



**UNIVERSIDADE ESTADUAL DA PARAÍBA
CAMPUS I – CAMPINA GRANDE/PB
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
DEPARTAMENTO DE BIOLOGIA
CURSO DE GRADUAÇÃO EM LICENCIATURA EM CIÊNCIAS BIOLÓGICAS**

THAYNARA SILVA RAMOS

**REDE DE INTERAÇÕES PROTEÍNA-PROTEÍNA DA ENZIMA LAAO EM
Gluconacetobacter diazotrophicus E SEUS EFEITOS EM PLANTAS HOSPEDEIRAS**

CAMPINA GRANDE-PB

2023

THAYNARA SILVA RAMOS

**REDE DE INTERAÇÕES PROTEÍNA-PROTEÍNA DA ENZIMA LAAO EM
Gluconacetobacter diazotrophicus E SEUS EFEITOS EM PLANTAS HOSPEDEIRAS**

Trabalho de Conclusão de Curso apresentando ao Curso de Ciências Biológicas da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de Graduação em Licenciatura em Ciências Biológicas.

Orientadora: Prof. Dra. Maria José Lima da Silva

Coorientadora: Dra. Nathalia Máira Cabral de Medeiros

CAMPINA GRANDE – PB

2023

É expressamente proibido a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano do trabalho.

R175r Ramos, Thaynara Silva.
Rede de interações proteína-proteína da enzima LAO em *Gluconacetobacter diazotrophicus* e seus efeitos em plantas hospedeiras [manuscrito] / Thaynara Silva Ramos. - 2023.
74 p.

Digitado.

Trabalho de Conclusão de Curso (Graduação em Ciências Biológicas) - Universidade Estadual da Paraíba, Centro de Ciências Biológicas e da Saúde, 2024.

"Orientação : Profa. Dra. Maria José Lima da Silva, Coordenação de Curso de Biologia - CCBS. "

"Coorientação: Profa. Dra. Nathalia Maira Cabral de Medeiros, . "

1. Crescimento. 2. Endofítica. 3. Interação. 4. L-aminoácido oxidase. I. Título

21. ed. CDD 570

THAYNARA SILVA RAMOS

**REDE DE INTERAÇÕES PROTEÍNA-PROTEÍNA DA ENZIMA LAAO EM
Gluconacetobacter diazotrophicus E SEUS EFEITOS EM PLANTAS HOSPEDEIRAS**

Trabalho de Conclusão de Curso apresentando ao Curso de Ciências Biológicas da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de Graduação em Licenciatura em Ciências Biológicas.

Aprovada em: 30___/11___/2023_____.

BANCA EXAMINADORA



Prof. Dra. Maria José Lima da Silva (Orientadora)
Universidade Estadual da Paraíba (UEPB)



Prof. Me. Izamara Gesiele Bezerra de Oliveira
Universidade Estadual da Paraíba (UEPB)



Prof. Dr. Carlos Henrique Salvino Gadelha de Meneses
Universidade Estadual da Paraíba (UEPB)

RESUMO

A bactéria *Gluconacetobacter diazotrophicus* é endofítica e fixadora de nitrogênio encontrada colonizando o interior de caules e raízes de culturas como arroz e cana-de-açúcar. As cepas de *G. diazotrophicus* são reconhecidas por estimular a produção de hormônios vegetais na planta hospedeira, incluindo a auxina. Pesquisas apontam que essa substância atua em sinergia com a fixação biológica de nitrogênio para potencializar o crescimento da cana-de-açúcar, quando em interação com *G. diazotrophicus*. O gene dessa bactéria foi identificado como responsável pela codificação da enzima L-aminoácido oxidase (LAAO), que tem um papel direto na biossíntese de ácido indolacético (IAA), uma variante de auxina através da via IPyA. Esta enzima é produzida pela bactéria em questão durante a interação com a planta hospedeira. O presente estudo tem como objetivo elucidar o papel da enzima Putative L-amino-acid oxidase de *Gluconacetobacter diazotrophicus* PAL5 no contexto das interações proteína-proteína, ressaltando nas implicações dessas interações no desenvolvimento e produtividade vegetal e identificar potenciais aplicações agroeconômicas. Para isso, foi utilizado a plataforma STRING e Uniprot para a análise da rede de interação proteína-proteína revelando um panorama complexo. Desse modo, foram identificadas proteínas com funções variadas, como a D-isomer specific 2-hydroxyacid dehydrogenase NAD-binding e a Beta-ketoacyl synthase, que enfatizam a multifuncionalidade de *G. diazotrophicus* em processos chave, como fixação de nitrogênio e síntese de ácidos graxos. A análise de genes próximos no genoma sugere uma regulação coordenada entre genes como quinona oxidoreductase subunidade I, L-amino-acid oxidase, citocromo c e endoribonuclease, com múltiplas implicações na resposta celular e adaptativa a condições ambientais. A análise dos clusters na rede de interação proteína-proteína revelou quatro grupos com características e funções distintas. O Grupo 1, com diversidade em processos biológicos, tem um papel crucial no metabolismo energético e biossíntese. O Grupo 2 foca no metabolismo do esteroide e lipídico, essencial para a estrutura celular e sinalização. O Grupo 3, onde a LAAO foi identificada, e o Grupo 4, sem associações conhecidas, indicam funções menos exploradas ou únicas na célula. Os resultados apontam para uma interconexão complexa entre a biossíntese de IPyA, o metabolismo energético e lipídico, e a eficiência na produção de IAA, essencial para o crescimento vegetal. O metabolismo energético e de lipídeos parece influenciar a atividade da LAAO e, por extensão, a capacidade da bactéria em auxiliar o crescimento das plantas. Este estudo destaca a importância de *G. diazotrophicus* como agente promotor de crescimento vegetal e abre caminhos para futuras investigações sobre suas

aplicações biotecnológicas na agricultura. Apesar disso, estudos adicionais são necessários para elucidar mais profundamente essas interações e explorar seu potencial biotecnológico.

Palavras-chave: crescimento; endofítica; interação; L-aminoácido oxidase.

ABSTRACT

The bacterium *Gluconacetobacter diazotrophicus* is an endophytic, nitrogen-fixing bacterium found colonizing the interior of stems and roots of crops such as rice and sugar cane. Strains of *G. diazotrophicus* are recognized for stimulating the production of plant hormones in the host plant, including auxin. Research indicates that this substance acts in synergy with biological nitrogen fixation to enhance the growth of sugarcane, when in interaction with *G. diazotrophicus*. The gene from this bacterium was identified as responsible for encoding the enzyme L-amino acid oxidase (LAAO), which has a direct role in the biosynthesis of indoleacetic acid (IAA), a variant of auxin through the IPyA pathway. This enzyme is produced by the bacteria in question during interaction with the host plant. The present study aims to elucidate the role of the enzyme Putative L-amino-acid oxidase from *Gluconacetobacter diazotrophicus* PAL5 in the context of protein-protein interactions, highlighting the implications of these interactions on plant development and productivity and identifying potential agro-economic applications. For this, the STRING and Uniprot platforms were used to analyze the protein-protein interaction network, revealing a complex panorama. In this way, proteins with varied functions were identified, such as D-isomer specific 2-hydroxyacid dehydrogenase NAD-binding and Beta-ketoacyl synthase, which emphasize the multifunctionality of *G. diazotrophicus* in key processes, such as nitrogen fixation and acid synthesis fatty. The analysis of nearby genes in the genome suggests a coordinated regulation between genes such as quinone oxidoreductase subunit I, L-amino-acid oxidase, cytochrome c and endoribonuclease, with multiple implications in the cellular and adaptive response to environmental conditions. Cluster analysis in the protein-protein interaction network revealed four groups with distinct characteristics and functions. Group 1, with diversity in biological processes, has a crucial role in energy metabolism and biosynthesis. Group 2 focuses on sterol and lipid metabolism, essential for cellular structure and signaling. Group 3, where LAAO was identified, and Group 4, with no known associations, indicate less explored or unique functions in the cell. The results point to a complex interconnection between IPyA biosynthesis, energy and lipid metabolism, and the efficiency in IAA production, essential for plant growth. Energy and lipid metabolism appear to influence LAAO activity and, by extension, the bacteria's ability to support plant growth. This study highlights the importance of *G. diazotrophicus* as a plant growth-promoting agent and opens avenues for future investigations into its biotechnological applications in agriculture. Despite this, additional studies are needed to further elucidate these interactions and explore their biotechnological potential.

Keywords: growth; endophytic; interaction; L-amino acid oxidase.

LISTA DE FIGURAS

FIGURA 1 – Rede de interação proteína-proteína da plataforma STRING.....	26
FIGURA 2 – Função Molecular (GO).....	35
FIGURA 3 – Clusters identificados e relacionados com a rede de interação de L-amino- acid oxidase	39
FIGURA 4 – Rede de interação da proteína L-amino-acid oxidase de <i>G. diazotrophicus</i> PA1 5, representando as vias KEGG	41
FIGURA 5 – Vizinhança gênica na bactéria <i>G. diazotrophicus</i>.....	46
FIGURA 6 – Vizinhança gênica da rede de interação da L-amino-acid oxidase com as demais proteínas	47
FIGURA 7 – Co-ocorrência entre genomas da rede de interação da L-amino-acid oxidase com as demais proteínas	50
FIGURA 8 – Cluster gerada pelo STRING	52
FIGURA 9 – Clusters que evidenciam processos biológicos	54
FIGURA 10 – Clusters evidencia as funções moleculares (GO).....	56
FIGURA 11 – Cluster evidencia as vias KEGG.....	58
FIGURA 12 – Clusters evidencia a localização sub-celular	59
FIGURA 13 – Clusters evidencia as palavras-chave	60
FIGURA 14 – Clusters evidenciam os domínios de proteínas (Pfam).....	62

LISTA DE TABELAS

Tabela 1 – Componentes e significado dos elementos envolvidos em uma rede de interação	19
Tabela 2 – Principais bancos de dados de redes de interação	20
Tabela 3 – Informações encontradas no STRING relacionados a interações proteína-proteína	21
Tabela 4 – Informações sobre a rede de proteínas gerada pela plataforma STRING	27
Tabela 5 – Informações gerais sobre as estatísticas de rede fornecidas pela plataforma STRING	32
Tabela 6 – Processo Biológico (Ontologia Gênica-GO).....	34
Tabela 7 – Clusters identificados e relacionados com a rede de interação L-amino-acid oxidase	36
Tabela 8 – Vias KEGG	39
Tabela 9 – Informações gerais sobre as proteínas do cluster	52

SUMÁRIO

1	INTRODUÇÃO	10
2	REVISÃO DE LITERATURA	12
2.1	Bactérias Promotoras de Crescimento Vegetal	12
2.1.1	Fixação Biológica de nitrogênio	12
2.2	Interação bactérias-planta	13
2.2.1	Gluconacetobacter diazotrophicus PA1 5	15
2.2.2	Auxina	16
2.3	Redes de Interação Proteína-Proteína	18
2.3.1	Componentes de uma rede de interação	18
2.3.2	STRING	20
2.3.3	Cluster	22
3	OBJETIVOS	24
4	MATERIAIS E METÓDOS	25
4.1	Identificação da Sequência e sua análise	25
4.2	Construção de interação	25
4.3	Rede de interação e Clusters	25
5	RESULTADOS E DISCUSSÃO	26
5.1	Rede de interação proteína-proteína - LAAO e seus parceiros	26
5.2	Análise da estatística da rede de interação	31
5.3	Análise do enriquecimento funcional na rede de interação para LAAO	33
5.4	Análise das interações entre os componentes do interactoma de L-amino-acid-oxidase	41
5.4.1	Base de dados	42
5.4.2	Vizinhança no Genoma	43
5.4.3	Coocorrência entre genomas	47
5.5	Cluster	51
6	CONCLUSÃO	66
	REFERÊNCIAS	68

1 INTRODUÇÃO

As plantas são colonizadas por uma diversidade de bactérias capazes de se multiplicarem e ocuparem diversos nichos ecológicos, incluindo aqueles presentes nas raízes. Esta interação resulta em uma relação simbiótica colaborativa (Galvão, 2012). Estas bactérias, conhecidas como bactérias promotoras de crescimento vegetal, exercem influência benéfica no desenvolvimento das plantas (Rodrigues et al., 2016). Elas atuam através de mecanismos como aumento da disponibilidade de nutrientes, produção de fitormônios, e atenuação dos efeitos nocivos de patógenos (Galvão, 2012).

As Bactérias Promotoras do Crescimento Vegetal (BPCVs) são frequentemente reconhecidas como biofertilizantes, especialmente quando promovem o crescimento vegetal através do aumento da disponibilidade ou do acesso a nutrientes pelas plantas (Galvão, 2012). Uma das funções mais críticas das BPCVs é a fixação biológica de nitrogênio, processo que fornece nitrogênio essencial às plantas. Esta atividade impacta significativamente a produtividade vegetal, influenciando a biomassa e a produção de grãos (Galvão, 2012).

Um exemplo notável dentro do grupo das BPCVs é a *Gluconacetobacter diazotrophicus*, uma bactéria endofítica fixadora de nitrogênio encontrada colonizando o interior de caules e raízes em culturas como arroz e cana-de-açúcar (Sabino, 2007; Rodrigues et al., 2016). Esta bactéria, atuando como modelo endofítico, tem sido fundamental para o estudo das interações planta-bactéria em hospedeiros não leguminosos. A simbiose entre *G. diazotrophicus* e suas plantas hospedeiras demonstra um impacto positivo na promoção do crescimento vegetal, conforme evidenciado por Rodrigues et al. (2016)

As cepas de *G. diazotrophicus* induzem a produção dos hormônios vegetais nas plantas a qual interação, isso inclui giberelinas, citocininas e auxinas. O hormônio giberelina atua regulando processos-chave nas plantas; como germinação de sementes, alongamento de raízes e brotos, floração e padrão de frutos (Binenbaum et al., 2018). As citocininas, promovem o desenvolvimento da parte aérea através do controle do ciclo celular do meristema apical do caule e também estimulam a fotossíntese através de diferentes mecanismos fisiológicos (Cortleven et al., 2019). Embora a função das auxinas na interação *G. diazotrophicus*-cana-de-açúcar não seja bem compreendida, estudos de inoculação sugeriram que a auxina bacteriana promove o crescimento da cana-de-açúcar em conjunto com a fixação biológica de nitrogênio (FBN) (Rodrigues et al., 2016).

A relação entre a bactéria *G. diazotrophicus* e a auxina, particularmente o ácido indol-3-acético (IAA), foi estudada em um trabalho publicado por Rodrigues et al. (2016). Este estudo

identificou que um gene específico, responsável pela codificação de uma enzima chamada L-amino-acid oxidase (LAAO), está diretamente envolvido na biossíntese de IAA na bactéria em questão. Importante ressaltar que o IAA é um fitormônio amplamente associado a efeitos benéficos na promoção do crescimento e na produtividade da cana-de-açúcar, que é uma das plantas hospedeiras de *G. diazotrophicus* (Rodrigues et al., 2016).

Aprofundando a análise da enzima L-amino-acid oxidase (LAAO), esta catalisa a oxidação de aminoácidos com configuração L, conduzindo à formação de ácidos alfa-ceto, amônia e peróxido de hidrogênio. A designação "L" em aminoácidos L origina-se da palavra latina "Laevus", que significa "esquerda", em referência à capacidade desses aminoácidos de girar a luz polarizada no sentido anti-horário. Esse fenômeno, denominado "atividade óptica", é quantificável por meio de um instrumento conhecido como polarímetro. É importante notar que a presença da LAAO não é restrita a bactérias; a enzima também foi identificada em uma gama de outros organismos, incluindo mamíferos, conforme documentado por Savino et al. (2021).

Além disso, a produção de peróxido de hidrogênio pela LAAO foi implicada na defesa contra bactérias e parasitas, como apontado por Castellano & Molinier-Frenkel (2017). Este aspecto é especialmente relevante para a relação simbiótica entre a bactéria *G. diazotrophicus* e suas plantas hospedeiras, uma vez que pode auxiliar na proteção das plantas contra diversos patógenos. Entretanto, é imperativo que pesquisas adicionais sejam conduzidas para elucidar completamente o papel da LAAO nesta complexa interação biológica.

