Detecting and validating influential organisms for rice growth: An ecological network approach

Masayuki Ushio[1,2,*,†], Hiroki Saito[3], Motoaki Tojo[4], Atsushi J. Nagano[5,6]

[1]Hakubi Center, Kyoto University, Kyoto 606-8501, Japan,[2]Center for Ecological Research, Kyoto University, Otsu 520-2113, Japan,[3]Tropical Agriculture Research Front, Japan International Research Center for Agricultural Sciences, Ishigaki, Okinawa 907-0002, Japan,[4]Graduate School of Agriculture, Osaka Metropolitan University, Sakai, Osaka 599-8531, Japan,^{5}Faculty of Agriculture, Ryukoku University, Otsu 520-2194, Japan,[6]Institute for Advanced Biosciences, Keio University, Tsuruoka 997-0017, Japan

Abstract

1. How to achieve sustainable food production while reducing environmental impacts is a major concern in agricultural science, and advanced breeding techniques are promising for achieving such goals. However, rice is usually grown under field conditions and inevitably influenced by surrounding ecological community members, and whether and how ecological communities influence the rice performance under field conditions has been underexplored.
2. In the present study, we demonstrate an ecological-network-based approach to detect potentially influential, previously ignored organisms for rice (*Oryza sativa*). First, we established small experimental rice plots, and measured rice growth and monitored ecological community dynamics intensively and extensively using quantitative environmental DNA analysis in 2017 in Japan. We detected more than 1000 species in the rice plots, and nonlinear time series analysis detected 52 potentially influential organisms with lower-level taxonomic information.
3. The results of the time series analysis were validated in 2019. In 2019, we focused on two species, *Globisporangium nunn* and *Chironomus kiiensis*, whose abundance was manipulated in artificial rice plots. The responses of rice, namely, the growth rate and gene expression patterns, were measured before and after the manipulation. We confirmed that, especially in the *G. nunn*-added treatment, rice growth rate and gene expression pattern were changed.
4. In the present study, we demonstrated that intensive monitoring of an agricultural system and the application of nonlinear time series analysis were helpful to identify influential organisms under field conditions. Although the effects of the manipulations were relatively small, the research framework presented here has future potential to harness the ecological complexity and utilize it in agriculture. Our proof-of-concept study would be an important basis for the further development of field-basis system management.

Keywords: Agricultural system; *Chironomus kiiensis*; Environmental DNA; *Globisporangium nunn*; Gene expression; Manipulation experiment; Rice; Unified information-theoretic causality

*Corresponding author: ong8181@gmail.com (MU)

†Present address (MU): Department of Ocean Science, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR, China

Introduction

Global food production supports energy requirements for human beings (Godfray *et al.* 2010), while the food production is a major driver of greenhouse gas emission and other environmental loads (Aleksandrowicz *et al.* 2016). How to achieve sustainable food production while reducing environmental impacts is thus a major concern in agricultural science (Lenaerts *et al.* 2019). Rice (*Oryza sativa*), one of the world's major staple crops, is an essential component of the diets and livelihood of over 3.5 billion people (Wing *et al.* 2018). A large number of studies have investigated how the performance of rice can be improved, and genomics-based breeding, combined with other advanced technologies such as high-throughput field-based phenotyping, is one of the most promising approaches to improve the performance (Wing *et al.* 2018).

While the advanced breeding techniques are promising, rice is usually grown under field conditions and inevitably influenced by the surrounding biotic and abiotic environment. Previous studies have investigated how meteorological and endogenous (e.g., plant age and genotype) variables influence the gene expression patterns (transcriptome dynamics) of rice under fluctuating field conditions (Nagano *et al.* 2012; Kashima *et al.* 2021), and have suggested that transcriptome dynamics can be predominantly explained by endogenous factors and abiotic variables such as ambient air temperature and solar radiation. Nonetheless, biotic variables such as insect herbivory and microbial mutualists/pathogens also play an important role in determining the transcriptome dynamics and productivity of crops (Cohen & Leach 2019; Savary *et al.* 2019). However, the dynamics of biotic variables (i.e., ecological community members) under field conditions are often difficult to predict because they often show more complex, nonlinear dynamics than abiotic variables (Hsieh *et al.* 2005). Thus, understanding whether and how biotic variables influence the rice performance (i.e., interspecific interactions) has been underexplored despite its importance for sustainable agriculture (Toju *et al.* 2018).

In ecological studies, biotic interactions have been studied theoretically and empirically for decades (May 1972; Allesina & Tang 2012; Ushio *et al.* 2018a). Ecologists have traditionally studied interspecific interactions based on observations and manipulations. For example, Paine (1966) measured the strength of interspecific interactions by removing predator in a rocky tidal system, and showed that the interspecific interactions have critical influences on the system dynamics. Thus far, interspecific interactions and their influences on the system dynamics have been studied as in Wootton & Emmerson (2005). However, despite the substantial contributions to ecology, these observation- and manipulation-based approaches have critical limitations: the identification of multitaxa species and the quantification of their abundance under field conditions are challenging, and the quantification of their interactions is even more difficult (but see Ushio 2022). Overcoming these difficulties and understanding how ecological community members influence the rice performance under field conditions will provide insights into how we can improve the rice performance and how rice responds to the ongoing and future anthropogenic impacts.

One of the promising approaches for overcoming the current limitations is to monitor the system frequently and detect interspecific interactions using time series data. Recent advances in empirical and statistical methods provide a practical way to achieve this goal. First, use of environmental DNA (eDNA) enables researchers to efficiently detect ecological community members under field conditions (Taberlet *et al.* 2018). Previous studies have shown that eDNA metabarcoding, an approach to comprehensively amplify and sequence DNAs belonging to a target taxa in an environmental sample, is a cost- and time-effective means to detect a large number of species (e.g., Miya *et al.* 2015), and the eDNA-based community data is

especially informative when it is obtained quantitatively (e.g., sequencing with internal spike-in DNAs: Ushio *et al.* 2018b; Ushio 2022). Second, nonlinear time series analytical tools enable researchers to reconstruct complex interaction networks (Schreiber 2000; Sugihara *et al.* 2012; Deyle *et al.* 2016; Chang *et al.* 2021), and previous studies adopting the approach have detected causality between many variables (e.g., Runge *et al.* 2019).

In the present study, we reconstructed the interaction network surrounding rice and detected potentially influential organisms for rice under field conditions. First, in 2017, we established small experimental rice plots, and monitored rice performance (i.e., growth rates) and ecological community dynamics intensively and extensively. Rice performance was quantified by measuring the growth rates (cm/day), and ecological community members were monitored by a quantitative eDNA metabarcoding approach (Ushio *et al.* 2018b; Ushio 2022). The monitoring was performed daily from 23 May 2017 to 22 September 2017 (122 consecutive days). Second, we analyzed the generated extensive time series containing 1197 species and rice growth rates and produced a list of 52 potentially influential species. Third, in 2019, we empirically tested the effects of two species that had been identified as potentially influential in 2017 using manipulative experiments. During the growing season in 2019, an Oomycetes species, *Globisporangium nunn* (syn. *Pythium nunn*), was added, and a midge species, *Chironomus kiiensis*, was removed from small rice plots. The rice responses (the growth rate and gene expression patterns) were measured before and after the manipulation. We confirmed that the two species, especially *G. nunn*, indeed had statistically clear effects on the rice performance.

Mateirals and Methods

In this section, we provide an overview of the methods. Full details of the methods are described in the Supporting Information (for the manipulative experiments and RNA expression analysis) and Ushio (2022) (https://ndownloader.figstatic.com/files/34067324 and https://ndownloader.figstatic.com/files/34067327) (for eDNA monitoring).

