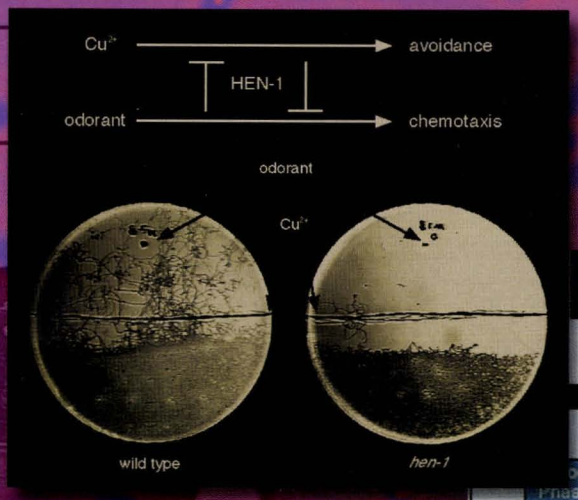
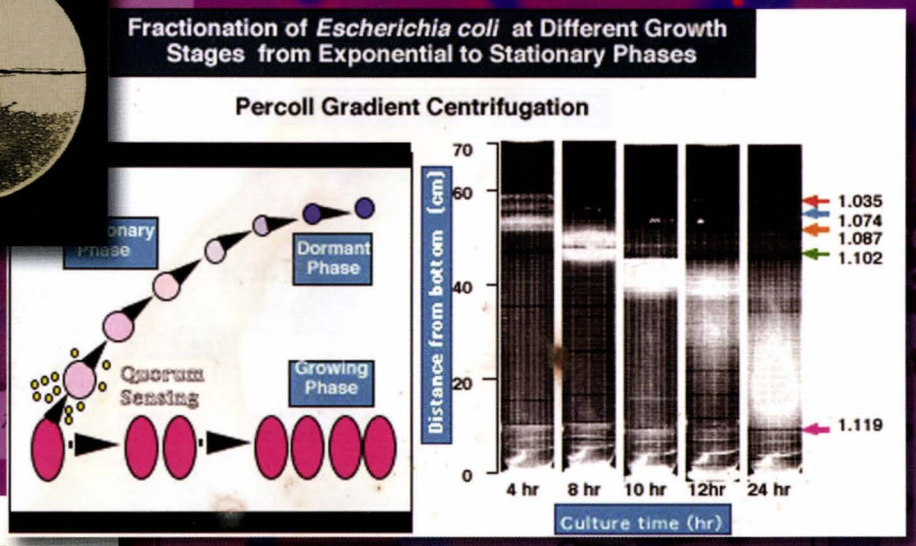
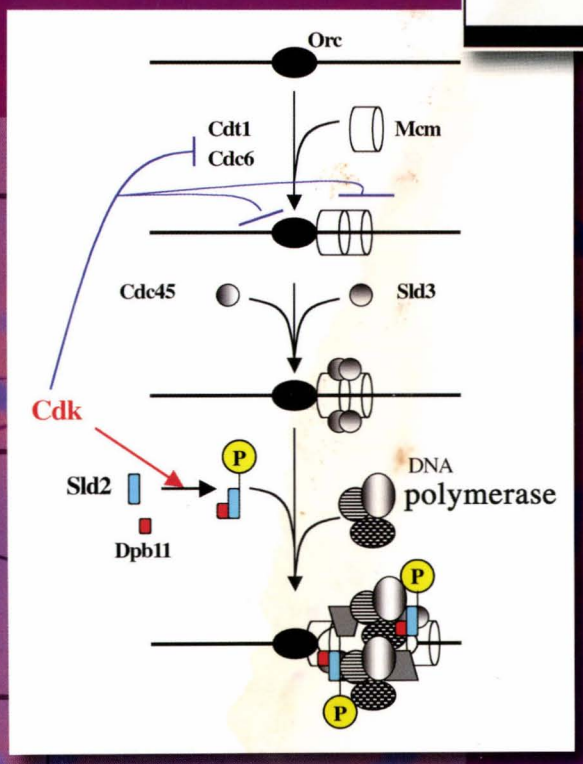
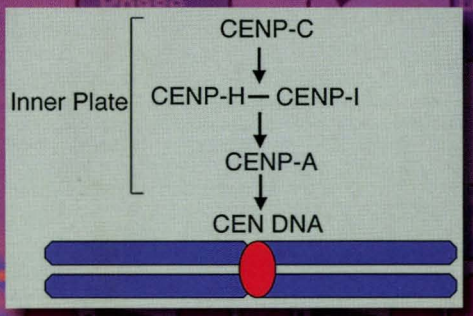
ANNUAL REPORT

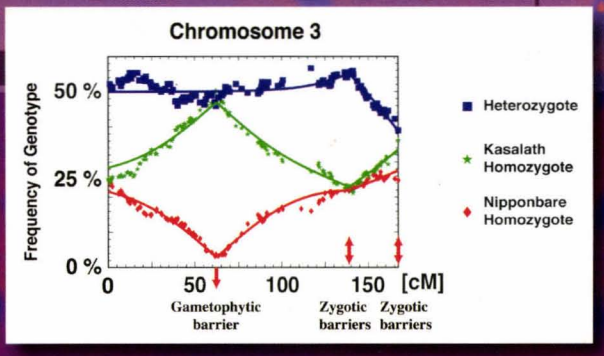
No.53
2002





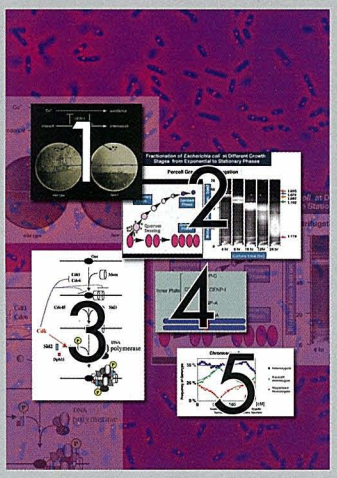






NATIONAL INSTITUTE OF GENETICS
http://www.nig.ac.jp

Highlights of 2002 (http://www.nig.ac.jp/hot/index.html)



1. **HEN-1, a secretory protein with an LDL receptor motif, regulates sensory integration and learning in C. elegans.**

HEN-1, a secretory protein with an LDL receptor motif, regulates sensory integration and learning in *C. elegans*
Ishihara,T., Iino,Y., Mohri,A., Mori,I., Gengyo-Ando,K., Mitani,S. Katsura,I.
Cell, *109*, 639-649 (2002)

2. **Discontinuous phenotype changes during bacterial differentiation**

Fractionation of E. coii Cell Populations of Different Stages during Growth Transition to Stationary Phase.Makinoshima, H., Nishimura, A. and Ishihama, A.
Mol. Microbiol. *43*, 2689-280 (2002)

3. **Identification of a substrate of S-phase-cyclin-dependent protein kinase**

S-Cdk-dependent phosphorylation of Sld2 essential for chromosomal DNA replication in budding yeast.Masumoto, H., Muramatsu, S., Kamimura, Y., and Araki, H.
Nature *415*, 651-655 (2002)

4. **CENP-I is essential for centromere function in vertebrate cells**

CENP-I is essential for centromere function in vertebrate cells.Ai Nishihashi, Tokuko Haraguchi, Yasushi Hiraoka, Toshimichi Ikemura, Vinciane Regnier, Helen Dodson, William C. Earnshaw, and Tatsuo Fukagawa
Dev.Cell, *2*, 463-476 (2002)

5. **A genome-wide survey of reproductive barriers**

A genome-wide survey of reproductive barriers in an intraspecific hybrid.Yoshiaki Harushima, Masahiro Nakagahra, Masahiro Yano, Takuji Sasaki and Nori Kurata.
Genetics, *159*, 883-892, (2001)

CONTENTS

Introduction

This is the 53rd annual report of the National Institute of Genetics (NIG). Japanese Government is deciding a reform plan to convert National Institutes as well as National Universities into a kind of "independent agencies". It is supposed to reduce governmental control and increase independence. We are, however, somewhat worrying about the reform, since less control may mean less public support in the future. In order to survive through the "political turmoil", it is more important than ever to enhance our scientific activities to receive world-wide recognition from both scientific and public communities.

The background of genetics is changing very rapidly through the increasing rate of genome sequencings. We must realize that the accumulation of the genome-wide knowledge about many biological species, combined with the rapid progress in computer and information science, will revolutionize genetics. NIG has an advantage in the "scientific turmoil", since we have wide spectra of research and database groups, and they are in close contact in every-day activity. NIG's mottoes are, "From Molecules to Individuals", "From Development to Evolution" and "From Experiment to Theory". We are supporting DDBJ (DNA Data Bank of Japan) as one of the three world-wide genome database centers. Another power of NIG is in the variety of studies with many model organisms. We have groups working with bacteria, yeast, nematode, Drosophila, zebra fish, medaka, mouse and man. Rice and Arabidopsis groups are also very active. Less popular but increasingly important organisms used here include planaria, hydra and chordate. The combinations of such a variety of scientific research will provide us an additional impetus in the post-genome sequence era.

We have a weak point, however. NIG has had no affiliated hospital, and has been strong in biology, but weak in medicine. Fortunately, Shizuoka Cancer Hospital has been open in the close location. It will add Cancer Research Center in a few years, and we have started to discuss about our future collaborations.

We welcome and appreciate any comments, advice and suggestions from everyone who reads the annual report.

Yoshiki Hotta, Director-General

STAFF (as of December 31, 2002)

Director-General

HOTTA, Yoshiki, D. Med.

Vice-Director

KOHARA, Yuji, D. Sc.

Members

1. Department of Molecular Genetics

SHIMAMOTO, Nobuo, D. Sc., Head of the Department

Division of Molecular Genetics

FUJITA, Nobuyuki, D. Sc.
MITSUZAWA, Hiroshi, D. Sc.
KIMURA, Makoto, D. Sc.

Division of Mutagenesis

YAMAO, Fumiaki, D. Sc., Associate Professor
KISHI, Tsutomu, D. Eng.
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
IMAI, Hirotami T., D. Sc., Associate Professor
UMEBAYASHI, Kyohei, D. Ag.

Division of Microbial Genetics

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

Division of Cytoplasmic Genetics

YOSHIKAWA, Takeo, D. Med., Adjunct Professor
YANAGIDA, Mitsuhiro, D. Sc., Professor (Adjunct)

3. Department of Developmental Genetics

HIROSE, Susumu, 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.

Division of Gene Expression

HIROSE, Susumu, D. Sc., Professor
UEDA, Hitoshi, D. Ag., Associate Professor
MINATO, Kiyoshi, M. Sc.

YAMADA, Masa-aki, D. Ag.

Division of Molecular and Developmental Biology

KAWAKAMI, Koichi, D. Sc., Associate Professor

Division of Physiological Genetics

KOSEKI, Haruhiko, D. Med., Professor (Adjunct)

KIYAMA, Ryoiti, D. Sci., Adjunct Professor

4. Department of Population Genetics

IKEMURA, Toshimichi, D. Sc., Head of the Department

Division of Population Genetics

SAITOU, Naruya, Ph. D., Professor

TAKANO, Toshiyuki, D. Sc., Associate Professor

Division of Evolutionary Gentics

IKEMURA, Toshimichi, D. Sc., Professor

FUKAGAWA, Tatsuo, D. Sc., Associate Professor (Adjunct)

Division of Theoretical Genetics

KONDO, Shigeru, Ph. D., D. Sc., Professor (Adjunct)

TAKAGI, Toshihisa, D. Eng., Professor (Adjunct)

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.

Division of Brain Fanction

HIRATA, Tatsumi, D. Med., Associate Professor

KAWASAKI, Takahiko, D. Sc.

Division of Applied Genetics

HIROCHIKA, Hirohiko, D. Sc., Adjunct Professor

TAKAGI, Nobuo, D. Sc., Professor (Adjunct)

6. Genetic Strains Research Center

SHIROISHI, Toshihiko, D. Sc., Head of the Center

Mammalian Genetics Laboratory

SHIROISHI, Toshihiko, D. Sc., Professor

Mouse Genomics Resource Laboratory

KOIDE, Tsuyoshi, D. Med., Associate Professor

Mammalian Development Laboratory

SAGA, Yumiko, D. Sc., Professor

KOKUBO, Hiroki, D. Sc.

Plant Genetics Laboratory

KURATA, Nori, D. Ag., Associate Professor

ITO, Yukihiro, D. Ag.

Microbial Genetics Laboratory

NISHIMURA, Akiko, D. Ag., Associate Professor

Invertebrate Genetics Laboratory

UEDA, Ryu, D. Sc., Professor

TAKAHASHI, Kuniaki, D. Sc.

Laboratory for Frontier Research

ISSHIKI, Takako, D. Sc., Associate Professor

7. Center for Genetic Resource Information

KOHARA, Yuji, D. Sc., Head of the Center

Genetic Informatics Laboratory

YAMAZAKI, Yukiko, D. Sc., Associate Professor

FUJITA, Masaya, D. Eng.

Genome Biology Laboratory

KOHARA, Yuji, D. Sc., Professor

ANDACHI, Yoshiki, D. Sc.

8. Structural Biology Center

KATSURA, 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

Multicelluar Organization Laboratory

KATSURA, Isao, D. Sc., Professor

ISHIHARA, Takeshi, D. Sc.

Biomolecular Structure Laboratory

SHIRAKIHARA, Yasuo, D. Sc., Associate Professor

Gene Network Laboratory

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.

Laboratory for Gene-Product Informatics

NISHIKAWA, Ken, D. Sc., Professor

FUKUCHI, Satoshi, D. Sc.

Laboratory for Gene Function Research

TATENO, Yoshio, Ph. D., D. Sc., Professor

FUKAMI-KOBAYASHI, Kaoru, Ph. D.

Laboratory for the Research and Development of Biological Databases

SUGAWARA, Hideaki, D. Eng., Professor

MIYAZAKI, Satoru, D. Sc.

10. Radioisotope Center

NIKI, Hironori, D. Med., 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

ISHIKAWA, Kenji, Head of the Department
TOMIYAMA, Yukio, Chief of the General Affairs Section
TAKAHASHI, Shouji, Chief of the Finance Section

COUNCIL (as of December 31, 2002)

Chairman

SUGIMURA, Takashi; President Emeritus, National Cancer Center

Vice-chairman

OSAWA, Shozo; Adviser, Biohistory Research Hall

Members (Alphabetical order)

HIROBE, Masaaki; President,University of Shizuoka
ISHII, Shiro; Member, Council for Science and Technology Policy, Cabinet Office
IWATSUKI, Kunio; Professor, University of the Air
KATSUKI, Motoya; Director-General, National Institute for Basic Biology
KODAIRA, Keiichi; President, The Graduate University for Advanced Studies
KURODA, Reiko; Professor, Graduate School of Arts and Science, The University of Tokyo
KYOGOKU, Yoshimasa; Director-General, Biological Information Research Center, National Insttute of Advanced Industrial Science and Technology
MATSUBARA, Ken-ichi; President, DNA Chip Research Inc.
MATSUO, Minoru; President, Nagoya University
MIURA, Kin-ichiro; Professor, Emeritus, National Institute of Genetics
MOHRI, Hideo; President, Okazaki National Research Institutes
OKADA, Masukichi; Vice-Director, International Institute for Advanced Studies
OSAKI, Hitoshi; Director-General, Center for National University Finance
OTSUKA, Eiko; Fellow, National Institute of Advanced Industrial Science and Technology
SHINJI, Isoya; President, Tokyo University of Agiculture
TOYOSHIMA, Kumao; Director, Sumitomo Hospital
TSUNEWAKI, Koichiro; President, Fukui Prefectural University
YAMANOUCHI, Kazuya; Senior Scientific Staff, Nippon Institute for Biological Science

ADVISORY COMMITTEE (as of December 31, 2002)

Chairman

KOHARA, Yuji; Professor, National Insutitute of Genetics

Vice-chairman

SEKIGUCHI, Mutsuo; Director, Biomolecular Engineering Research Institute

Outside Members (Alphabetical order)

GO, Michiko; Professor, Graduate School of Science, Nagoya University
HANAOKA, Fumio; Professor, Institute of Molecular and Cellular Biology, Osaka University
ISONO, Katsumi; Professor, Faculty of Science, Kobe University
ITO, Koreaki: Professor, Institute for Virus Research, Kyoto University
MATUURA, Etsuko; Professor, Faculty of Science, Ochanomizu University
OGAWA, Tomoko; Vice-President, Iwate College of Nursing
SASAZUKI, Takehiko; Director-General, International Medical Center of Japan Research Institute
SHINOZAKI, Kazuo; Chief Scientist, RIKEN Tsukuba Institute
TAJIMA, Fumio; Professor, Graduate School of Science, The University of Tokyo

Inside Menbers (Alphabetical order)

ARAKI, Hiroyuki; Professor, National Insutitute of Genetics
GOJOBORI, Takashi; Professor, National Insutitute of Genetics
HIROMI, Yasushi; Professor, National Insutitute of Genetics
HIROSE, Susumu; Professor, National Insutitute of Genetics
IKEMURA, Toshimichi; Professor, National Insutitute of Genetics
KATSURA, Isao; Professor, National Insutitute of Genetics
NISHIKAWA, Ken; Professor, National Institute of Genetics
SASAKI, Hiroyuki; Professor, National Insutitute of Genetics
SHIMAMOTO, Nobuo; Professor, National Insutitute of Genetics
SHIROISHI, Toshihiko; Professor, National Insutitute of Genetics

A. DEPARTMENT OF MOLECULAR GENETICS

A-a. DIVISION OF MOLECULAR GENETICS

(1) Comparative genomics of bacterial transcription factors

Nobuyuki FUJITA

Transcription factor is one of the largest and the most divergent protein families in bacteria. Genome-wide comparison of transcription factors among bacterial species, therefore, will be an important step for understanding bacterial strategy for survival, adaptation, and pathogenicity. Complete genomic sequences of 44 bacterial species were searched for bacteria-type transcription factors based on sequence similarity. Using known transcription factors, mainly from *Escherichia coli* and *Bacillus subtilis*, as initial queries, a total of ~6,000 candidate proteins were collected and classified into more than 50 protein families primarily according to the structure of DNA-binding domain. All bacterial species with medium to large genomic size, irrespective of their taxonomic group, share most of the transcription factor families, suggesting that these protein families have ancient phylogenetic origins. However, similarity and phylogenetic analysis of each protein family allowed unequivocal assignment of orthologs only within closely related bacteria because of the high degree of sequence diversity. Total number of transcription factors in each bacteria showed fairly good correlation with the genome size except for cyanobacterial and mycobacterial species in which the number of transcription factors is significantly smaller. Bacteria with smaller genomic size, typically less than 1.5 Mbp, retain only a few transcription factors. Similar survey of 11 archaeal and 2 eukaryotic genomes indicated that archaea share limited numbers of transcription factor families with bacteria, and virtually no bacterial transcription factor is conserved in eukaryotes. Possible horizontal gene transfer events were inferred by several methods including massive principal component analysis of codon usage pattern. The results suggest that horizontal gene transfer occurs at much higher frequency with the members of certain transcription factor families such as Xre and CadC families than those of other families. An integrated Web-based database of bacterial transcription factors, with identification and classification system for newly determined sequences, is being constructed.

(2) In vitro selection of transcription factor binding sequences from random genomic library

Nobuyuki FUJITA and Shizuko ENDO

Comparative genomics analysis suggests that Escherichia coli genome codes for nearly 270 transcription factors of which about 120 are of unknown functions. As a means to identify the target genes of these hypothetical factors, an in vitro selection system based on sequence-specific DNA binding of purified protein was set up. A plasmid library containing 200-300 bp random genomic DNA of *Escherichia coli* (100,000 independent clones) was constructed and used to generate a random population of PCR fragments. After several cycles of complex formation, copurification, and reamplification, high-affinity binding fragments were sequenced and mapped on the genome. In a pilot experiment with His-tagged CRP (cAMP receptor protein), 161 independent clones were mapped within 31 different regions on the chromosome of which 10 regions were located upstream of known CRP-regulated genes. All clones except 4 were intragenic, suggesting high selectivity and reliability of this method. This in vitro selection system should complement other screening methods such as DNA microarray analysis.

(3) Identification of proteins that interact with the Rpb7 subunit of RNA polymerase II in fission yeast

Hiroshi MITSUZAWA

RNA polymerase II (pol II) in eukaryotes consists of twelve subunits, Rpb1 through Rpb12. Rpb7 subunit is required for accurate transcription initiation from a promoter but is unnecessary for RNA synthesis itself. Rpb7 has RNA binding domains and binds to RNA *in*

vitro. To understand the role of Rpb7 in transcription initiation or other processes, we carried out a two-hybrid screen for proteins that interact with Rpb7 of the fission yeast *Schizosaccharomyces pombe*.

The screen has identified the *S. pombe* homolog of the *Saccharomyces cerevisiae* Nrd1, an RNA-binding protein implicated in 3'-end formation of small nucleolar RNAs transcribed by pol II. The *S. pombe* protein, named Seb1 for seven binding, was shown to be essential for cell viability. Rpb7 bound directly to Seb1 *in vitro*. *S. cerevisiae* Rpb7 also interacted with Nrd1, indicating that the interaction is conserved in evolution. Glu 166 and/or Asp 167 of *S. pombe* Rpb7, residues near the carboxyl terminus of the 172-amino acid protein, were found to be important for its interaction with Seb1. Based on recent structural studies, these residues are presumed to be located in close proximity to nascent RNA transcript extended from the active site of pol II. Our results suggest that Rpb7 subunit may anchor a processing factor to the pol II complex, thereby coupling RNA processing to transcription.

The screen has also identified actin and an actin-related protein as Rpb7-interacting proteins. The actin-related protein is the *S. pombe* homolog of *S. cerevisiae* Arp3. Interaction of Rpb7 with actin and an actin-related protein is interesting because they are known to be components of chromatin remodeling and histone acetyltransferase complexes and are implicated in transcription.

(4) Identification of TBP-associated factors (TAFs) in fission yeast

Hiroshi MITSUZAWA

The general transcription factor TFIID consists of the TATA-binding protein (TBP) and multiple TBP-associated factors (TAFs). TAFs have been proposed to function in core promoter recognition and transcription activation. We reported the first biochemical identification of TAFs in the fission yeast *S. pombe*. Genetic data suggest that two WD repeat TAFs, spTAF72 (spTAF5) and spTAF73 (spTAF5b), regulate genes involved in progression through the M phase of the cell cycle. We also identified a complex containing the *S. pombe* homolog of Gcn5 histone acetyltransferase (SAGA-like complex). Interestingly, spTAF72 but not spTAF73 was shared by this complex.

We have identified another *S. pombe* TAF, spTAF50 (spTAF6), as a protein that interacts with a carboxyl-terminal WD repeat-containing region of spTAF72. spTAF50 is the *S. pombe* homolog of histone H4-like TAFs such as human TAF80 (hsTAF6), *Drosophila* TAF60 (dmTAF6), and *S. cerevisiae* TAF60 (scTAF6). Like spTAF72, spTAF50 was present in both the TFIID and SAGA-like complexes.

Our results have revealed that *S. pombe* has TAF-containing complexes similar to those characterized in other organisms. *S. pombe* TFIID is unique, however, in that it contains two distinct WD repeat TAFs, one of which (spTAF73) is specific to TFIID and the other (spTAF72) is shared by the SAGA-like complex. In contrast, the human, *Drosophila* and *S. cerevisiae* TFIID complexes contain a single species of WD repeat TAF, which is shared by other histone acetyltransferase complexes. The presence of spTAF72 and spTAF73 in a single TFIID complex suggests that WD repeat TAFs in other organisms are present in two copies in TFIID. Support for this prediction has come from a recent analysis of *S. cerevisiae* TFIID.

(5) Interaction between RNA polymerase II and the carboxy-terminal domain phosphatase Fcp1 in the fission yeast *Schizosaccharomyces pombe*

Makoto KIMURA

RNA polymerase II (pol II), which synthesizes all mRNA, is a multiprotein complex consisting of twelve subunits, Rpb1-Rpb12. The largest subunit, Rpb1, contains an unique carboxy-terminal domain (CTD) which consists of 29 repeats in *S. pombe* of heptapeptide with the consensus of YSPTSPS. The CTD is a target of phosphorylation, and at least two phosphoisoforms of pol II, hypophosphorylated IIA and hyperphosphorylated IIO froms, coexist in cells. The hypophosphorylated CTD interacts with a transcription mediator complex, and the IIA from of pol II is recruited on to promoters to assemble preinitiation complexes with general transcription factors. On the other hand, the IIO form elongates mRNA, and the phosphorylated CTD is supposed to be a platform for the assembly of mRNA processing factors. The CTD is deduced to be phosphorylated by cycline-dependent kinase (CDK) 7 in general transcription factor TFIIH, and CDK8 in transcription elongation factor P-TEFb at transcription

initiation and elongation, respectively. After termination, the phosphorylated CTD should be dephosphorylated for another round of transcription.

To isolate and identify protein factors which associate and regulate the pol II, an *S. pombe* strain carrying a FLAG-tag sequence fused to the *rpb3* gene encoding the pol II Rpb3 subunit was constructed. From cell extract prepared as to contain non-engaged pol II, a complex containing IIA form of pol II was isolated by immuno-affinity purification with anti-FLAG antibody. In addition to the pol II subunits, this complex was found to contain three subunits (TFIIFα, TFIIFβ and Tfg3) of general transcription factor TFIIF, and *S. pombe* homologue of TFIIF-interacting CTD-phosphatase Fcp1, which is an only identified "CTD-phosphatase" in mammals and the budding yeast *Saccharomyces cerevisiae*. Though the Fcp1 had been isolated in these organisms, this type of complex, which must be a CTD-dephosphorylating complex, was purified for the first time, and therefore, Fcp1 was proved to be a genuine CTD-phosphatase which functions after transcription termination in cells. The intracellular contents of pol II, Fcp1 and Tfg3 in an exponentially growing *S. pombe* cell were estimated as 10,000, 3,000 and 20,000 molecules per cell, respectively. The same complexes could be purified from a FLAG-Fcp1 and a FLAG-Tfg3 strain, and the stoichiometry of Fcp1 in these complexes revealed that most of Fcp1 in cells exists in the pol II-complex. Cell extracts which contain only engaged pol II were also prepared from the FLAG-Rpb3 and the FLAG-Fcp1 strains, and a complex containing pol II and Fcp1 was purified by a similar method. Analysis of the complex revealed that though most of engaged pol II was IIO form, a small portion of it was IIA form associated with Fcp1. This finding supports the notion that conversion between the IIA and the IIO occurs during elongation also by the functions of Fcp1 and P-TEFb.

Recombinant *S. pombe* Fcp1 showed one thousand times higher hydrolysis activity for an artificial substrate p-nitrophenyl phosphate than the *S. cerevisiae* enzyme, and showed CTD-phosphatase activity *in vitro*. Since Fcp1 was reported to require a non-CTD docking site on pol II for CTD-dephosphorylation, I tried to identify the Fcp1-binding subunit of pol II. By chemical crosslinking, GST-pull down and affinity chromatography, the Fcp1-binding subunit was decided as the Rpb4 subunit, which plays regulatory roles in transcription initiation. To see if the Rpb4-Fcp1 interaction is crucial *in vivo*, a thiamine dependent *rpb4* shut-off system of *S. pombe* was constructed. On repression of the *rpb4* expression, the cell produced more IIA from of pol II, but the pol II-complex isolated with the anti-FLAG antibody contained less Fcp1 and more IIO form of pol II with the concomitant reduction of Rpb4. This result indicates the importance of Fcp1-Rpb4 interaction for the formation of Fcp1/TFIIF/pol II complex and CTD-dephosphorylation *in vivo*. Interaction between TFIIFα/β complex and Fcp1 was also detected by GST-pull down assay. However, the binding was weak in *S. pombe*, indicating the importance of the Rpb4-Fcp1 interaction for the complex formation. Tfg3 subunit of TFIIF did not bind to Fcp1.

To understand the regulation mechanism of CTD-dephosphorylation through the Fcp1-Rpb4 interaction, Fcp1 and Rpb4 were dissected. Wild type or deletion mutants of Fcp1 were expressed in *S. pombe* from plasmids and their assembly into the Fcp1/TFIIF/pol II complex was analyzed. The Fcp1 with N- or C-terminal short deletion did not assemble into the complex, suggesting that the formation of complex depends on the structure of Fcp1. The crystal structure of an archaeal Rpb4 homologue and several secondary structure prediction programs suggested that *S. pombe* Rpb4, which is essential for cell viability, contains seven α-helixes. Serial mutants of the *rpb4* gene with deletions by the helix unit were constructed and expressed constitutively in the *S. pombe rpb4* shut-off strain. The *rpb4* with N-terminal deletion supported the cell growth, the *rpb4* with C-terminal deletion showed a dominant negative effect, and other mutants indicated a lethal phenotype. The characters of these mutant Rpb4 is now being analyzed biochemically.

(6) Analysis of the function of Tfg3 subunit of general transcription factor TFIIF

Makoto KIMURA

Three subunits (TFIIFα, TFIIFβ and Tfg3) of general transcription factor TFIIF were identified in the Fcp1/TFIIF/pol II complex isolated above. The identification of *S. pombe* Tfg3 homologue was unexpected, because this subunit is contained *in S. cerevisiae* TFIIF but not in that from higher eukaryotes, and generally the molecular compositions of transcription factors in *S. pombe* are more similar to those in higher eukaryotes than in *S. cerevisiae*. The

importance of this subunit in transcription regulation was analyzed. Disruption of the *tfg3* gene in *S. pombe* indicated that the *tfg3* is non-essential for cell growth, but the disruptant showed *ts* phenotype as was reported in *S. cerevisiae*. In *S. serevisiae*, this protein is also reported as a subunit of another general transcription factor TFIID and two more complexes participating in transcription regulation. To see if the *S. pombe* Tfg3 is also a subunit of TFIID, TFIID complex was isolated from *S. pombe* strains containing FLAG-tagged *taf72* or *taf73* subunit gene of TFIID, and the isolated complexes were subjected to Western-blotting with anti-Tfg3 antibody. The isolated complexes were proved to contain the Tfg3. Furthermore, the TFIID complex isolated from the cells grown at higher temperature contained more Tfg3 than the cells grown at lower temperature, in spite that the intracellular content of Tfg3 in the former cells was less, suggesting that the Tfg3 plays some role at higher temperature. Direct interaction between Tfg3 and TATA-binding protein (TBP), which is a central component of TFIID, was also detected by GST-pull down assay, and the binding *in vitro* was also stronger at higher temperature. The function of Tfg3 is now being investigated.

Publications

1. Colland, F., Fujita, N., Ishihama, A. and Kolb, A.: The interaction between σ^S, the stationary phase σ factor, and the core enzyme of *Escherichia coli* RNA polymerase. *Genes Cells* **7**, 233-247, 2002.

2. Kimura, M., Suzuki, H. and Ishihama, A.: Formation of a carboxy-terminal domain phosphatase (Fcp1)/TFIIF/RNA polymerase II (pol II) complex in *Schizosaccharomyces pombe* involves direct interaction between Fcp1 and the Rpb4 subunit of pol II. *Mol. Cell.* Biol. **22**, 1577-1588, 2002.

3. Mitsuzawa, H. and Ishihama, A.: Identification of histone H4-like TAF in *Schizosaccharomyces pombe* as a protein that interacts with WD repeat-containing TAF. *Nucleic Acids Res.* **30**, 1952-1958, 2002.

4. Mitsuzawa, H.: Analysis of the general transcription factor TFIID using the fission yeast *Schizosaccharomyces pombe*. Tanpakushitsu Kakusan Koso **47**, 1931-1938, 2002.

5. Yamamoto, K., Ogasawara, H., Fujita, N., Utsumi, R. and Ishihama, A.: Novel mode of transcription regulation of divergently overlapping promoters by PhoP, the regulator of two-component system sensing external magnesium availability. *Mol. Microbiol.* **45**, 423-438, 2002.

A-b. DIVISION OF MUTAGENESIS

(1) Hyper-recombination and hyper-resistance to DNA damage in fission yeast defective in ubiquitin-conjugating enzyme Ubc7

Joon-Hyun PARK, Yufuko AKAMATSU,† Hiroshi IWASAKI† and Fumiaki YAMAO (†*Graduate School of Integrated Science, Yokohama City University*)

Inappropriate or unregulated homologous recombination can have deleterious effects and lead to genomic instability. In diploid eukaryotes, homologous recombination is essential during meiosis, but it is suppressed during mitosis by an as yet unknown mechanism. We identified and characterized Ubc7, one of the 14 ubiquitin-conjugating enzymes in *Schizosaccharomyces pombe*. Mutants defective in Ubc7 have enhanced mitotic recombination between tandem-repeated segments at *mat1* or *ade6*. The increased mitotic recombination was observed only in *ubc7*$^-$ cells which are competent for DSB-induced homologous recombination. Furthermore, Ubc7 mutants are hyper-resistant to UV radiation and MMS-induced DNA damage in cells carrying wild type Rhp51. These results suggest that mitotic homologous recombination is a two-step reaction in eukaryotes; the initiation step establishes competence for homologous recombination and DNA exchange, while the second step mediates the recombination itself. This study demonstrates that Ubc7 negatively regulates the latter step during mitosis in *S. pombe*.

(2) Regulatory network of the yeast cell growth by the SCF ubiquitin ligase

Tsutomu KISHI

The ubiquitin pathway is involved in the regulation of many basic cellular processes. The SCF ubiquitin ligase, one of the ubiquitin ligases in the

ubiquitin system, regulates cell cycle progression, and differenciation and development. Proteins which organize the SCF ubiquitin ligase have been identified and characterized. However, only a few proteins which are ubiquitinated by the ligase are identified. This study aims to identify proteins to be ubiquitinated by the SCF ubiquitin ligase in budding yeast, and to assess how their degradation contributes to cellular regulation.

(3) The ubiquitin-conjugating enzymes involved in ubiquitination of mitotic cyclin

Hiroaki SEINO

Cell cycle events are regulated by sequential activation and inactivation of Cdk kinases. Inactivation of mitotic Cdk kinases is mainly achieved by degradation of mitotic cyclin. Degradation of mitotic cyclin is important for mitotic transition. Mitotic cyclin is degraded by a ubiquitin/proteasome system. The ubiquitin ligase that is involved in ubiquitination of mitotic cyclin is APC/C (anaphase-promoting complex/ cyclosome). In Xenopus and clam oocytes, ubiquitin-conjugating enzymes that function with APC/C have been identified as two proteins, UBC4 and UBCx/E2-C. Previously we reported that the fission yeast ubiquitin-conjugating enzyme UbcP4/Ubc11 that was homologue of UBCx/E2-C was required for mitotic transition.

Furthermore we show other fission yeast ubiquitin-conjugating enzyme UbcP1/Ubc4 that is homologue of UBC4 is also required for mitotic transition in the same manner as UbcP4/Ubc11. Both ubiquitin-conjugating enzymes are revealed to be directly required for the degradation of mitotic cyclin Cdc13 by synchronization experiments. The activity of APC/C(Ste9) that was hyper activated by over-production of Ste9 required both ubiquitin-conjugating enzymes. In vivo analysis of ubiquitination of Cdc13 shown that the ubiquitin chains on Cdc13 were short in ubcP1 mutant cells and ubiquitinated Cdc13 was totally reduced in ubcP4 mutant cells. It is possible that UbcP4/Ubc11-pathway might initiate ubiquitination of Cdc13 and UbcP1/Ubc4-pathway might elongate the short ubiquitin chains on Cdc13. We conclude that these two ubiquitin-conjugating enzymes have distinct and essential function for degradation of mitotic cyclin Cdc13.

A-c. DIVISION OF NUCLEIC ACID CHEMISTRY

(1) Synthesis of a membrane protein with two transmembrane regions

Saburo AIMOTO (Institute for Protein Research, Osaka University)

About 30% of the human genome encodes membrane proteins, but much of the information concerning the structure and function of these membrane proteins 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. However, peptides that have been synthesized have only a single transmembrane domain. Membrane proteins with multiple transmembrane domains must be synthesized, for chemical synthesis to be recognized as an acceptable technique for membrane protein study. The strategy that was employed for the synthesis of a multiple trans-membrane protein was the thioester method. The purification of peptide having a transmembrane region(s), however, represents a challenge. Difficulties are frequently encountered during the RP-HPLC purification of membrane peptides due to their insolubility in the mobile phases, and irreversible adsorption to the column. Suitable purification conditions must be found for preparing membrane peptide in high purity. Overcoming the problems involved in the synthesis of a two transmembrane protein, we succeeded in the preparation of F_1F_0 ATP synthase subunit c.

(2) Synthesis of 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 potent protease inhibitors for this virus, the knowledge of the characteristics of HTLV-1 PR itself and its substrate specificities is critical. Searched the synthetic conditions, especially conditions related to the supression of epimerization during segment condensation, we succeeded in the synthesis of HTLV-1 PR that consisted of 125 amino acids.

(3) Preparation of peptide thioesters by an Fmoc solid phase method

Saburo AIMOTO (Institute for Protein Research, Osaka University)

In 1991, Hojo and Aimoto reported on a method for polypeptide synthesis, which employed *S*-alkyl peptide thioesters (peptide thioesters) as building blocks. Since then, improvements have been made to the thioester method, and it has been applied to the synthesis of a variety of polypeptides. Peptide thioesters are currently in use as intermediates not only in the thioester method, but also in the native chemical ligation method as well as related reactions used in polypeptide synthesis. However, a very basic problem that still remains to be solved involves the difficulty in the preparation of peptide thioesters. In particular, the synthetic yields of phosphopeptide thioesters are frequently low when they are prepared by a Boc solid-phase method using an MBHA resin. Furthermore, as phosphorylated amino acid residues are often located in acid regions in proteins, there is the additional problem of aspartyl residue conversion to succinimide derivatives during peptide assembly. To overcome the problems involved in the preparation of peptide thioesters, we investigated the possibility to synthesize peptide thioesters by an Fmoc solid phase method and found excellent conditions to obtain peptide thioesters.

Publications

1. Sato, T., Kawakami, T., Akaji, K., Konishi, H., Mochizuki, K., Fujiwara, T., Akutsu, H. and Aimoto, S.: Synthesis of a membrane protein with two transmembrane regions. *Journal of Peptide Science*, 8, 172-180, 2002.

2. Nishimura, I., Uetsuki, T., Kuwako, K., Hara, T., Kawakami, T., Aimoto, S. and Yoshikawa, K.: Cell death induced by a caspase-cleaved transmembrane fragment of the Alzheimer amyloid precursor protein. *Cell Death. Differ.*, **9**, 199-208, 2002.

3. Teruya, K., Kawakami, T., Akaji, K. and Aimoto, S.: Total synthesis of [L40I, C90A, C109A]-human T-cell leukemia virus type 1 protease. *Tetrahedron Letters*, **43**, 1487-1490, 2002.

4. Ohtani, M., Minakata, H. and Aimoto, S.: Potent antagonistic action of synthetic analogues of APGWGNamide, an antagonist of molluscan neuropeptide APGWamide. *Peptides*, **5**, 843-852, 2002.

5. Hasegawa, K., Sha, Y. L., Bang, J. K., Kawakami, T., Akaji, K. and Aimoto, S.: Preparation of phosphopeptide thioesters by Fmoc- and Fmoc(2-F)-solid phase synthesis. *Letters in Peptide Science*, 8, 277-284, 2002.

6. Smith, S. O., Eilers, M., Song, D., Crocker, E., Ying, W., Groesbeek, M., Metz, G., Ziliox, M. and Aimoto, S.: Implications of threonine hydrogen bonding in the glycophorin A transmembrane helix dimer. *Biophysical Journal*, **82**, 2476-2486, 2002.

7. Smith, S. O., Smith, C., Shekar, S., Peersen, O., Ziliox, M. and Aimoto, S.: Transmembrane interactions in the activation of the Neu receptor tyrosine kinase. *Biochemistry*, **41**, 9321-9332, 2002.

(1) Molecular mechanisms of the regulatory inactivation of DnaA protein

Kazuyuki FUJIMITSU, Masayuki SU'ETSUGU, Hironori KAWAKAMI, Satoshi NISHIDA, Tadahiro OHMURA, Tadashi UEDA and Tsutomu KATAYAMA

The ATP-DnaA protein opens duplex DNA at the *Escherichia coli* origin of replication, leading to a series of initiation reactions *in vitro*. When loaded on DNA, the DNA polymerase III sliding clamp stimulates hydrolysis of DnaA-bound ATP in the presence of the IdaB/Hda protein, thereby yielding ADP-DnaA, which

is inactive for initiation *in vitro*. This negative feedback regulation of DnaA activity is proposed to play a crucial role in the replication cycle. We found that the mutant protein DnaA R334A is inert to hydrolysis of bound ATP although its affinities for ATP and ADP remain unaffected. The ATP-bound DnaA R334A protein, but not the ADP form, initiates minichromosomal replication *in vitro* at a level similar to that seen with the wild-type DnaA. When expressed at moderate levels *in vivo*, DnaA R334A is predominantly in the ATP-bound form, unlike the wild-type and DnaA E204Q proteins, which *in vitro* hydrolyze ATP in a sliding clamp- and IdaB/Hda-dependent manner. Furthermore, DnaA R334A, but not the wild-type or the DnaA E204Q proteins, promotes overinitiation of chromosomal replication. This *in vivo* data supports a crucial role for bound nucleotides in regulating the activity of DnaA during replication. Based on a homology modeling analysis, we suggest that the R334 residue closely interacts with bound nucleotides. (see Ref. 1)

(2) Structure-function relationship analysis of DnaA DNA-binding domain

Takafumi IWURA, Takayuki OBITA, Masayuki SU'ETSUGU, Tadashi UEDA, Taiji IMOTO and Tsutomu KATAYAMA

DnaA protein binds specifically to a group of binding sites collectively called as DnaA boxes within the bacterial replication origin to induce local unwinding of duplex DNA. The DNA-binding domain of DnaA, domain IV, comprises the C-terminal 94 amino acid residues of the protein. We overproduced and purified a protein containing only this domain plus a methionine residue. This protein was stable as a monomer and maintained DnaA box-specific binding activity. We then analyzed its solution structure by CD spectrum and heteronuclear multi-dimensional NMR experiments. We established extensive assignments of the ^{1}H, ^{13}C, and ^{15}N nuclei, and revealed by obtaining combined analyses of chemical shift index and NOE connectivities that DnaA domain IV contains six α-helices and no β-sheets, consistent with results of CD analysis. Mutations known to reduce DnaA box-binding activity were specifically located in or near two of the α-helices. These findings indicate that the DNA-binding fold of DnaA domain IV is unique among origin-binding proteins. (see Ref. 3)

Publications

1. Nishida, S., Fujimitsu, K., Sekimizu, K., Ohmura, T., Ueda, T. and Katayama, T.: A nucleotide switch in *E. coli* DnaA protein initiates chromosomal replication: Evidence from a mutant DnaA protein defective in regulatory ATP hydrolysis *in vitro* and *in vivo*. *J. Biol. Chem.* **277** (17), 14986-14995, 2002.

2. Onogi, T., Ohsumi, K., Katayama, T. and Hiraga, S.: Replication-dependent recruitment of the b subunit of DNA polymerase III from cytsolic spaces to replication forks in *Escherichia coli*. *J. Bacteriol.* **184** (3), 867-870, 2002.

3. Obita, T., Iwura, T., Su'etsugu, M., Yoshida, Y., Tanaka, Y., Katayama, T., Ueda, T. and Imoto, T.: Determination of the secondary structure in solution of the *Escherichia coli* DnaA DNA-binding domain. *Biochem. Biophys. Res. Commun.* **299** (1), 42-48, 2002.

B. DEPARTMENT OF CELL GENETICS

B-a. DIVISION OF CYTOGENETICS

Membrane traffic: the intracellular transport network essential for the formation and maintenance of the multi-cellular system

Tamotsu YOSHIMORI

Eukaryotic cells are equipped inside them with a transportation system delivering molecules, so-called membrane traffic, which is operated by the protein-regulated membrane dynamics. By performing secretion, endocytosis, etc., membrane traffic is involved not only in housekeeping of each cell but also in various functions including formation of cell polarity and intercellular communication, which are essential for the multi-cellular system. We aim to unravel molecular mechanisms of membrane traffic and their roles in higher 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 received macromolecules taken up by endocytosis from outside. The cargo is then either sorted to lysosomes or recycled back to the plasma membrane. Each molecular basis underlying autophagy and the endosomal sorting transport are poorly understood.

