

Optimization of soil DNA extraction protocol using Na₂EDTA, SDS, heating, vortexing and CaCl₂, and its validation for metagenomic studies

Abstract

Aim: Several methods described previously for isolation and purification of soil DNA. Most of these protocols use combination of techniques or methods but the role and contribution of each individual method or component used is not clearly discussed. This study aims at analysing the effect of individual components used in extraction of DNA from soil, and finally to optimize soil DNA isolation protocol and its validation by using 16S rDNA sequence analysis.

Methods and Results: The soil was washed with anionic buffers before lysis step to reduce humic substances and release microbial cells from soil matrix, then the cells were lysed using combination of SDS, heating and vortexing and finally humic substances were removed using chemical flocculation. Pre-lysis washing of soil with 100 mmol l⁻¹ Na₂EDTA proved good for releasing microbial cells from soil matrix. Heating the soil sample at 75°C yielded good quantity (15.73 µg g⁻¹ soil) DNA followed by 2% SDS (10.28 µg g⁻¹ soil) and vortexing at 1400 rpm (8.94 µg g⁻¹ soil). Combination of heating, SDS and vortexing yielded 25 µg DNA per gram of soil. Different concentrations of chemical flocculants like AlNH₄(SO₄)₂, FeCl₃, CaCl₂ and MgCl₂ were used to reduce humic substances. Flocculation with 100 mmol l⁻¹ CaCl₂ removed 5.2 mg humic substances without significant loss of DNA. 16S rDNA sequence analysis of DNA extracted from soil reveals presence of all the common soil bacterial species indicating the protocol is unbiased.

Conclusion: Combination of chemical (SDS) and physical (heating and vortexing) methods yield good DNA whereas addition of enzyme (lysozyme) did not show significant effect on cell lysis. The digestion of isolated DNA with restriction enzyme and amplification of 16S rDNA using *Taq* DNA polymerase indicates the isolated DNA is pure enough for metagenomic analysis. 16S rDNA sequencing of soil DNA indicates that this protocol can extract good quality and quantity DNA from range of bacteria present in soil varying in their cell wall composition. 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.

Introduction

Soils are highly complex environment containing diverse microbial species. A gram of soil can contain between 10⁵ and 10⁶ 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. Molecular tools like PCR amplification followed by gel-based profiling or 16S rDNA sequencing hints at the microbial community composition, diversity and dynamics, whereas the whole metagenome sequencing helps us to study the functional capabilities of microbial species present in the sample without being culturing them (Krishnaraj and Pasha, 2014). Further, a potentially efficient gene could be isolated from unculturable species through metagenomic DNA, cloned and expressed in culturable species to produce agriculturally or industrially important protein (Krishnaraj and Pasha, 2014).

Good quality and quantity DNA from environmental sample is a fundamental step for all metagenomic studies. Several protocols are developed for Soil DNA isolation and these 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, Krishnaraj and Pasha, 2014). 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, 1995), column chromatography (Herrick et al, 1996) and chemical flocculation (Braid et al, 2003).

Many authors reported protocols using above methods or components, but most of them did not discuss the effect of each component used in isolation of DNA from soil. Because of lack of comparative studies, the efficiency of each method for cell lysis and purification is not clearly understood. Here we made an attempt to compare different methods for releasing microbial cells from soil matrix, lysis of cell and removal of humic substances from soil DNA, and finally an optimized soil DNA isolation protocol is presented. The optimised protocol is further demonstrated for its suitability for metagenomic studies by high throughput 16S rDNA sequencing analysis.

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 (Na_2EDTA) (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 Na_2EDTA (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 (Sigma Aldrich) 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 bacterial 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 rpm at 37°C for 0, 30, 60, 90, 120 and 180 minutes on thermomixer (Eppendorf, Germany). For heating, the sample was incubated at 75°C for 0, 30, 60, 90, 120 and 180

minutes in water bath. Sodium dodecyl sulphate (SDS) at the rate of 1, 2, 3, 4, 5 and 6% was used for chemical cell lysis in DNA isolation buffer at 65°C temperature. lysozyme (HiMedia, India) at the concentrations of 0, 25, 50, 75, 100, 125 and 150 mg was used for enzymatic lysis at 37°C.

