Marine Genomics 18 (2014) 83–85

Contents lists available at ScienceDirect

Marine Genomics

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

Genomics/Technical resources

Draft genome sequence of *Microbulbifer elongatus* strain HZ11, a brown seaweed-degrading bacterium with potential ability to produce bioethanol from alginate

Cong Sun [a], Yu-jie Chen [a], Xin-qi Zhang [b], Jie Pan [a], Hong Cheng [a], Min Wu [a,*]

[a] College of Life Sciences, Zhejiang University, Hangzhou, China
[b] School of Foresty and Biotechnology, Zhejiang Agriculture and Forestry University, Linan, China

ARTICLE INFO

Article history:
Received 7 May 2014
Received in revised form 23 May 2014
Accepted 24 May 2014
Available online 4 June 2014

Keywords:
Brown seaweed
Microbulbifer elongatus
Bioethanol
Alginate
Entner–Doudoroff pathway

ABSTRACT

Microbulbifer elongatus strain HZ11, was a new strain of *M. elongates* DSM 6810^T, which has the ability to degrade brown seaweeds such as *Laminaria japonica* into single cell detritus particles. Here we report a high quality draft genome of *M. elongatus* strain HZ11, which comprises 4,223,108 bp in 9 contigs with the G + C content of 56.70%. A total of 3293 protein-coding sequences were predicted, including nine alginate lyases (EC 4.2.2.3), five agarases (EC 3.2.1.81), 2-dehydro-3-deoxygluconate kinase (EC 2.7.1.45) and all enzymes involved in the Entner–Doudoroff pathway. Our results suggest that strain HZ11 has the potential ability to produce bioethanol from alginate with moderate genetic modification, which may significantly increase the yield of bioethanol from brown seaweed and the utilization rate of brown seaweeds.

1. Introduction

In East-Asia, brown seaweeds such as *Laminaria japonica*, *Undaria pinnatifida* and *Ecklonia kurome* have been traditional resources of foods or medicines for a long time. In the modern day, the brown seaweeds can also be an appropriate feedstock for bioconversion into bioethanol (John et al., 2011). *Microbulbifer elongatus* strain HZ11 was isolated from seawater of Zhoushan Islands of the East China Sea by direction isolation of the brown seaweed-degrading strain. With the secretion of many polysaccharidases such as alginate lyase, cellulose and amylase, strain HZ11 can degrade seaweed such as *L. japonica* into particles whose particle-size are less than 10 μm. For further research of brown seaweed saccharification and fermentation of bioethanol, we have determined the genome sequence of *M. elongatus* strain HZ11 (=CGMCC 6242).

M. elongatus HZ11 was cultivated on modified 2216 medium, which contains (per liter distilled water): yeast extract 5 g, peptone 1 g, ferric citrate 0.1 g, NaCl 19.45 g, $MgCl_2 \cdot 6H_2O$ 8.8 g, $CaCl_2 \cdot 2H_2O$ 1.8 g, KCl 0.55 g, $NaHCO_3$ 0.16 g, Na_2SO_4 3.24 g, KBr 0.08 g, $SrCl_2$ 34 mg, H_3BO_4 22 mg, $NaSiO_4$ 4 mg, NaF 2.4 mg, NH_4NO_3 1.6 mg, Na_2HPO_4 8 mg, pH 7.4 adjusted with NaOH, at 28 °C for 24 h. Genomic DNA was extracted using the method described by Marmur and Doty (1962). The genome was sequenced using paired-end sequencing technology (HiSeq 2000 system, Illumina, USA) (Bentley et al., 2008). The shotgun library was constructed with a 500 bp-span and a 2000 bp-span paired-end library. All clean reads were assembled into 19 scaffolds using the SOAPdenovo v.1.05 assembler (Li et al., 2010). After manual gap-filling steps and mapping to reference sequences, a high quality draft genome sequence with 9 contigs was obtained for further analysis.

Gene prediction was performed using Glimmer v. 3.02 (Delcher et al., 2007), and functions of the gene products were annotated by BLAST + (Camacho et al., 2009) using NCBI-nr protein (Sayers et al., 2012) and Swiss-Prot databases (Bairoch et al., 2004). The rRNA and tRNA genes were identified by using RNAmmer (Lagesen et al., 2007), tRNAscan-SE (Lowe and Eddy, 1997) and Rfam (Griffiths-Jones et al., 2003) database. Classification of predicted genes and pathways were analyzed by using COGs (Tatusov et al., 2000, 2001) and KEGG (Kanehisa and Goto, 2000) databases. The putative carbohydrate-active enzymes were analyzed by using CAZy (Lombard et al., 2014) and Pfam (Finn et al., 2014) databases.

