problems. Specifically, I use the distance-based Bayesian method for inferring the single most likely ancestral codon from 61 sense codons at each interior node. In the computer simulation and real data analysis, the modified SG method showed a higher overall efficiency of detecting positive selection than the original SG method particularly at highly polymorphic codon sites. These results indicate that the modified SG method is useful for inferring positive selection at codon sites where neutrality cannot be tested by the original SG method. I also discuss that the p-distance is preferable to the number of synonymous substitutions for inferring the phylogenetic tree in the SG method, and present a maximum likelihood method for detecting positive selection at single amino acid sites, which produced reasonable results in the real data analysis.

(28) Three-dimensional window analysis for detecting positive selection at structural regions of proteins

Yoshiyuki Suzuki

Detection of natural selection operating at the amino acid sequence level is important in the study of molecular evolution. Single site analysis and one-dimensional window analysis can be used to detect selection when the biological functions of amino acid sites are unknown. Single site analysis is useful when selection operates more or less constantly over evolutionary time, but less so when it operates temporarily. One-dimensional window analysis is more sensitive than single site analysis when the functions of amino acid sites in close proximity in the linear sequence are similar although this is not always the

case. Here I present a three-dimensional window analysis method for detecting selection given the three-dimensional structure of the protein of interest. In the three-dimensional structure, the window is defined as the sphere centered on the α-carbon of an amino acid site. The window size is the radius of the sphere. The sites whose α-carbons are included in the window are grouped for the neutrality test. The window is moved within the three-dimensional structure by sequentially moving the central site along the primary amino acid sequence. To detect positive selection, it may also be useful to group the surface-exposed sites in the window separately. Three-dimensional window analysis appears to be not only more sensitive than single site analysis and one-dimensional window analysis, but also provides similar specificity for inferring positive selection in the analyses of the hemagglutinin and neuraminidase genes of human influenza A viruses. This method, however, may fail to detect selection when it operates only on a particular site, in which case single site analysis may be preferred although a large number of sequences is required.

(29) Evolutionary process of amino Acid biosynthesis in corynebacterium at the whole genome level

Yousuke Nishio, Yoji Nakamura, Yoshihiro Usuda, Shinichi Sugimoto, Kazuhiko Matsui, Yutaka Kawarabayasi, Hisashi Kikuchi, Takashi Gojobori and Kazuho Ikeo

Corynebacterium glutamicum, which is the closest relative of *Corynebacterium efficiens*, is widely used for the large scale production of many kinds of amino acids, particularly glutamic acid and lysine, by fermentation. *Corynebacterium diphtheriae*, which is well known as a human pathogen, is also closely related to these two species of Corynebacteria, but it lacks such productivity of amino acids. It is an important and interesting question to ask how those closely related bacterial species have undergone such significant functional differentiation in amino acid biosynthesis. The main purpose of the present study is to clarify the evolutionary process of functional differentiation among the three species of Corynebacteria by conducting a comparative analysis of genome sequences. When Mycobacterium and Streptomyces were used as out groups, our comparative study suggested that the common ancestor of Corynebacteria already possessed almost all of the gene sets necessary for amino acid production. However, *C. diphtheriae* was found to have lost the genes responsible for amino acid production. Moreover, we found that the common ancestor of *C. efficiens* and *C. glutamicum* have acquired some of genes responsible for amino acid production by horizontal gene transfer. Thus, we conclude that the evolutionary events of gene loss and horizontal gene transfer must have been responsible for functional differentiation in amino acid biosynthesis of the three species of Corynebacteria.

PUBLICATIONS

Papers

1. Nakamura, Y., Itoh, T., Matsuda, H. and **Gojobori, T.** (2004). Biased biological functions of horizontally transferred genes on 324,653 open reading frames of 116 prokaryotic complete genomes. ***Nature Genetics*** *36*, 760-6.

2. Iwama, H. and **Gojobori, T.** (2004). Highly conserved upstream sequences for transcription factor genes and implications for the regulatory network. ***Proc. Natl. Acad. Sci. USA*** *101*, 17156-61.

3. Imanishi, T. and other 152 authors, ***Gojobori, T.** and Sugano S. (2004). Integrative annotation of 21,037 human genes validated by full-length cDNA clones. ***PLoS Biology*** *2*, 1-21. *Correspondence Author

4. Ogura, A., **Ikeo, K.** and **Gojobori, T.** (2004). Comparative analysis of gene expression for convergent evolution of camera eye between octopus and human. ***Genome Res.*** *14*, 1555-61.

5. Andrews, T.D. and **Gojobori, T.** (2004). Strong positive selection and recombination drive the antigenic variation of the PilE protein of the human pathogen neisseria meningitidis. ***Genetics*** *166*, 25-32.

6. Hanada, K., **Suzuki, Y.** and **Gojobori, T.** (2004). A large variation in the rates of synonymous substitution for RNA viruses and its relationship to a diversity of viral infection and transmission modes. ***Mol Biol Evol.*** *21*, 1074-80.

7. Jung Shan, H., Kobayashi, C., Agata, K., **Ikeo, K.** and **Gojobori, T.** (2004). Detection of apoptosis during planarian regeneration by the expression of apoptosis-related genes and TUNEL assay. ***GENE*** *333*, 15-25.

8. Komiyama, T., **Ikeo, K.** and **Gojobori, T.** (2004). The evolutionary origin of long-crowing chicken: its

evolutionary relationship with fighting cocks disclosed by the mtDNA sequence analysis. ***GENE*** *333*, 91-99.
9. Komiyama, T., **Ikeo, K.,** Tateno, Y. and **Gojobori, T.** (2004). Japanese domesticated chickens have been drived from Shamo, traditional fighting cocks. ***Mol. Phylogenet. Evo.*** *33*, 16-21.
10. Mano, S., Yasuda, N., Katoh, T., Tounai, K., Inoko, H., Imanishi, T., Tamiya, G. and **Gojobori, T.** (2004). Notes on the maximum likelihood estimation of haplotype frequescies. ***Ann Hum Genet.*** *68*, 257-64.
11. Matsumoto, T., Yukawa, W., Nozaki, Y., Nakashige, R., Shinya, M., Makino, S., Yagura, M., Ikuta, T., Imanishi, T., Inoko, H., Tamiya, G. and **Gojobori, T.** (2004). Novel algorithm for automated genotyping of microsatelites. ***Nucleic Acids Res.*** *32*, 6069-77.
12. Tanaka, T., Tateno, Y. and **Gojobori, T.** (2004). Evolution of vitamin B6 (Pyridoxine) metabolism by gain and loss of genes. ***Mol. Biol. Evol.***
13. Alexopoulos, H., Bottger, A., Fischer, S., Levin, A., Wolf, A., Fujisawa, T., Hayakawa, S., **Gojobori, T.**, Davies, J., David, C. and Bacon, J. (2004). Evolution of gap junctions: the missing link? ***Curr. Biol.*** *14 (20)*, R879-80.
14. Bellgard, M., Ye, J., **Gojobori, T.** and Appels, R. (2004). The bioinformatics challenges is comparative analysis of cereal genomes-an overview. ***Funct. Integr. Genomics*** *4*, 1-11.
15. Kadota, M., Nishigaki, R., Wang, C.C., Toda, T., Shirayoshi, Y., Inoue, T., **Gojobori, T., Ikeo, K.,** Rogers, M.S. and Oshimura, M. (2004). Containing a single human chromosome 21 in neuronal differentiation: an in vitro model of Down syndrome. ***Neuroscience*** *129 (2)*, 325-35.
16. Miyazaki, S., Sugawara, H., **Ikeo, K., Gojobori, T.** and Tateno, Y. (2004). DDBJ in the stream of various biological data. ***Nucleic Acids Res*** *32*, D31-4.
17. Nishio, Y., Nakamura, Y., Usuda, Y., Sugimoto, S., Matsui, K., Kawarabayashi, Y., Kikuchi, H., **Gojobori, T.** and **Ikeo, K.** (2004). Evolutionary process of the amino acids biosynthesis in Corynebacterium at the whole genome level. ***Mol. Biol. Evol.*** *21*, 1683-91.
18. **Suzuki, Y.** (2004). Negative selection on neutralization epitopes of poliovirus surface proteins: implications for prediction of candidate epitopes for immunization. ***Gene*** *328*, 127-133.
19. **Suzuki, Y.** (2004). New methods for detecting positive selection at single amino acid sites. ***J. of Mol. Evol.*** *59*, 11-19.
20. **Suzuki, Y.** and Nei, M. (2004). False positive selection identified by ML-based methods: examples from the Sig1 gene of diatom Thalassiosira weissflogii and the tax gene of a human T-cell lymphotropic virus. ***Molecular Biology and Evolution*** *21*, 914-921.
21. **Suzuki, Y.** (2004). Three-dimensional window analysis for detecting positive selection at structural regions of proteins. ***Mol Biol and Evol*** *21*, 2352-2359.
22. Toyoda, R., Kasai, A., Sato, S., Wada, S., Saiga, H., **Ikeo, K., Gojobori, T.**, Numakunai, T. and Yamamoto, H. (2004). Pigment cell lineage-specific expression activity of the ascidian tyrosinase-related gene. ***GENE*** *332*, 61-69.
23. Wang, C.C., Kadota, M., Nishigaki, R., Kazuki, Y., Shirayoshi, Y., Rogers, M.S., **Gojobori, T., Ikeo, K.** and Oshimura, M. (2004). Molecular hierarchy in neurons differentiated from mouse ES cells containing a single human chromosome 21. ***Biochem. Biophys. Res. Commun.*** *314*, 335-50.
24. 伊藤　剛，五條堀孝(2004)「ゲノミクス・プロテオミクスの新展開」第5節比較ゲノム(進化)微生物のゲノム進化解析，エヌ・ティ・エス，897-903.
25. 小笠原倫大，五條堀孝(2004)「病理診断における分子生物学一第一部16. バイオインフォマティクス」『病理と臨床』臨時増刊号，91-97.
26. 五條堀孝(2004)「総合化生命科学としての現代進化学」(巻頭言)『科学』10月号，1203.
27. 五條堀孝(2004)「バイオインフォマティクス一生命の多様性と進化に基づく生命科学の統合化にむけて」『総研大ジャーナル』5号，26-28.
28. 田中　剛，五條堀孝(2004)「特集　分子進化学の現在・ゲノムの比較解析と分子進化」『生体の科学』55巻，241-6.
29. 花田耕介，五條堀孝(2004)「ゲノムの再編成」第一章ヒトゲノムの全貌『Molecular Medicine』増刊号，中山書店，101-109.
30. 溝上雅史，花田耕介，五條堀孝(2004)「ウイルス性肝炎(上)」日本臨床62巻，76-80.

Reviews

31. 五條堀孝・Hwang Jung Shan・田中信彦，「ヒトにおけるヒドラ遺伝子」，生体の科学第55巻5号，2004年10月

EDUCATION

1. Y. Suzuki "Methods for detecting positive selection at single amino acid sites" The Graduate University of Advanced Studies, Hayama, April, 2004.

