Genome Mining Discovery of Hydrogen Production Pathway of *Klebsiella* sp. WL1316 Fermenting Cotton Stalk Hydrolysate

Yanbin Li

Anhui Polytechnic University

Qin Zhang (✉ jhtabszq@163.com)

Anhui Polytechnic University <https://orcid.org/0000-0001-5718-526X>

Zhanwen Liu

Tarim University

Hui Jiang

Tarim University

Qinghua Jia

Tarim University

Research Article

Keywords: Genome, Pathway, *Klebsiella* sp., Hydrogen production, Cotton stalk hydrolysate

Posted Date: August 23rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-808932/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at International Microbiology on February 11th, 2022. See the published version at <https://doi.org/10.1007/s10123-022-00241-0>.

Abstract

Genome sequencing was used to identify key genes for the generation of hydrogen gas through cotton stalk hydrolysate fermentation by *Klebsiella* sp. WL1316. Genome annotation indicated that the genome size was 5.2 Mb with GC content 57.6%. Xylose was metabolised in the pentose phosphate pathway via the conversion of xylose to xylulose in *Klebsiella* sp. WL1316. This strain contained diverse formate-hydrogen lyases and hydrogenases with gene numbers higher than closely related species. A metabolic network involving glucose, xylose utilisation, and fermentative hydrogen production was reconstructed. Metabolic analysis of key node metabolites showed that glucose and xylose metabolism influenced biomass synthesis and biohydrogen production. Formic acid accumulated during fermentation at 24–48 h but decreased sharply after 48 h, illustrating the splitting of formic acid to hydrogen gas during early-to-mid fermentation. The Krebs cycle was the main competitive metabolic branch of biohydrogen synthesis at 24 h of fermentation. Lactic and acetic acid fermentation and late ethanol accumulation competed the carbon skeleton of biohydrogen synthesis after 72 h of fermentation, indicating that these competitive pathways are regulated in middle-to-late fermentation (48–96 h). This study is the first to elucidate the metabolic mechanisms of mixed sugar utilisation and biohydrogen synthesis based on genomic information.

Introduction

Under the combined pressure of decreasing fossil energy reserves and increasing environmental pollution, making full use of lignocellulose and other renewable resources to develop clean energy has attracted the attention of the scientific community worldwide (Boodhun et al. 2017; Rodionova et al. 2017; Tian et al. 2019). Hydrogen energy is the cleanest renewable energy on Earth. Among current production methods, hydrogen produced by the dark fermentation of lignocellulosic materials by microorganisms is one of the most prominent strategies (Saratale and Oh 2015; Sivagurunathan et al. 2017; Kamaraj et al. 2020; Kumar et al. 2019; Shanmugam et al. 2020). In this method, the selection of the fermentation strain is of great importance. However, fermentative hydrogen production by single wild strains is relatively low. Therefore, measurements, such as metabolic pathway analysis and engineering modification, are needed for strains with clear genome information (Tran et al. 2014; Cha et al. 2013; Lu et al. 2016; Jo and Cha 2015; Lu et al. 2017; Lee et al. 2019; Zhang et al. 2020) in order to reconstruct the hydrogen production metabolic pathways and increase the yield of hydrogen produced by hydrogen-producing strains.

In recent years, with the development of genome sequencing, biosynthesis networks of biofuel-generating microorganisms have been analysed and regulated using genomics (Bruce et al. 2016; Deegan et al. 2019; Mekanik et al. 2019). Genome-scale information related to microbial synthesis metabolic pathways is necessary for the analysis the metabolic network for the accurate synthesis and metabolic regulation of biofuels. As early as 2003, Kalia et al. analysed the KEGG metabolic pathways of several hydrogen-generating microorganisms based on 84 genomic sequences and 92 partial genomic sequences to screen for new strains with a high efficiency for hydrogen production. In the field of bioenergy, several

reports have elucidated the biofuel synthesis pathway and regulatory mechanisms using genomics information. For example, the whole genome sequencing of *Clostridium butyrate* INCQS635, which is able to produce ethanol and butanol fuel, revealed the metabolic mechanism for the conversion of complex sugars into biofuels (Bruce et al., 2015). As for hydrogen-producing microbes, genomic information has been applied in the following ways: (1) to search for novel hydrogenase genes of highly efficient hydrogen-producing microbes (Kamalaskar et al. 2016); (2) to analyse metabolic pathways for microbial hydrogen synthesis (Khanna et al. 2013; Castro et al. 2013); (3) to reveal the regulatory genes for biohydrogen synthesis in wild and mutant hydrogen-producing strains via comparative omics analysis (Gökçe et al. 2017).

