Human Genome Center

Laboratory of DNA Information Analysis
Laboratory of Sequence Data Analysis
Laboratory of Genome Database

DNA情報解析分野
シークエンスデータ情報処理分野
ゲノムデータベース分野

Professor	Satoru Miyano, Ph.D.	教　授	理学博士	宮　野　　悟
Associate Professor	Seiya Imoto, Ph.D.	准教授	博士(数理学)	井　元　清　哉
Assistant Professor	Teppei Shimamura, Ph.D.	助　教	博士(工学)	島　村　徹　平
Project Assistant Professor	Atsushi Niida, Ph.D.	特任助教	博士(理学)	新井田　厚　司
Project Assistant Professor	Yuichi Shiraishi, Ph.D.	特任助教	博士(統計科学)	白　石　友　一
Associate Professor	Tetsuo Shibuya, Ph.D.	准教授	博士(理学)	渋　谷　哲　朗
Lecturer	Rui Yamaguchi, Ph.D.	講　師	博士(理学)	山　口　　類

The recent advances in biomedical research have been producing large-scale, ultra-high dimensional, ultra-heterogeneous data. Due to these post-genomic research progresses, our current mission is to implement computational/informatics strategy for systems biology and medicine towards translational bioinformatics for personalized genomic medicine. With this mission, we have been developing computational methods for understanding life as system and applying them to practical issues in medicine and biology.

1. Systems Cancer Research and Systems Biology

a. Statistical model-based testing to evaluate the recurrence of genomic aberrations

Niida A, Imoto S, Shimamura T, Miyano S

In cancer genomes, chromosomal regions harboring cancer genes are often subjected to genomic aberrations like copy number alteration and loss of heterozygosity. Given this, finding recurrent genomic aberrations is considered an apt approach for screening cancer genes. Although several permutation-based tests have been proposed for this purpose, none of them are designed to find recurrent aberrations from the genomic dataset without paired normal sample controls. Their application to unpaired genomic data may lead to false discoveries, because they retrieve pseudo-aberrations that exist in normal genomes as polymorphisms. We develop a new parametric method named parametric aberration recurrence test (PART) to test for the recurrence of genomic aberrations. The introduction of Poisson-binomial statistics allow us to compute small P-values more efficiently and precisely than the previously proposed permutation-based approach. Moreover, we extended PART to cover unpaired data (PART-up) so that there is a statistical basis for analyzing un-

paired genomic data. PART-up uses information from unpaired normal sample controls to remove pseudo-aberrations in unpaired genomic data. Using PART-up, we successfully predict recurrent genomic aberrations in cancer cell line samples whose paired normal sample controls are unavailable. This article thus proposes a powerful statistical framework for the identification of driver aberrations, which would be applicable to ever-increasing amounts of cancer genomic data seen in the era of next generation sequencing. Our implementations of PART and PART-up are available from http://www.hgc.jp/~niiyan/PART/manual.html.

b. Identifying regulational alterations in gene regulatory networks by state space representation of vector autoregressive models and variational annealing

Kojima K, Imoto S, Yamaguchi R, Fujita A[1] Yamauchi M, Gotoh N[2], Miyano S: [1]University of São Paulo [2]Molecular Targets Laboratory, Division of Molecular Therapy

In the analysis of effects by cell treatment such as drug dosing, identifying changes on gene network structures between normal and treated cells is a key task. A possible way for identifying the changes is to compare structures of networks estimated from data on normal and treated cells separately. However, this approach usually fails to estimate accurate gene networks due to the limited length of time series data and measurement noise. Thus, approaches that identify changes on regulations by using time series data on both conditions in an efficient manner are demanded.We developed a new statistical method that is based on the state space representation of the vector autoregressive model and estimates gene networks on two different conditions in order to identify changes on regulations between the conditions. In the mathematical model of our approach, hidden binary variables are newly introduced to indicate the presence of regulations on each condition. The use of the hidden binary variables enables an efficient data usage; data on both conditions are used for commonly existing regulations, while for condition specific regulations corresponding data are only applied. Also, the similarity of networks on two conditions is automatically considered from the design of the potential function for the hidden binary variables. For the estimation of the hidden binary variables, we derive a new variational annealing method that searches the configuration of the binary variables maximizing the marginal likelihood. For the performance evaluation, we use time series data from two topologically similar synthetic networks, and confirm that our proposed approach estimates commonly existing regulations as well as changes on regulations with higher coverage and precision than other existing approaches in almost all the experimental settings. For a real data application, our proposed approach is applied to time series data from normal Human lung cells and Human lung cells treated by stimulating EGF-receptors and dosing an anticancer drug termed Gefitinib. In the treated lung cells, a cancer cell condition is simulated by the stimulation of EGF-receptors, but the effect would be counteracted due to the selective inhibition of EGF-receptors by Gefitinib. However, gene expression profiles are actually different between the conditions, and the genes related to the identified changes are considered as possible off-targets of Gefitinib. From the synthetically generated time series data, our proposed approach can identify changes on regulations more accurately than existing methods. By applying the proposed approach to the time series data on normal and treated Human lung cells, candidates of off-target genes of Gefitinib are found. According to the published clinical information, one of the genes can be related to a factor of interstitial pneumonia, which is known as a side effect of Gefitinib.

c. Epidermal growth factor receptor tyrosine kinase defines critical prognostic genes of stage I lung adenocarcinoma

Yamauchi M[2], Yamaguchi R, Nakata A[2], Kohno T[10], Nagasaki M, Shimamura T, Imoto S, Saito A, Ueno K, Hatanaka Y, Yoshida R[8], Higuchi T[8], Nomura M[24], Beer DG[23], Yokota J[10], Miyano S, Gotoh N[2]: [8]Institute of Statistical Mathematics [23]University of Michigan [24]Tokyo Medical University

To identify stage I lung adenocarcinoma patients with a poor prognosis who will benefit from adjuvant therapy. Whole gene expression profiles were obtained at 19 time points over a 48-hour time course from human primary lung epithelial cells that were stimulated with epidermal growth factor (EGF) in the presence or absence of a clinically used EGF receptor tyrosine kinase (RTK)-specific inhibitor, gefitinib. The data were subjected to a mathematical simulation using the State Space Model (SSM). "Gefitinib-sensitive" genes, the expressional dynamics of which were altered by addition of gefitinib, were identified. A risk scoring model was constructed to classify high- or low-risk patients based on expression signatures of 139 gefitinib-sensitive genes in lung cancer using a training data set of 253 lung adenocarcinomas of North American cohort. The predictive ability of the risk scoring model was examined in independent cohorts of surgical specimens of lung cancer. The risk scoring model enabled the identification of high-risk stage IA and IB cases in another North

American cohort for overall survival (OS) with a hazard ratio (HR) of 7.16 (P = 0.029) and 3.26 (P = 0.0072), respectively. It also enabled the identification of high-risk stage I cases without bronchioalveolar carcinoma (BAC) histology in a Japanese cohort for OS and recurrence-free survival (RFS) with HRs of 8.79 (P = 0.001) and 3.72 (P = 0.0049), respectively. The set of 139 gefitinib-sensitive genes includes many genes known to be involved in biological aspects of cancer phenotypes, but not known to be involved in EGF signaling. The present result strongly re-emphasizes that EGF signaling status in cancer cells underlies an aggressive phenotype of cancer cells, which is useful for the selection of early-stage lung adenocarcinoma patients with a poor prognosis.

d. Cell cycle gene networks are associated with melanoma prognosis

Wang L[3], Hurley D[3], Watkins W[3], Araki H[3], Tamada Y[4], Muthukaruppan A[3], Ranjard L[3], Derkac E[3], Imoto S, Crampin E[3], Print C[3], Miyano S: [3]University of Auckland [4]Department of Computer Science, The University of Tokyo

Our understanding of the molecular pathways that underlie melanoma remains incomplete. Although several published microarray studies of clinical melanomas have provided valuable information, we found only limited concordance between these studies. Therefore, we took an in vitro functional genomics approach to understand melanoma molecular pathways. Affymetrix microarray data were generated from A375 melanoma cells treated in vitro with siRNAs against 45 transcription factors and signaling molecules. Analysis of this data using unsupervised hierarchical clustering and Bayesian gene networks identified proliferation-association RNA clusters, which were co-ordinately expressed across the A375 cells and also across melanomas from patients. The abundance in metastatic melanomas of these cellular proliferation clusters and their putative upstream regulators was significantly associated with patient prognosis. An 8-gene classifier derived from gene network hub genes correctly classified the prognosis of 23/26 metastatic melanoma patients in a cross-validation study. Unlike the RNA clusters associated with cellular proliferation described above, co-ordinately expressed RNA clusters associated with immune response were clearly identified across melanoma tumours from patients but not across the siRNA-treated A375 cells, in which immune responses are not active. Three uncharacterised genes, which the gene networks predicted to be upstream of apoptosis- or cellular proliferation-associated RNAs, were found to significantly alter apoptosis and cell number when over-expressed in vitro. This analysis identified co-expression of RNAs that encode functionally-related proteins, in particular, proliferation-associated RNA clusters that are linked to melanoma patient prognosis. Our analysis suggests that A375 cells in vitro may be valid models in which to study the gene expression modules that underlie some melanoma biological processes (e.g., proliferation) but not others (e.g., immune response). The gene expression modules identified here, and the RNAs predicted by Bayesian network inference to be upstream of these modules, are potential prognostic biomarkers and drug targets.

e. Computational gene network analysis reveals TNF-induced angiogenesis

Ogami K, Yamaguchi R, Imoto S, Tamada Y[4] Araki H[3], Print C[3], Miyano S

TNF (Tumor Necrosis Factor-α) induces HUVEC (Human Umbilical Vein Endothelial Cells) to proliferate and form new blood vessels. This TNF-induced angiogenesis plays a key role in cancer and rheumatic disease. However, the molecular system that underlies TNF-induced angiogenesis is largely unknown. We analyzed the gene expression changes stimulated by TNF in HUVEC over a time course using microarrays to reveal the molecular system underlying TNF-induced angiogenesis. Traditional k-means clustering analysis was performed to identify informative temporal gene expression patterns buried in the time course data. Functional enrichment analysis using DAVID was then performed for each cluster. The genes that belonged to informative clusters were then used as the input for gene network analysis using a Bayesian network and nonparametric regression method. Based on this TNF-induced gene network, we searched for sub-networks related to angiogenesis by integrating existing biological knowledge. k-means clustering of the TNF stimulated time course microarray gene expression data, followed by functional enrichment analysis identified three biologically informative clusters related to apoptosis, cellular proliferation and angiogenesis. These three clusters included 648 genes in total, which were used to estimate dynamic Bayesian networks. Based on the estimated TNF-induced gene networks, we hypothesized that a sub-network including IL6 and IL8 inhibits apoptosis and promotes TNF-induced angiogenesis. More particularly, IL6 promotes TNF-induced angiogenesis by inducing NF-κB and IL8, which are strong cell growth factors. Computational gene network analysis revealed a novel molecular system that may play an important role in the TNF-induced angiogenesis seen in cancer and rheumatic disease. This analysis suggests that Bayesian network analysis linked to functional annotation may be a powerful tool to provide insight into disease.

f. Functional clustering of time series gene expression data by Granger causality

Fujita A[1], Severino P[1], Kojima K, Sato JR[1], Patriota AG[1], Miyano S

A common approach for time series gene expression data analysis includes the clustering of genes with similar expression patterns throughout time. Clustered gene expression profiles point to the joint contribution of groups of genes to a particular cellular process. However, since genes belong to intricate networks, other features, besides comparable expression patterns, should provide additional information for the identification of functionally similar genes. In this study we performed gene clustering through the identification of Granger causality between and within sets of time series gene expression data. Granger causality is based on the idea that the cause of an event cannot come after its consequence. This kind of analysis can be used as a complementary approach for functional clustering, wherein genes would be clustered not solely based on their expression similarity but on their topological proximity built according to the intensity of Granger causality among them.

g. Vasohibin-1 is identified as a master-regulator of endothelial cell apoptosis using gene network analysis

Affara M[5], Sanders D[5], Araki H[3], Tamada Y[4], Dunmore B[5], Humphreys S[5], Imoto S, Savoie C[6], Kuhara S[7], Print C[3], Charnock-Jones DS[5], Miyano S: [5]University of Cambridge [6]GNI, Inc. [7]Kyushu University

Apoptosis is a critical process in endothelial cell (EC) biology and pathology, which has been extensively studied at protein level. Numerous gene expression studies of EC apoptosis have also been performed, however few attempts have been made to use gene expression data to identify the molecular relationships and master regulators that underlie EC apoptosis. Therefore, we sought to understand these relationships by generating a Bayesian gene regulatory network (GRN) model. ECs were induced to undergo apoptosis using serum withdrawal and followed over a time course in triplicate, using microarrays. When generating the GRN, this EC time course data was supplemented by a library of microarray data from EC treated with siRNAs targeting over 350 signalling molecules. The GRN model proposed Vasohibin-1 (VASH1) as one of the candidate master-regulators of EC apoptosis with numerous downstream mRNAs. To evaluate the role played by VASH1 in EC, we used siRNA to reduce the expression of VASH1. Of 10 mRNAs downstream of VASH1 in the GRN that were examined, 7 were significantly up-or down-regulated in the direction predicted by the GRN. Further supporting an important biological role of VASH1 in EC, targeted reduction of VASH1 mRNA abundance conferred resistance to serum withdrawal-induced EC death. We have utilised Bayesian GRN modelling to identify a novel candidate master regulator of EC apoptosis. This study demonstrates how GRN technology can complement traditional methods to hypothesise the regulatory relationships that underlie important biological processes.

h. XiP: a computational environment to create, extend and share workflows

Nagasaki M, Fujita A[1], Sekiya Y, Saito A, Ikeda E, Li C, Miyano S

XiP (eXtensible integrative Pipeline) is a flexible, editable and modular environment with a user-friendly interface that does not require previous advanced programming skills to run, construct and edit workflows. XiP allows the construction of workflows by linking components written in both R and Java, the analysis of high-throughput data in grid engine systems and also the development of customized pipelines that can be encapsulated in a package and distributed. XiP already comes with several ready-to-use pipeline flows for the most common genomic and transcriptomic analysis and ~300 computational components. XiP is open source, freely available under the Lesser General Public License (LGPL) and can be downloaded from http://xip.hgc.jp.

2. International Cancer Genome Consortium

a. Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators

Fujimoto A[9], Totoki Y[10], Abe T[9], Boroevich KA[9], Hosoda F[10], Hai Nguyen H[9], Aoki M[9], Hosono N[9], Kubo M[9], Miya F[9], Arai Y[10], Takahashi H[10], Shirakihara T[10], Nagasaki M, Shibuya T, Nakano K[9], Watanabe-Makino K[9], Tanaka H, Nakamura H[10], Kusuda J[11], Ojima H[10], Shimada K[12], Okusaka T[12], Ueno M[13], Shigekawa Y[13], Kawakami Y[14], Arihiro K[14], Ohdan H[14], Gotoh K[15], Ishikawa O[15], Ariizumi S[16], Yamamoto M[16], Yamada T[15], Chayama K[14], Kosuge T[12], Yamaue H[13], Kamatani N[9], Miyano S, Nakagama H[10], Nakamura Y[9,17], Tsunoda T[9], Shibata T[10], Nakagawa H[9]: [9]RIKEN [10]National Cancer Center Research Institute [11]National Institute of Biomedical Innovation [12]National Cancer Center Hospital [13]Wakayama Medical University [14]Hiroshima University School of Medicine [15]Osaka Medical Center for Cancer and Cardiovascular

Diseases [16]Tokyo Women's Medical University [17]Laboratory of Molecular Medicine

Hepatocellular carcinoma (HCC) is the third leading cause of cancer-related death worldwide. We sequenced and analyzed the whole genomes of 27 HCCs, 25 of which were associated with hepatitis B or C virus infections, including two sets of multicentric tumors. Although no common somatic mutations were identified in the multicentric tumor pairs, their whole-genome substitution patterns were similar, suggesting that these tumors developed from independent mutations, although their shared etiological backgrounds may have strongly influenced their somatic mutation patterns. Statistical and functional analyses yielded a list of recurrently mutated genes. Multiple chromatin regulators, including ARID1A, ARID1B, ARID2, MLL and MLL3, were mutated in ～50% of the tumors. Hepatitis B virus genome integration in the TERT locus was frequently observed in a high clonal proportion. Our whole-genome sequencing analysis of HCCs identified the influence of etiological background on somatic mutation patterns and subsequent carcinogenesis, as well as recurrent mutations in chromatin regulators in HCCs.