Removal of humic substances

Chemical flocculation using aluminium ammonium sulphate ($\text{AlNH}_4(\text{SO}_4)_2$), 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} Na_2EDTA , 1.5 mol l^{-1} NaCl and 1.25% PVPP). After treatment, the supernatant was mixed with equal volume of chloroform : 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.

Protocol optimized in brief

Two hundred mg soil (clay loam with 0.41% organic carbon, pH 7.55 and 58.59 me.100g^{-1} CEC) free from pebbles and debris was mixed with 1 ml of 100 mmol l^{-1} Na_2EDTA , 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 rpm 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. The upper aqueous layer was then transferred to fresh 1.5 ml centrifuge tube, 1/10th volume of 3M sodium acetate buffer (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 below flowchart (figure 1) represents the protocol in brief.

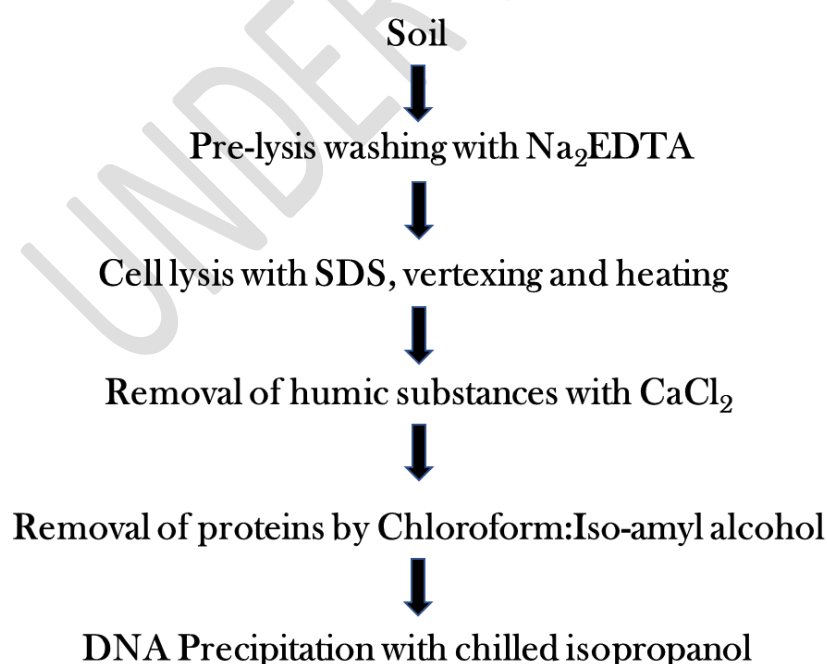


Figure 1. flowchart depicting brief steps involved in protocol for soil DNA extraction.

The amount of humic substances was measured using Cary 50 Bio UV-spectrophotometer at 320 nm with humic acid sodium salt (Sigma Aldrich) as standard (Arbeli and Fuentes, 2007, Olson and Morrow, 2012). DNA was analysed by agarose electrophoresis followed by documentation in Syngene G box gel documentation unit and quantified by Syngene Genetool software using Lambda DNA *HindIII* single digest (Bangalore Genie, India) 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

16S rDNA was amplified using PRBA338 and PRUN518 primers (Nakatsu *et al.*, 2000 and Pasha *et al.*, 2019). Each PCR reaction contained 1X PCR buffer, 1.2 mmol l⁻¹ MgCl₂, 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 along with 100 bp DNA ladder (Bangalore Genie, India).

Restriction digestion

Five hundred ng of metagenomic DNA was digested by 5 units of *EcoRI* and *HindIII* (Bangalore Genie, India) 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 DNA and Lambda DNA *HindIII* single digest (Bangalore Genie, India) marker as reference, 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 good quality sequences were phylogenetically classified using M5RNA annotation source of MG-RAST (Meyer *et al.*, 2008). Species richness and their diversity was 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 staining reaction information is not available were classified as others.