Klebsiella is a facultative anaerobic hydrogen-producing bacterium with a high hydrogen production potential. At present, research on hydrogen-producing bacteria has mainly focused on the optimisation of the fermentation conditions and the identification of cheap substrates. By contrast, studies on the analysis and regulation of its hydrogen-producing metabolic pathways remain relatively scarce. In general, there is a lack of conclusive results regarding the hydrogen production metabolism of *Klebsiella*. Therefore, the metabolic regulation of hydrogen production pathways in *Klebsiella* is not as well understood as in *E. coli* (Maeda et al. 2007; Soo et al. 2017), with most studies choosing to focus on the overexpression of key enzyme genes in specific metabolic pathways (Bai et al. 2012; Wu et al. 2014; Jawed et al. 2016). However, some strains of *Klebsiella* have undergone genome sequencing, including *Klebsiella oxytoca* SA2 with nitrogen fixation (Chen et al. 2013), *Klebsiella* sp. BRI6-2 with lignin degradation (Woo et al. 2014), *Klebsiella pneumoniae* with drug resistance (Hua et al. 2014), and *Klebsiella oxytoca* HKOPL1 isolated from faeces of giant pandas (Jiang et al. 2014). However, there have been no reports on the whole-genome sequencing of hydrogen-producing strains of *Klebsiella*, despite the fact that it would provide insights into its hydrogen production metabolic pathways.

In this study, the whole genome of *Klebsiella* sp. was sequenced with the aim of reconstructing the metabolic network for hydrogen production from cotton stalk hydrolysate to identify key metabolic pathways for glucose and xylose utilisation and biohydrogen synthesis. Moreover, the essential enzyme activities were measured and metabolic flux analysis was developed by detecting the metabolites in the node of key pathways to confirm the genome-scale reconstructed metabolic pathways and provide a new interpretation of the metabolic basis for hydrogen production from cotton stalk hydrolysate.

Materials And Methods

Microorganism

In this study, *Klebsiella* sp. WL1316 was used for genomic DNA extraction and fermentative hydrogen production with the ability to utilise cotton stalk hydrolysate for fermentative hydrogen production (Li et al. 2018; Li et al. 2020). The medium, comprised of 10 g/L xylose, 10 g/L glucose, 5 g/L beef extract, 10 g/L peptone, 5 g/L NaCl, 0.5 g/L KH_2PO_4 , and 1 g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, was used to activate the strain and cultivate the seed culture (Li et al. 2018).

Methods

Genomic DNA extraction, purification, and sequencing

Genomic DNA was extracted and purified using an Ezup Column Bacteria Genomic DNA Purification Kit (Sangon, Shanghai, China) following the manufacturer's instructions. The purified genomic DNA was sent to Beijing Novogene Biological Information Technology Co. Ltd. for whole-genome sequencing using the MiSeq sequencing platform. The sequencing data were deposited into the NCBI database under the BioProject accession number PRJNA611005.

Genome assembly and preliminary analysis

GC and depth correlation analysis and K-mer analysis were performed before assembly. Further, SOAPdenovo software was used to assemble the processed reads, and SOAPaligner sequence alignment software was used to align the reads with the assembly results. The *de novo* prediction of the assembly results was developed using the Glimmer 3.0 gene prediction software, and GeneMarkS software was used to predict the coding gene; tRNAscan software was used to predict the tRNA region and secondary structure; sRNA was predicted using Rfam software; RepeatMasker software was used to predict the scattered repeats; and TRF was used to search the tandem repeat sequences (Zerbino and Birney 2008; Tarailo-Graovac and Chen 2009; Ye et al. 2015; Xu et al. 2016).

Bioinformatics analysis based on the genome sequencing results of *Klebsiella* sp. WL1316

The genomic data of *Klebsiella* sp. WL1316 were blasted through the GO (Gene Ontology), COG (Cluster of Orthologous Groups of proteins), and KEGG (Kyoto Encyclopedia of Genes and Genomes) databases for functional gene prediction (Yu et al. 2006; Horton et al. 2007; Jamialahmadi et al. 2016). As a result, the KEGG metabolic pathways and functional proteins related to glucose metabolism and fermentative hydrogen production by of *Klebsiella* sp. were identified. Further comparative genomic analysis was performed using five bacteria, *Klebsiella pneumoniae* subsp. HS11286, *Klebsiella pneumoniae* subsp. NTUH-K2044, *Klebsiella oxytoca* KONIH1, *Enterobacter ludwigii* EcWSU1, and *Enterobacter* sp. R4-368, which were selected as comparative objects, of which the last two strains were reported to be hydrogen-producing bacteria. The genome characterisation of the five reference strains is illustrated in Supplementary Table 1. The genes encoding the hydrogenase functional proteins were manually mined based on the annotated genome data. The gene numbers of different functional proteins were statistically analysed, and a heatmap was constructed using Heml software.

Cultivation and sampling

After activation, a single colony strain was picked and inoculated into the seed culture medium. The seed was then agitated at a speed of 120 r/min for 15–16 h under 37°C. The OD₆₀₀ of the inoculum was adjusted to approximately 1.0, and inoculated into the fermentation medium at an inoculation size of 10% (v/v). The fermentation was carried out using a self-designed fermenter and following the optimised conditions previously reported by our group: initial sugar concentration of 40 g/L, fermentation temperature 37°C, initial pH 8.0 (Li et al. 2018).