The genome sequence of *M. elongatus* HZ11, with a genome size of 4,223,108 bp from 9 contigs, contains 56.70% G + C content. A total of 3293 coding sequences were predicted including 51 RNA genes and 904 hypothetical proteins.

The annotation results of genome suggest that strain HZ11 has large amount of genes related to brown seaweed degradation and polysaccharide utilization. As reported, brown seaweed is composed of several polysaccharides including alginate, laminarin, fucoidan and cellulose, among which, alginate composes 30–60% of the total sugars in brown

* Corresponding author. Tel./fax: +86 571 88206261.
E-mail address: wumin@zju.edu.cn (M. Wu).

http://dx.doi.org/10.1016/j.margen.2014.05.009
1874-7787/

seaweed (Chapman, 1970). In the protein-coding genes of strain HZ11, several genes encoding polysaccharidases relating to brown seaweed degradation were found, including nine alginate lyases (EC 4.2.2.3), one cellulase (EC 3.2.1.4) and two amylases (EC 3.2.1.1). Additionally, five agarases (EC 3.2.1.81) were also found, which is consistent to the phenotype of agar-liquefaction. Since agar is the typical component of red seaweed, strain HZ11 might also be able to degrade red seaweeds.

The analysis results of putative carbohydrate-active enzymes suggest that all nine putative alginate lyases (Alys) belong to four different polysaccharide lyases (PL) families. Five Alys are classified into PL7 family, where known activities are alginate lyase (Aly, EC 4.2.2.3) and G-specific alginate lyase (AlyG, EC 4.2.2.11); two Alys are classified into PL6 family, in which known activities are Aly and MG-specific alginate lyase (AlyMG, EC 4.2.2.—). In the PL6 family, only two Alys (Aly Q06365 and AlyMG AFC88009) were characterized, which have a mass of 44.5 kDa and 49.9 kDa respectively (Maki et al., 1993; Lee et al., 2012); one Aly is classified into PL17 family, which comprises Aly and oligoalginate lyase (Oal, EC 4.2.2.—). Currently, three-fourths of characterized Alys in PL17 family were Oals; the last Aly is classified into PL18 family that was known as Aly, AlyG and AlyMG. The neighbor-joining tree constructed by the amino acid sequences of alginate lyases also shows the same results (Fig. 1a). All five putative agarases (Agas) are classified into three different glycoside hydrolase (GH) families including GH16, GH86 and GH50. Two Agas are classified to GH50 family. In this family, almost all members are neoagarotetraose-producing Agas, which suggest that these two Agas may be neoagarotetraose-producing Agas. Additionally, three types of carbohydrate-binding modules (CBM) are found which may promote the association of the enzyme with the substrate (Boraston et al., 2004). In detail, CBM32 (or F5/8 type C domain) is related to some Alys in PL7 family; CBM16 (or CBM_4_9) is related to Alys in PL18 and PL6 families; CBM6 is related to Agas in GH16 and GH86 families.

Interestingly, our analysis also reveals that strain HZ11 contains all genes encoding the enzymes involved in the Entner–Doudoroff (ED) pathway, including glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49), 6-phosphogluconolactonase (EC 3.1.1.31), phosphogluconate dehydratase (EC 4.2.1.12), 2-dehydro-3-deoxyphosphogluconate aldolase (EC 4.1.2.14), pyruvate decarboxylase (EC 1.2.4.1) and alcohol dehydrogenase (EC 1.1.1.1), which imply the complete ED pathway is considered to exist (Conway, 1992). Moreover, the gene encoding 2-dehydro-3-deoxygluconate kinase (EC 2.7.1.45) was found, which