Results

Sodium phosphate and Na₂EDTA was found to extract 7.33 and 24.40 mg humic acid respectively from a gram of soil (table 1). There was no additional humic acid extracted with increase in the concentration (above 180 mmol l⁻¹) of Na₂EDTA. However, higher

concentration (above 180 mmol l⁻¹) of sodium phosphate buffer resulted in the extraction of more humic substances (table 1). 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 substances extraction increased with the increase in concentration of Na₂EDTA, 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 Na₂EDTA used in pre-lysis wash step. Hence, we used only 100 mmol l⁻¹ Na₂EDTA for pre-lysis wash.

Table 1. Effect of pre-lysis washing on DNA and humic acid concentration

Conc. Of Sod. Phosphate / Na ₂ EDTA 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	Na ₂ EDTA	Sodium phosphate	Na ₂ EDTA	Sodium phosphate	Na ₂ 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 ^c
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 vortexing at 1400 rpm. Significant difference in the yield of DNA was observed by heating at 75°C up to 90 minutes, vortexing at 1400 rpm for 90 minutes (table 2) and treating with 2% SDS (table 3). Along with DNA some amount of humic substances were also precipitated and the amount of humic substances precipitated was more with heating (table 2 and table 3). Treatment of soil slurry with different concentrations of lysozyme did not yield DNA (table 3).

Table 2. Effect of vortexing and heating on cell lysis and DNA yield

Incubation time (Minutes)	Vortexing at 1400 rpm 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

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

Table 3. 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 $p \leq 0.01$, as calculated by t test

Vertexing at 1400 rpm 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 rpm 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 (figure 2). However, large amount of humic substances (10.88 mg g⁻¹ of soil) was also extracted with this method.

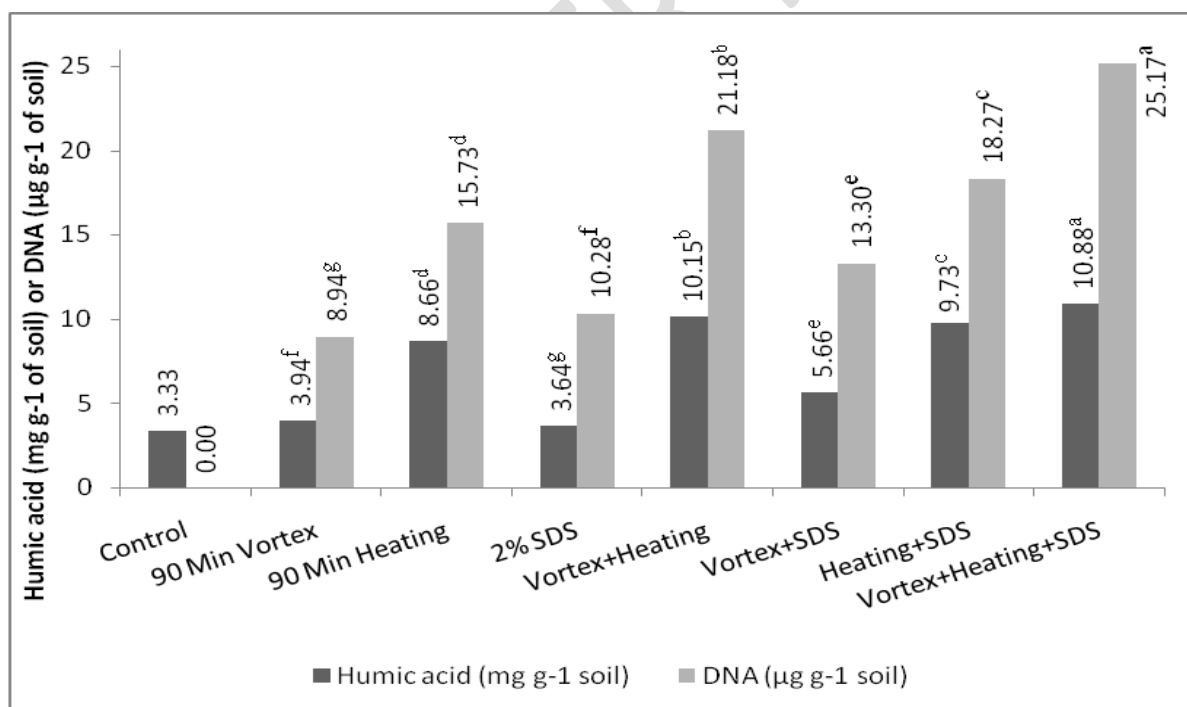


Figure 2. Effect of combination of methods on cell lysis. Heating at 75°C with vertexing (1400 rpm) in presence of 2% SDS for 90 minutes yield highest DNA followed by heating with vertexing and heating with SDS. Heating also extracts lot of humic acid. Values followed by letters indicate significant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