Measurement

The volume of the hydrogen gas was examined by 1 mol/L NaOH displacement in an inverted burette, and the biohydrogen concentration was measured using a hand-held hydrogen detector (KP810H20; Zhong'an Electronic Detection, Henan, China). The volume of hydrogen gas was measured at 6-h intervals and added after 24 h of daily hydrogen production. The hydrogen production volume measured every 24 h was accumulated and converted into the cumulative hydrogen production in the corresponding fermentation period. At the end of fermentation, the fermentative broth was collected and centrifuged at $8000 \times g$ for 10 min and filtered through syringe filters with a 0.22- μ m membrane before analysis. The glucose and xylose concentrations were measured using a glucose detection kit (Comin, Suzhou, China) and a xylose detection kit (ZZStandard, Shanghai, China). The concentrations of key node metabolites, such as pyruvic acid, citric acid, succinic acid, lactic acid, acetic acid, formic acid, and ethanol, were measured using detection kits (ZZStandard, Shanghai, China) using an automatic microplate reader (Multiskan FC; Thermo Scientific, Massachusetts, USA).

Results And Discussion

General features of the genome of *Klebsiella* sp. WL1316 and functional gene prediction

In this study, genome sequencing and characterisation were conducted on *Klebsiella* sp. WL1316 to gain a comprehensive understanding of the functional genes of this strain, especially those response to glucose and xylose metabolism and hydrogen production. The genome characteristics are listed in Table 1. The genome of *Klebsiella* sp. WL1316 is a chromosomal DNA with a total length of 5.2 Mb and a GC content of 57.6%. A total of 4915 coding genes were predicted, up to 86.7% of the whole genome, and 76.7% of the predicted coding genes encoded products possessed definite functions, while the rest were hypothesised proteins or unknown proteins. The functional protein-coding genes were annotated by aligning the sequences to the COG, GO, and KEGG databases. The COG and GO annotated protein-coding gene numbers reached 3768 and 3622, respectively, representing 76.7% and 73.7% of the total gene numbers, respectively. The COG and GO annotated functional categories are illustrated in Supplementary Fig. 1 and Supplementary Fig. 2. In addition, the KEGG annotated protein-coding genes reached 3408 genes, which were involved in 152 metabolic pathways. The matching genes were mined and discovered to locate them in the variable KEGG metabolic pathways, which covered the basic metabolic pathways,

such as carbohydrate metabolism, lipid metabolism, amino acid metabolism, energy metabolism, and some special metabolic pathways (Supplementary Fig. 3).

Table 1
Genome characteristics of *Klebsiella* sp.
WL1316

General features	
Genome size (Mb)	5.20
Gene number (#)	4,915
GC content (%)	57.6
N50 Length (bp)	136,619
N90 Length (bp)	51,942
Number of scaffolds	76
Properties of gene annotation	
Number of rRNA genes	5
Number of tRNA genes	76
Number of sRNA genes	41
COG annotation	3768
GO annotation	3622
KEGG annotation	3408

Annotation of the metabolic pathways for glucose and xylose utilisation and bioconversion

In a previous study, we reported that *Klebsiella* sp. WL1316 intrinsically possesses the ability to utilise glucose and xylose produced by lignocellulosic hydrolysate (Li et al. 2018). Therefore, we firstly focused on the sugar metabolic pathway, and found that glucose could be metabolised and subsequently converted in the fermentation process via the EMP and PPP pathways, as illustrated in Supplementary Fig. 4. However, there was no reference pathway for xylose metabolism that could be blasted from the KEGG database. Therefore, several functional proteins related to xylose transportation and metabolism were further mined based on the gene annotation results. As a result, we discovered that the *Klebsiella* sp. WL1316 was equipped with the ribose/xylose/arabinose/galactoside ABC-type transport system, as well as the D-xylose transport system ATP-binding component, D-xylose ABC transporter, and periplasmic D-xylose-binding protein, which are conducive to the transmembrane transport of xylose in bacteria.

Furthermore, in terms of xylose metabolism, this strain was found to possess xylose isomerase and xylulokinase, which isomerises xylose to xylulose and further phosphorylates xylulose-5-phosphate to enter the pentose phosphate pathway for metabolism. The schematic pathways are illustrated in Fig. 1.

Identification of the genes coding for key functional proteins for hydrogen production

