이학석사 학위논문

교란 빈도가 토양 미생물 군집의 구조와 기능에 미치는 영향: 메타지노믹스를 통한 접근

Effect of disturbance on structure and function of soil microorganisms: A metagenomic approach

2017 년 2 월

서울대학교 대학원

생명과학부

최 수 범

교란 빈도가 토양 미생물 군집의 구조와 기능에 미치는 영향: 메타지노믹스를 통한 접근

Effect of disturbance on structure and function of soil microorganisms: A metagenomic approach

지도 교수 Jonathan Miles Adams

이 논문을 이학석사 학위논문으로 제출함
2017 년 2 월

서울대학교 대학원
생명과학부
최 수 범

최수범의 이학석사 학위논문을 인준함
2017 년 2 월

위 원 장 이 은 주 (인)

부위원장 Jonathan Miles Adams (인)

위 원 Piotr Jablonski (인)

Effect of disturbance on structure and function of soil microorganisms: A metagenomic approach

Soobeom Choi

Advisor: Jonathan Miles Adams

A Thesis Submitted for the Partial Fulfillment of the

Degree of Master of Science, Biological Sciences

February 2017

Graduate School of Biological Sciences

Seoul National University

Abstract

There has been little study of effects of disturbance on soil biota combining closely controlled experimental conditions and DNA-based methods. A soil microcosm system was studied subject to disturbance by investigating the taxonomic composition, gene function and gene expression patterns using metagenomics. In this paper, effect of time successional and disturbance frequency on soil communities was studied during 24 weeks.

First, a soil microcosm system was set up where pots of soil were sampled at varying times following an initial simulated mass mortality event. Soil DNA was extracted at intervals up to 24 weeks after the event for studying time successional recovery, and the shotgun metagenomes were sequenced using NextSeq. Compared to the initial conditions, results showed: consistent, sequential changes in functional metagenome and community structure over time, indicating successional niche differentiation. As predicted, the early successional systems had greater abundance of genes associated with motility, but fewer genes relating to DNA/RNA/protein metabolism, cell division and cell cycle. However, contrary to predictions, there were no significant differences in cell signaling and virulence & defense, and less abundant stress related genes in late succession. The early successional system had lower taxonomic diversity but higher functional gene diversity. Over time, all these characteristics changed in the direction of a return towards the initial conditions before disturbance, although recovery 'stalled' at later stages.

Secondly, an experiment was set up where pots of soil were incubated over 24 weeks, undergoing ‘disturbance events’ at different frequencies. Soil DNA was extracted from the five different levels of disturbance frequencies, and the shotgun metagenomes were sequenced using NextSeq. Diversity of soil microbe had consistently decline by increasing disturbance but functional gene diversity increased. All these characteristics changed by disturbance frequency, less disturbed samples tended to be similar to initial conditions although distinction became less among disturbed samples. As predicted, there were general differences through frequent disturbance and the communities with high disturbance had greater abundance of genes associated with DNA/RNA metabolism, motility, cell division and cell cycle. However, contrary to predictions, protein metabolism, stress response related genes were more abundant in lower disturbed communities.

Results indicated that there is a predictable sequence of gene functions and taxa that accompany succession after soil disturbance and disturbance frequency within the communities, paralleling general patterns of ecosystem succession for large organisms.

Keyword : Disturbance, Succession, Soil microbe, Metagenomics , Microbial communities, Microbial ecology

Student Number : 2015-20449

List of Figures

List of Tables

Table of Contents

Chapter 1. Introduction

Disturbance – a physical event involving the death of a large proportion of a community in a short period of time – is widely regarded as an important influence on community structure of larger organisms. For instance, the effects of disturbance have been described for forest trees (Molino and Sabatier, 2001), herbaceous plants (Ikeda, 2003), algae (Lubchenco and Menge, 1978) and also corals (Connell, 1978), amongst others. Physical disturbance of ecosystems releases nutrients by breakdown of organisms, providing available resources and altering ecosystem function. It is also known to lead to changes in species richness (Connell, 1978; Tilman, 1994). After disturbance, 'lottery effects' provide ready access to resources of all species in similar niches, regardless of their effectiveness in steady competition. By exploiting random opportunities, it may be possible for more species to coexist in a disturbed community (Chesson and Warner, 1981). However, while disturbance may increase diversity in communities with moderate levels of disturbance, beyond a certain point more frequent disturbance causes a decline in diversity (Connell, 1978). By creating a range of niches for differing degrees of dispersal ability,

disturbance is often important in adding to community diversity (Connell, 1978; Tilman, 1994).

To investigate the effect of disturbance on diversity, community structure and ecosystem functioning, some studies have been conducted in both naturally or artificially disturbed systems (Huston, 1979; Paine and Levin, 1981; Sousa, 1984). However, in contrast to studies on larger organisms, the response of microbial communities to disturbance and primary successional environments has been relatively little studied – and most of what has been done has tended to focus on observation of naturally occurring processes rather than experimental manipulation of controlled conditions (Buckling et al., 2000; Fierer et al., 2003; Kang and Mills, 2004; Allison and Martiny, 2008; Langenheder and Székely, 2011). As in the study of microbial ecology in general, culture-independent methods of community analysis have provided new perspectives on the responses of microbial communities to disturbance, through both natural and anthropogenic processes (Griffiths et al., 2000; Shade et al., 2012).

Several studies in bacterial ecology have concentrated on newly created or drastically disturbed habitats, such as the infant gut (Trosvik et al., 2010), and rock pools (Langenheder and Székely, 2011) explaining the mechanisms of assembly and

reassembly during the colonization. In an earlier study using an experimental soil system, Laboratory of Geographical Ecology (Kim et al., 2013) studied the response of the assembly pattern of soil bacterial communities to repeated disturbance on a fine scale in a microcosm in incubated pots of soil. That study found that distinct, predictable bacterial assemblages occurred at each incubation time and frequency of disturbance, highlighting the potential role of niche differentiation of bacteria in relation to disturbance events, and its possible importance in producing the diversity of bacterial communities seen in nature.

Here, a similar system was used based instead upon a one-time nonspecific mass mortality events of soil biota, in small pots of soil kept in a laboratory growth chamber. The aim in this project was Ⅰ) to investigate the sequential changes which occur in community composition and functional genetics after a severe disturbance event affects a soil and Ⅱ) the response of microbial community structure and function to soil disturbance.

The work was structured around several main questions.

Ⅰ) What are the functional characteristics of early successional vs later successional soil communities?

The experiment explored changes in communities over a 168-day (24 week) time period. This, I propose, could be broadly

analogous to the process of ecological succession seen elsewhere in nature for larger organisms, for example following a forest fire or other disturbance event. Since many microbes have potential doubling times of the order of hours to weeks depending on the species (Schaechter et al., 1958; Clark and Maaløe, 1967; Cooper and Helmstetter, 1968; Brock, 1971; Fantes and Nurse, 1977), this 168-day experiment can be regarded as analogous to the studied time spans of secondary succession for plants or corals for example, where generation times vary from months to decades and successional systems span years to centuries (Connell, 1978; Myster and Pickett, 1992; Molino and Sabatier, 2001; Zhang et al., 2002). While I was unsure what time span successional changes in the recovering soil community might require to be completed, this 24 week study was intended as a 'first' glimpse on a time scale that might plausibly show significant changes. The time scale is also relevant to many agricultural systems, which are ploughed or sprayed on a cycle of 6 months or so (Lindén and Wallgren, 1993; Davies et al., 2001).

In the context of the broader background of community and ecosystem ecology for larger organisms, I made the following predictions for differences in metagenome functions for 'earlier' and 'later' successional bacterial communities. Many of these

predicted traits can be seen in terms of the 'r' versus 'K' dichotomy in ecology (MacArthur and Wilson, 2015), although here it applies to aggregate traits of the whole community, rather than focusing on individual species.

Cell division related genes will be more abundant in the early successional stage.

Cell division is an aspect of growth, particularly in the prokaryotic world. Typically in ecology of larger organisms, in the early successional stages with abundant space and release of nutrients from dead organisms, species grow fast and give many offspring, rapidly expanding their biomass (Odum, 1969; Connell and Slatyer, 1977; Bazzaz, 1979). Thus I expected this pattern to hold true in soil microbiota succession.

At the later successional time stage, 'housekeeping' genes associated with basic metabolic functions will become relatively less common than genes for other extra functions associated with nutrient acquisition and competition.

In secondary successional systems in general, mass death of organisms is accompanied by a release of nutrients which then become readily available for uptake (Odum, 1969). As living biomass increases, competition increases and usage of recalcitrant nutrients plus interference competition becomes more important

(Odum, 1969). I predicted that later successional stages in this soil systems would require greater numbers of genes associated with sequestering nutrients from the more recalcitrant amongst the polymers associated with dead cells and soil humus, and also with antagonism. In classical ecology, traits associated with nutrient acquisition and with interference competition are seen as being more abundant in later successional stages (Odum, 1969).

Genes related to cell-cell interactions will be more abundant in the later successional stages.

It is generally agreed in ecology that later successional ecosystems have more intense and species-specific mutualistic and antagonistic interactions (Odum, 1969). One example from forest successions would be the heavier reliance of plants on Ectomycorrhizal (EcM) fungi. I predicted similar trends towards intensity of (either positive or negative) organismic interactions in the late successional stage in the soil bacterial systems I was studied. The result was predicted that genes associated with antibiotic resistance, production of secondary compounds, quorum sensing and cell-cell recognition would become more important.

Viruses and anti-viruses defenses will become more abundant at later stages in the succession.

High mortality of natural populations from diseases is associated with stable and high density populations (Solomon, 1949; Burdon and Chilvers, 1982; Gilbert, 2002). The prediction was that in the early stages of succession after disturbance, a low total abundance of cells of all kinds should result in lower transmissivity and lower infection rates by viruses. As soil living biota build up in biomass, cell−cell neighbor distances should decrease (for all types of organism), resulting in greater infection rates. Likewise, strains of bacteria and archaea bearing CRISPR elements as anti−viral defenses (Horvath and Barrangou, 2010)will be more strongly selected in later stages, resulting in greater frequency of these in the metagenome.

Motility related genes will be more abundant in early succession.

In early successional organisms, dispersal is seen as a key trait in exploiting the environment fully by finding open space and resources to increase population (Nathan and Muller−Landau, 2000; Bell, 2012).Following the initial disturbance event, organisms from the 10 % of ‘living’ soil mixed back in will be surrounded by a large proportion of potentially colonizable soil rich in resources from killed biota. This should select for organisms with the means to disperse within soil (e.g. flagellae), and the greater abundance of

this trait should show up as genes for such characteristics as flagellae.

Stress and dormancy genes will become more abundant in later succession.

In secondary succession of larger organisms, resources are at first abundant, but become less abundant over time as they are sequestered in living biomass (Horn, 1974; Brown and Lugo, 1990; Guariguata and Ostertag, 2001). In late succession, not only nutrient shortage but interference competition between organisms becomes more common (Bazzaz, 1979). I predicted that likewise, there would be an increase in stress response genes related to both nutrient shortage and interference competition (e.g. antibiotic effects). I also predicted that dormancy genes would become more common, through selection of organisms able to survive in dormant form under nutrient shortage or interference competition.

Ⅱ) What are the taxonomic and functional characteristics among the samples treated with different disturbance frequency?

A form of relationship between disturbance frequency and taxonomic diversity in terms of richness and equability might be a classic 'humpbacked' curve in relation to disturbance frequency.

In traditional concept of disturbance, disturbance tends to prevent the few fittest species from dominating the community and

allows the weak to continue to coexist with the strong if disturbance is sufficiently frequent (Adams, 2010). However, if the community subjected to short periodic disturbance, it will not be able to recover enough to build up a healthy population again. Importance of intermediate levels of disturbance in promoting species richness is organized in Intermediate disturbance theory which explains the relationship between the amount of disturbance and species richness (Connell, 1978). This gradual changes of species diversity drawn a 'humpbacked' curve which have positive correlation with increasing in disturbance frequency up to a certain point, but beyond which more frequent disturbance, causes a decline in diversity. Previous study of relation between bacterial diversity and disturbance frequency in our group showed that bacterial diversity decreased with increasing disturbance frequency (Kim et al., 2013). In this study, I will use metagenomics to encompass all the microorganisms not only bacteria but soil Eukaryota, Archaea and Viruses to find out whether classic relation between disturbance frequency and species diversity will be shown in all the soil microbial community or not.

Groups of organisms usually regarded as 'copiotrophic' and 'fast growing' will be abundant in more frequently disturbed soil systems with discrete communities at each disturbance level.

Short generation times will favor these groups, and the ability of exploit high nutrient concentrations is important to this (Grime, 2006). The result was predicted that bacteria will increase in abundance relative to fungi. It was also predicted that Basidiomycetes, which are generally thought to grow more slowly but able to utilize recalcitrant substrates, will be replaced by Ascomycetes at higher disturbance frequencies. This is because Ascomycetes are able to grow at faster potential rates through utilizing relatively labile substrates (Wang et al., 2010). Metagenomic profiling reveals the entire spectrum of organisms, and their relative abundances, and this is an effective method of assessing changes in relative abundance or biomass.

Viruses and virus related genes will be more abundant and diverse in more disturbed systems

By analogy, an unstable system might be expected to have a lower relative abundance of viruses which will be slower to catch up their numbers with the population of host cells. However, there are also reasons for supposing that viruses could actually be more abundant in the fast-growing r-selected populations of more disturbed systems (Suttle, 2007). With dense populations of potential hosts (copiotrophic bacteria or fungi), whose strategy emphasizes rapid reproduction over defense and genetic re-

assortment, viruses in having very brief generation times may be better able to spread through their specific host populations causing outbreaks and population crashes. Thus, genes related to viruses and anti-viruses defenses will relatively increase.

That taxonomic diversity will be paralleled by functional diversity in relation to the range of cellular and biochemical gene functions definable by metagenomics gene databases.

Changes in taxonomic diversity causes changes of functional diversity. Another core principle of ecology is that diversity of species and taxonomic groups begets diversity of genetically determined functionality in ecosystems, and from this in turn comes greater stability against perturbation, and greater ability to recover quickly following disturbances. Various experiments in Minnesota (Tilman et al., 2001; Grime, 2006; Tilman et al., 2006) involving mixtures of prairie grassland plants in microcosm systems have suggested that for plants, greater diversity may induce greater stability and resilience, at least up to a point. The principle that greater diversity leads to greater functional diversity at a genetic level has never been effectively tested for microbial communities, even though their complete metagenomic profiles are amenable to comparing both taxonomic and genetic diversity.

That genes associated with 'housekeeping', motility and stress responses will become more abundant in highly disturbed system

That 'housekeeping' genes are continuously expressed to maintain cellular functions (Zhu et al., 2008; She et al., 2009; Eisenberg and Levanon, 2013). Their functions associated with DNA replication, RNA processing, protein metabolism and cell cycle will be relatively more common in the more frequently disturbed systems. In organisms with more of other genes involved in secondary metabolism or other special functions in their genome, the house keeping genes are 'diluted' as a proportion of the total population of genes in the metagenome. Thus it is predicted that genes associated with secondary metabolism and cell-cell interaction will be more common in less disturbed systems. Also, in more disturbed system, genes related to stress responding and motility will be more abundant. Since disturbance is thought to be important in differing degrees of dispersal ability through creating a range of niches (Tilman, 1994), communities under disturbance will have high rate of motility and stress responding genes to resist disturbance.

Chapter 2. Materials and methods

2.1. Preparation of soil microcosms for disturbance gradient study

The soil for this experiment was taken from an area approximately 30 m x 20 m in a fallow field located in Suwon, South Korea (Fig. 1, lat. 37° 16'N, long.126° 59'E) on the University Farm of Seoul National University in mid-June 2015. Soil was taken from the top 10 cm with large roots and stones removed, and it was thoroughly homogenized by sieving through a 2-mm mesh. The experiment itself was conducted in the laboratories of the School of Biological Sciences of Seoul National University. The sieved soil was sterilized by autoclaving twice for 90 minutes each time at 121 ℃, to provide a large pool of sterile soil for 'restocking' pots following disturbance. Soil sterility was validated by enumeration of heterotrophic bacteria by the most probable number (MPN) technique (Wertz et al., 2006; Wertz et al., 2007), and absence of significant quantities of intact DNA (less than 10 ng/μl) by failure of test PCR for bacteria 16S rRNA and fungal ITS region. Soil was frozen at −20 ℃ until ready for the experiment to begin.

Each pot in the experimental system was initially stocked with 450 g of moist sterilized soil and 50 g of moist unsterilized sieved soil from the same batch, which has been stored at 3 ℃ for one week prior to the experiment (Fig. 2A). This was intended to simulate a 'disturbance' event in which 90 % of soil biota had been killed, with the remaining 10 % going on to recolonize the whole bulk of the soil. The sterilized and unsterilized soil were thoroughly mixed by shaking in a sterile bag before being placed in the pot.

For studying time succession effect, I set three replicates of disturbed and incubated controls with different incubation times: 1-week, 2-week, 4-week, 8-week, 12-week and 24-week and non-disturbance control incubated 24-week. One 0-week sample was collected after the disturbance event (Fig. 2B). Also, to analysis disturbance effect, I set five different levels of disturbance frequencies: every 1-week, 2-week, 4-week, 8-week, 12-week, and no disturbance (Fig. 2C). The highest disturbance frequency (1-week treatment) set to 1.0 and no disturbance setting set to 0.0. The highest disturbance frequency (1-week treatment) is determined based on estimated mean turnover times of soil bacteria. In total 49 pots (15 ×15 ×13 cm, ca. 500 g of soil each) were arranged in a completely randomized design with each replicates.

Each soil-filled pot was placed with another empty pot upside down over it, forming a lid that was sealed around the edges with tape, to avoid as much as possible additional recruitment from direct cross contamination by dust from adjacent pots. Five holes (each 0.5 mm in diameter) were drilled in the top of each pot, enough to allow aeration but unlikely to allow soil dust to travel between pots.

The pots were kept in a growth chamber in darkness at 24 ℃ and all were watered every week with a uniform 50 ml of sterile distilled water to maintain a moderately damp soil. At each watering time, all pots were switched around randomly in terms of their positions within the growth chamber. The experiment was continued for 24 weeks and soil samples for DNA extraction were be taken from the mixture of each pot.

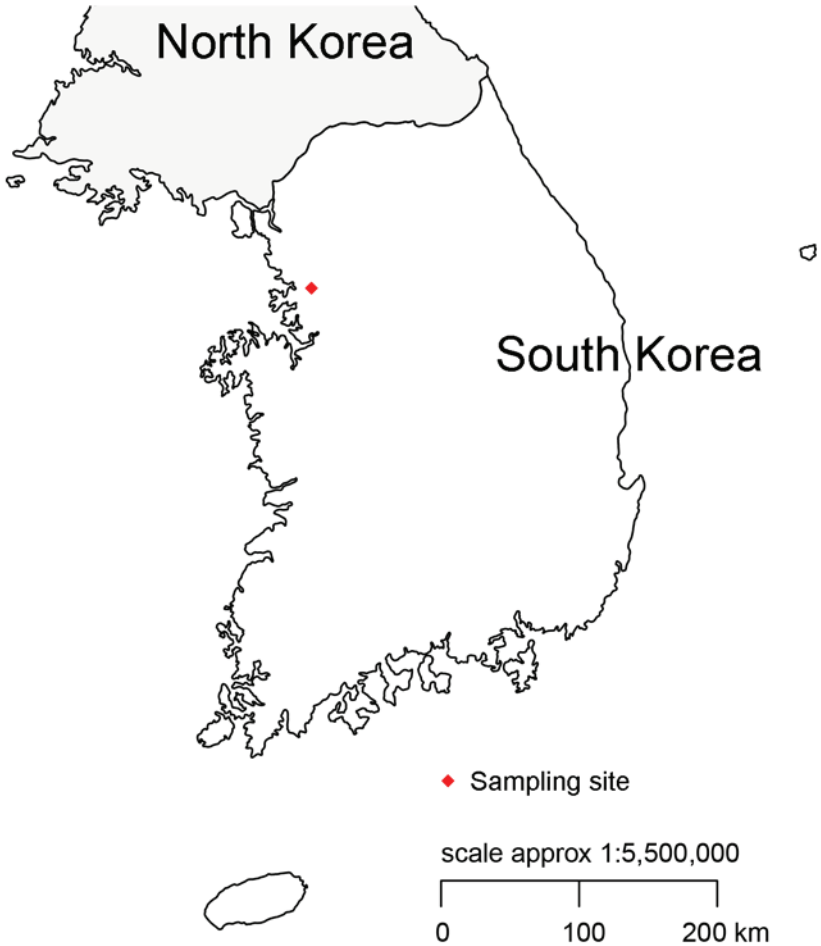
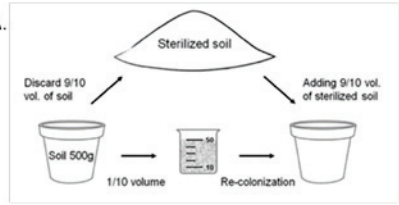


Figure 1. A map of sampling area generated by using software R version 3.1.2. University Farm of Seoul Nation University, Suwon, South Korea. Latitude: 37° 16'N, Longitude: 126° 59'E

A.



B.

Samples\week	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
0	D	C																							
1	D	C																							
2	D		C																						
4	D				C																				
8	D								C																
12	D												C												
24	D																								C
Nondisturbed																									C

C.

