Biotechnology Reports 28 (2020) e00530

Contents lists available at ScienceDirect

Biotechnology Reports

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

Complete genome sequencing and comparative genome characterization of the lignocellulosic biomass degrading bacterium *Pseudomonas stutzeri* MP4687 from cattle rumen

Maulik Patel[a,c,*], Hiral M. Patel[b], Nasim Vohra[d], Sanjay Dave[a]

[a] Department of Biotechnology, Hemchandracharya North Gujarat University, Patan, 384265, Gujarat, India
[b] P.G. Department of Biosciences, Vadtal Road, Sardar Patel University, Bakrol, 388315, Anand, Gujarat, India
[c] Laboratory of Renewable Resources Engineering and Department of Agricultural and Biological Engineering, Purdue University, West Lafayette, 47906, IN, United States
[d] Anand Agricultural University, Anand, 388110, Gujarat, India

ARTICLE INFO

Article history:
Received 2 February 2020
Received in revised form 9 September 2020
Accepted 9 September 2020

Keywords:
Pseudomonas stutzeri
Lignocellulosic Biomass
Whole genome
CAZymes

ABSTRACT

We report the complete genome sequencing of novel *Pseudomonas stutzeri* strain MP4687 isolated from cattle rumen. Various strains of *P. stutzeri* have been reported from different environmental samples including oil-contaminated sites, crop roots, air, and human clinical samples, but not from rumen samples, which is being reported here for the first time. The genome of P. stutzeri MP4687 has a single replicon, 4.75 Mb chromosome and a G + C content of 63.45%. The genome encodes for 4,790 protein coding genes including 164 CAZymes and 345 carbohydrate processing genes. The isolate MP4687 harbors LCB hydrolyzing potential through endoglucanase (4.5 U/mL), xylanase (3.1 U/mL), β-glucosidase (3.3 U/mL) and β-xylosidase (1.9 U/mL) activities. The pangenome analysis further revealed that MP4687 has a very high number of unique genes (>2100) compared to other *P. stutzeri* genomes, which might have an important role in rumen functioning.

1. Introduction

Lignocellulosic biomass (LCB) is abundantly available on the earth and its utilization as a source of carbon or in the form of energy generation would be beneficial due to its renewable nature. However, the utilization of biomass is quite challenging due to its resistance to biodegradation and its physiognomies [1]. In nature, there are few microorganisms that have effective capabilities to degrade these recalcitrant lignocellulosic biomasses into small sugars. Subsequently, these hydrolysed sugar molecules can be used by nearby microbial communities as an energy source [2]. Hence screening efficient LCB degrading microorganisms has various biotechnological applications. For instance, it could replace sugarcane or corn based glucose or fructose with agricultural biomass based sugars in saccharification industries. Moreover, the production of industrially important metabolites by LCB hydrolyzing microbes are advantageous and economically very important due to the reduction of input cost [3,4]. Therefore, scientists around the globe are working on screening novel LCB hydrolyzing microbes with industrial importance.

In search of industrially important microbes and genes various habitats like the deep sea, hot springs, frigid zones, high and low altitudes, the gut of various animals and insects have been explored [1,5,6]. Similarly, the rumen of herbivores is also one of the important habitats to explore for the LCB degrading bacteria [1]. The rumen, a part of herbivores stomach, is a complex and dynamic ecosystem comprising of a diverse group of bacteria including aerobic and anaerobic bacteria and fungi, methanogen, protozoa, and phages. These microbes are symbiotically involved in the breakdown of complex plant cell wall glycosides and provide energy to the host [7–9]. LCB digestion varies with the host to host, source of feed intake and environmental conditions [10]. The microbiome of the rumen is diverse, and in this sense is a good medium in which to probe for and recover various novel microorganisms that hydrolyze lignocellulosic biomass as well as caring out biotransformations which may have wider biotechnological application [11].

Pseudomonas stutzeri is a non-fluorescent, non-spore forming, flagellated, Gram-negative bacteria ubiquitously disseminated in the environment that is adaptable to a broad range of

* Corresponding author at: Laboratory of Renewable Resources Engineering and Department of Agricultural and Biological Engineering, Purdue University, West Lafayette, 47906, IN, United States
E-mail address: maulik8090@yahoo.co.in (M. Patel).

https://doi.org/10.1016/j.btre.2020.e00530

environments resulting in its ubiquitous presence and unusual features. These include hydrocarbon degradation, toleration of heavy metals, xenobiotic degradation, nitrification, denitrification and many more [12]. Hydrocarbon degrading *P. stutzeri* has been well documented but very few strains were reported with cellulolytic potential [13–15] and the role of *P. stutzeri* in rumen fermentation or LCB hydrolysis is not studied or documented.

The present work reports the LCB hydrolyzing potential of a novel strain of *P. stutzeri* MP4687 isolated from cattle rumen. Previous reports have assumed that high genome plasticity and natural transformation proficiency helps to support various environments in which *P. stutzeri* can live. We believe it would be interesting to explore the genes acquired by isolate MP4687 to survive in cattle rumen. Thus, the complete genome sequencing of the isolate MP4687 was carried out to understand the role of *P. stutzeri* in the rumen and on LCB hydrolysis.

2. Materials and methods

2.1. Isolation of LCB hydrolyzing bacteria from cattle rumen

A total of 500 mL of the cow rumen digesta were collected 3 h post morning feeding using a flexible stomach tube attached to the vacuum pump and further processed as described in previous study [1]. The 0.2 mL of processed rumen was taken and inoculated into the 50 ml of the flask containing 1gm of finely grinded wheat straw powder in minimal medium and incubated for 72 h at 50 °C and 180 rpm on an orbital shaker. After incubation, aliquots were taken serially diluted and plated onto the Carboxymethylcellulose (CMC) containing minimal medium plate for 24 h. The fast-growing bacterial colonies were transferred on to the new CMC containing minimal medium plates and the procedure was repeated for three times. Among the colonies, the fast and large growing bacterial colony was selected for further analysis. Furthermore, the enzyme activity was analyzed on the plate by incubating with 3 mL of 1% congo red followed by 15 min washing with 20 mL of 1 M NaCl or with Lugol's iodine. In both cases, the clear halo surrounding the colony indicates the enzymatic activity [16].

2.2. Lignocelluloses hydrolyzing enzyme activity studies

To study the lignocellulose hydrolyzing potential, the isolates were grown overnight in a nutrient medium containing 2% (w/v) grinded dry lignocellulosic biomass for 18 h at 50 °C and 180 rpm on an orbital shaker. The aliquots were taken from the grown culture and centrifuged at 8000 g for 10 minutes. The supernatant was used to perform enzyme assays. Endoglucanase and xylanase activities of the crude enzyme was measured by the dinitrosalicylic acid (DNS) method with CMC and birchwood xylan as substrate, respectively [17]. The β-glucosidase and β-xylosidase activities were performed using p-nitrophenyl-D-glucopyranoside and p-nitrophenyl-D-xylopyranoside as substrate, respectively. In all the assays, 10 μL of diluted enzyme solution was added to 90 μL of 0.1 M sodium phosphate buffer (pH 6), containing 1% (w/v) of each substrate. The enzymatic reactions were allowed to proceed at 50 °C for 15 min. The quantity of released reducing sugar and p-nitrophenol was determined by measuring the absorbance at 540 nm and 405 nm, respectively. One unit of enzymatic activity was defined as the amount of enzyme necessary to release 1 μmol of reducing sugar per minute, whereas for p-nitrophenyl-D-glucopyranoside and p-nitrophenyl-D-xylopyranoside, 1 unit of enzymatic activity was defined as the amount of the enzyme required to release 1 μmol of p-nitrophenol per minute. All the enzymatic activities were carried out in triplicate.