Field experimental setting in 2017

Five artificial rice plots were established using small plastic containers (90 × 90 × 34.5 cm; 216 L total volume; Risu Kogyo, Kagamigahara, Japan) in an experimental field at the Center for Ecological Research, Kyoto University, in Otsu, Japan (34° 58′ 18′′ N, 135° 57′ 33′′ E) (Figs. 1a and S1a,b). Sixteen Wagner pots were filled with commercial soil, and three rice seedlings (var. Hinohikari) were planted in each pot on 23 May 2017 and then harvested on 22 September 2017 (122 days). The containers (hereafter, "plots") were filled with well water. Daily rice growth was monitored by measuring rice leaf height of target individuals (red points in the right panel of Fig. 1a) every day using a ruler (the largest leaf heights were measured) (a time-lapse movie that shows rice growth is available here: https://doi.org/10.6084/m9.figshare.19029650.v1). Leaf SPAD was also measured, but not included in the present study. Climate variables (temperature, light intensity, and humidity) were also monitored at each rice plot.

Field monitoring of ecological communities in 2017

To monitor the ecological community, water samples were collected daily from the five rice plots. Approximately 200 ml of water in each rice plot was collected from the plot and taken to the laboratory within 30 minutes. The water was filtered using two types of Sterivex filter cartridges (ϕ 0.22-μm and ϕ 0.45-μm; Merck Millipore, Darmstadt, Germany). In total, 1220

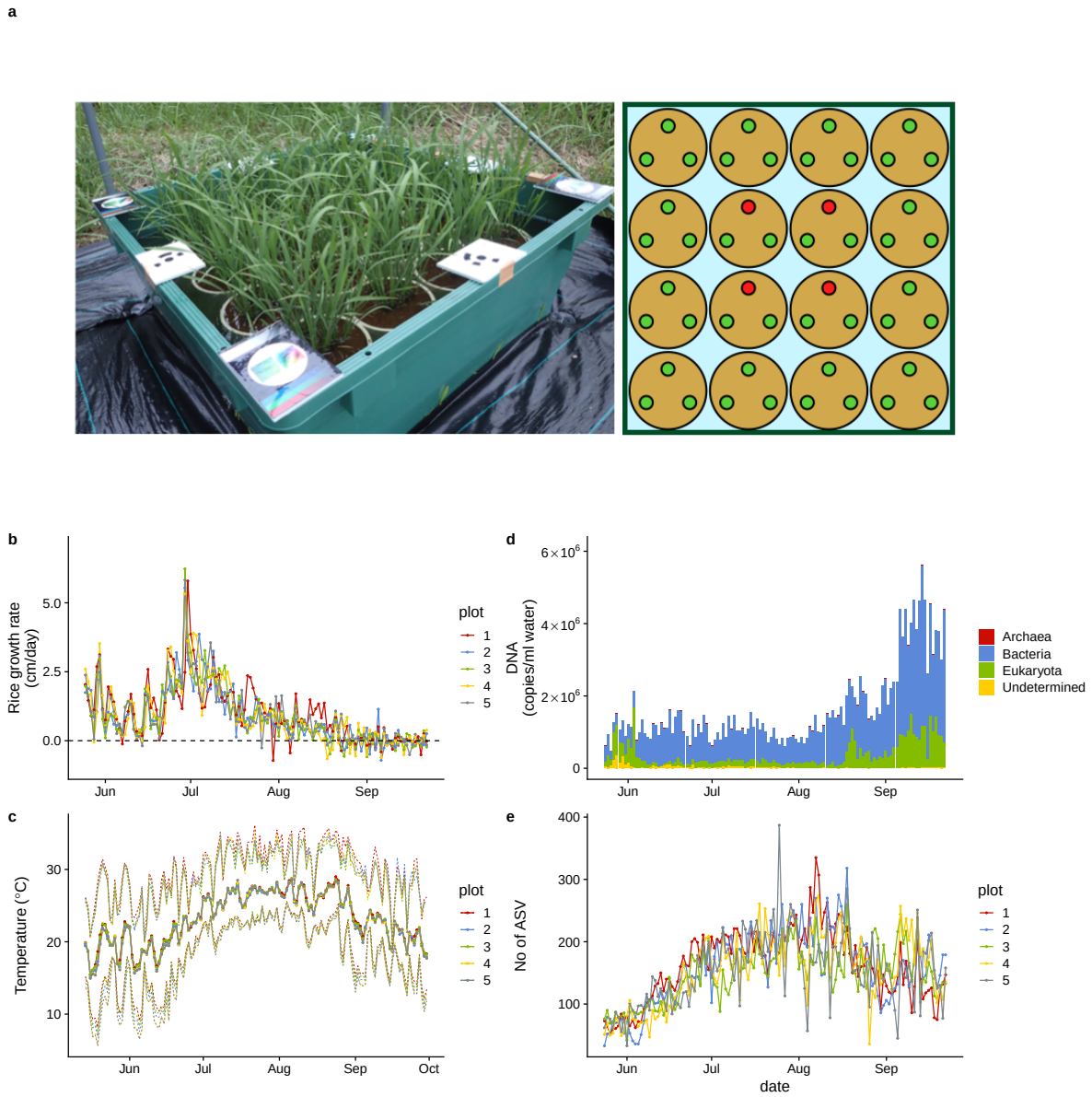

Figure 1. Rice plot, growth rate, air temperature, and ecological community dynamics. (**a**) 90 cm × 90 cm rice plot and Wagner pot alignments. Four rice individuals were grown in each pot (indicated by points in the right panel of **a**). Heights and SPAD of the target individuals (indicated by four red points in the right panel of **a**) were measured everyday during the monitoring period. (**b**) Rice growth rate (cm/day). (**c**) Daily mean air temperature measured at each rice plot. Upper and lower dotted lines indicate daily maximum and minimum air temperature. (**d**) Ecological community compositions and average DNA copy numbers per ml water (reported in Ushio 2022). (**e**) The number of amplicon sequence variants (ASVs) from each water sample (reported in Ushio 2022). For **b**, **c**, and **e**, different colors indicate data from different rice plots.

water samples (122 days × 2 filter types × 5 plots) were collected in addition to negative control samples during the census term.

eDNA was extracted from the filters and purified, and eDNA libraries were constructed using a standard two-step PCR protocol with four universal primers (16S rRNA, 18S rRNA, COI and ITS regions), and sequenced on the MiSeq system (Illumina, San Diego, CA, USA). Then, the sequence reads were analyzed using the Amplicon Sequence Variant (ASV) method implemented in "dada2" package (Callahan *et al.* 2016) of R (Team 2022). Taxonomic identification was performed using Claident (Tanabe & Toju 2013). After the assignment of the taxonomic information, we estimated the number of eDNA copies based on the spike-in standard DNAs (Ushio *et al.* 2018b).

Detections of potential causal relationships between rice growth and ecological community members

We detected potentially influential organisms for rice growth analyzing the quantitative, 1197-species eDNA time-series (Ushio 2022) and the daily rice growth rate obtained in 2017 using nonlinear time series analysis. We quantified information flow from eDNA time series to rice growth rate by the "unified information-theoretic causality (UIC)" method implemented in the "rUIC" package (Osada & Ushio 2021) of R. UIC tests the statistical clarity of information flow between variables in terms of transfer entropy (TE) (Schreiber 2000). Climate data was included as a conditional variable to quantify only species effects on rice growth (i.e., conditional UIC). By using UIC, we quantified TE from eDNA time series to rice growth (see Supporting Information for the detailed methods). We standardized the eDNA time series (copies/ml water) and rice growth rates (cm/day) to have zero means and a unit of variance before the analysis.

Field experimental setting in 2019

Based on the results in 2017, we performed a field manipulative experiment in 2019. First, nine artificial rice plots were established using plastic containers (42.2 × 32.0 × 30.0 cm; 40.5 L total volume; Sanko, Tokyo, Japan) in the same experimental field at Kyoto University. Three Wagner pots were filled with commercial soil, and three rice seedlings (var. Hinohikari) were planted in each pot on 20 May 2019 and then harvested on 20 September 2019 (124 days) (9 plots × 3 Wanger pots × 3 rice seedlings = 81 rice seedlings in total). In 2019, the rice growth was monitored once a week except during the period of the manipulation experiments, when the monitoring was daily (see Table S1). The containers (hereafter, "plots 1–9") were filled with well water. After we harvested rice, we counted the number of fertile and sterile grains and quantified rice yields.