Samples\week	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1 (24/24)	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	C
0.5 (12/24)	D		D		D		D		D		D		D		D		D		D		D		D		C
0.25 (6/24)	D				D				D				D				D				D				C
0.125 (3/24)	D								D								D								C
0.0833 (2/24)	D												D												C
Nondisturbed																									C

Figure 2. Experimental design. A. Figure of mixing 90 % of moist sterilized soil and 10 % of moist unsterilized sieved soil, B. Name of samples and incubation weeks after initial disturbance, C. Name of samples and weeks of disturbance events. D: disturbance, C: collection.

2.2. Chemical analysis and DNA extraction

Soil analyses were carried out at the National Instrumentation Center for Environmental Management (NICEM, South Korea). Soil pH, total nitrogen(TN) and total organic carbon (TOC) content were measured using standard protocols of the Soil Science Association of America (SSSA) (Allen et al., 1974).

Total DNA was extracted from all the samples (from both disturbance and culture study) using the Power Soil DNA extraction kit (MoBio Laboratories, Carlsbad, CA, USA) following manufacturer's instructions and stored at −20 ℃ until further processing.

2.3. Metagenomic sequencing and bioinformatics analyses

All the purified DNA samples were sequenced using a Nextseq500 paired−end of 2 x 150 bp (Illumina) at Celemics Incorporation (Seoul, Korea). In total 164,488,287 reads of sequence data were generated. Paired end metagenome sequences were annotated with the Metagenomics Rapid Annotation (MG−RAST) pipeline version 3.3 (Meyer et al., 2008). Phylogenetic information was extracted from the metagenomes using M5NR data

bases using BLASTX (e-value less than 1 X 10^{-5}, minimum percent identity was 60 % and sequence match length greater than 15 nucleotides). Functional profiles were generated with the SEED subsystems database (Overbeek et al., 2005) using a maximum e-value of 1e-5, a minimum identity of 60 %, and a minimum alignment length of 15 aa. These profiles were then normalized for differences in sequencing coverage by calculating percent distribution, prior to downstream statistical analysis. All of the sequence data are available under the MG-RAST project ID 18857 and 18393 (http://metagenomics.anl.gov/linkin.cgi?project=18857, http://metagenomics.anl.gov/linkin.cgi?project=18393).

16S rRNA sequences were extracted from the 50 metagenomes including 1 original soil samples, 19 samples for time successional study, 25 samples for disturbance effect and 5 non-disturbance control samples by using SortMeRNA (Kopylova et al., 2012). Total number of extracted 16S rRNA sequences was 653092 and processed using the Mothur (Schloss et al., 2009). Each length of all of sequences was 151. Homopolymers longer than 8 were removed from all of extracted sequences. Taxonomy was assigned using EzTaxon database (Chun et al., 2007) using classify command in Mothur. After the quality control, minimum value of

reads per sample was 1245. Each sample was subsampled to 1245 reads to adjust the minimum value of samples.

A nonmetric multidimensional scaling (NMDS) plot was used to visualize the structure among samples, using the taxonomic and functional abundance matrix using PRIMER v6. The plots were generated from Bray–Curtis similarity index matrices (M5NR taxonomic profile at family level, SEED subsystem function at level 3 and bacterial 16S rRNA gene at genus level).

Diversity measurements (Shannon's index) were calculated based on M5NR taxonomy at species level as concept of diversity is based on species level and SEED subsystem function at level 3 by using software R version 3.1.2 (Team, 2014). I plotted Envfit with the content of soil parameters and diversity of each samples and mapped sampling site (Fig. 1) by using software R version 3.1.2. Software SigmaPlot 10.0 was used for regression study of Shannon diversity index on each treatments.

Chapter 3. Results

3.1. Study Ⅰ. Time succession: Effect of time successional recovery after soil disturbance on structure and function of soil community

From the 23 separate pot soil DNA samples, including 1 original garden soil sample, 19 disturbed-incubated control samples, and 3 non-disturbance incubation samples, sequence count 71,263,671 of sequence data were generated. Phylogenetic information was extracted from the metagenomes using M5NR data bases using BLASTX and functional profiles were generated with the SEED Subsystems database. Information of samples and results of soil chemical analysis are shown in Table 1 and Table 2. A total of 27-40 % of the sequences were annotated as protein using E < 1 X 105 and 15-bp minimum alignment length on MG-RAST server.

Table 1. Total number of base pairs, sequencing reads.

MG-RAST ID	Sample name	Treatment	Sampling date	Total bp	Sequence count	Predicted proteins	Annotated proteins (%)
4707578.3	OR	Original soil	2015-06-16	627969701	3473797	2367047	37.41
4707563.3	DC0	Disturbed&Incubated0wk	2015-07-07	572161561	3204375	2198562	27.7
4707570.3	DC1_1	Disturbed&Incubated1wk	2015-07-14	575197473	3114726	2113170	39.44
4707583.3	DC1_2	Disturbed&Incubated1wk	2015-07-14	554356262	3093279	2048950	39.58
4707561.3	DC1_3	Disturbed&Incubated1wk	2015-07-14	545495218	3018065	1976633	38.14
4707566.3	DC2_1	Disturbed&Incubated2wk	2015-07-21	601812503	3274033	2145887	31.7
4707564.3	DC2_2	Disturbed&Incubated2wk	2015-07-21	570773518	3155286	2120126	40.64
4707582.3	DC2_3	Disturbed&Incubated2wk	2015-07-21	552567024	3026986	2032050	31.4
4707581.3	DC4_1	Disturbed&Incubated4wk	2015-08-04	652609696	3637570	2558238	41.92
4707574.3	DC4_2	Disturbed&Incubated4wk	2015-08-04	538489571	2992906	2078311	32.9
4707589.3	DC4_3	Disturbed&Incubated4wk	2015-08-04	691109486	3855896	2652703	41.28
4707584.3	DC8_1	Disturbed&Incubated8wk	2015-09-01	582908908	3257066	2226347	28.9
4707577.3	DC8_2	Disturbed&Incubated8wk	2015-09-01	609806249	3412911	2366742	28.1
4707572.3	DC8_3	Disturbed&Incubated8wk	2015-09-01	652423120	3660170	2532948	29.5
4707568.3	DC12_1	Disturbed&Incubated12wk	2015-09-29	676225075	3762861	2570042	37.52
4707585.3	DC12_2	Disturbed&Incubated12wk	2015-09-29	643579358	3497651	2447951	37.7
4707580.3	DC12_3	Disturbed&Incubated12wk	2015-09-29	650301895	3553533	2457049	33.58
4707567.3	DC24_1	Disturbed&Incubated24wk	2015-12-22	566112460	3108174	2027623	34.65
4707565.3	DC24_2	Disturbed&Incubated24wk	2015-12-22	634245205	3399243	2034141	37.23
4707576.3	DC24_3	Disturbed&Incubated24wk	2015-12-22	580058103	3246637	2218102	30.85
4702982.3	NC1	Nondisturbed&Incubated24wk	2015-12-22	251762454	1416604	955073	34.82
4702995.3	NC2	Nondisturbed&Incubated24wk	2015-12-22	295991344	1644333	1145926	35.92
4702971.3	NC3	Nondisturbed&Incubated24wk	2015-12-22	263241389	1457569	1008593	33.18

4702989.3	NC4	Nondisturbed&Incubated24wk	2015-12-22	268132713	1481176	1050460	34.93
4702980.3	NC5	Nondisturbed&Incubated24wk	2015-12-22	334106633	1848889	1283210	34.33
4702967.3	T1_1	Disturbance Frequency: 1	2015-12-22	570977883	3130138	1831559	50.62
4702972.3	T1_2	Disturbance Frequency: 1	2015-12-22	759134203	4209432	2352792	50.86
4702984.3	T1_3	Disturbance Frequency: 1	2015-12-22	1522527691	8297494	4529012	51.77
4702981.3	T1_4	Disturbance Frequency: 1	2015-12-22	2095879179	11428311	6173549	51.23
4702994.3	T1_5	Disturbance Frequency: 1	2015-12-22	838829333	4584694	2640227	50.76
4702974.3	T2_1	Disturbance Frequency: 0.5	2015-12-22	1382755860	7623170	4094773	45.44
4702991.3	T2_2	Disturbance Frequency: 0.5	2015-12-22	726420143	3941838	2532870	43.8
4702976.3	T2_3	Disturbance Frequency: 0.5	2015-12-22	681454104	3767325	2367424	40.83
4702992.3	T2_4	Disturbance Frequency: 0.5	2015-12-22	524935120	2872536	1885787	44.26
4702983.3	T2_5	Disturbance Frequency: 0.5	2015-12-22	379698417	2069903	1375371	41.78
4702988.3	T4_1	Disturbance Frequency: 0.25	2015-12-22	458323520	2513034	1697819	38.38
4702987.3	T4_2	Disturbance Frequency: 0.25	2015-12-22	393816984	2158172	1439964	40.04
4702985.3	T4_3	Disturbance Frequency: 0.25	2015-12-22	425897545	2339629	1537920	38.45
4702966.3	T4_4	Disturbance Frequency: 0.25	2015-12-22	584082373	3224188	2063751	40.38
4702978.3	T4_5	Disturbance Frequency: 0.25	2015-12-22	516699190	2860404	1867016	39.67
4702986.3	T8_1	Disturbance Frequency: 0.125	2015-12-22	370497677	2039288	1343589	39.66
4702977.3	T8_2	Disturbance Frequency: 0.125	2015-12-22	845842384	4697108	3067153	38.56
4702968.3	T8_3	Disturbance Frequency: 0.125	2015-12-22	266506149	1466354	999957	41.13
4702975.3	T8_4	Disturbance Frequency: 0.125	2015-12-22	344102058	1901086	1251784	37.61
4702970.3	T8_5	Disturbance Frequency: 0.125	2015-12-22	496934737	2748139	1838295	38.08
4702973.3	T12_1	Disturbance Frequency: 0.0833	2015-12-22	514420404	2823441	1893616	39.56
4702993.3	T12_2	Disturbance Frequency: 0.0833	2015-12-22	595898657	3301223	2136738	42.44
4702990.3	T12_3	Disturbance Frequency: 0.0833	2015-12-22	401251633	2208832	1497110	40.34
4702979.3	T12_4	Disturbance Frequency: 0.0833	2015-12-22	372902610	2056981	1312653	42.45
4702969.3	T12_5	Disturbance Frequency: 0.0833	2015-12-22	286633404	1631831	1066045	38.89

Table 2. Results of soil chemical analysis.

Sample name	Treatment	pH	TOC (%)	TN (%)
OR	Original soil	6.7	1.59	0.157
NC1	Nondisturbed&Incubated24wk	6	1.45	0.179
NC2	Nondisturbed&Incubated24wk	5.9	1.54	0.185
NC3	Nondisturbed&Incubated24wk	6.6	1.49	0.162
T1_1	Disturbance Frequency: 1	6.7	1.49	0.168
T1_2	Disturbance Frequency: 1	6.9	1.49	0.151
T1_3	Disturbance Frequency: 1	6.7	1.49	0.173
T4_1	Disturbance Frequency: 0.25	6.6	1.37	0.155
T4_2	Disturbance Frequency: 0.25	6.5	1.42	0.159
T4_3	Disturbance Frequency: 0.25	6.6	1.42	0.155
T12_1	Disturbance Frequency: 0.0833	6.1	1.46	0.180
T12_2	Disturbance Frequency: 0.0833	5.4	1.44	0.205
T12_3	Disturbance Frequency: 0.0833	6.0	1.46	0.176
Regression		P=0.0172, R^2=0.449, Coefficient: 0.72	-	-
DC0	Disturbed&Incubated0wk	6.7	1.46	0.155
DC1_1	Disturbed&Incubated1wk	6.6	1.46	0.159
DC1_2	Disturbed&Incubated1wk	6.7	1.51	0.16
DC1_3	Disturbed&Incubated1wk	6.6	1.52	0.156
DC4_1	Disturbed&Incubated4wk	6.4	1.59	0.163
DC4_2	Disturbed&Incubated4wk	6.3	1.48	0.172
DC4_3	Disturbed&Incubated4wk	6.3	1.48	0.177
DC12_1	Disturbed&Incubated12wk	6.3	1.54	0.167
DC12_2	Disturbed&Incubated12wk	6.3	1.48	0.163
DC12_3	Disturbed&Incubated12wk	6.3	1.43	0.167
DC24_1	Disturbed&Incubated24wk	5.9	1.38	0.19
DC24_2	Disturbed&Incubated24wk	5.9	1.42	0.189
DC24_3	Disturbed&Incubated24wk	6.5	1.43	0.156
Regression		P=0.0099, R^2=0.502, Coefficient: -0.01907	P=0.0155, R^2=0.459, Coefficient: -0.00423	-

3.1.1. Study Ⅰ. Time succession: Dominant microbial taxa

The majority of the metagenomic sequences among disturbance incubation samples were dominated by Bacteria (83.1 % on average) followed by Eukaryota (2.7 %), Archaea (0.9 %) and Viruses (0.1 %) according to the M5NR database.13.2 % of reads were unassigned. Bacteria and Archaea relative abundances a showed linear relationship to one another over time, with relative abundance of Bacteria decreasing but Archaea increasing (Fig. 3A, Fig. 3C). In case of original soil which was 100 % unsterilized sieved soil, the total community based on taxonomically assigned reads consisted of Bacteria 83.9 %, Eukaryota 1.4 %, Archaea 1 % and Viruses 0.01 % on average (unassigned 13.7 %). Samples of the non-disturbed control samples that were incubated for 24 weeks were also dominated by Bacteria (84.5 % on average), plus Archaea (2 %), Eukaryota (1.2 %) and Viruses (0.02 %).

Abundant bacterial phyla from original soil which was 100% unsterilized sieved soil, the total community consisted of Proteobacteria 45.18 %, Actinobacteria 18.08 %, Firmicutes 6.08 % , Bacteroidetes 5.53 %, Acidobacteria 5.47 %, Chloroflexi 4.56 %, Planctomycetes 3.91 %, Verrucomicrobia 2.97 %, Cyanobacteria 2.64 %, and Gemmatimonadetes 1.24 % and samples

incubated 24-week with non-disturbed had similar population (Table 3). The most abundant bacterial phylum amoung the metagenomic sequence data after the disturbance event was Proteobacteria (37.27 % on average), followed by Bacteroidetes (26.05 %), Firmicutes (12.97 %), Actinobacteria (6.10 %), Acidobacteria (3.22 %), Planctomycetes (2.72 %), Verrucomicrobia (2.62 %), Chloroflexi (2.06 %), Cyanobacteria (1.89 %) and Nitrospirae (1.05 %). Among 27 phyla, 17 of them had significant ($P < 0.05$) relationship with time flow.

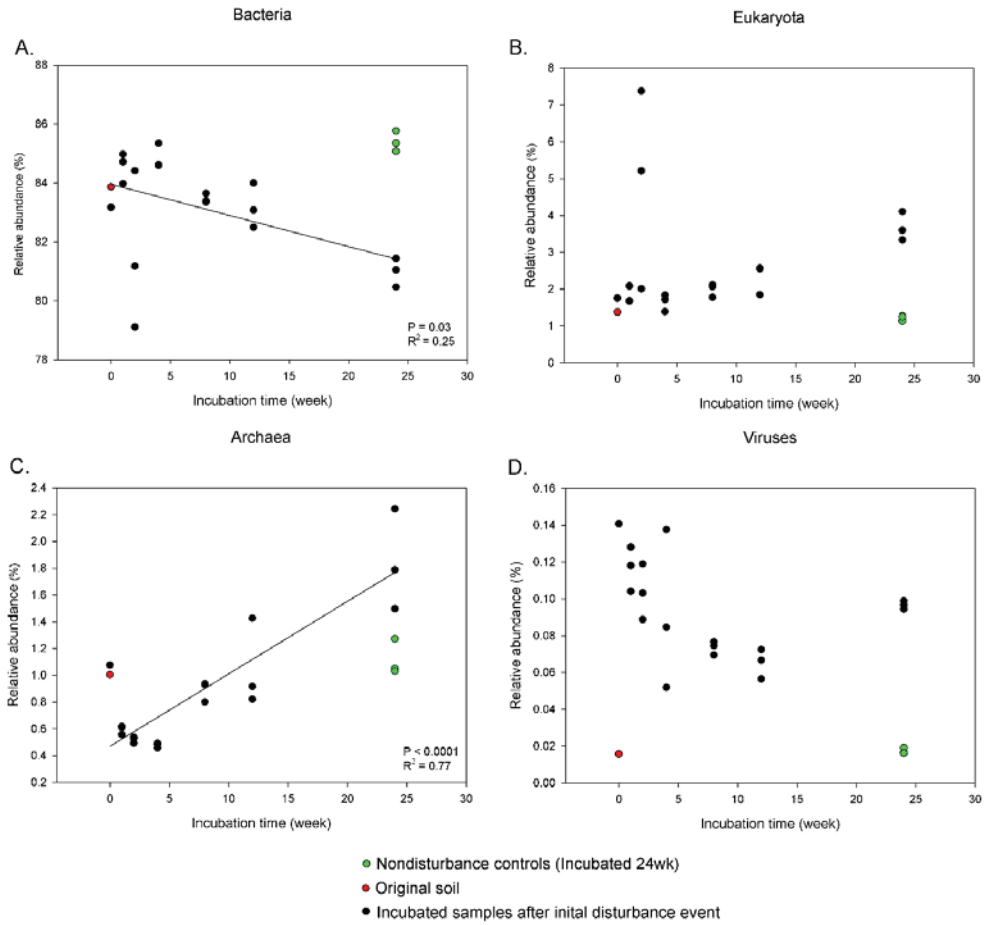


Figure 3. Relative abundance of each domain level assigned based on the M5NR database. Solid line represents linear regression fits to data. A. Relative abundance of Bacteria, B. Relative abundance of Eukaryota, C. Relative abundance of Archaea, D. Relative abundance of Viruses. The colored points were not included in the statistics.

Table 3. Relative abundance (%) (means ± SD) of bacterial phyla observed in shotgun metagenomics sequences.

phylum	Original soil	Nondisturbed control	Total average	Incubation control time							Regression	Coefficient
				0	1	2	4	8	12	24		
Poribacteria	0.07	0.08 ± 0.03	0.13 ± 0.13	0.08	0.04 ± 0.01	0.02 ± 0	0.07 ± 0.01	0.14 ± 0.01	0.17 ± 0.06	0.36 ± 0.18	P<0.0001, R^2=0.740	0.0141
Dictyoglomi	0.06	0.05 ± 0.02	0.05 ± 0.02	0.07	0.04 ± 0	0.04 ± 0	0.04 ± 0	0.05 ± 0.01	0.06 ± 0.01	0.07 ± 0.01	P<0.001, R^2=0.586	0.0015
Deinococcus-Thermus	0.82	0.71 ± 0.29	0.49 ± 0.13	0.81	0.34 ± 0.02	0.35 ± 0.03	0.48 ± 0.05	0.54 ± 0.02	0.56 ± 0.05	0.53 ± 0.15	x	
Gemmatimonadetes	1.24	0.95 ± 0.46	0.78 ± 0.53	1.17	0.21 ± 0.05	0.24 ± 0.09	1.36 ± 0.21	0.96 ± 0.31	1.19 ± 0.57	0.57 ± 0.39	x	
Firmicutes	6.08	5.09 ± 2.27	12.97 ± 5.95	6.56	19.2 ± 4.32	15.63 ± 6.39	5.77 ± 0.6	9.54 ± 1.9	13.05 ± 4.44	16.74 ± 5.44	x	
Fusobacteria	0.09	0.07 ± 0.03	0.11 ± 0.03	0.1	0.11 ± 0.01	0.09 ± 0.01	0.07 ± 0.01	0.11 ± 0.01	0.11 ± 0.03	0.15 ± 0.02	P<0.001, R^2=0.552	0.0025
Cyanobacteria	2.64	1.90 ± 0.83	1.89 ± 0.44	2.8	1.5 ± 0.06	1.42 ± 0.05	1.58 ± 0.07	2.08 ± 0.14	2.14 ± 0.16	2.31 ± 0.4	P=0.012, R^2=0.316	0.0304
Deferribacteres	0.09	0.07 ± 0.03	0.08 ± 0.02	0.1	0.05 ± 0	0.05 ± 0	0.06 ± 0.01	0.09 ± 0.01	0.09 ± 0.01	0.11 ± 0.01	P<0.0001, R^2=0.619	0.0024
Chrysiogenetes	0.03	0.02 ± 0.01	0.02 ± 0.01	0.03	0.02 ± 0	0.01 ± 0	0.02 ± 0	0.03 ± 0	0.03 ± 0	0.03 ± 0	P<0.01, R^2=0.391	0.0005
Acidobacteria	5.47	3.48 ± 1.37	3.22 ± 1.12	5.65	2.09 ± 0.23	1.99 ± 0.41	3.2 ± 0.22	4.21 ± 0.24	4.08 ± 0.39	2.92 ± 0.98	x	
Spirochaetes	0.27	0.21 ± 0.10	0.3 ± 0.06	0.3	0.26 ± 0.02	0.26 ± 0.01	0.23 ± 0.02	0.33 ± 0.03	0.31 ± 0.04	0.39 ± 0.04	P<0.0001, R^2=0.668	0.0060
Nitrospirae	0.66	0.69 ± 0.22	1.05 ± 0.93	0.73	0.3 ± 0.04	0.22 ± 0.04	0.66 ± 0.12	1.19 ± 0.1	1.41 ± 0.48	2.64 ± 1.19	P<0.0001, R^2=0.758	0.1001
Chlamydiae	0.12	0.10 ± 0.05	0.15 ± 0.08	0.14	0.09 ± 0.01	0.08 ± 0	0.09 ± 0.01	0.18 ± 0.01	0.21 ± 0.07	0.26 ± 0.04	P<0.0001, R^2=0.733	0.0080
Tenericutes	0.03	0.03 ± 0.01	0.04 ± 0.01	0.04	0.03 ± 0	0.03 ± 0	0.02 ± 0	0.04 ± 0	0.04 ± 0.01	0.05 ± 0.01	P<0.001, R^2=0.536	0.0011
Fibrobacteres	0.02	0.01 ± 0.01	0.02 ± 0	0.02	0.02 ± 0	0.02 ± 0	0.02 ± 0	0.02 ± 0	0.02 ± 0	0.02 ± 0	P<0.001, R^2=0.331	0.0003
Verrucomicrobia	2.97	1.69 ± 0.83	2.62 ± 0.85	3.38	1.71 ± 0.21	1.71 ± 0.6	2.77 ± 0.44	3.47 ± 0.42	3.08 ± 0.28	2.75 ± 1.18	x	

