Original Research Article

An improved protocol for isolation of soil microbial community DNA for metagenomic studies

Abstract

Aim: Several methods were described previously for isolation and purification of soil DNA. Most of these protocols use combination of techniques or methods but the effect of each method individually is not clear. This study aims at analysing the effect of individual components of soil DNA isolation on quantity and quality of soil DNA, and finally to optimize soil DNA isolation protocol and its validation by 16SrDNA sequence analysis.
Methods and Results: An attempt was made to study the effect of vortexig, heating, SDS and lysozyme on cell lysis. Different concentrations of chemical flocculants like $AlNH_4SO_4$, $FeCl_3$, $CaCl_2$ and $MgCl_2$ were used to reduce humic substances. Pre-lysis washing of soil with 100 mmol l^{-1} disodium EDTA proved good for releasing microbial cells from soil matrix. Heating the soil sample at 75 °C yielded good quantity DNA followed by 2 % SDS and vertexing at 1400 g. Flocculation with 100 mmol l^{-1} $CaCl_2$ reduces significant amount of humic substances in isolated DNA.
Conclusion: 16SrDNA sequencing of soil DNA indicates that this protocol is unbiased and can extract good quality and quantity DNA from range of microbes varying in their cell wall composition.
Significance: The optimised protocol is unbiased, very simple, does not need special equipments and many samples can be processed simultaneously.

Keywords: Metagenomics, Soil DNA, Humic acid, PCR, Protocol, Sequencing.

Comment [D1]: Check spelling vortexing

Introduction

Soils are highly complex environment containing diverse microbial species. A gram of soil can contain between 10^5 and 10^6 different bacterial species (Bates et al, 2011). Traditionally, microbes have been isolated, purified and characterized by their morphological and biochemical properties (Maron et al, 2011). Due to limited information about nutrient requirement and culture conditions, it is now known that less than 1% of the microbial species can be cultured in laboratory (Torsvik and Ovreas, 2002). Analysis of nucleic acid extracted directly from environmental sample allows the researcher to study both culturable and non-culturable microbes. Getting good quality and quantity DNA from environmental sample is a fundamental step for metagenomic studies. Many efforts have been made to optimize DNA extraction procedure from soil. Soil DNA isolation protocols are broadly classified into direct and indirect lysis. Indirect lysis involves separation of cells from soil matrix followed by cell lysis and DNA extraction (Holben et al, 1988). Cell lysis in the soil matrix followed by separation and purification of DNA from matrix and cell debris forms the direct lysis method (Ogram et al, 1987).

Releasing of microbial cells from soil particle, cell lysis and purification of soil DNA from contaminants like humic substances are critical and challenging. The procedures for soil DNA isolation belong to physical (Zhou et al, 1996 and Kuske et al, 1998), chemical (Tebbe and Vahjen, 1993 and Holben, 1994) and enzymatic (Malik et al, 1994 and DeGrange and Bardin, 1995) lysis of cell followed by purification using density centrifugation (Leff et al,

UNDER PEER REVIEW

995), column chromatography (Herrick et al, 1996) and chemical flocculation (Braid et al, 2003). Because of lack of comparative studies, the efficiency of each method for cell lysis and purification is not clearly understood. An attempt was made to compare different methods for releasing cells from soil matrix, lysis of cell and removal of humic substances from soil DNA. An optimised protocol was demonstrated for its suitability for metagenomic studies by high throughput 16SrDNA sequencing of soil metagenome.

Material and methods

Pre-lysis wash

To release microbial cells from soil matrix and to reduce the humic substances extracted with DNA, the soil sample was washed with different concentrations of sodium phosphate buffer (pH 7.0) and disodium EDTA (pH 7.0). Two hundred milligram of soil was mixed with 1 ml of 0, 20, 40, 60, 80, 100, 120, 140, 160 and 180 mmol l^{-1} sodium phosphate buffer (pH 7.0) and disodium EDTA (pH 7.0) separately and gently shaken for 1h in orbital incubator shaker at room temperature. The sample was centrifuged at 10,000 g for 5 minutes and the supernatant was collected. Quantity of humic acid was measured in spectrophotometer at 320 nm with humic acid sodium salt as standard (Arbeli and Fuentes, 2007, Olson and Morrow, 2012. The soil pellet was processed further to extract DNA.

Cell lysis

Physical, chemical and enzymatic methods and their combinations were used to lyse the microbial cells in soil. One ml of soil DNA isolation buffer (100 mmol l^{-1} Tris, 100 mmol l^{-1} Na_2EDTA, 1.5 mol l^{-1} Nacl and 1.25 % PVPP) was mixed with 200 mg soil and vortexed at 1400 g at 37°C for 0, 30, 60, 90, 120 and 180 minutes on thermomixer (Eppendorf). For heating, the sample was incubated at 75 °C for 0, 30, 60, 90, 120 and 180 minutes in water bath. 1, 2, 3, 4, 5 and 6 % sodium dodecyl sulphate (SDS) was used for chemical cell lysis in DNA isolation buffer at 65 °C temperature. 0, 25, 50, 75, 100, 125 and 150 mg lysozyme was used for enzymatic lysis at 37 °C.

Removal of humic substances

Chemical flocculation using aluminium ammonium sulphate ($AlNH_4SO_4$), calcium chloride ($CaCl_2$), ferric chloride ($FeCl_3$) and magnesium chloride ($MgCl_2$) were used to remove humic substances from the soil DNA. Concentrations ranging from 0 to 180 mmol l^{-1} in increments of 20 mmol l^{-1} was added in lysis buffer (100 mmol l^{-1} Tris, 100 mmol l^{-1} disodium EDTA, 1.5 mol l^{-1} NaCl and 1.25 % PVPP).

After treatment, the supernatant was mixed with equal volume of choloroform isoamyl alcohol (24:1) and centrifuged at 13,200 g. Aqueous layer was transferred to fresh tube, mixed with equal volume of chilled isopropanol and centrifuged at 13,200 g for 10 minutes at 4 °C.