I-b. Laboratory for Gene-Product Informatics
Ken Nishikawa Group

RESEARCH ACTIVITIES

(1) Eigenvalue analysis of amino acid substitution matrices reveals a sharp transition of the mode of sequence conservation in proteins

Akira R. Kinjo and Ken Nishikawa

The pattern of amino acid substitutions and sequence conservation over many structure-based alignments of protein sequences was analyzed as a function of percentage sequence identity. The statistics of the amino acid substitutions were converted into the form of log-odds amino acid substitution matrices to which eigenvalue decomposition was applied. It was found that the most important component of the substitution matrices exhibited a sharp transition at the sequence identity of 30-35%, which coincides with the twilight zone. Above the transition point, the most dominant component is related to the mutability of amino acids and it acts to disfavor any substitutions, whereas below the transition point, the most dominant component is related to the hydrophobicity of amino acids and substitutions between residues of similar hydrophobic character are positively favored. Implications for protein evolution and sequence analysis are discussed. See Ref. 1 for details.

(2) Estimation of the number of authentic orphan genes in bacterial genomes

Satoshi Fukuchi and Ken Nishikawa

Genome annotation produces a considerable number of putative proteins lacking sequence similarity to known proteins. These are referred to as "Orphans". The proportion of orphan genes varies among genomes, and is independent of genome size. In the present study, we show that the proportion of orphan genes roughly correlates with the isolation index of organisms (IIO), an indicator introduced in the present study, which represents the degree of isolation of a given genome as measured by sequence similarity. However, there are outlier genomes with respect to the linear correlation, consisting of those genomes that may contain excess amounts of orphan genes. Comparisons of genome sequences among closely related strains revealed that some of the annotated genes are not conserved, suggesting that they are ORFs occurring by chance. Exclusion of these non-conserved ORFs within closely related genomes improved the correlation between the proportion of orphan genes and the IIO values. Assuming that the correlation holds in general, this relationship was used to estimate the number of "authentic" orphan genes in a genome. Using this definition of authentic orphan genes, the anomalies arising from over-assignments, e.g., the percentages of structural annotations, were corrected for 16 genomes, including those of five archaea. See Ref. 2 for details.

(3) Alternative splice variants encoding unstable protein domains exist in the human brain

Keiichi Homma, Reiko F. Kikuno, Takahiro Nagase, Osamu Ohara and Ken Nishikawa

Alternative splicing has been recognized as a major mechanism by which protein diversity is increased without significantly increasing genome size in animals and has crucial medical implications, as many alternative splice variants are known to cause diseases. Despite the importance of knowing what structural changes alternative splicing introduces to the encoded proteins for the consideration of its significance, the problem has not been adequately explored. Therefore, we systematically examined the structures of the proteins encoded by the alternative splice variants in the HUGE protein database derived from long (>4 kb) human brain cDNAs. Limiting our analyses to reliable alternative splice junctions, we found alternative splice junctions to have a slight tendency to avoid the interior of SCOP domains and a strong statistically significant tendency to coincide with SCOP domain boundaries. These findings reflect the occurrence of some alternative splicing events that utilize protein structural units as a cassette. However, 50 cases were identified in which SCOP domains are disrupted in the middle by alternative splicing. In six of the cases, insertions are introduced at the molecular surface, presumably affecting protein functions, while in 11 of the cases alternatively spliced variants were found to encode pairs of stable and unstable proteins. The mRNAs encoding such unstable proteins are

much less abundant than those encoding stable proteins and tend not to have corresponding mRNAs in non-primate species. We propose that most unstable proteins encoded by alternative splice variants lack normal functions and are an evolutionary dead-end. See Ref. 3 for details.

(4) Construction and characterization of chimeric proteins composed of type-1 and type-2 periplasmic binding proteins MglB and ArgT

Kenji Kashiwagi, Kaoru Fukami-Kobayashi, Kiyotaka Shiba and Ken Nishikawa

The respective type-1 and type-2 periplasmic binding proteins (PBPs) MglB and ArgT are believed to have evolved from a common ancestor into siblings showing topological differences in their main chain connectivity. At first glance, they show similar structure. But, more detailed examination reveals that the chain connectivity of ArgT is more convoluted than that of MglB. Reflecting that complexity, the folding of ArgT is complicated and involves intermediate folds. On the other hand, the folding of MglB is a simple two-state transition. In the present study, we constructed and characterized several chimeras made up of various subdomains of MglB and ArgT with the aim of gaining insight into the evolution of protein folding and protein structure. Although these chimeras did not fold as compactly as their parental proteins, some did exhibit cooperative folding, which suggests that novel proteins with new connectivity and new folding pathways could have emerged at a fairly high rate throughout the evolution of proteins. See Ref. 4 for details.

PUBLICATIONS

Papers

1. Kinjo, A.R. and Nishikawa, K. (2004). Eigenvalue analysis of amino acid substitution matrices reveals a sharp transition of the mode of sequence conservation in proteins. *Bioinformatics, 20 (16)* 2504-2508.
2. Fukuchi, S. and Nishikawa, K. (2004). Estimation of the number of authentic orphan genes in bacterial genomes. *DNA Res., 11*, 219-231.
3. Homma, K., Kikuno, R.F., Nagase, T., Ohara, O. and Nishikawa, K. (2004). Alternative splice variants encoding unstable protein domains exist in the human brain. *J. Mol. Biol., 343*, 1207-1220.
4. Kashiwagi, K., Fukami-Kobayashi, K., Shiba, K. and Nishikawa, K. (2004). Construction and characterization of chimeric proteins composed of type-1 and type-2 periplasmic binding proteins MglB and ArgT. *Biosci. Biotechnol. Biochem., 68 (4)* 808-813.
5. Imanishi, T., Itoh, T., Suzuki, Y., O'Donovan, C. and Fukuchi, S. et al. (2004). Integrative annotation of 21,037 human genes validated by full-length cDNA clones. *PLoS Biology, 2 (6)*, 856-875.
6. Kinjo, A.R., Horimoto, K. and Nishikawa, K. (2005). Predicting absolute contact numbers of native protein structure from amino acid sequence. *Proteins, 58*, 158-165.
7. Kinjo, A.R. and Nishikawa, K. Recoverable one-dimensional encoding of protein three-dimensional structures. *Bioinformatics*, in press.

Reviews

8. 吉宗一晃, 福地佐斗志, 森口充瞭, 西川建(2004)「タンパク質から見た極限微生物の環境適応戦略」, バイオサイエンスとインダストリー, Vol.62, 17-22.
9. 福地佐斗志, 西川建(2004)「蛋白質構造解析プログラム・データベース」蛋白質核酸酵素増刊「バイオ高性能機器・新技術利用マニュアル」(小原収他編, 共立出版) Vol.49 (11), pp.1944-1948.

Database

GTOP(ゲノム中のタンパク質立体構造DB):
http://spock.genes.nig.ac.jp/~genome/gtop.html
PMD(変異タンパク質DB):
http://pmd.ddbj.nig.ac.jp/~pmd/pmd.html
TTDB(原核生物の転写因子DB):
http://spock.genes.nig.ac.jp/~ttdb/

EDUCATION

1. 福地佐斗志　第9回DDBJing講習会, 東京, 3月.
2. 西川建, 福地佐斗志, 金城玲　科学技術振興事業団主催ゲノムリテラシー講座「データベースを利用した蛋白質の立体構造予測」東京, 7月.
3. 西川建　立命館大学理工学部生命情報学科セミナー, 11月.

I-c. Laboratory for Gene Function Research
Yoshio Tateno Group

RESEARCH ACTIVITY

(1) Submission of microarray data to public repositories

Catherine A. Ball, Alvis Brazma, Helen Causton, Steve Chervitz, Ron Edgar, Pascal Hingamp, John C. Matese, Helen Parkinson, John Quackenbush, Martin Ringwald, Susanna-Assunta Sansone, Gavin Sherlock, Paul Spellman, Chris Stoeckert, Yoshio Tateno, Ronald Taylor, Joseph White and Neil Winegarden

What this work states is a change in the way in which we approach the publication of microarray-based studies. Both authors and journals have a responsibility to assure that the requisite data are available, and because submitting MIAME-compliant data can take considerable time and effort, this process should be factored into review and publication timelines. However, while this process may be time consuming and painful at first, we believe that the benefits of building an open repository of microarray data will far outweigh any initial disadvantages. As always, it is our sincere hope that these suggestions stimulate discussion within the community and that together we can arrive at a consensus that ensures that microarray data are widely and easily accessible. Finally we would like to urge the DDBJ, EBI, and NCBI to work together towards exchanging all MIAME-compliant microarray data[1].

(2) The origin of eukaryotes is suggested as the symbiosis of pyrococcus into proteobacteria by phylogenetic tree based on gene content

Tokumasa Horiike, Kazuo Hamada, Daisuke Miyata and Takao Shinozawa

Attempts were made to define the relationship among the three domains (eukaryotes, archaea, and eubacteria) using phylogenetic tree analyses of 16S rRNA sequences as well as of other protein sequences. Since the results are inconsistent, it is implied that the eukaryotic genome has a chimeric structure. In our previous studies, the origin of eukaryotes to be the symbiosis of archaea into eubacteria using the whole open reading frames (ORF) of many genomes was suggested. In these studies, the species participating in the symbiosis were not clarified, and the effect of gene duplication after speciation (in-paralog) was not addressed. To avoid the influence of the in-paralog, we developed a new method to calculate orthologous ORFs. Furthermore, we separated eukaryotic in-paralogs into three groups by sequence similarity to archaea, eubacteria (other than -proteobacteria), and -proteobacteria and treated them as individual organisms. The relationship between the three ORF groups and the functional classification was clarified by this analysis. The introduction of this new method into the phylogenetic tree analysis of 66 organisms (4 eukaryotes, 13 archaea, and 49 eubacteria) based on gene content suggests the symbiosis of pyrococcus into -proteobacteria as the origin of eukaryotes[2].

(3) Integrative annotation of 21,037 human genes validated by full-length cDNA clones

Tadashi Imanishi, Takeshi Itoh, Yutaka Suzuki, Claire O'Donovan, Satoshi Fukuchi, Kanako O. Koyanagi, Roberto A. Barrero, Yoshio Tateno, Zhu Chen, Michio Oishi, Peter Tonellato, Rolf Apweiler, Kousaku Okubo, Lukas Wagner, Stefan Wiemann, Robert L. Strausberg, Takao Isogai, Charles Auffray, Nobuo Nomura, Takashi Gojobori and Sumio Sugano. *et al.*

The human genome sequence defines our inherent biological potential; the realization of the biology encoded therein requires knowledge of the function of each gene. Currently, our knowledge in this area is still limited. Several lines of investigation have been used to elucidate the structure and function of the genes in the human genome. Even so, gene prediction remains a difficult task, as the varieties of transcripts of a gene may vary to a great extent. We thus performed an exhaustive integrative characterization of 41,118 full-length cDNAs that capture the gene transcripts as complete functional cassettes, providing an unequivocal report of structural and functional diversity at the gene level. Our international collaboration has validated 21,037 human gene candidates by analysis of high-quality full-length cDNA clones through curation using unified criteria. This led to the identification of 5,155 new gene candidates. It also manifested the most reliable way to control the quality of the cDNA clones. We have developed a human gene

database, called the H-Invitational Database (H-InvDB; http://www.h-invitational.jp/). It provides the following: integrative annotation of human genes, description of gene structures, details of novel alternative splicing isoforms, non-protein-coding RNAs, functional domains, subcellular localizations, metabolic pathways, predictions of protein three-dimensional structure, mapping of known single nucleotide polymorphisms (SNPs), identification of polymorphic microsatellite repeats within human genes, and comparative results with mouse full-length cDNAs. The H-InvDB analysis has shown that up to 4% of the human genome sequence (National Center for Biotechnology Information build 34 assembly) may contain misassembled or missing regions. We found that 6.5% of the human gene candidates (1,377 loci) did not have a good protein-coding open reading frame, of which 296 loci are strong candidates for non-protein-coding RNA genes. In addition, among 72,027 uniquely mapped SNPs and insertions/deletions localized within human genes, 13,215 nonsynonymous SNPs, 315 nonsense SNPs, and 452 indels occurred in coding regions. Together with 25 polymorphic microsatellite repeats present in coding regions, they may alter protein structure, causing phenotypic effects or resulting in disease. The H-InvDB platform represents a substantial contribution to resources needed for the exploration of human biology and pathology[3].