(1) Analysis of relationship between autophagy and the endosomal system

Atsuki NARA, Noboru MIZUSHIMA[1,2], Akitsugu YAMAMOTO[3], Yukiko KABEYA[1], Yoshinori OHSUMI[1] and Tamotsu YOSHIMORI ([1]National Institute for Basic Biology, [2]PRESTO, Japan Science and Technology Corporation, [3]Kansai Medical University)

We previously showed that mouse SKD1 AAA ATPase is involved in the sorting and transport from endosomes; cells overexpressing a dominant-negative mutant, SKD1^{E235Q} were defective in the endosomal transport to both the plasma membranes and lysosomes. On the other hand, in null mutants of yeast Vps4/Csc1, a homologue of SKD1, autophagy is impaired, suggesting that SKD1 functions not only in the endosomal system but also in autopgay directly or indirectly.To test the involvement of SKD1 in mammalian autophagy, we examined the effects of overexpression of SKD1^{E235Q} on autophagy in HeLa cells. Overexpression of SKD1^{E235Q} using an adenovirus delivery system caused a defect in the autophagy-dependent bulk protein degradation. Morphological observations suggested that this inhibition of autophagy results from an impairment of autophagosome-lysosome fusion. Emergence of the effect on endosomal transport preceded that on autophagy and SKD1^{E235Q} localizes exclusively to endosomes but not to autophagosomes, suggesting that the effect of the mutant protein on autophagy is indirect. We further demonstrated that the membrane traffic from endosomes to autophagosomes was also inhibited in the dominant-negative transfectants. These data suggest that SKD1-dependent endosomal membrane trafficking is required for the late stages of the autophagic pathway and is critical in autophagosome-lysosome fusion. The putative component(s) required for fusion with lysosomes may be supplied by the endosome-to-autophagosome membrane traffic; SKD1 is necessary for the transport. (Ref. 1)

(2) Molecular dissection of the endosomal sorting and transport by using semi-intact cells

Atsuki NARA and Tamotsu YOSHIMORI

In vitro reconstitution of membrane trafficking by using cell free systems or semi-intact cell systems is a powerful tool to resolve its molecular machinery; application of the technique to secretory pathway has provided many crucial information about it. To elucidate mechanisms of the sorting and transport in endosomes, we established semi-intact cells by treatment of HeLa cells with a bacterial toxin, Streptricin O, which forms micro pores on the plasma

membranes and allow us to access the cytoplasm directly. We succeeded to reconstitute the endosomes-to-lysosomes transport in these cells by adding the cytosol fraction and ATP-regenerating system. The transport was quantified by an assay system using avidin and biotin that we established previously. We are now investigating a role of SKD1 in the transport. We will then reconstitute the sorting of cargo in endosomes and identify molecules required for the event.

(3) Studies on role of autophagy in protection against abnormal protein aggregation within cells

Shisako SHOUJI, David H. PERLMUTTER[1], Akira KAKIZUKA[2] and Tamotsu YOSHIMORI ([1]University of Pittsburgh, [2]Kyoto University)

Intracelluar accumulation of aggregated abnormal protein are linked to many diseases, including neurodegeneration. It has been believed that cells avoid accumulating potentially toxic aggregates by degradation of abnormal protein mainly via ubiquitin-proteasome system. We hypothesized that autophagy is also involved in degradation of such protein, since induction of autophagy has been often observed in the unfolded protein diseases. To examine the possibility, we have used disease models of cultured cells; we overexpressed the expanded polyQ fragment, which causes Huntington disease, or mutant α_1 antitrypsin Z, which causes degeneration of liver cells. By analyzing these cells, we obtained preliminary data suggesting that autophagy reduces aggregation of the abnormal proteins by degrading part of them.

(4) Analysis of autophagic response to invasion of pathogenic bacteria into cells

Kayoko TSUDA, Ichiro NAKAGAWA[1], Atsuo AMANO[1] and Tamotsu YOSHIMORI ([1]Osaka University)

Some pathogenic bacteria invade into animal cells and survive or proliferate inside cells. Such bacteria often internalize to the membrane-bound compartments such as phagosomes and modulate them to avoid digestion in lysosomes. It was recently reported that Porphyromonas gingivalis appears in autophagosomes. We confirm this and, moreover, found that another pathogenic bacteria also colocalizes with LC3, an autophagosome membrane marker. Autophagy was induced by infection of both bacteria in cultured cells. We are examining whether autophagy attacks the bacteria or is used by them for their survival.

(5) Identification of a mammalian enzyme catalyzing the Apg12 conjugation essential for autophagosome formation

Noboru MIZUSHIMA[1,2], Yoshinori OHSUMI[1] and Tamotsu YOSHIMORI ([1]National Institute for Basic Biology, [2]PRESTO, Japan Science and Technology Corporation)

We previously reported that the Apg12-Apg5 conjugate, which is essential in autophagosome formation, is generated by a ubiquitin-like protein conjugation system. In yeast, Apg12, following activation by the E1-like Apg7, forms a thioester with Apg10 (E2-like). Apg12 is finally conjugated to Apg5 via an isopeptide bond. The Apg12 system is conserved among eukaryotes, although a mammalian counterpart of Apg10 has not yet been identified. We identified the mouse Apg10 ortholog (mApg10) by a yeast two-hybrid screen using the mouse Apg5 (mApg5) as bait. We demonstrated by a modified yeast two-hybrid assay that mApg10 mediates the conjugation of mApg12 and mApg5. The *in vivo* interaction of mApg12 with mApg10 in HeLa cells suggests that mApg10 is an Apg12-conjugating enzyme, likely serving as an Apg5-recognition molecule in the Apg12 system. This novel two-hybrid method, which we have named "conjugation-mediated yeast two-hybrid", proves to be a simple and useful technique with which to analyze protein-protein conjugation. (Ref. 5)

Publications

1. Nara, A., Mizushima, N., Yamamoto, A., Kabeya, Y., Ohsumi, Y. and Yoshimori, T.: SKD1 AAA ATPase-dependent endosomal transport is involved in autolysosome formation. *Cell Struct. Funct.* **27**, 29-37, 2002.

2. Suzuki, T., Nakagawa, M., Yoshikawa, A., Sasagawa, N., Yoshimori, T., Ohsumi, Y., Nishino, I., Ishiura, S. and Nonaka, I.: The first molecular evidence that autophagy relates rimmed vacuole formation in chloroquine myopathy. *J Biochem.* **131**, 647-51, 2002.

3. Ikonomov, OC., Sbrissa, D., Yoshimori, T., Cover, TL. and Shisheva, A.: PIKfyve kinase and SKD1 AAA ATPase define distinct endocytic compartments: only PIKfyve expression inhibits the cell-vacuolating activity of helicobacter pylori VacA toxin. *J.Biol. Chem* **277**, 46785-46790, 2002.

4. Fujita, H., Yamanaka, M., Imamura, K., Tanaka, Y., Nara, A., Yoshimori, T., Yokota, S. and Himeno, M.: A dominant negative form of the AAA ATPase SKD1/VPS4 impairs membrane trafficking out of endosomal/lysosomal compartments: Class E *vps* phenotype. *J. Cell Sci.* **116**, 401-414, 2002.

5. Mizushima, N., Yoshimori, T. and Ohsumi, Y.: Mouse Apg10 as an Apg12 conjugating enzyme: Analysis by the conjugation-mediated yeast two-hybrid method. *FEBS let.* **532**, 450-454, 2002.

B-b. DIVISION OF MICROBIAL GENETICS

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 genetic system of budding yeast, we have identified novel factors involving in chromosomal DNA replication and revealed their functions in chromosomal DNA replication. We have also focused on the S-phase checkpoint that monitors the status of DNA replication and arrest the cell cycle when DNA replication is inhibited since the Dpb11 protein we found seems to be related to the checkpoint. Furthermore, we have extended our study to duplication mechanism of chromatin structure accompanying DNA replication.

(1) Functional characterization of Slds which genetically interact with Dpb11 in budding yeast

Yoichiro KAMIMURA and Hiroyuki ARAKI

Chromosomal DNA replication in eukaryotic cells initiates from multiple origins which fire sequentially throughout the S phase; some fires early and others late. The pre-replicative complex (pre-RC) starts to assemble at origins from late M phase and DNA polymerases are recruited onto origins to initiate DNA synthesis during the S phase. The Dpb11 protein which forms a complex with essential DNA polymerase ε (Pol ε) is required for association of DNA polymerases with origins in DNA replication (Masumoto et al., Mol. Cell. Biol., **20**, 2809-2817, 2000) and for the control of the S-phase checkpoint (Araki et al., Proc. Natl. Acad. Sci. USA, **92**, 11791-11795, 1995). To identify factors interacting with Dpb11, we have isolated 10 *sld* (synthetic lethality with *dpb11-1*) mutations which fall into 6 complementation groups. We have cloned *SLD1*~*6* and found that *SLD1* is identical to *DPB3* encoding the third largest subunit of Pol ε, *SLD4* is identical to *CDC45* required for the initiation and elongation step of chromosomal DNA replication and *SLD6* is identical to *RAD53* which has crucial role for the cell cycle checkpoint. The *SLD2* gene encodes the essential protein which forms a complex with Dpb11 when S-CDK (Cyclin-dependent kinase) phosphorylates it (Kamimura et al., Mol. Cell. Biol., **18**, 6102-6109, 1998; ref. 1). Since the Sld2-Dpb11 complex formation is essential for DNA replication and probably for polymerases' association with origins, S-CDK regulates loading of DNA polymerases onto replication origins through complex formation between Sld2 and Dpb11. The Sld3 and Cdc45 proteins also form a complex required for further loading of replication proteins including DNA polymerases and Dpb11 (Kamimura et al., EMBO J., **20**, 2097-2107, 2001).

The remaining *SLD* gene, *SLD5* encodes an essential 34 kDa protein. Using *sld5-12*, a thermosensitive allele of *SLD5*, we isolated the *PSF1* (Partner of *Rom*) as a multicopy suppressor. Then the *PSF3* gene was isolated as a multicopy suppressor of the *psf1-1* thermosensitive mutation. In co-immunoprecipitation experiment, we found that Sld5 forms a complex, together with a novel protein, Psf2, as well as Psf1 and Psf3. We named this complex GINS (Go, Ichi, Ni, San; Five, one, two, three in Japanese). *PSF1*, *-2* and *-3* genes are all novel genes essential for cell growth. Since the amount of GINS was reduced in *sld5-12* cells and overproduction of Psf3 in *sld5-12* cells restored the cell growth and the amount of GINS, it is suggested that Psf1, -2 and -3 proteins function through the GINS complex. We thus examined the *psf1-1* and *sld5-12* mutant cells and found that they arrest with a dumbbell shape and are defective in DNA replication at the restrictive temperature. Therefore, the GINS complex is required for DNA replication.

Many replication-related proteins including Dpb11

associate with replication origins in chromatin immunoprecipitation assay (CHIP). Indeed, the CHIP assay revealed that Psf1 also associates with origins in S phase at the same time as Dpb11 or Pol ε. This association with origins depends on Sld3 and Dpb11. Conversely, Psf1 is required for the association of Dpb11 with origins. Moreover, two-hybrid assay showed the interaction between Psf1 and Dpb11 and Sld3. These results suggest that GINS is involved in a multi-protein complex with at least Dpb11 and Pol ε in S phase and intermediates the association between Sld3-Cdc45 on origins and Dpb11-Pol ε to assemble the replisome on replication origins.

(2) Molecular mechanism of the initiation of DNA replication regulated by S-CDK

Yon-Soo TAK, Yoichiro KAMIMURA and Hiroyuki ARAKI

The regulated assembly of replication factors on origins is crucial for the control of the initiation of replication. At the onset of S-phase, the activated S-CDK leads to tight association of Cdc45 with pre-RC. Furthermore, it is believed that S-CDK activates one or more replication factors including Sld2 that has six preferred CDK phsophorylation motifs. Sld2 forms a complex with Dpb11 predominantly in S phase and this complex is required for DNA replication. The mutant Sld2 lacking all the CDK phosphorylation sites fails to interact efficiently with Dpb11 and is defective in chromosomal DNA replication. These results suggest that S-CDK dependent phosphorylation of Sld2 enhance the complex formation with Dpb11 to function in the initiation of DNA replication (ref.1).

To understand how DNA replication is initiated by active S-CDK, we tried to clarify the relationship between the tight association of Cdc45 with pre-RC and the association of Dpb11-Sld2 complex with origins. Since the tight allele of thermosensitive *cdc45* mutation was not available, we first isolated thermosensitive alleles of *CDC45* defective in the initiation of DNA replication. One of the tight alleles, the *cdc45-26* mutation does not affect the Sld2-Dpb11 complex formation in S phase, but abolishes association of Sld2 and Dpb11 with early-origins. Conversely, chromatin-bound Cdc45 significantly reduces in *drc1-1* that is a temperature-sensitive allele of *SLD2* and *dpb11-26* mutant cells. These results suggest that, in addition to complex formation between Sld2 and Dpb11, S-CDK regulates the tight association of Cdc45 with pre-RC and the association of Dpb11-Sld2 complex with origins and their association is in mutually dependent manner.

(3) Functional analysis of the Dpb11 protein

Sachiko MURAMATSU and Hiroyuki ARAKI

To elucidate the function of Dpb11 further, we explored defect in *dpb11-1* cells. Previous studies show that Dpb11 is required for the origin association of DNA polymerases at the initiation step of DNA replication and moves rapidly away from origins in CHIP assay, suggesting that Dpb11 functions only for the initiation of DNA replication. However, *dpb11-1* cells that were arrested by hydroxyurea (HU), an inhibitor of DNA replication, and released at the restrictive temperature lost their viability quickly. Moreover, the cells that were released in medium with Nocodazole, an inhibitor of mitosis, from HU at the restrictive temperature did not restore viability. These results suggest that Dpb11 plays roles not only before HU block but also after HU block in S phase.

We therefore first addressed whether *dpb11-1* cells complete chromosomal DNA replication at the restrictive temperature. For this purpose, we performed pulsed field gel electrophoresis (PFGE) of chromosomal DNA obtained from the mutant cells. Replicating DNA does not enter the gel of PFGE by its complicated DNA. Then we found that chromosomal DNA of *dpb11-1* cells released from HU at restrictive temperature enters the gel as efficiently as that from the wild type cells. This result suggests that *dpb11-1* cells roughly complete chromosomal DNA replication.

We also examined sister chromatids separation in *dpb11-1* cells by visualizing tetR-GFP fusion proteins bound to tandem repeats of tetO integrated at near the cenV, because chromatid cohesion occurs in S phase. However, we did not find any difference between *dpb11-1* and the wild-type cells in the sister chromatid separation.

Further analysis of possible defect in *dpb11-1* cells is under way.

(4) Counteractions of DNA polymerase ε and yCHRAC for epigenetic inheritance of telomere position effect variegation (TPE)

Tetsushi IIDA and Hiroyuki ARAKI

Eukaryotic chromosomal DNA is packaged into chromatin structures: Most parts are in euchromatin and others are in heterochromatin. The chromatin structures are stably propagated during or soon after DNA replication. In budding yeast, when a wild-type gene is located near telomere, it is subjected to TPE that includes transcriptional silencing and provides heritable silent and expressed states and reversible switching between these epigenetic states. The silent state of telomeric gene is attributable to heterochromatin-like structure. To address how DNA replication couples with chromatin duplication, we monitored TPE in mutant cells lacking Dpb3 or Dpb4, subunits of Polε. Polε, one of replicative DNA polymerases, is composed of catalytic-subunit Pol2, Dpb2, Dpb3 and Dpb4. Although Pol2 and Dpb2 are essential for DNA replication, Dpb3 and Dpb4, which contain histone-fold motif related to chromatin metabolisms, are dispensable and their function is not clear.

In the assay of silencing with telomeric *URA3*, *dpb3*Δ, *dpb4*Δ and *pol2-11* C-terminal mutant of *POL2*, cells displayed partial defect of silencing, and this defect was most evident in *dpb3*Δ cells. The silencing defect of *pol2-11* cells was completely suppressed by simultaneous introduction of high copy *DPB3* and *DPB4*. Moreover, the silencing level in *dpb3*Δ *dpb4*Δ double mutant was similar to that in *dpb4*Δ, indicating that the *dpb4*Δ mutation is epistatic to the *dpb3*Δ mutation. In parallel, we also observed the expression of telomeric *ADE2*, which gave rise to red and white sector colonies in the wild-type strain because of mosaic gene silencing by TPE, and found that *dpb4*Δ mutant cells form non-sectoring light pink colonies and *dpb3*Δ mutant cells form white colonies. These results suggest that the mutations in Polε increase the switching frequency between alternative epigenetic states in TPE.

To monitor the switching between a silent and an expressed state in each cell division, we developed the single-cell telomeric silencing assay that distinguish between a silent state (off) and an expressed state (on) of telomeric *α2* gene in a single cell on the α-factor containing medium. Using this assay, we found that the switching rates for both directions increased in *dpb4*Δ cells, whereas in *dpb3*Δ cells the switching rate from off to on specifically increased. These results suggest that Dpb4 plays roles in stable inheritance of both silent and expressed states, while Dpb3 is only involved in stable inheritance of a silent state.

The epistasis of *dpb4*Δ to *dpb3*Δ, together with different switching patterns in these mutants, suggests that Dpb4 is shared by Polε and an unknown complex that plays counteracting roles for TPE. We thus purified protein complexes containing Dpb4 using anti-Flag antibody and the 5Flag-epitope tagged Dpb4 protein, and found that two distinct protein complexes, Polε and yCHRAC, share Dpb4. yCHRAC is a putative homologue of chromatin accessibility complex, CHRAC in higher eukaryotes, and composed of a WAC-motif protein Itc1, an ISWI-chromatin remodeling factor homologue Isw2, a novel histone-fold protein Dpb31 and Dpb4.

We next addressed whether yCHRAC counteracts Polε for TPE. In the assays with telomeric *URA3* and *ADE2*, the *itc1*Δ and *dpb31*Δ mutations enhanced telomere silencing and restored it in *dpb3*Δ cells to the level of *dpb4*Δ cells, whereas they did not affect it in *dpb4*Δ cells. Moreover, in contrast to the *dpb3*Δ mutation, the *dpb31*Δ mutation increased the switching rate from on to off, but did not affect that from off to on. Therefore, these results suggest that yCHRAC regulates stable inheritance of an expressed state of TPE, independent from Polε, while they counteract for TPE.

Taken all together, TPE at yeast telomeres is caused by not only simple competition between assembly of heterochromatin and establishment of active-chromatin but also by specific factors such as Polε and yCHRAC, which serve for stable inheritance of the epigenetic states.

Publications

1. Masumoto, H., Muramatsu, S., Kamimura, Y. and Araki, H.: S-Cdk-dependent phosphorylation of Sld2 essential for chromosomal DNA replication in budding yeast. *Nature* **415**, 651-655, 2002.

(1) Binding of oriC DNA protects DnaA from phospholipid-induced nucleotide release

Seiichi YASUDA

DnaA is an essential protein in the initiation of replication that occurs at a unique locus, oriC, on the chromosome in bacteria. DnaA binds to the oriC DNA and opens the double stranded DNA at an AT-rich region within 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 domains. Among them domain Ⅲ belongs to a group of the AAA+ family of ATPases which binds and hydrolyses ATP and other nucleotides. Domain IV at the C-terminus contains a DNA-binding motif, and is responsible for specific binding of DnaA to a conserved 9-base pair sequence called DnaA box in oriC.

Although ATP-bound, ADP-bound and nucleotide-free DnaA can all bind oriC DNA, only ATP-bound form is active in opening double stranded oriC. It has been known that acidic phospholipids such as cardiolipin release ATP and ADP from DnaA. It was suggested that phospholipids might be involved in rejuvenation of inactive, ADP-bound DnaA by releasing ADP, forming nucleotide-free DnaA that is readily activated by binding ATP.

I have been studying the properties of *Escherichia coli* DnaA and found that the release of ATP and ADP from DnaA by cardiolipin is inhibited by DNA. Covalently closed circular and linear DNA of oriC or non-oriC plasmid was equally effective in this inhibition. The inhibition was observed only when DNA was added before cardiolipin. Moreover, I found that cardiolipin inhibits oriC binding of DnaA only when oriC DNA is added after cardiolipin. Thus, binding of DNA to DnaA seems to protect DnaA from nucleotide release by cardiolipin. Although initial experiments with DNA of large size did not show any differences in the protection efficiency between one DNA species and another, the specificity to DNA sequence and size became evident when smaller DNA of defined sequence was used. Synthetic double stranded DNA of various lengths containing a random sequence or a DnaA box sequence was tested for binding to DnaA and protection of nucleotide release. As to DnaA box-containing DNA, 17-mer was fully active, 13-mer was intermediate, and 9-mer was completely inactive in both reactions. In contrast, even 21-mer of random sequence showed only low activities, and 13-mer was completely inactive in both reactions. This parallelism between the binding affinity to DnaA and the protection of nucleotide release observed with different DNA suggests that the nucleotide release by phospholipids is inhibited by binding of DNA *per se* and not by a steric effect of bound DNA.

A mutant DnaA with deletion of 14 amino acids at the C terminus was purified. This truncated DnaA bound ATP and oriC DNA as well as the wild type DnaA, but was inactive in *in vitro* replication in a crude protein extract. Release of bound ATP by cardiolipin was normal in this mutant protein, but was not protected by oriC binding. This suggests that the C-terminal sequence of DnaA is necessary for the protection of nucleotide release by DNA. According to the three dimensional structure of DnaA of an Alchae, *Aquifex aeolicus*, the C-terminal 14 amino acid sequence that was deleted in the above mutant constitutes a major part of the C-terminal alpha helix called α17 which lies on the surface of the DNA-binding domain, separated distantly from the ATP-binding domain. The function of α17 is not known, but it might somehow be involved in exchange of information between domains Ⅲ and IV that is essential for the initiation function of DnaA.

B-c. DIVISION OF CYTOPLASMIC GENETICS

(1) Linkage disequilibrium mapping of schizophrenia

Takeo YOSHIKAWA (Lab. for Molecular Psychiatry, RIKEN Brain Science Institute)

We have performed a genome scan of schizophrenia pedigrees (119 families, total of 357 individuals) using family-based association tests, at 5 cM marker density for chromosomes 6, 11 and 18, and at a 10 cM density for the remaining chromosomes. These searches have identified several genomic loci including the short arm of chromosome 18, as candidate intervals for harboring disease susceptibility genes. In 2002, we focused on chromosome 18p in detail, by typing 16 genetic markers from an approximate 14 Mb region. This analysis revealed that two markers

derived from the *C18orf1* gene were significantly associated with schizophrenia (Publication #12). We had previously cloned and reported *C18orf1* as a brain transcript encoding a type Ib integral membrane protein. We analyzed the gene for polymorphisms and found a novel, single nucleotide polymorphism (SNP) in the 5'-untranslated region. The association of this SNP with schizophrenia was confirmed in a differently recruited case-control panel. These results suggest that a true disease-causing variant may be located in the gene or an adjacent region. We are presently searching for functional polymorphisms that may predispose to schizophrenia.

(2) QTL (Quantitative Trait Loci) mapping of mouse prepulse inhibition

Takeo YOSHIKAWA

Prepulse inhibition (PPI) of the acoustic startle response is a form of sensory motor gating. It is defined as an inhibition of the startle response when a low intensity stimulus, the prepulse, precedes the startling stimulus. PPI is one of the few paradigms in which humans and rodents are tested in a similar manner. Deficits in PPI have been reported in schizophrenia, making PPI a useful endophenotype of schizophrenia. Some of the genes controlling PPI are expected to overlap with disease susceptibility genes. As preliminary experiments for mapping PPI-regulating genes in mice by QTL, we examined several parameters including: (1) mouse strain, (2) breeders, (3) apparatus for measuring PPI, (4) effects of age and sex of mice, (5) effect of seasons on both mouse purchase date and experiment date. In addition, inherent PPI experimental parameters such as the magnitude of background noise, prepulse and pulse were also examined. Now that we have determined the optimal settings, we are making F2 generation mice for QTL mapping.

(3) Gene expression analysis of an animal model for depression using GeneChips

Takeo YOSHIKAWA

Learned helplessness (LH) is a good animal model for chronic depression. By using LH rats and DNA microarray technology, we have performed a genome-wide search of known genes (~8,000 genes) perturbed in the depressive state and following antidepressant treatments. In 2002, we included ~16,000 new ESTs (expressed tag sequences) for examination, and conducted extensive computational work to obtain information on predictive protein characteristics of these ESTs that could be relevant to depression and responsiveness to drugs.

(4) Candidate gene analyses

Takeo YOSHIKAWA

Target genes are selected for both positional and functional relevance to schizophrenia, mood disorder or epilepsy. The analyzed genes in 2002 include those for adenylate cyclase type 9 (*ADCY9*) (for mood disorder; Publications #1, 6), cholecystokinin (for mood disorder; Publication #5), nicotinic acetylcholine receptor □6 subunit (*CHNRA6* gene) (for epilepsy; Publication #7), NR4A2 (for schizophrenia; Publication #10). We have identified either genetic polymorphisms or elucidated a regulatory mechanism (for the *CHNRA6* gene) within these genes and examined their roles in the development of neuropsychiatric disorders.

Publications

1. Toyota, T., Hattori, E., Meerabux, J., Yamada, K., Saito, K., Shibuya, H., Nankai, M. and Yoshikawa, T.: Molecular analysis, mutation screening, and association study of adenylate cyclase type 9 gene (*ADCY9*) in mood disorders. Am. J. Med. Genet. Neuropsychiatric Genet. **114**, 84-92, 2002.

2. Yoshikawa, T., Watanabe, A., Ishitsuka, Y., Nakaya, A. and Nakatani, N.: Identification of multiple genetic loci linked to the propensity for "behavioral despair" in mice. Genome Research **12**, 357-366, 2002.

3. Akanuma, N., Saitoh, O., Yoshikawa, T., Matsuda, H., Ishikura, N., Kato, M., Adachi, N. and Onuma, T.: Interictal schizophrenia-like psychosis in a patient with double cortex syndrome. The Journal of Neuropsychiatry and Clinical Neurosciences **14**, 210-213, 2002.

4. Reyes, G. D., Esterling, L. E., Corona, W., Ferraren, D., Rollins, D. Y., Padigaru, M., Yoshikawa, T., Monje, V. D. and Detera-Wadleigh, S. D.: Map of candidate genes and STSs on 18p11.2, a bipolar disorder and schizophrenia susceptibility region. Mol. Psychiatry

7, 337-339, 2002.

5. Hattori, E., Yamada, K., Ebihara, M., Toyota, T., Masahiro, M., Shibuya, H. and Yoshikawa, T.: An association study of the short tandem repeat in the 5' upstream region of the cholecystokinin gene with mood disorders in the Japanese population. Am. J. Med. Genet. Neuropsychiatric Genet. **114**, 523-526, 2002.

6. Toyota, T., Yamada, K., Saito, K., Detera-Wadleigh, S. D. and Yoshikawa, T.: Association analysis of adenylate cyclase type 9 gene using pedigree disequilibrium test in bipolar disorder. Mol. Psychiatry **7**, 450-452, 2002.

7. Ebihara, M., Ohba, H., Ohno, S. and Yoshikawa, T.: Genomic organization and promoter analysis of the human nicotinic acetylcholine receptor □6 subunit (*CHNRA6*) gene: *Alu* and other elements direct transcriptional repression. Gene **298**, 101-108, 2002.

8. Ohtsuki, H., Ishiguro, H., Detera-Wadleigh, S. D., Toyota, T., Shimizu, H., Yamada, K., Yoshitsugu, K., Hattori, E., Yoshikawa, T. and Arinami, T.: Association between serotonin 4 receptor gene polymorphisms and bipolar disorder in Japanese case-control samples and the NIMH Genetics Initiative Bipolar Pedigrees. Mol. Psychiatry **7**, 954-961, 2002.

9. Yoshitsugu, K., Meerabux, J. M. A., Asai, K. and Yoshikawa, T.: Fine mapping of an isodicentric Y chromosomal breakpoint from a schizophrenic patient. Am. J. Med. Genet. Neuropsychiatric Genet. 116B, 27-31, 2003.

10. Iwayama-Shigeno, Y., Yamada, K., Toyota, T., Shimizu, H., Hattori, E., Yoshitsugu, K., Fujisawa, T., Yoshida, Y., Kobayashi, T., Toru, M., Kurumaji, A., Detera-Wadleigh, S. and Yoshikawa, T.: (in press). Distribution of haplotypes derived from three common variants of the *NR4A2* gene in Japanese schizophrenic patients. Am. J. Med. Genet. Neuropsychiatric Genet.

11. Ebihara, M., Ohba, H., Hattori, E., Yamada, K. and Yoshikawa, T.: (in press). Transcriptional activities of cholecystokinin promoter haplotypes and their relevance to panic disorder susceptibility. Am. J. Med. Genet. Neuropsychiatric Genet.

12. Kikuchi, M., Yamada, K., Toyota, T., Itokawa, M., Hattori, E., Yoshitsugu, K., Shimizu, H. and Yoshikawa, T.: (in press). Two-step association analyses of the chromosome 18p11.2 region in schizophrenia detect a locus encompassing *C18orf1*. Mol. Psychiatry.

Oral Presentations

1. Yoshikawa, T.: QTL analysis of mouse depression models. XII World Congress of Psychiatry, Yokohama, Japan, August

2. Itokawa, M., Yamada, K., Yoshitsugu, K., Suga, I., Toru, M. and Yoshikawa, T.: Mutation analysis of the NMDA receptor subtype 2A gene. XII World Congress of Psychiatry, Yokohama, Japan, August

3. Yamada, K., Iwayama-Shigeno, Y., Yoshida, Y., Toyota, T., Itokawa, M., Hattori, E., Yoshitsugu, K. and Yoshikawa, T.: Evidence for transmission distortion in schizophrenia on 11q12.1 and 16P13.3. Xth World Congress on Psychiatric Genetics, Brussels, Belgium, October

4. Ebihara, M., Ohba, H., Hattori, E. and Yoshikawa, T.: Transcriptional and genetic re-analyses of cck promoter repeat suggest that the frequency of the low activity haplotype is decreased in panic disorder. American Society of Human Genetics 52nd Annual Meeting, Baltimore, USA. October

5. Nakatani, N., Aburatni, H., Iijima, Y., Iwayama-Shigeno, Y. and Yoshikawa, T.: Genome-wide expression analysis using animal model of depression. Society for Neuroscience 32nd Annual Meeting, Orlando, USA, November

C. DEPARTMENT OF DEVELOPMENTAL GENETICS

C-a. DIVISION OF DEVELOPMENTAL GENETICS

(1) Function of the nuclear receptor Seven-up in the Drosophila central nervous system

Makoto KANAI, Masataka OKABE and Yasushi HIROMI

The diversity of neurons and glial cells in the central nervous system is generated in two steps. First, the identities of about 30 Neuroblasts (NB) are specified by the spatial information provided by the segment polarity genes and the columnar genes. Second, the unique identities of the progeny of each NB are determined by the birth order within the lineage. Recently, Isshiki et al. (2001) showed that NBs sequentially express transcription factors HB, KR, PDM, CAS, and their progeny maintain the expression pattern of the transcription factor using it as a determinant of birth order identity. While the order of the expression of these transcription factors are identical in every NB examined, the mechanism that instructs when to switch the expression of these genes is unknown.

Nuclear receptor Seven-up (SVP) is a candidate for a factor that specify the temporal information for the switching. *svp* is expressed in almost all NBs, but only in certain subsections of the NB lineage. For example, some NB express *svp* at the time of NB formation, while others start its expression only after several cell divisions. This suggests that every NB has a unique *svp* expression pattern. To test the possibility that *svp* plays a role in regulating the expression sequence of HB, KR, PDM, CAS, we assayed the expression pattern of these transcription factors in several NB lineages. In certain NB lineages the number of HB, KR double-positive cells was increased in a *svp* null mutant compared to the wild type embryo, in which only early born cells coexpress HB and KR. This phenotype suggests that without *svp* function, the switching of HB to KR fails to occur properly within the NB lineage.

(2) Regulation of *prospero* expression in longitudinal glial cells

Yoshihiro YUASA and Yasushi HIROMI

In Drosophila, the longitudinal glia regulates formation of longitudinal axonal bundles and neuronal survival. A critical factor that enables the function of longitudinal glial cells is Prospero (PROS), a homeodomain transcription factor. PROS is expressed in a subset of longitudinal glial cells. In *pros* mutant embryos longitudinal glial cells form, but fail to differentiate and support axonal growth. We are analyzing the regulation of *pros* expression in the longitudinal glia. We focused on three transcription factors, homeodomain protein REPO, Ets transcriptional factor PNTP1 and Retained/Dead ringer (RETN/DRI), which is a member of the ARID (AT-rich interaction domain) family. These transcription factors are all expressed in longitudinal glial cells. We found that PROS expression was absent or reduced in mutants of any of these three transcription factors. While ectopic expression of each transcription factor alone had no effect on PROS expression, when three genes were simultaneously misexpressed, there was a dramatic increase in the number of PROS-positive cells in the epidermis. Within the upstream regulatory region of the *pros* gene there were consensus binding sites for all three transcriptional factors. Thus *pros* could be a direct target gene for these three transcription factors in the longitudinal glia.

(3) Role of Drosophila Spred in Ras signaling

Shu KONDO and Yasushi HIROMI

Negative regulation of Ras signaling is known to be essential for proper develpment and prevention of cancer formation. Spred is an intracellular protein originally identified in mammals as a negative regulator of Ras signaling. We have identified a Drosophila homolog of Spred. We found that 1) Spred can suppress Ras signaling when overexpressed in the developing eye, where the differentiation of any particular cell type depends on the activation of Ras signaling, 2) In S2 cells, overexpressed Spred is

localized to the plasma membrane and suppresses the activation of MAP kinase, 3) The antiserum raised against the Spred protein detects the antigen at the basal membrane of embryonic muscle attachment cells, where ras sinaling is highly active. Although the gain-of-function experiments all points to its potential role as a Ras antagonist, the Drosophila Spred mutant shows no detectable abnormalities, even in sensitive backgrounds where Ras signaling is either above or below the physiological level, hampering our understanding of its physiological role *in vivo*.

(4) Going through the segmental boundary: Changing the responsiveness to the segmentary repeated guidance cue presented by Frazzled

Masaki HIRAMOTO and Yasushi HIROMI

During axonal navigation growth cones must often alter their responsiveness to guidance cues. For example, commissural axons switch their response to the midline cue Netrin upon crossing the midline, from attractive to repulsive. This is achieved by the binding of the cytoplasmic domain of the Netrin receptor DCC to the cytoplasmic domain of the Slit receptor Robo, causing the desensitization of DCC to Netrin. Longitudinal axons also encounter a similar situation when they go thorough repeated structures, leaving and re-entering identical structural unit in each segment. The longitudinal pioneer neuron dMP2 initially turns posteriorly, using the Netrin cue positioned by the Drosophila DCC homolog Frazzled. However, dMP2growth cones stop tracking the boundary of Frazzled/Netrin-positive region after the turning event. We showed that the entry of dMP2 into the succeeding segment requires the change in the responsiveness to Frazzled/Netrin by turning on Robo expression. Thus, the midline crossing phenotype of dMP2 in *robo* mutant is due to a non-cell-autonomous effect of Frazzled.

(5) Intra-axonal patterining

Takeo KATSUKI, Masaki HIRAMOTO and Yasushi HIROMI

How positional information for axon guidance is provided remains as one of the central questions in neurobiology. We have previously shown that intra-axonal localization of a Drosophila guidance receptor Frazzled creates positional information by rearranging the distribution of the guidance molecule Netrin, and proposed the "capture/relocation model of guidance receptors. This raises a question on how the receptors themselves localize inside the neuron to create positional information. To address this issue, we compared the distribution of several neuronal proteins in embryonic Drosophila CNS and in primary culture. We found that several molecules exhibit uneven but regulated subcellular distribution even when neurons grow axons in culture, isolated from their normal environment; These results imply that neurons possess an intrinsic intracellular patterning system that can localize molecules in axonal segments on the cell surface. Neurons might utilize their characteristic structures with long processes to deliver wide and specific pattern of positional information within the nervous system. Live observations should provide us with insight into the mechanism of localization.

Publication

1. Umesono, Y., Hiromi, Y. and Hotta, Y.: Context-dependent utilization of Notch activity in Drosophila glial determination. *Development* **129**, 2391-2399, 2002.

2. Okano, H., Imai, T. and Okabe, M.: Musashi: a translational regulator of cell fates. *J. Cell Sci.* **115**, 1355-1359, 2002.

(1) Identification of New Peptides and Peptide Genes by Combining the *Hydra* Peptide Project and the *Hydra* EST Project

Toshitaka FUJISAWA, Chiemi FUJISAWA, Yasuharu TAKAKU, Shiho HAYAKAWA[1] and Takashi GOJOBORI[1]
([1]DNA Data Analysis laboratory, Center for Information Biology)

We have initiated the Hydra EST Project this year. The EST data have proved to be extremely valuable for the Hydra Peptide Project. First, we have identified a number of isoforms of peptide genes that encode new species of related as well as unrelated peptides. The expression patterns of these genes are also interesting. For example, neuropeptide genes clearly identified subpopulations among neurons that will lender a new

tool to examine mechanisms for neuron differentiation. Second, the genes encoding peptides that were identified by the Peptide Project can be easily obtained. Moreover, it is possible to postulate some of the post-translational modifications of the peptides from the precursor sequences. This vastly saves the time to know, for example, if the C-terminus is amidated or not. This was only possible by HPLC analyses. Thus, the EST Project is valuable for itself, but or mass also it will greatly speed up the ongoing Peptide Project.

Publications

1. Iwao, K., Fujisawa, T. and Hatta, M.: *Coral Reefs* **21**, 127-129, 2002.

2. Mochizuki, K., Fine, N. A., Fujisawa, T. and Gorovsky, *M. A.*: *Cell* **110(6)**, 689-99, 2002.

3. Fabila, Y., Navarro, L., Fujisawa, T., Bode, H. R. and Salgado, L. M.: *Mech. Dev.* **119(2)**, 157-64, 2002.

4. Morishita, F., Nitagai, Y., Furukawa, Y., Matsushima, O., Takahashi, T., Hatta, M., Fujisawa, T., Tsunamoto, S. and Koizumi, O.: *Peptides* **6542**, 1-10, 2003.

5. Fujisawa, T.: *Developmental Dynamics* (in press).

(2) The identification of peduncle of *Hydra* as a possible earliest heart in metazoan evolution

Hiroshi SHIMIZU and Toshitaka FUJISAWA

The heart is assumed to have evolved as the organ for pumping blood. Here we present evidence that *Hydra* a phylum Cnidaria has an activity to pump gastric fluid and the mechanism for the pumping is similar to that of blood pumping by the heart in higher metazoans. We find that *Hydra* peduncle stores the gastric fluid and upon its contraction pumps the fluid out of it to circulate throughout the coelenteron. We also find that Hydra RFamide III, a homologue of cardioexcitatory peptides, elevates the pumping activity. Further, the expression of *CnNk-2* which has 70% homology to *Nkx-2.5*, a marker gene of heart-cell differentiation, is expressed in the peduncle. These observations demonstrate that the pumping fluid by the peduncle has functional, neurological and genetic similarities to the pumping blood by the heart suggesting that the heart-like pumping appeared in Cnidaria far preceding the appearance of blood vascular system.

(3) The identification of diffuse nerve net in the body column of Hydra as the enteric nervous system

Hiroshi SHIMIZU and Toshitaka FUJISAWA

How nervous system appeared in metazoan evolution is a crucial issue in neurobiology. A widely accepted view is that the diffuse nerve net most conspicuously found in the body column of *Hydra* a phylum Cnidaria is the remnant of the ancestral nervous system and is functionally meager contributing at most to bidirectional conduction of stimulus. Here we present evidence that the body column nerve net is a functionally ample enteric nervous system. We find that the body column of *Hydra* when fed with Artemia undergoes a series of movements. Comparison of the movements with the digestive movements in mammalian digestive tract showed three similarities. First, the movement in the early stage of digestion was similar to the movement of mammalian esophagus termed esophageal reflex. Second, the movement in the middle stage was similar to that of stomach and intestine termed peristaltic reflex. Third, the movement in the final stage was similar to that of rectum termed defecation reflex. Isolated body column tissue showed movements similar to those in intact *Hydra* demonstrating that the movements are indeed reflexes which are not dominated by exogeneous nervous system. We also find that polyps which have no nerve cells are capable of undergoing peristaltic movement to a limited extent. These observations suggest that *Hydra's* diffuse nerve net is functionally equivalent to enteric nervous system of mammals which regulates digestive movements and that metazoans had obtained the capacity of these movements in the very early stage of their evolution. Because of the structural simplicity, the diffuse nerve net of *Hydra* could be a very good experimental system to elucidate the mechanism of digestive movements.

Publications

1. Shimizu, H.: *Comp Biochem Physiol A.* **131**, 669-674, 2002.

2. Shimizu, H., Zhang, X., Zhang, J., Leontovich, A., Fei, K., Yan, L. and Sarras, MP.: *Development* **129**, 1521-1532, 2002.

C-b. DIVISION OF GENE EXPRESSION

1. Research activities

In multicellular organisms, a single fertilized egg divides into multiple cells that give rise to tissues and organs. We are studying gene expression during embryonic and postembryonic development of *Drosophila*. Our current research projects are :

(1) Role of GAGA factor-dependent chromatin remodeling in *Hox* gene expression

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

(3) Functional analysis of transcriptional coactivator MBF1

(4) Mechanism of transcriptional regulation of FTZ-F1

(5) Role of FTZ-F1 during development

Following members participated in our research activities: Susumu Hirose (Professor), Hitoshi Ueda (Associate professor), Masa-aki Yamada and Kiyoshi Minato (Research associates), Qin-Xin Liu (Foreign researcher), Marek Jindra (JSPS Short term invited fellow), Kunihiro Matsumoto, Takahiro Nakayama and Tsukasa Shimojima (Postdoctoral fellows), Haruhisa Kawasaki, Yasunari Kayashima, Yasuo Agawa and Hirofumi Furuhashi (Graduate students), Yuko Takada and Kozue Uematsu (Research assistants). Our collaborators are; Yasushi Hiromi (Professor, Division of Developmental Genetics), Hiroshi Handa (Professor, Frontier Collaborative Resarch Center, Tokyo Institute of Technology), Tadashi Wada (Associate professor, Faculty of Bioscince and Biotechnology, Tokyo Institute of Technology), Koji Akasaka (Associate professor, Graduate Department of Gene Science, Hiroshima University), Katsuyuki Hamada (Faculty of Medicine, Ehime University).

Research activities in 2002 are supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Science, Sports, Culture and Technology of Japan to Hirose and Ueda, Grants for Collaborative Research from our Institute to Hirose, and a Grant from Naito Science Foundation to Hirose.

(1) Role of GAGA factor-dependent chromatin remodeling in *Hox* gene expression

Tsukasa SHIMOJIMA, Masahiro OKADA[1], Takahiro NAKAYAMA, Hitoshi UEDA, Hiroshi HANDA[2], Tadashi WADA[3] and Susumu HIROSE ([1]Present address, Laboratory of Molecular Genetics, RIKEN Tsukuba Institute, Tsukuba, Ibaragi-ken 305-0074, Japan; [2]Frontier Collaborative Research Center, and [3]Faculty of Bioscience and Biotechnology, Tokyo Institute of Technology, Nagatsuda, Midori-ku, Yokohama 226-8501, Japan)

The identity of the body segment of metazoans is determined by the expression patterns of *Hox* genes. *Hox* gene expression is initiated by the actions of many transcription factors during embryogenesis. Curiously, these expression patterns are maintained after disappearance of the transcription factors. The active or inactive state is memorized in the chromatin structure and transmitted to daughter cells through cell division. Although the mechanism of this epigenetic gene expression is not clear, genetic studies have identified many genes involved in the process. These genes are classified into two groups. *Pc* group genes mainly govern the maintenance of the inactive state, while *trx* group genes are mainly responsible for the maintenance of the active state. *Drosophila* GAGA factor is encoded by a *trx* group gene *Trithorax-like* and directs chromatin remodeling to its binding sites. We found that *Drosophila* FACT, a heterodimer of dSPT16 and dSSRP1, is associated with GAGA factor through its dSSRP1 subunit, binds to necleosomes and facilitates GAGA factor-directed chromatin remodeling. Moreover, genetic interactions between *Trithorax-like* and *spt16* implicate the GAGA factor-FACT complex in expression of *Hox* genes *Ultrabithorax, Sex combs reduced* and *Abdominal-B*. Chromatin immunoprecipitation experiments indicated presence of the GAGA factor-FACT complex in the regulatory regions of *Ultrabithorax*. These data illustrate a crucial role of the GAGA factor-FACT complex in the modulation of chromatin structure for regulation of gene expression (Oral presentations 1, 5, 6, 8-10, 14 and 18). We have initiated collaboration with Dr. T. Wada to extend studies on chromatin transcription.