Brief protocol

Two hundred mg soil (clay loam with 0.41 % organic carbon, pH 7.55 and 58.59 me.100g^{-1} CEC) was mixed with 1 ml of 100 mmol l^{-1} disodium EDTA, shaken for 1 h at room temperature and centrifuged at 10,000 g for 5 minutes. The supernatant was discarded, the soil pellet was dissolved in 1 ml of DNA extraction buffer (100 mmol l^{-1} Tris, 100 mmol l^{-1} Na_2EDTA, 1.5 mol l^{-1} Nacl and 1.25 % PVPP) containing 2 % SDS and 100 mmol l^{-1} $CaCl_2$ and vortexed at 1400 g for 1 h 30 minutes at 75 °C. The sample was centrifuged at 13,200 g at room temperature for 10 minutes, supernatant was transferred to fresh 2 ml centrifuge tube, mixed with equal volume of chloroform isoamyl alcohol (24:1) and centrifuged at 13,200 g.

UNDER PEER REVIEW

The upper layer was then transferred to 1.5 ml centrifuge tube, $1/10^{th}$ volume of sodium acetate (pH 7.5) and equal volume of chilled isopropanol was added, mixed and incubated overnight at -20°C. The DNA was collected by centrifugation at 13,200 g at 4 °C for 10 minutes, washed with 70 % alcohol, dried and dissolved in 50 µl of $T_{10}E_1$.

The amount of humic substances was measured using Cary 50 Bio UV-spectrophotometer at 320 nm with humic acid sodium salt as standard (Arbeli and Fuentes, 2007, Olson and Morrow, 2012). DNA was analysed by agarose electrophoresis followed by documentation in Synegene G box gel documentation unit and quantified by Syngene Genetool software using Lambda DNA *HindIII* single digest marker as reference. The effect of treatments used in pre-lysis wash, cell lysis and removal of humic substances was statistically analysed by t test.

PCR amplification

PCR amplification and restriction digestion was performed to check for the purity of the isolated metagenomic DNA. 16S rDNA was amplified using PRBA338 and PRUN518 primers (Nakatsu *et. al*, 2000). Each PCR reaction contained 1X PCR buffer, 1.2 mmol l^{-1} $MgCl_2$, 250 µmoles of each dNTP, 5 µM of each primer, 1 unit *Taq* DNA polymerase, and 100 ng template DNA. To check the inhibitory activity of humic acid, unpurified and purified DNA was used as template. Purified DNA was four fold diluted to check the effect of traces of humic acid on *Taq* DNA polymerase. The template DNA was denatured at 95 °C for 5 minutes followed by 32 cycles of denaturation at 94 °C for 50 seconds, primer annealing at 55 °C for 30 seconds and extension at 72 °C for 50 seconds. The amplified product was checked on 1.2 % agarose gel.

Restriction digestion

Five hundered ng of metagenomic DNA was digested by 5 units of *EcoR*I and *Hind*III restriction endonuclease at 37 °C for 2 hours under optimum conditions of buffer and pH. The digested DNA was separated in 0.7 % agarose gel along with undigested control, stained with ethidium bromide and documented in Syngene gel documentation unit.

Sequencing of 16SrDNA

Suitability of the above soil DNA isolation method for metagenomic studies was analysed by high throughput sequencing of 16S rDNA. Amplification of 16S rDNA was done as mentioned above. An aliquot (2µl) of amplified product was checked on 1.2 % agarose gel. The remaining amplified product was purified and sequenced using 316 chip on Ion torrent personal genome machine according to the manufacturer protocol. The Q20 reads generated by Ion torrent sequencer were deposited in MG-RAST database (Accession number 4564423). The quality sequence were phylogenetically classified using M5RNA annotation source of MG-RAST (Meyer *et al*, 2008). Species richness and their diversity were calculated by Species Prediction and Diversity Estimation (SPADE) (Chao and Shen, 2010) with cut off value of 10 and 200 bootstrapping. The species (based on sequence reads) were searched manually for Gram reaction and classified as Gram-positive, Gram-negative and Gram variable. The species whose Gram reaction information is not available were classified as others.

Results

Sodium phosphate and disodium EDTA was found to extract 7.33 and 24.40 mg humic acid respectively from a gram of soil (Table1). There was no additional humic acid extracted with increase in the concentration (above 180 mmol l^{-1}) of disodium EDTA. However, higher

UNDER PEER REVIEW

concentration (above 180 mmol l^{-1}) of sodium phosphate buffer resulted in the extraction of more humic substances (Table1). Along with humic substances, lot of debris were also precipitated when sodium phosphate buffer was used for washing the soil before cell lysis. Though the amount of humic subsances extraction increased with the increase in concentration of disodium EDTA, the difference was significant only up to 100 mmol l^{-1} (table 1). The amount of DNA extracted decreased drastically after 100 mmol l^{-1} of disodium EDTA used in pre-lysis wash step. Hence, we used only 100 mmol l^{-1} disodium EDTA for pre-lysis wash.

Table1. Effect of pre-lysis washing on DNA and humic acid concentration

Conc. Of Sod. Phosphate / Na_2EDTA used	Humic acid in supernatant (mg g^{-1} soil)		DNA concentration (µg g^{-1} soil)		Humic acid in DNA (mg g^{-1} soil)	
	Sodium phosphate	Disodium EDTA	Sodium phosphate	Disodium EDTA	Sodium phosphate	Disodium EDTA
00 mM	0.00	0.00	21.27	21.27	24.87	26.50
20 mM	1.50^{f}	8.03^{d}	19.73^{a}	27.63^{a}	23.23^{d}	17.17^{e}
40 mM	2.13^{e}	9.80^{c}	18.77^{a}	27.38^{a}	23.03^{d}	14.80^{d}
60 mM	2.83^{d}	10.83^{c}	17.23^{b}	27.03^{a}	22.10^{c}	11.87^{c}
80 mM	3.67^{c}	14.00^{b}	16.13^{b}	26.50^{a}	21.40^{c}	8.20^{b}
100 mM	4.13^{c}	20.53^{a}	14.57^{c}	25.10^{b}	20.17^{c}	5.37^{a}
120 mM	5.00^{b}	21.77^{a}	13.00^{c}	20.57^{c}	19.03^{b}	4.50^{a}
140 mM	6.10^{a}	22.33^{a}	12.40^{c}	17.00^{d}	18.50^{b}	3.93^{a}
160 mM	6.43^{a}	24.40^{a}	11.63^{c}	12.20^{e}	16.90^{a}	3.47^{a}
180 mM	7.33^{a}	23.80^{a}	10.17^{c}	7.10^{f}	16.13^{a}	2.70^{a}

