

Article

Not peer-reviewed version

---

# Complete Genome Sequence and Comparative Genomics of *Acetobacter cerevisiae* KSO5 (KACC 92352P) Provide Genome-Based Insights into Acid Tolerance

---

[Sun Hee Kim](#) , [Dong Min Han](#) , [Seong-Eui Yoo](#) , [Jin Ju Park](#) , [Chan Woo Kim](#) , [So-Young Kim](#) \*

Posted Date: 16 March 2026

doi: 10.20944/preprints202603.1090.v1

Keywords: *Acetobacter cerevisiae*; complete genome; oxidative fermentation; ethanol stress; vinegar fermentation



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

## Article

# Complete Genome Sequence and Comparative Genomics of *Acetobacter cerevisiae* KSO5 (KACC 92352P) Provide Genome-Based Insights into Acid Tolerance

Sun Hee Kim <sup>1</sup>, Dong Min Han <sup>2</sup>, Seong-Eui Yoo <sup>1</sup>, Jin Ju Park <sup>1</sup>, Chan Woo Kim <sup>1</sup> and So-Young Kim <sup>1,\*</sup>

<sup>1</sup> Fermented and Processed Food Research Division, Department of Food Sciences, NICS, RDA, Wanju, 55365, Korea

<sup>2</sup> Department of Life Science, Chung-Ang University, Seoul 06974, Republic of Korea

\* Correspondence: Correspondence: foodksy@korea.kr (S.-Y.K.); Tel.: +82-63-238-3610; Fax.: +82-63-238-3843

## Abstract

We report the first complete circular genome of *Acetobacter cerevisiae* KSO5, an indigenous strain isolated from Korean fruit vinegar, comprising a 3.3 Mb chromosome and two plasmids encoding 2,898 genes. Phylogenomics confirmed species assignment (average nucleotide identity, ANI 97%; digital DNA–DNA hybridization, dDDH 71%). Comparison with seven draft *A. cerevisiae* genomes revealed strain-specific genomic islands, mobile genetic elements and polymorphisms in stress-response pathways, with enrichment in acid-tolerance-associated functions, and highlighted plasmid-borne modules potentially linked to genetic stability. The genome encodes a periplasmic oxidative fermentation system with membrane-bound pyrroloquinoline quinone-dependent alcohol dehydrogenase (PQQ-ADH) and molybdopterin-dependent aldehyde dehydrogenase (Mo-ALDH), together with respiratory-chain components consistent with flexible aerobic metabolism. Three acetate-handling routes (efflux, acetyl-CoA conversion and an AarC branch) were also predicted, suggesting mechanisms to limit intracellular acetate accumulation. Consistent with these features, phenotyping under ethanol stress (5–10%) showed measurable growth and titratable acidity production up to 9% ethanol (late-stage peak acidity). These data provide a genomic and phenotypic basis for developing robust vinegar starter cultures.

**Keywords:** *Acetobacter cerevisiae*; complete genome; oxidative fermentation; ethanol stress; vinegar fermentation

## 1. Introduction

Vinegar production depends on the remarkable physiology of acetic acid bacteria (AAB), which oxidize reduced substrates in the periplasm and channel electrons to the respiratory chain while tolerating combined stresses of ethanol, oxygen gradients, and high acetic-acid levels [1]. Central to this lifestyle are membrane-bound dehydrogenase systems—most prominently a pyrroloquinoline quinone (PQQ)-dependent alcohol dehydrogenase (PQQ-ADH) and, typically, a membrane-bound molybdopterin (Mo)-dependent acetaldehyde dehydrogenase (Mo-ALDH; AldFGH-type)—that catalyze the two-step conversion of ethanol to acetate at the periplasmic face of the inner membrane, routing reducing equivalents into the ubiquinone pool and largely decoupling carbon oxidation from cytosolic NAD(H) balance, thereby enabling rapid, surface-confined oxidative fermentation [1–3]. The resulting electron flux is dissipated via branched terminal oxidases (e.g.,  $\text{bo}_3^-$  and bd-type

ubiquinol oxidases) and auxiliary respiratory modules, together sustaining ATP generation and redox homeostasis under acidic and microaerobic conditions typical of vinegar fermentations [4].

While *Acetobacter cerevisiae* has been recurrently isolated from food fermentations, a closed, reference-quality genome sequence has not been reported for this species in the literature, and publicly available resources have largely consisted of draft assemblies. Accordingly, comparative inferences have often drawn on related acetic acid bacteria such as *Gluconobacter oxydans*, which lacks respiratory complex I and exhibits a distinct respiratory architecture [5].

To address this gap, we generated a complete, closed circular genome for *A. cerevisiae* from a fruit-vinegar isolate (strain KSO5) and used it as a reference framework for species-focused comparative analysis. Leveraging high-confidence structural and functional annotation, we integrate comparative genomics with seven available *A. cerevisiae* draft genomes to (i) delineate strain-unique genomic islands and other mobile genetic elements, (ii) map single-nucleotide polymorphisms across conserved stress-response circuitry with emphasis on acid-tolerance pathways, and (iii) resolve plasmid-associated protein modules that may contribute to genetic stability and fermentation-relevant traits. To further connect genotype with phenotype, we also performed ethanol challenge experiments across a 5–10% ethanol range to evaluate strain performance under increasing ethanol stress. Together, these analyses provide a genome-enabled, systems-level view of *A. cerevisiae* physiology, linking respiratory architecture, carbohydrate oxidation, and stress tolerance to genotype variation, and establishing a foundation for hypothesis-driven functional studies of fermentation-relevant traits.

## 2. Materials and Methods

### 2.1. Bacterial Isolation

Samples were collected from farm-produced omija (magnolia berry, *Schisandra chinensis*) fruit vinegar in Gyeonggi Province, South Korea. A 100–200  $\mu$ L aliquot of the vinegar sample was spread onto YGC agar plates (5 g/L yeast extract, 30 g/L glucose, 10 g/L  $\text{CaCO}_3$ , 4% (v/v) ethanol, 2% (w/v) agar) and incubated at 30°C for 3 days under aerobic conditions. Representative colonies, characterized by a clear halo indicating the dissolution of  $\text{CaCO}_3$  in the medium due to acetic acid production, were selected from the plates. These colonies were further purified using fresh YGC agar plates and stored at -80°C in YG broth containing 80% glycerol for subsequent analysis.

### 2.2. Phenotyping Under Ethanol Stress

AAB strains were cultured aerobically (250 mL flasks containing 50 mL medium; inoculum standardized to  $\text{OD}_{660}$  0.5 and inoculated 1:1000, v/v) in yeast extract–glucose medium containing 1% (v/v) acetic acid and ethanol (5–10%, v/v) at 30 °C and 150 rpm for 10 days. Growth ( $\text{OD}_{660}$ ) and titratable acidity were measured on days 1, 4, 6, 8 and 10. The  $\text{OD}_{660}$  and titratable acidity time-course measurements were obtained from a single culture per condition; detailed procedures are provided in Supplementary Methods.

### 2.3. Characterization of Bacterial Isolates

The morphological characteristics of the bacterial isolates were observed using a Leica optical DM500 or stereo microscope EZ4 (Leica Microsystems, Wetzlar, Germany) and a ZEISS Gemini Scanning electron microscope (SEM) 300 (ZEISS, Jena, Germany) after culturing the AAB on YGC solid medium at 30°C for three days. The Gram reaction was determined using the standard Gram staining method. To further assess the physiological, biochemical, and enzymatic activities of the isolates, experiments were conducted using the API ZYM test kit (25200, BioMérieux, Marcy-l'Étoile, France) according to the manufacturer's instruction.

2.4. Phylogenetic Analysis

The 16S rRNA gene was amplified using primers 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-GGTTACCTTGTTACGACTT-3') and the amplicons were sequenced. Taxonomic assignment used BLASTn against the NCBI database. Phylogeny was inferred in MEGA v6 (megasoftware.net) by the neighbor-joining method with bootstrap resampling.

2.5. Genome Sequencing and Assembly

Genomic DNA from the KSO5 strain was extracted using the QIAamp DNA Mini Kit (Qiagen, Valencia, CA, USA) following the manufacturer's instructions. Whole-genome sequencing and analyses were performed by Macrogen (Seoul, Republic of Korea) in accordance with minimal standards for prokaryotic taxonomy. PacBio sequel subreads were assembled with the microbial assembly pipeline in SMRT Link v13.0.0.207600 (HGAP) using default parameters. For polishing, illumina paired-end raw reads were quality-filtered under Macrogen's internal QC criterion designed to maximize the precision of read-based error correction: only reads in which ≥90% of bases had a Phred quality score ≥30 (Q30; ~99.9% base-call accuracy) were retained, and reads failing this threshold were excluded to avoid spurious corrections during Pilon polishing. The assembly was then polished with Pilon v1.21 using the retained high-quality reads [6]. Gene prediction and primary annotation were performed with Prokka v1.14.6 (--compliant, --rnammer, --addgenes) [7], and protein-coding genes were functionally classified using the Clusters of Orthologous Groups (COGs) database. Assembly quality was documented by contiguity and completeness indicators: the final assembly deposited at NCBI under ASM4409463v1 (=GCA\_044094635.1) comprises one circular chromosome (CP172014.1; 3,257,599 bp) and two small circular plasmids (CP172015.1, 4,905 bp; CP172016.1, 4,820 bp), yielding a contig N50 of 3,257,599 bp and an L50 of 1; NCBI lists the assembly level as "Complete genome."

2.6. Comparative Genomics Analysis

Comparative genomics was performed at two tiers. Tier-1 (AAB-wide): To minimise assembly bias in genus-wide analyses, we included only complete genomes from *Acetobacter*, *Gluconobacter*, *Gluconacetobacter* and *Komagataeibacter* (n = 39 including KSO5; Table 1). Tier-2 (species-level, *A. cerevisiae*): For within-species comparisons we analyzed eight *A. cerevisiae* genomes (Table 2), comprising KSO5 (complete) and seven publicly available draft assemblies.

ANI (pyANI v0.2.12; ANIm) and pan-genome analysis (Prokka v1.14.6; Roary v3.13.0, -i 90, -e, -n) [8] were run separately for the two tiers; core-gene alignments informed FastTree v2.1.11 phylogenies [9], with visualisation in iTOL v5 [10]. NCBI Assembly (11 Feb 2025) entries at "Complete Genome" level from *Acetobacter*, *Gluconobacter*, *Gluconacetobacter*, and *Komagataeibacter* were retained; KSO5 (ASM4409463v1) was compared with 38 complete assemblies (n=39; Table 1). Average nucleotide identity (ANI) was computed with pyANI v0.2.12 (ANIm/MUMmer), with heatmaps from the pairwise matrix; species relatedness was interpreted at ANI ≥95–96% and dDDH ≥70%, and near outgroups from non-*Acetobacter* genera were included to stabilise rooting and test KSO5 assignment.