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

Kuniharu MATSUMOTO, Hirofumi FURUHASHI, Hitoshi UEDA, Koji AKASAKA[1] and Susumu HIROSE ([1]Graduate Department of Gene Science, Hiroshima University, Kagamiyama, Higashi-Hiroshima 739-8526, Japan)

Recent studies have established that chromatin structure is important for transcriptional regulation. Starting from analyses of DNase I hypersensitivity in 1980s, lines of evidence have been accumulated for changes in the chromatin structure, such as DNA methylation, chemical modifications of histones and chromatin remodeling during regulation of gene expression. However, besides transcription-driven supercoiling of DNA, our knowledge on conformation of chromatin DNA was elusive due to the lack of proper probes for analyses of DNA topology in vivo. To circumvent the situation, we have developed a method for visualization of negatively supercoiled DNA within the interphase genome. Using the methods, we observed unconstrained negative supercoils in many loci on polytene chromosomes of *Drosophila melanogaster* (Oral presentations 7 and 11).

Supercoiling factor (SCF) is a protein capable of introducing negative supercoils into DNA in conjunction with DNA topoisomerase II. Partial knock-out of the SCF function in *Drosophila melanogaster* by RNAi indicated that SCF is involved in superactivation of transcription on male X-chromosome (Oral presentations 4 and 12). In addition, we made studies on a *Drosophila* counterpart of sea urchin insulator binding protein GSAP in collaboration with Dr. K. Akasaka.

(3) Functional analysis of transcriptional coactivator MBF1

Qin-Xin LIU, Marek JINDRA, Hitoshi UEDA, Yasushi HIROMI[1] and Susumu HIROSE ([1]Division of Developmental Genetics)

During gene activation, the effect of binding of transcription factors to cis-acting DNA sequences is transmitted to RNA polymerase by means of coactivators. Although coactivators contribute to the efficiency of transcription, their developmental roles are poorly understood. We used *Drosophila* to conduct molecular and genetic dissection of an evolutionarily conserved but unique coactivator, Multiprotein bridging factor 1 (MBF1), in a multicellular organism. Through immunoprecipitation MBF1 was found to form a ternary complex including MBF1, TATA-binding protein (TBP) and a bZIP protein Tracheae defective (TDF)/Apontic. We have isolated a *Drosophila* mutant lacking the *mbf1* gene. In the *mbf1* mutant, no stable association between TBP and TDF is detectable, and transcription of a TDF-dependent reporter gene is reduced five fold. Although the null mutants of *mbf1* are viable, *tdf* becomes haplo-insufficient in *mbf1* deficient background, causing severe lesions in tracheae and the central nervous system, similar to those resulting from a complete loss of *tdf* function. These data demonstrate a critical role of MBF1 in the development of tracheae and central nervous system (Article 1, and Oral presentations 2 and 15)

(4) Mechanism of transcriptional regulation of the FTZ-F1 gene

Yasuo AGAWA, Susumu HIROSE and Hitoshi UEDA

Insect FTZ-F1 is a unique transcription factor, because it is induced after a pulse exposure of ecdysteroids and is expressed just before hatching, ecdysis and pupation in a stage specific manner. The stage specific expression of FTZ-F1 is necessary for larval ecdysis and metamorphosis, indicating that temporary precise regulation of the gene is important for the development of *Drosophila*. Factor I-4 has been identified as a sequence specific binding factor that binds to a *cis*-regulatory region of the FTZ-F1 gene. The factor is present during high ecdysteroid periods, so that it was thought as a repressor of the FTZ-F1 gene. However, reporter assays using *FTZ-F1* regulatory regions fused to *LacZ* showed that the recognition sequence of Factor I-4 is required for high-level expression of *FTZ-F1*. We purified Factor I-4 from embryonic extracts based on the affinity to its recognition sequence. Sequence analysis of the purified polypeptide revealed that it is a zinc finger protein carrying several other interesting motifs. Developmental Northern blotting analysis showed that several mRNA isofoms are expressed from the Factor I-4 gene and two of them are expressed during the high ecdysteroid periods around prepupa in coincidence with the developmental expression profile of Factor I-4

detected by gel mobility shift assays. These results suggest that the Factor I-4 gene belongs to an early gene in Ashburner's model and potentiates expression of the FTZ-F1 gene (Oral presentation 16 and 19).

(5) Role of FTZ-F1 during development (Regulation mechanism of gene expression during insect metamorphosis)

Haruhisa KAWASAKI[1], Yasunari KAYASHIMA, Masa-aki YAMADA, Susumu HIROSE and Hitoshi UEDA ([1]Present address, Advanced Research Institute for Science and Engineering, Waseda University)

Insect metamorphosis is an important event to change their morphology from larva to completely different adult. During the metamorphosis, many transcription factors are induced by ecdysteroids, but little is known how these factors regulate space specific expression of their target genes. To understand molecular mechanism underlying space specific gene expression during metamorphosis, regulatory regions of the *EDG84A* gene which is expressed in the anterior epidermis during the mid- to late prepupal periods depending on *Drosophila* temporal regulator FTZ-F1 were analyzed. Using transgenic fly reporter gene carrying various *cis*-regulaory regions fused to *LacZ*, we identified at least two positive and three negative elements for the space specific expression. The results suggested that spatial regulation is established by a combination of several transcription factors. Candidates of regulatory factors which bind to the regulatory regions were detected by gel mobility shift assays. By using computer search and Yeast two-hybrid assays, we identified a factor which binds to one of the positive regulatory elements. On the other hand, analysis of *cis*-regulatory region in the *EDG78E* gene which is expressed in whole epidermis during the mid- to late prepupal periods represented that FTZ-F1 is necessary for expression in the anterior epidermis but expression in the posterior epidermis is not completely dependent on FTZ-F1. The results suggested complex mechanisms that govern the temporal and spatial regulation of gene expression during metamorphosis (Article 2, and oral presentation 3, 17, 20 and 21).

During early metamorphosis larval salivary glands disappear by histolysis. For the histolysis, expression of FTZ-F1 during the mid- to late prepupal periods is necessary as revealed by persistence of larval salivary glands in *ftz-f1* mutant. Forced expression of Z1 isofom of BR-C or E74A during late prepupal period in *ftz-f1* mutant recovered the histolysis. These results suggest that BR-C and E74A work redundantly in the downstream of FTZ-F1 in the histolysis.

2. Publications

(1) Articles

1. Liu, Q.-X., Jindra, M., Ueda, H., Hiromi, Y. and Hirose, S.: *Drosophila* MBF1 is a co-activator for Tracheae Defective and contributes to the formation of tracheal and nervous systems. Development, in press.

2. Kawasaki, H., Hirose, S. and Ueda, H.: βFTZ-F1 dependent and inepedent activation of *Edg78E*, a pupal cuticle gene, during the early metamorphic period in *Drosophila melanogaster*. Develop. Growth Differ. **44**, 419-425, 2002.

(2) Oral presentations

1. Hirose, S., Nakayama, T., Shimojima, T., Okada, M. and Ueda, H.: Epigenetic gene expression through GAGA factor-dependent chromatin remodeling. 7th Asian Conference on Transcription, Kuala Lumpur, Malasia, July, 2002.

2. Liu, Q.-X., Jindra, M., Ueda, H., Hiromi, Y. and Hirose, S.: MBF1 serves as a transcriptional coactivator of Tracheae defective/Apontic during development of the trachea and central nervous system in *Drosophila melanogaster*. 43rd Annual Drosophila Research Conference, San Diego, California, USA, April, 2002.

3. Ueda, H., Kayashima, Y. and Hirose, S.: Mechanism of space specific gene expression in response to ecdysteroids-Transcriptional regulation of FTZ-F1 target gene EDG84A. 43rd Annual Drosophila Research Conference, San Diego, California, USA, April, 2002.

4. Furuhashi, H. and Hirose, S.: Functional analysis of supercoiling factor by using RNAi. 19th Workshop on Chromosome, Kobe, Japan, Jan, 2002.

5. Hirose, S.: GAGA factor-dependent chromatin remodeling and epigenetics. Seminar in Waseda University, Tokyo, Japan, March, 2002.

6. Hirose, S.: FACT is involved in epigenetic gene expression. Seminar in BERI, Osaka, Japan, June, 2002.

7. Matsumoto, K. and Hirose, S.: In situ analyses on negative supercoiling of DNA and transcription. NIG workshop on Biology of Chromatin, Mishima, Japan, November, 2002.

8. Hirose, S.: FACT is involved in epigenetic maintenance of gene expression. Seminar in Tsukuba University, Tsukuba, Japan, November, 2002.

9. Hirose, S.: Chromatin remodeling and maintenance of gene expression. 12th Workshop on Radiation Biology, Mito, Japan, November, 2002.

10. Hirose, S., Shimojima, T., Nakayama, T. and Ueda, H.: FACT is involved in epigenetic maintenance of gene expression. Symposium on Infrastructures of Chromatin Function, 25th Annual Meeting of Molecular Biology Society of Japan, Yokohama, Japan, December, 2002.

11. Matsumoto, K. and Hirose, S.: Visualization of DNA supercoiling and transcription in situ. 25th Annual Meeting of Molecular Biology Society of Japan, Yokohama, Japan, December, 2002.

12. Furuhashi, H. and Hirose, S.: Functional analysis of DNA supercoiling factor in Drosophila melanogaster. 25th Annual Meeting of Molecular Biology Society of Japan, Yokohama, Japan, December, 2002.

13. Ueda, H. and Hirose, S.: Chromosome pairing-dependent gene regulation in Drosophila. 25th Annual Meeting of Molecular Biology Society of Japan, Yokohama, Japan, December, 2002.

14. Nakayama, T., Shimojima, T. and Hirose, S.: Dynamic distribution of *Drosophila* GAGA factor-dFACT complex on *hsp70* gene. 25th Annual Meeting of Molecular Biology Society of Japan, Yokohama, Japan, December, 2002.

15. Liu, Q.-X., Ueda, H., Hiromi, Y. and Hirose, S.: Functional analysis of *Drosophila* TDF during the neuronal development. 25th Annual Meeting of Molecular Biology Society of Japan, Yokohama, Japan, December, 2002.

16. Agawa, Y., Hirose, S., Handa, H., Wada, T., Shinohara, A., Iwamatsu, A. and Ueda, H.: Mechanism of transcriptional regulation of the *FTZ-F1* gene in *Drosophila melanogaster.* 25th Annual Meeting of Molecular Biology Society of Japan, Yokohama, Japan, December, 2002.

17. Kyashima, Y., Hirose, S. and Ueda, H.: Candidates for regulator of space-specific expression of FTZ-F1-target gene *EDG84A* in *Drosophila.* 25th Annual Meeting of Molecular Biology Society of Japan, Yokohama, Japan, December, 2002.

18. Hirose, S.: Maintenance of gene expression through chromatin structure. Seminar in Tokyo Institute of Technology, Yokohama, Japan, December, 2002.

19. Agawa, Y., Hirose, S. and Ueda, H.: Regulation mechanism of the FTZ-F1 gene. Workshop on Insect 02 Hayama, Japan, March, 2002.

20. Kayashima, Y., Hirose, S. and Ueda, H.: Control mechanism to produce tissue specific gene expression during insect metamorphosis. Workshop on Insect 02 Hayama, Japan, March, 2002.

21. Ueda, H., Kayashima, Y. and Hirose, S.: Mechanism of gene expression in insects. Workshop on Insect 02 Hayama, Japan, March, 2002.

C-c. DIVISION OF MOLECULAR AND DEVELOPMENTAL BIOLOGY

(1) Genetic studies on vertebrate development using zebrafish as a model animal: Development of transposon-mediated insertional mutagenesis methods in zebrafish

Koichi KAWAKAMI

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. Taking advantage of these merits, large-scale chemical mutagenesis screens in zebrafish had been successfully carried out and hundreds of mutants defective in every aspect of early developmental processes had been collected. It has, however, not been easy to identify the mutated genes since most of the chemically induced mutations were point mutations and time-consuming positional cloning approaches are required to clone the genes. Therefore it is important to develop insertional mutagenesis methods in zebrafish, such as the P transposon system in Drosophila.

The *Tol2* element is a transposable element identified from the genome of the medaka fish. Previously we have found that the *Tol2* element encodes a fully functional transposase and have shown 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, that has ever been shown to encode a functional transposase, and thereby from which an autonomous element has been identified. We have been interested in developing novel insertional mutagenesis methods in zebrafish using the *Tol2* element. Using this *Tol2* transposon system, we successfully established a protocol to generate transgenic zebrafish very efficiently. We then made an exon trap construct with a promoter-less GFP gene using the *Tol2* transposon, and generated hundreds of insertions of the exon trap construct in the zebrafish genome. 10% of the insertions yielded "specific" GFP expression, indicating that the gene (exon) trap method using the *Tol2* transposon system is successfully working. In those fish, GFP was expressed the heart, forebrain, midbrain, midbrain-hindbrain boundary, hindbrain, notochord, floorplate, etc. Our current goal is to establish methods to identify genes expressed in a spatially and temporally regulated fashion and play an important role in development of the structure or the organ where the GFP expression is detected. Efforts are in progress toward this goal.

(2) Characterization of the *hagoromo* gene from cichlid fishes in the East Africa

Yohey TERAI,[1] Naoko MORIKAWA[1], Koichi KAWAKAMI and Norihiro OKADA[1] ([1]Tokyo Institute of Technology)

Previously we have developed an insertional mutagenesis method in zebrafish using a pseudotyped retrovirus. By performing this method, we isolated a dominant mutation *hagoromo* (*hag*), which caused disorganized stripe patterns on the skin of the adult zebrafish. In the *hag* mutant, the provirus was located within a novel gene, which we named the *hagoromo* gene, encoding a novel F-box/WD40-repeat protein. Interestingly, the mouse ortholog of the *hag* gene was mutated in the *Dactylaplasia* (*Dac*) mutant mouse. The *Dac* mutation is also dominant and causes defects in digit formation in fore- and hindlimbs. Thus, the *hag*/*Dac* locus is important for pattern formation in vertebrates but is involved in distinct morphogenetic events in different animals. In order to examine whether the *hag* gene is involved in divergence of pigment patterns on the skin of diverged fish species, we cloned homologs of the *hag* gene from different species of cichlid fishes in the East Africa. By comparing amino acid sequences of these Hag proteins, we found that the surface amino acid residues of the WD-repeat domain, which may interact with other proteins, changed rapidly as the fish species diverged. Thus, our studies suggested the changes in the amino acid sequences in the surface of the WD-repeat domain may be related to cichlid speciation by affecting their pigment patterns.

Publications

1. Terai, Y., Morikawa, N., Kawakami, K. and Okada, N.: Accelerated evolution of the surface amino acids in the WD-repeat domain encoded by the *hagoromo* gene in an explosively speciated lineage of east cichlid fishes. *Molecular Biology and Evolution* **19**, 574-578, 2002.

C-d. DIVISION OF PHYSIOLOGICAL GENETICS

Functional and structural analyses of mammalian Polycomb group comlexes

Haruhiko KOSEKI

Most of efforts of our group are paid to understand the functions of mammalian Polycomb group (PcG) multimeric comlexes during embryogenesis. We focused on identification and functional annotation of individual constituents of PcG comlexes.

(1) Ring1B

We have identified *Ring1B*, the product of the Ring1B gene, by means of its interaction with the Polycomb group protein Mel18. We describe biochemical and genetic studies directed to understand the biological role of Ring1B. Immunoprecipitation studies indicate that *Ring1B* form part of protein complexes containing the products of other Polycomb group genes, such as Rae28/Mph1 and M33 and that this complexes associate to chromosomal DNA. We have generated a mouse line bearing a hypomorphic *Ring1B* allele which shows posterior homeotic transformations of the axial skeleton and a mild derepression of some *Hox genes* (*Hoxb4, Hoxb6 and Hoxb8*) in cells anterior to their normal boundaries of expression in the mesodermal compartment. In contrast, the overexpression of Ring1B in chick embryos results in the repression of *cHoxb9* expression in the neural tube. These results, together with the genetic interactions observed in compound *Ring1B/Mel18* mutant mice are consistent with a role for Ring1B in the regulation of *Hox* gene expression by Polycomb group complexes.

(2) Scmh1

The product of the Scmh1 gene, a mammalian homologue of *Drosophila* Sex comb on midleg, interacts with two mammalian Polycomb group (PcG) proteins Mph2/Edr2 and rae28/Mph1 proteins through respective their SPM domains. Based on this observation we have generated a mutant allele of *Scmh1* designated as $Scmh1^{\Delta SPM}$, where the Scmh1 SPM domain has been deleted. $Scmh1^{\Delta SPM/\Delta SPM}$ mice exhibit posterior transformations of the axial skeleton although the penetrance is extremely low. In addition, $Scmh1^{\Delta SPM/\Delta SPM}$ primary embryonic fibroblasts exhibit premature entry into senescence and early mitotic arrest that is associated with overexpression of $p19^{ARF}$ and mis-segregation of chromosomes. A partially penetrant phenotype is also observed during spermatogenesis. Approximately half of the $Scmh1^{\Delta SPM/\Delta SPM}$ males are infertile which correlates with an accelerated apoptosis of spermatocytes in meiotic prophase I, and overexpression of p53, $p21^{WAF1/CIP1}$ and Bax.

Publications

1. Nakayama, T., Kasprowicz, D. J., Yamashita, M., Schubert, L. A., Gillard, G., Kimura, M., Didierlaurent, A., Koseki, H. and Ziegler, S. F.: The Generation of Mature, Single-Positive Thymocytes *In Vivo* is Dysregulated by CD69 Blockade or Overexpression. *J Immunol.* **168**, 87-94, 2002.

2. Miki, T., Suzuki, M., Shibasaki, T., Uemura,H., Sato, T., Yamaguchi, K., Koseki, H., Iwanaga, T., Nakaya, H. and Seino, S.: Mouse model of Prinzmetal angina by disruption of the inward rectifier *Kir6.1*. Nature Med. 8, 466-72, 2002.

3. Murakami, H., Okawa, A., Yoshida, H., Nishikawa, S., Moriya, H. and Koseki, H.: *Elbow knee synostosis (Eks)*: A new mutation on mouse Chromosome 14. *Mammalian Genome* **13**, 341-344, 2002.

4. O-Wang, J., Kajiwara, K., Kawamura, K., Kimura, M., Miyagishima, H., Koseki, H. and Tagawa, M. An essential role for REV3 in mammalian cell survival: absence of REV3 incuces p53-independent embryonic death. *Biochem Biophys Res Commun.* **293(3)**, 1132-1137, 2002.

5. Ohara, O., Nagase, T., Mitsui, G., Kohga, H., Kikuno, R., Hiraoka, S., Takahashi, Y., Kitajima, S., Saga, Y. and Koseki, H.: Characterizaion of size-fractionated cDNA libraries generated by the in vitro recombination-assisted method. *DNA Research.* **9**, 47-57, 2002.

6. Yamaki, M., Isono, K., Takada, Y., Abe, K., Akasaka, T., Tanzawa, H. and Koseki, H.: The mouse Edr2 (Mph2) gene has two forms of mRNA encoding 90- and 36-kDa polypeptides. *Gene*; **288**,103-110, 2002.

7. Suzuki, M., Mizutani-Koseki, Y., Fujimura, Y., Miyagishima, H., Kaneko, T., Takada, Y., Akasaka, T., Tanzawa, H., Takihara, Y., Nakano, M., Masumoto, H., Vidal, M., Isono, K. and Koseki, H.: Involvement of the Polycomb-group gene Ring1B in the specification of the anterior-posterior axis in mice. *Development* **129**, 4171-4183, 2002.

8. Zhang, Y., Miki, T., Iwanaga, T., Koseki, Y., Okuno, M., Sunaga, Y., Ozaki, N., Yano, H., Koseki, H. and Seino, S.: Otx3: identification, tissue expression and functional characterization of a novel member of the Otx family. *J Biol Chem.* **277**, 28065-28069, 2002.

9. Ishigami, A., Fujita, T., Handa, S., Shirasawa, T., Koseki, H., Kitamura, T., Enomoto, N., Sato, N., Shimosawa, T. and Maruyama, N.: Senescence marker protein-30 knockout mouse liver are highly susceptible to TNF- and Fas-mediated apoptosis. *Am. J. Pathol.* **161**, 1273-1281, 2002.

10. Akasaka, T., Takahashi, N., Suzuki, M., Koseki, H., Bodmer, R. and Koga, H.: MBLR, a new RING finger protein resembling mammalian Polycomb gene products, is regulated by cell cycle-dependent phosphorylation. *Genes Cells.* **7**, 835-850, 2002.

11. Takedam, U., Utani, A., Wu, J., Adachi, E., Koseki, H., Taniguchi, M., Matsumoto, T., Ohashi, T., Sato, M. and Shinkai, H.: Targeted Disruption of Dermatopontin Causes Abnormal Collagen Fibrillogenesis. *J Invest Dermatol.* **119**, 678-683, 2002.

12. Yuasa, S., Nakajima, M., Aizawa, H., Sahara, N., Koizumi, K., Sakai, T., Usami, M., Kobayashi, S., Kuroyanagi, H., Mori, H., Koseki, H. and Shirasawa, T.: Impaired cell cycle control of neuronal precursor cells in the neocortical primordium of presenilin-1-deficient mice. *J Neurosci Res* **70**, 501-513, 2002.

13. Shirasawa, T., Izumizaki, M., Suzuki, YI., Ishihara, A., Shimizu, T., Tamaki, M., Huang, F., Koizumi, KI., Iwase, M., Sakai, H., Tsuchida, E., Ueshima, K., Inoue, H., Koseki, H., Senda, T., Kuriyama, T. and Homma, I.: Oxygen affinity of hemoglobin regulates O2 consumption, metabolism, and physical activity. *J Biol Chem* in press, 2002.

(1) Nucleosome positioning and regulation of gene expression

Ryoiti KIYAMA[1], Yoshiaki ONISHI[1] and Yuko WADA-KIYAMA[2] ([1]Res. Center for Glycoscience, AIST and [2]Nippon Med. Sch.)

The cellular regulation of expression of a gene at a given time under specific stimulation is a central tissue in the study of differentiation, organogenesis, development and phenotypes as in diseases. The genetic information in individual cells should be expressed at specific times in the following steps: first, modulating a higher-order chromatin structure to change a repressive environment, second, binding factors to promoters and/or enhancers, and then modifying nucleosome structure for environmental settings and finally initiating transcription. Therefore, at the first step, transcription is initiated under the repressive environment including the packed chromatin, by binding and functioning of the very initial components in the transcriptional apparatus in the chromatin architecture. In the β-globin gene expression, an erythroid-specific enhancer binding protein NF-E2 is critically involved in remodeling the nucleosome structure over the DNase I hypersensitive site 2 (HS2) of the human β-globin locus control region (β-LCR). We obtained evidence of interaction of NF-E2 to the cognate motif at HS2 in a repressive state. First, the nucleosome containing the NF-E2 binding site showed characteristic rotational and translational phases. Second, the binding site had less affinity for the histone octamers than the nearby regions while showing higher accessibility to DNase I and micrococcal nuclease. Finally, the motif was recognized by the exogenous NF-E2 protein expressed in non-erythroid HeLa cells as shown by ligation-mediated PCR and chromatin immunoprecipitation assay. These lines of evidence indicate that NF-E2 interacts with the cognate motif on the nucleosome before chromatin is remodeled, which may be a rate-limiting step and critical for activation of globin gene expression. Taken together with our previous findings that the nucleosome over HS2 was aligned by a key nucleosome located at a distance of two nucleosomes from HS2 and DNA curvature had a significant role in this nucleosome positioning, these results suggest that not only transcription factor binding sites on the genome but also the chromatin structure around HS2 arranged by a specific DNA structure and changing its status is critical in the early steps of β-globin gene expression.

(2) A novel gene, *Kank*, is a potential growth suppressor of renal cell carcinoma

Badal C. ROY[1], Shubhashish SARKAR[1], Naoya HATANO[1], Kazuo GOHJI[2], Teiichiro AOYAGI[3] and Ryoiti KIYAMA[1] ([1]Res. Center for Glycoscience, AIST, [2]Dept. Urol., Osaka Med. College, [3]Tokyo Dental College)

We identified a potential tumor suppressor gene named *Kank* (for kidney ankyrin repeat-containing protein) on chromosome 9p24 by using a genome

subtraction method followed by comprehensive analysis of loss of heterozygosity (LOH). The gene contained an open reading frame for a protein of 1194 amino acids long with an estimated MW of 130 kD. It also contained some interesting motifs including three coiled-coil domains in the N-terminus and four ankyrin-repeats in the C-terminus. The protein was localized in the cytosol as revealed by immunostaining. The *Kank* gene showed LOH in about 25% RCC patients and its expression was suppressed in tumor tissues examined by RT-PCR, immunostaining and Western blotting. Interestingly, the genomic DNA from RCC tumors and several kidney cancer cell lines showed hypermethylation at CpG sites of the gene. The gene showed growth suppression both *in vivo* and *in vitro*. Overexpression of the *Kank* gene in expression-negative HEK293 and G-402 cells showed dramatic morphological changes which seemed to be mediated by the cytoskeletal changes. Our results demonstrate that Kank protein may function as a growth regulator through the cytoskeleton and may play an important role in cell migration as well as cell growth.

(3) Development and application of DNA microarrays using estrogen responsive genes

Ryoiti KIYAMA

Estrogen is a member of the steroid-thyroid hormones and regulates development and maintenance of female phenotype and behavior. The primary target of estrogen is the estrogen receptor and the estrogen-estrogen receptor complex binds to a number of secondary target genes, which eventually regulate production of growth factors as a part of hormone actions. The chemicals that mimic estrogen activity, so called endocrine disruptors, are becoming worldwide problems. To monitor estrogen activity in the environment as well as in the commercial products and foods and understand the mechanisms of the effects of endocrine disruptors on the human body, we developed a microarray-based technique by means of profiling estrogen responsive genes. Approximately 200 cDNAs from the genes and ESTs (expressed sequence tags) were selected after screening approximately 9,200 genes for the estrogen response and they were spotted on a custom microarray, the EstrArray. The genes and ESTs selected were classified into several groups according to their functions: cancer-related genes, the genes for growth and proliferation, structural genes and others. Using EstrArrays, we first investigated the time course of the responses of the genes in MCF-7 cells to an estrogen, 17 β-estradiol. The estrogen responsive genes were classified according to this result into the early and late responding types. The results were confirmed by Northern blot analysis. We then performed the assay for profiling gene expression in MCF-7 cells after chemical treatment with bisphenol A, nonylphenol, dioxin and others. The EstrArray system can be used in two different major fields: environmental monitoring, and diagnosis/drug screening. Both systems need accurate and high throughput analysis of cellular responses which can be monitored by the gene response. In the first field, gene expression profiling of endocrine disruptors will be a useful tool for monitoring effects of industrial chemicals, and natural and artificial biomaterials. In the second field, the EstrArray analysis of estrogen receptor-positive cancer cell lines from various tissues showed similar but not identical profiles, indicating the usefulness of this custom microarray system for monitoring estrogen responses in various tissue types, which is important for cancer therapy and production of anti-estrogen drugs.

Publications

1. Li, X.-M., Onishi, Y., Kuwabara, K., Rho, J.-y., Wada-Kiyama, Y., Sakuma, Y. and Kiyama, R.: Ligand-dependent transcriptional enhancement by DNA curvature between two half motifs of the estrogen response element in the human estrogen receptor α gene. *Gene* **294**, 279-290, 2002.

2. Kiyama, R. and Trifonov, E. N.: What positions nucleosomes? ---A model (Review). *FEBS Letters* **523**, 7-11, 2002.

3. Onishi, Y. and Kiyama, R.: Interaction of NF-E2 in the human β-globin locus control region before chromatin remodeling. *J. Biol. Chem.* **278**, 8163-8171, 2003.

4. Wanapirak, C., Kato, M., Onishi, Y., Wada-Kiyama, Y. and Kiyama, R.: Evolutionary conservation and functional synergism of curved DNA at the mouse ε- and other globin-gene promoters. *J. Mol. Evol* (in press).

5. Nadkarni, A., Sakaguchi, T., Takaku, H.,

Gorakshakar, A., Phanasgaonkar, S., Colah, R., Mohanty, D. and Kiyama, R.: A novel β0-thalassemia mutation at codon 55 (-A) and a rare 17-bp deletion at codons 126-131 in the Indian population. *Hemoglobin* **26**, 41-47, 2002.

6. Sarkar, S., Roy, B. C., Aoyagi, T., Gohji, K. and Kiyama, R.: A novel ankyrin repeat-containing gene Kank located at 9p24 is a growth suppressor of renal cell carcinoma. *J. Biol. Chem.* **277**, 36585-36591, 2002.

7. Inoue, A., Yoshida, N., Omoto, Y., Oguchi, S., Yamori, T., Kiyama, R. and Hayashi, S.: Development of cDNA microarray for expression profiling of estrogen-responsive genes. *J. Mol. Endocrinol.* **29**, 175-192, 2002.

8. Inoue, A., Hayashi, S.-i., Aoyagi, K., Nishigaki, M., Sasaki, H. and Kiyama, R.: Evaluation of tissue-specific response to estrogen activity by means of a reporter gene assay based on the differential use of promoters A to F of the human estrogen receptor α gene. *J. Pharmacol. Toxicol. Methods* **47**, 129-135, 2002.

D. DEPARTMENT OF POPULATION GENETICS

D-a. DIVISION OF POPULATION GENETICS

(1) Adaptive evolution of the IgA hinge region in primates

Kenta SUMIYAMA, Naruya SAITOU and Shintaroh UEDA[1]
([1]Univ. Tokyo, Grad. Sch. Science)

IgA is a major component that prevents the penetration of pathogenic bacteria into mucosal surfaces. The IgA antibody is cleaved at the IgA hinge region with high specificity by IgA-specific proteases produced by several pathogenic bacteria. We conducted a genomic sequence analysis of the IgA genes of a wide spectrum of primates, including the first intron and second exon, which consist of the hinge region and the CH2 domain, to find evidence of positive selection. Because the hinge region is quite small, we combined the largest collection of sequences that could be clearly aligned and evaluated the total numbers of synonymous and nonsynonymous substitutions on the phylogenetic tree. The nonsynonymous to synonymous substitution ratio (dN/dS test) showed that hominoids, Old World monkeys, and New World monkeys have dN/dS ratios of 5.4, 6.3, and 4.2, respectively. Fisher's exact probability tests showed statistical significance for the Old World monkey. Because the substitution rates of the flanking sequences are more or less similar to the synonymous rates of the hinge region, these high values of dN/dS should be the result of positive selection at the hinge region. Combining the high sequence variability in each population and the highly accelerated nonsynonymous substitution rates in the hinge region, we conclude that this unusual IgA evolution is a molecular evidence of adaptive evolution possibly caused by the host-parasite relationship. For details, see ref. (1).

(2) NJML+P: A hybrid algorithm of the maximum likelihood and neighbor-joining methods using parallel computing

Satoshi OOTA and Naruya SAITOU

The NJML method [2, 3] is a hybrid algorithm of the two well-known methods to reconstruct molecular phylogenetic trees: the neighbor-joining (NJ) method [4] and the maximum likelihood (ML) method [1]. The NJML method is considerably efficient both in reliability and speed comparing with the other existing ML-based methods. By giving appropriate parameters, the NJML method gradually approaches the exhaustive topology search, that is supposed to be the most accurate way to find the true phylogenetic tree. However, it is obvious that more exhaustive searches require more computational time. We have implemented the NJML method by using parallel computing to reduce the computational time, in which the bootstrap trials and the maximum likelihood estimation were parallelized. For details, see ref. (2).

(3) Extreme mtDNA homogeneity in continental Asian populations

Hiroki OOTA[1], Takashi KITANO, Feng JIN[2], Isao YUASA[3], Li WANG[2], Shintaroh UEDA[3], Naruya SAITOU and Mark STONEKING[1] ([1]Max Planck Institute for Evolutionary Anthropology, [2]Institute of Genetics, Beijing, [3]Univ. Tokyo, Grad. Sch. Science)

Mitochondrial DNA (mtDNA) variation in continental Asia has not been well-studied. Here, we report mtDNA HV1 sequences for 84 Xi'an and 82 Changsha Han Chinese, 89 Honshu Japanese, and 35 Vietnamese. Comparison of these sequences with other Asian mtDNA sequences reveals high variability within populations, but extremely low differentiation among Asian populations. Correlations between genetic distance and geographic distance, based on mtDNA and Y chromosome variation, indicate a higher migration rate in females than in males. This may reflect patrilocality, as suggested previously, but another plausible hypothesis is that the demographic expansion associated with the spread of agriculture in Asia may be responsible for the extreme genetic homogeneity in Asia. For details, see ref. (3).

(4) A large-scale analysis of human mitochondrial DNA sequences with special reference to the population history of East Eurasian

Hiroki OOTA[1], Naruya SAITOU and Shintaroh UEDA[2] ([1]Max Planck Institute for Evolutionary Anthropology, [2]Univ. Tokyo, Grad. Sch. Science)

Ancient DNA analysis is very powerful for the studies on past human populations. However, in most cases ancient DNA is extremely degraded into short fragments, and the information is limited because of the damaged state. A large-scale data analysis of human mitochondrial DNA (mtDNA) was carried out to assess validity of the short nucleotide sequence for closely related human populations. We retrieved mtDNA data from the DDBJ/EMBL/GenBank nucleotide sequence database and constructed a data set containing 414 distinct mtDNA types derived from 19 populations of East Eurasia and the surrounding area. A series of new procedures were applied and an mtDNA phylogenetic tree was constructed. Six major star-like clusters were observed in this tree, and the corresponding six radiation groups (I-VI) were characterized. Frequency distributions of each radiation group showed remarkable difference in each geographical area, suggesting that the short mtDNA nucleotide sequences were valuable in analyzing ancient human populations. The efficient procedure for data analysis will enhance the usefulness of ancient DNA data. Additionally, we discuss a possibility of two human migration routes from Africa to East Eurasia based on the mtDNA tree topology and the coalescence times in each radiation group. For details, see ref. (4).

(5) Understanding the dynamics of Spinocerebellar Ataxia 8 (SCA8) locus through a comparative genetic approach in humans and apes

Aida M. ANDRES[1], M. SOLDEVILA[1], Naruya SAITOU, F. CALAFELL[1], Jaume BERTRANPETIT and others ([1]Univ. Pompeu Fabra, Barcelona)

Spinocerebellar Ataxia 8 (SCA8) is a neurodegenerative disorder caused by expansion of a trinucleotide repeat. We undertake a comparative genetic analysis among human populations and primate species in the normal variation range, where forces that shaped present diversity can be recognised. We determinate number of repeats of the short tandem repeat through allele length sizing and sequencing methods. Human allele distributions are very similar among populations, ruling out ethnicity as a genetic risk for allele expansion. Primate comparison shows human-specific features, with longer human alleles due to a novel variable trinucleotide repeat, not present in non-human primates, which increased the disease-causing expansion likelihood. SCA8 seems to be a human specific disease. For details, see ref. (5).

(6) Helicobacter pylori in North and South America before Columbus

Yoshio YAMAOKA[1], Naruya SAITOU, David Y. GRAHAM[1] and others ([1]Baylor College of Medicine, Houston)

Abstract We present a molecular epidemiologic study, based on an analysis of vacA, cagA and cag right end junction genotypes from 1042 Helicobacter pylori isolates, suggesting that H. pylori was present in the New World before Columbus. Eight Native Colombian and Alaskan strains possessed novel vacA and/or cagA gene structures and were more closely related to East Asian than to non-Asian H. pylori. Some Native Alaskan strains appear to have originated in Central Asia and to have arrived after strains found in South America suggesting that H. pylori crossed the Bering Strait from Asia to the New World at different times. For details, see ref. (6).

(7) A phylogenomic study of the OCTase genes in Pseudomonas syringae pathovars: the horizontal transfer of the argK-tox cluster and the evolutionary history of OCTase genes on their genomes

Hiroyuki SAWADA[1], Shigehiko KANAYA[2], Masayuki TSUDA[3], Fumihiko SUZUKI[1], Kozi AZEGAMI[1] and Naruya SAITOU ([1]National Institute for Agro-Environmental Sciences, Tsukuba, [2]Nara Institute of Science and Technology, [3]Tohoku Univ.)

Phytopathogenic Pseudomonas syringae is subdivided into about 50 pathovars due to their conspicuous differentiation with regard to pathogenicity. Based on the results of a phylogenetic analysis of four genes (gyrB, rpoD, hrpL, and hrpS), Sawada et al. (1999) showed that the ancestor of P. syringae had diverged into at least three monophyletic

groups during its evolution. Physical maps of the genomes of representative strains of these three groups were constructed, which revealed that each strain had five rrn operons which existed on one circular genome. The fact that the structure and size of genomes vary greatly depending on the pathovar shows that P. syringae genomes are quite rich in plasticity and that they have undergone large-scale genomic rearrangements. Analyses of the codon usage and the GC content at the codon third position, in conjunction with phylogenomic analyses, showed that the gene cluster involved in phaseolotoxin synthesis (argK-tox cluster) expanded its distribution by conducting horizontal transfer onto the genomes of two P. syringae pathovars (pv. actinidiae and pv. phaseolicola) from bacterial species distantly related to P. syringae and that its acquisition was quite recent (i.e., after the ancestor of P. syringae diverged into the respective pathovars). Furthermore, the results of a detailed analysis of argK [an anabolic ornithine carbamoyltransferase (anabolic OCTase) gene] , which is present within the argK-tox cluster, revealed the plausible process of generation of an unusual composition of the OCTase genes on the genomes of these two phaseolotoxin-producing pathovars: a catabolic OCTase gene (equivalent to the orthologue of arcB of P. aeruginosa) and an anabolic OCTase gene (argF), which must have been formed by gene duplication, have first been present on the genome of the ancestor of P. syringae; the catabolic OCTase gene has been deleted; the ancestor has diverged into the respective pathovars; the foreign-originated argK-tox cluster has horizontally transferred onto the genomes of pv. actinidiae and pv. phaseolicola; and hence two copies of only the anabolic OCTase genes (argK and argF) came to exist on the genomes of these two pathovars. Thus, the horizontal gene transfer and the genomic rearrangement were proven to have played an important role in the pathogenic differentiation and diversification of P. syringae. For details, see ref. (7).

(8) Construction and analysis of a human-chimpanzee comparative clone map

Asao FUJIYAMA[1], Hidemi WATANABE[1], Atsushi TOYODA[1], Todd D. TAYLOR[1], Takehiko ITOH[1], Naruya SAITOU, Masahira HATTORI[1], Yoshiyuki SAKAKI[1] and others ([1]Riken Genomic Science Center, Yokohama)

The recently released human genome sequences provide us with reference data to conduct comparative genomic research on primates, which will be important to understand what genetic information makes us human. Here we present a first-generation human-chimpanzee comparative genome map and its initial analysis. The map was constructed through paired alignment of 77,461 chimpanzee bacterial artificial chromosome end sequences with publicly available human genome sequences. We detected candidate positions, including two clusters on human chromosome 21 that suggest large, nonrandom regions of difference between the two genomes. For details, see ref. (8).

Publications

1. Sumiyama, K., Saitou, N. and Ueda, S.: Adaptive evolution of the IgA hinge region in primates. *Molecular Biology and Evolution*, **19**, 1093-1099, 2002.

2. Oota, S. and Saitou, N.: NJML+P: A Hybrid Algorithm of The Maximum Likelihood and Neighbor-Joining Methods Using Parallel Computing. *Genome Informatics Series*, **13**, 434-435, 2002.

3. Oota, H., Kitano, T., Jin, F., Yuasa, I., Wang, L., Ueda, S., Saitou, N. and Stoneking, M.: Extreme mtDNA homogeneity in continental Asian populations. *American Journal of Physical Anthropology*, **118**, 146-153, 2002.

4. Oota, H., Saitou, N. and Ueda, S.: A large-scale analysis of human mitochondrial DNA sequences with special reference to the population history of East Eurasian. *Anthropological Science*, **110**, 293-312, 2002.

5. Andres, A. M., Soldevila, M., Saitou, N., Volpini, V., Calafell, F. and Bertranpetit, J.: Understanding the dynamics of Spinocerebellar Ataxia 8 (SCA8) locus through a comparative genetic approach in humans and apes. *Neuroscience Letters*, **336**, 143-146, 2002.

6. Yamaoka, Y., Orito, E., Mizokami, M., Gutierrez, O., Saitou, N., Kodama, T., Osato, M. S., Kim, J. G., Ramirez, F. C., Mahachai, V. and Graham, D. Y.: Helicobacter pylori in North and South America before Columbus. *FEBS Letter*, **517**, 180-184, 2002.

7. Sawada, H., Kanaya, S., Tsuda, M., Suzuki, F., Azegami, K. and Saitou, N. Phylogenomic study of

OCTase genes in Pseudomonas syringae pathovars: Horizontal transfer of argK-tox cluster and evolutionary history of OCTase genes on their genomes. *Journal of Molecular Evolution*, **54**, 437-457, 2002.

8. Fujiyama, A., Watanabe, H., Toyoda, A., Taylor, T. D., Itoh, T., Tsai, S.-F., Park, H.-S., Yaspo, M.-L,. Lehrach, H., Chen, Z., Fu, G., Saitou, N., Osoegawa, K., de Jong, P. J., Suto, Y., Hattori, M. and Sakaki, Y.: Construction and Analysis of a Human-Chimpanzee Comparative Clone Map. *Science*, **295**, 131-134, 2002.

9. Umetsu, K., Iwabuchi, N., Yuasa, I., Saitou, N., Clark, PF., Boxshall, G., Osawa, M. and Igarashi, K.: Complete mitochondrial DNA sequence of a tadpole shrimp (Triops cancriformis) and analysis of museum samples. *Electrophoresis*, **23**, 4080-4084, 2002.

10. Tateno, Y., Imanishi, T., Miyazaki, S., Fukami-Kobayashi, K., Saitou, N., Sugawara, H. and Gojobori, T.: DNA Data Bank of Japan (DDBJ) for genome scale research in life science. *Nucleic Acids Research*, **30**, 27-30, 2002.

(1) Genetic and molecular dissection of the within- and between-species variation in the sex comb tooth number in Drosophila

Haruki TATSUTA and Toshiyuki TAKANO-SHIMIZU

Sexual dimorphic characters are often among the first characters to change in the evolution of new species. Sex comb, a specific row of enlarged bristles on the first legs of males of some species of the subgenus *Sophophora*, is such an example. Indeed, the sex comb is a highly variable character between and within species. We first analyzed a within-species sex-comb-tooth number variation by the composite interval mapping for F2 samples between high (H) and low (L) sex-comb-tooth number lines of *D. simulans*, finding four significant QTL, two on each major autosome. Interestingly, one of them showed an overdominant effect. To study possible epistasis between QTL, we constructed the chromosome substitution lines and recombinant inbred lines. While we confirmed the significant effects of the two QTL on the third chromosome by using recombinant inbred lines, the results of QTL analysis for the second chromosome substitution lines differed between the two backgrounds. No QTL was found in the L background; we found three significant QTL in the H background, but their positions were different from the QTL positions estimated from the whole chromosome analysis. These results suggested that epistasis between QTL may have been important in the phenotypic evolution of this character.

