Bioinformatics Center
– Chemical Life Science –

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

Prof
OGATA, Hiroyuki
(D Sc)

Assist Prof
ENDO, Hisashi
(D Environmental Science)

Assist Prof
OKAZAKI, Yusuke
(D Sc)

PD
HIKIDA, Hiroyuki
(D Agr)

Guest Res Assoc
SHIH, Chi-Yu
(Ph D)

Proj Res
NGUYEN, Thi Tuyen

Reseachers (pt)
NISHIYAMA, Hiroki (D Pharm Sc)

Assist Techn Staff
ZHU, Mengdi

Students
PRODINGER, Florian (D3)
KANEKO, Hiroto (D3)
MENG, Lingjie (D3)
XIA, Jun (D3)
BAN, Hiroki (D1)
GARCIA, Kimberly Gacayon (M2)
ZHANG, Ruixuan (M2)
HASHIMOTO, Kentaro (M2)
KIJIMA, Soichiro (M2)
FANG, Yue (M1)
WU, Junyi (M1)
YAMAMOTO, Tomokazu (M1)
YANG, Qingwei (RS)
ZHANG, Liwen (RS)
CHEN, Jingjie (RS)
LIU, Wenwen (RS)
ZHAO, Hongda (RS)
IDE, Shiina (UG)
YAMADA, Kohei (UG)

Guest Res Assoc
SHIH, Chi-Yu (Ph D) National Taiwan Ocean University, Taiwan, 23 August, 2021–31 July, 2022

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

Zhang, R.; Endo, H.; Takemura, M.; Ogata, H., RNA Sequencing of Medusavirus Suggests Remodeling of the Host Nuclear Environment at an Early Infection Stage, *Microbiol. Spectr.*, e0006421, doi: 10.1128/Spectrum.00064-21 (2021).

Kijima, S.; Delmont, T. O.; Miyazaki, U.; Gaia, M.; Endo, H.; Ogata, H., Discovery of Viral Myosin Genes with Complex Evolutionary History within Plankton, *Front. Microbiol.*, **12,** 683294, doi: 10.3389/fmicb.2021.683294 (2021).

Meng, L.; Endo, H.; Blanc-Mathieu, R.; Hernández-Velázquez, R.; Kaneko, H.; Ogata, H., Quantitative Assessment of NCLDV–Host Interactions Predicted by co-Occurrence Analyses, *mSphere*, **6,** e01298-20 (2021).

Kaneko, H.; Blanc-Mathieu, R.; Endo, H.; Chaffron, S.; Delmont, T. O.; Gaia, M.; Henry, N.; Hernández-Velázquez, R.; Nguyen, C. H.; Mamitsuka, H.; Forterre, P.; Jaillon, O.; de Vargas, C.; Sullivan, M.B.; Suttle, C. A.; Guidi, L.; Ogata, H., Eukaryotic Virus Composition Can Predict the Efficiency of Carbon Export in the Global Ocean, *iScience*, **24,** 102002 (2021).

Endo, H.; Blanc-Mathieu, R.; Li, Y.; Salazar, G.; Henry, N.; Labadie, K.; de Vargas, C.; Sullivan, M. B.; Bowler, C.; Wincker, P.; Karp-Boss, L.; Sunagawa, S.; Ogata, H., Biogeography of Marine Giant Viruses Reveals Their Interplay with Eukaryotes and Ecological Functions, *Nat. Ecol. Evol.*, doi: 10.1038/s41559-020-01288-w (2020).

A Giant Virus Remodels the Host Nuclear Environment

Medusavirus, a recently isolated giant virus, has a nucleocytoplasmic replication cycle in amoebas during which the host nuclear membrane apparently remains intact, a unique feature among amoeba-infecting giant viruses. To investigate its infection strategy, we performed a time course RNA sequencing experiment. All viral genes were transcribed and classified into five temporal expression clusters. The immediate early genes (42 genes) were mostly (83 %) of unknown functions, frequently (95 %) associated with a palindromic promoter-like motif, and often (45 %) encoded putative nucleus-localized proteins. These results suggest massive reshaping of the host nuclear environment by viral proteins at an early stage of infection. Genes in the following clusters were assigned to various functional categories. The different expression profile between viral core histone genes and linker histone H1 gene suggests that they have distinct roles during the course of the virus infection. The transcriptional profile of the host *Acanthamoeba castellanii* genes was greatly altered postinfection. This work has been published in Zhang et al., *Microbiol. Spectr.*, 2021, doi: 10.1128/Spectrum.00064-21.

