2,5-Diketo-D-Gluconate Hyperproducing Gluconobacter sphaericus SJF2-1 with Reporting Multiple Genes Encoding the Membrane-Associated Flavoprotein-Cytochrome c Complexed Dehydrogenases

Haelim Son, Sang-Uk Han and Kyoung Lee *

Department of Bio Health Science, Changwon National University, Changwon 51140, Korea
* Correspondence: kyounglee@changwon.ac.kr; Tel.: +82-55-213-3486

Abstract: Gluconobacter sphaericus has not yet been used in biotransformation studies. In this study, G. sphaericus SJF2-1, which produces a diffusible pigment, was isolated from grape. The spent culture medium became dark black when the cells were grown in medium containing glucose and then autoclaved. This bacterium produced 2,5-diketo-D-gluconate (2,5-DKG) from D-glucose and D-gluconate. When 5% D-glucose was used, the conversion efficiency was approximately 52.4% in a flask culture. 2,5-DKG is a precursor of 2-keto-L-gulonic acid, which is a key intermediate in the industrial production of L-ascorbic acid. The complete genome sequence of G. sphaericus SJF2-1 was determined for the first time in the G. sphaericus species. The total size was 3,198,086 bp, with 2867 protein-coding sequences; one chromosome and six plasmids were identified. From the genome of SJF2-1, multiple genes homologous to those involved in the conversion of D-glucose to 2,5-DKG were identified. In particular, six different genes encoding membrane-associated flavoprotein-cytochrome c complexed dehydrogenase were identified and divided into two different lineages. This study suggests the potential of G. sphaericus SJF2-1 to mass-produce 2,5-DKG and other D-glucose oxidation products.

Keywords: Gluconobacter sphaericus; 2,5-diketo-D-gluconate; membrane-associated flavoprotein-cytochrome c complexed dehydrogenase; L-ascorbic acid; D-glucose biotransformation; D-gluconate

1. Introduction

L-ascorbic acid (vitamin C) is an essential nutrient for humans, acts as a coenzyme for collagen synthesis and intracellular energy metabolism, and has an antioxidant effect that removes radicals [1]. The global vitamin C market was valued at 1.55 billion USD in 2021 [2].

Currently, L-ascorbic acid is industrially manufactured in six steps, including two fermentation steps, by modifying the previous Reichstein–Grüssner process [3,4]. First, D-glucose is converted to D-sorbitol through a catalytic hydrogenation reaction, and then to L-sorbose using Gluconobacter oxydans in the first fermentation step. Next, the culture supernatant is pasteurized and used for the second fermentation process to produce the key intermediate for L-ascorbic acid synthesis, 2-keto-L-gulonic acid (2-KLG), using a mixed culture of Ketogulonicigenium vulgare and Bacillus megaterium or Bacillus cereus. The final chemical conversion stage includes the esterification and lactonization of purified 2-KLG to produce L-ascorbate (Figure 1). The production of 2-KLG has been challenged to improve the current process, including costly hydrogenation, long-term incubation, and additional sterilization steps. One of the challenges is a novel two-step fermentation process that involves the production of 2,5-diketo-D-gluconate (2,5-DKG) from D-glucose by biotransformation and subsequent additional fermentation with a bacterial strain that expresses regiospecific 2,5-DKG reductase to form 2-KLG (Figure 1). One example involves
the first fermentation using Erwinia sp. for 2,5-DKG production, and the fermentation step with Corynebacterium sp. that expresses NADPH-dependent 2,5-DKG reductase. This fermentation showed an 84.6% yield from D-glucose in 10-m³ conventional fermenters [5]. A similar example is the production of 2-KLG via a co-immobilized culture of G. oxydans ATCC 9937 and Corynebacterium sp. using D-gluconate as a starting substrate [6]. The two-stage process was further advanced by designing a single-strain one-step fermentation process using recombinant Erwinia herbicola, which expresses the 2,5-DKG reductase gene of Corynebacterium [7]. When supplied a total of 40 g/L of D-glucose to the culture medium of the recombinant strain, 49.4% of the D-glucose was converted to 2-KLG during 72-h fermentation [8]. For these one- and two-step processes to be industrially feasible, selecting a high-efficiency bacterial chassis for 2,5-DKG production is a key factor [9].

Some strains of Erwinia and Gluconobacter have been used in fermentation studies to produce 2,5-DKG from D-glucose [5,7,9,10]. The conversion of D-glucose to 2,5-DKG via 2-KG (2-keto-D-gluconate) by Gluconobacter strains is shown in Figure 1. The first three enzymes (E1–E3) are activated in the periplasmic space or on the periplasmic side of the cytoplasmic membrane to promote substrate accessibility and product diffusion from the cell. Through these reactions, the electrons form reduced ubiquinone (coenzyme Q) and are transferred to membrane-bound ubiquinol oxidase, where they reduce O₂ to form H₂O and generate energy [11]. In most cases, the activity of these enzymes is not coordinated, and intermediates, such as D-gluconate and 2-KG, often accumulate in the final culture medium [10,12]. This prevents the homogeneous production of 2,5-DKG. When using G. oxydans with a high concentration of glucose (10%), the conversion to 2,5-DKG of 66% for 120 h [10] and 21.6% for 60 h [9] have been reported. In the latter case, fed-batch fermentation was performed to increase the yield to 48% at 48 h. Tatumella citrea CICC 10802 (formerly Erwinia citrea) yielded 23.6 g/L with 10% glucose [9].

G. sphaericus NBRC 12467T, isolated from grapes, has been reported to produce 2,5-DKG [13,14]. However, the related genes and fermentation properties involved in biotransformation by this strain are not fully known, even though G. sphaericus is closely related to G. oxydans, which is widely used as a versatile biocatalyst in the manufacturing of L-ascorbic acid, dihydroxyacetone, miglitol, and gluconic acid and its derivatives [15,16]. Here, we present the possibility of using G. sphaericus SJF2-1 for the mass production of 2,5-DKG. Moreover, the first complete genome sequence of G. sphaericus in this study could further broaden the application range of this species for biotransformation.