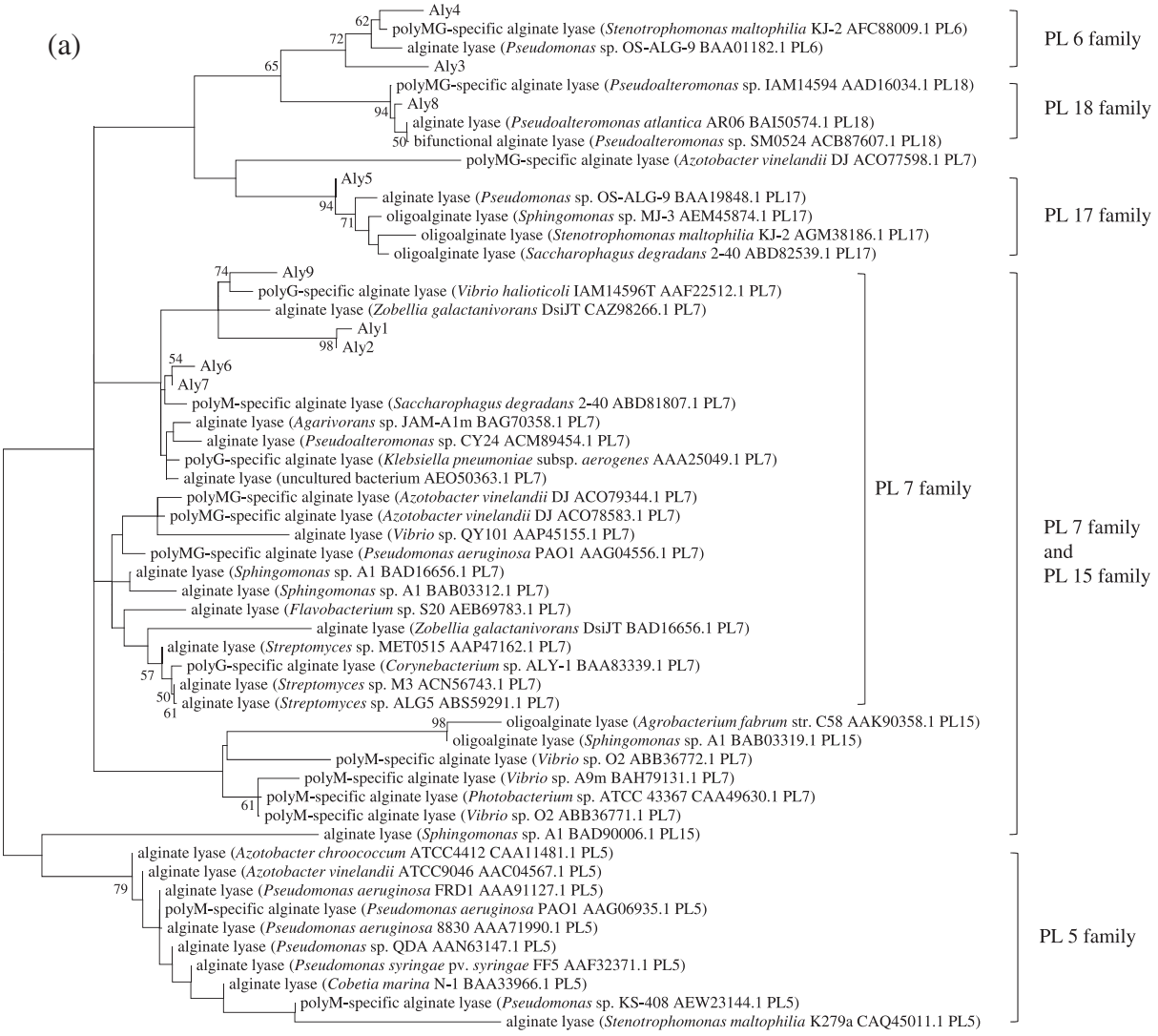

Fig. 1. (a) Neighbor-joining tree using the Jones–Taylor–Thornton (JTT) model based on the amino acid sequences of alginate lyases, showing the relationship between alginate lyases in *Microbulbifer elongatus* strain HZ11 and other bacterial characterized alginate lyases which belong to PL5, PL6, PL7, PL15, PL17 and PL18 families (based on CAZy database). Bootstrap values are based on 500 replicates; values ≥50% are shown. (b) Bacterial fermentation pathway of abundant fermentable components in brown seaweed and relevant enzymes. ADH: alcohol dehydrogenase; Aly: alginate lyase; GPD: glucose-6-phosphate 1-dehydrogenase; GNPD: phosphogluconate dehydratase; KDGK: 2-dehydro-3-deoxygluconate kinase; KDPGA: 2-dehydro-3-deoxyphosphogluconate aldolase; PDC: pyruvate decarboxylase; PGL: 6-phosphogluconolactonase; Oal: oligoalginate lyase; DEHU: 4-deoxy-L-erythro-5-hexoseulose uronic acid; GAP: 3-phosphate-glyceraldehyde; KDG: 2-keto-3-deoxy-gluconate; KDPG: 2-dehydro/keto-3-deoxy-phosphogluconate.