Table 1. Genome sequences used in the study.

| No. | Name                               | Length (bp) | Accession No. |
|-----|------------------------------------|-------------|---------------|
| 1   | <i>Acetobacter cerevisiae</i> KSO5 | 3,257,599   | CP172014      |
| 2   | <i>A. pasteurianus</i> 386B        | 2,818,679   | NC_021991.1   |
| 3   | <i>A. pasteurianus</i> CICC 22518  | 2,772,347   | NZ_CP39846.2  |
| 4   | <i>A. pasteurianus</i> SRCM101468  | 2,996,610   | NZ_CP021922.1 |
| 5   | <i>A. pasteurianus</i> SRCM101342  | 2,754,755   | NZ_CP021509.1 |
| 6   | <i>A. pasteurianus</i> NBRC 101655 | 2,902,389   | AP014881.1    |

|    |                                                         |           |                 |
|----|---------------------------------------------------------|-----------|-----------------|
| 7  | <i>A. oryzifermentans</i> DM                            | 3,127,455 | NZ_CP022374.1   |
| 8  | <i>A. oryzifermentans</i> SLV-7                         | 2,799,488 | NZ_CP011120.1   |
| 9  | <i>A. ascendens</i> SRCM101447                          | 2,901,846 | NZ_CP021524.1   |
| 10 | <i>A. persici</i> TMW2.1084                             | 3,230,507 | NZ_CP014687.1   |
| 11 | <i>A. orientalis</i> FAN1                               | 3,041,114 | AP018515.1      |
| 12 | <i>A. senegalensis</i> 108B                             | 3,889,881 | NZ_LN606600.1   |
| 13 | <i>A. tropicalis</i> BDGP1                              | 3,988,649 | NZ_CP022699.1   |
| 14 | <i>A. acetii</i> NBRC 14818                             | 3,596,270 | NZ_AP023410.1   |
| 15 | <i>A. acetii</i> JCM20276                               | 3,743,357 | NZ_AP023326.1   |
| 16 | <i>A. acetii</i> TMW2.1153                              | 3,725,037 | NZ_CP014692.1   |
| 17 | <i>Gluconacetobacter diazotrophicus</i> PA1 5           | 3,887,492 | NC_011365.1     |
| 18 | <i>Glu. diazotrophicus</i> PA1 5                        | 3,944,163 | NC_010125.1     |
| 19 | <i>Komagataeibacter medellinensis</i> NBRC 3288         | 3,136,818 | NC_016027.1     |
| 20 | <i>Kom. xylinus</i> DSM 2325                            | 3,353,346 | NZ_CP025269.1   |
| 21 | <i>Kom. xylinus</i> CGMCC 17276                         | 3,527,401 | NZ_CP041348.1   |
| 22 | <i>Kom. xylinus</i> CGMCC 2955                          | 3,563,314 | CP024644.1      |
| 23 | <i>Kom. xylinus</i> E25                                 | 3,447,725 | CP004360.1      |
| 24 | <i>Gluconobacter oxydans</i> 621H                       | 2,704,625 | NZ_LT900338.1   |
| 25 | <i>G. oxydans</i> 621H                                  | 2,702,173 | NC_006677.1     |
| 26 | <i>A. pasteurianus</i> GHA7                             | 2,927,634 | CP157844        |
| 27 | <i>A. syzygii</i> 9H-2                                  | 2,672,115 | GCA_000964225   |
| 28 | <i>A. pomorum</i> LHT 2458                              | 3,308,689 | GCA_002738225   |
| 29 | <i>A. oryzoeni</i> B6 <sup>T</sup>                      | 3,153,180 | GCF_004014775   |
| 30 | <i>A. pomorum</i> DSM 11825                             | 3,319,623 | GCF_025995455.1 |
| 31 | <i>A. ghanensis</i> LMG 23848                           | 2,843,474 | GCA_001499675   |
| 32 | <i>A. pasteurianus</i> LMG 1262                         | 2,982,262 | GCA_000285275   |
| 33 | <i>A. pasteurianus</i> subsp. <i>paradoxus</i> LMG 1591 | 3,216,032 | GCA_001766255   |
| 34 | <i>A. pasteurianus</i> subsp. <i>ascendens</i> LMG 1590 | 2,999,217 | GCA_001766235   |
| 35 | <i>A. okinawensis</i> JCM 25146                         | 3,166,244 | GCA_000613865   |
| 36 | <i>A. orleanensis</i> LMG 1583                          | 3,007,844 | GCF_001581005   |
| 37 | <i>A. malorum</i> LMG 1746                              | 3,833,476 | GCF_001580615   |
| 38 | <i>A. cerevisiae</i> LMG 1625                           | 3,088,073 | GCF_001580535   |
| 39 | <i>A. vaccinii</i> KACC 21233                           | 3,082,251 | GCA_008365315   |

Table 2. Genome information for *A. cerevisiae* registered in NCBI.

| No. | Strain name | Source                         | Assembly     | RefSeq          | level           | Scaffolds |
|-----|-------------|--------------------------------|--------------|-----------------|-----------------|-----------|
| 1   | KSO5        | Fruit vinegar                  | ASM4409463v1 | GCF_44094635.1  | Complete Genome | 3         |
| 2   | LMG 1625    | Beer                           | ASM158053v1  | GCF_001580535.1 | Contig          | 157       |
| 3   | DSM 14362   | Tokyo University <sup>1)</sup> | ASM2599619v1 | GCF_025996195.1 | Contig          | 218       |

|   |          |              |              |                 |        |     |
|---|----------|--------------|--------------|-----------------|--------|-----|
| 4 | R-83281  | Lambic beer  | ASM2415826v1 | GCF_024158265.1 | Contig | 128 |
| 5 | R-82823  | Lambic beer  | ASM2415830v1 | GCF_024158305.1 | Contig | 137 |
| 6 | R-82821  | Lambic beer  | ASM2415828v1 | GCF_024158285.1 | Contig | 144 |
| 7 | R-82820  | Lambic beer  | ASM2415831   | GCF_024158315.1 | Contig | 145 |
| 8 | LMG 1545 | Rice vinegar | ASM158110v1  | GCF_001581105.1 | Contig | 108 |
| 9 | LMG1608  | Beer         | ASM158107v1  | GCF_001581075.1 | Contig | 177 |

<sup>1</sup>This strain does not specify the isolation source and as DSM 14362 = LMG 1625 (type strain), analyses used LMG 1625.

2.7. Network Visualization of Plasmid-Associated Functional Modules in *A. cerevisiae* Strains

The plasmid module network was visualized in Cytoscape v3.10.2 using the Prefuse Force Directed layout and was slightly adjusted manually to improve high-resolution rendering. Functional modules were identified based on RefSeq annotations and domain assignments (e.g., COG, Pfam) from the plasmid-associated proteins in each genome (Table 3). The network layout was generated using the spring layout algorithm in networkx, which places nodes in a force-directed layout to minimize edge crossings and enhance interpretability. Node colors were assigned to differentiate strain identifiers (blue) from plasmid functional modules (orange).

Table 3. Plasmid-associated protein modules and genetic implications in *A. cerevisiae* strains.

| Strain            | Key Plasmid-Associated Protein Modules                            | Genetic Implication                                                                                                     |
|-------------------|-------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------|
| KSO5              | RepA (CyRepA1 family), CRISPR-associated primase-polymerase       | Autonomous plasmid replication; potential integration of CRISPR-mediated defense and recombination modules (This study) |
| LMG 1625          | Two RepA variants, CcdB toxin, HigB toxin, stabilization proteins | Toxin–antitoxin (TA) system–based plasmid stabilization (This study)                                                    |
| LMG 1545          | RepB, MobA, RepC, MobC                                            | Complete modules for plasmid partitioning, mobilization, and replication (This study)                                   |
| LMG 1608          | RepB, MobA                                                        | Focused on distribution and transfer rather than autonomous replication (This study)                                    |
| R-82820 / R-82821 | RepC, MobC, stabilization proteins                                | Minimal replication and mobilization module architecture (This study)                                                   |
| R-82823 / R-83281 | RepC, RepB, MobA                                                  | Canonical RepABC system combining replication initiation, partitioning, and transfer (This study)                       |

2.8. SNP Comparison of Proteins Involved in Acetic-Acid Resilience Mechanisms Across Eight *A. cerevisiae* Strains

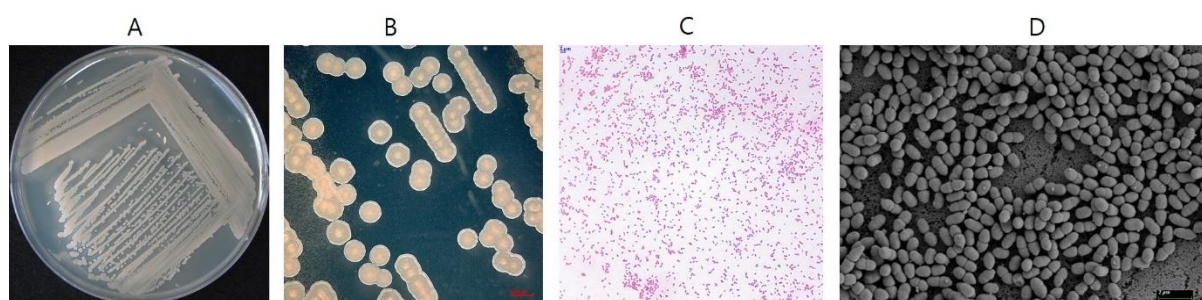
Orthologous coding sequences (CDSs) were aligned across eight *A. cerevisiae* strains, using the KSO5 allele as the reference, and SNPs were called from the resulting multiple sequence alignments (MSAs). Variant and codon coordinates were lifted over from KSO5 to the type strain LMG1625 to ensure coordinate comparability. Alignment quality was assessed by BLAST-based checks

(alignment length, mismatches, and gap openings), and SNP counts were verified to be consistent with BLAST mismatch counts. To represent each acetic-acid resilience layer, 1–5 sentinel genes were selected per mechanism (Supplementary Table 1). For enzymatic acetate metabolism (assimilation), Acs/AcsA, AckA, Pta, AarC, Mqo, and AcnA were included. For efflux/transport, OprM\_1 was selected. For chaperones/stress response, GroEL and DnaK were included. For ROS detoxification, SrpA was selected. The sentinel panel was analyzed in depth by quantifying per-codon amino-acid impacts (synonymous, nonsynonymous, nonsense, and stop-loss), mapping variants onto conserved domains, and extracting key variants. In parallel, the same metrics were applied in a genome-wide auxiliary screen to non-sentinel genes to capture broader patterns.

### 3. Results

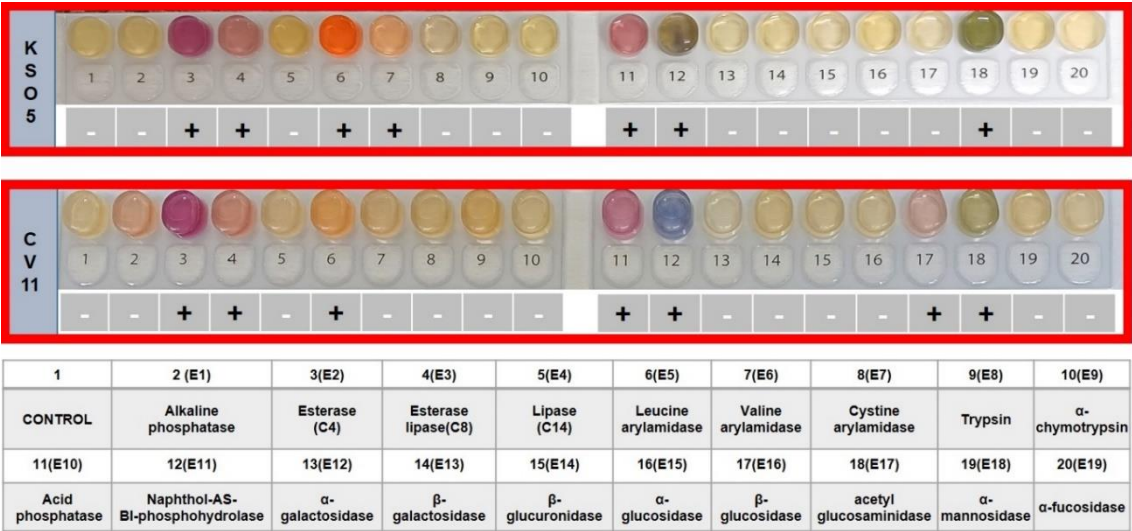
#### 3.1. Morphology and Physiology of *A. cerevisiae* KSO5

An acetic acid bacteria (AAB) strain, *A. cerevisiae* KSO5, was isolated from farm-produced omija (magnolia berry, *Schisandra chinensis*) fruit vinegar. The general characteristics of the strain are shown in Figure 1. Colonies grown on YGC solid medium were circular, convex, opaque, and cream-colored (Figure 1A-B), consistent with typical *Acetobacter* morphology. The cells were confirmed as Gram-negative (Figure 1C). Scanning electron microscopy (SEM) revealed that the cells were rod-shaped with oval ends, measuring 0.3–0.4  $\mu\text{m}$  in diameter and 0.9–1.2  $\mu\text{m}$  in length. The cells typically appeared in pairs, with some single cells observed (Figure 1D). No flagella, stalks, or prosthecae were observed.



**Figure 1.** Morphology of KSO5 strain. (A) Colonies grown on YGC agar medium. (B) Colonies observed under a stereomicroscope. (C) Gram-stained cells. (D) Scanning electron microscopy (SEM) images.

The KSO5 strain exhibited positive enzymatic activities for esterase (C4), esterase lipase (C8), leucine arylamidase, valine arylamidase, acid phosphatase, naphthol-AS-BI-phosphohydrolase, and acetyl-glucosaminidase (Figure 2). When compared to the closely related *A. malorum* CV11 (KACC 92076P), KSO5 showed positive valine arylamidase activity, whereas  $\beta$ -glucosidase activity was undetectable. Minor variations were observed in specific enzymatic activities, but overall, the enzymatic profiles of the KSO5 and CV11 strains were largely similar. The KSO5 strain has been deposited in the Korean Agricultural Culture Collection (KACC 92352P), part of the National Institute of Agricultural Sciences.