Values followed by different letters are significantly different between treatments at $p \leq 0.01$, as calculated by t test

The highest yield of DNA was obtained by heating the soil at 75 °C followed by incubating the soil with SDS and vertexing at 1400 g. Significant difference in the yield of DNA was observed by heating at 75 °C up to 90 minutes, vertexing at 1400 g for 90 minutes (Table2) and treating with 2 % SDS (table3). Along with DNA some amount of humic substances were also precipitated and the amount of humic substances precipitated was more with heating (Table2 and Table3). Treatment of soil slurry with different concentrations of lysozyme did not yield DNA (Table3).

Table2. Effect of vertexing and heating on cell lysis and DNA yield

Incubation time (Minutes)	Vertexing at 1400 g at room temperature		Heating at 75 °C	
	Humic acid (mg g^{-1} soil)	DNA (µg g^{-1} soil)	Humic acid (mg g^{-1} soil)	DNA (µg g^{-1} soil)
0	3.33	0.00	3.45	0.00
30	3.51	1.53^{c}	5.83^{a}	6.72^{c}
60	3.73	5.13^{b}	7.54^{b}	9.48^{b}
90	3.94	8.94^{a}	9.29^{c}	15.73^{a}
120	4.16	9.20^{a}	9.80^{d}	17.64^{a}
150	4.44	9.44^{a}	10.34^{e}	18.48^{a}
180	4.81	9.89^{a}	10.49^{e}	19.09^{a}

UNDER PEER REVIEW

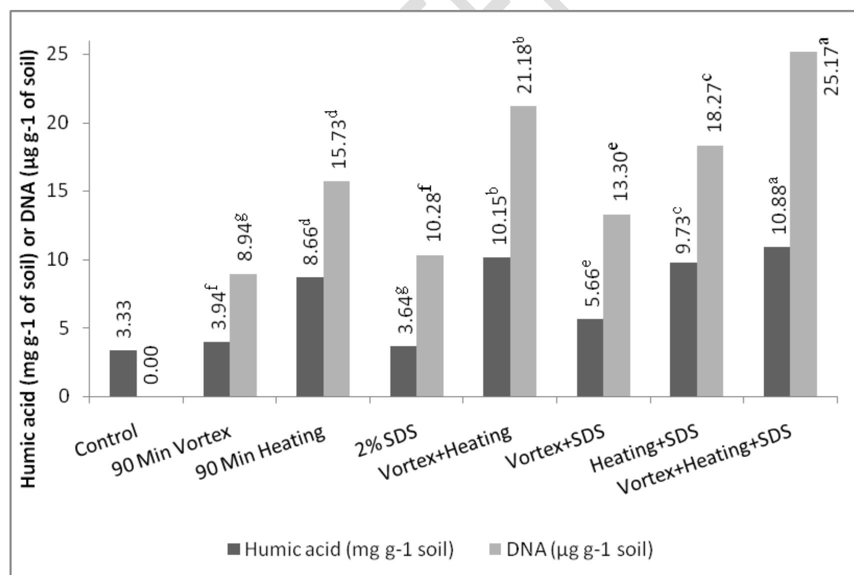
Values followed by different letters are significantly different between treatments at p ≤ 0.01, as calculated by t test

Table3. Effect of SDS and lysozyme on cell lysis and DNA yield

Sodium dodecyl sulphate (SDS)			Lysozyme		
Conc. of SDS (%)	Humic acid (mg g^{-1} soil)	DNA(µg g^{-1} soil)	Conc. of lysozyme (mg)	Humic acid (mg g^{-1} soil)	DNA(µg g^{-1} soil)
0.00	3.26	0.00	0.00	3.33	0.003
1.00	3.50	4.92[b]	25.00	3.85	0.005
2.00	3.64	10.28[a]	50.00	4.06	0.007
3.00	3.81	10.53[a]	75.00	3.94	0.006
4.00	4.00	10.71[a]	100.00	4.06	0.007
5.00	4.14	10.97[a]	125.00	4.10	0.008
6.00	4.26	11.12[a]	150.00	4.09	0.008

Values followed by different letters are significantly different between treatments at p ≤ 0.01, as calculated by t test

Vertexing at 1400 g for 90-minute, heating at 75 °C for 90 minute and incubation with 2 % SDS for 90 minutes yielded 8.94, 15.73 and 10.28 µg DNA g^{-1} of soil respectively with 3.94, 9.29 and 3.64 mg humic acid g^{-1} of soil respectively. The combination of three methods; vertexing at 1400 g with heating at 75 °C for 90 minutes in presence of 2 % SDS yielded the highest DNA (25 µg g^{-1} of soil) as compared to any of them alone and control (Fig1). However, large amount of humic substances (10.88 mg g^{-1} of soil) were also extracted with this method.



Figue1. Effect of combination of methods on cell lysis. Heating at 75°C with vertexing (1400 rpm) in presence of 2% SDS for 90 minutes yield highest DNA followed by heating with vertexing and heating with SDS. Heating also extracts lot of humic acid. Values followed by letters indicate significant difference as calculated by t test.

All the chemical flocculants used to remove humic substances also reduced the concentration of DNA. Aluminium ammonium sulphate ($AlNH_4SO_4$), magnesium chloride ($MgCl_2$) and ferric chloride ($FeCl_3$) reduces DNA concentration drastically without having much effect on the concentration of humic substances. Only calcium chloride ($CaCl_2$) yielded good concentration DNA with less of humic substances (Table4). There was significant reduction in humic substances even up to 180 mmol l^{-1} $CaCl_2$, but the concentration of DNA reduced drastically after 100 mmol l^{-1} $CaCl_2$ (Table4).