2. Materials and Methods

2.1. Strain Isolation and Bacteria Culture

The spoiled Campbell’s Early grape collected at Sangju, Korea, was scrubbed on YPD agar (1% yeast extract, 2% peptone, 2% glucose, and 1.5% agar; KisanBio, Seoul, Korea), and a bacterial colony that produced a brown diffusive pigment was selected and desig-
nated SJF2-1. This strain has been deposited in the Korean Collection for Type Cultures (KCTC 82305). The strain was routinely cultured in YPS medium (1% yeast extract, 2% peptone, and 2% D-sorbitol), on which the pigment did not produce. API 50 CH (bioMérieux, Craponne, France) was used, according to the manufacturer’s protocol, to test the acid production from 49 carbohydrates. For species identification of the isolate, the 16S rRNA gene was amplified by PCR with primers 27F (5′-AGAGTTTGATCCTGGCTCAG-3′) and 1492R (5′-GGTTACCTTGTTACGACTT-3′), and the nucleotide sequence obtained by Sanger sequencing was used for homology searches in the NCBI database. PCR was performed using a TRIO thermal cycler (Biometra, Goettingen, Germany). The amplification conditions were as follows: (1) at 95 °C for 5 min; (2) 30 cycles of 95 °C for 30 s, 53 °C for 30 s, and 72 °C for 1 min; and (3) 72 °C for 5 min. Sanger sequencing was performed at GenoTech Co. (Daejeon, Korea). As a medium for 2,5-DKG production, D-glucose was separately added to YPP (1% yeast extract, 2% peptone, 0.3% K₂HPO₄, 0.25% CaCO₃, 0.05% MgSO₄, 0.1% (NH)₂SO₄ medium. An amount of 0.1 mL of seed from YPS medium cultured for 36 h was inoculated into a 50 mL medium contained in 250 mL-Erlenmeyer flasks and incubated in a rotary shaking incubator (140 rpm) at 28 °C. The experiment was performed in triplicate. pH was measured using a LAQUAtwin pH meter (Horiba, Kyoto, Japan). Cell growth was monitored by measuring absorbance at 600 nm using a spectrophotometer (nano-MD model, Sinco Co., Seoul, Korea). Strain *Fusarium moniliforme* var. *subglutinans* was obtained from the Korean Collection for Type Cultures (KCTC 6149).

### 2.2. Analyses of Metabolites

Thin layer chromatography (TLC) was performed on plastic-backed silica gel 60 F254 (0.1 mm thickness; Sigma-Aldrich) using a mobile phase consisting of ethyl acetate: acetic acid: methanol: deionized water (10:2:2:4, v/v). Sugar and sugar acids were detected by spraying a color-developing agent (a mixture of 1 g diphenylamine, 1 mL aniline, 50 mL acetone, 7.5 mL phosphoric acid) [17], followed by drying and heating at 120 °C for 5 min.

For HPLC measurements, a YL9100 HPLC system (Young Lin Instrument Co., Anyang, Korea) with a YL9170 refractive index (RI) detector was used. The ion exclusion column Rezex ROA Organic Acid H⁺ (8%, 300 × 7.8 mm; Phenomenex, Torrance, CA, USA) was operated with 0.005 N H₂SO₄ at 30 °C at a flow rate of 1 mL/min using an isocratic mode for 20 min. However, 2,5-DKG is not commercially available. Because the response of the RI detector is less sensitive to sugar isomers [18], a calibration curve of 2-KG was used to roughly quantify 2,5-DKG. The calibration curve was obtained by plotting the concentrations of 2-KG (0.2%, 0.6%, 0.8%, 1.0%, and 1.2%) versus the area of the respective peaks using the Microsoft Excel program (version 2019). The curve yielded a straight-line relationship, with a coefficient of determination (R²) of 0.998. Other products were quantified and compared to those of the standard compounds. Because the D-glucose and D-gluconate peaks overlapped on the ion exclusion column, D-glucose was detected using a ZORBAX carbohydrate column (4.6 mm × 150 mm; Agilent, USA) with an RI detector. The HPLC conditions were set with: mobile phase, 75% ACN in H₂O; column temperature, 35 °C; injection volume, 10 µL; and flow rate, 1 mL/min for 10 min.

Mass analysis of the metabolites was carried out using a Q Exactive Plus Quadrupole-Orbitrap Mass Spectrometer (Thermo Scientific, Waltham, MA, USA). The samples were then loaded via direct infusion. The electrospray ionization spray voltage was 2.5 kV in the negative ion mode. The ion transfer capillary temperature was 320 °C, the sheath gas N₂ pressure was 15 (arbitrary units), the auxiliary gas pressure was 5 (arbitrary units), and the sweep gas pressure was 0. Mass spectra were recorded in the m/z range of 75–1000.

### 2.3. Genome Sequencing and Analyses

The genomic DNA of SJF2-1 was sequenced using the Illumina NovaSeq6000 sequencer (Illumina Technologies, Hayward, CA, USA) and the PacBio RS II platform (PacBio Sciences, Menlo Park, CA, USA) at DNAlink Co. (Seoul, Korea). For Illumina sequencing, a TruSeq DNA PCR-free 550-bp library kit (Illumina) was used for library preparation, and
the read length was 2 × 151. The quality of the raw sequencing data was verified using the FastQC program (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) [19] accessed on 20 December 2020. A total of 28,220,814 reads with 8523 Mbp with a mean Phred quality score of 35.22 were generated with a GC content of 57.30%. For PacBio sequencing, a DNA library was prepared using SMRTbell templates with an insert size of 20 kb, as previously described [20]. Raw reads were filtered using a single-molecule real-time (SMRT) portal (version 2.3.0). After subread filtering, 107,194 reads and 8823 Mbp were obtained, with a GC content of 53.95%. The reads were assembled de novo using a hybrid Nanopore-Illumina program in Unicycler version 0.4.9b [21].