Field manipulation experiments in 2019

Field manipulation experiments were performed using two potentially influential species, *Globisporangium nunn* (Oomycete; previously known as *Pythium nunn*) (Uzuhashi *et al.* 2010) and *Chironomus kiiensis* (midge species) (Fig. S2). The reasons why we focused on these two species are multifold. First, they can be relatively easily manipulated; *G. nunn* had already been isolated and cultivated and *C. kiiensis* is a relatively large insect species (several mm to cm). Second, the two species had not been reported to be pathogenic; using pathogen species under field conditions is not allowed in the experimental field because they might have adverse effects on nearby agricultural systems. As for *G. nunn*, we referred to the ASV found by the eDNA metabarcoding in 2017 as "putative *Globisporangium*" because there are differences in the ASV

sequence and the *G. nunn* type strain sequence. Detailed explanations about the difference in the sequences and reasons why we used this species are explained in "Notes on *Globiosporangium nunn*" in the Supporting Information.

We performed the rice plot manipulations three times at around noon on 24, 26, and 28 June 2019. There were three treatments, namely, "*Globisporangium nunn*-added" (GN), "*C. kiiensis-removed*" (CK), and control (CT) treatments, and there were three replicate plots for each treatment (3 treatments × 3 replicates = 9 plots). For the GN treatment, isolated and incubated *G. nunn* was mixed in vermiculite and approximately 200 ml of vermiculite was added to each Wagner pot. Vermiculite without *G. nunn* was added to Wagner pots in the two other treatments. For the CK treatment, midge larvae were removed instead of adding them because isolation and cultivation of this insect species are technically difficult and time- and effort-consuming. Midge larvae were removed using a ϕ 1.0 mm net. On average, 283 midge larvae were removed from each plot and the DNA of about 100 removed larvae was sequenced using Sanger sequencing, and all were confirmed as *C. kiiensis*. For the other treatments, the net was also inserted into plots, but midge larvae were not removed.

Rice growth and ecological community monitoring, and rice leaf sampling during the manipulation experiments

Before and after the manipulations, the rice plots were intensively monitored (Table S1). Rice performance (growth rate and SPAD) was monitored every day from 11 June to 12 July (32 days), as in 2017. Water samples for eDNA-based ecological community monitoring were collected every day from 18 June to 12 July (25 days), as in 2017. Rice leaf samples for RNA-seq were collected before and after the field manipulations. In total, 108 leaf samples were collected for RNA-seq analysis (Table S2).

Effects of the field manipulation experiment on rice growth and ecological communities

The differences in rice performance among the three treatments and total eDNA concentrations of ecological community members (copies/ml water) before and after the manipulation experiments were tested using linear a mixed model (LMM) and general linear mixed model (GLMM), respectively, using "lme4" package of R (Bates *et al.* 2015). For rice height, the effect of the treatment was tested with the random effects of rice individual and/or rice plot. The effects were separately analyzed for before/after the manipulation for visualization purposes (see the main figures), and the results of the analysis that merged the data before/after the treatments were also reported in the Supplementary Tables. Rice cumulative growth (i.e., growth 10 days after the manipulation; from 30 June to 10 July) was analyzed similarly, but the random effect of rice individuals was not included because there was only one cumulative value for each rice individual. For ecological community, the effects of the manipulation were separately tested for each treatment with the random effect of rice plot assuming a gamma error distribution for visualization purposes (see the main figures), and the results of the analysis that merged all treatments were reported in the Supplementary Tables. Differences in the ecological community compositions were visualized using t-distributed stochastic neighbor embedding (t-SNE) (Van der Maaten & Hinton 2008). In the present study, we used the term "statistical clarity" instead of "statistical significance" to avoid misinterpretations, according to the recommendations by Dushoff et al. (2019), when $P < 0.05$.

Quantifications of RNA expressions of rice leaves and differentially expressed genes (DEGs)

The leaf samples were ground under cryogenic conditions, and total RNA was extracted using the Maxwell 16 LEV Plant RNA Kit (Promega, Madison, WI, USA). Extracted RNA was quantified and 500 ng of RNA was used as the input of each sample for library preparation. Library preparation for RNA-sequencing was conducted using Lasy-Seq (Kamitani *et al.* 2019:) version 1.1 (https://sites.google.com/view/lasy-seq/). The library was sequenced using HiSeq X (Illumina, San Diego, CA, USA) with paired-end sequencing lengths of 150 bp. Obtained reads were trimmed and mapped onto the reference. The outputs of the sequence processing (i.e., "expected counts") were imported to R, and differentially expressed genes (DEGs) were detected using the "DESeq2" package (Love et al., 2014). We measured rice gene expression four times and the results were clearest at sampling event 2 (referred to as "S2" in Table S2) (Table S1; 5 days after the first field manipulation). Therefore, we report only the results of sampling event 2 in the present study. In the Supplementary Tables, we report a list of nuclear DEGs.

Results

Field monitoring of rice growth and ecological communities in 2017

Daily rice growth rate (cm/day in height) showed consistent patterns among the five rice plots (Fig. 1b). Daily mean air temperature also showed consistent patterns among the five plots (Fig. 1c). The daily growth rate reached the maximum during late June to early July, and the height of rice individuals did not increase after the middle of August (first headings appeared on 12 or 13 August in the five plots). During the monitoring period, we occasionally observed decreases in the rice heights due to mechanical damage or insect herbivores.

Detailed patterns of the ecological community dynamics and their implications were reported in Ushio (2022). Briefly, the total eDNA copy number increased late in the sampling period (Fig. 1d). In contrast, ASV diversity (a surrogate of species diversity) was highest in August and then decreased in September (Fig. 1e). Prokaryotes largely accounted for this pattern (Fig. 1d). In the previous study, a large ecological interaction network was reconstructed (Fig. S1c), and possible mechanisms of the ecological dynamics were discussed in detail.

Detections of potential causal relationships between rice growth and ecological community dynamics

Potential causal relationships among the daily rice growth rates, eDNA time series, and climate data were detected in terms of information flow (Fig. 2 and Table S3). Fig. 2a is a demonstration of the performance of the UIC analysis, showing the influence of air temperature on the rice growth. The left panel in Fig. 2a indicates that there were strong effects of air temperature on the rice growth (x-axis indicates the time-lag of the effects), while the right panel indicates that there were virtually no effects of the rice growth on air temperature. In total, 718 ASVs were identified as potentially causal (i.e., statistically clear information flow; $P < 0.05$), among which 52 ASVs were identified with lower-level taxonomic information (Table S3). The potentially causal ASVs belong to various taxa, including fungi and insects, with varying degree of the effects, measured as the information flow (Fig. 2b). Among the 52 ASVs, we particularly focused on two ASVs, *G. nunn* and *C. kiiensis* (see the eDNA dynamics of putative *G. nunn* and *C. kiiensis* in Fig. 2c, d), in the field manipulation experiments in 2019, as they are identified with lower-level taxonomic information, already incubated or isolated, and/or could be manipulated because of the relatively large body size.