Neste estudo, focando em aprofundar o entendimento da função da proteína putativa L-aminoácido oxidase (LAAO) em *G. diazotrophicus*, com ênfase especial em sua relação simbiótica com as plantas. Utilizamos técnicas para construir redes de interação específicas para a LAAO. Estas redes não só facilitam a identificação de proteínas associadas, mas também permitem a criação de grupos de proteínas, ou 'clusters', enriquecendo a análise funcional das interações.

O principal objetivo deste trabalho é elucidar a natureza da interação entre *G. diazotrophicus* PAL5 e plantas, investigando o papel da LAAO. Buscamos compreender como essa interação pode promover benefícios, como o aumento do crescimento e da produtividade vegetal. Tal compreensão é essencial, pois abre caminhos para avanços significativos e estratégias sustentáveis na agricultura. Adicionalmente, nossa metodologia, que se baseia na construção de redes de interação proteína-proteína, oferece uma ferramenta inovadora para explorar os mecanismos biológicos que fundamentam essa simbiose, contribuindo assim para o avanço em um campo de pesquisa de vital importância.

2 REVISÃO DE LITERATURA

2.1 Bactérias Promotoras de Crescimento Vegetal (BPCV)

As bactérias promotoras de crescimento vegetal (BPCV) são representadas por um grupo de microrganismos de vida livre, ao qual, tem o potencial de estimular o crescimento das plantas de maneira direta ou indireta (Glick et al., 2012). A promoção direta geralmente envolve a síntese de compostos pelas bactérias que facilitam a absorção de nutrientes pelas plantas (Glick et al., 1999; Galvão, 2012). Adicionalmente, estas bactérias podem ativar mecanismos de defesa nas plantas, aumentando sua resistência a doenças e estresses ambientais (Mariano et al., 2004).

Quando os BPCV promovem especificamente o crescimento vegetativo, por exemplo, aumentando o acesso das plantas aos nutrientes ou sua disponibilidade, ou ainda ampliando a área de superfície das folhas, são frequentemente classificados como biofertilizantes. Em contrapartida, esses microrganismos são conhecidos como agentes de controle biológico ou biopesticidas quando estimulam o crescimento vegetal controlando organismos prejudiciais. Além de proteger as plantas contra patógenos, os BPCV também fornecem nutrientes de fácil assimilação pelas plantas e atenuam estresses bióticos e abióticos (Rashid et al., 2016; Somers et al., 2004).

As BPCV mais conhecidas incluem membros do gênero *Azospirillum*, *Bacillus*, *Paenibacillus*, *Pseudomonas*, *Enterobacter*, *Burkholderia*, *Gluconacetobacter* e *Arthrobacter* (Girio, et al., 2015; Santos Dias & Santos). Algumas delas são encontradas na superfície das raízes enquanto outras invadem os tecidos das plantas sem causar sintomas aparentes de doença, quando são conhecidas como endofíticas (Galvão, 2012).

A transição para o tema subsequente conduz ao exame aprofundado de um aspecto crítico nas interações planta-microrganismo: a fixação biológica de nitrogênio. Este processo, executado por um subconjunto de BPCV, é indispensável para o ciclo global do nitrogênio e tem implicações significativas na nutrição vegetal. Ele influencia de maneira marcante o crescimento e a produtividade das plantas. Na próxima seção, será realizada uma exploração detalhada dos mecanismos subjacentes à fixação biológica de nitrogênio, com foco especial nos agentes microbiológicos envolvidos.

2.2.1 Fixação Biológica do nitrogênio

Certas bactérias são capazes de fixar nitrogênio atmosférico em plantas, um processo conhecido como fixação biológica de nitrogênio (FBN). Estas bactérias podem colonizar tanto

a rizosfera quanto os tecidos internos da planta, estabelecendo-se em nichos protegidos de concentrações elevadas de oxigênio e outros fatores adversos, condições essas que facilitam a expressão de seu potencial para a FBN (Marin et al., 1999). Adicionalmente, BPCV desempenham um papel vital na melhoria da nutrição vegetal, principalmente fornecendo nitrogênio, fósforo e ferro, nutrientes essenciais para o desenvolvimento das plantas.

Para que as bactérias sejam capazes de realizar a fixação biológica do nitrogênio (N^2), a presença da enzima nitrogenase é fundamental. Esta enzima desempenha um papel crucial ao catalisar a quebra da tripla ligação do N^2 atmosférico, reduzindo-o a amônia (NH^3). Esse processo é análogo à síntese de amônia na indústria, embora ocorra sob condições ambientais mais brandas. Bactérias que realizam este processo podem ser associativas, endofíticas ou simbióticas, dependendo do grau e do tipo de sua associação com as plantas. Estas relações variam desde uma associação não específica na rizosfera até uma simbiose intracelular complexa em órgãos especializados como os nódulos radiculares (Hungria et al., 2001).

Diversas bactérias têm a capacidade de fixar o nitrogênio atmosférico. Algumas das mais significativas incluem: *Azospirillum*, *Azotobacter*, *Gluconacetobacter*, *Enterobacter*, *Burkholderia* (Somers et al., 2004). Essas bactérias desempenham um papel fundamental na disponibilidade de nitrogênio para organismos vivos, contribuindo para a fertilidade do solo e para o crescimento saudável de plantas.

Além disso, a fixação biológica do nitrogênio é vital para a disponibilidade de nutrientes nas cadeias alimentares e desempenha um papel essencial na sustentabilidade dos ecossistemas, na produtividade agrícola e na manutenção da saúde do solo. Através desse processo, essas bactérias fornecem uma fonte de nitrogênio essencial para muitos sistemas agrícolas, reduzindo a dependência de fertilizantes químicos e promovendo práticas agrícolas mais sustentáveis.

Tendo delineado o papel fundamental das bactérias na fixação biológica do nitrogênio e sua contribuição para a nutrição vegetal, a fertilidade do solo, e a sustentabilidade agrícola, o foco agora se volta para a complexa interação planta-bactéria. Este tópico abordará a complexidade e a diversidade dessas interações, explorando como as bactérias estabelecem relações mutualísticas ou simbióticas com as plantas e como essas associações influenciam o crescimento e a saúde vegetal.

2.2 Interação bactéria-planta

Muitas bactérias têm a capacidade de colonizar plantas, multiplicando-se e ocupando diversos nichos nas raízes. Esta interação planta-bactéria pode manifestar-se de diferentes

formas. Em alguns casos, a presença bacteriana é neutra, não exercendo influência significativa no crescimento vegetal. Por outro lado, determinadas bactérias podem impactar negativamente o desenvolvimento das plantas, sendo então classificadas como patogênicas ou deletérias. Contrariamente, há bactérias que promovem o crescimento vegetal, exercendo um efeito positivo na saúde e no desenvolvimento das plantas; estas são conhecidas como bactérias promotoras de crescimento vegetal que já foram explanadas em tópico anterior deste trabalho (Galvão, 2012). Esta diversidade nas interações reflete a complexidade das relações ecológicas entre plantas e microorganismos, onde cada tipo de interação tem implicações distintas para a saúde das plantas e a produtividade agrícola.

As interações planta-microrganismo geralmente ocorrem principalmente em três lugares da planta: filosfera, endosfera e rizosfera. A filosfera está relacionada às partes aéreas (caule, folhas e flores ou frutos) e a endosfera com o sistema de transporte. A rizosfera, pode ser definida como qualquer volume de solo influenciado ou em associação com as raízes e o material produzido pelas plantas (Hinsinger et al., 2005). A extensão exata e a geometria desta zona não são fixas, variando em resposta a uma série de fatores, incluindo as condições ambientais e a natureza dos exsudados radiculares secretados pela planta. Estes exsudados, que incluem uma variedade de compostos orgânicos, desempenham um papel crucial na modulação das interações solo-raiz e na estruturação da comunidade microbiana da rizosfera (Hinsinger et al., 2005; Dzantor, 2007).

A rizosfera é a região do solo mais rica em bactérias do que todas as outras. Populações microbianas ao redor das raízes incluem bactérias, fungos, leveduras e protozoários. As concentrações de bactérias em substratos superiores podem conter até 10⁹ células por grama de solo. É uma área de intensa atividade biológica e química influenciada por compostos exsudados pela raiz, incluindo ácidos orgânicos, açúcares, aminoácidos e pequenos peptídeos, metabólitos secundários. Sendo comprovado que as comunidades bacterianas incluídas na planta estão associadas a nutrientes como carbono, fósforo, nitrogênio e enxofre, bem como a eliminação de toxinas e produção de fitohormônios ou antibióticos (Liu et al., 2017; Firdous et al., 2019).

Dentro do espectro das interações planta-bactéria, as bactérias endofíticas desempenham um papel notável, caracterizando-se por sua habilidade de colonizar os tecidos internos das plantas hospedeiras. Estas bactérias pertencem a diversos grupos taxonômicos, incluindo, mas não se limitando à, filos como Alfa-, Beta-, e Gamma-Proteobacteria, Actinobacteria, Firmicutes e Bacteroidetes (Bacon & Hinton, 2006; Miguel et al., 2021). A colonização por esses microorganismos pode ocorrer de diferentes maneiras: algumas espécies permanecem em

estado latente dentro dos tecidos da planta, enquanto outras se estabelecem e se multiplicam ativamente, afetando local ou sistemicamente o hospedeiro.

Esses microrganismos apresentam um elevado potencial para utilização na agricultura e na indústria, podendo ser empregados como vetores para introdução de genes de interesse em plantas, inibidores de patógenos e fontes de metabólitos primários e secundários de interesse. Assim, eles são importantes tanto na ciência básica quanto na aplicada, sendo de interesse fundamental o seu estudo.

Após explorar as interações planta-bactéria e a significativa influência que essas relações têm na saúde vegetal e na produtividade agrícola, torna-se imperativo aprofundar o estudo de um exemplo específico que exemplifica esta dinâmica complexa: o *Gluconacetobacter diazotrophicus* PA15. Este microrganismo, representando um caso de interação simbiótica com plantas, é reconhecido pela sua capacidade de fixar nitrogênio atmosférico, desempenhando um papel crucial no ciclo do nitrogênio e contribuindo significativamente para a nutrição vegetal. Na próxima seção, será abordado em detalhes o papel do *G. diazotrophicus* PA15 na agricultura, sua interação específica com as plantas hospedeiras, e o impacto dessa associação na eficiência agrícola e na sustentabilidade ambiental.

2.2.1 *Gluconacetobacter diazotrophicus* PAL 5

A bactéria *Gluconacetobacter diazotrophicus* PAL5, pertencente ao filo Proteobacteria, classe Alphaproteobacteria, ordem Rhodospirillales e família Acetobacteraceae, é capaz de colonizar o interior dos tecidos distribuindo-se nos espaços intercelulares e vasos do xilema de folhas, colmos e raízes (JAMES et al., 2001). A bactéria endofítica *G. diazotrophicus* que foi descrita por Cavalcante e Döbereiner em 1988 possui uma característica relevante que é sua capacidade de realizar a fixação biológica de nitrogênio, ou seja, converter o nitrogênio atmosférico em formas utilizáveis, como amônia.

Esta bactéria parece ser bem adaptada à vida dentro da cana-de-açúcar, uma vez que o seu melhor crescimento ocorre em pH 5,5 e na presença de 10% de sacarose, as mesmas condições encontradas dentro desta planta (Cavalcante & Döbereiner, 1988; Gillis & Kersters, 1989). *G. diazotrophicus* são bactérias Gram-negativas, com capacidade de fixar nitrogênio em condições microaerofílicas (Cavalcante & Döbereiner, 1998), isoladas em meio semi-sólido LGI-P (Reis & Olivares, 1994) que, após o período de sete a dez dias, apresentam uma película alaranjada, e o meio localizado abaixo da película torna-se incolor devido à assimilação do azul de bromotimol (Döbereiner et al., 1993). *G. diazotrophicus* é considerada um diazotrófico

aerotolerante no qual o oxigênio é fundamental para a geração de grandes quantidades de ATP requerida para a fixação de nitrogênio.

A *Gluconacetobacter diazotrophicus*, particularmente conhecida pela sua associação com plantas como a cana-de-açúcar, é um exemplo de bactéria endofítica. A cepa específica *Gluconacetobacter diazotrophicus* PAL.

5, isolada de raízes de cana-de-açúcar no Brasil, coloniza os tecidos internos das plantas, desempenhando um papel vital no fornecimento de nitrogênio às hospedeiras. Essa relação simbiótica é altamente benéfica, contribuindo significativamente para o crescimento e desenvolvimento das plantas e reduzindo a dependência de fertilizantes nitrogenados. Além de fixar nitrogênio, a *G. diazotrophicus* PAL5 também promove o crescimento vegetal por meio da produção de compostos bioativos, como fitormônios, e induzindo resistência a patógenos.

Destaca-se, em particular, a capacidade desta bactéria de produzir auxina, um fitormônio crucial para o crescimento das plantas. A biossíntese e o papel da auxina na interação *G. diazotrophicus*-planta serão explorados detalhadamente na próxima seção, ilustrando como este microrganismo contribui para a fisiologia vegetal.

2.2.2 Auxina

A produção de substâncias promotoras do crescimento vegetal pelas bactérias já foi relatada para muitas espécies bacterianas, contribuindo para os efeitos promotores do crescimento vegetal. As diferentes substâncias produzidas pelas BPCVs podem induzir mudanças na morfologia ou fisiologia das raízes que aumentam sua área superficial e taxa de respiração, desta forma, influenciando a captação de nutrientes e o crescimento vegetal, entre elas está a auxina (Beattie et al., 2006; Munõz, 2017).

As auxinas, cujo nome deriva do grego 'auxein', que significa 'crescer', são fitormônios essenciais que atuam como moléculas sinalizadoras móveis. Elas são transportadas ativamente ao longo de gradientes locais e por toda a planta, coordenando uma variedade de processos relacionados ao crescimento e desenvolvimento vegetal. A descoberta da auxina remonta aos anos 1930, quando foi determinado que o ácido indol-3-acético (AIA) era um representante principal desta classe de fitormônios. O aminoácido triptofano foi identificado como o precursor na biossíntese do AIA. Posteriormente, outras auxinas naturais, como o ácido fenilacético (Ácido Fenil-Acético) e o ácido 4-cloroindol-3-acético (Ácido 4-Cl-Indol-3-Acético), foram descobertas em plantas superiores. No entanto, o AIA permanece sendo a auxina mais abundante e com maior relevância fisiológica.

Na definição original, o termo 'auxina' englobava todas as substâncias químicas, tanto naturais quanto sintéticas, que promoviam o alongamento de coleótilos e segmentos de caules. Entretanto, as auxinas afetam muitos outros processos do desenvolvimento, além do alongamento celular (Taiz e Zeiger, 2009, p820). Assim, as auxinas podem ser definidas como compostos que, a nível vegetal, modulam diversos processos como as respostas trópicas à luz e à gravidade, formação de raízes, tropismos (fototropismo e geotropismo), florescimento (Davies, 1995).

A nível celular, as auxinas promovem a hiperpolarização da membrana plasmática e, provavelmente devido ao estímulo das H⁺-ATPase, levam à acidificação da parede celular, acarretando o seu afrouxamento. O turgor, o alongamento, a divisão e a diferenciação celular são outros fatores influenciados pelas auxinas a nível celular (Berleth e Sachs, 2001). O principal precursor das vias de biossíntese do IAA é o triptofano (trp), que é sintetizado em uma rota comum a três aminoácidos: triptofano, fenilalanina e tirosina. As vias de biossíntese dos aminoácidos aromáticos têm em comum a via do shiquimato (Pittard, 1996).

A maioria das rotas biossintéticas de IAA se iniciam a partir do precursor principal triptofano, entretanto, a possível presença de uma via independente desse hormônio não deve ser descartada. As rotas são nomeadas em função das moléculas intermediárias produzidas. É importante ressaltar que nem todas as vias são caracterizadas na mesma medida e que podem existir múltiplas vias em um mesmo organismo.

Em plantas, a síntese se inicia alguns dias após a germinação e a partir daí as principais fontes de IAA são as folhas jovens próximas ao ápice (Ljung et al., 2002). Além das folhas jovens, IAA também ocorre em órgãos que estão crescendo ativamente, tais como meristemas apicais da parte aérea, e frutos em desenvolvimento, embora também possa ser produzido em folhas maduras e nos ápices radiculares, onde o nível de produção é usualmente baixo (Davies, 1995).

