Bioinformatics Center
– Chemical Life Science –

https://cls.kuicr.kyoto-u.ac.jp/en/

Prof
OGATA, Hiroyuki
(D Sc)

Assoc Prof
ENDO, Hisashi
(D Environmental Science)

Assist Prof
OKAZAKI, Yusuke
(D Sc)

Assist Prof*
HIKIDA, Hiroyuki
(D Agr)

Distinguished Visiting Assoc Prof
HINGAMP, Pascal Michel
(Ph D)

Program-Specific Res
SATO, Takuya
(D Agr)

Program-Specific Res
MENG, Lingjie

Res (pt)
NECHES, Russell Young
(Ph D)

Guest Res Assoc
GONG, Chunjie
(D Agr)

Proj Res
NGUYEN, Thi Tuyen
OKUDA, Shiho
YAMAGISHI, Yuki

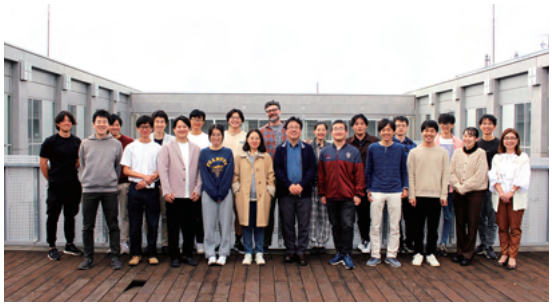
Assist Techn Staff
OGATA, Yuki

*New Research Field Development Project

Students
XIA, Jun (D3)
BAN, Hiroki (D3)
ZHANG, Ruixuan (D2)
YANG, Qingwei (D2)
WU, Junyi (D1)
LIU, Wenwen (D1)
ZHANG, Liwen (M2)
CHEN, Jingjie (M2)
FANG, Yue (M2)
ZHAO, Hongda (M1)
NAGASAKA, Koumei (M1)
TANG, Wei (M1)
KIKUYA, Saki (M1)
YU, Zhaoxi (RS)
SASAKI, Hiroto (UG)
HONDA, Naoki (UG)
QIN, Hongwei*
*Short-Term International Student

Scope of Research

We are interested in understanding the functioning and evolution of biological systems at varying scales from tiny microbes up to the Earth's environment, by leveraging rapidly accumulating big data in life science and bioinformatics approaches. We currently focus on 1) the evolution of viruses and their links to the origin of life, 2) microbial ecology in different ecosystems, and 3) the development of bioinformatics methods and biological knowledge resources for biomedical and industrial applications. To fuel these research activities, we take part in environmental sampling campaigns such as *Tara* Oceans. Our resources and developed tools are accessible through GenomeNet (www.genome.jp) to scientific communities and the public.



KEYWORDS GenomeNet Bioinformatics Environmental Genomics Virology Molecular Evolution

Recent Selected Publications

Gaïa M.; Meng L.; Pelletier E.; Forterre P.; Vanni C.; Fernandez-Guerra A.; Jaillon O.; Wincker P.; Ogata H.; Krupovic M.; Delmont, T. O., Mirusviruses Link Herpesviruses to Giant Viruses, *Nature*, **616,** 783-789 (2023).

Hikida H.; Okazaki Y.; Zhang R.; Nguyen T. T.; Ogata H., A Rapid Genome-Wide Analysis of Isolated Giant Viruses only Using MinION Sequencing, *Environ. Microbiol.*, **25 (11),** 2621-2635 (2023).

Ban H.; Sato S.; Yoshikawa S.; Yamada K.; Nakamura Y.; Ichinomiya M.; Sato N.; Blanc-Mathieu R.; Endo H.; Kuwata A.; Ogata H., Genome Analysis of Parmales, a Sister Group of Diatoms, Reveals the Evolutionary Specialization of Diatoms from Phago-Mixotrophs to Photoautotrophs, *Commun. Biol.*, **6,** 697 (2023).

Kaneko H.; Endo H.; Henry H.; Berney C.; Mahé F.; Poulain J.; Labadie K.; Beluche O.; El Hourany R.; Tara Oceans Coordinators, Chaffron S.; Wincker P.; Nakamura R.; Karp-Boss L.; Boss E.; Bowler C.; de Vargas C.; Tomii K.; Ogata H., Predicting Global Distributions of Eukaryotic Plankton Communities from Satellite Data, *ISME Commun.*, **3,** 101 (2023).

Meng L.; Delmont T. O.; Gaïa M.; Pelletier E.; Fernandez-Guerra A.; Chaffron S.; Neches R. Y.; Wu J.; Kaneko H.; Endo H.; Ogata H., Genomic Adaptation of Giant Viruses in Polar Oceans, *Nat. Commun.*, **14,** 6233 (2023).

