Salam *Bulletin of the National Research Centre* (2018) 42:10
https://doi.org/10.1186/s42269-018-0013-6

Bulletin of the National Research Centre

RESEARCH Open Access

# Detection of carbohydrate-active enzymes and genes in a spent engine oil-perturbed agricultural soil

Lateef Babatunde Salam

## Abstract

**Background:** The purpose of this study is to decipher the diverse carbohydrate metabolism pathways in a spent engine oil-perturbed agricultural soil, enunciate the carbohydrate-active enzymes and genes involved in the process, taxonomically classify the annotated enzymes and genes, and highlight the importance of the study for ecological and biotechnological processes.

**Results:** Functional analysis of the metagenome of spent engine oil (SEO)-contaminated agricultural soil (AB1) using the Kyoto Encyclopedia of Genes and Genomes (KEGG) GhostKOALA, Cluster of Orthologous Groups (COG) of proteins, the Carbohydrate-Active Enzymes (CAZy) database, and the NCBI's conserved domain database (CDD) revealed extensive metabolism of carbohydrates via diverse carbohydrate-active enzymes and genes. Enzymes and genes annotated for glycolysis/gluconeogenesis pathway, citric acid (TCA) cycle, pentose phosphate pathway, and pyruvate metabolism, among others, were detected, and these were not detected in the original agricultural soil (1S). Analysis of carbohydrate-active enzymes, using the CAZy database, showed 45 CAZy families with preponderance of glycoside hydrolases (GHs, 46.7%), glycosyltransferases (GTs, 24.4%), and carbohydrate-binding modules (CBMs, 15.5%). Taxonomic classification of the annotated enzymes and genes for carbohydrate metabolism using the GhostKOALA and CAZy databases revealed the predominance of the phylum *Proteobacteria* with the representative genera *Pseudomonas* (18%), *Sphingobium* (13.5%), and *Sphingomonas* (4.5%), respectively. Biotechnologically important enzymes such as xylanases, endoglucanases, α- and β-glucosidases and glycogen debranching enzymes were also retrieved from the metagenome.

**Conclusions:** This study revealed the presence of diverse carbohydrate-active enzymes and genes mediating various carbohydrate metabolism pathways in the SEO-perturbed soil metagenome. It also reveals the detection of biotechnologically important enzymes with potentials for industrial use.

**Keywords:** Spent engine oil, Agricultural soil, Soil microcosm, Illumina sequencing, Carbohydrate metabolism, Carbohydrate-active enzymes, Microbial enzymes and genes

## Background

Hydrolysis of carbohydrates by microorganisms is a combination of diverse biochemical processes responsible for their formation, degradation, and transformation. The metabolic pathways used and the metabolites produced are determined exclusively by the enzyme machinery of the microbial community in such environment (Ley et al. 2008; Xia et al. 2015).

Complex carbohydrates found in nature are catalyzed by a range of enzymes involved in their assembly (glycosyltransferases, GTs) and their breakdown (glycoside hydrolases, GHs; polysaccharide lyases, PLs; carbohydrate esterases, CEs; auxiliary activities, AAs), collectively called carbohydrate-active enzymes (CAZymes). In addition, the carbohydrate-binding modules (CBMs) assist in hydrolysis of polysaccharides by bringing the biocatalyst into close contact with its recalcitrant substrate (Lombard et al. 2014).

Glycoside hydrolases (GHs) are enzymes that cleave glycosidic bonds in glycosides, glucan, and glycoconjugates. Their industrial and biotechnological importance

Correspondence: babssalaam@yahoo.com
Microbiology Unit, Department of Biological Sciences, Al-Hikmah University, Ilorin, Kwara, Nigeria

is not in doubt as they play key roles in the development of biofuels (cellulases, xylanases, etc.) and disease research (Yuzwa et al. 2008; Abbott et al. 2009; Wilson 2009). They also play crucial role in global carbon cycling by allowing microorganisms in the soil to breakdown plant cells, releasing $CO_2$ aerobically and various fermentation products anaerobically (Bardgett et al. 2008). Glycosyltransferases (GTs) represent a subclass of enzymes that catalyze the synthesis of glycosidic linkages by the transfer of a sugar residue from a donor substrate to an acceptor. They play key roles in biosynthesis pathways of oligo- and polysaccharides, as well as protein glycosylation and formation of valuable natural products (Lairson et al. 2008; Schmid et al. 2016). The carbohydrate esterases (CEs) involve enzymes that catalyze the de-O or de-N-acylation to remove the ester embellishments from carbohydrates (Cantarel et al. 2009). The range of biological and biotechnological applications of CEs is diverse. For example, majority of CE families include members that catalyze the removal of acylated moieties of polysaccharides, which enhance their degradation and facilitate access of glycoside hydrolases, thus assisting in biomass saccharification and generation of renewable biofuels, sustainable materials, and green chemicals (Christov and Prior 1993; Gupta and Verma 2015). Furthermore, several CE families have been reported to contain enzyme targets for drug design and considerable potentials in biomedical applications (Nakamura et al. 2017).

In recent years, our understanding of microbial metabolism has been enhanced considerably by the application of next-generation sequencing (NGS) techniques. NGS-based metagenomics allows the study of microbial communities without prior culturing or marker gene amplification. This provides a relatively unbiased view of not only the microbial community structure but also the metabolic pathways of the community.

Hydrocarbon pollution generally imposed significant stress on microbial community in soil. As hydrocarbons are weak in nutrients consisting mostly of carbon and hydrogen, there is significant pressure on the microbial community to utilize stored and easily metabolized carbon sources as sources of energy to drive metabolism and biosynthetic processes (Demoling et al. 2007).

Previously, we have investigated the effects of SEO perturbation on the microbial community structure and functions of an agricultural soil using Illumina NGS sequencing technique. The study particularly showed the dominance of hydrocarbonoclastic organisms due to the SEO contamination and preponderance of functional genes related to hydrocarbon degradation, heavy metal tolerance and detoxification, oxidative stress, response to nutrient starvation, and many others (Salam et al. 2017). However, a curious observation is the detection of huge array of carbohydrate metabolism pathways mediated by carbohydrate-active enzymes and genes in the SEO-perturbed agricultural soil not detected in the original agricultural soil. These observations spur investigations into these unusual findings and the possible roles played by SEO presence and degradation.

## Methods

### Site description, microcosm set up, and determination of residual hydrocarbons

Soil samples were collected from an agricultural farm at Atere, Ilorin, Nigeria. The coordinates of the sampling site were latitude 8° 29′ 11.54″ N and longitude 4° 29′ 30.11″ E. The soil, which is dark-brown in color, is used to plant majorly maize and vegetables. Soil microcosm (agricultural soil, 1S; agricultural soil + SEO, AB1) set up, and incubation conditions have been described previously (Salam et al. 2017). Similarly, the physicochemical properties, the residual spent engine oil (SEO), and its various constituents as well as the heavy metal content of the soil microcosms have been reported previously (Salam et al. 2017).

### DNA extraction, library construction, sequencing, and metagenome properties

Genomic DNA used for metagenomic analysis was extracted directly from soil microcosms. 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 1S and AB1 microcosms 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 assay kit. 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 $2 \times 300$ 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 IS and AB1 metagenomes were

deposited on the MG-RAST server with the IDs 4,704,688.3 and 4,704,689.3, respectively.

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. The sequences were assembled into 2389 and 2948 unique contigs for soil microcosms 1S and AB1 with a total of 611,018 and 761,383 bp, an average length of 255 ± 63 and 258 ± 62 bp, and the GC content of 56 ± 4 and 62 ± 5%, respectively. After dereplication and quality control by the MG-RAST, the total number of sequence reads in AB1 metagenome reduced to 2828 sequence reads with 711,035 bp, an average length of 251 ± 54 bp, and a GC content of 62 ± 5% (Salam et al. 2017).

### Functional analyses of metagenomic read for carbohydrate metabolism

Gene calling was performed on the AB1 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), the Carbohydrate-Active Enzymes (CAZy) database, 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 CAZymes Analysis Toolkit (CAT) (Park et al. 2010) was also used to detect the carbohydrate-active enzymes in AB1 metagenome using a sequence similarity-based annotation. This is based on the similarity search of the protein sequences in AB1 sequence reads against the entire non-redundant sequences of the carbohydrate-active enzymes (CAZy) database (Park et al. 2010; Lombard et al. 2014). Sequence reads annotated for carbohydrate metabolism in AB1 metagenome was further elucidated 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.

## Results

Functional analyses of AB1 metagenome using GhostKOALA, COG, CAZy, and NCBI CDD databases revealed the presence of various carbohydrate metabolism pathways such as glycolysis/gluconeogenesis, citric acid (TCA) cycle, pentose phosphate pathway, and several others as well as carbohydrate-active enzymes (Table 1 and Table 2). It is instructive to note that only four genes encoding four different enzymes (β-glucuronidase, ribulose-5-phosphate 4-epimerase, 4-alpha-glucanotransferase, pyruvate dehydrogenase (quinone)) for carbohydrate metabolism were detected in 1S microcosm (Additional file 1).

### Glycolysis/gluconeogenesis

Functional characterization of AB1 metagenome using GhostKOALA and COG revealed 13 enzymes annotated for glycolysis/gluconeogenesis (Fig. 1). Interestingly, only three of the enzymes are glycolytic enzymes, though aldehyde dehydrogenase (EC: 1.2.1.3) a superfamily having glyceraldehyde-3-phosphate dehydrogenase as member was detected. These are glucokinase (EC: 2.7.1.2), glucose-6-phosphate isomerase (EC: 5.3.1.9), and phosphoglycerate kinase (EC: 2.7.2.3). Similarly, three gluconeogenic enzymes were also detected in AB1 metagenome. These are pyruvate carboxylase (EC: 6.4.1.1), fructose 1,6-bisphosphatase I (EC: 3.1.3.11), and glucose-6-phosphate isomerase, respectively.

The remaining six enzymes annotated for glycolysis/gluconeogenesis falls into two categories. Enzymes that catalyze the feeder pathways for glycolysis include phosphoglucomutase (EC: 5.4.2.2), and phosphomannomutase (EC: 5.4.2.8). The second category is enzymes that catalyze pyruvate, the metabolic product of glycolysis. These are pyruvate dehydrogenase E1 (EC: 1.2.4.1) and E2 (EC: 2.3.1.12) components, alcohol dehydrogenase (EC: 1.1.1.1), and acetyl-CoA synthetase (EC: 6.2.1.1), which catalyze the conversion of acetate to acetyl-CoA, respectively (Table 1, Fig. 1).

Taxonomic affiliation of the functional genes/enzymes annotated for glycolysis/gluconeogenesis revealed that the organisms belong to the phylum *Proteobacteria*. The class *γ-Proteobacteria* predominates, constituting 40% of the taxonomic affiliations with members such as *Pseudomonas*, *Stenotrophomonas*, and *Thioploca*. This is closely followed by *α-Proteobacteria*, contributing 33.3% of the organisms annotated with representatives such as *Caulobacter*, *Acidiphilium*, and *Methylobacterium*. *β-Proteobacteria* contributes 26.7% of the organisms annotated with members such as *Achromobacter* and *Massilia* (Table 1).