(2) DNA variation in the gustatory and olfactory receptor gene families of *Drosophila melanogaster*

Akira KAWABE and Toshiyuki TAKANO-SHIMIZU

DNA variation in 55 gustatory and 54 olfactory receptor gene candidates were analyzed for eight *Drosophila melanogaster* and one *D. simulans* inbred strains. The average nucleotide diversity (π) was 0.008, which is similar to those for other *D. melanogaster* genes. A significant correlation was observed between the amount of intra-specific DNA variations in synonymous sites and those in non-coding regions, and both of them were significantly correlated with the level of local crossover frequency. Local mutation rate or local crossover frequency or both might play a significant role in shaping silent-site variation. A significant correlation was also found between the levels of intra- and inter-specific replacement-site DNA variations, suggesting that the strength of selective constraint varies among loci, but is constant between the two species. In both the gustatory and olfactory genes, the ratio of the numbers of replacement changes to synonymous ones was significantly higher in inter-specific divergence than in intra-specific polymorphism. The clustered genes showing high similarity within each cluster had higher level of DNA polymorphism than the single-copy genes in isolation. Gene conversion might be responsible for the high DNA polymorphism in the clustered genes. At ten loci of 109 loci analyzed, we found 12 putative null alleles: small indels causing flame shift, large coding-sequence deletion, and nonsense mutations. All of them were polymorphic in a sample. Because we analyzed only half to one-third of entire coding sequences, a single *D. melanogaster* fly may have eight or more null alleles for the gustatory and olfactory genes.

I should add that Ms Yuriko Ishii has contributed significantly to the above two works.

Publications

1. Furuse, T., Takano-Shimizu, T., Moriwaki, K., Shiroishi, T. and Koide, T.: QTL analyses of spontaneous activity by using mouse strains from Mishima battery. *Mamm. Genome* **13**, 411-415, 2002.

D-b. DIVISION OF EVOLUTIONARY GENETICS

(1) Informatics for unveiling hidden genome signatures

Takashi ABE, Yuta ICHIBA, Yoko KOSAKA, Nanayo ISHIHARA and Toshimichi IKEMURA

With the increasing amount of available genome sequences, novel tools are needed for comprehensive analysis of species-specific sequence characteristics for a wide variety of genomes. We used an unsupervised neural network algorithm, a self-organizing map (SOM), to analyze di-, tri-, and tetranucleotide frequencies in a wide variety of prokaryotic and eukaryotic genomes. The SOM, which can cluster complex data efficiently, was shown to be an excellent tool for analyzing global characteristics of genome sequences and for revealing key combinations of oligonucleotides representing individual genomes. From analysis of 1-kb and 10-kb genomic sequences derived from 65 bacteria and from 6 eukaryotes, clear species-specific separations of major portions of the sequences were obtained with the di-, tri-, and tetranucleotide SOMs. The unsupervised algorithm could recognize, in most 10-kb sequences, the species-specific characteristics (key combinations of oligonucleotide frequencies) that are signature features of each genome. We were able to classify DNA sequences within one and between many species into subgroups that corresponded generally to biological categories. Because the classification power is very high, the SOM is an efficient and fundamental bioinformatic strategy for extracting a wide range of genomic information from a vast amount of sequences. (Abe et al. Genome Research, **13**, 693-702, 2003).

(2) A novel bioinformatic strategy for unveiling hidden genome signatures of eukaryotes: Self-organizing map of oligonucleotide frequency

Takashi ABE, Yuta ICHIBA, Jian-Ping SONG, Yoko KOSAKA, Kiyomi KITA and Toshimichi IKEMURA

In addition to protein-coding information, genome sequences contain a wealth of information of interest in many fields of biology from molecular evolution to genome engineering and biomedicine. G+C% is used as a fundamental characteristic of individual genomes but too simple a parameter to differentiate a wide variety of genomes. Because oligonucleotide frequencies vary significantly among genomes, they can be used to distinguish genomes and genomic regions. We used an unsupervised neural network algorithm, Kohonen's self-organizing map (SOM), to analyze di- and trinucleotide frequencies in 9 eukaryotic genomes (a total of 1.2 Gb); S. cerevisiae, S. pombe, C. elegans, A. thaliana, D. melanogaster, Fugu, and rice, as well as P. falciparum chromosomes 2 and 3, and human chromosomes 14, 20, 21, and 22. Each genomic sequence with different window sizes was encoded as a 16- and 64-dimensional vector giving relative frequencies of di- and trinucleotides, respectively. From analysis of a total of 120,000 nonoverlapping 10-kb sequences and overlapping 100-kb sequences with a moving step size of 10 kb, derived from a total of the 1.2 Gb sequences, clear species-specific separations of most genomic sequences were obtained with the SOMs. The unsupervised algorithm could recognize, in most of the 10-kb sequences, the species-specific characteristics (key combinations of oligonucleotide frequencies) that are signature representations of each genome. (Abe et al., Genome Informatics Series, **13**, 12-20, 2002).

(3) Functional analyses of centromere of higher vertebrate cells

Tatsuo FUKAGAWA, Ai NISHIHASHI, Atsushi OKAMURA, Tetsuya HORI, Masahiro NOGAMI, Yoshikazu MIKAMI, Yoko MIYAUCHI, Jian-Ping SONG, Kazuko SUZUKI, Mitsuko YOSHIKAWA, Yuko FUKAGAWA and Toshimichi IKEMURA

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 also generated human artificial mini-chromosomes using the telomere-directed breakage method in DT40 cells. We have combined human artificial mini-chromosomes and DT40 cells to study chromosome segregation. 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. In this year we focused on the following subjects to the analyses of centromeres.

I) We identified a novel constitutive centromere protein, CENP-I, which shows sequence similarity with fission yeast Mis6 protein, and we showed that CENP-I is a constitutive component of the centromere that colocalizes with CENP-A, -C and -H throughout the cell cycle in vertebrate cells. To determine the precise function of CENP-I, we examined its role in centromere function and assembly by generating a conditional loss-of-function mutant in the chicken DT40 cell line. In the absence of CENP-I, cells accumulated BubR1 at kinetochores and arrested at prometaphase with mis-aligned chromosomes for long periods of time. Eventually, checkpoint function appeared to be lost and cells reentered the cell cycle without undergoing cytokinesis. Immunocytochemical analysis of CENP-I-deficient cells demonstrated that both CENP-I and CENP-H are necessary for CENP-C, but not CENP-A, localization to the centromere, indicating that CENP-I plays an essential role in centromere assembly in vertebrate cells. These experiments define an assembly pathway for the vertebrate kinetochore in which binding of CENP-A is followed by mutually interdependent targeting of CENP-H and CENP-I, and then by binding of CENP-C. (Nishihashi et al. Dev. Cell, 2002).

II) Dissection of human centromeres is difficult because of the lack of landmarks within highly repeated DNA. We have systematically manipulated a single human X centromere using homologous recombination and *de novo* telomere formation to generate a large series of minichromosome deletion derivatives. On the basis of the molecular structure of the minichromosomes recovered, the asymmetric distribution of CENP-A, -C and -H relative to the DXZ1 α-satellite array and the identification of a single major site of topoisomerase II activity we suggest that the kinetochore demarcates a fixed region of limited extent within the many megabases of centromerically-localised repeat. Moreover, our findings suggest a fundamental role for topoisomerase II at the centromere. (Spence et al., EMBO J., 2002)

(4) Chromosome-wide assessment of replication timing for human chromosomes 11q and 21q: Genome-synteny breakpoints between mouse and human chromosomes at or near replication-transition regions

Yoshihisa WATANABE, Yuta ICHIBA and Toshimichi IKEMURA

Completion of the human genome sequencing will greatly accelerates the development of a new branch of evolutionary genetics. Comparison of the human genome with other mammalian genome, particularly identification of genome-synteny breakpoints, aids in understanding the evolutionary history of human and other mammalian genomes. We measured replication timing of the entire lengths of human chromosomes 11q and 21q. Megabase-sized zones that replicate early or late in S phase (thus early/late transition) were defined at the sequence level and many disease-related genes were identified in the timing-switch regions. We then compared human and mouse genomes on a chromosome level, and found that synteny breakage between the two genomes occurred primarily at or near the replication-transition regions found on human chromosomes 11q and 21q. The finding that synteny breakpoints between human and mouse genomes coincide to replication-transition regions shows that mechanisms of replication-timing transition may have been involved in evolutional processes to build up the present chromosome structures. (Watanabe et al. Human Molecular Genetics, 11, 13-21, 2002).

Publications

1. Watanabe, Y., Fujiyama, A., Ichiba, Y., Hattori, M., Yada, T., Sakaki, Y. and Ikemura, T.: Chromosome-wide assessment of replication timing for human chromosomes 11q and 21q: disease-related genes in

timing-switch regions. *Human Molecular Genetics* **11**, 13-21, 2002.

2. Nishihashi, A., Haraguchi, T., Hiraoka, Y., Ikemura, T., Regnier, V., Dodson, H., Earnshaw, W. C. and Fukagawa, T.: CENP-I is essential for centromere function in vertebrate cells. *Dev. Cell* **2**, 463-476, 2002.

3. Abe, T., Kanaya, S., Kinouchi, M., Ichiba, Y., Kozuki, T. and Ikemura, T.: Informatics for unveiling hidden genome signatures. *Genome Research*, **13**, 693-702, 2003.

4. Ohno, M., Fukagawa, T., Lee, L.S. and Ikemura, T.: Triplex-forming DNAs in the human interphase nucleus visualized in situ by polypurine/ polypyrimidine DNA probes and antitriplex antibodies. *Chromosoma* **111**, 201-213, 2002.

5. Abe, T., Kanaya, S., Kinouchi, M., Ichiba, Y., Kozuki, T. and Ikemura, T.: A novel bioinformatic strategy for unveiling hidden genome signatures of eukaryotes: Self-organizing map of oligonucleotide frequency. *Genome Informatics Series* **13**, 12-20, 2002.

6. Fukushima, A., Ikemura, T., Kinouchi, M., Oshima, T., Mori, H. and Kanaya, S.: Detection of periodicity in eukaryotic genomes on the basis of power spectrum analysis. *Genome Informatics Series* **13**, 21-29, 2002.

7. Fukushima, A., Ikemura, T., Kinouchi, M., Oshima, T., Kudo, Y., Mori, H. and Kanaya, S.: Periodicities in prokaryotic and eukaryotic genomes identified by power spectrum analysis. *Gene* **300**, 203-211, 2002.

8. Spence, M. J., Critcher, R., Ebersole, A. T., Fukagawa, T. and Farr, J. C.: Co-localisation of centromere activity, CEN proteins and a Topoisomerase II cleavage site to a fixed region of limited extent within the human X-satellite locus DXZ1. *The EMBO Journal* **21**, 5269-5280, 2002.

9. Hayashi, T., Seki, M., Maeda, D., Wang, W., Kawabe, Y., Seki, T., Saitoh, H., Fukagawa, T., Yagi, H. and Enomoto, T.: Ubc9 is essential for viability of higher eukaryotic cells. *Experimental Cell Research* **280**, 212-221, 2002.

E. DEPARTMENT OF INTEGRATED GENETICS

E-a. DIVISION OF HUMAN GENETICS

(1) Comparative analysis of an imprinted domain in the mouse and its orthologue in the chicken

Hiroyuki SASAKI, Takaaki YOKOMINE, Hisao SHIROHZU, Chikako SUDA, Wahyu PURBOWASITO[1], Masaoki TSUZUKI[2], Atsushi TOYODA[3], Masahira HATTORI[3] and Yoshiyuki SAKAKI[3] ([1]Kyushu Univ.; [2]Hiroshima Univ.; [3]GSC, RIKEN)

Genomic imprinting, an epigenetic gene-marking phenomenon in mammals, causes parent-specific monoallelic expression of a subset of genes. Imprinted genes tend to form clusters in the genome (imprinted domains), which may be related to the mechanism of imprinting or to the reason for the evolution of imprinting. As a step to understand the structural and functional characteristics of imprinted domains, we have characterized a 1-Mb imprinted domain in mouse chromosome 7F4/F5 and its orthologous domain in chicken chromosome 5 (~0.5 Mb). We 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 ancestor of mammals and birds and that the imprinting mechanism, which can affect multiple genes in the cluster, came in later during mammalian evolution (in preparation).

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

Hiroyuki SASAKI, Ko ISHIHARA, Yuzuru KATO, Hiroyasu FURUUMI and Wolf REIK[1] ([1]Babraham Inst.)

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) located upstream of *H19* is the primary cause for the *Igf2/H19* imprinting. We investigated the changes in *Igf2/H19* imprinting in uniparental mouse fetuses (ref. 2). We also identified an evolutionarily conserved CTCF-dependent insulator element at the centromeric boundary of the *Igf2/H19* sub-domain (ref. 3). Lastly, we characterized a mouse mutant called *minute*, which has an inversion breakpoint at about 20-kb downstream of *H19* (ref. 10). The study confirmed our previous observation that mesodermal enhancers, which are involved in the *Igf2/H19* imprinting in muscles and cartilages, are located in the downstream region of *H19*. These results provide the basis for understanding the organization of the imprinted domains and subdomains and its significance in the control of monoallelic gene expression.

(3) Role of *de novo* DNA methyltransferases *Dnmt3a/Dnmt3b* in the establishment of genomic imprinting

Hiroyuki SASAKI, Masahiro KANEDA, Takashi SADO, Naomi TSUJIMOTO, Kenji KUMAKI, Shoji TAJIMA[1], Masaki OKANO[2] and En LI[3] ([1]Osaka Univ.; [2]CDB, RIKEN; [3]Harvard Med. Sch.)

DNA methylation serves as an important gene-marking mechanism for the discrimination of the parental alleles of imprinted genes. To understand how the primary imprints are established in the male and female germlines, we studied the expression and localization of the *de novo* DNA methyltransferases *Dnmt3a* and *Dnmt3b* in the male and female gonads. We have found that *Dnmt3b* and a specific form of *Dnmt3a* are present in the nuclei of prospermatogonia (also called gonocytes) and growing oocytes (in preparation), in which the primary imprints are established. We have also started to analyze germline-specific knockout of *Dnmt3a* and *Dnmt3b* to know which DNA methyltransferase is responsible for the imprint establishment.

(4) Molecular pathology of human disorders associated with DNA methylation and imprinting

Hiroyuki SASAKI, Hisao SHIROHZU, Hiroyasu FURUUMI, Takeo KUBOTA[1], Kenji KUROSAWA[2] and Shoji TAJIMA[3] ([1]Natl. Center of Neurology and Psychiatry; [2]Kanagawa Children's Hospital; [3]Osaka Univ.)

We studied three Japanese cases with ICF (immunodeficiency, centromeric instability, facial anomalies) syndrome, an autosomal recessive disorder with hypomethylation of satellite DNA, and identified three novel mutations within the coding region of *DNMT3B* (ref. 5). We also reported that paternal disomy for chromosome 14 causes a distinctive malformation syndrome with abdominal muscular defects, skeletal anomalies and characteristic facies (ref. 4). We speculated that *DLK1*, a paternally expressed gene located on this chromosome, is associated with the malformation of the mesodermal tissues.

(5) Investigation on the possibility of Z chromosome dosage compensation in chicken

Hiroyuki SASAKI, Takaaki YOKOMINE, Asato KUROIWA[1], Yoichi MATSUDA[1] and Masaoki TSUDUKI[2] ([1]Hokkaido Univ.; [2]Hiroshima Univ.)

We studied whether chicken Z chromosome is subject to inactivation for gene dosage compensation in ZZ males. It was found that Z-linked genes are biallelically expressed in ZZ males (ref. 7). The results suggest that sex-chromosome inactivation of a mammalian X-chromosome type does not occur in the chicken.

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

Hiroyuki SASAKI, Shin-ichi MIZUNO[1], Tadafumi IINO[1], Hidetoshi OZAWA[1], Teruhisa OTSUKA[1], Kosuke TASHIRO[1] and Takashi GOJOHBORI ([1]Kyushu Univ.)

We carried out a collaborative research project 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. Development of the universal DNA chip system for practical use is now underway.

(7) Role of *de novo* methylatransferases in X chromosome inactivation

Takashi SADO, Masaki OKANO[1], En LI[2] and Hiroyuki SASAKI ([1]CDB, RIKEN; [2]Harvard Med. Sch.)

X chromosome inactivation was studied in ES cells and embryos deficient for both *Dnmt3a* and *Dnmt3b*. In double mutant male ES cells, *Xist* became ectopically activated upon differentiation. It appeared, however, that these cells did not undergo X-inactivation despite the association of *Xist* RNA with the X chromosome. In contrast, such ectopic association of *Xist* RNA was observed only in a subset of nuclei in double mutant embryos in both sexes. Interestingly, most of nuclei in female embryos contained a single *Xist* domain despite the lack of differential methylation at the promoter region of *Xist*. Furthermore, one of the two X chromosomes in female embryos replicated late in S phase, suggesting that X-inactivation had properly taken place in the absence of functional *de novo* methyltransferases. These results suggest that differential methylation is not necessarily the underlying mechanism for monoallelic expression of *Xist* and these *de novo* methyltransferases may not be required for the initiation of X-inactivation and propagation of the inactive state along the chromosome.

(8) Effect of *Tsix*-disruption on *Xist* expression

Takashi SADO and Hiroyuki SASAKI

Xist and its antisense partner, *Tsix*, are the key players in X inactivation, both of which encode a non-coding RNA. Targeted disruption of *Tsix* causes ectopic expression of *Xist* in the extraembryonic tissues upon maternal transmission, which subsequently causes embryonic lethality due to inactivation of both X chromosomes in females and a single X chromosome in males. *Tsix*, therefore, plays a crucial role in maintaining *Xist* silencing in *cis* and in the regulation of imprinted X-inactivation in the extraembryonic tissues. We examined the effect of *Tsix*-disruption on *Xist* expression in the embryonic lineage using embryonic stem (ES) cells as a model system. Upon differentiation, *Xist* is ectopically activated in a subset of nuclei of male ES cells harboring the Tsix-deficient allele on the single X chromosome. Such ectopic

expression, however, ceased during prolonged culture. We are tempted to speculate that surveillance by the counting mechanism somehow shuts off ectopic *Xist* expression before that X chromosome undergoes inactivation. Counting does not seem to be impaired in *Tsix*-deficient ES cells (ref. 8).

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

Kenichiro HATA and En LI[1] ([1]Harvard Med. Sch.)

Dnmt3L (DNA cytosine-5-methyltransferase 3-Like) encodes a protein of 421 amino acid residues. Dnmt3L 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. *Dnmt3L* mRNA was detected in undifferentiated ES cells and testes by Northern blot analysis. We have generated *Dnmt3L*-deficient mice by gene targeting (ref. 6). 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. The embryos derived from a [*Dnmt3a*−/−, *Dnmt3b*+/−] ovary showed the same aberrant methylation of imprinted genes. 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 (ref. 6).

Publications

1. Davies, K., Bowden, L., Smith, P., Dean, W., Hill, D., Furuumi, H., Sasaki, H., Cattanach, B. and Reik, W.: Disruption of mesodermal enhancers for *Igf2* in the minute mutant. *Development* **129**, 1657-1666, 2002.

2. Sotomaru, Y., Katsuzawa, Y., Hatada, I., Obata, Y., Sasaki, H. and Kono, T.: Upregulated expression of the imprinted genes *H19* and *Igf2r* in mouse uniparental fetuses. *J. Biol. Chem.* **277**, 12474-12478, 2002.

3. Ishihara, K. and Sasaki, H.: An evolutionarily conserved putative insulator element near the 3' boundary of the imprinted *Igf2/H19* domain. *Hum. Mol. Genet.* **11**, 1627-1636, 2002.

4. Kurosawa, K., Sasaki, H., Sato, Y., Yamanaka, M., Shimizu, M., Ito, Y., Okuyama, T., Matsuo, M., Imaizumi, K., Kuroki, Y. and Nishimura, G.: Paternal UPD14 is responsible for a distinctive malformation complex. *Am. J. Med. Genet.* **110**, 268-272, 2002.

5. Shirohzu, H., Kubota, T., Kumazawa, A., Sado, T., Chijiwa, T., Inagaki, K., Suetake, I., Tajima, S., Wakui, K., Miki, Y., Hayashi, M., Fukushima, Y. and Sasaki, H.: Three novel *DNMT3B* mutations in Japanese patients with ICF syndrome. *Am. J. Med. Genet.* **112**, 31-37, 2002.

6. Hata, K., Okano, M., Lei, H. and Li, E.: Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* **129**, 1983-93, 2002.

7. Kuroiwa, A., Yokomine, T., Sasaki, H., Tsudzuki, M., Tanaka, K., Namikawa, T. and Matsuda, Y.: Biallelic expression of Z-linked genes in male chickens. *Cytogenet. Genome Res. in press*, 2003.

8. Sado, T., Li, E. and Sasaki, H.: Effect of *Tsix*-disruption on *Xist* expression in male ES cells. *Cytogenet. Genome Res. in press*, 2003.

E-b. DIVISION OF AGRICULTURAL GENETICS

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

Tetsuji KAKUTANI, Tetsu KINOSHITA, Asuka MIURA, Yuki KINOSHITA and Masaomi KATO

Substantial part of large genome of flowering plants and vertebrates are repetitive sequences, such as transposable elements and their derivatives. Genomic regions rich in repeated sequences tend to be inactive in transcription and recombination and have condensed chromatin (heterochromatin). In those regions, cytosine residues of genomic DNA are often methylated at high frequency. Although function of

repeated sequences are largely unknown, uncontrolled activation of these sequences are presumed to be deleterious to genome stability. On the other hand, some of repetitive sequences such as centromeric repeats and telomeres are regarded as important for the chromosome function. We are studying control and biological function of repetitive sequences using DNA methylation mutants of Arabidopsis.

In Arabidopsis, genome sequencing has been completed for the first case in plants. In addition, many trans mutations affecting epigenetic states have been isolated in this plant. Arabidopsis *ddm1* (*decrease in DNA methylation*) mutation results in decrease in methylation and transcriptional de-repression in genomic repeat sequences. The *DDM1* gene encodes a protein similar to the chromatin-remodeling factor SWI2/SNF2. The most striking feature of *ddm1* mutation is that it induces a variety of developmental abnormalities by causing heritable change in other loci. The molecular basis has been clarified in two of the loci directly causing the developmental abnormalities (see below).

(2) Charcterization of transposons activated by DNA hypomethylation mutants

Asuka MIURA, Masaomi KATO, Yuki KINOSHITA and Tetsuji KAKUTANI

One of the *ddm1*-induced developmental abnormalities, *clam*, is characterized by lack of elongation in leaves, roots and shoots. This phenotype is heritable, but somatic sectors with normal phenotype were occasionally observed. The size and frequency of the sector differ from plant to plant. The phenotype was stabilized in some of the progeny families; no reversion sector was observed in such family. Through genetic mapping , we identified the gene responsible for the *clam* phenotype; it is the *DWF4* gene, which involved in synthesis of brassinolide, a plant growth regulator necessary for cell elongation. The unstable *clam* phenotype was induced by insertion of a novel endogenous Arabidopsis transposon, which we named *CACTA1*. This transposon transposes and increases in the copy number specifically in *ddm1* mutant background. These results suggest that gene silencing associated with DNA methylation is important for suppression of transposons.

As *DDM1* gene does not encode DNA methyltransferase but chromatin remodeling factor, it has not been proved that loss of DNA methylation is sufficient for the mobilization of *CACTA* elements. In order to clarify this point, we examined mobility of *CACTA* elements in mutants of DNA methyltransferase genes. 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). We examined the mobility of endogenous *CACTA* transposons in *met1*, *cmt3*, and *cmt3-met1* mutants. 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, despite significant induction of *CACTA* transcript accumulation. These results lead us to conclude that CG and non-CG methylation systems redundantly function for immobilization of transposons (Kato et al., 2003). *CMT3* gene and non-CG methylation in plants may have evolved as an additional epigenetic tag dedicated to transposon control.

(3) Inheritance of epigenetic developmental abnormality

Tetsuji KAKUTANI, Asuka MIURA and Tetsu KINOSHITA

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. 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 important for plant development. 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 absence of the potentiating *ddm1* mutation.

(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 has been completely unknown. To understand why the *FWA* gene is epigenetically controlled, we further examined expression of this gene 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. We are currently dissecting molecular mechanisms of this control in the endosperm using a series of Arabidopsis mutants affecting epigenetic states in trans.

(5) DNA hypomethylation mutation in rice

Tetsuji KAKUTANI and Asuka MIURA

Among plant species, Arabidopsis genome has extreme feature that it contains only small proportion of repeated sequences. We therefore extended our research by examining effect of de-repression of genomic repeated sequences in rice. Rice has more repeat sequences than Arabidopsis. We found a rice EST similar to Arabidopsis *DDM1* gene and generated a transgenic rice lines expressing that EST in antisense orientation. These transgenic lines shows reduced genomic DNA in centromeric repeats, repeat encoding rRNA and retroelement-like sequence Tos3.

Supported by MAFF rice genome project, we examined change in transcription pattern with microarray. Despite the global loss of genomic DNA methylation, expression level did not change in the mutant for most of the EST clones on the microarray. We are currently examining transcript level of endogenous rice transposons, most of which are not represented in the ESTs.

Publications

1. Kakutani, T.: Epi-alleles in plants: Inheritance of epigenetic information over generations. *Plant Cell Physiol.* **43**, 1106-1111, 2002.

2. Soppe, W., Jasencakova, Z., Hauben, A., Kakutani, T., Meister, A., Huang, M., Jacobsen, S., Schubert, I. and Fransz, P.: DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in Arabidopsis. *EMBO Journal.* **21**, 6549-6559, 2002.

3. Kato, M., Miura, A., Bender, J., Jacobsen, S.E. and Kakutani, T.: Role of CG and non-CG methylation in immobilization of transposons in Arabidopsis. *Current Biology* **13**, 421-426, 2003.

(1) Analysis of nucleosome assembly during DNA replication

Tomohiro KUBO[1], Akira SHIMIZU[1], Bruce STILLMAN[2] and Keiichi SHIBAHARA[3] ([1]Graduate School of Biostudies, Kyoto University, [2]Cold Spring Harbor Laboratory, USA, [3]Japan Science and Technology Corporation)

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 found to be involved in some process of the nucleosome assembly.

We recently developed an *in vitro* reconstitution system with human cell extract to analyze a mechanism of the DNA replication-dependent nucleosome assembly. This nucleosome assembly reaction was formerly analyzed in an ongoing replication reaction of SV40 *in vitro* system. This replication-coupled system has proven particularly useful in identifying CAF-1. However, its use to identify other activities involved in the replication-dependent nucleosome assembly has been limited, because biochemical fractionations of the human cell cytosolic extract (the S100 extract) used in the SV40 DNA replication (which contains all the necessary DNA replication factors, histones, and other chromatin assembly factors) often resulted in a severe loss of DNA replication activity. In order to overcome these limitations, we successfully used replicated DNA purified by a gel filtration spin column as a substrate for the nucleosome assembly reaction. The newly replicated and separated DNA was competent for the replication-dependent nucleosome assembly even after separation from free protein fractions in the S100 extract because the DNA remains

marked, at least by PCNA (Shibahara and Stillman, *Cell 96*, 575-585, 1999).

With this "two step system", we showed that N-terminal of histones H3 and H4 were not necessarily required for their association with CAF-1 and nucleosome assembly by CAF-1 (Shibahara *et al., Proc. Natl. Acad. Sci. USA* 97, 7766-7771, 2000). In addition, we showed that unidentified activities in the S100 extract were required for the nucleosome assembly in addition to CAF-1 and histones (unpublished data). We are now on the way to identify additional activities required for the DNA replication-dependent nucleosome assembly.

(2) Physiological implications of CAF-1 and ASF1 in *Arabidopsis*

Hidetaka KAYA[1,2], Takashi ARAKI[3] and Keiichi SHIBAHARA[4] ([1]Faculty of Medicine, Kyoto University, [2]Japan Society for the Promotion of Science, [3]Kyoto University, [4]Japan Science and Technology Corporation)

The genetic approach of *Arabidopsis* was a powerful tool to see physiological implications of CAF-1. We had analyzed loss-of-function mutants of *caf-1* (*fasciata*) in *Arabidopsis*. Those *caf-1* mutants displayed severely disturbed cellular and functional organization of both shoot apical meristem (SAM) and root apical meristem (RAM).

They also showed a varied pattern of distorted expression of both *WUS* and *SCR*, which play key roles in the organization of SAM and RAM, respectively. Asymmetrical cell division of a stem cell in RAM (Co/En initial) gives rise to two different daughter cells (a Co/En initial and a Co/En daughter cell), which then divides periclinally to generate a cortex and an endodermal cell. *SCR* is expressed only in the endodermal lineage. However, in some roots of the *caf-1* mutants, ectopic expression of *SCR::GFP* was observed in single cells or small groups of cells adjacent to the cells normally expressing *SCR::GFP*. On the other hand, the absence of expression was observed in single cells and small groups of adjacent cells including all types of cells that normally express *SCR::GFP*. The disturbances of SCR-GFP expression were varied from roots to roots and from plants to plants.

We proposed several models to explain those observations, and the most plausible was that CAF-1 ensures stable propagation of epigenetic states of chromatin through facilitating nucleosome assembly during DNA replication. Rapid re-formation of nucleosomes onto newly replicated DNA by CAF-1 would prevent transcriptional regulators from being targeted to the DNA in a non-regulated manner, thereby preventing random changes in gene expression patterns in daughter cells. We also isolated *asf1* mutants in *Arabidopsis* with similar phenotypes. We are now analyzing those mutants and/or crossed mutants from several angles to discuss the physiological implications of the DNA replication-dependent nuclesome assembly more precisely.

References

1. Kaya, H., Shibahara, K-i., Tasaka, K-I., Iwabuchi, M., Stillman, B. and Araki, T.: FASCIATA genes for chromatin assembly factor-1 in Arabidopsis maintain the cellular organization of apical meristems. *Cell.* **104**, 131-142, 2001.

E-c. DIVISION OF BRAIN FUNCTION

(1) Mosaic development of the olfactory cortex with Pax6-dependent and -independent components

Tatsumi HIRATA, Tadashi NOMURA[1], Yoshiko TAKAGI, Yasufumi SATO, Naomi TOMIOKA[2], Hajime FUJISAWA[2] and Noriko OSUMI[1] ([1]Department of Developmental Neurobiology, Tohoku University Graduate School of Medicine, [2]Division of Biological Science, Nagoya University Graduate School of Science)

One of the early events in the development of the central nervous system is regionalization of the neural tube into organized domains. The telencephalic primordium is divided into two major domains: the pallium and subpallium. In each domain, various genes are expressed in a specific manner and assign different developmental potentials to neural progenitors.

In the present study, we examined the development of the olfactory cortex, using the newly developed monoclonal antibody (mAb) 9-4c. The olfactory cortex is the target area of olfactory bulb axons and suggested to be derived from neural progenitors of various ventricular domains during development. The novel regional marker, mAb 9-4c, labeled neural progenitors at the pallio-subpallial boundary (PSB) and

their putative descendants in the deep layers of the olfactory cortex. A search for the antigen recognized by mAb 9-4c identified reticulon 1(Rtn 1), an endoplasmic reticulum (ER) protein. In the mutant embryos of PAX6, a pallium specific transcription factor, expression of this antigen was completely absent at the PSB, and the number of immunopositive cells in the olfactory cortex was markedly reduced. Nevertheless, the guidepost neurons of olfactory bulb axons, lot cells, developed relatively normally in the superficial layer of the olfactory cortex in the mutant embryos. These guidepost neurons are shown to originate in the pallium and eventually guide the initial projection of olfactory bulb axons. Subsequent projection of olfactory bulb axons in PAX6 mutant embryos also supported the mosaic nature of the olfactory cortex development; the axons initially projected via the normal pathway and formed the LOT bundle, but later failed to develop collateral projections over the wide olfactory cortical areas.

(2) Expression of Nogo protein by growing axons in the developing nervous system

Hirokazu TOZAKI, Takahiko KAWASAKI, Yoshiko TAKAGI and Tatsumi HIRATA

Development of the nervous system is characterized by the striking capacity for axonal outgrowth. The axons, however, lose the capacity with maturation of the central nervous system. Differences in the growing capabilities of embryonic and adult axons can be partly attributed to the growth-inhibitory environment of the adult nervous system. Previous studies *in vivo* and *in vitro* suggested the role of the central myelin in preventing axonal growth in the adult nervous system.

In the present study, we isolated a monoclonal antibodies (mAb) designated as NG1 that strongly bound to growing axons in the developing nervous system. The growth cones in cultured neurons were particularly strongly labeled with this mAb. Although these immunostaining patterns suggested the association of the antigen recognized by mAb NG1 with growing axons, a search for this antigen led to an unexpected finding; mAb NG1 recognizes Nogo, an axonal repulsive factor in the myelin. Previous studies show that Nogo is produced by oligodendrocytes in the adult nervous system and acts to prevent the growth of central axons. Our present finding, however, suggests another function of Nogo protein in the developing nervous system.

Publications

1. Hirata, T., Nomura, T., Takagi, Y., Sato, Y., Tomioka, N., Fujisawa, H. and Osumi, N.: Mosaic development of the olfactory cortex with Pax6-dependent and -independent components. *Dev. Brain Res.* **136**, 17-26, 2002.

2. Tozaki, H., Kawasaki, T., Takagi, Y. and Hirata, T.: Expression of Nogo protein by growing axons in the developing nervous system. *Mol. Brain Res.* **104**, 111-119, 2002.

E-d. DIVISION OF APPLIED GENETICS

(1) The *slender* rice mutant, with constitutively activated gibberellin signal transduction, has enhanced capacity for abscisic acid level

Akira IKEDA[1], Yutaka SONODA[1], Paolo VERNIERI[2], Pierdomenico PERATA[3], Hirohiko HIROCHIKA[4] and Junji YAMAGUCHI[1] ([1]Hokkaido University, [2]National Institute of Agrobiological Resources, [3]University of Pisa, University of Modena and Reggio Emilia, [4]National Institute of Agrobiological Sciences).

The *slender* rice (*slr1-1*) mutant, carrying a lethal and recessive single mutation, has a constitutive gibberellin (GA)-response phenotype and behaves as if it were saturated with GAs [Ikeda et al. (2001) *Plant Cell* 13, 999]. The *SLR1* gene, with sequence homology to members of the plant-specific GRAS gene family, is a mediator of the GA signal transduction process. In the *slender* rice, GA-inducible alpha-amylase was produced from the aleurone layer without applying GA. GA-independent alpha-amylase production in the mutant was inhibited by applying abscisic acid (ABA). Shoot elongation in the mutant was also suppressed by ABA, indicating that the *slender* rice responds normally to ABA. Interestingly, shoot ABA content was 10-fold higher in the mutant than in the wild type, while there was no difference in root ABA content. Expression of the *Rab16A* gene, which is known to be ABA inducible, was about 10-fold higher in shoots of the mutant than in those of the wild type. These results

indicate that constitutive activation of the GA signal transduction pathway by the *slr1-1* mutation promotes the endogenous ABA level.

(2) Isolation and characterization of a rice dwarf mutant with a defect in brassinosteroid biosynthesis

Masaki MORI[1], Takahito NOMURA[2], Hisako OOKA[3], Masumi ISHIZAKA[4], Takao YOKOTA[2], Kazuhiko SUGIMOTO[1], Ken OKABE[1], Hideyuki KAJIWARA[1], Kouji SATOH[1], Koji YAMAMOTO[3], Hirohiko HIROCHIKA[1] and Shoshi KIKUCHI[1] ([1]National Institute of Agrobiological Sciences, [2]Teikyo University, [3]Utsunomiya, Nagaoka University of Technology, [4]National Institute for Agro-Environmental Sciences)

We have isolated a new recessive dwarf mutant of rice (Oryza sativa L. cv Nipponbare). Under normal growth conditions, the mutant has very short leaf sheaths; has short, curled, and frizzled leaf blades; has few tillers; and is sterile. Longitudinal sections of the leaf sheaths revealed that the cell length along the longitudinal axis is reduced, which explains the short leaf sheaths. Transverse sections of the leaf blades revealed enlargement of the motor cells along the dorsal-ventral axis, which explains the curled and frizzled leaf blades. In addition, the number of crown roots was smaller and the growth of branch roots was weaker than those in the wild-type plant. Because exogenously supplied brassinolide considerably restored the normal phenotypes, we designated the mutant brassinosteroid-dependent 1 (*brd1*). Further, under darkness, brd1 showed constitutive photomorphogenesis. Quantitative analyses of endogenous sterols and brassinosteroids (BRs) indicated that BR-6-oxidase, a BR biosynthesis enzyme, would be defective. In fact, a 0.2-kb deletion was detected in the genomic region of *OsBR6ox* (a rice *BR-6-oxidase* gene) in the brd1 mutant. These results indicate that BRs are involved in many morphological and physiological processes in rice, including the elongation and unrolling of leaves, development of tillers, skotomorphogenesis, root differentiation, and reproductive growth, and that the defect of BR-6-oxidase caused the brd1 phenotype.

(3) Organ-specific alternative transcripts of KNOX family class 2 homeobox genes of rice

Yukihiro ITO[1], Hirohiko HIROCHIKA[2] and Nori KURATA[1] ([1]National Institute of Genetics, [2]National Institute of Agrobiological Sciences)

We identified three genes (HOS58, HOS59 and HOS66) of rice (Oryza sativa), which encode predicted proteins with the KNOX family class 2 homeodomain. These proteins contain three conserved domains, the KNOX domain, ELK domain and homeodomain, from an N-terminus to a C-terminus. In addition to similarity of predicted amino acid sequences, these genes showed a similar exon/intron structure. cDNA cloning and reverse transcription-polymerase chain reaction analyses of these genes indicated tissue-specific expression of alternative transcripts. The expression of the longer mRNAs of HOS58 (HOS58L) and HOS59L was detected in all organs examined such as roots, leaf blades, leaf sheaths, flowers and calli, whereas the shorter mRNAs of HOS58 (HOS58S) and HOS59S were expressed in leaf blades, leaf sheaths and flowers. The expression of HOS66L was detected in roots, leaf blades, leaf sheaths and flowers, whereas the expression of HOS66S was detected in roots and flowers. The alternative transcripts of HOS66 arose by use of alternative transcription start sites. The longer transcripts contained an exon 1which encodes an alanine/glycine-rich region, whereas the shorter ones lacked it. These results suggest that theexpression of the alternative transcripts is organ-specific, and their products have different degrees of abilities for activation or repression of transcription.

(4) An active DNA transposon family in rice

Ning JIANG[1], Zhirong BAO[2], Xiaoyu ZHANG[1], Hirohiko HIROCHIKA[3], Sean R. EDDY[2], Susan R. MCCOUCH[4] and Susan R. WESSLER[1] ([1]University of Georgia, [2]Washington University School of Medicine, [3]National Institute of Agrobiological Resources, [4]Cornell University)

The publication of draft sequences for the two subspecies of *Oryza sativa* (rice), *japonica* (cv. Nipponbare) and *indica* (cv. 93-11), provides a unique opportunity to study the dynamics of transposable elements in this important crop plant. Here we report the use of these sequences in a computational approach

to identify the first active DNA transposons from rice and the first active miniature inverted-repeat transposable element (MITE) from any organism. A sequence classified as a *Tourist*-like MITE of 430 base pairs, called *miniature Ping* (*mPing*), was present in about 70 copies in Nipponbare and in about 14 copies in 93-11. These *mPing* elements, which are all nearly identical, transpose actively in an *indica* cell-culture line. Database searches identified a family of related transposase-encoding elements (called *Pong*), which also transpose actively in the same cells. Virtually all new insertions of *mPing* and *Pong* elements were into low-copy regions of the rice genome. Since the domestication of rice *mPing* MITEs have been amplified preferentially in cultivars adapted to environmental extremes-a situation that is reminiscent of the genomic shock theory for transposon activation.

Publications

1. Ikeda, A., Sonoda, Y., Vernieri, P., Perata, P., Hirochika, H. and Yamaguchi, J.: The *slender* rice mutant, with constitutively activated gibberellin signal transduction, has enhanced capacity for abscisic acid level. *Plant Cell Physiol.* **43**, 974-979. 2002.

2. Mori, M., Nomura, T., Ooka, H., Ishizaka, M., Yokota, T., Sugimoto, K., Okabe, K., Kajiwara, H., Satoh, K., Yamamoto, K., Hirochika, H. and Kikuchi, S.: Isolation and characterization of a rice dwarf mutant with a defect in brassinosteroid biosynthesis. *Plant Physiol.* **130**, 1152-1161, 2002.

3. Ito, Y., Hirochika, H. and Kurata, N.: Organ-specific alternative transcripts of KNOX family class 2 homeobox genes of rice. Gene **17**, 41-47, 2002.

4. Jiang, N., Bao, Z., Zhang, X., Hirochika, H., Eddy, SR., McCouch, SR. and Wessler, SR.: An active DNA transposon family in rice. Nature **421**, 163-167, 2003.

(1) Non-random X Chromosome Inactivation in Mouse Embryos Carrying Searle's T(X;16)16H Translocation Visualized using X-linked *lacZ* and *GFP* Transgenes

Nobuo TAKAGI, Michihiko SUGIMOTO, Shinpei YAMAGUCHI, Mitsuteru ITO, Seong-Seng TAN[1] and Masaru OKABE[2] ([1]Howard Florey Inst., Univ. Melbourne, and [2]Genome Inform. Res. CTR, Osaka Univ.)

Only the morphologically normal X chromosome is inactivated in female mice heterozygous for Searle's X-autosome translocation, T(X;16)16H. Here we performed a visual study of the primary and secondary events that culminate in the completely non-random inactivation in female embryos having this translocation. The data we have obtained so far indicate that the initial choice of the future inactive X chromosome is biased: the degree of skewing is somewhere between 70:30% and 90:10% in favor of the morphologically normal X chromosome. The majority of genetically unbalanced cells that inactivate a translocated X chromosome are quickly eliminated from the embryo proper by E8.5, although the survival of such cells is sporadically observed thereafter. The initial non-random choice demonstrated in this study supports the contention that the T(X;16)16H translocation disrupts one of the loci involved in the randomness of the choice of the future inactive X chromosome. Although the *HMG-lacZ* transgene in the H253 stock mice is an excellent marker of X-inactivation, the present study suggests that it is infrequently de-repressed on the inactive X chromosome.

(2) Mapping of the breakpoint in the mouse T(X;16)16H translocation that causes non-random X chromosome inactivation.

Mitsuteru ITO and Nobuo TAKAGI

On the assumption that there is a novel element involved in the random choice of the future inactive X in the area near the T16H translocation, we attempted to map and clone the translocation breakpoint. Previous studies showed that the T16H breakpoint is located in a 4 cM region spanning from *G6pdx* to *Pola* about 8 cM proximal to the X-chromosome inactivation center. Using available maps such as the radiation hybrid map, genetic map and physical map of YACs and BACs, we made a detailed physical map near the translocation breakpoint. FISH assay with BAC clones as probes showed that the breakpoint is on the region that is roughly estimated about 900 kb. This region is physically continuous with seven BAC clones and one assembled contig. However, the analysis was hindered by the discouraging fact that six different BAC clones gave rise to two FISH signals one on 16^{X} and the other

on X^{16} chromosome. Further sequence analysis showed a possibility that the X chromosome was broken somewhere in the BACRP23-279G2 that is less than 250kb in length.

(3) New XX ES cell lines carrying *GFP* and *lacZ* transgenes on the maternal or the paternal X chromosome

Shirin FARIVAR, Seong-Seng TAN[1], Masaru OKABE[2] and Nobuo TAKAGI ([1]Howard Florey Inst., Univ. Melbourne, and [2]Genome Inform. Res. CTR, Osaka Univ.)

Three new ES cell lines carrying the *GFP* and *HMG-lacZ* transgenes on the same X chromosome were established. The first ES cell line (MC58) was derived from an XX blastocyst obtained from a GL female doubly homozygous for *GFP* and *HMG-lacZ* transgenes mated with a normal male (129/Sv). Two transgenes were, therefore, carried by the maternally derived X-chromosome. The second and third lines (GL1 and GL2) were established from blastocysts obtained from 129/Sv females mated with GL male mice. In these cell lines, the two transgenes were carried by the paternally inherited X-chromosome. Differentiation was induced by culturing MC58 cells in the absence of LIF in the bacteriological dishes to which ES cells do not adhere. On day 3 of differentiation, half the embryoid bodies were transferred to tissue culture dishes to the bottom of which embryoid bodies adhere and spread to enhance differentiation further. GFP-negative cells appeared first in the outer endoderm layer 4 days after the initiation of differentiation suggesting the occurrence of X-inactivation. A single late replicating X chromosome was identified in a low frequency of metaphase cells, but β-galactosidase negative cell was not observed at this stage. Loss of β-gal activity occurred about two days after the first occurrence of differential GFP fluorescence. Visual examination of X-inactivation thus becomes possible in differentiating ES cells may reveal here-to-fore hidden aspects of random and nonrandom X-inactivation.