Table4. Effect of chemical flocculants on concentration of DNA and humic acid

Comment [D2]: Give space after Table

Concentration (mM)	$AlNH_4SO_4$		$CaCl_2$		$FeCl_3$		$MgCl_2$	
	Humic acid (mg g^{-1} soil)	DNA (µg g^{-1} soil)	Humic acid (mg g^{-1} soil)	DNA (µg g^{-1} soil)	Humic acid (mg g^{-1} soil)	DNA (µg g^{-1} soil)	Humic acid (mg g^{-1} soil)	DNA (µg g^{-1} soil)
0	5.200	25.73	5.207	25.51	4.840	25.43	5.130	25.64
20	4.260^{i}	17.81^{a}	3.450^{h}	23.04^{a}	4.270^{i}	18.15^{a}	5.010^{g}	18.01^{a}
40	3.620^{h}	12.46^{b}	2.067^{g}	22.19^{a}	3.680^{h}	10.78^{b}	4.880^{f}	13.54^{b}
60	3.410^{g}	7.88^{c}	0.117^{f}	21.24^{a}	3.400^{g}	1.64^{c}	4.720^{e}	8.96^{c}
80	3.040^{f}	4.23^{d}	0.075^{e}	21.22^{a}	3.050^{f}	0.52^{d}	4.560^{d}	5.02^{d}
100	2.740^{e}	2.43^{e}	0.007^{d}	19.93^{b}	2.850^{e}	0.32^{d}	4.450^{d}	3.15^{e}
120	2.620^{d}	2.04^{e}	0.003^{c}	15.01^{c}	2.710^{d}	0.18^{d}	4.350^{d}	1.87^{f}
140	2.440^{c}	1.34^{f}	0.002^{b}	10.36^{d}	2.440^{c}	0.15^{d}	4.170^{c}	1.05^{g}
160	2.110^{b}	0.93^{f}	0.001^{b}	5.15^{e}	2.080^{b}	0.09^{d}	4.020^{b}	0.63^{h}
180	1.510^{a}	0.60^{g}	0.001^{a}	1.00^{f}	1.640^{a}	0.06^{d}	3.880^{a}	0.42^{h}

Values followed by different letters are significantly different between treatments at $p \leq 0.01$, as calculated by t test

PCR amplification and restriction digestion (Fig2 and 3) of soil DNA purified with 100 mmol l^{-1} $CaCl_2$ indicates that the extracted DNA is pure enough for enzymatic activity and can be used for metagenomic analysis.

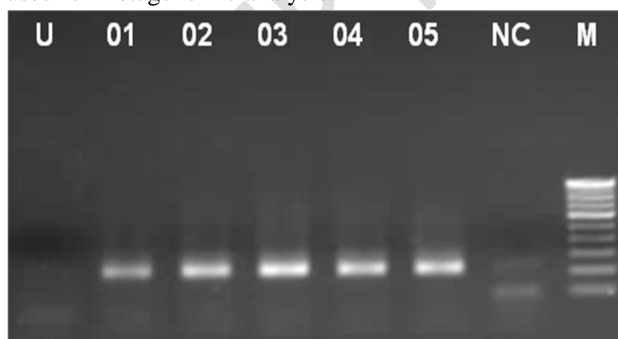


Figure2. PCR amplification of 16S rDNA from soil metagenomics. No amplification was observed in (lane U) DNA without using chemical flocculant, indicating humic acid inhibits *Taq* DNA polymerase activity. Lane 01 to 05 is purified DNA using CaCl2, 01, 02, 03, 04 and 05 are undiluted, one-fold, two-fold, three-fold and four-fold diluted DNA respectively. Amplification in all samples indicates the extracted DNA is pure enough for enzymatic activity. NC is no template control and M is 100bp DNA marker.

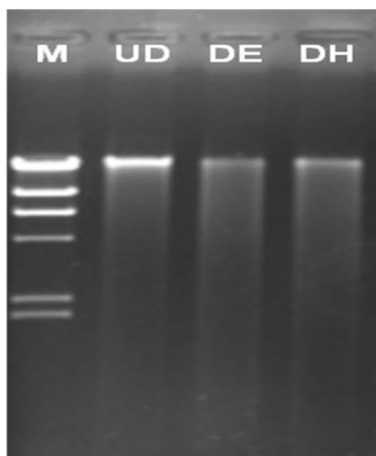


Figure3. Restriction digestion of metagenomic DNA. 5 unit of *EcoRI* (DE) and *HindIII* (DH) could able to digest 1 µg DNA in 3 hours indicating DNA is free from humic acid and can be used for metagenomic studies. UD is undigested DNA and M is lambda DNA HindIII single digest marker.

UNDER PEER REVIEW

High throughput sequencing of 16S rDNA amplified from isolated soil DNA identified 1558 species with the Shannon diversity index of 4.42 and 83 effective number of species. Around 35.7 % 16S rDNA sequences could not be classified to any of the known phylum of bacteria. Among classified sequences, 55.92, 30.92, 6.71, 3.35, 0.96, 0.64, 0.35, 0.27 and 0.24 % sequences belonged to *Proteobacteria, Actinobacteria, Fermicutes, Bacteroidetes, Cyanobacteria, Chloroflexi, Nitrospirae, Acidobacteria and Gemmatimonadetes* respectively (Table5).

Table5. Proportion of bacterial phyla in soil metagenome based on 16S r DNA sequencing

Phylum	Janssen, 2006		This study	
	Range (%)	Average (%)	With unclassified (%)	Without unclassified (%)
Proteobacteria	10-77	39.00	35.95	55.92
Actinobacteria	0-34	13.00	19.88	30.92
Acidobacteria	5-46	20.00	0.18	0.28
Verrucomicrobia	0-21	7.00	0.09	0.14
Bacteroidetes	0-18	5.00	2.15	3.35
Chloroflexi	0-16	3.00	0.41	0.64
Planctomycetes	0-8	2.00	0.03	0.03
Firmicutes	0-7	1.80	4.31	6.71
Gemmatimonadetes	0-4	2.00	0.15	0.24

Around 49.90 % and 46.73 % species belonged to Gram-positive and Gram-negative bacteria respectively with the ratio of 1.07 between Gram-positive and Gram-negative bacteria (Fig4).