### Citric acid (TCA) cycle

Functional characterization of AB1 metagenome revealed eight (8) enzymes annotated for citric acid (TCA cycle) (Fig. 2). Aside from detection of pyruvate dehydrogenase E1 (EC: 1.2.4.1) and E2 (EC: 2.3.1.12) dihydrolipollysine-residue acetyltransferase components, which catalyze the starting point of the TCA cycle, seven (7) other major enzymes of the citric acid cycle were detected in AB1 metagenome. These are aconitate

**Table 1** List of the enzymes/genes and microorganisms detected in AB1 microcosm involved in diverse metabolism of carbohydrates

| Pathway | Enzymes/genes | Microorganisms |
|---|---|---|
| Glycolysis/ gluconeogenesis | Glucokinase; glucose-6-phosphate isomerase; fructose 1,6-bispho sphatase; phosphoglycerate kinase; pyruvate dehydrogenase E1 component; pyruvate dehydrogenase E1 component α-subunit; pyruvate dehydrogenase E2 component (dihydrolipoamide ace tyltransferase); alcohol dehydrogenase; aldehyde dehydrogenase ($NAD^+$); acetyl-CoA synthetase; pyruvate carboxylase phospho mannomutase/phosphoglucomutase; glucose-6-phosphate 1-epimerase | *Thioploca*, *Achromobacter*, *Colwellia*, *Massilia*, *Thauera*, *Acidiphilium*, *Methylobacterium*, *Stenotrophomonas*, *Caulobacter*, *Comamonas*, *Paracoccus*, *Azotobacter*, *Pseudomonas* |
| Citric acid (TCA) cycle | Aconitate hydratase 2/2-methylisocitrate dehydratase; 2-oxoglutarate dehydrogenase E2 component (dihydrolipoamide succinyltransferase); succinyl-CoA:acetate-CoA transferase; succinate dehydrogenase/fumarate reductase, flavoprotein subunit; fumarate reductase, flavoprotein subunit; fumarate hydratase, class I; malate dehydrogenase; pyruvate dehydrogen ase E1 component; pyruvate dehydrogenase E1 component α-subunit; pyruvate dehydrogenase E2 component (dihydrolipoa mide acetyltransferase) | *Novosphingobium*, *Sphingobium*, *Azorhizobium*, *Candidatus Symbiobacter*, *Steroidobacter*, *Xanthomonas*, *Acidiphilium*, *Thauera*, *Methylobacterium*, *Stenotrophomonas* |
| Pentose phosphate pathway | Glucose-6-phosphate isomerase; glucose-6-phosphate 1-dehydrogenase; 6-phosphogluconolactonase; 6-phosphogluconate dehydrogenase; ribulose-phosphate 3-epimerase; xylulose-5-phosphate/fructose-6-phosphate phospho ketolase; ribokinase; phosphomannomutase/phosphoglucomu tase; fructose 1,6-bisphosphatase I | *Achromobacter*, *Brevundimonas*, *Xanthomonas*, *Nitrospira*, *Pseudanabaena*, *Methylophaga*, *Sinorhizobium*, *Ensifer*, *Pseudomonas*, *Caulobacter*, *Colwellia*, *Symbiobacterium* |
| Pentose and glucuronate interconversions | 2-Deoxy-D-gluconate-3-dehydrogenase; tagaturonate reductase; ribulose-phosphate-3-epimerase; L-arabinose isomerase; L-rham nose isomerase/sugar isomerase | *Chelativorans*, *Klebsiella*, *Pseudanabaena*, *Granulicella*, *Novosphingobium* |
| Fructose and mannose metabolism | Mannose-6-phosphate isomerase; phosphomannomutase/ phosphoglucomutase, L-rhamnose isomerase; L-rhamnose isomerase/sugar isomerase; fructose 1,6-bisphosphatase I; mannitol-1-phosphate/altronate dehydrogenases | *Raoultella*, *Pseudomonas*, *Klebsiella*, *Novosphingobium*, *Caulobacter*, *Colwellia* |
| Galactose metabolism | UDP-glucose-4-epimerase; phosphomannomutase/ phosphoglucomutase; glucokinase; D-tagatose 1,6-bisphosphate aldolase; α-glucosidase | *Niastella*, *Pseudomonas*, *Thioploca*, *Stenotrophomonas*, *Citrobacter* |
| Amino sugar and nucleotide sugar metabolism | β-N-Acetylhexosaminidase; anomeric MURNAc/GlcNAc kinase; UDP-N-acetylglucosamine-2-epimerase; N-acetylglucosamine-6-phosphate deacetylase; glucosamine-6-phosphate deaminase; α-N-arabinofuranosidase; UDP-4-amino-4-deoxy-L-arabinose formyl transferase/UDP-glucuronic acid dehydrogenase; UDP-glucuronate decarboxylase; glucokinase; glucose-6-phosphate isomerase; phosphomannomutase/phosphoglucomutase; UDP-glucose-4-epimerase; mannose-6-phosphate isomerase | *Azorhizobium*, *Magnetospira*, *Comamonadaceae* bacterium A1, *Terriglobus*, *Microterricola*, *Symbiobacterium*, *Sphingomonas*, *Kosakonia*, *Pseudomonas*, *Sphingobium*, *Thioploca*, *Achromobacter*, *Niastella*, *Raoultella* |
| Starch and sucrose metabolism | α-Glucosidase; β-glucosidase; trehalose-6-phosphate synthase; α, α-trehalase; maltose-α-D-glucosyltransferase/α-amylase; glycogen operon protein; α-D-glucan-1-α-D-glucosylmutase; UDP-glucuronate decarboxylase; isoamylase; glucokinase; glucose-6-phosphate isomerase; 4-α-glucanotransferase; cellulose synthase (UDP-forming); endoglucanase; endo-1,4-β-xylanase; phospho mannomutase/phosphoglucomutase, type II secretory pathway pullulanase PulA and related glycosidases (glycogen debranching enzyme) | *Citrobacter*, *Sphingobium*, *Agrobacterium*, *Azotobacter*, *Pseudomonas*, *Acidobacterium*, *Desulfurivibrio*, *Methylocella*, *Paraburkholderia*, *Sphingomonas*, *Thioploca*, *Achromobacter*, *Rhodospirillum*, *Terriglobus* |
| Pyruvate metabolism | Acetyl-CoA synthetase; pyruvate dehydrogenase E1 component; pyruvate dehydrogenase E1 component α-subunit; pyruvate de hydrogenase E2 component (dihydrolipoamide acetyltransferase); acetyl-CoA carboxylase carboxyl transferase α-subunit; acetyl-CoA carboxylase, biotin carboxylase subunit; aldehyde dehydrogenase ($NAD^+$); succinyl-CoA:acetate-CoA transferase; L-lactate dehydro genase (cytochrome); pyruvate dehydrogenase (quinone); glyoxy late/hydroxypyruvate reductase A; malate dehydrogenase; fumarate hydratase class I; fumarate reductase flavoprotein subunit; isopropylmalate synthase | *Azotobacter*, *Acidiphilium*, *Thauera*, *Methylobacterium*, *Stenotrophomonas*, *Dyella*, *Novosphingobium*, *Comamonas*, *Paracoccus*, *Azorhizobium*, *Sphingobium*, *Paraburkholderia*, *Xanthomonas*, *Steroidobacter*, *Frateuria* |
| Glyoxylate and dicarboxylate | Isocitrate lyase; malate dehydrogenase; aconitate hydratase 2/2-methylisocitrate dehydratase; acetoacetyl-CoA reductase; propionyl- | *Sphingobium*, *Sphingomonas*, *Beijerinckia*, *Pseudomonas*, *Altererythrobacter*, *Xanthomonas*, *Novosphingobium*, |

**Table 1** List of the enzymes/genes and microorganisms detected in AB1 microcosm involved in diverse metabolism of carbohydrates *(Continued)*

| Pathway | Enzymes/genes | Microorganisms |
|---|---|---|
| metabolism | CoA carboxylase β-chain; catalase; phosphoglycolate phosphatase; glycine dehydrogenase subunit 1 and 2; glyoxylate/ hydroxypyruvate reductase A; formate dehydrogenase major sub unit; formyltetrahydrofolate deformylase | *Caedibacter*, *Lysobacter*, *Caulobacter*, *Paraburkholderia*, *Rhizobium*, *Burkholderia* |
| Propanoate metabolism | Acetyl-CoA synthetase; acyl-CoA dehydrogenase; acrylyl-CoA reductase (NADPH); 3-hydroxyacyl-CoA dehydrogenase/enoyl-CoA hydratase/3-hydroxybutyryl-CoA epimerase/enoyl-CoA isomerase; acetyl-CoA carboxylase carboxyl transferase subunit alpha; acetyl-CoA carboxylase, biotin carboxylase subunit; propionyl-CoA carb oxylase β-chain; aconitate hydratase 2/2-methylisocitrate dehydratase | *Azotobacter*, *Alkanivorax*, *Desulfurispirillum*, *Pseudomonas*, *Dyella*, *Novosphingobium*, *Sphingobium* |
| Butanoate metabolism | 3-Hydroxyacyl-CoA dehydrogenase; 3-hydroxyacyl-CoA dehydro genase/enoyl-CoA hydratase/3-hydroxybutyryl-CoA epimerase/ enoyl-CoA isomerase; enoyl-CoA hydratase; butyryl-CoA dehydro genase; poly (3-hydroxybutyrate) depolymerase; acetoacetyl-CoA reductase; fumarate reductase flavoprotein subunit; succinate semialdehyde dehydrogenase/glutarate semialdehyde dehydro genase; succinyl-CoA:acetate CoA-transferase; acetolactate syn thase I/II/II large subunit | *Novosphingobium*, *Cupriavidus*, *Xanthomonas*, *Pseudomonas*, *Chloracidobacterium*, *Burkholderia*, *Sphingomonas*, *Caedibacter*, *Steroidobacter*, *Chromohalobacter*, *Azorhizobium*, *Klebsiella* |
| Sugar transport systems and signal transduction | ABC-type sugar transport system, ATPase component; ABC-type sugar transport systems, permease components; ABC-type sugar transport system, periplasmic component; ribose transport system, substrate binding protein; L-fucose:$H^+$ symporter permease; phosphotransferase system, HPr-related proteins; $Na^+$/melibiose symporter and related transporters; sugar phosphate permease; TRAP-type C4-dicarboxylate transport system, small permease component; aerobic C4-dicarboxylate transport protein; maltoporin (maltose/maltodextrin high affinity phage lambda receptor protein); C4-dicarboxylate ABC transporter; histidine kinase regulating citrate/malate metabolism; histidine kinase regulating C4-dicarboxylate transport system | *Bradyrhizobium*, *Roseobacter*, *Hamadaea*, *Dyella*, *Kluyvera*, *Desulfosporosinus*, *Luteibacter*, *Sphingobium*, *Enterobacter*, *Burkholdera*, *Paraburkholderia*, *Pseudoxanthomonas*, *Sphingomonas*, *Klebsiella*, *Novosphingobium*, *Pseudomonas*, *Bosea* |

hydratase (EC: 4.2.1.3), 2-oxoglutarate (α-ketoglutarate) dehydrogenase (EC: 2.3.1.61), and succinyl-CoA:acetate-CoA transferase (EC:2.8.3.18). Others include succinate dehydrogenase (EC: 1.3.5.1), fumarate reductase (EC: 1.3.5.4), fumarate hydratase (EC: 4.2.1.2), and malate dehydrogenase (EC: 1.1.1.37), respectively.

Taxonomic affiliation of the functional genes/enzymes annotated for TCA cycle revealed that the microorganisms belong to the phylum *Proteobacteria*. *α-Proteobacteria* predominates with 50% of the taxonomic affiliations. The representative genera include *Novosphingobium*, *Sphingobium*, and *Azorhizobium*. This is followed by *γ-Proteobacteria* (35.7%) with representative genera such as *Steroidobacter*, *Xanthomonas*, and *Stenotrophomonas*. *β-Proteobacteria* contributes 14.3% of the organisms annotated with members such as *Thauera* and *Candidatus Symbiobacter*, respectively (Table 1).