All the chemical flocculants used to remove humic substances also reduced the concentration of DNA. AlNH₄(SO₄)₂, and MgCl₂ FeCl₃ reduces DNA concentration drastically without having much effect on the concentration of humic substances. But CaCl₂ yielded good

concentration DNA with less of humic substances (table 4). There was significant reduction in humic substances even up to 180 mmol l⁻¹ CaCl₂, but the concentration of DNA reduced drastically after 100 mmol l⁻¹ CaCl₂ (table 4).

Table 4. Effect of chemical flocculants on concentration of DNA and humic acid

Concentration (mM)	AlNH ₄ (SO ₄) ₂		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 (figure 3 and figure 4) of soil DNA purified with 100 mmol l⁻¹ CaCl₂ indicates that the extracted DNA is pure enough for enzymatic activity and can be used for metagenomic analysis.

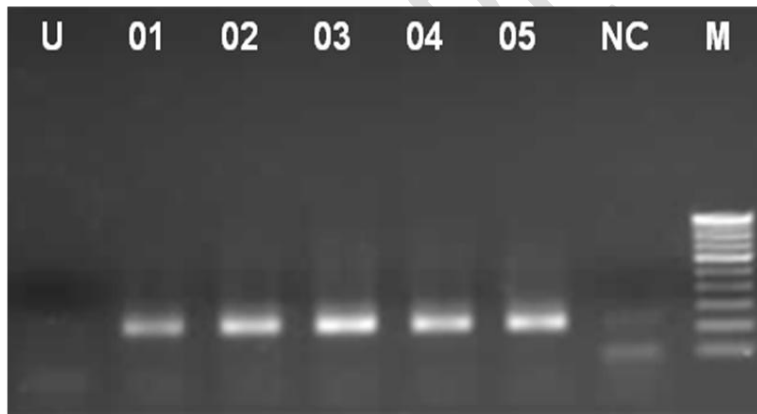


Figure 3. PCR amplification of 16S rDNA from soil metagenome. No amplification was observed in (lane U) DNA without using chemical flocculant, indicating humic acid inhibits *Taq* DNA polymerase activity. Lane 01 to 05 is purified DNA using CaCl₂, 01, 02, 03, 04 and 05 are undiluted, one-fold, two-fold, 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.

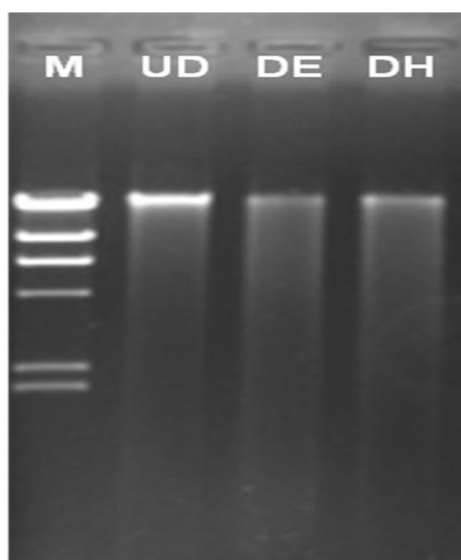


Figure 4. Restriction digestion of metagenomic DNA. 5 unit of *EcoRI* (DE) and *HindIII* (DH) could able to digest 1 μ g DNA in 3 hours indicating DNA is free from humic acid and can be used for metagenomic studies. UD is undigested DNA and M is lambda DNA *HindIII* single digest marker.

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*, *Firmicutes*, *Bacteroidetes*, *Cyanobacteria*, *Chloroflexi*, *Nitrospirae*, *Acidobacteria* and *Gemmatimonadetes* respectively (table 5).

