SJSS
SPANISH JOURNAL OF SOIL SCIENCE

Structural and functional metagenomic analyses of a tropical agricultural soil

Análisis metagenómicos estructurales y funcionales de un suelo agrícola tropical
Análises metagenómicas estruturais e funcionais de um solo agrícola tropical

AUTHORS

Lateef Babatunde Salam[@,1]
babssalaam@yahoo.com

Oluwafemi Sunday Obayori[2]

@ Corresponding Author

[1]Department of Biological Sciences, Al-Hikmah University, Ilorin, Kwara State, Nigeria.

[2]Department of Microbiology, Lagos State University, Ojo, Lagos State, Nigeria.

Received: 04.07.2018 | **Revised:** 09.12.2018 | **Accepted:** 21.01.2019

ABSTRACT

Understanding the intricate link between the soil microbiota and their metabolic functions is important for agricultural and ecological processes and could be used as a biomarker of soil health. To understand the relationship between soil microbial community structure and functions, a soil microcosm designated 2S (agricultural soil) was set up. Metagenomic DNA was extracted from the soil microcosm and sequenced using Miseq Illumina next generation sequencing and analysed for their structural and functional properties. Structural analysis of the soil microcosm by MG-RAST revealed 40 phyla, 78 classes, 157 orders, 273 families and 750 genera. *Actinobacteria* (54.0%) and *Proteobacteria* (17.5%) are the dominant phyla while *Conexibacter* (8.38%), *Thermoleophilum* (7.40%), and *Streptomyces* (4.14%) are the dominant genera. Further assignment of the metagenomics using Cluster of Orthologous Groups (COG), Kyoto Encyclopedia of Genes and Genomes (KEGG), GhostKOALA, and NCBI's CDD revealed diverse metabolic pathways utilized by the microbial community for the metabolism of carbohydrates, amino acids, lipids, biosynthesis of secondary metabolites and resistance to antibiotics. Taxonomic analysis of the annotated genes also revealed the preponderance of members of *Actinobacteria* and *Proteobacteria*. This study has established that members of the phyla *Actinobacteria* and *Proteobacteria* are the key drivers of the majority of important metabolic activities in the soil ecosystem and are thus an integral part of the soil microbial community.

RESUMEN

Existe una complicada relación entre la microbiota del suelo y sus funciones metabólicas en los procesos agrícolas y ecológicos que es necesario comprender con objeto de poder ser utilizada como biomarcador a la hora de establecer la salud del suelo. Para entender la relación entre la estructura y las funciones de la comunidad microbiana del suelo se llevó a cabo un microcosmo de suelo designado como 2S (suelo agrícola). Se extrajo el ADN metagenómico del microcosmo de suelo, se secuenció utilizando secuenciación de nueva generación con tecnología Miseq Illumina y se analizaron sus propiedades estructurales y funcionales. El análisis estructural del microcosmo por MG-RAST dio como resultado 40 filos, 78 clases, 157 órdenes, 273 familias y 750 géneros. Actinobacteria *(54,0%) y* Proteobacteria *(17,5%) fueron los filos dominantes mientras que* Conexibacter *(8,38%),* Thermoleophilum *(7,40%), y* Streptomyces *(4,14%) fueron los géneros dominantes. Posteriores estudios metagenómicos utilizando el Cluster de Grupos Ortólogos (COG), la Enciclopedia de Genes y Genomas de Kioto (KEGG), el GhostKOALA y*

DOI: 10.3232/SJSS.2019.V9.N1.01

el NCBI's CDD mostraron vías metabólicas diversas utilizadas por la comunidad microbiana para el metabolismo de carbohidratos, aminoácidos, lípidos, la biosíntesis de metabolitos secundarios y la resistencia a antibióticos. Los análisis taxonómicos de los genes anotados también revelaron el predominio de miembros de Actinobacteria *y* Proteobacteria. *Este estudio ha establecido que los miembros de* Actinobacteria *y* Proteobacteria *son factores clave de la mayoría de las actividades metabólicas en el ecosistema del suelo y son una parte integral de la comunidad microbiana edáfica.*

RESUMO

A compreensão da ligação complexa entre o microbioma do solo e as suas funções metabólicas é importante para os processos agrícolas e ecológicos, podendo ser usada como um biomarcador de avaliação da qualidade do solo. Para perceber a relação entre a estrutura e as funções da comunidade microbiana do solo foi realizado um ensaio em microcosmos com um solo agrícola e designado como 2S. Realizou-se a extração do ADN do solo do microcosmos, seguida da sequenciação de nova geração Miseq Illumina com o objetivo de analisar as suas propriedades estruturais e funcionais. A análise estrutural do solo do microcosmos por MG-RAST identificou 40 filos, 78 classes, 157 ordens, 273 famílias e 750 géneros. Actinobacteria *(54,0%) e* Proteobacteria *(17,5%) são os filos dominantes enquanto* Conexibacter *(8,38%),* Thermoleophilum *(7,40%), e* Streptomyces *(4,14%) são os géneros dominantes. Posteriores estudos bioinformáticos usando o Cluster de Grupos Ortólogos (COG), a Enciclopédia de Genes e Genomas de Kioto (KEGG), o GhostKOALA e o NCBI's CDD, foram identificadas diversas vias metabólicas utilizadas pela comunidade microbiana no metabolismo dos hidratos de carbono, aminoácidos, lípidos, biossíntese de metabolitos secundários e a resistência a antibióticos. As análises taxonómicas dos genes anotados também revelaram o predomínio de membros de* Actinobacteria *e* Proteobacteria. *Este estudo estabeleceu que os membros do filo* Actinobacteria *e* Proteobacteria *são fatores chave da maioria das atividades metabólicas no ecossistema solo e são uma parte integrante da comunidade microbiana do solo.*

KEY WORDS
Illumina next generation sequencing, microbial community structure, shotgun metagenomics, *Actinobacteria*, *Proteobacteria*.

PALABRAS CLAVE
Secuenciación de nueva generación con tecnología Illumina, estructura de la comunidad microbiana, metagenómica, *Actinobacteria*, *Proteobacteria*.

PALAVRAS-CHAVE
Sequenciação de nova geração com tecnologia Illumina, estrutura da comunidade microbiana, metagenómica, *Actinobacteria*, *Proteobacteria*.

1. Introduction

Microorganisms found in natural environments are responsible for most of the biological transformations that lead to formation of soil nutrients (Asadu et al. 2015; Costa et al. 2015). The role these organisms play in driving various biogeochemical cycles of elements, fixing carbon and nitrogen, mineralizing dead organic matter and protection of plants from biotic and abiotic stresses is replete in the literature. They play a very important role in the health and ecological balance of these environments. Their cardinal role in influencing soil function and determining overall soil quality in the terrestrial environment, including agricultural soil, cannot be overemphasized (Nannipieri et al. 2003; Wardle et al. 2004; Arias et al. 2005).

Furthermore, soil microbes are important in provisioning and regulating services provided by the soil ecosystem. This includes growth medium for plants, soil aggregation improvement and stability, buffering waterflow, recycling of waste and detoxification, increasing nutrient bioavailability to plants, filtering contaminants, biological control of pests, weeds and pathogens, carbon storage and regulation of greenhouse gases (Aislabie and Deslippe 2013; Rashid et al. 2016). Trivedi et al. (2016) noted that changes in microbial population or activity could precede detectable changes in soil physical and chemical properties, thereby providing an early sign of soil improvement or an early warning of soil degradation. Thus, shifts in microbial community structure and function could be used as biomarkers of soil ecological health.

Recognition of the essential role of microbial communities in life and in the functioning of the ecosystem has engendered aggressive efforts to discern their diversity and harness their hidden metabolic and genetic potentials (Mocali and Benedetti 2010; Paul et al. 2016). Traditionally, methods that involve culturing of microorganisms were employed for this purpose. However, such methods often reveal < 1% of the total soil populations, in most cases excluding some of the most important contributors to soil functions (Maier and Pepper 2009). One major factor responsible for this is the fastidious nature of some of these organisms or their slow growth relative to other members of the community, which overgrow and overshadow them on routine media (Cheung and Kinkle 2001). Therefore, in the last two decades, molecular approaches targeting 16S rRNA genes useful in taxonomic affiliation or marker genes specific for certain metabolic functions have become the favoured strategy for unravelling the structure, functional diversity and dynamics of microbial communities in soil and other environments (Stach and Burns 2002).

Some of the methods that have revolutionarily transformed research in this area include Denaturing Gradient Gel Electrophoresis-PCR (DGGE-PCR), Length Heterogeneity-PCR (LH-PCR), Clone Libraries and terminal Restriction Fragment Length Polymorphism (tRFLP) (Muyzer et al. 1993; Liu et al. 1997; Suzuki et al. 1998). The advances made notwithstanding, these methods are limited by the fact that they are laborious, time consuming and offer access to only a target group or populations within the broader microbial community (Paul et al. 2016). However, with advances in high-throughput sequencing technologies such as illumina next generation shotgun sequencing, construction of clone libraries is no longer necessary, greater yield of sequence data can be obtained and information can be provided about which organisms are present and what metabolic processes are possible in a community (Segata et al. 2013). Metagenomics using next generation sequencing demonstrably gives unprecedented insight into the genetic potentials of microbial communities as well as underrepresented populations (Handelsman 2004; Newby et al. 2009; Oulas et al. 2015). It is cheaper and enables the identification of novel molecules with significant functionalities and applications (Streit and Schmitz 2004; Bashir et al. 2014).

Although the metagenomic approach is increasingly gaining ascendancy as the gold standard for discerning microbial community structure and functions, much of the ecosystems in the world remain unexplored and even fewer have been investigated using this approach. In the Nigerian environment, there have been only a handful of reports employing this approach (Ogbulie and Nwaokorie 2016; Salam et al. 2017). However, globally, very few published works utilized sequence data generated from illumina next generation shotgun sequencing for functional characterization of environmental samples. Elucidation of microbial community structure and functions as designed in this study allows for quantitative and qualitative assessment of soil health, which is pivotal in the field of agriculture and soil ecology. Here, we report the use of illumina next generation shotgun sequencing to decipher the microbial community structure and functions of a tropical agricultural soil in Ilorin, Nigeria and determine the possible role of dominant members of the microbial community in driving core metabolic functions in the soil.

2. Materials and Methods

2.1. Sampling site description

Soil samples were collected from an Agricultural farm in Ilorin, Kwara State, Nigeria. The coordinates of the sampling site were latitude 8° 29' 18.11" N and longitude 4° 28' 58.19" E. The soil is dark-brown in colour and consists of fine, medium coarse silty sand (sand 72%, silt 27%, traces of fine gravel; Soil Survey Division Staff 1993) and is used to plant majorly millet, maize and vegetables.

2.2. Sampling and microcosms set up

Soil samples were collected at a depth of 10-12 cm with sterile trowel after clearing debris from the soil surface. Soil sample was passed through a 2-mm mesh size sieve. The sieved soil was thoroughly mixed in a large plastic bag to avoid variability among the results of replicate soil samples and used without air-drying. Sieved soil (2 kg) was measured and placed in open aluminum pans (37 cm x 14 cm x 7 cm). It was designated 2S. The set up (in triplicates) was incubated at room temperature (25 ± 3 °C) and flooded with 100 ml sterile distilled water to maintain moisture level of about 25%. Samples (in triplicates) were taken for physicochemical analysis. The physicochemistry of the soil sample has been described previously (Salam et al. 2014).

2.3. DNA extraction and Shotgun Metagenomics

Genomic DNA used for metagenomic analysis was extracted directly from 2S soil microcosm. Genomic DNA was extracted from the agricultural soil (2S) immediately after sampling to determine the microbial community structure of the soil. Genomic DNA were extracted from the sieved soil samples (0.25 g) using ZYMO soil DNA extraction Kit (Model D 6001, Zymo Research, USA) following the manufacturer's instructions. Genomic DNA concentration and quality was ascertained using NanoDrop spectrophotometer and electrophoresed on a 0.9% (w/v) agarose gel, respectively.