(4) Structural and functional differences in two cyclic bacteriocins with the same sequences produced by lactobacilli

Yasushi Kawai, Yasuyuki Ishii, Kensuke Arakawa, Koichiro Uemura, Boku Saitoh, Junko Nishimura, Haruki Kitazawa, Yukiko Yamazaki, Yoshio Tateno, Takatoshi Itoh and Tadao Saito

Lactobacillus gasseri LA39 and *L. reuteri* LA6 isolated from feces of the same human infant were found to produce similar cyclic bacteriocins (named gassericin A and reutericin 6, respectively) that cannot be distinguished by molecular weights or primary amino acid sequences. However, reutericin 6 has a narrower spectrum than gassericin A. In this study, gassericin A inhibited the growth of L. reuteri LA6, but reutericin 6 did not inhibit the growth of L. gasseri LA39. Both bacteriocins caused potassium ion efflux from indicator cells and liposomes, but the amounts of efflux and patterns of action were different. Although circular dichroism spectra of purified bacteriocins revealed that both antibacterial peptides are composed mainly of alpha-helices, the spectra of the bacteriocins did not coincide. The results of D- and L-amino acid composition analysis showed that two residues and one residue of D-Ala were detected among 18 Ala residues of gassericin A and reutericin 6, respectively. These findings suggest that the different D-alanine contents of the bacteriocins may cause the differences in modes of action, amounts of potassium ion efflux, and secondary structures. This is the first report that characteristics of native bacteriocins produced by wild lactobacillus strains having the same structural genes are influenced by a difference in D-amino acid contents in the molecules[4].

(5) Japanese domesticated chickens have been derived from Shamo traditional fighting cocks

Tomoyoshi Komiyama, Kazuho Ikeo, Yoshio Tateno and Takashi Gojobori

With the aim of elucidating the evolutionary origin of Japanese domesticated chickens, this study evolutionarily analyzed 85 chicken mtDNA sequences. Thirty-four various ornamental chickens, 42 fighting cocks (Shamo), and nine long-crowing chickens (Naganakidori) were included. Of the Shamo, 18 were sampled from Okinawa, while the remaining 24 were collected in other islands around Japan. In addition, three Southeast Asian Junglefowls were used as a reference to determine the common ancestor of Japanese domesticated chickens. A phylogenetic tree was constructed for the 88 mtDNA sequences revealing that the Shamo group from Okinawa clearly diverged from the other Japanese domesticated chickens studied. This strongly suggests that all Japanese domesticated chickens, including the ornamental varieties and Naganakidori, derived from the ancestors of the Shamo in Okinawa. To create novel varieties of ornamental chickens, intensive artificial selection is imposed on ancestral Shamo populations, resulting in profoundly differentiated Japanese domesticated chickens[5].

(6) DDBJ in the stream of various biological data

Satoru Miyazaki, Hideaki Sugawara, Kazuho Ikeo, Takashi Gojobori and Yoshio Tateno

In the past year we at DDBJ (http://www.ddbj.nig.ac.jp) have made a steady increase in the number of data submissions with a 50.6% increment in the number of bases or 46.5% increment in the number of entries. Among them the genome data of man, ascidian and rice hold the top three. Our activity has extended to providing a tool that enables sequence retrieval using regular expressions, and to launching our SOAP server and web services to facilitate the acquisition of proper data and tools from a huge number of biological data resources on websites worldwide. We have also opened our public gene expression database, CIBEX[6].

(7) Molecular chaperones: proposal of a systematic computer-oriented nomenclature and construction of a centralized database

Haitham Sghaier, Thuy Le Huyen Ai, Tokumasa Horiike and Takao Shinozawa

Molecular chaperones are a wide group of unrelated protein families whose role is to assist others proteins. Comparably, under environmental stress, stress proteins behave as biocatalysts of protein stabilization. Stress proteins include a large class of proteins that were originally termed heat shock proteins (HSPs) due to their initial discovery in tissues exposed to elevated temperatures. Many, but not all, stress proteins and HSPs are molecular chaperones. Moreover, not all HSPs are derivable from stress. HSPs are structurally diversified by the contribution of various domains having specific roles. HSPs have been grouped, mainly on the basis of their molecular masses, into specific families that include small HSPs (sHSPs)/ α-crystallins, HSP10s, HSP40s, HSP60s, HSP70s, HSP90s, HSP100s and HSP110s. The names of these major families are historical artefacts with limited information content. Using the current databases, names and proteic domains of many molecular chaperones in different species were analyzed. Although traditional names of HSPs are trivial, it is unrealistic to suggest replacing them, because they are preferred and widely used. Here we suggest that these traditional names be chaperoned, *in silico*, by a systematic nomenclature. Thus, for example, with the same intent of use of [trioxygen: O_3] for ozone, we propose here C7HSP70[Ehsa]ER-P11021 for GRP78 (78 kDa endoplasmic Human molecular chaperone in HSP70 superfamily with P11021 as its accession number in the database of the National Center for Biotechnology Information (NCBI)). The proposed systematic computer-oriented naming and classification method is designed for HSPs and also their partners based on the number of amino acids, domain structure, phylogenetic domain, localization in the cell and accession number as stated in the NCBI. *Arabidopsis thaliana* was analyzed as a model, because it contains a large number of various HSPs localized in several organelles. Overall, this naming system helps in building, optimizing and managing a novel online database entirely devoted to HSPs. The purported taxonomy, coupled with the newly constructed database, can contribute to studies involving large amounts of stored data on HSPs[7].

(8) Extensive analysis of ORF sequences from two different cichlid species in Lake Victoria provides molecular evidence for a recent radiation event of the Victoria species flock Identity of EST sequences between Haplochromis chilotes and Haplochromis sp. "Redtailsheller"

Masakatsu Watanabe, Naoki Kobayashi, Tadasu Shin-I, Tokumasa Horiike, Yoshio Tateno, Yuji Kohara and Norihiro Okada

The Lake Victoria Cichlid fishes have diverged very rapidly. The estimated 500 species inhabiting the lake are believed to have arisen within the last 14,000 years. The fishes' jaws and teeth have diverged markedly to adapt to different feeding behaviors and environments. To examine how the genomes of these fishes differentiated during speciation, we performed comparative analysis of expressed sequenced tag (EST) sequences. We constructed cDNA libraries derived only from the jaw portions of two cichlid species endemic to Lake Victoria. We sequenced 17,280 cDNA clones from Haplochromis chilotes and 9600 cDNA clones from Haplochromis sp. "Redtailsheller" and obtained 543 different genes common to both species. Of these genes, 441 were essentially identical between species and 102 contained base replacements in their open reading frame (ORF) or untranslated (UTR) regions.

Comparative analysis of 71selected sequences has revealed that while the degree of polymorphism is 0.0054/site for H. chilotes and 0.0047/site for H. sp. " Redtailsheller", genetic distance between the two species is 0.0031/site. The genetic distance particularly indicates that the two species diverged about 890,000 years ago[8].

(9) International public gene expression database (CIBEX) and data submission

Yoshio Tateno and Kazuho Ikeo

We have opened our gene expression database, CIBEX, to the public. CIBEX has been developed as an international public database with the aim of the collaboration with ArrayExpress at EBI and GEO at NCBI. The collaboration mainly means to share the annotation manual and to exchange the data collected and annotated among the three databases. The data collection will hopefully be promoted by the open letter issued by the MGED society to the editors of relevant journals[9].

(10) Analysis of biological networks in eukaryotes using the whole genome sequences

Tsuyoshi Tanaka and Takashi Gojobori

Since the whole genome sequencing of *Haemophilus influenzae* was completed in 1995, the number of species whose genomes were completely sequenced has steeply been increased. As of January 2005, the number of such species is more than 210 in the Genome Information Broker (GIB) of the Center for Information Biology and DNA Data Bank of Japan. The information on the whole genome sequences enables us to study the origins and evolutionary processes of various biological networks such as metabolic pathways and signal transductions. We analyzed biological networks such as amino acid metabolic pathways by conducting comparative analysis of the complete genome sequences of six eukaryotic species including man, fly, nematode, yeasts and plant, and found that a particular pathway had evolved independently in multiple lineages of the species studied[10].

(11) Microarray gene expression database

Kazuho Ikeo and Yoshio Tateno

As the international standardizations of microarray data description and data sharing have been promoted by the MGED society, researchers conducting microarray experiments are encouraged to submit their data to one of the international gene expression databases, ArrayEpress, GEO and CIBEX. CIBEX being developed by us is in compliance with the international standard, MIAME, and equipped with several search functions[11].

(12) Discovery and annotation of forty seven non-protein coding human RNAs

Roberto A. Barrero, Inna Dubchak, Charles Auffray, Laurens Wilming, Jun-ichi Takeda, Yutaka Suzuki, Erimi Harada, Marie-Anne Debily, Esther Graudens, John Quackenbush, Takuro Tamura, Dmitriy V. Ryaboy, Sandrine Imbeaud, Kazuho Ikeo, Peter Tonellato, Nobuo Nomura, Sumio Sugano, Tadashi Imanishi, Takashi Gojobori and Libin Jia

Non-coding RNAs (ncRNAs) play crucial roles in a variety of procsses including replication, transcriptional regulation, splicing, dosage compensation, genetic imprinting, translational regulation, and modulation of protein function. Here we report the discovery and annotation of ncRNAs from the human full-length cDNA dataset evaluated at the first International Human Full-length cDNA Annotation Meeting. A total of 1,485 cDNA transcripts, mapped onto 1,300 loci on the human genome, were found to encode putative open reading frames (ORFs) equal to or less than 80 amino acids (aa). To select putative ncRNAs all cDNA sequences were mapped to the human genome to study the genomic neighbourhood for the presence of ab initio predicted genes and neighbouring genes, and compared to Expressed Sequence Tag (EST) databases for supporting evidence. This method yielded 296 putative ncRNAs that were analyzed for conservation by determining mouse DNA and RNA sequence similarities. Putative ncRNAs with mouse ortholog support were further analyzed using QRNA. We found 47 ncRNAs containing a conserved RNA secondary structure. Of these, 60% were found to be expressed in up to eight

human tissues, implying that ncRNAs are seemingly tissue-specifically regulated.

