A model ecosystem of twelve cryopreservable microbial species allowing for a non-invasive approach

Kazufumi Hosoda[1*], Shigeto Seno[2], Naomi Murakami[1], Hideo Matsuda[2], Yutaka Osada[3], Michio Kondoh[4*]

[1]Institute for Transdisciplinary Graduate Degree Programs, Osaka University, 1-5 Yamadaoka, Suita, Osaka 565-0871, Japan.
[2]Graduate School of Information Science and Technology, Osaka University, 1-5 Yamadaoka, Suita, Osaka 565-0871, Japan.
[3]Fisheries Resources Institute, 2-12-4 Fukuura, Kanazawa, Yokohama, Kanagawa 236-8648, Japan.
[4]Graduate School of Life Sciences, Tohoku University, 6-3 Aoba, Aramaki, Aoba-ku, Sendai 980-8578, Japan

*Corresponding authors: Kazufumi Hosoda (hosoda@humanware.osaka-u.ac.jp), Michio Kondoh (michio.kondo.b8@tohoku.ac.jp)
Keywords: Synthetic ecosystem, Experimental model ecosystem, Microbial microcosm, Machine learning

Abstract

We developed a synthetic ecosystem of 12 cryopreservable microbial species with diverse interactions as an experimental "model ecosystem." We created a machine learning model that noninvasively distinguished the 12 species on micrographs enabling high-throughput measurements. Our synthetic ecosystems maintained a certain diversity for at least six months.

Main text

Experiments on ecosystems have intrinsic difficulties in handling, reproducibility, and experimental commonality compared with individual organisms, where large-scale experiments using some "model organisms" have been commonly conducted. Species-defined microbial experimental ecosystems, called gnotobiotic microcosms (Beyers and Odum, 1993, Nixon, 1969, Taub, 1969), have been used or recognized as experimental "model ecosystems" not only for understanding the species composition (Tanaka et al., 1994), community assembly (Drake, 1991, Fukami, 2015), trophic structure (Naeem and Li, 1998), and dynamics of complex communities in microbiota (Grosskopf and Soyer, 2014, Ponomarova and Patil, 2015, Tan et al., 2015), but also for the application to bioprocessing (Brenner et al., 2008, Shong et al., 2012) and to environmental problems such as the effect of chemicals or genetically-modified organisms (Benton et al., 2007, Fuma et al., 2005, Inamori, 2020).

A gnotobiotic microcosm consisting of organisms that can be axenically-culturable and cryopreservable is an ideal model ecosystem (Kawabata et al., 1995), in which the organism composition is freely designable and the experiments are completely reproducible. In fact, such gnotobiotic cryopreservable-microbial microcosms (GCMs) have been applied for the understanding of evolution (Momeni et al., 2011, Nakajima et al., 2009), historical contingency (Chuang et al., 2019, Hekstra and Leibler, 2012), and higher-order microbial interactions (Mickalide and Kuehn, 2019). However, previously-reported GCMs with trophic relationships included no more than three species of organisms: typically each single species of producer, decomposer, and predator. Such systems are too simple to study diversity-dependent characteristics of ecosystems, which are often targeted in theoretical studies of ecological community (May, 1972, Mougi and Kondoh, 2012).

In this study, we developed a GCM consisting of phylogenetically-diverse species with various interspecific interactions. Specifically, the GCM comprised twelve species of microorganisms including eukaryotic consumers and producers and prokaryotic consumers and producers (Table 1). All twelve species are axenically-culturable and cryopreservable, and their genomes are known. Diverse interspecific interactions can exist among these twelve species, including competition, mutualism, parasitism (benefit-harm interactions), and predation and herbivoration (foraging interactions).

Table 1. Species used.

		Species	Classification	Strain	Shape	Fluorescence
Decomposers	#0	*Escherichia coli*	Proteobacteria	DH1ΔilvE::(dsred.T3-cat)[a]	Rod	ex.485/em.515
	#1	*Staphylococcus epidermidis*	Firmicutes	JCM2414[b]	Spherical	(−)
	#2	*Schizosaccharomyces pombe*	Ascomycota	L972[c]	Rod	(−)
	#3	*Saccharomyces cerevisiae*	Ascomycota	BY25611[d]	Oval	ex.540/em.580
	#4	*Candida parapsilosis*	Ascomycota	JCM1612[b]	Oval/Filamentous	(−)
Predators	#5	*Dictyostelium discoideum*	Amoebozoa	AX2[e]	Amoeba	(−)
	#6	*Tetrahymena thermophila*	Ciliophora	CU427&CU428[f]	Oval	(−)
Producers	#7	*Anabaenopsis circularis*	Cyanobacteria	NIES21[g]	Filamentous	ex.615/em.655
	#8	*Synechocystis* sp.	Cyanobacteria	PCC6803[h]	Spherical	ex.615/em.655
	#9	*Raphidocelis subcapitata*	Chlorophyta	NIES35[g]	Crescent	ex.480/em.685
	#10	*Chlorella vulgaris*	Chlorophyta	NIES227[g]	Spherical	ex.480/em.685
	#11	*Chlamydomonas reinhardtii*	Chlorophyta	NIES2236[g]	Oval	ex.480/em.685