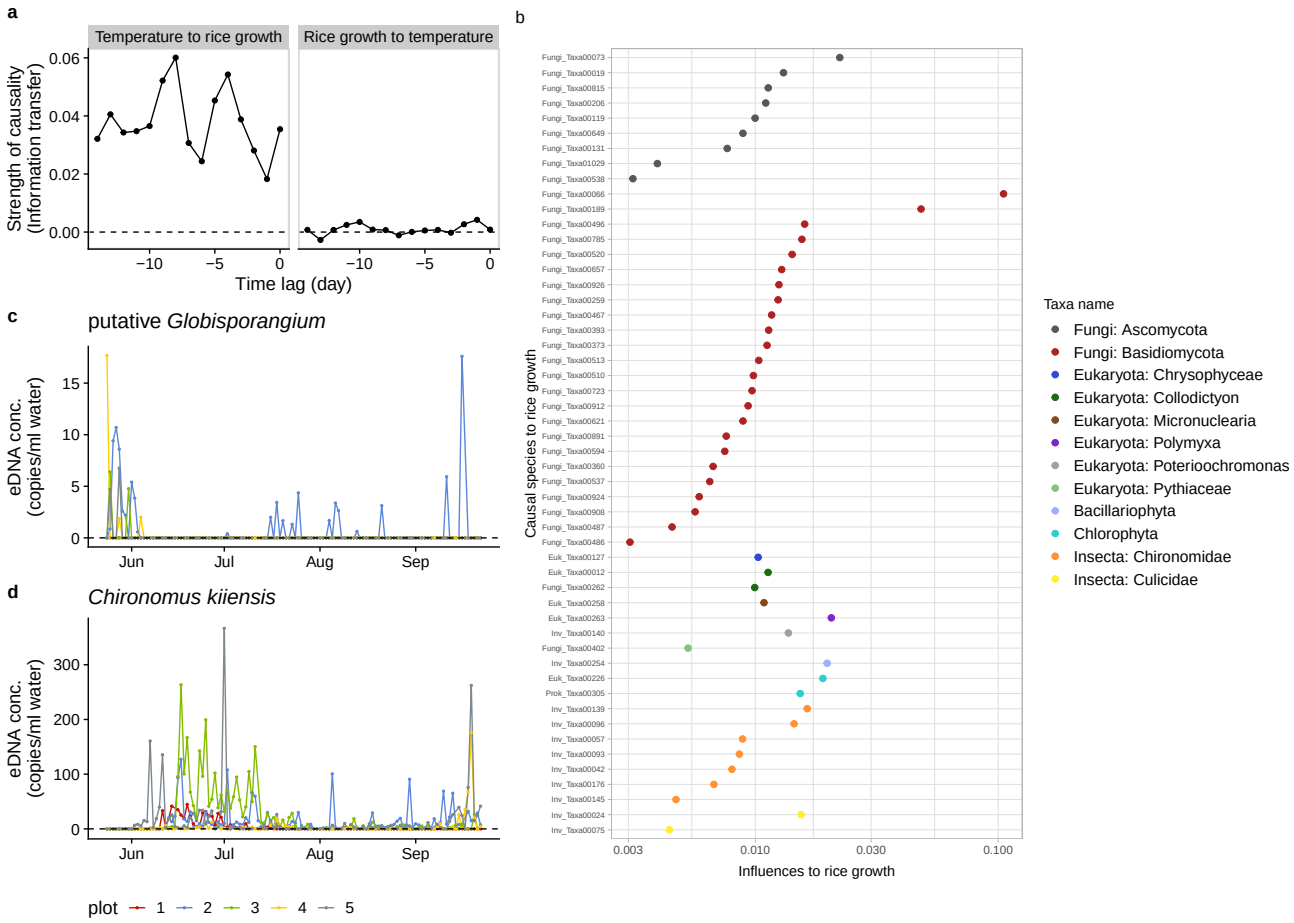

Figure 2. Information transfer between rice growth and ecological community members. (**a**) An example of the results of the unified information-theoretic causality analysis. Information transfer between air temperature and rice growth rates was quantified. Much higher information transfer was detected from air temperature to rice growth (left panel) compared with the opposite direction (right panel). (**b**) Strength of causal influence from ecological community members to rice growth. Colors indicate taxa assigned to ASVs. *y*-axis indicates ASV ID. Note that the prefix (e.g., "Fungi_") of the IDs corresponds a major target group of the primer and does not necessarily indicate the taxonomic group assigned to the ASV. (**c**) eDNA dynamics of putative *Globisporangium* (Fungi_Taxa00402 in Table S3). (**d**) eDNA dynamics of *Chironomus kiiensis* (total DNA copy numbers of five midge ASVs). For **c** and **d**, different colors indicate data from different rice plots.

The ecological community composition after the manipulation experiments in 2019

In 2019, we manipulated the abundance of *G. nunn* and *C. kiiensis* (the GN treatment [i.e., *G. nunn*-added treatment] and CK treatment [i.e., *C. kiiensis*-removed treatment], respectively) (Fig. 3a), and quantified the eDNA concentrations of the target species. As for the GN treatment, we detected 19 putative *Globisporangium* sequences belonging to the family Pythiaceae from the four marker genes (i.e., 16S, 18S, ITS and COI), most of which we could not identify with lower-level taxonomic information partially because of the conservative assignment algorithm of Claident. Therefore, we described them as "putative *Globisporangium* spp." because the eDNA concentrations of these Pythiaceae sequences were increased after we added *G. nunn* isolates. There was a statistically clear increase in the eDNA concentrations of putative *Globisporangium* spp. in the GN treatment after the manipulation ($P < 2.0 \times 10^{-16}$; Fig. 3b). In addition, there was a slight, but statistically clear increase in the eDNA concentrations of putative *Globisporangium* spp. in the CT treatment after the manipulation ($P = 0.032$; Fig. 3b). On the other hand, the eDNA concentrations of putative *Globisporangium* spp. in the CK treatment was not changed after the manipulation ($P > 0.05$). The eDNA concentrations of *C. kiiensis* were statistically clearly changed after the manipulation in the CT (slightly decreased; $P = 0.045$; Fig. 3c) and GN (increased; $P = 6.6 \times 10^{-6}$; Fig. 3c) treatments. In the CK treatment, the eDNA concentration of *C. kiiensis* was not changed ($P > 0.05$; Fig. 3c). These results were further confirmed by the t-SNE analysis of the overall community composition. Ecological communities of the GN treatment were generally separated from those of the CT treatment, while those of the CK treatment almost overlapped with those of the CT treatment (Fig. 3d). Alternative statistical modeling that simultaneously addressed the treatment effects and manipulation timing (i.e., before or after the manipulation) also showed qualitatively similar results (Table S4).

Rice performance after the field manipulation experiments in 2019

We compared the rice growth of the GN and CK treatments with that of the CT treatment before and after the manipulation experiments. In general, we did not find statistically clear differences among the treatment in the daily rice growth rate before and after manipulation ($P > 0.05$; Fig. 4a) (see Fig. S3 for the rice growth trajectory during the growing season). However, we found that the cumulative rice growth (during 10 days after the manipulation) increased in the GN and CK treatments after the manipulation, and the effect was statistically clear ($P = 0.016$ and $P = 0.021$ for the GN and CK treatments, respectively; Fig. 4b). Alternative statistical modeling that simultaneously addressed the treatment effects and manipulation timing (i.e., before or after the manipulation) also showed qualitatively similar results (Table S5). We did not find any statistically clear difference in the rice yields (e.g., the number of rice grains per head, the grain weights, or the proportion of fertile grains) among the treatments (Table S6).

Changes in rice gene expression pattern in 2019

We further explored changes in the rice performance before and after the manipulation experiments using RNA-seq. As a result of the RNA-seq analysis, 875,534,703 reads were generated (Table S2), and 8,105,902 reads per sample were used for the sequence analysis. DESeq2 analysis was performed after importing the results into R, and the analysis suggested that, when all three rice pot locations within each plot were merged (see Fig. 3a for the pot locations), there were differentially expressed genes in the GN treatment (Fig. 5a; see Table S7 for nuclear DEGs) while there were no differentially expressed genes in the CK treatment (Fig. 5b). However, when the data was separately analyzed for each Wagner pot location, we found that there were several differentially expressed genes even in the CK treatment (Fig. S4 and Table S8), suggesting that

