Science of the Total Environment 628–629 (2018) 94–102

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Microbial functional gene patterns related to soil greenhouse gas emissions in oil contaminated areas

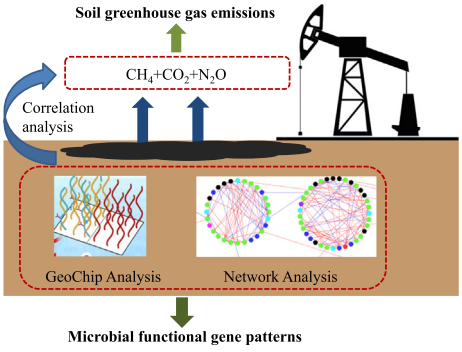
Juejie Yang, Guanghe Li, Yi Qian, Yunfeng Yang, Fang Zhang *

School of Environment and State Key Joint Laboratory of Environment Simulation and Pollution Control, Tsinghua University, Beijing 100084, China

HIGHLIGHTS

- Revealing the functional gene pattern response to oil contamination
- Analyzing the interactions between species in microbial cooperating network
- Identifying keystone genes related to soil GHG emissions under oil contamination
- Linking the microbial functional gene pattern to soil GHG emissions

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 3 December 2017
Received in revised form 1 February 2018
Accepted 1 February 2018
Available online xxxx

Editor: Zhen (Jason) He

Keywords:
Soil carbon dioxide
Methane
Nitrous oxide
Microbial functional gene
Molecular ecological network
Oil contamination

ABSTRACT

Linking microbial community structure to physiology and ecological processes is a critical focus of microbial ecology. To understand the microbial functional gene patterns related to soil greenhouse gas [carbon dioxide (CO_2), methane (CH_4) and nitrous oxide (N_2O)] emissions under oil contamination, we used functional gene array (GeoChip 5.0) analysis and network methods to investigate the feedback responses of soil microbial functional gene patterns and identify keystone genes in Shengli Oilfield, China. The microbial functional gene number, relative abundance and diversity involved in carbon degradation and nitrogen cycling decreased consistently with the reduced CO_2 and N_2O flux in oil contaminated soils, whereas the gene number and relative abundance of methane-production related genes increased with contamination. Functional molecular ecological networks were built based on random matrix theory, where network structures and properties showed significantly variation between oil contaminated and uncontaminated soils ($P < 0.05$). Network nodes, connectivity and complexity all reduced under oil contamination. The sensitive and the highest connective genes in the network were identified as keystone genes, based on Mann-Whitney *U* tests and network analysis. Our findings improved the understanding of the microbe-mediated mechanisms affecting soil greenhouse gas emissions.

1. Introduction

Soil contamination caused by crude oil exploration and related activities has drawn increasing attention worldwide. Oil contaminants could alter soil physical and chemical properties, influence terrestrial ecosystem function such as soil carbon and nitrogen cycling, and release large amount of CH_4 or CO_2 during crude oil biodegradation (Marin et al.,

* Corresponding author.
E-mail addresses: yang-jj13@mails.tsinghua.edu.cn (J. Yang), ligh@tsinghua.edu.cn (G. Li), qiany@tsinghua.edu.cn (Y. Qian), yangyf@tsinghua.edu.cn (Y. Yang), fangzhang@tsinghua.edu.cn (F. Zhang).

https://doi.org/10.1016/j.scitotenv.2018.02.007
0048-9697/

2005; Mayumi et al., 2013; Sihota et al., 2010). Soil greenhouse gas (GHG) emissions as a result of microbiological processes have also been affected by oil contamination (Agegnehu et al., 2016; Siddique et al., 2011), and our recent research found appreciably increased *in situ* soil CH_4 emissions and decreased CO_2 emissions from oil contaminated areas (Yang et al., 2017). Microorganisms in oilfields had less carbon source types to utilize and showed lower metabolic activity than those in uncontaminated areas (Zhen et al., 2015). Although recent studies have investigated the influence of oil contamination on microbial community involved in carbon degradation and nitrogen utilization (Hu et al., 2015; Trivedi et al., 2013), the microbial mechanism of soil GHG emissions in oil contaminated area remained mostly unclear.

Revealing the functional gene pattern and identifying keystone genes related to soil GHG emissions under oil contamination could improve our understanding of microbial mechanism in this ecological process. However gene pattern characterization and keystone gene identification are challenging, because of high diversity of microbial communities and uncultivated status of most microbes (Montoya et al., 2006; Morrison et al., 2016; Zhou et al., 2010). Available studies to date have demonstrated that information provided by functional gene array analysis and network methods was critical for comprehensive analysis of microbial mechanisms in major ecological processes (Coyte et al., 2015; Yue et al., 2015; Zhou et al., 2010). Some studies have proposed that species sensitive to environment change and highly connective in networks both referred to "keystone species/genes" (Deng et al., 2016; Hartmann et al., 2012), but keystone genes related to soil GHG emissions in oil contaminated areas have not been thoroughly investigated.