(4) A developmental biological study on the *t*-complex recessive lethal mutation *tclw5* in the mouse

Michihiko SUGIMOTO, Kuniya ABE[1], Seong-Seng TAN[2], Masaru OKABE[3] and Nobuo TAKAGI ([1]Riken BRC, [2]Howard Florey Inst., Univ. Melbourne, and [3]Genome Inform. Res. CTR, Osaka Univ.)

The mouse *t*-complex is a naturally occurring variant of the proximal region of chromosome 17 with four independent inversions. The *t*-haplotypes usually contain a recessive lethal mutation. In spite of remarkable progress in genetics from the first discovery of *t*-haplotypes in 1932, the mechanism of the lethality of any *t*-haplotypes is not yet understood. *tclw5* is a *t*-complex recessive lethal mutation in the *tw5*-haplotype. Since *tw5/tw5* embryos die soon after implantation, *tclw5* gene is thought to play an important role in embryogenesis. To understand the function of *tclw5* gene product,we tried to determine the primary site of *tclw5* gene action. Histological examination demonstrated that homozygotes of *tw5* died at the gastrulation stage due to extensive death of the embryonic ectoderm cells while the extraembryonic ectoderm and the visceral endoderm were less affected in these embryos. Histologically, *tw5/tw5* embryos were first distinguished from normally growing littermates at E5.5. At this stage, the visceral endoderm cells of *tw5/tw5* embryos were characterized by vacuoles smaller in number and size than normal littermates. By aggregation with diploid wild-type embryos, *tw5/tw5* cells were capable of contributing to all three germ layers. Thus, it is likely that *tclw5* is not cell autonomous lethal. Furthermore, the hypothesis that tissues other than the embryonic ectoderm are the primary sites of *tclw5* action was supported by the rescue of *tw5/tw5* embryos by aggregation with tetraploid embryos. Hence, it is probable that the product of *tclw5* is essential for the function of extraembryonic tissues. However, embryos thus rescued were arrested at E9.0 with the severely underdeveloped posterior region, indicating that *tclw5* is also necessary for development at midgestation.

Publications

1. Mizuno, H., Okamoto, I. and Takagi, N.: Developmental abnormalities in mouse embryos tetrasomic for chromosome 11: Apparent similarity to embryos functionally disomic for the X chromosome. *Genes and Genetic Systems*, **77**, 269-276, 2002.

2. Goto, Y., Matsui, J. and Takagi, N.: Developmental potential of mouse tetraploid cells in diploid tetraploid chimeric embryos. *International Journal of*

Developmental Biology, **46**, 741-745, 2002.

3. Takagi, N., Sugimoto, M., Yamaguchi, S., Ito, M., Tan, S.S. and Okabe, M.: Non-random X chromosome inactivation in mouse embryos carrying Searle's T(X,16)16H translocation visualized using X-linked *lacZ* and *GFP* transgenes. *Cytogenetic and Genome Research*, in press.

F. GENETIC STRAINS RESEARCH CENTER

F-a. MAMMALIAN GENETICS LABORATORY

(1) Developmental Genetics of Mouse Preaxial Polydacylous Mutations

Yukari YADA[1], Tomoko SAGAI, Sadao ISHIWA[1] and Toshihiko SHIROISHI ([1]Ochanomizu University)

The anteroposterior (A-P) axis patterning of developing limbs is controlled by formation of the zone of polarizing activity (ZPA), located at the posterior margin of limb buds. Sonic hedgehog (*Shh*) is expressed at the posterior margin of limb bud, and is thought to mediate the ZPA activity. Recently, it was reported that the A-P axis formation of limbs is not affected by *Shh* null mutation in the early stage of limb development. Thus, the initial A-P axis of limb is likely determined by *Shh*-independent signaling pathway.

To uncover the genetic network that controls the early phase of the A-P axis formation of limbs, we carried out molecular characterization of the preaxial polydacylous mutations, luxate (*lx*), X-linked polydactyly (*Xpl*) and Recombination induced mutation -4 (*Rim4*), which were affected in the A-P axis formation. Since causative genes for these mutants have not been identified, we examined the expression patterns of genes known to be involved in the A-P axis formation of limbs with the mutant embryos. In *Xpl* embryos, though the expression of *Shh* was normal in the posterior of limb buds at the initial stage, the ectopic expression of *Shh* was detected in the anterior side of the hindlimb buds. No ectopic expression of the other genes was detected before the ectopic *Shh* expression. Thus, *Xpl* gene likely acts in the downstream of the *Shh* signaling, and mediates the *Shh*-dependent signaling during limb morphogenesis. On the other hand, in *lx* and *Rim4*, the *Shh* expression domain in the posterior side of the limb buds was shifted anteriorly compared with that in wild-type embryos. Thus, in these mutants, the A-P axis formation of limbs is disrupted from the initial stage of limb development. Therefore, we examined the expression of *dHAND*, *Gli3* and *Alx4* in the *lx* and *Rim4* embryos. The result revealed that their expression was normal in *Rim4* embryos, but shifted anteriorly in the *lx* limb buds, indicating that *lx* gene acts in upstream of *dHAND*, *Gli3* and *Alx4* genes, while *Rim4* gene acts in downstream of them. Thus, it is most likely that *lx* controls the initial A-P axis formation through the *Shh* independent signal, and that based on the spatiotemporal expression patterns, the *Rim4* gene is involved in positioning of *Shh* induction in posterior margin of limb buds (For detail, see ref.1).

(2) Phylogenic conservation of a *cis*-acting regulator that controls polarized expression of Sonic hedgehog (*Shh*) in limb buds

Tomoko SAGAI, Hiroshi MASUYA[1], Masaru TAMURA, Kunihiko SHIMIZU[2], Yukari YADA[3], Shigeharu WAKANA[1], Yoichi GONDO[1], Tetsuo NODA[1] and Toshihiko SHIROISHI ([1]RIKEN GSC, [2]Nihon University, [3]Ochanomizu University)

Polarized expression of the Sonic hedgehog (*Shh*) gene in the posterior mesenchyme is common from teleost fin to tetrapod limb buds, and essential for limb pattern formation. But the molecular mechanism of its regulation is poorly understood.

We have performed the positional cloning of two preaxial polydactyl mutations, *Hx* and *M100081*, localized in the vicinity of *Shh* gene on chromosome 5. *M100081* is a novel mutant generated from the ENU mutagenesis in RIKEN GSC. The both mutants are characterized with the mirror-image digit duplication coupled with an ectopic expression of *Shh* in the anterior limb margin. In the 84kb critical region of *Hx*, a candidate gene, *Lmbr1*, was identified. However, any alteration was not found in the *Lmbr1* coding sequence of *Hx*, and the expression level of the *Lmbr1* was similar between *Hx* and the wild type embryos. We identified a sequence conserved between human and mouse in intron 5 of *Lmbr1*. It resides approximately 1Mbp from the *Shh* coding region in the mouse genome. Moreover, the sequence is highly conserved throughout the tetrapod lineage. The sequences of 500bp fragment from mammalian and reptile species and the sequence of a 300bp fragment from amphibian

species were highly conserved with the average matching rate exceeding 80% among species.

Sequence analysis of the two mutations, *Hx* and *M100081*, detected single base-substitution located in the conserved sequence. In order to clarify whether the sequence is a *cis*-acting regulator, we carried out a *cis*-trans test for *Hx* mutation using *Shh* knockout (KO) allele. As a result, in spite of the dominant phenotype of the *Hx* mutation, mice heterozygous for the recombinant chromosome harboring the *Hx* mutation and *Shh* KO allele on the same chromosome did not show the polydactylous phenotype. This result revealed that *Hx* mutation is attributed to an alteration of the *cis*-acting regulator of *Shh*.

In some vertevrate species, limb loss is often accompanied with trunk elongation and loss of regional differentiation in the axial skeleton. It has been reported that in pythons, a group evolved from the tetrapod lizards, and that forelimbs are absent, but hindlimb buds are formed during embryogenesis. It is also known that the *Shh* expression is not detected in python hindlimb buds, and there is no polarizing activity. If a function of the conserved sequence in intron 5 of *Lmbr1* is crucial for the polarized expression of *Shh* in the posterior mesenchyme of limb buds, this sequence may be lost in limbless species. We therefore examined whether several limbless snake species and a limbless newt retain the conserved sequence. Southern blot analysis using probes amplified for lizard and newt did not yielded any signal for the limbless species. Thus, it appears that loss of the conserved sequence is coupled with the loss of *Shh* signaling pathway, suggesting the conserved sequence has played a key role in tetrapod evolution. Considering the ectopic expression in *Hx* mutant and the highly conserved sequence in intron 5 of *Lmbr1*, the wild-type *cis*-element is likely to be involved in polarized *Shh* expression by repressing the *Shh* expression in the anterior mesenchyme of developing limb buds.

(3) Identification and characterization of the responsible gene of mouse skin mutations, Recombination-induced mutation 3 (*Rim3*) and Rex denuded (*Re^{den}*)

Masaru TAMURA, Shigekazu TANAKA, Tomoaki FUJII, Aya AOKI[1] and Toshihiko SHIROISHI. ([1]Department of Neurosurgery, Juntendo University, Izunagaoka Hospital)

Epidermis is composed of multiple cell layers, the outermost being the body surface. Only innermost, basal layer, contains living and multiplying cells. When a basal cell ceases to divide and begins its journey to the skin surface, it embarks on a program of terminal differentiation. However, the molecular mechanism by which epidermal stem cells differentiate is largely unknown.

Currently, we are interested in mechanisms of epidermal cell growth, differentiation and hair development. To study these subjects, we are using several mutant mice, such as Recombination-induced mutation 3 (*Rim3*) and Rex denuded (*Re^{den}*). *Rim3* and *Re^{den}* arose spontaneously in the congenic strain B10.BR (R228) and B10.129, respectively. *Rim3* exhibits epidermal hyperplasia, hyperkeratosis and abnormal hair development resembling those of *Re^{den}*. Linkage analysis and phenotype characterization indicated that *Rim3* and *Re^{den}* might be allelic. *Rim3* was mapped to the distal portion of mouse chromosome 11 between *Grb7* and *D11Mit14* (Sato et. al., 1998).

Recently, we isolated several genes in the critical region of *Rim3*. Sequence analysis of *Rim3* and *Re^{den}* mutations revealed that one of these genes is the causative gene for the two mutations. To examine expression pattern of this gene, we carried out RT-PCR and *in situ* hybridization analysis for several different tissues of adult mouse. We found that this gene is mainly expressed in hair follicules. Histological analysis of *Rim3* mutant mice at 3 months of age showed that the interfollicnler epidermis and permanent zone of the hair follicles are highly hyperplastic and hyperkeratinized. At this stage, the number of BrdU-possitive cell nuclei of *Rim3* was six- to sevenfold higher in the basal and the permanent zone of the hair follicle cells than those of control wild-type mice, indicating a higher proliferation rate of those cells in the *Rim3* skin. Immunohistochemical analysis with several antibodies such as anti-Keratin 14, -keratin 10 and -keratin 6 demonstrated that the hair follicle structure was converted into epidermis in *Rim3*, and that differentiation of epidermal keratinocytes was impaired in the mutant mice. These data suggested that function of causative gene of *Rim3* is to regulate cell growth and terminal differentiation of epidermis.

(4) Male-specific reproductive failure caused by X-chromosomal substitution between two mouse subspecies

Ayako OKA, Nobuo TAKAGI[1], Kiyotaka TOSHIMORI[2], Toshiyuki TAKANO[3], Akihiko MITA, Yoichi MIZUSHINA, Noriko SAKURAI, Hiromi YAMAMOTO and Toshihiko SHIROISHI ([1]Res. Cent. For Mol. Genet. and Grad. Schl. of Env. Earth Sci. Hokkaido Univ., Department of Anatomy and Developmental Biology(G1) Grad. Schl. of Medicine, Chiba Univ. [3]Division of Population Genetics, NIG)

Consomic strain is defined as an inbred strain, in which a certain chromosome of a strain is replaced as a whole by the same chromosome of another strain, leaving the rest of the former strain's chromosomes intact. We have set out construction of a consomic mouse strain, replacing X-chromosome of a standard laboratory strain, C57BL/6J (*Mus musculus domesticus*), by its homologue of a strain MSM (*Mus musculus molossinus*). During generating the X-chromosomal consomic strain, B6.MSM-ChrX, we found sterility only in the males. When eggs/embryos were collected from oviducts one day after the copulation with the B6.MSM-ChrX males, none of them developed to two-cell-stage embryos. This suggests that the sperms of B6.MSM-ChrX are defective in the fertilization. Alternatively, the embryos generated from the sperms of B6.MSM-ChrX are not able to develop beyond the first cleavage of the fertilized eggs. Histological analysis revealed that the B6.MSM-ChrX males have reduced testis weight with variable degree of degeneration in spermatogenesis and severe anomaly in the sperm head.

In order to map gene(s) responsible for the reduced fecundity of B6.MSM-ChrX males, we carried out quantitative trait loci (QTL) analysis with respect to the reduced testis weight and the sperm anomaly. We mapped a major QTL, responsible for the reduced testis weight, to the distal region of the X-chromosome. The other QTLs affecting sperm anomaly was mapped to three loci on the X chromosome. In the central region of the X-chromosome, where the highest likelihood ratio was located, several candidate genes have been mapped so far. One of them, *Halap-X* encoding haploid-specific alanin-rich acidic protein, is located to this region and expressed in the nucleiplasm of spermatids during the spermiogenesis. It likely functions for chromatin condensation in the sperm head. Comparative sequence analysis of *Halap-X* showed polymorphism of nucleotide sequence between C57BL/6J and MSM strains. It is most likely that mouse X-chromosome has crucial genes that control spermatogenesis through epistatic interaction with other genes on autosome(s) and /or the Y-chromosome. The sterility of the B6.MSM-ChrX males is possibly caused by incompatibility of MSM allele(s) of X-linked gene(s) and C57BL/6J allele(s) of the second gene(s) in autosome(s) and/or Y-chromosome. The disruption of the interaction may occur during the process of backcrossing to produce the consomic strains, and result in the sterility. The whole system uncovered in this study may represent reproductive isolation between two mouse subspecies, *Mus musculus domesticus* and *Mus musculus molossinus*.

Publication

1. Yada, Y., Makino, S., Chigusa-Ishiwa, S. and Shiroishi, T.: The mouse polydactylous mutation, luxate (*lx*), causes anterior shift of the anteroposterior border in the developing hindlimb bud. Int. J. Dev. Biol. **46**, 975-982, 2002.

2. Furuse, T., Blizard, D. A., Moriwaki, K., Miura, Y., Yagasaki, K., Shiroishi, R. and Koide, T.: Genetic diversity underlying capsaicin intake in the Mishima battery of mouse strains. Brain Res. Bull. **57**, 49-55, 2002.

3. Chan, B. W. H., Chan, Kwok-S., Koide, T., Yeung, Sau-m., Leung, M. B. W., Copp, A. J., Loeken, M. R., Shiroishi, T. and Shum, A. S. W.: Maternal Diabetes Increases the Risk if Caudal Regression Casused by Retinoic Acid. Diabetes, **51**, 2811-2816, 2002.

4. Isobe, T., Yoshino, M., Mizuno, K-I., Lindahl, K. F., Kiode, T., Gaudieri, S., Gojobori, T. and Shiroishi, T.: Molecular characterization of the Pb recombination hotspot in the mouse major histocompatibility complex class II region. Genomics, **80**, 229-235, 2002.

5. Furuse, T., Takano-Shimizu, T., Moriwaki, K., Shiroshi, T. and Koide, T.: QTL analyses of spontaneous activity by using mouse strains from Mishima battery. Mammal. Genet. **13**, 411-415, 2002.

F-b. MAMMALIAN DEVELOPMENTAL LABORATORY

(1) ecular mechanism of somite segmentation

Yu TAKAHASHI (National Institute of Health Sciences) and Yumiko SAGA

The somite is the first morphologically distinct segmental unit formed in a vertebrate embryo and give rise to metameric structures such as vertebrae, ribs and skeletal muscles. Each somite is subdivided into anterior (rostral) and posterior (caudal) compartments that differ in their properties and gene expression. The rostro-caudal polarity of a somite is established within the PSM prior to segmentation. Our aim is to understand the mechanisms underlying generation of a segmental pattern within the PSM. Previously, we cloned *Mesp2*, a gene encoding a novel bHLH transcription factor, MesP2, which is expressed in the rostral PSM. *Mesp2*-null mice exhibited defective somitogenesis due to a lack of rostral somitic compartment. By the genetic analysis, we have shown that MesP2 plays a critical role in establishment of rostro-caudal polariy of a somite by regulating *Dll1* expression via Notch signaling pathways.

One feature of the Notch signaling system in somitogenesis is that it utilizes several ligands. In mouse embryos, at least two Notch ligands, Dll1 and Dll3, are co-expressed in the PSM, and their expression domains are finally segregated into the rostral or caudal half of formed somites. Despite a large number of studies, the functional significance of two distinct Notch ligands in the PSM is not clear. Furthermore, the role of Psen1, a protein involved in nuclear translocation of the Notch intracellular domain, is not fully understood. Therefore, we genetically analyzed the role in rostro-caudal patterning triggered by Notch signaling mediated by either Dll1 or Dll3, the relationship between Notch signaling and Mesp2 function, and the involvement of Psen1 in Dll1- and Dll3-mediated Notch pathways. Our analyses revealed several novel findings: 1) Dll1, Dll3 and Mesp2 constitute a complex signaling network for rostro-caudal patterning in the anterior PSM. 2) Mesp2 function in the rostro-caudal pattering downstream of Dll1 or Dll3. 3) Psen1 is differently involved in Dll1-Notch and Dll3-Notch pathways. 4) Dll3-Notch signaling counteracts with the Psen1-dependent Dll1-Notch signaling. Based on these findings, we propose a new model, which is different from the previous concept that rostro-caudal patterning, i.e., formation of half-a-somite stripe pattern of gene expression, can be regarded as a result of stabilization of oscillating expression in the posterior PSM.

(2) Functional analyses of hesr1, 2 and 3

Hiroki KOKUBO and Yumiko SAGA

Notch signaling pathway has been implicated in development of the cardiovascular system. *hairy and enhancer of split* genes are known to be direct targets of Notch signaling pathway in *Drosophila* and vertebrates. We have cloned novel mouse hairy and E(spl) related subfamily genes, called *hesr -1, -2,* and *-3.* These genes are expressed in the developing heart, somites, limb bud and other tissues. To investigate the function of these genes, we have generated single gene-knockout mice for *hesr1, hesr2* or *hesr3.* Among them *hesr2*-knockout mouse showed abnormalities in heart morphogenesis. *hesr2* mutant mice are high lethal in early postnatal period with congenital heart defects, featured by the enlarged hearts. We successfully performed the transthoracic echocardiography in 5-day-old mice to reveal the heart dysfunction. A time motion (M-) mode echocardiogram of the left ventricle (LV) indicated the dilated LV chamber with markedly diminished in fractional shortening of the LV. The Doppler echocardiogram revealed the atrio-ventricular (AV) valve regurgitation. Morphological and histological analyses of homozygous hearts demonstrated that the heart dysfunction was caused by septal defects in subaortic ventricle and secundum atrium, and the regurgitation in the dysplastic and stenotic tricuspid valve. In addition, the heart rate in homozygous mice significantly decreased. We have noticed a drastic increase in the expression of s-myosin heavy chain, which could lead to a change of the heart rate. These observations strongly indicate that hesr2 plays an important role in formation of functional AV valves and thus, *hesr2* gene null mice provide a useful model system with which to elucidate genetic bases of the heart dysfunction, a common disease in human infants.

(3) The lineage analysis of Mesp1-expressing cells

Satoshi KITAJIMA (National Institute of Health Sciences) and Yumiko SAGA

MesP1 and MesP2 are transcription factors containing an almost identical bHLH motif. Cells lack both genes are unable to contribute to heart formation. A lineage study using Cre-lox system revealed that *Mesp1* is the earliest molecular marker expressed in heart precursor cells. We have conducted a detail lineage analysis using Cre-reportor line, Rosa-26. The heart is exclusively composed of cells derived from *Mesp1*-expressing mesoderm. However, we have found some cells are deprived of the β-gal staining, which means that these are derived from *Mesp1*-nonexpressing cells. Among them, neural crest cells have been shown to contribute to the conotrucus septum prior to and during overt septation of the outflow tract. In addition, we found several cell clusters in between primitive ventricles and trabecular component of both right and left ventricles, which distribution is reasonably similar to the cells defined as the cardiac conduction system. To define whether these cell are really contribute cardiac conduction system, we compared distribution of *Mesp1*-nonexpressing cells with cells in the conduction system using CCS-LacZ transgenic mouse, in which only cells in conduction system are stained by β-gal. The result strongly suggests that *Mesp1*-nonexpressing cells corresponds to cells in the conduction system. In addition, we used *Connexin-40* (*Cx40*) as a molecular marker for the conduction system. The *Cx40* signal and *Mesp1*-LacZ signal are segregated, which indicates that the *Mesp1*-nonexpressing cells are *Cx40*-positive. These observations strongly suggest that the origin of the conduction system is different from other cardiac cells.

(4) Cloning and functional analysis of mouse nanos genes

Masayuki TSUDA, Yumiko SASAOKA, Makoto KISO, Seiki HARAGUCHI (Shiga university of Medicine) and Yumiko SAGA

A *Drosophila* gene *nanos* encodes a RNA-binding zinc-finger protein that is shown to be involved in germ cell development by suppressing somatic gene expression. To know the involvement of mouse nanos proteins in germ cell development, we have isolated three mouse *nanos* genes (There are designated as *nanos1*, *nanos2* and *nanos3*). *nanos1* is maternally expressed in an unfertilized egg and the zygotic expression is observed in the neuronal cell lineage, especially in the adult hippocampus, but it was not expressed in developing germ cells. The knockout mouse developed without any significant abnormalities. Interestingly, the maternal expression was also observed for *nanos2*, and the expression pattern was quite similar to that of *nanos1*. Both *nanos2* and *nanos3* are expressed in male germ cells between embryonic days 13.5 and 15.5, and *nanos3* expression is also observed in the bipotential gonad of 11.5 dpc and in 12.5-dpc female gonad. In addition, *nanos3* is expressed in PGCs at as early as 9.5 dpc. The functions of these genes are examined by generating gene-knockout mice for *nanos2* and *nanos3*. A defect of *nanos2*-null mouse was observed only in male spermatogenesis. The adult testes showed reduced in size and weight and no germ cell was detected. In contrast, female gonad normally developed and the homozygous female was fertile, indicating no maternal defect due to the lack of *nanos2* expression in the oocyte. As expected, apoptotic cells were observed in the testes after 15.5 dpc and the apoptosis continued up to 18. 5 dpc, indicating nanos 2 is required for the maintenance of germ cells in the embryonic testes.

Reflecting its expression pattern, we found defects in both ovary and testes of *nanos3*-null mice. In adult male and female mutants, no germ cell was observed. In contrast to *nanos2*-null mice, germ cells were already missing in the genital ridge, at as early as 12.5 dpc, indicating that PGC failed to enter the genital ridge. The detailed mechanism leading to a loss of PGC is now under investigation.

Publications

1. Nomura-Kitabayashi, A., Takahashi, Y., Kitajima, S., Inoue, T., Takeda, H. and Saga, Y.: Hypomorphic Mesp allele distinguishes establishment of rostrocaudal polarity and segment border formation in somitogenesis. *Development*, **129**, 2473-2481, 2002.

2. Ohara, O., Nagase, T., Mitsui, G., Kohga, H., Kikuno, R., Hiraoka, S., Takahashi, Y., Kitajima, S., Saga, Y. and Koseki, H.: Characterization of size-fractionated cDNA libraries generated by the in vitro recombination-assisted method. *DNA Res.* **9**, 47-57,

2002.

3. Katoh, C., Kitajima, S., Saga, Y., Kanno, J., Horii, I. and Inoue, T.: Assessment of quantitative dual-parameter flow cytometric analysis for the evaluation of testicular toxicity usingcyclophosphamide- and ethinylestradiol-treated rats. J. Toxicol. *Sci.* **27**, 87-96, 2002.

F-c. MOUSE GENOMICS RESOURCE LABORATORY

(1) Quantitative trait analyses of sensitivities to heat evoked pain and capsaicin in mouse strains.

Tamio FURUSE[1], Kazumi YAGASAKI[1], Toshihiko SHIROISHI[2] and Tsuyoshi KOIDE ([1]Tokyo University of Agriculture and Technology, [2]Mammalian Genetics Laboratory, NIG)

Pain sensation is an essential alert for avoiding environmental danger and feeling tissue damage or illness. In spite of the essential role of pain sensation, surplus pain is stressful or unpleasant for animals. Heat evoked pain and capsaicin (8-methyl-N-vanillyl-6-nonenamide) sensation activate the multi functional pain receptor, VR1. Capsaicin is the main chemical component of hot chili peppers providing its hot taste and causes sensation in physiological pain system. In the previous annual report, we described about genetic diversity of sensitivity for capsaicin and pain in the stock of mouse strains, Mishima battery (1). In the present study, we conducted QTL analyses to characterize genetic loci involved in the different sensitivity to capsaicin and heat evoked pain between two mouse strains, C57BL/6 and KJR. In the study, 470 F2 mice were generated from a cross of F1 mice obtained from a cross of C57BL/6 and KJR. We applied fluid intake test of capsaicin solution and 52℃ hot plate test for the F2 progeny. In the mapping study, four significant QTLs for capsaicin sensitivity were detected on chr2, chr7, chr8 and two significant QTLs for heat sensitivity were detected on chr8 and 18.

(2) Analysis of *Loco1* and *Loco2* involved in spontaneous locomotor activity of wild-derived mouse strains

Juzoh UMEMORI, Tamio FURUSE[1], Hiromi SANO[3], Kazuto KOBAYASHI[3], Toshihiko SHIROISHI[2] and Tsuyoshi KOIDE ([1]Tokyo University of Agriculture and Technology, [2]Mammalian Genetics Laboratory, NIG, [3]Fukushima Med. Univ.)

It has been previously described that spontaneous activities of mice are dependent on a light/dark rhythm. Measuring spontaneous activity of mice revealed that most of mice are active during dark period, and the spontaneous activity is quite different among common laboratory inbred strains and wild mice-derived inbred strains. In these strains, mice of KJR strain were especially hyperactive in contrast to mice of BLG2 which were hypoactive. The F1 females from the cross between KJR and BLG2 performed the same levels of spontaneous activity as KJR mice, suggesting some dominance of the hyperactive phenotype. A genetic linkage analysis using the progeny from a backcross KJR to BLG2 revealed that two loci, *Loco1* and *Loco2*, are involved in spontaneous activity (2). We conducted phenotypic and molecular analyses to elucidate genetic mechanism of this hyperactive phenotype.

It has been proposed that hyperactivity is related to dopamine and dopaminergic system. Therefore we determined level of Dopamine (DA) and DA degradation products, DOPAC (dihydroxyphenylacetate) and HVA (Homovanilate) in midbrain and striatum. In the striatum of KJR, levels of DA, DOPAC and HVA were significantly higher than that of BLG2. These data suggested that *Loco1* and *Loco2* might be involved in elevated levels of DA, DOPAC, and HVA in striatum of KJR. We also compared the localization and expression levels of TH (tyrosine hydorogenase), DRD1 (Dopamine receptor D1A), and DRD2 (Dopamine receptor D2) between KJR and BLG2. As a result, expression levels of DRD2 in striatum and midbrain of BLG2 were significantly higher than that of KJR. This result suggested that D2DR is down regulated according to high level of dopamine in stiratum of KJR.

Publication

(1) Furuse, T., Blizard, D. A., Moriwaki, K., Miura, Y., Yagasaki, K., Shiroishi, T. and Koide, T.: Genetic diversity underlying capsaicin intake in the Mishima battery of mouse strains. Brain Research Bulletin **57**, 49-55, 2002.

(2) Furuse, T., Takano-Shimizu, T., Moriwaki, K., Shiroishi, T. and Koide, T.: QTL analyses of

spontaneous locomotor activity using mouse strains from Mishima battery. *Mammalian Genome* **13**, 411-415, 2002.

(3) Chan, B. W. H., Chan, K.-S., Koide, T., Yeung, S.-M., Leung, M. B. W., Copp, A. J., Loeken, M. R., Shiroishi, T. and Shum, A. S. W.: Maternal Diabetes Increases the Risk of Caudal Regression Caused by Retinoic Acid. Diabetes **51**, 2811-2816, 2002.

(4) Isobe, T., Yoshino, M., Mizuno, K., Fisher Lindahl, K., Koide, T., Gaudiery, S., Gojobori, T. and Shiroishi, T.: Molecular characterization of the Pb recombination hotspot in the mouse major histocompatibility complex class I region. Genomics **80**, 229-235, 2002.

F-d. PLANT GENETICS LABORATORY

(1) Genetic dissection of rice development

1-A. Regulation of expression of *KNOX* family class 1 homeobox genes of rice

Yukihiro ITO and Nori KURATA

To elucidate genetic programmes that control embryogenesis and regeneration of rice, we conducted cloning and structural and functional analyses of genes which encode transcription factors and protein kinases. We previously identified five *KNOX* family class 1 homeobox genes and analysed their expression patterns during early embryogenesis and regeneration by RT-PCR and *in situ* hybridization. We also found that constitutive expression of these genes is sufficient to maintain cells in a meristematic undifferentiated state. Since specific expression of *KNOX* family class 1 homeobox genes in a shoot meristem is essential for normal development of plants, we started to study its regulatory mechanisms.

We showed that the coding region of *OSH1* is involved in the regulation of its expression. Introduction of *OSH1* cDNA without a promoter sequence induced expression of both endogenous and introduced *OSH1* in a leaf and resulted in altered leaf morphology. This indicates that an extra copy of *OSH1* induces ectopic expression of the endogenous *OSH1* and the *OSH1* cDNA sequence has ability to induce its own expression in leaves (in preparation). In addition, we found that cytokinin which is known to induce shoot regeneration from calli induces expression of *OSH1* in calli and suspension culture cells of rice. Mechanisms involved in the induction and repression of *OSH1* expression and its relation to rice development will be examined.

1-B. Structural and functional analysis of rice *OsHAP3* genes

1-B-1, Identification and expression analysis of *OsHAP3* genes

Nori KURATA, Kazumaru MIYOSHI, Thirumurugan THIRUVENGADAM and Yukihiro ITO

We identified 11 genes that encode subunits of CAAT-box binding complex in rice by cDNA screening and database search. We classified these genes into two groups based on the deduced amino acid sequences, exon-intron structure and expression patterns. The first group consists of three genes named *OsHAP3A*, *OsHAP3B* and *OsHAP3C*, which are expressed ubiquitously. The second group consists of eight genes designated *OsHAP3D* to *OsHAP3K*, whose expression is restricted in specific meristem-containing organs at specific stages depending on genes. Detail analysis of expression of these genes by *in situ* hybridization and promoter-GUS fusion is underway.

1-B-2, Functional analysis of *OsHAP3A*

Yukihiro ITO, Kazumaru MIYOSHI and Nori KURATA

To examine roles of *HAP3s* on rice development, we generated antisense and RNAi transgenic rice plants of *OsHAP3A*, a member of the first group. The transformants showed reduced expression of *OsHAP3A* and had pale green leaves. This phenotype was inherited to the next generation, albeit low penetrance, and detail analysis was carried out with progenies in the next generation.

We examined the chlorophyll content in the third leaf blades of 2-week seedlings and showed that the chlorophyll content of the antisense plant was about half of that of the vector-transformed plant. Transmission electron microscope analysis showed that chloroplasts of the antisense plant were greatly affected. Development of tyrakoid membrane was perturbed and no accumulation of starch grain was

observed. Quantitative RT-PCR analysis showed that expression of nuclear-encoded photosynthesis genes such as *RBCS* and *cab* was reduced, while expression of chloroplast-encoded genes such as *RBCL* and *RPL23* was unaffected or rather up-regulated. We also examined the expression level of *OsHAP3B* and *OsHAP3C*, two genes highly related to *OsHAP3A* in sequence. In the antisense plants, expression of *OsHAP3B* and *OsHAP3C* were also reduced compared with the vector-transformed plant. These results suggest that *OsHAP3* genes are involved in the normal biogenesis of chloroplast through gene regulation (submitted). Considering that *LEC1*, a member of *NF-YB/HAP3* genes of *Arabidopsis*, is a critical regulator of embryogenesis, plant *NF-YB/HAP3* genes may have diverse functions ranging from chloroplast biogenesis to embryogenesis depending on genes.

1-B-3, Functional analysis of *OsHAP3D and OsHAP3E*

Thirumurugan THIRUVENGADAM, Yukihiro ITO, Kazumaru MIYOSHI and Nori KURATA

We have been generating transgenic rice plants in which the second group *OsHAP3* genes such as *OsHAP3D*, *OsHAP3E*, *OsHAP3F*, *OsHAP3G* and *OsHAP3H* are overexpressed or suppressed by antisense or RNAi. Among them, overexpressing plants of *OsHAP3D* and *OsHAP3E*, showing high similarity with each other, showed abnormal phenotypes. Transformed calli regenerated shoots with narrow leaves. After transfer to soil, the transformants were dwarf and had small leaves. These results suggest that *OsNF-YB6* is involved in normal development of rice, and its overexpression somehow modified rice development and resulted in abnormal phenotypes.

1-C. Genetic analysis of sporogenesis during sexual plant reproduction using sterile mutants of rice

Ken-Ichi NONOMURA, Toshiyuki FUKUDA, Mutsuko NAKANO, Mitsugu EIGUCHI, Akio MIYAO*, Hirohiko HIROCHIKA* and Nori KURATA (* Natl. Inst. Agrobiol. Sci.)

Genetic analyses for microsporogenesis and megasporogenesis is important to understand the system of plant meiosis and sexual reproduction. We identified and characterized three genes of rice, *MSP1*, *PAIR1*, and *PAIR2*, tagged by transposed *Tos17*. The *MSP1* (multiple sporocytes) gene encodes a leucin-rich-repeat receptor-like protein kinase. A loss-of-function mutation in the *MSP1* gives rise to excess number of both male and female sporocytes. In the anther, development of inner wall layer is disordered and then tapetum cells are completely absent. The *in situ* expression of the *MSP1* suggests that *MSP1* products play important roles in preventing the ectopic male and female sporogenesis during flower development (submitted).

The *PAIR1* gene encodes a novel protein with a putative coiled-coil domain, and the *PAIR2* gene is a homologue of *Arabidopsis ASY1*. A loss-of-function mutations in both genes showed failure in homologous chromosome pairing. Furthermore, in the *pair1* mutant, second meiotic cell division is also defected, resulting in formation of polyad spores instead of tetrad. The phenotype of *pair1* mutant resembles to a mutation in *Arabidopsis SYN1*, a gene encoding a cohesin protein. These facts raise a possibility that the genes identified in this study play important roles in meiotic chromosome pairing as cooperating with the cohesin complex.

1-D. Molecular and genetic analysis of *PLASTOCHRON* genes, timekeepers for shoot development in rice

Heterochronic mutations affecting the timing of developmental events may be of major significance in onotogeny and evolution. In plants, several heterochronic mutations that affect stepwise development of vegetative tissue and therefore alter shoot architecture have been identified. The recessive mutations at the rice *PLASTCHRON* loci (*PLA1* and *PLA2*) cause the short plastochron and ectopic expression of vegetative programs in the reproductive phase, indicating that the *PLA* genes play important roles in temporal regulation of shoot development. To clarify the mechanism operating in the time-controlling process of shoot development, we have been analyzing *PLA* genes.

1-D-1, Cloning and characterization of *PLA1* gene encoding a cytochrome P450 protein, CYP78A11.

Byoung-Ohg AHN, Kazumaru MIYOSHI, Yukihiro ITO, Jun-Ichi ITOH, Yasuo NAGATO and Nori KURATA (Graduate School of Agricultural and Life Science,

University of Tokyo)

In the previous work, we already identified strong candidate gene for *PLA1* that encodes CYP78A11, a member of cytochrome P450 proteins (submitted). Some members of the plant P450 family are known to be involved in important biochemical pathways such as biosynthesis of phenylpropanoids, lipids, glucosinolates, and plant growth regulators such as gibberellins, jasmonic acid, and brassinosteroids. The candidate P450 gene carried mutations in all *pla1* alleles (*pla1-1, -2, -3* and *pla1-4*). The predicted protein contains a hydrophobic region in the N-terminus and putative oxygen and heme-binding domains that are characteristic of microsomal cytochrome P450. Complementation experiment that the wild-type P450 gene rescued the *pla1*phenotype, confirmed that the gene was a real *PLA1* gene. Examination of temporal and spatial expression patterns of *PLA1* by *in situ* hybridization revealed that *PLA1* is expressed exclusively in young leaf but was not in shoot apical meristem (SAM). The expression pattern, potential enzymatic activity of the protein and the mutant phenotype of *PLA1* suggest that *PLA1* acts directly or indirectly in the synthesis of inhibitory substances in leaf primordia. *PLA1* could be a molecular clue for the inhibitory field theory, an oldest but the most widely accepted model for leaf initiation process.

1-D-2, Chromosomal location of *PLA2* gene

Kazumaru MIYOSHI, Yukihiro ITO, Taiji KAWAKATSU, Jun-Ichi ITOH, Yasuo NAGATO and Nori KURATA (Graduate School of Agr. and Life Sci., University of Tokyo)

The *pla2* mutant shows similar but more severe phenotype to that of *pla1*. The double mutant (*pla1-2/pla2*) took the phenotype of *pla2*, indicating that *PLA2* is epistatic to *PLA1*. Thus, *PLA2* gene would function upstream of *PLA1* gene in shoot development. Elucidation of functional relationship between these genes would help to clarify the mechanism for the time-controlling process of shoot development. As the first step to isolate *PLA2* gene, we mapped the *PLA2* gene on an approximate chromosomal location. F2 progeny analysis using 179 *pla2* homozygous plants revealed that *PLA2* is located on the long arm of chromosome 1 within the interval of 3.6 cM between the marker L819 and S10562. Further fine mapping and cloning are in progress.

1-E. Isolation and sequence analysis of receptor-like protein kinase genes of rice

Yukihiro ITO and Nori KURATA

We started to isolate and characterize protein kinase genes as candidates to transduce positional information which is postulated to be important for plant development. We focused on a receptor-like protein kinase gene such as *SOMATIC EMBRYOGENESIS RECEPTOR KINASE* (*SERK*) of carrot and isolated its homologues from rice. We have cloned two genes and those cDNAs were named *ORK1* and *ORK6*. The genome clone of *ORK1* showed unusual structure and *ORK6* encodes a receptor-like protein kinase with leucine-rich repeats in a receptor domain. The amino acid sequence of *ORK6* is highly similar to plant receptor-like protein kinases such as SERK, CLAVATA1 and ERECTA of *Arabidopsis*, which are known to function in a shoot meristem. The *ORK6* promoter fused to GUS directed GUS activity in a root tip, a lamina joint and a callus. Overexpression of *ORK6* or introduction of antisense or RNAi constructs of *ORK6* into rice showed no phenotypic change under normal growth condition. Analysis of regeneration is underway using calli derived from these transgenic plants.

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

Yukihiro ITO, Kazumaru MIYOSHI, Thirumurugan THIRUVENGADAM, Mitsugu EIGUCHI and Nori KURATA

To isolate valuable mutants defective in various steps of rice development and to clone the corresponding genes, we are generating enhancer trap lines of rice. We employ an enhancer trap system used in *Arabidopsis* with some modifications as described in the previous reports in 2000 and 2001. This system proved to transpose well the Ds-GUS element by Ac-transposase in rice.

We continued screening of transposants from the four *Ds-GUS* lines by the chlorsulfuron and hygromycin resistance. Out of 18,020 F2 plants screened in two years, 3,711 were judged to be transposants. We

also crossed with other 22 *Ds-GUS* lines with the 35S-AcTPase line, and screening of transposants was carried out. Out of 6,590 F2 plants screened by PCR, 1,874 plants (28%) were judged to be transposants. Using these transposants, we also determined the suitable concentration of chlorsulfuron for screening

Screening of about 800 transposant lines for embryonic mutants, we identified two lines with an abnormal embryo size. One line produced seeds with a greater embryo. Seed size and differentiation of embryonic organs seemed normal. The other line produced seeds with a smaller embryo. The size of embryos was greatly different from seed to seed, and in an extreme case a seed had no embryo. However, linkage analysis indicated that these genes are not tagged by *Ds-GUS*.

We have also been screening enhancer trap lines using GUS gene as a reporter. A total of 60 lines which showed *GUS* expression in any organ are now examined to detect flanking sequences of the trapped sites. GUS staining pattern and flanking sequence data of these enhancer trap lines are being collected to house in the rice development database which we collaboratively construct as a rice functional genomics database.

(3) Analysis of nuclear and chromosomal factors that maintain genome function

3-A. Finding of various plant nuclear proteins using yeast nuclear transportation trap system.

Kazuki MORIGUCHI and Nori KURATA

Nucleus is the most important intracellular organelle, in which replication and gene expression are regulated by accepting signals from environment as the terminal place and responding them as the first place. Although the genome projects of *Arabidopsis* and *Oryza* have been finished, the large-scale functional analyses of these genes are just at its beginning on their assignment of functions.

We have been trying a comprehensive analysis of plant nuclear proteins by yeast nuclear transportation trap (NTT) system using *Oryza sativa* cDNA libraries. At present, about 520 gene species were obtained. Based on homology search results, 33% of them encodes nuclear and nuclear/cytoplasmic proteins, but 43% and 24% of them are shown to encode functionally unknown and generally thought as non-nuclear proteins, respectively. By observing the localization of a subset of those fusion proteins with GFP, we confirmed that nuclear localizing proteins were actually enriched not only in those unknown function proteins group but also even in those non-nuclear proteins group. These results showed that it is difficult to predict nuclear proteins only by in silico analysis based on the previous data. We also analyzed expression profiles of the isolated nuclear protein genes in several organs. At present, 43 and 29 genes, which are up-regulated at meiotic stage and early embryo enriched panicles, respectively, have been found by microarray analysis. Relating to these genes, two approaches will be applied. One is the analysis of nuclear-cytoplasmic shuttling proteins, which were found during our NTT screening. We have been generating transgenic rice strains introduced GFP-tagged those genes. The other is the analysis of the plant nuclear proteins network using yeast two-hybrid system. We are mainly focusing on plant nuclear matrix proteins, those content has been thought to be quite different from those in mammals.

3-B. Construction and introduction of rice artificial chromosomes

Tadzunu SUZUKI, Ken-Ichi NONOMURA and Nori KURATA

Generation of transgenic rice plants which harbor artificial chromosomes will be a useful tool not only to elucidate factors necessary for maintaining chromosome architecture and basal function but also in introducing large genomic fragments into rice under controlled condition. To create rice artificial chromosomes, the candidate DNA construction and application of several transformation methods on rice cells have been carried out. The candidate DNA fragments of 40-100 kb derived from centromere regions of rice were integrated into bacterial artificial chromosomes (BAC). A YAC vector carrying 300 kb chromosome 5 centromeric region, telomere repeats and selection markers had also constructed.

The transfer method of huge size DNAs into plant cells have not been established. For introduction of the constructed artificial chromosome vectors, two methods of lipofection and particle bombardment system were examined here. Trials of lipofection method using small

plasmids or BAC vectors and rice cultured cells, showed it worked at least for transient expression of the small plasmid. Particle delivery system was shown more effective. Several lines of calli, which survived on a selection media, were obtained by particle bombardment introduction of BAC vectors, although the growth of these calli was slower than that of the calli transformed with control vectors. RT-PCR analysis showed that only one line expressed the marker gene while the others did not or less. This might suggest that the marker gene was suppressed by the centromeric repeats inserted in the vector, or hopefully that the construct existed unstably as an episomal state. Further examination wil be needed.