According to the literature, *Klebsiella* is a facultative anaerobic bacterium that is able to generate hydrogen gas via mixed acid fermentation pathway, in which the key enzyme for hydrogen production is formate-hydrogen lyase (Jawed et al. 2016). In addition, hydrogenase is also an important enzyme that catalyses the hydrogen production process. As such, we analysed the hydrogenase and formate-hydrogen lyase-related functional proteins of *Klebsiella* sp. WL1316 based on genome annotation results. As shown in Fig. 2, this hydrogen-producing bacterium possesses diverse hydrogenase and formate-hydrogen lyase related functional proteins. The hydrogenase functional proteins included the protein subunits of hydrogenase, the components of the Fe-S cluster and nickel-iron ion binding protein in the hydrogenase active center, and the assembly protein in the metal ion binding center of hydrogenase. The hydrogenase-related gene cluster is illustrated in Fig. 2a. The formate-hydrogen lyase functional proteins included multiple protein subunits, formate-hydrogen lyase regulatory proteins and transcriptional activators, and formate-hydrogen lyase Fe-S protein (Fig. 2b). The distributions of gene clusters of hydrogenase and formate-hydrogen lyase were all located in the genome (Fig. 2c). The annotated functional proteins and their Gene_ids are listed in Supplementary Table 2. The diversity of hydrogenase and hydrogen formate lyase indicated that the strain had a strong hydrogen production potential at the nucleic acid and protein levels. Hydrogenase is the most important functional enzyme that catalyses the hydrogen production process in hydrogen-producing bacteria. Therefore, we further compared the difference in the category and number of hydrogenase functional proteins between *Klebsiella* sp. WL1316 and the five reference bacteria based on the bacterial genome information (Fig. 3). Figure 3a reveals that *Klebsiella* sp. WL1316 in this study obtained abundant hydrogenase functional proteins, while the hydrogenase3 large subunit was dominant in *Klebsiella pneumoniae* subsp. NTUH-K2044, *Klebsiella oxytoca* KONIH1, *Enterobacter* sp. R4-368. Clustering analysis indicated the category diversity of hydrogenase functional proteins in *Klebsiella* sp. WL1316 was more similar to the two reported hydrogen-producing bacteria *Enterobacter ludwigii* EcWSU1 and *Enterobacter* sp. R4-368. The statistical results of gene numbers of hydrogenase functional proteins indicated that *Klebsiella* sp. WL1316 contained the highest gene number of hydrogenase functional proteins, which is close to the hydrogen-producing bacterium *Enterobacter* sp. R4-368 (Fig. 3b). This result indicated that *Klebsiella* sp. WL1316 could be a hydrogen-producing bacterium with a high hydrogen production potential, with more hydrogenase functional proteins than the five reference bacteria.

Construction of the genome scale metabolic model for hydrogen production

Based on the above analysis of glucose metabolic pathways, xylose metabolism-related enzymes and functional proteins, as well as enzymes and functional proteins related to biological hydrogen synthesis, the metabolic network for hydrogen production by *Klebsiella* sp. WL1316 from the fermentation of cotton stalk hydrolysate was reconstructed. This metabolic network covers not only the glucose and xylose metabolic pathways, but also the biohydrogen synthesis pathway and some competitive metabolic branches, as illustrated in Fig. 4. The biohydrogen synthesis pathway included the mixed acid fermentation pathway and hydrogenase catalytic pathway, while the competitive metabolic branches included the succinic, lactic, acetic acids, and ethanol generation pathways, which acted as the branches of the mixed acid fermentation pathway.

Metabolic analysis for hydrogen production of *Klebsiella* sp. WL1316 from the fermentation of cotton stalk hydrolysate

Pyruvic acid is an important intermediate metabolite of *Klebsiella* sp. WL1316 for the metabolism of glucose and xylose in cotton stalk hydrolysate. This strain enters the mixed acid pathway for fermentative hydrogen production, which was initiated by the decarboxylation of pyruvate. Therefore, pyruvate decarboxylase (PDC) is the key enzyme involved in this metabolic pathway. Supplementary Fig. 5 shows that PDC activity was relatively high after 24 h of fermentation, indicating that this enzyme promoted pyruvate decarboxylation, and then entered the fermentation pathway. Thereafter, PDC activity decreased gradually to 96 h, and the PDC activity declined to a relatively low level, almost 0, after 120 h of fermentation. The lactate dehydrogenase (LDH) activity was low at 24–48 h and increased rapidly to the highest level ($0.027 \text{ U}/10^4 \text{ cells}$) at 72 h, indicating that high LDH activity in 72-h fermentation may lead to a large accumulation of lactic acid and compete for the distribution of metabolic carbon skeleton in the biohydrogen synthesis pathway. Such a competitive effect for biohydrogen synthesis controlled by LDH was consistent with that reported by Jung et al. (2012). The aldehyde dehydrogenase (ALDH) activity reached the highest level after fermenting 48–96 h, indicating that the metabolic pathway of acetic acid and ethanol synthesis mainly competed for the metabolic carbon skeleton of the biohydrogen synthesis in the mid-to-late fermentation stage.

The metabolic analysis was first developed toward the key metabolites in the conversion of sugars, such as glucose, xylose, and xylulose (Fig. 5). Glucose and xylose consumption rates decreased with fermentation time during the fermentation stage, and the results showed that glucose was metabolised effectively, which promoted the synthesis and fermentation of cell material. According to the biohydrogen synthesis metabolic network of *Klebsiella* sp. WL1316, xylose was metabolised by converting xylose to xylulose through the PPP pathway, therefore the generation rate of xylulose was further monitored. Although the generation rates of xylulose were lower than consumption rates of xylose during the whole

fermentation stage, its generation rates also gradually decreased, indicating that xylulose was also rapidly metabolised via the PPP pathway, which was provided as a carbon skeleton for the fermentation process.