The quality of the assembled genome sequences was evaluated using CheckM v1.1.3 [22]. Gene predictions and annotations were provided by NCBI using the best-placed reference protein set and GeneMarkS-2+ of the NCBI Prokaryotic Genome Annotation Pipeline 6.1 [23]. The BLAST average nucleotide identity (ANIb) values between the SJF2-1 strain and type species in the database were retrieved from the JSpeciesWS server (https://jspecies.ribohost.com/jspeciesws/) [24] (accessed on 19-August-2022). Digital DNA–DNA hybridization (dDDH) values were calculated by applying the genome-to-genome distance calculator (GGDC 3.0) using formula 2 (https://ggdc.dsmz.de/ggdc.php (accessed on 19 August 2022)) [25].

Genome comparisons based on deduced amino acid sequences were conducted using Rapid Annotations using Subsystems Technology (RAST) [26]. The cumulative GC skew and GC content were depicted using the CGview program with sliding window sizes of 1000-bp and 100-bp steps [27]. The signal peptides and transmembrane domains of proteins were searched for using SignalP-5.0 [28] and TMHMM [29], respectively. The identities between the proteins were calculated by alignment of deduced amino acid sequences using the ClustalW Multiple alignment program (1000 bootstraps of NJ tree) and BLOSUM 62 in BioEdit Alignment Editor software version 7.0.5.3. Phylogenetic trees of proteins were constructed in MEGA_64 using Maximum likelihood analysis with the Jones–Taylor–Thornton model [30].

3. Results and Discussion
3.1. Identification of Strain SJF2-1

Strain SJF2-1 was isolated from spoiled Campbell Early grapes with a very sour taste, as described in Materials and Methods. During incubation on YPD agar for three days, a diffusible light brown pigment was formed by SJF2-1. SJF2-1 was Gram-negative and short-rod (0.8–1.2 µm × ~0.5 µm) under microscopy (Figure 2a,b).

![Figure 2](image_url)

**Figure 2.** Culture properties of SJF2-1. (a) Production of a diffusible light brown pigment by SJF2-1 grown on YPD agar. (b) Cells under a phase-contrast microscope. Bar, 1 µm. (c) A black-colored product formed by autoclaving the YPD culture solution of SJF2-1. (d) Growth inhibition of *Fusarium moniliforme* var. *subglutinans* (white mycelia) by SJF2-1 on YPD agar.

In the API 50 CH test, SJF2-1 produced acid from L-arabinose, D-xylose, D-galactose, D-glucose, D-mannose, D-melibiose, D-fucose, and potassium 5-keto-D-gluconate (5-KG). The nucleotide sequence of the 16S rRNA gene (1020 bp, NCBI accession number OP077118)
showed more than 99.9% similarity to those of *Gluconobacter cerevisiae*, *Gluconobacter albidus*, *Gluconobacter cadivus*, *Gluconobacter sphaericus*, and *Gluconobacter kondonii*. Interestingly, when the 3-d cultured YPD medium was autoclaved, the culture medium changed color to dark black (Figure 2c). The same color was formed even when the pH was adjusted to neutral before autoclaving. This phenomenon was also reported in *Pantoea citrea*, which is known to cause pink disease in pineapples [31]. In this study, the black color was caused by 2,5-DKG produced by *P. citrea*. Furthermore, SJF2-1 inhibited the growth of *Fusarium moniliforme* var. *subglutinans* (Figure 2d). This could be due, in part, to the pH drop to around 2, showing an example of a survival strategy of the 2,5-DKG-producing *G. sphaericus* in competition with the other species in sugar-rich environments.

3.2. Biotransformation of Glucose to 2,5-DKG and Metabolite Analyses

First, the bioconversion pathway and substrate preference involved in the production of 2,5-DKG by SJF2-1 were investigated by TLC and HPLC analyses of culture supernatants after 96 h of incubation. Cells were cultured in the YPP medium with 1% substrates of D-glucose, D-gluconate, 2-KG, and 5-KG.

In TLC analysis, standard chemicals were visualized using a color-developing spray agent with a spot loading of 40 µg, except for a ten times higher amount of D-gluconate for visible color development. In addition, there were differences in the color of each chemical. D-glucose, 2-KG, and 5-KG are blue; 2,5-DKG is deep blue; and D-gluconate is red (Figure 2a). The substrates D-glucose, D-gluconate, and 2-KG produced a single dominant product with the same color. The supernatant of D-glucose was analyzed by LC/MS, and the largest mass peak, with an m/z of 191.02 [M-H]⁺, expected for 2,5-DKG, was identified (Figure 2b). Thus, the products produced from D-gluconate and 2-KG were concluded to be 2,5-DKG. When 5-KG was used as the substrate, the color of the product on TLC was different from that of 2,5-DKG. Additionally, we found that the 2,5-DKG spot on TLC disappeared from the black-colored sample produced by autoclaving.

The results of HPLC with the Rezex ROA Organic Acid H+ column in a mix of triplicates are presented in Figure 3c. D-glucose (10 mg/mL) was completely consumed during 96 h of incubation, and one product peak in the HPLC chromatogram appeared at 4.50 min. We considered 2,5-DKG based on TLC and LC/MS results, as described above. The amount of 2,5-DKG was approximately 6.8 mg/mL (68% yield). Under the same incubation conditions, D-gluconate was consumed by 45.2% with an accumulation of 2-KG (1.8 mg/mL) and 2,5-DKG (2.2 mg/mL). 2-KG was less consumed, by 13.8%, yielding 0.07 mg/mL of 2,5-DKG. 5-KG was most consumed, by 89.2%. A product peak was observed at the same retention time as 2,5-DKG, but it was considered to be a different substance from 2,5-DKG by TLC analysis. Accordingly, D-glucose was the most effective substrate for the production of 2,5-DKG by SJF2-1, and D-gluconate was converted to 2,5-DKG through 2-KG, as presented in Figure 1. In contrast, 2-KG and 5-KG are not effective substrates for biotransformation. This result may reflect the differential expression of genes related to 2,5-DKG production by substrates.

