Human Genome Center

Division of Digital Genomics
デジタル・ゲノミクス分野

Professor	Natsuhiko Kumasaka, Ph.D.		教　授	博士(理学)	熊　坂　夏　彦
Assistant Professor	Yuji Miyatake, Ph.D.		助　教	博士(医学)	宮　竹　佑　治

Genome-wide association studies (GWAS) are a powerful approach for identifying genetic variants and genes underlying common, complex traits such as diabetes and human height. As of 2025, the GWAS Catalogue has documented more than one million genetic associations across over 5,000 complex traits. The Division of Digital Genomics aims to systematically identify genetic associations underlying common complex traits and to elucidate their molecular mechanisms by integrating cutting-edge molecular biology assays with advanced mathematical and statistical methodologies.

1. Discovery of genetic determinants for child health and development

Natsuhiko Kumasaka

Understanding how genetic and environmental factors jointly influence human health—particularly during early life—is essential for predicting long-term health outcomes. This project leverages the Japan Environment and Children's Study (JECS), a population-based prospective birth cohort, to conduct large-scale genetic analyses using questionnaire data, biological samples, and physical measurements collected longitudinally from parents and their children starting from pregnancy.

JECS is a nationwide birth cohort established by the Ministry of the Environment, Government of Japan, to evaluate the effects of environmental exposures on children's health and development. More than 100,000 pregnant women were recruited across 15 geographically distinct regional centers, capturing the broad genetic diversity of the Japanese population. Detailed questionnaire data and biological and physical measurements have been collected, with follow-up surveys conducted every six months for approximately 70% of participating children until four years of age.

My role is to perform large-scale genome-wide association studies across a wide range of child health and developmental phenotypes and to disseminate the results publicly. We identified 5,685 common genomic loci, of which 2,218 passed the phenome-wide significance threshold ($P = 4.4 \times 10^{-11}$). Approximately 15–17% of these loci represent novel associations not previously reported, including potential therapeutic targets.

In addition, longitudinal GWAS of body mass index (BMI) using Gaussian process regression identified 153 novel dynamic genetic associations across childhood development. These results enable prediction of individual growth trajectories based on genetic background. The findings have been compiled into a flagship manuscript entitled "Genome-wide association study on longitudinal and cross-sectional traits of child health and development in a Japanese population," which has been posted on medRxiv and is currently under revision at Nature.

2. Development of a novel in-vitro approach to validate and elucidate underlying molecular mechanisms of genetic associations discovered through GWAS

Yuji Miyatake, Kotoe Katayama[1], Seiya Imoto[1], Kazuaki Yokoyama[2], Yasuhito Nannya[2], Atsushi Fukuda[3], Natsuhiko Kumasaka
[1]Division of Health Medical Intelligence, Human Genome Center, IMSUT, [2]Department of Hematology/Oncology, IMSUT, [3]School of Medicine, Tokai University

Although GWAS have identified hundreds of thousands of genetic associations for complex traits, more than 90% of GWAS variants reside in non-coding regions of the genome. This presents a major challenge in pinpointing causal variants, target genes, and regulatory mechanisms. Moreover, the relevant cell types and cellular states in which these variants exert their effects often remain unknown.

Expression quantitative trait locus (eQTL) mapping provides a powerful framework for linking non-coding variants to gene regulation. Recent advances in single-cell genomics further enable the identification of specific cell types and states in which GWAS variants modulate gene expression. However, conventional eQTL studies are often costly and labor-intensive, particularly when sample acquisition is challenging, such as in human brain tissue or differentiated neurons derived from large numbers of pluripotent stem cell lines.

An emerging alternative combines single-cell RNA sequencing with massively parallel CRISPR interference (CRISPRi) screening, enabling eQTL mapping through targeted perturbations rather than natural genetic variation. While this approach substantially reduces experimental burden, it has not yet been widely applied to brain cell types relevant to neuropsychiatric or neurodegenerative diseases.

To address these limitations, we integrate stem cell biology, CRISPR-based gene regulation, and advanced machine learning. We have developed a novel CRISPR perturbation framework using human pluripotent stem cell–derived midbrain organoids, coupled with Gaussian process–based modeling. We established GASPACHO (GAuSsian Processes for Association mapping leveraging Cell HeterOgeneity), a computational framework for mapping eQTLs across continuous cellular states. In this context, the presence or absence of a guide RNA is treated analogously to genotype variation.

We have generated inducible CRISPRi and CRISPRa iPSC lines expressing dCas9-KRAB or dCas9-p300 under doxycycline control, enabling stable culture and precise temporal control of perturbations. As of December 2025, we have established midbrain organoid systems suitable for studying Parkinson's disease GWAS loci and sequenced approximately 500,000 cells across multiple time points. Integration with public human fetal midbrain single-cell datasets revealed partial overlap with developmental transition states and mature neuronal populations. Using differentiation state scores derived from reference data, exploratory analyses indicate differentiation-dependent effects of PD risk loci. Ongoing work applies nonlinear Gaussian process latent variable models to identify high-confidence eQTLs and elucidate molecular mechanisms underlying PD pathogenesis.