Chlorobi	0.53	0.38 ± 0.17	0.61 ± 0.11	0.58	0.53 ± 0.03	0.54 ± 0.03	0.49 ± 0.04	0.67 ± 0.06	0.66 ± 0.08	0.77 ± 0.03	P<0.0001, R^2=0.695	0.0109
Lentisphaerae	0.07	0.05 ± 0.02	0.07 ± 0.02	0.08	0.05 ± 0	0.05 ± 0.01	0.07 ± 0	0.09 ± 0.01	0.08 ± 0	0.08 ± 0.02	P=0.018, R^2=0.286	0.0012
Aquificae	0.22	0.18 ± 0.07	0.2 ± 0.06	0.22	0.14 ± 0	0.13 ± 0	0.16 ± 0.01	0.23 ± 0.02	0.24 ± 0.02	0.3 ± 0.02	P<0.0001, R^2=0.784	0.0068
Proteobacteria	45.18	36.21±13.78	37.27 ± 9.94	44.45	31.99 ± 5.87	31.43 ± 9.25	52.36 ± 1.62	39.06 ± 7.14	38.3 ± 7.56	28.12 ± 8.07	x	
Elusimicrobia	0.03	0.02 ± 0.01	0.03 ± 0.01	0.03	0.02 ± 0	0.02 ± 0	0.02 ± 0	0.03 ± 0	0.03 ± 0	0.04 ± 0.01	P<0.001, R^2=0.503	0.0007
Actinobacteria	18.08	19.33±8.49	6.1 ± 3.12	15.12	3.96 ± 1.01	5.77 ± 1.75	8.19 ± 3.25	6.03 ± 1.51	6.06 ± 2.4	3.6 ± 1.21	x	
Chloroflexi	4.56	3.56 ± 1.62	2.06 ± 0.85	4.52	1.41 ± 0.11	1.33 ± 0.15	1.65 ± 0.12	2.15 ± 0.15	2.38 ± 0.1	2.65 ± 1.03	x	
Planctomycetes	3.91	2.43 ± 1.20	2.72 ± 1.23	4.17	1.25 ± 0.18	1.27 ± 0.41	3.26 ± 0.08	3.93 ± 0.21	3.55 ± 0.72	2.57 ± 1.26	x	
Synergistetes	0.13	0.11 ± 0.05	0.08 ± 0.02	0.12	0.06 ± 0	0.05 ± 0	0.07 ± 0.01	0.09 ± 0.01	0.1 ± 0	0.1 ± 0.03	P=0.018, R^2=0.289	0.0016
Bacteroidetes	5.53	3.09 ± 1.45	26.05 ± 9.88	7.56	33.83 ± 4.53	36.47 ± 6.34	16.57 ± 3.99	23.81 ± 7.53	21.09 ± 6.75	30.72 ± 7.9	x	
Thermotogae	0.22	0.18 ± 0.08	0.16 ± 0.04	0.13	0.12 ± 0	0.12 ± 0	0.13 ± 0.02	0.17 ± 0.01	0.22 ± 0.04	0.19 ± 0.01	P<0.001, R^2=0.519	0.0037

3.1.2. Study Ⅰ. Time succession: Community composition in terms of functional genes and taxonomy.

To visualize both taxonomic family-level and functional metagenome profiles, I generated nonmetric multidimensional scaling (NMDS) plots for summarizing taxonomic and functional information. In terms of community taxonomic composition from the metagenome, the NMDS plot showed that replicates clearly clustered by time of incubation, in terms of taxonomic profiles (Fig. 4). In Fig. 4A, the plot was generated at family level based on M5NR taxonomic profile. Nondisturbed control samples, sieved but not partially sterilized and stored alongside the other pots at 25 ° C for the 24 week period, clustered nearby the original soil in both functional and taxonomic terms. Of the experimental treatments which had gone through the 90 % by volume sterilization, the earliest samples were most distinct from the original soil (and the controls). Over time, they moved back across the NMDS plot in the direction of the original/control soil, but after the 8 week stage recovery (as judged by the NMDS) the change seems to slow or stall.

Structure of the bacterial community showed a similar trend. At the genus level, calculated from 16S rRNA genes extracted from metagenomes, early stage (weeks 1-4) bacterial communities were

clustered furthest from original soil, and on the NMDS space shifted over time back in the direction of the original and control soils (Fig. 4C). However, unlike the total taxonomic composition, change in bacterial community on the NMDS did not stall at 12 weeks – although the rate of change appeared to slow down.

A very similar pattern also occurred for the functional gene composition of the soil biota, based on the metagenomes. In terms of functional level 3, based on the SEED functional profile, the earliest samples after the disturbance event likewise clustered furthest from original soil on the NMDS at first, then moved back in the direction of the original soil (Fig. 4B). By the 4-week stage, this movement in terms of NMDS slowed and stalled. Changes of functional genes by time thus showed a similar trend to the taxonomic community changes.

An Envfit was done for analysis for the most important environmental factors in influencing the NMDS using R. NMDS plot and Envfit of shotgun metagenomic sequences based on M5NR taxonomic profile at family level were the most strongly effected by TOC ($P = 0.036$, Fig. 5A). Also, community composition in terms of functional genes of all biota based on M5NR taxonomic profile (Fig. 5B) and communities from metagenome-extracted bacterial 16S

rRNA genes (Fig. 5C) were influenced by TOC ($P = 0.041$, $P = 0.0140$).

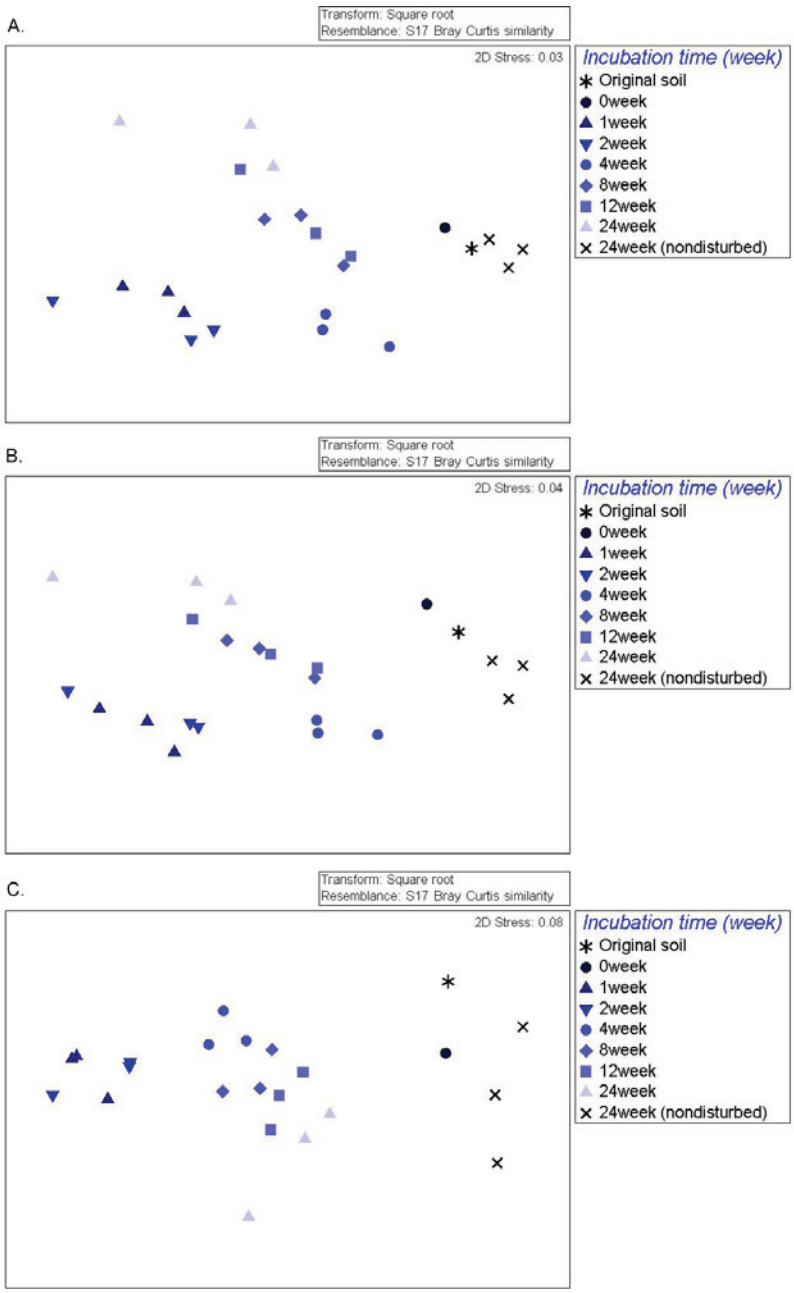


Figure 4. NMDS plot generated using weighted pairwise Unifrac distances between samples. A. NMDS plot of shotgun metagenomic sequences based on M5NR taxonomic profile at family level, B. NMDS plot of shotgun metagenomic sequences based on SEED functional profile at subsystem level 3, C. NMDS plot of 16S rRNA bacterial genus level which extracted from metagenome.

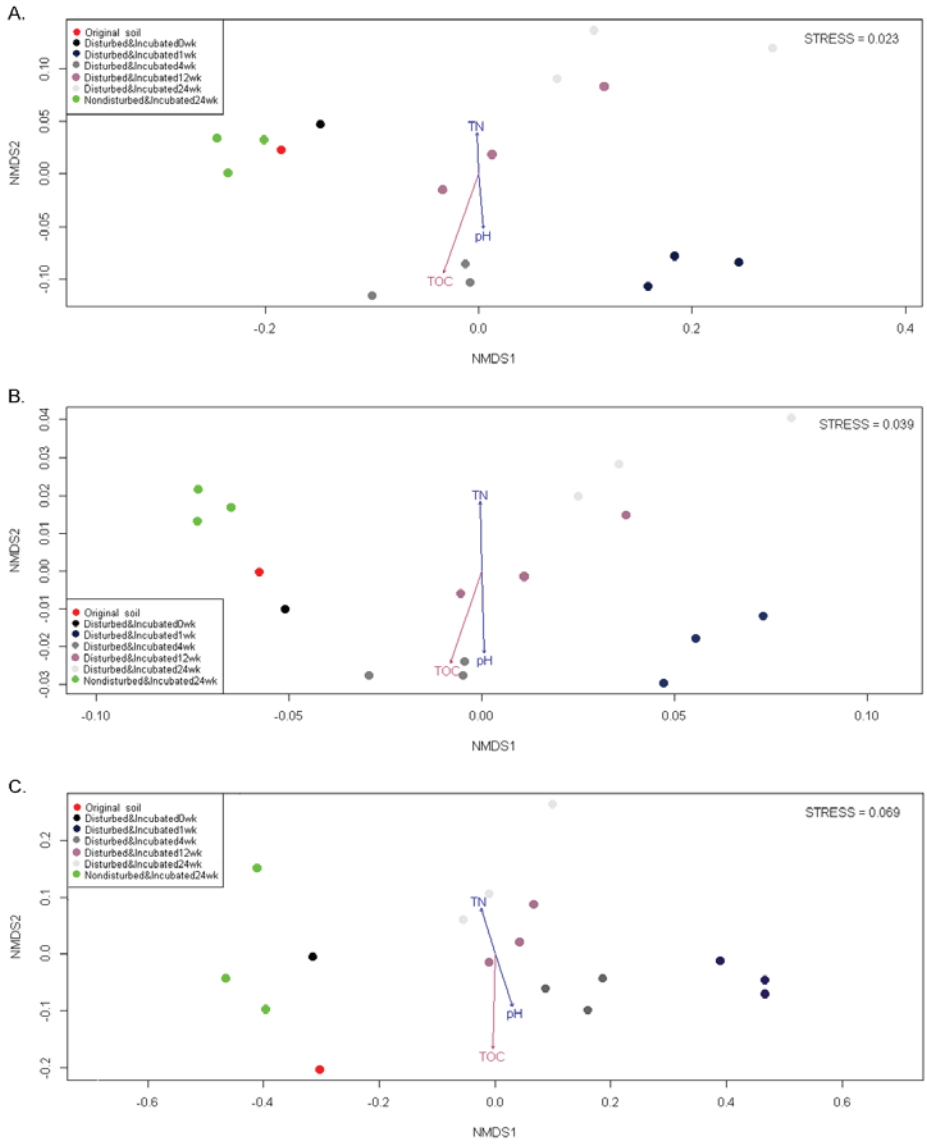


Figure 5. NMDS plot generated using weighted pairwise Unifrac distances between samples and Envfit with pH, TOC and TN. Important environmental factor marked red. A. NMDS plot and Envfit of shotgun metagenomic sequences based on M5NR taxonomic profile at family level, B. NMDS plot and Envfit of shotgun metagenomic sequences based on SEED functional profile at subsystem level 3, C. NMDS plot and Envfit of 16S rRNA bacterial genus level which extracted from metagenome.

3.1.3. Study Ⅰ. Time succession: Alpha diversity.

Diversity was calculated based on Shannon diversity index. Shannon diversity index was calculated based on M5NR taxonomy at species level as concept of diversity is based on species level and SEED subsystem function at level 1, 2 and 3 by using software R. In terms of alpha-diversity, the species level diversity (for taxonomically described published species) detected in the metagenomes was not affected by incubation time ($P > 0.05$, Fig. 6A). However, total functional gene Shannon diversity index, calculated for SEED subsystems level 1, 2 and 3, showed a significant declining trend of diversity over time ($P < 0.05$, Fig. 6B, Fig. 6C, Fig. 6D). Among 24-week incubated samples, non-disturbed samples tended to have a Shannon diversity index similar to the original soil, compared to disturbed & incubated samples in general (Fig. 6).

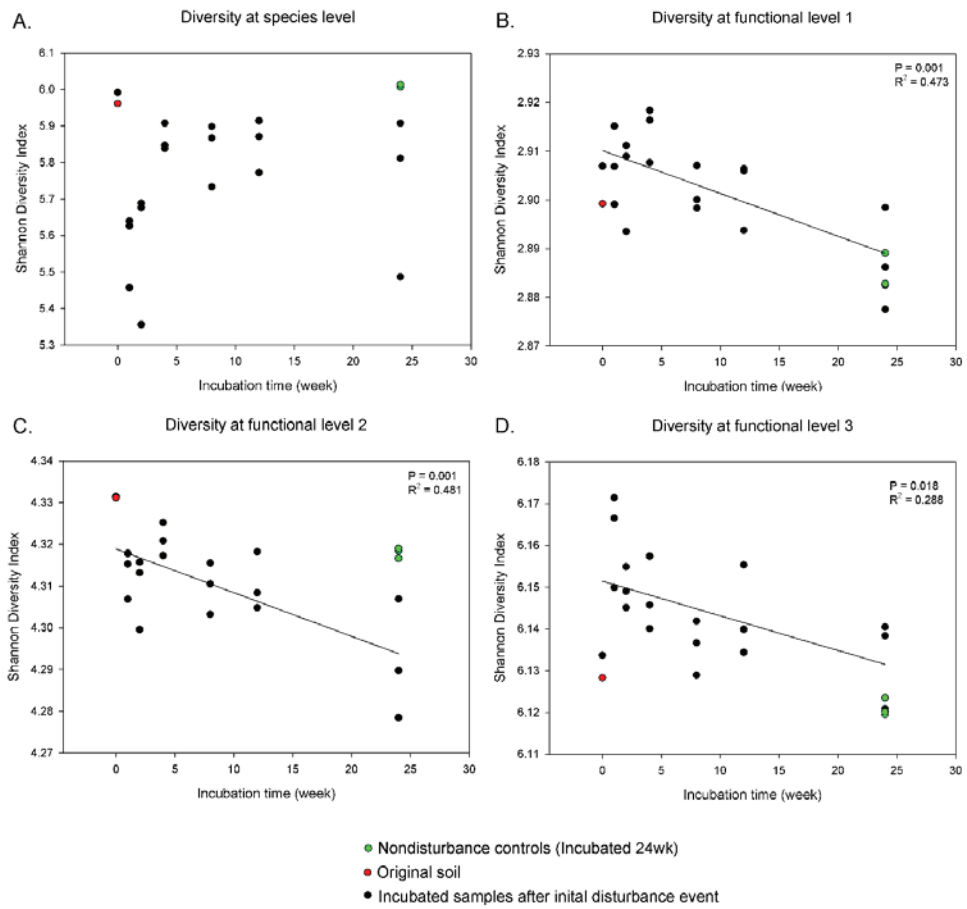


Figure 6. Taxonomic (M5NR database) and Functional diversity (SEED database) at soil microbial communities by incubation time based on Shannon diversity index. Solid line represents linear regression fits to data. A. Diversity at species level from all kingdoms, B. Diversity at functional level 1, C. Diversity at functional level 2, D. Diversity at functional level 3. Colored points were not included in the statistics.

3.1.4. Study Ⅰ. Time succession: Predominant changes in functional gene categories

To examine the effect of time succession after the disturbance event on soil microbial functions, the functional profile of shotgun metagenomic sequences was analysed using the SEED database, and the metagenomic sequences were distributed into 28 functional gene categories (level 1 SEED subsystems).

Of the 28 functional gene categories, 14 categories differed significantly ($P < 0.05$) in relation to time since disturbance (Fig. 7, Fig. 8). At functional level 1, 6 gene categories which represent genes related to protein metabolism ($P < 0.01$, Fig. 7A), DNA metabolism ($P < 0.001$, Fig. 7B), RNA metabolism ($P < 0.05$, Fig. 7C), miscellaneous ($P < 0.05$, Fig. 7D), clustering-based subsystems ($P < 0.001$, Fig. 7E) and cell division and cell cycle ($P < 0.01$, Fig. 7F) all showed an increase in relative abundance over time (Fig. 7). Functions which are related to protein, DNA and RNA metabolism are one of important functions of the 'housekeeping' gene category, which are constitutively expressed to maintain cellular function (Zhu et al., 2008; She et al., 2009; Eisenberg and Levanon, 2013). Thus, following disturbance, the relative abundance of housekeeping genes in this category showed an increasing trend.

More narrowly defined functions classified as level 2 under protein and RNA metabolism individually did not show any significant relationship with time since disturbance ($P > 0.05$) and showed this trend only when combined on level 1 (Table 4). However, 5 separate functions classified under DNA metabolism had increasing trend over time (Table 1). DNA repair, metabolism, replication, recombination and CRISPs had strong ($P < 0.001$) positive correlations with time. Likewise, in functional level 2, genes related to miscellaneous, cell division and cell cycle did not show any time-related trend, but 22 functions related to clustering-based subsystems showed significant changes over time. There were various functions which had positive changes over time such as shiga toxin cluster, two related proteases, putrescine/GABA utilization cluster-temporal, to add to SSs, ribosome-related cluster, D-tyrosyl-tRNA(Tyr) deacylase (EC 3.1.) cluster, carbohydrates, pyruvate kinase associated cluster, cytochrome biogenesis, fatty acid metabolic cluster, chemotaxis, response regulators, molybdopterin oxidoreductase, hypothetical associated with RecF, probably pyrimidine biosynthesis-related, TldD cluster and custering-based subsystems. However, there were negative relationships between time and abundance for recX and regulatory cluster, Lysine Biosynthesis, Tricarboxylate transporter, DNA

polymerase III epsilon cluster, 'probably GTP or GMP signaling related', Hypothetical protein possible functionally linked with Alanyl-tRNA synthetase and Sulfatases and sulfatase modifying factor 1.

At functional level 1 using the SEED database, 8 gene categories related to membrane transport ($P < 0.01$, Fig. 8A), motility and chemotaxis ($P < 0.05$, Fig. 8B), metabolism of aromatic compounds ($P < 0.001$, Fig. 8C), stress response ($P < 0.05$, Fig. 8D), fatty acids, lipids and Isoprenoids ($P < 0.01$, Fig. 8E), potassium metabolism ($P < 0.001$, Fig. 8F), sulfur metabolism ($P< 0.001$, Fig. 8G), nitrogen Metabolism ($P < 0.05$, Fig. 8H) showed declining relative abundance over time (Fig. 8). Detailed functions classified as level 2 under nitrogen metabolism did not had significant relationship with time ($P > 0.05$) and had trend only as combined on level 1 (Table 5). Under motility and chemotaxis, functional level 2 classification of social motility and nonflagellar swimming in bacteria and flagellar motility in prokaryota showed significantly ($P < 0.05$) decreasing trends after the initial disturbance event. Functional level 2 genes for heat shock responses, oxidative stress, acid stress, and stress response also showed decreasing patterns under the broader stress response label.