### Pentose phosphate pathway (PPP)

Functional analysis of AB1 metagenome revealed ten (10) enzymes annotated for the pentose phosphate pathway (Fig. 3). Out of these, four enzymes glucose 6-phosphate isomerase, phosphomannomutase, phosphoglucomutase, and fructose 1,6-bisphosphatase are glycolytic/gluconeogenic enzymes, with possible roles in the generation of glucose-6-phosphate, the starting substrate in pentose phosphate pathway. The remaining six (6) enzymes are PPP enzymes. These include glucose 6-phosphate dehydrogenase (EC: 1.1.1.49), 6-phosphogluconolactonase (EC: 3.1.1.31), and 6-phosphogluconate dehydrogenase (EC: 1.1.1.44). Others include ribulose phosphate 3-epimerase (EC: 5.1.3.1), xylulose 5-phosphate phosphoketolase (EC: 4.1.2.9), fructose 6-phosphate phosphoketolase (EC: 4.1.2.22), and ribokinase (EC: 2.7.1.15), which phosphorylate ribose to ribose-5-phosphate.

Taxonomic affiliation of the annotated PPP enzymes revealed that the phylum *Proteobacteria* dominates (62.5%) with representative members from *α*-(*Brevundimonas*, *Sinorhizobium*, *Ensifer*), *β*-(*Achromobacter*), and *γ-Proteobacteria* (*Xanthomonas*, *Methylophaga*, *Colwellia*, etc.) classes. Other phyla detected are *Nitrospirae* (*Nitrospira*), *Cyanobacteria* (*Pseudanabaena*), and *Firmicutes* (*Symbiobacterium*) each constituting 12.5% of the taxonomic affiliation (Table 1).

### Starch and sucrose metabolism

Functional characterization of AB1 metagenome detected 18 enzymes annotated for starch and sucrose metabolism (Fig. 4). These include α-glucosidase (EC: 3.2.1.20), β-glucosidase (EC: 3.2.1.21), and trehalose 6-phosphate

**Table 2** List of AB1 microcosm sequences annotated for carbohydrate-active enzymes (CAZy) and their taxonomic affiliations

| Sequence ID | Microorganisms | CAZy Family/enzymes |
|---|---|---|
| NODE_6017_length_143_cov_1.160839_1_211_- | *Sphingobium yanoikuyae* | AA3<br>GMC family oxidoreductase |
| NODE_2589_length_175_cov_1.097143_1_243_- | *Burkholderia cenocepacia* | AA3<br>GMC family oxidoreductase |
| NODE_5399_length_154_cov_2.253247_1_222_- | *Pseudomonas* sp. LFM046 | AA3<br>GMC family oxidoreductase |
| NODE_5122_length_210_cov_1.723809_1_278_- | *Myxococcus fulvus* | AA5<br>Galactose oxidase; *N*-acetylneuraminic acid mutarotase |
| NODE_6741_length_192_cov_2.000000_17_260_- | *Dyella jiangningensis* | AA5<br>Cell envelope biogenesis protein OmpA |
| NODE_6279_length_159_cov_10.132075_17_227_- | *Stenotrophomonas panacihumi* | AA6<br>RidA family protein |
| NODE_5545_length_215_cov_3.297674_1_283_- | *Burkholderia multivorans* | AA6<br>NAD(P)H-dependent oxidoreductases |
| NODE_846_length_148_cov_3.243243_1_215_+ | *Burkholderia cenocepacia* | CBM13<br>Fe-S protein assembly chaperone HscA |
| NODE_2953_length_216_cov_2.287037_1_284_- | *Luteibacter* sp. 329MFSha | CBM2<br>Polyisoprenoid binding protein Yce1 |
| NODE_2277_length_196_cov_2.520408_1_263_+ | *Terriglobus saanensis* SP1PR4 | CBM32\|GH64<br>β-Galactosidase/β-glucuronidase |
| NODE_6409_length_192_cov_1.015625_1_260_- | *Pseudomonas aeruginosa* M18, *Pseudomonas aeruginosa* PAO1, *Pseudomonas aeruginosa* DK2, *Pseudomonas aeruginosa* LESB58, *Pseudomonas aeruginosa* UCBPP-PA14, *Pseudomonas aeruginosa* NCGM2.S1, *Pseudomonas aeruginosa* PA7, *Pseudomonas aeruginosa* B136–33 | CBM48\|GH13<br>Glycogen debranching enzyme, Glgx |
| NODE_1616_length_151_cov_5.688742_1_219_- | *Desulfobacteriales* bacterium S3730MH5, *Desulfurivibrio alkaliphilus* AHT2 | CBM48\|GH13<br>Glycogen debranching enzyme, Glgx |
| NODE_4632_length_223_cov_1.408072_1_291_- | *Methylocella silvestris* BL2 | CBM48\|GH13<br>Glycogen debranching enzyme, Glgx |
| NODE_5593_length_203_cov_1.793103_1_271_- | *Zymobacter palmae* | CBM50<br>N-Acetylmuramoyl-L-alanine amidase |
| NODE_5996_length_211_cov_1.436019_1_278_+ | *Massilia* sp. WF1 | CBM50<br>Bifunctional metallophosphatase/5′-nucleotidase |
| NODE_3645_length_144_cov_1.305556_1_212_+ | *Ralstonia* sp. 25mfcol4.1 | CBM50<br>Bifunctional metallophosphatase/5′-nucleotidase |
| NODE_6551_length_305_cov_1.183607_1_373_- | *Terriglobus* sp. TAA 43 | CBM6<br>Trehalose utilization; bacterial ThuA like proteins |
| NODE_2354_length_166_cov_1.963855_1_234_- | *Caulobacter crescentus* CB15 | CE10<br>S9 family peptidase |
| NODE_1193_length_268_cov_1.029851_1_336_- | *Pseudomonas* sp. ATCC 13867 | CE10<br>Alpha/beta hydrolase |
| NODE_782_length_156_cov_4.596154_1_224_- | *Sphingomonas sanxanigenens* | CE4<br>Polysaccharide deacetylase |
| NODE_7975_length_185_cov_1.070270_1_253_- | *Edaphobacter aggregans* | CE9<br>N-Acetylglucosamine-6-phosphate deacetylase |
| NODE_5120_length_169_cov_1.633136_1_237_- | *Enterobacter cloacae* complex sp. GN02208 | GH1<br>Mannose-6-phosphate isomerase |
| NODE_3720_length_211_cov_1.113744_1_279_- | *Pseudomonas jinjuensis* | GH1 |

**Table 2** List of AB1 microcosm sequences annotated for carbohydrate-active enzymes (CAZy) and their taxonomic affiliations *(Continued)*

| Sequence ID | Microorganisms | CAZy Family/enzymes |
|---|---|---|
| | | Polysaccharide biosynthesis protein |
| NODE_784_length_167_cov_6.101797_1_234_+ | Candidate division NC10 bacterium | GH1<br>Glucokinase |
| NODE_5198_length_149_cov_1.000000_1_217_+ | *Pseudomonas* sp. ATCC 13867 | GH1<br>Polysaccharide biosynthesis protein |
| NODE_3278_length_268_cov_1.171642_1_336_- | *Pseudomonas* sp. ATCC 13867 | GH1<br>dTDP-4-dehydrorhamnose reductase |
| NODE_5480_length_258_cov_1.608527_1_326_+ | *Verrucomicrobia* bacterium TMED102 | GH1<br>UDP-glucose-4-epimerase |
| NODE_4670_length_180_cov_13.094444_1_248_- | *Pseudomonas nitroreducens* | GH1<br>Polysaccharide biosynthesis protein |
| NODE_5370_length_265_cov_1.532076_1_333_- | *Novosphingobium panipatense* | GH10<br>Endo-1,4-beta-xylanase |
| NODE_1568_length_228_cov_1.947368_1_296_+ | *Sphingobium yanoikuyae* | GH102<br>Membrane-bound lytic murein transglycosylase A (Mlt A) |
| NODE_6766_length_185_cov_1.437838_1_253_- | *Acidobacterium capsulatum* ATCC 51196 | GH13<br>α-Amylase; maltose α-D-glucosyltransferase |
| NODE_5915_length_232_cov_1.948276_1_300_- | *Pseudomonas aeruginosa* M18, *Pseudomonas aeruginosa* DK2, *Pseudomonas aeruginosa* B136–33, *Pseudomonas aeruginosa* PAO1, *Pseudomonas aeruginosa* NCGM2.S1, *Pseudomonas aeruginosa* LESB58 | GH13<br>Maltose α-D-glucosyltransferase, trehalose synthase, α-amylase |
| NODE_5876_length_208_cov_1.855769_1_276_- | *Klebsiella pneumoniae* NTUH-K2044, *Klebsiella pneumoniae* subsp. pneumoniae 1084, *Klebsiella pneumoniae* subsp. pneumoniae MGH 78578 ATCC 700721, *Klebsiella pneumoniae, Klebsiella pneumoniae* KCTC 2242 | GH13\|CBM34<br>Maltodextrin glucosidase |
| NODE_7762_length_298_cov_1.144295_1_366_- | *Klebsiella pneumoniae* subsp. pneumoniae HS11286, *Klebsiella pneumoniae* NTUH-K2044, *Klebsiella variicola* At-22, *Klebsiella pneumoniae* subsp. pneumoniae 1084, *Klebsiella pneumoniae* KCTC 2242, *Klebsiella pneumoniae* 342 | GH23<br>Transglycosylase, Slt family |
| NODE_2599_length_298_cov_1.399329_1_366_- | *Pantoea* sp. AS1 | GH23<br>Lytic transglycosylase F |
| NODE_2200_length_253_cov_7.509881_1_321_- | *Azoarcus* sp. BH72 | GH23<br>Transglycosylase |
| NODE_5069_length_206_cov_1.009709_1_218_+ | *Paracoccus alcaliphilus* | GH23<br>Transglycosylase SLT domain-containing protein |
| NODE_662_length_247_cov_1.692308_1_315_+ | *Sphingobium yanoikuyae* | GH28<br>Exo-poly-alpha-D-galacturonosidase |
| NODE_3908_length_235_cov_2.161702_1_303_- | *Sphingobium yanoikuyae* | GH3<br>β-Glucosidase |
| NODE_1878_length_262_cov_3.122137_1_330_+ | *Sphingobium yanoikuyae* | GH3<br>β-Glucosidase |
| NODE_2573_length_202_cov_1.816832_1_270_- | *Dyella japonica* | GH3<br>Serine hydrolase |
| NODE_8591_length_292_cov_1.068493_1_360_- | *Dyella jiangningensis* | GH3<br>Serine hydrolase |
| NODE_6030_length_152_cov_3.302632_1_220_- | *Dyella japonica* | GH3<br>Serine hydrolase |
| NODE_3562_length_235_cov_1.365957_1_303_+ | *Sphingobium yanoikuyae* | GH3<br>β-Glucosidase |
| NODE_2535_length_220_cov_1.418182_17_288_- | *Enterobacter cloacae* EcWSU1 | GH31<br>α-Glucosidase |
| NODE_5102_length_288_cov_1.451389_1_356_+ | *Novosphingobium lindaniclasticum* | GH32<br>Sulfoxide reductase heme-binding subunit |