Comment [D3]: Give space after Table

Comment [D4]: Use small letter (g)in case of Gram-positive or Gram-negative…check throughout the manuscript

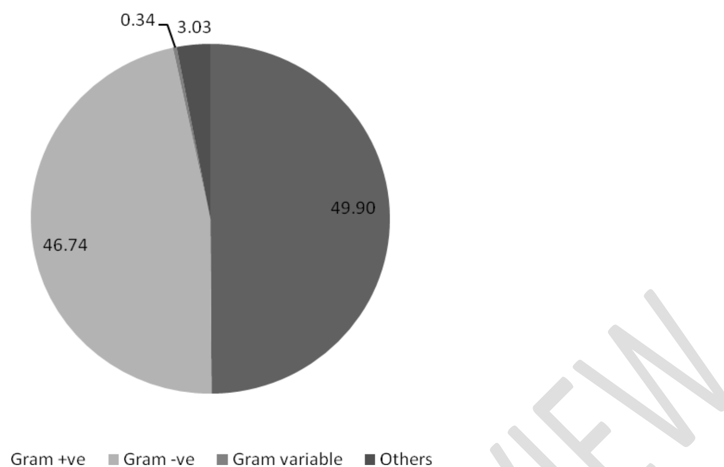


Figure4. Proportion of Gram-positive and Gram-negative bacteria in soil metgenome. 49.9 %, 46.74 %, 0.34 % and 3.03 % 16S r DNA sequences belong to Gram-positive, Gram-negative, Gram variable and others respectively. The ratio between Gram-positive and Gram-negative bacteria is 1.07.

Discussion

Good quality and quantity DNA is very much essential for the culture independent genomic analysis of environmental sample. There are several protocols to extract DNA from soil, but none of them analysed effect of individual methods and cannot be accepted as standard for different soil types and among scientific community (Frostegard *et al,* 1999 and Schneegurt *et al.,* 2003). Extraction of microbial community DNA from soil has many challenges. Soil texture and its organic matter content pose serious problem to the extraction of microbial community DNA from soil (Zhou *et al.,* 1996). High levels of clay particles inhibit the cell lysis and the extracted DNA is adsorbed back to the clay particles (Frostegard *et al.,* 1999).

More than 80% of the microbial cells stick on to the soil particles rich in clay and organic material (Van Es *et al,* 1984). Adhesion of bacterial cell to the soil particle is complex phenomena involving characteristics of cell, soil particle and the liquid phase (Ling *et al,* 2002). Clay particle, organic matter and bacterial cells have net negative charges and the bacterial cells adhere to soil particle by cation bridging involving polyvalent cations (Huang, 2004). Different concentrations of sodium phosphate buffer (pH 7.0) and disodium EDTA (pH 7.0) was used to disperse and dislodge the microbial cells adhered to soil particles. It is assumed that the anions of these buffers compete for the adsorption sites on clay particle and organic matter and hence release the microbial cells and their DNA by breaking the cat ion bridging (He *et al,* 2005).

Disodium EDTA was found to be more efficient in releasing the cells and removing humic substances as evidenced by the good concentration DNA with less humic acid observed in this study. EDTA is known to destabilise the cell by chelating divalent cat ions (Moore *et al,* 2004). This may be the reason for loss of DNA with the higher concentration of disodium EDTA in pre-lysis buffer. Pre-lysis wash is also known to remove persisting extracellular DNA from dead organisms (Pietramellara et al., 2009). Since the amount of DNA extracted was less with more of humic substances and some insoluble, sodium phosphate was not used

UNDER PEER REVIEW

in pre-lysis and lysis buffer. Addition of SDS to sodium phosphate buffer forms potassium salt of SDS which is insoluble and precipitates along with DNA.

Extraction of DNA from soil microbial community is challenging as the soil harbour diverse microbes varying in cell size, shape and their cell wall composition (Portillo *et al,* 2013). Most of the cell lysis protocols in metagenomics have bias (Delmont *et al,* 2011). Physical, chemical and enzymatic methods and their combination were used for cell lysis. Heating at 75 °C was the most efficient method followed by SDS and vertexing at 1400 g. Longer heating causes more cell disruption and release of DNA (Packard *et al,* 2013), but the quantity of DNA obtained by incubation at 75 °C for more than 90 minutes is non-significant.

Thicker peptidoglycan layer in the cell wall of Gram-positive bacteria makes them difficult to lyse (Packard *et al,* 2013). Enzymatic (lysozyme) lysis of glycosidic linkages of glycan polymer is the standard method (Mahalanabis *et al,* 2009). But in this study lysozyme did not yield DNA. Li *et al* (2013) reported the inhibition in activity of lysozyme by humic and fulvic acid. Even after washing the soil before cell lysis, 5.87 mg humic substances were still remained in soil pellet which might have inhibited the activity of lysozyme. Though many soil DNA isolation protocols used lysozyme (Yeates, 1998 and Moore et al, 2004) along with other methods of cell lysis, but their experiment did not clearly show the effect of only lysozyme on cell lysis. Good amount of DNA from Gram-positive bacteria (*Bacillus subtilis*) was isolated by using only SDS (Shahriar et al, 2011). SDS is a strong anionic detergent which dissolves proteins and lipids. Since cell wall of some Gram-positive bacteria contains significant amount of non-peptidoglycan amino acids and covalently attached proteins (Navarre and Schneewind, 1999), it is expected that the extracted DNA belongs to both Gram-positive and Gram-negative bacteria.