2.3. Biochemical and morphological characterization of the isolate

To study the potential of different kind of carbohydrates and sugar utilization potential, the isolated bacterial culture were examined for biochemical parameters mainly sugar utilization profile (API Kit, Biomerieux, Germany) [18]. Furthermore, the isolate was microscopically and physiologically characterized.

2.4. DNA extraction, genome sequencing, assembly, and annotation

Genomic DNA was extracted with the bacterial DNA isolation kit (Genetix India Limited) following the manufacturer's instructions. The isolated DNA was quantified by a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) and the DNA quality was assayed with Agilent Bioanalyzer (Agilent, Santa Clara, CA, USA). Whole genome sequencing was performed on the Illumina NextSeq500 platform [19]. The paired-end sequencing library was prepared using the Illumina NextSeq Nano DNA library preparation kit. For library preparation, approximately 200 ng of DNA was taken and fragmented using Covaris M220 to generate the mean fragment size of 350bp. Raw data produced by Illumina NextSeq500 platform were initially

Table 1
Detail of *P.stutzeri* strains with the accession number and pangenome based genes information.

Genome no.	Strain name	Organism name / Accession no.	No. of core genes	No. of accessory genes	No. of unique genes	No. of exclusively absent genes
1	A1501	ASM1378v1	2197	1583	150	12
2	DSM 4166	ASM19510v1	2197	1761	120	11
3	CGMCC 1.1803	ASM21960v1	2197	1566	114	6
4	CCUG 29243	ASM26754v1	2197	1737	188	2
5	DSM 10701	ASM27916v1	2197	1181	267	88
6	RCH2	ASM32706v1	2197	1726	108	5
7	28a24	ASM59047v1	2197	1486	350	83
8	19SMN4	ASM66191v1	2197	1815	204	5
9	SLG510A3-8	ASM103864v1	2197	1690	84	52
10	273	ASM164819v1	2197	1742	353	26
11	DW2-1	ASM300165v1	2197	1585	93	2
12	1W1-1A	ASM300849v1	2197	1659	100	5
13	SGAir0442	ASM304714v2	2197	1617	141	6
14	PheN2	ASM584400v1	2197	1266	152	21
15	XL272	ASM974018v1	2197	1190	152	4
16	PM101005	ASM978955v1	2197	1645	464	28
17	NCTC10450	44858_G01	2197	1580	90	10
18	NCTC10475	55147_C01	2197	1654	88	2
19	MP4687	PS_MP4687	2197	1762	2114	2

processed using CLC Genomics Workbench version 9., (CLC bio, Aarhus, Denmark). The adapters were removed and the reads were trimmed based on quality (limit = 0.02), then the resulting reads were assembled using the de Bruijn graph method of de novo assembly [20]. The assembly was further annotated using automated annotation with MicroScope 3.14.1 to confirm the coding genes [21]. The coding genes were further blasted against several public databases including GO (Gene Ontology) (http://geneontology.org/), KEGG (Kyoto Encyclopedia of Genes and Genomes) (https://www.genome.jp/kegg/pathway.html), COG (Clusters of Orthologous Groups) (https://www.ncbi.nlm.nih.gov/COG/), and NR (Non-Redundant Protein the Database databases) (https://www.ncbi.nlm.nih.gov/protein). A whole genome Blast search (E-value less than 1e-5, minimal alignment length percentage larger than 40%) was performed against above databases and the protein sequences were predicted. To investigate the presence of complex plant based carbohydrate hydrolyzing genes, the genome was annotated against three CAZyme databases and consistent results among two of them were considered as appropriate positive results (https://www.cazy.org/).

2.5. Comparative genomic analysis of P. stutzeri genomes

The pangenome analysis approach was used to identify strain-specific genomic features in a genome and to determine the genomic diversity among the P. stutzeri strains. In order to get this information, Bacterial Pan Genome Analysis (BPGA) tool was used [22]. More than 250 genome sequences of *P. stutzeri* are available in the NCBI database, however, only 18 are complete sequences (Table 1). The complete genome sequences and related proteins fasta files were downloaded from NCBI databases and used for BPGA analysis. BPGA further processed these files to clustering of genes into families by USEARCH with a 50% sequence similarity as a cutoff. The pangenome profile analysis was performed with the 18 complete genomes of *P.stutzeri* from the database and isolate MP4687 genome. Using the pangenome sequence extraction module, core, accessory, and unique gene families were extracted. The unique genes (singleton) specific to each genome were identified and extracted for further analysis. The pangenome functional analysis module was used to find the COGs and KEGG pathway distribution. Evolutionary analysis done by BPGA, based on concatenated core gene alignment using a binary pan-matrix was used for the generation of a phylogenetic tree. Furthermore the genome was investigated for presence of genomic islands (GEI), transposons or other insertional sequences using MicroScope 3.14.1 platform [21].

2.6. Gene Bank Accession Number

The sequence data of the whole genome of isolate MP4687 has been deposited to the NCBI database with the accession number **WWNS00000000** and BioProject number PRJNA598454.

3. Results and Discussion

3.1. Isolation of LCB hydrolyzing bacteria and enzyme analysis

In total 19 different bacterial strains growing on LCB containing minimal medium plates were selected based on morphological and physiological characteristics. These strains were further cultured on to CMC plates a couple of times and the fastest growing strain along with larger halo of endoglucanase activity was selected for further study. One strain, MP4687, outperformed among all in terms of growth and size of the colony and cellulase enzyme activity, therefore it was chosen to study further its LCB hydrolyzing potential. To study the LCB hydrolyzing potentials of MP4687, aliquots were taken from 2% LCB containing grown medium and endoglucanase, xylanase, β-glucosidase, and β-xylosidase activities were performed at 50 °C and pH 6 to mimic optimum rumen conditions [23]. The enzyme analysis showed 4.57 ± 0.34 U/mL, 3.1 ± 0.29 U/mL, 3.3 ± 0.2 U/mL and 1.9 ± 0.26 U/mL of endoglucanase, xylanase, β-glucosidase, and β-xylosidase activities respectively, with the crude enzyme. The endoglucanase and xylanase activities convert the complex LCB biomass into cello-oligosaccharides and xylooligosaccharides. While these oligosaccharides are further converted into glucose and xylose by β-glucosidase and β-xylosidase enzymes respectively. Thus,all four types of activities necessary for the complete decomposition of LCB into sugars are present [24]. The results indicated that the isolate MP4687 secrets all necessary enzymes to hydrolyze LCB into sugars. Various strains of bacteria belong to Bacillus sp., Pseudomonas sp., Streptomyces sp. and Aeromonas sp. have been reported with similar activities [4,25,26]. While, similar activities was reported into the various fungal strains isolated from agriculture waste or forest. The crude extract of white rot fungi Pleurotus ostreatus, Pleurotus eryngii, Trametes versicolor,

Table 2
Sugar utilization profile of *P. stutzeri* MP4687.