3. Development of novel statistical approaches to map genetic associations using Gaussian Processes

Yuji Miyatake and Natsuhiko Kumasaka

Gaussian processes (GPs) provide a flexible and powerful framework for modeling nonlinear biological phenomena. This project focuses on developing GP-based methods for genetic association mapping across continuous cellular states and longitudinal outcomes. We are currently developing a robust quasi-Poisson generalized linear mixed model framework that enables sensitive detection of dynamic genetic effects for both Gaussian and non-Gaussian outcomes, while ensuring appropriate statistical calibration.

4. A Genotype-Tissue Expression project using Pan-cancer data

Teru Ebihara[3], Yuji Miyatake, Masahiro Yoshida[4], Kotoe Katayama[1], Seiya Imoto[1] and Natsuhiko Kumasaka
[3]Department of Otorhinolaryngology, The Jikei University School of Medicine, Tokyo, Japan.
[4]Division of Respiratory Diseases, Department of Internal Medicine, The Jikei University School of Medicine, Tokyo, Japan.

The Genotype–Tissue Expression (GTEx) Project has systematically characterized genetic regulation of gene expression across multiple human tissues, but its findings are predominantly based on European populations. To address this gap, we analyze data generated under the AMED Innovative Cancer Therapeutics Program, comprising over 8,000 tissue samples from more than 7,000 Japanese cancer patients with matched RNA sequencing and whole-genome sequencing data.

We are mapping tissue- and cell-type–specific eQTLs by integrating bulk RNA-seq data with state-of-the-art single-cell–based deconvolution methods.

This effort will produce the largest publicly available tissue- and cell-type–resolved eQTL catalogue in an Asian population.

Publications

Kumasaka N, Mezawa H, Sakamoto K, Sato M, Saito-Abe M, Akiba K, Hattori A, et al. Genome-wide association study on longitudinal and cross-sectional traits of child health and development in a Japanese population. medRxiv. doi:10.1101/2025.02.18.25322458, 2025.

Miyaji Y, Yang L, Harama D, Saito-Abe M, Sato M, Sakamoto K, Nishizato M, Kumasaka N, Mezawa H, Ohya Y, Yamamoto-Hanada K; JECS Group. Trajectories of IgE, IgA and IgG Levels in Preschool-Aged Children. Clin Exp Allergy. 55:808–819, 2025.

Yang L, Sato M, Saito-Abe M, Miyaji Y, Harama D, Sakamoto K, Nishizato M, Kumasaka N, Mezawa H, Yamamoto-Hanada K, Ohya Y, Fukami M; JECS Group. Association Between Particulate Matter Exposure and Thyroid Hormone Levels in Early Childhood. J Clin Endocrinol Metab. doi:10.1210/clinem/dgaf446, 2025.

Panousis NI, El Garwany O, Knights A, Rop JC, Kumasaka N, Imaz M, Boquete Vilarino L, Tsingene A, Tokolyi A, Barnett A, Gomez C, Anderson CA, Gaffney DJ. Gene expression QTL mapping in stimulated iPSC-derived macrophages. Nat Commun. 16:7204, 2025.

El Garwany O, Panousis NI, Knights A, Kumasaka N, Imaz M, Boquete Vilarino L, Tsingene A, Tokolyi A, Cotobal Martin C, Alegbe T, Krzak M, Barnett A, Raine T, Anderson CA, Gaffney DJ. Splicing QTL mapping in stimulated macrophages associates low-usage splice junctions with immune-mediated disease risk. Nat Commun. 16:7205, 2025.

Yang L, Sato M, Saito-Abe M, Miyaji Y, Harama D, Sakamoto K, Nishizato M, Kumasaka N, Mezawa H, Yamamoto-Hanada K, Ohya Y, Fukami M; JECS Group. Maternal blood metal levels during pregnancy and childhood obesity. Environ Res. 286:122755, 2025.

Harama D, Saito-Abe M, Pak K, Iwamoto S, Sato M, Miyaji Y, Sakamoto K, Mezawa H, Nishizato M, Yang L, Kumasaka N, Ohya Y, Fukami M, Yamamoto-Hanada K; JECS Group. Not All Rhinitis Follows the Atopic March. Allergy. doi:10.1111/all.70103, 2025.

Izumi K, Motosugi N, Sugiyama A, Kurosaki N, Iida Y, Higashiseto M, Yokoyama K, Sasaki A, Izumi A, Kiumra T, Yamada M, Umezawa A, Akutsu H, Ishimoto H, Kumasaka N, Fukuda A. Upregulation of the lncRNA XACT sustains pluripotency and drives germ cell tumor-like programs. bioRxiv. doi:10.1101/2025.10.02.679492, 2025.

Yanagiya R, Kato H, Ninomiya A, Ueno M, Kanamori A, Miyatake Y, Oka M, Ishii K, Matsuura T, Nakagawa S, Hirao A, Onizuka M, Kotani A. Revised model for cell cycle regulation by iron: differential roles between transferrin and ferritin. Redox Biol. 85:103727, 2025.