3. Statistical/Algorithmic Data Analysis Methods for Gene Expression Data, and Next-Generation Sequence Data, and Medical Record Data

a. Population model-based inter-diplotype similarity measure for accurate diplotype clustering

Onuki R[18], Yamada R[18], Yamaguchi R, Kanehisa M[18], Shibuya T: [18]Kyoto University

Classification of the individuals' genotype data is important in various kinds of biomedical research. There are many sophisticated clustering algorithms, but most of them require some appropriate similarity measure between objects to be clustered. Hence, accurate inter-diplotype similarity measures are always required for classification of diplotypes. In this article, we propose a new accurate inter-diplotype similarity measure that we call the population model-based distance (PMD), so that we can cluster individuals with diplotype SNPs data (i.e., unphased-diplotypes) with higher accuracies. For unphased-diplotypes, the allele sharing distance (ASD) has been the standard to measure the genetic distance between the diplotypes of individuals. To achieve higher clustering accuracies, our new measure PMD makes good use of a given appropriate population model which has never been utilized in the ASD. As the population model, we propose to use an hidden Markov model (HMM)-based model. We call the PMD based on the model the HHD (HIT HMM-based Distance). We demonstrate the impact of the HHD on the diplotype classification through comprehensive large-scale experiments over the genome-wide 8930 data sets derived from the HapMap SNPs database. The experiments revealed that the HHD enables significantly more accurate clustering than the ASD.

b. A filter based feature selection algorithm using null space of covariance matrix for DNA microarray gene expression data

Sharma A, Imoto S, Miyano S

We developed a new filter based feature selection algorithm for classification based on DNA microarray gene expression data. It utilizes null space of covariance matrix for feature selection. The algorithm can perform bulk reduction of features (genes) while maintaining the quality information in the reduced subset of features for discriminative purpose. Thus, it can be used as a pre-processing step for other feature selection algorithms. The algorithm does not assume statistical independency among the features. The algorithm shows promising classification accuracy when compared with other existing techniques on several DNA microarray gene expression datasets.

c. A between-class overlapping filter-based method for transcriptome data analysis

Sharma A, Imoto S, Miyano S

Feature selection algorithms play a crucial role in identifying and discovering important genes for cancer classification. Feature selection algorithms can be broadly categorized into two main groups: filter-based methods and wrapper-based methods. Filter-based methods have been quite popular in the literature due to their many advantages, including computational efficiency, simplistic architecture, and an intuitively simple means of discovering biological and clinical aspects. However, these methods have limitations, and the classification accuracy of the selected genes is less accurate. In this paper, we propose a set of univariate filter-based methods using a between-class overlapping criterion. The proposed techniques have been compared with many other univariate filter-based methods using an acute leukemia dataset. The following properties have been examined: classification accuracy of the selected individual genes and the gene subsets; redundancy check among selected genes using ridge regression and LASSO methods; similarity and sensitivity analyses; functional analysis; and, stability analysis. A comprehensive experiment shows promising results for our proposed techniques. The

univariate filter based methods using between-class overlapping criterion are accurate and robust, have biological significance, and are computationally efficient and easy to implement. Therefore, they are well suited for biological and clinical discoveries.

d. Forecasting Japan's physician shortage in 2035 as the first full-fledged aged society

Yuji K[19], Imoto S, Yamaguchi R, Matsumura T[20], Murashige N[20], Kodama Y[20], Miyano S, Imai K[19], Kami M[20]: [19]Department of Internal Medicine, Research Hospital [20]Division of Social Communication System for Advanced Clinical Research

Japan is rapidly becoming a full-fledged aged society, and physician shortage is a significant concern. The Japanese government has increased the number of medical school enrollments since 2008, but some researchers warn that this increase could lead to physician surplus in the future. It is unknown how many physicians will be required to accommodate future healthcare needs. We simulated changes in age/sex composition of the population, fatalities (the number of fatalities for the consecutive five years), and number of physicians from 2010 to 2035. Two indicators were defined: fatalities per physician and fatalities by physician working hour, based on the data of the working hours of physicians for each tuple of sex and age groups. We estimated the necessary number of physicians in 2035 and the number of new physicians to maintain the indicator levels in 2010. The number of physicians per 1,000 population is predicted to rise from 2·00 in 2010 to 3·14 in 2035. The number of physicians aged 60 years or older is expected to increase from 55,375 (20% of physicians) to 141,711 (36%). In 2010 and 2035, fatalities per physician were 23·1 and 24·0 for the total population, and 13·9 and 19·2 for 75 years or older, respectively. Fatalities per physician working hour are predicted to rise from 0·128 to 0·138. If working hours are limited to 48 hours per week in 2035, the number of fatalities per physician working hour is expected to be 0·196, and the number of new physicians must be increased by 53% over the current pace. The number of physicians per population continues to rise, but the estimated supply will not fulfill the demand for healthcare in the aging society. Strategies to increase the number of physicians and improve working conditions are urgently needed.

e. Does Twitter trigger bursts in signature collections?

Yamaguchi R, Imoto S, Kami M[20], Watanabe K[21], Miyano S, Yuji K[19]

The quantification of social media impacts on societal and political events is a difficult undertaking. The Japanese Society of Oriental Medicine started a signature-collecting campaign to oppose a medical policy of the Government Revitalization Unit to exclude a traditional Japanese medicine, "Kampo," from the public insurance system. The signature count showed a series of aberrant bursts from November 26 to 29, 2009. In the same interval, the number of messages on Twitter including the keywords "Signature" and "Kampo," increased abruptly. Moreover, the number of messages on an Internet forum that discussed the policy and called for signatures showed a train of spikes. Methods and Findings: In order to estimate the contributions of social media, we developed a statistical model with state-space modeling framework that distinguishes the contributions of multiple social media in time-series of collected public opinions. We applied the model to the time-series of signature counts of the campaign and quantified contributions of two social media, i.e., Twitter and an Internet forum, by the estimation. We found that a considerable portion (78%) of the signatures was affected from either of the social media throughout the campaign and the Twitter effect (26%) was smaller than the Forum effect (52%) in total, although Twitter probably triggered the initial two bursts of signatures. Comparisons of the estimated profiles of the both effects suggested distinctions between the social media in terms of sustainable impact of messages or tweets. Twitter shows messages on various topics on a time-line; newer messages push out older ones. Twitter may diminish the impact of messages that are tweeted intermittently. The quantification of social media impacts is beneficial to better understand people's tendency and may promote developing strategies to engage public opinions effectively. Our proposed method is a promising tool to explore information hidden in social phenomena.

f. Connection between traditional medicine and disease

Katayama K, Yamaguchi R, Imoto S, Matsuura K[21], Watanabe K[21], Miyano S: [21]Center for Kampo Medicine, Keio University School of Medicine

In Japanese traditional medicine, "Monshin" plays an important role. "Monshin" is a questionnaire that asked the patient's lifestyle and subjective symptoms. Specialists decide traditional herbal medicine by using of "Monshin". In this research, we connect "Monshin" to disease through building the Network.

g. A microarray analysis of gnotobiotic mice indicating that microbial exposure during the neonatal period plays an essential role in immune system development

Yamamoto M[21], Yamaguchi R, Muanakata K[21], Takashima K[21], Nishiyama M[21], Hioki K[21], Ohnishi Y[21], Nagasaki M, Imoto S, Miyano S, Ishige A[21], Watanabe K[21]

Epidemiological studies have suggested that the encounter with commensal microorganisms during the neonatal period is essential for normal development of the host immune system. Basic research involving gnotobiotic mice has demonstrated that colonization at the age of 5 weeks is too late to reconstitute normal immune function. In this study, we examined the transcriptome profiles of the large intestine (LI), small intestine (SI), liver (LIV), and spleen (SPL) of 3 bacterial colonization models−specific pathogen-free mice (SPF), ex-germ-free mice with bacterial reconstitution at the time of delivery (0WexGF), and ex-germ-free mice with bacterial reconstitution at 5 weeks of age (5WexGF)−and compared them with those of germ-free (GF) mice. Hundreds of genes were affected in all tissues in each of the colonized models; however, a gene set enrichment analysis method, MetaGene Profiler (MGP), demonstrated that the specific changes of Gene Ontology (GO) categories occurred predominantly in 0WexGF LI, SPF SI, and 5WexGF SPL, respectively. MGP analysis on signal pathways revealed prominent changes in toll-like receptor (TLR)- and type 1 interferon (IFN)-signaling in LI of 0WexGF and SPF mice, but not 5WexGF mice, while 5WexGF mice showed specific changes in chemokine signaling. RT-PCR analysis of TLR-related genes showed that the expression of interferon regulatory factor 3 (Irf3), a crucial rate-limiting transcription factor in the induction of type 1 IFN, prominently decreased in 0WexGF and SPF mice but not in 5WexGF and GF mice. The present study provides important new information regarding the molecular mechanisms of the so-called "hygiene hypothesis".

h. ChopSticks: High-resolution analysis of homozygous deletions by exploiting concordant read pairs

Yasuda T, Suzuki S, Nagasaki M, Miyano S

Structural variations (SVs) in genomes are commonly observed even in healthy individuals and play key roles in biological functions. To understand their functional impact or to infer molecular mechanisms of SVs, they have to be characterized with the maximum resolution. However, high-resolution analysis is a difficult task because it requires investigation of the complex structures involved in an enormous number of alignments of next-generation sequencing (NGS) reads and genome sequences that contain errors. We propose a new method called ChopSticks that improves the resolution of SV detection for homozygous deletions even when the depth of coverage is low. Conventional methods based on read pairs use only discordant pairs to localize the positions of deletions, where a discordant pair is a read pair whose alignment has an aberrant strand or distance. In contrast, our method exploits concordant reads as well. We theoretically proved that when the depth of coverage approaches zero or infinity, the expected resolution of our method is asymptotically equal to that of methods based only on discordant pairs under double coverage. To confirm the effectiveness of ChopSticks, we conducted computational experiments against both simulated NGS reads and real NGS sequences. The resolution of deletion calls by other methods was significantly improved, thus demonstrating the usefulness of ChopSticks. ChopSticks can generate high-resolution deletion calls of homozygous deletions using information independent of other methods, and it is therefore useful to examine the functional impact of SVs or to infer SV generation mechanisms.

Publications

1. C Affara M, Sanders D, Araki H, Tamada Y, Dunmore B, Humphreys S, Imoto S, Savoie C, Miyano S, Kuhara S, Print C, Charnock-Jones DS. Vasohibin-1 is identified as a master-regulator of endothelial cell apoptosis using gene network analysis. BMC Genomics. 14(1): 23, 2013.
2. Bowe A, Onoder T, Sadakane K, Shibuya T. Succinct de Bruijn graphs. The 12th Workshop on Algorithms in Bioinformatics, Lecture Notes in Computer Science. 7534: 225-235, 2012.
3. Chalkidis G, Tremmel G, Ray W, Bartlett C, Miyano, S, Nagasaki M. Reverse engineering complex disease networks by information flow. IEEE Biomedical Engineering. In press.
4. Fujimori S, Hino K, Saito A, Miyano S, Miyamoto-Sato E. PRD: A protein-RNA interaction database. Bioinformation. 8(15): 729-730, 2012.
5. Fujimori S, Hirai N, Masuoka K, Oshikubo T, Yamashita T, Washio T, Saito A, Nagasaki M, Miyano S, Miyamoto-Sato E. IRView: a data-

base and viewer for protein interacting regions. Bioinformatics. 28(14): 1949-1950, 2012.
6. Fujimoto A, Totoki Y, Abe T, Boroevich KA, Hosoda F, Hai Nguyen H, Aoki M, Hosono N, Kubo M, Miya F, Arai Y, Takahashi H, Shirakihara T, Nagasaki M, Shibuya T, Nakano K, Watanabe-Makino K, Tanaka H, Nakamura H, Kusuda J, Ojima H, Shimada K, Okusaka T, Ueno M, Shigekawa Y, Kawakami Y, Arihiro K, Ohdan H, Gotoh K, Ishikawa O, Ariizumi S, Yamamoto M, Yamada T, Chayama K, Kosuge T, Yamaue H, Kamatani N, Miyano S, Nakagama H, Nakamura Y, Tsunoda T, Shibata T, Nakagawa H. Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. Nature Genetics. 44(7): 760-764, 2012.
7. Fujita A, Severino P, Kojima K, Sato JR, Patriota AG, Miyano S. Functional clustering of time series gene expression data by Granger causality. BMC Systems Biology. BMC Systems Biology. 6: 137, 2012.
8. Hurley D, Araki H, Tamada Y, Dunmore B, Sanders D, Humphreys S, Affara M, Imoto S, Yasuda K, Tomiyasu Y, Tashiro K, Savoie C, Cho V, Smith S, Kuhara S, Miyano S, Charnock-Jones DS, Crampin EJ, Print CG. Gene network inference and visualization tools for biologists: application to new human transcriptome datasets. Nucleic Acids Res. 40(6): 2377-2398, 2012.
9. Ishimaru S, Mimori K, Yamamoto K, Inoue H, Imoto S, Kawano S, Yamaguchi R, Sato T, Toh H, Iinuma H, Maeda T, Ishii H, Suzuki S, Tokudome S, Watanabe M, Tanaka JI, Kudo SE, Sugihara KI, Hase K, Mochizuki H, Kusunoki M, Yamada K, Shimada Y, Moriya Y, Barnard GF, Miyano S, Mori M. Increased risk for CRC in diabetic patients with the nonrisk allele of SNPs at 8q24. Ann Surg Oncol. 19(9): 2853-2858, 2012.
10. Kawano S, Shimamura T, Niida A, Imoto S, Yamaguchi R, Nagasaki M, Yoshida R, Print C, Miyano S. Identifying Gene pathways associated with cancer characteristics via sparse statistical methods. IEEE/ACM Transactions on Computational Biology and Bioinformatics. 9(4): 966-972, 2012.
11. Katayama K, Yamaguchi R, Imoto S, Matsuura K, Watanabe K, Miyano S. Analysis of questionnaire for Traditional Medical and develop decision support system. Proc. 2012 International Workshop on Biomedical and Health Informatics. IEEE Computer Society Press. In press.
12. Katayama K, Yamaguchi R, Imoto S, Matsuura K, Watanabe K, Miyano S. Symbolic hierarchical clustering for pain vector. Intelligent Decision Technologies. 16: 17-124, 2012.
13. Katayama K, Yamaguchi R, Imoto S, Matsuura K, Watanabe K, Miyano S. Connection between traditional medicine and disease. ACM SIGHIT Record. 2(1): 21-21. 2012.
14. Kojima K, Imoto S, Yamaguchi R, Fujita A, Yamauchi M, Gotoh N, Miyano S. Identifying regulational alterations in gene regulatory networks by state space representation of vector autoregressive models and variational annealing. BMC Genomics. 13 (Suppl 1): S6, 2012.
15. Komatsu M, Yoshimaru T, Matsuo T, Kiyotani K, Miyoshi Y, Tanahashi T, Rokutan K, Yamaguchi R, Saito A, Imoto S, Miyano S, Nakamura Y, Sasa M, Shimada M, Katagiri T. Molecular features of triple negative breast cancer cells by genome-wide gene expression profiling analysis. Int J Oncol. 42(2): 478-506, 2013.
16. Kunishima S, Okuno Y, Yoshida K, Shiraishi Y, Sanada M, Muramatsu H, Chiba K, Tanaka H, Miyazaki K, Sakai M, Ohtake M, Kobayashi R, Iguchi A, Takahashi Y, Miyano S, Saito H, Kojima S, Ogawa S. ACTN1 is a novel causative gene for congenital macrothrombocytopenia. American Journal of Human Genetics. In press.
17. Mimura I, Nangaku M, Kanki Y, Tsutsumi S, Inoue T, Kohro T, Yamamoto S, Fujita T, Shimamura T, Suehiro J, Taguchi A, Kobayashi M, Tanimura K, Inagaki T, Tanaka T, Hamakubo T, Sakai J, Aburatani H, Kodama T, Wada Y. Dynamic change of chromatin conformation in response to hypoxia enhances the expression of GLUT3 (SLC2A3) by cooperative interaction of hypoxia-inducible factor 1 and KDM3A. Mol Cell Biol. 32(15): 3018-32, 2012.
18. Nagasaki M, Fujita A, Sekiya Y, Saito A, Ikeda E, Li C, Miyano S. XiP: a computational environment to create, extend and share workflows. Bioinformatics. 29(1): 137-139, 2013.
19. Niida A, Imoto S, Shimamura T, Miyano S. Statistical model-based testing to evaluate the recurrence of genomic aberrations. Bioinformatics. 28(12): i115-i120, 2012.
20. Ogami K, Yamaguchi R, Imoto S, Tamada Y, Araki H, Print C, Miyano S. Computational gene network analysis reveals TNF-induced angiogenesis. BMC Systems Biology. 6 (Suppl 2): S 12, 2012.
21. Okayama H, Kohno T, Ishii1 Y, Shimada Y, Shiraishi K, Iwakawa R, Furuta K, Tsuta K, Shibata T, Yamamoto S, Watanabe S, Sakamoto H, Kumamoto K, Takenoshita S, Gotoh N, Mizuno H, Sarai A, Kawano S, Yamaguchi R, Miyano S, Yokota J. Identification of genes up-regulated in ALK-positive and EGFR/KRAS/ALK-negative lung adenocarcinomas. Cancer Res. 72(1): 100-111, 2012.
22. Onuki R, Yamada R, Yamaguchi R, Kanehisa M, Shibuya T. Population model-based interdiplotype similarity measure for accurate diplotype clustering. J Comp Biol. 19(1): 55-67, 2012.