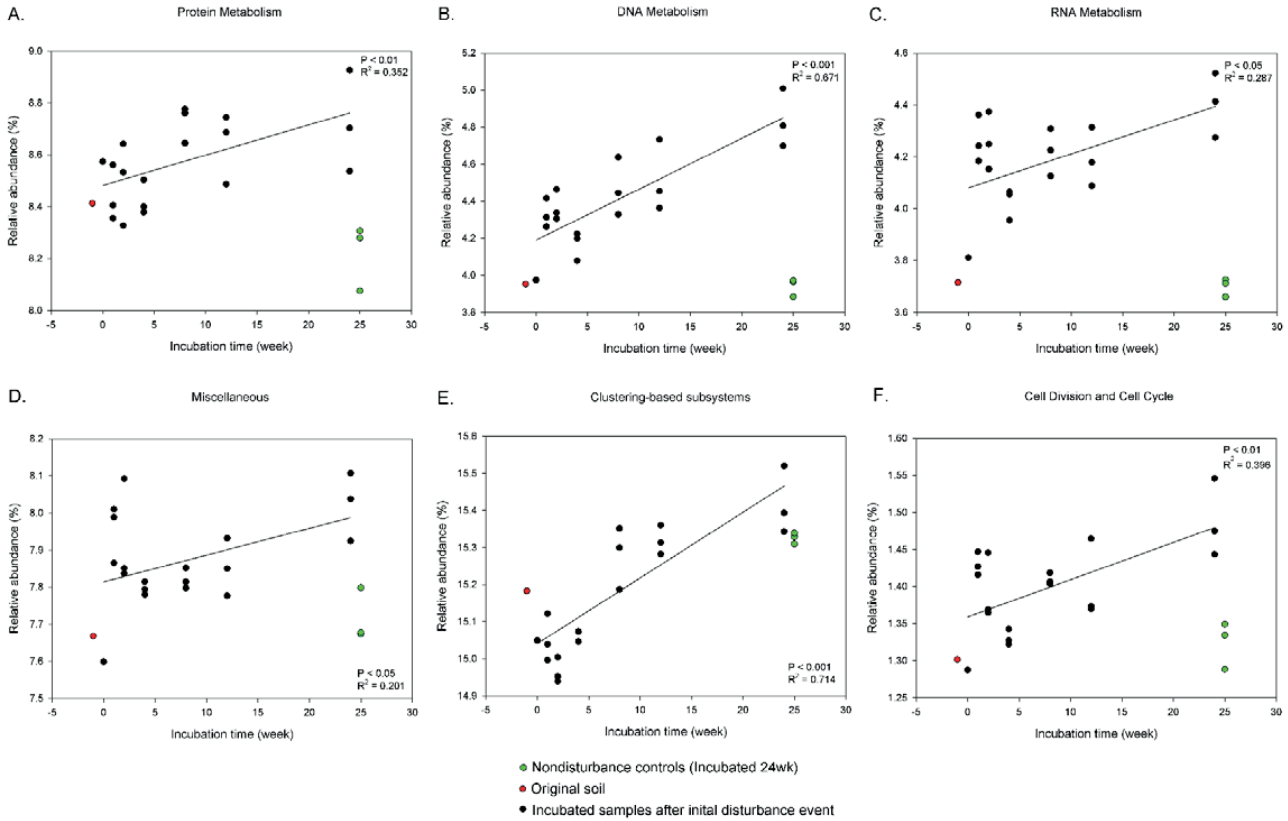


Figure 7. Relationship between incubation time and relative abundance of functional gene categories at subsystem level 1 (SEED database). Solid line represents linear regression fits to data. A. Protein metabolism, B. DNA metabolism, C. RNA metabolism, D. Miscellaneous, E. Clustering-based subsystems, F. Cell division and cell cycle. Original soils were marked red and non-disturbance control samples were marked green. Colored points were not included in the statistics.

Table 4. Functional abundance analysis of level 2 under functional gene level 1 (SEED database) which had positive correlation with incubation time. Only classification levels having $P < 0.05$ were shown.

Function classification	Function classification	Regression		
Level 1	Level 2	P	R^2	Coefficient
Protein Metabolism	-	-	-	-
DNA Metabolism	DNA repair	<0.001	0.571	0.00925
	DNA Metabolism	0.001	0.470	0.00596
	DNA replication	<0.001	0.687	0.00892
	DNA recombination	<0.001	0.519	0.00078
	CRISPs	<0.001	0.760	0.00126
RNA Metabolism	-	-	-	-
Miscellaneous	-	-	-	-
Clustering-based subsystems	recX and regulatory cluster	0.003	0.406	-0.00020
	Shiga toxin cluster	0.004	0.390	0.00062
	Two related proteases	0.040	0.226	0.00043
	Lysine Biosynthesis	0.042	0.221	-0.00229
	Tricarboxylate transporter	0.007	0.354	-0.00148
	Putrescine/GABA utilization cluster-temporal,to add to SSs	0.006	0.367	0.00270
	Ribosome-related cluster	0.001	0.464	0.00056
	D-tyrosyl-tRNA(Tyr) deacylase (EC 3.1.) cluster	0.007	0.357	0.00080
	Carbohydrates - #1	<0.001	0.753	0.00031
	DNA polymerase III epsilon cluster	0.011	0.322	-0.00033
	Pyruvate kinase associated cluster	0.004	0.396	0.00280
	Cytochrome biogenesis	0.046	0.215	0.00151
	Probably GTP or GMP signaling related	0.011	0.326	-0.00138
	Fatty acid metabolic cluster	0.040	0.225	0.00148
	Chemotaxis, response regulators	<0.001	0.661	0.00061
	Hypothetical protein possible functionally linked with Alanyl-tRNA synthetase	0.007	0.356	-0.00107
	Molybdopterin oxidoreductase	0.001	0.497	0.00028
	Sulfatases and sulfatase modifying factor 1 (and a hypothetical)	0.001	0.460	-0.00143
	Hypothetical associated with RecF	0.038	0.229	0.00026
	Probably Pyrimidine biosynthesis-related	0.001	0.494	0.00052
	TldD cluster	0.004	0.388	0.01092
	Clustering-based subsystems	0.006	0.365	0.00399
Cell Division and Cell Cycle	Cell Division and Cell Cycle	0.006	0.365	0.00399

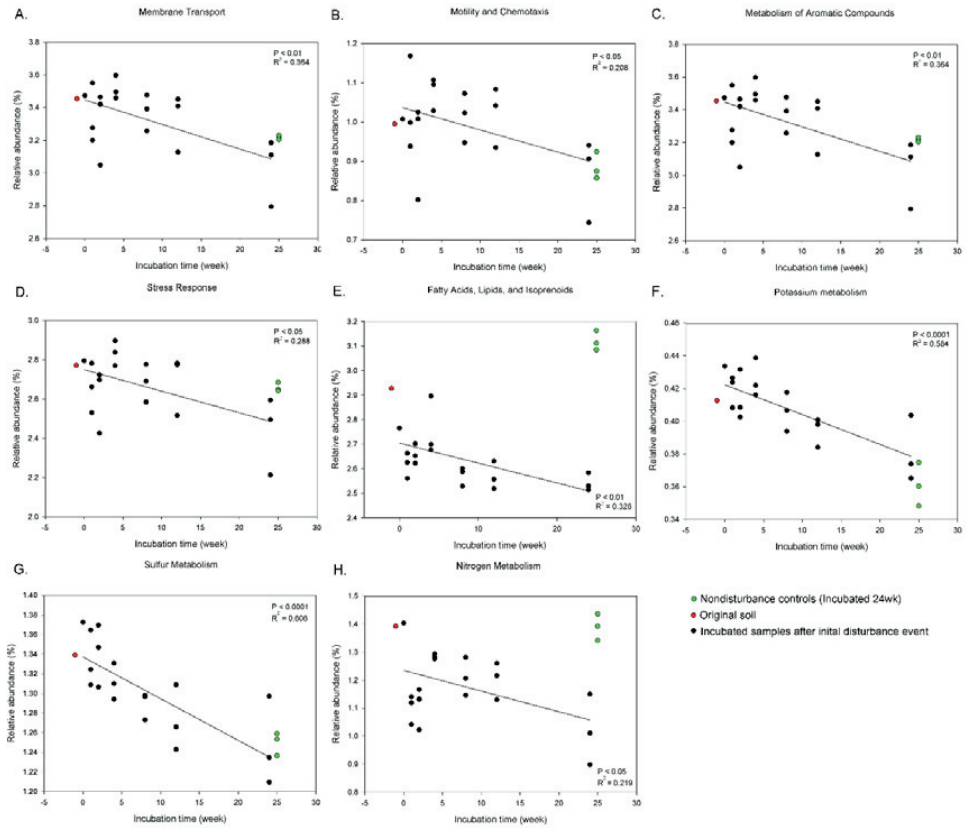


Figure 8. Relationship between incubation time and relative abundance of functional gene categories at subsystem level 1 (SEED database). Solid line represents linear regression fits to data. A. Membrane transport, B. Motility and Chemotaxis, C. Metabolism of Aromatic Compounds, D. Stress Response, E. Fatty Acids, Lipids and Isoprenoids, F. Potassium metabolism, G. Sulfur Metabolism, H. Nitrogen Metabolism. Colored points were not included in the statistics.

Table 5. Functional abundance analysis of level 2 under functional gene level 1(SEED database) which had negative correlation with incubation time. Only classification levels having $P < 0.05$ were shown.

Function classification	Function classification	Regression		
Level 1	Level 2	P	R^2	Coefficient
Membrane transport	Protein translocation across cytoplasmic membrane	0.007	0.352	0.00182
	Protein secretion system, Type V	<0.001	0.569	-0.00016
	Membrane Transport	0.011	0.321	-0.00703
	Protein secretion system, Type III	0.016	0.297	-0.00022
	Protein secretion system, Type VII (Chaperone/Usher pathway, CU)	0.005	0.384	-0.00011
Motility and Chemotaxis	Social motility and nonflagellar swimming in bacteria	0.008	0.351	-0.00004
	Flagellar motility in Prokaryota	0.024	0.264	-0.00469
Metabolism of Aromatic Compounds	Metabolism of central aromatic intermediates	<0.001	0.775	-0.00624
	Peripheral pathways for catabolism of aromatic compounds	0.002	0.437	-0.00768
	Metabolism of Aromatic Compounds	0.001	0.482	-0.00272
Stress Response	Heat shock	0.020	0.278	0.00166
	Oxidative stress - #1	0.006	0.364	-0.00620
	Acid stress	0.034	0.238	-0.00062
	Stress Response	<0.001	0.661	-0.00279
Fatty Acids, Lipids and Isoprenoids	Phospholipids	0.010	0.333	-0.00197
	Fatty Acids, Lipids, and Isoprenoids	0.014	0.307	-0.00303
	Fatty acids	0.015	0.299	-0.00554
	Triacylglycerols	0.010	0.331	-0.00012
	Isoprenoids	0.049	0.209	0.00249
Potassium metabolism	Potassium metabolism	<0.001	0.588	-0.00173
Sulfur Metabolism	Organic sulfur assimilation	0.003	0.413	-0.00313
Nitrogen Metabolism	-	-	-	-

3.2. Study Ⅱ. Disturbance frequency: Effect of disturbance frequency on structure and function of soil communities.

From the 31 separate pot soil DNA samples, including 1 original soil samples, 25 disturbed samples, and 5 non-disturbance incubation samples, sequence count 101,216,919 of sequence data were generated. Phylogenetic information was extracted from the metagenomes using M5NR data bases using BLASTX and functional profiles were generated with the SEED Subsystems database. Information of samples and results of soil chemical analysis are shown in Table 1 and Table 2. A total of 33-52 % of the sequences were annotated as protein using $E < 1 \times 10^{5}$ and 15-bp minimum alignment length on MG-RAST server.

3.2.1. Study Ⅱ. Disturbance frequency: Dominant microbial taxa.

The majority of the metagenomic sequences among samples which disturbed with different frequencies were dominated by Bacteria (86.6 % on average) followed by Eukaryota (1.1 %), Archaea (0.5 %) and Viruses (0.4 %) according to the M5NR database.11.8 % of reads were unassigned. Bacteria and Viruses had sigmoidal correlation with disturbance frequency ($P < 0.01$, Fig.

9A) ($P < 0.0001$, Fig. 9D) while Eukaryota and Archaea declined continuously ($P < 0.0001$, Fig. 9B) ($P < 0.0001$, Fig. 9C). In case of original soil which was 100 % unsterilized sieved soil, the total community based on taxonomically assigned reads consisted of Bacteria 83.9 %, Eukaryota 1.4 %, Archaea 1 % and Viruses 0.01 % on average (Fig. 9, unassigned 13.7 %).

Abundant bacterial phyla from original soil which was 100% unsterilized sieved soil, the total community consisted of Proteobacteria 45.18 %, Actinobacteria 18.08 %, Firmicutes 6.08 % , Bacteroidetes 5.53 %, Acidobacteria 5.47 %, Chloroflexi 4.56 %, Planctomycetes 3.91 %, Verrucomicrobia 2.97 %, Cyanobacteria 2.64 %, and Gemmatimonadetes 1.24 % (Table 6). The most abundant bacterial phylum amoung the metagenomic sequence data undergo disturbance treatments for 24 weeks was Proteobacteria (54.29 % on average), followed by Actinobacteria (15.30 %), Bacteroidetes (10.08 %), Firmicutes (4.85 %), Planctomycetes (2.49 %), Acidobacteria (2.10 %), Verrucomicrobia (2.03 %), Gemmatimonadetes (1.68 %), Chloroflexi (1.67 %) and Cyanobacteria (1.31 %). Among 28 phyla, 23 of them had significant ($P < 0.05$) relationship with disturbance frequency, by showing strong impact of disturbance to soil microbial communities. In general, regression analysis showed most of bacteria phyla had

negative coefficient but relative abundance of bacteria did not changed because of Proteobacteria and Bacteroidetes increased.

Relation between disturbance frequency and relative abundance of Eukaryota (1.1 % on average at total community) and Archaea (0.5 % on average) phyla were calculated (Table 7, Table 8). Among 52 phyla under Eukaryota, 22 phyla had significant ($P < 0.05$) relation with disturbance frequency (Table 7). Abundant fungal phyla Basidiomycota and Ascomycota were both decreased in increasing disturbance event. Among 6 phyla of Archaea, 5 phyla showed significant relation ($P < 0.02$, Table 8).

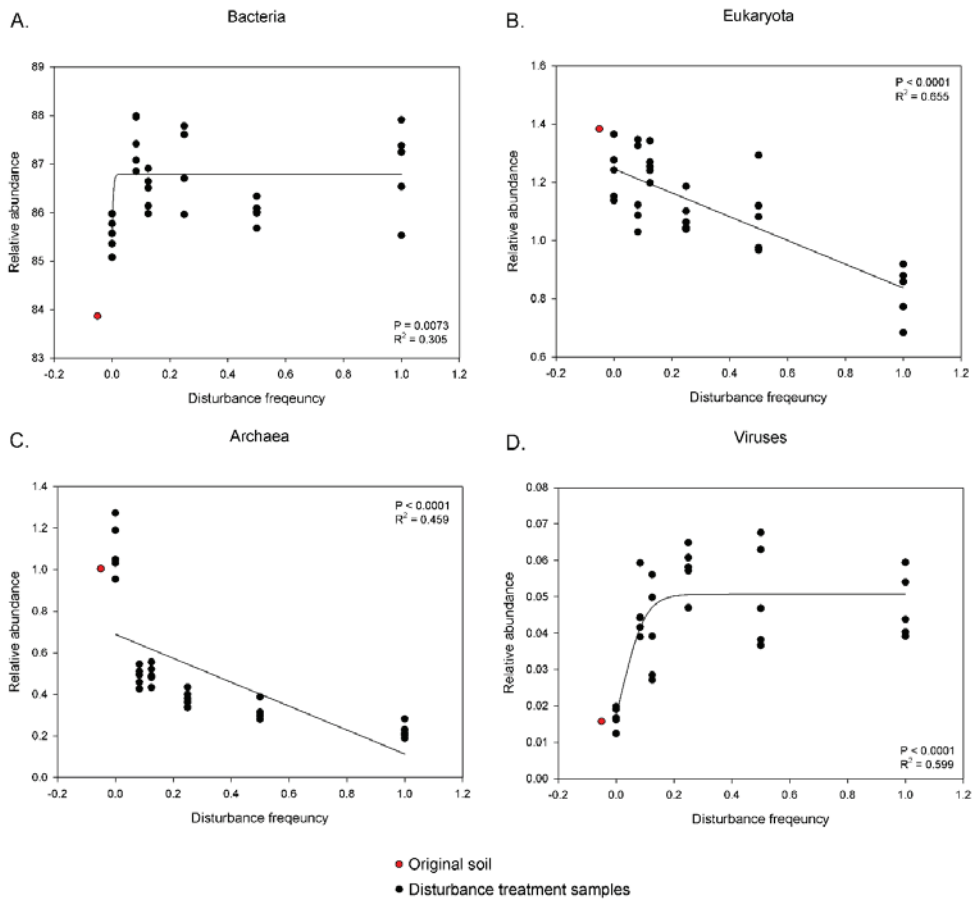


Figure 9. Relative abundance of each domain level assigned based on the M5NR database. A. Relative abundance of Bacteria, B. Relative abundance of Eukaryota, C. Relative abundance of Archaea, D. Relative abundance of Viruses. Colored point was not included in the statistics.

Table 6. Relative abundance (%) (means ± SD) of bacterial phyla observed in shotgun metagenomics sequences.

Phylum	Original soil	Disturbance freqeuncy							Regression		
		0	0.0833	0.125	0.25	0.5	1	Average	P	R^2	Coefficient
Poribacteria	0.07	0.08 ± 0.02	0.07 ± 0.03	0.07 ± 0.01	0.01 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	0.04 ± 0.04	0.000	0.550	-0.07713
Dictyoglomi	0.06	0.06 ± 0.02	0.03 ± 0.00	0.03 ± 0.00	0.03 ± 0.00	0.02 ± 0.00	0.02 ± 0.00	0.03 ± 0.01	0.000	0.465	-0.02901
Deinococcus-Thermus	0.82	0.79 ± 0.23	0.58 ± 0.08	0.59 ± 0.09	0.52 ± 0.06	0.41 ± 0.04	0.29 ± 0.03	0.53 ± 0.19	0.000	0.585	-0.41590
Gemmatimonadetes	1.24	1.09 ± 0.48	2.17 ± 0.70	2.30 ± 0.61	1.87 ± 0.61	1.66 ± 0.27	0.95 ± 0.47	1.68 ± 0.71	0.033	0.152	-0.80140
Firmicutes	6.08	5.37 ± 1.67	5.17 ± 0.53	4.84 ± 0.46	5.11 ± 0.40	3.94 ± 0.29	4.65 ± 0.77	4.85 ± 0.89	-		
Fusobacteria	0.09	0.07 ± 0.02	0.06 ± 0.00	0.07 ± 0.01	0.06 ± 0.00	0.06 ± 0.01	0.05 ± 0.00	0.06 ± 0.01	0.005	0.245	-0.01848
Cyanobacteria	2.64	2.04 ± 0.62	1.25 ± 0.15	1.43 ± 0.23	1.25 ± 0.16	1.08 ± 0.08	0.80 ± 0.08	1.31 ± 0.47	0.000	0.447	-0.90229
Deferribacteres	0.09	0.07 ± 0.02	0.05 ± 0.01	0.06 ± 0.01	0.05 ± 0.01	0.04 ± 0.00	0.03 ± 0.00	0.05 ± 0.02	0.000	0.521	-0.03460
Chrysiogenetes	0.03	0.02 ± 0.01	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.00	0.01 ± 0.00	0.02 ± 0.00	0.001	0.327	-0.00817
Acidobacteria	5.47	3.64 ± 1.03	2.10 ± 0.24	2.38 ± 0.51	1.93 ± 0.41	1.55 ± 0.11	1.04 ± 0.19	2.10 ± 0.94	0.001	0.327	-0.00817
Spirochaetes	0.27	0.22 ± 0.07	0.16 ± 0.02	0.21 ± 0.03	0.20 ± 0.03	0.20 ± 0.01	0.15 ± 0.02	0.19 ± 0.04	-		
Nitrospirae	0.66	0.72 ± 0.16	0.62 ± 0.23	0.60 ± 0.12	0.14 ± 0.03	0.10 ± 0.01	0.07 ± 0.01	0.37 ± 0.30	0.000	0.562	-0.65543
Chlamydiae	0.12	0.09 ± 0.03	0.08 ± 0.01	0.10 ± 0.01	0.08 ± 0.01	0.09 ± 0.03	0.06 ± 0.01	0.09 ± 0.02	0.021	0.176	-0.02496

Tenericutes	0.03	0.03 ± 0.01	0.02 ± 0.00	0.03 ± 0.00	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.01	0.004	0.262	-0.00814
Fibrobacteres	0.02	0.01 ± 0.00	0.01 ± 0.00	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	0.012	0.205	-0.00538
Verrucomicrobia	2.97	1.74 ± 0.59	1.34 ± 0.19	2.27 ± 0.59	2.51 ± 0.41	2.76 ± 0.41	1.56 ± 0.58	2.03 ± 0.68	-		
Chlorobi	0.53	0.41 ± 0.12	0.36 ± 0.04	0.43 ± 0.07	0.39 ± 0.07	0.40 ± 0.02	0.30 ± 0.04	0.38 ± 0.08	0.021	0.177	-0.09196
Lentisphaerae	0.07	0.05 ± 0.02	0.04 ± 0.01	0.06 ± 0.01	0.06 ± 0.01	0.05 ± 0.00	0.04 ± 0.01	0.05 ± 0.01	-		
Aquificae	0.22	0.19 ± 0.05	0.14 ± 0.01	0.15 ± 0.02	0.13 ± 0.03	0.11 ± 0.01	0.08 ± 0.01	0.13 ± 0.04	0.000	0.578	-0.09356
Proteobacteria	45.18	38.57 ± 10.32	53.70 ± 4.93	57.41 ± 5.93	53.65 ± 2.89	59.25 ± 0.82	63.13 ± 2.50	54.29 ± 9.33	0.000	0.377	16.54700
Elusimicrobia	0.03	0.02 ± 0.01	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.01	0.01 ± 0.00	0.01 ± 0.00	0.02 ± 0.01	0.000	0.447	-0.01133
Actinobacteria	18.08	22.68 ± 7.60	20.57 ± 3.38	14.07 ± 3.98	16.46 ± 5.61	8.28 ± 1.45	9.75 ± 6.54	15.30 ± 7.12	0.001	0.352	-12.19400
Chloroflexi	4.56	3.99 ± 1.37	1.82 ± 0.28	1.61 ± 0.27	1.27 ± 0.10	0.77 ± 0.06	0.53 ± 0.07	1.67 ± 1.27	0.000	0.431	-2.40380
Planctomycetes	3.91	2.50 ± 0.86	2.44 ± 0.51	2.81 ± 0.36	2.89 ± 0.23	2.23 ± 0.07	2.09 ± 0.48	2.49 ± 0.53	0.041	0.140	-0.57290
Synergistetes	0.13	0.12 ± 0.04	0.08 ± 0.01	0.08 ± 0.01	0.07 ± 0.01	0.06 ± 0.01	0.04 ± 0.01	0.07 ± 0.03	0.000	0.520	-0.05991
unclassified (derived from Bacteria)	0.89	0.78 ± 0.23	0.59 ± 0.05	0.63 ± 0.07	0.52 ± 0.04	0.44 ± 0.02	0.59 ± 0.21	0.59 ± 0.16	-		
Bacteroidetes	5.53	3.30 ± 1.10	6.41 ± 0.76	7.59 ± 1.14	10.61 ± 1.52	16.31 ± 1.49	16.27 ± 1.59	10.08 ± 5.11	0.000	0.739	12.67650
Thermotogae	0.22	0.20 ± 0.06	0.12 ± 0.01	0.13 ± 0.02	0.11 ± 0.01	0.09 ± 0.01	0.07 ± 0.01	0.12 ± 0.05	0.000	0.470	-0.09492

Table 7. Relative abundance (%) (means ± SD) of Eukaryota phyla observed in shotgun metagenomics sequences.