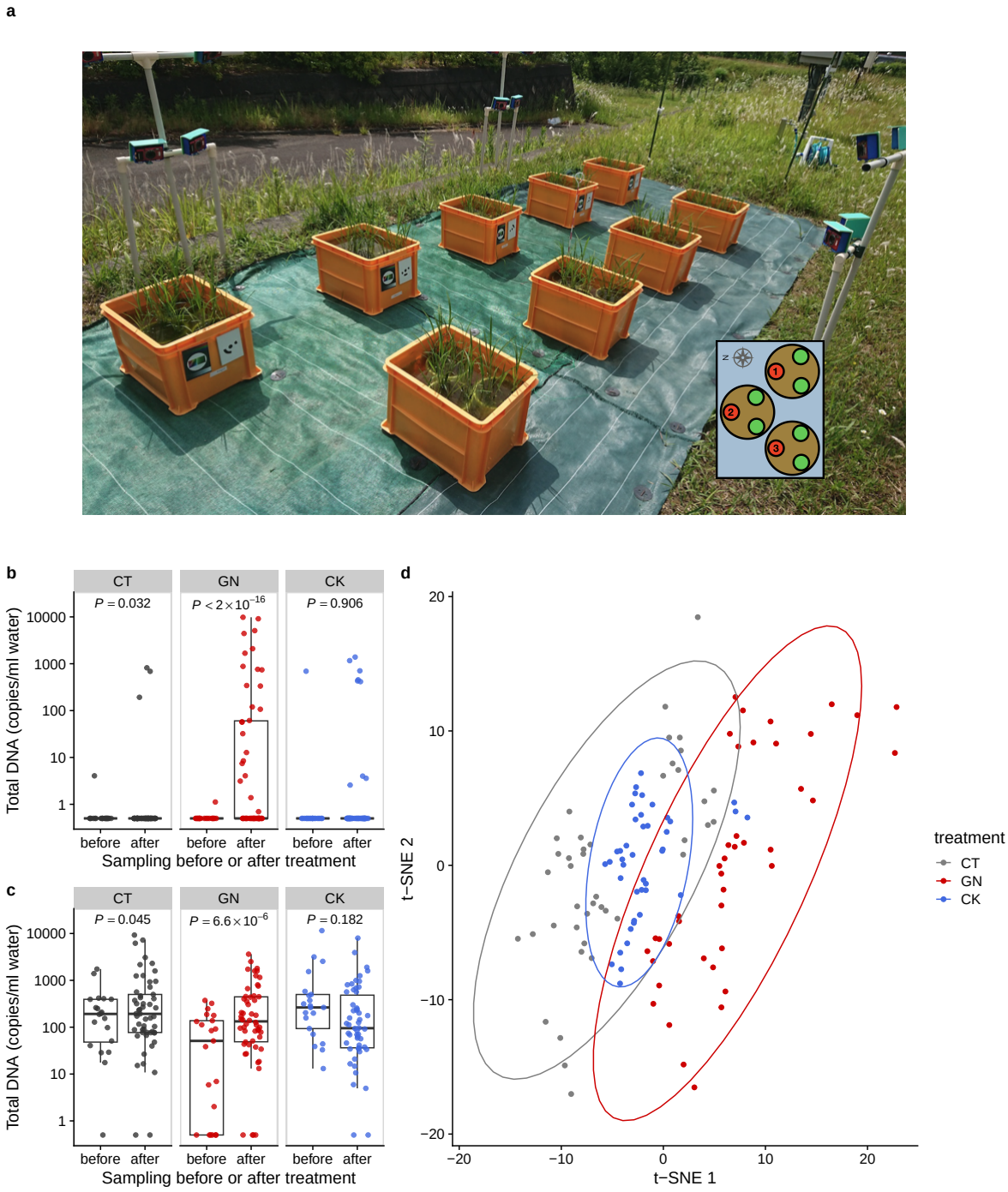

Figure 3. The manipulation experiment performed in 2019 and ecological community compositions before and after the manipulation. (**a**) Setting of the manipulation experiment in 2019. The inset shows three individuals (red and green points) in each Wagner pot. Heights and SPAD of the red individuals were measured. Total eDNA copy numbers of (**b**) putative *Globisporangium* spp. and (**c**) midge (*Chironomus kiiensis*) in the rice plots. (**d**) Overall community compositions after the manipulation. Gray, red, and blue indicate CT (control), GN (*Globisporangium nunn* added), and CK (*Chironomus kiiensis* removed) treatments, respectively. Each ellipse indicates the overall distribution of each treatment data.

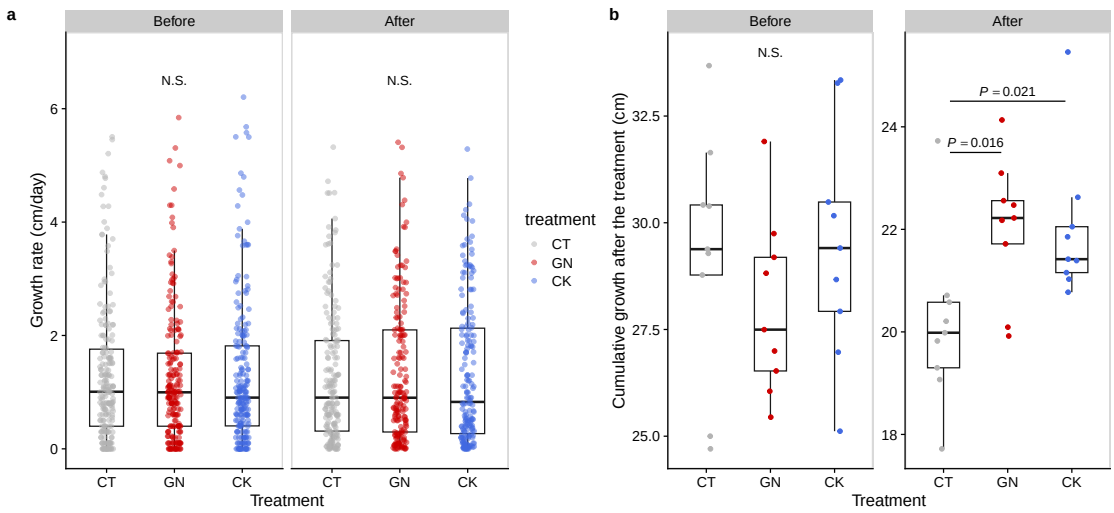

Figure 4. Ecological community compositions before and after the manipulation experiment in 2019. (**a**) Growth rates and (**b**) cumulative growth of the rice individuals in the three treatment (CT = control; GN = *Globisporangium nunn* added; CK = *Chironomus kiiensis* removed) before and after the manipulation.

there was a pot location-specific effect (probably due to the differences in the microclimate) and that the CK treatment also had small effects on the rice gene expression. In Fig. 6, we show examples of differentially expressed genes. In the GN treatment, the expressions of three genes, Os12g0504050, Os11g0184900, and Os01g0678500 ($P < 0.0001$, negative binomial GLMM; Fig. 6a–c), were increased, while those of three other genes (Os01g0642200, Os08g0162800, and Os03g0285700) were decreased ($P < 0.0001$, negative binomial GLMM; Fig. 6d–f). These DEGs are mostly related to the photosynthetic system and stress responses (Table S7).

Discussion

In the present study, we demonstrated a novel research framework to integrate the ecological network concept and agricultural practices to screen potentially influential organisms. We performed intensive field monitoring in 2017, analyzed the data using nonlinear time series analysis, and based on the results, we performed field manipulation experiments and evaluated the effect of the manipulation. Throughout the study, we detected statistically clear effects of the manipulation on the rice growth and gene expression patterns especially for the GN treatment, proving that our proposed framework might work as a novel approach to comprehensively detect potentially beneficial (or detrimental), previously unknown ecological community members for an agricultural system (in our case, rice). Below, we discuss the results, advantages, disadvantages, potential limitations, and future perspectives of our approach.