23. Sahli M, Shibuya T. Arapan-S: A Fast and Highly Accurate Whole-Genome Assembly Software for Viruses and Small Genomes. BMC Research Notes. 5: 243, 2012.
24. Sahli M, Shibuya T. Max-Shift BM and Max-Shift Horspool: practical fast exact string matching algorithms. J Information Processing. 20(2): 419-425, 2012.
25. Sahli M, Shibuya T. Argan-an artificial sequencing tool for simulated data and experimental. Proc. The 4th International Conference on Bioinformatics and Biomedical Technology (ICBBT 2012). 29: 196-199, 2012.
26. Sahli M, Shibuya T. An algorithm for classifying DNA reads. Proc. International Conference on Bioscience, Biochemistry and Bioinformatics (ICBBB 2012). 31: 59-63, 2012.
27. Sahli M, Shibuya T. Qamar-A more accurate DNA sequencing error correcting algorithm. Proc. International Conference on Bioscience, Biochemistry and Bioinformatics (ICBBB 2012). 31: 53-58, 2012.
28. Saito MM, Imoto S, Yamaguchi R, Miyano S, Higuchi T. Identifiability of local transmissibility parameters in agent-based pandemic simulation. Proc. 15th International Conference on Information Fusion. IEEE Computer Society Press. 2466-2471, 2012.
29. Saito MM, Imoto S, Yamaguchi R, Miyano S, Higuchi T. Parallel agent-based simulator for influenza pandemic. Lecture Notes in Computer Science. 7068: 361-370, 2012.
30. Sharma A, Imoto S, Miyano S. A filter based feature selection algorithm using null space of covariance matrix for DNA microarray gene expression data. Current Bioinformatics. 7(3): 289-294, 2012.
31. Sharma A, Imoto S, Miyano S.A between-class overlapping filter-based method for transcriptome data analysis. J Bioinformatics and Computational Biology. 10(5): 1250010, 2012.
32. Sharma A, Imoto S, Miyano S. A top-r feature selection algorithm for microarray gene expression data. IEEE/ACM Transactions on Computational Biology and Bioinformatics. 9(3): 754-64, 2012.
33. Sharma A, Imoto S, Miyano S, Sharma V. Null space based feature selection method for gene expression data, International Journal of Machine Learning and Cybernetics. 3(4): 269-276, 2012.
34. Sharma A, Paliwal KK, Imoto S, Miyano S. Principal component analysis using QR decomposition. International Journal of Machine Learning and Cybernetics. In press.
35. Shiraishi Y, Sato Y, Chiba K, Okuno Y, Nagata Y, Yoshida K, Shiba N, Hayashi Y, Kume H, Homma Y, Sanada M, Ogawa S, Miyano S. An empirical Bayesian framework for somatic mutation detection from cancer. Nucleic Acids Res. In press.
36. Takatsuno Y, Mimori K, Yamamoto K, Sato T, Niida A, Inoue H, Imoto S, Kawano S, Yamaguchi R, Toh H, Iinuma H, Ishimaru S, Ishii H, Suzuki S, Tokudome S, Watanabe M, Tanaka JI, Kudo SE, Mochizuki H, Kusunoki M, Yamada K, Shimada Y, Moriya Y, Miyano S, Sugihara K, Mori M. The rs6983267 SNP is associated with MYC transcription efficiency, which promotes progression and worsens prognosis of colorectal cancer. Ann Surg Oncol. In press.
37. Tamura T, Sone M, Nakamura Y, Shimamura T, Imoto S, Miyano S, Okazawa H.A restricted level of PQBP1 is needed for the best longevity of Drosophila. Neurobiol Aging. 34(1): 356.e11-20. 2013.
38. Terashi G, Shibuya T, Takeda-Shitaka M. LB3D: a protein 3D substructure search program based on the lower bound of a RMSD value. J Comp Biol. 19(5): 493-503, 2012.
39. Wang L, Hurley D, Watkins W, Araki H, Tamada Y, Muthukaruppan A, Ranjard L, Derkac E, Imoto S, Miyano S, Crampin E, Print C. Cell cycle gene networks are associated with melanoma prognosis. PLoS One. 7(4): e34247, 2012.
40. Yamaguchi R, Imoto S, Kami M, Watanabe K, Miyano S, Yuji K. Does Twitter trigger bursts in signature collections? PLoS One. In press.
41. Yamamoto M, Yamaguchi R, Muanakata K, Takashima K, Nishiyama M, Hioki K, Ohnishi Y, Nagasaki M, Imoto S, Miyano S, Ishige A, Watanabe K. A microarray analysis of gnotobiotic mice indicating that microbial exposure during the neonatal period plays an essential role in immune system development. BMC Genomics. 13: 335, 2012.
42. Yamauchi M, Yamaguchi R, Nakata A, Kohno T, Nagasaki M, Shimamura T, Imoto S, Saito A, Ueno K, Hatanaka Y, Yoshida R, Higuchi T, Nomura M, Beer DG, Yokota J, Miyano S, Gotoh N. Epidermal growth factor receptor tyrosine kinase defines critical prognostic genes of stage I lung adenocarcinoma. PLoS One. 2012; 7 (9): e43923.
43. Yasuda T, Suzuki S, Nagasaki M, Miyano S. ChopSticks: High-resolution analysis of homozygous deletions by exploiting concordant read pairs. BMC Bioinformatics. 13(1): 279, 2012.
44. Yokobori T, Iinuma H, Shimamura T, Imoto S, Sugimachi K, Ishii H, Iwatsuki M, Ota D, Ohkuma M, Iwaya T, Nishida N, Kogo R, Sudo T, Tanaka F, Shibata K, Toh H, Sato T, Barnard GF, Fukagawa T, Yamamoto S, Nakanishi H, Sasaki S, Miyano S, Watanabe T, Kuwano H, Mimori K, Pantel K, Mori M. Plastin3 is a novel marker for circulating tumor cells undergoing

the epithelial-mesenchymal transition and is associated with colorectal cancer prognosis. Cancer Res. In press.
45. Yuji K, Imoto S, Yamaguchi R, Matsumura T, Murashige N, Kodama Y, Miyano S, Imai K, Kami M. Forecasting Japan's physician shortage in 2035 as the first full-fledged aged society. PLoS One. 7(11): e50410, 2012.

Human Genome Center

Laboratory of Molecular Medicine
Laboratory of Genome Technology
ゲノムシークエンス解析分野
シークエンス技術開発分野

Professor	Yusuke Nakamura, M.D., Ph.D.	教　授	中　村　祐　輔
Associate Professor	Koichi Matsuda, M.D., Ph.D.	准教授	松　田　浩　一
Assistant Professor	Ryuji Hamamoto, Ph.D.	助　教	浜　本　隆　二
Assistant Professor	Hitoshi Zembutsu, M.D., Ph.D.	助　教	前　佛　　均
Assistant Professor	Chizu Tanikawa, Ph.D.	助　教	谷　川　千　津

The major goal of our group is to identify genes of medical importance, and to develop new diagnostic and therapeutic tools. We have been attempting to isolate genes involving in carcinogenesis and also those causing or predisposing to various diseases as well as those related to drug efficacies and adverse reactions. By means of technologies developed through the genome project including a high-resolution SNP map, a large-scale DNA sequencing, and the cDNA microarray method, we have isolated a number of biologically and/or medically important genes, and are developing novel diagnostic and therapeutic tools.

1. Genes playing significant roles in human cancer

Yusuke Nakamura, Koichi Matsuda, Ryuji Hamamoto, Hitoshi Zembutsu, Chizu Tanikawa, Yuji Urabe, Jiaying Lin, Zhenzhong Deng, Paulisally Hau Yi Lo, Yusei Funauchi, Yousef Salama Mahmmoud, Masami Tanaka, Mitsuko Nakashima, Hyun-Soo Cho, Reem Abdelrahim Ibrahim, Kang Daechun, Su-Youn Chung, Osman W Mohammed, Takashi Fujitomo, Seham Elgazzar,

(1) Epigenetics

Regulation of histone modification and chromatin structure by the p53-PADI4 pathway

Histone proteins are modified in response to various external signals, however their mechanisms are still not fully understood. Citrullination is a post-transcriptional modification which converts arginine in protein into citrulline. Here we show *in vivo* and *in vitro* citrullination of arginine 3 residue of histone H4 (cit-H4R3) in response to DNA damage through the p53-PADI4 pathway. We also observed DNA damage-induced citrullination of Lamin C. Cit-H4R3 and citrullinated Lamin C are located around fragmented nuclei in apoptotic cells. Ectopic expression of PADI4 led to chromatin decondensation and promoted DNA cleavage, while *Padi4*$^{-/-}$ mice exhibited resistance to radiation-induced apoptosis in the thymus. Furthermore, the level of cit-H4R3 was negatively correlated with p 53 protein expression and with tumor size in non-small cell lung cancer tissues. Our findings reveal that cit-H4R3 would be an "apoptotic histone code" to detect damaged cells and induce nuclear fragmentation, which plays a crucial role in carcinogenesis.

Enhanced HSP70 lysine methylation promotes proliferation of cancer cells through activation of Aurora kinase B.

Although heat-shock protein 70 (HSP70), an evolutionarily highly conserved molecular chaperone, is known to be post-translationally modified in various ways such as phosphorylation, ubiquitination and glycosylation, physiological significance of lysine methylation has never been elucidated. Here we identify dimethylation of HSP70 at Lys-561 by SETD1A. Enhanced HSP70 methylation was detected in various types of human cancer by immunohistochemical analysis, although the methylation was barely detectable in corresponding non-neoplastic tissues. Interestingly, methylated HSP70 predominantly localizes to the nucleus of cancer cells, whereas most of the HSP70 protein locates to the cytoplasm. Nuclear HSP70 directly interacts with Aurora kinase B (AURKB) in a methylation-dependent manner and promotes AURKB activity in vitro and in vivo. We also find that methylated HSP70 has a growth-promoting effect in cancer cells. Our findings demonstrate a crucial role of HSP70 methylation in human carcinogenesis..

Histone lysine methyltransferase SETD8 promotes carcinogenesis by deregulating PCNA expression.

Although the physiologic significance of lysine methylation of histones is well known, whether lysine methylation plays a role in the regulation of nonhistone proteins has not yet been examined. The histone lysine methyltransferase SETD8 is overexpressed in various types of cancer and seems to play a crucial role in S-phase progression. Here, we show that SETD8 regulates the function of proliferating cell nuclear antigen (PCNA) protein through lysine methylation. We found that SETD8 methylated PCNA on lysine 248, and either depletion of SETD8 or substitution of lysine 248 destabilized PCNA expression. Mechanistically, lysine methylation significantly enhanced the interaction between PCNA and the flap endonuclease FEN1. Loss of PCNA methylation retarded the maturation of Okazaki fragments, slowed DNA replication, and induced DNA damage, and cells expressing a methylation-inactive PCNA mutant were more susceptible to DNA damage. An increase of methylated PCNA was found in cancer cells, and the expression levels of SETD8 and PCNA were correlated in cancer tissue samples. Together, our findings reveal a function for lysine methylation on a nonhistone protein and suggest that aberrant lysine methylation of PCNA may play a role in human carcinogenesis.

(2) Lung cancer

Critical function for nuclear envelope protein TMEM209 in human pulmonary carcinogenesis.

Therapeutic targets for more effective and less toxic treatments of lung cancer remain important. Here we report the identification of the integral nuclear envelope protein TMEM209 as a critical driver of human lung cancer growth and survival. TMEM 209 expression was normally limited to testis, but we found that it was widely expressed in lung cancer, in which it localized to the nuclear envelope, Golgi apparatus, and the cytoplasm of lung cancer cells. Ectopic overexpression of TMEM209 promoted cell growth, whereas TMEM209 attenuation was sufficient to block growth. Mass spectrometric analysis identified the nucleoporin protein NUP205 as a TMEM209-interacting protein, stabilizing NUP 205 and increasing the level of c-Myc in the nucleus. Taken together, our findings indicate that TMEM209 overexpression and TMEM209-NUP205 interaction are critical drivers of lung cancer proliferation, suggesting a promising new target for lung cancer therapy.

(3) Breast cancer

Development of an orally-administrative MELK-targeting inhibitor that suppresses the growth of various types of human cancer.

We previously reported MELK (maternal embryonic leucine zipper kinase) as a novel therapeutic target for breast cancer. MELK was also reported to be highly upregulated in multiple types of human cancer. It was implied to play indispensable roles in cancer cell survival and indicated its involvement in the maintenance of tumor-initiating cells. We conducted a high-throughput screening of a compound library followed by structure-activity relationship studies, and successfully obtained a highly potent MELK inhibitor OTSSP167 with IC50 of 0.41 nM. OTSSP167 inhibited the phosphorylation of PSMA1 (proteasome subunit alpha type 1) and DBNL (drebrin-like), which we identified as novel MELK substrates and are important for stem-cell characteristics and invasiveness. The compound suppressed mammosphere formation of breast cancer cells and exhibited significant tumor growth suppression in xenograft studies using breast, lung, prostate, and pancreas cancer cell lines in mice by both intravenous and oral administration. This MELK inhibitor should be a promising compound possibly to suppress the growth of tumor-initiating cells and be applied for treatment of a wide range of human cancer.

(4) Prostate cancer

CLCA2 as a p53-inducible senescence mediator.

p53 is a tumor suppressor gene that is frequently mutated in multiple cancer tissues. Activated p53 protein regulates its downstream genes and subsequently inhibits malignant transformation by inducing cell cycle arrest, apoptosis, DNA repair, and senescence. However, genes involved in the p53-mediated senescence pathway are not yet fully elucidated. Through the screening of two genome-wide expression profile data sets, one for cells in which exogenous p53 was introduced and the other for senescent fibroblasts, we have identified chloride channel accessory 2 (CLCA2) as a p53-inducible senescence-associated gene. CLCA2 was remarkably induced by replicative senescence as well as oxidative stress in a p53-dependent manner. We also found that ectopically expressed CLCA2 induced cellular senescence, and the down-regulation of CLCA2 by small interfering RNA caused inhibition of oxidative stress-induced senescence. Interestingly, the reduced expression of CLCA2 was frequently observed in various kinds of cancers including prostate cancer, whereas its expression was not affected in precancerous prostatic intraepithelial neoplasia. Thus, our findings suggest a crucial role of p53/CLCA2-mediated senescence induction as a barrier for malignant transformation.

(5) Liver cancer

Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators.