The results of biotransformation using 5% D-glucose to investigate the fermentation pattern during the 96-h incubation period are shown in Figure 4. D-glucose, a substrate, was rapidly consumed between 24 and 48 h. The initial high OD₆₀₀ value was attributed to insoluble CaCO₃, and the OD₆₀₀ values during the first 48-h of incubation were a combination of CaCO₃ dissolution and increased cell density. D-gluconate, as an intermediate metabolite, accumulated maximally at 48 h, and was relatively slowly converted to 2,5-DKG. 2-KG was not accumulated. Finally, approximately 26.2 mg/mL of 2,5-DKG (52.4% yield) was produced. At the end of the incubation period, the pH decreased to 2.0, and the OD₆₀₀ value was 6.9. Notably, after 48 h, cell growth and biotransformation occurred, even at pH 2.5 or lower.
with 2867 protein-coding sequences, and one chromosome and six plasmids were identified. For strain SJF2-1, cultured in 50 mL of YPP + 5\% D-glucose medium as described in Materials and Methods. Sampling times are indicated. The results are from three independent experiments.

### 3.3. Genome Features of Strain SJF2-1 and Comparative Genome Analysis

Using Illumina and PacBio platforms, the nucleotide sequence of the complete genome of SJF2-1 was determined with 2870.0 \times genome coverage. The total size was 3,198,086 bp, with 2867 protein-coding sequences, and one chromosome and six plasmids were identified. The total GC content was 58.2\%. Analysis using CheckM showed 99.50\% completeness and 0\% strain heterogeneity. Table 1 shows the general characteristics of the SJF2-1 genome. In the JSpeciesWS server, the highest ANIb value was 97.95\% (88.32\% coverage) for \textit{G. sphaericus} NBRC 12467\textsuperscript{T}, followed by 89.69\% (73.36\% coverage) for \textit{G. oxydans} LMG1408\textsuperscript{T}. Additionally, the dDDH values (Formula 2) were 86.1\% and 41.7\%, respectively. These genome-based comparisons indicated that SJF2-1 belongs to the species \textit{G. sphaericus}, based on the species threshold values of ANIb and dDDH of 95\%–96\% and 70\%, respectively [32].

**Figure 3.** Analyses of D-glucose and sugar acid formed by \textit{Gluconobacter sphaericus} SJF2-1. (a) TLC analysis (1: D-glucose 40 \mu g, 2: D-gluconate 400 \mu g, 3: 2-KG 40 \mu g, 4: 5-KG 40 \mu g, 5: YPP + 1\% D-glucose, 6: YPP + 1\% D-gluconate, 7: YPP + 1\% 2-KG, 8: YPP + 1\% 5-KG); (b) LC/MS analysis of product formed by D-glucose; (c) HPLC chromatograms of standards in red and culture supernatants in black with indicated substrate.

**Figure 4.** Time course of D-glucose fermentation by \textit{Gluconobacter sphaericus} SJF2-1. Cells were cultured in 50 mL of YPP + 5\% D-glucose medium as described in Materials and Methods. Sampling times are indicated. The results are from three independent experiments.
There are five unpublished genome sequences of this species in the NCBI genome database (accessed on 30 September 2022), two (NBRC 12467\textsuperscript{T} and LMG1414\textsuperscript{T}) from the same type of strain, and three (Dm-14, Dm-21, and Dm-28) from the same source of apple dumpster by the same research group. Genome sequences were at the contig level. In this study, the complete genome of this species was identified for the first time, and our study is the first to analyze the genes of this species. The cumulative GC mol% and GC skew plots are shown in Figure 5.

### Table 1. General features of the \textit{Gluconobacter sphaericus} SJF2-1 genome.

| Genome Name | GenBank Accession | Genome Size (bp) | GC (%) | No. of Coding Sequence | No. of rRNA | No. of tRNA | No. of Gene | No. of Pseudogene |
|-------------|-------------------|-----------------|--------|------------------------|------------|------------|-------------|------------------|
| Chromosome  | CP068419.1        | 2,844,550       | 58.5   | 2518                   | 12         | 56         | 2693        | 103              |
| pSJF2-a     | CP068420.1        | 162,291         | 55.7   | 177                    | - *        | -          | 193         | 16               |
| pSJF2-b     | CP068421.1        | 109,107         | 55.2   | 102                    | -          | -          | 115         | 13               |
| pSJF2-c     | CP068422.1        | 44,465          | 53.0   | 31                     | -          | -          | 39          | 8                |
| pSJF2-d     | CP068423.1        | 22,858          | 49.3   | 19                     | -          | -          | 29          | 10               |
| pSJF2-e     | CP068424.1        | 9589            | 52.0   | 13                     | -          | -          | 13          | -                |
| pSJF2-f     | CP068425.1        | 5226            | 56.3   | 7                      | -          | -          | 7           | -                |

* Not present.

Figure 5. Genome comparison of \textit{Gluconobacter sphaericus} SJF2-1 (reference) with indicated \textit{Gluconobacter} strains conducted using RAST (outside tracks), and GC skew and GC content plots of SJF2-1 (inside). Tracks from the \textit{Gluconobacter} strains represent pairwise BLAST comparisons between the ORFs of each strain to the ORFs of the SJF2-1 genome, with percentages of identity represented with different colors shown in the legend.