		Disturbance freqeuncy							Regression		
Phylum	**Original soil**	**0**	**0.0833**	**0.125**	**0.25**	**0.5**	**1**	**Average**	**P**	**R2**	**Coefficient**
Entoprocta	0.00005	0.00000 ± 0.00000	0.00003 ± 0.00004	0.00001 ± 0.00002	0.00001 ± 0.00003	0.00004 ± 0.00003	0.00001 ± 0.00002	0.00002 ± 0.00003	-		
Echiura	0.00000	0.00000 ± 0.00000	0.00005 ± 0.00008	0.00007 ± 0.00004	0.00006 ± 0.00007	0.00002 ± 0.00003	0.00002 ± 0.00002	0.00004 ± 0.00005	-		
Glomeromycota	0.00025	0.00022 ± 0.00007	0.00017 ± 0.00012	0.00023 ± 0.00023	0.00023 ± 0.00017	0.00032 ± 0.00009	0.00017 ± 0.00007	0.00022 ± 0.00014	-		
Porifera	0.00065	0.00060 ± 0.00023	0.00068 ± 0.00038	0.00073 ± 0.00021	0.00081 ± 0.00040	0.00058 ± 0.00018	0.00048 ± 0.00010	0.00065 ± 0.00027	-		
Bryozoa	0.00000	0.00005 ± 0.00007	0.00002 ± 0.00004	0.00001 ± 0.00003	0.00000 ± 0.00000	0.00001 ± 0.00002	0.00003 ± 0.00002	0.00002 ± 0.00004	-		
Acanthocephala	0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00006 ± 0.00006	0.00001 ± 0.00002	0.00000 ± 0.00000	0.00001 ± 0.00001	0.00001 ± 0.00003	-		
Phaeophyceae	0.00920	0.00949 ± 0.00099	0.00677 ± 0.00164	0.00823 ± 0.00074	0.01120 ± 0.00293	0.00877 ± 0.00259	0.00704 ± 0.00231	0.00858 ± 0.00239	-		
unclassified (derived from Fungi)	0.00204	0.00250 ± 0.00105	0.00198 ± 0.00065	0.00250 ± 0.00145	0.00214 ± 0.00045	0.00191 ± 0.00103	0.00125 ± 0.00043	0.00205 ± 0.00094	0.022	0.174	-0.00113
Platyhelminthes	0.00303	0.00310 ± 0.00075	0.00259 ± 0.00072	0.00316 ± 0.00051	0.00273 ± 0.00022	0.00284 ± 0.00015	0.00185 ± 0.00036	0.00271 ± 0.00064	0.001	0.342	-0.00108
Placozoa	0.00597	0.00549 ± 0.00029	0.00455 ± 0.00036	0.00504 ± 0.00057	0.00548 ± 0.00065	0.00476 ± 0.00049	0.00328 ± 0.00050	0.00477 ± 0.00088	0.000	0.532	-0.00186
Blastocladiomycota	0.00040	0.00043 ± 0.00021	0.00034 ± 0.00010	0.00039 ± 0.00016	0.00043 ± 0.00011	0.00024 ± 0.00012	0.00016 ± 0.00007	0.00033 ± 0.00016	0.001	0.326	-0.00027
Chlorophyta	0.06322	0.05698 ± 0.00206	0.04450 ± 0.00273	0.04900 ± 0.00116	0.05060 ± 0.00556	0.04570 ± 0.00395	0.03440 ± 0.00356	0.04690 ± 0.00766	0.000	0.610	-0.01728
Cnidaria	0.02517	0.02062 ± 0.00286	0.02420 ± 0.00200	0.03270 ± 0.00160	0.03770 ± 0.00644	0.04910 ± 0.00317	0.03910 ± 0.00770	0.03390 ± 0.01050	0.000	0.366	0.01839
Onychophora	0.00005	0.00002 ± 0.00005	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00001 ± 0.00002	0.00002 ± 0.00003	0.00001 ± 0.00001	0.00001 ± 0.00003	-		

Apicomplexa	0.01398	0.01400 ± 0.00163	0.01500 ± 0.00139	0.01580 ± 0.00096	0.01640 ± 0.00129	0.01400 ± 0.00104	0.00952 ± 0.00182	0.01410 ± 0.00260	0.000	0.526	-0.00545
Hemichordata	0.00378	0.00349 ± 0.00071	0.00294 ± 0.00053	0.00329 ± 0.00055	0.00297 ± 0.00024	0.00298 ± 0.00047	0.00181 ± 0.00023	0.00291 ± 0.00070	0.000	0.515	-0.00145
Echinodermata	0.00716	0.00677 ± 0.00064	0.00489 ± 0.00034	0.00634 ± 0.00070	0.00550 ± 0.00087	0.00629 ± 0.00071	0.00452 ± 0.00054	0.00572 ± 0.00102	0.009	0.218	-0.00137
Pinguiophyceae	0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00001 ± 0.00003	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00001 ± 0.00002	0.00000 ± 0.00001	-		
Eustigmatophyceae	0.00000	0.00005 ± 0.00007	0.00009 ± 0.00009	0.00004 ± 0.00005	0.00000 ± 0.00000	0.00002 ± 0.00003	0.00000 ± 0.00000	0.00003 ± 0.00006	-		
Nemertea	0.00000	0.00002 ± 0.00004	0.00004 ± 0.00007	0.00004 ± 0.00004	0.00002 ± 0.00003	0.00001 ± 0.00002	0.00002 ± 0.00003	0.00003 ± 0.00004	-		
Euglenida	0.00075	0.00061 ± 0.00019	0.00048 ± 0.00019	0.00088 ± 0.00006	0.00079 ± 0.00024	0.00058 ± 0.00007	0.00036 ± 0.00005	0.00062 ± 0.00023	0.010	0.216	-0.00030
Basidiomycota	0.04621	0.03977 ± 0.00414	0.03120 ± 0.00299	0.03400 ± 0.00246	0.03350 ± 0.00314	0.03180 ± 0.00233	0.02170 ± 0.00202	0.03200 ± 0.00607	0.000	0.647	-0.01408
Xanthophyceae	0.00234	0.00187 ± 0.00045	0.00022 ± 0.00015	0.00011 ± 0.00009	0.00013 ± 0.00011	0.00011 ± 0.00007	0.00006 ± 0.00005	0.00042 ± 0.00069	0.011	0.211	-0.00091
Aurearenophyceae	0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	-		
Arthropoda	0.07576	0.08684 ± 0.02166	0.06320 ± 0.00482	0.06880 ± 0.00421	0.06330 ± 0.00777	0.06610 ± 0.00559	0.04360 ± 0.00648	0.06530 ± 0.01590	0.000	0.454	-0.03098
Nematoda	0.02850	0.02311 ± 0.00318	0.02590 ± 0.01013	0.02550 ± 0.00319	0.02510 ± 0.00429	0.06360 ± 0.06095	0.03360 ± 0.01269	0.03280 ± 0.02760	-		
Xenoturbellida	0.00005	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00004 ± 0.00010	0.00001 ± 0.00003	0.00000 ± 0.00000	0.00001 ± 0.00001	0.00001 ± 0.00004	-		
Chaetognatha	0.00000	0.00005 ± 0.00007	0.00000 ± 0.00000	0.00003 ± 0.00004	0.00000 ± 0.00000	0.00001 ± 0.00003	0.00000 ± 0.00000	0.00002 ± 0.00004	-		
Brachiopoda	0.00015	0.00004 ± 0.00006	0.00005 ± 0.00006	0.00016 ± 0.00006	0.00009 ± 0.00009	0.00004 ± 0.00003	0.00004 ± 0.00003	0.00007 ± 0.00007	-		
Neocallimastigomycota	0.00040	0.00026 ± 0.00011	0.00010 ± 0.00006	0.00019 ± 0.00011	0.00016 ± 0.00008	0.00013 ± 0.00006	0.00018 ± 0.00008	0.00017 ± 0.00009	-		
Streptophyta	0.20399	0.17817 ± 0.00674	0.18400 ± 0.00711	0.19800 ± 0.01248	0.21900 ± 0.01021	0.22300 ± 0.02314	0.21100 ± 0.02816	0.20200 ± 0.02300	0.010	0.213	0.03063
Chytridiomycota	0.00035	0.00023 ±	0.00111 ±	0.00108 ±	0.00195 ±	0.00228 ±	0.00169 ±	0.00139 ±	-		

		0.00014	0.00072	0.00080	0.00110	0.00079	0.00183	0.00116			
Rhombozoa	0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00003 ± 0.00007	0.00001 ± 0.00003	0.00000 ± 0.00001	0.00001 ± 0.00001	0.00001 ± 0.00003	-		
Mollusca	0.00085	0.00152 ± 0.00106	0.00098 ± 0.00024	0.00134 ± 0.00043	0.00107 ± 0.00034	0.00102 ± 0.00011	0.00062 ± 0.00016	0.00109 ± 0.00054	0.012	0.206	-0.00071
Ctenophora	0.00000	0.00005 ± 0.00006	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00001 ± 0.00003	-		
Annelida	0.00020	0.00044 ± 0.00030	0.00029 ± 0.00015	0.00014 ± 0.00006	0.00023 ± 0.00013	0.00022 ± 0.00003	0.00014 ± 0.00006	0.00024 ± 0.00017	0.040	0.143	-0.00019
Kinorhyncha	0.00000	0.00000 ± 0.00000	0.00010 ± 0.00017	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00002 ± 0.00007	-		
Nematomorpha	0.00000	0.00000 ± 0.00000	0.00002 ± 0.00005	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00001	0.00000 ± 0.00002	-		
Tardigrada	0.00005	0.00007 ± 0.00007	0.00007 ± 0.00008	0.00001 ± 0.00003	0.00003 ± 0.00003	0.00002 ± 0.00002	0.00004 ± 0.00005	0.00004 ± 0.00005	-		
Priapulida	0.00000	0.00003 ± 0.00006	0.00001 ± 0.00003	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00001 ± 0.00001	0.00001 ± 0.00003	-		
Ascomycota	0.51238	0.43766 ± 0.05858	0.45400 ± 0.13882	0.43000 ± 0.08372	0.25500 ± 0.03622	0.22300 ± 0.06383	0.16100 ± 0.03791	0.32700 ± 0.13900	0.000	0.579	-0.30485
Bacillariophyta	0.03641	0.02178 ± 0.00197	0.01160 ± 0.00160	0.01290 ± 0.00150	0.01530 ± 0.00224	0.01300 ± 0.00263	0.00991 ± 0.00178	0.01410 ± 0.00427	0.001	0.308	-0.00685
Cryptomycota	0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00001 ± 0.00002	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00001	-		
Rotifera	0.00134	0.00065 ± 0.00063	0.00080 ± 0.00033	0.00091 ± 0.00033	0.00063 ± 0.00010	0.00065 ± 0.00017	0.00080 ± 0.00021	0.00074 ± 0.00033	-		
unclassified (derived from Metazoa)	0.00000	0.00000 ± 0.00000	0.00001 ± 0.00002	0.00002 ± 0.00004	0.00006 ± 0.00001	0.00003 ± 0.00004	0.00002 ± 0.00002	0.00002 ± 0.00003	-		
Chromerida	0.00005	0.00000 ± 0.00000	0.00002 ± 0.00005	0.00000 ± 0.00000	0.00005 ± 0.00005	0.00004 ± 0.00004	0.00002 ± 0.00002	0.00002 ± 0.00003	-		
Chordata	0.19096	0.18698 ± 0.01557	0.15500 ± 0.00612	0.20200 ± 0.05683	0.16900 ± 0.01914	0.18100 ± 0.01277	0.12700 ± 0.01339	0.17000 ± 0.03450	0.004	0.266	-0.05136
unclassified (derived from Eukaryota)	0.14614	0.12849 ± 0.00745	0.14300 ± 0.00841	0.15600 ± 0.00988	0.16300 ± 0.01432	0.14200 ± 0.00774	0.10700 ± 0.01921	0.14000 ± 0.02160	0.002	0.300	-0.03421

Myzostomida	0.00000	0.00000 ± 0.00000	0.00001 ± 0.00003	0.00000 ± 0.00000	0.00001 ± 0.00003	0.00001 ± 0.00002	0.00000 ± 0.00001	0.00001 ± 0.00002	-		
Haplosporidia	0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00000	0.00000 ± 0.00001	0.00000 ± 0.00000	0.046	0.135	0.00000
Sipuncula	0.00000	0.00007 ± 0.00011	0.00002 ± 0.00003	0.00001 ± 0.00002	0.00003 ± 0.00003	0.00002 ± 0.00002	0.00001 ± 0.00002	0.00003 ± 0.00005	-		
Microsporidia	0.00144	0.00195 ± 0.00036	0.00146 ± 0.00030	0.00164 ± 0.00033	0.00147 ± 0.00028	0.00141 ± 0.00026	0.00110 ± 0.00019	0.00151 ± 0.00037	0.000	0.363	-0.00065

Table 8. Relative abundance (%) (means ± SD) of Archaea phyla observed in shotgun metagenomics sequences.

Phylum	Original soil	Disturbance freqeuncy							Regression		
		0	0.0833	0.125	0.25	0.5	1	Average	P	R^2	Coefficient
Thaumarchaeota	0.16808	0.19872 ± 0.05639	0.04566 ± 0.01106	0.03052 ± 0.0032	0.00463 ± 0.0027	0.00805 ± 0.00795	0.00306 ± 0.00069	0.04844 ± 0.07338	0.002	0.287	-0.11352
Crenarchaeota	0.14330	0.17471 ± 0.02897	0.06632 ± 0.00383	0.06411 ± 0.00586	0.0349 ± 0.01985	0.05898 ± 0.05914	0.0247 ± 0.00394	0.07062 ± 0.05609	0.002	0.301	-0.08886
Korarchaeota	0.00751	0.00721 ± 0.00126	0.00295 ± 0.00044	0.00316 ± 0.00046	0.00201 ± 0.00119	0.00333 ± 0.00271	0.00141 ± 0.00048	0.00334 ± 0.00225	0.004	0.265	-0.00334
Euryarchaeota	0.66817	0.70304 ± 0.05076	0.36409 ± 0.03412	0.38954 ± 0.04223	0.249 ± 0.14068	0.44736 ± 0.45843	0.19095 ± 0.02972	0.39074 ± 0.2458	0.020	0.179	-0.30060
unclassified (derived from Archaea)	0.01587	0.01433 ± 0.00202	0.00777 ± 0.00172	0.0084 ± 0.00105	0.00502 ± 0.00283	0.00909 ± 0.00823	0.00362 ± 0.00115	0.00804 ± 0.00487	0.008	0.223	-0.00664
Nanoarchaeota	0.00070	0.00091 ± 0.00059	0.00035 ± 0.00022	0.00082 ± 0.00031	0.00038 ± 0.00041	0.00072 ± 0.00066	0.00033 ± 0.00033	0.00059 ± 0.00047	-		

3.2.2. Study Ⅱ. Disturbance frequency: Community composition in terms of functional genes and taxonomy.

To visualize both taxonomic family-level and functional metagenome profiles, nonmetric multidimensional scaling (NMDS) plots were generated for summarizing taxonomic and functional information. In terms of community taxonomic composition from the metagenome, the NMDS plot showed that replicates clearly clustered by disturbance frequency, in terms of taxonomic profiles (Fig. 10). In Fig. 10A, the plot was generated at family level based on M5NR taxonomic profile. Non-disturbed control samples, sieved but not partially sterilized and stored alongside the other pots at 25 ° C for the 24 week period clustered nearby the original soil in both functional and taxonomic terms (Fig. 10A, Fig. 10B). Of the experimental treatments which had gone through different frequency of disturbance by volume sterilization, the samples with the highest disturbance frequency (frequency 1) were most distinct from the original soil and non-disturbed soil. By declining disturbance frequency, communities moved toward in the direction of the original soil. However, after frequency 0.25, there were little differences among samples although disturbance was occurred less. The samples subjected to disturbance once in a 12 weeks still

tended to closer to other disturbance treatment samples than original soil.

Structure of the bacterial community showed a similar trend. At the genus level, calculated from bacterial 16S rRNA genes extracted from metagenomes, samples which subjected to the highest disturbance had the most distinctive communities from original soil and clustered farthest on the NMDS (Fig. 10C). The communities with lower disturbance frequency composed nearby original/non-disturbance samples. Also, soil bacteria showed clear differences with a low frequency of disturbance 0.0833, although the samples had 12 weeks of recovery time.

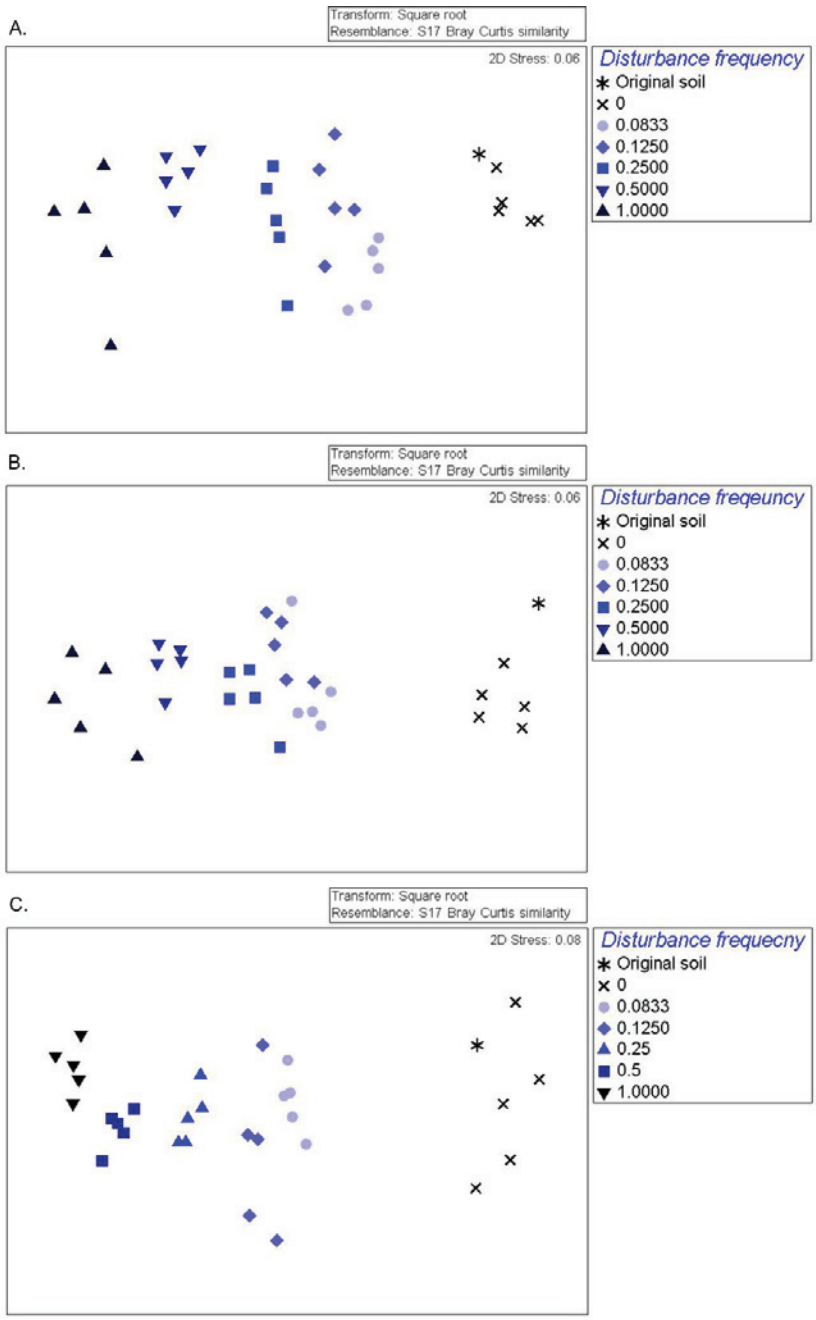


Figure 10. NMDS plot generated using weighted pairwise Unifrac distances between samples. A. NMDS plot of shotgun metagenomic sequences based on M5NR taxonomic profile at family level, B. NMDS plot of shotgun metagenomic sequences based on SEED functional profile at subsystem level 3, C. NMDS plot of 16S rRNA bacterial genus level which extracted from metagenome.