**Figure 2.** Enzyme activity of AC KSO5 determined using the API ZYM kit. Enzyme activities were assessed based on the hydrolysis of 19 substrates, following the interpretation criteria provided by the API manufacturer (<http://apiweb.biomerieux.com>). Symbol: +, positive; -, negative; CV11, *A. malorum* CV11 (KACC 92076P).

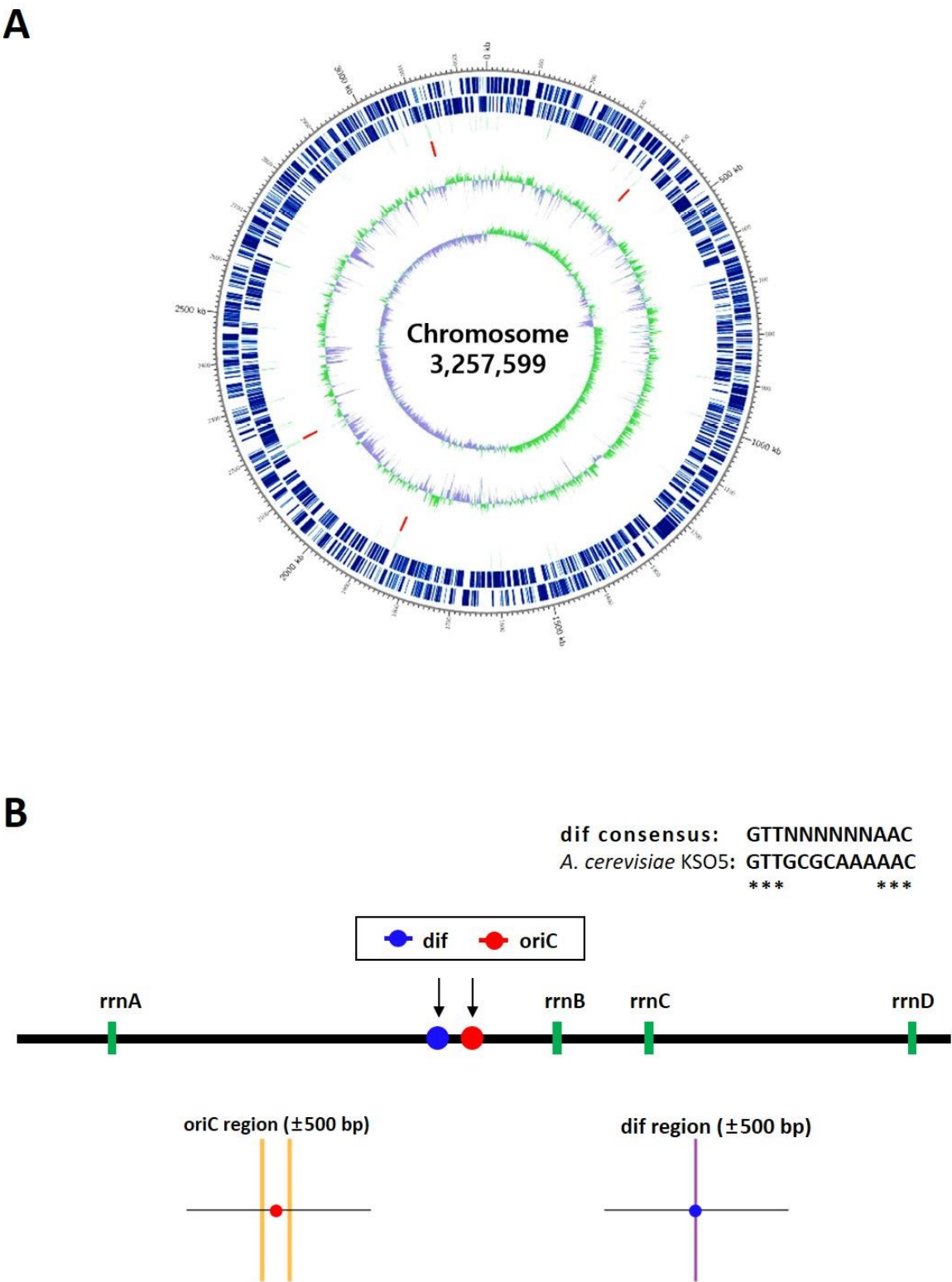
3.2. Genome Features of *A. cerevisiae* KSO5

The complete genome of strain KSO5 was determined and deposited in NCBI GenBank (accession no. CP172014–CP172016). It comprises one circular chromosome (3,257,599 bp; GC content, 57.8%; 2,889 protein-coding sequences [CDSs]; 54 tRNA genes; and 12 rRNA genes arranged in four rRNA operons) and two circular plasmids (KSO5\_P1, 4,905 bp, 56.2% GC, 5 CDSs; and KSO5\_P2, 4,820 bp, 56.1% GC, 4 CDSs) (Figure 3A; Table 4). In total, the genome spans 3,267,324 bp with an overall GC content of 57.8% and encodes 2,898 protein-coding genes, in addition to the 54 tRNA genes and 12 rRNA genes noted above (Table 4). The genomic GC content falls within the range reported for members of the genus *Acetobacter*.

**Table 4.** General genomic features of the chromosome and plasmids of the strain KSO5.

| Assortment | Name      | NCBI<br>Accession<br>no. | Length    | GC<br>(%) | Depth | Circular | CDS   | tRNA | rRNA |
|------------|-----------|--------------------------|-----------|-----------|-------|----------|-------|------|------|
| Chromosome | KSO5_Chrr | CP172014                 | 3,257,599 | 57.8      | 294.0 | YES      | 2,889 | 54   | 12   |
| Plasmids   | KSO5_P1   | CP172015                 | 4,905     | 56.2      | 14.3  | YES      | 5     | 0    | 0    |
|            | KSO5_P2   | CP172016                 | 4,820     | 56.1      | 28.2  | YES      | 4     | 0    | 0    |
| Total      |           |                          | 3,267,324 | 57.8      | 293.2 |          |       |      |      |

The origin of chromosomal replication (oriC) in AC KSO5 was predicted at position 1,582,749 bp, based on the coincidence of the GC skew maximum [11,12], clustering of at least seven *Alphaproteobacteria*-type DnaA-box motifs (TTATCCACA and variants) [13,14], and the presence of an AT-rich region adjacent to replication initiation genes (*dnaA*, *dnaN*, *recF*, *gyrB*) [13,15]. The putative replication termination site (dif) was identified at position 1,531,352 bp, showing perfect alignment to the *Alphaproteobacteria* consensus dif sequence (5'-GTTN{6}AAC-3') and proximity to the *xerC/xerD* recombinase genes [14]. These loci were further validated by cumulative GC skew analysis, revealing the oriC and dif positions at the skew maximum and minimum, respectively [11,12]. Four *rrn* operons were mapped at positions: *rrnA*, 385,463–390,694 bp; *rrnB*, 1,848,948–1,854,179 bp; *rrnC*, 2,189,505–2,194,736 bp; *rrnD*, 3,114,803–3,120,034 bp [15] (Figure 3B).

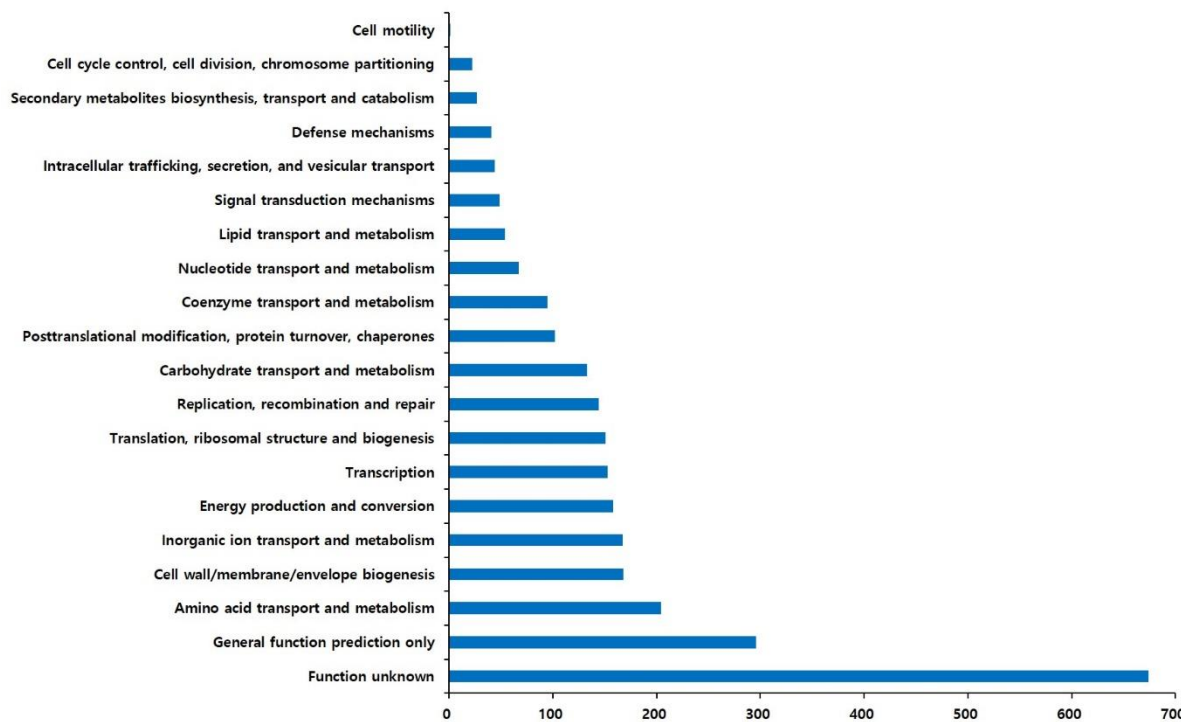


**Figure 3.** Features of the KSO5 chromosome. (A) Circular map (outer→inner): coordinates; CDS +/-; tRNA; rRNA; GC content; GC skew. (B) Linearised genome (3,257,599 bp) highlighting oriC and dif with rrn operons; insets magnify ±500 bp motifs (orange for DnaA-box, purple for dif site).

3.3. Gene Functional Analysis of *A. cerevisiae* KSO5

The complete genome harbours 2,898 coding sequences (CDSs) in total; of these, 2,722 CDSs were assigned to at least one COG category. In the COG classification, “general function prediction only” (296) was most abundant, followed by amino acid transport and metabolism (204), cell

wall/membrane/envelope biogenesis (168), inorganic ion transport and metabolism (167), energy production and conversion (158), transcription (153), translation, ribosomal structure and biogenesis (151), replication, recombination and repair (144), carbohydrate transport and metabolism (133), and post-translational modification, protein turnover and chaperones (102) (Figure 4). This profile suggests a metabolism centred on membrane/ion transport and energy conversion. We identified 674 hypothetical proteins (~25%) and 176 no-hit CDSs (~6%); the unclassified fraction is much lower than the 38% reported for *Acetobacter* species.



