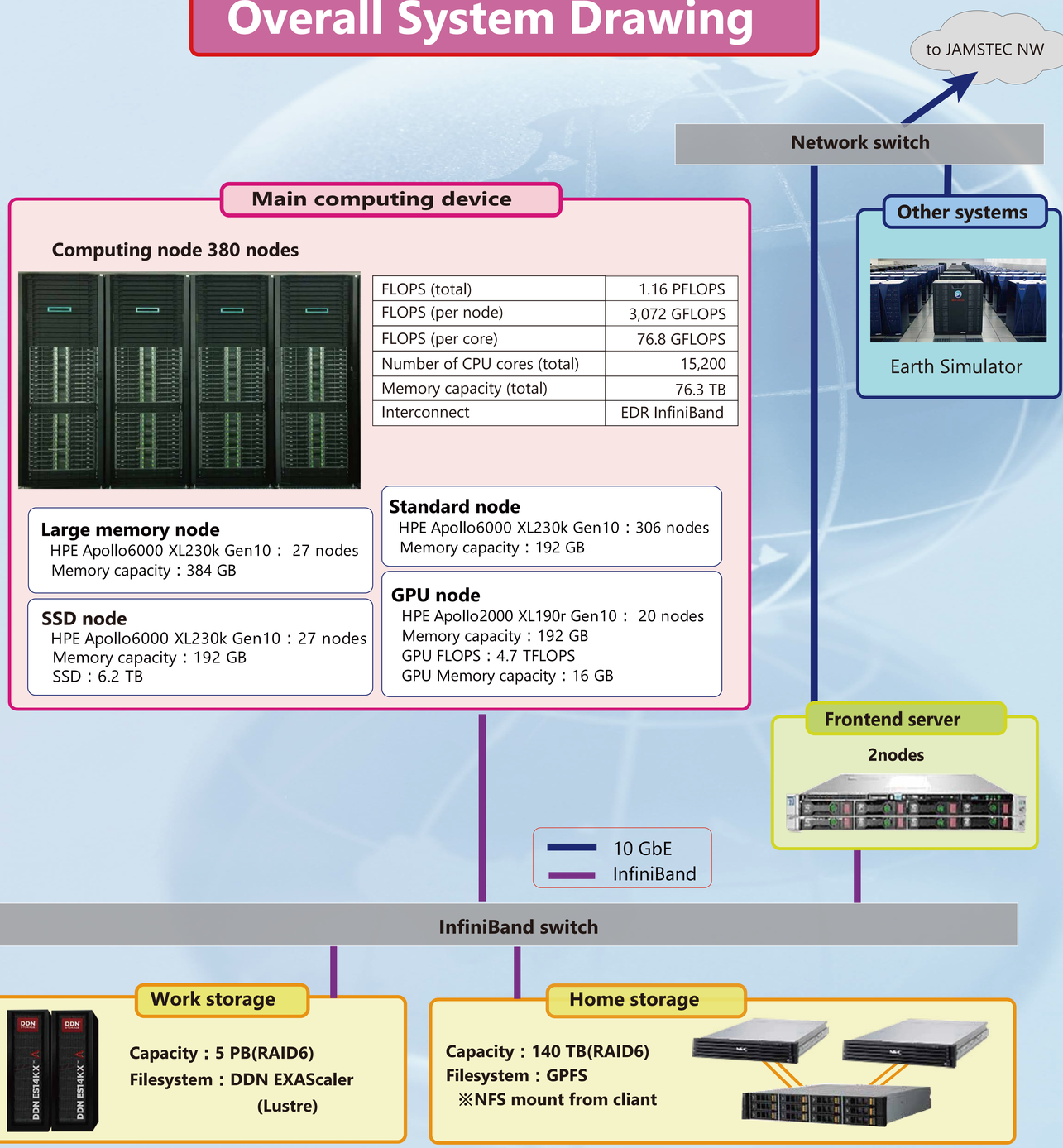
Data Analyzer System

DA System Overview

HPE Apollo 6000 and HPE Apollo 2000
CPU：Intel Xeon 2.4 GHz、20 core
Number of nodes ：380
Number of CPU sockets：760
Number of cores ：15,200
Total memory capacity ：76 TB
Home storage capacity ：140 TB
Work storage capacity ：5 PB