3.2.3. Study Ⅱ. Disturbance frequency: Alpha diversity.

Diversity was calculated based on Shannon diversity index. Shannon diversity index was calculated based on M5NR taxonomy at species level as concept of diversity is based on species level and SEED subsystem function at level 1, 2 and 3 by using software R. In terms of alpha-diversity, the species level diversity (for taxonomically described published species) detected in the metagenomes showed a trend of continuously decrease with increasing disturbance frequency ($P < 0.0001$, Fig. 11A). Shannon diversity index of functional gene showed different relationship with disturbance frequency in each level. In gene classification at functional level 1 and 3 had positive sigmoidal correlation with disturbance frequency ($P < 0.0001$, Fig. 11B) ($P < 0.0001$, Fig. 11D). However, diversity at functional level 2 was reduced by increasing disturbance frequency ($P < 0.001$, Fig. 11C).

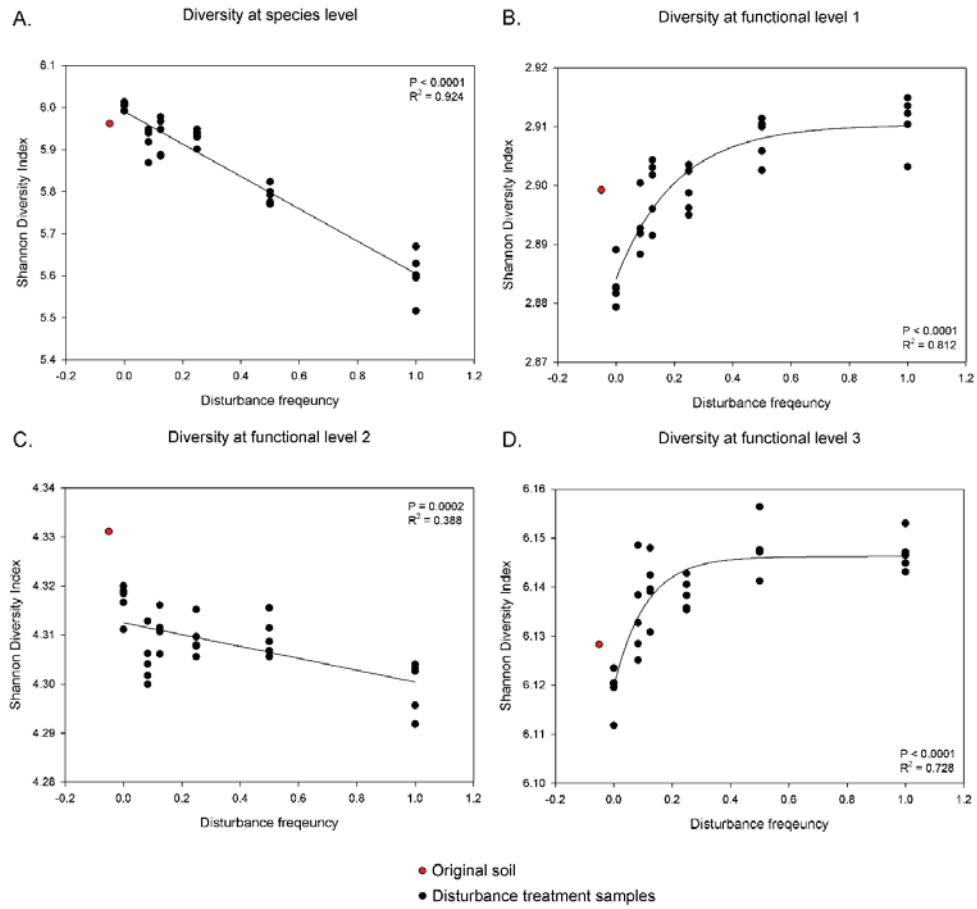


Figure 11. Taxonomic (M5NR database) and Functional diversity (SEED database) at soil microbial communities by disturbance frequency based on Shannon diversity index. Solid line represents linear, sigmoidal regression fits to data. A. Diversity at species level from all kingdoms, B. Diversity at functional level 1, C. Diversity at functional level 2, D. Diversity at functional level 3. Colored point was not included in the statistics.

3.2.4. Study Ⅱ. Disturbance frequency: Predominant changes in functional gene categories.

To examine the effect of disturbance frequency on soil microbial functions, the functional profile of shotgun metagenomic sequences was analysed using the SEED database, and the metagenomic sequences were distributed into 28 functional gene categories (level 1 SEED subsystems). Of the 28 functional gene categories, 24 categories differed significantly ($P < 0.05$) in relation to time since disturbance (Fig. 12, Fig. 13). More number of functional genes at level 1 classification was related to disturbance frequency than disturbance time (Fig. 7, Fig. 8).

At functional level 1, 12 gene categories had positive correlation with increasing disturbance frequency (Fig. 12). Iron acquisition and metabolism ($P < 0.0001$, Fig. 12A), potassium metabolism ($P < 0.0001$, Fig. 12B), phosphorus metabolism ($P < 0.01$, Fig. 12C), motility and chemotaxis ($P < 0.0001$, Fig. 12D), miscellaneous ($P < 0.0001$, Fig. 12E), Cell wall and capsule ($P < 0.0001$, Fig. 12F), dormancy and sporulation ($P < 0.01$, Fig. 12G), regulation and cell signaling ($P < 0.05$, Fig. 12H), RNA metabolism ($P < 0.0001$, Fig. 12I), DNA metabolism ($P < 0.01$, Fig. 12J), cell division and cell cycle ($P < 0.001$, Fig. 12K), virulence, disease and defense ($P = 0.001$, Fig. 12L) showed an increase in relative

abundance through increase in disturbance frequency. Functions which are related to RNA/DNA metabolism, cell division and cycle are one of important functions of the 'housekeeping' gene category, which are constitutively expressed to maintain cellular function (Zhu et al., 2008; She et al., 2009; Eisenberg and Levanon, 2013). Thus, the relative abundance of housekeeping genes in this category showed an increasing trend by increasing disturbance.

To understand more narrowly defined functions under 12 positively related gene level 1, functional genes classified as level 2 were analysed (Table 9). Among 30 functions at level 2, 21 functions were positively correlated to disturbance frequency and 9 functions were negatively correlated. Functional gene categories at level 2 under genes related to iron acquisition and metabolism, potassium metabolism, phosphorus metabolism, miscellaneous, dormancy and sporulation, cell division and cell cycle had positive correlation. However, although there were positive relation at level 1, some detailed functions under the level 1 had negative relation. 9 functional genes at level 2 among 30 functions were identified as genes which had negative relation. They were the functions related to social motility and nonflagellar swimming in bacteria ($P < 0.05$) under the genes related to motility, cell wall of gram−positive cell ($P < 0.05$) and mycobacteria ($P < 0.0001$) under function of cell

wall and capsule, RNA metabolism genes ($P < 0.001$), proteolytic pathway ($P < 0.01$) under the regulation and cell signaling genes, CRISPs ($P < 0.01$) under DNA metabolism category, toxins and superantigens ($P < 0.01$), virulence, disease and defense ($P < 0.05$), detection ($P < 0.05$) under virulence, disease and defense genes.

At functional level 1 using the SEED database, 12 gene categories related to protein metabolism ($P < 0.01$, Fig. 13A), carbohydrates ($P < 0.0001$, Fig. 13B), fatty acids, lipids, and isoprenoids ($P < 0.01$, Fig. 13C), nucleosides and nucleotides ($P < 0.05$, Fig. 13D), respiration ($P < 0.0001$, Fig. 13E), amino acids and derivatives ($P < 0.01$, Fig. 13F), nitrogen metabolism ($P < 0.001$, Fig. 13G), photosynthesis ($P < 0.0001$, Fig. 13H), metabolism of aromatic compounds ($P < 0.0001$, Fig. 13I), sulfur metabolism ($P < 0.01$, Fig. 13J), stress response ($P < 0.01$, Fig. 13K) and clustering based subsystems ($P < 0.05$, Fig. 13L) had negative relation with increasing disturbance frequency. At functional level 2, totally 83 functions which were distributed under 12 functional gene level 1 categories were identified as a significant relation with disturbance frequency (Table 10). 30 functional classifications at level 2 had positive relation and 53 functions had negative relation. Especially, functional genes related to clustering-based subsystems and carbohydrates contained 42 functions and 10 functional genes at

level 2 respectively which all had significant relations with disturbance frequency ($P < 0.05$).

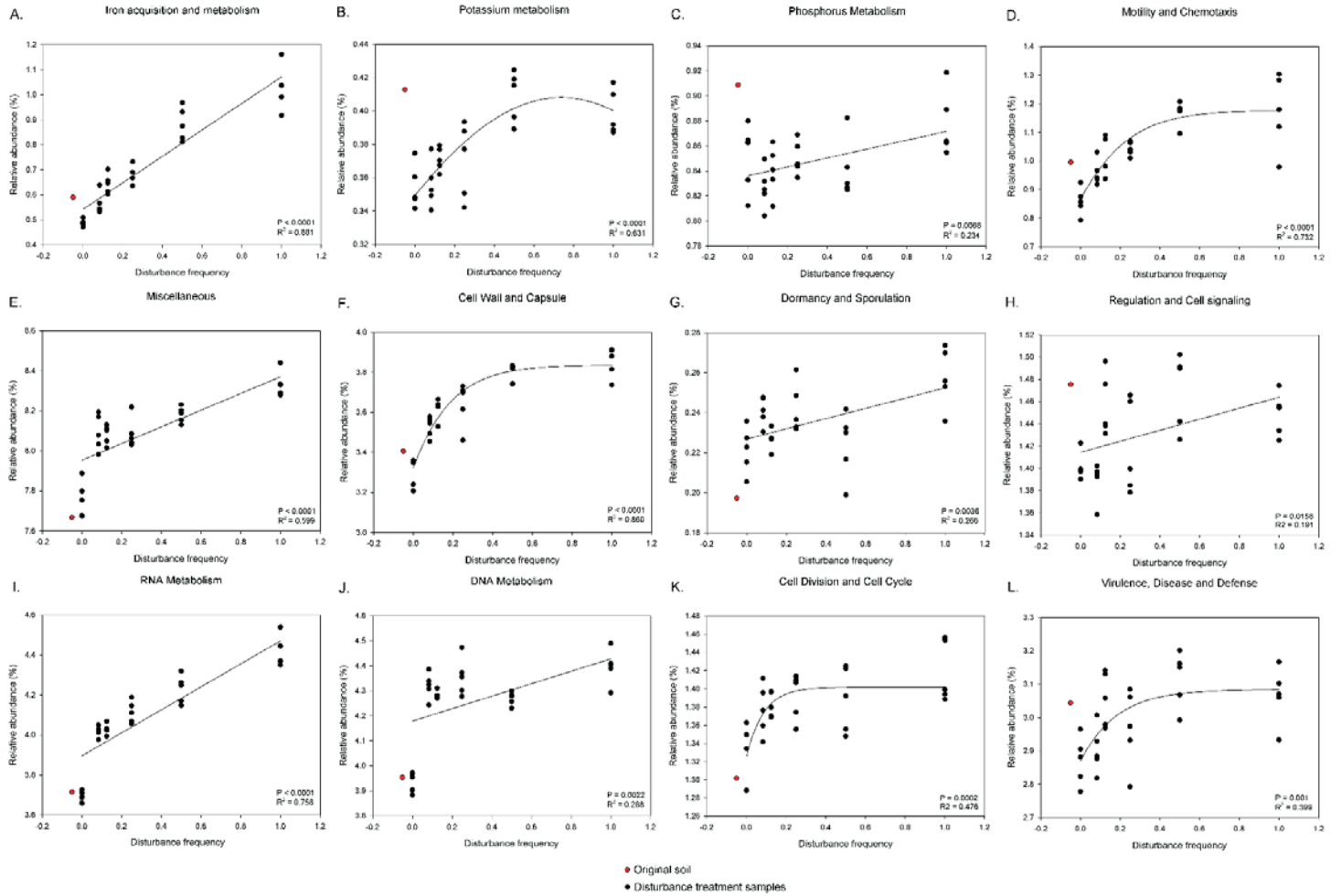

A.
Iron acquisition and metabolism
B.
Potassium metabolism
C.
Phosphorus Metabolism
D.
Motility and Chemotaxis
E.
Miscellaneous
F.
Cell Wall and Capsule
G.
Dormancy and Sporulation
H.
Regulation and Cell signaling
I.
RNA Metabolism
J.
DNA Metabolism
K.
Cell Division and Cell Cycle
L.
Virulence, Disease and Defense
Relative abundance (%)
Disturbance frequency
Original soil
Disturbance treatment samples

Figure 12. Relationship between disturbance frequency and relative abundance of functional gene categories at subsystem level 1 (SEED database). Solid line represents linear or sigmoidal regression fits to data. A. Iron acquisition and metabolism, B. Potassium metabolism, C. Phosphorus metabolism, D. Motility and chemotaxis, E. Miscellaneous, F. Cell wall and capsule, G. Dormancy and sporulation, H. Regulation and cell signaling, I. RNA metabolism, J. DNA metabolism, K. Cell division and cell cycle, L. Virulence, disease and defense.

Table 9. Functional abundance analysis of level 2 under functional gene level 1 (SEED database) which had positive correlation with disturbance frequency. Only classification levels having P < 0.05 were shown.

Function classification Level 1	**Function classification Level 2**	**P**	**R2**	**Coefficient**
Iron acquisition and metabolism	Iron acquisition and metabolism	4.56E-15	0.8921	0.47025
Potassium metabolism	Potassium metabolism	1.74E-05	0.4885	0.046585
Phosphorus Metabolism	Phosphorus Metabolism	0.001918	0.2952	0.03696
Motility and Chemotaxis	Social motility and nonflagellar swimming in bacteria	0.01446	0.1954	-0.0008
	Flagellar motility in Prokaryota	7.18E-07	0.5902	0.22821
	Motility and Chemotaxis	0.001508	0.3064	0.07519
Miscellaneous	Miscellaneous	0.00631	0.2375	0.02772
	Plant-Prokaryote DOE project	2.37E-07	0.6207	0.39424
Cell Wall and Capsule	Gram-Positive cell wall components	0.02058	0.1771	-0.00864
	Cell wall of Mycobacteria	2.63E-07	0.618	-0.07501
	Cell Wall and Capsule	5.06E-06	0.5304	0.18497
	Gram-Negative cell wall components	5.72E-09	0.7081	0.33165
Dormancy and Sporulation	Spore DNA protection	0.02134	0.1752	0.000813
	Dormancy and Sporulation	0.002276	0.287	0.023475
RNA Metabolism	RNA Metabolism	0.00025	0.3856	-0.01122
	RNA processing and modification	9.89E-09	0.6966	0.55602
Regulation and Cell signaling	Proteolytic pathway	0.006494	0.236	-0.00099

	Regulation of virulence	3.61E-09	0.7174	0.039932
	Regulation and Cell signaling	0.009359	0.2176	0.054207
DNA Metabolism	DNA repair	0.00093	0.3285	0.16771
	DNA uptake, competence	0.009241	0.2183	0.010834
	DNA replication	0.001609	0.3034	0.13763
	CRISPs	0.001316	0.3127	-0.00806
Cell Division and Cell Cycle	Cell Division and Cell Cycle	0.000384	0.3675	0.065796
Virulence, Disease and Defense	Toxins and superantigens	0.004519	0.254	-0.00278
	Virulence, Disease and Defense	0.01989	0.1788	-0.02283
	Resistance to antibiotics and toxic compounds	0.001601	0.3036	0.20224
	Bacteriocins, ribosomally synthesized antibacterial peptides	6.26E-07	0.5941	0.009187
	Detection	0.02793	0.1611	-0.01721
	Invasion and intracellular resistance	0.02303	0.1712	0.002992

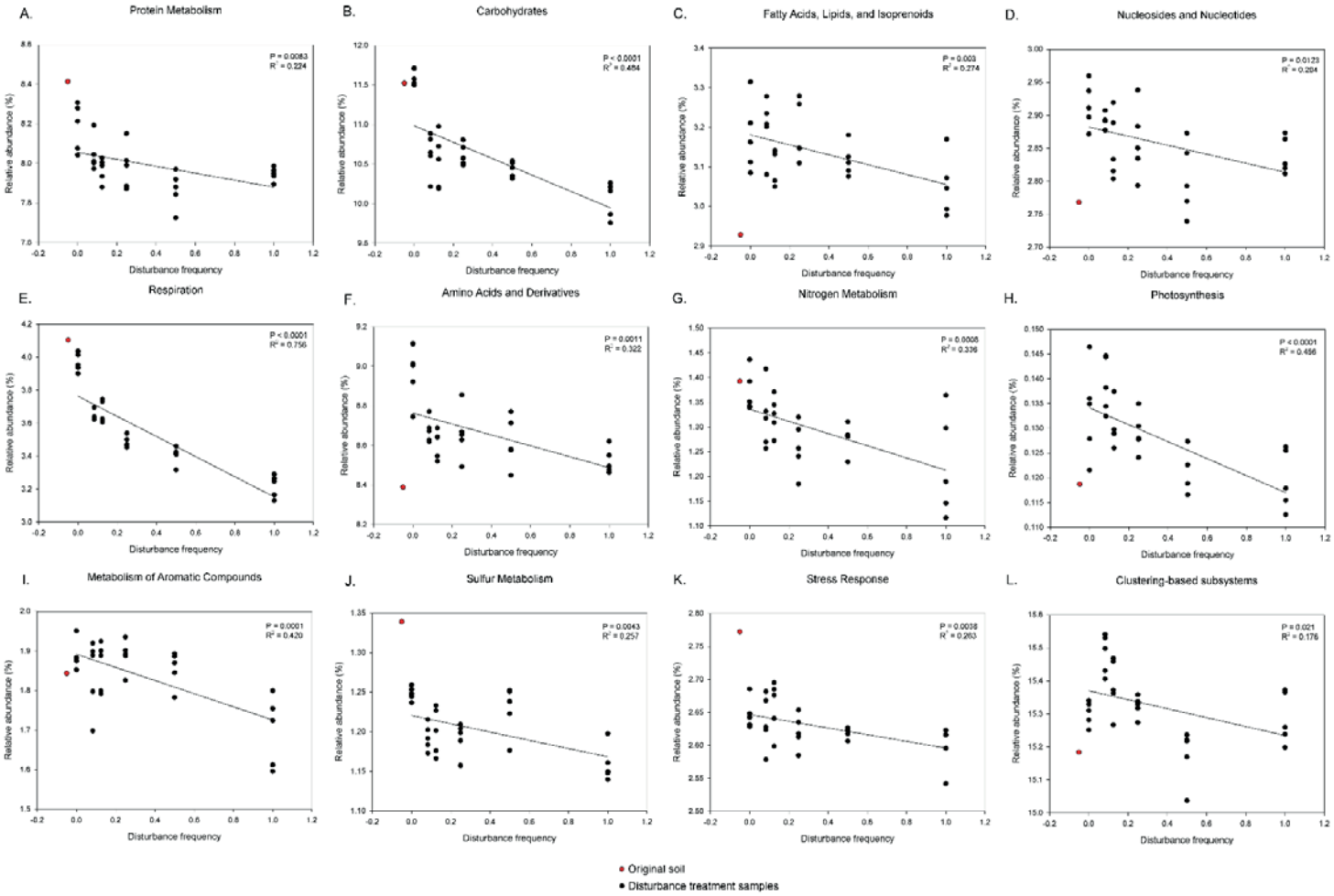

A.
Protein Metabolism
P = 0.0083
R² = 0.224
B.
Carbohydrates
P < 0.0001
R² = 0.464
C.
Fatty Acids, Lipids, and Isoprenoids
P = 0.003
R² = 0.274
D.
Nucleosides and Nucleotides
P = 0.0123
R² = 0.204
E.
Respiration
P < 0.0001
R² = 0.756
F.
Amino Acids and Derivatives
P = 0.0011
R² = 0.322
G.
Nitrogen Metabolism
P = 0.0008
R² = 0.336
H.
Photosynthesis
P < 0.0001
R² = 0.456
I.
Metabolism of Aromatic Compounds
P = 0.0001
R² = 0.420
J.
Sulfur Metabolism
P = 0.0043
R² = 0.257
K.
Stress Response
P = 0.0038
R² = 0.263
L.
Clustering-based subsystems
P = 0.021
R² = 0.176
Relative abundance (%)
Disturbance frequency
Original soil
Disturbance treatment samples

Figure 13. Relationship between disturbance frequency and relative abundance of functional gene categories at subsystem level 1 (SEED database). Solid line represents linear or sigmoidal regression fits to data. A. Protein metabolism, B. Carbohydrates, C. Fatty acids, lipids, and isoprenoids, D. Nucleosides and nucleotides, E. Respiration, F. Amino acids and derivatives, G. Nitrogen metabolism, H. Photosynthesis, I. Metabolism of aromatic compounds, J. Sulfur metabolism, K. Stress response, L. Clustering based subsystems.