Na via do ácido indol-3-pirúvico (IPyA), o triptofano é inicialmente convertido em IPyA, processo que pode ocorrer através de uma transaminação mediada por uma trp aminotransferase ou de uma oxidação catalisada por uma L-aminoácido oxidase. Posteriormente, o IPyA é descarboxilado pela enzima indol-3-piruvato descarboxilase (IPDC), formando indol-3-acetaldeído (IAAld). Este, por sua vez, é oxidado a IAA, ácido indol-3-acético, seja pela IAAld oxidase ou pela aldeído desidrogenase (Costacurta e Vanderleyden, 1995; Patten e Glick, 1996). Em um estudo de inoculação, Rodrigues et al. (2016) sugeriram que a promoção do crescimento da cana-de-açúcar por *G. diazotrophicus* pode estar relacionada à liberação deste fitormônio.

2.3 Redes de Interação Proteína-Proteína

A bioinformática é uma área interdisciplinar que combina a biologia, a computação e a informática para estudar e analisar dados biológicos complexos. Desempenhando um papel fundamental na compreensão dos processos biológicos ao aplicar métodos computacionais avançados para analisar e interpretar informações genéticas, moleculares e celulares.

Dentro desse contexto, as redes de interação proteína-proteína são uma representação visual e computacional das interações entre as proteínas em um organismo. Essas redes são construídas a partir de dados experimentais ou preditivos e fornecem uma representação gráfica das interações físicas e funcionais que ocorrem entre as proteínas. Na perspectiva da bioinformática, a análise de redes de interação de proteínas permite uma compreensão abrangente e sistemática das relações entre as proteínas, revelando padrões, complexidade e propriedades emergentes dos sistemas biológicos (Bruckner, 2012).

A construção e análise de redes de interação de proteínas envolvem o uso de técnicas bioinformática, como a integração de dados de interação proteína-proteína obtidos de experimentos, a previsão de interações por meio de algoritmos computacionais e a visualização dos resultados por meio de ferramentas especializadas. Essa abordagem bioinformática permite explorar as interações entre as proteínas em uma escala global, identificar proteínas-chave, módulos funcionais e vias de sinalização e compreender a organização e a dinâmica dos sistemas biológicos (Walport et al., 2021).

A importância das redes de interação de proteínas na compreensão dos processos biológicos é significativa. Essas redes, de significativa importância na compreensão dos processos biológicos, oferecem uma visão integrada das interações proteína-proteína, ajudando a revelar como as proteínas se comunicam, coordenam suas funções e participam de processos celulares e metabólicos. Ao identificar e analisar os nós (proteínas) e as conexões (interações) da rede, é possível descobrir proteínas centrais que desempenham papéis-chave na regulação e na transmissão de informações. Além disso, as redes de interação de proteínas são ferramentas valiosas para a descoberta de novos alvos terapêuticos, o estudo de doenças complexas e o desenvolvimento de medicamentos (Bruckner, 2012).

2.3.1 Componentes de uma rede de interação

Em uma interação proteína-proteína, existem vários componentes essenciais que desempenham papéis específicos na estrutura e funcionalidade da rede. A seguir, estão representados na (Tabela 1) os principais componentes de uma rede de interação de proteínas e seus significados.

Tabela 1 - Componentes e significação dos elementos envolvidos em uma rede de interação.

Componente	Definição
Proteínas	São os elementos essenciais de uma rede de interação, desempenhando várias funções importantes nas células.
Nós	Representam as proteínas individuais na rede de interação proteína-proteína.
Arestas	São as conexões que representam as interações entre as proteínas.
Interações físicas	Relacionam-se às ligações físicas diretas entre as proteínas na rede, envolvendo ligações entre domínios estruturais, sítios de ligação ou outras regiões específicas das proteínas
Interações químicas	Referem-se a associações indiretas entre as proteínas, com base nas funções ou participação em vias metabólicas ou processos biológicos em comum.
Grau	É o número de conexões (arestas) que o nó possui.
Módulos	São grupos de proteínas altamente interconectadas dentro da rede.
Centralidade	Trata-se de uma medida que avalia a importância relativa de um nó na rede.

Fonte: Oliveira, 2023.

Para construir redes de interação proteína-proteína, é essencial buscar informações detalhadas sobre estas interações. Estas informações são geralmente armazenadas em bancos de dados, que compilam dados oriundos de uma variedade de fontes, incluindo estudos científicos, experimentos de laboratório e predições baseadas em modelos computacionais. Estes repositórios proporcionam uma riqueza de dados sobre as interações entre proteínas em diversos organismos e sob variadas condições ambientais. O acesso a esses bancos de dados é fundamental para entender as complexas redes de interações proteicas e suas implicações biológicas.

A seguir na Tabela 2, é destacado alguns dos principais bancos de dados e repositórios utilizados na pesquisa de interação proteína-proteína.

Tabela 2 -Principais bancos de dados de redes de interação.

Banco de dados	Descrição	Site
STRING (Search Tool for the Retrieval of Interacting Genes/Proteins)	Banco de dados que integra informações sobre interações proteína-proteína, associações funcionais e redes de interação proteínas.	https://string-db.org/
BioGRID (Biological General Repository for Interaction Datasets)	Repositório que abriga informações sobre interações de proteínas.	https://thebiogrid.org/
IntAct	Banco de dados sobre interações proteína-proteína. Incluindo dados de diferentes tipos de experimentos.	https://www.ebi.ac.uk/intact/
MINT (Molecular Interaction Database):	Banco de dados que oferece informações detalhadas sobre interações proteína-proteína, e interações moleculares.	https://mint.bio.uniroma2.it/

Fonte: Oliveira, 2023.

Neste estudo, o foco será no banco de dados STRING, que será detalhado mais adiante. O STRING é uma ferramenta essencial para a análise de interações proteína-proteína, fornecendo informações valiosas que serão exploradas em profundidade nas seções subsequentes.

2.3.2 STRING

A plataforma "STRING: functional protein association networks" é uma ferramenta poderosa para a análise de interações de proteínas e associações funcionais. Visando integrar

as informações e consolidar dados de associação proteína-proteína previstos para muitos organismos. As conexões no STRING incluem interações diretas (físicas), bem como interações indiretas (funcionais). Além disso, a plataforma coleta e reavalia dados experimentais disponíveis sobre interações proteína-proteína e importa vias conhecidas e complexos proteicos de bancos de dados selecionados.

Para utilizar a plataforma STRING, os usuários podem acessar o site oficial em <https://string-db.org/> e inserir as proteínas de interesse. A STRING oferece uma variedade de recursos, incluindo a visualização de redes de interação proteína-proteína, análise funcional, predição de interações e integração com outras ferramentas bioinformáticas, facilitando uma compreensão abrangente das interações e funções proteicas.

Para isso, o banco de dados STRING utiliza várias fontes de dados para obter informações sobre as associações entre as proteínas, os quais serão destacados na Tabela 3.

Tabela 3 - Informações encontradas no STRING relacionados a interações proteína-proteína.

Informações encontradas no STRING	Descrição
Dados Experimentais/Bioquímicos	São informações obtidas por meio de experimentos ou abordagens bioquímicas para investigar interações proteína-proteína e ligações funcionais.
Associação em banco de dados	Possui informações sobre interações proteína-proteína e anotações funcionais de bases de dados específicas, enriquecendo a compreensão das proteínas em contextos biológicos.
Publicações de referência	Agrupa proteínas com base em sua menção em resumos de literatura científica indexados no PubMed, fornecendo evidências de possíveis ligações funcionais entre elas.
Vizinhança gênica	Representa as interações previstas com base na proximidade genética, ou seja, a co-localização de genes no genoma.

Co-expressão	Indica que dois ou mais genes têm padrões de expressão semelhantes em condições experimentais, tecidos ou estágios de desenvolvimento.
Co-ocorrência entre genomas	São interações previstas com base na co-ocorrência de genes em diferentes organismos e experimentos.
Fusões genéticas	Representa as interações previstas com base em fusões gênicas, onde dois genes adjacentes se fundem em um único gene.

Fonte: Oliveira, 2023.

Esses recursos do STRING auxiliam na investigação das interações proteína-proteína e fornecem informações valiosas sobre a função e o contexto biológico das proteínas (Jensen et al., 2009).

2.3.3 Cluster

Dentro das análises realizadas pela plataforma STRING, a funcionalidade de 'cluster' desempenha um papel crucial ao fornecer informações sobre as funções biológicas e moleculares das proteínas. Esta ferramenta permite o agrupamento de proteínas que apresentam interações e similaridades funcionais significativas. Os clusters identificados na rede de interação proteína-proteína representam grupos de proteínas que estão fortemente conectadas e desempenham funções semelhantes dentro de um contexto biológico específico (Szklarczyk et al., 2015).

Utilizando a opção 'cluster', é possível identificar agrupamentos de proteínas que interagem frequentemente ou que formam complexos proteicos específicos. Estes clusters são particularmente valiosos para mapear vias metabólicas, processos biológicos e funções celulares comuns às proteínas do grupo. A análise de clusters facilita a compreensão das relações funcionais em um contexto biológico mais amplo, oferecendo insights sobre a organização e dinâmica das redes proteicas.

Além disso, ao agrupar proteínas em clusters, o STRING emprega algoritmos avançados de análise de rede para identificar sub-redes ou módulos densamente interconectados dentro da rede global de interação proteína-proteína. Estes clusters não apenas revelam a organização

funcional das proteínas em um contexto biológico, mas também fornecem informações valiosas sobre vias metabólicas, interações celulares e processos fisiológicos envolvidos (Szkłarczyk et al., 2015)."

Embora o algoritmo K-means seja frequentemente empregado na criação de clusters no contexto do STRING, outros algoritmos, como o DBSCAN (Density- Based Spatial Clustering of Applications with Noise), podem ser mais apropriados em determinadas situações. A escolha do algoritmo depende das características do conjunto de dados e das necessidades específicas do problema em questão (Szkłarczyk et al., 2015).

O K-means é um dos algoritmos utilizados na área de clusterização, que é uma técnica de análise de dados. Tendo como objetivo agrupar objetos ou dados similares em conjuntos chamados clusters, facilitando a identificação de padrões, estruturas e relações dentro dos dados (Oliveira, 2023).

O algoritmo K-means é baseado na ideia de encontrar K centróides (pontos centrais) que representam os clusters. Ele começa selecionando aleatoriamente K pontos como centróides iniciais. Em seguida, ele atribui cada objeto ao cluster cujo centróide está mais próximo. Após essa atribuição, os centróides são recalculados com base nos objetos pertencentes a cada cluster. O processo é repetido até que os centróides não se movam significativamente ou até que um critério pré-definido seja atingido (Oliveira, 2023).

Em suma, a funcionalidade de 'cluster' na plataforma STRING, complementada pela aplicação de algoritmos como K-means e DBSCAN, é indispensável para desvendar a complexidade das interações proteína-proteína. O emprego desses algoritmos permite a identificação eficiente de padrões, estruturas e relações intrínsecas aos conjuntos de dados proteicos. Essas técnicas de clusterização são cruciais para agrupar objetos ou dados similares, facilitando significativamente a identificação de padrões e relações funcionais nas redes proteicas. Essa análise aprofundada de clusters não apenas esclarecem mais as vias metabólicas e processos celulares, mas também abre caminho para novas perspectivas e descobertas no campo da bioinformática e biologia molecular.

3 OBJETIVOS

Objetivo Geral: Elucidar o papel da enzima Putative L-amino-acid oxidase de *Gluconacetobacter diazotrophicus* PA15 no contexto das interações proteína-proteína, focando nas implicações dessas interações no desenvolvimento e produtividade vegetal e identificar potenciais aplicações agroeconômicas.

Objetivos Específicos:

1. Utilizar abordagens bioinformáticas para identificar proteínas parceiras que possam interagir com a Putative L-amino-acid oxidase, com o intuito de explorar a complexidade das interações proteína-proteína.
2. Empregar a plataforma STRING para construir uma rede de interações entre a Putative L-amino-acid oxidase e suas proteínas parceiras, visando elucidar as relações funcionais em um contexto biológico.
3. Enriquecer a rede de interações criadas e delimitar clusters, permitindo um aprofundamento das relações funcionais entre os componentes da rede.
4. Investigar as funções biológicas das proteínas que interagem na rede construída, com especial atenção às implicações dessas interações no desenvolvimento vegetal e aumento de produtividade vegetal.

4 MATERIAIS E MÉTODOS

4.1 Identificação da sequência e sua análise

No contexto deste estudo, a sequência da proteína Putative L-amino-acid oxidase foi obtida no NCBI (The National Center for Biotechnology Information - <https://www.ncbi.nlm.nih.gov/>) com o número de acesso no GenBank: CAP56399. Para obter informações mais abrangentes sobre a proteína, realizou-se uma análise no UniProt (<https://www.uniprot.org/>) utilizando o número de identificação B5ZEA3.

4.2 Construção de interação

A partir da sequência fasta do gene Putative L-amino-acid oxidase, foi construída uma rede de interação proteína-proteína utilizando a plataforma STRING (<https://string-db.org/>). Essa rede foi analisada para identificar os componentes envolvidos na interação com a proteína Putative L-amino-acid oxidase, bem como descrever as interações observadas.

As interações encontradas em relação à proteína Putative L-amino-acid oxidase foram verificadas e descritas, fornecendo informações valiosas sobre as possíveis vias e processos biológicos envolvidos na função da proteína Putative L-amino-acid oxidase.

4.3 Rede de interação e Clusters

A metodologia empregada neste estudo envolveu a construção de uma rede de interação utilizando a plataforma STRING. Preliminarmente, foi criada a estrutura básica da rede, em seguida adicionados nós extras para enriquecer suas conexões. Esse procedimento foi repetido três vezes, resultando em uma rede mais complexa e abrangente. Posteriormente, foi aplicado o algoritmo “Kmeans clustering” disponível na plataforma STRING, com intuito de agrupar os elementos da rede em quatro clusters distintos. Essa abordagem possibilitou a identificação e análise de padrões de interação entre os componentes da rede, assim como a identificação de grupos específicos de nós com características semelhantes.

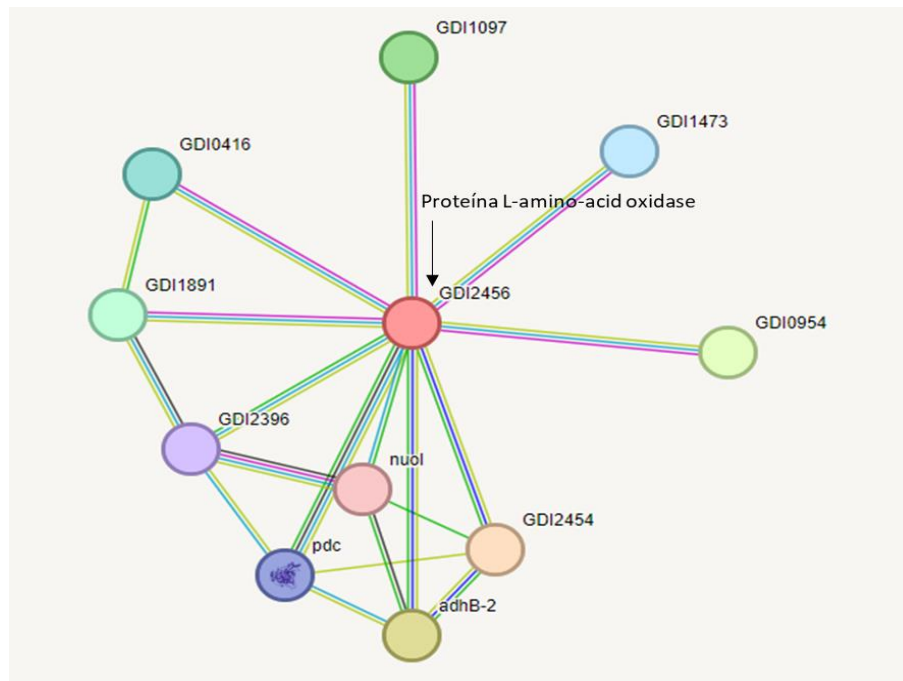
5 RESULTADOS E DISCUSSÃO

5.1 Rede de interação proteína-proteína - LAAO e seus parceiros.

A rede de interações resultante da análise conduzida por meio da plataforma String apresenta a proteína focal deste estudo, a "Putative L-amino-acid oxidase precursor" (LAAO), bem como seus parceiros potenciais de interação. A representação gráfica dessa rede é ilustrada na Figura 1, na qual a LAAO é centralizada e é notadamente identificada pela coloração vermelha.