Table 5. 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 (figure 5).

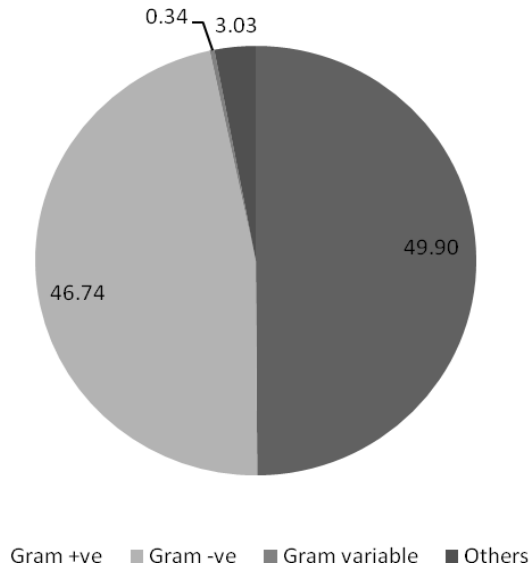


Figure 5. Proportion of gram positive and gram negative bacteria in soil metgenome. 49.9%, 46.74%, 0.34% and 3.03% 16S r DNA sequences belong to gram positive, gram negative, gram variable and others respectively. The ratio between gram positive and gram negative bacteria is 1.07.

Discussion

Good quality and quantity DNA is very much essential for the culture independent genomic analysis of environmental sample. There are several protocols available to extract DNA from soil, but none of them analysed effect of individual component or method 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 Na₂EDTA (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). Na₂EDTA 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 Na₂EDTA in pre-lysis buffer. Pre-lysis wash is also known to remove persisting extracellular DNA from dead organisms (Pietramellara *et al.*, 2009). The amount of DNA extracted was less with more of contaminating humic substances and also formation of insoluble precipitation in the tube when sodium phosphate was used for pre-lysis washing. Hence, it is not used in pre-lysis and lysis buffer.

Extraction of DNA from soil bacterial community is challenging as the soil harbour diverse bacteria varying in cell size, shape and their cell wall composition (Portillo *et al.*, 2013). The cell lysis protocols in metagenomics could be biased due to above reasons (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 vortexing at 1400 rpm. 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) and hence enzymatic (lysozyme) lysis of glycosidic linkages of glycan polymer is the standard method (Mahalanabis *et al.*, 2009) for extraction of DNA from bacteria. 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 mention the effect of using only lysozyme on cell lysis. This study reveals that amount of DNA extracted with only lysozyme (even at 150 mg concentration) is negligible (0.008 µg g⁻¹ soil) as compared to other methods used here like heating, SDS and vortexing. This could be due to inhibitory effect of humic substance on the activity of lysozyme as there was 5.87 mg humic substances present in soil pellet used for lysozyme treatment. Li *et al.* (2013) reported the inhibition in activity of lysozyme by humic and fulvic acid.

Since good amount of DNA is obtained in this study even without using lysozyme, we analysed whether the protocol is biased with gram positive bacteria. Cell wall of gram-negative bacteria tend to break easily than gram positive bacteria (Schneegurt *et al.*, 2003). If the protocol is inefficient then it could be problem with lysis of gram positive bacterial cell. 16S rDNA sequence analysis shows similar proportion of gram positive (49.90%) and gram negative (46.74%) bacteria in this study sample. There are few studies which reported isolation of DNA from gram positive bacteria even without using lysozyme (Martzy *et al.*, 2019, Bollet *et al.*, 1955) Yuan *et al.* (2012) reported that lysozyme alone is not efficient in extracting DNA from human microbiome (especially gram positive bacteria). Shahriar *et al.* (2011) isolated good amount of DNA from gram positive bacteria (*Bacillus subtilis*) using only SDS. 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 combination of heating and SDS extracted DNA from both gram positive and gram negative bacteria in this study.