Table 10. Functional abundance analysis of level 2 under functional gene level 1(SEED database) which had negative correlation with disturbance frequency. Only classification levels having P < 0.05 were shown.

Function classification Level 1	Function classification Level 2	P	R2	Coefficient
Protein Metabolism	Selenoproteins	6.05E-09	0.7069	-0.07402
	Protein biosynthesis	0.000797	0.3355	0.14737
	Protein processing and modification	5.02E-08	0.6598	-0.1551
Carbohydrates	Aminosugars	0.000106	0.4208	0.04563
	Sugar alcohols	0.000133	0.4117	-0.06334
	Polysaccharides	3.29E-09	0.7193	-0.11456
	One-carbon Metabolism	3.63E-10	0.7597	-0.26445
	Organic acids	0.01619	0.1896	-0.03922
	Di- and oligosaccharides	3.22E-08	0.6703	-0.36618
	Central carbohydrate metabolism	4.77E-07	0.6017	-0.50242
	CO2 fixation	1.31E-06	0.5726	-0.09871
	Carbohydrates	0.001961	0.2941	-0.19112
	Fermentation	3.70E-06	0.5405	-0.10962
Fatty Acids, Lipids, and Isoprenoids	Fatty acids	0.000183	0.3986	-0.10357
	Triacylglycerols	0.000375	0.3685	0.004123
Nucleosides and Nucleotides	Pyrimidines	0.01257	0.2026	-0.03108
	Nucleosides and Nucleotides	1.01E-06	0.5802	-0.06649

	Detoxification	1.28E-09	0.7373	0.052005
Respiration	Sodium Ion-Coupled Energetics	0.007009	0.2322	-0.00468
	Reverse electron transport	0.004642	0.2527	-0.00174
	Electron donating reactions	2.42E-09	0.7253	-0.42088
	Respiration	1.57E-05	0.4921	-0.09059
	Electron accepting reactions	4.05E-09	0.7151	-0.0916
Amino Acids and Derivatives	Glutamine, glutamate, aspartate, asparagine; ammonia assimilation	5.26E-06	0.5292	-0.07292
	Amino Acids and Derivatives	5.13E-08	0.6593	-0.06805
	Arginine; urea cycle, polyamines	0.01336	0.1995	-0.06831
	Aromatic amino acids and derivatives	1.75E-06	0.5638	0.085156
	Branched-chain amino acids	1.99E-05	0.4837	-0.18088
	Proline and 4-hydroxyproline	5.29E-05	0.4476	0.053582
	Histidine Metabolism	3.92E-05	0.4589	0.054793
Nitrogen Metabolism	Nitrogen Metabolism	0.001627	0.3029	-0.10257
Photosynthesis	Photosynthesis	0.000329	0.3741	-0.01344
Metabolism of Aromatic Compounds	Anaerobic degradation of aromatic compounds	0.01108	0.2091	-0.0168
	Peripheral pathways for catabolism of aromatic compounds	0.000181	0.3992	-0.08872
Sulfur Metabolism	Organic sulfur assimilation	0.004347	0.2559	-0.02912
	Inorganic sulfur assimilation	0.01981	0.1791	-0.0234
Stress Response	Detoxification - #1	8.23E-06	0.5143	0.01778
	Acid stress	0.006127	0.2389	0.009071
	Osmotic stress	0.000236	0.3882	-0.06305

	Periplasmic Stress	0.004712	0.2519	0.013116
Clustering-based subsystems	recX and regulatory cluster	0.006397	0.2368	0.007637
	Putative GGDEF domain protein related to agglutinin secretion	3.93E-11	0.7947	0.014642
	Sarcosine oxidase	0.000551	0.3518	-0.01498
	Nucleotidyl-phosphate metabolic cluster	4.01E-06	0.5379	-0.06934
	Hypothetical lipase related to Phosphatidate metabolism	0.01767	0.185	-0.00514
	alpha-proteobacterial cluster of hypotheticals	0.04526	0.1356	-0.0093
	Methylamine utilization	1.82E-05	0.4868	-0.08511
	Shiga toxin cluster	0.000103	0.422	0.039451
	Translation	3.18E-06	0.5453	0.035347
	Two related proteases	1.58E-08	0.6864	-0.04092
	Oxidative stress	7.17E-06	0.5189	0.005617
	Lysine Biosynthesis	0.000634	0.3457	0.008756
	Tricarboxylate transporter	0.002075	0.2915	-0.04091
	Three hypotheticals linked to lipoprotein biosynthesis	1.10E-05	0.5043	0.018066
	Ribosome-related cluster	5.18E-06	0.5297	0.026652
	Carotenoid biosynthesis	0.000762	0.3375	-0.01705
	Probably organic hydroperoxide resistance related hypothetical protein	4.40E-07	0.6039	0.012491
	D-tyrosyl-tRNA(Tyr) deacylase (EC 3.1.-.-) cluster	0.000192	0.3966	0.012642
	Carbohydrates - #1	0.0456	0.1352	-0.00806
	Proteasome related clusters	0.01501	0.1935	-0.00098
	Catabolism of an unknown compound	0.00014	0.4095	-0.02

Related to Menaquinone-cytochrome C reductase	6.68E-05	0.4387	-0.00523
proteosome related	0.00067	0.3432	-0.04018
DNA polymerase III epsilon cluster	0.00032	0.3754	0.014216
Probably Ybbk-related hypothetical membrane proteins	0.002081	0.2913	0.007259
Cytochrome biogenesis	9.47E-08	0.6444	-0.06279
Fatty acid metabolic cluster	0.04833	0.1321	-0.01678
Biosynthesis of galactoglycans and related lipopolysacharides	1.61E-10	0.7731	-0.09138
Chemotaxis, response regulators	0.02894	0.1592	0.002313
Hypothetical protein possible functionally linked with Alanyl-tRNA synthetase	0.003841	0.2619	0.011613
Putative asociate of RNA polymerase sigma-54 factor rpoN	0.00131	0.3129	-0.01977
Ribosomal Protein L28P relates to a set of uncharacterized proteins	1.61E-05	0.4911	-0.01684
DNA metabolism	1.35E-10	0.7759	0.013153
Choline bitartrate degradation, putative	0.005756	0.242	-0.00773
Hypothetical in Lysine biosynthetic cluster	7.08E-08	0.6516	0.036955
CRISPRs and associated hypotheticals	0.04359	0.1376	-0.00309
Putative Isoquinoline 1-oxidoreductase subunit	0.000201	0.3947	-0.02478
Hypothetical Related to Dihydroorate Dehydrogenase	9.13E-06	0.5108	0.010678
Hypothetical associated with RecF	0.00205	0.292	0.006931
TldD cluster	6.85E-08	0.6523	0.040388
Clustering-based subsystems	8.17E-05	0.4309	0.29865
Isoprenoid/cell wall biosynthesis: PREDICTED UNDECAPRENYL DIPHOSPHATE PHOSPHATASE	0.000408	0.3649	0.023633

Chapter 4. Discussion

4.1. Study Ⅰ. Time succession: Effect of time successional recovery after soil disturbance, on structure and function of soil community

4.1.1. Study Ⅰ. Time succession: Functional characteristics of early successional versus later successional soil communities

Motility.

As predicted, the early successional systems tended to have greater relative abundance of genes associated with motility (Fig. 8B). Extreme motility is a characteristic utilized by many early successional organisms, for example plants which rely on wind dispersal of seeds, or plants which rely on birds consuming berries and then dropping the seeds (Howe and Smallwood, 1982; Loveless and Hamrick, 1984). In later succession, the necessity of having function of motility reduced as if in situ growth and persistence of plants is more important than dispersal mechanisms on seeds (Walker and Chapin, 1987).

Active (as opposed to passive) motility in microorganisms takes various forms (Henrichsen, 1972; Harshey, 2003), but in the functional gene category characterized here it is dominated by genes for both flagellar and nonflagellar swimming, mainly in

bacteria which dominate the system (Table 5). Social motility is one of well-characterized group activities of bacterial(Shapiro, 1998; Velicer and Yuen-tsu, 2003). Surface-induced cooperative motilities are known to be widespread among bacteria (Harshey, 1994; Rashid and Kornberg, 2000; Verstraeten et al., 2008; Oberholzer et al., 2010).

Housekeeping genes.

Greater relative abundance of genes related to basic 'housekeeping' functions was predicted for the earlier stages of successional recovery of the soil system. However, the opposite pattern was found that housekeeping genes which related to basic metabolisms were more abundant in late successional stages (Fig. 7A, 7B, 7C). Increase of those basic metabolisms can be related to increase of cell division and cell cycle (Fig. 7F). After the dispersal by having high motility functions in early successional stages, microorganisms increased genes which related to cell division. However, the roll of housekeeping genes includes functions on 'amino acids and derivatives', 'cofactors, vitamins, prosthetic groups, pigments', 'nucleosides and nucleotides', 'protein metabolism'. Those functions showed variable patterns according to subdivision.

Stress response genes.

I had predicted increased abundance of stress response genes in later succession, in relation to the lower nutrient availability, and under increased interference competition. The main gene families identified under this category by MG-RAST were responses to oxidative stress, which have action in producing antioxidant like Glutathione and protection from reactive oxygen species. However, contrary to the prediction, there was a clear trend in the opposite direction: stress response genes were more abundant in the earliest time slices after disturbance and became progressively less common (Fig. 8D). It is unclear what it is about the earliest stages of succession in this system that leads to more of these stress response genes, or whether this indeed means that the early successional environment is in some way 'stressful'. In successional systems involving larger organisms, connectivity and mutualism are thought to increase over time as the system 'matures' (Odum, 1969) - by extension, in this microcosm successional system it is possible that stressful aspects of the early succession might be lack of connectivity and support networks due to longer cell-cell spacing, or some particular aspect of colonizing soil or organic matter particles that are devoid of living cells. One possibility is osmotic stress caused by release of ions and other solutes from dead cells, killed during autoclaving.

Cell signaling and virulence genes.

Also contrary to predictions, there were no significant differences in relative abundance of cell signaling and virulence & defense genes over time. In successional systems involving larger organisms, connectivity between organisms in both mutualistic and antagonistic interactions is seen as increasing in later succession (Odum, 1969). However, if the groups of genes I pinpointed are any guide, it appears that such interactions are equally important throughout the successional time scale I studied.

Viruses and anti-viruses defenses.

I had predicted that viruses, and the corresponding CRISPR elements which defend against them, would become progressively more abundant over successional time. However, both showed no significant differences during the successional time series.

Succession in functional gene assemblages.

Plotted on an NMDS, there is a clear and consistent progression in the total assemblage of functional genes (Fig. 4B). At first, in Week 1, the assemblage was the most different from the original soils, but over time it progressively became more similar, returning towards the original soils. However, as time passes, the rate of change on the NMDS slowed. By the end of the experiment, at week 24, the change had apparently ‘stalled’ or is not evident.

This is in contrast with the control microcosms which had not been through the 90% population reduction disturbance event and had been stocked with 100% unsterilized sieved soil. These undisturbed microcosms remained very close to the original soil on the NMDS. It appears, then, that the '90% killed' disturbance event has a dramatic effect on soil biota functional gene assemblage, which lasts at least 6 months, and (extrapolating from the slow rate of change), possibly much longer. This difference must be brought about initially by the presence of a large volume of colonizable soil rich in nutrients from dead soil biota, and the opportunities and challenges presented by this. It provides a potential analog for changes which might occur in nature or agricultural systems following drastic disturbance events. It appears that timescale of soil 'recovery' to the original state, 'engineering resilience' for biogeochemical functions is of the order of years as shown by earlier studies that measured processes (Holling, 1996; Gunderson, 2000). This study indicates how pervasive such changes are in terms of functional genes, from the lack of complete recovery likely persisting on the timescale longer than I measured here. By contrast, less drastic physical disturbance of a soil (in this experiment, the initial sieving and then storage for a week at 3 ℃) appears to have very little effect on soil functional gene profile after 24 weeks – as revealed

by the soil’s resemblance to the initial state at time 0, and also its close resemblance to the garden soil on day of harvest, even after a further 6 months’ incubation.

4.1.2. Study Ⅰ. Time succession: Changes in taxonomic community composition and diversity

Community composition.

As other studies have shown, I found a predictable succession of communities of microbes over time (Fig. 4A, 4C). At the broadest level, bacteria became less abundant and Archaea became more abundant (Fig. 4A, 4C). The increase in Archaea over time may represent their slow rate of natural increase rather than a fundamental adaptation to the late successional environment. Presumably they are mostly playing a role in NH4 oxidation (Schleper et al., 2005; Leininger et al., 2006) and their build up in abundance may reflect the exploitable supply of ammonia in the soil. With absence of new primary production from plants, heterotrophic bacteria by contrast may become less abundant in the soil. In case of disturbance by soil salinity and agricultural land uses, Archaea showed high diversity (Walsh et al., 2005; Roesch et al., 2007).

At taxonomic level of phyla, amongst the bacteria, there is a shift from early Proteobacteria to later Bacteroidetes over time

(Table 3). This may relate to pH and TOC, since pH and TOC decreased over the time sequence ($P < 0.01$, $P < 0.05$ respectively, Table 2). With Envfit on the NMDS, with in the community composition, TOC had significant influence on the community composition in the result from shotgun metagenomic sequences based on M5NR taxonomic profile and pH and TN influence the communities based on the extracted bacterial 16S rRNA genus level (Fig. 5A, Fig. 5B).

There is a clear succession of total soil biota community composition, as judged from the metagenome reads that can be assigned taxonomically (Fig. 4A), the NMDS shows that communities with longer incubation time after disturbance treatment became more similar, returning towards the original soils. However, as time passes the rate of change on the NMDS slows. By the end of the experiment, at week 24, the change has apparently 'stalled' or is not evident. This feature was also observed in composition of functional genes (Fig. 4B). As with the functional genes, the stalling of taxonomic composition change in the community also provides a potential analog for changes which might occur in nature or agricultural systems following drastic disturbance events. In natural disturbed systems regrowth of plants and renewed animal activity might accelerate return of the soil

community to its original state. However, field observations also suggest very long time scale for recovery of the soil community to its original state following fire disturbance – as much as a decade (Xiang et al., 2014).

Effects of disturbance on diversity.

Diversity at the species level based on the M5NR database recovered after the initial disturbance event (Fig. 6A), unlike the diversity at different functional levels (Fig 6B, 6C, 6D). At the 1-week stage after initial disturbance, Shannon diversity was lower than on day 0, but increased until the 4-week stage (Fig. 6A). This species level diversity at species level was then maintained until the end of experiment.

Although taxonomic diversity recovered and maintained itself after the 4-week stage, diversity of functional genes remained lower than day 0, decreased over time (Fig 6B, 6C, 6D). In this cultured microcosm experiments which exclude abiotic/biotic interactions, this implies that microbial communities impacted by disturbance event change their functions gradually over time and can be different from original soil after a 6 month period.

4.2. Study Ⅱ. Disturbance frequency: Effect of disturbance frequency on structure and function of soil community

4.2.1. Study Ⅱ. Disturbance frequency: Changes in taxonomic community diversity and composition.

Community composition.

There was predictable sequential changes of communities of microbes based on the difference in frequency of disturbance (Fig. 10A, Fig. 10C). As predicted, bacteria and viruses became more abundant ($P < 0.01$, Fig. 9A, Fig. 9D) and Eukaryota and Archaea became less abundant ($P < 0.0001$, Fig. 9B, Fig. 9C) at the broadest taxonomic level. Short generation times will favor copiotrophic groups (Grime, 2006), in here, bacteria. Bacteria tend to grow faster than Eukaryota, by having advantages on asexual reproduction and short generation time relative to large organisms (Cooper, 2012).

It appears from this study that members of some phyla of bacteria tend to play different ecological 'roles' in exploiting resources available in the soil. 23 phyla under bacteria showed strong relation with disturbance (Table 6). Among 23 phyla, only Proteobacteria and Bacteroidetes increased by increasing disturbance frequency with big rate. So, Proteobacteria and

Bacteroidetes could be regarded as 'r' selected in plant ecology (Grime, 2006), or 'copiotrophs,' by the terminology of Fierer et al. (2007). Especially Bacteroidetes were identified as a disturbance tolerant phyla in previous study (Kim et al., 2013). Also, Actinobacteria, Gemmatimonadetes, Acidobacteria, Chloroflexi and Planctomycetes were could be consider as a slow growing/disturbance sensitive groups and may be compared to 'K' selected groups among the phyla which had more than 1 % of relative abundance on average (Table 6). Acidobacteria and Planctomycetes were discussed as a 'K' selected groups in previous study (Kim et al., 2013) and this result supports the study.

Although relative abundance of Eukaryota was 1.1 %, 52 phyla were identified (Table 7). 20 phyla had negative correlation with disturbance frequency, and 2 phyla, Onychophora and Streptophyta had positive correlation. Increase of Streptophyta can be explained by strategy of Streptophyta as a 'r' selected, since descendants of streptophyte algae was contributed on the colonization of terrestrial habitats (Gensel, 2008; Becker, 2013). Relative abundance of Ascomycetes and Basidiomycetes in fungi had decline trend contrary to expectation. Although Ascomycota has a fast evolutionary rate (Wang et al., 2010), the period of disturbance might be short for fungi to adapt.

With low populated groups, less than 1 % relative abundance, identified phyla of Archaea showed generally decreased (Table 8). Also, viruses increased when disturbance frequency increased as I predicted. Although classification of M5NR database did not reveal lower level classification, derived from viruses but unclassified viruses had increased, sigmoidal relation with disturbance frequency as I predicted. This result might be because viruses could actually be more abundant in the fast-growing r-selected populations of more disturbed systems (Suttle, 2007).

There is a clear sequential changes on total soil biota community composition, as judged from the metagenome reads that can be assigned taxonomically (Fig. 10A), the NMDS shows that less disturbed communities became more similar the original soils. However, distance between treatment samples and original status on NMDS was not proportional, as disturbed soils composed closely together although the disturbance frequency was low. This feature was also observed in composition of functional genes (Fig. 10B).

4.2.2. Study Ⅱ. Disturbance frequency: The relation between taxonomic diversity and functional diversity.

Effects of disturbance on diversity.

The experiment reveals that increasing disturbance frequency has a pervasive effect on soil microbial diversity. This is despite the fact that in some respects, attributes of the soil and its community stayed remarkably constant irrespective of disturbance. Along the disturbance gradient, soil microbial diversity at the species level based on the M5NR database declined with increasing disturbance frequency, at the range of frequencies used here ($P < 0.0001$, Fig. 11A). This contrasts with many classical studies of ecological systems on larger organisms for example, of plants, corals and algae, where a 'humpbacked' diversity curve is reported (Connell, 1978; Molino and Sabatier, 2001; Ikeda, 2003). It is unclear whether widening the range of frequencies of disturbance would have produced such a curve, but what
the declining trend does seem to make clear is that the remarkably high diversity of soil bacterial communities is sensitive to physical disturbance as our group studied on soil bacterial diversity (Kim et al., 2013).

According to the view of Grime (Grime, 2006), the declining diversity I observed would be explained in terms of the demands on population survival and the need for rapid growth and reproduction at high disturbance, where fewer lineages have successfully managed to evolve the complex adaptations needed. There is no

sign of the opposite effect, whereby the release from competition caused by disturbance increases diversity. This may indicate that at least in the soil community microcosm studied here, competition is not of great enough importance in limiting the number of species which can survive together.

Diversity at different functional levels based on SEED database had distinct relations depending on the classification. Functional level diversity at 1 and 3 were tended to increase along with disturbance frequency by having sigmoidal regression ($P < 0.0001$, Fig. 11B, Fig. 11D). However, diversity at functional level 2 classification had negative correlation like diversity at species level.

When I compare the diversity of the lowest taxonomic level and functional level, which is species level and functional classification level 3, changes in taxonomic diversity causes changes of functional diversity inversely in this microbial microcosm study. In expectation, diversity of species begets diversity of genetically determined functionality in ecosystems and from this in turn comes greater stability against perturbation (Naeem and Li, 1997). In this study at least, I found species diversity loss did not influence on loss of function.

4.2.3. Study Ⅱ. Disturbance frequency: Functional characteristics of different disturbance frequencies.

Housekeeping genes.

I found relatively high rate of genes related to RNA/DNA metabolism, cell division and cell cycle as predicted there will be relatively high portion of 'housekeeping gene' functions in the more frequently disturbed system (Fig. 12I, Fig. 12J, Fig. 12K). Increase of RNA/DNA metabolisms can be related to increase of cell division and cell cycle (Fig. 12K). However, since the role of housekeeping genes includes various functions not only RNA/DNA metabolisms, but 'protein metabolism (Fig. 13A)', 'nucleosides and nucleotides (Fig. 13D)', 'amino acids and derivatives (Fig. 13F)' which showed negative correlation with disturbance frequency. The housekeeping genes showed variable patterns according to subdivision also and hard to define relation with disturbance in this experiment.

Motility.