In the RAST analysis, a comparison of orthologous proteins based on the predicted protein sequences between closely related strains showed that the genes in the genome of SJF2-1 exhibited high amino acid sequence similarities to those in NBRC 12467\textsuperscript{T} and...
Dm-14 (Figure 5). For instance, of 3294 predicted proteins in SJF2-1, 1002 and 1016 proteins represented 100% identical sequences to those in NBRC 12467\textsuperscript{T} and Dm-14, respectively. The latter two strains share 1492 proteins with identical sequences. In contrast, 243 proteins in SJF2-1 do not match those in strains NBRC 12467\textsuperscript{T} and Dm-14. For example, 43 out of 53 genes in the IGS75\_01300 to 1565 gene cluster were only present in SJF2-1 as a group of genes encoding hypothetical proteins with unknown domains (Figure 5). This gene cluster was also not found in the genome of \textit{G. oxydans}. The plasmids contained fewer homologous genes compared to chromosomes between other \textit{G. sphaericus} (Dm-14 and NBRC 12467\textsuperscript{T}) and \textit{G. oxydans} (621H and DSM 3504) strains (Figure 5).

At the genomic level, \textit{G. sphaericus} is most closely related to \textit{G. oxydans}; however, the genetic makeup and homologous gene sequence similarities between the two species differed significantly. For example, in RAST analysis, only 180 genes showed more than 99% homology with \textit{G. oxydans} 621H\textsuperscript{T} [33], and 764 genes were present only in SJF2-1. Thus, SJF2-1 possesses fermentative properties that are different from those of \textit{G. oxydans}, which are widely used in industrial applications. As a representative of this species, the genome of SJF2-1 was registered in the KEGG database (accession no. T07169), which will aid in metabolic studies of this species.

3.4. Genome Mining for Genes Related to Biotransformation of Glucose to 2,5-DKG

3.4.1. Genes Related to PQQ (Pyrroloquinoline Quinone)-Dependent Glucose Dehydrogenase

During the oxidation of glucose to 2,5-DKG, the oxidation of D-glucose to D-gluconate by \textit{Gluconobacter} is catalyzed by membrane-bound PQQ-dependent dehydrogenase, which first forms D-glucono-1,5-lactone from D-glucose [34]. This enzyme is a monomeric protein (approximately 87 kDa) that requires PQQ and a divalent cation, such as Mg\textsuperscript{2+}, for the reconstitution of the active enzyme from the apoenzyme [35]. Thus, the D-glucono-1,5-lactone formed is converted to D-gluconate either spontaneously or by a membrane-bound lactonase [36]. In \textit{G. oxydans} 621H, the GOX0265 gene (808 aa residues) encodes a PQQ-dependent glucose dehydrogenase [37]. As a result of searching for the homologous genes in the genome of SJF2-1, the IGS75\_01835 gene (808 aa residues) showed the most homologous sequence to the GOX0265 gene, showing 97.5% identity in the deduced amino acid sequence (Figure 6). Both proteins have five transmembrane domains in the N-terminal region. In the genome of SJF2-1, IGS75\_03370 was also found as a quinoprotein, with the same membrane topology and 37.4% amino acid sequence identity as GOX0265. These two enzymes belong to the dehydrogenases of the glucose/quinate/shikimate family.

Other genes encoding quinoprotein dehydrogenases include two methanol/ethanol family dehydrogenases, IGS75\_06000 (757 aa) and IGS75\_08895 (693 aa), and glycerol/sorbitol dehydrogenase IGS75\_05190 (741 aa) in the SJF2-1 genome. The genes IGS75\_06000 and IGS75\_08895 were accompanied by genes encoding small subunits. No transmembrane domains were identified in these three proteins. In SignalP 5.0 analysis, they possess a Sec/SPI signal peptide, indicating that the quinoprotein dehydrogenases encoded in these three genes have different topologies and membrane localizations from the two quinoprotein dehydrogenases mentioned above.

3.4.2. Genes Related to Membrane-Bound Gluconate 2-Dehydrogenase

D-gluconate is converted to 2-KG by membrane-bound gluconate 2-dehydrogenase. This enzyme has been purified and characterized from \textit{G. dioxyacetonicus} IFO 3271, and is composed of three subunits: the FAD covalently linked large (flavoprotein) subunit (ca. 64 kDa), the cytochrome c subunit (ca. 45 kDa), and the small subunit (ca. 21 kDa) [38]. \textit{G. oxydans} 621H is encoded by a gene cluster of gox1232 (small subunit, 236 aa)-gox1231 (flavoprotein subunit, 592 aa)-gox1230 (cytochrome c subunit, 437 aa) [39]. A search for homologous genes in the SJF2-1 genome based on the amino acid sequence of GOX1231 revealed that IGS75\_07935 was most homologous to GOX1232, at 79.3% identity, followed by IGS75\_00745, at 59.8% identity. These genes were included in the gene clus-
Microorganisms 2022, 10, x FOR PEER REVIEW 9 of 13

ters IGS75_07930 (210 aa)-IGS75_07935 (592 aa)-IGS75_07940 (442 aa) and IGS75_00750 (242 aa)-IGS75_00745 (592 aa)-IGS75_00740 (435 aa), respectively. The identities of the deduced amino acid sequences encoded by these genes compared with those of G. oxydans are shown in Figure 6. In SignalP 5.0 analysis, the Tat/SPI signal peptide was observed in small subunits (IGS75_07930 and IGS75_00750), and the Sec/SPI signal peptide was observed in cytochrome c subunits (IGS75_07940 and IGS75_00740). No signal sequence by SignalP 5.0 analysis was observed in the flavoprotein subunits (IGS75_07935 and IGS75_00745). No transmembrane domains were observed in any of the three subunits. The cytochrome c subunits were found to contain three heme c-binding CXXCH sequence motifs [40]. The large flavoprotein subunits contain GAGWAG and GFGWVG residues in the N-terminal region, corresponding to the FAD-binding motif, GXGXXG [41]. The subunits of gluconate-2-dehydrogenase in G. oxydans 621H were also found to possess the same signal peptides, cofactor-binding motifs, and no transmembrane domains.