**Figure 4.** The genes of the *A. cerevisiae* KSO5 genome in COG functional categories.

3.4. Comparative Genomics and Phylogenomic Placement of *A. cerevisiae* KSO5

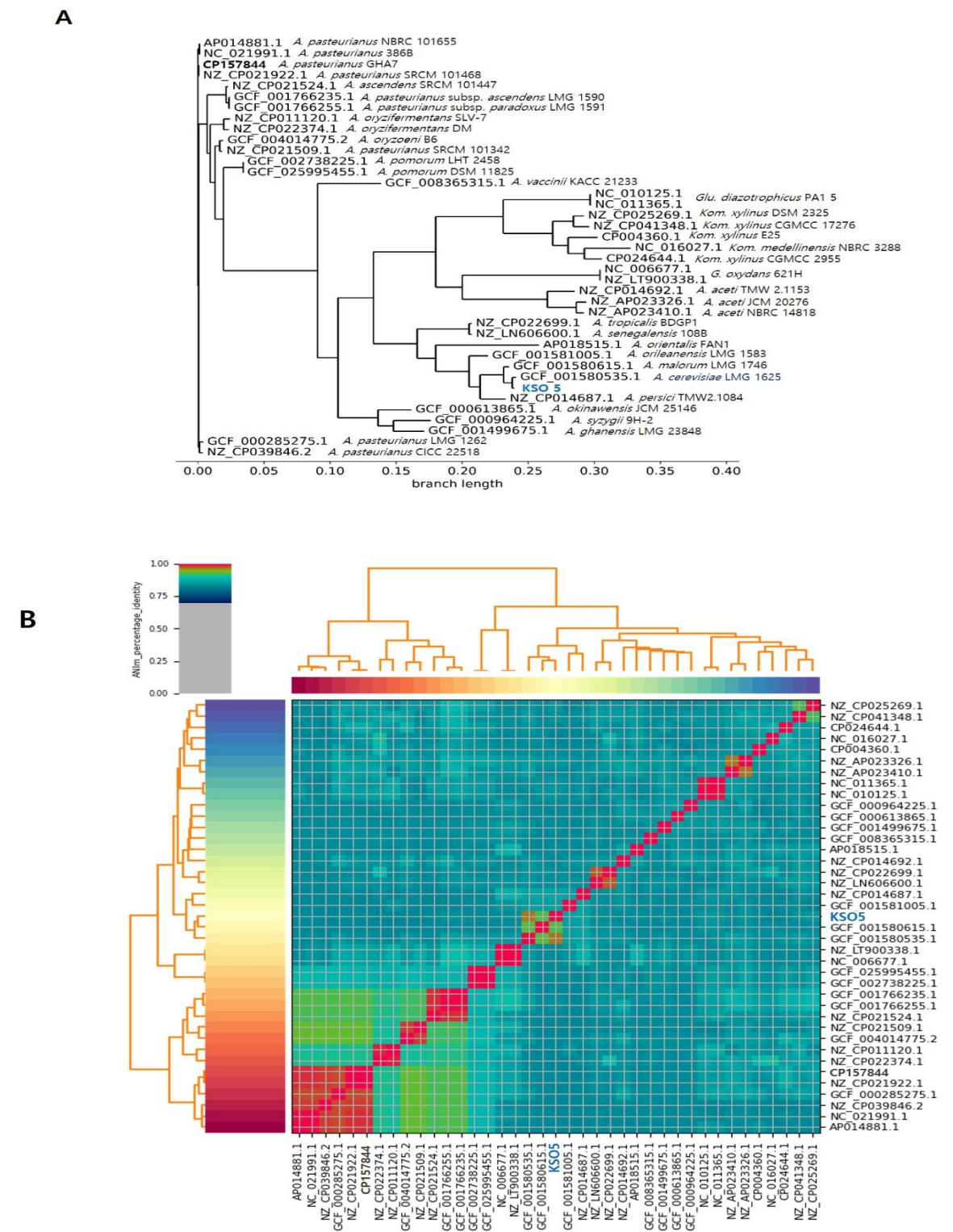
To determine the taxonomic placement of strain KSO5, we first compared its 16S rRNA gene sequence with those of *Acetobacter* strains; KSO5 showed the highest sequence similarity to *A. cerevisiae* LMG 1625<sup>T</sup> (99.71%; 16S rRNA gene phylogeny provided in Supplementary Figure S1), and the 16S rRNA gene sequence has been deposited in the NCBI GenBank database under accession number PP478110.

We then performed phylogenomic analyses across 39 genomes (Table 1): pan-genome clustering identified six core genes (present in 99–100% of genomes), 2,435 shell genes (15–95%), and 54,294 cloud genes (<15%), with core functions primarily associated with ribosomal proteins (30S/50S), acyl carrier proteins, and elongation factors (Supplementary Figure S2–S4).

Because the phylogeny in Figure 5A was inferred from the Roary gene presence–absence matrix, the dominant split reflected differences in gene-repertoire similarity rather than sequence divergence. In this framework, genomes separated into an *A. pasteurianus* complex (n = 16) and a broader lineage containing KSO5 (n = 23). The *A. pasteurianus* complex formed a more cohesive gene-content cluster, whereas the KSO5-containing lineage encompassed a broader taxonomic span and therefore displayed greater gene-content diversity.

Taken together, the gene presence–absence phylogeny placed KSO5 as a sister strain to the type strain LMG 1625<sup>T</sup> (Figure 5A). This relationship was further supported by ANI and dDDH analyses: KSO5 showed 97% ANI to *A. cerevisiae* LMG 1625<sup>T</sup>, exceeding the commonly accepted species delineation threshold (95–96%; Figure 5B; Supplementary Excel Data S1), whereas its ANI to *A. malorum* LMG 1746 was 93%, below the species boundary. Consistently, dDDH between KSO5 and

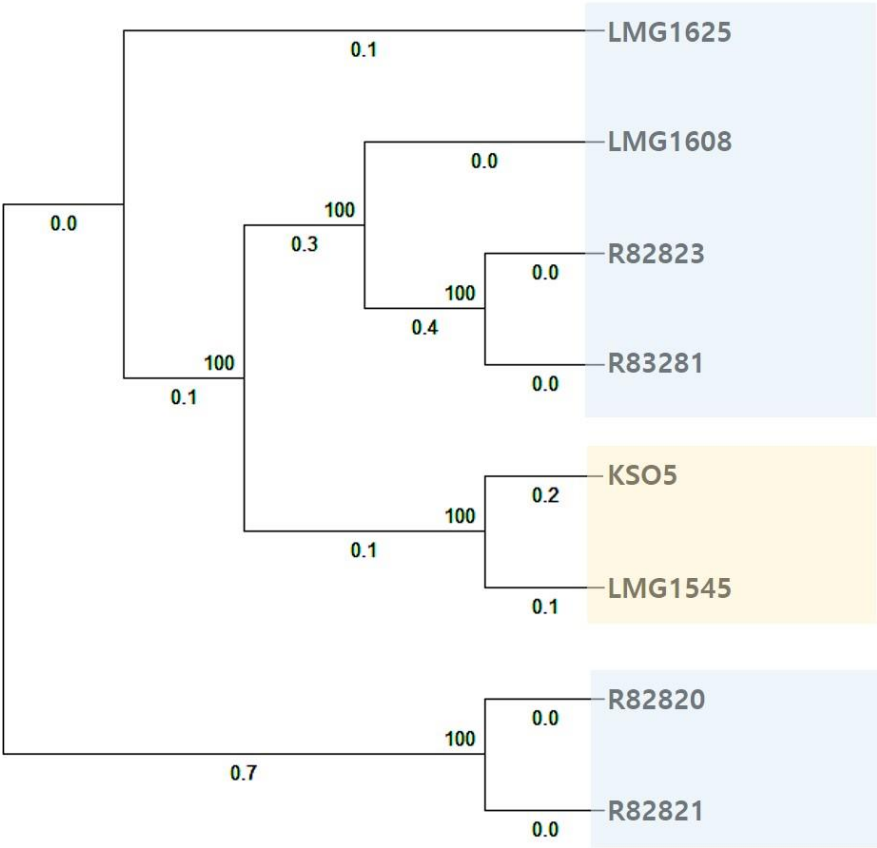
*A. cerevisiae* LMG 1625<sup>T</sup> was 71%, meeting the  $\geq 70\%$  species cutoff (Supplementary Excel Data S2). Collectively, these genomic indices confirm the assignment of KSO5 to *A. cerevisiae* (Figure 5B; Supplementary Excel Data S1–S2).



**Figure 5.** Phylogenomics of KSO5 from gene clusters. (A) presence-absence tree for 39 *Acetobacter* genomes (branch lengths shown). (B) ANI heatmap across coding sequences; all pairwise genome comparisons computed (Table 1).

3.5. Comparative Phylogenomic Analysis of *A. cerevisiae* KSO5

Across eight *A. cerevisiae* genomes, the pan-genome comprised 2,012 core, 2,024 shell, and 1,572 cloud genes. An accessory gene presence–absence phylogeny was inferred from the binary matrix of all non-core gene clusters (shell and cloud genes) across the eight genomes (Figure 6), thereby reflecting overall genome-content similarity rather than sequence divergence at conserved loci. In this genome-content framework, KSO5 clustered as a sister strain to LMG 1545, whereas the remaining strains formed two additional clades (R82820/21 and R82823/83281), with LMG 1608 and LMG 1625 branching within the broader set of beer-associated isolates. Notably, the KSO5–LMG 1545 grouping corresponded to the two vinegar-derived strains, while most beer-associated strains grouped outside this clade. However, because the current dataset was limited to eight genomes and did not include systematic phenotyping, this pattern was reported as a difference in accessory gene repertoires rather than as evidence of predictable, mechanism-specific ecological differentiation. Consistent with accessory-genome turnover, the strain set showed variation in mobile genetic element–associated functions (e.g., transposases, integrases, and phage-related genes), which was consistent with differences in the variable gene repertoire captured by the genome-content analysis.



**Figure 6.** Phylogeny of eight *A. cerevisiae* strains from accessory gene presence–absence. Terminal labels are strain names. Light-blue shading marks strains isolated from beer (LMG1625, LMG1608, R82823, R83281, R82820, and R82821), and pale-yellow shading marks the two vinegar-derived strains—KSO5 from fruit vinegar and LMG1545 from cereal vinegar.

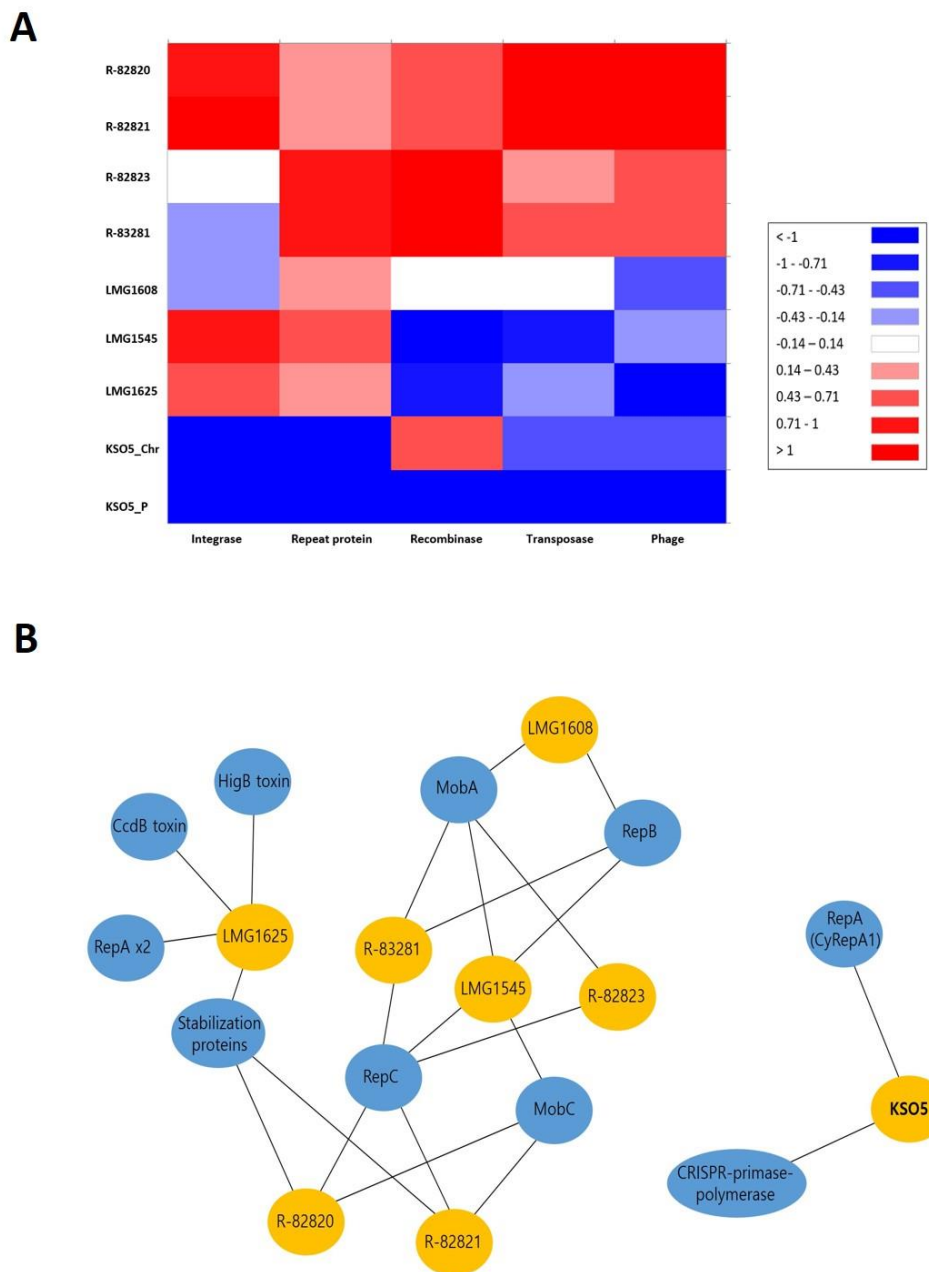
3.6. Comparative Profiling of Mobile Genetic Elements and Plasmid-Associated Protein Modules with Genetic Implications in *A. cerevisiae* KSO5 and Related Strains

To further contextualize the genome-content structure observed in Figure 6, we next quantified mobile genetic elements (MGEs) and plasmid-associated protein modules across the eight genomes.