Biochemical test								
No	Sugar	+/-	No	Sugar	+/-	No	Sugar	+/-
1	Glycerol	-	17	Inositol	-	33	Inulin	-
2	Erythritol	-	18	D-Mannitol	+	34	D-Melezitose	+
3	D-Arabinose	-	19	D-Sorbitol	+	35	D-reffinose	-
4	L-Arabinose	-	20	Methyl α-D-Mannopyranoside	-	36	Amidon(starch)	-
5	D-Ribose	+	21	Methyl α-D-Glucopyranoside	-	37	Glycogen	-
6	D-Xylose	+	22	N-Acetyl Glucosamine	+	38	Xylitol	-
7	L-Xylose	+	23	Amygdalin	-	39	Gentiobiose	+
8	D-Adonitol	-	24	Arbutin	+	40	D-Turanose	+
9	Methylβ-D-Xylopyranoside	+	25	Esculin Ferric citrate	+	41	D-Lyxose	-
10	D-Galactose	+	26	Salicin	+	42	D-Tagatose	+
11	D-Glucose	+	27	D-Celiobiose	+	43	D-Fucose	-
12	D-Fructose	+	28	D-Maltose	+	44	L-Fucose	-
13	D-Mannose	+	29	D-Lactose (bovine origin)	+	45	D-Arbitol	-
14	L-Sorbose	+	30	D-Melibiose	+	46	L-Arbitol	-
15	L-Rhomnose	+	31	D-Saccharose (Sucrose)	+	47	Potassium Gluconate	-
16	Dulcitol	-	32	D-Trehalose	+	48	Potassium2-KetoGluconate	-
						49	Potassium 5-keto Gluconate	-

Pycnosporus sanguineus and Phanerochaete chrysosporium were showed enzyme activites in similar range or less on pretreated pulps of sorghum straw [27]. The rumen is a natural fermenter having a role in plant polysaccharide processing. Hence the bacteria of rumen origin should have the plant polysaccharides processing potential [28]. The isolate MP4687 has shown excellent response to that in the LCB hydrolyzing study. A similar result of various plant polysaccharides hydrolyzing potential was reported in *Streptococcus gallolyticus* isolated from the bovine rumen [29].

3.2. Biochemical and morphological characterization of MP4687

The isolate MP4687 was microscopically examined and revealed as Gram-negative rod-shaped bacteria. The bacterial colonies were pale in color, wrinkled, dry and coherent in nature and some resemble coral structure in appearance. The obtained characteristics resemble characteristics of *Pseudomonas* sp. reported elsewhere [12]. The commercial kit API 20NE based identification confirmed the MP4687 belongs to *P. stutzeri*. Moreover, due to diverse characteristics reported in *P. stutzeri* species and to understand the utilization of sugar profile in the rumen, we performed various sugar biochemical analyses. The isolate MP4687 was positive for gluconate, D-glucose, D-maltose, starch, acetate, butyrate, isobutyrate, isovalerate, propionate, fumarate, glutarate, glycolate, glyoxylate, hydroxybutyrate, itaconate, oxaloacetate, 2-oxoglutarate, pyruvate, succinate, DL-lactate, DL-malate, L-malate, malonate, D-alanine, L-alanine, L-asparagine, L-glutamate, L- glutamine, L-isoleucine, L-leucine, L-proline and hydrolysis of L-alanine. The obtained results were in accordance with the previously reported results for *P. stutzeri* [30]. While they had also reported that *P. stutzeri* was unable to utilize mannose, xylose, cellobiose, lactose, fucose, L-rhamnose, D-ribose, D-sucrose, L-sorbose, D-tagatose, D-turanose and salicin, the isolate MP4687 is able to utilize those in our studies (Table 2). The utilization of such sugars advocates its functionality in the rumen, a rich pool of biomass derived sugars and supports its adaptation behavior to rumen conditions. The adaptation to the rumen environment might be due to the natural transformation capabilities of *P. stutzeri* [31,32]. Natural transformation is the process of naked DNA uptake by bacteria from the surrounding environment and integrates it into its genome. Furthermore, it has been studied that *P. stutzeri* can be transformed with DNA absorbed in the soil, DNA associated with minerals or DNA bound to marine sediments, even in presence of DNases and native microorganisms [12,33,34]. Therefore, it is envisaged that in a rumen, isolate MP4687 has taken up some genes responsible for LCB hydrolyzing and cello-oligosaccharides utilization with time.

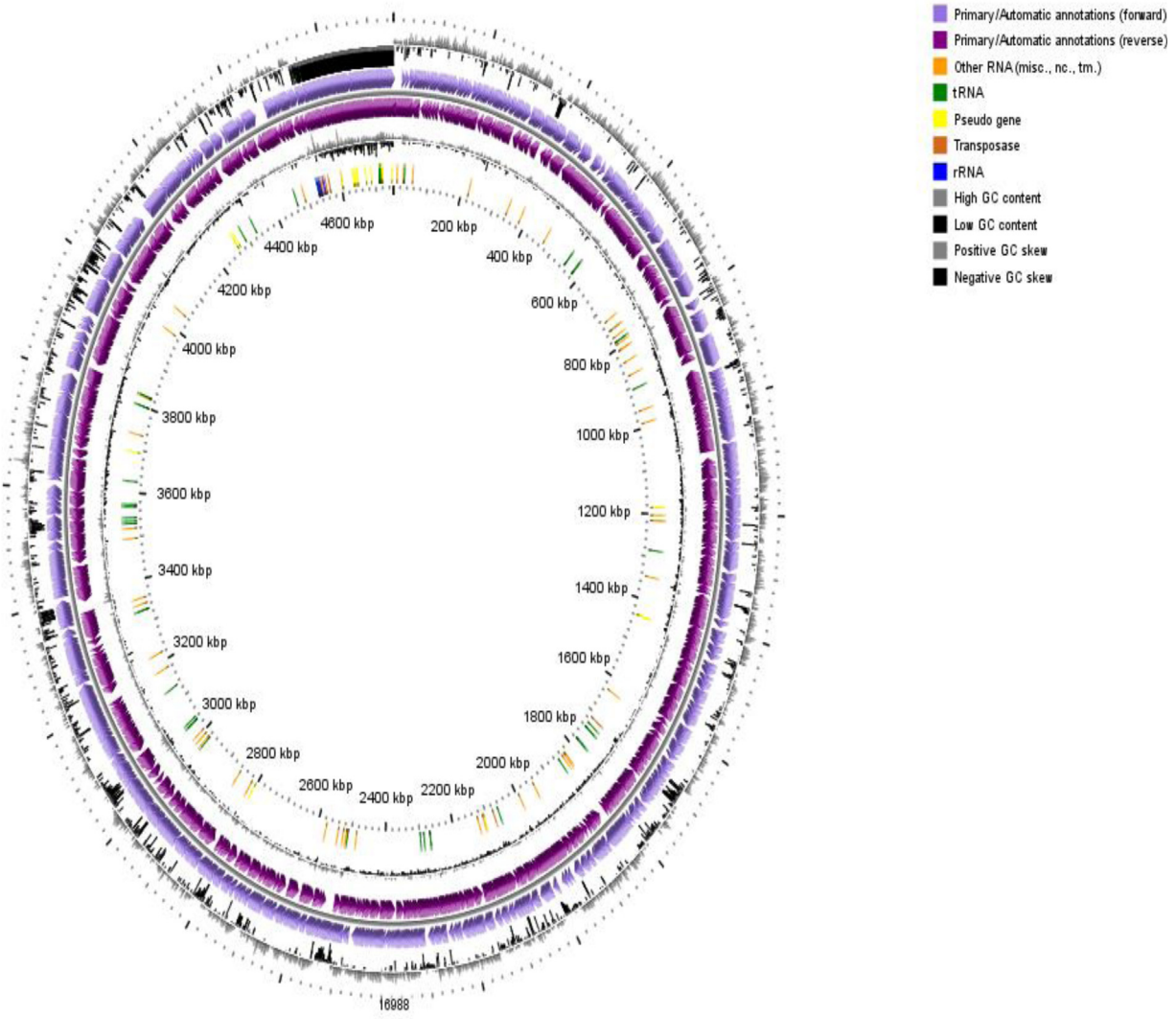

Fig. 1. Circular genome map of *P. stutzeri* MP4687.