**Table 2** List of AB1 microcosm sequences annotated for carbohydrate-active enzymes (CAZy) and their taxonomic affiliations *(Continued)*

| Sequence ID | Microorganisms | CAZy Family/enzymes |
|---|---|---|
| | | YedZ |
| NODE_2470_length_195_cov_1.123077_1_263_- | *Pseudomonas* sp. ATCC 13867 | GH33<br>Dihydrodipicolinate synthase |
| NODE_3807_length_220_cov_1.659091_1_288_- | *Pseudomonas* spp. | GH37<br>Alpha, alpha-trehalase |
| NODE_3648_length_147_cov_3.619048_1_214_+ | *Sphingobium* sp. AP50 | GH37<br>Alpha, alpha-trehalase |
| NODE_6376_length_211_cov_1.677725_1_279_- | *Sphingobium yanoikuyae* | GH43<br>Arylsulfatase |
| NODE_3174_length_220_cov_1.795455_1_288_- | *Caulobacter segnis* ATCC 21756 | GH43<br>Glycosyl hydrolase |
| NODE_4474_length_145_cov_2.613793_1_213_- | *Sphingobium yanoikuyae* | GH51<br>Alpha-N-arabinofuranosidase, alpha-L-arabinofuranosidase |
| NODE_5021_length_187_cov_1.946524_1_255_- | *Pseudomonas* spp. | GH53<br>Diguanylate cyclase |
| NODE_806_length_180_cov_1.127778_1_248_- | *Mycobacterium* spp. | GH63<br>Glycogen debranching protein |
| NODE_3706_length_225_cov_1.666667_17_293_- | *Cupriavidus* sp. amp6 | GH72<br>GTP-binding protein YchF |
| NODE_2527_length_263_cov_1.615970_1_331_- | *Sphingomonas* sp. NFR15 | GH78<br>Alpha-L-rhamnosidase |
| NODE_5291_length_236_cov_1.822034_1_304_- | *Acidobacterium capsulatum* ATCC 51196 | GH79<br>Chain A, crystal structure of β-glucuronidase |
| NODE_7967_length_218_cov_2.000000_1_286_- | *Terriglobus* sp. TAA 43 | GH8<br>Endoglucanase |
| NODE_8032_length_165_cov_1.351515_1_232_+ | *Terriglobus* sp. TAA 43 | GH92<br>Alpha-mannosidase |
| NODE_878_length_314_cov_4.114650_17_382_- | *Klebsiella pneumoniae* | GH92<br>Short-chain dehydrogenase/reductase SDR |
| NODE_6075_length_277_cov_1.241877_1_345_- | *Novosphingobium panipatense* | GH92<br>Metallophosphoesterase |
| NODE_5945_length_252_cov_1.484127_1_320_+ | *Enterobacteriaceae* bacterium DC416, *Pantoea cypripedii* | GT1<br>Zeaxanthin glucosyltransferase |
| NODE_4367_length_123_cov_3.170732_1_191_+ | *Pseudomonas* spp. | GT2<br>Glucan biosynthesis glucosyltransferase H |
| NODE_6688_length_238_cov_1.373950_1_306_- | *Luteibacter* spp. | GT2<br>Sensor histidine kinase |
| NODE_429_length_307_cov_1.420195_1_375_- | *Sphingobium yanoikuyae* | GT2<br>Cellulose synthase catalytic subunit (UDP-forming) |
| NODE_8608_length_193_cov_1.663212_1_261_- | *Cupriavidus pauculus* | GT2<br>DNA-binding response regulator |
| NODE_8082_length_308_cov_1.019480_1_376_- | *Acidobacteria* bacterium | GT2<br>Cysteine-tRNA ligase |
| NODE_8019_length_166_cov_1.590361_1_234_- | *Bacillus* sp. LK2 | GT2<br>Non-ribosomal peptide synthetase |
| NODE_7554_length_274_cov_1.547445_1_342_- | *Pseudomonas* spp. | GT2<br>Non-ribosomal peptide synthetase |
| NODE_6340_length_262_cov_1.774809_1_330_- | *Acinetobacter baumannii* | GT2<br>DNA-binding response regulator |
| NODE_6666_length_155_cov_1.922581_1_223_- | *Pseudomonas* spp. | GT2<br>Diguanylate cyclase |
| NODE_4265_length_277_cov_1.346570_168_345_- | *Rhizobium* sp. NT-26 | GT20<br>Trehalose-6-phosphate synthase |

**Table 2** List of AB1 microcosm sequences annotated for carbohydrate-active enzymes (CAZy) and their taxonomic affiliations *(Continued)*

| Sequence ID | Microorganisms | CAZy Family/enzymes |
|---|---|---|
| NODE_1784_length_280_cov_1.378571_1_348_- | *Sphingomonas* sp. MM-1 | GT28<br>Undecaprenyl-diphosphomuramoyl pentapeptide beta-N-acetylglucosaminyltransferase |
| NODE_6862_length_207_cov_1.608696_1_275_- | *Phenylobacterium* spp. | GT28<br>Rod shape-determining protein-RodA |
| NODE_2164_length_148_cov_1.925676_1_216_- | *Azospirillum brasilense* Sp245 | GT4<br>UDP-galactopyranose mutase |
| NODE_3242_length_217_cov_1.447005_1_285_- | *Geobacter* sp. M18 | GT41<br>Glycosyltransferase |
| NODE_4396_length_252_cov_1.111111_17_320_- | *Pseudoxanthomonas spadix* BD-a59 | GT49<br>C4-dicarboxylate transport protein |
| NODE_8448_length_151_cov_1.350993_1_219_- | *Sphingobium yanoikuyae* | GT49<br>Dicarboxylate/amino acid:cation symporter |
| NODE_272_length_160_cov_1.768750_1_228_- | *Achromobacter* spp. | GT5<br>Histidine kinase |
| NODE_1519_length_282_cov_1.148936_1_350_+ | *Sphingobium* spp. | GT51<br>Penicillin-binding protein |
| NODE_7662_length_265_cov_1.418868_1_333_- | *Methylobacterium radiotolerans* JCM 2831 | GT51<br>Penicillin-binding protein 1A |
| NODE_5621_length_311_cov_1.035370_1_378_+ | *Sphingomonas* spp. | GT51<br>Penicillin-binding protein |
| NODE_2467_length_181_cov_1.049724_1_249_- | *Pseudomonas aeruginosa* PAO1, *Pseudomonas aeruginosa* B136–33, *Pseudomonas aeruginosa* DK2, *Pseudomonas aeruginosa* UCBPP-PA14, *Pseudomonas aeruginosa* M18, *Pseudomonas aeruginosa* LESB58, *Pseudomonas aeruginosa* NCGM2.S1 | GT51<br>Penicillin-binding protein 1B |
| NODE_6756_length_244_cov_1.627049_1_312_- | *Salinispora pacifica* | GT51<br>Penicillin-binding protein 1A |
| NODE_6722_length_178_cov_2.000000_1_246_+ | *Bacillus* spp. | GT51<br>Penicillin-binding protein |
| NODE_2550_length_228_cov_1.763158_1_296_- | *Terriglobus roseus* DSM 18391 | GT83<br>4-Amino-4-deoxy-L-arabinose transferase |
| NODE_742_length_196_cov_2.811224_1_264_- | *Caulobacter* spp. | GT83<br>Phospholipid carrier-dependent gly cosyltransferase; 4-amino-4-deoxy-L-arabinose transferase |
| NODE_6274_length_182_cov_7.450550_1_250_- | *Pseudomonas* spp. | GT9<br>Bifunctional heptose 7-phosphate kinase/heptose 1-phosphate adenyltransferase |
| NODE_8074_length_252_cov_1.190476_143_320_- | *Rhodoplanes* sp. Z2-YC6860 | GT9<br>Lipopolysaccharide heptosyltransferase I |

synthase (EC: 2.4.1.15). Others include trehalase (EC: 3.2.1.28), maltose α-D-glucosyltransferase (EC: 5.4.99.16), α-amylase (EC: 3.2.1.1), and UDP-glucuronate decarboxylase (EC: 4.1.1.35). Furthermore, various enzymes participating in cellulose and hemicellulose degradation and D-xylose degradation as well as starch and glycogen debranching enzymes among others were detected in AB1 metagenome (Table 1).

Taxonomic characterization of the annotated enzymes/genes for starch and sucrose metabolism revealed the dominance of *Protebacteria* phylum (91%) with representative members from α-(*Sphingobium*, *Agrobacterium*, *Beijerinckia*, *Methylocella*, etc.), β-(*Paraburkholderia*, *Achromobacter*), γ-(*Citrobacter*, *Azotobacter*, *Pseudomonas*, etc.), and δ-*Proteobacteria* (*Desulfurivibrio*) classes. *Acidobacteria* (9%) is the other phylum detected with representative members such as *Acidobacterium* and *Terriglobus* (Table 1).

### Pyruvate metabolism

Functional characterization of the AB1 metagenome revealed 13 enzymes annotated for pyruvate metabolism

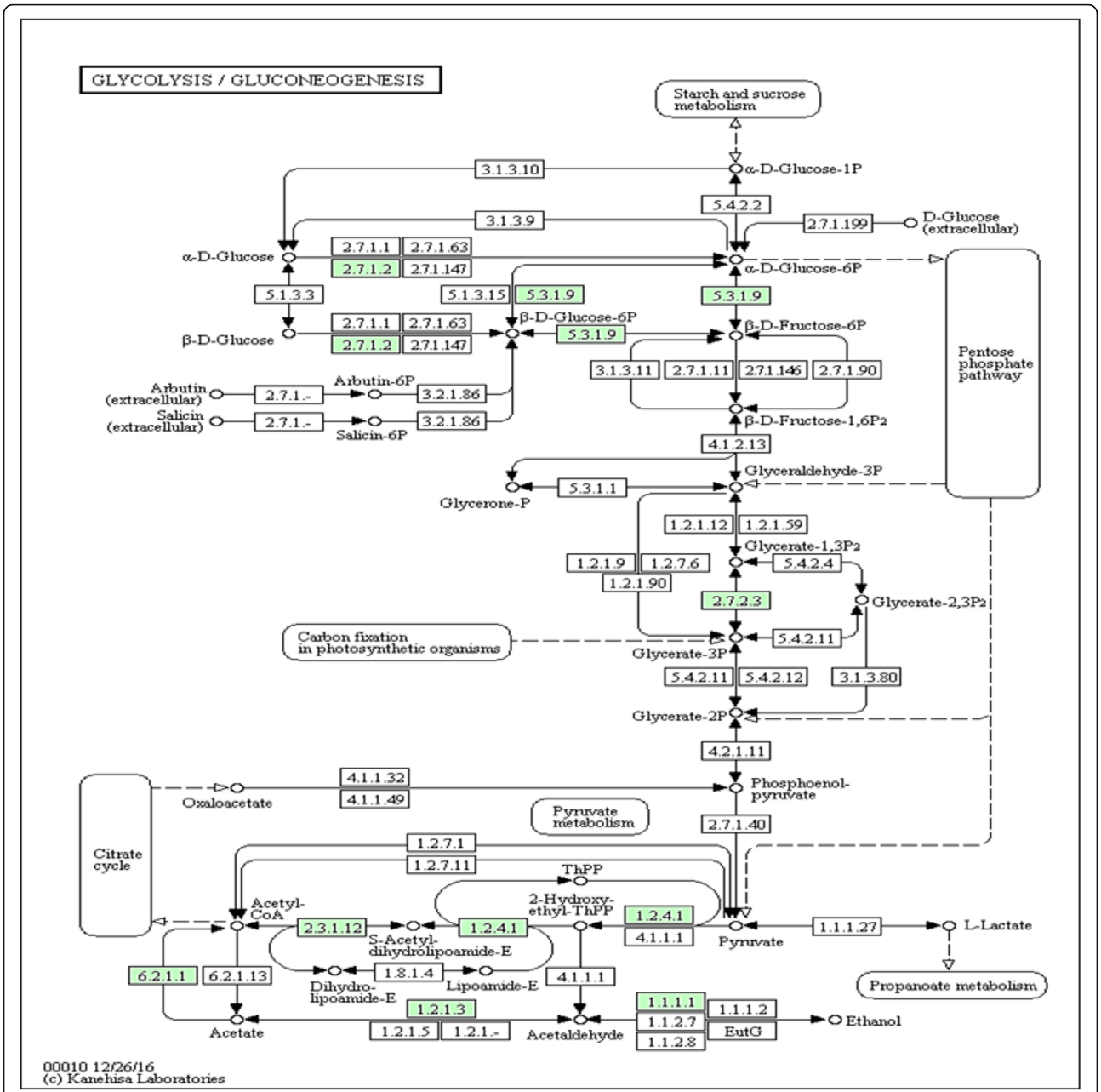


**Fig. 1** Pathway of glycolysis/gluconeogenesis performed using the KEGG pathway in KEGG Orthology. EC numbers in green are the enzymes identified in AB1 microcosm

(Fig. 5). These include Acetyl-CoA synthetase (EC: 6.2.1.1), pyruvate dehydrogenase E1 (EC: 1.2.4.1) and E2 (EC: 2.3.1.12) components, and acetyl-CoA carboxylase (EC: 6.4.1.2). Others include succinyl-CoA:acetate-CoA transferase (EC:2.8.3.18), L-lactate dehydrogenase (EC: 1.1.2.3), pyruvate dehydrogenase (quinone) (EC: 1.2.5.1), and glyoxylate/hydroxypyruvate reductase A (EC: 1.1.1.79 1.1.1.81). Furthermore, other pyruvate metabolism enzymes and enzyme participating in L-leucine biosynthesis such as isopropylmalate synthase are also annotated (Table 1).