Shotgun metagenomics of 2S microcosm was prepared using the Illumina Nextera XT sample processing kit and sequenced on a MiSeq. Genomic DNA (50 ng) were fragmented and tagmented and unique indexes were added using reduced-cycle PCR amplification consisting 8 cycles of 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s, and a final extension at 72 °C for 5 min before cooling to 4 °C. Constructed metagenomic libraries were purified with Agencourt AMPure XP beads and quantified with Quant-iT PicoGreen. The library size and quality were validated on Agilent Technologies 2100 Bioanalyzer. Libraries were normalized, pooled in equal volumes and run on a 600 cycles MiSeq Reagent kit v3 (Illumina Inc., San Diego, CA). All samples were multiplexed and sequenced in a single lane on the MiSeq using 2 x 300 bp paired-end sequencing, which generates 20 Mb of data for each sample. Sequence reads were generated in < 65 h, while image analysis and base calling were performed directly on MiSeq. The sequences of 2S metagenome were deposited on the MG-RAST server with the ID 4704691.3. Sequences generated from the microcosm set up were assembled individually by VelvetOptimiser v2.2.5 and the contigs generated were fed into the MG-RAST metagenomic analysis pipeline.

2.4. Taxonomic characterization of metagenomic reads

The taxonomic characterization of the metagenomic reads was determined using the MG-RAST server. A BLAT similarity search for the longest cluster representative in the metagenomes was performed against the MG-RAST M5rna database, which integrated SILVA, Greengenes and RDP databases. The abundance data were identified through the lowest common ancestor (LCA) with the default parameters $1e^{-05}$ as the maximum e-value, 60% as the minimum identity, and a minimum alignment length of 15 as cutoff. In addition, a rarefaction curve was generated for 2S sequence reads according to the data obtained from MG-RAST server and various diversity indices of the metagenome was determined using Mothur v. 1.30.2 (Schloss et al. 2009).

2.5. Functional analyses of metagenomic reads

Gene calling was performed on the 2S contigs using FragGeneScan (Rho et al. 2010) to predict open reading frames (ORFs), which were functionally annotated using KEGG GhostKOALA, the Clusters of Orthologous Groups of proteins (COG) (Tatusov et al. 2001), and the NCBI's conserved domain database (CDD; Marchler-Bauer et al. 2015).

In GhostKOALA, each query gene is assigned a taxonomic category according to the best-hit gene in the Cd-hit cluster supplemented version of the non-redundant pangenome dataset (Kanehisa et al. 2016). In addition, the ORFs were functionally annotated and assigned to the COG database that compares protein sequences encoded in complete genomes, representing major phylogenetic lineage. The 2S metagenome sequence reads was further functionally annotated using the NCBI's conserved domain database (CDSEARCH/cdd v 3.15) using the default blast search parameters. The CDD is a protein annotation resource that consists of a collection of well annotated multiple sequence alignment models for ancient domains and full-length proteins.

3. Results

3.1. Physicochemistry of the soil microcosm

The physicochemical properties of 2S soil microcosm showed that the pH was slightly above neutral (7.20 ± 0.01) while the organic matter content was 1.10%. The total nitrogen, phosphorus and potassium content of the agricultural soil were 0.09%, 3.80 mg/kg, and 0.94 mg/kg respectively.

3.2. General characteristics of 2S metagenome

Illumina sequencing of the DNA from 2S microcosm resulted in 1,512,514 sequence reads. These sequences were assembled into 36,503 unique contigs with a total of 5,779,527 bp, an average length of 158 ± 37 bp, and GC content of 58 ± 4%. After dereplication and quality control by the MG-RAST, the total number of unique contigs in 2S metagenome reduced to 27,320 with 4,197,699 bp, an average length of 154 ± 26 bp, and the GC content of 58 ± 4%. Statistical analysis of the rarefaction curve was conducted to assess the species richness and abundance of the metagenome. As shown in the rarefaction curve, there are 18,484 unique species (phylotypes) in the 2S microcosm with 52% coverage at species delineation (0.03; 97%). Analysis of species abundance and richness revealed that the Shannon index (*H'*) is 9.63, Simpson's index (*D*) and Simpson's reciprocal index (*1/D*) are 0.00005 and 19941.53, Chao index was 47974.87, and Shannon and Simpson's evenness are 0.981 and 1.08, respectively (**Figure 1**).

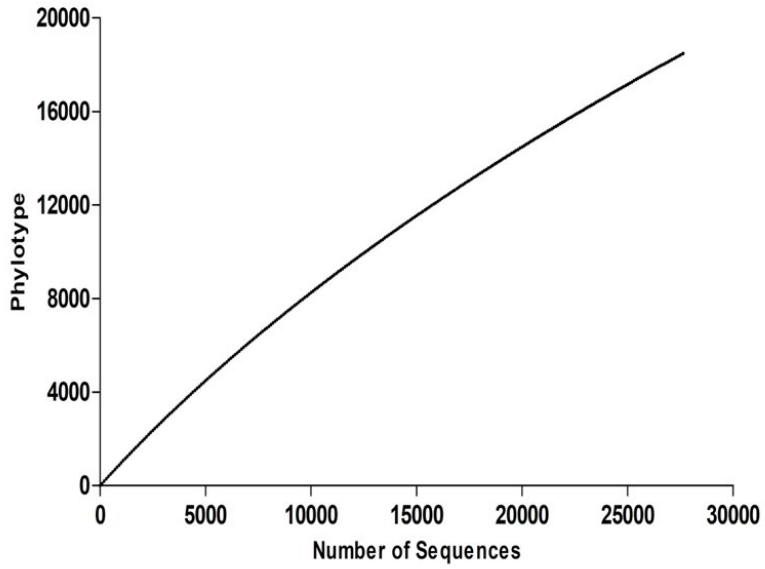

Figure 1. Rarefaction curve of number of unique sequences recovered vs. number of clones sequenced for 2S microcosm. The number of unique sequences (phylotypes) at species delineation (0.03, 97%) is 18,484, the coverage is 51% and the Shannon, Chao, and Simpson's reciprocal indices are 9.63, 47974.87 and 19941.53, respectively.

3.3. Structural diversity of 2S metagenome

Structural analysis of the biota of 2S soil metagenome was conducted using the taxonomic profiles generated by MG-RAST. The three domains, *Bacteria*, *Archaea* and *Eukarya* were duly represented. As expected, the domain *Bacteria* is predominant contributing 96.51% of the domain architecture of the soil biota. This is followed by *Eukarya* and *Archaea* contributing 3.47% and 0.02%, respectively. Other members of the soil biota detected are viruses and unclassified sequence reads that cannot be placed in any of the three domains.

In phylum delineation of 2S metagenome, 40 phyla were retrieved. The predominant phyla were *Actinobacteria* (54.0%), *Proteobacteria* (17.5%), *Firmicutes* (8.5%), *Chloroflexi* (3.6%), and *Planctomycetes* (3.2%), respectively. Other phyla detected in 2S metagenome and their corresponding % abundance is presented in **Figure 2**.

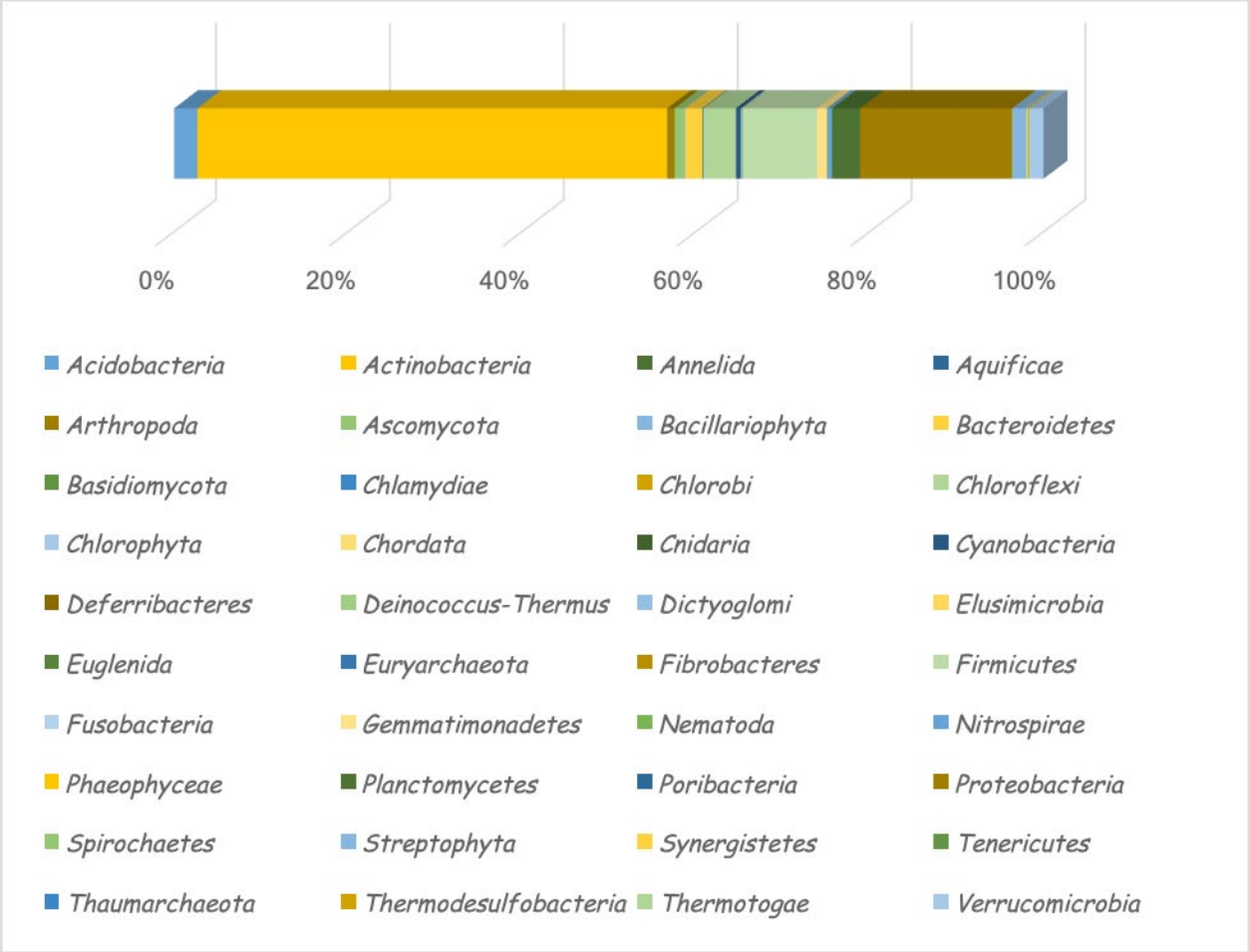

Figure 2. Phylum classification of 2S metagenome. Unclassified reads were not used for the analysis. All the 40 phyla detected in 2S metagenome were used.

In class delineation of 2S metagenome, 78 classes were retrieved from the metagenome. The predominant classes were *Actinobacteria* (55.96%), *Alphaproteobacteria* (5.48%), *Gammaproteobacteria* (4.84%), *Bacilli* (4.83%), and *Deltaproteobacteria* (4.04%), respectively. Other classes detected in 2S metagenome and their corresponding % abundance is presented in **Figure 3**.