Table 3
Genome characteristics of the *P. stutzeri* MP4687.

Issue	Number
Genome Size (bp):	4752154
Number of contigs	292
Largest contig size	852816
GC content (%):	63.45
Gene number:	4796
Gene average length (bp):	898.52
Average intergenic length	121.57
Pseudogenes	6
tRNA gene	60
rRNA gene	5
Genomic regions in tandem duplication	60
Redundancy (%)	2.99
Protein coding density (%)	90.14

3.3. Genome characteristics and functional annotation of P. stutzeri-*MP4687*

The whole genome sequencing of isolate MP4687 was carried out to get insights of its role in the rumen. The sequencing data revealed that MP4687 has a single chromosome of 4.75Mb with 63.45% GC content and does not contain any extrachromosomal DNA (Fig. 1). The genome of MP4687 was predicted with 4913 genomic objects, which includes 4790 coding genes (CDS), 6 pseudogenes, 60 tRNA, 5 rRNA (2-16 s rRNA, 2-23 s rRNA, and 1-5 s rRNA), 51-miscRNA and 1-tmRNA (Table 3). The genome has 29100 undetermined bases, 2.99 repeated regions and 90.14% protein coding density. Overall, the putative genes have 3153 distinct functions. The genes include 1991 genes with a SEED annotation ontology across 1487 distinct SEED functions. The cluster of orthologous groups (COG) database annotation revealed that 79.48 % of the CDSs are classified in at least one COG group (3812 CDSs / 4796) while 87.95 % of the CDSs are classified in at least one EGGNOG group (4218 CDSs / 4796). Furthermore, KEGG annotation revealed that a total of 345 genes involved in carbohydrate metabolism and 128 genes for xenobiotic degradation and metabolism are found in the MP4687 genome. In addition to that, 19 genes involved in nitrogen metabolism, including all major genes for nitrification and denitrification, has been predicted in the KEGG pathway. Additionally, 30 genes of methane metabolism pathways and 13 genes for sulfur metabolism pathway have been predicted. The whole genome-based phylogenetic analysis further revealed that isolate MP4687 was closely related to *Pseudomonas stutzeri* and share more than 99% 16S rDNA based similarity (Fig. 2).

The CAZymes are very important for rumen functioning and bacterial survival in rumen conditions. The CAZymes annotation pipeline annotated a total of 164 carbohydrate processing enzymes in isolate MP4687, with 91, 115 and 112 CAZymes identified by HMMER, Hotpot and DIAMOND respectively. However, among them, 60 CAZymes were commonly predicted by all three databases, while 7,8 and 7 CAZymes were predicted by DIAMOND and HMMER, HMMER and Hotpep, and Hotpep and DIAMOND respectively. There were 12, 32 and 38 CAzymes predicted individually by HMMER, Hotpep, and DIAMOND databases respectively. The ven diagram of predicted CAZymes is given in Fig. 3. The detailed analysis of predicted Glucoside hydrolases (GHase) reveled the presence of total 40 GHase in the genome of MP4687. Among them, 13 GHase belonged to GH13 family followed by 5 GHase from GH23 family, 3 GHase from GH1 family, 2 GHase from GH3, GH5 and GH103 families each and single GHase from GH2, GH19, GH29, GH30, GH35, GH36, GH37, GH39, GH59, GH73, GH77, and GH136 families were predicted (Fig. 3b). The presence of such a variety numbers of GHase belongs to various GHase families that were not reported earlier in the *P. stutzeri* genome [35]. The enzyme products of GH1, and GH30 families having specific role in the processing of xylan to xylose which are plant based oligosachharides, hence their presence into the genome of MP4687 presume the role of this bacteria into biomass hydrolysing and the rumen functioning of cattle [36,37]. It is assumed that many of the genes associated with these specific

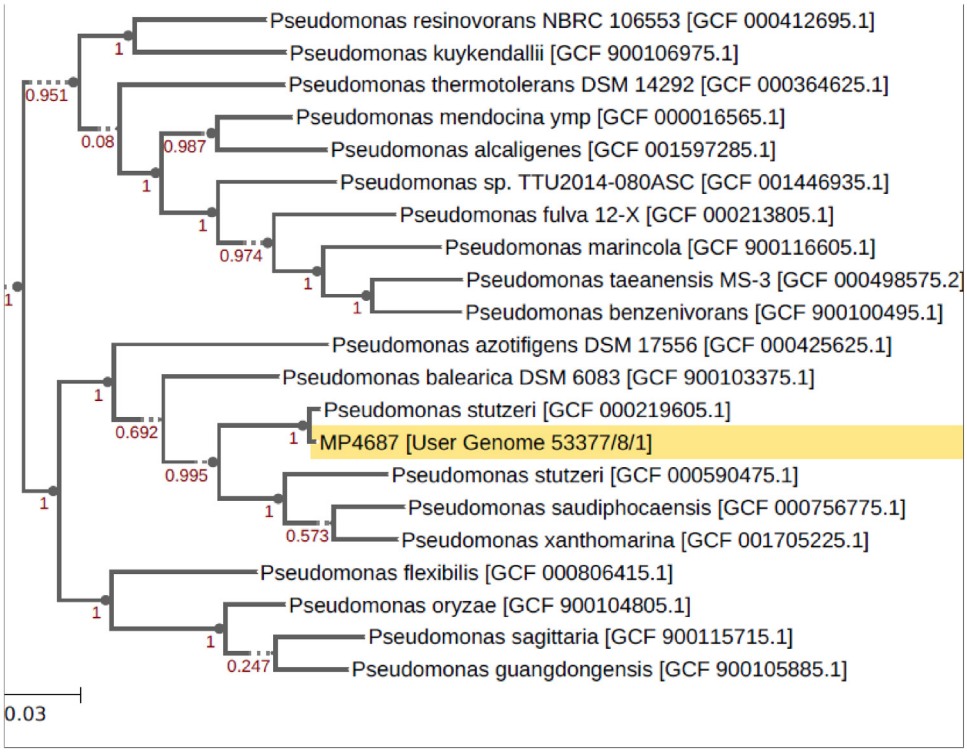

Fig. 2. Phylogenetic tree of MP4687 and its close relative strains.