(a) Constructed from DH1 (Hosoda et al., 2011)
(b) Provided by Japan Collection of Microorganisms (JCM), RIKEN BRC, Japan
(c) A kind gift from Dr. Kojiro Ishii of Kochi University of Technology
(d) BY4741ho::Kan-TEFIIpr-GFP-ADH1ter (Breslow et al., 2008), provided by the National Bio-Resource Project (NBRP), Japan
(e) DBS0235521, provided by the Dicty Stock Center, Northwestern University, USA
(f) A conjugation progeny of CU427 and CU428, provided by Tetrahymena Stock Center, Cornell University
(g) Provided by NIES collection, Japan
(h) ATCC27184, provided by the American Type Culture Collection (ATCC), USA

A prominent feature of this experimental system is that populations of all twelve species are noninvasively measurable using a machine learning on micrographs of microcosms. Noninvasive measurements has several advantages: (i) applicability to completely closed systems, (ii) high-throughput capacity by omitting sample collection, and (iii) avoidance of system-volume reduction and/or population disturbance due to sample collection. Specifically, for microcosm experiments, we prepared 50 μL of culture solution in each well of a 384-well plate (#142761; Nunc, USA) and sealed the plate with a heat adsorption seal (#4ti-05481; 4titude, UK). Before microscopy, the plate was shaken by inversion, spun down, and left for 3 h. We scanned each well using an inverted microscope Ti-E with the Perfect Focus System and the High Content Analysis (Nikon, Japan), and took bright-field and fluorescent-field micrographs of each well's center from the bottom using a 60X objective lens. We used YOLOv3 (Redmon and Farhadi, 2018) for detecting the organisms (details are shown in supplementary data, Text S1).

With this method, all twelve species were correctly detected and identified (Figure 1A). The errors were less than 1% for 9 of twelve species. The errors greater than 1% for both protozoa (#5 and #6 in Table 1) were due to the false detection of tiny particles from protozoan as cocci. This misidentification can be avoided in the future by fluorescently labeling cocci. In any case, both protozoa were rarely dense enough for these particles to be detected and they were rarely erroneously detected as cocci in microcosm experiments.

Figure 1B shows the standard curve of each species. We found that there was a range where the counted amount correlates with the input amount for all species. Such range was roughly 10^4–10^6/mL (500–50,000/well or 1–100/microscopic field), which can be applied in semi-open microcosm experiments shown below (except the protozoa). Unlike most bioanalytical methods, the lower detection limit mainly depends on