Isolation and Whole Genome Sequencing of Giant Viruses from Lake Biwa, Japan

Giant viruses are double-stranded DNA viruses with extremely large genomes and particles reaching 2.5 megabases and 1.5 µm, respectively. These viruses are ubiquitous in the environment and important drivers for nutrient cycles. Currently, various giant viruses have been isolated, but whole-genome sequences of these viruses are limited because of sequencing costs. Here, we evaluated accuracy of giant virus genome assembly by MinION sequencing, which enables rapid and low-cost sequencing. As MinION sequencing produces error-prone reads, the assembly process generally requires correction by other sequencing platforms. However, recent studies assembled high-quality microbial genomes by MinION sequencing alone. We confirmed that genome assembly constructed by MinION sequencing alone is highly accurate for giant viruses with over 99.98% identity to the reference genome by re-sequencing a prototype giant virus. As a proof of concept, we further sequenced five giant viruses isolated from Lake biwa, Japan, by using MinION sequencing. Comparison between newly assembled genomes and reference genomes revealed that these isolates represent new species of marseillevirus, pithovirus, and mimivirus. Overall, we propose that genome assembly by MinION sequencing alone is an effective approach for a genome-wide analysis of isolated giant viruses. This research was published in a journal, *Environmental microbiology* (doi: 10.1111/1462-2920.16476).

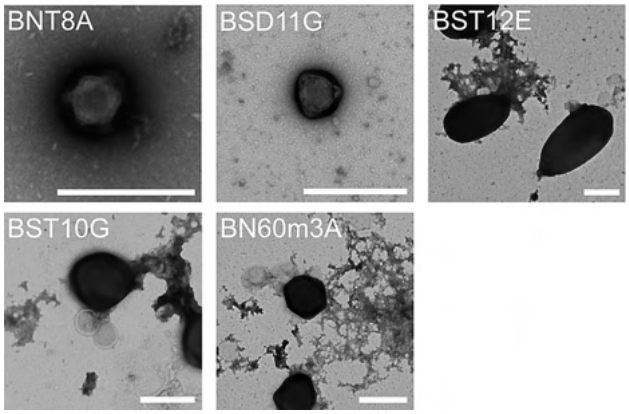


Figure 1. Isolated giant viruses observed under transmission electron microscopy. Names of isolates were shown top left. BNT8A and BSD11G are marseilleviruses. BST12E is a pithovirus. BST10G and BN60m3A are mimiviruses. Bars indicate 500 nm. This figure is modified from that published in Hikida et al. (2023) *Environ. Microbiol.* under the CC BY 4.0 license.

A Giant Endogenous Viral Element: Evidence of Recent Infection of a Double-Stranded DNA Virus in a Fungus

Fungal virosphere is dominated by RNA viruses, with few single-stranded DNA viruses. So far, no double-stranded DNA virus has been identified in fungi. *Rhizophagus irregularis* is a species of arbuscular mycorrhizal fungus, belonging to the class *Glomeromycetes*. We searched viral signals within chromosome-level genomic assemblies of five different strains of this fungi species. On chromosome 8 of a strain 4401, we discovered a 1.5-megabase region, which showed distinct features from other chromosomal regions with strong viral signals. This viral region harbors five *Nucleocytoviricota* marker genes. Our phylogenetic analysis revealed that these genes are closely related to *Asfarviridae* sequences (Fig. 1). These marker genes remain in single copy and show the same tree topology, which suggest that this viral region originated from a single viral integration event. In the chromosomes of other strains, this viral region is absent. Also, the content of repetitive sequences and transposable elements in this viral region is lower compared to other genomic regions. These findings suggest that the viral region was inserted after the divergence of five *R. irregularis* strains. This work has been published in Zhao et al., Virus Evolution, 2023, doi: 10.1093/ve/vead064.

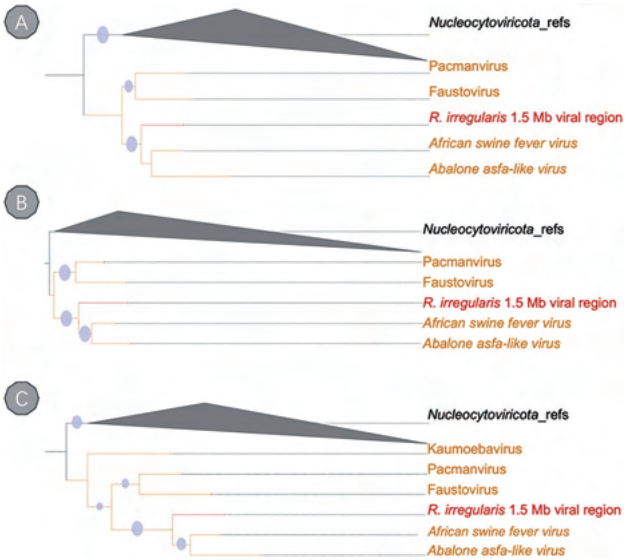


Figure 2. Phylogenetic trees based on core genes of viruses of the phylum *Nucleocytoviricota* based on maximum-likelihood frameworks. Blue circles represent supports that passed confidence cutoff for branches. Yellow represents asfarviruses, and red indicates sequences from the viral region. (a) Concatenated tree of three polymerases. (b) Tree of mRNA capping enzyme (c) Tree of viral late transcription factor 3.