Taxonomic affiliation of the annotated enzymes revealed that the microorganisms belong to the phylum *Proteobacteria* with representative members of α-(*Acidiphilium*, *Paracoccus*, *Novosphingobium*, etc.), and γ-*Proteobacteria* (*Dyella*, *Steroidobacter*, *Frateuria*, etc.) constituting 80% of the genera while the remaining 20% belong to β-*Proteobacteria* (*Thauera*, *Comamonas*, etc.). Other functionally annotated carbohydrate metabolic pathways detected in AB1 metagenome are listed in Table 1.

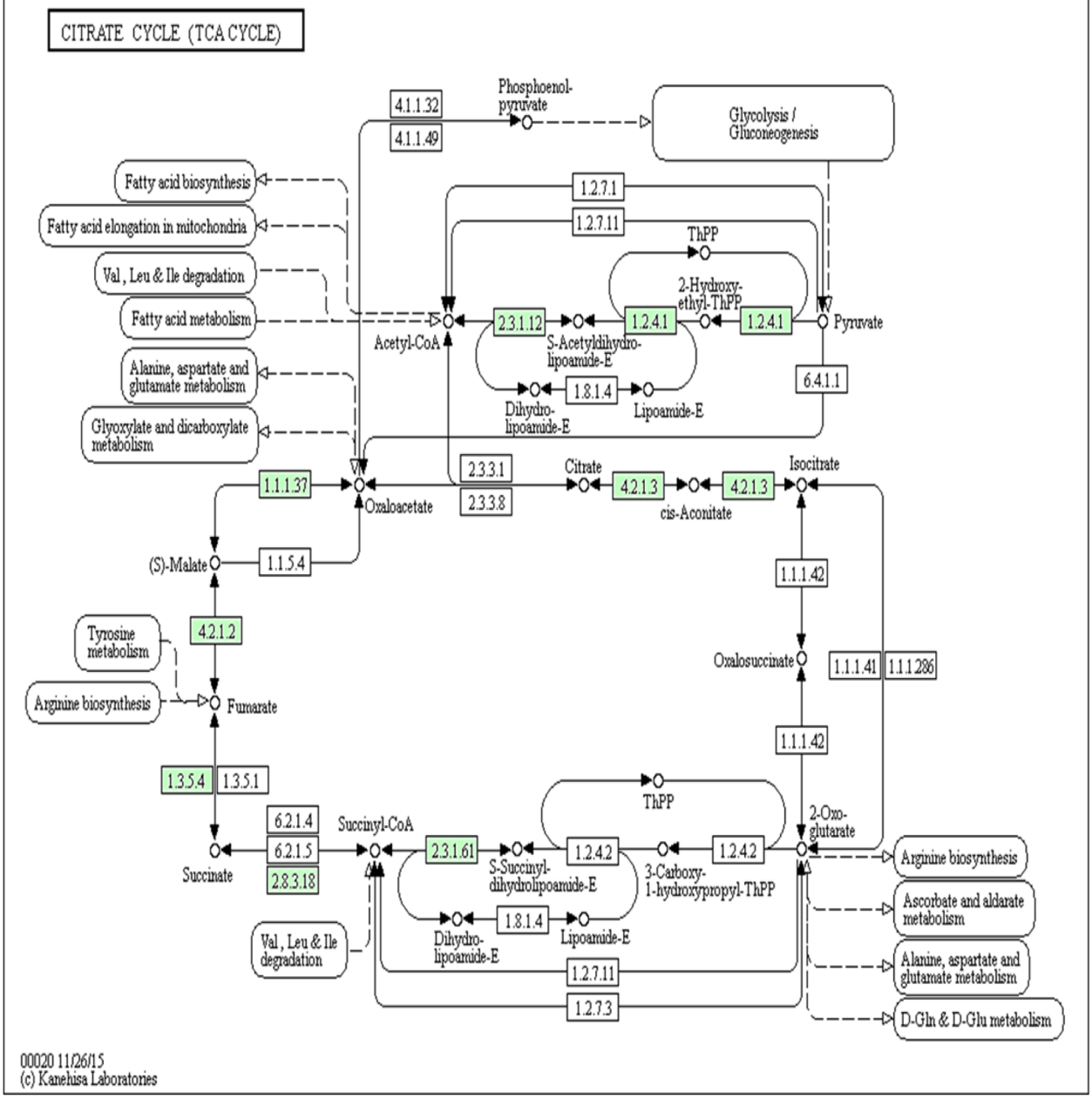


**Fig. 2** Pathway of citric acid (TCA) cycle performed using the KEGG pathway in KEGG Orthology. EC numbers in green are the enzymes identified in AB1 microcosm

### Annotation of AB1 metagenome for carbohydrate-active enzymes

Functional characterization of AB1 metagenome for carbohydrate-active enzymes using the CATS system revealed 89 sequences with 45 CAZy families (Table 2). Majority (21; 46.7%) of the carbohydrate-active enzymes belong to GH (glycoside hydrolase) families. Aside GH families, other carbohydrate-active enzyme families detected are GTs (glycosyltransferases; 11, 24.4%), and the CBMs (carbohydrate-binding modules; 7, 15.5%). Furthermore, CEs (carbohydrate esterases) and the AAs (auxiliary activities) were also detected. Of these, the predominant CAZy families detected in AB1 metagenome are GT2 (10%), GH1 (7.9%), GH3 (6.7%), GH13 (6.7%), and GT51 (6.7%), respectively. Taxonomic characterization of the CAZy families showed the

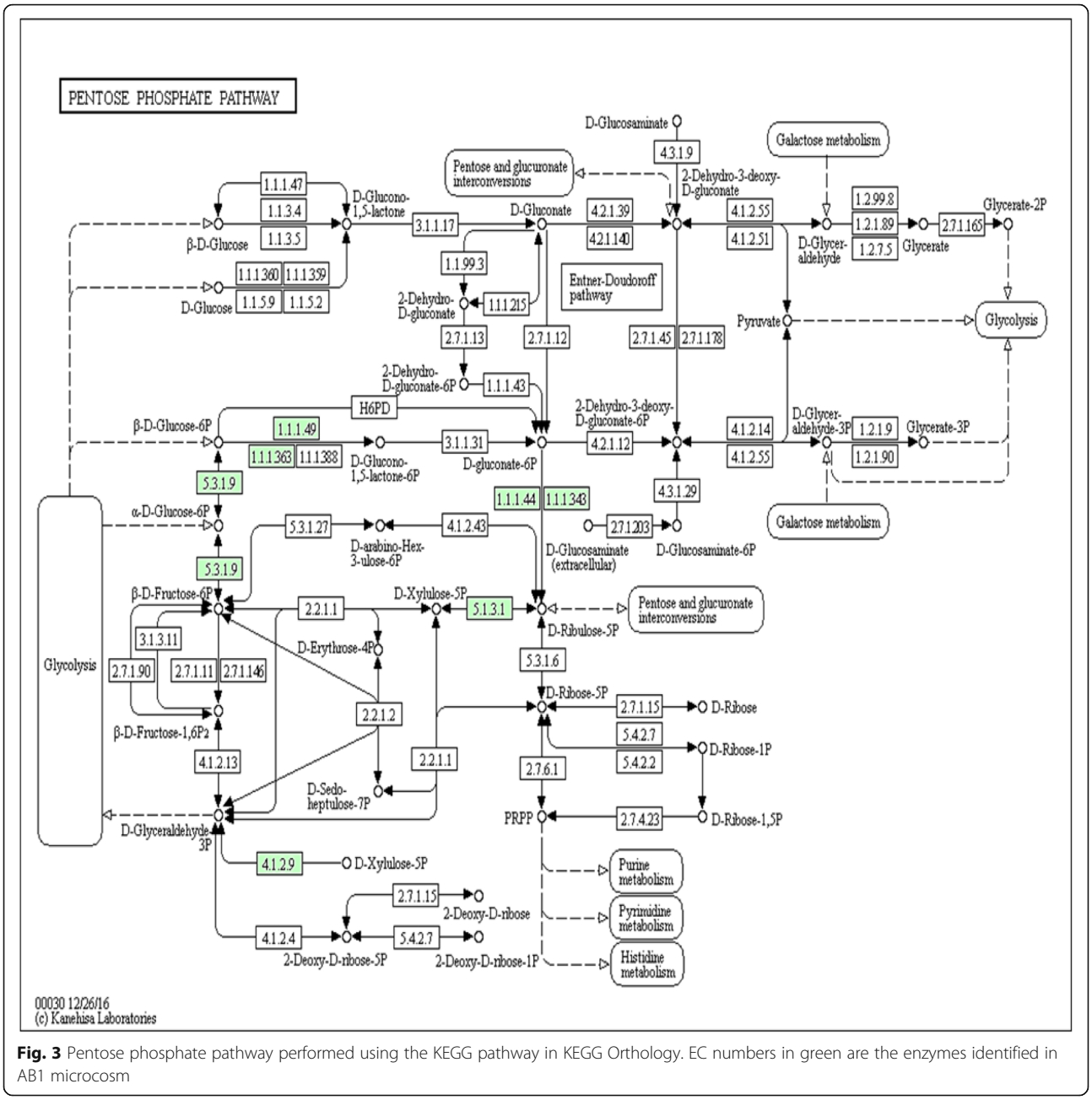


**Fig. 3** Pentose phosphate pathway performed using the KEGG pathway in KEGG Orthology. EC numbers in green are the enzymes identified in AB1 microcosm

preponderance of *Pseudomonas* (18%), *Sphingobium* (13.5%), *Terriglobus* (5.6%), and *Sphingomonas* species (4.5%), respectively.

## Discussion

This study was necessitated by the surprise detection of diverse carbohydrate metabolism pathways in SEO-perturbed soil not detected in the original agricultural soil. This might be attributed to the stress imposed by SEO contamination of the agricultural soil. The stress may force the microbial community to catabolize the energy-rich stored carbohydrates such as cellulose, starch, glycogen, and others to easily metabolized carbohydrate such as glucose as sources of carbon and energy to drive metabolic processes (Demoling et al. 2007). This assertion is strengthened by the fact that most of the carbohydrate-active enzymes detected in AB1 microcosm performed catabolic functions, breaking down stored carbohydrates to easily metabolizable monomers.