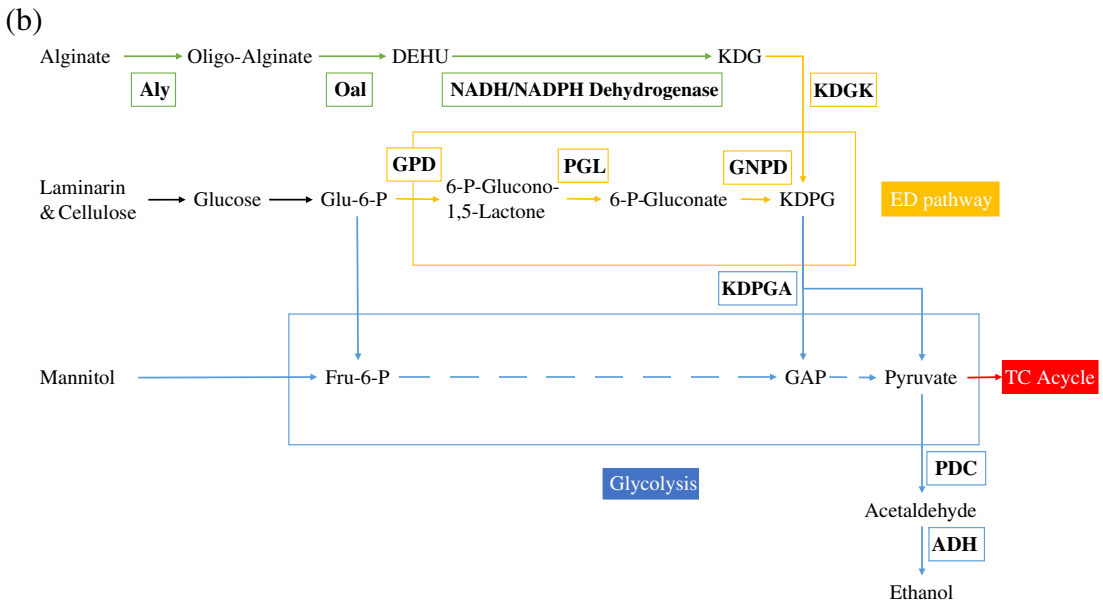

Fig. 1 (*continued*).

plays an important role in the connection of alginate depolymerization and ED pathway (Fig. 1b, Preiss and Ashwell, 1962a,b). Based on the analysis, strain HZ11 has the potential ability to produce bioethanol directly from alginate by ED pathway with moderate genetic modification such as delete *pflB-focA*, *frdABCD*, and *ldhA* from the genome to divert carbon flux away from fermentative by-products (Wargacki et al., 2012). Additionally, an engineered microbial platform and a synthetic yeast platform were reported as genetic modification strains to produce ethanol from brown seaweeds by using the similar pathway above (Wargacki et al., 2012; Enquist-Newman et al., 2014). Up to now, most reported bioethanol transferred from brown seaweeds were produced from mannitol or glucan including cellulose and laminarin (Yanagisawa et al., 2011; Lee et al., 2013; Wang et al., 2013). Hence, by developing the fermentation of alginate which is the most abundant component in brown seaweeds, strain HZ11 may significantly increase the yield of bioethanol from brown seaweeds and the utilization rate of brown seaweeds (Wargacki et al., 2012).

This Whole Genome Shotgun project of *M. elongatus* HZ11 (=CGMCC 6242) has been deposited at DDBJ/EMBL/GenBank database under the accession JELR00000000.

Acknowledgment

This work was supported by Research Program of the National Natural Science Foundation of China (Grant no.: 31170001).