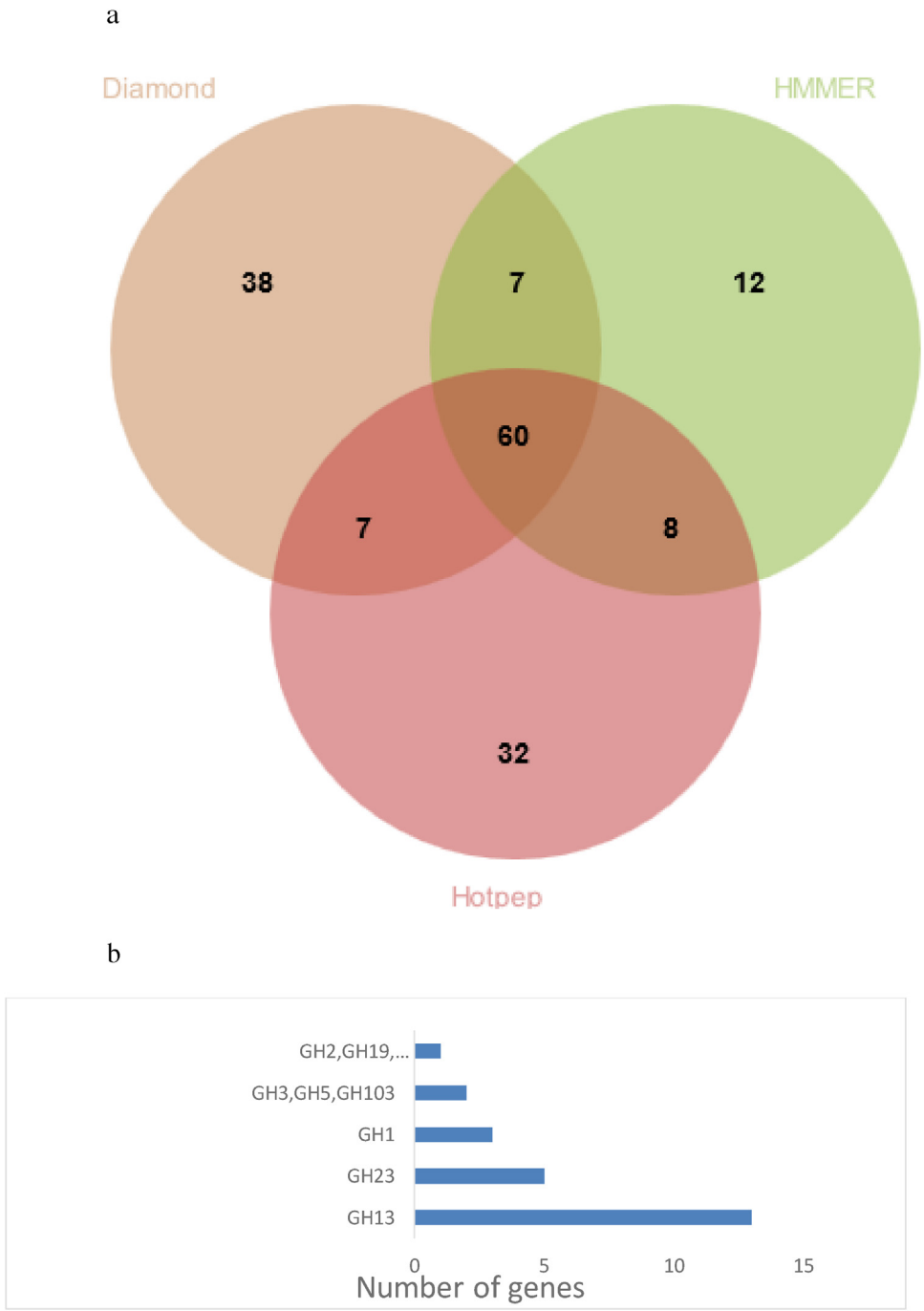

Fig. 3. a. Ven diagram of predicted CAZymes of isolate MP4687. B. Number of CAZymes families predicted in isolate MP4687 genome.

functions were transferred from other ruminal bacterial species to MP4687, to confer additional power to survive in the rumen environment [29].

In addition to the cellulose hydrolysing potential, the presence of lignin degradation capbilities would boost the LCB hydrolysing of biomass. Various oxidoreductase enzymes associated with xenobiotics and recalcitrant compound degradation were reported in the *Pseudomonas* species [12,38]. Recently, multicoper oxidase associated with lignin depolymrization were identified [39]. To that, lacase activity of CopA from *P. stutzeri* was determined recently [40]. CopA is a TAT-secreted multicopper oxidase (MCO) that oxidise various aromatic compounds including lignin in the presence or absence of Cu(II). They also studied lignin transformation activity, where 2,6-dimethoxyphenol (DMP) was oxidised into 6-acetyl-2,5-dihydroxy-1,4-naphthoquinone, 2,6-dimethoxybenzene-1,4-diol, 4-hydroxy-3-methoxybenzoic acid, and 4-hydroxy-3,5-dimethoxybenzoic acid. Thus it is assumed that the product of copA gene has major role in the lignin transformation. In context to that we evaluated the presence of such genes in the MP4687 genome. Interestingly, two copy of copA gene and various oxigenase enzymes involved in a aromatic hydrocarbon degradation were found. The presence of this genes in the genome of

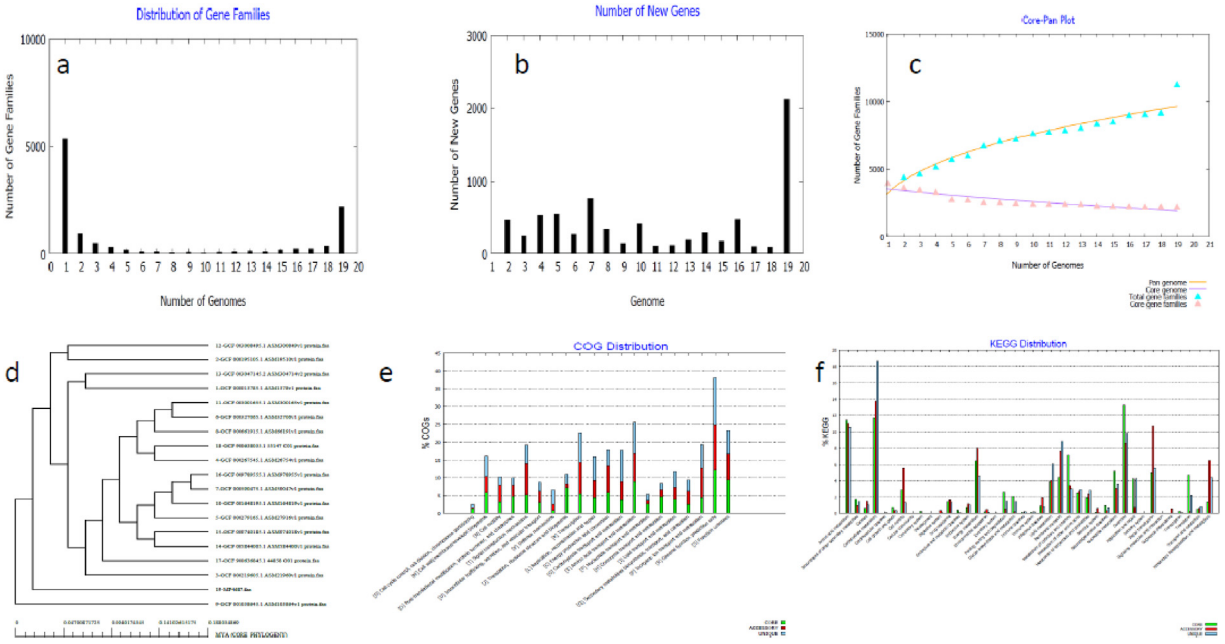

Fig. 4. Pangenome analysis of *P. stutzeri* genomes.
(**a**) The gene family frequency spectrum. (**b**) New gene family distribution after the sequential addition of each genome to the analysis. (**c**) The pangenome profile trends obtained using clustering tools- USEARCH. (**d**) Phylogenetic tree based on core genes of genomes. (**e**) COG distribution of core, accessory, and unique genes. (**f**) KEGG distribution of core, accessory, and unique genes.