Comparative genomic profiling of mobile genetic elements (MGEs) across eight *A. cerevisiae* strains revealed that KSO5 possesses a distinctive genomic signature characterized by a low abundance of MGEs in its plasmid and a moderate abundance in its chromosome (KSO5\_Chr) (Figure 7A). After normalizing all MGE tallies to counts per Mbp and confirming the trend per effective Mbp (i.e., genome size  $\times$  BUSCO completeness) (Supplementary Excel Data S3), KSO5 exhibits a moderate chromosomal MGE density but a minimal plasmid signal. On the closed chromosome (three circular contigs: chromosome + two plasmids; contig N50 = 3,257,599 bp; L50 = 1), we detected 18 transposases ( $\approx 5.45/\text{Mbp}$ ), 2 integrases ( $\approx 0.61/\text{Mbp}$ ), 7 phage-related genes ( $\approx 2.12/\text{Mbp}$ ), and 8 recombinases ( $\approx 2.42/\text{Mbp}$ ), with no “repeat-protein” annotations. The plasmid (KSO5\_P) contains only six transposases and one recombinase and lacks integrases, phage-related proteins, and repeat proteins, consistent with its structurally conservative nature. In a cross-strain comparison, KSO5 sits mid-range: it is below the MGE-rich R-82820/-82821 strains (e.g., transposases = 42; integrases = 8–9; phage-related = 29), but above strains with lower counts in specific categories (e.g., transposases in LMG1545). Its recombinase count ( $n = 8$ ) is relatively high—comparable to R-82820/-82821 but lower than R-82823/-R83281 ( $n = 10$ ). The absence of repeat-protein annotations in KSO5, compared with LMG1545/LMG1625 (14–16), should not be over-interpreted as a global reduction in repetitive DNA without sequence-level repeat analyses. Importantly, the conclusion remains robust regardless of assembly quality. While all non-KSO5 assemblies are fragmented drafts (128–177 contigs; N50/genome = 0.014–0.027), the key MGE categories—repeats, integrases, and phage-related genes—were consistently detected across these drafts. Thus, the observed contrasts are interpreted as differences in MGE load (reported as counts per Mbp and checked per effective Mbp), rather than as detection failures or true gene loss.

At the plasmid-module level, the KSO5 strain encodes a CyRepA1-family RepA plus a CRISPR-associated primase–polymerase (PrimPol-like) (Figure 7B and Table 3). Other strains display distinct strategies: LMG1625 carries dual RepA variants together with CcdB/HigB and multiple stabilization proteins; LMG1545 carries RepB/RepC with MobA/MobC (RepABC–MobAC-like); LMG1608 shows a lean RepB+MobA set; R-82820/-R82821 retain a minimal RepC+MobC architecture with stabilization proteins; and R-82823/-R83281 harbor RepB+RepC with MobA (RepABC-like). This network highlights the diversity of plasmid structures. The LMG1625 strain employs a toxin-antitoxin system for plasmid stabilization, while the KSO5 strain possesses a unique CRISPR-primase module potentially associated with adaptive immunity and replication regulation.

Overall, this data indicates moderate chromosomal mobility and a conservative plasmid structure for the KSO5 strain, supporting a stable yet adaptable genetic profile.

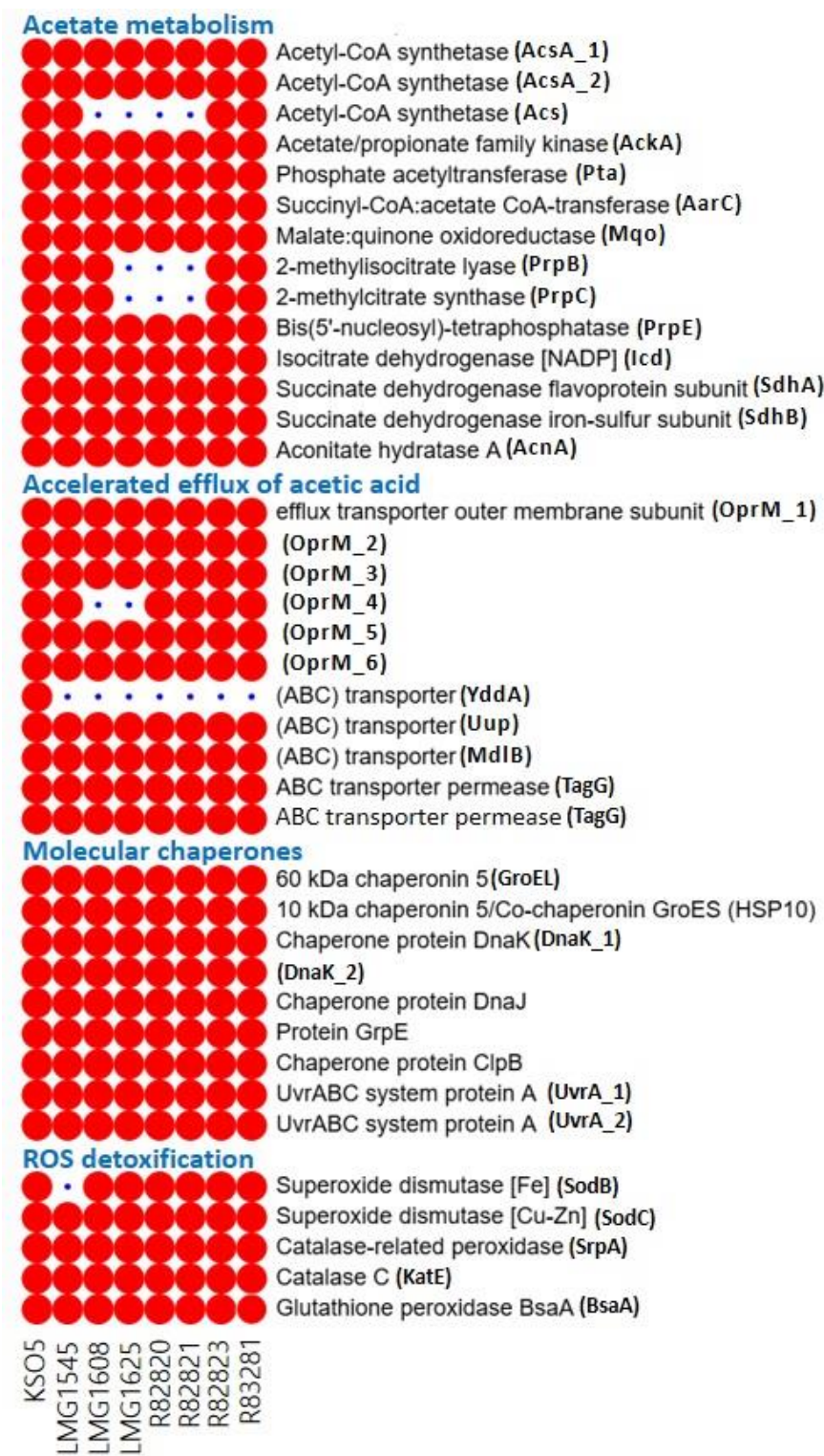


**Figure 7.** Distribution of mobile genetic element-related proteins and plasmid-associated protein modules in *A. cerevisiae* strains. (A) Heatmap visualizing the relative abundance of each protein type per strain. Warmer colors (red) indicate higher counts; cooler colors (blue) indicate lower counts. KSO5\_P is represented predominantly in the coolest color tones, while KSO5\_Chtr exhibits lighter shades for integrases and repeat proteins, in contrast to the warmer tones of high-MGE strains. (B) Blue nodes represent *A. cerevisiae* strains, while orange nodes indicate plasmid functional modules identified through genome annotation. The modules include replication initiators (RepA, RepB, RepC), mobilization proteins (MobA, MobC), stabilization proteins, toxin–antitoxin system components (CcdB, HigB), and a CRISPR-associated primase-polymerase. Edges indicate the presence of a given module in the corresponding strain.

3.7. Genetic Architecture of Acetic-Acid Resilience Across *A. cerevisiae* KSO5

AAB routinely face high organic-acid loads, especially acetic acid, which sharply depresses growth and metabolism [16]. To survive, AAB mount a multilayered defense that (i) leverages enzymatic acetate metabolism (assimilation), (ii) accelerates export through membrane transporters, including ABC systems and proton-motive-force-driven pumps, (iii) preserves proteome integrity by

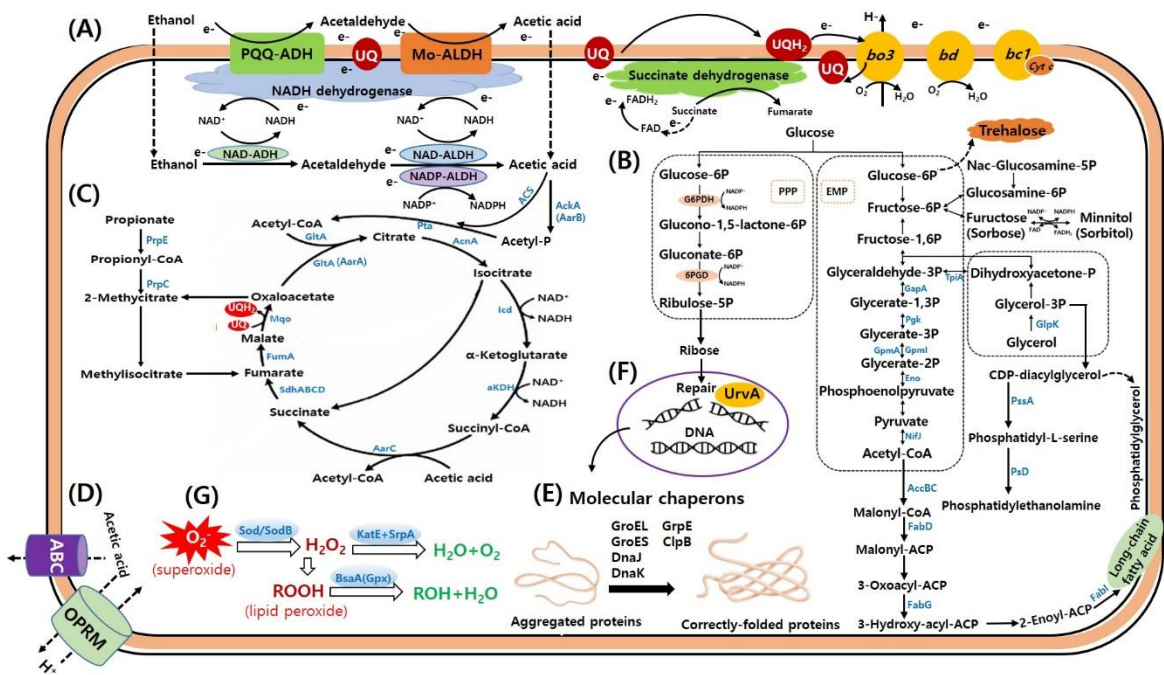
inducing stress chaperones such as DnaK and GroEL, and (iv) mitigates oxidative damage by detoxifying reactive oxygen species through enzymes such as superoxide dismutase, catalase, and peroxidases [17]. Consistent with this framework, the circular genome of *A. cerevisiae* KSO5 encodes extensive complements across all four layers, indicating a robust, genome-encoded capacity for acid tolerance [1,18] (Figure 8).



**Figure 8.** Pangenome-based presence-absence heatmap of acid-tolerance gene modules across eight *A. cerevisiae* genomes. The heatmap summarizes gene content (red = present; blue = absent) for four acid-tolerance modules across eight genomes. Columns represent strains.