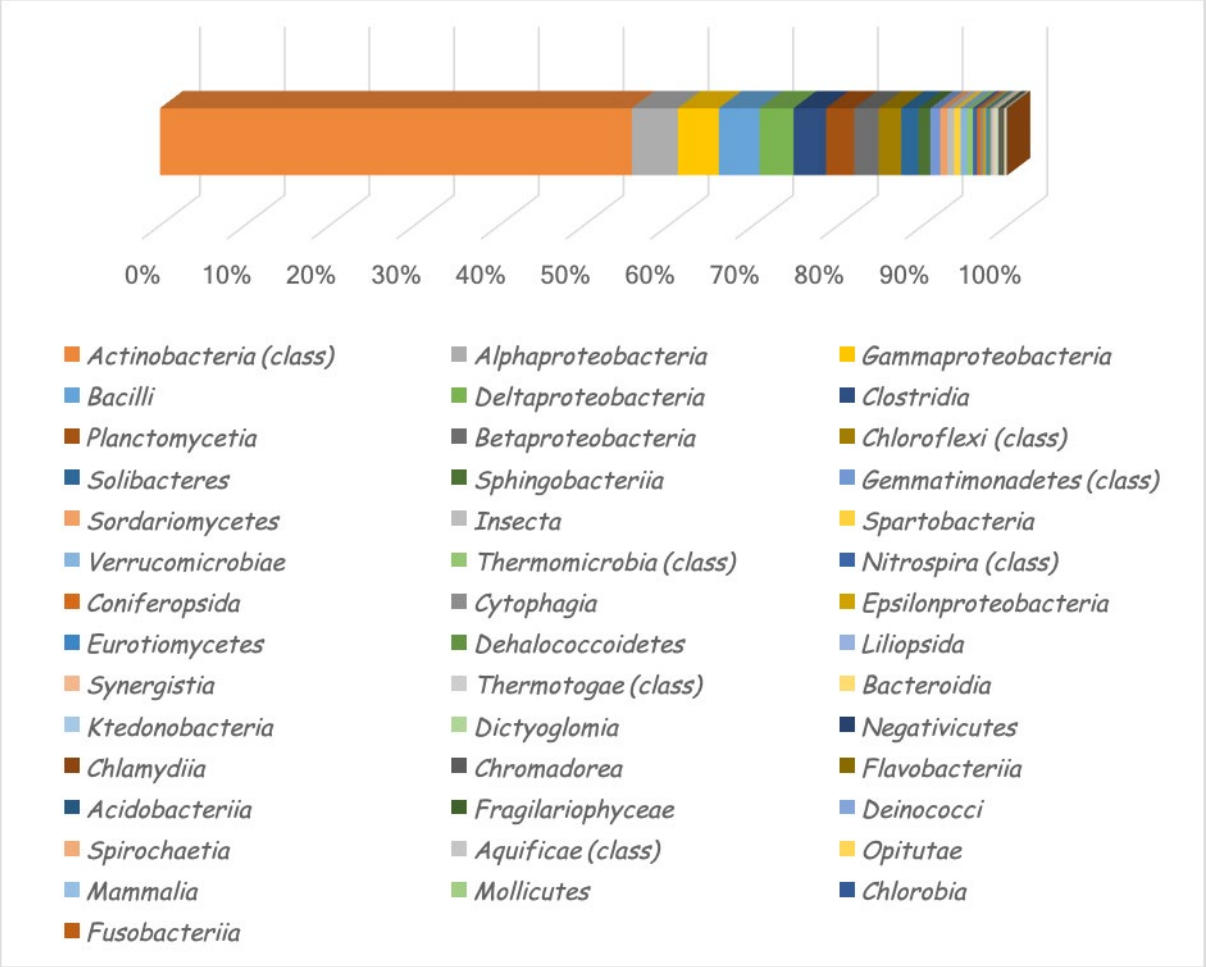

Figure 3. Class delineation of sequence reads from 2S soil microcosm. Unclassified sequence reads were not used. Of the 78 classes recovered from 2S metagenome, only 43 that have sequence reads of > 10 were used for this analysis.

Order delineation of 2S metagenome revealed 157 orders retrieved from the metagenome. The predominant orders were *Actinomycetales* (36.51%), *Solirubrobacterales* (8.45%), *Thermoleophilales* (7.47%), *Bacillales* (4.54%), and *Acidimicrobiales* (3.92%), respectively (**Table S1**, additional information). In family delineation, 273 families were retrieved from the metagenome. The predominant families were *Conexibacteraceae* (8.69%), *Thermoleophilaceae* (7.68%), *Micromonosporaceae* (5.94%), *Pseudonocardiaceae* (4.71%), and *Streptomycetaceae* (4.35%), respectively (**Table S2**, additional information).

In genus delineation, 750 genera were retrieved from 2S metagenome. The predominant genera are *Conexibacter* (8.38%), *Thermoleophilum* (7.40%), *Streptomyces* (4.14%), *Frankia* (3.10%), and *Acidimicrobium* (3.08%), respectively. Other genera recovered from 2S metagenome and their corresponding % abundance is presented in **Figure 4**.

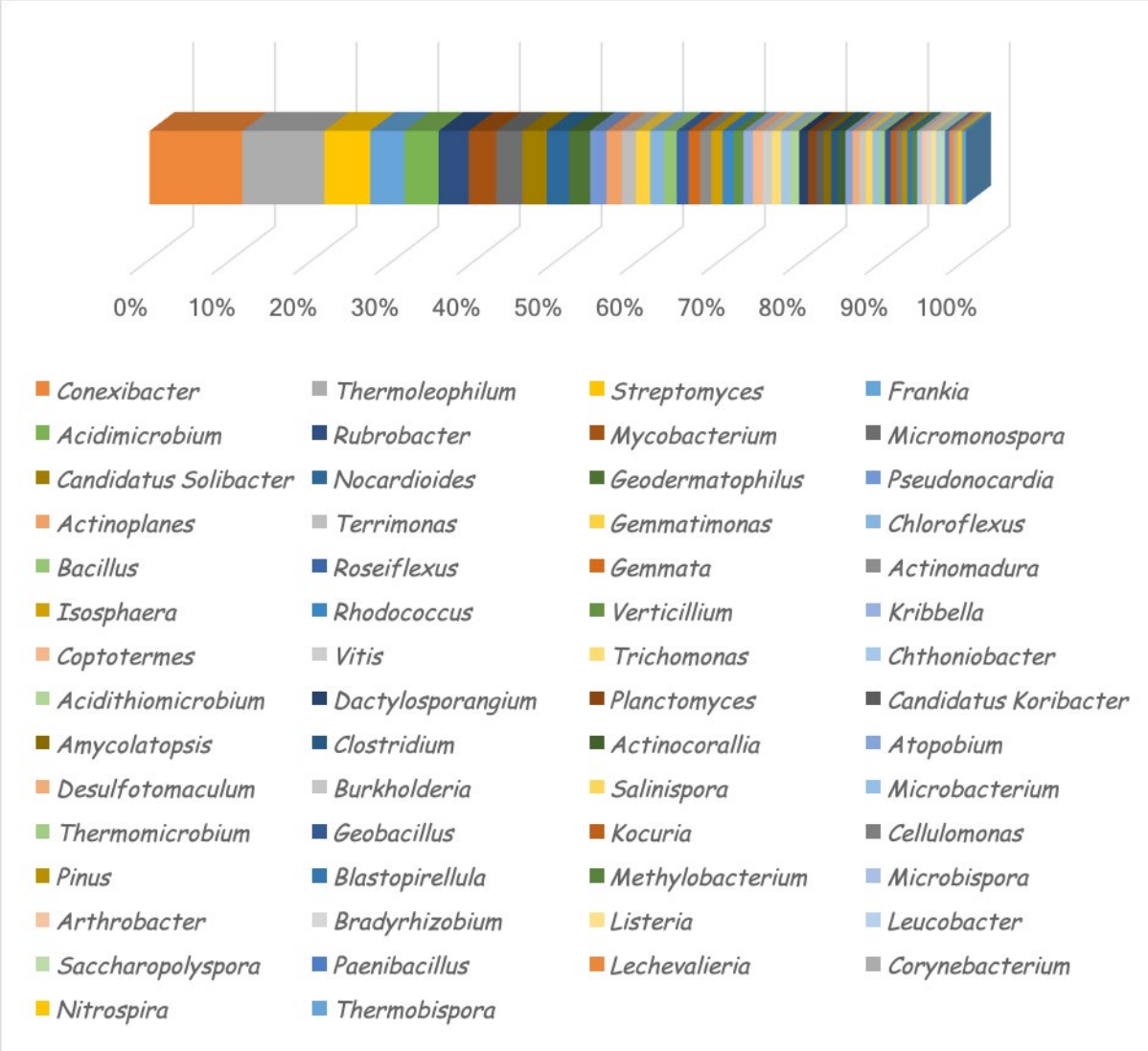

Figure 4. Genus delineation of sequence reads from 2S soil microcosm. Unclassified sequence reads were not used. Of the 750 classified genera recovered from 2S metagenome, only 58 that have sequence reads of > 100 were used for this analysis.

3.4. Functional diversity of 2S metagenome

Carbohydrate metabolism

In carbohydrate metabolism, genes coding for five (5) enzymes were annotated for glycolysis /gluconeogenesis. Five (5) genes were also annotated for TCA cycle. In pentose phosphate pathway, four (4) genes were annotated. Furthermore, functional annotation of 2S metagenome revealed genes that code for nine (9) enzymes for starch and sucrose metabolism. In pyruvate metabolism, genes for pyruvate dehydrogenase E1 component alpha subunit (EC 1.2.4.1), and pyruvate ferredoxin oxidoreductase beta subunit (EC 1.2.7.1) were annotated. Others include acetyl-CoA carboxylase (EC 6.4.1.2), acetyl-CoA/propionyl-CoA carboxylase, biotin carboxylase, biotin carboxyl carrier protein (EC 6.4.1.2 6.4.1.3 6.3.4.14), and acetyl-CoA C-acetyltransferase (EC 2.3.1.9) and several others. Other pathways annotated for carbohydrate metabolism, the genes involved and the microorganisms implicated are presented in **Table 1**.

Taxonomic characterization of the annotated genes for carbohydrate metabolism revealed the predominance of members of the

phylum *Actinobacteria*. Other phyla duly represented were *Proteobacteria* (particularly *Alphaproteobacteria*), *Firmicutes*, *Cyanobacteria* and *Deinococcus-Thermus*.

Amino acid metabolism

Functional characterization of alanine, aspartate and glutamate metabolism revealed genes that code for six (6) enzymes. Annotated for glycine, serine and threonine metabolism include genes for choline dehydrogenase (EC 1.1.99.1), sarcosine oxidase (EC 1.5.3.1), cystathionine beta-synthase (EC 4.2.1.22), and tryptophan synthase alpha chain (EC 4.2.1.20) among others. Annotated for valine, leucine, isoleucine and lysine metabolism include genes for 3-hydroxyisobutyrate dehydrogenase (EC 1.1.1.31), beta-alanine-pyruvate transaminase (EC 2.6.1.18), 3-iso-propylmalate dehydrogenase (EC 1.1.1.85) and diaminopimelate decarboxylase (EC 4.1.1.20) and others. Functional characterization of the genes annotated for histidine, phenylalanine, tyrosine and tryptophan metabolism revealed the presence of histidinol dehydrogenase (EC 1.1.1.23), tryptophan 2,3-dioxygenase (EC 1.13.11.11), 3-phosphoshikimate 1-carboxy-vinyltransferase (EC 2.5.1.19), and prephenate dehydrogenase (EC 1.3.1.12), respectively (**Table 1**).

Taxonomic characterization of the annotated genes for amino acid metabolism revealed the predominance of members of the phyla *Proteobacteria* and *Actinobacteria*. Other phyla with interesting representations are *Verrucomicrobia*, *Bacteroidetes*, *Chlamydiae*, *Euryarchaeota*, *Cyanobacteria*, *Firmicutes*, *Deinococcus-Thermus*, and *Basidiomycota*.

Lipid metabolism

Annotations of genes, which code for enzymes responsible for metabolism of fatty acid, glycerolipid, glycerophospholipid and sphingolipid revealed the presence of 3-oxoacyl-(acyl-carrier-protein) synthase I (EC 2.3.1.41), 3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35), glycerol kinase (EC 2.7.1.30), alpha-galactosidase (EC 3.2.1.22), and phosphatidate cytidylyltransferase (EC 2.7.7.41), which synthesize CDP-diacylglycerol from glecerol-3-phosphate and so on (**Table 1**).

Taxonomic characterization of the annotated enzymes for lipid metabolism revealed the dominance of member of the phyla *Proteobacteria* (*Alphaproteobacteria* class) and *Actinobacteria*.