Hepatocellular carcinoma (HCC) is the third leading cause of cancer-related death worldwide. We sequenced and analyzed the whole genomes of 27 HCCs, 25 of which were associated with hepatitis B or C virus infections, including two sets of multicentric tumors. Although no common somatic mutations were identified in the multicentric tumor pairs, their whole-genome substitution patterns were similar, suggesting that these tumors developed from independent mutations, although their shared etiological backgrounds may have strongly influenced their somatic mutation patterns. Statistical and functional analyses yielded a list of recurrently mutated genes. Multiple chromatin regulators, including ARID1A, ARID1B, ARID2, MLL and MLL3, were mutated in ~50% of the tumors. Hepatitis B virus genome integration in the TERT locus was frequently observed in a high clonal proportion. Our whole-genome sequencing analysis of HCCs identified the influence of etiological background on somatic mutation patterns and subsequent carcinogenesis, as well as recurrent mutations in chromatin regulators in HCCs

2. Pharmacogenetics

A genome-wide association study identifies locus at 10q22 associated with clinical outcomes of adjuvant tamoxifen therapy for breast cancer patients in Japanese.

Although many association studies of polymorphisms in candidate genes with the clinical outcomes of breast cancer patients receiving adjuvant tamoxifen therapy have been reported, genetic factors determining individual response to tamoxifen are not fully understood. To identify genetic polymorphisms associated with clinical outcomes of patients with tamoxifen treatment, we conducted a genome-wide association study (GWAS). We studied 462 Japanese patients with hormone receptor-positive, invasive breast cancer receiving adjuvant tamoxifen therapy. Of them, 240 patients were analyzed by genome-wide genotyping using the Illumina Human610-Quad BeadChips, and two independent sets of 105 and 117 cases were used for replication studies. In the GWAS, we detected significant associations with recurrence-free survival at 15 single-nucleotide polymorphisms (SNPs) on nine chromosomal loci (1p31, 1q41, 5q33, 7p11, 10q22, 12 q13, 13q22, 18q12 and 19p13) that satisfied a genome-wide significant threshold (log-rank $P = 2.87 \times 10(-9) - 9.41 \times 10(-8)$). Among them, rs 10509373 in C10orf11 gene on 10q22 was significantly associated with recurrence-free survival in the replication study (log-rank $P = 2.02 \times 10(-4)$) and a combined analysis indicated a strong association of this SNP with recurrence-free survival in breast cancer patients treated with tamoxifen (log-rank $P = 1.26 \times 10(-10)$). Hazard ratio per C allele of rs10509373 was 4.51 [95% confidence interval (CI), 2.72 − 7.51; $P = 6.29 \times 10(-9)$]. In a combined analysis of rs10509373 genotype with previously identified genetic makers, CYP2D6 and ABCC2, the number of risk alleles of these three genes had cumulative effects on recurrence-free survival among 345 patients receiving tamoxifen monotherapy (log-rank $P = 2.28 \times 10(-12)$). In conclusion, we identified a novel locus associated with recurrence-free survival in Japanese breast cancer patients receiving adjuvant tamoxifen therapy.

A genome-wide association study identifies four genetic markers for hematological toxicities in cancer patients receiving gemcitabine therapy.

Objective: Genetic factors are thought to be one of the causes of individual variability in the adverse reactions observed in cancer patients who received gemcitabine therapy. However, genetic factors determining the risk of adverse reactions of gemcitabine are not fully understood.

PATIENTS AND METHODS: To identify a genetic factor(s) determining the risk of gemcitabine-induced leukopenia/neutropenia, we conducted a

genome-wide association study, by genotyping over 610 000 single nucleotide polymorphisms (SNPs), and a replication study in a total of 174 patients, including 54 patients with at least grade 3 leukopenia/neutropenia and 120 patients without any toxicities.

RESULTS: We identified four loci possibly associated with gemcitabine-induced leukopenia/neutropenia [rs11141915 in DAPK1 on chromosome 9q21, combined $P = 1.27 \times 10$, odds ratio (OR) = 4.10; rs 1901440 on chromosome 2q12, combined $P = 3.11 \times 10$, OR = 34.00; rs12046844 in PDE4B on chromosome 1p31, combined $P = 4.56 \times 10$, OR = 4.13; rs 11719165 on chromosome 3q29, combined $P = 5.98 \times 10$, OR = 2.60]. When we examined the combined effects of these four SNPs, by classifying patients into four groups on the basis of the total number of risk genotypes of these four SNPs, significantly higher risks of gemcitabine-induced leukopenia/neutropenia were observed in the patients having two and three risk genotypes ($P = 6.25 \times 10$, OR = 11.97 and $P = 4.13 \times 10$, OR = 50.00, respectively) relative to patients with zero or one risk genotype.

CONCLUSION: We identified four novel SNPs associated with gemcitabine-induced severe leukopenia/neutropenia. These SNPs might be applicable in predicting the risk of hematological toxicity in patients receiving gemcitabine therapy

Impact of four loci on serum tamsulosin hydrochloride concentration.

Tamsulosin hydrochloride is one of the most potent drugs for treatment of benign prostatic hyperplasia (BPH), however, the efficacy of tamsulosin hydrochloride varies among individuals. In this study, we measured the maximum serum concentration (Cmax) of tamsulosin hydrochloride in 182 of BPH patients and found remarkable individual variability. To investigate the genetic factors that regulate pharmacokinetics of tamsulosin hydrochloride, we conducted a genome-wide association study in these 182 BPH patients. As a result, rs 16902947 on chromosome 5p13.2, rs7779057 on 7q 22.3, rs35681285 on 7p21.2 and rs2122469 on 8p21.3 indicated possible associations with Cmax of tamsulosin hydrochloride ($P = 1.29 \times 10(-7)$, $2.15 \times 10(-7)$, $4.35 \times 10(-7)$ and $7.03 \times 10(-7)$, respectively), although these single-nucleotide polymorphisms (SNPs) did not reach the genome-wide significance threshold after Bonferroni correction. As these associated SNPs showed additive effects on serum tamsulosin hydrochloride concentration, we defined the 'Cmax prediction index' based on genotypes of these SNPs. This index clearly associated with Cmax values ($P = 4.5 \times 10(-6)$), indicating the possible roles of these four variants in tamsulosin hydrochloride pharmacokinetics. Our findings would partially explain the variability of the response to the tamsulosin hydrochloride treatment.

3. Genome-wide association study

(1) cancer susceptibility gene

Common variants at 11q12, 10q26 and 3p11.2 are associated with prostate cancer susceptibility in Japanese.

We have previously reported multiple loci associated with prostate cancer susceptibility in a Japanese population using a genome-wide association study (GWAS). To identify additional prostate cancer susceptibility loci, we genotyped nine SNPs that were nominally associated with prostate cancer ($P < 1 \times 10(-4)$) in our previous GWAS in three independent studies of prostate cancer in Japanese men (2,557 individuals with prostate cancer (cases) and 3,003 controls). In a meta-analysis of our previous GWAS and the replication studies, which included a total of 7,141 prostate cancer cases and 11,804 controls from a single ancestry group, three new loci reached genome-wide significance on chromosomes 11q12 (rs1938781; $P = 1.10 \times 10(-10)$; FAM 111A-FAM111B), 10q26 (rs2252004; $P = 1.98 \times 10(-8)$) and 3p11.2 (rs2055109; $P = 3.94 \times 10(-8)$). We also found suggestive evidence of association at a previously reported prostate cancer susceptibility locus at 2p11 (rs2028898; $P = 1.08 \times 10(-7)$). The identification of three new susceptibility loci should provide additional insight into the pathogenesis of prostate cancer and emphasizes the importance of conducting GWAS in diverse populations.

A genome-wide association study identifies SNP in DCC is associated with gallbladder cancer in the Japanese population.

Gallbladder cancer (GC) is a relatively uncommon cancer with higher incidence in certain areas including Japan. Because of the difficulty in diagnosis, prognosis of GC is very poor. To identify genetic determinants of GC, we conducted a genome-wide association study (GWAS) in 41 GC patients and 866 controls. Association between each single-nucleotide polymorphism (SNP) with GC susceptibility was evaluated by multivariate logistic regression analysis conditioned on age and gender of subjects. SNPs that showed suggestive association ($P < 1 \times 10(-4)$) with GC were further examined in 30 cases and 898 controls. SNP rs7504990 in the DCC (deleted in colorectal cancer, 18q21.3) that encodes a netrin 1 receptor achieved a combined P-value of $7.46 \times 10(-8)$ (OR = 6.95; 95% CI = 3.43 − 14.08). Subsequent imputation analysis identified multiple SNPs with similarly strong associations in an adjacent genomic region, where loss of heterozygosity was reported in GC and other cancers. Re-

duced expression of DCC was indicated to be associated with the poorly differentiated histological type, increased proliferation and metastasis through loss of adhesiveness. However, due to the limited sample size investigated here, further replication study and functional analysis would be necessary to further confirm the result of the association.

(2) other diseases

A genome-wide association study identifies two susceptibility loci for duodenal ulcer in the Japanese population

Through a genome-wide association analysis using a total of 7,035 duodenal ulcer cases and 25,323 controls of Japanese populations, we identified two susceptibility loci at the *Prostate stem cell antigen* (*PSCA*) on 8q24 and at the *ABO blood group* (*ABO*) on 9q34. A C-allele of rs2294008 at *PSCA* increased the risk of duodenal ulcer (odds ratio (OR) of 1.84 with P value of 3.92×10^{-33}) in a recessive model, while it decreased the risk of gastric cancer (OR of 0.79 with P value of 6.79×10^{-12}) as reported previously[1]. A T-allele of SNP rs2294008 created the upstream translational initiation codon and affects the protein localization from cytoplasm to cell surface. SNP rs505922 on *ABO* also associated with duodenal ulcer in a recessive model (OR of 1.32 with P value of 1.15×10^{-10}). Our finding implies the crucial roles of genetic variations on the pathogenesis of duodenal ulcer.

Common variant near the endothelin receptor type A (EDNRA) gene is associated with intracranial aneurysm risk.

The pathogenesis of intracranial aneurysm (IA) formation and rupture is complex, with significant contribution from genetic factors. We previously reported genome-wide association studies based on European discovery and Japanese replication cohorts of 5,891 cases and 14,181 controls that identified five disease-related loci. These studies were based on testing replication of genomic regions that contained SNPs with posterior probability of association (PPA) greater than 0.5 in the discovery cohort. To identify additional IA risk loci, we pursued 14 loci with PPAs in the discovery cohort between 0.1 and 0.5. Twenty-five SNPs from these loci were genotyped using two independent Japanese cohorts, and the results from discovery and replication cohorts were combined by meta-analysis. The results demonstrated significant association of IA with rs 6841581 on chromosome 4q31.23, immediately 5' of the endothelin receptor type A with $P = 2.2 \times 10(-8)$ [odds ratio (OR) = 1.22, PPA = 0.986]. We also observed substantially increased evidence of association for two other regions on chromosomes 12q 22 (OR = 1.16, $P = 1.1 \times 10(-7)$, PPA = 0.934) and 20 p12.1 (OR = 1.20, $P = 6.9 \times 10(-7)$, PPA = 0.728). Although endothelin signaling has been hypothesized to play a role in various cardiovascular disorders for over two decades, our results are unique in providing genetic evidence for a significant association with IA and suggest that manipulation of the endothelin pathway may have important implications for the prevention and treatment of IA

Meta-analysis identifies multiple loci associated with kidney function-related traits in east Asian populations.

Chronic kidney disease (CKD), impairment of kidney function, is a serious public health problem, and the assessment of genetic factors influencing kidney function has substantial clinical relevance. Here, we report a meta-analysis of genome-wide association studies for kidney function-related traits, including 71,149 east Asian individuals from 18 studies in 11 population-, hospital- or family-based cohorts, conducted as part of the Asian Genetic Epidemiology Network (AGEN). Our meta-analysis identified 17 loci newly associated with kidney function-related traits, including the concentrations of blood urea nitrogen, uric acid and serum creatinine and estimated glomerular filtration rate based on serum creatinine levels (eGFRcrea) ($P < 5.0 \times 10(-8)$). We further examined these loci with in silico replication in individuals of European ancestry from the KidneyGen, CKDGen and GUGC consortia, including a combined total of ~110,347 individuals. We identify pleiotropic associations among these loci with kidney function-related traits and risk of CKD. These findings provide new insights into the genetics of kidney function.

References

1. N. Kumasaka, M. Aoki, Y. Okada, A. Takahashi, K. Ozak, T. Mushiroda, T. Hirota, M. Tamari, T. Tanaka, Y. Nakamura, N. Kamatani, and M. Kubo: Haplotypes with copy number and single nucleotide polymorphisms in CYP2A 6 locus are associated with smoking quantity in a Japanese population. Plos One, 7: e44507, 2012
2. H.-S. Cho, G. Toyokawa, Y. Daigo, S. Hayami, K. Masuda, N. Ikawa, Y. Yamane, K. Maejima, M. Yoshimatsu, T. Tsunoda, H.I. Field, J.D. Kelly, D.E. Neal, B.A.J. Ponder, Y. Maehara, Y. Nakamura, and R. Hamamoto: The JmjC domain-containing histone demethylase KDM3

A Is a positive regulator of the G1/S transition in cancer cells via transcriptional regulation of the HOXA1 gene. Int. J. Cancer, 131: 179-189, 2012

3. F. Innocenti, K. Owzar, N.L. Cox, P. Evans, M. Kubo, H. Zembutsu, C. Jiang, D. Hollis, T. Mushiroda, L. Li, P. Friedman, L. Wang, H. Hurwitz, K.M. Giacomini, H.L. McLeod, R.M. Goldberg, R.L. Schilsky, H.L. Kindler, Y. Nakamura, and M.J. Ratain: A genome-wide association study of overall survival in pancreatic cancer patients treated with gemcitabine in CALGB 80303. Clinical Cancer Research, 18: 577-584, 2012
4. M. Maimbo, K. Kiyotani, T. Mushiroda, C. Masimirembwa, and Y. Nakamura: CYP2B6 genotype is a strong predictor of systemic exposure to efavirenz in HIV-infected Zimbabweans. European Journal of Clinical Pharmacology, 68: 267-271, 2012
5. N. Akuta, F. Suzuki, M. Hirakawa, Y. Kawamura, H. Sezaki, Y. Suzuki, T. Hosaka, M. Kobayashi, M. Kobayashi, S. Saitoh, Y. Arase, K. Ikeda, K. Chayama, Y. Nakamura, and H. Kumada: Amino acid substitution in HCV core/NS 5A region and genetic variation near IL28B gene affect treatment efficacy to interferon plus ribavirin combination therapy. Intervirology, 55: 231-241, 2012
6. K. Kiyotani, T. Mushiroda, Y. Nakamura, and H. Zembutsu: Pharmacogenomics of tamoxifen: roles of drug metabolizing enzymes and transporters. Drug Metabolism and Pharmacokinetics, 27: 122-131, 2012
7. J. Koinuma, H. Akiyama, M. Fujita, M. Hosokawa, E. Tsuchiya, S. Kondo, Y. Nakamura, and Y. Daigo: Characterization of an Opa interacting protein 5 (OIP5) involved in lung and esophageal carcinogenesis. Cancer Science, 103: 577-586, 2012
8. K. Kiyotani, S. Uno, T. Mushiroda, A. Takahashi, M. Kubo, N. Mitsuhata, S. Ina, C. Kihara, Y. Kimura, H. Yamaue, K. Hirata, Y. Nakamura, and H. Zembutsu: Genome-wide association study identifies four genetic markers for hematological toxicities in cancer patients receiving gemcitabine therapy. Pharmacogenetics and Genomics, 22: 229-235, 2012
9. Y. Okada, M. Kubo, H. Ohmiya, A. Takahashi, N. Kumasaka, N. Hosono, S. Maeda, W. Wen, R. Dorajoo, M.-J. Go, W. Zheng, N. Kato, J.-Y. Wu, Q. Lu, the GIANT consortium, T. Tsunoda, K. Yamamoto, Y. Nakamura, N. Kamatani, and T. Tanaka: Common variants at CDKAL1 and KLF9 are associated with body mass index in East Asian populations. Nature Genetics, 44: 302-306, 2012
10. W. Wen, Y.S. Cho, W. Zheng, R. Dorajoo, Y. Nakamura, et al.: Meta-analysis of genome-wide association studies in East Asians identifies novel genetic variants associated with body mass index. Nature Genetics, 44: 307-311, 2012
11. R. Abo, S. Hebbring, Y. Ji, H. Zhu, Z.-B. Zeng, A. Batzler, G.D. Jenkins, J. Biernacka, K. Snyder, M. Drews, O. Fiehn, B. Fridley, D. Schaid, N. Kamatani, Y. Nakamura, M. Kubo, T. Mushiroda, R. Kaddurah-Daouk, D.A. Mrazek, and R.M. Weinshilboum: Merging pharmacometabolomics with pharmacogenomics using "1000 Genomes" SNP imputation: Selective serotonin reuptake inhibitor response pharmacogenomics. Pharmacogenetics and Genomics, 22: 247-253, 2012
12. H. Nakagawa, S. Akamatsu, R. Takata, A. Takahashi, M. Kubo, and Y. Nakamura: Prostate cancer genomics, biology, and risk assessment through genome-wide association studies. Cancer Science, 103: 607-613, 2012
13. K. Kiyotani, T. Mushiroda, T. Tsunoda, T. Morizono, N. Hosono, M. Kubo, Y. Tanigawara, C. K. Imamura, D.A. Flockhart, F. Aki, K. Hirata, Y. Takatsuka, M. Okazaki, S. Ohsumi, T. Yamakawa, M. Sasa, Y. Nakamura, and H. Zembutsu: A Genome-wide association study identifies locus at 10q22 associated with clinical outcomes of adjuvant tamoxifen therapy for breast cancer patients. Human Molecular Genetics, 21: 1665-1672, 2012
14. H. Yoshioka, S. Yamamoto, H. Hanaoka, Y. Iida, P. Paudya, T. Higuchi, H. Tominaga, N. Oriuchi, H. Nakagawa, Y. Shiba, K. Yoshida, R. Osawa, T. Katagiri, T. Tsunoda, Y. Nakamura, and K. Endo: In vivo therapeutic effect of CDH 3 / P-cadherin-targeting radioimmunotherapy. Cancer Immunology, Immunotherapy, 61: 1211? 1220, 2012
15. C. Tanikawa, Y. Urabe, K. Matsuo, M. Kubo, A. Takahashi, H. Ito, K. Tajima, N. Kamatani,Y. Nakamura, and K. Matsuda: Genome wide association study identified two susceptible loci for duodenal ulcer in Japanese population. Nature Genetics, 44: 430-441, 2012
16. Y. Urabe, C. Tanikawa, A. Takahashi, Y. Okada, T. Morizono, T. Tsunoda, N. Kamatani, K. Kohri, K. Chayama, M. Kubo, Y. Nakamura, and K. Matsuda: Genome-wide association study of nephrolithiasis in Japanese population identifies novel susceptible loci at 5q35.3, 7p14.3 and 13q 14.1. PLoS Genetics, 8: e1002541, 2012
17. S.-K. Low, A. Takahashi, P.-C. Cha, H. Zembutsu, N. Kamatani, M. Kubo, and Y. Nakamura: Genome-wide association study for intracranial aneurysm in Japanese population identifies three candidate susceptible loci and a functional genetic variant at EDNRA. Human Molecular Genetics, 21: 2102-2110, 2012
18. C. Tanikawa, M. Espinosa, A. Suzuki, K. Masuda, K. Yamamoto, E. Tsuchiya, K. Ueda,