3.7.1. Enzymatic Acetate Metabolism (Assimilation)

*A. cerevisiae* KSO5 encoded a broad intracellular acetate activation and assimilation repertoire, including three acetyl-CoA synthetases (AcsA\_1, AcsA\_2, Acs), the AckA–Pta route, and the AarC–Mqo-associated modified CAC framework, together with TCA nodes (Icd, SdhAB, AcnA) and prpB/prpC and a prpE-annotated locus (Figures 8 and 9C). Across the eight strains, these loci appeared predominantly red in the heatmap, with only sporadic blue gaps in non-KSO5 isolates, indicating a species-level backbone while positioning KSO5 as fully equipped at the gene-content level. This genomic pattern was consistent with prior functional studies implicating acetate activation/assimilation systems (e.g., Acs and the AckA–Pta route) and an AarC–Mqo-associated modified CAC in acetate oxidation and acetic-acid resistance in acetic acid bacteria [19,20]. In aggregate, the heatmap profile supported the interpretation that acetic-acid resilience in KSO5 reflected distributed metabolic capacity across acetate activation and central-carbon flux, rather than dependence on a single determinant [21].



**Figure 9.** An illustration of the key metabolic pathways and regulatory networks involving the expressed proteins. Detailed protein annotation data were displayed in Supplementary Table 1, Supplementary Table 2, and Supplementary Table 3.

3.7.2. Accelerated Efflux of Acetic Acid

A second pillar of acid tolerance in *A. cerevisiae* KSO5 was active acetic-acid extrusion (Figure 9D). Two transporter classes were commonly associated with this process: proton-motive-force-dependent efflux systems (e.g., RND/SMR/MFS families operating with TolC/OprM-like outer-membrane channels) and ATP-binding cassette (ABC) transporters [22–24]. To summarize the gene-content context of this efflux layer, we examined transporter loci that were (i) annotated as OprM/TolC-like outer-membrane efflux channels or ABC exporter components and (ii) resolved as distinct orthologous clusters in the pan-genome presence–absence matrix. This filtering retained six OprM-family paralogs (oprM\_1–oprM\_6) and five ABC transporter components (YddA, Uup, MdlB, TagG, and TagG permease) (Figure 8). In previously characterized bacterial efflux architectures, OprM-family proteins are typically described as outer-membrane conduits of tripartite assemblies that operate together with proton-motive-force-energized inner-membrane transporters, whereas ABC transporters are generally defined as ATP-driven exporters [25–28]. In KSO5, all loci within this module scored “present,” a pattern that was consistent with a gene repertoire compatible with acetic-

acid efflux. Across the eight genomes, these loci were largely conserved, with only sporadic absences in non-KSO5 strains, a distribution that was indicative of a species-level backbone underlying the potential for acetic-acid extrusion (Figure 8).

3.7.3. Stress Response Molecular Chaperones

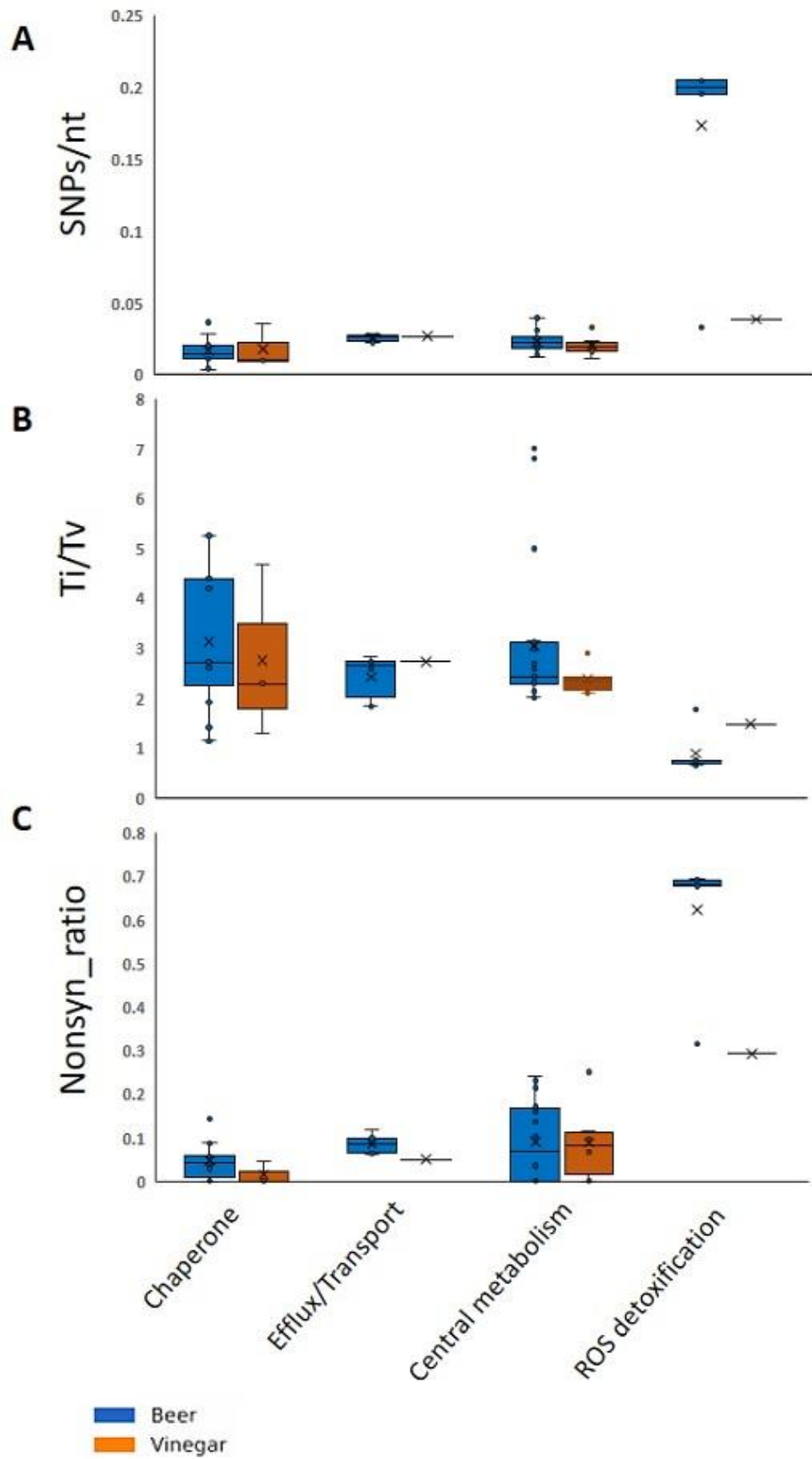
The proteostasis/repair arm is universally conserved across all eight *A. cerevisiae* genomes: GroES–GroEL, DnaK–DnaJ–GrpE, ClpB, and the DNA-repair factor UvrA show uninterrupted presence in every strain (Figure 8; Figure 9E–F). This species-wide invariance aligns with functional evidence in AAB—GroES/EL induction under acetic acid, ethanol, and heat correlates with improved viability [29], while dnaK, dnaJ, grpE, and clpB are upregulated during acetic-acid fermentation [30,31]; overexpressing uvrA further enhances resistance [32]. Together, the all chaperone/repair profile indicates that KSO5 does not depend on unique gene content in this module; instead, acid-tolerance differences among strains are likely quantitative/regulatory rather than content-driven.

3.7.4. Oxidative-Stress Detoxification (ROS Defense)

The oxidative-stress detoxification repertoire is conserved across the eight *A. cerevisiae* genomes, with one exception: LMG1545 lacks SodB. In contrast, KSO5 encodes a complete set of core ROS-defense enzymes, including superoxide dismutase (Sod), catalase (KatE), the catalase-related peroxidase (SrpA), and glutathione peroxidase (BsaA) (Figure 8; Supplementary Table 1). Collectively, these enzymes are well suited to mitigate the elevated reactive oxygen species generated during membrane-bound oxidative fermentation (e.g., superoxide and hydrogen peroxide), thereby limiting oxidative damage to lipids, proteins, and DNA and potentially enhancing fitness under high acetic-acid loads (Figure 9G).

3.7.5. SNP Comparison of Proteins Involved in Acetic Acid Resilience Mechanisms Across Eight *A. cerevisiae* Strains

Deliverables were compiled as Figure 10, which summarized mechanism-wise distributions of SNPs/nt, Ti/Tv, and the nonsynonymous ratio across Beer- and Vinegar-origin lineages (Table 2). In the chaperone/stress-response category, median SNPs/nt remained low in both lineages (Beer 0.0141 vs Vinegar 0.0098); Ti/Tv was higher in Beer (3.46 vs 2.29), and the nonsynonymous fraction was near zero (0.040 vs 0.000). In efflux/transport, SNPs/nt was similarly low in both groups (0.0267 in each), Ti/Tv was nearly indistinguishable (2.66 vs 2.73), and the nonsynonymous fraction was modestly higher in Beer (0.0845 vs 0.0488). In enzymatic acetate metabolism (assimilation), SNPs/nt values were close between lineages (Beer 0.0219 vs Vinegar 0.0189), Ti/Tv remained transition-biased and similar (2.43 vs 2.35), and nonsynonymous fractions were low in both groups (Beer 0.0685 vs Vinegar 0.0817). By contrast, the ROS detoxification category—represented by SrpA in the sentinel panel—showed higher between-lineage separation across metrics: Beer exhibited higher SNP density (0.200 vs 0.0384), lower Ti/Tv (0.740 vs 1.471), and a higher nonsynonymous fraction (0.681 vs 0.293). Overall, across this sentinel set, chaperone, efflux/transport, and central-carbon modules displayed comparatively low variation between lineages, whereas the strongest differences were concentrated in the ROS detoxification–represented category (Figure 10), consistent with heterogeneous variability across mechanisms rather than a uniform shift across the entire acid-resilience repertoire.



**Figure 10.** Mechanism-wise distributions of sequence variation. Boxplots summarize per-strain metrics from eight *A. cerevisiae* genomes, computed on MSA with the KSO5 allele as reference (CDSs were trimmed to the common in-frame length when needed), examining three metrics by mechanism for the Beer vs. Vinegar

(=LMG1545) source groups. (A) SNPs per nucleotide (SNPs/nt) across each CDS. (B) Transition/transversion ratio (Ti/Tv). (C) Proportion of nonsynonymous codons among variable codons ( $\text{nonsynonymous} \div [\text{synonymous} + \text{nonsynonymous} + \text{nonsense} + \text{stop-loss}]$ ). Mechanism assignments follow the sentinel manifest. Box-and whisker plots were generated with internal elements and outliers shown, including a mean marker and the median.