MP4687 confers its lignin depolymerisation potential during the LCB hydrolysis.

3.4. *Comparative genomic analysis of* P. stutzeri *genomes*

The genome of isolate MP4687 and 18 complete whole genomes of *P. stutzeri* available in the NCBI database were used for bacterial pangenome analysis (BGAP). The analysis was carried out based on the all translated protein sequence comparison. The pangenome of *P. stutzeri* based on power regression function depicted open type, suggesting that continuous new genes are being added to the pangenome and hence, with the addition of new genome, leads to an increase in accessory and unique genes number and decrease in core gene numbers (Fig. 4). The detailed numbers of core genes, accessory genes, number of absence genes and singleton specific to each genome are given in Table 1. The distribution of gene families and new genes within the pangenome specific to the genome is given in Fig. 4a and b. The number of unique genes/singleton presence in each genome advocates the high transformation potential of the genomes and their ubiquitous presence [12]. The higher number of unique genes that have been observed into the isolate MP4687 indicates its special features among the other *P. stutzeri* species. The phylogenetic tree developed based on pangenome is given in Fig. 4d. The phylogenetic tree determined isolate MP4687 is closely related to the GCF 001038645.1 strain of *P. stutzeri* Genome available in the database. That GCF 001038645.1 strain was isolated from the oil-contaminated site in Shengli Oilfield, China and reported to contain a complete set of benzoate degradation genes [38]. The protein sequences associated with core, accessory and unique genes were identified and annotated with COG and KEGG databases (Fig. 4e & f). The obtained data indicated a higher number of genes predicted by COG is involved in general functions which is followed by genes involved in metabolic functions. In addition the higher number of unique genes predicted in the studies are also belongs metabolism and the least numbers are in cellular processing and signaling. To that, KEGG prediction has further revealed that more number of unique genes are of carbohydrates metabolism activities. It is postulated that a higher number of carbohydrate metabolism genes might be due to its capability to metabolize various complex carbon sources [41]. Similarly, a higher number of carbohydrate metabolism genes in MP4687 support its function in the rumen to process various plant-based polysaccharides into energy.

In the previous reports, high genomic plasticity associated with the *P. stutzeri* was reported [12], hence, we studied the comparative genomic plasticity in MP4687 strain. The study revealed 79 - 81% genes were in synteny with closest *P. stutzeri* strains in the database and determined 20 Region of Genomic Plasticity (RPG) ranging of (5-20 kb) in size in genome of isolate MP4687 (Fig. 5). Among them 6 RPGs were region specific, 7 RPGs were tRNA-miscRNA, SIGI (Score-Based Identification of Genomic Islands)-IVOM (Interpolated Variable Order Motifs) specific, 1 RPG was mobile genetic element – misc RNA specific and remaining were either SIGI or IVOM or both specific determined [42,43]. All RPGs togather, represents nearly 20% plasticity in the MP4687 genome. To that, RPG7 and RPG13 codes for putative dienelactone hydrolase and amidohydrolase2 respectively. The both enzymes have potential ligninase / lacase activity. In addition two CALIN (clusters of attC sites lacking integron-integrases) type integron elements were found in the genome of MP4687 and codes for potential beta-D-xylosidase 2 (BXL2) enzyme which poses xylan 1,4-beta-xylosidase activity. Furthermore, the detailed examination of uniqe genes associated with isolate MP4687 revealed that it codes two putative alpha / beta hydrolases potentially have role in xylanase and lacase activities individually or togather and a putative epoxide hydrolase enzyme having role in lignin degradation. Thus the presence of lignin and cellooligosaccharide hydrolyzing genes in the MP4687 confers its LCB hydrolysing potential in the rumen.

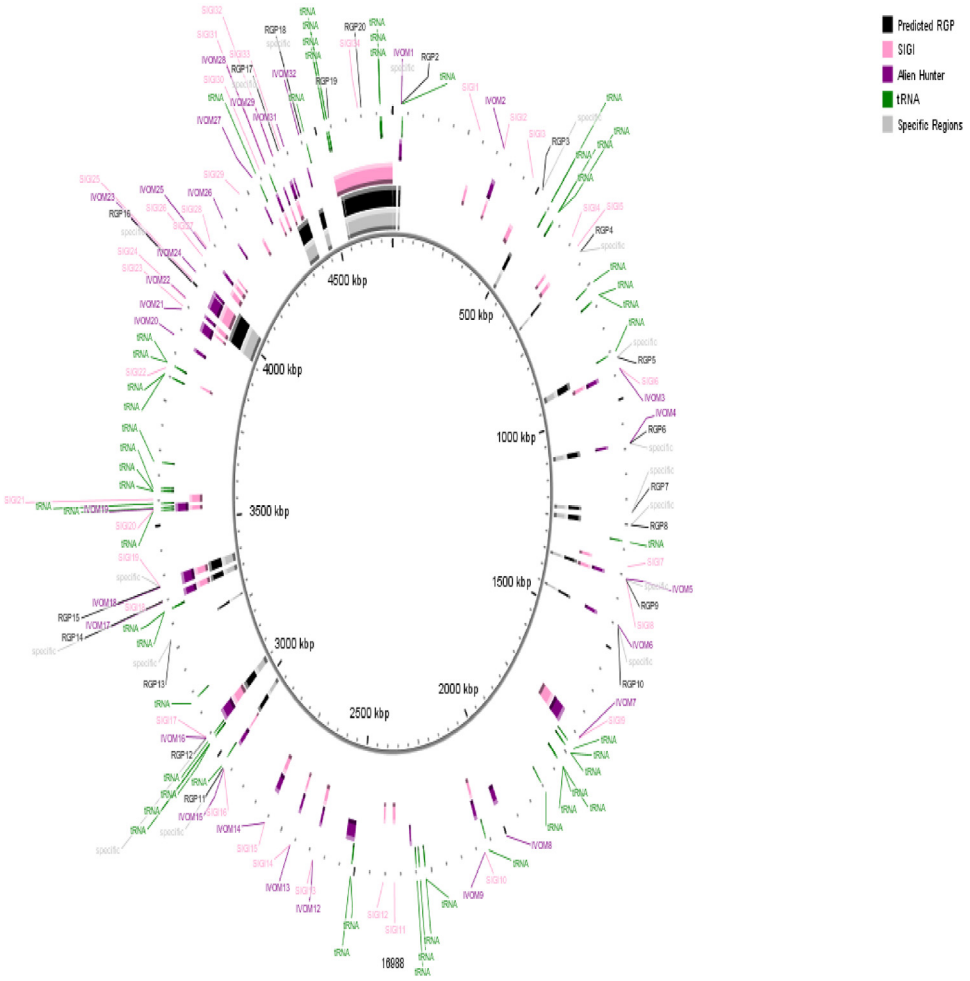

Fig. 5. Map depicting genomic plasticity (RPG) in the genome of MP4687.