Y. Daigo, Y. Nakamura, and K. Matsuda: Regulation of histone modification and chromatin structure by the p53-PADI4 pathway. Nature Communications,: DOI 10.1038, 2012

19. P.-C. Cha, H. Zembutsu, A. Takahashi, M. Kubo, N. Kamatani, and Y. Nakamura: A genome-wide association study (GWAS) identifies SNP in DCC is associated with gallbladder cancer (GC) in the Japanese population. Journal of Human Genetics, 57: 235-237, 2012
20. S. Akamatsu, R. Takata, C.A. Haiman, A. Takahashi, T. Inoue, M. Kubo, M. Furihata, N. Kamatani, J. Inazawa, G.K. Chen, L.L. Marchand, L.N. Kolonel, T. Katoh, Y. Yamano, M. Yamakado, H. Takahashi, H. Yamada, S. Egawa, T. Fujioka, B.E. Henderson, T. Habuchi, O. Ogawa, Y. Nakamura, and H. Nakagawa: Common variants at 11q12, 10q26 and 3p11.2 are associated with prostate cancer susceptibility in Japanese. Nature Genetics, 44: 426-429, 2012
21. W.-C. Chang, C.-H. Lee., T. Hirota, L.-F. Wang, S. Doi, A. Miyatake, T. Enomoto, K. Tomita, M. Sakashita, T. Yamada, S. Fujieda, K. Ebe, H. Saeki, S. Takeuchi, M. Furue, W.-C. Chen, Y.-C. Chiu, W.P. Chang, C.-H. Hong, E. Hsi, S.-H. H. Juo, H.-S. Yu, Y. Nakamura, and M. Tamari: ORAI1 genetic polymorphisms associated with the susceptibility of atopic dermatitis in Japanese andTaiwanese populations. PLoS ONE, 7: e29387, 2012
22. W. Osman, Y. Okada, Y. Kamatani, M. Kubo, K. Matsuda, and Y. Nakamura: Association of common variants in TNFRSF13B, TNFSF13, and ANXA3 with serum levels of non-albumin protein and immunoglobulin isotypes in the Japanese population. PLoS ONE, 7: e32683, 2012
23. H.H. Nguyen, R. Takata, S. Akamatsu, D. Shigemizu, T. Tsunoda, M. Furihata, A. Takahashi, M. Kubo, N. Kamatani, O. Ogawa, T. Fujioka, Y Nakamura, and H. Nakagawa: IRX4 at 5p15 suppresses prostate cancer growth through the interaction with vitamin D receptor, conferring prostate cancer susceptibility. Human Molecular Genetics, 21: 2076-2085, 2012
24. Y. Okada, C. Terao, K. Ikari, Y. Kochi, K. Ohmura, A. Suzuki, T. Kawaguchi, E. Stahl, F. Kurreman, N. Nishida, H. Ohmiya, K. Myouzen, M. Takahashi, T. Sawada, Y. Nishioka, M. Yukioka, T. Matsubara, S. Wakitani, R. Teshima, S. Tohma, K. Takasugi, K. Shimada, A. Murasawa, S. Honjo, K. Matsuo, H. Tanaka, K. Tajima, T. Suzuki, T. Iwamoto, Y. Kawamura, H. Tanii, Y. Okazaki, T. Sasaki, P.K. Gregersen, L. Padyukov, J. Worthington, K.A. Siminovitch, M. Lathrop, A. Taniguchi, A. Takahashi, K. Tokunaga, M. Kubo, Y. Nakamura, N. Kamatani, T. Mimori, R.M. Plenge, H. Yamanaka, S. Momohara, R. Yamada, F. Matsuda, and K. Yamamoto: Meta-analysis identifies nine new loci associated with rheumatoid arthritis in the Japanese population. Nature Genetics, 44: 511-516, 2012"
25. W. Jongjaroenpraser, T. Phusantisampan, S. Mahasirimongkol, T. Mushiroda, N. Hirankarn, T. Snabboon, S. Chanprasertyotin, P. Tantiwong, S. Soonthornpun, P. Rattanapichart, S. Mamanasiri, T. Himathongkam, B. Ongphiphadhanakul, A. Takahashi, N. Kamatani, M. Kubo, and Y. Nakamura: A genome-wide association study identifies novel susceptibility loci for thyrotoxic hypokalemic periodic paralysis. Journal of Human Genetics, 57: 301-304, 2012
26. M. Aoki, N. Hosono, S. Kakata, Y. Nakamura, N. Kamatani, and M. Kubo: New pharmacogenetic test for detecting HLA-A*31:01 allele using invaderPlus assay. Pharmacogenetics and Genomics, 22: 441-446, 2012
27. K. Kiyotani, T. Mushiroda, C.K. Imamura, Y. Tanigawara, N. Hosono, M. Kubo, M. Sasa, Y. Nakamura, and H. Zembutsu: Dose-adjustment study of tamoxifen based on CYP2D6 genotypes in Japanese breast cancer patients. Breast Cancer Res Treat, 131: 137-145, 2012
28. C. Tanikawa, H. Nakagawa, Y. Furukawa, Y. Nakamura, and K. Matsuda: Title CLCA2 as a p53-inducible senescence mediator. Neoplasia, 14: 141-147, 2012
29. Y. Yamaguchi-Kabata, T. Tsunoda, N. Kumasaka, A. Takahashi, N. Hosono, M. Kubo, Y. Nakamura, and N. Kamatani: Genetic differences in the two main groups of the Japanese population based on autosomal SNPs and haplotypes. Journal of Human Genetics, 57: 326-334, 2012
30. Y. Onouchi, K. Ozaki, J.C. Burns, C. Shimizu, M. Terai, H. Hamada, T. Honda, H. Suzuki, T. Suenaga, T. Takeuchi, N. Yoshikawa, Y. Suzuki K. Yasukawa, R. Ebata, K. Higashi, T. Saji, Y. Kemmotsu, S. Takatsuki, K. Ouchi, F. Kishi, T. Yoshikawa, T. Nagai, K Hamamoto, Y. Sata, A. Honda, H. Kobayashi, J. Sato, S. Shibuta, M. Miyawaki, K. Oishi, H. Yamaga, N. Aoyagi, S. Iwahashi, R. Miyashita, Y. Murata, K. Sasago, A. Takahashi, N. Kamatani, M. Kubo, T. Tsunoda, A. Hata, Y. Nakamura, T. Tanaka, and Japan Kawasaki Disease Genome Consortium & U.S. Kawasaki Disease Genetics Consortium: A genome-wide association study identifies three new loci for Kawasaki disease. Nature Genetics, 44: 517-521, 2012
31. W. Osman, S.-K. Low, A. Takahashi, M. Kubo, and Y. Nakamura: A Genome-wide association study in the Japanese population confirms 9p21 and 14q23 as susceptibility loci for primary open angle glaucoma. Human Molecular Genetics, 21: 2836-2842, 2012
32. S. Mahasirimongkol, H. Yanai, T. Mushiroda,

W. Promphittayarat, S. Wattanapokayakit, J. Promjai, R. Yuliwulandari, N. Wichukchinda, A. Yowang, N. Yamada, P. Kantipong, A. Takahashi, M. Kubo, P. Sawanpanyalert, N. Kamatani, Y. Nakamura, and K. Tokunaga: Genome-wide association studies of tuberculosis in Asians identify distinct at-risk locus for young tuberculosis. Journal of Human Genetics, 57: 363-367, 2012

33. M Liu, L. Wang, T. Bongartz, J.R. Hawse, S.N. Markovic, D.J. Schaid, T. Mushiroda, M. Kubo, Y. Nakamura, N. Kamatani, P.E. Goss, J.N. Ingle, and R.M. Weinshilboum: Aromatase inhibitors, estrogens and musculoskeletal pain: estrogen-dependent T-cell leukemia 1A (TCL1 A) gene-mediated regulation of cytokine expression. Breast Cancer Research, 9: R41, 2012
34. M. Imamura, S. Maeda, T. Yamauchi, K. Hara, K. Yasuda, T. Morizono, A. Takahashi, M. Horikoshi, M. Nakamura, H. Fujita, T. Tsunoda, M. Kubo, H. Watada, H. Maegawa, M. Okada-Iwabu, M. Iwabu, N. Shojima, T. Ohshige, S. Omori, M. Iwata, H. Hirose, K. Kaku, C. Ito, Y. Tanaka, K. Tobe, A. Kashiwagi, R. Kawamori, M. Kasuga, N. Kamatani, Diabetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium, Y. Nakamura, and T. Kadowaki: A single nucleotide polymorphism in ANK1 is associated with susceptibility to type 2 diabetes in Japanese populations. Human Molecular Genetics, 21: 3042-3049, 2012
35. P.T. Ellinor, K.L. Lunetta, C.M. Albert, N.L. Glazer, M.D. Ritchie, A.V. Smith, D.E. Arking, M. Müller-Nurasyid, B.P. Krijthe, S.A. Lubitz, J. C. Bis, M.K. Chung, K. Ozaki, J.D. Roberts, J.G. Smith, A. Pfeufer, M.F. Sinner, K. Lohman, J. Ding, N.L. Smith, J.D. Smith, M. Rienstra, K.M. Rice, D.R. Van Wagoner, J.W. Magnani, R. Wakili, S. Clauss, J.I. Rotter, G. Steinbeck, L.J. Launer, R.W. Davies, M. Borkovich, T.B. Harris, H. Lin, U. Völker, H. Völzke, D.J. Milan, A. Hofman, E. Boerwinkle, L.Y. Chen, E.Z. Soliman, B.F. Voight, G. Li, A. Chakravarti, M. Kubo, U. Tedrow, L.M. Rse, P.M. Ridker, D. Conen, T. Tsunoda, T. Furukawa, N. Sotoodehnia, S. Xu, N. Kamatani, D. Levy, Y. Nakamura, B. Parvez, S. Mahida, K.L. Furie, J. Rosand, R. Muhammad, B.M. Psaty, T. Meitinger, S. Perz, H.-E. Wichmann, J.C.M. Witteman, W.H.L. Kao, S. Kathiresan, D.M. Roden, A.G. Uitterlinden, F. Rivadeneira, B. McKnight, M. Sjögren, A.B. Newman, Y. Liu, M.H. Gollob, O. Melander, T. Tanaka, B.H.C. Stricker, S.B. Felix, A. Alonso, D. Darbar, J. Barnard, D.I. Chasman, S.R. Heckbert, E.J. Benjamin, V. Gudnason, and S. Kääb: Meta analysis identifies six new susceptibility loci for atrial fibrillation. Nature Genetics, 44: 670-688, 2012
36. R. Osawa, T. Tsunoda, S. Yoshimura, T. Watanabe, M. Miyazawa, M. Tani, K. Takeda, H. Nakagawa, Y. Nakamura, and H. Yamaue: Identification of HLA-A24-restricted novel T cell epitope peptides derived from P-cadherin and kinesin family member 20A. Journal of Biomedicine and Biotechnology,: doi:10.1155/2012/848042, 2012
37. M. Takawa, H.-S. Cho, S. Hayami, G. Toyokawa, M. Kogure, Y. Yamane, Y. Iwai, K. Maejima, K. Ueda, A. Masuda, N. Dohmae, H.I. Field, T. Tsunoda, T. Kobayashi, T. Akasu, M. Sugiyama, S. Ohnuma, Y. Atomi, B.A.J. Ponder, Y. Nakamura, and R. Hamamoto: Histone lysine methyltransferase SETD8 promotes carcinogenesis by deregulating PCNA expression. Cancer Research, 72: 3217-3227, 2012
38. H.-S. Cho, S. Hayami, G. Toyokawa, K. Maejima, Y. Yamane, T. Suzuki, N. Dohmae, M. Kogure, D. Kang, D.E. Neal, B.A.J. Ponder, H. Yamaue, Y. Nakamura, and R. Hamamoto: RB1 methylation by SMYD2 enhances cell cycle progression through an increase of RB1 phosphorylation. Neoplasia, 14: 476-, 2012
39. A. Fujimoto, Y. Totoki, T. Abe, K. A. Boroevich, F. Hosoda, H. H. Nguyen, M. Aoki, N. Hoshono, M. Kubo, F. Miya, Y. Arai, H. Takahashi, T. Shirakihara, M. Nagasaki, T. Shibuya, K. Nakano, K. Watanabe-Makino, H. Tanaka, H. Nakamura, J. Kusuda, H. Ojima, K. Shimada, T. Okusaka, M. Ueno, Y. Shigekawa, Y. Kawakami, K. Arihiro, H. Ohdan, K. Gotoh, O. Ishikawa, S. Ariizumi, M. Yamamoto, T. Yamada, K. Chayama, T. Kosuge, H. Yamaue, N. Kamatani, S. Miyano, H. Nakagama, Y. Nakamura, T. Tsunoda, T. Shibata, and H. Nakagawa: Whole genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. Nature Genetics, 44: 760-764, 2012
40. W. Obara, R. Ohsawa, M. Kanehira, R. Takata, T. Tsunoda, K. Yoshida, K. Takeda, T. Katagiri, Y. Nakamura, T. Fujioka: Cancer peptide vaccine therapy developed from ongoantigens that were identified through genome-wide expression profile analysis for bladder cancer Japanese Journal of Clinical Oncology, 42: 591-600, 2012
41. K. Okuno, F. Sugiura, K. Itoh, Y. Yoshida, T. Tsunoda, Y. Nakamura: Recent advances in active specific cancer vaccine threatment for colorectal cancer Current Pharmacology and Biotechnology, 13: 1439-1445, 2012
42. A. Toyama, A. Suzuki, T. Shimada, C. Aoki, Y. Aoki, Y. Umino, Y. Nakamura, D. Aoki, and TA. Sato: Proteomic characterization of ovarian cancers identifying annexin-A4, phosphoserine aminotransferase, cellular retinoic acid-binding protein 2, and serpin B5 as histology-specific biomarkers. Cancer Science, 103: 747-755, 2012