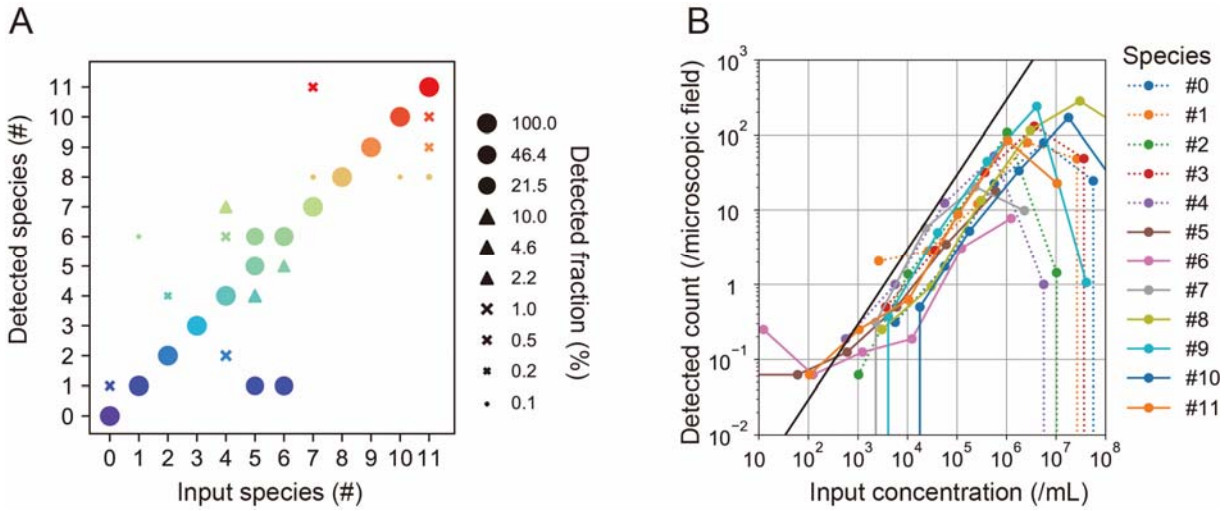

Figure 1. Detection of each species using the machine learning model. The data used here are different from that used for training. (A) Detection accuracy. The fractions of detected species are shown on the vertical axis when we tested monocultures of species on the horizontal axis. We used the data of 36 replicated cultures for a single species. (B) Calibration curve for each of the twelve species. The mean value of 16 replicated wells is plotted. The black line represents the value when the input organisms are ideally counted. Details such as the deviation for each plot are shown in supplementary data (Figure S2).

the presence of the organisms in the micrograph, as the noise at the blank is very little. The detection efficiency (difference from the black line) within the range was greater than 10% for most species. The efficiency depends not only on the machine learning but also on the fraction of organisms present at the center of the well's bottom (Figure S2). At a high concentration, the detected counts decreased steeply because the organisms overlap in three dimensions and are indistinguishable.

For the microcosm experiments, we tested a closed system (organisms were mixed and kept closed) and a semi-open system (the system was opened once every two weeks and transferred to a new environment at a dilution of 1/10 with fresh medium). We used a liquid medium BG11HLB, which contains half the concentration of BG-11 medium for cyanobacteria (Allen, 1968) and 1/100 concentration of LB medium for bacteria (Bertani, 1951). We placed the 384-well plate on a white LED panel (TH-224X170SW; CCS, Japan) in an incubator at 23°C and irradiated at an intensity of 50 $\mu mol \cdot m^{-2} \cdot s^{-2}$ for 12 h intervals.

In the closed systems, we found experimental conditions that maintained at least 6 diverse species for 6 months. It was impossible to measure the population using the machine learning due to a high concentration in these conditions. Instead, we used a fluorescence plate reader (Varioskan Flash; Thermo, USA) to quantify the concentration of Cyanobacteria, Chlorophyta, and *E. coli* (red fluorescent protein-labeled) from the corresponding fluorescence (Table 1). We also quantified the concentration of *T. thermophila* (a predator) from the temporal variation of low-magnification micrographs (using a 4X objective lens, whose viewfield captured the entire well) because there was a correlation between *T. thermophila* concentration and the temporal variation since *T. thermophila* is large and swims. We mixed the twelve species at concentrations that were 1/200 of the saturation concentration in monocultures. In 3 of the 64 tested wells, *E. coli* (bacterial decomposer), *T. thermophila* (eukaryotic predator), Cyanobacteria (bacterial producers), and Chlorophyta (eukaryotic producers) were maintained. We confirmed the existence of *C. parapsilosis*, *Synechocystis sp.*, *R. subcapitata*, and *C. vulgaris*, in addition to *E. coli* and *T. thermophila* (thus at least 6 species) in all 3 wells by observing the wells immediately after shaking by inversion. The timecourse of one of the three wells is shown in Figure 2A. Moreover, we found that *E. coli*, *T. thermophila*, Cyanobacteria, and Chlorophyta were maintained in all 64 wells tested when we excluded *C. parapsilosis* (*i.e.*, the mixture of 11 species, as a more stable condition).

In semi-open systems, the machine learning quantification worked in most conditions tested. When twelve species were mixed at the same initial concentration of organisms as in closed systems, the predators were absent in all 64 tested wells after 6 months. In the 11-species microcosm (except *C. parapsilosis*), *T.*