Monitoring of ecological community using eDNA and rice growth

Our framework started from monitoring ecological community members using quantitative eDNA metabarcoding (Ushio 2022). We successfully detected more than 1000 ASVs and described detailed temporal dynamics of the ASVs [Fig. 1d,e; Ushio (2022)]. We included spike-in standard DNAs whose concentrations were known a priori, and our previous study had shown that, in almost all cases, there were clear positive linear correlations between the sequence reads and the copy numbers of the spike-in standard DNAs (Fig. S3 in Ushio 2022). This suggested that our quantitative eDNA metabarcoding reasonably quantified the concentrations of eDNA, which can be a “rough” index of species abundance in the artificial rice plots. Some time

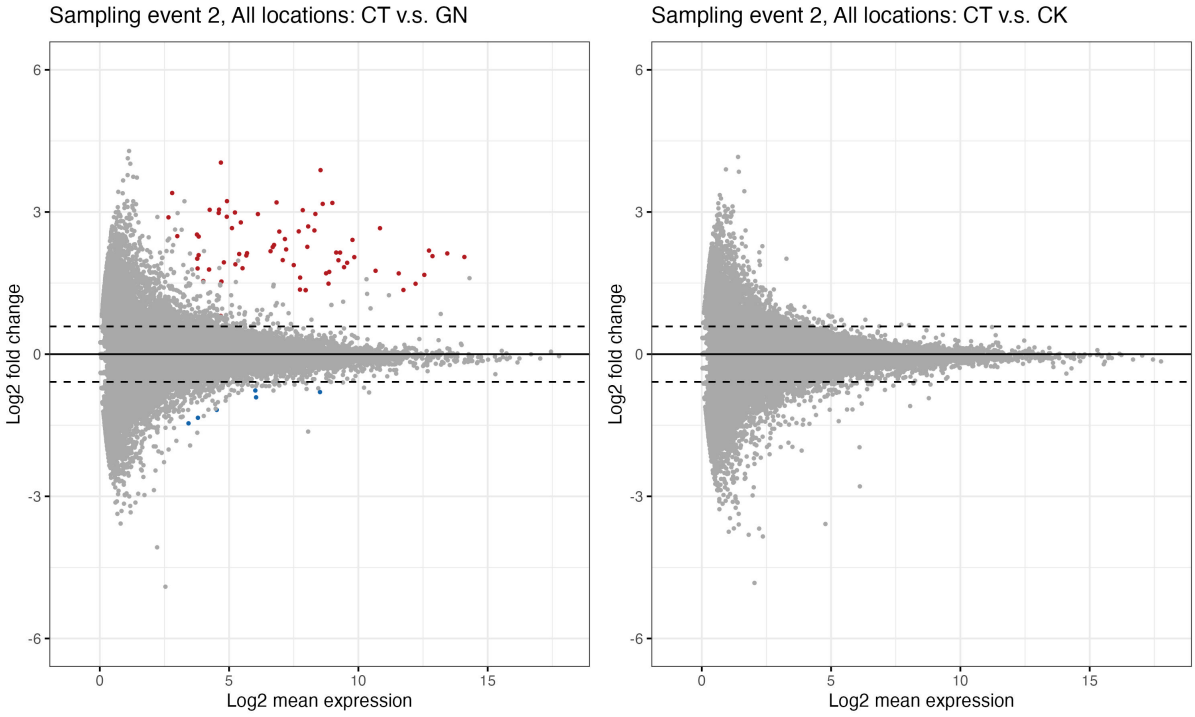

Figure 5. Differential expression genes analysis. (**a**) *Globisporangium nunn*-added and (**b**) *Chironomus kiiensis*-removed treatment. Red and blue points indicate significant up- and down-regulated genes, respectively. Upper and lower dashed lines indicate $\log_2(1.5)$ and $-\log_2(1.5)$, respectively.

series-based causal inference methods required quantitative data (i.e., relative abundance data that is typical for amplicon sequencing is not applicable) (Schreiber 2000; Sugihara *et al.* 2012; Osada & Ushio 2021), and our method justifies the use of the advanced time series method to infer the causal relationship between rice growth and ecological community members.

Rice growth was monitored by a manual measurement, e.g., rice height measurements made using rulers (Fig. 1b). Although the measurements successfully captured the seasonal dynamics of the rice growth, one might speculate whether only a single variable is sufficient to capture the rice performance. Indeed, it is true that assessing more variables, for example, gene expression (Nagano *et al.* 2019) and multispectral image monitoring during the whole growth season (Candiago *et al.* 2015), would be potentially useful to fully capture the trajectory of the rice growth. Nonetheless, we suggest that very frequent monitoring of a single variable is sufficient to reconstruct and predict the dynamics of the whole system dynamics (Takens 1981). Thus, we suggest that our monitoring approach can resolve the dynamics at least as a proof-of-concept study, but taking more variables at fine time resolution is certainly beneficial for more accurate delineation of the system dynamics.

Detections of potentially influential organisms using nonlinear time series analysis

By using UIC (Osada & Ushio 2021), we detected 718 potentially influential ASVs. Short read amplicon sequencing and conservative taxa assignment algorithm of Claident (Tanabe & Toju 2013) did not allow to assign the species name to most of the ASVs, and we could assign the species name only for 52 ASVs. Among them, we detected a previously known detrimental fungus, *Penicillium citrinum*, that may colonize rice and produce a toxin called citrinin that could cause human health problems (Ostry *et al.* 2013; Ali 2018), which partially supported the validity of our time series analysis. However, we could not find information regarding the

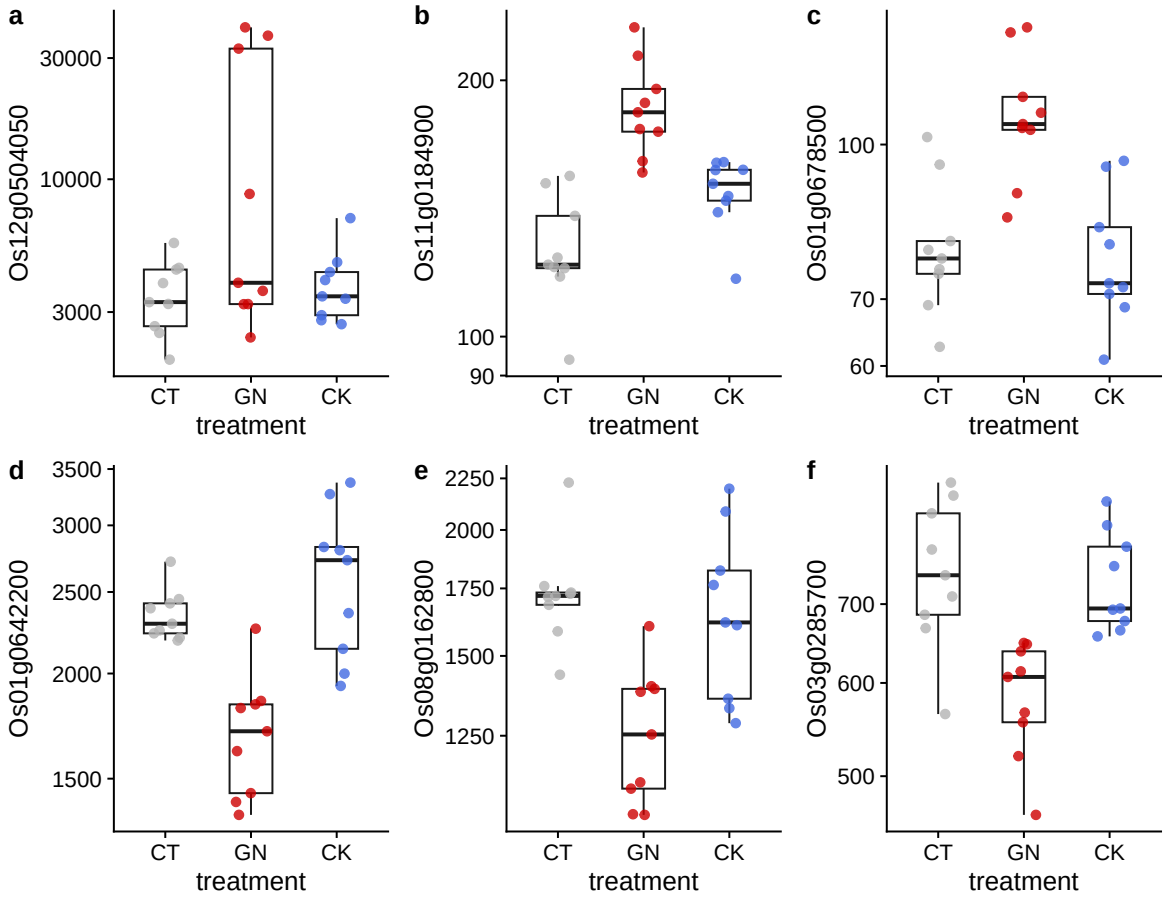

Figure 6. Examples of differentially expressed genes after the manipulation experiment. Results of (**a**) Os12g0504050, (**b**) Os11g0184900, (**c**) Os01g0678500, (**d**) Os01g0642200, (**e**) Os08g0162800, and (**f**) Os03g0285700 are presented. *y*-axis represents DESeq2-normalized read counts. Gray, red, and blue indicate CT (control), GN (*Globisporangium nunn* added), and CK (*Chironomus kiiensis* removed) treatments, respectively. The gene expressions of the GN treatment in all six genes are statistically clearly different from those of the other two treatments ($P < 0.0001$) except for GN v.s. CK in **b** ($P = 0.0087$) and GN v.s. CK **e** ($P = 0.00014$).

relationship between rice performance and most of the potentially influential ASVs by searching with the species name. Thus, at least at the moment, it is unknown whether the species in Table S3 really include influential organisms, and the effect of these species should be verified in future studies. Targeting ASVs with a large TE value would be a promising direction to find previously unknown beneficial/detrimental organisms for rice under field conditions.