![Figure 6. D-glucose oxidation pathway to 2,5-DKG and homologous genes encoding enzymes of the catalytic steps by SJF2-1. The genes in the arrows have been functionally demonstrated in different species, and the SJF2-1 genes in the boxes are homologous to the respective genes in the arrows. Gene clusters with the highest homology are shown in bold. Numbers in parentheses indicate the percentage of amino acid sequence identity compared to the genes in the arrows.](image-url)

In addition, IGS75_07930-IGS75_07935-IGS75_07940 showed 96.0%-99.5%-95.2% amino acid sequence identities with those encoded in the respective gndXYZ, which showed the highest activity among three isofunctional gluconate 2-dehydrogenases in G. oxydans NBRC 3293 [42]. The high amino acid sequence identities between the proteins in these gene clusters indicate that IGS75_07930-IGS75_07935-IGS75_07940 is most responsible for functioning as gluconate 2-dehydrogenase.
3.4.3. Genes Related to the Membrane-Bound 2-KG Dehydrogenase

The reaction of 2-KG to 2,5-DKG is catalyzed by 2-KG dehydrogenase, which consists of three membrane-bound subunits, similar to gluconate 2-dehydrogenase [38]. In the case of G. oxydans NBRC3293, the kgdABC (synonym kgdSLC) cluster encodes a small subunit (flavoprotein)-cytochrome c subunit with 222 aa, 552 aa, and 476 aa residues, respectively [12]. In the SJF2-1 genome, the IGS75_07910 (201 aa)-IGS75_07905 (552 aa)-IGS75_07900 (476 aa) gene cluster was most homologous to kgdABC. The identity of the amino acid sequence between each subunit is shown in Figure 6. In SignalP 5.0 analysis, three subunits from these two gene clusters showed the same results as the membrane-bound gluconate 2-dehydrogenase. However, cytochrome c subunits contain one transmembrane domain at the C-terminus, which is not observed in membrane-bound gluconate 2-dehydrogenase.

In the SJF2-1 strain, there were three more gene clusters homologous to the kgdABC genes, and the identity of the amino acid sequence of each subunit is shown in Figure 6. Usually, in the gene cluster IGS75_00810, the cytochrome c subunit gene precedes the small subunit gene. The gene clusters containing IGS75_13945 and IGS75_14610 were located in the pSJF2-a and pSJF2-b plasmids, respectively, which may have caused differences in expression due to the gene copy number.

3.4.4. Phylogeny of the Membrane-Bound Flavoprotein-Cytochrome c Complexed Dehydrogenases

Analysis of the phylogeny according to the amino acid sequence of the subunits comprising D-gluconate 2-dehydrogenase and 2-keto-D-gluconate dehydrogenase families showed that each subunit is clustered into a separate clade according to two enzyme branches, as shown in Figure 7. Subunits with the same function maintained comparable distances between gene clusters. Previous studies have also shown that large subunits can be phylogenetically distinguished according to their two enzyme branches [43]. Compared to the phylogenetic distances for flavoprotein and cytochrome c subunits, longer distances between smaller subunits were identified, which was reflected by the low sequence similarity between them. Thus, it was predicted that the small subunits would contribute more to the stabilization of the protein structure than to the catalyst function. To date, the protein structures of these enzymes have not yet been elucidated.

![Figure 7](image-url)

**Figure 7.** Rooted phylogenetic tree of the subunits of the membrane-associated flavoprotein-cytochrome c complexed dehydrogenases using the Maximum likelihood algorithm. Trees are scaled by the amino acid distance between proteins. Bootstrap values with 1000 rapid bootstrap replicates are shown. The protein IDs of the gene clusters are shown. (a) Small subunit; (b) flavoprotein (large subunit); (c) cytochrome c subunit. GDase and KGDase for gluconate 2-dehydrogenase and 2-keto-D-gluconate dehydrogenase families, respectively.
4. Conclusions

The genus *Gluconobacter*, belonging to the acetic acid bacteria, is capable of performing oxidative fermentation with various types of sugars or alcohols. They express multiple dehydrogenases with high activity in the periplasmic space, and intracellularly for oxidation. Owing to incomplete oxidation, the conversion from the substrate to the cell mass is relatively small, yielding a high accumulation of the product. Therefore, *G. oxydans* is used in the production of industrial chemicals.

There have been no studies on biotransformation using *G. sphaericus*. The yield of D-glucose biotransformation to 2,5-DKG by *G. sphaericus* SJF2-1 was approximately 52.4% with 5% D-glucose during 96-h fermentation in flask culture. The final product is relatively homogenous. Compared with other studies with different species, our results are considered to be one of the highest yields in batch fermentation. The biotransformation efficiency could be greatly improved in the fermenter under controlled culture conditions. In addition, by using recombinant strains overexpressing the rate-limiting membrane-associated gluconate 2-dehydrogenase and mutants lacking intracellular catabolism of the intermediates, the yield may be further enhanced. For more applications of SJF2-1, mutants defective in membrane-associated gluconate 2-dehydrogenase or 2-KG dehydrogenase could be used for the production of industrially valuable D-gluconate and 2-KG.

Author Contributions: Conceptualization, K.L.; methodology, H.S. and S.-U.H.; software, K.L. and H.S.; validation, H.S., S.-U.H. and K.L.; formal analysis, H.S., S.-U.H. and K.L.; investigation, K.L.; resources, K.L.; data curation, H.S. and K.L.; writing—original draft preparation, K.L.; writing—review and editing, K.L.; visualization, H.S. and K.L.; supervision, K.L.; project administration, K.L.; funding acquisition, K.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Changwon National University in 2021–2022.