The AA (auxiliary activities) families detected in AB1 microcosm are AA3, AA5, and AA6, respectively. AA3 belong to the glucose-methanol-choline (GMC) oxidoreductases family. The functions of the enzymes include cellulose, hemicellulose, and lignin biodegradation

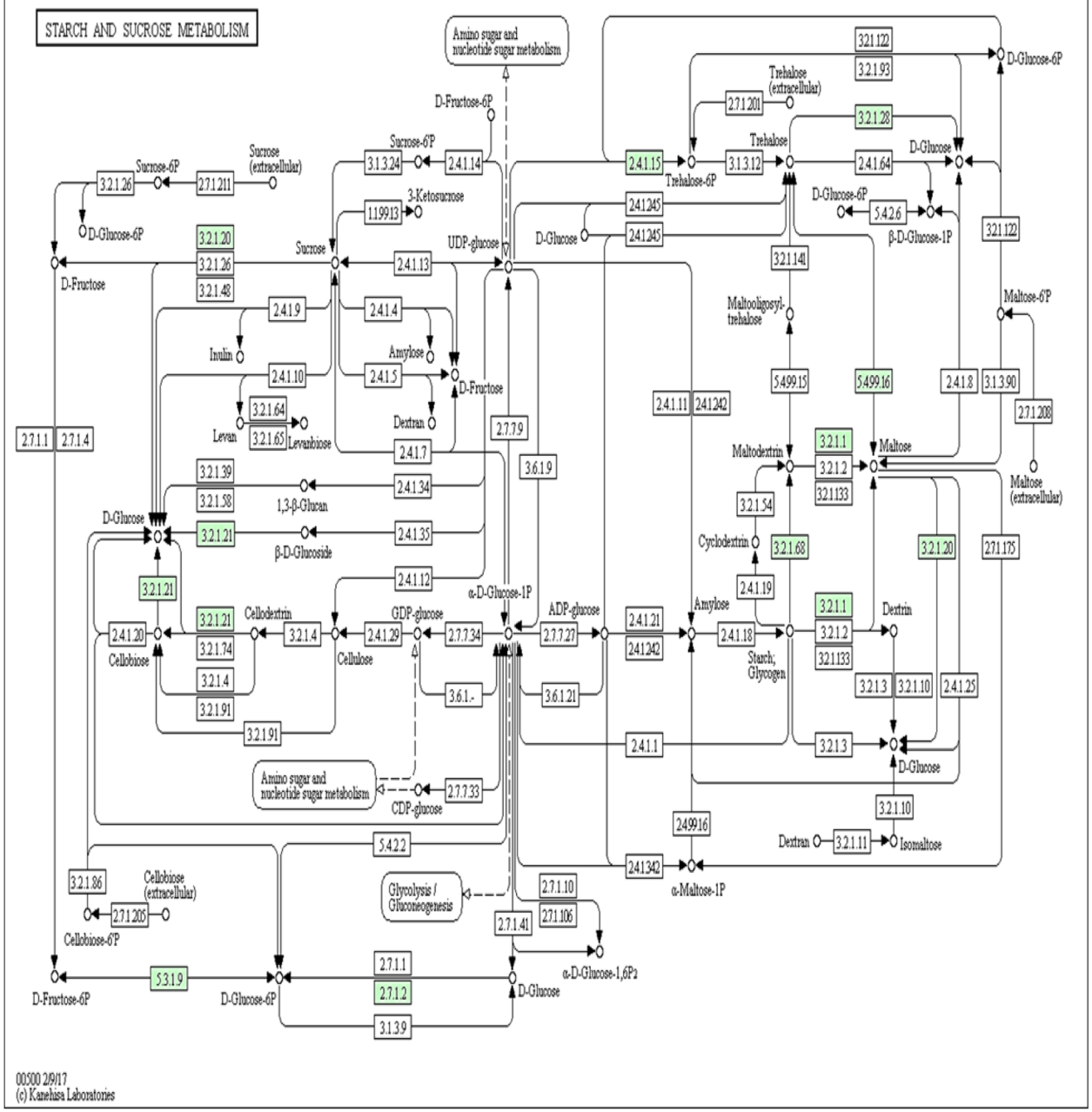


**Fig. 4** Pathway for starch and sucrose metabolism performed using the KEGG pathway in KEGG Orthology. EC numbers in green are the enzymes identified in AB1 microcosm

(CDH), oxidative dehydrogenation of several aromatic and aliphatic polyunsaturated alcohols leading to concomitant reduction of $O_2$ to $H_2O_2$ (AO), and oxidation of the hydroxyl group at the C1 position of sugars (GOX) (Eriksson et al. 1986; Kremer and Wood 1992; Martinez et al. 2005; Hernández-Ortega et al. 2012). Also retrieved from AB1 microcosm is AA5 with galactose oxidase, *N*-acetylneuraminic acid mutarotase, and cell envelope biogenesis protein, OmpA as representatives (Table 2). Galactose oxidases (GAO) are copper-containing enzymes that catalyze the oxidation of a primary alcohol and reduction of $O_2$ to $H_2O_2$ (Whittaker 2003). The enzyme catalyzes the oxidation of a wide range of carbohydrates (including galactose) and primary alcohols into the corresponding aldehydes, with the reduction of $O_2$ into $H_2O_2$ (Levasseur et al. 2013).

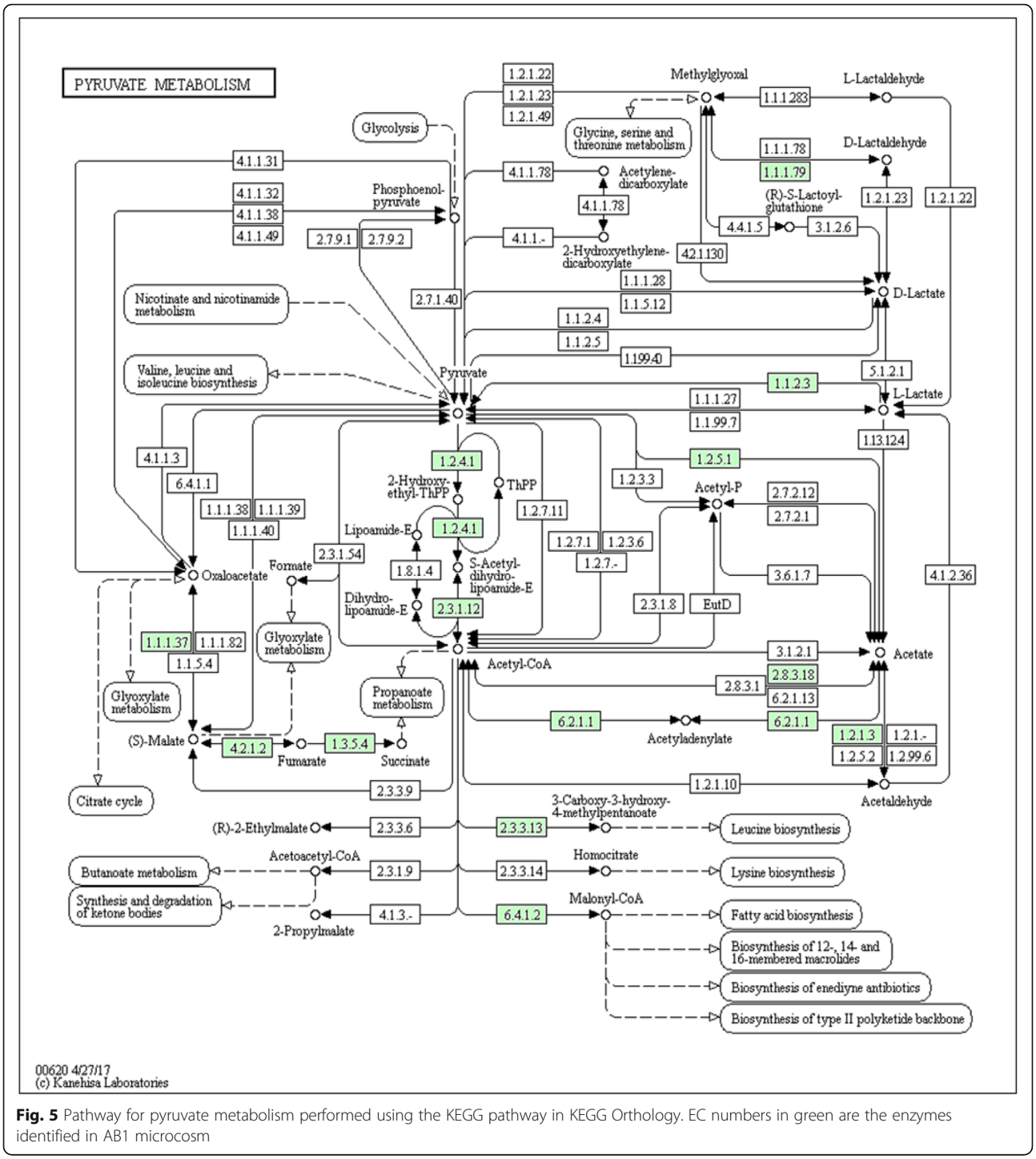


**Fig. 5** Pathway for pyruvate metabolism performed using the KEGG pathway in KEGG Orthology. EC numbers in green are the enzymes identified in AB1 microcosm

The detection of NAD(P)H-dependent oxidoreductases as representative enzymes of AA6 family is not surprising as all experimentally characterized proteins of AA6 family are 1,4-benzoquinone reductases, a member of NAD(P)H-dependent oxidoreductases. 1,4-Benzoquinone reductase is an intracellular enzyme that catalyzes the reduction of methoxylated lignin-derived quinones to hydroquinones via a ping-pong steady state kinetic mechanism (Brock and Gold 1996; Akileswaran et al. 1999).

A carbohydrate-binding module (CBM) is a contiguous amino acid sequence within a carbohydrate-active enzyme with a discrete fold having carbohydrate-binding activity (Shoseyov et al. 2006). CBMs contribute in the binding of the target substrates to carbohydrate-degrading enzymes, and its catalytic modules is widely believed to

perform three general roles defined as a proximity effect, a targeting function, and a disruptive function (Din et al. 1994; Bolam et al. 1998; Ali et al. 2001; Boraston et al. 2004; Herve et al. 2010). In this study, seven (7) CBM families namely CBM 2, 6, 13, 32, 34, 48, and 50 were detected in AB1 microcosm. CBM2 has polyisoprenoid-binding protein Yce1 as its representative. The primary function of CBM2 is to bind cellulose or xylan and bring it to proximity of cellulose- and xylan-degrading enzymes. Polyisoprenoid-binding proteins could be involved in substrate oxidation in order to facilitate the breakdown of the plant cell wall.

CBM13, classified as type-C CBM, with a pocket-like sugar-binding site generally recognizes one or two monosaccharide units within a polysaccharide (Fujimoto 2013). It shows specificity for the backbone of xylan, but it is found not only in xylanases (EC 3.2.1.8), but also in other GHs, several lyases, and glycosyltransferases (Fujimoto 2013). The detected CBM50 have two representatives, N-acetylmuramoyl-L-alanine amidase and bifunctional metallophosphatase/5′-nucleotidase. N-Acetylmuramoyl-L-alanine amidase is an autolysin that hydrolyzes the amide bond between N-acetylmuramoyl and L-amino acids in certain cell wall glycopeptides. The enzyme cleaves the septum to release daughter cells after cell division (Herbold and Glaser 1975; Ward et al. 1982). Bifunctional metallophosphatase/5′-nucleotidase (EC 3.6.1.45) catalyzes the degradation of periplasmic UDP-glucose to uridine, glucose-1-phosphate, and inorganic phosphate. Other CBMs detected in combination with GHs are CBM32, CBM64, and CBM48, which played prominent roles in facilitating the degradation of the complex carbohydrates by bringing the substrates in close contact to the enzymes.