In 2019, we focused on two species, *G. nunn* and *C. kiiensis* (Fig. S2). At the time of that analysis, these choices were reasonable: *G. nunn* was previously reported to be beneficial for vegetables (Kobayashi *et al.* 2010) and did not have any clear detrimental effects on rice (which was necessary in order not to spread potential pathogens to nearby farms), and *C. kiiensis* (midge) is a common insect species in the study region. These species may have effects on water quality (e.g., the concentration of phosphorus) through their feeding behavior (Kawai *et al.* 2003). However, they have never been focused on in terms of rice growth management, and thus, they provided an opportunity to test the validity of our approach in the present study.

Field manipulation experiments in 2019

In 2019, we added incubated *G. nunn* in the GN treatment or removed larvae of *C. kiiensis* in the CK treatment (Fig. 3). Although the CK treatment did not change the community composition statistically clearly, the GN treatment changed the community compositions (Fig. 3b–d). In the GN treatment, the abundance of putative *Globisporangium* spp. clearly increased (Fig. 3b), suggesting that the introduction of *G. nunn* was successful. However, we also observed increases in the putative *Globisporangium* spp. abundance in some of the CT and CK samples. The rice plots were located close to each other (50–60 cm apart; Fig. 3a) and thus immigration might have occurred between the GN and the other plots. As for *C. kiiensis*, we did not observe a statistically clear decrease in *C. kiiensis* eDNA in the CK treatments (Fig. 3c), suggesting that *C. kiiensis* individuals or their eDNA persisted even after the removal treatment. This might suggest that the treatment was not successful, or that quantitative eDNA metabarcoding failed to capture the actual decrease in their abundance. On the other hand, we observed statistically clear increases in *C. kiiensis* eDNA in the CT and GN treatments. This response was unexpected, but the effects of the increase in *G. nunn* might be transmitted to other species through the interaction networks which were reconstructed in the previous study (Ushio 2022).

As for the rice growth, although these manipulations were mild (the GN treatment) or relatively unsuccessful (the CK treatment) compared with other conventional manipulations such as the addition of fertilizers and insecticides, we surprisingly found statistically clear effects of the manipulations under field conditions (Fig. 4b). The addition of *G. nunn* increased the cumulative rice growth and changed the gene expression pattern after the manipulations statistically clearly. The removal of *C. kiiensis* also affected the rice growth (Fig. 4b), and the gene expression patterns were changed when the pot locations were taken into account (Fig. S4c,f and Table S8), suggesting that the effects of ecological community members might be situation dependent (e.g., microclimate-dependent). Unfortunately, detailed mechanisms of these effects cannot be elucidated in the present study, but some inferences could be made from the gene expression patterns. For example, nuclear DEGs included genes related to photosynthesis and stress responses (Table S7 and S8). These changes were observed relatively quickly (less than 5 days after the manipulation; see Table S1) and thus they could be direct effects from *G. nunn*. Up-regulations of photosynthesis-related genes (e.g., Os08g0433350, Os04g0473150, Os12g0504050, and Os01g0791033 in Table S7) might have contributed to the increase in the cumulative rice heights. Stress response genes such as Os11g0184900 might also have contributed to the changes in the rice performance, but further studies are required to elucidate

the mechanisms by which ecological community members affect the rice performance under field conditions.

Potential limitations and future perspectives

Although the present study has shown that our framework is potentially useful for detecting potentially influential, previously ignored organisms for rice growth under field conditions, there are potential limitations that should be acknowledged.

One limitation is the quantitative capacity of the eDNA metabarcoding. The quantitative eDNA metabarcoding approach used in this study provides reasonable estimates of eDNA concentrations (Ushio *et al.* 2018b; Ushio 2022), and the eDNA concentrations usually positively correlate with the number or biomass of individuals. However, there are species- and environment-specific eDNA dynamics such as degradation and release processes (Dejean *et al.* 2011; Maruyama *et al.* 2014), which may bias the estimation of absolute abundance. Although our causality detection might not be strongly influenced by the species-specific bias (i.e., time series were standardized before the analysis), accurate estimations of the absolute abundance will improve the accuracy of analyses. Identifications and enumerations of individuals using eDNA would be a fascinating direction to improve the accuracy of the species abundance estimations.

Another limitation is that the rice monitoring in this study was not comprehensive. While intensive monitoring was conducted during the rice growing season, only a limited number of variables were measured for a limited number of individuals. To overcome this limitation, future studies should consider applying a more advanced monitoring approach, such as unamended aerial vehicle (Fujiwara *et al.* 2022), multispectral camera (Candiago *et al.* 2015), long-term gene expression monitoring (Nagano *et al.* 2019) and/or neural network-based image analysis (Toda & Okura 2019).

Lastly, while we used an advanced nonlinear time series analysis to detect potential causal organisms for rice growth (Osada & Ushio 2021), there is still room for improvement. Nonlinear time series analysis is also a rapidly developing field, and many new techniques are emerging for causal inference (Runge *et al.* 2019; Suzuki *et al.* 2022), quantification of interaction strength (Chang *et al.* 2021) and near future forecasting (Chen *et al.* 2020). Applying these techniques to agricultural data may help to identify previously ignored beneficial organisms under field conditions.

Conclusions

In this study, we demonstrated that intensive monitoring of an agricultural system and the application of nonlinear time series analysis are helpful for identifying potentially influential organisms under field conditions. Although the effects of the two species were relatively small, the research framework presented here has future potential. The intensive monitoring, nonlinear time series analysis, and in situ validation of the statistical results may be applicable for other agricultural and aquatic systems. For example, the application of our framework to fishery aquaculture systems may be helpful for detecting potential pathogens of fish. Also, if the continuous monitoring of soil ecological communities becomes possible, our approach can be extended to other crops and vegetables. An automated, real-time monitoring system combined with advanced machine learning methods would provide a promising tool for increasing the accuracy and applicability of our approach. In conclusion, we proposed a novel framework to integrate the ecological network concept and agricultural practices to explore and detect

potentially influential organisms. Our proof-of-concept study would be an important basis for the further development of field-basis system management.

Endnotes

Data and analysis code availability

Computer codes used in the present study are available in Github (https://github.com/ong8181/rice-ecolnet-2019). Sequence data were deposited in DDBJ Sequence Read Archives (DRA). The accession numbers are as follows: DRA009658, DRA009659, DRA009660 and DRA009661 for eDNA data of ecological communities in 2017 (Ushio 2022), DRA015682, DRA015683, DRA015685, and DRA015686 for eDNA for eDNA data of ecological communities in 2019, and DRA015706 for rice RNA-seq data.

Acknowledgements

We thank Asako Kawai and Mutsumi Kato for assistance in field monitoring and DNA library preparations and Yutaka Osada for advice on the data analysis. This research was supported by PRESTO (JPMJPR16O2) from the Japan Science and Technology Agency (JST), KAKENHI (B) 20H03323 (to MU), and the Hakubi Project in Kyoto University.

Conflict of Interest Statements

The authors declare no conflicts of interest.

Author Contributions

MU conceived the idea and designed the research; MU, HS, and MT performed field monitoring and experiments; MU and AJN performed laboratory experiments; MU analyzed the data; MU wrote the first draft and completed the final manuscript with contributions from all coauthors.

References

Aleksandrowicz, L., Green, R., Joy, E.J.M., Smith, P. & Haines, A. (2016). The impacts of dietary change on greenhouse gas emissions, land use, water use, and health: A systematic review. *PloS One*, 11, e0165797.