Energy metabolism

Functional characterization of 2S metagenome for energy metabolism revealed energy generation of the microbial community via diverse routes such as oxidative phosphorylation, photosynthesis and carbon fixation. It also revealed the presence of F-ATPases, the prime producers of ATP, using the proton generated by oxidative phosphorylation or photosynthesis. Also detected are genes coding for enzymes that participated in Calvin cycle (Fructose bisphosphate aldolase, class II), reductive TCA cycle (pyruvate ferredoxin oxidoreductase beta subunit; succinate dehydrogenase/fumarate reductase, iron sulfur subunit; fumarate reductase flavoprotein subunit; isocitrate dehydrogenase), and pathways used by microorganisms for carbon fixation (**Table 2**). Other carbon fixation pathways detected in 2S metagenome are 3-hydroxypropionate bi-cycle, hydroxypropionate/hydroxybutyrate cycle and dicarboxylate/hydroxybutyrate cycle, respectively.

Nitrogen metabolism in 2S metagenome is mediated by nitrate reductase/nitrite oxidoreductase, alpha subunit (EC 1.7.5.1 1.7.99.4). This multifunctional enzyme participated in three metabolic pathways of nitrogen metabolism. These are dissimilatory nitrate reduction, denitrification, and complete nitrification (comammox). Other enzymes/genes annotated for nitrogen metabolism catalyse the transport of nitrogen and biosynthesis of nitrogen-rich amino acids (**Table 2**). In methane metabolism, detected genes include formylmethanofuran dehydrogenase subunit C (EC 1.2.99.5), formate dehydrogenase, Fructose bisphosphate aldolase, class II, and pyruvate ferredoxin oxidoreductase beta subunit, which participated in CO_2 fixation via the reductive TCA cycle.

Table 1. List of the genes (coding for various enzymes) and microorganisms detected in 2S microcosm involved in diverse metabolism of carbohydrates, amino acids, and lipids

Metabolism[a]	Enzyme/Genes[a]	Microorganisms[a]
Carbohydrate Metabolism		
Glycolysis/Gluconeogenesis	Pyrophosphate-fructose-6-phosphate 1-phosphotransferase; fructose-bisphosphate aldolase, class II; pyruvate dehydrogenase E1 component alpha subunit; pyruvate ferredoxin oxidoreductase beta subunit; 6-phospho-beta-glucosidase	**Actinobacteria** (*Xylanimonas*, *Cellulomonas*, *Kytococcus, Rubrobacter, Actinoplanes*); **Firmicutes** (*Thermobacillus, Enterococcus*)
Citrate cycle (TCA cycle)	Isocitrate dehydrogenase; succinate dehydrogenase/fumarate reductase, iron-sulfur subunit; fumarate reductase flavoprotein subunit; pyruvate dehydrogenase E1 component alpha subunit; pyruvate ferredoxin oxidoreductase beta subunit	**Actinobacteria** (*Olsenella, Eggerthella, Cellulomonas, Kytococcus, Rubrobacter*); **Alphaproteobacteria** (*Oligotropha*)
Pentose phosphate pathway	Glucose-6-phosphate 1-dehydrogenase; 2-dehydro-3-deoxy-phosphogluconate aldolase; fructose-bisphosphate aldolase, class II; Pyrophosphate--fructose-6-phosphate 1-phosphotransferase	**Betaproteobacteria** (*Verminephrobacter*); **Alphaproteobacteria** (*Methylobacterium*); **Gammaproteobacteria** (*Escherichia*); **Firmicutes** (*Thermobacillus*); **Actinobacteria** (*Xylanimonas*)
Pentose and glucuronate interconversion	Ribulose 5-phosphate 4-epimerase; xylulokinase	**Firmicutes** (*Leuconostoc*); **Deinococcus-Thermus** (*Truepera*); **Actinobacteria** (*Streptosporangium*)
Fructose and mannose metabolism	Mannose-6-phosphate isomerase; mannose-1-phosphate guanylyltransferase; Pyrophosphate--fructose-6-phosphate 1-phosphotransferase; mannitol 2-dehydrogenase; fructose-bisphosphate aldolase, class II	**Cyanobacteria** (*Oscillatoria*); **Firmicutes** (*Kyrpidia, Thermobacillus*); **Actinobacteria** (*Xylanimonas*); **Alphaproteobacteria** (*Methylobacterium*)
Galactose metabolism	Galactokinase; UDP-glucose-4-epimerase; UDP-galactopyranose mutase; alpha galactosidase; galactose oxidase	**Actinobacteria** (*Brachybacterium, Microbacterium, Streptomyces*); **Cyanobacteria** (*Cylindrospermum*); **Firmicutes** (*Lactococcus*)
Starch and sucrose metabolism	Beta-glucosidase; cellulose synthase (UDP forming); 6-phospho-beta-glucosidase; glycogen phosphorylase; 4-alpha-glucanotransferase; maltokinase; isoamylase; (1-4)-alpha-D-glucan-1-alpha-D-glucosylmutase; trehalose-6-phosphate phosphatase	**Firmicutes** (*Paenibacillus, Enterococcus*); **Actinobacteria** (*Modestobacter, Actinoplanes, Saccharomonospora, Aeromicrobium, Brachybacterium, Nocardiopsis, Ilumatobacter*); **Alphaproteobacteria** (*Gluconobacter, Granulibacter, Methylocella, Hyphomicrobium*); **Gammaproteobacteria** (*Thioalkalimicrobium*); **Cyanobacteria** (*Arthrospira*)
Amino sugar and nucleotide sugar metabolism	Bifunctional UDP-N-acetylglucosamine pyrophosphorylase/glucosamine 1-phosphate N-acetyltransferase; UDP-N-acetylmuramate dehydrogenase; UDP-arabinose-4-epimerase; galactokinase; UDP-glucose-4-epimerase; UDP-galactopyranose mutase; mannose-1-phosphate guanylyltransferase; Mannose-6-phosphate isomerase	**Actinobacteria** (*Rhodococcus, Brachybacterium*); **Firmicutes** (*Lactococcus, Kyrpidia*); **Alphaproteobacteria** (*Hirschia, Altererythrobacter, Rhodospirillum*); **Cyanobacteria** (*Cylindrospermum, Oscillatoria*)
Pyruvate metabolism	pyruvate ferredoxin oxidoreductase beta subunit; pyruvate dehydrogenase E1 component alpha subunit; acetyl-CoA carboxylase, biotin carboxylase subunit; acetyl-CoA/propionyl-CoA carboxylase, biotin carboxylase, biotin carboxyl carrier protein; fumarate reductase flavoprotein subunit; acetyl-CoA C-acetyltransferase	**Actinobacteria** (*Rubrobacter, Cellulomonas, Kytococcus, Brevibacterium, Eggerthella, Kocuria*); **Alphaproteobacteria** (*Shinella*)
Glyoxylate and dicarboxylate metabolism	Acetyl-CoA C-acetyltransferase; acetyl-CoA/propionyl-CoA carboxylase, biotin carboxylase, biotin carboxyl carrier protein; glutamine synthetase; formate dehydrogenase iron sulfur subunit	**Actinobacteria** (*Kocuria, Brevibacterium, Corynebacterium*); **Alphaproteobacteria** (*Acidiphilium*)

Table 1. List of the genes (coding for various enzymes) and microorganisms detected in 2S microcosm involved in diverse metabolism of carbohydrates, amino acids, and lipids

Metabolism[a]	Enzyme/Genes[a]	Microorganisms[a]
Propionate metabolism	Pyruvate ferredoxin oxidoreductase beta subunit; acetyl-CoA carboxylase, biotin carboxylase subunit; acetyl-CoA/propionyl-CoA carboxylase, biotin carboxylase, biotin carboxyl carrier protein; 4-aminobutyrate aminotransferase/(S)-3-amino-2-methylpropionate transaminase; beta-alanine---pyruvate transaminase; acetyl-CoA C-acetyltransferase	**Actinobacteria** (*Rubrobacter, Brevibacterium, Kocuria*); **Alphaproteobacteria** (*Shinella, Polymorphum, Blastochloris*)
Butanoate metabolism	Acetyl-CoA C-acetyltransferase; 3-hydroxyacyl-CoA dehydrogenase; crotonyl-CoA reductase; succinate dehydrogenase/fumarate reductase, iron-sulfur subunit; fumarate reductase flavoprotein subunit; 4-aminobutyrate aminotransferase/(S)-3-amino-2-methylpropionate transaminase; pyruvate ferredoxin oxidoreductase beta subunit; acetolactate synthase I/II/II large subunit	**Actinobacteria** (*Kocuria, Kitasatospora, Eggerthella, Rubrobacter, Blastococcus*); **Alphaproteobacteria** (*Rhizobiales bacterium NRL2, Xanthobacter, Oligotropha, Polymorphum*)
Amino acid Metabolism		
Alanine, aspartate and glutamate metabolism	Alanine dehydrogenase, 4-aminobutyrate aminotransferase/(S)-3-amino-2-methylpropionate transaminase; glutamate synthase (NADPH/NADH) small chain; glutamate dehydrogenase; RHH-type transcriptional regulator, proline utilization regulon repressor/proline dehydrogenase/delta 1-pyrroline-5-carboxylate dehydrogenase; glutamine synthetase	**Verrucomicrobia** (*Metylacidiphilum*); **Firmicutes** (*Paenibacillus, Syntrophothermus*); **Bacteroidetes** (*Arachidicoccus*); **Deltaproteobacteria** (*Syntrophus, Haliangium, Sorangium*); **Alphaproteobacteria** (*Polymorphum, Aureimonas*); **Gammaproteobacteria** (*Halothiobacillus, Enterobacteriaceae bacterium* strain FGI 57); **Haptophyta** (*Emiliania*); **Actinobacteria** (*Streptomyces, Corynebacterium*)
Glycine, serine and threonine metabolism	Choline dehydrogenase; sarcosine oxidase, alpha subunit; cystathionine beta-synthase; cystathionine gamma-lyase; tryptophan synthase alpha chain; L-2,4-diaminobutyrate decarboxylase	**Gammaproteobacteria** (*Erwinia*); **Alphaproteobacteria** (*Starkeya, Caulobacteraceae bacterium*); **Actinobacteria** (*Micrococcus, Thermobifida, Brevibacterium*)
Cysteine and methionine metabolism	Cysteine synthase A; cystathionine gamma-lyase; cystathionine beta-synthase; glutamate-cysteine ligase; thiosulfate/3-mercaptopyruvate sulfurtransferase	**Gammaproteobacteria** (*Pantoea*); **Alphaproteobacteria** (*Caulobacteraceae bacterium, Pelagibacterium, Agrobacterium*); **Actinobacteria** (*Micrococcus*); **Deltaproteobacteria** (*Archangium*)
Valine, leucine isoleucine metabolism	Acetyl-CoA/propionyl-CoA carboxylase, biotin carboxylase, biotin carboxyl carrier protein; 3-hydroxyisobutyrate dehydrogenase; beta-alanine-pyruvate transaminase; 4-aminobutyrate aminotransferase/(S)-3-amino-2-methylpropionate transaminase; acetyl-CoA C-acetyltransferase; 3-isopropylmalate dehydrogenase; acetolactate synthase I/II/II large subunit	**Actinobacteria** (*Brevibacterium, Leifsonia, Kocuria, Blastococcus*); **Alphaproteobacteria** (*Blastochloris, Polymorphum*); **Firmicutes** (*Alicyclobacillus*)
Lysine metabolism	Diaminopimelate decarboxylase; UDP-N-acetylmuramoyl-L-alanyl-D-glutamate-2,6-diaminopimelate ligase; acetyl-CoA C-acetyltransferase	**Actinobacteria** (*Microterricola, Jonesia, Kocuria*); **Firmicutes** (*Halobacillus*); **Deltaproteobacteria** (*Chondromyces*); **Chlamydiae** (*Waddlia*)
Arginine and proline metabolism	Arginase, glutamine synthetase, glutamate dehydrogenase; glutamate N-acetyltransferase/amino acid N-acetyltransferase; creatinine amidohydrolase; N-methylhydantoinase B; amidase; proline dehydrogenase; RHH-type transcriptional regulator, proline utilization regulon repressor/proline dehydrogenase/delta 1-pyrroline-5-carboxylate dehydrogenase	**Alphaproteobacteria** (*Roseobacter, Ketogulonicigenium, Jannaschia*); **Actinobacteria** (*Corynebacterium, Slackia, Streptomyces, Nocardiopsis, Actinoplanes*); **Haptophyta** (*Emiliania*); **Basidiomycota** (*Ustilago*); **Euryarchaeota** (*Salinarchaeum*)