(4) Genome-wide analysis of reproductive barriers and positional cloning of a reproductive barrier

4-A. Interactions of reproductive barriers

Yoshiaki HARUSHIMA and Nori KURATA

Genetic mechanisms for isolation of 'species' are called as reproductive barriers and these include hybrid incompatibility, hybrid inviability, hybrid sterility, hybrid breakdown, etc. The distortions of allele frequencies from Mendelian expectation in progeny of inter- or intra- species hybrid due to hybrid sterility genes, hybrid breakdown genes and gametophytic competition genes have been often observed. We have studied reproductive barriers between Japonica and Indica cultivars. We previously mapped all reproductive barriers causing allele frequency distortions from Mendelian expectation in Nipponbare (Japonica) and Kasalath (Indica) F_2 populations. These mapped reproductive barriers would be resulted from interactions between alleles, between segregating different loci within a gametophyte or a zygote, between genome and cytoplasmic factors, and between parent genome and progeny genome. To detect interactive loci within a segregating gametophyte or zygote in F_2 population, a simple χ^2 test for independence of allele frequencies for all pairs of 1055 DNA markers covering the entire genome were performed. A high value of χ^2 would be expected for linked marker pairs and for marker pairs in the vicinity of interactive reproductive barriers. The contour plots of the χ^2 results are symmetrical about the diagonal and high χ^2 values in that region are due to linkage between markers. There were regions showing high χ^2 values other than in the diagonal region. The χ^2 values show peaks. Some DNA markers of interaction peaks were also the closest markers to the reproductive barriers identified by regression analysis, other interactive loci did not induce apparent segregation distortion. To verify relevance between the barriers and the interaction peaks, favorable or unfavorable genotype combinations were investigated. Several zygotic barriers might be resulted from the interactions detected in this study; however there were no relevance between the gametophytic barriers and the interaction peaks. To investigate reproducibility of the interaction peaks, we performed the independence test of the non-linked marker combinations with the high χ^2 values by another population from the same cross. Some interactive peak that close to reproductive barrier were reproducible; however most of the combinations of the markers that were not close to the reproductive barriers did not show high χ^2 values in the different population.

4-B. 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 F_2 of Nipponbare-Kasalath hybrid by positional cloning, and elucidation of the molecular nature of the individual reproductive barriers. We have clarified the aimed gene was a male gametophyte gene that interact with maternal locus on chromosome 6. In other words, the pollen with Kasalath genotype at the gametophyte gene preferentially fertilized by 94% probability in maternal plant that is heterozygote or Kasalath homozygote at the interactive locus on chromosome 6. For detailed mapping the gametophyte gene, we have selected plants with recombination in 1.9 cM interval from 1000 F_2 and 473 backcross plants and retrieved their genomic DNA of the selected plants from bulked young leaves of their selfed seedling. Considering the genotypes of the interactive locus, the genetic map of the male gametophyte gene is being constructed by the dosage analysis using the selected population. The physical map of this region was constructed using Nipponbare BACs and PACs. A PAC that cover the

candidate region was sequenced and physical length of the candidate region was about 50 kb and there was no detectable insertion or deletion between Nipponbare and Kasalath by Southern analysis. To narrow down the candidate region, we found the flanking PCR markers of the candidate region by the sequence information, selected recombinants from another 2600 F_2 plants by the flanking PCR markers, and mapped the aimed gene precisely. We narrowed down the candidate region to 5.3 Kb. Sequencing of the corresponding Kasalath genome, and survey of gene expression in pollen from the candidate region are in progress.

(5) A new resource work for wild rice species involved in the activity as a center of national bioresource project for rice.

Nori KURATA, Toshie MIYABAYASHI, Yukiko YAMAZAKI*, Mitsugu EIGUCHI and Ken-Ichi NONOMURA (* Genetic Strains Information Lab.)

In 2002, the national bioresource project started for representative genetic and biological resources under the MEXT. The plant genetics laboratory and experimental farm act as a center for rice genetic resources project. In this project, we deal with about 2000 wild rice accessions including 9 genomes and over 20 species collected all over the world. Propagation, seed storage, distribution and database construction for wild rice species are progressed. Details are shown in the report of Experimental Farm.

Publications

1. Harushima, Y., Nakagahra, M., Yano, M., Sasaki, T. and Kurata, N.: Diverse variation of reproductive barriers in three intraspecific rice crosses. Genetics. 313-322, 2002.

2. Ahn, B. O., Miyoshi, K., Itoh, J. I., Nagato, Y. and Kurata, N.: A genetic and physical mapping of the region containing *PLASTOCHRON1*,a *heterochronic* gene, in rice (*Oryza sativa*, L). Theor. Appl. Genet. **105**, 654-659, 2002.

3. Ito, Y., Hirochika, H. and Kurata, N.: Organ specific alternative transcripts of KNOX family class 2 homeobox genes of rice. Gene, **288**, 41-47, 2002.

4. Dunford, R.P., Yano, M., Kurata, N., Sasaki, T., Huestis, G., Rocheford, T. and Laurie, D.A.: Comparative mapping of the barley *Ppd-H1* photoperiod response gene region, which lies close to a junction between two rice linkage segments. Genetics, **161**, 825-834, 2002.

5. Kurata, N., Nonomura, K-I. and Harushima, Y.: Rice genome organization: the centromere and genome interactions *Annals of Botany* **90**, 427-435, 2002.

6. Yamazaki, Y., Yoshimura, A., Nagato, Y. and Kurata, N.: Recent advances in Oryzabase (Integrated Rice Database). *Rice Genetics Newslet.* **19**, 7-8, 2002.

7. Kurata, N. and Fukui, K.: Chromosome research in genus Oryza. in Monograph in genus Oryza. Nanda JS. Ed, Springer (in press).

F-e. MICROBIAL GENETICS LABORATORY

(1) Timing of cell division

Akiko NISHIMURA

During the cell cycle,of *E. coli*, several fundamental events take place through strictly periodic processes, although the synthetic patterns of DNA, cell mass increase, and membrane in the cell cycle are completely different from each other. As a result, two identical daughter cells are produced. The cell must have mechanisms coordinating each event. Cell division, for example, occurs with constant period (20 min), and the coefficient of variation in cell size at the time of cell division is relatively small, at about 10%. These facts suggest that the *E. coli* must have mechanisms to the initiation of cell division at a specific cell length, and that the consistency of the cell cycle process is based on coordination mechanisms between various intracellular reactions and cell division. DNA replication and cell division also follow a strict timetable before two identical daughter cells can be produced. If cells were to divide too early, progeny would be smaller than normal and more importantly, replicating DNA strands could be torn apart and/or annucleoid cells could be produced. The cell must therefore have mechanisms for coordinating each event. I have isolated *cfc* mutants, which uncouple DNA replication and cell division. *E. coli* cells possessing *cfc*

mutations divide before reaching the optimal size at which *cfc*$^+$ cells divide, and many small cells with a single nucleoid are produced. Mutations affect the timing of cell division, but not the initiation mass for chromosome replication. This is due to a reduction in the period between nuclear division and cell division, and a compensatory increase in the interval between cell division and initiation of the next round of DNA replication. These mutations relieve cells from division arrest caused by the inhibition of DNA replication. As such, *cfc* mutants are defective in the signaling mechanisms that couple DNA replication and cell division. I have concluded that Ap4A forms part of this signaling mechanism. I have also isolated the Ap4A binding protein, which is structurally well conserved from human to bacteria. This displays nucleotide binding domains, and interacts with tubulin in eucaryotic cells. I demonstrated that knockout of this gene produces longer rod cells and over-expression of this gene produces numerous small cells. To analyze the molecular mechanisms of timing in cell division, I am investigating the mode of expression of this gene in a synchronized cell cycle, localization of gene products within cells, *in vitro* function of gene products, and relationships with Ap4A.

(2) 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 constructed a whole set of mobile plasmid clones of ORFs, facilitating genetic studies of *E. coli*. Using these clones, we have identified 229 cell division genes by the method of complementation. Of total identified, 5% were known fts genes, 27% were novel genes, 25% were involved in the key reaction such as DNA replication and cell division, 19% were involved in cellular processes such as ion transport and signal transduction, and 24% were involved in metabolism. These results support our speculation that the consistency of the cell cycle process is based on coordination mechanisms between various intracellular reactions and cell division. Using these mutants, we are planning to investigate the hierarchy of expression for these genes, and the effects of their mutations on global regulation of cells, using DNA chip techniques.

(3) Guanosine 5'-diphosphate 37-diphosphate (ppGpp) synthetic activities on *Escherichia coli* SpoT domains.

Chizuko FUJITA, Akiko NISHIMURA, Ryoko IWAMOTO and Kenji IKEHARA

Escherichia coli SpoT protein with 702 amino acid residues is a bifunctional enzyme catalyzing both guanosine 5'-diphosphate 3'-diphosphate (ppGpp) degradation and its synthesis. First, we investigated how many domains are comprised in SpoT protein, by limited hydrolysis of the protein with serine proteases, α-chromotrypsin and elastase. Based on the results, we deduced that SpoT protein is composed of two major domains, an N-terminal half domain from Met1 to Phe373 and a C-terminal half domain from Glu374 to Asn702 (C-terminal end). In addition, by a further α-chymotrypsin digestion two cleaved sites were found at Arg196 in the N-terminal half domain (D12) and at Lys475 in the C-terminal half domain (D34), respectively, to produce four minor domains, D1, D2, D3 and D4. Next, plasmids expressing the major two domains (D12 and D34) and four minor domains (D1, D2, D3 and D4) were constructed. Consequently, the deduced SpoT minor domains as well as the major domains were expressed as stable protein units except for D4. D4 may be also folded into a stable protein in *E. coli* cells, since high expression of D4 from a plasmid results in host cell lethality. *E. coli relA*$^-$, *spot* double null strains expressing D1, D2 and D12 recovered cell growth in M9 minimal medium, whereas the transformants of D3, D4 and D34 did not grow in the minimal medium. It indicates that ppGpp synthetic activities could be restricted in the N-terminal half domain (D12, D1 and D2). For details, see Ref. 3.

Pubrications

1. Fujishima, H., Nishimura, A., Wachi, M., Hirasawa, T., Teraoka, H., Nishimori, k., Kawabata, T., Nishikawa, K. and Nagai, K.: The *kdsA* mutations affect FtsZ-ring formation in *Escherichia coli* K-12.

Microbiol. **148**, 103-112, 2002.

2. Makinoshima, H., Nishimura, A. and Ishihama, A.: Fractionation of *E. coli* cell populations of different stages during growth transition to stationary phase. Mol. Microbiol. **43**, 269-280, 2003.

3. Fujita, C., Nishimura, A., Iwamoto, R. and Ikehara, K.: Guanosine 5'-diphosphate 3'-diphosphate (ppGpp) synthetic activities on *Escherichia coli* SpoT domains. Biosci. Biotechnol. Biochem. **66**, 1515-1523, 2002.

4. Makinoshima, H., Aizawa, S., Hayashi, H., Miki, T., Nishimura, A. and Ishihama, A.: Growth phase-coupled alterations in cell structure and function of *Escherichia coli*. J. Bacteriol. **185**, 1338-1345, 2002.

F-f. INVERTEBRATE GENETICS LABORATORY

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

Ryu UEDA, Misako TANIGUCHI, Kaoru SAIGO[1] and Kuniaki TAKAHASHI ([1]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 and has no complex organs such as an alimentary canal, cardiovascular system, or central nervous system, 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 13,500 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 13,500 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 13,500 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 2002, over 4000 transformation vectors had been constructed, 2500 of which have been successfully introduced into the fly. Considering the pace of this work, the construction of vectors for over 11,000 genes (accounting for 80% of the entire genome) should be accomplished during the next fiscal year. Completion of the transformation will require about one additional year. Clearly then, we will be able to obtain a powerful tool with which to determine the function of the entire fly genome in the near future.

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 1617 genes tested, 51.7% 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 30 groups. The usefulness of inducible RNAi for investigating gene function in *Drosophila* is being revealed in many aspects. We published 3 papers in 2002 using RNAi flies.

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

Publications

1. Inaki, M., Kojima, T., Ueda, R. and Saigo, K.: Requirements of high levels of Hedgehog signaling activity for medial-region cell fate determination in *Drosophila* leg: identification of *pxb*, a putative Hedgehog signaling attenuator gene repressed along the anterior-posterior compartment boundary. Mech. Develop. **116**, 3-18, 2002.

2. Leulier, F., Vidal, S.S., Saigo, K., Ueda, R. and Lemaitre, B.: Inducible expression of double strand RNA reveals a role of dFADD in the regulation of antibacterial response in *Drosophila*. Curr. Biol. **12**, 996-1000, 2002.

3. Hayashi, S., Ito, K., Sado, Y., Taniguchi, M., Akimoto, A., Takeuchi, H., Aigaki, T., Matsuzaki, F., Nakagoshi, H., Tanimura, T., Ueda, R., Uemura, T., Yoshihara, M. and Goto, S.: GETDB, a database compiling expression patterns and molecular locations of a collection of Gal4 enhancer traps. Genesis, **34**, 58-61, 2002.

4. Doi, N., Zenno, S., Ueda, R., Ohki-Hamazaki, H., Ui-Tei, K. and Saigo, K.: Requirement of Dicer and eIF2C translation initiation factors for short-interfering-RNA-mediated gene silencing in mammalian cells. Curr. Biol. **13**, 41-46, 2003.

F-g. LABORATORY FOR FRONTIER RESEARCH

Lineage Development in the Central Nervous System

Takako Isshki

Generation of multiple cell types during embryogenesis is orchestrated temporally as well as spatially. Yet we know relatively little about how different cell fates are generated over time. Neural progenitors often generate diverse neurons and glia in stereotyped order, and the cell fate of a neuron/glia is tightly linked to its birth order within a parental progenitor lineage. Our group focuses on elucidating the molecular mechanisms of birth order-dependent specification of neural cells within a lineage by employing a molecular genetic approach using the *Drosophila* central nervous system as a model system. Previously, we have reported that most *Drosophila* neuroblasts sequentially express the transcription factors Hunchback - Krüppel -Pdm-Castor over time, while differentiated daughter cells maintain the expression profile of the transcription factor present at their birth. Moreover, we showed that Hunchback and Krüppel, and possibly Pdm and Castor regulate temporal identity within a lineage. These findings have provided the clues to study the molecular mechanisms of temporal specification, and raised many new questions, such as how the expressions of the temporal

identity genes are regulated.

The regulatory genomic region driving the transcription of *hunchback* in the CNS was identified. This region can be divided into two parts: the part for driving Hb expression in most neuroblasts, and the part for maintaining Hb expression in differentiated daughter cells. Dissection and molecular analysis of the region should reveal what kind of enhancers are present, how transcriptional regulation within the CNS is coordinated. Furthermore, to identify the new factors responsible for driving the sequential expression of the "temporal identity" genes revealed so far, I started performing EMS mutagenesis, screening for alterations in several neuroblast lineages, with emphasis on the neuroblast 7-1 and 7-3 lineages.

Previous study showed that Hb and Kr could regulate the expressions of the subsequent temporal identity genes, suggesting that Hb and Kr themselves are components of the molecular clock keeping time in neuroblasts. In castor mutants, neuroblasts produce extra daughter cells with mid-born cell fates, thus progression from the middle to late stage of neuroblast lineage during development is severely perturbed or prolonged. This indicates that Castor also participates in the molecular clock as a factor accelerating the clock.

G. CENTER FOR GENETICS RESOURCE INFORMATION

G-a. GENETIC INFORMATICS LABORATORY

(1) PEC

Takehiro YAMAKAWA, Junichi KATO and Yukiko YAMAZAKI

PEC (Profiling of Escherichia coli Chromosome) database is an integrated public resource database which provides useful information to characterize the gene function of Escherichia coli. A unique characteristic of the PEC is the gene classification based on essentiality for cell growth. All of the Escherichia coli genes have been classified into three groups, (1) gene essential for cell growth (essential), (2) those dispensable for cell growth (non-essential), and (3) those unknown to be essential or non-essential (unknown), mainly using information from deletion mutants and journal articles. Another unique characteristic is the comparative gene analysis among bacterial genomes. PEC provides the list of numbers of homologues genes found in other bacteria genomes analyzed three times with different e-values, so that species-specific genes and/or orthologous genes are easily detected. In the database, each gene has publicly accessible link to the relevant strain stocks. The database is based on ObjectStore (Object Oriented Database Management System) and is set out with a comprehensive interface.

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

(2) CARD R-DB

Takehiro YAMAKAWA, Hideki KATO, Naomi NAKAGATA, Kenichi YAMAMURA and Yukiko YAMAZAKI

CARD R-DB is the genetically engineered mice database of the Center for Animal Resources and Development (CARD), Kumamoto University, established as a core center for the production, cryopreservation and supply of mice.

The database system has been implemented with a direct online data submission and validation system. Each entry belongs to at least one group and one entry consists of strain-specific and/or gene specific information. Although the data type of most items is free text type, strain and gene names strictly follow internationally approved vocabulary. We have developed data submission and validation tools by which the user can easily submit their data through the Internet. The administrator can validate the information right after the submission and then automatically update the database on the same web screen. The database is managed through the relational database system (Oracle) with Java servlet applications. The CARD R-DB is accessible at http://cardb.cc.kumamoto-u.ac.jp/transgenic/.

(3) JMSR

Takehiro YAMAKAWA and Yukiko YAMAZAKI

JMSR (Japan Mouse Strain Resource) database provides a common gateway of mouse genetic resources from different organizations. The current database is compiled of information from four organizations including Kumamoto University, National Institute of Genetics, Tokushima University and the Japan Animal Strain Committee. Each data consists of minimal items such as (1) Strain name, (2) Strain group, (3) Stock type, (4) Genes, (5) Model disease name and/or application field and (6) URL address from where full information of the certain strain can be retrieved. We use the free PostgreSQL relational database in order to make the system freely available under an Apache open source license. JSMR expects new members to join in the near future. Data distribution in XML format with the associated document type definition (DTD) file is also an ongoing project. JMSR provides web access at http://www.shigen.nig.ac.jp/mouse/jmsr/.

(4) Oryzabase

Takehiro YAMAKAWA, Shingo UENO, Nori KURATA, Atsushi YOSHIMURA, Hikaru SATO, Yasuo NAGATO and Yukiko YAMAZAKI

Oryzabase represents an integrated rice science database, containing
(1) Rice Genetic Resources, (2) Mutant Phenotype Collection,(3)Gene Dictionary, (4) Genetic Maps, (5) Developmental stage and gene expression, (6) Basic Information and (7)Useful site information.

Current database compiles ca. 11000 strains maintained by 25 stock centers. Oryzabase also provides the "Core Collection of Wild Rice" to show representative wild rice accessions, which cover 9 different genomes (AA, BB, CC, BBCC, CCDD, EE, FF, GG, and HHJJ) and 18 different species. All collections have been characterized and presented with image data. Mutant phenotype collection is currently classified into 5 groups such as Coloration, Heterochrony, Reproductive organ, Seed, and Vegetative organ. Oryzabase also provides a web-based data submission system by which users can upload the mutant phenotype image files and submit them to the database.

The gene information of Oryzabase is based on the report of the Committee on Gene symbolization, Nomenclature and Linkage Groups.

Ontology is one powerful idea to specify a certain concept and useful especially to compare concepts common between different fields. Among many biological ontology projects recently established, the Gene Ontology (GO) Project is the most popular and wide spread because major databases such as Flybase, SGD, MGI, TAIR, WormBase, PomBase, RGD, EBI, TIGR have already joined. Gramene database, a comparative mapping resource for grains maintained by Cornell University, now proposes three rice ontologies, gene ontology (GO), traits ontology (TO) and plant ontology (PO) that defines the plant anatomy and/ or plant developmental stages. Coordinating the Gramene's efforts, all genes of the current dictionary were assigned to the above ontologies (GO/TO) in the Oryzabase. The assignment of GO for newly coming genes obtained from genomic sequences as well as PO assignment is an ongoing project in collaboration with rice researchers and other rice databases. To make this collaboration easier, a comprehensive viewer called "GOALL", that will give researchers an overview of the current GO and help them understand how ontology works and how it is useful, has been developed.

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

(5) KOMUGI

Takehiro YAMAKAWA, Takashi ENDO, Yasunari OGIHARA, Hitoshi TSUJIMOTO, Taihachi KAWAHARA, Tetsuro SASAKUMA and Yukiko YAMAZAKI

KOMUGI was originally established as a wheat genetic resource database by collecting strain stock information from several Japanese universities and grew to become an integrated wheat information database. The current version of KOMUGI database consists of (1) strain resource information, (2) EST and DNA resource information, (3) wheat gene catalogue, (4) composite maps, (5) wheat information service (WIS), and (6) DNA sequence collection of wheat (Triticeae, Avena, Secale and soon.). The gene catalogue is published every 5 years and only supplemental data is annually published. To make an up-to-date gene catalogue available through the Internet, we have developed a management system of a gene catalogue called "MacGene". Through this, by which the catalogue keeper can easily submit and/or update the data automatically. The KOMUGI database is available to the public through the Internet at http://www.shigen.nig.ac.jp/wheat/komugi/.

Publications

(1) Matsuoka, Y., Yamazaki, Y., Ogihara, Y. and Tsunewaki, K.: Whole chloroplast genome comparison of rice, maize, and wheat implications for chloroplast gene diversification and phylogeny of cereals. *Mol. Biol. Evol.* **19**, 2084-2091, 2002.

(2) Kitakami, H., Kanbara, T., Mori, Y., Kuroki, S. and Yamazaki, Y.: Modified PrefixSpan Method for Motif Discovery in Sequence Databases. *PRICAI 2002, LNAI 2417*, 482-491, 2002.

(3) Kanbara, T., Mori, Y., Kitakami, H., Kuroki, S. and Yamazaki, Y.: Discovering Motifs in Amino Acid Sequences using a Modified PrefixSpan Method. *Currents in Computational Molecular Biology* 96-97, 2002.

G-b. GENOME BIOLOGY LABORATORY

Genome Biology Laboratory, Center for Genetic Resource Information

(1) NEXTDB: The nematode expression pattern database

Tadasu SHIN-I, Ikuko SUGIURA, Masumi OBARA, Wakako SHIMIZU, Yuko SUZUKI, Jean THIERRY-MIEG[1], Danielle THIERRY-MIEG[1], Yutaka SUZUKI[2], Sumio SUGANO[2], Kazuko OISHI, Masako SANO, Hisayo NOMOTO, Shinobu HAGA, Satoko NISHIZAKA, Hiroko HAYASHI, Fumiko OHTA, Sachiko MIURA, Keiko HIRONO, Kyoko NAKATA, Hiroko UESUGI, Yasuko SUGIYAMA and Yuji KOHARA ([1]NCBI, NIH, [2]Institute of Medical Sciences, U. Tokyo)

We continued to update 'NEXTDB' that integrated all the information of ESTs, gene expression patterns and gene functions of *C.elegans* which were produced and analyzed in this laboratory. Images of whole mount *in situ* hybridization for mRNA were taken by microscopes equipped with DIC (differential interference contrast) apparatus, loaded to the Sun workstation to process and arranged properly in the database. Then, they were annotated with respect to developmental stages and expression patterns. Current version of NEXTDB contains EST information of 136,467 cDNA clones (5'-EST: 126,431, 3'-EST: 116,616) that have been classified to 11,903 unique cDNA groups, *in situ* images of 9,183cDNA groups and RNAi phenotype images of 295 groups. The cDNA clones are freely available for academic use. Thus far, the total 23,460 clones have been sent to the total 4,829 requests from 26 countries. NEXTDB is available at URL: http://nematode.lab.nig.ac.jp/. Many collaborations based on NEXTDB have been conducted (See Ref. 2, 3, 6, 9, 12)

(2) Analysis of protein distributions of maternal genes by antibody staining in *C.elegans*.

Yumiko UETA, Masahiro ITO, Hiroko OCHI, Chihiro HIJIKATA, Keiko HIRONO, Yoshinori UKAI[1], Midori SHINOHARA[1], Hisayoshi TORII[2], Emi SHIMIZU[2], Miki TAKENAKA[2], Yoshitaka IBA[3], Yoshikazu KUROSAWA[3] and Yuji KOHARA ([1]Institute for Antibodies, [2]MBL, [3]Fujita Health Univ.)

The accumulation of mRNA expression patterns of *C.elegans* that are being produced in this lab provides unique sets of genes that are expressed in specific developmental stages, cell-lineages, tissues and so on. The gene expression data led us systematic analyses to elucidate the molecular mechanisms governing the biological process. Since we are particularly interested in early embryogenesis, we have focused on a set of genes whose mRNA are maternally supplied and disappear by early stage of gastrulation or localize to specific cells in early embryos. We found 477 (about 8.8 %) such genes when the expression analysis of about 5000 genes was finished. To analyze the subset genes systematically with respect to the distribution of the protein products, we started to raise mouse antisera against the 477 maternal gene products using bacterially-expressed partial proteins. Until now, we obtained antisera against 282 gene products, and immuno-stained embryos and gonads using the antisera.

The spatial immunostaining patterns were as follows; nucleus (32%), cytoplasm (35%), plasma membrane (19%), nuclear membrane (10%) and P-granules (20%). The temporal staining patterns were as follows; the staining signals were first observed from the mitotic region in gonad (32%), oogenesis (12%) and from 2-cell stage (11%) . We confirmed the specificity of some of the antibodies by testing if the staining signal disappeared in RNAi-treated specimen.

(3) Roles of POS-1 and PIP-1 in the translational regulation of maternal mRNA in *C. elegans*

Norihito KISHIMOTO, Ken-ichi OGURA[1] and Yuji KOHARA ([1]Yokohama City U.)

Translatinal control of maternal mRNA is important for cell fate specification in early embryos of *C. elegans*. For example, GLP-1 protein, a Notch family receptor, is translated in anterior cells but not in posterior cells, although *glp-1* mRNA is present in all of the cells. Several regulatory regions in *glp-1* 3'UTR (untranslated region) have been identified as spatial control region (SCR) and temporal control region (TCR). We found that POS-1 specifically interacted with SCR and negatively regulated the translation of *glp-1* mRNA. A maternal factor POS-1 regulates cell fate determination of early embryos (Tabara *et al.*, Development 126 1-11 1999). POS-1 has

TIS11-type (CCCH) zinc-finger motifs, and is present in cytoplasm of posterior blastomeres, being a temporary component of P granules. PIP-1 (POS-1 interacting protein) has an RNP-type RNA-binding motif and is present in cytoplasm of oocytes and early embryos, also being a temporary component of P granules. We found that PIP-1 positively regulates maternal *glp-1* mRNA translation. Based on these findings, we proposed a combinatorial mechanism for the translation regulation by the two proteins, POS-1 and PIP-1.

To examine the roles of POS-1 and PIP-1 in translational control of maternal mRNA, we studied interactions between POS-1, PIP-1 and 3'UTRs of various maternal mRNAs by using the yeast tri-hybrid system. Other RNA-binding proteins —MEX-1, MEX-3, MEX- 6— were also studied. Futhermore, we mapped the protein binding region in POS-1, PIP-1 and another POS-1 interacting protein. Based on the results, we propose that the cell fate specification involves cascades of tranlational control in which POS-1 and PIP-1 appear to control poly(A) tail length and regulate translation of maternal mRNA. (See Ref. 14)

(4) Systematic identification of *cis*-regulatory elements of the genes that are expressed in specific nerve cells

Hiroshi KAGOSHIMA, Akiko KAMAMOTO and Yuji KOHARA

To understand the mechanism of transcriptional regulation, we have taken a simple strategy. We have been trying to extract consensus regulatory sequence of the genes which share spatial and temporal expression pattern in a certain cell.

There are a large number of GFP reporters that are known to show cell-specific expression in particular set of cells in *C. elegans*. One can accordingly expect that the same transcription machinery might regulate the genes which are expressed in a certain cell. If so, their regulatory core regions should share common sequences. To test this idea, we checked the expression patterns of a series of deletion promoter::GFP reporters in transgenic worms, and identified minimal enhancer "core" for the cell-specific expression.

We currently focus on head neuron-specific promoters. Thus far, we narrowed down 7 promoters for the expression in the thermosensing neuron AFD; for example, 30 bp sequence of *gcy-8* (guanylyl cyclase), 80 bp sequence of *nhr-38* (nuclear hormone receptor). Interestingly, within these sequences we found the binding consensus for the transcription factor OTX1, which has important roles for head development in fly and vertebrate. It has been revealed that ttx-1 (OTX1 homolog of *C. elegans*) regulates both *gcy-8* and *nhr-38* expression in AFD (Satterlee et al. 2001). Systematic promoter analysis also demonstrated that some deletion cause ectopic expression in additional neurons. This observation suggests negative regulatory mechanism seems to participate in the cell specification.

(5) Systematic RNAi Analysis of Genes Expressed in *C. elegans* Primordial Germ Cells, Z2 and Z3

Ichiro KAWASAKI, Junko KAJIWARA, Michiko SERIZAWA, Masahiro ITO and Yuji KOHARA

Germ lineage is unique and distinctive in several fundamental aspects from the other somatic lineages. In *C. elegans*, embryonic germline precursor cells (P_1 to P_4) are generated through 4-rounds of unequal divisions and set aside from somatic blastomeres early in embryogenesis. Although the embryonic germline cells inherit and accumulate many maternal mRNAs, they are transcriptionally inactive partly due to the negative regulation by *pie-1* function. The germ line-age becomes first transcriptionally active in morphogenesis-stage embryos after P_4 is divided into the two primordial germ cells, Z2 and Z3, when PIE-1 protein is already gone. After hatching, the germ lineage becomes one of the most active lineages in the organism. Z2 and Z3 divide throughout larval development, giving rise to ~2,000 germ cells in an adult hermaphrodite. Then they progressively undergo meiosis, spermatogenesis, and oogenesis to produce gametes that generate the next generation.

In this lab, a systematic whole mount *in situ* hybridization analysis has been performed against the non-redundant cDNA groups that had been identified and classified in our EST project. Through the analysis, thus far, about 9,000 cDNA groups (genes) have been annotated for their temporal and spatial expression patterns. By screening germline specific expression patterns, we have identified that as many as 700 genes are enriched or specifically expressed in Z2 and Z3 or in their immediate daughter cells either in morphogenesis-stage embryos or in L1, L2-stage

larvae. Assuming that some of those Z2, Z3-expressing genes play pivotal roles in activation of post-embryonic germline-developmental programs, we performed systematic RNAi analysis against those 700 genes using L1-soaking method. Among them, 115 genes showed strong sterile phenotypes. Furthermore, 34 of them (25 are novel) exhibited Glp (Germ Line Proliferation defective) phenotype. We have started analyzing these new Glp genes more closely to understand their functions and interplays at molecular levels.

(6) Effect of LIM binding factor LDB-1 on the function of T-box transcription factor TBX-9 in *C. elegans*

Yoshiki ANDACHI

tbx-9 is a member of the T-box family of transcription factors. *tbx-9* is expressed in a few precursor cells of intestine, body-wall muscle and hypodermis. A *tbx-9* deletion mutation causes disorganization in embryogenesis that predominantly occurs in posterior part of the body, though the penetrancy was very low.

To further elucidate the function of *tbx-9*, I searched for its binding by the yeast two-hybrid system, and found LDB-1 as one of the interacting proteins. *ldb-1* is a orthologue of vertebrate LIM domain binding factor *CLIM-2/NLI/Ldb1* and is expressed in all the blastomeres of early stage embryo. Two *Ldb1* monomers form a homodimer and each monomer binds to LIM transcription factor or *Pitx* transcription factor, so that they constitute a complex of transcription factors. Thus fur, relation between T-box factor and *Ldb1* factor has not been reported. Their interaction was confirmed by the in vitro pulldown assay. RNAi experiments with *ldb-1* ds-RNA caused 3-times higher penetrancy of the *tbx-9* phenotype in *tbx-9* mutant but no phenotype in the N2 wild type. These results suggest that *tbx-9* and *ldb-1* work for the same function. It is speculated that *tbx-9* transcription factor forms with another transcription factor through the mediation of *ldb-1* dimers to synergistically regulate the expression of their target genes.

(7) Computer Simulation of the Cellular Arrangement Using Physical Model in Early Cleavage of the Nematode *C. elegans*

Atsushi KAJITA, Masayuki YAMAMURA[1] and Yuji KOHARA ([1] Tokyo Institute of Technology)

The ultimate goal of bioinformatics is to reconstruct biological systems in the computer. Since biological systems have many levels, it is important to focus on an appropriate level. In our first application of computer modeling to the early development of the nematode *C. elegans*, we focus on the cellular arrangement in early embryos. This plays a very important role in cell fate determination by cell-cell interaction, and is regarded as a system, one level higher than the system of gene regulation within cells. It is largely restricted by physical conditions that seemed feasible to model by computer.

We constructed a computer model of the C. elegans embryo, currently up to the 4-cell stage, using a deformable and dividable triangulated network. The model is based solely on cellular-level dynamics. We found that the optimal ranges of three parameters that affect the elongation of dividing cells led, in computer simulations, to almost the same cellular arrangements as in real embryos. The nature of the model and the relationship with real embryos were examined.

In addition, we modeled new physical phenomena of cell division, cell rounding and stiffening; we then combined them with already modeled phenomena, contractile ring contraction and cell elongation. We investigated effectiveness of the new model on the cellular arrangement by computer simulations. We found that cell rounding and stiffening only during the period of cell division were effective to generate almost identical cellular arrangements to in real embryos. Since cells could be soft during the period between cell divisions, implementation of the new model resulted in cell shapes similar to real embryos. The nature of the model and its relationship to real embryos are discussed. (See Ref. 10, 11)

(8) Toward a four-dimensional database of development of the nematode *Ceanorhabditis elegans*

Masahiro ITO, Yohei MINAKUCHI, Michiko SERIZAWA and Yuji KOHARA

One of the most challenging targets in the post genomics era is computer simulation of development. *C. elegans* is thought to be the most suitable system for this purpose since the complete list of cell components has been identified and expression patterns and functions of all genes are being analyzed systematically. We wish to integrate such information and to create a four-dimensional (3D + time course) database of gene expression and development, based on which better computer modeling and simulation would be achieved.

First, we need a good framework for such 4D database. For this purpose, we made a computer graphics (CG) of early embryogenesis up to 86-cell stage based on the image data taken on 4D microscope. 4D microscope can record DIC (differential interference contrast) images of live embryos at different focal planes and along time course. First, we reconstructed 3D structure from a series of images of optically sliced section. Since it was very difficult for conventional image processing software to extract the cell contours from the DIC images, we adopted a manual procedure: we traced the contours of individual cells and nuclei by hand on thousands of original images of the 4D microscopic data; then, we reconstructed 3D structures by stacking them and making relationships among the contour data. Resulting 3D structures of embryos at various developmental stages were connected with each other by an interpolation method to generate a computer graphics. The CG provides various information of early embryogenesis such as individual cell volume, extent of cell-cell contact, and so on, which are difficult to measure on the original microscopic images. We are using the CG as a framework of our 4D database of *C.elegans* development.

Second, onto the CG framework, we planned to incorporate expression data of many genes at subcellular level. However, the assignment of 3D distribution of gene products (protein, mRNA) of embryos at various developmental stages is difficult and tedious. We need to automate this process. For this purpose, we developed a new system, named SPI, for superimposing fluorescent confocal microscopic data onto the computer graphics framework.

The scheme of this system is as follows: (1) acquirement of serial sections of fluorescent confocal images with triple colors, that is, DAPI for nuclei, Cy-3 for an internal marker and Cy-5 for the query gene product; (2) identification of several features of the stained embryos such as the contour, developmental stage and the position of the internal marker; (3) selection of computer graphics images of the corresponding stage for template matching; (4) superimposition of serial sections onto the computer graphics; (5) assignment of the position of the query gene products. The Snakes algorithm was used for the identification of embryo contour. Detection accuracy of embryo contour was 92.1% when it was applied to embryos of 2 to 28-cell stage. The accuracy of the method of developmental stage prediction was 74.6% for embryos at 2-12 cell stage. Since the accuracy for embryos at later stage was unsatisfactory due to experimental noise effect, we adopted manual judgment by scientist only the later stage embryos. Finally, our system chose optimal CG and performed the superposition and the assignment of the gene product distribution successfully. Initial 4D gene expression database containing 56 maternal genes was established. (Manuscript submitted)

(9) Automatic system for creation of cell shape model in *C. elegans* embryogenesis

Hideaki HIRAKI and Yuji KOHARA

C. elegans is an excellent model organism for the studies of development as its body is transparent and consists of a small number of cells. Cell to cell interactions play critical roles in early embryogenesis, therefore, it is very important to have information about the arrangements of cells, cell shapes and the contact among them. A polygon model of the early embryogenesis of a wild type embryo was built in our laboratory based on a time series of 3-dimensional images of differential interference contrast (DIC) microscopy. This type of computer model is very useful for comparative studies among mutant embryos, however, building of the model takes long time and much effort because it is difficult to extract the cell boundaries in DIC images automatically. Indeed, the outlines of the cells were traced by hand to build the current model. Therefore, in this work we are developing an automated system to build such a computer model of the cells in embryos based on a time series of confocal microscopic images of embryos whose plasma membrane is stained with a fluorescent dye. The system can fully automatically construct the computer models of 1-24 cell stages from an assured

seed sets of 24 cell stage. Using the overlapping information in the time series of the constructed cell shape models, the system can reconstruct virtually correct cell lineage. Now we are trying to process a real dataset of 24-200 cell stages, and are planning to apply this system to compare the cellular arrangements and the cell-to-cell contacts among mutant embryos and the embryos from other species closely related to *C. elegans*.

(10) Towards comparative genomics: Genomic/EST sequencing

Kazuko OISHI, Shinobu HAGA, Hisayo NOMOTO, Masako SANO, Satoko NISIZAKA, Fumiko OHTA, Hiroko HAYASHI, Sachiko MIURA, Tomomi MORISHITA, Tomoko ENDO, Motoyo TAMIYA, Noriko HASEGAWA, Masumi MIZUKOSHI, Etsuko YOKOYAMA, Junko MIYAMOTO, Shigeko IIYAMA, Tadasu SHIN-I, Shigeru SAITO, Kumiko KAWAGUCHI, Naoko SAKAMOTO, Yasuko SUGIYAMA and Yuji KOHARA

As a core facility of the group grant "Genome" (supported by MEXT and consisted of about 300 labs), we have established a DNA sequencing center. Currently our capacity is 8 million reads per year. Thus far, we have performed the followings in collaboration with the members of the group grant (in the parenthesis);

(a) EST: the nematode *Caenorhabditis elegans*, 180,864 reads (Yuji Kohara, NIG), Ascidian *Ciona intestinalis*, 684,096 reads (Nori Satoh, Kyoto U.), Ascidian *Ciona savignyo*, 7,300 reads (Nori Satoh, Kyoto U.), Ascidian *Halocynthia roretzi*, 73,986 reads (Kazuhiro Makabe, Kyoto U.), Hagfish *Eptatretus burgeri*, 62,208 reads (Masanori Kasahara, the Graduate U. for Advanced Studies), Medaka fish *Oryzias latipes* , 268,320 reads (Hiroyuki Takeda, U. Tokyo and Akira Kudo, Tokyo Inst. Tech.), ciclid fish *Haplochromis chilotes*, 50,112 reads (Norihiro Okada, Tokyo Insti. Tech.), frog *Xenopus laevies*, 146,496 reads (Makoto Motii, NIBB), slime mold *Dictyostelium discodeum*, 183,648 reads (Yoshimasa Tanaka, Tsukuba U.), read algae *Cyanidioschyzon merolae*, 76,704 reads (Tsuneyoshi Kuroiwa, U. Tokyo), moss *Physcomitrella patens subsp. patens*, 104,064 reads (Mitsuyasu Hasebe, NIBB), barley, 147,648 reads (Kazuhiro Satoh, Okayama U.), wheat *Triticim aestivum cv. Chinese Spring*, 276,096 reads (Yasunari Ogiwara, Yokohama City U.), Japanese morning glory, 30,720 reads (Eiji Nitasaka, Kyushu U.). High quality sequences were obtained with 70-90% of the reads (depending on the samples) and deposited in DDBJ. The analyzed cDNA clones are also distributed from this laboratory.

(b) Genome (BAC, fosmid clones): nematode *Pristionchus pacifficus* 2 BACs (Yuji Kohara, NIG and Ralf Sommer, MPI, Germany), fly *Drosophila sechelia* 2 cosmids (Toshiro Aigaki, Tokyo Metoropolitan U.), Acidian *Ciona intestinalis* 16 BACs (Nori Satoh, Kyoto U.), medaka fish *Oryzias latipes*, 5 BACs (HOX region) (Hiroshi Hori, Nagoya U.), Chimpanzee, 12 BACs (Naruya Saitou, NIG), Gorilla, 9 fosmids (Naruya Saitou , NIG), liverwort *Marchantia polymorpha L* 11 PACs (Kanji Ohyama, Kyoto U.).

(c) Genome (whole genome): read algae *Cyanidioschyzon merolae* 16Mb (Tsuneyoshi Kuroiwa, U. Tokyo), Ascidian *Ciona intestinalis* 160Mb (Nori Satoh, Kyoto U.)

The main target at the moment is the whole genome shotgun sequencing of Medaka fish *Oryzias latipes* (800Mb). (See Ref. 1, 4, 5, 7, 8, 15)

(11) The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins.

Dan ROKHSAR Lab[1], Yuji KOHARA Lab, Michael LEVINE Lab[2], Nori SATOH Lab[3] and Ascidian research community ([1]DOE Joint Genome Institute, USA, [2]UC Berkeley, USA, [3] Kyoto U.)

The first chordates appear in the fossil record at the time of the Cambrian explosion, nearly 550 million years ago. The modern ascidian tadpole represents a plausible approximation to these ancestral chordates. Here we present the draft genome sequence of the most thoroughly analyzed ascidian, *Ciona intestinalis*. The *Ciona* genome contains protein-coding genes, similar to the number in other invertebrates, but only half the gene number found in vertebrates. Vertebrate gene families are typically found in simplified form in Ciona, suggesting that ascidians contain the basic ancestral complement of genes involved in cell signaling and development. The ascidian genome has also acquired a number of lineage-specific innovations, including a group of genes engaged in cellulose metabolism that are related to those in bacteria and fungi.
http://genome.jgi-psf.org/ciona4/ciona4.home.html (See Ref. 13)

Publications

1. Kawashima, T., Kawashima, S., Kohara, Y., Kanehisa, M. and Makabe, K.W.: Update of MAGEST: Maboya Gene Expression patterns and Sequence Tags. Nucleic Acids Research, **30**, 119-120, 2002.

2. Morita, K., Flemming, A.J., Sugihara, Y., Mochii, M., Suzuki, Y., Yoshida, S., Wood, W.B., Kohara, Y., Leroi, A.M. and Ueno, N.: A *C. elegans* TGF-β, DBL-1, controls the expression of LON-1, a PR-related protein, that regulates polyploidization and body length. EMBO J. **21**, 1063-1073, 2002.

3. Mallo, G.V., C.L., Kurz, Couillault, C., Pujol, N., Granjeaud, S., Kohara, Y. and Jonathan J. Ewbank, J.J.: Inducible Antibacterial Defence System in *C.elegans*. Current Biology **12**, 1209-1214, 2002.

4. Fujiwara, S., Maeda, Y., Shin-i, T., Kohara, Y., Takatori, N., Satou, Y. and Satoh, N.: Gene expression profiles in *Ciona intestinalis* cleavage-stage embryos. Mechanisms of Development **112**, 115-127, 2002.

5. Ogasawara, M., Sasaki, A., Metoki, H., Shin-i, T., Kohara, Y., Satoh, N. and Satou, Y.: Gene expression profiles in young adult *Ciona intestinalis*. Dev. Genes Evol., **212**, 173-185, 2002.

6. Iwahashi, J., Kawasaki, I., Kohara, Y., Gengyo-Ando, K., Mitani, S., Ohshima, Y., Hamada, N., Hara, K., Kashiwagi, T. and Toyoda, T.: *Caenorhabditis elegans* reticulon interacts with RME-1 during embryogenesis. BBRC **293**, 698-704, 2002.

7. Satou, Y., Takatori, N., Fujiwara, S., Nishikata, T., Saiga, H., Kusakabe, T., Shin-i, T., Kohara, Y. and Satoh, N.: *Ciona intestinalis* cDNA projects: expressed sequence tag analyses and gene expression profiles during embryogenesis. Gene **287**, 83-96, 2002.

8. Kusakabe, T., Yoshida, R., Kawakami, I., Kusakabe, R., Mochizuki, Y., Yamada, L., Shin-i, T., Kohara, Y., Satoh, N., Tsuda, M. and Satou, Y.: Gene espression profiles in tadpole larvae of *Ciona intestinalis*. Developmental Biology **242**, 188-203, 2002.

9. Ichimiya, H., Hino, O., Kohara, Y. and Ishii, N.: VBP-1 is necessary for morphogenesis in *Caenorhabditis elegans*. Oncology Reports **10**, 293-295, (2003).

10. Kajita, A., Yamamura, M. and Kohara, Y.: Computer Simulation of the Cellular Arrangement in Early Cleavage of the Nematode *C. elegans*. Bioinformatics **19**, 704-716, 2003.