The carbohydrate esterases (CEs) are widely used as biocatalysts in industrial processes and biotechnology (Jaeger and Reetz 1998; Bornscheuer 2002; Jaeger and Eggert 2002). In this study, only three (3) CE families—CE4, CE9, CE10—were detected in AB1 microcosm. However, the CE10 family has been nullified since most of its members appeared to be esterases active against non-carbohydrate substrates (Nakamura et al. 2017). In AB1 microcosm, CE4 is represented by polysaccharide deacetylase (Table 2), the general name given to members belonging to this family. Polysaccharide deacetylases catalyze the *N*- or *O*-deacetylation of xylan (Aspinall 1959), chitin (Tsigos et al. 2000), and peptidoglycans (Johannsen 1993). In AB1 microcosm, CE9 is represented by *N*-acetylglucosamine-6-phosphate deacetylase (EC 3.5.1.25) (Table 2). This enzyme is involved in the first step in the biosynthesis of amino sugar nucleotides. It catalyzes the hydrolysis of the *N*-acetyl group of *N*-acetylglucosamine-6-phosphate to yield glucosamine-6-phosphate and acetate (Vincent et al. 2004; Ferreira et al. 2006; Nakamura et al. 2017).

Glycoside hydrolases (GHs) are a widespread group of enzymes, which hydrolyze the glycosidic bond between two or more carbohydrates or between a carbohydrate and a non-carbohydrate moiety. In this study, 21 GH families were detected in AB1 microcosm. Six of these families (GH1, 10, 51, 53, 72, 79) belong to the GH-A clan (type I classical $(\beta/\alpha)_8$ glycosidases) (Naumoff 2011). GH1 is represented in AB1 microcosm by mannose-6-phosphate isomerase (catalyze the interconversion of fructose-6-phosphate and mannose-6-phosphate), polysaccharide biosynthesis protein (implicated in the production of polysaccharides), and glucokinase (catalyze the phosphorylation of glucose to glucose-6-phosphate). Others include dTDP-4-dehydrorhamnose reductase (catalyze the reduction of dTDP-6-deoxy-L-lyxo-4-hexulose to dTDP-L-rhamnose, the precursor of L-rhamnose, a component of the lipopolysaccharide core and several O antigen polysaccharides) and UDP-glucose-4-epimerase (facilitate the interconversion of UDP-glucose and UDP-galactose, which are precursors of glucose- and galactose-containing exopolysaccharides (EPS)) (Rahim et al. 2000; Fortina et al. 2003; Yoo et al. 2011).

The detection of endo-1,4-β-xylanase as the representative of GH10 in AB1 microcosm further affirms extensive deconstruction of complex plant polysaccharide in the agricultural soil in spite of spent engine oil contamination. The enzyme degrades the linear polysaccharide beta-1,4-xylan into xylose, thereby breaking down hemicellulose, one of the major components of plant cell walls (Beg et al. 2001). Functionally, similar to endo-1,4-β-xylanase are the alpha-N-arabinofuranosidase and alpha-L-arabinofuranosidase annotated for GH51, which act on alpha-L-arabinofuranosides, alpha-L-arabinans containing (1,3) and/or (1,5)-linkages, arabinoxylans, and arabinogalactans resulting in degradation of lignocelluloses (Numan and Bhosle 2006). The sole representative of GH53 in AB1 microcosm is diguanylate cyclase, which catalyzes the breakdown of guanosine triphosphate to diphosphate and cyclic di-3′,5′-guanylate (cyclic di-GMP). Cyclic di-GMP is an important messenger known for the control of cellulose biosynthesis as well as other diverse bacterial cellular functions such as virulence, motility, bioluminescence, adhesion, secretion, community behavior, biofilm formation, and cell differentiation (Hecht and Newton 1995; Galperin et al. 2001; D'Argenio and Miller 2004; Whiteley and Lee 2015).

The deconstruction of complex carbohydrates (e.g., cellulose, starch, glycogen, and xylan), mostly by microorganisms, releases short metabolizable oligosaccharides to the environment. This contributes to the functioning of an ecosystem and is essential for global carbon cycling (Berlemont and Martiny 2016). The detection of diverse

complex carbohydrate deconstruction enzymes spanning several GH families such as GH13 (GH-H clan), GH31 (GH-D clan), GH3, GH37, and 63 (GH-G clan), and GH8 (GH-M clan) reflects the importance of these enzymes to the AB1 ecosystem and global carbon cycling. The short metabolizable oligosaccharides generated are fed into the central carbohydrate metabolic pathways, which not only generate energy yielding products but also produce precursor metabolites for other pathways (Neidhardt et al. 1990; Noor et al. 2010).

The facts elucidated in the preceding paragraph underscore the importance of these pathways in the normal functioning of the AB1 soil microcosm ecosystem in spite of spent engine oil contamination. The precursor metabolites are the source materials for all biomass components made by the microbial community in the microcosm: amino acids, RNA and DNA molecules, lipids, lipopolysaccharides, and peptidoglycan monomers. The cells cannot do without these components. Thus, there is a dire need for activation of enzyme machinery needed for these essential pathways.

Glycosyltransferases (GTs) are enzymes that catalyze the synthesis of glycosidic linkages by the transfer of a sugar residue from a donor substrate to an acceptor (Lairson et al. 2008; Ardèvol and Rovira 2015). The structure of GTs adopts one of threefold, namely GT-A, GT-B, and GT-C (Gloster 2014). In this study, the GT2 family represents the GT-A fold (inverting). Enzymes/proteins annotated for GT2 family in AB1 microcosm are glucan glucosyltransferase H, sensor histidine kinase, cellulose synthase catalytic subunit (UDP-forming), cysteine-tRNA ligase, non-ribosomal peptide synthetase, and DNA-binding response regulator. These enzymes are involved in biosynthesis of osmoregulated periplasmic glucans, signal transduction pathways upstream of many processes including various metabolic, virulence and homeostatic pathways, synthesis of cellulose, cysteine metabolism, aminoacyl-tRNA biosynthesis, among others (Wolanin et al. 2002; Strieker et al. 2010; Fuhs et al. 2015; Fuhs and Hunter 2017).

Of the eleven (11) GT families retrieved from AB1 microcosm, seven (7) belong to the GT-B fold. They are GT1, GT20, GT28, GT4, GT41, GT5, and GT9. Aside from GT20, GT4, and GT5 having retaining functions, the others possess inverting functions (Schmid et al. 2016). The enzymes annotated for the GT families in GT-B and GT-C (GT83) folds perform diverse functions in AB1 microcosm. This include glycan biosynthesis (GT20), peptidoglycan biosynthesis and maintenance of the rod cell shape and elongation of the lateral wall of the cell (GT28), lipopolysaccharide biosynthesis and bacterial outer membrane biogenesis (GT4, GT9), and lipopolysaccharide metabolism and cell wall/membrane/envelope biogenesis (GT83) (Henriques et al. 1998; van Heijenoort 2001; Bretscher et al. 2005).

It is noteworthy that most of the enzymes/genes annotated for AB1 microcosm using the CAZy database were also detected using the GhostKOALA, COG, and NCBI's CDD databases. This further affirms the importance of using diverse metagenomic analysis tools to characterize environmental metagenomes. The deployment of these tools allows for unbiased and comprehensive detection and elucidation of targeted genes and enzymes.

Taxonomic classification of the AB1 sequences annotated for enzymes/genes involved in carbohydrate metabolism revealed the preponderance of *Proteobacteria* phylum. This is not surprising as the phylum is a compendium of members that are metabolically diverse with ability to utilize a wide range of organic and inorganic compounds and of living under diverse environmental conditions. The dominance of *Pseudomonas* species may be attributed to its metabolic versatility and genetic plasticity. The genus is reputed to possess remarkable nutritional versatility with majority growing on more than 50 different substrates, some on more than 100 employing various pathways (Stanier et al. 1966; Palleroni and Doudoroff 1972; Palleroni 1993). Their ability to thrive in habitats with temperature range of 4–42 °C, a pH between 4 and 8, and containing simple or complex organic compounds explains their ubiquity, survivability, and dominance in diverse soil environments (Moore et al. 2006). Furthermore, the detection of various carbohydrate-active enzymes and genes cutting across various phyla and not limited to related phylogenetic groups is a clear indication of involvement of lateral and horizontal gene transfer within the microbial community.

## Conclusions

In summary, while the effects of SEO contamination on the microbial community structure and function in AB1 microcosm are profound, the microbial community maintains essential metabolic and cellular processes. The stress imposed on the agricultural soil by SEO contamination allowed the shift of the microbial community to utilization of stored and easily metabolized carbohydrates. The deconstruction of complex carbohydrate compounds in the soil liberates metabolizable oligosaccharides, which generates products (via diverse metabolic pathways) that are precursor metabolites for the biosynthesis of nucleotides, amino acids, lipids, lipopolysaccharides, and peptidoglycan monomers needed as building blocks and cellular constituents. Assemblage of complex carbohydrates through enzymatic formation of glycosidic linkages mediated by glycosyltransferases assists in construction of new cells and its constituents while the detection of industrial and biotechnologically

important enzymes such as xylanases, endoglucanases, and α- and β-glucosidases revealed the soil as important reservoir of these enzymes. The use of various metagenomics analysis tools enunciated in this study not only reveal the diverse carbohydrate metabolic pathways in AB1 microcosm but also enrich our understanding of the various carbon fluxes in SEO-contaminated agricultural soil.

## Additional file

**Additional file 1:** Carbohydrate metabolism enzymes/genes in AB1 microcosm detected by KEGG GhostKOALA and COG Databases. (DOCX 44 kb)

### Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files.

### Author's contributions

LBS designed the work, performed the experiments, analyzed the data, and wrote the manuscript. The author read and approved the final manuscript.

### Ethics approval and consent to participate

Not applicable

### Consent for publication

Not applicable

### Competing interests

The author declares that he has no competing interest.