A survey of the literature showed that key node metabolites, such as succinic, lactic, acetic, formic acids, and ethanol, are typically present in the mixed acid fermentation pathway (Jenni et al. 2013; Jung et al. 2012; Yoshida et al. 2005; Lu et al. 2009; Zhao et al. 2009). Therefore, the time-dependent generation profiles of these key node metabolites were analysed (Fig. 6). Pyruvic acid is the most important metabolic intermediate, which initiates the mixed acid fermentation pathway, especially for hydrogen gas generation. Pyruvate is also a key node metabolite for mixed acid fermentation. In this study, the pyruvate node obtained a high generation rate at 24 h of fermentation, and then decreased rapidly, indicating that pyruvate underwent rapid decarboxylation transformation, with more carbon skeletons flowing to the fermentation process. The splitting of formic acid is the most important hydrogen-producing pathway for facultative anaerobes (Zhao et al. 2009; Lu et al. 2009; Yoshida et al. 2005). The results revealed that the generation rate of formic acid was relatively low at 24 h of fermentation, but reached a peak value at 48 h, indicating that a large amount of formic acid accumulated between 24 and 48 h, and that the metabolic carbon skeleton mostly flowed to the mixed acid fermentation pathway at this fermentation stage. Subsequently, the generation rate of formic acid decreased sharply, indicating that formic acid may be used for biohydrogen synthesis in large quantities during the mid-to-late fermentation stage. The shift in citric acid concentration revealed that there was a high generation rate after 24–48-h fermentation at the citric acid metabolic node, which enabled many of the metabolic carbon skeletons to originate from the flow of pyruvate nodes to the TCA cycle. The generation rates of succinic acid were relatively low during the whole fermentation stage, indicating that succinic acid accumulated less under anaerobic fermentation conditions, indicating that it may compete for the metabolic carbon skeleton of biohydrogen synthesis with relatively weak competitiveness. The generation rate of the lactic acid node plays an important role in the whole fermentation stage, which was enhanced sharply after fermenting for 48 h, and reached a peak value at 72 h, indicating that the lactic acid production pathway may be the main competitive branch for biohydrogen synthesis. In addition, the generation rates of the ethanol and acetic acid in the same metabolic branch also significantly affected biohydrogen synthesis. The generation rate of the ethanol remained at a high level during the whole fermentation stage and decreased gradually with the fermentation time but increased to a certain extent at 120 h of fermentation. This indicates that ethanol may accumulate more in the late fermentation period at this time, and it may be an important competitive branch for biohydrogen synthesis. The metabolic flux of acetic acid was higher during the fermentation period of 48–72 h, indicating that the acetic acid production was higher in this fermentation stage, which may be strongly competitive for the metabolic carbon skeleton. In general, acetic acid was produced in a large quantity after 48 h of fermentation, and ethanol was accumulated in large quantities in this metabolic branch, indicating that this metabolic branch is an important branch in competing for the carbon skeleton of biohydrogen synthesis after 48 h of fermentation.

Conclusions

Overall, this study obtained novel results with regards to three aspects: (1) demonstrating the glucose and xylose metabolic pathways at the gene level of *Klebsiella* sp. WL1316; (2) revealing the high effective hydrogen production potential of *Klebsiella* sp. WL1316 from the gene diversity level, and confirming the occurrence of multiple pathways affecting the efficient production of hydrogen gas that are commonly found in facultative anaerobes; (3) constructing a genome scale of the metabolic network model (GSMM) for hydrogen production, and verifying the occurrence of the competitive and hydrogen biosynthesis pathways in *Klebsiella* sp. via measurement of consumption or generation rates, which is crucial for understanding the hydrogen production potential of this strain. On this basis, improvements in the reducing sugar utilisation pathways and channelling the metabolic flux of the competitive and hydrogen biosynthesis pathways will be needed to fine-tune the metabolic capabilities of this strain for highly efficient hydrogen production. Hence, the GSMM constructed in this study could play a pivotal role in the design and acceleration of the metabolic engineering approaches to harness the potential of this hydrogen-producing bacterium in subsequent studies.

Declarations

Ethics approval and Consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

Genome sequence used in this study has been deposited in the NCBI SRA with the BioProject accession number PRJNA611005.

Competing interests

The authors declare that they have no conflict of interest.

Funding

This work was supported by the National Natural Science Foundation of China (Grant No. 21406150), the Key Research and Development Project of Anhui Province in China (Grant No. 201904a07020003), the Talent Start-up Fund of Anhui Polytechnic University in China (Grant No. 2018YQQ029) , and the Top Young and Middle-aged Talents Program of Anhui Polytechnic University in China (Grant No.

S022019014) to conduct the genome sequencing and metabolic analysis experiments of *Klebsiella* sp. WL1316.

Author's contribution

Y. L. and Q. Z. conceived the presented idea, handled the data and drafted the manuscript. Q. Z. and Z. Liu developed the theory and performed the computations. Y. L. and H. J. carried out the fermentation experiments and examined the key metabolites and enzyme activities. H. J. and Q. J. verified the analytical methods. All authors discussed the results and contributed to the final manuscript. All authors read and approved the manuscript.

Acknowledgements

Not applicable.

Authors' information

Q Zhang, Professor, mainly engaged in researches of industrial microbial genomics and metabolic engineering; Y. Li, Professor, mainly engaged in lignocellulosic biorefinery studies.