(13) Evolutionary rate of enzymes in the metabolic network

Tsuyoshi Tanaka, Kazuho Ikeo and Takashi Gojobori

An enzyme interacts not only with the other proteins but also with low-weight molecules called substrates in the metabolic network. To understand an evolutionary process of interactions of enzymes, we studied the relationship between the evolutionary rate of the enzyme and these interacting partners. When we focused on the 498 enzymes in *Saccharomyces cerevisiae* that have orthologous pairs in *Ashbya gossypii*, we discovered the significant negative correlation between the evolutionary rate of the enzyme and the number of interacting proteins (protein-protein interaction; PPI). On the other hand, we found no correlation between the evolutionary rate of the enzyme and the number of interacting substrates (protein-substrate interaction; PSI). Therefore, we conclude that the number of interacting proteins is the most affective to the evolutionary rate of the enzyme compared with that of the other interacting partners such as the substrates.

(14) Development of a method for constructing a phylogenetic tree using a comprehensive orthologous gene cluster, and phylogenetic analysis of cyanobacteria

Tokumasa Horiike and Yoshio Tateno

Phylogenetic trees are constructed using DNA, RNA or amino acid sequences for estimating evolutionary relationships of genes or species. Currently, there are two problems with the tree construction. One is that horizontal gene transfer disturbs the estimation of the true relationships of genes or species. The other is that the construction sometimes depends on the choice of sequences. We occasionally observe that changing one sequence to another erroneously alters the reconstructed tree. Therefore, we are developing a method of the tree construction which is to reduce the interference caused by the two problems. In the method we can incorporate all available prokaryotic ORFs. We will then clarify the phylogenetic position of cyanobacteria by applying our method to as many pertinent sequences as possible.

(15) We are in collaboration with Prof. Tadao Saito of Tohoku University and his laboratory on the function and evolution of glucosidase and galactosidase genes in *Lactobcillus*

Tadao Saito, Yukiko Yamazaki and Yoshio Tateno

(16) We are in collaboration with Prof. Shintou Eguchi of the Institute of Mathematical Statistics and his laboratory on the statistical analyses of SNP and gene expression data

Shinto Eguchi, Kazuho Ikeo and Yoshio Tateno

PUBLICATIONS

Papers

1. Ball, CA., Brazma, A., Causton, H., Chervitz, S., Edgar, R., Hingamp, P., Matese, JC., Parkinson, H., Quackenbush, J., Ringwald, M., Sansone, SA., Sherlock, G., Spellman, P., Stoeckert, C., Tateno, Y., Taylor, R., White, J. and Winegarden, N. (2004). Submission of microarray data to public repositories. PLoS Biol. *2*, 1276-1277, also Microbiology *150*, 3522-3524 and Environ Health Perspect *112*, A666-A667.
2. Horiike, T., Hamada, K., Miyata, D. and Shinozawa, T. (2004). The origin of eukaryotes is suggested as the symbiosis of pyrococcus into γ-proteobacteria by phylogenetic tree based on gene content. J. Mol. Evol. *59*, 606-619.
3. Imanishi, T., Itoh, T., Suzuki, Y., O'Donovan, C., Fukuchi, S., Koyanagi, KO., Barrero, RA., Tateno, Y., Chen, Z., Oishi, M., Tonellato, P., Apweiler, R., Okubo, K., Wagner, L., Wiemann, S., Strausberg, RL., Isogai, T., Auffray, C., Nomura, N., Gojobori, T. and Sugano, S. *et al.* (2004). Integrative annotation of 21,037 human genes validated by full-length cDNA clones. PLoS Biol. *2*, 256-275.
4. Kawai, Y., Ishii, Y., Arakawa, K., Uemura, K., Saitoh, B., Nishimura, J., Kitazawa, H., Yamazaki, Y., Tateno, Y., Itoh, T. and Saito, T. (2004). Structural and functional differences in two cyclic bacteriocins with the same sequences produced by lactobacilli. Appl Environ Microbiol. *70*, 2906-2911.
5. Komiyama, T., Ikeo, K., Tateno, Y. and Gojobori, T. (2004). Japanese domesticated chickens have been

derived from Shamo traditional fighting cocks. Mol Phylogenet Evol. *33*, 16-21.
6. Miyazaki, S., Sugawara, H., Ikeo, K., Gojobori, T. and Tateno, Y. (2004). DDBJ in the stream of various biological data. Nucleic Acids Res. *32*, D31-D34.
7. Sghaier, H., Ai, TL., Horiike, T. and Shinozawa, T. (2004). Molecular chaperones: proposal of a systematic computer-oriented nomenclature and construction of a centralized database. In Silico Biol. *16*, 0025.
8. Watanabe, M., Kobayashi, N., Shin-I, T., Horiike, T., Tateno, Y., Kohara, Y. and Okada, N. (2004). Extensive analysis of ORF sequences from two different cichlid species in Lake Victoria provides molecular evidence for a recent radiation event of the Victoria species flock Identity of EST sequences between Haplochromis chilotes and Haplochromis sp. "Redtailsheller". Gene *343*, 263-269.
9. 舘野義男，池尾一穂(2004)国際公共遺伝子発現データベース(CIBEX)とデータの登録，蛋白質核酸酵素*49*, 2679-2683.
10. 田中剛，五條堀孝(2004)ゲノムの比較解析と分子進化，生体の科学*55*, 241-246.
11. 池尾一穂，舘野義男(2004)マイクロアレイ遺伝子発現データベース，ゲノミックスプロテオミックスの新展開(今中忠行監修)，pp828-836，NTS出版.

Database

12. DDBJ is operated by DNA Data Analysis, Gene Function Research, Gene-Product Informatics, Research and Development of Biological Databases and Gene Expression Laboratories, and Division of Population Genetics. DDBJ collects, annotates and publishes DNA sequence data, and exchanges the data with EMBL Bank and GenBank on a daily basis. In addition, DDBJ edits the data published by DDBJ, EMBL Bank and GenBank together four times a year and publishes as a release. In 2004 DDBJ published the following four releases.

Release 57	March., 04	32,693,678 entries
		38,008,449,840 bases
Release 58	June., 04	34,917,581 entries
		39,812,635,108 bases
Release 59	Sept., 04	37,926,117 entries
		44,416,752,273 bases
Release 60	Dec., 04	40,583,945 entries
		42,245,956,937 bases

13. Dr. Y. Tateno attended the 17th International Nucleotide Sequence Databases Collaborative Meeting, and the 15th International Nucleotide Sequence Databases Advisors Meeting held at EBI, Hinxton, UK, May, 2004.
14. Dr. Y. Tateno attended the MGED Board Meeting, Toronto, Canada, September, 2004.
15. Dr. Y. Tateno was invited to be a chairperson of a session at the UniProt Meeting, Antibes, France, September, 2004.

EDUCATION

1. Dr. Y. Tateno co-organized the 3rd Japan-Korea Bioinformatics Training Course at NIG, March, 2004.
2. Dr. Y. Tateno co-organized the SOKENDAI international lecture and gave a lecture on DNA and Protein Databases at the Olympic Hotel, Shanghai, China, October, 2004.

SOCIAL CONTRIBUTIONS AND OTHERS

学会活動

1. Dr. Y. Tateno serves the Microarray Gene Expression Data Society as an Advisory Board Member (2001-), and attended the Board Member Meeting, Toronto, Canada, September, 2004.
2. Dr. Y. Tateno serves the Society of Evolutionary Studies as a Member of the Committee of Genetics under the Science Council of Japan.
3. Dr. Y. Tateno serves the Genetics Society of Japan as a Member of Editorial Board.

I-d. Laboratory for Research and Development of Biological Databases Hideaki Sugawara Group

RESEARCH ACTIVITIES

(1) Information systems for molecular biology and its related disciplines

1) From Web services to a Bioportal

Yasumasa Shigemoto†, Haruka Sakai, Takashi Abe, Satoru Miyazaki†† and Hideaki Sugawara (†Hitachi soft, ††Tokyo Univ. of Sci.)

The publicly available bioinformatics resources, comprising databases and analytical tools, have

expanded in recent years. While the information environment for life sciences has gradually become more abounding, it is still difficult to combine multiple, heterogeneous bioinformatics resources for a specific research purpose. To set up and run an integrated system, it is often necessary to write and update custom programs. In addition, different research groups continually write programs that have overlapping functions. We need an information environment that is conducive to efficient and appropriate bioinformatics resource utilization for a wide range of users. Therefore, the Center for Information Biology and DNA Data Bank of Japan, in alliance with the National Institute of Informatics (NII) and the Mitsubisi Research Institute, Inc. (MRI) have started a three years long project since 2003, "Research and Development of the New Generation of Bio-portal", to enhance the information environment for the relevant user communities. In this project, the Laboratory for Research and Development of Biological Databases is responsible for the development of biological Web services. The project site is open at http://www.bioportal.jp/ in 2004. From there, a Web page for links to sites complete genome sequence and annotation are also prepared and accessible, in addition to the biological Web services. The former is named "Genome Menu".

2) Expansion of Genome Information Broker (GIB)

Masaki Hirahata, Naoto Tanaka, Takashi Abe, Satoru Miyazaki† and Hideaki Sugawa (†Tokyo Univ. of Sci.)

GIB was originally created for the retrieval and analysis of E. coli genomic information in a set. We implemented microbial genome data into GIB whenever genome sequencing was completed and the data is made open to the public. At the GIB Web page (http://gib.genes.nig.ac.jp/), key word search, homology search, links to DBGET, KEGG and GTOP and visualization of the data are available for more than 200 strains as of December 2004. We have utilized XML, CORBA and a distributed database in order to cope with the explosion of microbial genome information.

(2) Information systems on microbes[1)]

1) WFCC-MIRCEN World Data Centre for Microorganisms (WDCM)

Yasumasa Shigemoto†, Junko Nagaya and Hideaki Sugawara (†Fujitsu)

WFCC and MIRCEN stand for World Federation for Culture Collections and Microbial Resource Centers network respectively. The laboratory is the host of WDCM and maintains the World Directory of microbial resource centers. The on-line World Directory contains the detailed information of 469 centers in 65 countries and also the list of their holdings. Any culture collection is able to register, update and delete the information at http://www.wdcm.org/. WDCM could promote the update of the data by culture collections funded by the American Society for Microbiology and UNESCO.

2) Development of an e-Workbench for Biological Classification and Identification (InforBIO)

Naoto Tanaka, Kouji Koorikawa†, Takashi Abe, Satoru Miyazaki†† and Hideaki Sugawara (†Hitachi soft, ††Tokyo Univ. of Sci.)

We continued the development of an e-Workbench named InforBIO by use of JAVA, XML and a relational database management system in the public domain. We have distributed InforBIO to several laboratories that study microbes and improved the utility and robustness of InforBIO based on the feedback (http://lilium.genes.nig.ac.jp/index_e.html).

3) An information system for pathogenic microorganisms

Masaki Hirahata, Naoto Tanaka, Yasumasa Shigemoto† and Hideaki Sugawa (†Fujitsu)

We participated in a national project for the resource center of pathogenic microorganisms. Our role is to develop an information system for pathogenic fungi and actinomycetes, and also a portal site for pathogenic microorganisms in general (http://www.wdcm.org/byogen/).

(*) The information system on pathogenic microor-

ganisms has been supported by Special Coordination Funds for Promoting Science and Technology.