Table 1. List of the genes (coding for various enzymes) and microorganisms detected in 2S microcosm involved in diverse metabolism of carbohydrates, amino acids, and lipids

Metabolism[a]	Enzyme/Genes[a]	Microorganisms[a]
Histidine metabolism	ATP phosphoribosyltransferase regulatory subunit; phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase; histidinol dehydrogenase; gamma-glutamyl hercynylcysteine S-oxide hydrolase	**Alphaproteobacteria** (*Halocynthiibacter*); **Firmicutes** (*Alicyclobacillus*); **Betaproteobacteria** (*Methylobacillus*); **Cyanobacteria** (*Calothrix, Crinalium*); **Actinobacteria** (*Mycobacterium*)
Phenylalanine, tyrosine and tryptophan metabolism	Amidase; tryptophan 2,3-dioxygenase; acetyl-CoA C-acetyltransferase; quinate dehydrogenase (quinone); 3-phosphoshikimate 1-carboxyvinyltransferase; tryptophan synthase alpha chain; prephenate dehydrogenase	**Betaproteobacteria** (*Burkholderia*); **Cyanobacteria** (*Trichormus*); **Deinococcus-Thermus** (*Deinococcus*); **Alphaproteobacteria** (*Micavibrio, Dinoroseobacter*); **Actinobacteria** (*Kocuria, Salinispora, Thermobifida, Nocardiopsis*)
Lipid Metabolism		
Fatty acid metabolism	Acetyl-CoA carboxylase, biotin carboxylase subunit; acetyl-CoA/propionyl-CoA carboxylase, biotin carboxylase, biotin carboxyl carrier protein; 3-oxoacyl-(acyl-carrier-protein) synthase I; acetyl-CoA C-acetyltransferase; 3-hydroxyacyl-CoA dehydrogenase	**Alphaproteobacteria** (*Shinella, Ketogulonicigenium, Rhizobiales bacterium, Xanthobacter*); **Actinobacteria** (*Brevibacterium, Kocuria*); **Betaproteobacteria** (*Ottowia*)
Glycerolipid, glycerophospholipid and sphingolipid metabolism	Glycerol kinase; alpha-galactosidase; ethanolamine ammonia-lyase small subunit; phosphatidate cytidylyltransferase; neutral ceramidase	**Alphaproteobacteria** (*Defluviimonas, Caulobacteraceae bacterium*); **Actinobacteria** (*Microbacterium, Frankia, Gordonia*)

[a]KEGG GhostKOALA, COG and NCBI's CDD were used for functional annotation of 2S metagenome.

Table 2. List of the genes (coding for various enzymes) and microorganisms detected in 2S microcosm involved in energy metabolism, biosynthesis of secondary metabolites, antibiotic resistance and quorum sensing

Metabolism[a]	Enzyme/Genes[a]	Microorganisms[a]
Energy Metabolism		
Oxidative phosphorylation	NADH-quinone oxidoreductase subunit J; NADH-quinone oxidoreductase subunit M; NADH-quinone oxidoreductase subunit N; succinate dehydrogenase/fumarate reductase, iron sulfur subunit; fumarate reductase flavoprotein subunit; cytochrome c oxidase assembly protein subunit 15; F-type H+-transporting ATPase gamma subunit	**Actinobacteria** (*Saccharomonospora, Catenulispora, Eggerthella, Actinosynnema, Propionibacterium*); **Gammaproteobacteria** (*Raoultella, Klebsiella*); **Alphaproteobacteria** (*Oligotropha*); **Green algae** (*Volvox*); **Betaproteobacteria** (*Achromobacter*); **Bacteroidetes** (*Flavobacteriaceae bacterium*)
Photosynthesis	Photosystem I subunit XI; F-type H^+-transporting ATPase gamma subunit	**Cyanobacteria** (*Pleurocapsa*); **Actinobacteria** (*Propionibacterium*); **Betaproteobacteria** (*Achromobacter*)
Carbon fixation	Fructose bisphosphate aldolase, class II; pyruvate ferredoxin oxidoreductase beta subunit; succinate dehydrogenase/fumarate reductase, iron sulfur subunit; fumarate reductase flavoprotein subunit; acetyl-CoA C-acetyltransferase; isocitrate dehydrogenase; acetyl-CoA carboxylase, biotin carboxylase subunit	**Firmicutes** (*Thermobacillus*); **Actinobacteria** (*Rubrobacter, Eggerthella, Kocuria, Olsenella*); **Alphaproteobacteria** (*Oligotropha, Shinella*)
Methane metabolism	Formate dehydrogenase iron sulfur subunit; Fructose bisphosphate aldolase, class II; formylmethanofuran dehydrogenase subunit C; pyruvate ferredoxin oxidoreductase beta subunit	**Alphaproteobacteria** (**Acidiphilium, Hyphomicrobium**); **Firmicutes** (**Thermobacillus**); **Actinobacteria** (**Rubrobacter**)
Methane metabolism	Formate dehydrogenase iron sulfur subunit; Fructose bisphosphate aldolase, class II; formylmethanofuran dehydrogenase subunit C; pyruvate ferredoxin oxidoreductase beta subunit	**Alphaproteobacteria** (*Acidiphilium, Hyphomicrobium*); **Firmicutes** (*Thermobacillus*); **Actinobacteria** (*Rubrobacter*)
Sulfur metabolism	Sulfate adenylyltransferase subunit 2; thiosulfate/3-mercaptopyruvate sulfurtransferase; cysteine synthase A	**Actinobacteria** (*Lentzea*); **Alphaproteobacteria** (*Agrobacterium*); **Gammaproteobacteria** (*Pantoea*)
Biosynthesis of secondary metabolites and antibiotic resistance		
Type I polyketide structures	AmphN, nysN, fscP, pimG, cytochrome P450 monooxygenase; candicidin polyketide synthase FscD	**Actinobacteria** (*Pseudonocardia, Streptomyces*)
Biosynthesis of ansamycins	5-deoxy-5-amino-3-dehydroquinate synthase	**Actinobacteria** (*Streptomyces*)
Biosynthesis of type II pokyketide backbone	Minimal PKS chain-length factor (CLF/KS beta)	**Actinobacteria** (*Catenulispora*)
Biosynthesis of tetracycline, monobactam, streptomycin and novobiocin	Acetyl-CoA catrboxylase, biotin carboxylase subunit; acetyl-CoA/propionyl-CoA carboxylase, biotin carboxylase, biotin carboxyl carrier protein; Minimal PKS chain-length factor (CLF/KS beta); sulfate adenylyltransferase subunit 2; dTDP-4-dehydrorhamnose 3,5-epimerase; prephenate dehydrogenase	**Alphaproteobacteria** (*Shinella*); **Actinobacteria** (*Brevibacterium, Catenulispora, Lentzea, Nocardiopsis*); **Cyanobacteria** (*Geminocystis*)

Table 2. List of the genes (coding for various enzymes) and microorganisms detected in 2S microcosm involved in energy metabolism, biosynthesis of secondary metabolites, antibiotic resistance and quorum sensing

Metabolism[a]	Enzyme/Genes[a]	Microorganisms[a]
Beta-lactam resistance	Beta-lactamase class C; oligopeptide transport system substrate-binding protein; oligopeptide transport system ATP-binding protein; acrB, mexB, adeJ, smeE, mtrD, cmeB, multidrug efflux pump; penicillin-binding protein 1A; penicillin-binding protein 2A; cell division protein Ftsl (penicillin-binding protein 3)	**Gammaproteobacteria** (*Cronobacter, Pantoea, Plautia stali symbiont*); **Cyanobacteria** (*Gloeobacter*); **Aphaproteobacteria** (*Rhodoplanes, Parvibaculum, Sphingomonas*); **Betaproteobacteria** (*Nitrosomonas, Polaromonas*), **Epsilonproteobacteria** (*Arcobacter*); **Deltaproteobacteria** (*Geobacter*); **Firmicutes** (*Geobacillus*)
Vancomycin, cationic antimicrobial peptide (CAMP) and fluoroquinolone resistance	Alanine racemase; D-alanyl-D-alanine carboxypeptidase; acrB, mexB, adeJ, smeE, mtrD, cmeB multidrug efflux pump; repressor LexA	**Alphaproteobacteria** (*Bradyrhizobium, Parvibaculum*); **Actinobacteria** (*Stackebrandtia, Cellulomonas*); **Firmicutes** (*Bacillus selenitireducens*)
Quorum sensing	Trans-2,3-dihydro-3-hydroxyanthranilate isomerase; two-component system, OmpR family, KDP operon response regulator KdpE; branched-chain amino acid transport system substrate-binding protein; Fur family transcriptional regulator, zinc uptake regulator; oligopeptide transport system substrate-binding protein; oligopeptide transport system ATP-binding protein; peptide/nickel transport system substrate-binding protein; peptide/nickel transport system ATP-binding protein; fused signal recognition particle receptor; signal recognition particle subunit SRP54; lantibiotic biosynthesis protein	**Gammaproteobacteria** (*Serratia, Pantoea*); **Alphaproteobacteria** (*Caulobacteraceae bacterium, Bradyrhizobium, Magnetospirillum, Defluviimonas, Rhodoplanes, Rhodospirillum*); **Actinobacteria** (*Ilumatobacter, Pseudonocardia, Streptomyces*); **Cyanobacteria** (*Gloeobacter*); **Euryarchaeota** (*Haloterrigena, Halomicrobium*); **Deinococcus-Thermus** (*Truepera*)

[a]KEGG GhostKOALA, COG and NCBI's CDD were used for functional annotation of 2S metagenome.

Sulfur metabolism in 2S metagenome revealed the presence of sulfate adenylyltransferase subunit 2 (EC 2.7.7.4) implicated in assimilatory and dissimilatory sulfate reduction; thiosulfate/3-mercaptopyruvate sulfurtransferase (2.8.1.1 2.8.1.2), which facilitate the transfer of a sulfur ion from a donor to cyanide or to other thiol compounds. Also detected is cysteine synthase A (EC 2.5.1.47), which catalyse the formation of cysteine from O-acetyl-serine and hydrogen sulphide with the concomitant release of acetic acid.

Taxonomic characterization of annotated enzymes for energy metabolism revealed the dominance of members of *Proteobacteria* and *Actinobacteria* phyla. Other phyla with representations include *Cyanobacteria*, *Firmicutes* and *Bacteroidetes* (**Table 2**).

Biosynthesis of secondary metabolites and antibiotic resistance

Functional characterization of 2S metagenome showed the presence of type I and type II polyketide synthases and backbone such as candicidin polyketide synthase FscD, amphN, nysN, fscP, pimG, cytochrome P450 monooxygenase, Minimal PKS chain-length factor (CLF/KS beta); genes for enzymes that catalyse the biosynthesis of ansamycins, tetracyclines, monobactam, streptomycin and novobiocin, among others. Also detected were various genes for enzymes implicated in resistance to beta-lactam antibiotics, vancomycin, fluoroquinolones and cationic antimicrobial peptide (CAMP) (**Table 2**). Taxonomic characterization of the detected genes for biosynthesis of secondary metabolites revealed the predominance of the phylum *Actinobacteria*. However, for antibiotic resistance genes, the phylum *Proteobacteria* dominates.

4. Discussion

Soil microbial community structure and functions are highly influenced by a combination of factors, including soil structure and available nutrient, and these could be veritable predictors of soil health. It is not surprising; therefore, that one major objective of metagenomic studies is to gain insight into soil elemental cycles (Myrold et al. 2013). Nitrogen, phosphorus and potassium are essential nutrients that play a major role in crop production and rapid depletion of these nutrients because of intensive and sometimes poor agricultural practices has occasioned reliance on chemical fertilizers to amend the nutrient deficiency (Sharma et al. 2013; Rashid et al. 2016).