Os detalhes específicos dos parceiros da proteína em análise, conforme retratados na rede de interações (Figura 1), são abordados na Tabela 4. A referida tabela proporciona informações descritivas abrangendo o nome das proteínas, suas respectivas funções e, adicionalmente, suas identificações (IDs) nos bancos de dados Uniprot e String.

Figura 1 - Rede de interação proteína-proteína da plataforma String LAAO.



Fonte: Elaborada pela autora a partir dos dados do STRING (2023)

Cada esfera representa um nó, ou seja, uma proteína. A esfera vermelha central é a proteína alvo deste estudo (Putative L-amino-acid oxidase precursor). Já as outras esferas coloridas indicam as proteínas na qual L-amino-acid oxidase realiza interações.

Tabela 4 - Informações sobre a rede de proteínas gerada pela plataforma STRING.

Tabela de Proteínas			
ID String	ID Uniprot	Nome da Proteína	Função
GDI2456	B5ZEA3	Putative L-amino-acid oxidase	Molecular: Atividade Oxirredutase Genes DA L -aminoácido oxidase estão envolvidos na biossíntese bacteriana de IAA.
GDI0416	B5ZJG8	D-isomer specific 2-hydroxyacid dehydrogenase NAD-binding	Ligação NAD
GDI1097	A9HD25	Proteína não caracterizada - Hipotética	-
GDI2396	B5ZDU8	Beta-cetoacil sintase	Função molecular: atividade de 3-oxoacil-[acil-carrier-protein] sintase; Atividade oxirredutase; Ligação de Fosfopantetéia; Ligação de íon de zinco; Processo biológico: Processo biossintético de ácidos graxos
GDI2464 nuol	B5ZEB0	NADH-quinona oxidoreductase subunidade I	Atividade catalítica
GDI1473	B5ZE44	D-isomer specific 2-hydroxyacid dehydrogenase NAD-binding	Função molecular: Ligação NAD.

GDI0954	A9HC45	Proteína hipotética, ainda não tem uma função caracterizada.	-
GDI1891	B5ZJX5	Pyridoxal-dependent decarboxylase	Cofator; Função molecular: Atividade de carboxi-liase; ligação de fosfato de piridoxal; Processo metabólico de aminoácidos.
GDI2454	B5ZEA1	Endoribonuclease L-PSP	Atividade catalítica
GDI0172 pdc	B5ZEK7	Thiamine pyrophosphate protein TPP binding domain protein	Atividade carboxi-liase; Ligação de íons de magnésio; Ligação pirofosfato de tiamina; Processo metabólico de moléculas pequenas.
GDI2455 adhB-2	B5ZEA2	Citocromo c classe I	Cofator; Componente celular: respirassoma. Função molecular: Atividade de transferência de eletrons; Ligação heme; Ligação íon Ferro.

Fonte: Elaborada pela autora a partir dos dados do STRING (2023).

Sendo representado o código no string e uniprot. Bem como o nome da proteína e suas respectivas funções.

Os parágrafos seguintes se dedicam à elaboração ampliada das descrições pertinentes aos mencionados parceiros, os quais foram previamente apresentados na Tabela 4 e, por conseguinte, na rede de interações (Figura 1). Essa abordagem terá o propósito de estabelecer correlações substanciais entre esses parceiros e a LAAO (B5ZEA3), incluindo a análise de sua importância no contexto da bactéria *Gluconacetobacter diazotrophicus*.

As proteínas B5ZJG8 e B5ZE44, presente na rede de interação acima, são caracterizadas como D-isomer specific 2-hydroxyacid dehydrogenase NAD-binding (figura 1, tabela 4). Tal

enzima catalisa a redução reversível de cetonas a hidroxiácidos correspondentes pela oxidação simultânea do dinucleotídeo de nicotinamida e adenina (NAD⁺). Além disso, essa enzima é um membro da família de desidrogenases específicas para D-isômeros de 2-hidroxiácidos (2HADHs) que contém uma ampla gama de oxidoreduases com vários papéis metabólicos, bem como aplicações biotecnológicas (Matelska et al., 2023).

Em relação às bactérias *G. diazotrophicus*, a enzima D-isomer specific 2-hydroxyacid dehydrogenase NAD-binding desempenha um papel crucial no metabolismo do D-lactato, que é gerado durante o processo de fixação de nitrogênio (Matelska et al., 2023). É importante esclarecer que, embora esta enzima e a Putative L-amino-acid oxidase (LAAO) não apresentem uma relação direta, ambas são classificadas como flavoenzimas e atuam na catalisação da desaminação oxidativa de aminoácidos. Deste modo, embora as funções específicas dessas enzimas possam diferir, suas atividades catalíticas são essencialmente similares em termos bioquímicos.

Outra proteína que figura na rede de interações e que poderia ser um potencial parceiro de interação com a L-amino-acid oxidase (LAAO) é a B5ZDU8, também conhecida como Beta-ketoacyl synthase. Esta enzima desempenha um papel fundamental na síntese de ácidos graxos para uso pelo organismo, esses usos incluem armazenamento de energia e criação de membranas celulares. Além de estar envolvida em uma variedade de outras funções metabólicas e celulares. As enzimas da classe β -cetoacil-ACP sintases são responsáveis pela reação de alongamento inicial na via metabólica em questão. Elas utilizam acetil-CoA como molécula inicial, ou primer, para alongar o malonil-ACP em duas unidades de carbono. Posteriormente, estas enzimas catalisam o alongamento do substrato acil-ACP graxo em duas unidades adicionais de carbono (Pettersen et al., 2019).

Avançando na análise, a proteína B5ZEA2, que integra a rede de interações e é classificada dentro da categoria Cytochrome c, desempenha um papel crítico na transferência de elétrons e tem uma ampla distribuição entre bactérias. Esta enzima participa de uma série de processos bioquímicos fundamentais, como a respiração celular e a neutralização de peróxido de hidrogênio (H₂O₂). Adicionalmente, o citocromo c comumente interage com outras enzimas redox, servindo como um ponto crucial de entrada ou saída de elétrons nos respectivos ciclos catalíticos (Bertini et al., 2006).

No contexto das bactérias *G. diazotrophicus*, a relevância do citocromo c (B5ZEA2) se torna ainda mais evidente. Estudos revelaram que mutações nos genes responsáveis pela maturação do citocromo c resultaram em cepas bacterianas capazes de produzir apenas 5% da auxina gerada pela cepa selvagem. Esses dados sugerem que uma proteína do citocromo c pode

estar implicada na produção de ácido indol-3-acético (IAA) por *G. diazotrophicus* (Rodrigues et al., 2016). Contudo, mais investigações são necessárias para elucidar completamente este papel.

Outro possível parceiro de interação com LAAO identificado na rede é a proteína B5ZEB0, também conhecida como NADH-quinona oxidoreductase subunidade I. Essa enzima de membrana integral, que possui múltiplas subunidades, está presente tanto em bactérias como em organelas eucarióticas. O Complexo I desta enzima é responsável pela transferência reversível de elétrons do NADH solúvel para a quinona ligada à membrana, acoplando esta reação à geração de uma força motriz de prótons (PMF). Em eucariontes, essa enzima contribui para cerca de 40% do PMF utilizado na síntese de ATP. Estudos também apontam que, em bactérias, o Complexo I exerce diversas funções, influenciando tanto as condições de crescimento como processos metabólicos variados (Spero et al., 2015).

A seguir, temos a proteína B5ZEA1 indicada na rede de interação, também identificada como Endoribonuclease L-PSP (proteína solúvel em ácido perclórico de fígado). Esta proteína faz parte da família YER057c/YjgF/UK114. A família de proteínas YER057c/YjgF/UK114 é altamente conservada nos três domínios da vida, mas atualmente não possui uma função bioquímica consensual (Mistiniene et al., 2003; Burman et al., 2007; Christopherson et al., 2008). Acredita-se que membros dessa família estejam envolvidos na biossíntese de isoleucina (Yeo7, Ibm1, aldR) e purina (YjgF), bem como na degradação anaeróbica de treonina (tdcF) e manutenção do DNA mitocondrial (Ibm1) (Mistiniene et al., 2003; Burman et al., 2007; Christopherson et al., 2008). Essa família de proteínas forma um homotrímero com um sítio de ligação a polipeptídeos (Lambrecht et al., 2010). No entanto, a função exata dessas proteínas ainda não foi determinada.

Além disso, a proteína B5ZJX5, classificada como Pyridoxal-dependent decarboxylase, é outro elemento relevante na rede de interação analisada. Essa coenzima possui uma versatilidade catalítica notável e está envolvida em uma ampla variedade de reações bioquímicas, como transaminação, descarboxilação, desaminação, interconversão de aminoácidos L e D, e racemização (Kumar, 2016).

Outra proteína que faz parte da rede de interação é a B5ZEK7, indicada como Thiamine pyrophosphate protein TPP binding domain protein. O pirofosfato de tiamina (TPP) é uma coenzima derivada da vitamina B1 (tiamina). A síntese de TPP em eucariotos requer tiamina pirofosfoquinase (TPK), que catalisa a transferência de um grupo pirofosfato do ATP para a tiamina. A TPP é essencial para os processos metabólicos centrais, incluindo a formação de acetil CoA a partir da glicose e o ciclo de Krebs (Timm., et al 2001). Tiamina é um

microelemento essencial sintetizado por plantas e microrganismos. A forma ativa da tiamina, o pirofosfato de tiamina (TPP), desempenha um papel proeminente nas atividades metabólicas, particularmente como cofator enzimático (Subki et al., 2020).

Por fim, as proteínas A9HC45 e A9HD25, indicadas na rede de interação, são classificadas como enzimas hipotéticas. Por ainda não terem funções claramente definidas, isso sublinha a necessidade de pesquisas adicionais para elucidar seus respectivos papéis bioquímicos e funcionais.

As conclusões baseadas na análise dos parceiros de interação da enzima Putative L-amino-acid oxidase (LAAO) em *Gluconacetobacter diazotrophicus* revelam um panorama complexo de interações proteicas com implicações significativas no metabolismo e na fisiologia bacteriana que pode ter seus efeitos na interação com o organismo vegetal. As proteínas identificadas, como a D-isomer specific 2-hydroxyacid dehydrogenase NAD-binding e a Beta-ketoacyl synthase, demonstram a multifuncionalidade de *G. diazotrophicus*, a qual processos como a fixação de nitrogênio e a síntese de ácidos graxos são cruciais. Embora a relação direta entre a LAAO e estas proteínas não seja explicitamente clara, a natureza complementar de suas funções sugere um panorama metabólico interconectado que potencializa o crescimento vegetal e a produtividade.

Adicionalmente, a integração de proteínas como o Cytochrome c e a NADH-quinona oxidoreductase subunidade I na rede de interações ressalta a complexidade das vias metabólicas envolvidas na respiração celular e na geração de energia. A presença de enzimas como a Endoribonuclease L-PSP e a Pyridoxal-dependent decarboxylase, com funções ainda não totalmente elucidadas, abre caminhos para novas investigações que podem revelar aspectos desconhecidos da biologia de *G. diazotrophicus*. Estes achados, juntamente com a descoberta de enzimas hipotéticas, sublinham a importância de explorar a rede de interações proteicas para uma compreensão mais completa das capacidades metabólicas e potencial biotecnológico de *G. diazotrophicus*. Os tópicos posteriores terão como escopo aprofundar e delimitar as interações compartilhadas com as proteínas aqui identificadas, além de analisar a confiabilidade estatística da rede de interações geradas.

5.2 Análise da estatística da rede de interação

A análise estatística da rede de interação, disponível na seção 'Análises' da plataforma STRING, oferece informações cruciais para decifrar a complexidade inerente aos sistemas biológicos. Tais informações estão expostas abaixo na Tabela 5.

Tabela 5 - Informações gerais sobre as estatísticas de rede fornecidas pela plataforma STRING.

Estatística de rede	
Número de nós: 11	Número esperado de arestas: 18
Número de arestas: 19	Valor P de enriquecimento PPI: 0,424
Grau médio do nó: 3,45	Média coeficiente de agrupamento local: 0,761

Fonte: Elaborada pela autora a partir dos dados do STRING (2023).

Na tabela 5 é evidenciado o número de nós presentes na rede de interação, nesse caso existem 11 nós demonstrado na rede. Tais nós fazem referência ao número total de proteínas (ou genes) presente na rede gerada a partir da proteína analisada.

A análise estatística forneceu informações sobre o número de arestas, conforme apresentado na Tabela 5. Em uma rede de interação, as arestas simbolizam as relações entre os nós, que, neste contexto, representam proteínas ou genes. Cada aresta estabelece uma ligação entre dois nós, indicando uma interação específica entre eles. Os detalhes sobre os tipos dessas interações serão analisados e descritos posteriormente neste trabalho.

No caso da rede de interação proteína-proteína em estudo, foram identificadas 19 arestas. Por outro lado, o "número esperado de arestas" refere-se ao total de arestas que se espera encontrar em uma rede aleatória com o mesmo número de nós. Na rede ilustrada na Figura 1, esse número é de 18 arestas. Isso sugere uma variação mínima, de apenas uma aresta, entre o número observado e o número esperado de interações na rede.

Corroborando com esses dados, a tabela 5 fornece o 'valor-p de enriquecimento de PPI'. Este parâmetro avalia se as interações dentro da rede são mais frequentes do que o que seria esperado ao acaso. Um valor-p baixo (inferior a 0,05) geralmente indica que as interações na rede são significativas. O número 0,424 sugere que a rede de interações de proteínas não é estatisticamente enriquecida em comparação com uma rede aleatória, o que significa que as interações observadas podem ocorrer ao acaso.

A Tabela 5 também apresenta dados sobre o grau médio de nó, que é uma medida do número médio de conexões que cada nó estabelece dentro da rede. Este valor é especificado como 3,45 na tabela. Um grau médio de nó elevado pode ter significância biológica, indicando nós altamente interconectados. Contudo, é importante observar que um valor de grau excessivamente alto pode ser um sinal de ruído nos dados. No caso da rede de interação que

estamos estudando, essa preocupação não se aplica, uma vez que o grau médio observado não foi excessivamente alto.

Além disso, a tabela 5 apresenta o coeficiente de agrupamento local médio da rede analisada. Tal grau indica como os nós em uma rede tendem a agrupar-se, de modo que possibilite identificar conjuntos funcionalmente relacionados das proteínas que compõem a rede de interação. Os valores mais próximos de 1 deste coeficiente indicam um alto nível de agrupamento. A rede que está sendo descrita possui o valor 0,761 como coeficiente local médio, indicando que, em média, as proteínas na rede têm uma alta probabilidade de formar grupos de interações.

Com base nesses dados, fica evidente que a rede não apresenta um número significativamente maior de interações do que o inicialmente esperado. Isso pode sugerir que o conjunto de proteínas analisado é relativamente pequeno ou que consiste essencialmente em uma coleção aleatória de proteínas pouco interconectadas. No entanto, isso não exclui a possibilidade de que as proteínas selecionadas sejam biologicamente significativas. Pode ser que as proteínas da bactéria em questão ainda não tenham sido amplamente estudadas, o que limitaria nosso conhecimento sobre suas possíveis interações ou parceiros protéicos. Assim, essas interações ainda podem não estar registradas no banco de dados STRING. Essa situação ressalta a importância de expandir o conjunto de dados em análise ou de explorar outras fontes de informação para validar as interações biológicas identificadas.

5.3 Análise do enriquecimento funcional na rede de interação para LAAO

O estudo das redes de interação biológica oferece uma visão holística dos complexos sistemas biológicos, permitindo-nos entender como as diferentes entidades (tais como genes, proteínas ou metabólitos) interagem entre si para formar caminhos ou complexos funcionais. No entanto, a mera identificação dessas interações muitas vezes não é suficiente para fornecer um entendimento completo de sua relevância biológica. É aqui que a análise de enriquecimento funcional se torna crucial.

O conceito de "enriquecimento funcional" refere-se à identificação estatística de funções biológicas, processos ou estruturas celulares que estão representados em um conjunto particular de entidades, em comparação com o que seria esperado em uma distribuição aleatória. Essa análise não só ajuda a validar os resultados da rede de interação, como também proporciona insights adicionais sobre os papéis funcionais e os mecanismos subjacentes dessas associações na rede.