(3) Applications of IT to the International Nucleotide Sequence Database[2)]

1) Development of Open Annotation System

Satoru Miyazaki†, Takashi Abe and Hideaki Sugawara (†Tokyo Univ. of Sci.)

A number of the complete genome sequences have been submitted to INSD since 1995. The annotation information, however, is not consistent among genome sequencing teams. In addition, researchers outside of the team might have more information and knowledge on some genes and biological molecules. Therefore, it is quite important to develop the system which allows any expert to evaluate the annotation given by the team to attach more valuable information. As a new feature of INSD, we develop so-called "Open Annotation System (OASYS)" as an annotation editor in the distributed environment on the Internet.

(*) OASYS project has been supported by BIRD of Japan Science and Technology Corporation (JST)

2) Exhaustive evaluation of microbial genome information by use of GRID

Takehide Kosuge, Toshihisa Okido, Yasumasa Shigemoto†, Masaki Hirahata, Naoto Tanaka, Yuzuru Maruyama, Takashi Abe, Satoru Miyazaki†† and Hideaki Sugawara (†Fujitsu, ††Tokyo Univ. of Sci.)

Tsunami of biological data and multiple views of the data analysis require an expandable and flexible information environment. GRID computing is expected to be the solution. We prepared a computational environment composed of 5 sites in OBIGrid and succeeded in analyzing horizontal gene transfer and clusters of ORFs of more than 100 microbial genomes that were stored in the Genome Information Broker as of May, 2003. This scheme is being applied to more than 300 thousands ORFs of genomic sequences of 124 microbial species. In 2004, we evaluated the results of the analysis and have developed site to diffuse the result to the public. We also applied the workflow to all the microbial genome sequences that were publicly available by September 2004.

(4) Genomics

1) Development of the H-Invitational Database

Yasumasa Shigemoto†, Satoru Miyazaki†† and Hideaki Sugawara (†Fujitsu, ††Tokyo Univ. of Sci.)

We performed an exhaustive integrative characterization of 41,118 full-length cDNAs that capture the gene transcripts as complete functional cassettes, providing an unequivocal report of structural and functional diversity at the gene level. Our international collaboration has validated 21,037 human gene candidates by analysis of high-quality full-length cDNA clones through curation using unified criteria. We have developed a human gene database, called the H-Invitational Database (H-InvDB; http://www.h-invitational.jp/). The H-InvDB platform represents a substantial contribution to resources needed for the exploration of human biology and pathology.

2) Splicing Profile Based Protein Categorization between Human and Mouse Genomes

Åke Västermark†, Yasumasa Shigemoto††, Takashi Abe and Hideaki Sugawara (†Univ. of Oxford, ††Fujitsu)

We compared gene structures of human and mouse to explore the relationships of functions of genes and exon-intron structures. The central question is whether protein function is more correlated with splicing profiles than sequence similarity, or not. To approach this question, a splicing profile similarity (SPS) index, which measures relative exon length discrepancy, was devised. Arbitrary human proteins were compared, in terms of SPS and amino acid sequence similarity, to their 1) mouse orthologues and 2) human paralogues, which epitomise functional equivalence and non-equivalence, respectively, to methodically elucidate the global relationship between a) biological function, b) splicing profile similarity, and c) sequence similarity. Protein function is more correlated with splicing profile similarity than sequence similarity as demonstrated by the fact that human-mouse orthologues (HMOs) display significantly higher splicing profile similarity than do human-human paralogues (HHPs), despite the mutual sequence similarity between these two categories. This finding

indicates that splicing profile-based protein categorisation is biologically meaningful[4].

3) Phylogenetics Analyses of Environmental Samples on the Basis of Self-Organizing Map (SOM)

Takashi Abe, Toshimichi Ikemura† and Hideaki Sugawara (†SOKEN-DAI)

Metagenomic approach, which is the genome analysis on a mixture of uncultured microorganisms, has been recently developed to search for novel and industrially useful genes and to study microbial diversity in a wide variety of environments. We previously modified the conventional SOM for genome informatics to make the learning process and resulting map independent of the order of data input[5), 6)]. In the present study, we developed the SOM as a novel bioinformatics strategy to capture and visualize microbial diversity and relative abundance of microorganisms within an environmental sample. First we constructed SOMs of tri- and tetranucleotide frequencies in 1- and 5-kb sequence fragments from prokaryotic genomes for which complete sequence is available. The sequences could be classified primarily according to species and to 11 major phylogenetic groups without information regarding the species. For example, 88% of 5-kb sequences were classified into the correct phylogenetic group. Importantly, the classification could be done without orthologous sequence sets, and therefore, SOM was especially useful to analyze novel sequences from poorly characterized species for industrial applications and scientific studies. With the SOM method, all non-rRNA sequences in the Database that were from unidentified or uncultured bacteria and longer than 1 kb were classified into major phylogenetic groups[7)]. The present method can also be developed as a tool for surveys of pathogenic microorganisms in environmental and clinical samples that can not be cultured easily and in sterilized samples.

PUBLICATIONS

Papers

1. Sugawara, H., Abe, T., Tanaka, N. and Miyazaki, S. (2004). Encounter of microbiology with the data science in the phase called post-genome sequencing. Soil microorganisms. *58 (2)*, 57-67.

2. Miyazaki, S., Sugawara, H., Ikeo, K., Gojobori, T. and Tateno, Y. (2004). DDBJ in the stream of various biological data. Nucleic Acids Research. *32*, D31-D34.

3. Imanishi, T., Itoh, T., Suzuki, Y., O'Donovan, C., Fukuchi, S., Koyanagi, K., Barrero, R., Tamura, T., Yamaguchi-Kabata, Y., Tanino, M., Yura, K., Miyazaki, S., Ikeo, K., Homma, K., Kasprzyk, A., Nishikawa, T., Hirakawa, M., Thierry-Mieg, J., Thierry-Mieg, D., Ashurst, J., Jia, L., Nakao, M., Thomas, M., Mulder, N., Karavidopoulou, Y., Jin, L., Kim, S., Yasuda, T., Lenhard, B., Eveno, E., Suzuki, Y., Yamasaki, C., Takeda, J., Gough, C., Hilton, P., Fujii, Y., Sakai, H., Tanaka, S., Amid, C., Bellgard, M., Bonaldo, Mde, F., Bono, H., Bromberg, S., Brookes, A., Bruford, E., Carninci, P., Chelala, C., Couillault, C., De Souza, SJ., Debily, M., Devignes, M., Dubchak, I., Endo, T., Estreicher, A., Eyras, E., Fukami-Kobayashi, K., Gopinath, G.., Graudens, E., Hahn, Y., Han, M., Han, Z., Hanada, K., Hanaoka, H., Harada, E., Hashimoto, K., Hinz, U., Hirai, M., Hishiki, T., Hopkinson, I., Imbeaud, S., Inoko, H., Kanapin, A., Kaneko, Y., Kasukawa, T., Kelso, J., Kersey, P., Kikuno, R., Kimura, K., Korn, B., Kuryshev, V., Makalowska, I., Makino, T., Mano, S., Mariage-Samson, R., Mashima, J., Matsuda, H., Mewes, H., Minoshima, S., Nagai, K., Nagasaki, H., Nagata, N., Nigam, R., Ogasawara, O., Ohara, O., Ohtsubo, M., Okada, N., Okido, T., Oota, S., Ota, M., Ota, T., Otsuki, T., Piatier-Tonneau, D., Poustka, A., Ren, S., Saitou, N., Sakai, K., Sakamoto, S., Sakate, R., Schupp, I., Servant, F., Sherry, S., Shiba, R., Shimizu, N., Shimoyama, M., Simpson, AJ., Soares, B., Steward, C., Suwa, M., Suzuki, M., Takahashi, A., Tamiya, G., Tanaka, H., Taylor, T., Terwilliger, J., Unneberg, P., Veeramachaneni, V., Watanabe, S., Wilming, L., Yasuda, N., Yoo, H., Stodolsky, M., Makalowski, W., Go, M., Nakai, K., Takagi, T., Kanehisa, M., Sakaki, Y., Quackenbush, J., Okazaki, Y., Hayashizaki, Y., Hide, W., Chakraborty, R., Nishikawa, K., Sugawara, H., Tateno, Y., Chen, Z., Oishi, M., Tonellato, P., Apweiler, R., Okubo, K., Wagner, L., Wiemann, S., Strausberg, R., Isogai, T., Auffray, C., Nomura, N., Gojobori, T. and Sugano, S. (2004). Integrative annotation of 21,037 human genes validated by full-length cDNA clones. PLoS Biol., *2 (6)*, e162.

4. Vastermark, A., Shigemoto, Y., Abe, T. and Sugawara, H. (2004). Splicing Profile-based Protein Categorization between Human and Mouse Genome by use of DDBJ Web Services. Genome Informatics *15*, 13-20.

5. Abe, T., Kanaya, S., Kinouchi, M. and Ikemura, T. (2004). Genome Informatics for Unveiling Hidden Genome Signatures. Proceedings of the Institute of Statistical Mathematics *52*, 207-215.

6. Abe, T., Kanaya, S., Kinouchi, M., Kosaka, Y. and Ikemura, T. (2004). Novel bioinformatics for unveiling hidden characteristics in genome sequences and searching in silico for genetic signal sequences. Proceeding of The 8th World Multi-Conference on Systemics, Cybernetics and Informatics.

7. Abe, T., Ikemura, T., Kanaya, S., Kinouchi, M. and Sugawara, H. (2004). Novel genome informatics for unveiling hidden signatures in genome sequences: self-organizing map (SOM) of oligonucleotide frequencies. Proceedings of Information-Based Induction Sciences, 94-99.

Books

8. Sugawara, H. (2004). Tsunami of data: Data resources and utilization. Microbial Genetic Resources and Biodiscovery. Kurtboke, I. and Swings, J. ed., (National Library of Australia), 40-56.

Databases

9. Japanese Bio-portal site (Jabion), http://www.bioportal.jp/

10. Genome Information Broker, http://gib.genes.nig.ac.jp/

11. WFCC-MIRCEN World Data Centre for Microorganisms (WDCM), http://www.wdcm.org/

12. The portal site for pathogenic microorganisms, http://www.wdcm.org/byogen/

13. e-Workbench for Biological Classification and Identification, http://lilium.genes.nig.ac.jp/index_e.html

14. H-Invitational Database, http://www.h-invitational.jp/

EDUCATION

1. Dr. H. Sugawara was invited to give a lecture on "Databases are the key to bioinformatics" at the 2nd Open Symposium of Joint Research with Wakayama Pref., Wakayama, 2004 (in Japanese).

2. Dr. H. Sugawara was invited to give a lecture on "Invitation to information biology" at Campus system Research Group of Private Universities, Hamanako, August 2004 (in Japanese).

3. Dr. T. Abe was invited to give a lecture on "Genome analysis by PC-cluster." at Working group of scientific system, Tokyo, August, 2004 (in Japanese).

学会活動

1. Dr. H. Sugawara organized the 2nd International Conference on Biodata Interoperability, Tokyo, June, 2004.

2. Dr. H. Sugawara organized International program committee of the International Congress for Culture Collections, Tsukuba, October, 2004.