References

1. Bai L, Wu X, Jiang L, Liu J, Long M (2012) Hydrogen production by over-expression of hydrogenase subunit in oxygen-tolerant *Klebsiella oxytoca* HP1. Int J Hydrogen Energ 37(17):13227–13233. <https://doi.org/10.1016/j.ijhydene.2012.03.048>
2. Boodhun BSF, Mudhoo A, Kumar G, Kim SH, Lin CY (2017) Research perspectives on constraints, prospects and opportunities in biohydrogen production. Int J Hydrogen Energ 42(45):27471–27481. <https://doi.org/10.1016/j.ijhydene.2017.04.077>
3. Bruce T, Leite FG, Miranda M, Thompson CC Jr, Faber NP, Thompson M FL (2016) Insights from genome of *Clostridium butyricum* INCQS635 reveal mechanisms to convert complex sugars for biofuel production. Arch Microbiol 198(2):115–127. <https://doi.org/10.1007/s00203-015-1166-4>
4. Castro JF, Razmilic V, Gerdtzen ZP (2013) Genome based metabolic flux analysis of *Ethanoligenens harbinense* for enhanced hydrogen production. Int J Hydrogen Energ 38(3):1297–1306. <https://doi.org/10.1016/j.ijhydene.2012.11.007>
5. Cha M, Chung D, Elkins JD, Guss AM, Westpheling J (2013) Metabolic engineering of *Caldicellulosiruptor bescii* yields increased hydrogen production from lignocellulosic biomass. Biotechnol Biofuels 6:85. <https://doi.org/doi:10.1186/1754-6834-6-85>

6. Chen M, Lin L, Zhang Y, Sun L, An Q (2013) Genome Sequence of *Klebsiella oxytoca* SA2, an Endophytic nitrogen fixing bacterium isolated from the pioneer grass *Psammochloa villosa* [J]. Genome Announcements 1(4):e00601–e00613. <https://doi.org/10.1128/genomeA.00601-13>
7. Deegan Y, Sargsyan A, Hovhannisyan N, Babayan B, Petrikov K, Vainstein M (2019) Analysis of genome sequence and trehalose lipid production peculiarities of the thermotolerant *Gordonia* strain. J Basic Microbiol 10:1–7. <https://doi.org/10.1002/jobm.201900439>
8. Gökçe A, Öztürk Y, Çakar ZP (2017) Comparative genomic analysis of two heat-resistant *Rhodobacter capsulatus* mutants with different hydrogen production levels reveals mutations related to hydrogen production. Int J Hydrogen Energ 42(32):20529–20539. <https://doi.org/10.1016/j.ijhydene.2017.07.005>
9. Horton P, Park K-J, Obayashi T, Fujita N, Harada H, Adams-Collier C, Nakai K (2007) WoLF PSORT: protein localization predictor. Nucleic Acids Res 35(S2): 585–587. <https://doi.org/10.1093/nar/gkm259>
10. Hua X, Chen Q, Li X, Feng Y, Ruan Z, Yu Y (2014) Complete genome sequence of *Klebsiella pneumoniae* sequence type 17, a multidrug-resistant strain isolated during tigecycline treatment. Genome Announcements 2(6):e01337–e01314. <https://doi.org/10.1128/genomeA.01337-14>
11. Jamialahmadi O, Motamedian E, Hashemi-Najafabadi S (2016) BiKEGG: a COBRA toolbox extension for bridging the BiGG and KEGG databases. Mol Biosyst 12(11): 3459–3466. <https://doi.org/10.1039/c6mb90040b>
12. Jawed M, Pi J, Xu L, Zhang H, Hakeem A, Yan Y (2016) Enhanced H₂ production and redirected metabolic flux via overexpression of *fhIA* and *pncB* in *Klebsiella* HQ-3 strain. Appl Biochem Biotechnol 178(6):1113–1128. <https://doi.org/10.1007/s12010-015-1932-4>
13. Jiang J, Tun HM, Mauroo NF, Ma AY, Chan SY, Leung FC (2014) Complete genome sequence and comparative genome analysis of *Klebsiella oxytoca* HKOPL1 isolated from giant panda feces. BMC Research Notes 7(1):827. <https://doi.org/10.1186/1756-0500-7-827>
14. Jenni JS, Antti L, Tommi A, Olli Y, Matti TK, Ville S (2013) Prospecting hydrogen production of *Escherichia coli* by metabolic network modeling. Int J Hydrogen Energ 38(27):11780–11789. <https://doi.org/10.1016/j.ijhydene.2013.07.002>
15. Jo BH, Cha HJ (2015) Activation of formate hydrogen-lyase via expression of uptake [NiFe]-hydrogenase in *Escherichia coli* BL21(DE3). Microb Cell Fact 14:151. <https://doi.org/10.1186/s12934-015-0343-0>
16. Jung MY, Ng CY, Song H, Lee J, Oh MK (2012) Deletion of lactate dehydrogenase in *Enterobacter aerogenes* to enhance 2,3-butanediol production. Appl Microbiol Biotechnol 95: 461–469. <https://doi.org/10.1007/s00253-012-3883-9>
17. Kalia VC, Lal S, Ghai R, Mandal M, Chauhan A (2003) Mining genomic databases to identify novel hydrogen producers. Trends Biotechnol 21(4):152–156. [https://doi.org/10.1016/S0167-7799\(03\)00028-3](https://doi.org/10.1016/S0167-7799(03)00028-3)