As tabelas geradas por essas análises estão apresentadas abaixo. No entanto, é fundamental a prévia descrição de duas colunas de grande importância que correspondem a valores estatísticos relacionados ao enriquecimento. A primeira delas é denominada 'Força'. Essa medida descreve o tamanho do efeito do enriquecimento, revelando o quão forte é a associação entre o termo biológico específico e as proteínas na rede sob estudo. Quanto maior o valor nessa coluna, mais significativo é o enriquecimento. Um valor positivo indica que o termo em questão está mais presente na rede estudada do que seria esperado ao acaso.

A segunda coluna refere-se à 'Taxa de Descoberta Falsa', indicando a significância do enriquecimento. Quanto menor o valor nessa coluna, mais robusto é o resultado. Trata-se de uma medida estatística que incorpora correções para prevenir falsas descobertas de associações. Valores baixos apontam que o enriquecimento é estatisticamente confiável e menos suscetível a interpretações equivocadas.

Um dos resultados apresentados referente a rede de interação da proteína aqui estudada se refere ao Gene Ontology (GO) (Tabela 6). GO é um recurso extensivamente utilizado em bioinformática que oferece um vocabulário padronizado para descrever genes e proteínas em termos de suas funções biológicas associadas, processos celulares em que estão envolvidos e os componentes celulares onde se localizam. Em relação ao GO, observa-se a classificação das proteínas da rede de acordo com a função molecular, nesse quesito destaca-se a atividade remodeladora da cromatina dependente de ATP, encontrada em duas enzimas hipotéticas (GDI1097) e (GDI0954) conforme destacada nas esferas em vermelho na figura 2.

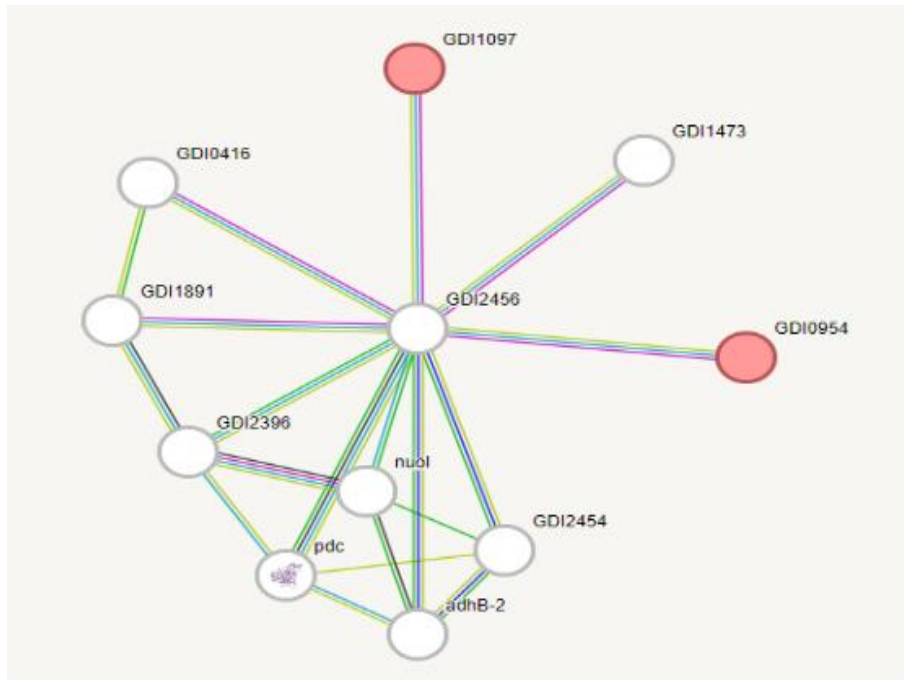
Tabela 6 - Processo biológico (Ontologia gênica-GO).

Enriquecimento funcional da rede de interação para LAAO			
Função molecular (Gene Ontology)			
Termo - GO	Descrição	Força	Taxa de descoberta falsa
IR: 0140658	Atividade remodeladora da cromatina dependente de ATP	2.54	0,0321

Fonte: Elaborada pela autora a partir dos dados do STRING (2023).

Nesta tabela é possível ver quais processos a rede de interação estão envolvidas com a L-amino-acid oxidase.

Figura 2 - Função molecular (Gene Ontology).



Fonte: Elaborada pela autora a partir dos dados do STRING.

A cor vermelha corresponde às proteínas que estão envolvidas na atividade remodeladora da cromatina dependente de ATP.

Na plataforma STRING, na seção de análise, a referida "Agrupamento Local na Rede" é representada na (Tabela 7) e diz respeito à identificação de agrupamentos de proteínas que interagem entre si em regiões específicas de uma rede de interações de proteínas. Esses conjuntos de proteínas são comumente denominados clusters e denotam uma concentração de interações entre as proteínas em áreas localizadas dentro da rede. Os "Agrupamentos Locais na Rede" fazem referência a conjuntos de proteínas que já são previamente conhecidos por interagirem entre si e que estão pré-determinados ou caracterizados como agrupamentos na rede de interações de proteínas, disponíveis no banco de dados da plataforma STRING. Esses agrupamentos representam conjuntos de proteínas que possuem relações biológicas ou funcionais específicas e são identificados com base em evidências experimentais ou em dados já conhecidos. A presença desses clusters indica uma probabilidade maior de que as proteínas contidas em cada cluster estejam envolvidas em funções biológicas semelhantes ou participem de vias metabólicas ou processos celulares específicos.

Tabela 7 - Clusters identificados e relacionados com a rede de interação L-amino-acid oxidase.

Agrupamentos Locais na Rede		
Descrição	Força	Taxa de descoberta
Misto, incluindo Proteína contendo domínio CBS CBSX3, domínio CBS e Amino oxidase	2.31	0,00085
Misto, incluindo 2-hidroxiácido desidrogenase específico do isômero D, domínio de ligação a NAD e hidrolase semelhante haloácido dehalogenase	1,73	0,0076
Fosforilação oxidativa e ligação Heme	1.29	0,0088

Fonte: Elaborada pela autora a partir dos dados do STRING.

Um dos cluster identificados na rede de interação da proteína putativa L-aminoácido oxidase foi “Proteína contendo domínio CBS CBSX3, domínio CBS e Amino oxidase” sendo representado na Figura 3 pela coloração verde as proteínas L-amino-acid oxidase (GDI2456), Endorribonuclease (GDI2454) e Citocromo C (GDI2455). O domínio proteico CBS, também conhecido como Cystathionine Beta-Synthase, é um domínio protéico altamente conservado ao longo da evolução e está presente no proteoma de arqueobactérias, procariotos e eucariotos. Esses domínios CBS são frequentemente encontrados em repetições em tandem e estão presentes em proteínas com diversas funções, tanto em compartimentos citosólicos como de membrana, desempenhando papéis variados, tais como enzimas metabólicas, quinases e canais iônicos (Ignoul & Eggermont, 2005). O domínio CBS é importante por sua capacidade de ligação a nucleotídeos, como o ATP, e pode ocorrer em diferentes repetições dentro de uma única proteína. Sua função principal envolve a regulação de moléculas de adenosina trifosfato (ATP) e/ou sulfato. Esse domínio desempenha um papel fundamental em várias vias metabólicas, incluindo a regulação da produção de adenosina trifosfato (ATP) e a resposta a mudanças nas concentrações de adenosina trifosfato (ATP) e sulfato (Ignoul & Eggermont, 2005).

A relação do domínio CBS em bactérias foi abordada em um estudo conduzido por Huynh et al. (2016). Este estudo indicou que esse domínio desempenha um papel na regulação da importação de solutos compatíveis, como a carnitina, em resposta ao estresse osmótico. Essa regulação é mediada pelo segundo mensageiro c-di-AMP, que interage com o domínio CBS.

Essa interação é de suma importância para a capacidade da bactéria se adaptar às mudanças nas condições ambientais.

Os domínios CBS estão correlacionados com atividade de crescimento e desenvolvimento da planta, mediante a via regulação do sistema de tioredoxina, bem como na resposta ao estresse abiótico e biótico (Mou et al., 2015). No mesmo estudo, um domínio CBS contendo o gene da proteína, *OsCBSX3*, é funcionalmente caracterizado na resistência do arroz contra *Magnaporthe oryzae* (*M. oryzae*), atuando como regulador positivo.

O segundo cluster identificado na análise de enriquecimento está está identificado na figura 3 pela coloração azul as duas proteínas 2-hydroxyacid dehydrogenase (GDI0416) e (GDI1473) bem como a Pyridoxal-dependent decarboxylase (GDI1891) sendo relacionadas ao agrupamento misto, de 2-hidroxiácido desidrogenase específico do isômero D. A família de proteínas das d-2-hidroxiácido desidrogenases inclui várias enzimas, como a d-lactato desidrogenase, a d-glicerato desidrogenase, a d-3-fosfoglicerato desidrogenase e a proteína H de resistência à vancomicina. Essa família de proteínas pode estar envolvida em diversas vias metabólicas, incluindo a degradação de aminoácidos, biossíntese de aminoácidos e o ciclo do ácido cítrico (TCA) (Grant, 1989).

No contexto de plantas e bactérias, as 2-hidroxiácido desidrogenases desempenham papéis importantes em várias vias metabólicas. Em plantas, essas enzimas podem estar envolvidas na degradação de aminoácidos, sendo essenciais para a mobilização de nitrogênio durante o desenvolvimento das plantas. Já em bactérias, essas enzimas podem estar relacionadas à fermentação e à utilização de diferentes fontes de carbono (Kumsab et al., 2020).

Na área de plantas, a d-2-hidroxiácido desidrogenases tem um papel significativo no metabolismo de ácidos graxos e pode estar envolvida na produção de energia durante a germinação de sementes e no desenvolvimento de plântulas. Além disso, ela pode desempenhar um papel na degradação de aminoácidos específicos (Kumsab et al., 2020).

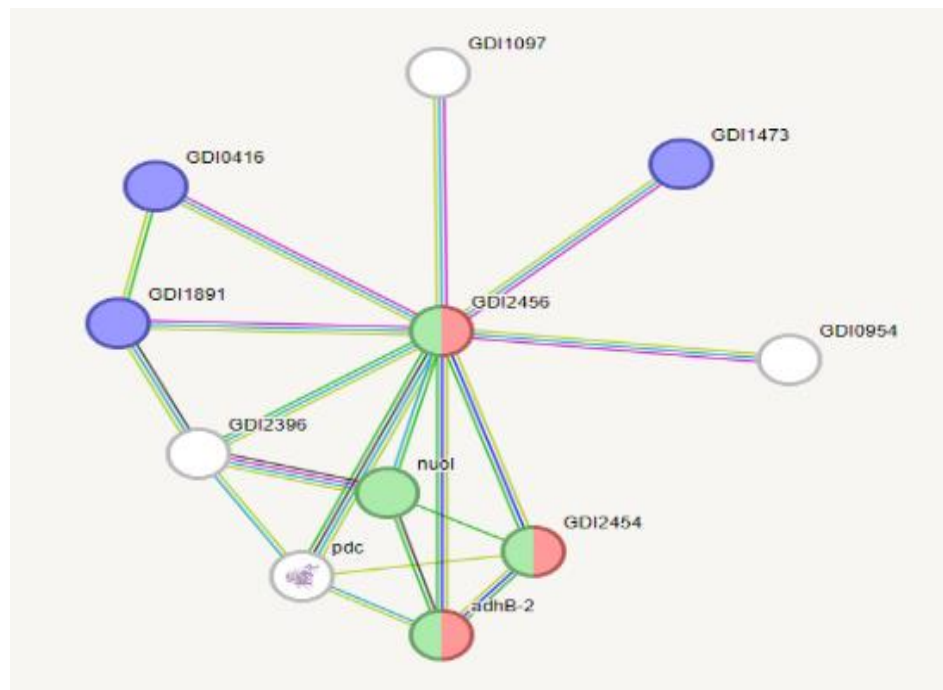
Em bactérias, as D-2-hidroxiácido desidrogenases são cruciais para a degradação de substratos específicos, como o D-2-hidroxiácido da isoleucina, um aminoácido. Essas enzimas podem participar de vias metabólicas relacionadas à utilização de fontes alternativas de carbono ou à produção de compostos intermediários necessários para o crescimento bacteriano.

O último Cluster identificado está relacionado a fosforilação oxidativa e ligação do grupo heme, indicados na figura 3 pela coloração vermelha as proteínas são Putative L-amino-acid oxidase (GDI2456), Endorribonuclease (GDI2454), NADH-quinona oxireductase Subunidade I (GDI2464) e Citocromo C (GDI2455). Esse grupo heme é um anel heterocíclico

de porfirina (tetrapirrol) que coordena no centro um átomo de ferro em sua forma reduzida (Fe^{2+}). O ferro coordenado pelo heme gera um centro redox ativo que confere a este grupo prostético, quando ligado a proteínas, a capacidade de transferir elétrons em citocromos da cadeia de transportes de elétrons; transportar oxigênio; atuar no ciclo catalítico de enzimas (Lima, 2020).

Em conclusão, a análise dos clusters da rede de interação da proteína putativa L-aminoácido oxidase revelou múltiplas conexões envolvendo domínios protéicos altamente conservados como o CBS e enzimas chave como as 2-hidroxiácido desidrogenases, assim como grupos relacionados à fosforilação oxidativa e à ligação do grupo heme. Cada um desses clusters representa uma faceta importante de sistemas biológicos complexos, seja na regulação metabólica, na resposta ao estresse ambiental ou na transferência de elétrons. O domínio CBS, por exemplo, ressalta a complexidade evolutiva e funcional da rede, já que sua presença e funcionalidade se estendem desde arqueobactérias até eucariotos, influenciando diversas vias metabólicas e mecanismos de regulação. As 2-hidroxiácido desidrogenases, por outro lado, têm implicações diretas tanto no metabolismo bacteriano quanto vegetal, ligando assim as redes de interações em diferentes organismos. Finalmente, a presença do grupo heme em um dos clusters sugere o envolvimento dessas proteínas em processos vitais de transferência de elétrons e transporte de oxigênio. Estas múltiplas facetas não apenas reforçam a complexidade intrínseca dessas redes protéicas, mas também sublinham a necessidade de estudos futuros para decifrar as interações funcionais e significados biológicos mais amplos.

Figura 3 - Clusters identificados e relacionados com a rede de interação de L-amino-acid oxidase.



Fonte: Elaborada pela autora a partir dos dados do STRING.

Cada cor representa um agrupamento de proteínas que interagem entre si em um local específico. Cor verde - Misto, incluindo Proteína contendo domínio CBS CBSX3, domínio CBS e Amino oxidase; Cor azul - Misto, incluindo 2-hidroxiácido desidrogenase específico do isômero D, domínio de ligação a NAD e hidrolase semelhante haloácido dehalogenase; Cor vermelha - Fosforilação oxidativa e ligação heme.

Outro resultado da análise de enriquecimento está relacionado às vias do KEGG (tabela 8), que se referem a um conjunto de informações sobre vias metabólicas e de sinalização presentes no banco de dados KEGG (Kyoto Encyclopedia of Genes and Genomes). O KEGG é uma fonte de dados amplamente utilizada que oferece informações detalhadas sobre as interações e funções de genes e proteínas.

Tabela 8 - Vias KEGG..

Vias KEGG		
Descrição	Força	Taxa de descoberta falsa
Metabolismo da tirosina	1,8	0,0367
Metabolismo do triptofano	1,72	0,0367

Metabolismo de fenilalanina	1,69	0,0367
Metabolismo da glicina, serina e treonina	1.36	0,0366
Biossíntese de metabólitos secundários	0,66	0,0367
Vias metabólicas	0,38	0,0367

Fonte: Elaborada pela autora a partir dos dados do STRING.

Nesta tabela é possível ver quais vias fazem parte das proteínas analisadas.