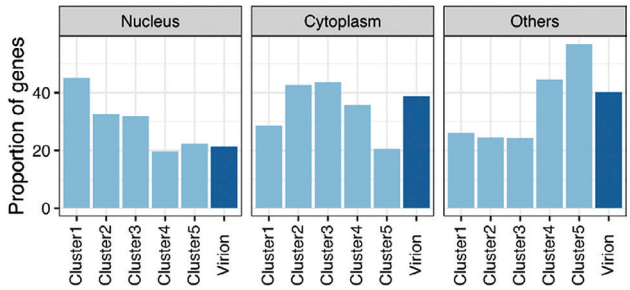

Figure 1. Predicted subcellular locations of viral proteins.

A Motor Protein Gene Discovered in Giant Viruses

Nucleocytoplasmic large DNA viruses (NCLDVs) are a group of giant viruses. A typical phage particle is 150 nm and encapsulates a tens of thousands base pairs (bp) genome. In contrast, Mimivirus, a member of NCLDVs, is 750 nm in diameter and possesses a genome over 1,000 kbp. Although most of phage genes are essential for their replication and morphogenesis, NCLDVs sometimes encode genes, such as actin genes, that are apparently non-essential for their replication. In this study, we discovered myosin genes, named "virmyosin", from NCLDV genomes. Our phylogenetic analysis revealed that some *Imitervirales* and *Phycodnaviridae* encode virmyosin. Especially, virmyosin-encoding *Imitervirales* are scattered in the phylogeny of *Imiternirales* family. Further investigation suggested that virmyosin of *Imitervirales* might have been recruited from marine plankton. Additional study showed an incongruence between the phylogenetic tree of NCLDVs and that of virmyosins, which suggested that horizontal gene transfers occurred not only between NCLDVs and eukaryotes but also between members of NCLDVs for the actin gene. We predicted the function of virmyosin in host cells based on previous studies. It has been known that myosins are motor proteins to walk along actin filaments, located at the peripheric side of cytoplasm. Some classes of myosin carry specific materials such as influenza virus. We thus consider that virmyosins may function as transporters of NCLDV particles. This work has been published in Kijima et al., *Front. Microbiol.*, 2021, doi: 10.3389/fmicb.2021.683294.

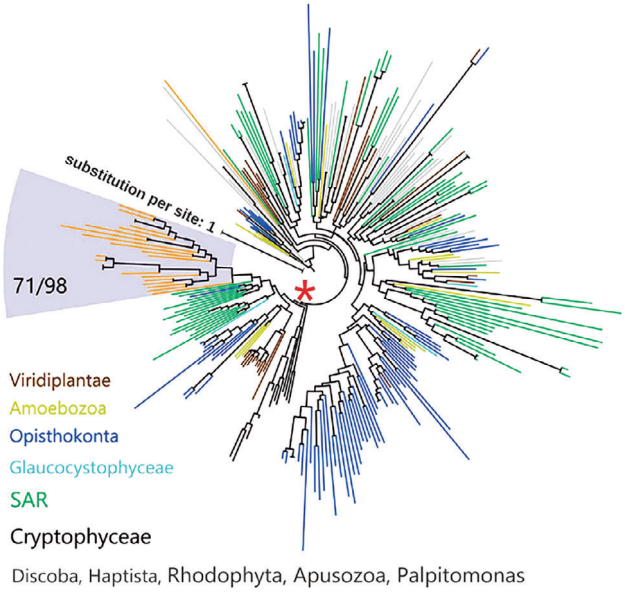

Figure 2. Phylogenetic tree of myosin genes of NCLDVs and eukaryotes, based on a maximum-likelihood framework. Color of each branch represents its taxonomic group (orange branches correspond to virmyosins).