References

Bairoch, A., Boeckmann, B., Ferro, S., Gasteiger, E., 2004. Swiss-Prot: juggling between evolution and stability. Brief Bioinform. 5, 39–55.
Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., et al., 2008. Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456, 53–59.
Boraston, A.B., Bolam, D.N., Gilbert, H.J., Davies, G.J., 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem. J. 382, 769–781.
Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinformatics. 10, 421.
Chapman, V.J., 1970. Seaweeds and their uses. Seaweeds and their uses.
Conway, T., 1992. The Entner–Doudoroff pathway: history, physiology and molecular biology. FEMS Microbiol. Rev. 9, 1–27.
Delcher, A.L., Bratke, K.A., Powers, E.C., Salzberg, S.L., 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics. 23, 673–679.
Enquist-Newman, M., Faust, A.M., Bravo, D.D., et al., 2014. Efficient ethanol production from brown macroalgae sugars by a synthetic yeast platform. Nature 505, 239–243.
Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E.L., Tate, J., Punta, M., 2014. Pfam: the protein families database. Nucleic Acids Res. 42, D222–D230.
Griffiths-Jones, S., Bateman, A., Marshall, M., Khanna, A., Eddy, S.R., 2003. Rfam: an RNA family database. Nucleic Acids Res. 31, 439–441.
John, R.P., Anisha, G.S., Nampoothiri, K.M., Pandey, A., 2011. Micro and macroalgal biomass: a renewable source for bioethanol. Bioresour. Technol. 102, 186–193.
Kanehisa, M., Goto, S., 2000. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28, 27–30.
Lagesen, K., Hallin, P., Rodland, E.A., Staerfeldt, H.H., Rognes, T., Ussery, D.W., 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 35, 3100–3108.
Lee, S., Choi, S.H., Lee, E.Y., Kim, H.S., 2012. Molecular cloning, purification, and characterization of a novel polyMG-specific alginate lyase responsible for alginate MG block degradation in *Stenotrophomas maltophilia* KJ-2. Appl. Microbiol. Biotechnol. 95, 1643–1653.
Lee, J.Y., Li, P., Lee, J., Ryu, H.J., Oh, K.K., 2013. Ethanol production from *Saccharina japonica* using an optimized extremely low acid pretreatment followed by simultaneous saccharification and fermentation. Bioresour. Technol. 127, 119–125.
Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z., Li, Y., Li, S., Shan, G., Kristiansen, K., Li, S., Yang, H., Wang, J., Wang, J., 2010. De novo assembly of human genomes with massively parallel short read sequencing. Genome Res. 20, 265–272.
Lombard, V., Golaconda, R.H., Drula, E., Coutinho, P.M., Henrissat, B., 2014. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 42, D490–D495.
Lowe, T.M., Eddy, S.R., 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25, 955–964.
Maki, H., Mori, A., Fujiyama, K., Kinoshita, S., Yoshida, T., 1993. Cloning, sequence analysis and expression in *Escherichia coli* of a gene encoding an alginate lyase from *Pseudomonas* sp. OS-ALG-9. J. Gen. Microbiol. 139, 987–993.
Marmur, J., Doty, P., 1962. Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. J. Mol. Biol. 5, 109–118.
Preiss, J., Ashwell, G., 1962a. Alginic acid metabolism in bacteria. I. Enzymatic formation of unsaturated oligosaccharides and 4-deoxy-L-erythro-5-hexoseulose uronic acid. J. Biol. Chem. 237, 309–316.
Preiss, J., Ashwell, G., 1962b. Alginic acid metabolism in bacteria. II. The enzymatic reduction of 4-deoxy-L-erythro-5-hexoseulose uronic acid to 2-keto-3-deoxy-D-gluconic acid. J. Biol. Chem. 237, 317–321.
Sayers, E.W., Barrett, T., Benson, D.A., et al., 2012. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 40, D13–D25.
Tatusov, R.L., Galperin, M.Y., Natale, D.A., Koonin, E.V., 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. 28, 33–36.
Tatusov, R.L., Natale, D.A., Garkavtsev, I.V., Tatusova, T.A., Shankavaram, U.T., Rao, B.S., Kiryutin, B., Galperin, M.Y., Fedorova, N.D., Koonin, E.V., 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res. 29, 22–28.
Wang, J., Kim, Y.M., Rhee, H.S., Lee, M.W., Park, J.M., 2013. Bioethanol production from mannitol by a newly isolated bacterium, *Enterobacter* sp. JMP3. Bioresour. Technol. 135, 199–206.
Wargacki, A.J., Leonard, E., Win, M.N., Regitsky, D.D., Santos, C.N.S., Kim, P.B., Cooper, S.R., Raisner, R.M., Herman, A., Sivitz, A.B., Lakshmanaswamy, A., Kashiyama, Y., Baker, D., Yoshikuni, Y., 2012. An engineered microbial platform for direct biofuel production from brown macroalgae. Science 335, 308–313.
Yanagisawa, M., Nakamura, K., Ariga, O., Nakasaki, K., 2011. Production of high concentrations of bioethanol from seaweeds that contain easily hydrolyzable polysaccharides. Process Biochem. 46, 2111–2116.