Another important problem in soil DNA extraction is co-extraction of humic substances with DNA. Similarity in physico-chemical properties of humic acid and DNA makes it difficult to separate humic acid from DNA (Lakay *et al*., 2007). The extracted DNA was dark in colour and no enzymatic acitivity (PCR amplification and restriction digestion) was observed. Humic acid inhibits PCR amplification, restriction digestion and transformation efficiency (Tsai and Olson, 1992). Many methods have been used to reduce the concentration of humic substances from soil DNA, we tried chemical flocculation method of Braid *et al* (2003) which is simplest and does not require special equipments to purify soil DNA from humic substances. Addition of calcium chloride to the crude DNA extracted from variety of soils differing in clay content, moisture and pH reduced humic acid content (Sagova-Mareckova et al. 2008). In this study, $CaCl_2$ was found most effective in removing humic substance with retaining good concentration of DNA. This observation contradicts the Schulze-Hardy rule (Schulze, 1882 and Hardy, 1900), according to which trivalent cations are more effective in coagulating humic substances followed by divalent and monovalent.

16S rDNA sequence reads of the isolated DNA were classified to large number of bacterial species with good Shannon diversity index, indicating the protocol is unbiased and is able to extract DNA from wide range of species. Around 1/3rd sequences cannot be classified to any of the known species but belong to bacteria, indicating the potential efficiency of protocol to extract DNA from diverse species. Among classified sequences, the distribution pattern at phylum level is similar to other soil metagenomic studies (Acosta-Martinez et al, 2008; Fierer et al, 2007 and Acosta-Martinez et al, 2010). Janssen (2006) surveyed 32 libraries of 16S rDNA from published articles of different bulk soils to assess the biases in metagenomic

UNDER PEER REVIEW

studies. Except for *Acidobacteria* the proportion of each phylum observed in this study is within the range of Janssen's observation (Table5). This difference in *Acidobacteria* could be due to bias in DNA extraction, PCR amplification or characteristic of sample (soil type and land management practice). Since most of the species belonging to *Acidobacteria* are Gram-negative, inefficiency of the protocol to lyse cell may not be the reason for less *Acidobacteria* observed.

Phylum	Per cent coverage			
	0 mismatch	1 mismatch	2 mismatch	3 mismatch
Actinobacteria	77.30	94.20	96.80	97.60
Aquificae	10.40	11.60	93.30	97.00
Bacteroidetes	90.30	96.00	97.40	97.90
Caldiserica	7.10	78.80	96.50	96.50
Chlamydiae	0.00	68.30	73.80	96.00
Chlorobi	69.10	93.50	97.40	97.80
Chloroflexi	25.50	64.20	92.30	94.30
Chrysiogenetes	87.50	100.00	100.00	100.00
Deferribacteres	75.00	83.90	89.70	92.40
Deinococcus-Thermus	87.60	96.60	98.20	98.70
Dictyoglomi	29.40	100.00	100.00	100.00
Elusimicrobia	64.30	94.70	97.20	97.50
Fibrobacteres	83.40	91.20	92.00	92.60
Fusobacteria	81.50	93.00	95.10	96.30
Gemmatimonadetes	73.60	96.40	97.90	98.30
Lentisphaerae	3.10	90.60	96.00	97.40
Nitrospira	38.60	95.70	97.60	97.90
Planctomycetes	1.60	29.90	35.70	83.60
Proteobacteria	91.00	96.60	97.70	98.10
Spirochaetes	21.00	89.50	92.80	93.50
Synergistetes	78.90	97.40	98.80	99.20
Tenericutes	53.40	90.40	97.40	98.00
Thermodesulfobacteria	85.10	97.90	100.00	100.00
Thermotaga	0.00	0.00	0.00	0.00
BRC1	14.60	51.40	96.60	98.90
OD1	3.10	34.30	63.80	95.80
OP11	0.70	3.50	12.70	72.20
SR1	73.10	77.60	94.20	95.50
TM7	2.50	92.10	95.70	98.10
WS3	45.50	95.60	97.40	97.70
Armatimonadetes	4.40	21.60	80.90	90.90
Verrucomicrobia	2.00	2.50	95.70	97.80
Acidobacteria	90.10	97.90	98.50	98.70
Firmicutes	82.50	95.70	97.40	98.00
Cyanobacteria/Chloroplast	59.10	93.30	96.70	97.20
Overall coverage	47.42	76.94	90.09	95.92

Figure5. Heatmap showing coverage rate of PRBA338 and PRUN518 primer pairs with different phyla of bacterial domain. The primer sequence in combination was searched for similar sequences with 0, 1, 2 and 3 mismatches in RDP databse using ProbeMatch algorithm. Red colour is lowest coverage and green colour is highest coverage.

UNDER PEER REVIEW

The coverage rate of the primer pair PRBA338 and PRUN518 for *Acidobacteria* is more than other dominant phyla observed in this study (*Actinobacteia*, *Proteobacteria* and *Fermicutes*) (Fig5) indicating the observed difference is not due to bias in amplification. We believe the less proportion of *Acidobacteria* observed in this study is due to the characteristic of sample. Species belonging to *Acidobacteria* are known to be predominant in extreme and contaminated soils (Ward *et al,* 2009) and their abundance is negatively correlated to the organic carbon content (Fierer et al, 2007). The soil used for this study is from organic farm known to contain fewer toxic compounds and is considered to be healthy and support beneficial copiotrophic microbes.

Cell wall of Gram-negative bacteria tend to break easily than Gram-positive bacteria (Schneegurt et al, 2003). We further checked whether the optimized protocol is biased towards particular group of cells (Gram-positive and Gram-negative cells) by classifying 16S rDNA sequences. In this study, 49.9, 46.74, 0.34 and 3.03 % of the 16S rDNA sequences belong to Gram-positive, Gram-negative, Gram variable and others respectively. Our results are similar to Hackl *et al* (2004) where 49% of the Austrian pine forest 16S rDNA clone belongs to G+C rich Gram-positive bacteria. The ratio between Gram-positive and Gram-negative bacteria observed in this study is similar to Qi-chun *et al* (2007). These observations clearly indicate that the protocol developed for soil microbial community DNA isolation is unbiased and can extract good quantity DNA free from humic substances.