Interconnect：EDR InfiniBand 100Gbps
Specifications of computing nodes
Standard node：306 nodes
Large memory node：27 nodes memory 384 GB
(capacity is double.)
SSD node：27 nodes、SSD 6TB
GPU node：20 nodes、NVIDIA TESLA P100

Overall System Drawing



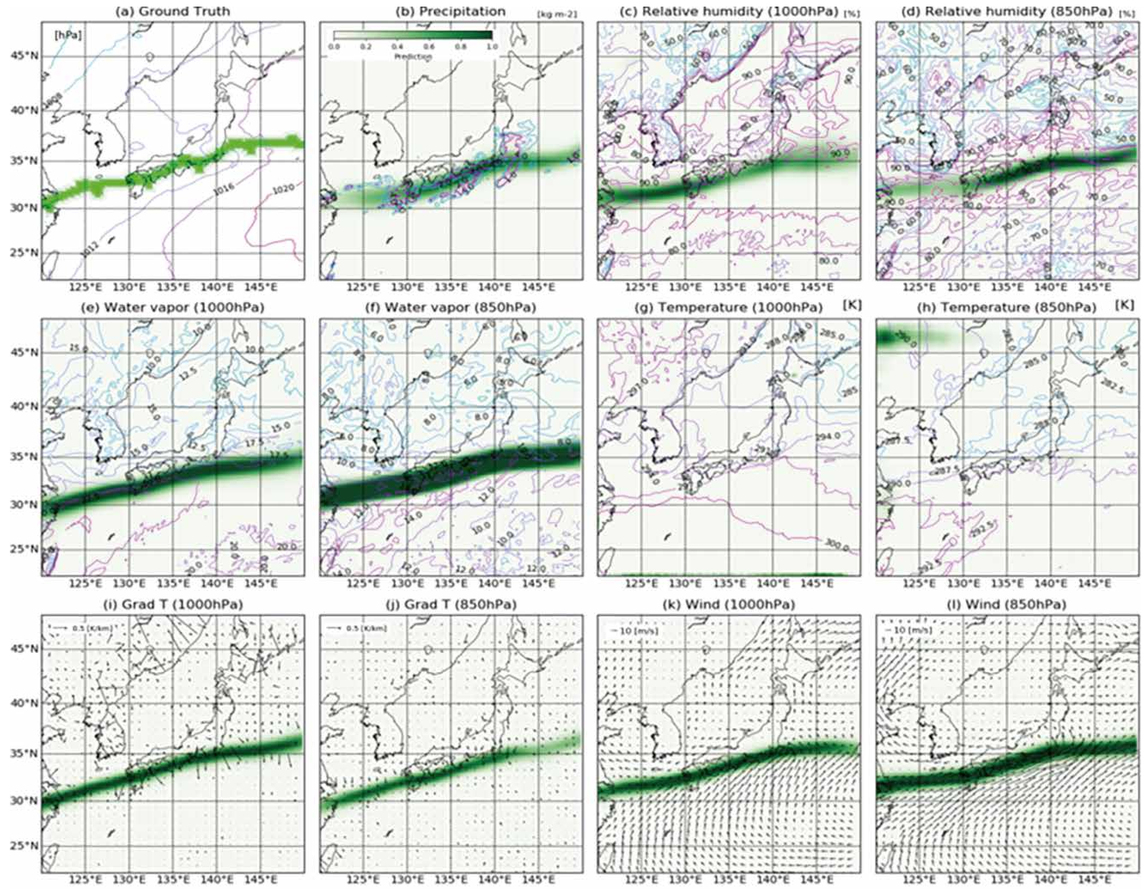
Data Analyzer System

Research Results

Meteorology and Deep Learning

Daisuke Matsuoka Information Engineering Program (IEP), Research Institute for Value-Added Information Generation (VAiG), JAMSTEC

Feature extraction from observational /simulation data using deep learning.
By training a huge amount of data, deep learning approach shows good performance for detection of precursors of tropical cyclones (left) and automatic drawing of stationary fronts (right).