## 4. Discussion

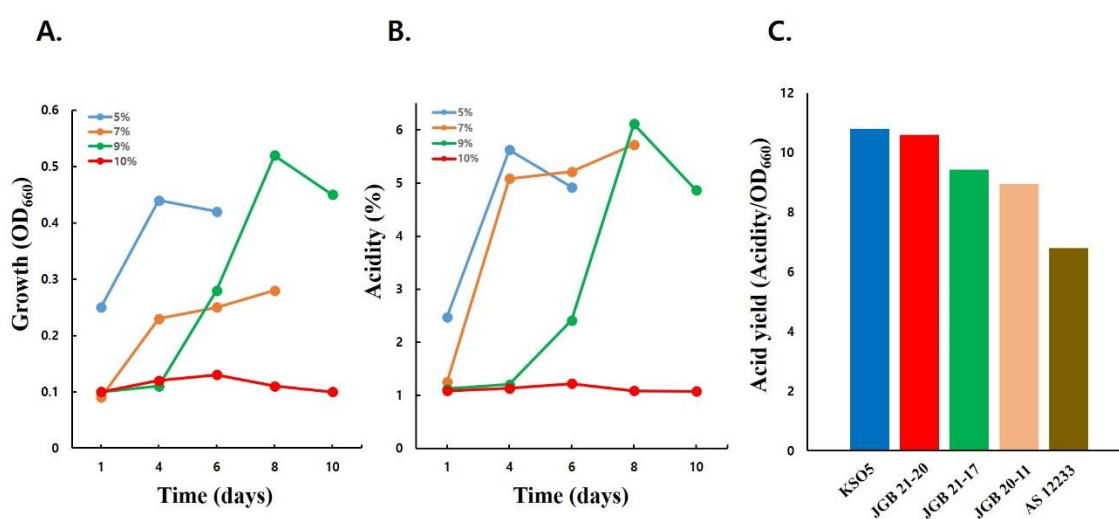
The complete genome of *A. cerevisiae* KSO5 (a 3.3-Mb chromosome and two plasmids) supports a robust genomic model of acetic acid tolerance, built on a tightly constrained backbone of central metabolism and essential protein-folding chaperones, overlaid by adaptable peripheral modules that regulate stress responses and the cell-environment interface. KSO5 harbored a respiratory and redox architecture compatible with rapid periplasmic oxidative fermentation under acid stress, including multiple PQQ-dependent alcohol and molybdopterin aldehyde dehydrogenase systems (PQQ-ADH/Mo-ALDH) that may help decouple carbon oxidation from cytosolic NAD(H) balance (Figure 9A; Supplementary Table 2). Notably, the membrane-bound PQQ-ADH module appeared to comprise the major AdhA/AdhB components without a clearly identifiable AdhS-like small subunit. This configuration resembled the two-subunit (AdhA/AdhB) membrane-bound ADH system reported in *Komagataeibacter* and *Gluconacetobacter*, rather than the typical three-subunit (AdhA/AdhB/AdhS) structure described in *Acetobacter* and *Gluconobacter* [3]. In parallel, KSO5 retained a putative Mo-ALDH-type module comprising a molybdopterin-dependent aldehyde dehydrogenase subunit, a 2Fe-2S-binding electron-transfer protein, and a cytochrome c subunit, consistent with the three-subunit architecture reported for membrane-bound ALDH complexes in acetic acid bacteria [3]. The genome also encoded PqqB, PqqC, PqqD, and PqqE, whereas a clearly annotated PqqA precursor peptide was not detected, suggesting a near-complete PQQ biosynthetic module with an unresolved small-peptide annotation rather than clear absence of cofactor biosynthetic capacity. The genome further encoded dual bo3- and bd-type ubiquinol oxidases (with bd duplication), a complete NADH dehydrogenase complex I (NuoA-N; notably absent in *Gluconobacter*), succinate dehydrogenase (SdhAB) [33], and an expanded bc1 complex (PetABC) with extra copies of petB/petC. In addition, the cytochrome-c maturation locus was largely conserved as CcmABCEFGH, although CcmG was not clearly identified in the current annotation [34]. Together, these features suggest that KSO5 maintains a highly configured electron-transfer network suited to fluctuating oxygen availability and persistent acid stress during vinegar fermentation.

Carbohydrate oxidation in KSO5 is equally versatile (Figure 9B; Supplementary Table 3). Consistent with acetic acid bacteria (AAB) physiology, the pentose phosphate pathway (PPP) is strongly represented (G6PDH, 6PGD), securing NADPH and ribulose-5-phosphate for anabolism and redox homeostasis [35–37]. Unlike many AAB that lack phosphofructokinase, KSO5 also encodes a complete Embden–Meyerhof–Parnas (EMP) pathway, allowing metabolic rerouting of fructose-6-phosphate and triose phosphates as fermentation conditions shift. Periplasmic PQQ-dependent glucose dehydrogenase accelerates extracellular glucose oxidation, rapidly acidifying the medium and conferring niche advantage. Additional polyol and sugar-acid routes—FAD-dependent sorbitol dehydrogenase (D-sorbitol→L-sorbose→D-fructose), mannitol oxidation via polyol oxidoreductase, and glycerol utilization via GlpF/GlpK and FAD/NAD-linked G3PDHs—feed EMP intermediates and enhance carbon throughput [38–40]. Osmotolerance and envelope resilience are supported by trehalose metabolism (OtsA/OtsB/TreA) [41] and large-conductance mechanosensitive channels (MscL/MscS/MscK) [42], while abundant transporters (e.g., ~50 ABC, multiple symporters/permeases) ensure broad uptake of sugars, polyols, organic acids, and ions to sustain flux through PPP/EMP, the TCA cycle, and oxidative phosphorylation.

These respiratory and metabolic layers integrate with three genome-encoded acetate-handling strategies that mitigate intracellular acid: (i) PMF-driven and ABC-mediated efflux; (ii) conversion to acetyl-CoA via acetic acid kinase/phosphotransacetylase/acetyl-CoA synthetase, channeling acetate

into central metabolism; and (iii) a specialized TCA branch via succinate CoA:acetate CoA-transferase (AarC), previously linked to secondary growth and acetic-acid resistance in AAB. Membrane homeostasis genes (ACC, FabD/FabG/FabI; phosphatidylserine and phosphatidylethanolamine synthesis) and cell-wall remodeling capacity (e.g., N-acetylglucosaminidase activity) further harden the envelope against ethanol/acid insults [43].

Consistent with the genomic predictions, phenotypic evaluation of KSO5 alongside four reference AAB from the Korean Agricultural Culture Collection (three *A. pasteurianus* strains, and one *Acetobacter* strain; Supplementary Methods) showed that KSO5 retained growth and titratable acidity production up to 9% ethanol, whereas both traits were markedly impaired at 10% ethanol (Figure 11; Supplementary Figure 5 and 6). Because the OD<sub>660</sub> and acidity time-course data were obtained from a single culture per condition (n = 1), these results should be interpreted as observational, trend-level evidence rather than as a basis for statistical comparison. Even so, the delayed rise in acidity observed under 9% ethanol was consistent with an adaptation phase under combined ethanol–acid stress, followed by activation of oxidative fermentation (Figure 11A and B).



**Figure 11.** Growth, acid production, and representing biomass-normalized acid production of *A. cerevisiae* KSO5 under ethanol stress. (A) Growth of KSO5 under different ethanol concentrations, (B) Acidity of KSO5 under different ethanol concentrations, (C) Comparison of biomass-normalized acid production efficiency among *Acetobacter* strains under ethanol stress (9% ethanol). The apparent acid-yield index was calculated as final acidity divided by final OD<sub>660</sub>. Individual growth curves and acid production profiles are shown in Supplementary Figures S5 and S6.

Under the 9% ethanol condition, KSO5 showed a high acid-yield index when acidity was normalized to OD<sub>660</sub>, indicating enhanced biomass-normalized acid production efficiency relative to its absolute final acidity (Figure 11C). In AAB, ethanol oxidation is driven by the respiratory chain-linked membrane-bound ADH/ALDH system [3,39,44]. Therefore, the relatively high acidity-to-biomass ratio observed in KSO5 may reflect preferential maintenance of oxidative metabolism and acidification capacity under ethanol stress, rather than superiority in biomass accumulation itself. Since this index was calculated as acidity/OD<sub>660</sub>, it should be interpreted as an operational proxy for metabolic and ethanol-oxidation efficiency rather than as a direct kinetic or stoichiometric yield parameter.

Taken together, considering the genome-based functional annotations together with the phenotypic observations, KSO5 appears capable of maintaining oxidative fermentation and acetate-handling functions even under high ethanol stress. However, this performance seems to be constrained near the upper stress limit, particularly at 10% ethanol. Additional independent replicate

experiments will be required to more rigorously define strain-level differences in rate, yield, and overall productivity.

## 5. Conclusion

The complete circular genome of *Acetobacter cerevisiae* KSO5 provides a genetic framework for ethanol oxidation and adaptation to acid stress, including genes associated with PQQ-ADH/Mo-ALDH-based oxidative fermentation, respiratory chain components, and candidate acetate-handling functions. Consistent with these genomic features, phenotypic evaluation under ethanol stress showed that KSO5 maintained both growth and acidity production up to 9% ethanol, but exhibited a clear threshold-like decline in both traits at 10%. Taken together, these findings provide genome-to-phenotype evidence supporting the potential utility of KSO5 in vinegar fermentation environments where ethanol and acid stress act as major limiting factors, and offer a basis for future starter-strain selection and process design. However, additional replicate experiments and pilot-scale validation will be required to establish quantitative productivity gains and statistically supported strain-level performance differences.

**Supplementary Materials:** The following supporting information can be downloaded at the website of this paper posted on Preprints.org. Supplementary Methods: Phenotyping under ethanol stress. Supplementary Figs S1–S6: Fig. S1, 16S rRNA Phylogenetic tree; Fig. S2, Proportional distribution of the pangenome components; Fig. S3, Upset plot depicting the number of gene clusters shared among samples; Fig. S4, Heatmap and phylogenetic tree based on presence/absence of gene clusters; Fig. S5, Growth of KSO5 alongside four reference AAB under different ethanol concentrations; Fig. S6, Acidity of KSO5 alongside four reference AAB under different ethanol concentrations. Supplementary Data S1 (.xlsx): KSO5\_ANI\_percentage. Supplementary Data S2 (.xlsx): DDH\_KSO5. Supplementary Data S3 (.xlsx): eight strains AC\_Normalization\_Quality\_Factors. Supplementary Table 1 (.docx): Mechanisms involved in acetic acid tolerance in *A. cerevisiae* KSO5. Supplementary Table 2 (.docx): Membrane-bound dehydrogenases- and respiratory chain-related proteins encoded in the genome of AC KSO5. Supplementary Table 3 (.docx): Proteins involved in carbohydrates oxidation in the genome of AC KSO5.

**Author Contributions:** This work was carried out in collaboration with all authors. S.H.K. and S.-Y.K. designed the study. S.H.K. performed the statistical analysis, wrote the protocol, and prepared the first draft of the manuscript. D.M.H. conducted a comparative analysis of the genome of eight strains of *A. cerevisiae* and SNP analysis of genes related to acid tolerance. S.-E.Y. captured the microscopic images of the samples. C.W.K. and J.J.P. managed the study analyses and conducted the literature review. All authors have read and approved the final version of the manuscript.

**Funding:** This study was supported by 2026 the RDA Fellowship Program of (Agricultural Science and Technology Development (Project No. PJ017474042026) and the National Institute of Crop and Food Science, Rural Development Administration, Republic of Korea.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The original contribution presented in this study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