43. Y. Okada, K. Shimane, Y. Kochi, T. Tahira, A. Suzuki, K. Higasa, A. Takahashi, T. Horita, T. Atsumi, T. Ishii, A. Okamoto, K. Fujio, M. Hirakata, H. Amano, Y. Kondo, S. Ito, K. Takada, A. Mimori, K. Saito, M. Kamachi, Y. Kawaguchi, K. Ikari, OW Mohammed, K. Matsuda, C. Terao, K. Ohmura, K. Myouzen, N. Hosono, T. Tsunoda, N. Nishimoto, T. Mimori, F. Matsuda, Y. Tanaka, T. Sumida, H. Yamanaka, Y. Takasaki, T. Koike, T. Horiuchi, K. Hayashi, M. Kubo, N. Kamatani, R. Yamada, Y. Nakamura, and K. Yamamoto: A genome-wide association study identified AFF1 as a susceptibility locus for systemic lupus eyrthematosus in Japanese. PLoS Genetics, 8: e 1002455, 2012
44. H. Ochi, CN. Hayes, H. Abe, Y. Hayashida, T. Uchiyama, N. Kamatani, Y. Nakamura, and K. Chayama:. Toward the establishment of a prediction system for the personalized treatment of chronic hepatitis C. Journal of Infectious Disease, 205: 204-210, 2012
45. Y. Okada, X. Sim, M.J. Go, C.-H. Chen, D. Gu, F. Takeuchi, A. Takahashi, S. Maeda, T. Tsunoda, P. Chen, S.-C. Lim1, T.-Y Wong, J. Liu 1, T.L. Young, T. Aung, M. Seielstad, Y.-Y. Teo, Y.J. Kim, J.-Y. Lee, B.-G. Han, D. Kang, F.-J. Tsai, L.-C. Chang, S.-J. C. Fann, Y.-T. Chen, H. Mei, D.C. Rao, J.E. Hixson, S. Chen, T. Katsuya, M. Isono, T. Ogihara, J.C. Chambers, W. Zhang, J.S. Kooner, the KidneyGen consortium, the CKDGen consortium, the GUGC consortium, K. Yamamoto, M. Kubo, Y. Nakamura, N. Kamatani, N. Kato, J. He, J.-Y. Wu, Y. S. Cho, E.-S. Tai, and T. Tanaka: Genome-wide meta-analysis identifies multiple loci associated with kidney function-related traits in east Asian populations. Nature Genetics, 44: 904-909, 2012
46. M.G. Dunlop, S.E. Dobbins, S.M. Farrington, A. M. Jones, C. Palles, N. Whiffin, A. Tenesa, S. Spain, P. Broderick, L.-Y. Ooi, E. Domingo, C. Smillie, M. Henrion, M. Frampton, L. Martin, G. Grimes, M. Gorman, C. Semple, Y.P. Ma, E. Barclay, J. Prendergast, J.-B. Cazier, B. Olver, S. Penegar, S. Lubbe, I. Chander, L.G. Carvajal-Carmona, S. Ballereau, A. Lloyd, J. Vijayakrishnan, L. Zgaga, I. Rudan, E. Theodoratou, The Colorectal Tumour Gene Identification (CORGI) Consortium, J.M. Starr, I. Deary, I. Kirac, D. Kovacevi, L.A. Aaltonen, L. Renkonen-Sinisalo, J.-P. Mecklin, K. Matsuda, Y. Nakamura, Y. Okada, S. Gallinger, D.J. Duggan, D. Conti, P. Newcomb, J. Hopper, M.A. Jenkins, F. Schumacher, G. Casey, D. Easton, M. Shah, P. Pharoah, A. Lindblom, T. Liu, The Swedish Low-Risk Colorectal Cancer Study Group, C.G Smith, H. West, J.P. Cheadle, The COIN Collaborative Group, R. Midgley, D.J. Kerr, H. Campbell, I.P. Tomlinson, and R.S. Houlston: Common variation near CDKN1A, POLD3 and SHROOM2 influences colorectal cancer risk. Nature Genetics, 44: 770-777, 2012
47. T. Masuzawa, Y. Fujiwara, K. Okada, A. Nakamura, S. Takiguchi, K. Nakajima, H. Miyata, M. Yamasaki, Y. Kurokawa, R. Osawa, K. Takeda, K. Yoshida, T. Tsunoda, Y. Nakamura, M. Mori, and Y. Doki: Phase I/II study of S-1 plus cisplatin combined with peptide vaccines for human vascular endothelial growth factor receptor 1 and 2 in patients with advanced gastric cancer. International Journal of Oncology, 41: 1297-1304, 2012
48. K. Kusakabe, N. Ide, Y. Daigo, T. Itoh, K. Higashino,Y. Okano, G. Tadano, Y. Tachibana, Y. Sato, M. Inoue, T. Wada, M. Iguchi, T. Kanazawa, Y. Ishioka, K. Dohi, S. Tagashira, Y. Kido, S. Sakamoto, K. Yasuo, M. Maeda, T. Yamamoto, M. Higaki, T. Endoh, K. Ueda, T. Shiota, H. Murai, and Y. Nakamura : Diaminopyridine-based potent and selective MPS1 kinase inhibitors binding to an unusual flipped-peptide conformation. Medical Chemistry Letters, 3: 560?564, 2012
49. T. Fujitomo, Y. Daigo, K. Matsuda, K. Ueda, and Y. Nakamura: Critical function for nuclear envelope protein TMEM209 in human pulmonary carcinogenesis Cancer Research, 72: 4110-4118, 2012
50. K. Shiraishi, H. Kunitoh, Y. Daigo, A. Takahashi, K. Goto, H. Sakamoto, S. Ohnami, Y. Shimada, K. Ashikawa, A. Saito, S. Watanabe, K. Tsuta, N. Kamatani, T. Yoshida, Y. Nakamura, J. Yokota, M. Kubo, and T. Kohno: A genome-wide association study identifies two new susceptibility loci for lung adenocarcinoma in the Japanese population. Nature Genetics, 44: 900-903, 2012
51. M.-H. Nguyen, K. Ueda, Y. Nakamura, and Y. Daigo: Identification of a novel oncogene MMS 22L involved in lung and esophageal carcinogenesis. International Journal of Oncology, 41: 1285-1296, 2012
52. K. Kono, H. Iinuma, Y. Akutsu, H. Tanaka, N. Hayashi, Y. Uchikado, T. Noguchi, H. Fujii, K. Okinaka, R. Fukushima, H. Matsubara, M. Ohira, H. Baba, S. Natsugoe, S. Kitano, K. Yoshida, T. Tsunoda, and Y. Nakamura: Multicenter, Phase II clinical trial of cancer vaccination for advanced esophageal cancer with three peptides derived from novel cancer-testis antigens. Journal of Trasnlational Medicine, 10141, 2012
53. B.E. Himes, X. Jiang, R. Hu, A.C. Wu, J.A. Lasky-Su, B.J. Klanderman, J. Ziniti, J. Senter-Sylvia, J.J. Lima, C.G. Irvin, S.P. Peters, D.A. Meyers, E.R. Bleecker, M. Kubo, M. Tamari, Y. Nakamura, S.J. Szefler, R.F. Lemanske Jr., R.S. Zeiger, R.C. Strunk, F.D. Martinez, J.P. Hanra-

han, G.H. Koppelman, D.S. Postma, M.A.E. Nieuwenhuis, J.M. Vonk, R.A. Panettieri Jr., A. Markezich, E. Israel, V.J. Carey, K.G. Tantisira, A.A. Litonjua, Q. Lu, and S.T. Weiss: Genome-wide association analysis in asthma subjects identifies SPATS2L as a novel bronchodilator response gene. PLoS Genetics, 8e1002824, 2012

54. M. Takahashi, Y. Furukawa, H. Shimodaira, M. Sakayori, T. Moriya, Y. Moriya, Y. Nakamura. and C. Ishioka: Aberrant splicing caused by a MLH1 splice donor site mutation found in a young Japanese patient with Lynch syndrome Familial Cancer, 11559-564, 2012
55. Y. Ji, J.M. Biernacka, S. Hebbring, Y Chai, G.D. Jenkins, A. Batzler, K.A. Snyder, M.S. Drews, Z. Desta, D. Flockhart, T. Mushiroda, M. Kubo, Y. Nakamura, N. Kamatani, D. Schaid, R.M. Weinshilboum and D.A. Mrazek: Pharmacogenomics of selective serotonin reuptake inhibitor treatment for major depressive disorder: genome-wide associations and functional genomics. Pharmacogenomics Journal, doi: 10.1038/tpj.2012.32, 2012
56. R.M. Baldwin, K. Owzar, H. Zembutsu, A. Chhibber, M. Kubo, C. Jiang, D. Watson, R.J. Eclov, J. Mefford, H.L. McLeod, P.N. Friedman, C.A. Hudis, E.P. Winer, E.M. Jorgenson, J.S. Witte, L.N. Shulman, Y. Nakamura, M.J. Ratain and D.L. Kroetz: A Genome-Wide Association Study Identifies Novel Loci for Paclitaxel-Induced Sensory Peripheral Neuropathy in CALGB 40101 Clinical Cancer Research, 185099-5109, 2012
57. T. Hirota, A. Takahashi, M. Kubo, T. Tsunoda, K. Tomita, M. Sakashita, T. Yamada, S. Fujieda, S. Tanaka, S. Doi, A. Miyatake, T. Enomoto, C. Nishiyama, N. Nakano, K. Maeda, K. Okumura, H. Ogawa, S. Ikeda, E. Noguchi, T. Sakamoto, N. Hizawa, K. Ebe, H. Saeki, T. Sasaki, T. Ebihara, M. Amagai, S. Takeuchi, M. Furue, Y. Nakamura, and M. Tamari Genome-wide association study identifies eight new susceptibility loci for atopic dermatitis in the Japanese population Nature Genetics, 441222-1226, 2012
58. V. Kumar, P.H.Y. Lo, H. Sawai, N. Kato, A. Takahashi, Z. Deng, Y. Urabe, H. Mbarek, K. Tokunaga, Y. Tanaka, M. Sugiyama, M. Mizokami, R. Muroyama, R. Tateishi, M. Omata, K. Koike, C. Tanikawa, N. Kamatani, M. Kubo, Y. Nakamura, and K. Matsuda Soluble MICA and a MICA variation as possible prognostic biomarkers for HBV-induced hepatocellular carcinoma Plos ONE, 7e44743, 2012
59. S. Elgazzar, H. Zembutsu, A. Takahashi, M. Kubo, F. Aki, K. Hirata, Y. Takatsuka, M. Okazaki, S. Ohsumi, T. Yamakawa, M. Sasa, T. Katagiri, Y. Miki, and Y. Nakamura: A genome-wide association study identifies a genetic variant in the SIAH2 locus associated with hormonal receptor positive breast cancer in Japanese Journal of Human Genetics, doi: 10.1038/jhg.2012.108, 2012
60. H.-S. Cho, T. Shimazu, G. Toyokawa, Y. Daigo, Y. Maehara, S. Hayami, A. Ito, K. Masuda, N. Ikawa, H. I. Field, E. Tsuchiya, S. Ohnuma, B.A. J. Ponder, M. Yoshida, Y. Nakamura, and R. Hamamoto: Enhanced HSP70 lysine 561 methylation promotes proliferation of cancer cells through activation of aurora kinase B Nature Communications, 3DOI: 10.1038/ncomms2074, 2012
61. S. Akamatsu, A. Takahashi, R. Takata, M. Kubo, T. Inoue, T. Morizono, T. Tsunoda, N. Kamatani, C.A. Haiman, P. Wan, G.K. Chen, L. Le Marchand, L.N. Kolonel, B.E. Henderson, T. Fujioka, T. Habuchi, Y. Nakamura, O. Ogawa, and H. Nakagawa: Reproducibility, Performance, and Clinical Utility of a Genetic Risk Prediction Model for Prostate Cancer in Japanese PLOS ONE, 7e46454, 2012
62. A. Toyama, H. Nakagawa, K. Matsuda, T. Sato, Y. Nakamura, and K. Ueda: Quantitative structural characterization of local N-glycan microheterogeneity in therapeutic antibodies by energy-resolved oxonium ion monitoring. Analytical Chemistry, 849655-9662, 2012
63. D.R. Nyholt, S.-K. Low, C.A. Anderson, J.N. Painter, S. Uno, A.P. Morris, S. MacGregor, S.D. Gordon, A.K. Henders, N.G. Martin, J. Attia, E. G. Holliday, M. McEvoy, R.J. Scott, S.H. Kennedy, S.A. Treloar, S.A. Missmer, S. Adachi, K. Tanaka, Y. Nakamura, K.T. Zondervan, H. Zembutsu, and G.W. Montgomery: Multi-ethnic GWA meta-analysis identifies new endometriosis risk loci Nature Genetics, 441355-1359, 2012
64. K.M. Giacomini, S.W. Yee, M.J. Ratain, R.M. Weinshilboum, N. Kamatani and Y. Nakamura: Pharmacogenomics and patient care: one size does not fit all. Science Translational Medicine, 4: 153ps18, 2012
65. R. Takata, K. Matsuda, J. Sugimura, W. Obara, T. Fujioka, K. Okihara, N. Takaha, T. Miki, S. Ashida, K. Inoue, C. Tanikawa, T. Shuin, S. Sasaki, Y. Kojima, K. Kohri, M. Kubo, M. Yamaguchi, Y. Ohnishi, and Y. Nakamura Impact of four loci on serum tamsulosin hydrochloride concentration. Journal of Human Genetics, in press: doi: 10.1038/jhg.2012.126, 2012
66. T. Fukawa,, M. Ono, T. Matsuo, H. Uehara, T. Miki, Y. Nakamura, H. Kanayama, and T. Katagiri: DDX31 regulates the p53-HDM2 pathway and rRNA gene transcription through its interaction with NPM1 in renal cell carcinomas Cancer Research, 72: 5867-5877, 2012
67. E. Danese, M. Montagnana, J.A. Johnson, A.E. Rettie, C.F. Zambon, S.A. Lubitz, G. Suarez-Kurtz, L.H. Cavallari, L. Zhao, M. Huang, Y. Nakamura, T. Mushiroda, M.K. Kringen, P. Bor-

giani, C. Ciccacci, N.T. Au, T. Langaee, V. Siguret, M.-A. Loriot, H. Sagreiya, R.B. Altman, M. H.A. Shahin, S.A. Scott, S.I. Khalifa, B. Chowbay, I.M. Suriapranata, M. Teichert, B.H. Stricker, M. Taljaard, M.R. Botton, J.E. Zhang, M. Pirmohamed, X. Zhang, J.F. Carlquist, B.D. Horne, M.T.M. Lee, V. Pengo, G.C. Guidi, P. Minuz, and C. Fava Impact of the CYP4F2 p.V 433M polymorphism on coumarin dose requirement: systematic review and meta-analysis Clinical Pharmacology & Therapeutics, 92: 746-756, 2012