18. Kamaraj M, Ramachandran KK, Aravind J (2020) Biohydrogen production from waste materials: benefits and challenges. *Int J Environ Sci Technol* 17: 559–576. <https://doi.org/10.1007/s13762-019-02577-z>
19. Kamalaskar L, Kapse N, Pore S, Dhakephalkar AP, Ranade DR, Dhakephalkar PK (2016) Genome sequence and gene expression studies reveal novel hydrogenases mediated hydrogen production by *Clostridium biohydrogenum* sp. nov., MCM B-509^T. *Int J Hydrogen Energ* 41(28):11990–11999. <https://doi.org/10.1016/j.ijhydene.2016.05.101>
20. Khanna N, Ghosh AK, Huntemann M, Deshpande S, Han J, Chen A, Kyrpides N, Mavrommatis K, Szeto E, Markowitz V (2013) Complete genome sequence of *Enterobacter* sp. IIT-BT 08: A potential microbial strain for high rate hydrogen production. *Stand Genomic Sci* 9(2):359–369. <https://doi.org/10.4056/sigs.4348035>
21. Kumar G, Mathimani T, Rene ER, Pugazhendhi A (2019) Application of nanotechnology in dark fermentation for enhanced biohydrogen production using inorganic nanoparticles. *Int J Hydrogen Energ* 44(26):13106–13113. <https://doi.org/10.1016/j.ijhydene.2019.03.131>
22. Lee SH, Kim MS, Kang SG, Lee HS (2019) Biohydrogen production of obligate anaerobic archaeon *Thermococcus onnurineus* NA1 under oxic conditions via overexpression of *frhAGB*-encoding hydrogenase genes. *Biotechnol Biofuels* 12:24. <https://doi.org/10.1186/s13068-019-1365-3>
23. Li Y, Zhang Q, Deng L, Liu Z, Jiang H, Wang F (2018) Biohydrogen production from fermentation of cotton stalk hydrolysate by *Klebsiella* sp. WL1316 newly isolated from wild carp (*Cyprinus carpio* L.) of the Tarim River basin. *Appl Microbiol Biotechnol* 102:4231–4242. <https://doi.org/10.1007/s00253-018-8882-z>
24. Li Y, Zhu Q, Ding P, You S, Zhang Q, Jiang H (2020) Effects of Fe⁰ and Ni⁰ nanoparticles on hydrogen production from cotton stalk hydrolysate using *Klebsiella* sp. WL1316: Evaluation of size and concentration of the nanoparticles. *Int J Hydrogen Energ* 45:6243–6253. <https://doi.org/10.1016/j.ijhydene.2020.01.012>
25. Lu Y, Zhao H, Zhang C, Lai Q, Xing XH (2009) Perturbation of formate pathway for hydrogen production by expressions of formate hydrogen lyase and its transcriptional activator in wild *Enterobacter aerogenes* and its mutants. *Int J Hydrogen Energ* 34:5072 – 5079. <https://doi.org/10.1016/j.ijhydene.2009.04.025>
26. Lu Y, Zhao H, Zhang C, Xing XH (2016) Insights into the global regulation of anaerobic metabolism for improved biohydrogen production. *Bioresource Technol* 20035–20041. <https://doi.org/10.1016/j.biortech.2015.10.007>
27. Maeda T, Sanchez-Torres V, Wood TK (2007) Enhanced hydrogen production from glucose by metabolically engineered *Escherichia coli*. *Appl Microbiol Biotechnol* 77:879–890. <https://doi.org/10.1007/s00253-007-1217-0>
28. Mekanik M, Motamedian E, Fotovat R, Jafarian V (2019) Reconstruction of a genome-scale metabolic model for *Auxenochlorella protothecoides* to study hydrogen production under