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 activity (PCR amplification and restriction digestion) was observed. Humic acid inhibits PCR amplification, restriction digestion and transformation efficiency (Tsai and Olson, 1992). Many methods are in use 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. We tested different concentration of AlNH₄(SO₄)₂, MgCl₂, FeCl₃ and CaCl₂ to reduce humic substances in DNA. In this study, CaCl₂ was found most effective in removing humic substance with retaining good concentration of DNA. Similarly, Sagova-Mareckova *et*

al. (2008) observed addition of CaCl₂ to the crude DNA extracted from variety of soils differing in clay content, moisture and pH reduced humic acid content. 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..

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

Figure 6. 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 database using ProbeMatch algorithm. Red colour is lowest coverage and green colour is highest coverage.

The optimized protocol was further validated by high throughput sequencing of 16S rDNA of DNA extracted from soil. 16S rDNA sequence reads of the isolated DNA were classified to large number of bacterial species with Shannon diversity index of 4.42, indicating the protocol 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 even which are yet unculturable in lab. 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). We analysed and compared our 16S rDNA sequence data with Janssen (2006), who surveyed 32 libraries of 16S rDNA from published articles of different bulk soils to assess the biases in metagenomic studies. Except for *Acidobacteria* the proportion of each phylum observed in this study is within the range of Janssen's observation (table 5).

The less representation of *Acidobacteria* in 16S rDNA sequence data could be due to bias in DNA extraction, PCR amplification or characteristic of sample (soil type and land management practices). Since most of the species belonging to *Acidobacteria* are gram negative (Dedysh *et al.*, 2017), inefficiency of the protocol to lyse cell may not be the reason for less *Acidobacteria* observed as they can be lysed easily (Schneegurt *et al.*, 2003). The bias in PCR amplification is also ruled out as the coverage rate of the primer pair PRBA338 and PRUN518 for *Acidobacteria* is more than other dominant phyla observed in this study (*Actinobacteria*, *Proteobacteria* and *Fermicutes*) (figure 6). We believe the less proportion of *Acidobacteria* observed in this study could be due to the characteristic of soil sample used in this study. 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 in this study is from organic farm with good organic carbon, known to contain fewer toxic compounds and is considered to be healthy and support beneficial copiotroph microbes (Lori *et al.*, 2017). This is also supported by presence of higher proportion of *Proteobacteria* (55.92%) and *Actinobacteria* (30.92%) in this study (table 5) (Fierer *et al.*, 2007).

Conclusion

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 Na₂EDTA before cell lysis reduces the co-precipitation of humic substances and increased DNA concentration. We observed addition of lysozyme does not yield DNA and can be excluded while extracting DNA from soils rich in humic substances. Heating with vortexing in presence of SDS proved to be most efficient method for cell lysis. Chemical flocculation using CaCl₂ is the most efficient in removing humic substances with retaining good quantity DNA. The amount of CaCl₂ required to remove or reduce humic substances need to be adjusted based on the amount of humic substances present in soil. Finally, we demonstrated the validity of optimized protocol for soil metagenomic studies by largescale 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. This protocol is very simple, does not need special technique/equipment or hazardous chemicals like phenol and many samples can be processed simultaneously.

Acknowledgement

This work is a part of All India Network Project on Soil Biodiversity and Biofertilizers (AINPSBB) and the authors would like to thank Indian Council of Agriculture Research for funding this work.