68. M. Komatsu, T. Yoshimaru, T. Matsuo, K. Kiyotani, Y. Miyoshi, T. Tanahashi, K. Rokutan, R. Yamaguchi, A. Saito, S. Imoto, S. Miyano, Y. Nakamura, M. Sasa, M. Shimada, and T. Katagiri: Molecular features of triple negative breast cancers by genome-wide gene expression profiling analysis. International Journal of Oncology, doi: 10.3892/ijo.2012.1744 2012
69. H.E. Wheeler, E.R. Gamazon, C. Wing, U.O. Njiaju, C. Njoku, R.M. Baldwin, K. Owzar, C. Jiang, D. Watson, I. Shterev, M. Kubo, H. Zembutsu, E. Winer, C. Hudis, L.N. Shulman, Y. Nakamura, M.J. Ratain, D.L. Kroetz, N.J. Cox, M.E. Dolan for the Cancer and Leukemia Group B: Integration of cell line and clinical trial genome-wide analyses implicates multiple loci in paclitaxel-induced peripheral neuropathy. Clinical Cancer Research, in press:, 2012
70. N. Franceschini, F.J. van Rooij, B.P. Prins, M.F. Feitosa, Karakas M, Eckfeldt JH, Folsom AR, Kopp J, Vaez A, Andrews JS, Baumert J, Boraska V, Broer L, Hayward C, Ngwa JS, Okada Y, Polasek O, Westra HJ, Wang YA, Del Greco MF, Glazer NL, Kapur K, Kema IP, Lopez LM, Schillert A, Smith AV, Winkler CA, Zgaga L; The LifeLines Cohort Study, Bandinelli S, Bergmann S, Boban M, Bochud M, Chen YD, Davies G, Dehghan A, Ding J, Doering A, Durda JP, Ferrucci L, Franco OH, Franke L, Gunjaca G, Hofman A, Hsu FC, Kolcic I, Kraja A, Kubo M, Lackner KJ, Launer L, Loehr LR, Li G, Meisinger C, Nakamura Y, Schwienbacher C, Starr JM, Takahashi A, Torlak V, Uitterlinden AG, Vitart V, Waldenberger M, Wild PS, Kirin M, Zeller T, Zemunik T, Zhang Q, Ziegler A, Blankenberg S, Boerwinkle E, Borecki IB, Campbell H, Deary IJ, Frayling TM, Gieger C, Harris TB, Hicks AA, Koenig W, O'Donnell CJ, Fox CS, Pramstaller PP, Psaty BM, Reiner AP, Rotter JI, Rudan I, Snieder H, Tanaka T, van Duijn CM, Vollenweider P, Waeber G, Wilson JF, Witteman JC, Wolffenbuttel BH, Wright AF, Wu Q, Liu Y, Jenny NS, North KE, Felix JF, Alizadeh BZ, Cupples LA, Perry JR, Morris AP. Discovery and Fine Mapping of Serum Protein Loci through Transethnic Meta-analysis Am. J. Human Genetics, 91: 744-753, 2012
71. S. Chung, H. Suzuki, T. Miyamoto, N. Takamatsu, A. Tatsuguchi, K. Ueda, K. Kijima, Y. Nakamura and Y. Matsuo: Development of an orally-administrative MELK-targeting inhibitor that suppresses the growth of various types of human cancer Oncotarget, in press:, 2012
72. F. Innocenti, K. Owzar, N.L. Cox, P. Evans, M. Kubo, H. Zembutsu, C. Jiang, D. Hollis, T. Mushiroda, L. Li, P. Friedman, L.W. Wang, D. Glubb, H. Hurwitz, K. Giacomini, H.L. McLeod, R.M. Goldberg, R.L. Schilsky, H.L. Kindler, Y. Nakamura, and M. Ratain: A genome-wide association study of overall survival in pancreatic cancer patients treated with gemcitabine in CALGB 80303 Clinical Cancer Research, 18: 577-584, 2012
73. A. Köttgen, E. Albrecht, A. Teumer, V. Vitart, J. Krumsiek et al. Genome-wide association analyses identify 18 new loci associated with serum urate concentrations Nature Genetics, doi: 10.1038/ng.2500, 2012
74. I. Cheng, G.K. Chen, H. Nakagawa, J. He, P. Wan, C.C. Laurie, J. Shen, X. Sheng, L.C. Pooler, A.H. Crenshaw, D.B. Mirel, A. Takahashi, M. Kubo, Y. Nakamura, Yusuke) 7; A.A. Olama, S. Benlloch, J.L. Donovan, M. Guy, F.C. Hamdy, Z. Kote-Jarai, D.E. Neal, L.R. Wilkens, K.R. Monroe, D.O. Stram, K. Muir, R.A. Eeles, D. Easton, L.N. Kolonel, B.E. Henderson, L. Le-Marchand, and C.A. Haiman; Evaluating genetic risk for prostate cancer among Japanese and Latinos. Cancer Epidemiology Biomarkers & Prevention, 21: 2048-2058, 2012
75. K. Yamazaki, J. Umeno, A. Takahashi, A. Hirano, T.A. Johnson, N. Kumasaka, T. Morizono, N. Hosono, T. Kawaguchi, M. Takazoe, T. Yamada, Y. Suzuki, H. Tanaka, S. Motoya, M. Hosokawa, Y. Arimura, Y. Shinomura, T. Matsui, T. Matsumoto, M. Iida, T. Tsunoda, Y. Nakamura, N. Kamatani and M. Kubo: A Genome-wide association study Identifies 2 susceptibility loci for Crohn's disease in a Japanese population. Gastroenterology, in press: doi: pii: S0016-5085(12)01847-1., 2012
76. E.T. Betcheva, A.G. Yosifova, T. Mushiroda, M. Kubo, A. Takahashi, S.K. Karachanak, I.T. Zaharieva, S.P. Hadjidekova, I.I. Dimova, R.V. Vazharova, D.S. Stoyanov, V.K. Milanova, T. Tolev, G. Kirov, N. Kamatani, D.I. Toncheva, and Y. Nakamura: Whole-genome-wide association study in the Bulgarian population reveals HHAT as schizophrenia susceptibility gene. Psychiatric Genetics, 23: 11-19, 2013
77. D. Kang, H.-S. Cho, G. Toyokawa, M. Kogure, Y. Yamane, Y. Iwai, S. Hayami, T. Tsunoda, H. I. Field, K. Matsuda, D.E. Neal, B. A.J. Ponder, Y. Maehara, Y. Nakamura, and R. Hamamoto: The Histone Methyltransferase Wolf-Hirschhorn Syndrome Candidate 1-like 1 (WHSC1L1) Is In-

volved in Human Carcinogenesis Genes Chromosomes and Cancer, 52: 126-139, 2013
78. R. Nakano, T. Maekawa, H. Abe, Y. Hayashida, H. Ochi, T. Tsunoda, H. Kumada, N. Kamatani, Y. Nakamura and K. Chayama: Single-nucleotide polymorphisms in GALNT8 are associated with the response to interferon therapy for chronic hepatitis C Journal of General Virology, 9481-89, 2013
79. Y. Kamada, N. Kinoshita, Y. Tsuchiya, K. Kobayashi, H. Fujii, N. Terao, K. Kamihagi, N. Koyama, S. Yamada, Y. Daigo, Y. Nakamura, N. Taniguchi, E. Miyoshi: Reevaluation of a lectin antibody ELISA kit for measuring fucosylated haptoglobin in various conditions Clinica Chimica Acta, in press, 2013
80. K. Shimane, Y. Kochi, A. Suzuki, Y. Okada, T. Ishii, T. Horita, K. Saito, A. Okamoto, N. Nishimoto, K. Myouzen, M. Kubo, M. Hirakata, T. Sumida, Y. Takasaki, R. Yamada, Y. Nakamura, N. Kamatani and K. Yamamoto An association analysis of HLA-DRB1 with systemic lupus erythematosus and rheumatoid arthritis in a Japanese population: effects of *09:01 allele on disease phenotypes. Rheumatology, in press, 2012
81. Y. Urabe, H. Ochi, N. Kato, V. Kumar1, A. Takahashi, R. Muroyama, N. Hosono3, M. Otsuka, R. Tateishi, P.H.Y. Lo, C. Tanikawa, M. Omata, K. Koike, D. Miki, H. Abe, N. Kamatani, J. Toyota, H. Kumada, M. Kubo, K. Chayama, Y. Nakamura, and Koichi Matsuda A genome-wide association study of HCV induced liver cirrhosis in the Japanese population identifies novel susceptibility loci at MHC region. Journal of Hepatology, in press, 2013
82. A. Aarnink, H.J. Garchon, Y. Okada, T. Takahashi, K. Matsuda, M. Kubo, Y. Nakamura, and B, Blancher Comparative analysis in cynomolgus macaque identifies a novel human MHC locus controlling platelet blood counts independently of BAK1 Journal of Thrombosis and Haemostasis, in press, 2013

Human Genome Center

Laboratory of Functional Analysis *In Silico*
機能解析イン・シリコ分野

Professor	Kenta Nakai, Ph.D.	教　授　理学博士	中　井　謙　太
Assistant Professor	Ashwini Patil, Ph.D.	助　教　理学博士	パティル，アシュウィニ

The mission of our laboratory is to conduct computational ("in silico") studies on the functional aspects of genome information. Roughly speaking, genome information represents what kind of proteins/RNAs are synthesized under which conditions. Thus, our study includes the structural analysis of molecular function of each gene product as well as the analysis of its regulatory information, which will lead us to the understanding of its cellular role represented by the networks of inter-gene interactions.

1. Promoter structure modeling of co-expressed genes in distinct cell types

Yosvany Lopez and Kenta Nakai

The understanding of the mechanisms of transcriptional regulation remains a great challenge for molecular biologists in the post-genome era. At the transcriptional level, DNA-binding proteins (transcription factors) modulate the expression of genes by binding to their specific DNA regulatory elements in nearby genomic regions. Nowadays, the identification and characterization of these components is valuable because the presence or absence of transcription factors binding sites (TFBSs) seems to be responsible for the complexity of gene regulation in every living organism. Based on the assumption that regulatory regions of (at least a part of) those genes showing similar expression profiles should share some common structural characteristics, we are attempting to explain how the binding of transcription factors is carried out in promoter regions of co-regulated genes in specific biological tissues. We are working on a database of co-expressed genes in *Arabidopsis thaliana* (ATTED-II) because, among other reasons, the intergenic regions of plant genes are smaller than that of higher organisms. We have found new patterns of motif combinations capable of describing promoter structures of co-expressed genes in five different cell types. The distance of motifs from translation start sites on both DNA strands are two characteristics taken into account. We will incorporate further features to achieve a promoter structure modeling as broad as possible and thus apply it to regulatory regions of co-expressed genes in different human tissues.

2. Analysis of Transcription Start Sites of mice dendritic cells after LPS stimulation

Kuo-Ching Liang, Yutaro Kumagai[2], Yutaka Suzuki[1], Shizuo Akira[2], and Kenta Nakai: [1]Graduate School of Frontier Sciences, [2]WPI-iFReC, Osaka University

Using mouse dendritic cell TSS-Seq data for samples collected at 0hr, 0.5hr, 1hr, 2hrs, 3hrs, 4hrs, 6 hrs, 8hrs, 16hrs and 24hrs after LPS stimulation, we have been able to classify genes into different classes based on how their TSS tag distribution changed between sample taken before stimulation (0hr) and samples taken after stimulation (0.5hr to 24hr). By classifying genes based on their TSS tag distribution changes, we may be able to discover genes that undergo usage of alternative promoter regions after an immune response. We group genes

that exhibit statistically significant changes in TSS tag distributions into two major categories: Class 1 where the number of tag counts for a gene increases due to LPS stimulation, but does not show a shift in the location of the dominant peak; and Class 2 where there is a shift in the dominant peak of TSS tags. For genes in Class 2, the shift of the dominant peak of TSS tags offers an indication of possible alternative promoter usage, which can be strengthened by comparison with the changes in RNA-Seq reads coverage in the exon regions over time. In our analysis, we have found correlations between the classification of a gene's TSS distribution changes to the dependency of its activation to MyD88 and/or TRIF. In addition, we have also found evidence for alternative promoter usage and possible alternative splicing due to LPS stimulation in genes such as IL6. Furthermore, we have also observed some associations between immune-related genes in Class 1 and Class 2, and the time it takes for the time-course expression to reach its peak value after LPS stimulation.

3. Massive-scale RNA-Seq analysis of mouse early embryo development

Sung-Joon Park, Makiko Komata[3], Fukashi Inoue[4], Kaori Yamada[4], Kenta Nakai, Miho Ohsugi[4], Katsuhiko Shirahige[3]: [3]IMCB, U. Tokyo, [4]Division of Oncology

To understand the mechanism of biological events observed during the early mammalian embryo development, we performed transcriptome analysis on mouse oocyte, one-cell, two-cell, and four-cell embryonic stages using approximately 10,000×2 high-quality cells per stage, which is an unprecedented experimental scale. In addition, we prepared parthenotes, as well as mouse embryonic fibroblasts. Over 130 million short reads sequenced by SOLiD system were analyzed by Tophat and Cufflinks software coupled with a recursive mapping strategy, where unmapped reads are truncated and mapped again. Consequently, we obtained the compendium of transcriptome landscape during the development. The landscape revealed gene expression patterns and their functions that are not suggested from previous studies. By incorporating a large range of publicly available datasets, we ascertained the great impact of ncRNAs immediately after fertilization, as well as the behavior of nascent and sperm-born transcripts at one-cell stage. Remarkably, we discovered unique promoter architecture from transcripts newly synthesized at one-cell stage, which suggests that specific transcription factors strictly regulate the onset of zygotic development. The transcriptome profile we established here is an important resource to help enhance our understanding of the mammalian embryo development.

4. Inferring gene regulatory network from RNA-seq data of mouse hematopoietic stem and progenitor cells

Sung-Joon Park, Terumasa Umemoto[5], Masayuki Yamamoto[5], Kenta Nakai: [5]Institute of Advanced Biomedical Engineering and Science, Tokyo Women's Medical University

Hematopoietic stem cell (HSC) is an ideal model for the study of mechanisms underlying multilineage cell differentiation and self-renewal. Despite significant efforts to understand hematopoiesis, it is still largely unknown how a HSC controls its self-renewal and how it is developed to a hematopoietic progenitor cells (HPC). To answer these questions, we predicted key transcription factors (TFs) by a regression model that uses *in silico* TFBS detection and *in vivo* gene expression level in HSCs and HPCs. Then, we modeled TF-gene interactions that reveal similar and different regulations for genes differentially expressed in HSCs and HPCs. Here, we used 10,000×3 $CD34^-$KSL HSCs and 10,000×3 $CD34^+$KSL HPCs from C57BL/6J mice, and profiled gene expression level by SOLiD system. In addition, we performed RT-PCR and a microarray experiment. The profiles suggested that 904 genes are differentially expressed (1.5 fold change). A linear regression model predicted these gene expressions with TFs that bind to TRANSFAC TFBSs ($R = 0.7 \sim 0.85$). Consequently, our model suggested a candidate of key TFs to govern the pluripotency and self-renewal. One of them was validated by transplantation. To analyze more comprehensively, we are extending this approach to the direction that incorporates public datasets, such as DNA methylation, histone modification, ChIP-seq of TFs, etc. This extension will give more detailed views of hematopoiesis.

5. Comprehensive analysis of transcription initiation patterns, promoter architecture and function in *Ciona intestinalis*

Rui Yokomori, Kotaro Shimai[6], Koki Nishitsuji[7], Yutaka Suzuki[1], Takehiro Kusakabe[6] and Kenta Nakai: [6]Fac. of Sci. and Engineering, Konan Univ, [7]Grad. Schl. of Life Sci., Univ. of Hyogo

Transcription is known to start from multiple positions and the distribution of transcription start sites (TSSs) can be classified into four types: NSP (Narrow and Sharp Peak), WSP (Wide and Sharp Peak), MP (Multi Peak), and BP (Broad Peak). However, the biological meaning of each type is poorly understood. In this study, we identified promoters across the *Ciona intestinalis* genome and

analyzed their TSS distributions (transcription initiation patterns), promoter architectures and functions using 5 different tissues. In addition, we identified trans-spliced accepter sites. As a result, we found many novel candidate promoter and accepter sites including those that show tissue-specificity. TSS distribution analysis showed that TATA-box was significantly enriched in NSP type promoters compared to the other 3 types. Also, ribosomal protein genes were significantly associated with NSP type and 97% of them possessed TATA-less promoters. We also found two potentially interesting unknown motifs showing positional bias. One of them was specific to NSP type promoters. Therefore, it may involve the determination of transcription start site as well as TATA- box.

6. Analysis of EJC contribution to the efficient mRNA splicing involved in mitotic cell-cycle progression

Shunichi Wakabayashi, Kazuhiro Fukumura[8], Kenta Nakai, Kunio Inoue[8], and Hiroshi Sakamoto[8]: [8]Gradu. Sch. Sci., Kobe U.