Data Availability Statement: The assembled genome sequence of *G. sphaericus* SJF2-1 was deposited in GenBank under the accession numbers CP068419–CP068425 for chromosomes and plasmids. Raw sequence data used for assembly were deposited in GenBank under SRA accession numbers SRX9780842 (Illumina seq) and SRX9198931 (PacBio RSII).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sauberlich, H.E. Pharmacology of vitamin C. *Annu. Rev. Nutr.* 1994, 14, 371–391. [CrossRef] [PubMed]
2. Vitamin C Market Size to Worth Around USD 15.7 Billion by 2030. *Precedence Research* July 05, 2022. Available online: https://www.globenewswire.com/en/news-release/2022/07/05/2473878/0/en/Vitamin-C-Market-Size-to-Worth-Around-USD-15-7-Billion-by-2030.html (accessed on 20 December 2020).
3. Wang, P.; Zeng, W.; Xu, S.; Du, G.; Zhou, J.; Chen, J. Current challenges facing one-step production of L-ascorbic acid. *Biotechnol. Adv.* 2018, 36, 1882–1899. [CrossRef] [PubMed]
4. Pappenberger, G.; Hohmann, H.-P. Industrial production of L-ascorbic acid (vitamin C) and D-isoascorbic acid. In *Biotechnology of Food and Feed Additives*; Zorn, H., Czermak, P., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 143–188.
5. Sonoyama, T.; Tani, H.; Matsuda, K.; Kageyama, B.; Tanimoto, M.; Kobayashi, K.; Yagi, S.; Kyotani, H.; Mitsushima, K. Production of 2-keto-L-gulonic acid from D-glucose by two-stage fermentation. *Appl. Environ. Microbiol.* 1982, 43, 1064–1069. [CrossRef] [PubMed]
6. Aiguo, J.; Peiji, G. Synthesis of 2-keto-L-gulonic acid from gluconic acid by co-immobilized *Gluconobacter oxydans* and Corynebacterium sp. *Biotechnol. Lett.* 1998, 20, 939–942. [CrossRef]
7. Anderson, S.; Marks, C.B.; Lazarus, R.; Miller, J.; Stafford, K.; Seymour, J.; Light, D.; Rastetter, W.; Estell, D. Production of 2-keto-L-gulonate, an intermediate in L-ascorbate synthesis, by a genetically modified *Erwinia herbicola*. *Science* 1985, 230, 144–149. [CrossRef]
8. Grindley, J.; Payton, M.; Van de Pol, H.; Hardy, K. Conversion of glucose to 2-keto-L-gulonate, an intermediate in L-ascorbate synthesis, by a recombinant strain of *Erwinia citrius*. *Appl. Environ. Microbiol.* 1988, 54, 1770–1775. [CrossRef] [PubMed]
9. Li, G.; Shan, X.Y.; Zeng, W.; Yu, S.; Zhang, G.; Zhou, J. The efficient production of 2, 5-diketo-D-gluconic acid by reducing browning levels during *Gluconobacter oxydans* ATCC 9937 fermentation. *Front. Bioeng. Biotechnol.* 2022, 10, 918277. [CrossRef] [PubMed]
10. Qazi, G.; Parshad, R.; Verma, V.; Chopra, C.; Buse, R.; Träger, M.; Onken, U. Diketo-glucanate fermentation by *Gluconobacter oxydans*. *Enzyme Microb. Technol.* 1991, 13, 504–507. [CrossRef]

11. He, Y.; Xie, Z.; Zhang, H.; Liebl, W.; Toyama, H.; Chen, F. Oxidative fermentation of acetic acid bacteria and its products. *Front. Microbiol.* 2022, 13, 879246. [CrossRef]

12. Katoaka, N.; Matsutani, M.; Yakushi, T.; Matsushita, K. Efficient production of 2, 5-diketo-D-gluconate via heterologous expression of 2-ketogluconate dehydrogenase in *Gluconobacter japonicus*. *Appl. Environ. Microbiol.* 2015, 81, 3532–3560. [CrossRef]

13. Aneyama, M. *Gluconobacter oxydans* subsp. *sphaericus*, new subspecies isolated from grapes. *Int. J. Syst. Evol. Microbiol.* 1975, 25, 365–370.

14. Malimas, T.; Yukphan, P.; Takahashi, M.; Muramatsu, Y.; Kaneyasu, M.; Potacharoen, W.; Tanasupawat, S.; Nakagawa, Y.; Tanticharoen, M.; Yamada, Y. *Gluconobacter sphaericus* (Aneyama 1975) comb. nov. a brown pigment-producing acetic acid bacterium in the Alphaproteobacteria. *J. Gen. Appl. Microbiol.* 2008, 54, 211–220. [CrossRef] [PubMed]

15. Gupta, A.; Singh, V.K.; Qazi, G.; Kumar, A. *Gluconobacter oxydans*: Its biotechnological applications. *J. Mol. Microbiol. Biotechnol.* 2001, 3, 445–456.

16. da Silva, G.A.R.; Oliveira, S.S.d.S.; Lima, S.F.; do Nascimento, R.P.; Baptista, A.R.d.S.; Fiaux, S.B. The industrial versatility of *Gluconobacter* sp. strain JH63, isolated from human skin. *Microbiol. Resour. Announc.* 2020, 9, e00059–20. [CrossRef]

17. Bailey, R.; Bourne, E. Colour reactions given by sugars and diphenylamine-aniline spray reagents on paper chromatograms. *J. Chromatogr.* 1960, 4, 206–213. [CrossRef]

18. Zielinski, A.A.F.; Braga, C.M.; Demiate, I.M.; Beltrame, F.L.; Nogueira, A.; Wosiacki, G. Development and optimization of a HPLC-RF method for the determination of major sugars in apple juice and evaluation of the effect of the ripening stage. *Food Sci. Technol.* 2014, 34, 38–43. [CrossRef]

19. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data. 2010. Available online: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed on 20 December 2020).