**Acknowledgments:** The financial support provided by the Rural Development Administration is gratefully acknowledged.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Mamlouk, D.; Gullo, M. Acetic acid bacteria: Physiology and carbon sources oxidation. *Indian J. Microbiol.* **2013**, *53*, 377-384. <https://doi.org/10.1007/s12088-013-0414-z>
2. Chinnawirotpisan, P.; Theeragool, G.; Limtong, S.; Toyama, H.; Adachi, O.O.; Matsushita, K. Quinoprotein alcohol dehydrogenase is involved in catabolic acetate production, while NAD-dependent alcohol dehydrogenase in ethanol assimilation in *Acetobacter pasteurianus* SKU1108. *J. Biosci. Bioeng.* **2003**, *96*, 564-571. [https://doi.org/10.1016/S1389-1723\(04\)70150-4](https://doi.org/10.1016/S1389-1723(04)70150-4)
3. Miah, R.; Nina, S.; Murate, T.; Kataoka, N.; Matsutani, M.; Ano, Y.; Matsushita, K.; Yakushi, T. Dissection and Reconstitution Provide Insights into Electron Transport in the Membrane-Bound Aldehyde Dehydrogenase Complex of *Gluconacetobacter diazotrophicus*. *J. Bacteriol.* **2022**, *204*, e00558-21. <https://doi.org/10.1128/jb.00558-21>
4. Illegheems, K.; De Vuyst, L.; Weckx, S. Complete genome sequence and comparative analysis of *Acetobacter pasteurianus* 386B, a strain well-adapted to the cocoa bean fermentation ecosystem. *BMC Genom.* **2013**, *14*, 526. <https://doi.org/10.1186/1471-2164-14-526>
5. Prust, C.; Hoffmeister, M.; Liesegang, H.; Wiezer, A.; Fricke, W.F.; Ehrenreich, A.; Gottschalk, G.; Deppenmeier, U. Complete Genome Sequence of the Acetic Acid Bacterium *Gluconobacter oxydans*. *Nat. Biotechnol.* **2005**, *23*, 195-200. <https://doi.org/10.1038/nbt1062>
6. Walker, B.J.; Abeel, T.; Shea, T.; Priest, M.; Abouelliel, A.; Sakthikumar, S.; Cuomo, C.A.; Zeng, Q.; Wortman, J.; Young, S.K.; Earl, A.M. Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. *PLoS ONE* **2014**, *9*, e112963. <https://doi.org/10.1371/journal.pone.0112963>
7. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinform.* **2014**, *30*, 2068-2069. <https://doi.org/10.1093/bioinformatics/btu153>
8. Page, A.J.; Cummins, C.A.; Hunt, M.; Wong, V.K.; Reuter, S.; Holden, M.T. G.; Fookes, M.; Falush, D.; Keane, J.A.; Parkhill, J. Roary: Rapid Large-Scale Prokaryote Pan Genome Analysis. *Bioinform.* **2015**, *31*, 3691-3693. <https://doi.org/10.1093/bioinformatics/btv421>
9. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree 2 –approximately maximum-likelihood trees for large alignments. *PLoS ONE* **2010**, *5*, e9490. <https://doi.org/10.1371/journal.pone.0009490>
10. Letunic, I.; Bork, P. Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **2021**, *49*(W1), W293–W296. <https://doi.org/10.1093/nar/gkab301>
11. Grigoriev, A. Analyzing genomes with cumulative skew diagrams. *Nucleic Acids Res.* **1998**, *26*, 2286–2290. <https://doi.org/10.1093/nar/26.10.2286>
12. Lobry, J.R. Asymmetric substitution patterns in the two DNA strands of bacteria. *Mol. Biol. Evol.* **1996**, *13*, 660–665. <https://doi.org/10.1093/oxfordjournals.molbev.a025626>
13. Gao, F.; Zhang, C.T. Ori-Finder: A web-based system for finding oriCs in unannotated bacterial genomes. *BMC Bioinform.* **2008**, *9*, 79. <https://doi.org/10.1186/1471-2105-9-79>
14. Kono, N.; Arakawa, K.; Tomita, M. Comprehensive prediction of chromosome dimer resolution sites in bacterial genomes. *BMC Genom.* **2011**, *12*, 19. <http://www.biomedcentral.com/1471-2164/12/19>
15. Mackiewicz, P.; Zakrzewska-Czerwińska, J.; Zawilak, A.; Dudek, M.R.; Cebat, S. Where does bacterial replication start? Rules for predicting the oriC region. *Nucleic Acids Res.* **2004**, *32*, 3781–3791. <https://doi.org/10.1093/nar/gkh699>
16. Kersters, K.; Lisdiyanti, P.; Komagata, K.; Swings, J. The family *Acetobacteraceae*. In *The genera Acetobacter, Acidomonas, Asaia, Gluconacetobacter, Gluconobacter, and Kozakia*, 2nd ed.; Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.H., Stackebrandt, E.; The Prokaryotes, Springer, New York, 2006; pp. 163–200. [https://doi.org/10.1007/0-387-30745-1\\_9](https://doi.org/10.1007/0-387-30745-1_9)
17. Yang, H.; Chen, T.; Wang, M.; Zhou, J.; Liebl, W.; Barja, F.; Chen, F. Molecular Biology: Fantastic Toolkits to Improve Knowledge and Application of Acetic Acid Bacteria. *Biotechnol. Adv.* **2022**, *58*, 107911. <https://doi.org/10.1016/j.biotechadv.2022.107911>
18. Azuma, Y.; Hosoyama, A.; Matsutani, M.; Furuya, N.; Horikawa, H.; Harada, T.; Hirakawa, H.; Kuhara, S.; Matsushita, K.; Fujita, N.; Shirai, M. Whole-Genome Analyses Reveal Genetic Instability of *Acetobacter pasteurianus*. *Nucleic Acids Res.* **2009**, *37*, 5768–5783. <https://doi.org/10.1093/nar/gkp612>

19. Yang, H.; Yu, Y.; Fu, C.; Chen, F. Bacterial acid resistance toward organic weak acid revealed by RNA-seq transcriptomic analysis in *Acetobacter pasteurianus*. *Front Microbiol.* **2019**, *10*, 1616. <https://doi.org/10.3389/fmicb.2019.01616>
20. Mullins, E.A.; Francois, J.A.; Kappock, T.J. A specialized citric acid cycle requiring succinyl-coenzyme A (CoA): acetate CoA-transferase (AarC) confers acetic acid resistance on the acidophile *Acetobacter aceti*. *J. Bacteriol.* **2008**, *190*, 4933–4940. <https://doi.org/10.1128/jb.00405-08>
21. Hua, S.; Wang, Y.; Wang, L.; Zhou, Q.; Li, Z.; Liu, P.; Wang, K.; Zhu, Y.; Han, D.; Yu, Y. Regulatory Mechanisms of Acetic Acid, Ethanol, and High Temperature Tolerances of Acetic Acid Bacteria during Vinegar Production. *Microb. Cell Fact.* **2024**, *23*, 324. <https://doi.org/10.1186/s12934-024-02602-y>
22. Nakano, S.; Fukaya, M. Analysis of proteins responsive to acetic acid in *Acetobacter*: Molecular mechanisms conferring acetic acid resistance in acetic acid bacteria. *Int. J. Food Microbiol.* **2008**, *125*, 54–59. <https://doi.org/10.1016/j.ijfoodmicro.2007.05.015>
23. Qiu, X.; Zhang, Y.; Hong, H. Classification of acetic acid bacteria and their acid resistant mechanism. *AMB Expr.* **2021**, *11*, 29. <https://doi.org/10.1186/s13568-021-01189-6>
24. Kim, S.H.; Jang, H.W.; Park, J.J.; Nam, D.G.; Lee, S.J.; Yeo, S.H.; Kim, S.Y. Antibiotic resistance in acetic acid bacteria originating from vinegar. *Antibiotics* **2024**, *13*, 626. <https://doi.org/10.3390/antibiotics13070626>
25. Li, X.Z.; Nikaido, H. Efflux-Mediated Drug Resistance in Bacteria: An Update. *Drugs* **2009**, *69*, 1555–1623. <https://doi.org/10.2165/11317030-000000000-00000>
26. Neuberger, A.; Du, D.; Luisi, B.F. Structure and mechanism of bacterial tripartite efflux pumps. *Res. Microbiol.* **2018**, *169*, 401–413. <https://doi.org/10.1016/j.resmic.2018.05.003>
27. Higgins, C.F. Multiple molecular mechanisms for multidrug resistance transporters. *Nature* **2007**, *446*, 749–757. doi:10.1038/nature05630
28. Davidson, A.L.; Dassa, E.; Orelle, C.; Chen, J. Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol Mol. Biol. Rev.* **2008**, *72*, 317–364. DOI: 10.1128/MMBR.00031-07
29. Wang, B.; Shao, Y.; Chen, T.; Chen, W.; Chen, F. Global insights into acetic acid resistance mechanisms and genetic stability of *Acetobacter pasteurianus* strains by comparative genomics. *Sci. Rep.* **2015**, *5*, 18330. <https://doi.org/10.1038/srep18330>
30. Andrés-Barrao, C.; Saad, M. M.; Chappuis, M. L.; Boffa, M.; Perret, X.; Ortega Pérez, R.; Barja, F. Proteome Analysis of *Acetobacter pasteurianus* during Acetic Acid Fermentation. *J. Proteomics* **2012**, *75*, 1701–1717. <https://doi.org/10.1016/j.jprot.2011.11.027>
31. Okamoto-Kainuma, A.; Ishikawa, M.; Nakamura, H.; Fukazawa, S.; Tanaka, N.; Yamagami, K.; Koizumi, Y. Characterization of *rpoH* in *Acetobacter pasteurianus* NBRC3283. *J. Biosci. Bioeng.* **2011**, *111*, 429–432. <https://doi.org/10.1016/j.jbiosc.2010.12.016>
32. Zheng, Y.; Wang, J.; Bai, X.; Chang, Y.; Mou, J.; Song, J.; Wang, M. Improving the Acetic Acid Tolerance and Fermentation of *Acetobacter pasteurianus* by Nucleotide Excision Repair Protein UvrA. *Appl. Microbiol. Biotechnol.* **2018**, *102*, 6493–6502. <https://doi.org/10.1007/s00253-018-9066-6>
33. Greenfield, S.; Claus, G.W. Nonfunctional tricarboxylic acid cycle and the mechanism of glutamate biosynthesis in *Acetobacter suboxydans*. *J. Bacteriol.* **1972**, *112*, 1295–1301. <https://doi.org/10.1128/JB.112.3.1295-1301.1972>
34. Sanders, C.; Turkarslan, S.; Lee, D. W.; Onder, O.; Kranz, R. G.; Daldal, F. The Cytochrome c Maturation Components CcmF, CcmH, and CcmI Form a Membrane-Integral Multisubunit Heme Ligation Complex. *J. Biol. Chem.* **2008**, *283*, 29715–29722. <https://doi.org/10.1074/jbc.M805413200>
35. Adler, P.; Frey, L. J.; Berger, A.; Bolten, C. J.; Hansen, C. E.; Wittmann, C. The Key to Acetate: Metabolic Fluxes of Acetic Acid Bacteria under Cocoa Pulp Fermentation-Simulating Conditions. *Appl. Environ. Microbiol.* **2014**, *80*, 4702–4716. <https://doi.org/10.1128/AEM.01048-14>
36. García-García, I.; Cañete-Rodríguez, A. M.; Santos-Dueñas, I. M.; Jiménez-Hornero, J. E.; Ehrenreich, A.; Liebl, W.; García-Martínez, T.; Mauricio, J. C. Biotechnologically Relevant Features of Gluconic Acid Production by Acetic Acid Bacteria. *Acetic Acid Bact.* **2017**, *6*, 6458. <https://doi.org/10.4081/aab.2017.6458>
37. Yin, H.; Zhang, R.; Xia, M.; Bai, X.; Mou, J.; Zheng, Y.; Wang, M. Effect of Aspartic Acid and Glutamate on Metabolism and Acid Stress Resistance of *Acetobacter pasteurianus*. *Microb. Cell Fact.* **2017**, *16*(1), 109. <https://doi.org/10.1186/s12934-017-0717-6>

38. Cummins, J.T.; King, T.E.; Cheldelin, V.H. The biological oxidation of sorbitol. *J. Biol. Chem.* **1957**, *224*, 323-329. [https://doi.org/10.1016/S0021-9258\(18\)65031-8](https://doi.org/10.1016/S0021-9258(18)65031-8)
39. 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. <https://doi.org/10.3389/fmicb.2022.879246>
40. Shinagawa, E.; Matsushita, K.; Adachi, O.; Ameyama, M. Purification and characterization of D-sorbitol dehydrogenase from membrane of *Gluconobacter suboxydans* var.  $\alpha$ . *Agric. Biol. Chem.* **1982**, *46*, 135-141. <https://doi.org/10.1271/bbb1961.46.135>
41. Argüelles, J.C. Physiological roles of trehalose in bacteria and yeasts: a comparative analysis. *Arch. Microbiol.* **2000**, *174*, 217-224. <https://doi.org/10.1007/s002030000192>
42. Booth, I.R.; Edwards, M.D.; Black, S.; Schumann, U.; Miller, S. Mechanosensitive channels in bacteria: signs of closure? *Nat. Rev. Microbiol.* **2007**, *5*, 431-440. <https://doi.org/10.1038/nrmicro1659>
43. Eberlein, C.; Baumgarten, T.; Starke, S.; Heipieper, H.J. Immediate response mechanisms of gram-negative solvent-tolerant bacteria to cope with environmental stress: Cis-trans isomerization of unsaturated fatty acids and outer membrane vesicle secretion. *Appl. Microbiol. Biotechnol.* **2018**, *102*, 2583-2593. <https://doi.org/10.1007/s00253-018-8832-9>
44. Song, J.; Wang, J.; Wang, X.; Zhao, H.; Hu, T.; Feng, Z.; Lei, Z.; Li, W.; Zheng, Y.; Wang, M. Improving the Acetic Acid Fermentation of *Acetobacter pasteurianus* by Enhancing the Energy Metabolism. *Front. Bioeng. Biotechnol.* **2022**, *10*, 3389. <https://doi.org/10.3389/fbioe.2022.815614>
45. Kim, S.H.; Kim, J.Y.; Jeong, W.S.; Gwon, H.M.; Kim, S.Y.; Yeo, S.H. Culture and Function-Related Characteristics of Six Acetic Acid Bacterial Strains Isolated from Farm-Made Fermented Vinegars. *Korean J. Food Preserv.* **2022**, *29*, 142-156. <https://doi.org/10.11002/kjfp.2022.29.1.142>
46. Kim, S.H.; Jeong, W.S.; Kim, S.Y.; Yeo, S.H. Quality and Functional Characterization of Acetic Acid Bacteria Isolated from Farm-Produced Fruit Vinegars. *Fermentation* **2023**, *9*, 447. <https://doi.org/10.3390/fermentation9050447>
47. Tamura, K.; Nei, M.; Kumar, S. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 11030-11035. <https://doi.org/10.1073/pnas.0404206101>
48. Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725-2729. <https://doi.org/10.1093/molbev/mst197>

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.