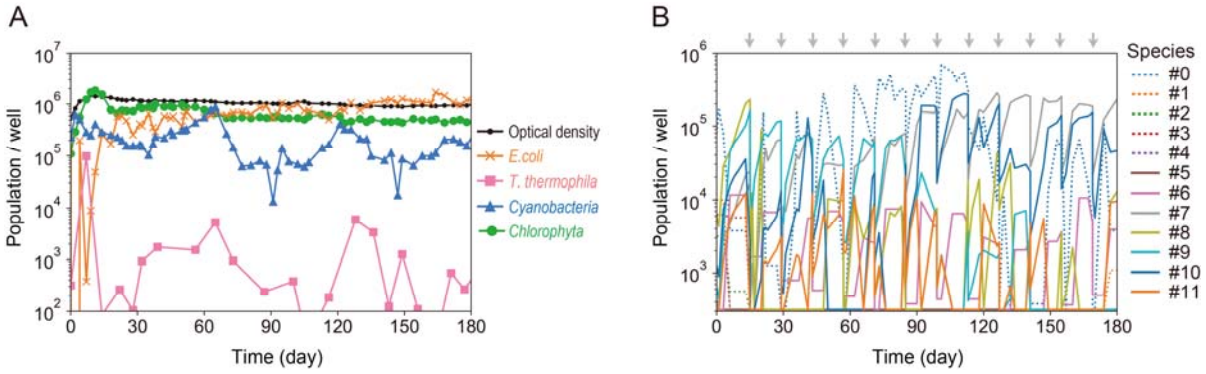

Figure 2. Population dynamics of microcosms. (A) Representative example of the closed systems starting from 12 species. The values of optical density, Cyanobacteria fluorescence, and Chlorophyta fluorescence were linearly converted to values corresponding to the population of *C. parapsilosis*, *Synechocystis sp.*, and *C. vulgaris*, respectively. (B) Representative example of the semi-open systems starting from 11 species. Gray arrows depicted at the top of the graph show the time point of subculturing, conducted approximately once every 2 weeks. The population of each species was determined by the combination of 3 noninvasive measurements using the machine learning, the fluorescence plate reader, and the temporal variation of micrographs (see Text S2 for details).

thermophila (a predator) was maintained in all 64 wells tested, and at least one Cyanobacteria and Chlorophyta were maintained in 9 wells. Furthermore, in this 11-species microcosm, when the initial concentration of all organisms was 1/10, *T. thermophila* and at least one Cyanobacteria and Chlorophyta were maintained in 28 of 72 wells tested. Of these 28 wells, *E. coli*, *A. circularis*, *Synechocystis sp.*, *R. subcapitata*, and *C. vulgaris* were observed in 2, 18, 21, 6, and 28 wells, respectively. Bacteria were not detected in most of the wells where *T. thermophila* survived, but *E. coli* was detected in all wells tested when 8 of these wells were transferred to the LB agar plate, indicating that *E. coli* was present below the detection limit. Figure 2B shows the population dynamics of one of the 28 wells. We confirmed the existence of *E. coli*, *T. thermophila*, *A. circularis*, *Synechocystis sp.*, and *C. vulgaris* by the machine learning and eye (the detection of *C. reinhardtii* by the machine learning, shown in Figure 2B, would be a false positive due to enlarged *C. vulgaris*). We checked that we obtained similar results by mixing these 6 species from the beginning.

In conclusion, we developed a GCM with diversity where all species were noninvasively quantified. We found experimental conditions that allowed the maintenance of at least 6 and 5 species for 6 months in closed and semi-open systems, respectively. Our GCM enables us to test large numbers of ecosystems with many species combinations and environmental conditions at once, reproducibly. We can use large data obtained for time series analysis and comparison with mathematical dynamics models. Moreover, as all species are cryopreservable, we can study the evolution of each organism and symbiotic pairs within the microcosm by measuring genetic and phenotypic changes. Our GCM alone is not enough for understanding all natural ecosystems, but is useful for highlighting features that underlie ecosystems with certain conditions. In the future, we will update our GCM with improvements in diversity, measurement accuracy, and experimental convenience as a standard experimental model ecosystem.

Acknowledgments

We thank Drs. Chikara Furusawa, Kojiro Ishii, and Toshiyuki Nakajima for their valuable advice, Mr. Yoshinori Hiratani and Dr. Tatsuya Nakamura for their technical assistance. We are also grateful to Dr. Isamu Taguchi and the other 108 supporters for their financial support through academist crowdfunding. This research was supported in part by JSPS KAKENHI Grant Numbers JP18H04821, JP20H05533, JP20H04868, and JP20K06825, and the "Program for Leading Graduate Schools" of the ministry of education, culture, sports, science and technology (MEXT) Japan. The authors declare no conflicts of interest.