11. Kajita, A., Yamamura, M. and Kohara, Y.: Physical Modeling of the Cellular Arrangement in *C. elegans* Early Embryo: Effect of Rounding and Stiffening of the Cells. Proceedings of the 13th International Conference on Genome Informatics (GIW 2002), **13**, 224-232, 2002.

12. Matsuno, T., Ura, K., Sonoda, R., Kohara, Y., Uesugi, H., Arizono, K., Iguchi, T. and Tominaga, N.: Sensing of chemical substances using gene expression patterns in *Caenorhabditis elegans*. Sensors and Materials, **14**, 395-406, 2002.

13. Dehal, P., Satou, Y., Campbell, RK., Chapman, J., Degnan, B., De Tomaso, A., Davidson, B., Di Gregorio, A., Gelpke, M., Goodstein, D., Harafuji, N., Hastings, K., Ho, I., Hotta, K., Huang, W., Kawashima, T., Lemaire, P., Martinez, D., Meinertzhagen, IA., Necula, S., Nonaka, M., Putnam, N., Rash, S., Saiga, H., Satake, M., Terry, A., Yamada, L., Wang, H., Awazu, S., Azumi, K., Boore, J., Branno, M., Chin-bow, S., DeSantis, R., Doyle, S., Francino, P., Keys, D., Haga, S., Hayashi, H., Hino, K., Imai, KS., Inaba, K., Kano, S., Kobayashi, K., Kobayashi, M., Lee, B., Makabe, KW., Manohar, C., Matassi, G., Medina, M., Mochizuki, Y., Mount, S., Morishita, T., Miura, S., Nakayama, A., Nishizaka, S., Nomoto, H., Ohta, F., Oishi, K., Rigoutsos, I., Sano, M., Sasaki, A., Sasakura, Y., Shoguchi, E., Sin-I, T., Spagnuolo, A., Stainier, D., Suzuki, MM., Tassy, O., Takatori, N., Tokuoka, M., Yagi, K., Yoshizaki, F., Wada, S., Zhang, C., Hyatt, PD., Larimer, F., Detter, C., Doggett, N., Glavina, T., Hawkins, T., Richardson, P., Lucas, S., Kohara, Y., Levine, M., Satoh, N. and Rokhsar, D.: The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. Science **298**, 2157-2167, 2002.

14. Ogura, K., Kishimoto, N., Mitani, S., Gengyo-Ando, K. and Kohara, Y.: Translational control of maternal *glp-1* mRNA by POS-1 and its interacting protein SPN-4 in *Caenorhabditis elegans*. Development **130**, 2495-2503, 2003.

15. Ohta, N., Matsuzaki, M., Misumi, O., Miyagishima, S., Nozaki, H., Tanaka, K., Shin-i, T., Kohara, Y. and Kuroiwa, T.: Complete sequence and analysis of the plastid genome of the unicellular red alga *Cyanidioschyzon merolae*. DNA Research **10**, 67-77, 2003.

H. STRUCTURAL BIOLOGY CENTER

H-a. BIOLOGICAL MACROMOLECULES LABORATORY

(1) Thin Layer Illumination Microscopy as a Novel Technique for Single Molecule Fluorescence Imaging Inside Cells

Makio TOKUNAGA

Thin-layer illumination fluorescence microscopy, a novel microscopy for single molecule imaging, has been further developed. Specimens were illuminated with a thin-layered laser beam with the thickness of about 3 to 5 micrometers. This microscopy has a significant advantage of low background and high sensitivity, so it allowed us to visualize single fluorescent molecules inside cells up to the depth of about 10 μm from the glass-medium interface. Despite of imaging inside cells, clear images of single GFP molecules were obtained. The resolving power, i.e. the resolution distinguishing two point images, of 70 nm has been achieved using image analysis of images of single nuclear pores. The present microscopy is also useful for observation of living cell, which is sensitive to the illumination light, and for time-lapse observation over a long period.

(2) Single Molecule Imaging of Nucleocytoplasmic Transport in Cells and Quantitative Analysis of Interaction with Nuclear Pores

Makio TOKUNAGA and Naoko IMAMOTO[1,2] ([1]Gene Network Laboratoy, [2]Riken)

Single molecules of green fluorescent protein (GFP)-tagged importin beta, a carrier protein, and GFP-tagged cargo protein are clearly visualized on the nuclear envelope. These images are almost as clearly as those on cover glass surfaces. At higher concentrations of GFP-tagged proteins, images of nuclear pores were obtained, and image analysis resolved single nuclear pore images. Quantitative analysis of single-molecule and single-pore images enabled us to obtain kinetic parameters of the interactions with the NPC, i.e., the retention time, the number of bound molecules and the binding constant. It was found that there were two kinds of binding site of importin beta on NPC in the absence of cargo and Ran-GTP, weak binding sites and strong ones. The former gathers many molecules up to a maximum of approximately one hundred molecules per NPC. The latter vanished in the presence of cargo and in the absence of Ran-GTP, suggesting that accessibility to the strong binding site holds the key of the transport. Thus, single molecule imaging has opened a new way to obtain quantitative information on kinetics of molecular interactions in cells.

(3) p105: a novel protein component of mRNA-transporting granules in neuronal dendrites

Nobuyuki SHIINA, Kazumi SHINKURA and Makio TOKUNAGA

mRNA transport and subsequent local translation of the mRNAs in neuronal dendrites are widely believed to be essential for local regulation of synaptic efficacy. The dendritic mRNAs are transported in densely packed granules containing the mRNAs and ribosomes, called RNA granules.

p105 is a novel protein we originally identified as a centrosome- and microtubule-associated protein. Expression of p105 protein in cultured cells induces formation of cytoplasmic granules containing ribosomes, translation factors (EF-1 alpha) and mRNAs. In this year, we showed that p105 was expressed highly in neurons. These characteristics of p105 protein suggested that the p105-induced granules might be equivalent to the RNA granules in neurons. Immunostaining of rat hippocampal slices showed that p105 was associated with granules containing ribosomes and mRNAs in neuronal dendrites. Staufen, a protein for mRNA transport, and FMRP, a protein for translational regulation, were identified as the components of the granule. Microtubule- and actin-dependent motors, kinesin and myosin, were also the components of the granule. Furthermore, mRNAs associated with the granule were identified. Many of the mRNAs, such as CaM kinase II alpha, CREB and

BDNF mRNAs, were the key molecules for long-term synaptic plasticity. These lines of evidence indicated that p105 is a novel protein associated with mRNA-transporting granules which might be responsible for synaptic plasticity in neuronal dendrites.

(4) Single molecule measurement of intermolecular and intramolecular interactions using subpiconewton intermolecular force microscopy

Michio HIROSHIMA and Makio TOKUNAGA

Intermolecular force microscopy (IFM) is able to measure forces of inter- and intra-molecular interactions at single molecular level. This microscopy has achieved the force resolution of subpiconewton with ultrasensitive cantilevers. The cantilever position is controlled with nanometer accuracy by the feedback system using laser radiation pressure, which also reduces thermal fluctuation of the cantilever.

Using IFM, we have measured unzipping forces between single complementary base pairs of a synthesized DNA oligomer. To resolve unzipping forces of single base pairs, both the force sensitivity of piconewton and the position control with nanometer accuracy are required. Any current method of single molecule detection does not fulfill this requirement, whereas IFM does. The observed force vs. extension curves showed a series of peaks. Each peak is 10-15 pN in amplitude and approximately 1 nm in width. The peak amplitude varied periodically, and the periodic pattern was consistent with the periodic sequence of the DNA oligomer. These results suggested that the observed peaks corresponded to unzipping forces of single base pairs.

(5) Mechanism for passive force generation of invertebrate connectin revealed by single molecule measurement

Michio HIROSHIMA, Atushi FUKUZAWA[1], Koscak MARUYAMA[1], Sumiko KIMURA[1] and Makio TOKUNAGA
([1]Department of Biology, Chiba University)

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 (Fukuzawa et al., EMBO J. 20: 4826-4835 (2001). I-Connectin can be extended up to 3.5 μm upon stretch of giant sarcomeres. 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. In the present study, the force measurement of the single recombinant SEK polypeptide containing biotinylated BDTC and GST tags at the N and C termini, respectively, were performed by intermolecular force microscopy (IFM), a refined AFM system. The force versus extension curves were well fit to the wormlike chain (WLC) model and the obtained persistence length of 0.37 ± 0.01 nm ($n=11$) 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.

(6) A Novel *in vitro* Assay System of Nucleocytoplasmic Transport

Atsuhito OKONOGI, Michio HIROSHIMA, Nobuyuki SHIINA, Naoko IMAMOTO[1,2] and Makio TOKUNAGA
([1]Gene Network Laboratoy, [2]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. At first, agarose plates were modified with glutation. Then glutation-agarose plates 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. This new cell-free system has marked advantages: 1) Solutions in both sides, pseudo-cytoplasmic and pseudo-nucleoplasmic sides, can be replaced independently. 2) The shape of the nuclear envelope can be changed as one likes, for example, a vertical plane and a horizontal plane. 3) The system doesn't contain organelle or cellular structures except the nuclear envelope. Using the *in vitro* system together with single molecule techniques or fluorescence resonance energy transfer should provide an innovative and powerful tool to investigate molecular mechanisms of transport.

(7) Molecular Imaging of Translation Initiation Factors for Their Transport and the Local Translation in Neuronal Dendrites

Hiraku MIYAGI, Nobuyuki SHIINA and Makio TOKUNAGA

The learning and memory is widely believed to be associated with local plastic changes of synapses, which requires the translation of new proteins. Some of the proteins are believed to be translated locally at the activated dendritic synapses. We are going to visualize translation initiation factors to investigate how they are transported to synaptic sites in dendrites and when and where the translation is initiated locally in the dendritic synapses.

Translation initiation is triggered by binding of eukaryotic translation initiation factor (eIF) 4G with eIF4E. Imaging of the binding should reveal the timing and location of the protein translation after the synaptic activation. We constructed vectors for expression of eIF4E-GFP, CFP, YFP and eIF4G-GFP, CFP, YFP fusion proteins. The eIF4E and eIF4G fusion proteins were bound to endogenous eIF4G and eIF4E, respectively, in CHO cells, suggesting that the fusion proteins function *in vivo*. These constructs will be transfected into cultured neurons. We are planning to observe the dynamics of the fusion proteins using single molecule microscopy, and the binding of eIF4E and eIF4G fusion proteins by some methods such as FRET (Fluorescence Resonance Energy Transfer).

Publications

1. Hirakawa, Y., Suzutoh, M., Ohnishi, H., Shingaki, T., Erying, E.M., Tokunaga, M. and Masujima, T.: Analysis of the Nano-Kinetic Movement of a Single DNA by a Pin-Fiber Video Scope. Analytical Sciences **18**, 1293-1294, 2002.

2. Fukuzawa, A., Hiroshima, M., Maruyama, K., Yonezawa, N., Tokunaga, M. and Kimura, S.: Single Molecule Measurement of Elasticity of Serine-, Glutamate- and Lysine-Rich Repeats of Invertebrate Connectin Reveals That Its Elasticity Is Caused Entropically by Random Coil Structure. J. Mus. Res. Cell Mot., **23**, 449-453, 2002.

H-b. MOLECULAR BIOMECHANISM LABORATORY

(1) The promoter arrest of *E. coli* RNA polymerase and the effect of the Gre factors

Shouji YAGI[1], Motoki SUSA[1], Tomoko KUBORI[1], Hiroki NAGAI[1] and Nobuo SHIMAMOTO[1] (1Structural Biology Center, National Institute of Genetics and and Department of Genetics, School of Life Science, The Graduate University for Advanced Studies)

Transcription initiation has been conventionally assumed to be a sequential process that is composed of several steps, although additional steps can exist in parallel. The in vitro initiation from T7A1 promoter in particular apparently behaves in a manner that can be fully explained by the sequential pathway. However, the initiation from λP_RAL promoter has been shown to follow the branched pathway, where a part of the enzyme-promoter complex is arrested at the promoter, raising the question as to which mechanism is general. We found that a moribund complex, which is generated in the arrested branch, is formed at the T7A1 promoter especially in low salt condition, indicating that the initiation mechanism for this promoter is also branched. The results of DNA footprinting suggested that holoenzyme in moribund complex is dislocated on DNA from its position of productive complex. However, only a small fraction of the binary complex becomes arrested at the promoter and the interconversion between subspecies of binary complex is apparently more reversible than at the λP_RAL promoter, explaining the reason why the reaction pathway looks sequential. These findings suggest a generality of the branched pathway mechanism, which would resolve contradictory observations that have been reported for various promoters (Susa et al., 2002).

We next addressed the question whether the branched pathway has physiological significance in *E. coli*, because the both promoters we have examined are bacteriophage promoters. The clue is the effect of GreA and GreB on promoter arrest, which was previously found in the initiation at the λP_RAL promoter. These two proteins relieve the arrest at the promoter by increasing the conversion of moribund subspecies of binary complex into productive one. By using genomic DNA array we selected candidate genes whose expressions are decreased in a strain with

dusrupted *greA* and *greB*. By confirmation with Northerm blotting, I discovered that mRNA levels of at least 3 genes, *cspA*, *rpsA* and *atp*, are decreased by disruption of *greA* and *greB*. Their transcriptions were enhanced in a purified reconstitution system by adding GreA or GreB. Among them, *atp* operon was examined most detail. The formation of moribund complex at the *atp* promoter, the major promoter of *atp* operon, was confirmed by the most sensitive kinetic assay. The transcription showed littele pause and GreA and GreB were shown to enhance transcription through initiation efficiency by mitigating the promoter arrest. These findings suggest the pleiotropic existence of the branched pathway in *E. coli* cell. They also provide a new concept that the Gre factors are bonafide initiation factors, although they were discovered as elongation factors that mitigate elongation arrest in vitro.

(2) Inactivation of σ^{70} by oligomerization and identification of the role of its spacer region

Usha PADMANABHAN[1], Hiroki NAGAI[1], Taciana KASCIUKOVIC[2], Richard S. HAYWARD[2], Yumiko SATO[1] and Nobuo SHIMAMOTO[1] ([1]Nat. Inst. of Genet./Grad. School of Adv. Stud., [2]Inst. Of Cell and Molec. Biol., Edinburgh Univ., Scotland)

We have found that σ^{70}, the major σ factors of *E. coli*, forms aggregate in vivo and in vitro at a high temperature within physiological condition. The oligomeric σ^{70} has little transcriptional activity and the oligomerization in vitro showed a sharp temperature dependence. We have constructed a strain with a disrupted *rpoD* (σ^{70}), and plasmid born σ^{70} supports its growth. We have observed a positive correlation between the intracellular concentration of σ^{70} and upper limit of growing temperature. This raises a possibility that σ^{70} is a molecular thermometer.

The major σ factors of proteobacteria mostly have a big spacer region between the conserved regions 1 and 2 which is not conserved in eubacteria. In *E. coli* σ^{70} this region has extensive acidic patches which may be concerned with the property of σ^{70} to readily interact with nonspecific and specific surfaces. An *rpoD*-disrupted strain was constructed to test the viability of strains expressing plasmid-borne mutated *rpoD* or another sigma factor. *E. coli rpoS* and *M. tubeculosis sigA* failed to complement the disruption, and we are now testing *B. subtilis sigA* and others. The mutant σ^{70} lacking the spacer region of aa130-374 complemented the disruption at 30 and 25°C, proving that the region is not essential for growth at low temperature. At all tested temperature this protein predominantly exists as aggregate which are in equilibrium with a small fraction of monomer. Therefore, the role of the spacer region is the maintenance of active monomeric form.

(3) Existence of tracking a groove of DNA by RNA polymerase during its sliding

Kumiko SAKATA-SOGAWA and Nobuo SHIMAMOTO
Structural Biology Center, National Institute of Genetics, Yata 1111 Mishima, Shizuoka 411-8540, Japan.

As the start of transcription of a gene, most RNA polymerase molecules first bind to sites other than promoters where the enzyme initiates transcription. They then reach promoters by sliding along DNA. Sliding of proteins along DNA is one-dimensional Brownian motion, the existence of which has been proved for *E. coli* RNA polymerase and several other DNA binding proteins by single-molecule dynamics as well as kinetics. There are two possible modes of movement during sliding. In the first mode, a protein molecule slides along the axis of DNA helix and thus contacts to one side of DNA. In the second, a protein molecule tracks a groove of DNA and makes a helical motion around the axis. We addressed the question whether or not RNA polymerase molecule track a groove of DNA during sliding.

We first developed a devise for detecting the rotational movement of a DNA molecule around its helical axis (Shimamoto, 2002). A nonspecific DNA is fixed to a 900 nm polystyrene bead coated with streptavidin by several biotin residues with no rotational freedom. A 90 nm fluorescent particle is covalently attached to the polystyrene bead to visualize the rotational motion of the DNA-fixed beads under a fluorescent microscope. The polystyrene bead is held with optical tweezers at a distance from the surface so that it can rotate almost freely. We can examine the existence of the groove-tracking with this devise. When a DNA molecule fixed to a bead binds to an RNA polymerase molecules moving in a direction, its sliding with a groove-tracking should generate a rotation in one direction of DNA molecule that can be detected with the devise. In this configuration, we moved RNA

polymerase relative to the DNA by shifting the slide continuously in one direction. If RNA polymerase tracks a groove, it should rotate DNA and the rotation can be detected as the movement of the fluorescent bead.

We have observed rotational motions of the beads that could be caused by groove tracking of RNA polymerase in a small number of cases. Such motions were observed only when the distances between beads and the surface are smaller than the length of DNA. However, the number of such observation is not enough to conclude groove tracking. A short lifetime of sliding complex, about 1 sec in solution, complicates the estimation of the frequency of rotational motions. Further quantitative determination proved the existence of groove tracking.

(4) Applicability of thermodynamics to equilibria in biology

Nobuo SHIMAMOTO and Jun-ichi TOMIZAWA

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.

Publications

1. Shimamoto, N.: Movements of RNA polymerase along DNA. In Methods Enzymology (San Diego, Elsevier Science), pp. in press, 2002.

2. Susa, M., Sen, R. and Shimamoto, N.: Generality of the Branched Pathway in Transcription Initiation by E. coli RNA Polymerase., J Biol Chem **277**, 15407-15412, 2002.

Meetings

1. Shimamoto, N.: "What are necessary for simulating biological events?" Temporal order of biological events: Orders and disorders." Kitazato University, January 17, 2002.

2. Shimamoto, N.: "The principles of nano-world: A pitfall in nanobiology and nano-technology." Nano-technology: A wonder, KAST Science Seminar, Kawasaki, February 28, 2002.

3. Shimamoto, N.: "The principles of biological molecular machines that are different from those of the conventional machines: A pitfall in nanobiology." Towards the boundary of Material Science and Biological Science, Okazaki, March 4, 2002.

4. Sakata-Sogawa, K. and Shimamoto, N.: "Single-molecule detection of groove-tracking of DNA by RNA polymerase during sliding." The 7th Asian Conference on Transcription, Kuala Lunpur, Malaysia, July 23-27, 2002.

5. Ampaabeng, K. G., Padmanabhan, U. and Shimamoto, N.: "Monitoring the state of sigma-70 in vivo." The 7th Asian Conference on Transcription, Kuala Lunpur, Malaysia, July 23-27, 2002.

6. Shimamoto, N.: "Effect of being molecules uncovered form single-molecule based observations." The summer school for young biophysicists, Kobe, July 29, 2002.

7. Shimamoto, N. and Sakata-Sogawa, K.: "Single-molecule detection of groove-tracking of DNA by RNA polymerase during sliding." The annual meeting of Bio-imaging Society of Japan, Nagoya, November 1, 2002.

8. Sakata-Sogawa, K. and Shimamoto, N.: "Single-molecule detection of groove-tracking of DNA by RNA polymerase during sliding." The annual meeting of Biophysical Society of Japan, Nagoya, November 3, 2002.

9. Shimamoto, N. and Miyawaki, A.: Workshop "A new biology with new probe techniques." The annual meeting of Molecular Biology Society of Japan, Yokohama, December 13, 2003.

10. Sakata-Sogawa, K. and Shimamoto, N.: "Detection of tracking a DNA groove by RNA polymerase during sliding using DNA single molecule manipulation." The annual meeting of Molecular Biology Society of Japan, Yokohama, December 13, 2003.

H-c. MULTICELLULAR ORGANIZATION LABORATORY

(1) Analysis of synthetic dauer-constitutive mutants in the nematode *Caenorhabditis elegans*

Kiyotaka OHKURA, Tomoko YABE, Akiko KAMAMOTO, 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 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 formation as the output. We found that mutations in more than 50 known genes show synthetic dauer-constitutive (Sdf-c) 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 Sdf-c phenotype and the pattern of suppression of the Sdf-c 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 isolated and mapped 44 mutations that show the Sdf-c 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 mapped 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. Some of the mutant genes have been cloned. *sdf-13* and *sdf-9* genes encode a homolog of the mammalian Tbx2 transcription factor and a tyrosine phosphatase-like molecule, respectively.

In the year 2002, the following results were obtained.

(a) The dauer larva of *sdf-9* mutants was different from the normal dauer larva but resembled the dauer-like larva of *daf-9* and *daf-12* dauer-constitutive mutants. Like these mutants, the dauer-constitutive phenotypes of *sdf-9* mutants were greatly enhanced by cholesterol deprivation. Epistasis analyses, together with the relationship between *sdf-9* mutations and *daf-9* expression, suggested that SDF-9 increases the activity of DAF-9 or helps the execution of the DAF-9 function. SDF-9 was expressed in two head cells in which DAF-9 is expressed. By their position and by genetic mosaic experiments, we identified these cells as XXXL/R cells, which were known as embryonic hypodermal cells and whose function at later stages was unknown. Killing of the *sdf-9*-expressing cells in the wild type first-stage larva induced formation of the dauer-like larva. Since this study on SDF-9 and former studies on DAF-9 showed that the functions of these proteins are related to steroids, XXXL/R cells seem to play a key role in the metabolism or function of a steroid hormone(s) that acts in dauer regulation.

(b) We cloned *sdf-14* gene and found that it was the same as *mrp-1* gene, which was formerly identified by its homology to multidrug resistance protein genes in other organisms. Thus, we found a function of *mrp-1* concerning dauer larva formation. An *sdf-14*::GFP fusion gene that can rescue *sdf-14* mutant phenotypes was expressed in neurons, pharynx and intestine. The expression of *sdf-14* gene in each of them by extrinsic promoters resulted in partial suppression of the Sdf phenotype, while expression in two of the three tissues/organs resulted in more efficient suppression. It seems that expression in all the three parts contributes to the wild type phenotype. Since there are three isoforms of SDF-14 made by alternative splicing, we are now investigating the difference between these isoforms in function and expression.

(2) 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(ut236)* mutant showed much weaker tendency to cross the Cu^{2+} barrier when migrating toward attractive odorants than the wild type, although that the *hen-1(ut236)* mutant has 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. 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 homolog) 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(ut236)* mutant also shows 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(ut236)* mutant showed 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(ut236)* mutant shows normal thermotaxis under a well-fed condition, it does 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(ut236)* animals can sense starvation, but we could not find any abnormality in the behavior after simple starvation. These results indicate that the *hen-1(ut236)* shows 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.

We also studied on another mutant *ut235*. This mutant behaved essentially normal in many other responses to starvation, such as the change of locomotion speed on food by starvation. Interestingly, the double mutant *ut235;hen-1* showed a preference to avoidance of copper ion, regardless of starvation. We are now trying to clone the gene for this mutation by positional cloning.

In 2002, we isolated and analyzed the deletion allele of *hen-1, tm501*. The *hen-1(tm501)* was isolated by

screening pools of TMP/UV mutagenized worms using PCR (collaboration with Dr. Gengyo-Ando and Dr. Mitani, Tokyo Women's University of Medicine). It has a small deletion within the *hen-1* gene and seems a null allele. The behavioral analyses of the *hen-1(tm501)* revealed that its phenotypes are almost indistinguishable from those of *hen-1(ut236)* in the sensory integration and in the behavioral plasticity.

Molecular genetic analyses of the HEN-1 suggest that the HEN-1 functions as a neuronal modulator for sensory integration and learning. To elucidate the molecular mechanisms of this neuromodulation, we start 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 the HEN-1 specifically binds a subpopulation of the primary culture cells in *C. elegans*. By using this system, we are trying to expression cloning of the HEN-1 receptor from an expression cDNA library of the primary culture cells of *C.elegans*.

(3) *C. elegans* mutants in the associative learning of odorants and food

Ichiro TORAYAMA, Takeshi ISHIHARA and Isao KATSURA

The nematode *C. elegans* provides a good system for the study of associative learning. However, the mechanism of this learning looks different from that of classical conditioning, because (a) the unconditional stimulus is usually limited to food or starvation, and because (b) the learning is efficient, if the conditional stimulus is presented at the same time as but not before the unconditional stimulus. To elucidate the molecular mechanism of associative learning in *C. elegans* and food/starvation signaling in such learning, we are isolating and characterizing mutants that show abnormality in the associative learning of butanone and food/stravation. It is known that butanone attracts wild type worms 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 decreased the efficiency of chemotaxis after conditioning with food and butanone, while others are attracted efficiently by butanone only after conditioning with food and butanone. We are mapping *ut305* and *ut306*, which belong to the former category. Interestingly, *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* was mapped in the righthand part of the linkage group X, while *ut306* near the center of the linkage group V. We are narrowing down the regions of the mutations and trying to rescue the mutant phenotypes by microinjection of genomic DNA fragments from these regions.

(4) Fluoride-resistant mutants of the nematode *Caenorhabditis elegans*

Akane OISHI, Minoru KAWAKAMI[1], Akiko KAMAMOTO, Takeshi ISHIHARA and Isao KATSURA ([1]Gothenburg University)

Fluoride-resistant *(flr)* mutations of *C. elegans* are recessive and grouped into two categories: class 1 mutations, which map in *flr-1*, *flr-3* and *flr-4*, and class 2 mutations, which map in *flr-2* and *flr-5* (Katsura, I. *et al.*: Genetics **136**, 145-154, 1994). Class 1 *flr* mutations show diverse phenotypes: slow growth, short defecation cycle periods, frequent skip of the expulsion step of defecation, and synthetic abnormality in dauer formation, besides strong resistance to fluoride ion. 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-terminal side. 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 *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. We therefore think that class 1 *flr* genes constitute a regulatory system that acts in the intestine and that controls many food-related functions. Class 2 *flr* mutations, which confer weak resistance to fluoride ion, suppress the slow growth and dauer formation abnormality but not the defecation abnormality and strong fluoride-

resistance of class 1 *flr* mutations. Hence, it seems that the regulatory pathway bifurcates after class 1 genes, and that class 2 genes control only one of the two branches. Of the class 2 genes, *flr-2* encodes a secretory protein belonging to the gremlin/DAN/cerberus family. On the basis of these results, we have made a model in which a signal controlled by class 1 genes is transmitted from the intestine and represses the action of class 2 gene products in the head nervous system.

In the year 2002, we studied expression of a functional *flr-2*::GFP fusion gene, using an improved GFP. The expression was detected in many pharyngeal neurons and some head and tail neurons. This method was more sensitive than the antibody staining, which was carried out a few years ago. We started investigating whether FLR-2 binds to a subpopulation of *C. elegans* primary culture cells.

Publications

1. Ishihara, T., Iino, Y., Mohri, A., Mori, I., Gengyo-Ando, K., Mitani, S. and Katsura, I.: HEN-1, a novel secretory protein with an LDL receptor motif, regulates sensory integration and learning in *C. elegans*. Cell **109**, 639-649, 2002.

2. Kuhara, A., Inada, H., Katsura, I. and Mori, I.: Negative regulation and gain control of sensory neurons by the *C. elegans* calcineurin TAX-6. Neuron **33**, 751-763, 2002.

H-d. BIOMOLECULAR STRUCTURE LABORATORY

(1) Crystallographic Study of F1-ATPase

Yasuo SHIRAKIHARA, Aya SHIRATORI and Chikako SHIRAKIHARA

F1-ATPase is the catalytic sector in ATP synthase that is responsible for ATP production in living cells. F1 has molecular mass of 380,000 dalton and a subunit composition of $\alpha 3\beta 3\gamma\delta\varepsilon$. 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. We have recently moved to the structural study of the $\alpha 3\beta 3\gamma\delta\varepsilon$ sub-assembly of thermophilic F1, after solution of the structure of the $\alpha 3\beta 3$ sub-assembly and committed attempts to crystallize the $\alpha 3\beta 3\gamma$ sub-assembly. Employing the $\alpha 3\beta 3\gamma\varepsilon$ sub-assembly that show catalytic properties very similar to those of F1, we aim to detect structural changes caused by different nucleotide occupancy, which should provide with structural basis for understanding the rotational catalysis mechanism.

The most recent structure we have obtained is that of the $\alpha 3\beta 3\gamma\varepsilon$ sub-assembly with Mg-free ADP bound to only one β-subunit. The structure is unique, as this is the first F1 structure that has one nucleotide-filled catalytic site; all the previous structures have two or more nucleotide-filled catalytic sites. Interpretation of the structure was not straightforward due to lack of bound Mg ion; there has not been much biochemical evidence on the Mg-free state. However, based on such sparse evidence collected together, we now believe that the structure represents a weakly inhibited state of F1 with the $\alpha 3\beta 3\gamma$ portion in a state not far from an active state but with the ε-subunit in an inhibitory state. One of the features of the structure is that this structure lacks the most active β-subunit conformation that is characterized as having the tightly closed catalytic interface and has been observed in all the previous F1 crystal structures with two or three nucleotide-bound β-subunits. The feature may explain the very slow catalysis in the uni-site mode (defined as having one nucleotide-bound catalytic site). This structural difference in the $\alpha 3\beta 3$ part appears to be coupled to the substantial structural differences in the stalk, including penetration of the C-terminal helices of the ε-subunit into the $\alpha 3\beta 3$ cavity. The new conformation of the C-terminal helices of the ε-subunit explains well the weak ε-inhibition in the thermophilic F1.

We have further examined the preparation method for the $\alpha 3\beta 3\gamma\varepsilon$ sub-assembly, because the recent preparations (over 1 year period) had refused to crystallize, especially in presence of Mg ion. Combination of a controlled heat treatment with chromatography using hydrophobic and ion-exchange columns produced a preparation that restored the original crystallization capability. Neither an additional second heat treatment nor gel-filtration chromatography gave significant benefits. Using crystals from our new preparation, further cooling experiments were performed. The previous difficulty in cooling the crystals grown in presence of Mg has gone, partly because of benefits of using the better preparation and partly because of our new cryo conditions. Further structural study is in progress.

These structure studies were done in collaboration with 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, Aya SHIRATORI, Hisako INOUE and Megumi SEKI

More than 160 transcription factors have been identified in *E.coli*, and they control transcription of their target gene(s) in either positive or negative manners, by binding to both their specific DNA sequence and the transcription apparatus. Sixty-five such transcription factors have been purified in one step Ni-NTA column chromatography using carefully constructed over-expression systems in Ishihama laboratory. In Niki laboratory, a number of novel proteins, that are judged to play roles in the genome and plasmid partitioning from their localization patterns, have been prepared using similarly constructed over-expression systems. Setting these transcription and genome partitioning factors as targets for structural study, we have started comprehensive structure determination in the Protein 3000 project (the sub-field of 'transcription and translation').

We have developed an efficient and reliable method for screening crystallization conditions against the large number of the provided proteins, based on our long-time experiences in protein crystallization. Using the system, 34 proteins have been screened so far (30 transcription factors: AhpC, ArcA, AcrR, CueR, DeoR, IlvY, KdpE, SdiA, SoxR, SoxS, EmrR, InfB, LeuO, Mlc, MntR, RscB, RstA,Thil,TufB, AtoC,BaeR, CheB, QseB, YehT, YfhA, CpxR, DcuR, NarP, EvgA, RcsB and 4 genome & plasmid partitioning factors, RecR,PspA,YcbW,YgfE). In spite of an initial success of finding small crystals of ArcA and SdiA, most of the proteins did not crystallize possibly due to the conformational flexibility of the transcription factors that commonly have the two-domain architecture (DNA binding and 'receiver' domains). Crystals of selenomethionyl derivative of ArcA and heavy atom derivative (Hg,Pt,Ir,Os) of SdiA were prepared for data collection at PF for MAD (Multi-wavelength Anomalous Dispersion) analysis. However all the crystals gave poor diffraction patterns with the best result of 8 Å resolution data for the native SdiA crystals. We plan to crystallize the transcription factors complexed with the cognate DNA, and their constituent domains.

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) Crystallographic Study of the Transcription Activator, PhoB, complexed with cognate DNA

Kazuyasu SHINDOH, Katsumi MAENAKA and Yasuo SHIRAKIHARA

PhoB Protein is a positive transcriptional activator to regulate the genes in the phosphate (*pho*) regulon of *E. coli* that are induced by phosphate starvation. PhoB acts by binding to the *pho* box in the promoter region, which is the consensus sequence shared by the regulatory regions of the genes in the regulon. The N-terminal region of PhoB is a regulatory domain, whereas the C-terminal region (PhoBC) has a DNA binding activity to the *pho* box specifically. In order to resolve the detailed interaction between PhoB and *pho* box DNA, we have made the crystallographic study of PhoBC complexed with the *pho* box DNA.

We have got the crystals of PhoBC complexed with 16mer pho box DNA oligomer (ACAGCTGTCATAAATC) in the condition of 0.1 M Ammonium Acetate, 0.01 M Mg Acetate, 0.025 M Sodium Cacodylate pH 6.5, 15% w/v Polyethylene Glycol 8000, at 25°C. The crystals have the P1 space group with the cell dimension, a=37.99 Å, b= 48.98Å, c=70.30Å, α=108.57°, β=120°, γ=104.37° (2 molecule per asymmetric unit). The diffraction data was collected at Photon Factory and Spring8 (around 2.8 Å resolution). We have already solved the crystal structure of free PhoBC by the MAD data from the selenomethionyl derivative of PhoBC. Based on the model of free PhoBC, a full structure determination by molecular replacement is in progress.

This work has been done in collaboration with Hideyasu Okamura, Yoshifumi Nishimura, Yokohama City University and Kozo Makino, Research Institute for Microbial Diseases, Osaka University.

(4) Structural and functional studies of immunoglobulin (Ig)-like receptors

Katsumi MAENAKA, Tsukasa SHIMOJIMA, Kimie AMANO, Tomoko MOTOHASHI and Yasuo SHIRAKIHARA

The immunoglobulin(Ig) superfamily domains are very common among the cell surface receptors in immune system, and so it is important to clarify the molecular recognition of Ig-like cell surface receptors towards various ligands. We are studying Ig-like transcript (ILTs) receptor family.

ILTs are expressed mainly on myelomonocytic cells and some of these can recognize MHCs on target cells to mediate an inhibitory or activating signals. ILT2 and ILT4 studied here can bind to various MHC class I molecules, shown by cellular based assays, but the detail interactions in protein level are not clear. We produced all recombinant proteins (ILT2, ILT4 and MHC ligands) by *E.coli* expression and refolding system. Surface Plasmon resonance (SPR) assays clearly showed that all receptors bind a wide range of MHC class I (e.g. HLA-A, -B, -C, and G) with low affinities (K_D ($10^{-4} \sim 10^{-6}$M) and fast kinetics, which is similar to other receptors involved in cell-cell interactions. Unexpectedly ILT2 and ILT4 can bind to nonclassical MHC, HLA-G most preferentially in comparison to other MHC ligands. Since HLA-G is expressed on unique tissue, placenta, thymus and some tumours, this result may suggest a predominant role of ILTs on regulation of immune cells in these special areas. We have also done the thermodynamics for these interactions.

This work has been done in collaboration with Anton van der Merwe, Yvonne Jones, (University of Oxford), and Izumi Kumagai, Kouhei Tsumoto, Mitsunori Shiroishi (Tohoku University).

(5) Structural analysis of glutaminase from *Micrococcus luteus K-3*

Yasuo SHIRAKIHARA and Aya SHIRATORI

Glutaminase (EC 3.5.1.2) catalyzes the hydrolytic degradation of L-glutamine to L-glutamic acid. The salt-tolerant glutaminase from marine *Micrococcus luteus* K-3 has an unusual property that its maximum activity is observed at about 2M salt, and also has industrial relevance that this enzyme may be more efficient than the conventional ones in the soy sauce fermentation that is carried out in the high-salt environments.

After analysis of MAD data using the Se-Met derivatives for the intact form (2.6Å resolution) and the truncated form (2.4Å resolution) of the enzyme done last year, model building of the truncated form was done and now refinement has started. We plan to utilize the model in determining the structure of the intact enzyme by molecular replacement.

The structural study was done in collaboration with Kazuaki Yoshimune and Mitsuaki Moriguchi at Oita University, and Mamoru Wakayama at Ritsumeikan University.

Publications

Shindoh, K., Maenaka, K., Akiba, T., Okamura, H., Nisimiura, Y., Makino, K. and Shirakihara, Y.: Crystallization and preliminary X-ray diffraction studies on the DNA-binding domain of the transciptional activator protein PhoB from *Escherichia coli*:, Acta Crystallogr D **58**, 1862-1864, 2002.

I. CENTER FOR INFORMATION BIOLOGY AND DNA DATA BANK OF JAPAN

I-a. LABORATORY FOR DNA DATA ANALYSIS

(1) Genetic variation of Japanese gamecock and its implication to the cultural background of cockfighting.

Tomoyoshi KOMIYAMA, Kazuho IKEO and Takashi GOJOBORI

The tradition of cockfighting is widespread throughout the world. Chickens, including gamecocks were domesticated from the Red Jungle fowl (*Gallus gallus gallus*) in Southeast Asia at an early stage of human prehistory. Originally cockfighting may have been held for the purpose of religious ceremony or entertainment. There is no doubt that the gamecock has evolved together with the human culture of cockfighting a long time ago. In Japan, there is a group of gamecock called the "Shamo" which are used specifically for cockfighting. From archive study, we found that gamecocks were already known at least from approximately 1,100 AD. However, the geographic distribution of cockfighting and the influx route of gamecocks into Japan are totally unclear. The molecular evolutionary study of gamecocks is obviously to gain useful understanding of the process of distribution of the culture itself. Through the study of gamecocks, we aim to throw new light on the evolution of human culture. For the purpose of this study, we collected blood samples of gamecocks from 11 different prefectures in Japan. A phylogenetic tree was constructed using a total of 46 mt DNA (D-loop, 1100bp) sequences including the data from the International Nuclear DNA database (DDBJ/EMBL/GenBank). The phylogenetic analysis showed that the Japanese Shamos were separated into two different groups with high bootstrap value: one group contained only individuals from the island of Okinawa and the other group is composed of samples from Kyushu and mainland Japan. We collected the three morphological traits (height, weight, shank length) and the molecular samples of these various Shamo. The results of the morphological studies were consistent with our phylogenetic analysis. In particular, the Okinawa group showed bigger values of the height, weight and shank length than the other Shamo groups. This observation suggests that Japanese Shamos might be of two independent origins. It is also known that different names are used for Shamo in Japan. In Okinawa Island, for example, "taucii" is used as a name pronouciation of which is quite similar to Chinese. In the other part of Japan, other word is used for Shamo, and it is very similar to the old name meaning "Thailand." These observations also supports the theory of separate origins of Japanese Shamos; the origins of China and Southeast Asia.

(2) EST analysis of a normalized library derived from 5-day regenerating planarian.

Jung Shan HWANG, Katsuhiko MINETA, Kazuho IKEO and Takashi GOJOBORI

Planarian has been widely known for its regenerative power. An adult planarian can be cut into 200-300 pieces and each piece would eventually regenerate into an individual. Approximately 10-14 days are required for the whole regenerating process. Although planarian regeneration had been studied for more than 200 years, no researcher has a clue of what determines the developmental plasticity in planarian and how did it arise during the evolution. At the current stage, one possible approach to be used is to disclose all candidate molecules that involve in the process of regeneration and then make a comparison to the developmental/regenerative process of other organisms.

A normalized cDNA library was constructed from 5-day regenerating body parts which were cut non-specifically at different body levels. Approximately 5000 clones were isolated and the number of non-redundant clones was reduced to its one third. By using the non-redundant ESTs, we identified genes that participate in the regenerative process by comparing them with homologous genes from the databank. From the studies, we (1) categorized clones that are related to planarian regeneration into several classes, (2) identified clones which show significant

similarity with those related to the development of higher organisms, and (3) examined the expression patterns of clones by whole mount in situ hybridization and some clones potentially can be used as molecular markers. A 5-day regeneration database of planarian will be set up in the near future.

(3) Analysis of evolutionary relationships between interacting proteins in *S. cerevisiae*

Takashi MAKINO, Kazuho IKEO, Yoshiyuki SUZUKI and Takashi GOJOBORI (National Institute of Genetics, Laboratory for DNA Data Analysis)

Recently, large scale approaches have predicted many new protein interactions in *S. cerevisiae*. However, there are not protein interactions datasets in the other organisms. Therefore we analyzed evolutionary relationships between interaction of proteins by comparing the proteins of *S. pombe* with those of *S. cerevisiae*. Core, *S. cerevisiae* interactions in PPI database DIP that are assessed of their accuracy, is used for this analysis. Each protein included in Core, has been subdivided into functional categories using a catalogue of known and predicted protein functions at database MIPS. To investigate the distribution of interactions, we compared *S. pombe* with *S. cerevisiae* in amino acid sequences and estimated the evolutionary distances between each homologues.

The results indicate that the proteins of related functions interact with each other that tend to be exposed to similar selection.

(4) Estimation of loss of functions in metabolic pathways by comparison of 94 species

Tsuyoshi TANAKA, Kazuho IKEO and Takashi GOJOBORI

There are two different mechanisms, which are the gain of functions and the loss of functions, for the evolution of metabolic pathways. However, the reports related to the loss of functions are few compared with the gain of functions, the gain of function is not enough to explain the evolution of metabolic pathways. Then, to evaluate more detail effect on the loss of functions in metabolic pathways, we estimated the number of the loss of functions during the evolution. We used 94 species and 1507 types of Enzymatic Reactions (ERs) for estimating the number of the loss of functions in metabolic pathways. Each ER is corresponding to a particular EC number. To evaluate the loss of functions by comparing the existence of ERs among species, we examined how many ERs existed in each species and how many ERs species shared. As a result, the number of ERs in each species was ranging from 101 to 744. Moreover, only 5 out of total ERs were commonly observed in all 94 species and about 80% of total ERs existed only in a half of them. We also estimated the number of ERs which existed in ancestral species by using phylogenetic trees and one assumption. We assumed that an emergence of functions of an ER was only once during evolution of 94 species. As a result, the number of ERs in ancestral species was speculated as more than 377. Moreover, when we summed up the number of loss of functions of ERs during the evolution of 94 species, there were 10792 times in 1042 kinds of ERs. These results suggested that the ERs in metabolic pathways changed among species and the loss of functions in ERs has occurred many times during the evolution of metabolic pathways.

Hisakazu IWAMA and Takashi GOJOBORI

Center for Information Biology and DNA Data Bank of Japan, National Institute of Genetics, Mishima-shi

Neurotransmitter receptors (neuroreceptors) are classified into two types, G protein-coupled receptors (GPCRs) and ligand-gated ion channels. The former occupies a small part of the large GPCR superfamily, whereas the latter consists of three superfamilies. In these superfamilies, humans and rodents share almost the same set of neuroreceptor genes. This neuroreceptor gene set is good material to examine the degree of selective constraint exerted on each member gene of a given superfamily. If there are any neuroreceptor genes under the degree of selective constraint that is very different from that of the other member genes, they may have influenced the functional features characteristic of human neural activities. With the aim of identifying such neuroreceptor genes, we collected sequence data of orthologous neuroreceptor genes for humans, mice, and rats by database searches. This data set included ortholog pairs for 141 kinds of neuroreceptor genes, which covered almost the whole set of neuroreceptor genes known to be expressed in the human brain. The degree of selective constraint was estimated by computing the ratio (d_N/d_S) of the

number of nonsynonymous substitutions to that of synonymous substitutions. We found that the d_N/d_S ratio ranged widely and its distribution fitted a gamma distribution. In particular, we found that four neuroreceptor genes are under the significantly relaxed selective constraint. They are an ionotropic glutamate receptor subunit NMDA-2C, two $GABA_A$ receptor subunits, i.e., $GABA_A$-ε and $GABA_A$-θ, and a dopamine receptor D_4. Interestingly, these neuroreceptors have been reported to be associated with cognitive abilities such as memory and learning, and responsiveness to novel stimuli. These cognitive abilities can influence the behavioral features of an individual. Thus, it suggests that the relaxed-constraint neuroreceptor genes have evolved, assuring that the nervous system responds to a variety of stimuli with proper flexibility.