## Publisher's Note



Received: 2 July 2018 Accepted: 28 September 2018
Published online: 11 October 2018

### References

Abbott DW, Macauley MS, Vocadlo DJ, Boraston AB (2009) *Streptococcus pneumoniae* endohexosaminidase D: structural and mechanistic insight into substrate-assisted catalysis in family 85 glycoside hydrolases. J Biol Chem 284(17):11676–11689

Akileswaran L, Brock BJ, Cereghino JL, Gold MH (1999) 1,4-benzaquinone reductase from *Phanerochaete chrysosporium*: cDNA cloning and regulation of expression. Appl Environ Microbiol 65(2):415–421

Ali MK, Hayashi H, Karita S, Goto M, Kimura T, Sakka K, Ohmiya K (2001) Importance of the carbohydrate-binding module of *Clostridium stercorarium* Xyn10B to xylan hydrolysis. Biosci Biotechnol Biochem 65:41–47

Ardèvol A, Rovira C (2015) Reaction mechanisms in carbohydrate-active enzymes: glycoside hydrolases and glycosyltransferases. Insights from ab initio quantum mechanics /molecular mechanics dynamic simulations. J Am Chem Soc 137:7528–7547

Aspinall GO (1959) In: Melville LW (ed) Structural chemistry of the hemicelluloses, Adv Carbohydr Chem. Academic Press, pp 429–468

Bardgett RD, Freeman C, Ostle NJ (2008) Microbial contributions to climate change through carbon cycle feedbacks. ISME J 2(8):805–814

Beg QK, Kapoor M, Mahajan L, Hoondal GS (2001) Microbial xylanases and their industrial applications: a review. Appl Microbiol Biotechnol 56(3–4):326–338

Berlemont R, Martiny AC (2016) Glycoside hydrolases across environmental microbial communities. PLoS Comput Biol 12(12):e1005300

Bolam DN, Ciruela A, McQueen-Mason S et al (1998) *Pseudomonas* cellulose-binding domains mediate their effects by increasing enzyme substrate proximity. Biochem J 331(Pt 3):775–781

Boraston AB, Bolam DN, Gilbert HJ, Davies GJ (2004) Carbohydrate-binding modules: fine tuning polysaccharide recognition. Biochem J 382:769–781

Bornscheuer UT (2002) Microbial carboxyl esterases: classification, properties and application in biocatalysis. FEMS Microbiol Rev 26(1):73–81

Bretscher LE, Morrell MT, Funk AL, Klug CS (2005) Purification and characterization of the L-Ara4N transferase protein ArnT from *Salmonella typhimurium*. Protein Expr Purif 46(1):33–39

Brock BJ, Gold MH (1996) 1,4-benzaquinone reductase from the basidiomycete *Phanerochaete chrysosporium*: spectral and kinetic analysis. Arch Biochem Biophys 331:31–40

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The carbohydrate-active enzymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res 37(Database):D233–D238

Christov LP, Prior BA (1993) Esterases of xylan-degrading microorganisms: production, properties, and significance. Enzym Microb Technol 15(6): 460–475

D'Argenio DA, Miller SI (2004) Cyclic di-GMP as a bacterial second messenger. Microbiology 150(Pt 8):2797–2502

Demoling F, Figueroa D, Bååth E (2007) Comparison of factors limiting bacterial growth in different soils. Soil Biol Biochem 39:2485–2495

Din N, Damude HG, Gilkes NR, Miller RC Jr, Warren RA, Kilburn DG (1994) C1-cx revisited: intramolecular synergism in a cellulase. Proc Natl Acad Sci U S A 91: 11383–11387

Eriksson KE, Pettersson B, Volc J, Musílek V (1986) Formation and partial characterization of glucose-2-oxidase, a $H_2O_2$ producing enzyme in *Phanerochaete chrysosporium*. Appl Microbiol Biotechnol 23:257–262

Ferreira FM, Mendoza-Hernandez G, Castañeda-Bueno M, Aparicio R, Fischer H, Calcagno ML, Oliva G (2006) Structural analysis of N-acetylglucosamine-6-phosphate deacetylase apoenzyme from *Escherichia coli*. J Mol Biol 359(2): 308–321

Fortina MG, Ricci G, Mora D, Guglielmetti S, Manachini PL (2003) Unusual organization for lactose and galactose gene clusters in *Lactobacillus helveticus*. Appl Environ Microbiol 69(6):3238–3243

Fuhs SR, Hunter T (2017) Phisphorylation: the emergence of histidine phosphorylation as a reversible regulatory modification. Curr Opin Cell Biol 45:8–16

Fuhs SR, Meisenhelder J, Aslanian A et al (2015) Monoclonal 1- and 3-phosphohistidine antibodies: new tools to study histidine phosphorylation. Cell 162:198–210

Fujimoto Z (2013) Structure and function of carbohydrate-binding module families 13 and 42 of glycoside hydrolases, comprising a β-trefoil fold. Biosci Biotechnol Biochem 77(7):1363–1371

Galperin MY, Nikolskaya AN, Koonin EV (2001) Novel domains of the prokaryotic two-component signal transduction systems. FEMS Microbiol Lett 203(1):11–21

Gloster TM (2014) Advances in understanding glycosyltransferases from a structural perspective. Curr Opin Struct Biol 28:131–141

Gupta A, Verma JP (2015) Sustainable bioethanol production from agro-residues: a review. Renew Sust Energ Rev 41:550–567

Hecht GB, Newton A (1995) Identification of a novel response regulator required for the swarmer-to-stalked-cell transition in *Caulobacter crescentus*. J Bacteriol 177(21):6223–6229

Henriques AO, Glaser P, Piggot PJ, Moran CR Jr (1998) Control of cell shape and elongation by the rodA gene in *Bacillus subtilis*. Mol Microbiol 28(2):235–247

Herbold DR, Glaser L (1975) *Bacillus subtilis* N-acetylmuramic acid L-alanine amidase. J Biol Chem 250(5):1676–1682

Hernández-Ortega A, Ferreira P, Martínez AT (2012) Fungal aryl-alcohol oxidase: a peroxide-producing flavoenzyme involved in lignin degradation. Appl Microbiol Biotechnol 93:1395–1410

Herve C, Rogowski A, Blake AW, Marcus SE, Gilbert HJ, Knox JP (2010) Carbohydrate binding modules promote the enzymatic deconstruction of intact plant cell walls by targeting and proximity effects. Proc Natl Acad Sci U S A 107:15293–15298

Jaeger KE, Eggert T (2002) Lipases for biotechnology. Curr Opin Biotechnol 13(4): 390–397

Jaeger KE, Reetz MT (1998) Microbial lipases form versatile tools for biotechnology. Trends Biotechnol 16(9):396–403

Johannsen L (1993) Biological properties of bacterial peptidoglycan. APMIS 101(1–6):337–344

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

Kremer SM, Wood PM (1992) Cellobiose oxidase from *Phanerochaete chrysosporium* as a source of Fenton's reagent. Biochem Soc Trans 20:110S
Lairson LL, Henrissat B, Davies GJ, Withers SG (2008) Glycosyltransferases: structures, functions, and mechanisms. Annu Rev Biochem 77:521–555
Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B (2013) Expansion of the enzymatic repertoire of the CAZy database to integrate auxillary redox enzymes. Biotechnol Biofuels 6:41
Ley RE, Lozupone CA, Hamady M, Knight R, Gordon JI (2008) Worlds within worlds: evolution of the vertebrate gut microbiota. Nat Rev Microbiol 6: 776–788
Lombard V, Golaconda RH, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42: D490–D495
Marchler-Bauer A, Derbyshire MK, Gonzales NR et al (2015) CDD: NCBI's conserved domain database. Nucleic Acids Res 43(D):222–D226
Martinez AT, Speranza M, Ruiz-Duenas FG et al (2005) Biodegradation of lignocellulosics: microbial, chemical, and enzymatic aspects of fungal attack of lignin. Int Microbiol 8:195–204
Moore ERB, Tindall BJ, VAPM DS, Pieper DH, Ramos J-L, Palleroni NJ (2006) Nonmedical: *Pseudomonas*. Prokaryotes 6:646–703
Nakamura AM, Nascimento AS, Polikarpov I (2017) Structural diversity of carbohydrate esterases. Biotechnol Res Innovation. https://doi.org/10.1016/j.biori.2017.02.001
Naumoff DG (2011) Hierarchical classification of glycoside hydrolases. Biochem Mosc 76(6):622–635
Neidhardt FC, Ingraham JL, Schaechter M (1990) Physiology of the bacterial cell: a molecular approach. Sinauer Associates, Sunderland
Noor E, Eden E, Milo R, Alon U (2010) Central carbon metabolism as a minimal biochemical walk between precursors for biomass and energy. Mol Cell 39: 809–820
Numan MT, Bhosle NB (2006) Alpha-L-arabinofuranosidases: the potential application in biotechnology. J Ind Microbiol Biotechnol 33(4):247–260
Palleroni NJ (1993) *Pseudomonas* classification. A new case history in the taxonomy of gram-negative bacteria. Antonie Van Leeuwenhoek 64:231–251
Palleroni NJ, Doudoroff M (1972) Some properties and subdivisions of the genus *Pseudomonas*. Annu Rev Phytopathol 10:73–100
Park BH, Karpinets TV, Syed MH, Leuze MR, Uberbacher EC (2010) Cazymes analysis toolkit (CAT): web service for searching and analysing carbohydrate active enzymes in a newly sequenced organism using CAZY database. Glycobiology 20(12):1574–1584
Rahim R, Burrows LL, Monteiro MA, Perry MB, Lam JS (2000) Involvement of the rml locus in core oligosaccharide and O polysaccharide assembly in *Pseudomonas aeruginosa*. Microbiology 146(Pt 11):2803–2814
Rho M, Tang H, Ye Y (2010) FragGeneScan: predicting genes in short and error-prone reads. Nucleic Acid Res 38:20–191
Salam LB, Obayori OS, Nwaokorie FO, Suleiman A, Mustapha R (2017) Metagenomic insights into effects of spent engine oil perturbation on the microbial community composition and function in a tropical agricultural soil. Environ Sci Pollut Res 24:7139–7159
Schmid J, Heider D, Wendel NJ, Sperl N, Sieber V (2016) Bacterial glycosyltransferases: challenges and opportunities of a highly diverse enzyme class towards tailoring natural products. Front Microbiol 7:182
Shoseyov O, Shani Z, Levy I (2006) Carbohydrate binding modules: biochemical properties and novel applications. Microbiol Mol Biol Rev 70:283–295
Stanier RY, Palleroni NJ, Doudoroff M (1966) The aerobic pseudomonads: a taxonomic study. J Gen Microbiol 43:159–271
Strieker M, Tanovic A, Marahiel MA (2010) Nonribosomal peptide synthetases: structures and dynamics. Curr Opin Struct Biol 20(2):234–240
Tatusov RL, Natale DA, Garkavtsev IV et al (2001) The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res 29:22–28
Tsigos I, Martinou A, Kafetzopoulos D, Bouriotis V (2000) Chitin deacetylases: new, versatile tools in biotechnology. Trends Biotechnol 18(7):305–312
van Heijenoort J (2001) Recent advances in the formation of the bacterial peptidoglycan monomer unit. Nat Prod Rep 18(5):503–519
Vincent F, Yates D, Garman E, Davies GJ, Brannigan JA (2004) The three-dimensional structure of the N-acetylglucosamine-6-phosphate deacetylase, NagA, from *Bacillus subtilis*: a member of the urease superfamily. J Biol Chem 279(4):2809–2816
Ward JB, Curtis CA, Taylor C, Buxton RS (1982) Purification and characterization of two phage PBSX-induced lytic enzymes of Bacillus subtilis 168: an N-acetylmuramoyl-L-alanine amidase and an N-acetylmuramidase. J Gen Microbiol 128(6):1171–1178
Whiteley CG, Lee DJ (2015) Bacterial diguanylate cyclases: structure, function and mechanism in exopolysaccharide biofilm development. Biotechnol Adv 33(1): 124–141
Whittaker JW (2003) Free radical catalysis by galactose oxidase. Chem Rev 103: 2347–2363
Wilson DB (2009) Cellulases and biofuels. Curr Opin Biotechnol 20(3):295–299
Wolanin PW, Thomason PA, Stock JB. Histidine protein kinases: key signal transducers outside the animal kingdom. Genome Biol. 2002;3(10): reviews 3013.1–3013.8
Xia Y, Chin FYL, Chao Y, Zhang T (2015) Phylogeny-structured carbohydrate metabolism across microbiomes collected from different units in wastewater treatment process. Biotechnol Biofuels 8:172
Yoo HG, Kwon SY, Karki S, Kwon HJ (2011) A new route to dTDP-6-deoxy-L-talose and dTDP-L-rhamnose: dTDP-L-rhamnose 4-epimerase in *Burkholderia thailandensis*. Bioorg Med Chem Lett 21(13):3914–3917
Yuzwa SA, Macauley MS, Heinonen JE et al (2008) A potent mechanism-inspired O-GlcNAcase inhibitor that blocks phosphorylation of tau *in vivo*. Nat Chem Biol 4(8):483–490