Ali, N. (2018). Co-occurrence of citrinin and ochratoxin a in rice in asia and its implications for human health. *Journal of the Science of Food and Agriculture*, 98, 2055–2059.

Allesina, S. & Tang, S. (2012). Stability criteria for complex ecosystems. *Nature*, 483, 205–8.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A. & Holmes, S.P. (2016). DADA2: High-resolution sample inference from illumina amplicon data. *Nature Methods*, 13, 581–583.

Candiago, S., Remondino, F., De Giglio, M., Dubbini, M. & Gattelli, M. (2015). Evaluating multispectral images and vegetation indices for precision farming applications from UAV images. *Remote Sensing*, 7, 4026–4047.

Chang, C.-W., Miki, T., Ushio, M., Ke, P.-J., Lu, H.-P., Shiah, F.-K., *et al.* (2021). Reconstructing large interaction networks from empirical time series data. *Ecology Letters*, 24, 2763–2774.

Chen, P., Liu, R., Aihara, K. & Chen, L. (2020). Autoreservoir computing for multistep ahead prediction based on the spatiotemporal information transformation. *Nature Communications*, 11, 4568.

Cohen, S.P. & Leach, J.E. (2019). Abiotic and biotic stresses induce a core transcriptome response in rice. *Scientific Reports*, 9, 6273.

Dejean, T., Valentini, A., Duparc, A., Pellier-Cuit, S., Pompanon, F., Taberlet, P., *et al.* (2011). Persistence of environmental DNA in freshwater ecosystems. *PLOS ONE*, 6, e23398.

Deyle, E.R., May, R.M., Munch, S.B. & Sugihara, G. (2016). Tracking and forecasting ecosystem interactions in real time. *Proceedings of the Royal Society B: Biological Sciences*, 283, 20152258.

Dushoff, J., Kain, M.P. & Bolker, B.M. (2019). I can see clearly now: Reinterpreting statistical significance. *Methods in Ecology and Evolution*, 10, 756–759.

Fujiwara, R., Kikawada, T., Sato, H. & Akiyama, Y. (2022). Comparison of remote sensing methods for plant heights in agricultural fields using unmanned aerial vehicle-based structure from motion. *Frontiers in Plant Science*, 13.

Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., *et al.* (2010). Food security: The challenge of feeding 9 billion people. *Science*, 327, 812–818.

Hsieh, C., Glaser, S.M., Lucas, A.J. & Sugihara, G. (2005). Distinguishing random environmental fluctuations from ecological catastrophes for the north pacific ocean. *Nature*, 435, 336–340.

Kamitani, M., Kashima, M., Tezuka, A. & Nagano, A.J. (2019). Lasy-seq: A high-throughput library preparation method for RNA-seq and its application in the analysis of plant responses to fluctuating temperatures. *Scientific Reports*, 9, 7091.

Kashima, M., Sakamoto, R.L., Saito, H., Ohkubo, S., Tezuka, A., Deguchi, A., *et al.* (2021). Genomic basis of transcriptome dynamics in rice under field conditions. *Plant and Cell Physiology*, 62, 1436–1445.

Kawai, K., Kawai, T. & Imabayashi, H. (2003). A comparison of improvemental ability of water quality among five chironomid species of the genus chironomus. *Medical Entomology and Zoology*, 54, 37–42.

Kobayashi, S., Uzuhashi, S., Tojo, M. & Kakishima, M. (2010). Characterization of pythium nunn newly recorded in japan and its antagonistic activity against p. Ultimum var. ultimum. *Journal of General Plant Pathology*, 76, 278–283.

Lenaerts, B., Collard, B.C.Y. & Demont, M. (2019). Review: Improving global food security through accelerated plant breeding. *Plant Science*, 287, 110207.

Maruyama, A., Nakamura, K., Yamanaka, H., Kondoh, M. & Minamoto, T. (2014). The release rate of environmental DNA from juvenile and adult fish. *PLOS ONE*, 9, e114639.

May, R.M. (1972). Will a large complex system be stable? *Nature*, 238, 413–414.

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J.Y., Sato, K., *et al.* (2015). MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: Detection of more than 230 subtropical marine species. *Royal Society Open Science*, 2, 150088.

Nagano, A.J., Kawagoe, T., Sugisaka, J., Honjo, M.N., Iwayama, K. & Kudoh, H. (2019). Annual transcriptome dynamics in natural environments reveals plant seasonal adaptation. *Nature Plants*, 5, 74–83.

Nagano, A.J., Sato, Y., Mihara, M., Antonio, B.A., Motoyama, R., Itoh, H., *et al.* (2012). Deciphering and prediction of transcriptome dynamics under fluctuating field conditions. *Cell*, 151, 1358–1369.

Osada, Y. & Ushio, M. (2021). R package for unified information-theoretic causality.

Ostry, V., Malir, F. & Ruprich, J. (2013). Producers and important dietary sources of ochratoxin a and citrinin. *Toxins*, 5, 1574–1586.

Paine, R.T. (1966). Food web complexity and species diversity. *The American Naturalist*, 100, 65–75.

Runge, J., Bathiany, S., Bollt, E., Camps-Valls, G., Coumou, D., Deyle, E., *et al.* (2019). Inferring causation from time series in earth system sciences. *Nature Communications*, 10, 1–13.

Savary, S., Willocquet, L., Pethybridge, S.J., Esker, P., McRoberts, N. & Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nature Ecology & Evolution*, 3, 430–439.

Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, 85, 461–464.

Sugihara, G., May, R., Ye, H., Hsieh, C., Deyle, E., Fogarty, M., *et al.* (2012). Detecting causality in complex ecosystems. *Science*, 338, 496–500.

Suzuki, K., Matsuzaki, S.S. & Masuya, H. (2022). Decomposing predictability to identify dominant causal drivers in complex ecosystems. *Proceedings of the National Academy of Sciences*, 119, e2204405119.

Taberlet, P., Bonin, A., Zinger, L. & Coissac, E. (2018). *Environmental DNA: For biodiversity research and monitoring*. OUP Oxford.

Takens, F. (1981). Detecting strange attractors in turbulence. In: *Dynamical systems and turbulence* (eds. Rand, D. & Young, L.-S.). pp. 366–381.

Tanabe, A.S. & Toju, H. (2013). Two new computational methods for universal DNA barcoding: A benchmark using barcode sequences of bacteria, archaea, animals, fungi, and land plants. *PloS one*, 8, e76910.

Team, R.C. (2022). R: A language and environment for statistical computing.

Toda, Y. & Okura, F. (2019). How convolutional neural networks diagnose plant disease. *Plant Phenomics*, 2019.

Toju, H., Peay, K.G., Yamamichi, M., Narisawa, K., Hiruma, K., Naito, K., *et al.* (2018). Core microbiomes for sustainable agroecosystems. *Nature Plants*, 4, 247–257.

Ushio, M. (2022). Interaction capacity as a potential driver of community diversity. *Proceedings of the Royal Society B: Biological Sciences*, 289, 20212690.

Ushio, M., Hsieh, C., Masuda, R., Deyle, E.R., Ye, H., Chang, C.-W., *et al.* (2018a). Fluctuating interaction network and time-varying stability of a natural fish community. *Nature*, 554, 360–363.

Ushio, M., Murakami, H., Masuda, R., Sado, T., Miya, M., Sakurai, S., *et al.* (2018b). Quantitative monitoring of multispecies fish environmental DNA using high-throughput sequencing. *Metabarcoding and Metagenomics*, 2, e23297.

Uzuhashi, S., Kakishima, M. & Tojo, M. (2010). Phylogeny of the genus pythium and description of new genera. *Mycoscience*, 51, 337–365.

Van der Maaten, L. & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9, 2579–2605.

Wing, R.A., Purugganan, M.D. & Zhang, Q. (2018). The rice genome revolution: From an ancient grain to green super rice. *Nature Reviews Genetics*, 19, 505–517.

Wootton, J.T. & Emmerson, M. (2005). Measurement of interaction strength in nature. *Annual Review of Ecology, Evolution, and Systematics*, 36, 419–444.