Ethic: NA

Consent: NA

References

- Acosta-Martinez V, Dowd S, Sun Y, Allen V. Tag-encoded pyrosequencing analysis of bacterial diversity in a single soil type as affected by management and land use. *Soil Biol Biochem.* 2008;40:2762–2770
- Acosta-Martinez V, Dowd SE, Bell CW, Lascano R, Booker JD, et al. Microbial community composition as affected by dryland cropping systems and tillage in a semiarid sandy soil. *Diversity.* 2010;2:910–931.
- Arbeli Z, Fuentes CL. Improved purification and PCR amplification of DNA from environmental samples. *FEMS Microbiol Lett.* 2007;272:269-275.
- Bates ST, Berg-Lyons D, Caporaso JG, Walters WA, Knight R, Fierer N. Examining the global distribution of dominant archaeal populations in soil. *ISME J.* 2011;5:908–917.
- Bollet C, Gevaudan MJ, de Lamballerie X, Zandotti C, de Micco P. A simple method for the isolation of chromosomal DNA from Gram positive or acid-fast bacteria. *Nucleic Acids Res.* 1991;19(8):1955.
- Braid MD, Daniels LM, Kitts CL. Removal of PCR inhibitors from soil DNA by chemical flocculation. *J. Microbiol Methods.* 2003;52:389–93
- Bruce KD, Horns WD, Hobman JL, Osborn AM, Strike P, Ritchie DA. Amplification of DNA from native populations of soil bacteria by using the polymerase chain reaction. *Appl Environ Microbiol.* 1992;58:3413–3416.
- Chao A, Shen TJ. Program SPADE (Species Prediction And Diversity Estimation). Program and User's Guide. 2010; published at <http://chao.stat.nthu.edu.tw>
- Dedysh SN, Kulichevskaya IS, Huber KJ, Overmann J. Defining the taxonomic status of described subdivision 3 Acidobacteria: the proposal of Bryobacteraceae fam. nov. *Int J Syst Evol Microbiol.* 2016;67:498-501.
- DeGrange V, Bardin R. Detection and counting of *Nitrobacter* populations in soil by PCR. *Appl Environ Microbiol.* 1995;61:2093–2098
- Delmont TO, Robe P, Clark I, Simonet P, Vogel TM. Metagenomic comparison of direct and indirect soil DNA extraction approaches. *J Microbiol Methods.* 2011;86:397–400
- Fierer N, Breitbart M, Nulton J, Salamon P, Lozupone C, Jones R, Robeson M, Edwards RA, Felts B, Rayhawk S. et al. Metagenomic and small-subunit rRNA analyses reveal the genetic diversity of bacteria, archaea, fungi, and viruses in soil. *Appl Environ Microbiol.* 2007;73:7059–7066
- Frostegard A, Courtois S, Ramiisse V, Clerc S, Bernillon D. et al. Quantification of bias related to the extraction of DNA directly from soils. *Appl Environ Microbiol.* 1999;65: 5409–5420
- Hardy WB. A preliminary investigation of the conditions which determine the stability of irreversible hydrosols. *Proc Roy Soc Lon.* 1900;66:110–125.
- He J, Xu Z, Hughes J. Pre-lysis washing improves DNA extraction from a forest soil. *Soil Biol Biochem.* 2005;37:2337–2341

Herrick JB, Miller DN, Madsen EL, Ghiorse WC. Extraction, purification, and amplification of microbial DNA from sediments and soils, In Burke JF. editors. PCR: essential techniques. John Wiley & Sons; 1996.

Holben WE, Jansson JK, Chelm BK, Tiedje JM. DNA probe method for the detection of specific microorganisms in the soil bacterial community. *Appl Environ Microbiol.* 1988;54:703–711.

Huang PM. Soil mineral–organic matter–microorganism interactions: fundamentals and impacts. *Adv Agron.* 2004;82:391-472.

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

Janssen PH. Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Appl Environ Microbiol.* 2006;72:1719–1728

Krishnaraj PU, Pasha MA. Metagenomics concepts, tools and applications. In *Environment Science and Engineering Vol. 7: Instrumentation, Modelling and Analysis*. Studium Press, USA. 2016;272-307

Kuske CR, Banton KL, Adorada DL, Stark PC, Hill KK, Jackson PJ. Small-scale DNA sample preparation method for field PCR detection of microbial cells and spores in soil. *Appl Environ Microbiol.* 64:1998;2463-72.

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

Leff LG, Dana JR, McArthur JV, Shimkets LJ. Comparison of methods of DNA extraction from stream sediments. *Appl Environ Microbiol.* 1995;61:1141-1143.

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

Ling TY, Achberger EC, Drapcho CM, Bengtson RL. Quantifying adsorption of an indicator bacteria in soil-water system. *Transactions of the American Society of Agricultural Engineers.* 2002;45:669-674.

Lori M, Symnackik S, Mäder P, De Deyn G, Gattinger A. Organic farming enhances soil microbial abundance and activity—A meta-analysis and meta-regression. *PLoS ONE.* 2017;12(7): e0180442.