20. Lee, K.; Badaya, S.K.; Singh, R.; Lim, J.Y. Complete genome sequence of *Gordonia* sp. strain JH63, isolated from human skin. *Microbiol. Resour. Announc.*: 2020, 9, e00059–20. [CrossRef]

21. Wick, R.R.; Judd, L.M.; Gorrie, C.L.; Holt, K.E. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PloS Comput. Biol.* 2017, 13, e1005595. [CrossRef]

22. Parks, D.H.; Imelfort, M.; Skennerton, C.T.; Hugenholtz, P.; Tyson, G.W. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 2015, 25, 1043–1055. [CrossRef]

23. Tatusova, T.; DiCuccio, M.; Badretdin, A.; Chetvernin, V.; Nawrocki, E.P.; Lomsadze, A.; Pruitt, K.D.; Borodovsky, M.; Ostertag, J. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* 2016, 44, 6614–6624. [CrossRef]

24. Richter, M.; Rossello-Mora, R.; Oliver Glockner, F.; Peplies, J. JSpeciesWS: A web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* 2016, 32, 929–931. [CrossRef]

25. Meier-Kolthoff, J.P.; Carbasse, J.S.; Peinado-Ortale, R.L.; Göker, M. TYGS and LPSN: A database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res.* 2021, 50, D801–D807. [CrossRef]

26. Aziz, R.K.; Bartels, D.; Best, A.A.; DeLong, M.; Disz, T.; Edwards, R.A.; Formosa, A.; Glass, E.M.; Kubal, M. The RAST server: Rapid annotations using subsystems technology. *BMC Genom.* 2008, 9, 75. [CrossRef]

27. Grant, J.R.; Stothard, P. The CGView Server: A comparative genomics tool for circular genomes. *Nucleic Acids Res.* 2008, 36, W181–W184. [CrossRef]

28. Almagro Armenteros, J.J.; Tsirigos, K.D.; Sønderby, C.K.; Petersen, T.N.; Winther, O.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotech.* 2019, 37, 420–423. [CrossRef]

29. Krogh, A.; Larsson, B.; Von Heijne, G.; Sonnhammer, E.L. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol.* 2001, 305, 567–580. [CrossRef]

30. Kumar, S.; Tamura, K.; Nei, M. MEGA: Molecular evolutionary genetics analysis software for microcomputers. *Bioinformatics* 1994, 10, 189–191. [CrossRef]

31. Pujol, C.J.; Kado, C.I. Genetic and biochemical characterization of the pathway in *Pantoea citrea* leading to pink disease of pineapple. *J. Bacteriol.* 2000, 182, 2230–2237. [CrossRef]

32. Chun, J.; Oren, A.; Ventosa, A.; Christensen, H.; Arahal, D.R.; da Costa, M.S.; Rooney, A.P.; Yi, H.; Xu, X.W.; De Meyer, S.; et al. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 2018, 68, 461–466. [CrossRef]

33. Prust, C.; Hofmeister, M.; Liesegang, H.; Pieczek, A.; Fricke, W.F.; Ehrenreich, A.; Gottschalk, G.; Deppenmeier, U. Complete genome sequence of the acetic acid bacterium *Gluconobacter oxydans*. *Nat. Biotech.* 2005, 23, 195–200. [CrossRef]

34. Pronk, J.; Levering, P.; Oljive, W.; Van Dijken, J. Role of NADP-dependent and quinoprotein glucose dehydrogenases in gluconic acid production by *Gluconobacter oxydans*. *Enzyme Microb. Technol.* 1989, 11, 160–164. [CrossRef]

35. Anthony, C.; Ghosh, M. The structure and function of the PQQ-containing quinoprotein dehydrogenases. *Prog. Biophys. Mol. Biol.* 1998, 69, 1–21. [CrossRef]

36. Shinagawa, E.; Aso, Y.; Yakushi, T.; Adachi, O.; Matsushita, K. Solubilization, purification, and properties of membrane-bound D-glucono-δ-lactone hydrdrolase from *Gluconobacter oxydans*. *Biosci. Biotechnol. Biochem.* 2009, 73, 241–244. [CrossRef]

37. Meyer, M.; Schweiger, P.; Deppenmeier, U. Effects of membrane-bound glucose dehydrogenase overproduction on the respiratory chain of *Gluconobacter oxydans*. *Appl. Microbiol. Biotech.* 2013, 97, 3457–3466. [CrossRef]
38. Shinagawa, E.; Matsushita, K.; Adachi, O.; Ameyama, M. D-Gluconate dehydrogenase, 2-keto-D-gluconate yielding, from *Gluconobacter dioxyacetonicus*: Purification and characterization. *Agric. Biol. Chem.* **1984**, *48*, 1517–1522. [CrossRef]

39. Elfari, M.; Ha, S.-W.; Bremus, C.; Merfort, M.; Khodaverdi, V.; Herrmann, U.; Sahm, H.; Görisch, H. A *Gluconobacter oxydans* mutant converting glucose almost quantitatively to 5-keto-D-gluconic acid. *Appl. Microbiol. Biotech.* **2005**, *66*, 668–674. [CrossRef]

40. Barker, P.D.; Ferguson, S.J. Still a puzzle: Why is haem covalently attached in c-type cytochromes? *Structure* **1999**, *7*, R281–R290. [CrossRef]

41. Dym, O.; Eisenberg, D. Sequence-structure analysis of FAD-containing proteins. *Protein Sci.* **2001**, *10*, 1712–1728. [CrossRef]

42. Kataoka, N.; Saichana, N.; Matsutani, M.; Toyama, H.; Matsushita, K.; Yakushi, T. Characterization of 3 phylogenetically distinct membrane-bound D-gluconate dehydrogenases of *Gluconobacter* spp. and their biotechnological application for efficient 2-keto-D-gluconate production. *Biosci. Biotechnol. Biochem.* **2022**, *86*, 681–690. [CrossRef]

43. Qin, Z.; Yu, S.; Chen, J.; Zhou, J. Dehydrogenases of acetic acid bacteria. *Biotechnol. Adv.* **2022**, *54*, 107863. [CrossRef]