3. Dr, H. Sugawara organized Program committee of Genome Informatics 2004, Yokohama, December, 2004.

4. Task Force of Biological Resource Centers, OECD Working Party for Biotechnology (Vice-chair).

5. World Federation for Culture Collections, Executive board member and journal editor.

6. 極限微生物学会(評議員)

7. 日本情報知識学会(理事)

I-e. Laboratory for Gene-Expression Analysis Kousaku Okubo Group

RESEARCH ACTIVITIES

(1) Expression profiling of human genes

(1a) Data integration (BodyMapII): "Do you know how much of our genes have reliable expression patterns on the net?"

Osamu Ogasawara and Kousaku Okubo

Unexpectedly small difference in gene numbers among multicellular organisms, precisely determined though whole genome sequence supports the idea that the complexity in our body is evolutionary achieved by sophistication in expression controls of genes. The anatomically comprehensive genome-wide gene expression profile is key data to appreciate such sophistication coded in our genome. Moreover, availability of such data opened up opportunities to explore the dependence of constitutive expression patterns on other features of genes and genomes, which may eventually leads to the understanding of coding principles in our genome.

Despite the frequent use of the term 'genome-wide profiling' and wealth of expression data in the public domain, it is still not explicit what fraction of our genes are provided with anatomical expression patterns (COVERAGE) and to what extent different data set

agrees in terms of tissue distribution (ACCURACY). In order for rational design of studies of human transcriptome as a whole, we started to integrate data from multiple different platforms on the framework of latest human genome. The preliminary data is opened to the public in collaboration with integrated database team at JBIRC. (https://www.jbirc.aist.go.jp/hinv/h-angel/)

(1b) Data generation and analysis; "Are you satisfied by the present resolution of anatomical expression data?"

Makiko Otsuji and Koichi Itoh

In order to functionally relate genes in co-expression cluster or to deduce promoter sequences through alignment of upstream regions of them, the resolution of expression pattern should be at the level of cells or homogeneous cell populations. The vast majority of the public data do not meet this criterion, except for those from induction experiments with cell lines. Moreover, the majority of the target genes for drug development such as receptors for signaling molecules and channels localize in the minority cells in the complex organs rather than in homogeneous cell population. We are taking several different approaches to generate such data in streamline in the organs with complex cell population such as brain and kidney.

(2) Knowledge encoding and computation with gene functions (BOB): "Are you confident in your massive data interpretation?"

Kousaku Okubo, Koichi Itoh and Osamu Ogasawara

With the advent in high throughput genome-wide measurement, hypothesis generation on gene functions by systematic and integrative interpretation of the accumulating data is anticipated. For the last few years, various statistical analysis techniques have been employed in extraction of global patterns from the massive data, in the forms of gene clusters and networks. However, at present, even with intensive use of web-based knowledge bases, human interpretation will not match machine-aided data production in neither speed nor scope. Automation of interpretation process, at least in part, appears essential for systematic and efficient hypothesis generation.

There are three steps in automation of interpretation process; (1) encode biomedical knowledge into computable form, (2) interpret data using encoded knowledge, and (3) represent the interpretation results. The pioneering approaches in the step (1) were all declarations (KEGG, GO). In general, some inherent limitations are known in declarative approach. First of all, continuous revising and updating efforts by experts are inevitable. Secondly, the manually declared structure has usually low dimension and representation power is poor. For example GO provides only three 'aspects' in the representation of functional relations in genes. And thirdly, expert's declaration is apt to be inconsistent when target domain becomes wider. Although such limitation may causes little problem in biochemistry and cell biology, they become serious in the field of medical biology where gene functions have been described by relating their roles with so many types of concepts from behavior to chemical compounds. These limitations make the effectiveness of declaration approach in the medical problems still elusive.

Rationale for our approach: A major process in biomedical interpretation of these data is to determine whether if there is a unique and common functional feature, within any part of data-driven structures. For this purpose, experts recall features of genes from every aspect. Such a process appears so dependent on expert's flexible recalling and thinking that machine cannot possibly do. In medical biology, however, we assumed that each aspect mostly represents a certain topic in biomedicine, such as 'pathogenesis of a disease' and 'molecular mechanism of an organs function.' If there is a book covering essentially all fundamental topics in medical biology, and if each page contains a list of all relevant genes, interpretation would be achieved by computing fitness of a given gene cluster to each page (meaningfulness) and by returning the title of the page that fits best as a meaning of the cluster.

We aimed to create such a book, 'BOB (Biomedical OminiBook)', by concept-based structuring of pages in biomedical textbooks and description of gene functions. We applied a concept based indexing technique, latent semantic indexing (LSI), for this purpose. LSI was developed to overcome the problems in term-matching-based-document-searching caused by the fact that similar concept may represented by texts with different term combinations and vice versa. In LSI, terms and documents are structured in a

'semantic' vector space, based on the global patterns in term-document association data, beforehand. In searching, users' queries are mapped onto this space as pseudo-documents. For each queries, the document vectors having supra-threshold cosine value to the query vector are returned.

In our application, we took textbooks in place of documents to be searched. Started from term-page association data, provided in the index sections of textbooks, we created a high-dimensional vector space wherein objects (term or page) are placed according to their relevance. Then we prepared a gene vector, that corresponds to a 'query vector' in original LSI, by counting occurrence of textbook terms in corresponding molecular database entry including abstracts for cited papers. Thus prepared gene vectors were mapped onto the textbook space similarly to the mapping of users queries. Resulted space contains vector representation of three different classes of objects; term meanings, page contents, and gene functions, arranged according to their conceptual relevance. High dimensionality of this space is expected to allow discriminative representation of many aspects of relations without inconsistency. Using this space, we may scale the functional relevance of any pairs of genes, with which 'meaningfulness' of any given gene cluster can be calculated. In addition, the title for the page-vector nearest to the center of a gene cluster will explain the biomedical 'meaning' of the meaningful cluster.

Publications

Papers

1. Michibata, H., Chiba, H., Wakimoto, K., Seishima, M., Kawasaki, S., Okubo, K., Mitsui, H., Torii, H. and Imai, Y. (2004). Identification and characterization of a novel component of the cornified envelope, cornifelin. Biochem Biophys Res Commun., 318, 803-13.

2. Hishiki, T., Ogasawara, O., Tsuruoka, Y. and Okubo, K. (2004). Indexing anatomical concepts to OMIM Clinical Synopsis using the UMLS Metathesaurus. In Silico Biol., 4, 31-54.

3. Chiba, H., Michibata, H., Wakimoto, K., Seishima, M., Kawasaki, S., Okubo, K., Mitsui, H., Torii, H. and Imai, Y. (2004). Cloning of a gene for a novel epithelium-specific cytosolic phospholipase A2, cPLA2delta, induced in psoriatic skin. J Biol Chem., 279, 12890-7.

Books

4. Kaimori, J., Takenaka, M. and Okubo, K. (2004). Quantification of Gene Expression in Mouse and Human Renal Proximal Tubules., 「Laser Capture Microdissection Methods and Protocols」(Grame I. Murray and Stephanie Curran ed.) Methods in Molecular Biology 293, Humana Press, 209-220.

5. 大久保公策，川本祥子(2004)「ゲノム語とオントロジーという名の形而上学—進化と医学」,科学　Vol.74 No.10, 1254-1257.

EDUCATION

1. Dr. Okubo gave a seminar on "The computational interpretation of biological data and its significance" at National Institute of Genetics, Oct., 2004 (in Japanese).

2. Dr. Okubo gave a seminar on "How can we (biologists) make good use of biological knowledge available?" for DDBJing and TERAKOYA, Tokyo, Mar., 2004 (in Japanese).

3. Dr. Okubo was invited to give a seminar on " Medical Science in the post-genome period- knowledge engineering in medicine" for Department of Biomolecular Sciences Saga Medical School Faculty of Medicine at Saga University, Mar., 2004 (in Japanese).

SOCIAL CONTRIBUTIONS AND OTHERS

1. BodyMap: http://bodymap.genes.nig.ac.jp/
2. Databese: http://www.jbirc.aist.go.jp/hinv/h-angel/
3. Databese: http://bodymap.ims.u-tokyo.ac.jp
4. JST領域探索プログラム「ゲノムと言語」コオーガナイザー
5. 日本学術会議　ゲノム科学研究連絡委員会　委員
6. (社)バイオ産業情報化コンソーシアム　タンパク質機能解析・活用pj研究推進委員会　委員

J. RADIOISOTOPE CENTER

RESEARCH ACTIVITIES

(1) Chromosome partitioning mechanism of *Escherichia coli*; Involvement of a cisacting site on the chromosome

Yoshiharu Yamaichi and Hironori Niki

During replication of the *Escherichia coli* chromosome, the replicated Ori domains migrate towards opposite cell poles, suggesting that a cis-acting site for bipolar migration is located in this region. To identify this cis-acting site, a series of mutants was constructed by splitting subchromosomes from the original chromosome. One mutant, containing a 720 kb subchromosome, was found to be defective in the bipolar positioning of *oriC*. The creation of deletion mutants allowed the identification of *migS*, a 25 bp sequence, as the cis-acting site for the bipolar positioning of *oriC*. When migS was located at the replication terminus, the chromosomal segment showed bipolar positioning. *migS* was able to rescue bipolar migration of plasmid DNA containing a mutation in the SopABC partitioning system. Interestingly, multiple copies of the *migS* sequence on a plasmid in trans inhibited the bipolar positioning of *oriC*. Taken together, these findings indicate that *migS* plays a crucial role in the bipolar positioning of *oriC*. In addition, real-time analysis of the dynamic morphological changes of nucleoids in wild-type and *migS* mutants suggests that bipolar positioning of the replicated *oriC* contributes to nucleoid organization.

(2) Partitioning mechanism of F plasmid; elucidation of motive force for bipolar migration of the plasmid DNA

Toshiyuki Hatano and Hironori Niki

In a bacterial cell, replicated chromosomal segments including the replication origin and the F plasmid actively move from midcell toward cell poles and are located at cell quarter positions. The Sop ABC partitioning system of F plasmid has been characterized genetically and biochemically to elucidate molecular mechanism of DNA partitioning. The SopA protein is a hypothetical motor protein because of its walker-type ATPase activity. Moreover, some proteins homologous to SopA change subcellular location periodically. We have constructed an active SopA-YFP and observed oscillation of YFP fluorescence from one cell pole to the other with an interval of a few minutes. Simultaneously migration of the plasmid DNA was detected using fluorescent labeling technique with lacO array and LacI-CFP. The analyses of movement of both the fluorescence suggest that SopA should not be the motor protein to move the plasmid DNA. Another new possibility is that the SopA protein could govern the direction of movement of the plasmid DNA.

(3) Subcellular localization of *Escherichia coli* proteins in living cells

Yasuyuki Ogata, Toshiyuki Hatano and Hironori Niki

Bacterial cells are too small to see details of their structure using a light microscope because of the ultimate limit to the resolution of the microscope set by the wavelength of visible light. In recent years, the development of methods for the specific labeling and imaging of individual cell components and the reconstruction of their three-demensional architecture allows us to understand the structural organization of bacterial cells. Cytoskeletal filaments that bacteria were long thought to lack were found by means of these methods and visualization of DNA polymerase of the living *Bacillus subtilis* cells by tagging the catalytic subunit with fluorescent protein supported the model in which the polymerase is stationary like a factory and DNA is pulled through.