- anaerobiosis using multiple optimal solutions. *Int J Hydrogen Energ* 44(5):2580–2591. <https://doi.org/10.1016/j.ijhydene.2018.12.049>
29. Rodionova MV, Poudyal RS, Tiwari I, Voloshin RA, Zharmukhamedov SK, Nam HG, Zayadan BK, Bruce BD, Hou HJM, Allakhverdiev SI (2017) Biofuel production: challenges and opportunities. *Int J Hydrogen Energ* 42: 8450–8461. <https://doi.org/10.1016/j.ijhydene.2016.11.125>
 30. Saratale GD, Oh MK (2015) Improving alkaline pretreatment method for preparation of whole rice waste biomass feedstock and bioethanol production. *RSC Adv* 5:91171–91179. <https://doi.org/10.1039/C5RA17797A>
 31. Shanmugam S, Haric A, Pandeyd A, Mathimanie T, Felix T, Pugazhendhi A (2020) Comprehensive review on the application of inorganic and organic nanoparticles for enhancing biohydrogen production. *Fuel* 270: 117453. <https://doi.org/10.1016/j.fuel.2020.117453>
 32. Sivagurunathan P, Kumar G, Mudhoo A, Rene ER, Saratale GD, Kobayashi T, Xu K, Kim SH, Kim DH (2017) Fermentative hydrogen production using biomass: an overview of pretreatment methods, inhibitor effects and detoxification experiences. *Renew Sustain Energy Rev* 77:28–42. <https://doi.org/10.1016/j.rser.2017.03.091>
 33. Soo CS, Yap WS, Hon WM, Jung JY, Park JM (2017) Improvement of hydrogen yield of ethanol-producing *Escherichia coli* recombinants in acidic conditions. *Electron J Biotechnol* 26:27–32. <https://doi.org/10.1166/jno.2011.1182>
 34. Tarailo-Graovac M, Chen N (2009) Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr Protoc Bioinform* 25(1): 4.10.1–4.10.14. <https://doi.org/10.1002/0471250953.bi0410s25>
 35. Tian H, Li J, Yan M, Tong YW, Wang CH, Wang X (2019) Organic waste to biohydrogen: A critical review from technological development and environmental impact analysis perspective. *Appl Energy* 256:113961. <https://doi.org/10.1016/j.apenergy.2019.113961>
 36. Thiele I, Palsson B (2010) A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature Protoc* 5(1):93–121. <https://doi.org/10.1038/nprot.2009.203>
 37. Tran KT, Maeda T, Wood TK (2014) Metabolic engineering of *Escherichia coli* to enhance hydrogen production from glycerol. *Appl Microbiol Biotechnol* 98:4757–4770. <https://doi.org/10.1007/s00253-014-5600-3>
 38. Wu XB, Huang GF, Bai LP, Long MN, Chen QX (2014) Enhanced hydrogen production from xylose and bamboo stalk hydrolysate by overexpression of xylulokinase and xylose isomerase in *Klebsiella oxytoca* HP1. *Int J Hydrogen Energ* 39: 221–230. <https://doi.org/10.1016/j.ijhydene.2013.10.078>
 39. Wu Y, Hao Y, Wei X, Shen Q, Ding X, Wang L, Zhao H, Lu Y (2017) Impairment of NADH dehydrogenase and regulation of anaerobic metabolism by the small RNA RyhB and NadE for improved biohydrogen production in *Enterobacter aerogenes*. *Biotechnol Biofuels* 10:248. <https://doi.org/10.1186/s13068-017-0938-2>
 40. Woo HL, Ballor NR, Hazen TC, Fortney JL, Simmons B, Davenport KW, Goodwin L, Ivanova N, Kyrpides NC, Mavromatis K, Woyke T, Jansson J, Kimbrel J, DeAngelis KM (2014) Complete genome

- sequence of the lignin-degrading bacterium *Klebsiella* sp. strain BRL6-2. Stand Genomic Sci 9:19. <https://doi.org/10.1186/1944-3277-9-19>
41. Xu N, Ye C, Chen X, Liu J, Liu L, Chen J (2016) Genome sequencing of the pyruvateproducing strain *Candida glabrata* CCTCC M202019 and genomic comparison with strain CBS138. Sci Rep 6:34893. <https://doi.org/10.1038/srep34893>
 42. Ye C, Qiao W, Yu X, Ji X, Huang H, Collier JL, Liu L (2015) Reconstruction and analysis of the genome-scale metabolic model of *Schizochytrium limacinum* SR21 for docosahexaenoic acid production. BMC Genome 16: 799. <https://doi.org/10.1186/s12864-015-2042-y>
 43. Yoshida A, Nishimura T, Kawaguchi H, Inui M, Yukaw H (2005) Enhanced hydrogen production from formic acid by formate hydrogen lyase-overexpressing *Escherichia coli* strains. Appl Environ Microb 71(11):6762–6768. <https://doi.org/10.1128/AEM.71.11.6762-6768.2005>
 44. Yu CS, Chen YC, Lu CH, Hwang JK (2006) Prediction of protein subcellular localization. Proteins: Struct Funct Bioinf 64(3): 643–651. <https://doi.org/10.1002/prot.21018>
 45. Zerbino DR, Birney E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18: 821–829. <https://doi.org/10.1101/gr.074492.107>
 46. Zhang Q, You S, Li Y, Qu X, Jiang H (2020) Enhanced biohydrogen production from cotton stalk hydrolysate of *Enterobacter cloacae* WL1318 by overexpression of the formate hydrogen lyase activator gene. Biotechnol Biofuels 13:94. <https://doi.org/10.1186/s13068-020-01733-9>
 47. Zhao HX, Ma K, Lu Y, Zhang C, Wang LY, Xing XH (2009) Cloning and knockout of formate hydrogenlyase and H₂-uptake hydrogenase genes in *Enterobacter aerogenes* for enhanced hydrogen production. Int J Hydrogen Energ 34: 186–194. <https://doi.org/10.1016/j.ijhydene.2008.10.025>

Figures

Figure 1

Schematic of pathways for glucose and xylose transport and metabolism in *Klebsiella* sp. WL1316

Figure 2

Hydrogenase- and formate-hydrogen lyase-related genes in *Klebsiella* sp.: a Direction and relative position of hydrogenase-related gene cluster; b Direction and relative position of formate-hydrogen lyase-related gene cluster; c Distributions of gene clusters of hydrogenase and formate-hydrogen lyase in the genome

Figure 3

a Heatmap and b statistical number of hydrogenase-relating functional proteins in *Klebsiella* sp. WL1316 and the five reference bacteria

Figure 4

The metabolic network for fermentative hydrogen production involving glucose and xylose utilisation, biohydrogen synthesis, and several competitive metabolic pathways

Figure 5

Metabolic analysis of glucose, xylose, and xylulose for biohydrogen production in *Klebsiella* sp. WL1316

Figure 6

Metabolic analysis of the key node metabolites for biohydrogen production process in *Klebsiella* sp. WL1316

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFile.docx](#)