4. Conclusion

In this article, we report the complete genome sequencing of *P. stutzeri* MP4687 isolated from cattle rumen. Being a rumen isolate, the lignocellulosic biomass hydrolyzing potential was examined. The isolate MP4687 showed decent biomass hydrolyzing potential with endoglucanase, xylanase, β-glucosidase, and β-xylosidase activities in compared to other reported microbes. The detailed genome analysis has further confirmed the presence of a higher number of carbohydrate processing genes in the genome. In addition, the CAZymes analysis showed the presence of additional GHase in the genome of *P. stutzeri* MP4687 compared to other *P. stutzeri* genomes. Hence, the detailed analysis of LCB hydrolyzing potential can be studied and optimized to explore its prospective in the biofuel industry and biorefineries.

Declaration of Competing Interest

- All authors declare that the presented work in the manuscript is original done by them.
- All authors mutually agree to submit the manuscript to "Journal of Biotechnology Reports".
- Authors declare that the manuscript has not been previously submitted to "Journal of Biotechnology Reports" or elsewhere for publication.
- Authors declare that there is no conflict of interest.

CRediT authorship contribution statement

Maulik Patel: Conceptualization, Data curation, Formal analysis, Methodology, Investigation, Validation, Visualization, Writing - original draft. **Hiral M. Patel:** Formal analysis, Methodology, Writing - review & editing. **Nasim Vohra:** Methodology. **Sanjay Dave:** Supervision, Writing - review & editing.

Acknowledgment

The authors wish to thank Dr. Michael Ladisch and Eduardo Ximenes for their internal review of this article.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.btre.2020.e00530.

References

[1] M. Patel, H.M. Patel, S. Dave, Determination of bioethanol production potential from lignocellulosic biomass using novel Cel-5m isolated from cow rumen metagenome, Int. J. Biol. Macromol. (2019).
[2] A. Pennacchio, V. Ventorino, D. Cimini, O. Pepe, C. Schiraldi, M. Inverso, V. Faraco, Isolation of new cellulase and xylanase producing strains and