Genes related to motility became more abundant with higher frequency of disturbance as expected (Fig. 12D). Samples undergo disturbance every week were remained in the early successional-like stage during experiment time. As I discussed with time successional experiment 1-week samples earlier in this paper,

extreme motility is a characteristic utilized by many early successional organisms, for example plants which rely on wind dispersal of seeds, or plants which rely on birds consuming berries and then dropping the seeds (Howe and Smallwood, 1982; Loveless and Hamrick, 1984). Both samples with disturbance frequency 1 and incubated 1-week after disturbed samples showed high rate of motility related genes by supporting traditional view of dispersal. This dispersal might be connected with gene functions like cell division and DNA/RNA metabolisms with colonization. Under the function of motility, 3 functional classification at level 2 had significant correlation which were functions on social motility and nonflagellar swimming in bacteria, flagellar motility in Prokaryota and motility and chemotaxis (Table 7). Active (as opposed to passive) motility in microorganisms takes various forms (Henrichsen, 1972; Harshey, 2003), but in the functional gene category characterized here it is dominated by genes for both flagellar and nonflagellar swimming, mainly in bacteria. Social motility is one of well-characterized group activities of bacterial(Shapiro, 1998; Velicer and Yuen-tsu, 2003). Surface-induced cooperative motilities are known to be widespread among bacteria by commonly using flagella (Harshey, 1994; Rashid and Kornberg, 2000; Verstraeten et al., 2008; Oberholzer et al., 2010).

Stress response genes.

Contrary to predictions, stress response genes identified under this category by MG-RAST were relatively decreased by increasing disturbance frequency (Fig. 13K). However, the 3 gene families as classified at level 2 under stress responding genes increased out of 4 functions. Only genes related to osmotic stress were decreased and detoxification, acid stress and periplasmic stress responding genes had positive coefficient values.

In successional systems involving larger organisms, connectivity and mutualism are thought to increase over time as the system 'matures' (Odum, 1969) – by extension, in this microcosm successional system it is possible that stressful aspects of the early succession might be lack of connectivity and support networks due to longer cell-cell spacing, or some particular aspect of colonizing soil or organic matter particles that are devoid of living cells.

Virulence, Disease and Defense.

Functional gene classification at level 1 based on SEED databases showed genes related to virulence, disease and defense were increasing with frequent disturbance (Fig. 12L). This increased abundance can be connected with increased abundance of viruses (Fig. 9D). When I analysed at functional classification at

level 2, functions on resistance to antibiotics and toxic compounds, bacteriocins, ribosomally synthesized antibacterial peptides, Invasion and intracellular resistance had positive coefficient values. With dense populations of microbes whose strategy emphasizes rapid reproduction over defense and genetic re-assortment, viruses in having very brief generation times may be better able to spread through their specific host populations causing outbreaks and population crashes. Thus, this might be the reason why genes related to viruses and anti-viruses defenses was increased as I predicted.

Chapter 5. Conclusion

This study gives some perspectives on a commonplace, important, but little studied system – disturbed soil. The results revealed strong, predictable changes in both taxonomic composition and functional gene profile of the soil biota, in response to a nonselective disturbance event that had killed most of the soil biota.

Interestingly, it is apparent from this experiment that taxonomic 'diversity' – as identified from the metagenome – in this system does not relate to functional diversity, as defined in terms of genes present. This is in contrast with the indications of microcosm studies on larger organisms which have suggested that greater species diversity leads to greater empirical functional diversity – also leading to greater resilience (Tilman, 1994). Thus, for soil microbiota these two characteristics of the community which would be expected to be closely linked, are in fact decoupled.

In some respects, this miniature successional system agreed with my predictions based upon observation of 'classic' successions involving large organisms. For example, all the biotic characteristics of the soil changed in the direction of a return towards the initial conditions before disturbance, although recovery

apparently ‘stalled’ at later stages. Also, for example, more abundant functional motility genes were more abundant in early succession, as predicted. However, in other respects the predictions based on analogy with successional systems of large organisms were not supported: for example, functional rate of cell division and housekeeping genes related to protein/DNA/RNA were instead more abundant at later stages.

Also, studying effect of disturbance effect on microcosm system helped us to understand responses of microorganism under disturbance in terms of taxonomic and functional differences. The concepts of copiotrophic microbes and functional reactions of soil community agreed on my predictions based upon observation of ‘classic’ successions involving large organisms. For example, all the biotic characteristics of the soil changed clustered together by the disturbance frequency and the less disturbed communities were placed nearer to the initial status and ‘copiotrophic’ organisms were more abundant in more frequently disturbed soil system. Also, in terms of functions, functions related to ‘housekeeping’, motility, viruses and anti-viruses defenses were more abundant as predicted. However, in other respects the predictions based on analogy with successional systems of large organisms were not supported: for example, changes of taxonomic diversity did not

follow the traditional humpbacked curve but inversely proportion to the frequency of disturbance.

It would be interesting to consider in further investigations what is different about a soil system as compared to a classical ecological succession system involving large organisms – and whether it varies fundamentally from one soil type or disturbance event to another. It is also clear that disturbance of the type I simulated here, involving death of the majority of the soil biota, has a drastic effect that persists months, and perhaps years (as the system had still not reached the original or control state by the end of this 6 month experiment).

It would be very interesting to use PCR based sequencing of 16S genes to study the roles of stochasticity in bacterial community composition the disturbed pots systems. It would also be interesting to link the functional and diversity changes seen in an experimental system such as this to biogeochemical functions of the soil such as respiration, nitrogen mineralization, or the ability to metabolise pollutants. As such the findings could begin to have wider ecological and indeed practical importance, giving clues to the underlying mechanisms of spatial and temporal changes in the functional behavior of soils. Longer term microcosm studies with a different functional and compositional changes through variety of types of

disturbance events, plus observation of the trajectory of population dynamics over time, would also be interesting.

Bibliography

Adams, J., 2010. Species richness: patterns in the diversity of life. Springer Science & Business Media.

Allen, S.E., Grimshaw, H.M., Parkinson, J.A., Quarmby, C., 1974. Chemical analysis of ecological materials. Blackwell Scientific Publications.

Allison, S.D., Martiny, J.B., 2008. Resistance, resilience, and redundancy in microbial communities. Proceedings of the National Academy of Sciences 105, 11512-11519.

Bazzaz, F., 1979. The physiological ecology of plant succession. Annual review of ecology and systematics 10, 351-371.

Becker, B., 2013. Snow ball earth and the split of Streptophyta and Chlorophyta. Trends in Plant science 18, 180-183.

Bell, W.J., 2012. Searching behaviour: the behavioural ecology of finding resources. Springer Science & Business Media.

Brock, T.D., 1971. Microbial growth rates in nature. Bacteriological Reviews 35, 39.

Brown, S., Lugo, A.E., 1990. Tropical secondary forests. Journal of tropical ecology 6, 1-32.

Buckling, A., Kassen, R., Bell, G., Rainey, P.B., 2000. Disturbance and diversity in experimental microcosms. Nature 408, 961-964.

Burdon, J., Chilvers, G., 1982. Host density as a factor in plant disease ecology. Annual Review of Phytopathology 20, 143-166.

Chesson, P.L., Warner, R.R., 1981. Environmental variability promotes coexistence in lottery competitive systems. American naturalist, 923-943.

Chun, J., Lee, J.-H., Jung, Y., Kim, M., Kim, S., Kim, B.K., Lim, Y.-W., 2007. EzTaxon: a web-based tool for the identification of

prokaryotes based on 16S ribosomal RNA gene sequences. International journal of systematic and evolutionary microbiology 57, 2259-2261.

Clark, D.J., Maaløe, O., 1967. DNA replication and the division cycle in Escherichia coli. Journal of molecular biology 23, 99-112.

Connell, J.H., 1978. Diversity in tropical rain forests and coral reefs. Science 199, 1302-1310.

Connell, J.H., Slatyer, R.O., 1977. Mechanisms of succession in natural communities and their role in community stability and organization. American naturalist, 1119-1144.

Cooper, S., 2012. Bacterial growth and division: biochemistry and regulation of prokaryotic and eukaryotic division cycles. Elsevier.

Cooper, S., Helmstetter, C.E., 1968. Chromosome replication and the division cycle of Escherichia coliBr. Journal of molecular biology 31, 519-540.

Davies, M., Smith, K., Vinten, A., 2001. The mineralisation and fate of nitrogen following ploughing of grass and grass-clover swards. Biology and Fertility of Soils 33, 423-434.

Eisenberg, E., Levanon, E.Y., 2013. Human housekeeping genes, revisited. Trends in Genetics 29, 569-574.

Fantes, P., Nurse, P., 1977. Control of cell size at division in fission yeast by a growth-modulated size control over nuclear division. Experimental cell research 107, 377-386.

Fierer, N., Bradford, M.A., Jackson, R.B., 2007. Toward an ecological classification of soil bacteria. Ecology 88, 1354-1364.

Fierer, N., Schimel, J., Holden, P., 2003. Influence of drying-rewetting frequency on soil bacterial community structure. Microbial ecology 45, 63-71.

Gensel, P.G., 2008. The earliest land plants. Annual Review of Ecology, Evolution, and Systematics, 459-477.
Gilbert, G.S., 2002. Evolutionary ecology of plant diseases in natural ecosystems. Annual Review of Phytopathology 40, 13-43.
Griffiths, B., Ritz, K., Bardgett, R.D., Cook, R., Christensen, S., Ekelund, F., Sørensen, S.J., Bååth, E., Bloem, J., De Ruiter, P., 2000. Ecosystem response of pasture soil communities to fumigation-induced microbial diversity reductions: an examination of the biodiversity–ecosystem function relationship. Oikos 90, 279-294.
Grime, J.P., 2006. Plant strategies, vegetation processes, and ecosystem properties. John Wiley & Sons.
Guariguata, M.R., Ostertag, R., 2001. Neotropical secondary forest succession: changes in structural and functional characteristics. Forest ecology and management 148, 185-206.
Gunderson, L.H., 2000. Ecological resilience--in theory and application. Annual review of ecology and systematics, 425-439.
Harshey, R.M., 1994. Bees aren't the only ones: swarming in Gram-negative bacteria. Molecular microbiology 13, 389-394.
Harshey, R.M., 2003. Bacterial motility on a surface: many ways to a common goal. Annual Reviews in Microbiology 57, 249-273.
Henrichsen, J., 1972. Bacterial surface translocation: a survey and a classification. Bacteriological Reviews 36, 478.
Holling, C.S., 1996. Engineering resilience versus ecological resilience. Engineering within ecological constraints, 31-44.
Horn, H.S., 1974. The ecology of secondary succession. Annual review of ecology and systematics, 25-37.
Horvath, P., Barrangou, R., 2010. CRISPR/Cas, the immune system of bacteria and archaea. Science 327, 167-170.

Howe, H.F., Smallwood, J., 1982. Ecology of seed dispersal. Annual review of ecology and systematics 13, 201-228.

Huston, M., 1979. A general hypothesis of species diversity. American naturalist, 81-101.

Ikeda, H., 2003. Testing the intermediate disturbance hypothesis on species diversity in herbaceous plant communities along a human trampling gradient using a 4-year experiment in an old-field. Ecological Research 18, 185-197.

Kang, S., Mills, A.L., 2004. Soil bacterial community structure changes following disturbance of the overlying plant community. Soil Science 169, 55-65.

Kim, M., Heo, E., Kang, H., Adams, J., 2013. Changes in soil bacterial community structure with increasing disturbance frequency. Microbial ecology 66, 171-181.

Kopylova, E., Noé, L., Touzet, H., 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28, 3211-3217.

Langenheder, S., Székely, A.J., 2011. Species sorting and neutral processes are both important during the initial assembly of bacterial communities. The ISME journal 5, 1086-1094.

Leininger, S., Urich, T., Schloter, M., Schwark, L., Qi, J., Nicol, G., Prosser, J., Schuster, S., Schleper, C., 2006. Archaea predominate among ammonia-oxidizing prokaryotes in soils. Nature 442, 806-809.

Lindén, B., Wallgren, B., 1993. Nitrogen mineralization after leys ploughed in early or late autumn. Swedish Journal of Agricultural Research (Sweden).

Loveless, M.D., Hamrick, J.L., 1984. Ecological determinants of genetic structure in plant populations. Annual review of ecology and systematics, 65-95.

Lubchenco, J., Menge, B.A., 1978. Community-Development and Persistence in a Low Rocky Inter-Tidal Zone. Ecological Monographs 48, 67-94.

MacArthur, R.H., Wilson, E.O., 2015. Theory of Island Biogeography.(MPB-1). Princeton University Press.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., 2008. The metagenomics RAST server-a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC bioinformatics 9, 1.

Molino, J.-F., Sabatier, D., 2001. Tree diversity in tropical rain forests: a validation of the intermediate disturbance hypothesis. Science 294, 1702-1704.

Myster, R.W., Pickett, S., 1992. Dynamics of associations between plants in ten old fields during 31 years of succession. Journal of Ecology, 291-302.

Naeem, S., Li, S., 1997. Biodiversity enhances ecosystem reliability. Nature 390, 507-509.

Nathan, R., Muller-Landau, H.C., 2000. Spatial patterns of seed dispersal, their determinants and consequences for recruitment. Trends in ecology & evolution 15, 278-285.

Oberholzer, M., Lopez, M.A., McLelland, B.T., Hill, K.L., 2010. Social motility in African trypanosomes. PLoS Pathog 6, e1000739.

Odum, E.P., 1969. The strategy of ecosystem development. Sustainability: Sustainability 164, 58.

Overbeek, R., Begley, T., Butler, R.M., Choudhuri, J.V., Chuang, H.-Y., Cohoon, M., de Crécy-Lagard, V., Diaz, N., Disz, T., Edwards, R., 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic acids research 33, 5691-5702.

Paine, R.T., Levin, S.A., 1981. Intertidal landscapes: disturbance and the dynamics of pattern. Ecological monographs 51, 145-178.

Rashid, M.H., Kornberg, A., 2000. Inorganic polyphosphate is needed for swimming, swarming, and twitching motilities of Pseudomonas aeruginosa. Proceedings of the National Academy of Sciences 97, 4885-4890.

Roesch, L.F., Fulthorpe, R.R., Riva, A., Casella, G., Hadwin, A.K., Kent, A.D., Daroub, S.H., Camargo, F.A., Farmerie, W.G., Triplett, E.W., 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. The ISME journal 1, 283-290.

Schaechter, M., Maaløe, O., Kjeldgaard, N.O., 1958. Dependency on medium and temperature of cell size and chemical composition during balanced growth of Salmonella typhimurium. Microbiology 19, 592-606.

Schleper, C., Jurgens, G., Jonuscheit, M., 2005. Genomic studies of uncultivated archaea. Nature Reviews Microbiology 3, 479-488.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and environmental microbiology 75, 7537-7541.

Shade, A., Read, J.S., Youngblut, N.D., Fierer, N., Knight, R., Kratz, T.K., Lottig, N.R., Roden, E.E., Stanley, E.H., Stombaugh, J., 2012.

Lake microbial communities are resilient after a whole-ecosystem disturbance. The ISME journal 6, 2153-2167.

Shapiro, J.A., 1998. Thinking about bacterial populations as multicellular organisms. Annual Reviews in Microbiology 52, 81-104.

She, X., Rohl, C.A., Castle, J.C., Kulkarni, A.V., Johnson, J.M., Chen, R., 2009. Definition, conservation and epigenetics of housekeeping and tissue-enriched genes. BMC genomics 10, 1.

Solomon, M., 1949. The natural control of animal populations. The Journal of Animal Ecology, 1-35.

Sousa, W.P., 1984. The role of disturbance in natural communities. Annual review of ecology and systematics 15, 353-391.

Suttle, C.A., 2007. Marine viruses—major players in the global ecosystem. Nature Reviews Microbiology 5, 801-812.

Team, R.C., 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna,

Austria. Version 3.1. 2. 2014. Online: http://www. R-project. org.

Tilman, D., 1994. Competition and biodiversity in spatially structured habitats. Ecology 75, 2-16.

Tilman, D., Reich, P.B., Knops, J., Wedin, D., Mielke, T., Lehman, C., 2001. Diversity and productivity in a long-term grassland experiment. Science 294, 843-845.

Tilman, D., Reich, P.B., Knops, J.M., 2006. Biodiversity and ecosystem stability in a decade-long grassland experiment. Nature 441, 629-632.

Trosvik, P., Stenseth, N.C., Rudi, K., 2010. Convergent temporal dynamics of the human infant gut microbiota. The ISME journal 4, 151-158.

Velicer, G.J., Yuen-tsu, N.Y., 2003. Evolution of novel cooperative swarming in the bacterium Myxococcus xanthus. Nature 425, 75-78.

Verstraeten, N., Braeken, K., Debkumari, B., Fauvart, M., Fransaer, J., Vermant, J., Michiels, J., 2008. Living on a surface: swarming and biofilm formation. Trends in microbiology 16, 496-506.

Walker, L.R., Chapin, F.S., 1987. Interactions among processes controlling successional change. Oikos, 131-135.

Walsh, D.A., Papke, R.T., Doolittle, W.F., 2005. Archaeal diversity along a soil salinity gradient prone to disturbance. Environmental Microbiology 7, 1655-1666.

Wang, H., Guo, S., Huang, M., Thorsten, L.H., Wei, J., 2010. Ascomycota has a faster evolutionary rate and higher species diversity than Basidiomycota. Science China Life Sciences 53, 1163-1169.

Wertz, S., Degrange, V., Prosser, J.I., Poly, F., Commeaux, C., Freitag, T., Guillaumaud, N., Roux, X.L., 2006. Maintenance of soil functioning following erosion of microbial diversity. Environmental Microbiology 8, 2162-2169.

Wertz, S., Degrange, V., Prosser, J.I., Poly, F., Commeaux, C., Guillaumaud, N., Le Roux, X., 2007. Decline of soil microbial diversity does not influence the resistance and resilience of key soil microbial functional groups following a model disturbance. Environmental Microbiology 9, 2211-2219.

Xiang, X., Shi, Y., Yang, J., Kong, J., Lin, X., Zhang, H., Zeng, J., Chu, H., 2014. Rapid recovery of soil bacterial communities after wildfire in a Chinese boreal forest. Scientific reports 4.

Zhang, J., Zhao, H., Zhang, T., Zhao, X., 2002. Niche dynamics of main populations of plants communities in the restoring succession

process in Horqin Sandy Land. Acta Ecologica Sinica 23, 2741–2746.

Zhu, J., He, F., Song, S., Wang, J., Yu, J., 2008. How many human genes can be defined as housekeeping with current expression data? BMC genomics 9, 1.

국문초록 (Abstract in Korean)

지금까지 토양 생물군의 교란에 대한 반응은 정밀하게 통제된 실험과 DNA 에 기초한 방법으로는 많이 연구되지 않았다. 본 연구에서는 토양 microcosm 에 인위적인 교란을 준 후 분류학상의 구성, 유전자의 기능과 발현의 상관 관계를 조사하였다. 이를 위해 총 24 주 동안 시간 흐름에 따른 교란 이후의 반응과 교란의 빈도에 따른 토양 군집의 회복과 영향을 알아보는 실험이 진행되었다.

먼저 토양 microcosm 시스템에 90 %의 멸균 토양을 섞는 초기 교란을 준 후 시간차를 두고 6 번의 샘플 채취를 통해 일련의 시간 흐름에 따른 토양 군집의 회복을 알아보는 실험이 진행되었다. 토양 DNA 는 24 주 동안 일정한 시간 간격을 두고 추출되었으며 각 샘플의 전체 메타지놈 데이터는 NextSeq 을 통해 시퀀싱되었다. 초기 조건과 비교해보았을 때, 토양 미생물 군집의 구조와 기능에서 시간 경과에 따라 일정하고 순차적인 변화가 관찰되었고 이는 시간에 따른 생태 자리(niche)의 변화를 보여주었다. 예상했던 것처럼 초기 천이 과정에 있는 샘플들은 운동성과 관련된 기능을 가진 유전자들이 상대적으로 많이 발견되었고 DNA/RNA/단백질 대사, 세포 분열과 세포 주기와 관련된 유전자들은 적게 발견되었다. 그러나 세포 신호 전달, 독성과 그 방어에 관련된 유전자들은 시간의 흐름에 유의미한 변화를 보이지 않았고, 환경 스트레스와 관련된 유전자들도 천이 과정 말기로 갈수록 상대적으로 줄어들었다. 초기 천이 과정에서는 낮은 분류학적 다양성을

가지고 있었지만 유전자 기능의 다양성은 높게 관찰되었다. 시간이 지남에 따라 관찰된 특징들은 말기에 그 변화 폭이 줄어들기는 하지만 교란을 주기 전인 초기 조건으로 돌아가는 방향성을 보였다.

두 번째로, 토양에 24 주 동안 서로 다른 빈도로 교란을 주는 실험을 진행하였다. 토양 DNA 는 5 단계의 교란 빈도 차이를 가진 샘플들에서 추출되었고 전체 메타지놈 데이터는 Nextseq 을 통해 시퀀싱되었다. 토양 미생물 군집의 다양성은 교란 빈도가 증가함에 따라 감소하였지만 기능 유전자들의 다양성은 교란 빈도에 따라 증가하였다. 관찰된 특징들은 교란 빈도에 영향을 받아 변화를 나타냈으며 변화 폭은 점차 줄어들었지만 낮은 빈도의 교란을 준 샘플일수록 초기 조건과 유사한 구성을 가졌다. 예상한 것처럼 교란 빈도에 따라 차이가 발생했으며 높은 교란을 경험한 군집은 DNA/RNA 대사, 운동성, 세포 분열과 세포 주기와 관련된 유전자가 상대적으로 높게 관찰되었다. 그러나, 예상과는 다르게 단백질 대사와 스트레스에 관련된 유전자의 경우에는 낮은 교란 빈도를 준 군집에서 더 많이 발견되었다.

이 결과는 토양 생태계 교란 이후 시간에 따른 변화와 교란 빈도에 따라 유전자 기능과 생물 분류군의 예상 가능한 변화를 나타내고 있으며, 큰 생물체에서 적용되었던 생태학적 천이의 양식과 비교되었다.

주요어 : 교란, 천이, 토양미생물, 메타지노믹스, 미생물 군집, 미생물 생태학

학번 : 2015-20449