Understanding the interactions among species/populations and their influence on ecological process are key issues in ecology (Montoya et al., 2006). The random matrix theory (RMT)-based network method is one of the commonly used correlation-based relevance network methods to examine the complex interactions among microbes and identify keystone species (Steele et al., 2011; Wu et al., 2016). Related approach has been applied to show the changes of functional gene expression in soil microbial communities with elevated atmospheric CO_2 conditions, or during groundwater uranium bioremediation (Deng et al., 2012; Deng et al., 2016; Zhou et al., 2011). A similar approach has been used to investigate interactions among root-associated microbes in greenhouse microcosms (Shi et al., 2016). Microbial networks were constructed in oil contaminated areas using functional genes involved in alkane and polycyclic aromatic hydrocarbon degradation to investigate the interaction of indigenous microbes (Liang et al., 2016). Therefore, this method provides a useful tool for the study of soil functional genes related to GHG emissions with the impact of oil contamination.

Here, in order to understand the patterns of microbial functional genes and their correlations with soil GHG emissions in oil contaminated areas, we conducted field sampling and measurements in Shengli Oilfield in Yellow River Delta (YRD). We applied a comprehensive functional gene array (GeoChip 5.0) and random matrix theory (RMT)-based network method to compare composition and correlation networks of microbial functional genes between the contaminated and uncontaminated soils. *In situ* CO_2, CH_4 and N_2O emissions and the alteration against the uncontaminated soils were measured near oil wells for the correlation analysis between microbial factors and soil GHG emissions. Potential keystone genes were further identified based on the sensitivity under oil contamination and highest connectivity in the network.

2. Materials and methods

2.1. Study sites and soil sample collections

The soil and gas samples were collected for use from Shengli Oilfield in Yellow River Delta (YRD), located in Shandong Province on the Bohai coast of Northern China (36°55′–38°12′ N, 118°07′–119°18′ E; Fig. S1). The region is dominated by salt tolerant species with mean annual air temperature of 11.9 °C and rainfall precipitation of 596.2 mm (Nie et al., 2009). The main soil type is sandy clay loam, with little impact of tide (Xie et al., 2014). Contaminated soil samples were collected near oil wells without vegetation in which contamination occurred several years ago. The total petroleum hydrocarbons (TPHs) were detected only in 0–20 cm throughout the soil profile of 0–100 cm. The range of TPH in soil samples among oil wells were 100–3634 mg kg^{-1}. Crude oil had been fully dispersed in soils and oil contamination could not be identified by eyes. Bare soils were exposed due to oil exploitation related activities and were likely to be contaminated. Therefore uncontaminated soil samples (no TPH detected) were collected from undisturbed pristine soils >20 m away from the oil wells. The dominant aboveground vegetation was *Phragmites australis*, which was manually removed before soil sampling. These soil samples (14 contaminated soil samples and 8 uncontaminated samples) were collected from the surface (0–20 cm) in August 2015 and used for soil property analysis and microbial functional gene (GeoChip 5.0) experiments. Soil samples for GeoChip analysis were kept at −80 °C until DNA extraction.

The soil redox potential (ORP), water content, and pH were *in situ* measured with multiple electrodes (FJA-6 Portable ORP Meter, China) and the time-domain reflectometry (TDR). The soil ORP (440–520 mV) and water content (20.3–28.8%) had no significant differences between contaminated and uncontaminated soils (Table S1). The pH was always >7.9 in oil contaminated soils, higher than those of the uncontaminated soils (7.5). The soil bulk density was analyzed by metal cylinders (5 cm in diameter, 5 cm in length) and used to estimate the soil porosity (Sasal et al., 2006). Other soil physical and chemical properties, including soil texture, total carbon (TC), organic matter (OM), total nitrogen (TN), total phosphorus (TP) and electrical conductivity (EC) were measured according to the recommended soil testing procedures (Lu, 1999). The soil TC (1.1–2.9%) and porosity (40.8–52.5%) had no significant differences between contaminated and uncontaminated soils. Higher values of OM and EC were observed in oil contaminated soils than those of the uncontaminated soils ($P < 0.05$). However, the TN and TP of oil contaminated soils were significantly lower than those of the uncontaminated soils ($P < 0.05$, Table S1).

2.2. Soil gas collection and concentration measurements

In situ soil GHG fluxes were measured on the daily basis in sunny days in August 2015 and average values were used for analysis. Gas samples (250 mL) were collected in stainless steel square boxes (0.5 m × 0.5 m × 0.3 m) (without a top or bottom) from 11:00 am at time intervals of 0, 5, 10, 15, 20, and 25 min both in the oil contaminated and uncontaminated soils. All samples were taken to the lab within 24 h of collection. Gas samples were selected to determine the CO_2, CH_4 and N_2O concentrations using a gas chromatograph (7890A GC, Agilent, Santa Clara, CA, USA, fitted with a flame ionization detector and a thermal conductivity detector for CO_2, CH_4 and N_2O). The differences in soil gas fluxes from contaminated and uncontaminated soils were analyzed by independent samples *t*-test using the SAS software (SAS Institute Inc., Cary, NC, USA).

2.3. DNA extraction and GeoChip analysis

All samples (14 contaminated samples and 8 uncontaminated samples) were analyzed for microbial functional gene patterns using GeoChip 5.0. GeoChip 5.0 contains over 57,000 oligonucleotide probes, covering over 144,000 gene sequences from 393 functional gene families involved in carbon, nitrogen, sulfur, phosphorus cycling and others (Wang et al., 2014). DNA was extracted from 5 g of each soil sample by freeze-grinding method as previously described (Zhou et al., 1996). Agarose gel electrophoresis was used for the purification of crude DNA followed by successive extractions with phenol, chloroform and butanol. DNA purity was determined by an ND-1000 spectrophotometer (Nanodrop Inc.,