Mahalanabis M, AlMuayad H, Kulinski MD, Altman D, Klapperich CM. Cell lysis and DNA extraction of gram-positive and gram-negative bacteria from whole blood in a disposable microfluidic chip. *Lab Chip.* 2009;9:2811-2817.

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

Maron PA, Mougel C, Ranjard L. Soil microbial diversity: methodological strategy, spatial overview and functional interest. *C R Biol.* 2011;334:403-411.

Martzy R, Bica-Schröder K, Pálvölgyi ÁM. *et al.* Simple lysis of bacterial cells for DNA-based diagnostics using hydrophilic ionic liquids. *Sci Rep.* 2019;9:13994.

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

Nakatsu CH, Torsvik V, Ovreas L. Soil community analysis using DGGE of 16S rDNA polymerase chain reaction products. *Soil Sci Soc Am J.* 2000;64:1382-1388.

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

Ogram A, Sayler GS, Barkay T. The extraction and purification of microbial DNA from sediments. *J Microbiol Methods*. 1987;7:57-66.

Olson ND, Morrow JB. DNA extract characterization process for microbial detection methods development and validation. *BMC Res Notes*. 2012;5:668. doi:10.1186/1756-0500-5-668.

Packard MM, Wheeler EK, Alocilja EC, Shusteff M. Performance Evaluation of Fast Microfluidic Thermal Lysis of Bacteria for Diagnostic Sample Preparation. *Diagnostics*. 2013;3:105-116.

Pasha MA, More SA, Krishnaraj PU. In silico and denaturing gradient gel electrophoresis-based evaluation of 16S rDNA primers for metagenomic studies. *Res. J. Biotech* 2019;14(2):54-60.

Picard C, Ponsonnet C, Paget E, Nesme X, Simonet P. Detection and enumeration of bacteria in soil by direct DNA extraction and polymerase chain reaction. *Appl Environ Microbiol*. 1992;58:2717-2722.

Pietramellara G, Ascher J, Borgogni F, Ceccherini MT, Guerri G, Nannipieri P. Extracellular DNA in soil and sediment: fate and ecological relevance. *Biol Fert Soils*. 2009;45:219-235.

Portillo MC, Leff, JW, Lauber CL, Fierer N. Cell size distributions of soil bacterial and archaeal taxa. *Appl Environ Microbiol*. 2013;79:7610-7617.

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

Sagar K, Singh SP, Goutam KK, Konwar BK. 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*. 2014;97:68-73.

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

Schneegurt MA, Dore S, Kulpa CF. Direct extraction of DNA from soils for studies in microbial ecology. *Curr Issues Mol Biol*. 2003;5:1-8.

Schulze HJ. Schwefelarsen im wasseroger Losung Prakt. *Chem*. 1882;25:431-452.

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

Tebbe CC, Vahjen W. 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*. 1993;59:2657-2665.

Torsvik VL. Isolation of bacterial DNA from soil. *Soil Biol Biochem*. 1980;12:5-21.

Trevors JT, Lee H, Cook S. Direct extraction of DNA from soil. *Microb Releases*. 1992;1:111-115.

Tsai Y, Olson BH. Rapid method for separation of bacterial DNA from humic substances in sediments for polymerase chain reaction. *Appl Environ Microbiol*. 1992;58:2292-2295.

Van Es FB, Laanbroek HJ, Veldkamp H. Microbial ecology: an overview, In Codd GA. editors. *Aspects of microbial metabolism and ecology*. New York: Academic Press, Inc. 1984

Ward NL, Challacombe JF, Janssen PH, Henrissat B, Coutinho PM, Wu M, Xie G, Haft DH, Sait M, et al. Three genomes from the phylum Acidobacteria provide insight into the lifestyles of these microorganisms in soils. *Appl Environ Microbiol*. 2009;75:2046-2056.

Yuan S, Cohen DB, Ravel J, Abdo Z, Forney LJ. Evaluation of Methods for the Extraction and Purification of DNA from the Human Microbiome. *PLoS ONE*. 2012;7(3): e33865.

Zhou J, Bruns MA, Tiedje JM. DNA recovery from soils of diverse composition. *Appl Environ Microbiol*. 1996;62:316-322.