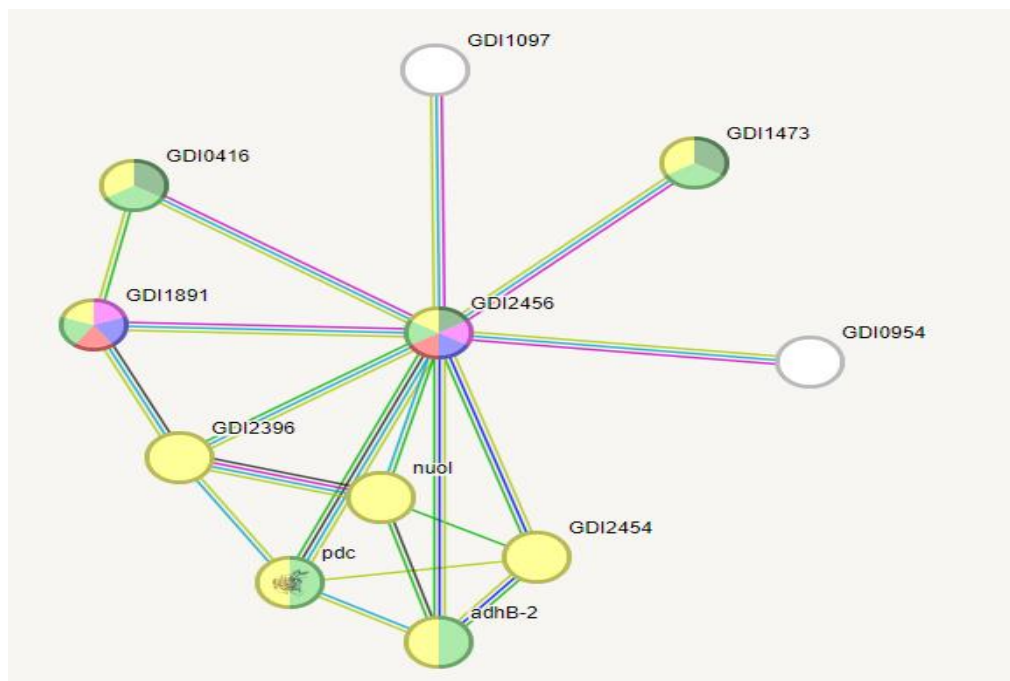
Na análise da via KEGG, constatou-se que a enzima L-aminoácido oxidase (GDI2456) tem atuação em todas as vias metabólicas representadas no KEGG, conforme ilustrado na Figura 4 pela esfera central que exibe diversas colorações, cada uma correspondente a uma via específica listada na tabela anterior. Em relação ao metabolismo de glicina, serina e treonina, assim como à biossíntese de metabólitos secundários e outras vias metabólicas, essas áreas são abordadas pelas duas enzimas 2-hydroxyacid dehydrogenase (GDI1473) e (GDI0416). Já a proteína Pyridoxal dependent decarboxylase (GDI1891) está envolvida no metabolismo de tirosina, triptofano e fenilalanina, além de também participar da biossíntese de metabólitos secundários e outras vias metabólicas. As enzimas Beta cetoacil sintase (GDI2396), NADH-quinona oxireductase (GDI2464) e Endorribonuclease (GDI2454) atuam unicamente em vias metabólicas gerais. Por último, as proteínas pdc Thiamine pyrophosphate protein (GDI0172) e Citocromo C (GDI2455) são ativas na biossíntese de metabólitos secundários e em outras vias metabólicas.

O metabolismo de aminoácidos como tirosina, triptofano, fenilalanina, glicina, serina e treonina, identificados na análise de enriquecimento, desempenha um papel crucial na produção de proteínas fundamentais, incluindo hormônios vegetais. Em particular, o metabolismo do triptofano (Trp) serve como precursor principal para a biossíntese de ácido indolacético (IAA), uma variante de auxina. Este aminoácido é sintetizado por uma rota inicialmente compartilhada com outros dois aminoácidos: fenilalanina e tirosina (Galvão, 2012; Gomes, 2020).

Ainda referente a via KEGG vale destacar a tirosina, que faz parte de uma importante família de enzimas de sinalização celular. Além disso, serina e treonina têm sido descritas em diversas bactérias patogênicas como enzimas essenciais em vias de sinalização que dependem de fosforilação (Manegatti, 2014). Este conjunto de informações sublinha a importância desses aminoácidos e suas vias metabólicas associadas, tanto em contextos vegetais quanto bacterianos.

Por fim, com relação à coluna chamada ‘força’ encontrada na tabela 8, é possível verificar que os valores das vias do metabolismo de tirosina (1,8), triptofano (1,72) fenilalanina (1,69), glicina, serina e treonina (1,36). São relativamente altos e positivos, informando que o enriquecimento pode ser considerado significativo. Mostrando que há uma maior presença desse termo na rede em relação ao esperado por acaso. Já no que concerne aos valores da taxa de descoberta falsa, todas as vias possuem o valor (0,0367), exceto uma que tem o valor (0,0366) inferindo que o enriquecimento é estatisticamente confiável.

Figura 4 - Rede de interação da proteína putative L-amino-acid oxidase de *G. diazotrophicus* PA1 5, representando as vias KEGG.



Fonte: Elaborada pela autora a partir dos dados do STRING (2023).

Cada cor corresponde a uma via KEGG: Amarela - vias metabólicas; verde claro - biossíntese de metabólitos secundários; vermelho - metabolismo de fenilalanina; azul - metabolismo de triptofano; rosa - metabolismo de tirosina; verde escuro - metabolismo de serina, glicina e treonina.

5.4 Análise das interações entre os componentes do interactoma de L-amino-acid-oxidase

No contexto da rede de interação proteína-proteína analisada pelo STRING, os resultados forneceram perspectivas valiosas sobre as interações específicas (arestas) entre a proteína em estudo e suas parceiras associadas. Essas informações permitem fazer inferências significativas, que serão detalhadamente apresentadas e discutidas a seguir:

5.4.1 Base de dados

Em relação às evidências em banco de dados para interações entre parceiros e a L-amino-acid oxidase, os resultados indicam a ausência de banco de dados específicos em *G. diazotrophicus*. No entanto, em outros organismos foi encontrada a presença de interações entre parceiros em bancos de dados relacionadas a proteínas homólogas das proteínas da rede aqui estudada. Observou-se que as proteínas homólogas da rede—tais como putative L-amino acid oxidase, citocromo c e putative descarboxilase de tirosina—estão envolvidas em vias metabólicas e na biossíntese de metabólitos secundários em *Acetobacter ascendens*. Em *Burkholderia stagnalis*, também foi identificada as proteínas putative L-amino-acid oxidase e putative descarboxilase de tirosina, que por sua vez, estão envolvidas nas mesmas vias metabólicas mencionadas anteriormente para *A. ascendens*.

Ao examinar as características de *Acetobacter ascendens* e *Burkholderia stagnalis*, é importante considerar suas semelhanças e diferenças com *Gluconacetobacter diazotrophicus*, especialmente em termos de interação com plantas. *A. ascendens* é notável por sua habilidade em oxidar etanol em ácido acético, um processo vital na indústria de alimentos, particularmente na produção de vinagre (Qiu et al., 2021). Contudo, diferentemente de *G. diazotrophicus*, não existem evidências claras de uma relação simbiótica específica entre *A. ascendens* e plantas. Esta lacuna destaca uma diferença fundamental no metabolismo e na interação ambiental de *A. ascendens*.

Burkholderia stagnalis, por outro lado, faz parte do versátil gênero *Burkholderia*, reconhecido por seu potencial em promover o crescimento de plantas por meio de diversos mecanismos. Apesar das informações específicas sobre a simbiose de *B. stagnalis* com plantas serem limitadas, o gênero *Burkholderia* é conhecido por contribuir para o crescimento e proteção das plantas, através de mecanismos como a produção de hormônios vegetais, exopolissacarídeos e sensor de densidade (quorum sensing) (Santos et al., 2022). Estas interações são fundamentais para a agricultura e gestão ambiental, demonstrando a ampla gama de interações ecológicas das *Burkholderia* em comparação com *G. diazotrophicus*.

A análise evolutiva destas bactérias revela adaptações únicas a nichos ecológicos específicos. *G. diazotrophicus* evoluiu para uma associação simbiótica estreita com plantas, especializando-se na fixação de nitrogênio. Em contraste, *A. ascendens* e *B. stagnalis* desenvolveram nichos ecológicos e metabólicos distintos, refletindo adaptações evolutivas a ambientes específicos e necessidades metabólicas variadas, influenciando suas interações com plantas e outros organismos no ecossistema.

Com base nos resultados e análises realizadas, é viável propor a hipótese de que as proteínas putative L-amino acid oxidase, citocromo c e putative descarboxilase de tirosina em *Gluconacetobacter diazotrophicus*, e seus equivalentes em *Acetobacter ascendens* e *Burkholderia stagnalis*, desempenham um papel fundamental na adaptação dessas bactérias ao ambiente. Tal adaptação pode influenciar a interação dessas bactérias com plantas, especialmente através do papel dessas proteínas nas vias metabólicas e na biossíntese de metabólitos secundários, principalmente relacionando o gênero *Burkholderia* com a *G. diazotrophicus*. Estas proteínas estão potencialmente envolvidas na produção de hormônios vegetais, antibióticos ou pigmentos, influenciando significativamente a interação bacteriana com o ambiente e outras espécies. A compreensão desses mecanismos pode ser crucial para o avanço de aplicações biotecnológicas e agrícolas.

5.4.2 Vizinhança no Genoma

A Figura 5 aborda os resultados referente à vizinhança gênica, cujo conceito se refere à posição espacial de genes situados em regiões adjacentes do genoma. Este aspecto é particularmente relevante no contexto bacteriano por duas razões principais: o tamanho compacto do genoma bacteriano e a existência de operons. No genoma bacteriano, o espaço é um recurso escasso e altamente organizado. Assim, genes que são vizinhos frequentemente participam de vias metabólicas similares ou estão envolvidos em funções biológicas relacionadas. Além disso, a presença de operons—unidades de transcrição múltipla onde vários genes são regulados como um único conjunto—acrescenta uma camada adicional de complexidade e eficiência na regulação gênica.

Dessa forma, o estudo da vizinhança gênica em bactérias não é apenas uma questão de mapeamento genético, mas também uma ferramenta valiosa para entender as redes de regulação e as vias metabólicas desses organismos. Isso pode fornecer perspectivas importantes para a biologia funcional e potenciais aplicações em diversas áreas nas ciências.

Ao analisar o score fornecido pelo STRING, é possível determinar a qualidade da interação entre as proteínas. Quanto mais próximo de 1, melhor é a pontuação (score). Nesse contexto, as proteínas Endoribonuclease GDI2454 (score 0.979) (cor bege), Citocromo C GDI2455 (score 0.978) (cor verde) apresentaram os melhores resultados. Indicando uma alta probabilidade de interação, estando próximas um do outro no genoma (figura 5). Salientando, que ambas estão adjacentes a *laao* em *G. diazotrophicus*, assim como é apresentado na figura 5.

No contexto da relação entre as proteínas GDI2455 e GDI2454 e *G. diazotrophicus* o estudo de Rodrigues et al. (2016) analisou o contexto genômico e revelou que o gene *laao* está localizado adjacente aos genes *cccA* e *ridA*, que correspondem aos genes GDI2455 (citocromo c) e GDI2454 (Endoribonuclease) respectivamente da rede de interação proteína-proteína aqui estudada, corroborando com o que é visualizado na figura 5. Esses genes são previstos para codificar, respectivamente, um putativo citocromo C e uma proteína RidA, conhecida como desaminase A intermediária reativa. Estes dados corroboram com a hipótese de que a expressão de *ridA*, *cccA* e *laao* são necessários para a biossíntese de IAA através da via IPyA, sendo esta a principal via para a síntese de IAA em *G. diazotrophicus* (Rodrigues et al., 2016).

Outra proteína que faz parte do genoma encontrado em *G. diazotrophicus* é *nuol* conhecida como NADH-quinona oxidoreductase subunidade I (GDI2464) (cor rosa), ao lado LAAO (vermelha) (figura 5). Essa enzima de membrana integral, que possui múltiplas subunidades, está presente tanto em bactérias como em organelas eucarióticas. O Complexo I desta enzima é responsável pela transferência reversível de elétrons do NADH solúvel para a quinona ligada à membrana, acoplando esta reação à geração de uma força motriz de prótons (PMF). Estudos também apontam que, em bactérias, o Complexo I exerce diversas funções, influenciando tanto as condições de crescimento como processos metabólicos variados (Spero et al., 2015).

A análise da proximidade física dos genes quinona oxidoreductase subunidade I, L-amino-acid oxidase, citocromo c e endoribonuclease no genoma de *Gluconacetobacter diazotrophicus* sugere uma regulação coordenada. Esta hipótese abrange múltiplos aspectos, incluindo regulação transcricional coordenada, participação em vias metabólicas complementares, resposta adaptativa a condições ambientais e implicações evolutivas, que serão detalhados a seguir.

No que se refere à regulação transcricional coordenada, a proximidade desses genes pode indicar mecanismos de regulação similares ou compartilhados. Por exemplo, eles podem estar sob o controle de um promotor comum ou de elementos regulatórios que respondem a sinais ambientais semelhantes. Isso implica que, diante de condições específicas como estresse, disponibilidade de nutrientes ou sinais emitidos pela planta hospedeira, esses genes podem ser ativados ou reprimidos simultaneamente.

Quanto à participação em vias metabólicas complementares, a coexistência desses genes pode indicar uma função metabólica compartilhada ou complementar dentro da bactéria. A quinona oxidoreductase subunidade I e o citocromo c, envolvidos na cadeia de transporte de elétrons, desempenham um papel sinérgico na geração de energia. A L-amino-acid oxidase,

relacionada à produção de metabólitos secundários, pode interagir com a via de transporte de elétrons, enquanto a endoribonuclease, atuando na regulação da expressão gênica, pode ajustar a resposta celular às mudanças no metabolismo energético.

A resposta adaptativa a condições ambientais, representada pela co-localização desses genes, sugere uma adaptação evolutiva para uma resposta rápida e coordenada a mudanças ambientais. Em um contexto simbiótico, esses genes podem ser fundamentais na otimização da interação bactéria-plantas, ajustando o metabolismo bacteriano para beneficiar a planta hospedeira sob condições específicas.

As implicações evolutivas sugerem que a organização genômica desses genes pode refletir uma evolução conjunta, devido ao seu papel colaborativo no sucesso da bactéria em ambientes específicos, incluindo a simbiose com plantas.

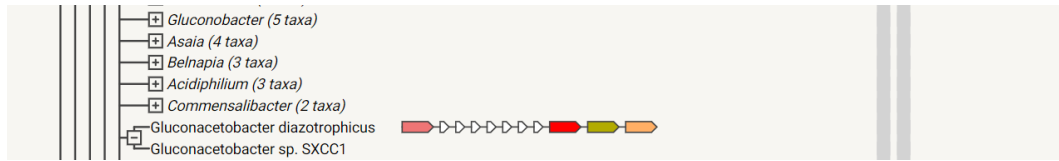
Em síntese, a proximidade e possível regulação coordenada desses genes no genoma de *G. diazotrophicus* indicam uma conexão funcional vital para a eficiência metabólica da bactéria e, conseqüentemente, para a simbiose com a planta hospedeira. Compreender essa correlação fornece insights importantes sobre as estratégias metabólicas e adaptativas das bactérias em ambientes simbióticos.

Esta interação genética em bactérias simbiontes pode aprimorar a eficiência energética, o fornecimento de nutrientes e a regulação gênica, criando condições favoráveis ao crescimento das plantas. A relação simbiótica entre bactérias e plantas é complexa, onde cada gene e enzima desempenha um papel significativo na saúde e no crescimento vegetal. Por exemplo, a quinona, atuando na cadeia de transporte de elétrons, e o citocromo c, transportando elétrons entre os complexos da cadeia respiratória, são essenciais para a geração eficiente de ATP. Em simbiose, os genes que codificam essas proteínas são cruciais para a eficiência energética da bactéria, influenciando positivamente o crescimento e a saúde da planta hospedeira.

Adicionalmente, a L-amino-acid oxidase, ao catalisar a oxidação de aminoácidos L, libera amônia, um nutriente valioso para as plantas, enquanto as endoribonucleases, envolvidas no processamento do RNA e na regulação da expressão gênica, podem influenciar a regulação de genes que afetam a interação simbiótica. Em relação a LAAO. Assim, essas enzimas podem otimizar o crescimento e a saúde das plantas.

Portanto, o gene quinona oxidoreductase subunidade I, em conjunto com L-amino-acid oxidase, citocromo c e endoribonuclease, pode ser essencial para o crescimento das plantas. No entanto, mais estudos são necessários para elucidar completamente essas hipóteses.

Figura 5 - Representa a vizinhança gênica na bactéria *G. diazotrophicus*.



Fonte: Elaborada pela autora a partir dos dados do STRING.