(5) The genome sequence and structure of rice chromosome 1.

Sasaki T, Matsumoto T, Yamamoto K, Sakata K, Baba T, Katayose Y, Wu J, Niimura Y, Cheng Z, Nagamura Y, Antonio BA, Kanamori H, Hosokawa S, Masukawa M, Arikawa K, Chiden Y, Hayashi M, Okamoto M, Ando T, Aoki H, Arita K, Hamada M, Harada C, Hijishita S, Honda M, Ichikawa Y, Idonuma A, Iijima M, Ikeda M, Ikeno M, Ito S, Ito T, Ito Y, Ito Y, Iwabuchi A, Kamiya K, Karasawa W, Katagiri S, Kikuta A, Kobayashi N, Kono I, Machita K, Maehara T, Mizuno H, Mizubayashi T, Mukai Y, Nagasaki H, Nakashima M, Nakama Y, Nakamichi Y, Nakamura M, Namiki N, Negishi M, Ohta I, Ono N, Saji S, Sakai K, Shibata M, Shimokawa T, Shomura A, Song J, Takazaki Y, Terasawa K, Tsuji K, Waki K, Yamagata H, Yamane H, Yoshiki S, Yoshihara R, Yukawa K, Zhong H, Iwama H, Endo T, Ito H, Hahn JH, Kim HI, Eun MY, Yano M, Jiang J, Gojobori T.

Rice Genome Research Program, National Institute of Agrobiological Sciences, 1-2, Kannondai 2-chome, Tsukuba, Ibaraki 305-8602, Japan.
tsasaki@nias.affrc.go.jp

The rice species Oryza sativa is considered to be a model plant because of its small genome size, extensive genetic map, relative ease of transformation and synteny with other cereal crops. Here we report the essentially complete sequence of chromosome 1, the longest chromosome in the rice genome. We summarize characteristics of the chromosome structure and the biological insight gained from the sequence. The analysis of 43.3 megabases (Mb) of non-overlapping sequence reveals 6,756 protein coding genes, of which 3,161 show homology to proteins of Arabidopsis thaliana, another model plant. About 30% (2,073) of the genes have been functionally categorized. Rice chromosome 1 is (G + C)-rich, especially in its coding regions, and is characterized by several gene families that are dispersed or arranged in tandem repeats. Comparison with a draft sequence indicates the importance of a high-quality finished sequence.

(6) Exhaustive search for T-box genes in the whole genome shotgun sequence of *Ciona intestinalis*

Kohji HOTTA, Hiroki TAKAHASHI[1], Yasuaki MOCHIZUKI[2], Takashi GOJOBORI and Nori SATOH[2]

1. *Department of Developmental Biology, National Institute for Basic Biology, 38 Nishigonaka, Myodaiji, Okazaki 444-8585, JAPAN*
2. *Department of Zoology, Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan*

T-box genes encode a family of transcription factors that share an evolutionally conserved T DNA binding domain. T-box genes were found among various animals, nematode, fly, sea-urchin, ascidians, amphioxus, zebrafish, *Xenopus*, newt, chick and mammals, and they make multiple T-box gene family. Several T-box subfamilies were known (ex. Brachyury, TBX4/5, TBX2/3, T-brain, TBX6). Members of the T-box gene family play important roles in the development of both vertebrate and invertebrate embryos, including specification of the mesoderm as well as heart and limb morphogenesis. To know what kind of T-box genes exist in primitive chordate ascidian is important for understanding the evolutionary process of the T-box genes and functional divergence in vertebrates. In this study, we performed exhaustive search against the *Ciona intestinalis* whole genome shotgun sequence data (total 900Mbp; National Institute of Genetics, Japan and total 450Mbp; Joint Genome Institute, USA). As a result, seven different types of 10 T-box genes were found in the *Ciona intestinalis* genome. Among them, six predicted genes have sequence homology with the known *ciona* T-box genes (i.e. *Ci-Bra, Ci-TBX6, Ci-VegTR, Ci-Tbx2/3, Ci-Tbx1, Ci-mT*). The other four T-box genes were

found as novel T-box genes (*Ci-Tbx20, Ci-TBX6a, Ci-TBX6c, Ci-TBX6d*) in the *Ciona intestinalis* genome. A Phylogenetic analysis of these genes suggested that the evolutionary origin of some of these genes is quite old, probably before the divergence between deuterostomes and protostomes. However, there is no gene which belongs to Tbx4/5 subfamily in *Ciona* genome as well as in the other protostomes genomes. On April 8~10, some of us took part in the *Ciona* Genome Annotation Jamboree (Department of Energy, Joint Genome Institute, USA). In this Jamboree, we could confirm above-mentioned data. In addition, we found that TBX6b,TBX6c and TBX6d genes were tandemly aligned on a single scaffold.

This study was reported at 'Molecular Evolution' a Meeting on Evoluion, Genomics, and Bioinformatics (13-16 June, Sorrento Italy). See Dehal *et al.*, (Science 298:2157-67, 2002) for details.

Publications

1. Katayama, K., Shirato-Horikoshi, H., Kojima, S., Kageyama, T., Oka, T., Hoshino, F. B., Fukushi, S., Shinohara, M., Uchida, K., Suzuki, Y., Gojobori, T. and Takeda, N.: Phylogenetic analysis of the complete genome of 18 norwalk-like viruses, Virology, **299**, 225-239, 2002.

2. Iwama, H. and Gojobori, T.: Identification of Neurotransmitter Receptor Genes under Significantly Relaxed Selective Constraint by Orthologous Gene Comparisons between Humans and Rodents, Molecular Biology and Evolution, **19 No.11**, 1891-1901, 2002.

3. Tanaka, Y., Hanada, K., Mizokami, M., Yeo, A. E. T., Shih, J. W.-K., Gojobori, T. and Alter, H. J.: A Comparison of the molecular clock of hepatitis C virus in the United States and Japan predicts that hepatocellular carcinoma incidence in the United States will increase over the next two decades, Proc. Natul. Acad. Sci., **99**, 15584-15589, 2002.

4. Fantom Consortium (Okazaki, Y., Gojobori, T., *et al.*), Riken Genome Exploration Research Group Phase I & II Team, Mouse Genome Sequencing Consortium, Scientific management: Hayashizaki, Y.: Analysis of the Mouse Transcriptome based upon Functional Annotation of 60,770 full length cDNAs. Nature **420**, 563-573, 2002.

5. Elmerot, C., Arnason, U., Gojobori, T. and Janke, A.: The mitochondrial genome of the pufferfish, Fugu rubripes, and ordinal teleostean relationships, Gene, **295**, 163-172, 2002.

6. Sasaki, T., Yamamoto, K., Sakata, K., Baba, T., Katayose, Y., Wu, J., Niimura, Y., Cheng, Z., Nagamura, Y., Kanamori, H., Hosokawa, S., Masukawa, M., Arikawa, K. and Antonio, B. A.: The genome sequence and structure of rice chromosome 1, Nature, **420**, 312-316, 2002.

7. Cebria, F., Kobayashi, C,. Umesono, Y., Nakazawa, M., Mineta, K., Ikeo, K., Gojobori, T., Itoh, M., Taira, M., Alvarado, A. and Agata, K.: FGFR-related gene nou-darake restricts brain tissues to the head region of planarians., Nature, **419**, 620-624, 2002.

8. Cebria, F., Kudome, T., Nakazawa, M., Mineta, K., Ikeo, K., Gojobori, T. and Agata, K.: The expression of neural-specific genes reveals the structural and molecular complexity of the planarian central nervous system, Elsevier, Mechanisms of Development **116**, 199-204, 2002.

9. K., Gojobori, T. and Agata, K.: The expression of neural-specific genes reveals the structural and molecular complexity of the planarian central nervous system, Elsevier, Mechanisms of Development **116**, 199-204, 2002.

10. Isobe, T., Yoshino, M., Mizuno, K.-I., Lindahl, K. F., Koide, T., Gaudieri, S., Gojobori, T. and Shiroishi, T.: Molecular Characterization of the Pb Recoombination Hotspot in the Mouse Major Histocompatibility Complex Class II Region, Genomics, **80**, 229-235, 2002.

11. Cebria, F., Nakazawa, M., Mineta, K., Ikeo, K., Gojobori, T. and Agata, K.: Dissecting planarian central nervous system regeneration by the expression of neural-specific genes, Develop. Growth Deffer, **44**, 135-146, 2002.

12. Kunihiro, S., Kawanishi, Y., Sano, M., Naito, K., Matsuura, Y., Tateno, Y., Gojobori, T., Yamagata, Y., Abe, K. and Machida, M.: A polymerase chain reaction-based method for cloning novel members of a gene family using a combination of degenerate and

inhibitory primers, Gene, **289**, 177-184, 2002.

13. Blair, J. E., Ikeo, K., Gojobori, T. and Hedges, B.: The Evolutionary Position of Nematodes, BMC Evolutionary Biology, 2:7, http://www.biomedcentral.com/1471-2148/2/7, 2002.

14. Akashi, H. and Gojobori, T.: Metabolic efficiency and amino acid composition in the proteomes of Escherichia coli and Bacillus subtilis, Proc. Natl. Acad. Sci., **99**, 3695-3700, 2002.

15. Ogihara, Y., Isono, K., Kojima, T., Endo, A., Hanaoka, M., Shiina, T., Terachi, T., Utsugi, S., Murata, M., Mori, N., Takumi, S., Ikeo, K., Gojobori, T., Murai, R., Murai, K., Matsuoka, Y., Ohnishi, Y., Tajiri, H. and Tsunewaki, K.: Structural features of a wheat plastome as revealed by complete sequencing of chloroplast DNA, Mol Genet Genomics, **266**, 740-746, 2002.

16. Tateno, Y., Imanishi, T., Miyazaki, S., Fukami-Kobayashi, K., Saitou, N., Sugawara, H. and Gojobori, T.: DNA Data Bank of Japan (DDBJ) for genome scale research in life science, Nucl. Acids. Res., **30**, 27-30, 2002.

17. Niimura, Y. and Gojobori, T.: In silico chromosome staining: reconstrustion of Giemsa bands from the whole human genome sequence, Proc. Natl. Acad. Sci., **99**, 797-802, 2002.

18. Ogawa, K., Ishihara, S., Saito, Y., Mineta, K., Nakazawa, M., Ikeo, K., Gojobori, T., Watanabe, K. and Agata, K.: Induction of a noggin-like gene by ectopic DV interaction during planarian regeneration. Dev Biol. 2002 Oct 1; 250(1): 59-70. PMID: 12297096 [PubMed-indexed for MEDLINE]

19. Callaerts, P., Lee, PN., Hartmann, B., Farfan, C., Choy, DW., Ikeo, K., Fischbach, KF. and Gehring, WJ.: de Couet HG. HOX genes in the sepiolid squid Euprymna scolopes: implications for the evolution of complex body plans. Proc Natl Acad Sci U S A. 2002 Feb 19; 99(4): 2088-93. PMID: 11842209 [PubMed-indexed for MEDLINE]

I-b. LABORATORY FOR GENE-PRODUCT INFORMATICS

Prof. Ken NISHIKAWA
Assist. Prof. Satoshi FUKUCHI

At present the Nishikawa laboratory consist of 14 members: Professor Ken Nishikawa, Assistant Professor Satoshi Fukuchi, two post-doctral research fellows (Drs. Homma and Nagashima), 8 technical assistants (Mses. Mimura, Yamamoto, Kuromaru, Abe, Kubota, Hongo, Ito and Suzuki), and a secretary (Ms. Sugiyama). We aim at theoretical work on protein molecules, such as prediction of three-dimensional structures from amino acid sequences, theoretical basis of protein stability, and molecular dynamics of proteins. Recently we have extended our research fields to genome analysis based on protein three-dimensional structures. We have developed two databases, Protein Mutant Database (PMD) and Genomes to Protein Structures and Functions (GTOP), and three computer tools for analysis of amino acid sequences and structures of proteins, LIBRA (Protein structure-sequence compatibility analysis), MATRAS (Protein tertiary structure comparison), and SSThread (Prediction of protein secondary structure using threading).

(1) GTOP: A database of protein structures predicted from genome sequences

Takeshi KAWABATA[1], Satoshi FUKUCHI, Keiichi HOMMA, Motonori OTA[2], Jiro ARAKI[3], Takehiko ITO[3], Nobuyuki ICHIYOSHI[3] and Ken NISHIKAWA ([1]Graduate School of Information Science, Nara Instiute of Science and Technology [2]Global Scientific Information and Computing Center, Tokyo Institute of Technology, [3]Mitsubishi Res. Ins.)

Large-scale genome projects generate an unprecedented number of protein sequences, most of them are experimentally uncharacterized. Predicting the 3D structures of sequences provides important clues as to their functions. We constructed the Genomes TO Protein structures and functions (GTOP) database, containing protein fold predictions of a huge number of sequences. Predictions are mainly carried out with the homology search program PSI-BLAST, currently the most popular among high-sensitivity

profile search methods. GTOP also includes the results of other analyses, e.g. homology and motif search, detection of transmembrane helices and repetitive sequences. We have completed analyzing the sequences of 101 organisms, with the number of proteins exceeding 430,000 in total. GTOP uses a graphical viewer to present the analytical results of each ORF in one page in a 'color-bar' format. The assigned 3D structures are presented by Chime plug-in or RasMol. The binding sites of ligands are also included, providing functional information. The GTOP server is available at http://spock.genes.nig.ac.jp/~genome/gtop.html. See Ref. 1 for details.

(2) A systematic investigation identifies a significant number of probable pseudogenes in the *Escherichia coli* genome

Keiichi HOMMA, Satoshi FUKUCHI, Takeshi KAWABATA[1], Motonori OTA[2] and Ken NISHIKAWA ([1]Graduate School of Information Science, Nara Instiute of Science and Technology [2]Global Scientific Information and Computing Center, Tokyo Institute of Technology)

Pseudogenes are dysfunctional open reading frames (ORFs) with a high homology to known genes. Although many pseudogenes have been reported to exist in the genomes of eukaryotes and obligate parasites, they have not received detailed analysis in free-living bacteria. In particular, pseudogenes have never been identified in *Eschericia coli*. Comparison of the genomes of *E. coli* strains K-12 and O157 identified a number of ORFs significantly shorter than their orthologs in the other strain. To select pseudogenes from these ORFs, we eliminated those possibly arising from of sequence errors, splitting gene candidates and those of prophage origin. The process yielded 101 and 116 ORFs in K-12 and O157 strains, respectively. Examinations of predicted three-dimensional structures of the protein products revealed that more than 90% of them are unable to fold properly and thus dysfunctional. We therefore conclude that a significant number of pseudogenes exist in *Eschericia coli*. As the genomes of free-living bacteria are turned over at a high rate, with a constant inflow of horizontal transfer genes, we propose that superfluous genes often become pseudogenes before eventual elimination. See Ref 2 for details.

(3) Identification of amino acids involved in protein structural uniqueness: Implication for de novo protein design

Yasuhiro ISOGAI[1], Motonori OTA[2], Anna ISHII[3], Manabu ISHIDA[4] and Ken NISHIKAWA ([1]RIKEN Wako, [2]Global Scientific Information and Computing Center, Tokyo Institute of Technology, [3]Department of Physics, Faculty of Science, Gakushuin University, [4]Department of Life Sciences, Faculty of Science, Himeji Institute of Technology)

Structural uniqueness is characteristic of native proteins and is essential to express their biological functions. The major factors that bring about the uniqueness are specific interactions between hydrophobic residues and their unique packing in the protein core. To find the origin of the uniqueness in their amino acid sequences, we analyzed the distribution of the side chain rotational isomers (rotamers) of hydrophobic amino acids in protein tertiary structures and derived deltaS(contact), the conformational-entropy changes of side chains by residue-residue contacts in each secondary structure. The deltaS(contact) values indicate distinct tendencies of the residue pairs to restrict side chain conformation by inter-residue contacts. Of the hydrophobic residues in alpha-helices, aliphatic residues (Leu, Val, Ile) strongly restrict the side chain conformations of each other. In beta-sheets, Met is most strongly restricted by contact with Ile, whereas Leu, Val and Ile are less affected by other residues in contact than those in alpha-helices. In designed and native protein variants, deltaS(contact) was found to correlate with the folding-unfolding cooperativity. Thus, it can be used as a specificity parameter for designing artificial proteins with a unique structure. See Ref. 4 for details.

Publications

1. Kawabata, T., Fukuchi, S., Homma, K., Ota, M., Araki, J., Ito, T., Ichiyoshi, N. and Nishikawa, K.: GTOP: A database of protein structures predicted from genome sequences. *Nucl. Acids Res.*, **30**, 294-298, 2002.

2. Homma, K., Fukuchi, S., Kawabata, T., Ota, M. and Nishikawa, K.: A systematic investigation identifies a significant number of probable pseudogenes in the *Escherichia coli* genome. *Gene*, **294**, 25-33, 2002.

3. Nishikawa, K.: Editorial: Information concept in biology. *Bioinformatics*, **18**, 649-651, 2002.

4. Isogai, Y., Ota, M., Ishii, A., Ishida, M. and Nishikawa, K.: Identification of amino acids involved in protein structural uniqueness: Implication for de novo protein design. *Prot. Engineer.* **15**, 555-560, 2002.

5. Fujishima, H., Nishimura, A., Wachi, M., Takagi, H., Hirasawa, T., Teraoka, H., Nishimori, K., Kawabata, T., Nishikawa, K. and Nagai, K.: kdsA mutations affect FtsZ-ring formation in *Escherichia coli* K-12. *Microbiology*, **148**, 103-112, 2002.

6. Fukami-Kobayashi, K., Tateno, Y. and Nishikawa, K.: Parallel evolution of ligand specificity between LacI/GalR family repressors and periplasmic sugar-binding proteins. *Molecular Biology and Evolution*, in press.

7. Kashiwagi, K., Shiba, K., Fukami-Kobayashi, K., Noda, T., Noguchi, H. and Nishikawa, K.: Characterization of folding pathways of the type-1 and type-2 periplasmic binding proteins MglB and ArgT. *J. Biochem*, in press.

8. Fukuchi, S., Yoshimune, K., Wakayama, M., Moriguchi, M. and Nishikawa, K.: Unique amino acid composition of proteins in halophilic bacteria. *J. Mol. Biol.*, in press.

9. Nakashima, H., Fukuchi, S. and Nishikawa, K.: Compositional changes in RNA, DNA and proteins for bacterial adaptation to higher and lower temperatures. *J. Biochem.*, in press.

I-c. LABORATORY FOR GENE FUNCTION RESEARCH

(1) Evolution of the MHC class I genome region in man and chimpanzee

Yoshio TATENO, Kaoru FUKAMI-KOBAYASHI, Hidetoshi INOKO[1], Takashi SHIINA[1] and Muneo YAMAZAKI[2]
([1]Tokai University, [2]Fujiya Bioscience Institute)

MHC class I genome region contains not only MHC genes, A, B, C and others, but also MHC related genes, MICA and MICB and repetitive sequences, LINEs and SINEs. As well known, this region is remarkably dynamic in view of evolution. We can observe almost every evolutionary event in this region; mutation, selection, transposition, duplication, deletion and insertion. We have paid special attention to evolutionary aspects of the B and C loci and MICA and MICB loci. Analysis of the human MHC class I genome region has shown that the two pairs of the MHC and MIC genes were each produced by genome fragment duplication. We have been estimating the evolutionary time of these duplications by comparing genome regions around HLA-B, HLA-C, MICA and MICB loci with the orthologous regions in the chimpanzee genome. We calculated the duplication time by using LINEs in the duplicated genome fragments. Those repetitive sequences can be used as a molecular clock, because all of them are fragmentary without their original function and thus considered to have evolved by neutral mutation. In the case of B and C loci, we confirmed that chimpanzee has essentially orthologous genome structures around these loci to those of man, indicating that the duplication of B and C loci predates the divergence of human and chimpanzee lineages. Sequence analysis of LINEs revealed that the duplication had occurred after the divergence of apes and old world monkey lineages. On the other hand, we found that chimpanzee has only one MIC locus. Genome structure around the gene indicated that the chimpanzee has a fused MIC gene by deletion of a region between MICA and MICB loci. We confirmed this evolutionary scenario by analyzing the LINE sequences around the MIC loci of man and chimpanzee. We are now analyzing the MHC class I genome region in rhesus monkey, to extend our study and trace the evolution of this region further back.

(2) Parallel evolution of ligand specificity between LacI/GalR family repressors and periplasmic sugar-binding proteins

Kaoru FUKAMI-KOBAYASHI, Yoshio TATENO and Ken NISHIKAWA

The bacterial LacI/GalR family repressors such as lactose operon repressor (LacI), purine nucleotide synthesis repressor (PurR), and trehalose operon repressor (TreR) consist of not only the N-terminal helix-turn-helix DNA-binding domain but also the C-terminal ligand-binding domain that is structurally homologous to periplasmic sugar-binding proteins.

These structural features imply that the repressor family evolved by acquiring the DNA-binding domain in the N-terminal of an ancestral periplasmic binding protein (PBP). Phylogenetic analysis of the LacI/GalR family repressors and their PBP homologues revealed that the acquisition of the DNA-binding domain occurred first in the family, and ligand specificity then evolved. The phylogenetic tree also indicates that the acquisition occurred only once before the divergence of the major lineages of eubacteria, and that the LacI/GalR and the PBP families have since undergone extensive gene duplication/loss independently along the evolutionary lineages. Multiple alignments of the repressors and PBPs furthermore revealed that repressors and PBPs with the same ligand specificity have the same or similar residues in their binding sites. This result, together with the phylogenetic relationship, demonstrates that the repressors and the PBPs individually acquired the same ligand specificity by homoplasious replacement, even though their genes are encoded in the same operon.

Publications

1. Tateno, Y., Imanishi, T., Miyazaki, S., Fukami-Kobayashi, K., Saitou, N., Sugawara, H. and Gojobori, T.: DNA Data Bank of Japan (DDBJ) for genome scale research in life science. Nucleic Acids Res **30**: 27-30, 2002.

2. Kunihiro, S., Kawanishi, Y., Sano, M., Naitoh, K., Matsu-ura, S., Tateno, Y., Gojobori, T., Yamagata, Y., Abe, K. and Machida, M.: A PCR-based method for cloning novel members of a gene family by a combination of degenerate and inhibitory primers. Gene **289**: 177-184, 2002.

3. Fukami-Kobayashi, K., Schrieber, D. R. and Benner, S. A.: Detecting compensatory covariation signals in protein evolution using reconstructed ancestral sequences., J. Mol. Biol. **319**: 729-743, 2002.

4. Fukami-Kobayashi, K.: How to make good use of CLUSTALW. Protein, Nucleic Acid and Enzyme **47**: 1237-1239, 2002.

5. Miyazaki, S., Sugawara, H., Gojobori, T. and Tateno, Y.: DNA Data Bank of Japan (DDBJ) in XML, Nucleic Acids Res **31**: 13-16, 2003.

6. Fukami-Kobayashi, K., Tateno, Y. and Nishikawa, K.: Parallel evolution of ligand specificity between LacI/GalR family repressors and periplasmic sugar-binding proteins. Mol Biol Evol (in press), 2003.

7. Miyazaki, S. and Tateno, Y.: DNA Data Bank of Japan as an Indispensable Public Database, In: Population and Community, B. M. Knoppers ed, Kluwer International Press (in press), 2003.

8. Brazma, A., Ikeo, K. and Tateno, Y.: Standardization of microarray experiment data, Protein, Nucleic Acid and Enzyme (in Japanese) (in press), 2003.

I-d. LABORATORY FOR RESEARCH AND DEVELOPMENT OF BIOLOGICAL DATABASES

1. Information systems for molecular data

1) Web services as the programmatic interfaces

It is very easy now for biologists to search, browse and utilize a number of data resources, i.e., databases and data analysis tools one by one thanks to the Internet and the World Wide Web. However, it is still difficult to use the data resource in an integrated way specific to a work flow for the research and development. In order to improve the situation, we introduced the Web services in addition to XML technology that we had implemented into the nucleotide sequence database in 2001.

The Web browser such as Internet Explorer and Netscape provides human-friendly interfaces. By contrast, the Web services provide the programmatic interfaces, i.e., program-friendly interfaces. A use program written in Java and Perl is able to search Web services that provide functions in need and bind to the Web services. In other words, the user program can integrate various Web services according to the work flow. The Web services that we developed are available at http://xml.nig.ac.jp/

2) Expansion of Genome Information Broker (GIB)

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 100 organisms as of December 2002. 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) WFCC-MIRCEN World Data Centre for Microorganisms (WDCM)

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 62 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/. We also used XML technology to organize the Web page.

2) Development of an e-Workbench for Biological Classification and Identification (InforBIO)

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.

3) An information system for pathogenic microorganisms[1]

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.

3. Applications of IT to the International Nucleotide Sequence Database

1) Development of Open Annotation System[2]

More than 100 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.

1 The information system on pathogenic microorganisms has been supported by Special Coordination Funds for Promoting Science and Technology

2 OASYS project has been supported by BIRD of Japan Science and Technology Corporation (JST)

2) Improvement of functions of INSD

We are in charge of developing tools for the construction, maintenance and dissemination of INSD. The following tools are studied and improved this year:
- the automatic scheduler to diffuse entries on the date specified by submitters
- a tool to compress the whole data set of INSD
- non-redundant daily update data set of INSD
- standalone application to check the translation of CDS in JAVA

Publication

1. Sugawara, H. (ed): Bioinformatics (*in Japanese*), Kyoritsu Shuppan Co., Ltd. (Tokyo), 2002.

2. Miyazaki, S. and Sugawara, H.: Networking of biological resource centers: WDCM experiences, *Data Science Journal*, **1** (2), 102-107, 2002.

3. Sugawara, H.: What can the pubic database of DDBJ provide? (*in Japanese*), *Genome-igaku*, **2** (3), 61-69, 2002.

4. Sugawara, H.: Agglomeration of Life Sciences and IT -Bioinformaitcs (*in Japanese*), Biological resources access (Watanabe M. and Nimura, S (ed)), Toyo-keizai-shinpou-sha (Tokyo), 205-238, 2002.

5. Sugawara, H.: Concept and progress of biological resource centers in OECD (in Japanese), Microbiol. Cult. Coll, **17** (2), 81-87, 2002.

6. Fumoto, M., Miyazaki, S. and Sugawara, H.: Genome Information Broker (GIB): data retrieval and comparative analysis system for completed microbial genomes and more, *Nucleic Acid Research*, **30** (1), 66-68, 2002.

7. Tateno, Y., Imanishi, T., Fukami-Kobayashi, K., Saitou, N., Sugawara, H. and Gojobori, T.: DNA Data Bank of Japan (DDBJ) for genome scale research in life science, *Nucleic Acid Research*, **30** (1), 27-30, 2002.

J. RADIOISOTOPE CENTER

Yoshiharu YAMAICHI, Katsunori YATA, Yasuyuki OGATA and Hironori NIKI

We focus on chromosome dynamics of the prokaryote for its genome integrity. We are studying the proteins and the DNA sites responsible for the regulation of prokaryotic DNA segregation using a combination of genetics, molecular, biochemical, cell-biological, and genomic approaches in Escherichia coli. Prokaryotes are not known to have a eukaryotic-like mitotic apparatus, and little is known about the mechanisms controlling chromosome partitioning. We visualized bacterial chromosomal DNA and plasmid DNA in cells by using fluorescence in situ hybridization (FISH).

We have already reveled subcellular localization of different DNA segments on the *E. coli* chromosome during the cell division cycle by using fluorescence in situ hybridization. These chromosome segments are located within the cell in the same order as a genetic map. A large chromosomal region including replication origin, *oriC* shows bipolar localization after replication, which we call the Ori domain. Then, we guessed that probably a chromosomal partitioning site was located within the Ori domain.

Now we know complete genome sequences of many kinds of bacteria. It is not impossible to identify replication origin and terminus based on only the sequence data. In contrast, we have not succeeded in finding a cis-acting site for chromosome partitioning, or a bacterial centromere. To identify a cis-acting site for chromosome partitioning, we have constructed a new mutant, in which a circular chromosome was split into two circular chromosomes. In this mutant, we were able to study an effect of a cis-acting segment on chromosome segregation. A chromosome-split mutant, in which 740 kb of the chromosome region [84.7-89.4 min] in the Ori domain is separated from the original chromosome, showed irregular localization of nucleoids and defective in bipolar migration of the chromosomal segment including *oriC*. These results suggest that a cis-acting site for bipolar migration of the Ori domain is located on this chromosomal region. Furthermore, we have analyzed that nucleoid localization and *oriC* migration in a series of chromosome-split mutants, and deletion mutants. Finally, we identified a site that was responsible for bipolar migration of the *oriC* segment as a 25 bp sequence on 89.1 min of the *E. coli* map.

Publication

1. Kodama, K., Kobayashi, T., Niki, H., Hiraga, S., Oshima, T., Mori, H. and Horiuchi, T.: Amplification of Hot DNA segments in Escherichia coli, Mol Microbiol **45**, 1575-1588, 2002.

K. EXPERIMENTAL FARM

(1) Development of new experimental strains and reevaluation of the genetic stocks of rice

Ken-Ichi NONOMURA, Mitsugu EIGUCHI, Toshie MIYABAYASHI and Nori KURATA

Making rice genetic stocks good resources and research works employing those resources carried out in the experimental farm are cooperative works between the experimental farm and the plant genetics laboratory. From 1998, Yukihiro ITO (Plant Genetics Lab.) and we have been developing a new rice genetic stock of enhancer trap lines aiming to generate over ten thousands lines, then to characterize, propagate and distribute them as rice resources for functional genomics. Other works are more or less depending on the experiments using plant resources; e.g. many numbers of transgenic rice, chemically induced mutant strains, strains of insertional mutagenesis and so on. These materials grown in the greenhouses and the exprimental farm are indispensable for plant genetics and genomics.

For details, see the reports of plant genetics lab.

(2) A new resource work for wild rice species involved in the activity as a center of national bioresource project for rice

Nori KURATA, Toshie MIYABAYASHI, Yukiko YAMAZAKI*, Mitsugu EIGUCHI and Ken-Ichi NONOMURA, (*Genetic Strains Information Lab.)

In 2002, the national bio-resource project was started for representative genetic and biological resources under the MEXT. The plant genetics laboratory and experimental farm act as a center for rice genetic resources project. In this project, we deal with about 2000 wild rice accessions including 9 genomes and over 20 species collected all over the world. We planned to propagate and store all seeds of these strains for distribution to rice research community till the end of 2006. In 2002, 160 accessions of wild- and cultivated- rice seeds were renewed and 559 lines were distributed to 24 institutes. Out of 2000 accessions, we chose 289 lines as core collections ranking from one to three, according to their typical characters, habitat and distribution in genome/species. Characterization of these strains for phenotypes and molecular features, are processed to house in the rice database "Oryzabase" with their photograps. The rice national bio-resource project consists of five university members including our laboratory. As a center of rice resource project, we are responsible for the manegement to make rice genetic resources powerful tools for further genomic researches. We are collaboratively progress the refinement of our rice resources. In this, we extract useful characters from all genetic stocks involved in the project as electric data and stored them in the "Oryzabase". Information about these strains can be accessed at the web site of "Oryzabase", which is a comprehensive rice genome resources database at http://www.shigen.nig.ac.jp/rice/oryzabase/index.html. Construction of the Oryzabase is undertaken by Yukiko Yamazaki, a manager of all data from organisms being dealt in the national bio-resource project. Addition of new information and revision of the "Oryzabase" will be performed occasionally to include more useful information for rice basic research under the guidance of the rice genetic resources committee in Japan (chair :N. Kurata).

Publications

1. Harushima, Y., Nakagahra, M., Yano, M., Sasaki, T. and Kurata, N.: Diverse variation of reproductive barriers in three intraspecific rice crosses. Genetics. 313-322, 2002.

2. Ahn, B. O., Miyoshi, K., Itoh, J. I., Nagato, Y. and Kurata, N.: A genetic and physical mapping of the region containing *PLASTOCHRON1*,a *heterochronic* gene, in rice (*Oryza sativa*, L). Theor. Appl. Genet. 105: 654-659, 2002.

3. Ito, Y., Hirochika, H. and Kurata, N.: Organ specific alternative transcripts of KNOX family class 2 homeobox genes of rice. Gene, 2002.

4. Dunford, R.P., Yano, M., Kurata, N., Sasaki, T., Huestis, G., Rocheford, T. and Laurie, D.A.: Compar-

ative mapping of the barley Ppd-H1 photoperiod response gene region, which lies close to a junction between two rice linkage segments. Genetics, **161**, 825-834, 2002.

5. Kurata, N., Nonomura, K-I. and Harushima, Y.: Rice genome organization focusing on centromere and genome interaction studies. Annals of Botany **90**, 427-435, 2002.

6. Yamazaki, Y., Yoshimura, A., Nagato, Y. and Kurata, N.: Recent advances in Oryzabase (Integrated Rice Database). Rice Genetics Newslet. **19**, 7-8, 2002.

7. Kurata, N. and Fukui, K.: Chromosome research in genus Oryza. in Monograph in genus Oryza. Nanda JS. Ed, Springer (in press).

L. TECHNICAL SECTION

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 *E. coli* 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 technical officials.

Publications

1. Davies, K., Bowden, L., Smith, P., Dean, W., Hill, D., Furuumi, H., Sasaki, H., Cattanach, B. and Reik, W.: Disruption of mesodermal enhancers for *Igf2* in the minute mutant. *Development* **129**, 1657-1668, 2002.

2. Hayashi, S., Ito, K., Sado, Y., Taniguchi, M., Akimoto, A., Takeuchi, H., Aigaki, T., Matsuzaki, F., Nakagoshi, H., Tanimura, T., Ueda, R., Uemura, T., Yoshihara, M. and Goto, S.: GETDB, a database compiling expression patterns and molecular locations of a collection of Gal4 enhancer traps. *Genesis* **34**, 58-61, 2002.

3. Masumoto, H., Muramatsu, S., Kamimura, Y. and Araki, H.: S-Cdk-dependent phosphorylation of Sld2 essential for chromosomal DNA replication in budding yeast. *Nature* **415**, 651-655, 2002.

Oral presentations

1. Furuumi, H.: Detection of methylated DNA by bisulfite genomic sequencing. The 13th Annual Meeting for Technical Engineers, Okazaki, February, 2002.

2. Miura, A., Watanabe, K., Kato, M., Kotani, H., Kakutani, T.: Mobile endogenous *CACTA* family elements in *Arabidopsis*. The 74th Annual Meeting of the Genetics Society of Japan, Fukuoka, October, 2002.

Poster presentations

1. Kiso, M.: Mouse ES cell culture and Generation of Gene targeted mouse. The 13th Annual Meeting for Technical Engineers, Okazaki, February, 2002.

2. Muramatsu, S. and Araki, H.: Functional analysis of the *DPB11* gene in *Saccharomyces cerevisiae*. The 25th Annual Meeting of the Molecular Biology Society of Japan, Yokohama, December, 2002.

3. Saka, K., Matsumoto, N. and Nishimura, A.: Isolation of a whole set of cell division genes in *Escherichia coli*. The 13th Annual Meeting for Technical Engineers, Okazaki, February, 2002.

4. Saka, K., Matsumoto, N. and Nishimura, A.: Global regulation of cell division in *Escherichia coli*. The 25th Annual Meeting of the Molecular Biology Society of Japan, Yokohama, December, 2002.

ABSTRACTS OF DIARY FOR 2002

Biological Symposium

Jan.11 Cell Cycle Control and Chromosome Partition (Mitsuhiro YANAGIDA)

Jan.21 The Vitamin D Receptor: Molecular Variation and Functional Outcomes (Alexander P KOUZMENKO)

Jan.25 The ultimate altruistic individuals: phylogeny, evolution, ecology, physiology and molecular biology of aphid soldiers (Kazuma FUKATSU)

Feb.14 Evolutionary Distance Estimation under Heterogeneous Substitution Pattern (Koichiro TAMURA)

Feb.20 Synchronization as a survival tool of the cellular slime mold and its generality (Seido NAGANO)

Feb.25 Molecular functions of two TERMINAL FLOWER genes of Arabidopsis (Koji GOTO)

Mar.4 Protein scaffolds and cell polarity in Drosophila (Elisabeth KNUST)

Mar.4 Cell type specification in the developingcentral nervous system of the zebrafish (Jose CAMPOS-ORTEGA)

Mar.13 Molecular dissection of S allele which controlls self-imcompatibility in Brassica (Masao WTANABE)

Mar.18 Regulatory mechanisms and function of the MAP kinase signal transduction pathways (Eisuke NISHIDA)

Mar.22 MEC3, MEC1 and DDC2 functions as telomere checkpoint (Shinichiro ENOMOTO)

Apr.3 Designing Artificial Transcription Factors (Aseem Z.ANSAR)

May.7 Stress and developmental regulation of RNA polymerase Ⅱ C-terminal domain phosphorylation (Oliver BENSAUDE)

June.17 Rice Functional Genomics and Bioinformatics at IRR (Richard M.BRUSKIEWICH)

June.20 Functional analysis of the canoe gene in Drosophila morphogenesis (Kuniaki TAKAHASHI)

July.3 Cell cycle regulation by myb oncogene (Masahiro OKADA)

July.4 The establishment of the stem cell fate in Drosophila germline (Miho ASAOKA)

July.4 Dnmt3L is required for genomic imprinting during oogenesis (Kenichiro HATA)

July.9 Molecular Genetic Analysis of Plant Sex Chromosome and Application of the Microscopic Manipulation to Genetics (Sachihiro MATSUNAGA)

July.9 Fundamental Mechanism of Axon Growth of Neuron Cells (Junichi YUASA)

July.12 Chromatin Inheritance during cellular proliferation (Keiichi SHIBAHARA)

Aug.16 Molecular mechanism of programmed cell death and its role in the determination of cell fate (Hirotaka KANUKA)

Aug.26 Partition of replicated sister chromosomes in Escherichia coli (Sota HIRAGA)

Aug.27 Membrane traffic and sterol-roles in post-Golgi protein sorting- (Kyohei UMEBAYASHI)

Aug.28 Structure and dynamics of self-organization: Examples for self-driving structures composed of macromolecules (Shin-ichiro, M. NOMURA)

Aug.28 1. Activation machanism of protease complexes 2. Observation of plant organelle by atomic-force microscope (Takafumi YAMADA)

Aug.29 Cell Cycle Control by Ubiquitin-dependent Protein Degradation (Hiroyuki YAMANO)

Aug.29 Maternal-effect mutations affecting the early development in zebrafish (Yasuyuki KISHIMOTO)

Aug.29 Part Ⅰ: T-box gene interactions mediate regional specification of the zebrafish mesoderm; Part Ⅱ: a new midline gene required to establish all left-right asymmetri es (Kazuyuki HOSHIJIMA)

Sep.2 Interplay of a vacuolar pathogen Legionella pneumophila and its phagocytic host cells (Hiroki NAGAI)

Sep.18 Mechanism and evolution of genomic regulation of the Dlx3-Dlx7 cluster (Kenta SUMIYAMA)

Sep.18 The Pttern of Polymorphism on Human Chromosome 21 (Hideki INNAN)

Sep.24 The structure of single-nucleotide variation in overlapping regions of human genome sequence

(Stephen T. SHERRY)

Sep.30 Small RNAs in genome rearragnement in Tetrahymena (Kazufumi MOCHIZUKI)

Sep.30 Hydra and the evolution of the bilateral body plan (Hans MEINHARDT)

Oct.7 Structural determination of biomolecules by solid NMR (Kiyonori TAKEKOSHI)

Oct.7 Poetry and prose of genetic sequences (Edward N. TRIFONOV)

Oct.9 Understanding the persistence of a long linkage disequilibrium block in the human genome: lessons fromthe GBA region of the human genome (Jaume BERTRANPETIT)

Oct.9 The new tide of the chemical synthesis of proteins: Its fusion with biology (Sabro AIMOTO)

Oct.10 Inducing and overcoming heterochromatin mediated silencing in mammals (Richard FESTENTEIN)

Oct.15 Functional analysis of the isthmin gene encoding a novel secretion protein (Isato ARAKI)

Oct.17 Distinct histone methylation patterns define active and silent chromatin states (Ken-ithi NOMA)

Nov.18 Mitosis and Endoreduplication: Alternative Cell Cycle Strategies in Plant Development (Yukiko MIZUKAMI)

Nov.19 Coactivator MBF1 supports AP-1 activity and extends lifetime under oxidative stress (Marek JINDRA)

Nov.28 Chromatin regulation at imprinting-control regions in the mouse (Robert FEIL)

Nov.28 Looking back at my research life (Tomoko OHTA)

Dec.5 Fishing for Novel Genes (John E. DOWLING)

Dec.9 Man? Mouse? Mole? Some lessons on imprinting from rare clinical disorders (David T. BONTHRON)

Dec.16 Mechanisms Preventing Re-replication in Saccharomyces cerevisiae. (Joachim LI)

Dec.18 Characterization of PR-Set7, a nucleosomal histone H4 lysine 20-specific methyltransfease (Kenichi NISHIOKA)

FOREIGN VISITORS IN 2002

Jan.21	Alexander P. Kouzmenko, Bone and Mineral Research Program, Garvan Institute of Medical Research, Sydney, Australia
Mar.4	Elisabeth Knust, Institut fuer Genetik, Heinrich-Heine Universitaet Duesseldorf
Mar.4	Jose Campos-Ortega, University of Cologne, Institute for Developmental Biology
Mar.22	Shinichiro Enomoto, University of Minnesota
Apr.3	Aseem Z. Ansari, Department of Biochemistry & Genome Center University of Wisconsin-Madison USA
May.7	Olivier Bensaude, Regulation de l'Expression Genetique, CNRS, Ecole Normale Superieure
June.17	Richard M. Bruskiewich, International Rice Research Institute (IRRI)
July.4	Miho ASAOKA, Department of Cell Biology Duke University Medical Center
July.4	Kenichiro HATA, Department of Medicine Harvard Medical School Cardiovascular Research Center Massachusetts General Hospital
Aug.29	Hiroyuki Yamano, Cancer Research UK London Research Institute Clare Hall Laboratories
Aug.29	Kazuyuki Hoshijima, Eccles Institute of Human genetics, Department of Human Genetics, University of Utah
Sep.2	Hiroki Nagai, Section of Microbial Pathogenesis, Yale School of Medicine
Sep.18	Kenta Sumiyama, Frank Ruddle laboratory, Yale University
Sep.18	Hideki Innan, Human Genetics Center, University of Texas Health Science Center at Houston
Sep.24	Stephen T. Sherry, National Center for Biotechnology Information, National Institute of Health
Sep.30	Kazufumi Mochizuki, Department of Biology, University of Rochester
Sep.30	Hans Meinhardt, Max Plank Institute, Tuebingen, Germany.
Oct.7	Edward N. Trifonov, Genome Diversity Center, Institute of Evolution, University of Haifa, Haifa, Israel
Oct.9	Jaume Bertranpetit, Unitat de Biologia Evolutiva. Universitat Pompeu Fabra. Spain
Oct.10	Richard Festenstein, MRC Clinical Sciences Center Hammersmith Hospital London, UK
Oct.15	Isato Araki, Max-Planck-Institute of Molecular Cell Biology and Genetics (MPI-CBG)
Oct.17	Ken-ithi Noma, Cold Spring Harbor Laboratory
Nov.18	Yukiko Mizukami, University of California, Berkeley, Department of Plant and Microbial Biology
Nov.19	Marek JINDRA, University of South Bohemia, Czech Republic
Nov.28	Robert Feil, Institute of Molecular Genetics, CNRS
Dec.5	John E. Dowling, Harvard University
Dec.9	David T. Bonthron, University of Leeds Molecular Medicine Unit St. James's University Hospital, UK
Dec.16	Joachim Li, Department of Microbiology & Immunology University of California, San Francisco
Dec.18	Kenichi Nishioka, University of Medicine and Dentistry of New Jersey

AUTHOR INDEX

Annual Report of the National Institute of Genetics. No.53

Issued by Director-General : Yoshiki Hotta, D.Med.
National Institute of Genetics
Yata 1111 Mishima, Shizuoka-Ken, 411-8540 Japan
Tel. 055-981-6718
Fax. 055-981-6715
Edited by Hironori Niki & Yasushi Hiromi
Published by National Institute of Genetics