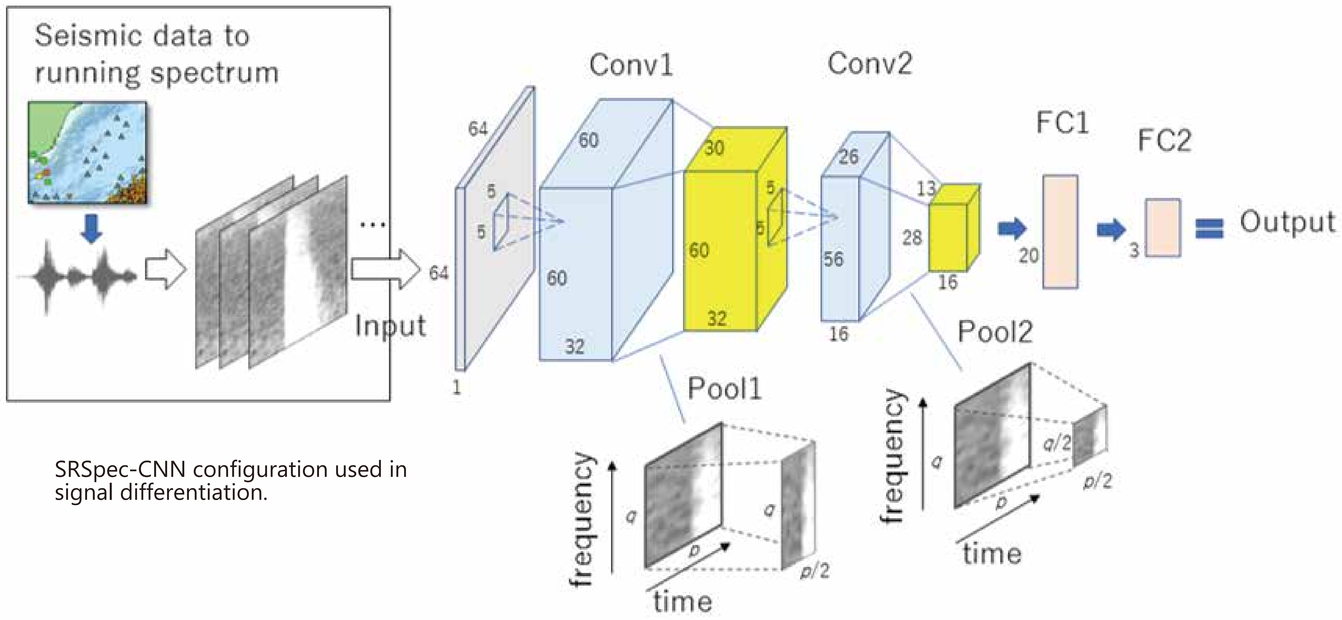


Data Analyzer System

Research Results

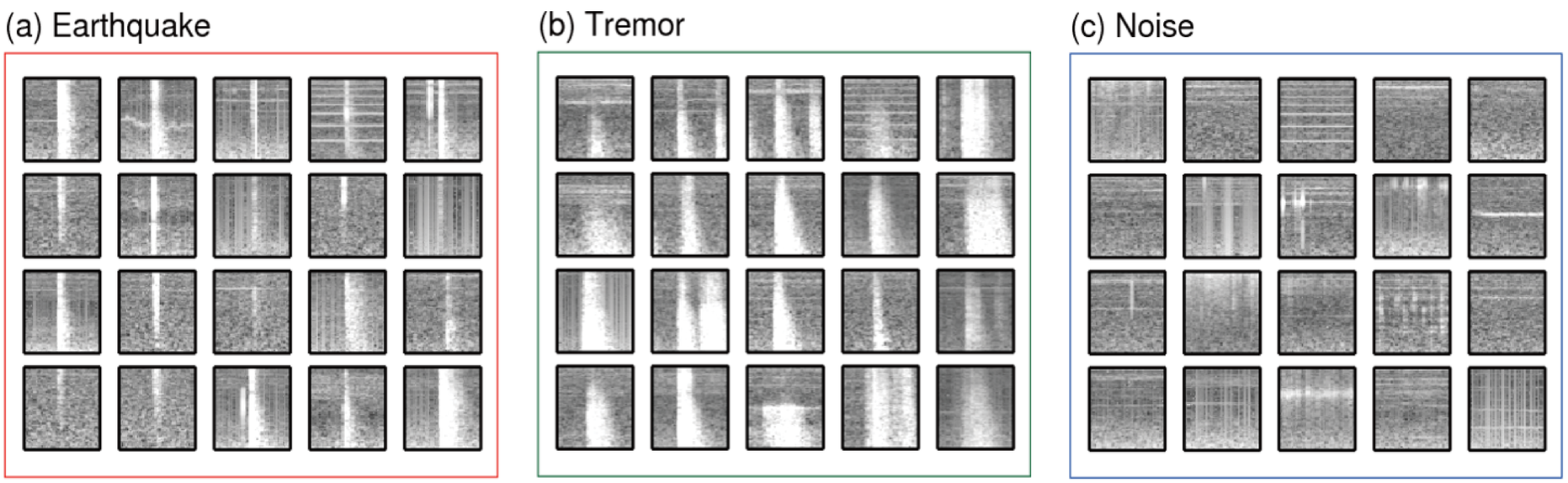
Earthquake and Deep Learning

Daisuke Sugiyama Information Engineering Program (IEP), Research Institute for Value-Added Information Generation (VAiG), JAMSTEC



SRSpec-CNN configuration used in signal differentiation.

Developing a New Method "SRSpec-CNN" of Automatically and Accurately Discriminating between Regular Earthquakes and Low-frequency Tremor Signals Using Artificial Intelligence.