Several methods were described previously for isolation and purification of soil DNA. Most of these protocols used combination of techniques or methods, hence the effect of each method individually is not clearly known. We tested different methods and analysed their efficiency. Washing the soil with disodium EDTA before cell lysis released microbial cells and reduces the co-precipitation of humic substances. Heating with vertexing in presence of SDS proved to be most efficient method for cell lysis. Chemical flocculation using $CaCl_2$ is the most efficient in removing humic substances with retaining good quantity DNA. The amount of $CaCl_2$ required to remove humic substances need to be adjusted based on the amount of humic substances present in soil. This protocol is very simple, does not need special technique/equipment and many samples can be processed simultaneously. Finally, we demonstrated the suitability of optimized protocol for soil metagenomic studies by sequencing of 16S rDNA from soil metagenome. We found that this protocol is unbiased and can extract good quality and quantity DNA from range of microbes varying in their cell wall composition.

UNDER PEER REVIEW

References

Acosta-Martinez, V., Dowd, S., Sun, Y. and Allen V. (2008) Tag-encoded pyrosequencing analysis of bacterial diversity in a single soil type as affected by management and land use. *Soil Biol Biochem* **40**, 2762–2770

Acosta-Martinez, V., Dowd, S.E., Bell, C.W., Lascano, R., Booker, J.D., et al. (2010) Microbial community composition as affected by dryland cropping systems and tillage in a semiarid sandy soil. *Diversity* **2**, 910–931.

Arbeli, Z. and Fuentes, C.L. (2007) Improved purification and PCR amplification of DNA from environmental samples. *FEMS Microbiol Lett* **272**, 269-275.

Bates, S.T., Berg-Lyons, D., Caporaso, J.G, Walters, W.A., Knight, R. and Fierer, N. (2011) Examining the global distribution of dominant archaeal populations in soil. *ISME J* **5**, 908–917.

Braid, M.D., Daniels, L.M. and Kitts, C.L. (2003) Removal of PCR inhibitors from soil DNA by chemical flocculation. *J. Microbiol Methods* **52**, 389–93

Bruce, K.D., Horns, W.D., Hobman, J.L., Osborn, A.M., Strike, P. and Ritchie, D.A. (1992) Amplification of DNA from native populations of soil bacteria by using the polymerase chain reaction. *Appl Environ Microbiol* **58**, 3413–3416.

Chao, A. and Shen, T.J. (2010) Program SPADE (Species Prediction And Diversity Estimation). Program and User's Guide published at http://chao.stat.nthu.edu.tw

DeGrange, V. and Bardin. R. (1995) Detection and counting of *Nitrobacter* populations in soil by PCR. *Appl Environ Microbiol* **61**, 2093–2098

Delmont, T.O., Robe, P., Clark, I., Simonet, P. and Vogel, T.M. (2011) Metagenomic comparison of direct and indirect soil DNA extraction approaches. *J Microbiol Methods* **86**, 397–400

Fierer, N., Breitbart, M., Nulton, J., Salamon, P., Lozupone, C., Jones, R., Robeson, M., Edwards, R.A., Felts, B., Rayhawk, S. et al. (2007) Metagenomic and small-subunit rRNA analyses reveal the genetic diversity of bacteria, archaea, fungi, and viruses in soil. *Appl Environ Microbiol* **73**, 7059–7066

Frostegard, A., Courtois, S., Ramisse, V., Clerc, S., Bernillon, D., et al. (1999) Quantification of bias related to the extraction of DNA directly from soils. *Appl Environ Microbiol* **65**, 5409–5420

Hackl, E., Zechmeister-Boltenstern, S., Bodrossy, L. and Sessitsch, A. (2004) Comparison of diversities and compositions of bacterial populations inhabiting natural forest soils. *Appl Environ Microbiol* **70**, 5057–5065.

Hardy, W.B. (1900) A preliminary investigation of the conditions which determine the stability of irreversible hydrosols. *Proc Roy Soc Lon* **66**,110–125.

He, J., Xu, Z. and Hughes, J. (2005) Pre-lysis washing improves DNA extraction from a forest soil. *Soil Biol Biochem* **37**, 2337–2341

Herrick, J.B., Miller, D.N., Madsen, E.L. and Ghiorse, W.C. (1996) Extraction, purification, and amplification of microbial DNA from sediments and soils, In PCR: essential techniques ed Burke, J.F. pp. 130–133. New York, N.Y: John Wiley & Sons.

Holben, W.E., Jansson, J.K., Chelm, B.K. and Tiedje, J.M. (1988) DNA probe method for the detection of specific microorganisms in the soil bacterial community. *Appl Environ Microbiol* **54**,703–711

Huang, P.M. (2004) Soil mineral–organic matter–microorganism interactions: fundamentals and impacts. *Adv Agron* **82**, 391-472.

Huysman, F. and Verstraete, W. (1993) Water facilitated transport of bacteria in unsaturated soil columns: influence of cell surface hydrophobicity and soil properties. *Soil Biol Biochem* **25**, 83-90

Janssen, P.H. (2006) Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Appl Environ Microbiol* **72**, 1719–1728

Kuske, C.R., Banton, K.L., Adorada, D.L., Stark, P.C., Hil,l K.K. and Jackson, P.J. (1998) Small-scale DNA sample preparation method for field PCR detection of microbial cells and spores in soil. *Appl Environ Microbiol* **64**, 2463–72.

Lakay, F.M., Botha, A. and Prior, B.A. (2007) Comparative analysis of environmental DNA extraction and purification methods from different humic acid-rich soils. *J Appl Microbiol* **102**(1), 265-73

Leff, L.G., Dana, J.R., McArthur, J.V. and Shimkets, L.J. (1995) Comparison of methods of DNA extraction from stream sediments. *Appl Environ Microbiol* **61**, 1141–1143

UNDER REVIEW

Li, Y., Tan, W., Koopal, L.K., Wang, M., Liu, F. and Norde, W. (2013) Influence of soil humic and fulvic acid on the activity and stability of lysozyme and urease. *Environ Sci Technol* **47**(10), 5050-5056

Ling, T.Y., Achberger, E.C., Drapcho, C.M. and Bengtson, R.L. (2002) Quantifying adsorption of an indicator bacteria in soil-water system. *Transactions of the American Society of Agricultural Engineers* **45**, 669–674.