The pH value (7.2) of the soil is close to neutral, which is considered favourable for most microorganisms, particularly bacteria that are essential for soil fertility and health (Lauber et al. 2009). The amounts of these nutrients recorded in 2S soil are very low compared with the values acceptable for fertile agricultural soils. However, these values are similar to those earlier reported by Salam et al. (2017) with N, P, K values of 0.04, 3.21 and 0.89 mg/kg respectively. Low levels of these nutrients in soils may also reflect a high level of microbial activity with demand on these nutrients, especially in soil with high organic matter (Ilori et al. 2015).

The identification of *Actinobacteria, Proteobacteria, Firmicutes, Chloroflexi*, and *Planctomycetes* as the predominant phyla among the identifiable 40 phyla in the sequences recovered from 2S soil is in consonance with previous reports, which had shown this group as preponderant members in agricultural soils (Trivedi et al. 2016; Salam et al. 2017; Yin et al. 2017). Cheema et al. (2015) identified twelve known bacterial groups in an unpolluted agricultural soil with the predominant ones being *Proteobacteria* (41%) and *Actinobacteria* (34%).

It is quite noteworthy that in our study, *Actinobacteria* alone accounted for 54.0% of the bacterial phyla identified. This is not surprising as members of the group are the largest taxonomic group among the recognized lineages of bacteria

(Ludwig et al. 2012) well adapted to oligotrophic conditions and are known to produce cocktails of enzymes which have growth promoting effects on plants (Sathya et al. 2017). They are reputable colonizers of soils and have also been linked with soil organic matter production and humus formation, perhaps on account of their ability to produce extracellular enzymes for degradation of macromolecules such as lignin, cellulose, chitin and starch (Silva et al. 2013).

The *Proteobacteria*, which constitute the second most abundant phyla in this study is notable not only as the largest group of Gram-negative bacteria but also as one with a huge morphological, physiological and metabolic diversity (Spain et al. 2009). They are highly represented in soils receiving high rate of carbon input and boasts of a significantly high number of unculturable or yet to be cultured species (Spain et al. 2009; Fierer et al. 2012). The *Firmicutes* are Gram-positive bacteria with many of the terrestrial species demonstrating high level of hardihood as spore formers. Members of this group may be favoured in agricultural soil systems because of their remarkable ability to resist desiccation because of sharp variation in soil surface temperature (Montecchia et al. 2015).

It is noteworthy that in terms of class delineation, all classes of *Proteobacteria* were represented in the 2S metagenome, with *Alphaproteobacteria* (5.48%), *Gammaproteobacteria* (4.84%) and *Deltaproteobacteria* (4.04%) as the most prominent. Some of the *Alphaproteobacteria* are particularly important as plant mutualists and pathogens and are equally well represented among primary producers (Williams et al. 2007). The prominence of *Gammaproteobacteria* in 2S metagenome may be attributed to the fact that this group that harbors the richest number of genera among bacterial classes and exhibits broad tropism, aerobicity and temperature tolerance, with symbiotic relationship towards plants, invertebrates and vertebrates (Garrity et al. 2005; Williams et al. 2010).

The predominance of the families belonging to *Actinobacteria* among the classified reads of the 2S metagenome further buttresses the importance of the *Actinobacteria* as the major player in the 2S soil biodiversity. *Micromonosporaceae*, well represented as the third most abundant family identified in the 2S metagenome have many strains that have been isolated from soil samples from both temperate and tropical environments (Ara and Kudo 2007; Monciardini et al. 2009) and roots of plants (Wang et al. 2013; Matsumoto et al. 2014). As endophytes of several crop plants, they have recently gained attention because of their importance in nitrogen fixation (Trujillo et al. 2015). Their possible role in the degradation of organics due to several cellulases they produce is well highlighted in the literature (de Menezes et al. 2008, 2012). It is therefore not surprising that they are part of the predominant clade in this study.

Conexibacter, which is the genus with the most abundant sequence read in the metagenome, is a very small genus of soil organisms with only two species so far characterized. They are aerobic and motile and tolerate pH in both acidic and alkali ranges and psychrophilic and mesophilic temperature (Monciardini et al. 2003; Seki et al. 2012). Their predominance in the agricultural soil may be connected with their ability to withstand the diurnal and seasonal changes in temperature characteristic of such tropical soil. The fact that they are also motile may also have played a role in their prompt colonization of wider reaches of the soil than other non-motile members of the community.

The presence of *Streptomyces* (in the family *Streptomycetaceae*) as one of the most abundant classified reads is not unexpected as previous reports from diverse environments had shown similar trend (Lipson and Schmidt 2004; Koberl et al. 2011; Shivlata and Satyanarayana 2015; Salam et al. 2017). Members of this genus are not only known to be highly ubiquitous but are mostly responsible for the decomposition of dead plant matter and maintenance of soil fertility (Thakur et al. 2007; Seipke et al. 2011). Their dominant role in this regard may be attributed to their ability to efficiently degrade cellulose and other complex plant material and production of antibiotics (Mellouli et al. 2003; Chater 2006).

A general review of the functional diversity of the 2S metagenome revealed that core metabolic pathways for carbohydrates, amino acid and lipid metabolism are well represented as

indicated by the annotation of the genes for the key enzymes of these pathways. Furthermore, the *Actinobacteria*, which constitutes the predominant phylum in the metagenome in the community, was also revealed as the predominant group annotated for most enzyme/ genes for carbohydrate metabolism and was second only to the *Proteobacteria* in annotation of genes for amino acids and lipid metabolism (**Table 1**).

It had previously been noted that while it might be difficult to correlate individual functional genes with community structures, overall functional attributes of soil microbial communities could be broadly predicted based on the taxonomic or phylogenetic structure of the communities (Fierer et al. 2012). Therefore, the *Actinobacteria* and *Proteobacteria*, which constitute the major phyla in the soil community structure are likely to be the key functional players. Screening of forest soil samples in Brazil for functional key enzymes showed that the *Actinobacteria* and *Proteobacteria* were the predominant players (Pacchioni et al. 2014). Qin et al. (2016) unequivocally highlighted the fact that the *Actinobacteria* play diverse functional roles in the soil environment, which makes them the key drivers of the microbial community structure and functions.

Glycolysis is the major pathway used by most microorganisms to break down hexose sugars (Moats et al. 2002). The decarboxylation of pyruvate and CoA activation to produce Acetyl CoA, a key precursor in the TCA cycle and lipid metabolism, is a very important precondition for aerobic metabolism of carbohydrates. The annotation of the genes for some of the key enzymes in these processes is not surprising as the predominant families in the microbiome are aerobic organisms.

Also, of interest is the annotation of the enzymes of the pentose phosphate pathway. Usually a subsidiary pathway of hexose breakdown, it is a very important pathway, which provide the precursors for riboses found in other sugars and nucleic acids and generates NADPH, which is a major source of reducing power in biosynthetic reactions for many prokaryotes and eukaryotes (Zubay et al. 1995). It is also the source of other precursors and sugars such as erythrose-phosphate which is the precursor of aromatic amino acids and sedoheptulose-phosphate which is required in small amounts by most cells.

The annotation of enzymes for starch and sucrose metabolism, particularly enzyme such as β-glucosidase suggests that these carbohydrates are a major source of metabolizable primary substrates in the community. Furthermore, our result underscores the presupposed role of the *Actinobacteria* as the group mainly responsible for degradation of organic matters. Members of the *Actinobacteria* are known to possess a variety of cellulases and other enzyme for metabolism of carbohydrate hetero- and homopolymers. Wibberg et al. (2016) reported that *Streptomyces recticuli* has an estimated 456 genes for proteins involved in the utilization of cellulose and other complex and simple carbohydrate substrates.

A cursory look at the genes annotated for amino acids metabolism show that generally, members of the phylum *Proteobacteria* were more predominant players than other phyla including *Actinobacteria* even though it is second in the entire community (**Table 1**). This is in line with previous reports, which highlighted the predominance of genes annotated for this group in soil metagenomic functions related to amino acid degradation, ammonia assimilation and nitrate and nitrite ammonification (Li et al. 2014; Castaneda and Barbosa 2017). Also noteworthy is the annotation of genes for the major enzymes of ammonia assimilation including glutamate dehydrogenase, alanine dehydrogenase, proline dehydrogenase, glutamate synthase and glutamine synthetase and the preponderance of these genes in less prominent phyla of the 2S metagenome such as *Verrucomicrobia, Firmicutes* and *Bacteroidetes*.

The characterization of *Basidiomycota* among the genes annotated for the amino acids arginine and proline (**Table 1**) is interesting but equally explicable. The predominant species among these, *Ustilago maydis*, is edible smut fungus and an endophytic parasite of grass (Pan et al. 2008), which is rich in free amino acids content including lysine, glycine, valine, leucine and glutamic acid and other unusual amino acids (Lizarraga-Guerra and López 1996). Thus, this fungus would appear to be an important

contributor to the metabolism of these amino acids in the 2S microbiome.

Lipid metabolism is generally well represented in various ecosystems as lipid derivatives are an important part of the cell envelope as the lipid bilayer (phospholipids) or cell wall components. Thus, it is widely used as an important community structure marker and measure of soil health (Bossio et al. 1998; Calderón et al. 2001; Islam et al. 2009; Kunihiro et al. 2014; Jiang et al. 2016; Castaneda and Barbosa 2017). The enzymes annotated for metabolism of fatty acids, glycerolipid, glycerophospholipid and sphingolipid in the 2S soil metagenome are in the classes that have been well elucidated in the literature. The irreversible conversion of acetyl-CoA to malonyl-CoA by the enzyme acetyl-CoA carboxylase (**Table 1**) which is a multicomponent enzyme system is a universal link between the carbohydrates metabolism and fatty acid synthetic pathways (Zubay et al. 1995). The dominance of members of the phyla *Proteobacteria* and *Actinobacteria* in the enzyme annotated for lipid metabolism is in consonance with their dominant role within the 2S soil metagenome.

Energy metabolism by the 2S community encompassed oxidative phosphorylation, photosynthesis, carbon fixation and methane metabolism (**Table 2**). Genes for oxidative phosphorylation were most abundantly annotated for members of the *Actinobacteria* and *Proteobacteria*. This is reasonable as they are equally the most abundant groups in the community. Both groups are known to be well equipped with a cocktail of genes with functional propensity for cycling of carbon, nitrogen, sulfur and other elements in the biogeochemical cycle (Trujillo et al. 2015; Shivlata and Satyanarayana 2015).

The major transformations of nitrogen are nitrogen fixation, nitrification, denitrification, anammox and ammonification and these have been shown to be all highly dependent on diverse assemblage of microorganisms in the environment (Costa et al. 2015). In the 2S metagenome analysis, it was revealed that nitrate reductase/nitrate oxidoreductase alpha subunit played a key role in the community. Since this multicomponent enzyme system is involved in dissimilatory nitrate reduction, denitrification and comammox, it is a pointer to the fact that certain amount of the ammonia is generated by either ammonification or nitrogen fixation or ammonia fertilizer is converted to nitrate which is more readily available to plants but at the same time prone to leaching from the soil.

The annotation of the genes for the transport of nitrogen and biosynthesis of nitrogen-rich amino acids is indicative of certain role for nitrogen fixers even though they are not a major player in the 2S soil based on the most abundant taxonomic affiliation detected in the 2S metagenome. Free-living organisms present in bulk soil and spanning the phyla *Cyanobacteria*, *Proteobacteria*, *Archaea* and *Firmicutes* are well documented as nitrogen-fixers and are known to contribute to soil nutrient availability (Rashid et al. 2015).

The three predominant enzymes whose genes were annotated for sulfur metabolism in the 2S metagenome, namely sulfate adenyltransferase subunit 2, thiosulfate/ 3-mercaptopyruvate sulfurtransferase and cysteine synthase are core enzymes of the sulfur cycle coupled in organic system. Sulfur is an important component of sulfur containing amino acid cysteine and methionine and equally important in the synthesis of CoA enzyme as well as synthesis of vitamins (Kertesz et al. 2007). Although microbial sulfur oxidation is a dynamic process occupying an important place in the biogeochemical cycle of sulfur (Yousuf et al. 2014), majority of sulfur in soil is bound to organic molecules and is not readily available to plants unless saprobes present in the soil degrade such organic molecules (Kertesz and Mirleau 2004).