Escherichia coli genome has been fully sequenced and the genes of unknown function revealed. To estimate their function, we used a plasmid library, archive clone library, that expresses fusion genes containing open reading flame (ORF) of *E. coli* W3110 strain and green fluorescent protein (GFP) gene and we observed their subcellular localizations in the living *E. coli* cells. These proteins were classified into five categories: (i) at cell periphery, (ii) in whole cell, (iii) spot, (iv) filament, and (v) on nucleoid.

We subsequently investigated the localizations of

the proteins of known function and found the correlations between subcellular localization and protein function. Most of the cell peripheral proteins have either trans membrane elements or signal peptides. Proteins with spots at cell quarter positions like a replication factory include replication machineries and recombination enzymes and proteins with several spots participate in transcription. Some proteins with a ring at mid cell are involved in cell division. Nucleoid proteins contain nucleoid constituents such as histone-like proteins and various enzymes involved in DNA-related events including DNA supercoiling, recombination, and repair.

We further replaced stop codon of these genes with *gfp* gene by the method of Warner and confirmed subcellular localization of these proteins.

(4) Stationary phase-induced illegitimate recombination in *Escherichia coli*

Yasuyuki Ogata and Hideo Ikeda

In nature, bacterial cells are usually in stationary stages, in which the cells alter their genetic inheritance. Adaptive mutations are induced when *Escherichia coli* cells are exposed to conditions of stationary-phase starvation, resulting in a variety of genetic changes, some of which allow the cells to survive under such conditions. Gene amplification of the *lac* operon of *E. coli* also occurs during stationary-phase starvation. Whether or not stationary-phase starvation induces chromosomal rearrangements has remained obscure.

We found that illegitimate recombination is induced when growth phase is changed from log to stationary phase. Illegitimate recombination occurred more frequently in early-stationary phase than in mid-log phase during formation of λ *bio*-transducing phage. Moreover, we found that illegitimate recombination increased in the *tag alkA1* double mutant in early-stationary phase, thereby implying that the alkylated lesion may be responsible for the stationary phase-induced illegitimate recombination.

When we did a quantitative analysis of deletion formation using a miniF-based plasmid, we found that the stationary phase-induced illegitimate recombination can be also detected in this assay system. In addition, we indicated that the frequency of deletion of the *tag alkA1* double mutant was higher than that of its isogenic wild-type strain in early-stationary phase.

Since illegitimate recombination results in chromosomal rearrangements including deletion, duplication, insertion, or translocation, it seems likely that stationary phase may induce chromosomal rearrangements. Thus, when bacterial cells confront unfavorable conditions for their growth, the cells seem to increase diverse mutations as if they struggle for adaptation to the environment. If the situation takes a favorable turn, a suitable mutant clone is likely to grow preferentially, which may lead to evolution.

PUBLICATIONS

Papers

1. Yamaichi, Y. and Niki, H. (2004). *migS*, a cis-acting site that affects bipolar positioning of oriC on the Escherichia coli chromosome. *EMBO J. 23*, 221-233.
2. Ishida, T., Akimitsu, N., Kashioka, T., Hatano, M., Kubota, T., Ogata, Y., Sekimizu, K. and Katayama, T. (2004). DiaA, a novel DnaA-binding protein, ensures the timely initiation of *E. coli* chromosome replication. *J. Biol. Chem., 279*, 45546-45555.
3. Ikeda, H., Shiraishi, K. and Ogata, Y. (2004). Illegitimate recombination mediated by double-strand break and end-joining in *Escherichia coli. Adv. Biophys., 38*, 3-20.

Reviews

4. 仁木宏典(2004)「原核生物の染色体の分配起点」蛋白質，核酸，酵素*49*, 2017-2023.

Books

5. 仁木宏典，ゲノミクス・プロテオミクスの新展開 生物情報の解析と応用，エヌ・ティ・エス，47-56.

K. EXPERIMENTAL FARM

RESEARCH ACTIVITIES

Experimaental Farm is responsible for preparation of all kinds of rice strains necessary for the studies of genetic resources and of functional genomics. All works in the Experimental Farm have been carried out as collaborative works with the Plant Genetics Lab. For details, see the reports of plant genetics lab.

(1) Examination of genetic and phenotypic nature of newly prepared core collections for wild rice comparative genomics

Toshie Miyabayashi, Mitsugu Eiguchi, Ken-Ichi Nonomura and Nori Kurata

For making rice genetic stocks good resources, we developed core collection of wild rice. Out of 2,000 accessions composed of twenty-one wild rice species in nine genomes, several accessions from each species were selected to prepare core collection. About 300 selected lines had already sub-grouped to Rank1, Rank2 and Rank3 categories, and a part of them were characterized for their phenotypes to record them in the Oryzabase, a rice comprehensive database.

Comparative genomics among nine rice genomes will reveal important aspects on evolution and genetic diversity in rice. For finding out genome specific genes or far-related genes different from AA, comparison of BB and CC genome species with the cultivated AA species have been started. Searches for structural and functional diversity between genes of different genomes are performed by cDNA clone sequencing and by microarray expression profile analysis. Some details obtained in 2004 were presented in the report of Plant Genet. Lab.

In addition, to estimate genome sizes of each genome and species, strains in the core collection were examined for their DNA content by the simple flow-cytometric analysis. Differences of DNA content among genomes and among species would give some relationships to the molecular nature of genetic diversity.

PUBLICATIONS

Papers

1. Miyoshi, K., Ahn, B-O., Kawakatsu, T., Ito, Y., Itoh, J-I., Nagato, Y. and Kurata, N. (2004). *PLASTOCHRON1*, a timekeeper of leaf initiation in rice, encodes cytochrome P450. Proc. Natl. Acad. Sci. USA *101*, 875-880.

2. Ito, Y., Chujo, A., Eiguchi, M. and Kurata, N. (2004). Radial axis differentiation in a globular embryo is marked by *HAZ*, a PHD-finger homeobox gene of rice. Gene *331*, 9-15.

3. Nonomura, K.I., Nakano, M., Murata, K., Miyoshi, K., Eiguchi, M., Miyao, A., Hirochika, H. and Kurata, N. (2004). The insertional mutation of rice *PAIR2* gene, the ortholog of Arabidopsis *ASY1*, caused a defect in homologous chromosome pairing in meiosis. Mol. Genet. Genomics *271*, 121-129.

4. Nonomura, K.I., Nakano, M., Fukuda, T., Eiguchi, M., Miyao, A., Hirochika, H. and Kurata, N. (2004). The novel gene *HOMOLOGOUS PAIRING ABERRATION IN RICE MEIOSIS 1* of rice encodes a putative coiled-coil protein required for homologous chromosome pairing in meiosis. Plant Cell *16*, 1008-1020.

5. Salina, E.A., Adonina, I., Vatolina, T. and Kurata, N. (2004). A comparative analysis of the composition and organization of two subtelomeric repeat families of *Aegilops speltoides* Tausch.and related species. Genetica *122*, 227-237.

6. Moriguchi, K., Suzuki, T., Ito, Y., Yamazaki, Y., Niwa, Y. and Kurata, N. (2005). Functional isolation of novel nuclear proteins showing a variety of sub-nuclear localizations. Plant Cell, *17*, 389-403.

Reviews

7. Itoh, J.I., Nonomura, K.I., Ikeda, K., Yamaki, S., Inukai, Y., Yamagishi, H., Kitano, H. and Nagato, Y. (2005). Rice plant development: from zygote to spikelet. Plant Cell Physiol. 46, 23-47.

8. Kurata, N., Miyoshi, K., Nonomura, K.I., Yamazaki, Y. and Ito Y. (2005). Rice mutants and genes related to organ development, morphogenesis and physiological traits. Plant Cell Physiol. 46, 48-62.

L. TECHNICAL SECTION

RESEARCH ACTIVITIES

The Technical section supports the research activity of NIG in many fields. For example, we support the Radioisotope center and Experimental farm, and the Genetic strain research center where genetic and transgenic strains of mice, rice, flies, and fishes are produced, maintained, and distributed to other institutions. Members of the staff additionally support many types of experiments in various laboratories, such as developmental, cell, and molecular genetics. We continually update our technical services through attending the training and annual meetings for the technological staffs.

PUBLICATIONS

Papers

1. Kakutani, T., Kato, M., Kinoshita, T. and Miura, A. (2005). Control of Development and Transposon Movement by DNA Methylation in *Arabidopsis thaliana*. Cold Spring Harbor Symposia on Quantitative Biology *69* in press.

2. Kinoshita, T., Miura, A., Choi, Y., Kinoshita, Y., Cao, X., Jacobsen, S.E., Fischer, R.L. and Kakutani, T. (2004). One-way control of *FWA* imprinting in *Arabidopsis* endosperm by DNA methylation. Science *303*, 521-523.

3. Kubota, T., Furuumi, H., Kamoda, T., Iwasaki, N., Tobita, N., Fujiwara, N., Goto, Y., Matsui, A., Sasaki, H. and Kajii, T. (2004). ICF syndrome in a girl with DNA hypomethylation but without detectable DNMT3B mutation. Am. J. Med. Genet. *129A*, 290-293.

4. Miura, A., Kato, M., Watanabe, K., Kawabe, A., Kotani, H. and Kakutani, T. (2004). Genomic localization of endogenous mobile CACTA family transposons in natural variants of *Arabidopsis thaliana*. MGG *270*, 524-532.

Books

5. 古海弘康，佐々木裕之(2004)シリーズ最新医学講座I・転写因子12：エピジェネティクス制御とその異常. 臨床検査Vol.48 No.13, 1673-1679.

6. 古海弘康，佐々木裕之(2004)哺乳類のメンデル遺伝するエピジェネティクス. 佐々木裕之編：エピジェネティクス. Springer Reviewsシリーズ, シュプリンガー・フェアラーク東京，129-134.