Mahalanabis, M., ALMuayad, H., Kulinski, M.D., Altman, D. and Klapperich, C.M. (2009) Cell lysis and DNA extraction of gram-positive and gram-negative bacteria from whole blood in a disposable microfluidic chip. *Lab Chip* **9**, 2811–2817

Malik, M.., Kain, J., Pettigrew, C. and Ogram A. (1994) Purification and molecular analysis of microbial DNA from compost. *J Microbiol Methods* **20**, 183–196.

Maron, P.A., Mougel, C. and Ranjard, L. (2011) Soil microbial diversity: methodological strategy, spatial overview and functional interest. *C R Biol* **334**, 403–411.

Moore, E., Arnscheidt, A., Kruger, A., Strompl, C. and Mau, M. (2004) Simplified protocols for the preparation of genomic DNA from bacterial cultures. Molecular Microbial Ecology Manual, Second Edition **101**, 3-18.

Nakatsu, C.H., Torsvik, V. and Ovreas, L. (2000) Soil community analysis using DGGE of 16S rDNA polymerase chain reaction products. *Soil Sci Soc Am J* **64**, 1382–1388.

Navarre, W.W. and Schneewind, O. (1999) Surface Proteins of Gram-Positive Bacteria and Mechanisms of Their Targeting to the Cell Wall Envelope. *Microbiol Mol Biol Rev* **63**, 174–229

Ogram, A., Sayler, G.S. and Barkay. T. (1987) The extraction and purification of microbial DNA from sediments. *J Microbiol Methods* **7**, 57–66.

Olson, N.D. and Morrow, J.B. (2012) DNA extract characterization process for microbial detection methods development and validation. *BMC Res Notes* **5**, 668. doi:10.1186/1756-0500-5-668.

Packard, M.M, Wheeler, E.K., Alocilja, E.C. and Shusteff, M. (2013) Performance Evaluation of Fast Microfluidic Thermal Lysis of Bacteria for Diagnostic Sample Preparation. *Diagnostics* **3**, 105-116

Picard, C., Ponsonnet, C., Paget, E., Nesme, X. and Simonet, P. (1992) Detection and enumeration of bacteria in soil by direct DNA extraction and polymerase chain reaction. *Appl Environ Microbiol* **58**, 2717–22.

Pietramellara, G., Ascher, J., Borgogni, F., Ceccherini, M.T., Guerri, G. and Nannipieri, P. (2009) Extracellular DNA in soil and sediment: fate and ecological relevance. *Biol Fert Soils* **45**, 219–235.

Portillo, M.C., Leff, J.W., Lauber, C.L. and Fierer, N. (2013) Cell size distributions of soil bacterial and archaeal taxa. *Appl Environ Microbiol* **79**, 7610-7617

Qi-chun, Z., Guang-huo, W. and Huai-ying, Y. (2007) Phospholipid fatty acid patterns of microbial communities in paddy soil under different fertilizer treatments. *J Environ Sci* **19**, 55–59

Rajendhran, J. and Gunasekaran, P. (2008) Strategies for accessing soil metagenome for desired applications. *Biotechnol Adv* **26**, 576-590.

Sagar, K., Singh, S.P., Goutam, K.K. and Konwar, B.K. (2014) Assessment of five soil DNA extraction methods and a rapid laboratory-developed method for quality soil DNA extraction for 16S rDNA-based amplification and library construction. *J Microbiol Methods* **97**, 68-73

Sagova-Mareckova, M., Cermak, L., Novotna, J., Plhackova, K., Forstova, J. and Kopecky, J. (2008) Innovative methods for soil DNA purification tested in soils with widely differing characteristics. *Appl Environ Microbiol* **74**, 2902–2907.

UNEDITED REVIEW

Schneegurt, M.A., Dore, S. and Kulpa, C.F. (2003) Direct extraction of DNA from soils for studies in microbial ecology. *Curr Issues Mol Biol* **5**, 1–8.

Schulze, H.J. (1882) Schwefelarsen im wasseroger Losung Prakt. *Chem* **25**, 431–452

Shahriar, M., Haque, R., Kabir, S., Dewan, I. and Bhuyian, M.A. (2011) Effect of Proteinase-K on Genomic DNA Extraction from Gram-positive Strains. *J Pharm Sci* **4**(1), 53-57

Tebbe, C.C. and Vahjen, W. (1993) Interference of humic acids and DNA extracted directly from soil in detection and transformation of recombinant DNA from bacteria and a yeast. *Appl Environ Microbiol* **59**, 2657–65.

Torsvik, V.L. (1980) Isolation of bacterial DNA from soil. *Soil Biol Biochem* **12**, 15-21.

Trevors, J.T., Lee, H. and Cook. S. (1992) Direct extraction of DNA from soil. *Microb Releases* **1**, 111–115.

Tsai, Y. and Olson, B. H. (1992) Rapid method for separation of bacterial DNA from humic substances in sediments for polymerase chain reaction. *Appl Environ Microbiol* **58**, 2292–2295

Van Es, F.B., Laanbroek, H.J. and Veldkamp, H. (1984) Microbial ecology: an overview, In Aspects of microbial metabolism and ecology ed. Codd, G.A. pp. 1–33. New York: Academic Press, Inc.

Ward, N.L., Challacombe, J.F., Janssen, P.H., Henrissat, B., Coutinho, P.M., Wu, M., Xie, G., Haft, D.H., Sait, M., et al. (2009) Three genomes from the phylum Acidobacteria provide insight into the lifestyles of these microorganisms in soils. *Appl Environ Microbiol* **75**, 2046–2056

Zhou, J., Bruns, M.A. and Tiedje, J.M. (1996) DNA recovery from soils of diverse composition. *Appl Environ Microbiol* **62**, 316–322

UNDER PEER REVIEW