

Group Members: Davon Perry Nugroho (2702240380), Fadya Annura Manan (2702311894), Izhar Octafirlian Susilo (2702222144)

1. A. **P5CS1**: CDS = 45..2234 -> in bp = 2234 - 45 + 1 = 2190 bp.
P5CS2: CDS = 11..2161 -> in bp = 2161 - 11 + 1 = 2151 bp.
- B. **P5CS1 protein**: 2190 bp / 3 bp/amino acid = 730 amino acids.
P5CS2 protein: 2151 bp / 3 bp/amino acid = 717 amino acids.
2. P5CS1:

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	Saccharum hybrid cultivar delta 1-pyrroline-5-carboxylate synthetase 1 (P5CS1) mRNA, complete cds	Saccharum hybr...	4340	4340	100%	0.0	100.00%	2350	KF178299.1
<input checked="" type="checkbox"/>	PREDICTED: Miscanthus floridulus delta-1-pyrroline-5-carboxylate synthase 2 (LOC136450733), transcript v...	Miscanthus flori...	4037	4037	100%	0.0	97.66%	3378	XM_066451328.1
<input checked="" type="checkbox"/>	PREDICTED: Sorghum bicolor delta-1-pyrroline-5-carboxylate synthase 2 (LOC110433516), mRNA	Sorghum bicolor	3982	3982	99%	0.0	97.59%	2910	XM_021455806.1
<input checked="" type="checkbox"/>	Sorghum bicolor delta 1-pyrroline-5-carboxylate synthetase 1 (P5CS1) mRNA, complete cds	Sorghum bicolor	3932	3932	99%	0.0	97.21%	2587	GQ377719.2
<input checked="" type="checkbox"/>	PREDICTED: Miscanthus floridulus delta-1-pyrroline-5-carboxylate synthase 2 (LOC136450733), transcript v...	Miscanthus flori...	3699	3920	97%	0.0	97.94%	3308	XM_066451329.1
<input checked="" type="checkbox"/>	Zea mays uncharacterized LOC100280719 (LOC100280719), mRNA	Zea mays	3637	3637	100%	0.0	94.69%	2889	NM_001153639.1
<input checked="" type="checkbox"/>	PREDICTED: Setaria viridis delta-1-pyrroline-5-carboxylate synthase 2 (LOC117857150), mRNA	Setaria viridis	3070	3070	98%	0.0	90.89%	2678	XM_034739658.2

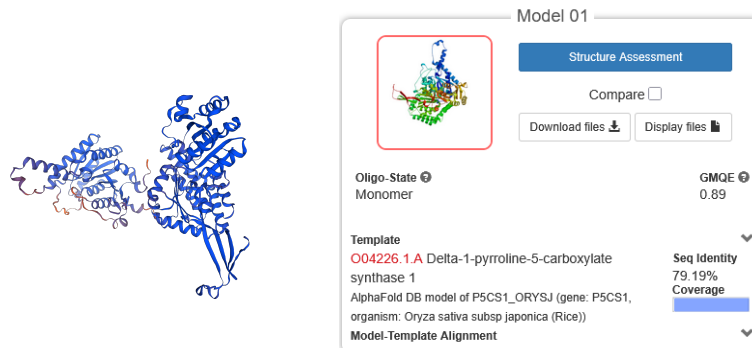
P5CS2:

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	Saccharum hybrid cultivar delta 1-pyrroline-5-carboxylate synthetase 2 (P5CS2) mRNA, complete cds	Saccharum hybr...	4048	4048	100%	0.0	100.00%	2244	KF178300.1
<input checked="" type="checkbox"/>	Saccharum arundinaceum delta-1-pyrroline-5-carboxylate synthetase mRNA, complete cds	Triplidium arundi...	3920	3920	100%	0.0	98.75%	2563	EU113257.1
<input checked="" type="checkbox"/>	Saccharum officinarum delta-1-pyrroline-5-carboxylate synthetase (P5CS) mRNA, complete cds	Saccharum offic...	3793	3793	100%	0.0	97.60%	2719	EF155655.1
<input checked="" type="checkbox"/>	Saccharum hybrid cultivar GT21 delta-1-pyrroline-5-carboxylate synthetase mRNA, complete cds	Saccharum hybr...	3781	3781	96%	0.0	98.98%	2151	KJ546350.1
<input checked="" type="checkbox"/>	Saccharum spontaneum pyrroline-5-carboxylate synthetase 2 (P5CS2) mRNA, complete cds	Saccharum spo...	3778	3778	96%	0.0	98.84%	2159	KX714117.1
<input checked="" type="checkbox"/>	Saccharum officinarum delta-1-pyrroline-5-carboxylate synthetase (P5CS) mRNA, complete cds	Saccharum offic...	3767	3767	96%	0.0	98.84%	2151	EU005373.2

Reason for similarity:

These might happen because they are closely related and belong to the same plant family, Poaceae (the grass family). They likely evolved from a common ancestor, which explains the high sequence similarity. These species also often grow in similar environments and face similar stresses that need to survive by relying on proline production, which is controlled by the P5CS gene.

3. P5CS1



Predicted P5CS1 tertiary structure is globular with multiple domains, rich in alpha-helices and likely beta-sheets. Modeled by SWISS-MODEL using a high-identity (79.19%), high-coverage (99%) AlphaFold template from *Oryza sativa* (O04226, GMQE 0.89).

P5CS2

Group Members: Davon Perry Nugroho (2702240380), Fadya Annura Manan (2702311894), Izhar Octafirlian Susilo (2702222144)



Predicted P5CS2 tertiary structure is globular with multiple domains, containing many alpha-helices and beta-sheets. Modeled by SWISS-MODEL using a very high-identity (87.71%), high-coverage (99%) AlphaFold template from *Oryza sativa* (O04226, GMQE 0.91).

database can protein data be obtained through websites like Protein Data Bank (PDB) (<https://www.rcsb.org/>), UniProt (<https://www.uniprot.org/>), NCBI Protein Database (<https://www.ncbi.nlm.nih.gov/protein/>), AlphaFold Protein Structure Database (<https://alphafold.ebi.ac.uk/>)

Predicting protein shape from a nucleotide sequence:

First, the gene's coding sequence (CDS) is translated into its amino acid sequence using the genetic code and online bioinformatics tools. The resulting amino acid sequence is then input into protein structure prediction tools (software or web servers). Common methods include homology modeling, which uses known structures of similar proteins as templates (with SWISS-MODEL using AlphaFold structures).

4. The protein synthesized is 1-pyrroline-5-carboxylate synthetase, a bifunctional enzyme that is useful in proline biosynthesis (an amino acid). P5CS has glutamate kinase (GK) and γ -glutamyl phosphate reductase (GPR) activities, which convert glutamate into proline and speed up (catalyze) the process. Proline acts as a protective substance that protect the cellular components, including as an osmoprotectant by maintaining water balance during drought, high salinity (a condition where there's too much salt in the soil), and other abiotic stresses such as extreme temperatures, nutrient deficiency, and harmful UV radiation. It also stabilizes proteins, membranes, and scavenging reactive oxygen species (ROS) to help plants tolerate environmental stress and can be degraded to provide energy and reduce power during the recovery from plant's stress.
5. A. Heuristic Database Searching
Used BLASTn, a heuristic algorithm, to quickly identify sequences similar to sugarcane P5CS genes in other plant species.
B. Homology, Similarity, and Identity
 - Homology: Inferred evolutionary relationships between P5CS genes across species.
 - Similarity and Identity: Quantified the extent of sequence conservation using BLASTn results.C. Local and Global Alignment
 - Local Alignment: Employed by BLASTn to find regions of high similarity within sequences.
 - Global Alignment: Could be used for aligning entire sequences to assess overall similarity, though not directly applied here.