Author contributions

Kazufumi Hosoda: Conceptualization, Methodology, Investigation, Writing - Original Draft, Project administration. **Shigeto Seno** and **Hideo Matsuda**: Software, Investigation, Writing - Review & Editing. **Naomi Murakami**: Investigation. **Yutaka Osada**: Conceptualization, Writing - Review & Editing. **Michio Kondoh**: Conceptualization, Writing - Review & Editing, Supervision. All authors commented on previous versions of this manuscript and approved the final manuscript.

References

Allen, M.M., 1968. Simple conditions for growth of unicellular blue-green algae on plates1, 2. J. Phycol. 4, 1-4.

Benton, T.G., Solan, M., Travis, J.M.J., Sait, S.M., 2007. Microcosm experiments can inform global ecological problems. Trends Ecol. Evol. 22, 516-521.

Bertani, G., 1951. Studies on Lysogenesis .1. The Mode of Phage Liberation by Lysogenic Escherichia-Coli. J. Bacteriol. 62, 293-300.

Beyers, R.J., Odum, H.T., 1993. Ecological microcosms, Springer-Verlag, New York.

Brenner, K., You, L., Arnold, F.H., 2008. Engineering microbial consortia: a new frontier in synthetic biology. Trends Biotechnol. 26, 483-489.

Breslow, D.K., Cameron, D.M., Collins, S.R., Schuldiner, M., Stewart-Ornstein, J., Newman, H.W., Braun, S., Madhani, H.D., Krogan, N.J., Weissman, J.S., 2008. A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. Nat. Methods. 5, 711-718.

Chuang, J.S., Frentz, Z., Leibler, S., 2019. Homeorhesis and ecological succession quantified in synthetic microbial ecosystems. Proceedings of the National Academy of Sciences of the United States of America. 116, 14852-14861.

Drake, J.A., 1991. Community-Assembly Mechanics and the Structure of an Experimental Species Ensemble. Am. Nat. 137, 1-26.

Fukami, T., 2015. Historical Contingency in Community Assembly: Integrating Niches, Species Pools, and Priority Effects. Annu Rev Ecol Evol S. 46, 1-23.

Fuma, S., Takeda, H., Takaku, Y., Hisamatsu, S., Kawabata, Z., 2005. Effects of dysprosium on the species-defined microbial microcosm. Bull. Environ. Contam. Toxicol. 74, 263-272.

Grosskopf, T., Soyer, O.S., 2014. Synthetic microbial communities. Curr. Opin. Microbiol. 18, 72-77.

Hekstra, D.R., Leibler, S., 2012. Contingency and Statistical Laws in Replicate Microbial Closed Ecosystems. Cell. 149, 1164-1173.

Hosoda, K., Suzuki, S., Yamauchi, Y., Shiroguchi, Y., Kashiwagi, A., Ono, N., Mori, K., Yomo, T., 2011. Cooperative adaptation to establishment of a synthetic bacterial mutualism. PLoS One. 6, e17105.

Inamori, Y., 2020. Microcosm Manual for Environmental Impact Risk Assessment, Springer Singapore.

Kawabata, Z., Matsui, K., Okazaki, K., Nasu, M., Nakano, N., Sugai, T., 1995. Synthesis of a Species-Defined Microcosm with Protozoa. Journal of Protozoology Research. 5, 23-26.

May, R.M., 1972. Will a Large Complex System Be Stable. Nature. 238, 413-&.

Mickalide, H., Kuehn, S., 2019. Higher-Order Interaction between Species Inhibits Bacterial Invasion of a Phototroph-Predator Microbial Community. Cell Systems. 9, 521-+.

Momeni, B., Chen, C.C., Hillesland, K.L., Waite, A., Shou, W.Y., 2011. Using artificial systems to explore the ecology and evolution of symbioses. Cell. Mol. Life Sci. 68, 1353-1368.

Mougi, A., Kondoh, M., 2012. Diversity of Interaction Types and Ecological Community Stability. Science. 337, 349-351.

Naeem, S., Li, S.B., 1998. Consumer species richness and autotrophic biomass. Ecology. 79, 2603-2615.

Nakajima, T., Sano, A., Matsuoka, H., 2009. Auto-/heterotrophic endosymbiosis evolves in a mature stage of ecosystem development in a microcosm composed of an alga, a bacterium and a ciliate. BioSyst. 96, 127-135.

Nixon, S.W., 1969. A Synthetic Microcosm. Limnol. Oceanogr. 14, 142-&.

Ponomarova, O., Patil, K.R., 2015. Metabolic interactions in microbial communities: untangling the Gordian knot. Curr. Opin. Microbiol. 27, 37-44.