The revelation of the presence of type I and type II polyketide synthases and several genes involved in the synthesis of various antibiotics as well as resistance to antibiotics in the 2S metagenome is in consonance with prevailing knowledge on antibiotic functionalities in soil microbial community. Production of antibiotics, which often give competitive advantage to the producers and resistance to antibiotics are common attributes of soil microorganism, some of which are known to thrive solely on antibiotics (Dantas et al. 2008; Allen et al. 2010). It has been shown that consonant with the revelation

on preponderance in soil of uncultured majority of diverse species; natural product diversity is potentially much larger than appreciated from culture-based studies (Reddy et al. 2012).

Since polyketides synthases are not only involved in the synthesis of antibiotics but are also known to be involved in formation of the polyketide backbones of immunosuppressive agents and anticancer compounds (Vandova et al. 2017), it is likely that apart from the antibiotic biosynthetic enzymes that were annotated, the 2S metagenome also contains yet to be identified functionalities for other value-added natural biosynthetic products. One of the polyketide synthases genes identified, the type II PKS gene, is characterized by high amino acids sequence homology and conserved sequence expression and the enzymes it encodes are involved in the synthesis of structurally complex molecules with potent functionalities (Wawrik et al. 2005; Selvin et al. 2016).

Members of the phylum *Actinobacteria* were revealed as the predominant producers of secondary metabolites in the 2S soil community. This is not surprising as they are the major natural source of antimicrobial agents and the number one antibiotic producing genus, *Streptomyces*, is a member of this group (Chater 2006). Indeed, in recent times, using a shotgun metagenomic approach, many rare members of the phylum with novel biosynthetic functionalities for compounds with potential for use as drugs and other biotechnological processes have been revealed in unique soil environments (Azman et al. 2015).

The annotation of genes for resistance to antibiotics in the 2S metagenome is not surprising, as antibiotic resistance is a widespread phenomenon in the environment and recent finding using metagenomic approach have shown that it is increasingly becoming important, particularly in cultivated lands and environmental compartments influenced by human activities (Durso et al. 2012; Xiong et al. 2014). The *Proteobacteria*, which constitute the phylum with the highest taxonomic annotation for antimicrobial resistance in the 2S metagenome are known to exhibit a plethora of mechanisms for resistance to antibiotics produced in the environment. This may be connected with their pliability and propensity for horizontal gene transfer, which is crucial to the evolution of resistance to antimicrobials. Such resistance may account for their co-existence with *Actinobacteria* as the two dominant phyla in the soil microbiome.

5. Conclusions

In summary, shotgun metagenomic analysis of the metagenome of a tropical agricultural soil in Ilorin, Nigeria revealed that the phyla *Actinobacteria* and *Proteobacteria* were the predominant members of the community and were equally responsible for the diverse metabolic potentials and functionalities that constituted the major drivers of the ecosystem. Whereas the majority of the microbial community participated in heterotrophic metabolism of both an assimilatory and dissimilatory nature, the community also contained a diverse minority of autotrophic metabolic potentials, which were annotated for several rare but cultured species as well as some very common taxa. Generally, identification of functions for key metabolic pathways of elemental cycles indicated a system of intricately linked microbiomes. Furthermore, the community exhibited remarkable richness of both antimicrobial production potentials and resistance functions, which deserve further investigation in the quest for new antibiotics as well as the drive towards resolving the challenges posed by emerging resistance to antimicrobial agents.

REFERENCES

• Aislabie J, Deslippe JR. 2013. Soil microbes and their contribution to soil services. In: Dymond JR, editor. Ecosystem services in New Zealand-conditions and trends. Lincoln: Manaaki Whenua Press. p. 143-161.

• Allen HK, Donato J, Wang HH, Cloud-Hansen KA, Davis J, Handelsman J. 2010. Call of the wild: Antibiotic resistance genes in natural environments. Nat Rev Microbiol. 8(4):251-259.

• Ara I, Kudo T. 2007. *Luedemannella* gen. nov., a new member of the family *Micromonosporaceae* and description of *Luedemannella helvata* sp. nov. and *Luedemannella flava* sp. nov. J Gen Appl Microbiol. 53:39-51.

• Arias ME, González-Pérez JA, González-Vila FJ, Ball AS. 2005. Soil health-a new challenge for microbiologists and chemists. Int Microbiol. 8:13-21.

• Asadu CLA, Nwafor IA, Chibuike GU. 2015. Contributions of microorganisms to soil fertility in adjacent forest, fallow and cultivated land use Types in Nsukka, Nigeria. Int J Agric Forest 5(3):199-204.

• Azman A-S, Othman I, Velu SS, Chan K-G, Lee L-H. 2015. Mangrove rare *actinobacteria*: taxonomy, natural compound, and discovery of bioactivity. Front Microbiol. 6:856. doi:10.3389/fmicb.2015.00856.

• Bashir Y, Singh SP, Konwar BK. 2014. Metagenomics: an application-based perspective. Chinese J Biol. doi:10-1155/2014/146030 1-7.

• Bossio DA, Scow KM, Gunapala N, Graham KJ. 1998. Determinants of soil microbial communities: effects of agricultural management, season, and soil type on phospholipid fatty acid profiles. Microb Ecol. 36:1-12.

• Calderón FJ, Jackson LE, Scow KM, Rolston DE. 2001. Short-term dynamics of nitrogen, microbial activity, and phospholipid fatty acids after tillage. Soil Sci Soc Am J. 65:118-126.

• Castaneda LE, Barbosa O. 2017. Metagenomic analysis exploring taxonomic and functional diversity of soil microbial communities in Chilean vineyards and surrounding native forests. Peer J. 5:e3098. doi:10.7717/peerj.3098.

• Chater KF. 2006. *Streptomyces* inside-out: a new perspective on the bacteria that provide us with antibiotics. Phil Trans Royal Society B. 361:761-768.

• Cheema S, Lavania M, Lal B. 2015. Impact of petroleum hydrocarbon contamination on the indigenous microbial community. Ann Microbiol. 66:359-369.

• Cheung P-Y, Kinkle BK. 2001. *Mycobacterium* diversity and pyrene mineralisation in petroleum-contaminated soil. Appl Environ Microbiol. 67:2222-2229.

• Costa PS, Reis MP, Ávila MP, Leite LR, de Araújo FMG, Salim ACM, Oliveira G, Barbosa F, Chartone-Souza E, Nascimiento AMA. 2015. Metagenome of a microbial community inhabiting a metal-rich tropical stream sediment. Plos ONE 10(3):e0119465. doi:10.1371/journal.pone.0119465.

• Dantas G, Sommer MOA, Oluwasegun RD, Church GM. 2008. Bacteria subsisting on antibiotics. Science 320(5872):100-103.

• de Menezes AB, Lockhart RJ, Cox MJ, Allison HE, McCarthy AJ. 2008. Cellulose degradation by *Micromonospora* recovered from freshwater lakes and classification of these actinomycetes by DNA gyrase B gene sequencing. Appl Environ Microbiol. 74:7080-7084. doi:10.1128/AEM. 01092-08.

• de Menezes AB, McDonald JE, Allison HE, McCarthy AJ. 2012. Importance of *Micromonospora* spp. as colonizers of cellulose in freshwater lakes as demonstrated by quantitative reverse transcriptase PCR of 16S rRNA. Appl Environ Microbiol. 78:3495-3499.

• Durso LM, Miller DN, Wienhold B. 2012. Distribution and quantification of antibiotic resistant genes and bacteria across agricultural and non-agricultural metagenomes. PLoS ONE 7(11):e48325. doi:10.1371/journal.pone.0048325.

• Fierer N, Leff JW, Adams BJ, Nielsen UN, Bates ST, Lauber CL, Owens S, Gilbert JA, Wall DH, Caporaso G. 2012. Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. Proc Nat Acad Sci USA. 109:21390-21395.

• Garrity GM, Bell JA, Lilburn TG. 2005. Class III. *Gammaproteobacteria* class. nov. In: Brenner DJ, Krieg NR, Staley JT, Garrity GM, editors. Bergey's manual of systematic bacteriology, 2nd ed., Vol. 2. New York, NY: Springer. p. 1.

• Handelsman J. 2004. Metagenomics: application of genomics to uncultured microorganisms. Microbiol Mol Biol Rev. 68:669-668.

• Ilori MO, Amund OO, Obayori OS, Omotayo AE. 2015. Microbial population and physico-chemical dynamics of a soil ecosystem upon petroleum contamination. J Sci Res Dev. 15:25-33.

• Islam MR, Trivedi P, Palaniappan P, Reddy MS, Sa T. 2009. Evaluating the effect of fertilizer application on soil microbial community structure in rice-based cropping system using fatty acid methyl esters (FAME) analysis. World J Microbiol Biotechnol. 25:1115-1117.

• Jiang Y, Xiong X, Danska J, Parkinson J. 2016. Metatranscriptomic analysis of diverse microbial communities reveals core metabolic pathways and microbiome-specific functionality. Microbiome 4:2. doi:10.1186/s40168-015-0146x.

• Kanehisa M, Sato Y, Morishima K. 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. J Mol Biol. 428(4):726-731.

• Kertesz MA, Fellows E, Schmalenberger A. 2007. Rhizobacteria and plant sulfur supply. Adv Appl Microbiol. 62:235-268.

• Kertesz MA, Mirleau R. 2004. The role of soil microbes in plant sulfur nutrition. J Exp Botany 55:1939-1945.

• Köberl M, Müller H, Ramadan EM, Berg G. 2011. Desert farming benefits from microbial potential in arid soils and promotes diversity and plant health. PLoS ONE 6(9):e24452. doi:10.1371/journal.pone.0024452.

• Kunihiro T, Veuger B, Vasquez-Cardenas D, Pozzato L, Le Guitton M, Moriya K, Kuwae M, Omori K, Boschker HTS, van Oevelen D. 2014. Phospholipid-derived fatty acids and quinones as markers for bacterial biomass and community structure in marine sediments. PLoS ONE 9(4):e96219. doi:10.1371/journal.pone.0096219.

• Lauber CL, Hamady M, Knight R, Fierer N. 2009. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. Appl Environ Microbiol. 75(15):5111-5120.

• Li X, Rui J, Xiong J, Li J, He Z, Zhou J, Yannarell AC, Mackie RI. 2014. Functional potential of soil microbial communities in the maize rhizosphere. PLoS ONE 9(11):e112609.

• Lipson DA, Schmidt SK. 2004. Seasonal changes in alpine soil bacterial community in the Colorado Rocky Mountains. Appl Environ Microbiol. 70(5):2867-2879.

• Liu WT, Marsh TL, Cheng H, Forney LJ. 1997. Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Appl Environ Microbiol. 63:4516-4522.

• Lizarraga-Guerra R, López MG. 1996. Content of free amino acids in Huitlacoche (*Ustilago maydis*). J Agric Food Chem. 44(9):2556-2559.

• Ludwig W, Euzeby J, Schumann P, Busse H-J, Trujillo ME, Kämpfer P, Whitman WB. 2012. Roadmap of the phylum *Actinobacteria*. In: Goodfellow M, Kämpfer P, Busse H-J, Trujillo ME, Suzuki K-I, Ludwig W, Whitman WB, editors. Bergey's Manual of Systematic Bacteriology. 2nd Ed. New York, NY: Springer. p. 1-28.

• Maier RM, Pepper IL. 2009. Earth environment. In: Maier RM, Pepper IL, Gerba CP, editors. Environmental Microbiology. 2nd Ed. Burlington: Academic Press. p. 57-82.

• Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH. 2015. CDD: NCBI's conserved domain database. Nucleic Acids Res. 43(D):222-226.

• Matsumoto A, Kawaguchi Y, Nakashima T, Iwatsuki M, Omura S. Takashi Y. 2014. *Rhizocola helleborigen*. nov., sp. nov., an actinomycete of the family *Micromonosporaceae* containing 3,4-dihydroxydiaminopimelic acid in the cell wall peptidoglycan. Int J Syst Microbiol. 64:2706-2711.