SRSpec-CNN was used to differentiate between signals from low-frequency tremors and regular earthquakes recorded via seismographs of the Dense Oceanfloor Network System for Earthquakes and Tsunamis (DONET) observation system, deployed along the Nankai Trough. The results demonstrated a remarkably high accuracy of 99.5%.



Examples of running spectral images used to differentiate signals.

Data Analyzer System

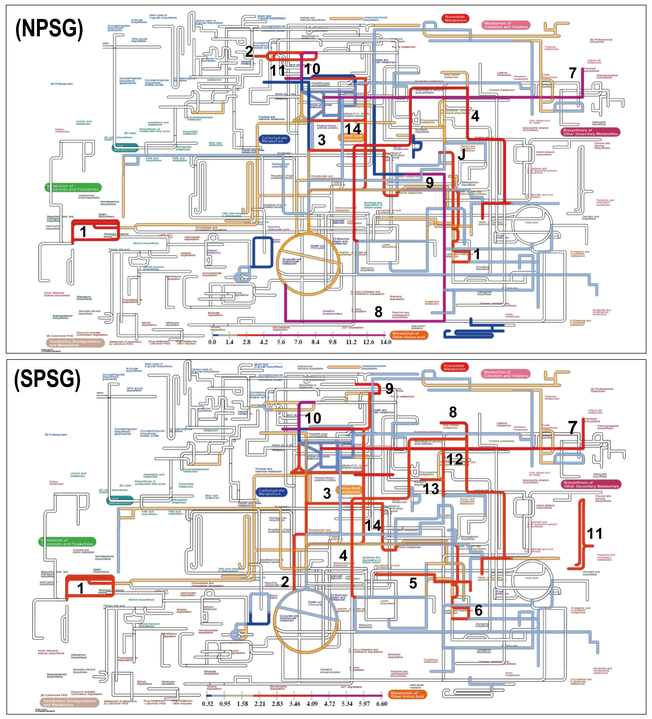
Research Results

Microbial Community and Their Function

Hideto Takami Marine Biodiversity and Environmental Assessment Research Center (BioEnv) , Research Institute for Global Change (RIGC) , JAMSTEC