Redmon, J., Farhadi, A., 2018. YOLOv3: An Incremental Improvement, pp. arXiv:1804.02767.

Shong, J., Jimenez Diaz, M.R., Collins, C.H., 2012. Towards synthetic microbial consortia for bioprocessing. Curr. Opin. Biotechnol. 23, 798-802.

Tan, J., Zuniga, C., Zengler, K., 2015. Unraveling interactions in microbial communities - from co-cultures to microbiomes. Journal of Microbiology. 53, 295-305.

Tanaka, N., Inamori, Y., Murakami, K., Akamatsu, T., Kurihara, Y., 1994. Effect of Species Composition on Stability and Reproductivity of a Small-Scale Microcosm System. Water Science and Technology. 30, 125-131.

Taub, F.B., 1969. A Biological Model of a Freshwater Community - a Gnotobiotic Ecosystem. Limnol. Oceanogr. 14, 136-&.

Supplementary material

A model ecosystem of twelve cryopreservable microbial species allowing for a non-invasive approach

Kazufumi Hosoda, Shigeto Seno, Naomi Murakami, Hideo Matsuda, Yutaka Osada, Michio Kondoh

This supplementary material includes

- Text S1 and S2
- Table S1
- Figures S1, S2, and S3

Text S1. Detailed information about the machine learning model.

Dataset preparation for machine learning based model

To build a model for detecting individual microorganisms in microcosm images, dataset for supervised learning was annotated by combinations of classical image processing methods. Wells with only a single species were prepared and microbial detection was performed using combinations of preprocessing, filters and segmentation algorithms. Since each species has a different shape (rod/oval/filamentous) and property (signals on the transmitted light/fluorescence), we customized the methods depending on the type of species. Although not all individuals were detected, thousands of bounding boxes (locations of microorganism individuals) were obtained. The input channels and the combinations of algorithms are shown in Table S1 and examples of the segmentation results are shown in Figure S1.

Training a model for microorganism detection

In this study, we used publicly available object detection network framework YOLOv3 and performed retraining using our dataset to detect and classify microorganisms. The code used for the YOLOv3 was cloned from the keras-yolo3 (https://github.com/qqwweee/keras-yolo3), a Keras implementation of YOLOv3 (Tensorflow backend). The weights of network were downloaded from the YOLO website (https://pjreddie.com/darknet/yolo/) and converted to the Keras model. Images of microorganisms were cropped to 512 × 512 from the original 2048 × 2048 pixels and information of bounding boxes was set appropriately. We performed retraining for 100 epochs using Adam optimizer and standard data augmentation (batch size = 32; learning rate = 10^{-3} for the first 50 epochs, then 10^{-4}).

Text S2. Detailed information about quantifying the population of 12 species.

We measured the microcosms using the fluorescence plate reader approximately once every 3 or 4 days, and also before and after subculturing on the same day. In addition, we took micrographs approximately once every 1 or 2 weeks. The micrographs were required to obtain the population of *T. thermophila* from the temporal variation of the image and also to obtain the population of other species using the machine learning model. We combined data from the plate reader and micrographs to obtain population dynamics at the resolution of the plate reader as follows.

We reconstructed the microscopy data by applying the most recent previous microscopy data each time the measurement under the microscope is missing (Figure S3A). The value after subculturing was set to 1/10 of the value before subculturing. These reconstructed values were directly used as the population for non-fluorescent species (#1–6; Figure S3B), including *S. cerevisiae*, which was green fluorescent protein (GFP)-labeled but not detected by the plate reader because the fluorescence was too weak. The population of *E. coli* can be measured directly using either the plate reader or the machine learning model, and the measurement using the machine learning model has a greater sensitivity but is not applicable for dense microcosms (giving a value lower than the truth; Figure 1B). Therefore, we applied the larger value of the population determined by the plate reader or the machine learning model to use the value measured by the machine learning for the part when the plate reader could not detect. For Cyanobacteria and Chlorophyta, the total amount of each population was measured using the plate reader, and the composition among each of them was determined by the population ratio using the machine learning model.