• Mellouli L, Mehdi RB, Sioud S, Salem M, Bejar S. 2003. Isolation, purification and partial characterization of antibacterial activities produced by a newly isolated *Streptomyces* sp. US24 strain. Res Microbiol. 154:345-352.

• Moats GA, Foster JW, Spector MP. 2002. Microbial Physiology, 4th Edition. New York: Wiley-Liss Inc. 714 p.

• Mocali S, Benedetti A. 2010. Exploring research frontiers in microbiology: the challenges of metagenomics in soil microbiology. Res Microbiol. 161(6):497-505. doi:10.1016/j.resmic.2010.04.010.

• Monciardini P, Cavaletti, L, Ranghetti A, Schumann P, Rohde M, Bamonte R, Sosio M, Mezzelani A, Donadio S. 2009. Novel members of the family *Micromonosporaceae*, *Rugosimonospora acidiphila* gen. nov., sp. nov. and *Rugosimonospora africana* sp. nov. Int J Syst Evol Microbiol. 59:2752-2758.

• Monciardini P, Cavaletti L, Schumann P, Rohde M, Donadio S. 2003. *Conexibacter woesei* gen. nov., sp. nov., a novel representative of a deep evolutionary line of descent within the class *Actinobacteria*. Int J Syst Evol Microbiol. 53(Pt 2):569-576.

• Montecchia MS, Tosi M, Soria MA, Vogrig JA, Sydorenko O, Correa OS. 2015. Pyrosequencing reveals changes in soil bacterial communities after conversion of Yungas forests to agriculture. PLoS One 10:e0119426. doi:10.1371/journal.pone.0119426.

• Muyzer G, de Waal EC, Uitterlinden AG. 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl Environ Microbiol. 59:695-700.

• Myrold DD, Zeglin LH, Jansson JK. 2013. The potential of metagenomics approaches for understanding soil microbial processes. Soil Sci Soc Am J. 78:3-10.

• Nannipieri, P, Ascher J, Ceccherini MT, Landi L, Pietramellara G, Renella G. 2003. Microbial diversity and soil functions. Eur J Soil Sci. 54:655-670.

• Newby DT, Marlowe EM, Maier RM. 2009. Nucleic acid-based methods of analysis. In: Maier RM, Pepper IL, Gerba CP, editors. Environmental Microbiology. 2nd Ed. Burlington: Academic Press. p. 243-28.

• Ogbulie TE, Nwaokorie FO. 2016. Molecular diversity of microbes with probable degradative genes in agriculture soil contaminated with Bonny Light crude oil. J Ecosyst Ecograph S5:002. doi:10.4172/2157-7625.S5-002.

• Oulas A, Pavloudi G, Polymanakou P, Pavlopoulus GA, Papanikolaou N, Kotoulas G, Arvanitidis C, Iliopoulus I. 2015. Metagenomics: Tools and insights for analyzing next-generation sequencing data derived from biodiversity studies. Bioinform Biol Insights 9:75-88.

• Pacchioni RG, Carvalho FM, Thompson CE, Faustino AL, Nicolini F, Pereira TS, Silva RC, Cantão ME, Gerber A, Vasconcelos AT, Agnez-Lima LF. 2014. Taxonomic and functional profiles of soil samples from Atlantic forest and Caatinga biomes in northeastern Brazil. MicrobiologyOpen 3(3):299-315.

• Pan JJ, Baumgarten AM, May G. 2008. Effects of host plant environment and *Ustilago maydis* infection on the fungal endophyte community of maize (*Zea mays*). New Phytologist 178:147-156.

• Paul S, Cortez Y, Vera N, Villena GK, Gutiérrez-Correa M. 2016. Metagenomic analysis of microbial community of an Amazonian geothermal spring in Peru. Genomic Data 9:63-66.

• Qin S, Li W-J, Dastager SG, Hozzein WN, editors. 2016. *Actinobacteria* in special and extreme habitats: diversity, function, roles and environmental adaptations. Lausanne: Frontiers Media. doi:10.3389/978-2-88945-013-8.

• Rashid MI, Mujawar LH, Shahzad T, Almeelbi T, Ismail IMI, Oves M. 2016. Bacteria and fungi can contribute to nutrients bioavailability and aggregate formation in degraded soils. Microbiol Res. 183:26-41.

• Rashid MH, Young JP, Everall I, Clercx P, Willems A, Santhosh Braun M, Wink M. 2015. Average nucleotide identity of genome sequences supports the description of *Rhizobium lentis* sp. nov., *Rhizobium bangladeshense* sp. nov., and *Rhizobium binae* sp. nov., from lentil (Len culinaris) nodules. Int J Syst Evol Microbiol. 65(9):3037-3045.

• Reddy BV, Kallifidas D, Kim JH, Charlop-Powers Z, Feng Z, Brady SF. 2012. Natural product biosynthetic gene diversity in geographically distinct soil microbiomes. Appl Environ Microbiol. 78(10):3744-3752.

• Rho M, Tang H, Ye Y. 2010. FragGeneScan: predicting genes in short and error-prone reads. Nucleic Acid Res. 38:20-191.

• Salam LB, Ilori MO, Amund OO, Numata M, Horisaki T, Nojiri H. 2014. Carbazole angular dioxygenation and mineralization by bacteria isolated from hydrocarbon-contaminated tropical African soil. Environ Sci Pollut Res. 21(15):9311-9324. doi:10.1007/s11356-014-2855-2.

• Salam LB, Obayori OS, Nwakorie FO, Suleiman A, Mustapha R. 2017. Metagenomic insight into effects of spent oil perturbation on the microbial community composition and function in a tropical agricultural soil. Environ Sci Pollut Res. 24:7139-7159. doi:10.1007/s11356-017-8364-3.

• Sathya A, Vijayabharathi R, Gopalakrishnan S. 2017. Plant growth-promoting *actinobacteria*: a new strategy for enhancing sustainable production and protection of grain legumes. 3 Biotech 7(2):102. doi:10.1007/s13205-017-0736-3.

• Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol. 75(23):7537-7541.

• Segata N, Boernigen D, Tickle TL, Morgan XC, Garrett WS, Huttenhower C. 2013. Computational meta'omics for microbial community studies. Mol Syst Biol. 9:666. doi:10.1038/msb.2013.22.

• Seipke RF, Kaltenpoth M, Huchings MI. 2011. *Streptomyces* as symbionts: an emerging and widespread theme? FEMS Microbiol Rev. 36 (2012):862-876.

• Seki T, Matsumoto A, Omura S, Takahashi Y. 2012. *Conexibacter arvalis* sp. nov. isolated from cultivated a field sample. Int J Syst Evol Microbiol. 62(10):2400-2404.

• Selvin J, Sathiyanarayanan G, Lipton AN, Al-Dhabi NA, Valan Arasu M, Kiran GS. 2016. Ketide Synthase (KS) domain prediction and analysis of iterative Type II PKS Gene in Marine Sponge-Associated *Actinobacteria* producing biosurfactants and antimicrobial agents. Front Microbiol. 7:63. doi:10.3389/fmicb.2016.00063.

• Sharma SB, Sayyed RZ, Trivedi MH, Gobi TA. 2013. Phosphate solubilizing microbes: sustainable approach for managing phosphorus deficiency in agricultural soils. SpringerPlus 2:587. doi:10.1186/2193-1801-2-587.

• Shivlata L, Satyanarayana T. 2015. Thermophilic and alkaliphilic *Actinobacteria*: biology and potential applications. Front Microbiol. 6:1014. doi:10.3389/fmicb.2015.01014.

• Silva MS, Sales AN, Magalhaes-Guedes KT, Dias DR, Scwan RF. 2013. Brazillian cerrado soil *actinobacteria*l ecology. BioMed Research International, Vol 2013, Article ID 503805.10 p. doi:10.1155/2013/503805.

• Soil Survey Division Staff. 1993. Soil Survey Manual. USDA Handbook No. 18. Washington, D.C., USA: US Government Printing Office. p. 63-65.

• Spain AM, Krumholz LR, Elshahed MS. 2009. Abundance, composition, diversity and novelty of soil Proteobacteria. ISME Journal 9:992-1000.

• Stach JEM, Burns RG. 2002. Enrichment versus biofilm culture: a functional and phylogenetic comparison of polycyclic aromatic hydrocarbon-degrading microbial communities. Environ Microbiol. 4(3):159-182.

• Streit WR, Schmitz RA. 2004. Metagenomics- key to the uncultured microbes. Curr Opin Microbiol. 7:492-498.

• Suzuki M, Rappe MS, Giovanoni SJ. 1998. Kinetic bias in estimates of coastal picoplankton community structure obtained by measurement of small subunit rRNA gene PCR amplicon length heterogeneity. Appl Environ Microbiol. 64:4522-4529.

• Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS, Kiryutin B, Galperin MY, Fedorova ND, Koonin EV. 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res. 29(1):22-28.

• Thakur D, Yadav A, Gogoi BK, Bora TC. 2007. Isolation and screening of *Streptomyces* in soil of protected forest areas from the states of Assam and Tripura, India, for antimicrobial metabolites. Journal de Mycologie Médicale 17:242-249.

• Trivedi P, Delgado-Baquerizo M, Anderson IC, Singh BK. 2016. Response of soil properties and microbial communities to agriculture: implications for primary productivity and health indicators. Front Plant Sci. 7:990. doi:10.3389/fpls.2016.00990.

• Trujillo ME, Riesco R, Benito P, Carro L. 2015. Endophytic *Actinobacteria* and the interaction of *Micromonospora* and nitrogen-fixing plants. Front Microbiol. 6:1341. doi:10:3389/fmicb.2015.01341.

• Vandova GA, O'Brien RV, Lowry B, Robbins TF, Fischer CR, Davis RW, Khosla C, Harvey CJB, Hillenmeyer ME. 2017. Heterologous expression of diverse propionyl-CoA carboxylases affects polyketide production in *Escherichia coli*. The Journal of Antibiotics 70:859-863.

• Wang X, Jia F, Liu C, Zhao J, Wang L, Shen Y, Wang J, Zhang J, Li C, Xiang W. 2013. *Xiangella phaseoli* gen. nov., sp. nov., a member of the family *Micromonosporaceae*. Int J Syst Evol Microbiol. 63:2138-2145.

• Wardle DA, Bardgett RD, Klironomos JN, Seta¨la¨ H, van der Putten WH, Wall DH. 2004. Ecological linkages between aboveground and belowground biota. Science 304(5677):1629-1633.

• Wawrik B, Kerkhof L, Zylstra GJ, Kukor JJ. 2005. Identification of unique type II polyketide synthase genes in soil. Appl Environ Microbiol. 71(5):2232-2238.

• Wibberg D, Al-Dilaimi A, Busche T, Wedderhoff I, Schrempf H, Kakinowski J, Ortiz de Orué Lucana D. 2016. Complete genome sequence of *Streptomyces reticuli*, an efficient degrader of crystalline cellulose. J Biotechnol. 222:13-14.

• Williams KP, Gillespie JJ, Sobral BW, Nordberg EK, Snyder EE, Shallom JM, Dickerman AW. 2010. Phylogeny of gammaproteobacteria. J Bacteriol. 192(9):2305-2314.

• Williams KP, Sobral BW, Dickerman AW. 2007. A robust species tree for the *Alphaproteobacteria*. J Bacteriol. 189:4578-4586.

• Xiong W, Sun Y, Ding X, Zhang Y, Zeng Z. 2014. Antibiotic resistance genes occurrence and bacterial community composition in the Liuxi River. Front Environ Sci. 2:61. doi:10.3389/fenvs.2014.00061.

• Yin C, Mueth N, Hulbert S. 2017. Bacterial community on wheat grown under long-term conventional tillage and no-till in the Pacific Northwest of the United States. Phytobiomes J. 1(2):83-90.

• Yousuf B, Kumar R, Mishra A, Jha B. 2014. Unravelling the carbon and sulphur metabolism in coastal soil ecosystems using comparative cultivation-independent genome-level characterisation of microbial communities. PLoS ONE 9(9):e107025. doi:10.1371/journal.pone.0107025.

• Zubay GL, Parson WW, Vance DE. 1995. Principles of Biochemistry. Dubuque: Wm. C. Brown Publishers. 863 p.