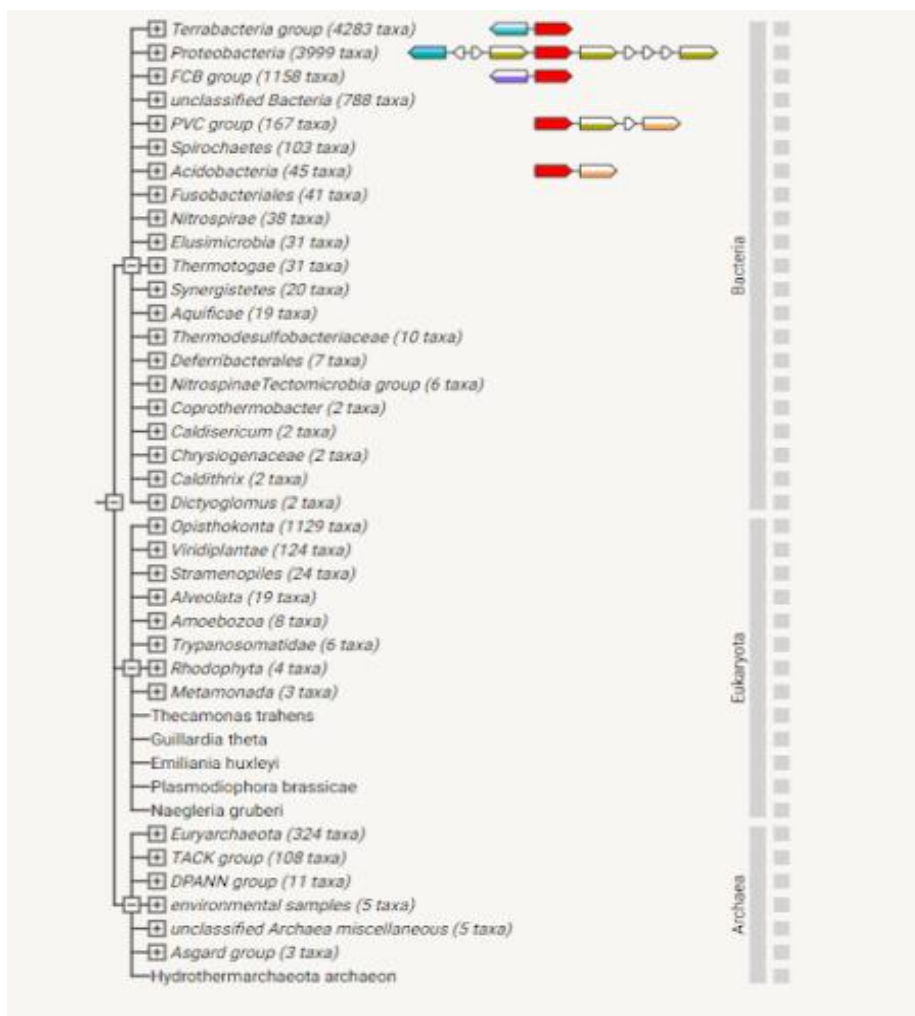
Cor vermelha - Putative L-amino-acid oxidase; Cor verde - Citocromo C; Cor bege - Endorribonuclease; Cor rosa - Quinona Oxiredutase subunidade I.

No estudo de três filos bacterianos distintos, foram identificadas duas homólogas enzimas associadas ao homólogo da LAAO: A 2-hydroxyacido deshydrogenase e Beta-ketoacyl sintase (Figura 6). No contexto planta-bactéria, as 2-hidroxiácido desidrogenases assumem papéis importantes em diversas vias metabólicas. Nas plantas, estas enzimas são cruciais para a degradação de aminoácidos, sendo fundamentais na mobilização de nitrogênio durante o desenvolvimento vegetal. Em bactérias, estão relacionadas à fermentação e ao uso de diferentes fontes de carbono. Assim, as 2-hidroxiácido desidrogenases são essenciais tanto para o metabolismo bacteriano quanto vegetal, estabelecendo um elo entre as redes de interações metabólicas nestes organismos (Kumsab et al., 2020).

A proximidade genética dessas enzimas sugere uma possível coordenação na expressão de proteínas que operam em vias metabólicas interconectadas ou complementares. Por exemplo, enquanto a 2-hidroxiácido desidrogenase e a LAAO atuam no metabolismo de aminoácidos, a beta-ketoacyl sintase é fundamental na síntese de ácidos graxos. Essa associação indica uma possível interdependência metabólica, na qual os produtos ou intermediários de uma via podem influenciar a outra. Além disso, a regulação coordenada desses genes pode ser vital para a adaptação das bactérias a variações ambientais, como mudanças na disponibilidade de nutrientes.

Este panorama amplia a compreensão sobre a complexidade e a amplitude das interações bacterianas no desenvolvimento vegetal. Cada enzima apresenta um papel fundamental, seja em processos metabólicos específicos ou em uma gama de funções biológicas, contribuindo para o crescimento e desenvolvimento das plantas.

Figura 6- Vizinhança da rede de interação da L-amino-acid oxidase com as demais proteínas.



Fonte: Elaborada pela autora com base nos dados do STRING.

Cor vermelha - Putative L-amino-acid oxidase; Cor bege - Endorribonuclease; Cor verde - Citocromo C; Cor azul - 2-hydroxyacido deshydrogenase; Cor lilás - Beta-ketoacyl sintase.

5.4.3 Coocorrência entre genomas

Na análise de co-ocorrência, avalia-se a presença simultânea de genes específicos — e, consequentemente, das proteínas que eles codificam — em diferentes espécies ou genomas. A premissa subjacente é que, se dois genes aparecem frequentemente juntos em diversos organismos, é provável que as proteínas codificadas por eles tenham alguma forma de interação. Conforme ilustrado na Figura 7, a intensidade da cor serve como um indicador da similaridade entre os genes, e assim, da probabilidade de co-ocorrência nos genomas analisados. A ausência

de um "quadrado" completo na referida figura sugere que ocorreu alinhamento parcial do gene, indicando que a correspondência não é completa.

A figura 5, mostra que a co-ocorrência ocorreu nos três domínios. No entanto, observou-se uma presença dos genes analisados com uma coloração mais intensa predominante em Proteobacteria, uma vez que esse filo é o qual a bactéria *G. diazotrophicus* faz parte.

Dentre as proteínas analisadas, as que apresentaram melhor score, foram novamente: GDI2454 - Endoribonuclease (0,979) e GDI2455 - Citocromo C (0,978). Esses scores indicam uma alta probabilidade de co-ocorrência dessas proteínas nos genomas estudados, sugerindo relação funcional entre elas. Como já foi mencionado na vizinhança gênica, esses genes são previstos para codificar, respectivamente, um putativo citocromo C que é uma proteína RidA. Conforme citado anteriormente no trabalho de Rodrigues et al. (2016) a expressão de *ridA*, *cccA* e *laao* são necessários para a biossíntese de IAA através da via IPyA. Além disso, é possível visualizar na figura 7 que os genes GDI2454 e GDI2455 estão presentes nos três domínios da vida, mas, possuem uma maior similaridade entre os grupos do domínio 'bacteria'.

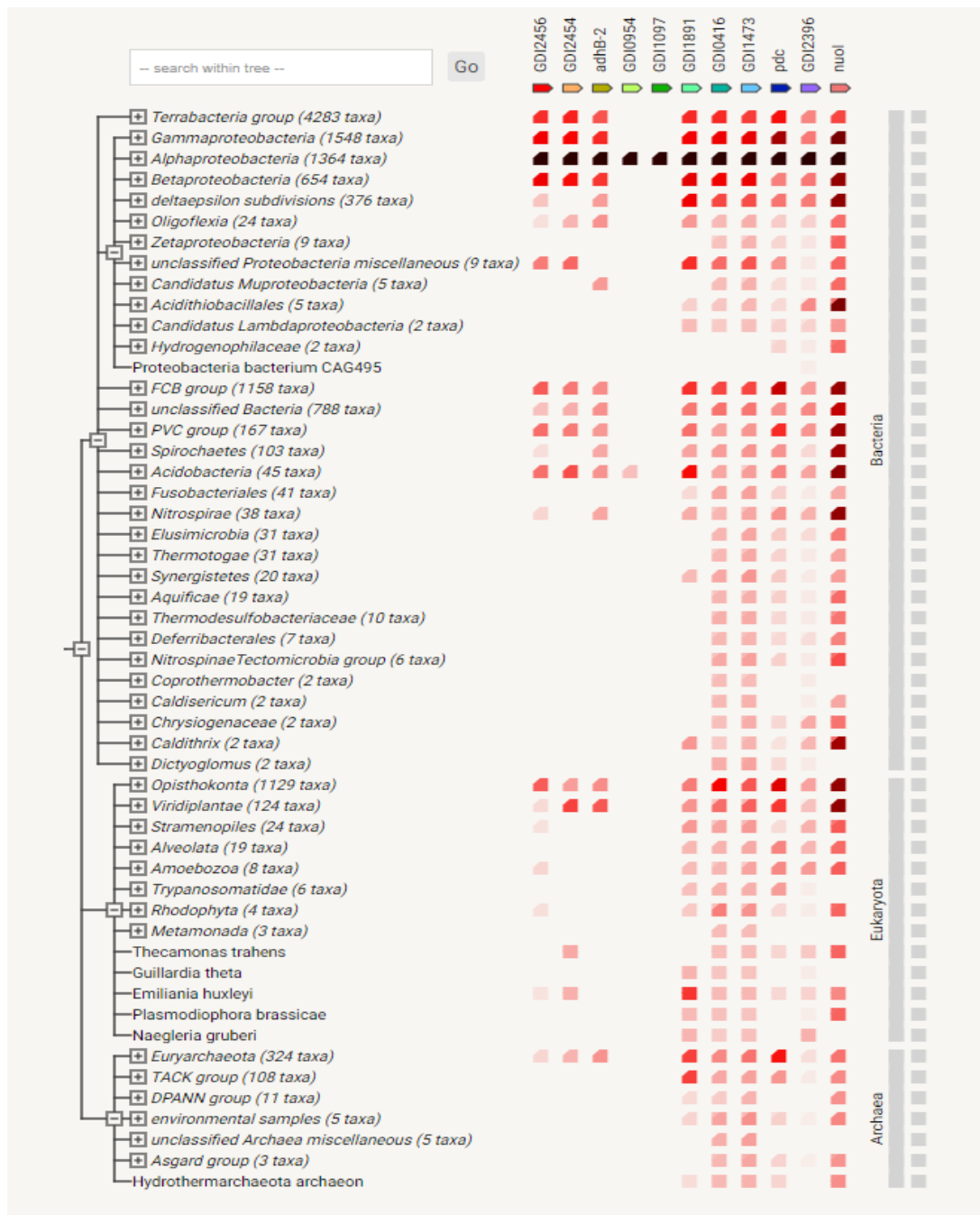
As enzimas GDI0954 e GDI1097, identificadas com alta similaridade exclusivamente em *Gluconacetobacter diazotrophicus*, conforme ilustrado na Figura 7, são atualmente classificadas como hipotéticas na rede de interação proteína-proteína, uma vez que ainda não foram caracterizadas funcionalmente. A presença marcante dessas enzimas apenas em *G. diazotrophicus* sugere que podem ser genes ou proteínas únicas para essa espécie de bactéria, o que lhes confere um interesse particular. Sua associação na rede de interações analisada neste estudo reforça a necessidade de investigar mais profundamente suas funções.

As proteínas Thiamine pyrophosphate (GDI0172), as duas 2-hydroxyacid dehydrogenase (GDI0416 e GDI1473), Pyridoxal-dependent decarboxylase (GDI1891), Beta-ketoacyl synthase (GDI2396) e Quinone oxidoreductase subunit I (GDI2464) são amplamente distribuídas nos três domínios da vida, com uma similaridade notável entre os grupos do domínio Bacteria. Esta distribuição abrangente é atribuída às suas funções essenciais em processos bioquímicos, justificando sua conservação evolutiva. Thiamine Pyrophosphate é uma coenzima crucial para o metabolismo de açúcares e aminoácidos, vital para a respiração celular e produção de energia. A 2-Hydroxyacid Dehydrogenase participa no metabolismo de aminoácidos e ácidos graxos, essencial para a síntese de biomoléculas e geração de energia. Pyridoxal-Dependent Decarboxylase é importante na síntese de neurotransmissores e aminoácidos, enquanto a Beta-Ketoacyl Synthase desempenha um papel chave na biossíntese de ácidos graxos.

Por fim, a Quinone Oxidoreductase Subunit I é fundamental na cadeia de transporte de elétrons para a respiração celular e produção de ATP, ademais destaca-se por sua representação com coloração mais intensa em diversos filos presentes nos três domínios. Visto que, são sugeridas como responsáveis pela translocação de carga e ligação da quinona, isso sublinha a importância do gene *nuol* nos referidos domínios (Marreiros, 2016). A preservação dessas proteínas ressalta sua importância nos mecanismos vitais dos organismos vivos.

A análise de co-ocorrência na rede de interação de LAAO em *Gluconacetobacter diazotrophicus* e outros organismos revela conexões significativas entre certas proteínas, evidenciando sua função essencial e preservação evolutiva. A recorrente co-presença de genes como Endoribonuclease e Citocromo C em múltiplos domínios da vida sugere uma relação funcional entre eles, embora estejam envolvidos na biossíntese de IAA em *G. diazotrophicus*, o que não necessariamente implica uma função idêntica em outros seres. A exclusividade de certas enzimas em *G. diazotrophicus* indica uma especialização adaptativa, merecendo investigações adicionais. Ademais, a distribuição generalizada de proteínas cruciais em todos os domínios da vida destaca sua relevância bioquímica para a sobrevivência e evolução de organismos vivos.

Figura 7 - Co-ocorrência entre genomas da rede de interação da L-amino-acid oxidase com as demais proteínas.



Fonte: Elaborada pela autora a partir dos dados do STRING.

GDI2456 = L-amino-acid oxidase; GDI2454 = Endorribonuclease, ADhb-2= Citocromo C; GDI0954 = Hipotética; GDI1097 = Hipotética; GDI1891 = Pyridoxal-dependent decarboxylase; GDI0416 e GDI1473 = D-isomer specific 2-hydroxyacid dehydrogenase; PDC == Thiamine Pyrophosphato protein; GDI2396 = Beta ketoacil Syntase; nuol = NADH-quinona oxiredutase subunidade I.

5.5 Cluster

Levando em consideração a análise inicial da rede de interação proteína-proteína exposto nas seções anteriores, foi realizado um enriquecimento seguido da aplicação do algoritmo de agrupamento k-means para geração de clusters, uma ferramenta disponibilizada pelo software STRING. Este processo foi adotado com o objetivo de aprofundar a análise e identificar grupos funcionais relevantes dentro da rede de interações gerada para a enzima LAAO. Como resultado, quatro grupos distintos foram identificados: o Cluster 1, composto por 6 genes; o Cluster 2, com 12 genes; o Cluster 3, incluindo 20 genes; e o Cluster 4, que abrange 4 genes, conforme detalhado na Tabela 9 e ilustrado na Figura 8.

A análise detalhada da rede revelou que esses grupos específicos tendem a interagir mais frequentemente entre si e com outros elementos. Tal estrutura modular na rede é evidenciada pela formação de clusters, onde os genes de cada grupo estão relacionados a funções ou processos biológicos e metabólicos similares.

Os dados obtidos da análise de clustering demonstraram que a rede era constituída por 41 nós e 287 arestas, indicando uma densidade de conexões notável. A média de grau dos nós foi de 14, sugerindo um bom nível de interação entre os componentes da rede. Esse padrão sugere a presença de nós altamente interconectados, que podem ter uma significância biológica relevante. Além disso, o coeficiente médio de clustering local foi de 0.668, indicando uma forte propensão para a formação de sub-redes ou grupos densamente conectados dentro da rede maior.

Quanto ao número esperado de arestas, este foi de 134, evidenciando que a rede observada possui um número de conexões significativamente maior do que o esperado por mero acaso. Esse achado é corroborado por um valor de p-valor extremamente baixo ($< 1.0e-16$) obtido no teste de enriquecimento de interações proteína-proteína (PPI enrichment), o que sinaliza que as conexões na rede são estatisticamente significativas e não aleatórias.