Table S1. Segmentation methods for each species

Species	Channel	Method
Escherichia coli	Fluorescence	thresholding (Otsu), morphology (opening, closing)
Staphylococcus epidermidis	Fluorescence	edge detection (DoG filter), thresholding (Otsu), morphology (dilation, fill holes)
Schizosaccharomyces pombe	Fluorescence	edge detection (DoG filter), thresholding (Otsu), morphology (dilation, fill holes)
Saccharomyces cerevisiae	Transmitted light	circle detection (canny filter, hough gradient method), watershed segmentation
Candida parapsilosis	Transmitted light	circle detection (canny filter, hough gradient method), watershed segmentation
Dictyostelium discoideum	Transmitted light	edge detection (DoG filter), thresholding (Otsu), morphology (dilation, fill holes)
Tetrahymena thermophila	Transmitted light	thresholding (local thresholding), morphology (opening, closing, fill holes), watershed segmentation
Anabaenopsis circularis	Fluorescence	Frangi filter, thresholding (Otsu), morphology (closing)
Synechocystis sp. PCC 6803	Fluorescence	circle detection (canny filter, hough gradient method), watershed segmentation
Pseudokirchneriella subcapitata	Fluorescence	thresholding (Otsu), morphology (opening, closing)
Chlorella vulgaris	Fluorescence	circle detection (canny filter, hough gradient method), watershed segmentation
Chlamydomonas reinhardtii	Fluorescence	circle detection (canny filter, hough gradient method), watershed segmentation

Figures

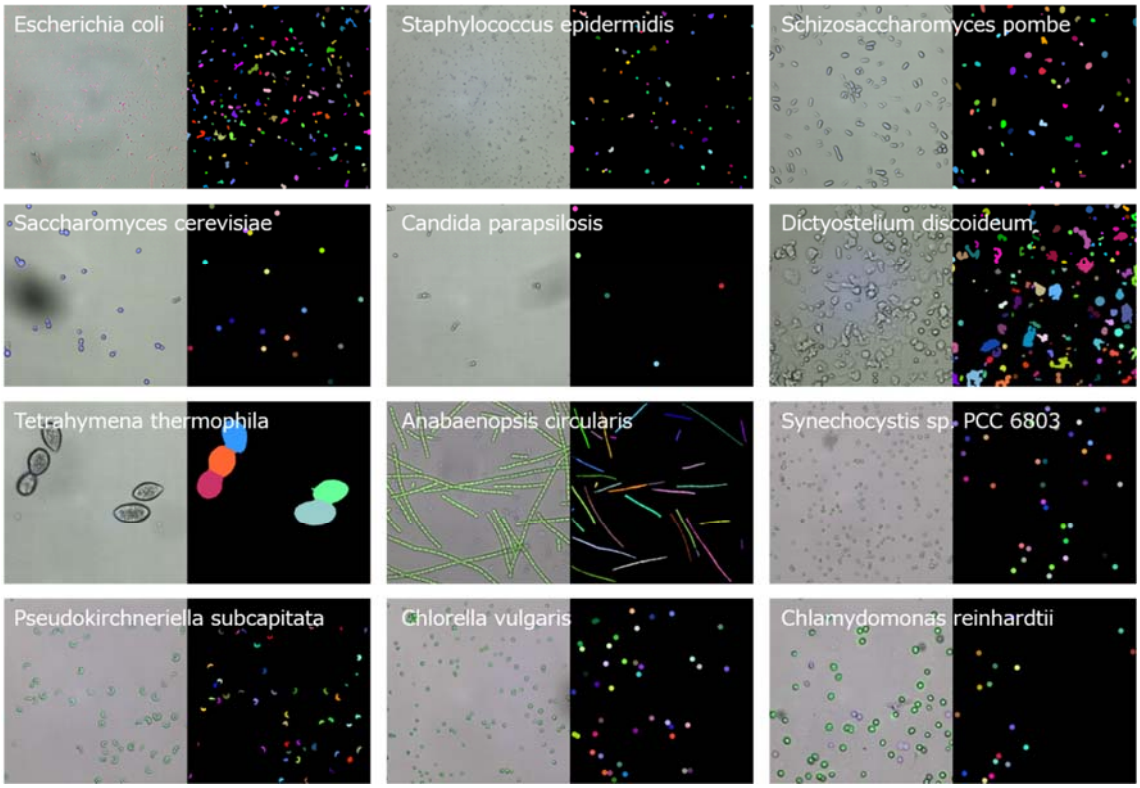

Figure S1. Segmentation results by classical methods. For each species, the left side is a representative micrograph of bright-field and fluorescent-field superimposed (one side is approximately 222 μm). The right side shows segmentation into each individual. The mask of each individual is painted in the same color.