application to lignocellulosic biomasses hydrolysis and succinic acid production, Bioresour. Technol. 259 (2018) 325–333.
[3] M.C.S. de Barcelos, C.H. Carvalho e Silva, C.L. Ramos, G. Molina, Microorganisms for Cellulase Production: Availability, Diversity, and Efficiency, in: M. Srivastava, N. Srivastava, P.W. Ramteke, P.K. Mishra (Eds.), Approaches to Enhance Industrial Production of Fungal Cellulases, Springer International Publishing, Cham, 2019, pp. 53–69.
[4] Z. Qi, Y. Zhu, H. Guo, Y. Chen, Y. Zhao, Y. Zhou, X. Wang, Y. Yang, W. Qin, Q. Shao, Production of glycoprotein bioflocculant from untreated rice straw by a CAZyme-rich bacterium, Pseudomonas sp. HP2, J. Biotechnol. 306 (2019) 185–192.
[5] D.E. Marco, F. Abram, Editorial: Using Genomics, Metagenomics and Other "Omics" to Assess Valuable Microbial Ecosystem Services and Novel Biotechnological Applications, Front. Microbiol. 10 (2019).
[6] A.R. Uria, D.S. Zilda, Chapter One - Metagenomics-Guided Mining of Commercially Useful Biocatalysts from Marine Microorganisms, in: S.-K. Kim, F. Toldrá (Eds.), Adv. Food Nutr. Res., Academic Press, 2016, pp. 1–26.
[7] R.E. Hungate, The rumen and its microbes, Elsevier, 2013.
[8] S.A. Huws, C.J. Creevey, L.B. Oyama, I. Mizrahi, S.E. Denman, M. Popova, R. Muñoz-Tamayo, E. Forano, S.M. Waters, M. Hess, I. Tapio, H. Smidt, S.J. Krizsan, D.R. Yáñez-Ruiz, A. Belanche, L. Guan, R.J. Gruninger, T.A. McAllister, C.J. Newbold, R. Roehe, R.J. Dewhurst, T.J. Snelling, M. Watson, G. Suen, E.H. Hart, A. H. Kingston-Smith, N.D. Scollan, R.M. do Prado, E.J. Pilau, H.C. Mantovani, G.T. Attwood, J.E. Edwards, N.R. McEwan, S. Morrisson, O.L. Mayorga, C. Elliott, D.P. Morgavi, Addressing Global Ruminant Agricultural Challenges Through Understanding the Rumen Microbiome: Past, Present, and Future, Front. Microbiol. 9 (2018).
[9] I. Mizrahi, The role of the rumen microbiota in determining the feed efficiency of dairy cows. Beneficial microorganisms in multicellular life forms, Springer, 2012, pp. 203–210.
[10] A. Belanche, M. Doreau, J.E. Edwards, J.M. Moorby, E. Pinloche, C.J. Newbold, Shifts in the rumen microbiota due to the type of carbohydrate and level of protein ingested by dairy cattle are associated with changes in rumen fermentation, The Journal of nutrition 142 (2012) 1684–1692.
[11] A.H. Kingston-Smith, T.E. Davies, P. Rees Stevens, L.A.J. Mur, Comparative Metabolite Fingerprinting of the Rumen System during Colonisation of Three Forage Grass (Lolium perenne L.) Varieties, PLoS One 8 (2013)e82801.
[12] J. Lalucat, A. Bennasar, R. Bosch, E. García-Valdés, N.J. Palleroni, Biology of Pseudomonas stutzeri, Microbiol. Mol. Biol. Rev. 70 (2006) 510–547.
[13] A. Augustine, I. Joseph, Four novel strains of cellulolytic symbiotic bacteria isolated and characterized from GI tract of marine fishes of various feeding habits, Biocatalysis and Agricultural Biotechnology 16 (2018) 706–714.
[14] R. Dutoit, M. Delsaute, L. Collet, C. Vander Wauven, D. Van Elder, R. Berlemont, A. Richel, M. Galleni, C. Bauvois, Crystal structure determination of Pseudomonas stutzeri A1501 endoglucanase Cel5A: the search for a molecular basis for glycosynthesis in GH5_5 enzymes, Acta Crystallographica Section D: Structural Biology 75 (2019) 605–615.
[15] S. Rabodonirina, R. Rasolomampianina, F. Krier, D. Drider, D. Merhaby, S. Net, B. Ouddane, Degradation of fluorene and phenanthrene in PAHs-contaminated soil using Pseudomonas and Bacillus strains isolated from oil spill sites, J. Environ. Manage. 232 (2019) 1–7.
[16] J.H. Carder, Detection and quantitation of cellulase by Congo red staining of substrates in a cup-plate diffusion assay, Anal. Biochem. 153 (1986) 75–79.
[17] G.L.J.A.c. Miller, Use of dinitrosalicylic acid reagent for determination of reducing sugar, Anal. Chem. 31 (1959) 426–428.
[18] S. Hati, M. Patel, B.K. Mishra, S. Das, Short-chain fatty acid and vitamin production potentials of Lactobacillus isolated from fermented foods of Khasi Tribes, Meghalaya, India, Ann. Microbiol. 69 (2019) 1191–1199.
[19] V. Bouchez, S.L. Baines, S. Guillot, S. Brisse, Complete Genome Sequences of Bordetella pertussis Clinical Isolate FR5810 and Reference Strain Tohama from Combined Oxford Nanopore and Illumina Sequencing, Microbiology resource announcements 7 (2018) e01207–01218.
[20] A.C. Nicholson, C.A. Gulvik, A.M. Whitney, B.W. Humrighouse, J. Graziano, B. Emery, M. Bell, V. Loparev, P. Juieng, J. Gartin, C. Bizet, D. Clermont, A. Criscuolo, S. Brisse, J.R. McQuiston, Revisiting the taxonomy of the genus Elizabethkingia using whole genome sequencing, optical mapping, and MALDI-TOF, along with proposal of three novel Elizabethkingia species: Elizabethkingia bruuniana sp. nov., Elizabethkingia ursingii sp. nov., and Elizabethkingia occulta sp. nov, Antonie Van Leeuwenhoek 111 (2018) 55–72.
[21] D. Vallenet, A. Calteau, M. Dubois, P. Amours, A. Bazin, M. Beuvin, L. Burlot, X. Bussell, S. Fouteau, G. Gautreau, A. Lajus, J. Langlois, R. Planel, D. Roche, J. Rollin, Z. Rouy, V. Sabatet, C. Médigue, MicroScope: an integrated platform for the annotation and exploration of microbial gene functions through genomic, pangenomic and metabolic comparative analysis, Nucleic Acids Res. 48 (2019) D579–D589.
[22] N.M. Chaudhari, V.K. Gupta, C. Dutta, BPGA- an ultra-fast pangenome analysis pipeline, Sci. Rep. 6 (2016) 24373.
[23] F.N. Owens, M. Basalan, Ruminal fermentation. Rumenology, Springer, 2016, pp. 63–102.
[24] S.J. Horn, G. Vaaje-Kolstad, B. Westereng, V. Eijsink, Novel enzymes for the degradation of cellulose, Biotechnology for Biofuels 5 (2012) 45.
[25] L. Bandounas, N.J. Wierckx, J.H. de Winde, H.J. Ruijssenaars, Isolation and characterization of novel bacterial strains exhibiting ligninolytic potential, BMC Biotechnol. 11 (2011) 94.
[26] Y.-C. Chang, D. Choi, K. Takamizawa, S. Kikuchi, Isolation of Bacillus sp. strains capable of decomposing alkali lignin and their application in combination with lactic acid bacteria for enhancing cellulase performance, Bioresour. Technol. 152 (2014) 429–436.
[27] W.S. Cardoso, P.V. Queiroz, G.P. Tavares, F.A. Santos, F. Soares, M. Kasuya, J.H. Queiroz, Multi-enzyme complex of white rot fungi in saccharification of lignocellulosic material, Braz J Microbiol. 49 (2018) 879–884.
[28] B.A. White, R. Lamed, E.A. Bayer, H.J. Flint, Biomass Utilization by Gut Microbiomes, Annu. Rev. Microbiol. 68 (2014) 279–296.
[29] C. Rusniok, E. Couvé, V. Da Cunha, R. El Gana, N. Zidane, C. Bouchier, C. Poyart, R. Leclercq, P. Trieu-Cuot, P. Glaser, Genome Sequence of *Streptococcus gallolyticus*: Insights into Its Adaptation to the Bovine Rumen and Its Ability To Cause Endocarditis, J. Bacteriol. 192 (2010) 2266–2276.
[30] R. Rosselló-Mora, J. Lalucat, W. Dott, P. Kämpfer, Biochemical and chemotaxonomic characterization of Pseudomonas stutzeri genomovars, J. Appl. Bacteriol. 76 (1994) 226–233.
[31] C.A. Carlson, L.S. Pierson, J. Rosen, J.L. Ingraham, Pseudomonas stutzeri and related species undergo natural transformation, J. Bacteriol. 153 (1983) 93–99.
[32] J. Sikorski, N. Teschner, W. Wackernagel, Highly different levels of natural transformation are associated with genomic subgroups within a local population of Pseudomonas stutzeri from soil, Appl. Environ. Microbiol. 68 (2002) 865–873.
[33] E. Paget, P. Simonet, Development of engineered genomic DNA to monitor the natural transformation of Pseudomonas stutzeri in soil-like microcosms, Can. J. Microbiol. 43 (1997) 78–84.
[34] J. Sikorski, S. Graupner, M.G. Lorenz, W. Wackernagel, Natural genetic transformation of Pseudomonas stutzeri in a non-sterile soil, Microbiology 144 (1998) 569–576.
[35] V. Lombard, H. Golaconda Ramulu, E. Drula, P.M. Coutinho, B. Henrissat, The carbohydrate-active enzymes database (CAZy) in 2013, Nucleic Acids Res. 42 (2013) D490–D495.
[36] S. Comtet-Marre, N. Parisot, P. Lepercq, F. Chaucheyras-Durand, P. Mosoni, E. Peyretaillade, A.R. Bayat, K.J. Shingfield, P. Peyret, E. Forano, Metatranscriptomics reveals the active bacterial and eukaryotic fibrolytic communities in the rumen of dairy cow fed a mixed diet, Front. Microbiol. 8 (2017) 67.
[37] J. Šimůnek, J. Killer, H. Sechovcová, R. Pechar, V. Rada, P. Švec, I. Sedláček, Characterization of a xylanolytic bacterial strain C10 isolated from the rumen of a red deer (Cervus elaphus) closely related of the recently described species Actinomyces succiniciruminis, A. glycerinitolerans, and A. ruminicola, Folia Microbiol. (Praha) 63 (2018) 391–399.
[38] B. Hu, Y. Nie, S. Geng, X.-L. Wu, Complete genome sequence of the petroleum-emulsifying bacterium Pseudomonas stutzeri SLG510A3-8, J. Biotechnol. 211 (2015) 1–2.
[39] G. de Gonzalo, D.I. Colpa, M.H. Habib, M.W. Fraaije, Bacterial enzymes involved in lignin degradation, Journal of Biotechnology 236 (2016) 110–119.
[40] C.R. Strachan, R. Singh, D. VanInsberghe, K. Ievdokymenko, K. Budwill, W.W. Mohn, L.D. Eltis, S.J. Hallam, Metagenomic scaffolds enable combinatorial lignin transformation, Proceedings of the National Academy of Sciences 111 (2014) 10143–10148.
[41] A. Fosses, M. Maté, N. Franche, N. Liu, Y. Denis, R. Borne, P. de Philip, H.-P. Fierobe, S. Perret, A seven-gene cluster in Ruminiclostridium cellulolyticum is essential for signalization, uptake and catabolism of the degradation products of cellulose hydrolysis, Biotechnology for Biofuels 10 (2017) 250.
[42] G.S. Vernikos, J. Parkhill, Interpolated variable order motifs for identification of horizontally acquired DNA: revisiting the Salmonella pathogenicity islands, Bioinformatics 22 (2006) 2196–2203.
[43] S. Waack, O. Keller, R. Asper, T. Brodag, C. Damm, W.F. Fricke, K. Surovcik, P. Meinicke, R. Merkl, Score-based prediction of genomic islands in prokaryotic genomes using hidden Markov models, BMC bioinformatics 7 (2006) 142.