BIOLOGICAL SYMPOSIUM 2004

Jan.9 初期遺伝子・神経遺伝子・老化遺伝子：その起源と多様性（Nozomu Mori）
Jan.13 DNA Sequencing: A Tool for Biomedical Research（Shih-Feng Tsai）
Jan.19 染色体分配，これから何を研究する？（Mitsuhiro Yanagida）
Jan.21 The Master Regulator for Entry into Sporulation in Bacillus subtilis becomes a Cell-Specific Transcription Factor after Asymmetric Division（Masaya Fujita）
Jan.28 Mammalian micro RNA genes and their downstream targets（Roberto Barrero）
Feb.3 Sequencing the Rice Genome − Where do we go from here?（Rod, A. Wing）
Feb.26 Chemistry of Methylation: from DNA to Histone（Xiaodong Cheng）
Feb.26 脂肪細胞の形質転換と生活習慣病（Takashi Kadowaki）
Feb.27 Real time footprinting techniques reveal the mechanism of promoter recognition by E. coli RNA polymerases（Malcolm Buckle）
Mar.3 巨核球・血小板の特異的形態形成機構とその医療への応用（Yuka Nagata）
Mar.15 大腸菌多剤排出膜輸送体のX線結晶構造解析（Satoshi Muarakami）
Mar.19 Cooperation between RNA polymerase molecules in transcription elongation（Evgeny Nudler）
Mar.23 SEEDLINGS, CHICKS AND CHILDREN: A DARWINIAN PEDIGREE FOR EUGENICS (James Moore)
Mar.24 Molecular genetic basis of the ABO system & expression profiling of glycosyltransfer ase genes（Fumiichiro Yamamoto）
Mar.24 次世代遺伝子発現プロフィール解析技術：HiCEP（Masumi Abe）
Mar.25 Implications of Small Molecule Activator and Inhibitors of Histone Acetyltransferase p300 in Chromatin Therapy（Tapas K. Kundu）
Apr.8 Meiotic chromosome morphogenesis and function in C. elegans（Kentaro Nabeshima）
Apr.8 Regulation of spindle orientation by apc tumor suppressor and centrosome in asymmetric stem cell division in Drosophila（Yukiko Yamashita）
Apr.21 ショウジョウバエの膜脂質と温度応答性：温度嗜好性変異体の樹立と解析（Shingo Umeda）
Apr.30 ショウジョウバエ神経幹細胞の非対称分裂におけるヘテロ3量体Gタンパク質の役割（Naoyuki Fuse）
May.12 Licensing mitotic entry: mechanisms of concerted phosphorylation and down-regulation of Swe1 by multiple kinases（Satoshi Asano）
May.13 Drosophila male GSCs（germline stem cells）における BMP シグナルの役割（Eihachiro Kawase）
May.18 Follicle-Stimulating Hormone（FSH） and Testis Functions in Fish（Rudiger W. Schulz）
May.24 ゲノムタイリングアレイを用いたゲノム動態解析（Katsuhiko Shirohige）
May.31 FGF signaling during mesoderm migration in the Drosophila gastrula（Arno Mueller）
Jun.3 大腸菌表層ストレス応答に関わる膜プロテアーゼ YaeL の機能とその調節（Yoshinori Akiyama）
Jul.5 エピジェネティックな不均等性とゲノムインプリンティング（Kozo Mitsuya）
Jul.7 植物の茎が重力方向に応じて運動するメカニズム（Masao Tasaka）
Jul.14 複雑数理モデリングとその生命情報システムへの応用（Kazuyuki Aihara）
Jul.15 Navigating Massive Data Sets via Local Clustering（Michae E. Houle）
Jul.27 BRU1, a novel link between DNA damage response and epigenetic regulation in Arabidopsis（Shin Takeda）
Aug.9 ヒトおよびメダカにおける複合形質の遺伝学的解析（Minori Shinya）
Aug.9 眼の特異化と形態形成におけるゼブラフィッシュ Vax 転写因子の機能（Masaya Takeuchi）
Aug.11 自殖生物シロイヌナズナができるまで〜ダーウィン仮説への分子集団遺伝学からの解析〜(Kentaro Shimizu)
Aug.26 チンパンジー22番染色体との比較による21番染色体における正の淘汰領域の遷移確率行列の推定：ヒトとチンパンジーを分け隔てたもの（Satoshi Ota）
Aug.30 The origin and evolution of vertebrate mineralized skeleton（Kazuhiko Kawasaki）
Sep.1 アデノウィルスベクターによる神経細胞の誕生日特異的遺伝子導入法を用いた，小脳形成機構の解析（Mithsuhiro Hashimoto）
Sep.1 In vivo imaging of Xenopus embryo early stages development using Quantum Dots encapsulated

in phospholipid micelles (Benoit Dubertret)

Sep.13 Tsukushi functions as a novel organizer inducer by inhibition of BMP activity in cooperation with chordin (Kunimasa Ota)

Sep.16 脊椎動物頭部形態の進化 (Shigeru Kutaratani)

Sep.17 ダイス (ダイズ) はなぜ黄色いか？一種皮で自発的に誘導されるカルコンシンターゼ遺伝子の PTGS について (Mineo Senda)

Sep.21 小細体品質管理の分子機構 (Kazuhiro Nagata)

Sep.24 Dictyostelium and its genome project (Robert Kay)

Sep.27 ユビキチン化とエンドサイトーシスによる膜貫通型レセプター Notch の制御 (Shigeo Hayashi)

Sep.30 CDK 依存的 BRCA2 リン酸化による組み換え修復制御 (Fumiko Esashi)

Oct.1 Tbx 遺伝子群の進化と形態形成メカニズムの進化 (Toshihiko Ogura)

Oct.7 ショウジョウバエにおける生殖質アッセンブリーと生殖細胞形成機構 (Akira Nakamura)

Oct.21 G9a / GLP 複合体による in vivo ヒストンメチル化制御 (Makoto Tachibana)

Nov.5 Specification and assembly of locomotion in Drosophila (Michael Bate)

Nov.15 嗅索ガイダンスにおけるホスファチジルイノシトール４－リン酸５－キナーゼの役割 (Masakazu Yamazaki)

Nov.17 Looking for the Genetic Origins of the Human Capacity for Language (Karl C. Diller)

Nov.22 軸索誘導遺伝子 lola が多様な遺伝子産物を生み出すメカニズム (Takayuki Horiuchi)

Nov.22 Geminin-Cdt1 balance as a guardian against genomic instability (Anindya Dutta)

Nov.24 セプター型チロシンキナーゼ PVR とその下流シグナル伝達系による JNK の活性化機構 (Satoshi Ishimaru)

Nov.24 ショウジョウバエを利用した神経回路網形成の遺伝学的解析 (Mitsuhiko Kurusu)

Nov.25 ゲノム解析から見た生殖細胞の特徴 (Masami Nozaki)

Nov.26 RNAi 経路における Argonaute 蛋白質の機能解析 (Katsutomo Okamura)

Dec.6 SPATIAL VARIATION IN GENE FREQUENCY (Aravinda Shakravarati)

Dec.6 Control of dendrite development (Yuh-Nung Jan)

Dec.7 Histone Modifications in Polycomb Silencing, X-inactivation and Cancer (Yi Zhang)

Dec.7 Regulation of Nucleosome Assembly during DNA Replication (Alain Verreault)

Dec.13 Bi-orienting chromosomes on the mitotic spindle (Tomoyuki U. Tanaka)

Dec.16 Splicing profile-based protein categorization (Ake Johannes Vastermark)

Dec.16 Cell cycle dependent acetylation of histone H3 lysine 56 contributes to double strand break repair in Saccharomyces cerevisiae (Hiroshi Masumoto)

Dec.20 THE HUMAN SGT1 – HSP90 COMPLEX IS REQUIRED FOR THE ASSEMBLY OF KINETOCHORE PROTEIN COMPLEXES (Katsumi Kitagawa)

Dec.20 The role of SMC1 in DNA damage induced signaling pathways (Risa Kitagawa)

FOREIGN VISITORS IN 2004

Jan.13	Shih-Feng Tsai	Division of Molecular & Genomic Medicine, National Health Research Institutes, Taipei 115, TAIWAN
Jan.21	Masaya Fujita	Department of Molecular and Cellular Biology Harvard University
Jan.28	Roberto Barrero	Japan Biological Information Research Center, JBIRC-Tokyo
Feb.3	Rod, A. Wing	Arizona Genomics Institute, Department of Plant Science, The University of Arizona
Feb.26	Xiaodong Cheng	Biochemistry Department Emory University School of Medicine Atlanta, GA, USA
Feb.27	Malcolm Buckle	Enzymologie et cinétique structurale, LBPA, UMR 8113 du CNRS, Ecole Normale Supérieure Cachan cedex France
Mar.19	Evgeny Nudler	Department of Biochemistry, New York University School of Medicine, New York, USA
Mar.23	James Moore	The Open University UK
Mar.24	Fumiichiro Yamamoto	The Burnham Institute, La Jolla, CA, USA
Mar.25	Tapas K. Kundu	Transcription and Disease Laboratory, Molecular Biology and Genetics Unit, Jawaharlal Nehru Center for Advanced Scientific Research, India.
Apr.8	Yukiko Yamashita	Stanford University
May.12	Satoshi Asano	National Institutes of Health
May.13	Eihachiro Kawase	Stowers Institute for Medical Research, Kansas City, USA
May.18	Rudiger W. Schulz	Dept Endocrinology, Fac Biology, University of Utrecht, The Netherlands
May.31	Arno Mueller	Institute for Genetics, Heinrich-Heine University, Dusseldorf
Jun.7	Masataka Okabe	MRC Centre for Developmental Neurobiology, King's College London
Jul.15	Michae E. Houle	National Institutes of Informatics Data mining (search, clustering, classification), design and analysis of algorithms, visualization, combinatorial geometry
Jul.27	Shin Takeda	Lab of Plant Genetics University of Geneva Switzerland
Aug.11	Kentaro Shimizu	North Carolina State University / Graduate School of Science, Kyoto University
Aug.30	Kazuhiko Kawasaki	Department of Anthropology, Pennsylvania State University
Sep.1	Benoit Dubertret	Laboratoire d'Optique Physique, ESPCI, Paris, France
Sep.24	Robert Kay	MRC Laboratory of Molecular Biology, Cambridge, England
Sep.30	Fumiko Esashi	Cancer Research UK, London Research Institute Clare Hall Laboratories
Nov.5	Michael Bate	Department of Zoology Cambridge UK and National Centre for Biological Sciences, Bangalore, India
Nov.11	Hans R. Bode	Dept. of Developmental and Cell Biology, Univ. of California, Irvine
Nov.17	Karl C. Diller	University of New Hampshire USA
Nov.21	David Liberles	Comptatioal biology unit, University of Bergen
Nov.22	Anindya Dutta	University of Virginia Health Sciences Center
Nov.24	Mitsuhiko Kurusu	California Institute of Technology Kai Zinn Laboratory
Nov.24	Shozo Yokoyama	Dept. of Biology Rollins Research Center Emory University
Dec.4	Antoine Blancher	Laboratoire d'immunologie, Hopitale Rangueil, Toulouse France
Dec.6	Aravinda Shakravarati	Henry J. Knott Professor & Director McKusick-Nathans Institute of Genetic Medicine Johns Hopkins University School of Medicine
Dec.6	Yuh-Nung Jan	University of California, San Francisco
Dec.7	Yi Zhang	University of North Carolina at Chapel Hill NC, USA
Dec.7	Alan Verreault	Cancer Research UK, London Research Institute, Clare Hall Laboratories

Dec.13	Tomoyuki U. Tanaka	School of Life Sciences, University of Dundee, Wellcome Trust Biocentre Dundee UK
Dec.15	Gerry Weinmaster	Biological Chemistry Department David Geffen School of Medicine at UCLA
Dec.16	Ake Johannes Vastermark	Bioinformatics Group, Department of Statistics, University of Oxford
Dec.16	Hiroshi Masumoto	Cancer Research UK, London Research Institute, Clare Hall Laboratories
Dec.20	Katsumi Kitagawa	Department of Molecular Pharmacology, St. Jude Children's Research Hospital, Memphis USA
Dec.20	Risa Kitagawa	Department of Hematology-Oncology St. Jude Children's Research Hospital, Memphis U.S.A

Annual Report of the National Institute of Genetics. No.55

Issued by Director-General : Yuji Kohara, D.Sc.
Research Organization of Information and Systems
National Institute of Genetics
Yata 1111 Mishima, Shizuoka-Ken, 411-8540 Japan
Tel. 055-981-6718
Fax. 055-981-6715
Edited by Ryu Ueda
Published by National Institute of Genetics