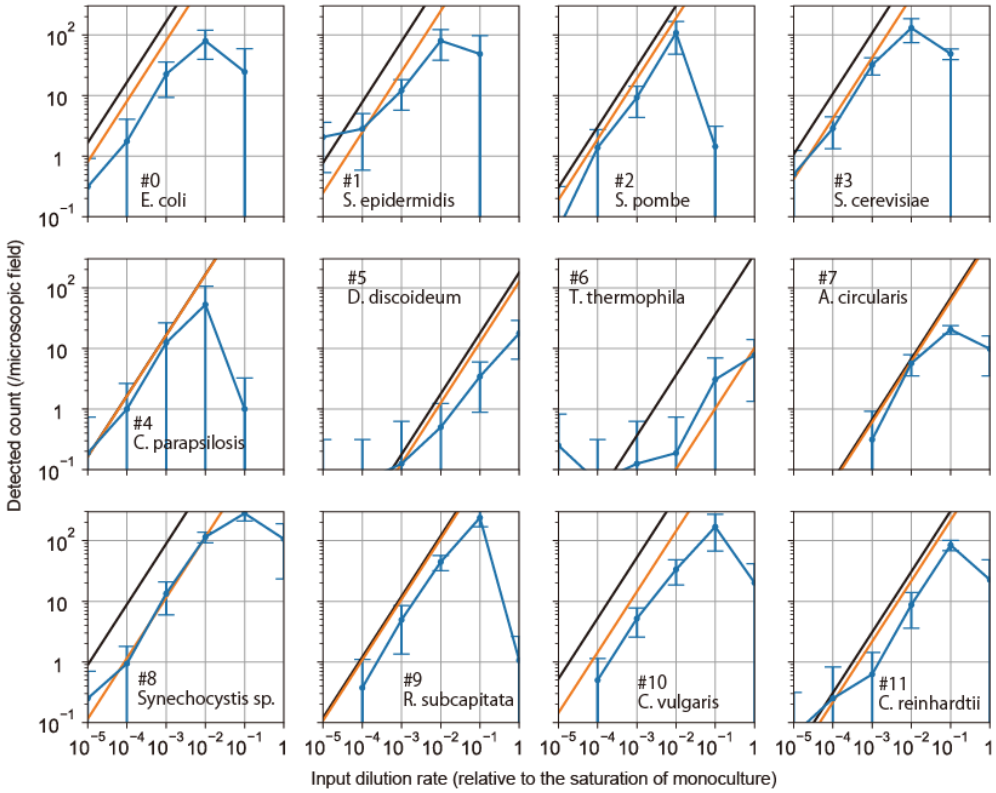

Figure S2. Calibration curves of measurements using the machine learning model for each species. We prepared full-growth monocultures of each species. These monocultures were diluted with BG11H medium. The horizontal axes show the dilution rate. Blue circles with lines show the mean detected count in each micrograph of 16 replicated wells, and the error bars show the standard deviation. We also counted the number of individuals in each micrograph with eyes at the appropriate dilution rate for counting (10–100 individuals in each micrograph). The orange lines show the values counted with eyes, extended to all dilution rate. The black lines show the expected values calculated from the input concentration if all input individuals are assumed to sink to the well's bottom and are evenly distributed on the bottom. Thus, the difference between the orange and black lines show the fraction of individuals that sunk to the bottom and entered the microscopic view, which is independent of the detection using the machine learning.

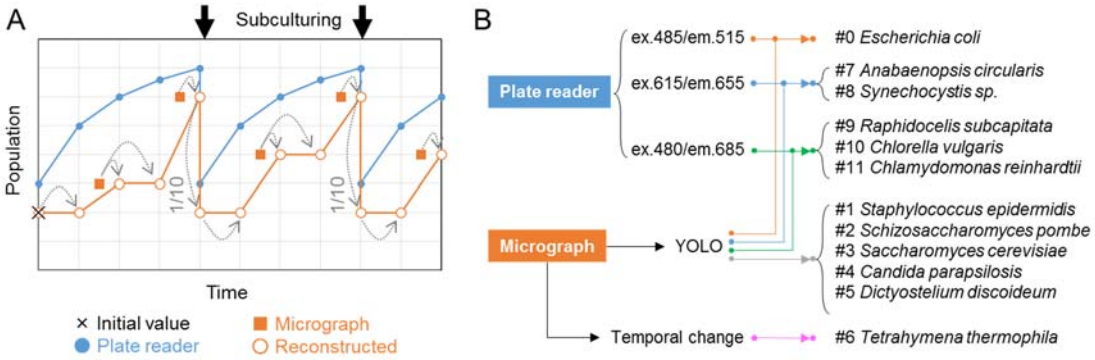

Figure S3. Detailed information about quantifying the population of the twelve species. See Text S2 for explanation.