Prokaryotic communities in the particle-associated and free-living assemblages of surface seawater at the North and South Pacific subtropical gyres and eastern equatorial Pacific regions were analysed using metabolic and physiological potential evaluator (Genomemaple). We show the particle-associated and free-living fractions could be discriminated from one another by their taxonomic compositions inferred from ribosomal proteins and relative module abundance. Module functions that were more abundant among particle-associated assemblages than the free-living assemblages were shared between both subtropical gyres, and their taxonomic compositions were similar. Bacterial transport systems associated with adhesive molecules were more abundant in the particle- associated assemblages, which adhered to particulate organic matter. Bacterial regulatory system elements for C4-dicarboxylate transport were also abundant among particle-associated assemblages, suggesting Proteobacteria depend on diatoms as a nutrient source. Based on our findings, we recommend that community structure analyses be conducted with ribosomal proteins instead of problematic 16S rRNA genes.

Genomaple system used in this study is now available at the following web site.
(https://maple.jamstec.go.jp/maple/maple-2.3.1/).

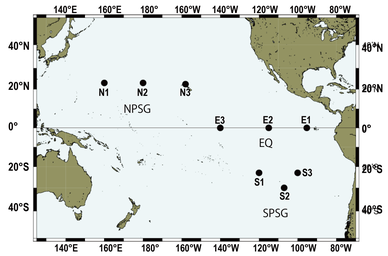


(P) Small catefory	ID	map	Module function	PA/FL ratio NPSG	PA/FL ratio SPSG
(Amino acid metabolism)	M00024	13	Phenylalanine biosynthesis	-	2.46
Aromatic a.a. metabolism	M00025	12	Tyrosine biosynthesis	-	2.01
BCAA metabolism	M00036	5	Leucine degradation	-	2.22
Serine & threonine metabolism	M00020	3	Serine biosynthesis, glycerate>3P => serine	2.98	2.27
Cysteine & methionine metabolism	M00035	6	Methionine degradation	2.62	2.39
	M00338	5	Cysteine biosynthesis	2.41	-
Histidine metabolism	M00045	9	Histidine degradation	5.14	-
Other a.a. metabolism	M00027	8	GABA (gamma-Aminobutyrate) shunt	6.70	-
(Carbohydrate metabolism)	M00008	14	Entner-Doudoroff pathway	2.62	3.05
Central carbohydrate metabolism	M00003	2	Gluconeogenesis	-	2.20
Other carbohydrate metabolism	M00549	10	Nucleotide sugar biosynthesis	6.77	6.31
	M00554	11	Nucleotide sugar biosynthesis	2.14	-
	M00632	2	Galactose degradation	2.48	-
	M00631	9	D-galacturonate degradation (bacteria)	-	2.91
Cofactor & vitamin biosynthesis	M00124	7	Pyridoxal biosynthesis	5.64	3.27
	M00122	11	Cobalamin biosynthesis	-	2.55
	M00119	4	Pantothenate biosynthesis	3.49	2.53
Pyrimidine metabolism	M00046	8	Pyrimidine degradation	-	3.35
Terpenoid backbone biosynthesis	M00364	1	C10-C20 isoprenoid biosynthesis, bacteria	3.27	2.91