The exon junction complex (EJC) deposited onto spliced mRNAs upstream of exon-exon junctions plays important roles in post splicing gene expression events. However, it is unclear whether the EJC has a direct role in pre-mRNA splicing. We performed RNA-seq and RT-PCR using HeLa cells in which one of the core components of EJC, Y14 was depleted. We calculated rates of accurately spliced mRNAs for all introns in all genes using RNA-seq data, and analyzed Y14 affected introns and genes with *in silico* approach. We confirmed effects of Y14 experimentally using RT-PCR. We revealed that Y 14 is required for efficient and accurate splicing of group of transcripts, in which short intron-containing genes involved in mitotic cell-cycle progression are enriched. Tethering assay of EJC components also supported these results. Moreover, we showed that one of the EJC core associated factors, RNPS1, is required for efficient splicing of a subset of EJC regulated introns, but not for that of another subset. These results indicate that the EJC is required for efficient pre-mRNA splicing of a specific subset of introns in mitotic cell cycle related genes.

7. Construction of disease related protein-protein network for the identification of inter-disease relationships and candidate pathways

Saaya Tsutsué, Ashwini Patil and Kenta Nakai

Recent developments of statistical methods for social networks, together with accumulation of Omics data have gradually opened the door to predict the potential candidates for future medical application through translational research. Recent innovation of various high-throughput technologies, which have made available large amounts of data, have made it difficult to identify the actual relationships within the data itself. Ordinary statistical network methods, while useful, do not facilitate the biological interpretation of data. Therefore, in order to extract significant information from the accumulated data, we need to establish a novel approach using both social network analysis, combined with other bioinformatics analysis. We created a large human protein-protein interaction network by combining interactions from several databases and assigned disease annotations from OMIM to the proteins. We then defined a new score (GSPA) using social network analyses to expand 25 disease classified protein network data to detect inter-disease relationships and potential candidate in pathways.

8. A study of the innate immunity interactome dynamics

Asmaa Elzawahry, Ashwini Patil, Yutaro Kumagai[2], Yutaka Suzuki[1], Kenta Nakai

Immune system of host is responsible for defense against invading pathogens. Protein-protein Interactions (PPIs) are essential components of the immune response. So far, protein-protein interactions have been curated as a static interaction map. However, interactions involved in the immune response are dynamic. This study focuses on interactome dynamics during immune response by combining time-series gene expression data with protein-protein interactions followed by protein complex identification. We identified differences in the interactome during immune response by constructing differential networks and identifying protein modules which were up/down regulated for each stage. In addition, we identified core interactions that are conserved in interactome throughout immune response. We defined Interaction protein ratio (IPR) and Pair wise differential matrix (PWDM) to assess differences between network maps. To get a comprehensive view of the TLR interactome network we investigated the TLR subnetwork and found that S100A8 is down regulated in dendritic cells after LPS stimulation. Combining time-series gene expression data with protein-protein interactions data revealed interactome dynamics during immune response. Identified protein complexes during interactome dynamics have a crucial role not only in innate immunity but also in other biological processes and pathways like pathways in cancer, circadian rhythm and p53 pathway.

9. A study of the maintenance of chemical composition in poorly conserved intrinsically disordered regions and its utility in classification.

Ashwini Patil, Harry Amri Moesa, Shunichi Wakabayashi and Kenta Nakai

Despite the importance of intrinsically disordered regions in proteins, there is currently no classification system available for these regions. We studied the levels of conservation of known and predicted disordered regions in eukaryotes and proposed a method to classify them based on their conservation and their residue content i.e. chemical composition. We found that disordered regions showing low conservation often show high maintenance of chemical composition (i.e. similar fraction of charged, polar and hydrophobic residues), which may be used to maintain their disorderliness in order to stay functional despite high rates of evolution. We also found that highly conserved disordered regions can be classified into distinct groups based on their chemical composition which are associated with distinct functions suggesting a possible functional classification of disordered regions.

10. A study of the pathways activated during the immune response using a comprehensive interaction and regulatory network

Ashwini Patil, Kuo-Ching Liang, Yutaro Kumagai[1], Shizuo Akira[1], and Kenta Nakai

The immune response is triggered at multiple levels and it is no longer sufficient to study its activation using protein-protein interactions or transcription regulatory information alone. In this study, we propose to identify the pathways activated during the immune response with the help of a comprehensive interaction and transcription regulatory network in combination with the time-series expression data from LPS stimulated mouse dendritic cells. A high-confidence network was constructed using protein-protein interactions from Innatedb and HitPredict, transcription regulatory data from TRANSFAC and pathway information from KEGG. Expression levels of several thousand genes were obtained at 10 time-points. The level and time of expression of each gene was used as a measure of its importance in the response. Using a strategy based on network flows, we propose to identify novel genes and their inter-relationships during the activation of the immune response. Thus, we hope to uncover new components in currently known pathways, as well as previously unknown pathways that function in the immune response.

11. Prediction of Ub-independent proteasome degradation substrates using their unstructured initiation regions

Ashwini Patil, Tomonao Inobe[9], Shunichi Wakabayashi, Kenta Nakai: [9]Toyama University

The presence of a poly-Ubiquitin (Ub) tag facilitates the degradation of proteins by the proteasome. In recent years, several proteasome substrates have been identified that can be degraded in the absence of the poly-Ub tag. It is thought that an unstructured initiation region in these proteins plays an important role in such degradation. Using the unstructured initiation regions of known Ub-independent proteasomal substrates, we identified its distinctive features. We then used the defined features to predict unstructured initiation regions in proteasome substrates that facilitate degradation in the absence of the Ub-tag. We further intend to experimentally confirm our predictions.

12. Project for constructing an IT infrastructure for accelerating the clinical application of regenerative medicine technologies

Kenta Nakai, Masanori Akiyama[10], Tatsutoshi Nakahata[11], Norio Nakatsuji[12], Kohji Nishida[13], Hideyuki Okano[14], Masayo Takahashi[15], Akihiro Umezawa[16], Masayuki Yamato[17], and Shinya Yamanaka[11]: [10]PARI, U. Tokyo, [11]CiRA, Kyoto U., [12]IFMS, Kyoto U., [13]Grad. Sch. Med., Osaka U., [14]Sch. Med. Keio U., [15]CDB, RIKEN, [16]Nat. Res. Inst. Child Health and Development, [17]Tokyo Women's Med. U.

This project aims to achieve the earliest possible clinical application of regenerative medicine technologies using safe, effective and high-quality human stem cells by building a collaborative platform among researchers through a network centric research infrastructure. Also in the project, the basis of an "open innovation" environment allowing research institutes to continuously create innovative technologies will be developed. To support this activity, an advisory committee has been formed with members from various sectors.

Publications

Patil, A., Teraguchi, S., Dinh, H., Nakai, K., and Standley, D.M. Functional annotation of intrinsically disordered domains by their amino acid content using IDD Navigator. Pacific Symposium

on Biocomputing 17: 164-175, 2012.
Yamashita, R., Sugano, S., Suzuki, Y., and Nakai, K. DBTSS: database of transcriptional start sites progress report in 2012. Nucl. Acids Res. 40(Database Issue): D150-154, 2012.
Kimura, K., Koike, A., and Nakai, K. A Bit-parallel dynamic programming algorithm suitable for DNA sequence alignment. J. Bioinformatics and Computational Biology. 10: 1250002.
Kraut, D.A., Israeli, E., Schrader, E.K., Patil, A., Nakai, K., Nanavati, D., Inobe, T., Matouschek, A. Sequence- and species-dependence of proteasomal processivity. ACS Chem Biol. 7(8): 1444-53, 2012.
Morita, S., Takahashi, R.U., Yamashita, R., Toyoda, A., Horii, T., Kimura, M., Fujiyama, A., Nakai, K., Tajima, S., Matoba, R., Ochiya, T., Hatada, I. Genome-Wide Analysis of DNA Methylation and Expression of MicroRNAs in Breast Cancer Cells. Int J Mol Sci. 13(7): 8259-72, 2012.
Moesa, H.A., Wakabayashi, S., Nakai, K., Patil, A. Chemical composition is maintained in poorly conserved intrinsically disordered regions and suggests a means for their classification. Mol Biosyst. 8(12): 3262-73, 2012.
Ohkura, N., Hamaguchi, M., Morikawa, H., Sugimura, K., Tanaka, A., Ito, Y., Osaki, M., Tanaka, Y., Yamashita, R., Nakano, N., Huehn, J., Fehling, H.J., Sparwasser, T., Nakai, K., Sakaguchi, S. T cell receptor stimulation-induced epigenetic changes and Foxp3 expression are independent and complementary events required for Treg cell development. Immunity 37(5): 785-99, 2012.
Ozawa, M., Sakatani, M., Yao, J., Shanker, S., Yu, F., Yamashita, R., Wakabayashi, S., Nakai, K., Dobbs, K.B., Sudano, M.J., Farmerie, W.G., Hansen, P.J. Global gene expression of the inner cell mass and trophectoderm of the bovine blastocyst. BMC Dev Biol. 12: 33, 2012.
Kusakabe, R., Tani, S., Nishitsuji, K., Shindo, M., Okamura, K., Miyamoto, Y., Nakai, K., Suzuki, Y., Kusakabe, T.G., Inoue, K. Characterization of the compact bicistronic microRNA precursor, miR-1/miR-133, expressed specifically in *Ciona* muscle tissues. Gene Expr Patterns 1: 43-50, 2013.
Makita, Y. and Nakai, K. Bacillus subtilis transcriptional network. In (Babu, M. ed.) Bacterial Gene Regulation and Transcriptional Networks, Horizon Sci. Press., (ISBN: 978-1-908230-14-0), in press. (2013)

Human Genome Center

Department of Public Policy
公共政策研究分野

Professor	Kaori Muto, Ph.D.	教　授	保健学博士	武藤香織
Assistant Professor	Yusuke Inoue, Ph.D.	助　教	社会医学博士	井上悠輔
Project Assistant Professor	Hyunsoo Hong, Ph.D.	特任助教	学術博士	洪　賢秀
Project Assistant Professor	Ayako Kamisato, M.A.	特任助教	法学修士	神里彩子
Project Assistant Professor	Yuichi Maru, M.A.	特任助教	法学修士	丸　祐一

The Department of Public Policy works to achieve three major missions: public policy studies of translational research, its application, and its impact on society; research ethics consultation for scientists to comply with ethical guidelines and to build public trust; and development of "minority-centered" scientific communication. By conducting qualitative and quantitative social science study and policy analysis, we facilitate discussion of challenges arising from advances in medical sciences. Furthermore, we study specific ethics issues related to construction of a human biological substances collection, and related to vaccination policy. We also held a Sci/Art Exhibition titled as "Office Bacteria−the Space" at the Medical Science Museum as one outreach activity undertaken via art.

1. Biobank Japan Project (BBJP) and its ethical, legal and social implications

The Biobank Japan Project (BBJP) is a disease-focused biobanking project since 2003. Biobank Japan consists of donated DNA, sera and clinical information from 200,000 patients of 66 hospitals in Japan, financially supported by the MEXT (Ministry of Education, Culture, Sports, Science, and Technology). Informed consent, which ensures the autonomous decisions of participants, is believed to be practically impossible for the biobanking project in general.

We have conducted interview studies of research coordinators ($n = 50$) since 2010 and we have just compiled our self-evaluation report on informed consent process of this project. As a means to maintain the participants' trust of the project, research coordinators who had been specially trained for the BBJP have played important roles. We have worked steadily to complete analyses of the research coordinators of the BBJP. At the beginning of the BBJP, their primary roles were recruitment. After the end of recruitment, their roles shifted to the tracing of participants to extract clinical information and to input it into the database. However, their support and encouragement of participants complemented the contents of the initial consent process and reinforced participants' incentive to continue in their role. The results of this study will contribute to improved quality control and better communication between administrators of long-term research projects and the project participants.

We also have provided research ethics consultation for conducting a follow-up study without consent, to obtain permission from 1,100 municipalities to access cause of death of 44,698 participants.

2. Research ethics consultation for multi-center cancer genome research

We have been commissioned to provide research ethics consultation to several big projects promoting medical sciences. We have started to provide re-

search ethics consultation for a new and comprehensive cancer genome research project for 34 clinical seeds of 64 designated institutions, called as Project for Development of Innovative Research on Cancer Therapeutics, headed by Tetsuo Noda and financially supported by the MEXT (Ministry of Education, Culture, Sports, Science, and Technology).

- We have addressed a basic research ethics policy to use stored and newly acquired samples from cancer patients for this research project.
- We have developed a research ethics management system and collected all protocols and consent forms. We have checked all documents and made sure that all IRB approvals would be acceptable ones.
- We provided educational opportunities to all researchers on ethical guidelines.

3. The agendas for discussion on human and animal chimera for scientific research

In July 2010, the MEXT (Ministry of Education, Culture, Sports, Science, and Technology) received the first notification of the procedure for creating human-to-animal chimeric embryos by inserting human induced pluripotent stem cells (iPS cells) to animal blastocysts, headed by Professor Hiromitsu Nakauchi of the IMSUT. This case drew attention to the necessity of reviewing the regulations of human-to-animal chimeric embryos under the Act on Regulation of Human Cloning Techniques and the Guidelines for Specified Embryo. We have conducted a questionnaire survey and a focus group interview regarding human animal chieras, for stimulating societal discussions.

4. Research in progress

We have been conducting other studies as described below.

- Ethical, legal and social implications of commercial genetic/genomic testing services in eastern Asia
- Development and evaluation of communication methods with participants of Biobank Japan and other long-term studies
- Analysis of roles of research coordinators for better recruitment and for building trust from participants
- Ethical, legal and social implications of stem cell studies including animal-human chimeric embryos and iPS cell banking
- Bench-side research ethics consultation and quality assurance of research ethics committees
- Science communication through art and ethical challenges of biomaterial art
- Ethics issues in collecting human biological substances for constructing a research infrastructure
- Vaccination policy

Publications

1. Izumi Ishiyama, Tetsuro Tanzawa, Maiko Watanabe, Tadahiko Maeda, Kaori Muto, Akiko Tamakoshi, Akiko Nagai, Zentaro Yamagata: Public attitudes to the promotion of genomic crop studies in Japan: correlations between genomic literacy, trust, and favourable attitude. Public Understanding of Science. 21(4): 498-516, 2012.
2. 丸祐一訳　スティーブン・ワイズ『コモン・ローにおける人以外の一部動物の基本権』(嶋津格編『個体と権利』(人文社会科学研究科研究プロジェクト報告書第243集), pp. 46–50, 2012.
3. 丸祐一訳「カリフォルニア州　終末期の選択肢を知る権利法Right to Know End-f-Life Options Act」,「ニューヨーク州　緩和ケア情報提供法 Palliative Care Information Act」, 平成23年度科学研究費補助金基盤研究B No. 23320001研究グループ『生命倫理研究資料集Ⅵ　世界における終末期の意思決定に関する原理・法・文献の批判的研究とガイドライン作成』(富山大学) pp. 122—130, 2012.
4. 井上悠輔. 欧州連合(EU)における臨床研究規制. 年報医事法学　27：70–80, 2012.
5. 井上悠輔. 臨床研究と利益相反. 年報医事法学　27：99–106, 2012.
6. 佐藤未来子, 井上悠輔, 武藤香織. ゲノム医療時代に必要なリテラシーとは何か？　生物の科学　遺伝66：456–459, 2012.
7. 「オーダーメイド医療実現化プロジェクト」社会との接点ワーキンググループ編著(武藤香織, 洪賢秀, 張瓊方, 永井亜貴子, 山下恭司). 『「オーダーメイド医療実現化プロジェクト」におけるインフォームド・コンセント～メディカル・コーディネーターの経験から』, 文部科学省委託事業報告書, 5–75, 2012.
8. 武藤香織, 柊中智恵子. 遺伝相談の医療化再考. インターナショナルナーシングレビュー35(3)：68–73, 2012.
9. 別府宏圀, 武藤香織. 臨床試験への患者参画＋資料：公開シンポジウム記録「臨床試験への患者参画」, 臨床評価　40(1)：53–70, 2012.
10. 井上悠輔「医学研究と利益相反」,『医学研究』(シリーズ生命倫理学第15巻, 笹栗俊之, 武藤香織編), 152–170, 丸善出版, 2012.
11. 武藤香織「倫理審査委員会」,『医学研究』(シリーズ生命倫理学第15巻, 笹栗俊之, 武藤香織編), 52–59, 丸善出版, 2012.