(C) Small catefory	ID	Module function	PA/FL ratio NPSG	PA/FL ratio SPSG
Mineral & organic ion transport	M00189	Molybdate transport system	3.62	3.55
ABC-2 type & other transport systems	M00250	Lipopolysaccharide transport system	2.63	3.26
	M00256	Cell division transport system	2.36	3.13
	M00253	Sodium transport system	5.63	-
	M00259	Heme transport system	2.84	-
Saccharide, polyol, & lipid transport system	M00212	Ribose transport system	2.22	-
Metallic cation, iron-siderophore & vitamin B12 transport	M00240	Iron complex transport system	3.43	-
	M00319	Manganese/zinc/iron transport system	-	3.70
Phosphate & amino acid transport system	M00236	Polar amino acid transport system	2.03	-
Bacterial secretion system	M00330	Adhesin protein transport system	2.11	2.15
Photosynthesis	M00597	Anoxygenic photosystem II	2.73	2.71
RNA processing	M00394	RNA degradosome	11.15	2.97
Proteasome	M00343	Archaeal proteasome	2.36	-
Phosphotransferase system (PTS)	M00273	PTS system, fructose-specific II	5.00	-

(F) Small catefory	ID	Module function	PA/FL ratio NPSG	PA/FL ratio SPSG
Two-component regulatory system	M00447	(CpxA-CpxRenvelope stress response)	5.20	3.36
	M00504	DctB-DctD (C4-dicarboxylate transport)	14.11	3.06
	M00434	PhoR-PhoB (phosphate starvation)	-	2.11
	M00501	PilS-PilR (type 4 fimbriae synthesis)	6.83	-
	M00475	BarA-UvrY (central carbon metabolism)	5.04	-
	M00499	HydH-HydG (metal tolerance)	4.92	-
	M00459	VicK-VicR (cell wall metabolism)	3.24	-

Comparison of relative abundance ratios between PA and FL fractions. PA/FL is the ratio of abundance of the PA fraction to that of the FL fraction. Abundance ratios were calculated based on the relative abundance of the modules commonly completed in both regions, and only modules sho-wing 2-fold or higher abundance ratios are highlighted in this figure. Light-blue shading indicates the modules commonly having a higher ratio in both the NPSG and SPSG regions. Warm and cold colours indicate higher and lower abundance ratios, respectively, in the coarse-grained meta-bolic map. This map was created based on the KEGG Atlas map.



Location map of sampling sites.

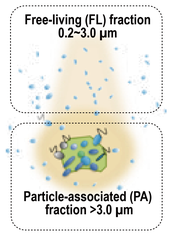
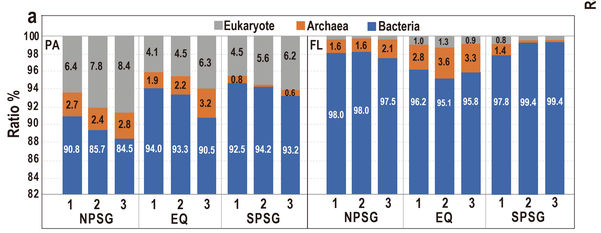
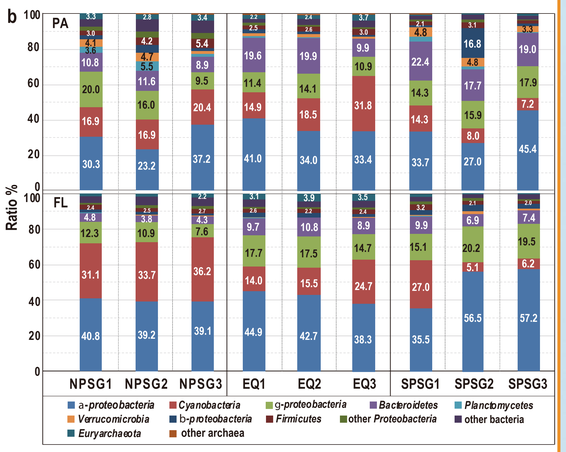


Diagram of sampling fractions





Microbial community structure analysis of sea surface water in Pacific Ocean based on ribosomal proteins using Genomaple system.
a. Comminity structure at dimain level, b. Prokaryotes