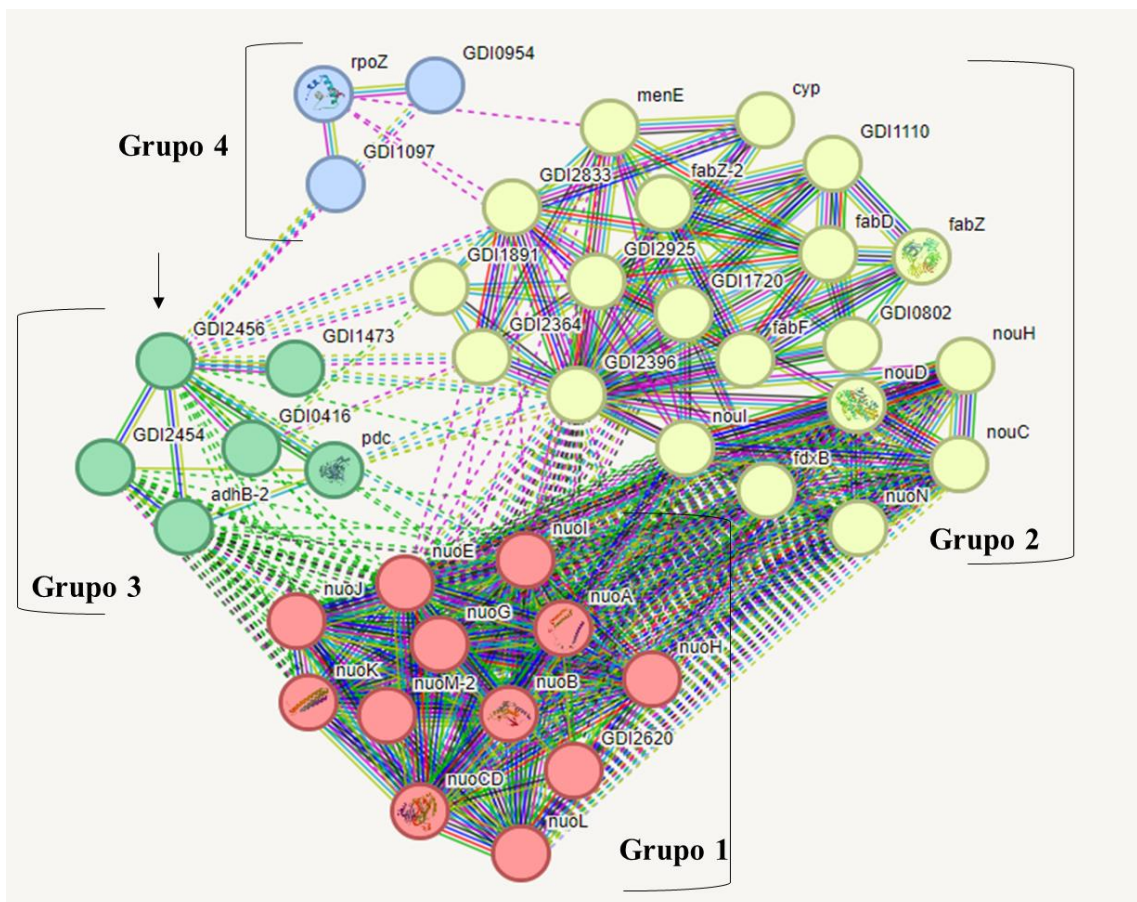
Esses resultados indicam que a rede de interação estudada possui uma organização estrutural definida, com grupos de nós altamente conectados e um número de conexões maior do que o esperado aleatoriamente. Estas descobertas são cruciais para entender os padrões de interação e a funcionalidade dos componentes da rede, sugerindo que o enriquecimento observado implica uma ligação biológica entre as proteínas, atuando de forma coesa como um grupo.

Tabela 9 - Informações gerais sobre as proteínas do cluster.

ID Cluster	Genes do grupo (quantidade)	Nome das Proteínas
Grupo 1	12	GDI2620, nuoA, nuoB, nuoCD, nuoE, nuoG, nuoH, nuoI, nuoJ, nuoK, nuoL, nuoM-2
Grupo 2	20	GDI0802, GDI1110, GDI1720, GDI1891, GDI2364, GDI2396, GDI2833, GDI2925, cyp, fabD, fabF, fabZ, fabZ-2, fdxB, menE, noC, noD, noH, noI, nuoN
Grupo 3	6	GDI0416, GDI1473, GDI2454, GDI2456, adhB-2, pdc
Grupo 4	3	GDI0954, GDI1097, rpoZ

Fonte: Elaborada pela autora a partir dos dados do STRING (2023).

Figura 8 - Cluster gerada pelo STRING.



Fonte: Elaborada pela autora a partir dos dados do STRING (2023).

Apresenta os quatro grupos que foram formados, a seta evidencia a proteína L-amino-acid oxidase.

O STRING revelou que esses grupos estão envolvidos em várias categorias, incluindo processos biológicos, funções moleculares, agrupamentos locais de rede, vias KEGG, localização sub-celular, palavras-chave anotadas (uniProt) e domínios proteicos (Pfam).

Os processos biológicos de GO mostrou que estas proteínas estão envolvidas nos seguintes processos (taxa de descoberta falsa): Processo metabólico de esterol (0,0498); Síntese de ATP acoplada ao transporte de elétrons ($6.64e-05$); Processo metabólico de ácidos graxos ($6.42e-06$); Cadeia respiratória de transporte de elétrons (0,00010); Respiração aeróbica ($3.02e-07$); Processo biosintético de ácidos graxos (0,00078); Respiração celular ($1.93e-08$); Cadeia de transporte de elétrons ($2.63e-09$); Geração de metabólitos precursores e energia ($1.72e-11$); Processo metabólico lipídico (0,00010); Processo metabólico celular ($5.52e-05$); Processo metabólico ($1.26e-06$) e Processo celular (0,0141). A Figura 9 apresenta uma visualização esquemática desses processos.

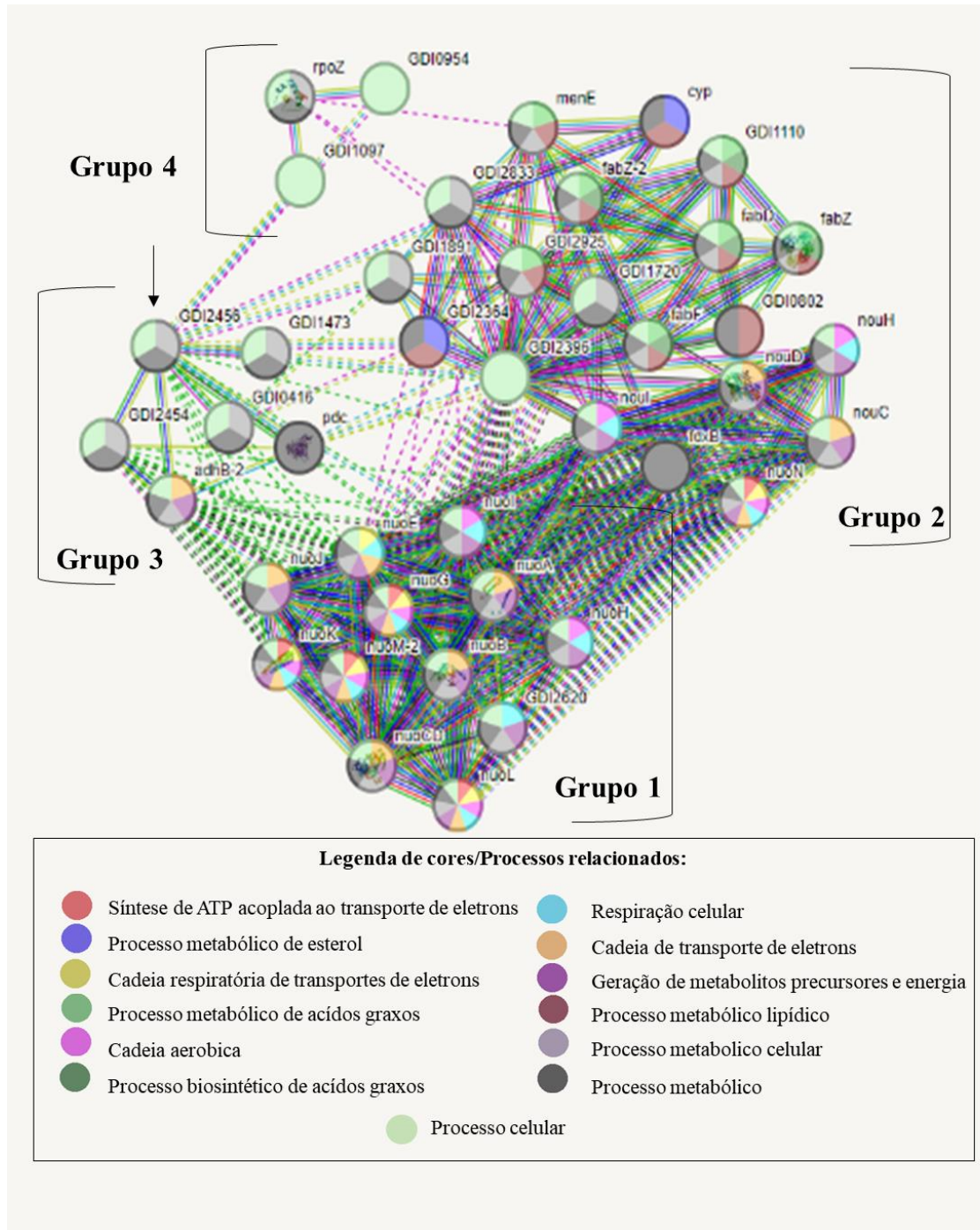
Na análise dos clusters, a avaliação dos Processos Biológicos do GO (Gene Ontology) revelou que todos os grupos (1, 2, 3 e 4) estão envolvidos em processos celulares e metabólicos. Dentro dessa análise, o Grupo 1 se distingue pela sua diversidade, abrangendo uma ampla gama dos processos biológicos identificados. Esta diversidade sugere uma multiplicidade de funções celulares que as proteínas deste grupo podem desempenhar.

Por outro lado, o Grupo 2 apresenta uma especificidade mais acentuada. Este grupo inclui apenas duas proteínas envolvidas no processo metabólico do esterol, indicando uma especialização nessas funções metabólicas particulares. Além disso, este grupo também é caracterizado pela presença de proteínas atuantes no metabolismo lipídico, incluindo processos metabólicos e biosintéticos de ácidos graxos.

É relevante destacar que os Grupos 1 e 2 têm em comum proteínas que participam ativamente no processo de respiração celular. Este achado sublinha uma função compartilhada entre esses grupos, fundamental para a bioenergética celular.

Já o Grupo 3 apresenta proteínas envolvidas nos processos metabólicos e celulares, uma característica comum a todos os clusters analisados. Uma proteína notável neste grupo é a adhB-2, que está associada à cadeia de transporte de elétrons e à geração de metabólitos precursores e energia. Esta associação indica um papel importante do Grupo 3 nos processos bioenergéticos essenciais da célula, potencialmente relacionado à eficiência e à regulação do metabolismo energético.

Figura 9 - Clusters que evidenciam processos biológicos.



Fonte: Elaborada pela autora a partir dos dados do String (2023).

A seta evidencia a proteína L-amino-acid oxidase.

As funções moleculares (GO) elencadas foram (processo/taxa falsa de descoberta): Atividade oxidorrredutase, atuando em NAD(P)H, quinona ou composto similar como acceptor (1.39e-16); Atividade catalítica (9.39e-06); Ligação de íons de ferro (0,00019); Atividade da aciltransferase (0,0084); Atividade da 3-hidroxipalmitoil-[acil-proteína transportadora] desidratase (0,0222); Atividade remodeladora da cromatina dependente de ATP (0,0222);

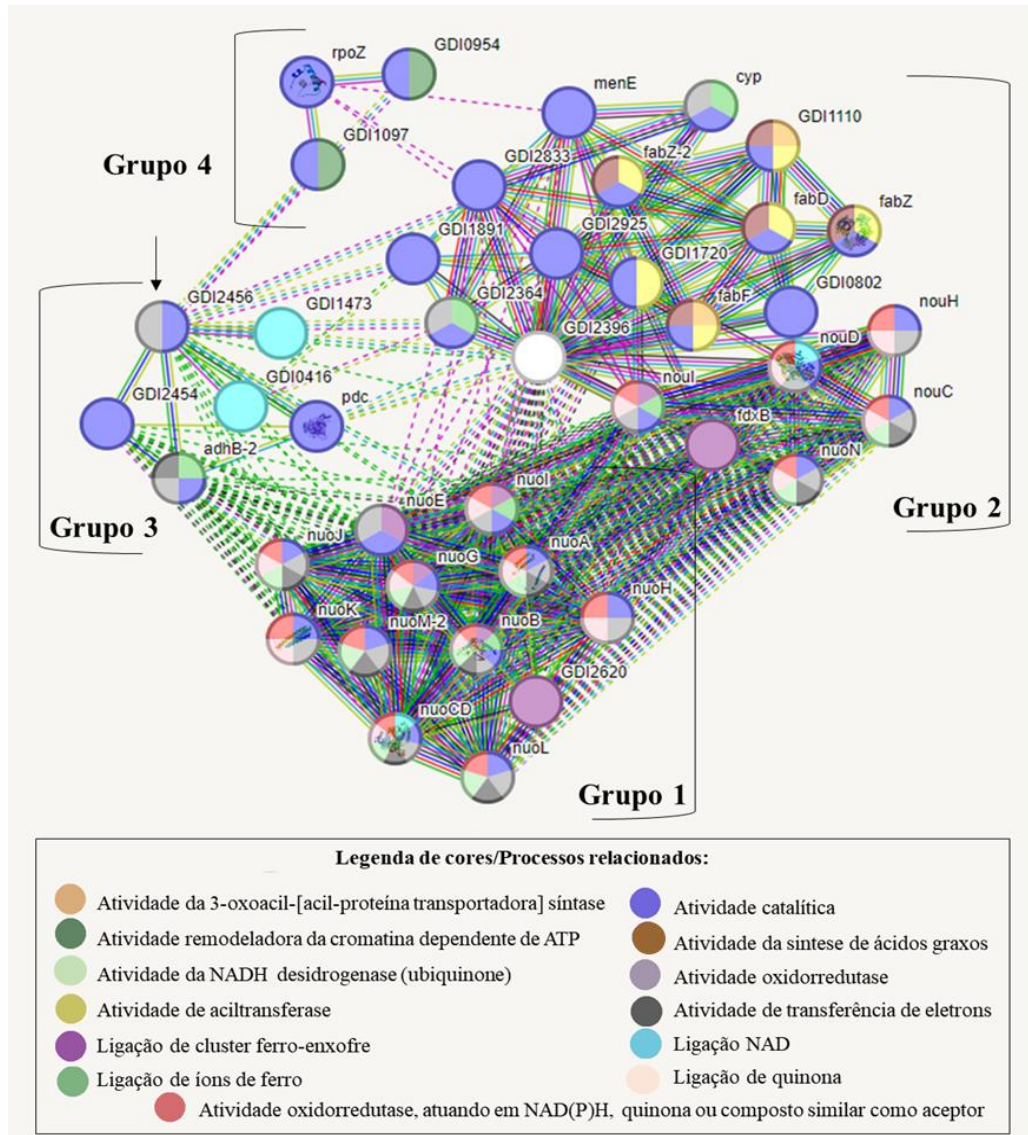
Ligação NAD (0,0222) e Atividade da 3-oxoacil-[acil-proteína transportadora] sintase (0,0275); ; Ligação de cluster ferro-enxofre (0,0084); Atividade da sintase de ácidos graxos ($5.28e-06$); Atividade oxidoredutase ($5.33e-07$); Atividade de transferência de elétrons ($2.81e-07$) e Atividade da NADH desidrogenase (ubiquinona) ($9.01e-10$); Ligação de quinona ($7.68e-15$), sendo encontradas na figura 10.

Na análise do Cluster, que envolve funções moleculares do GO, tornou-se possível inferir que os grupos (1, 2, 3 e 4) estão envolvidos com atividade catalítica. Além disso, o Grupo 1 se destaca por sua grande diversidade de funções moleculares. Destacando que são capazes de participar de várias funções celulares.

Além disso, o Grupo 2 está envolvido com algumas funções moleculares que não são encontradas nos outros grupos como: Atividade da 3-oxoacil-[acil-proteína transportadora] síntese; Atividade de Aciltransferase e Atividade de síntese de ácidos graxos. Ademais, esse grupo possui uma proteína que não está relacionada com nenhuma função molecular.

Em sequência, o Grupo 3 possui uma especificidade mais reduzida apresentando uma relação com; Ligação NAD; Atividade catalítica; Atividade oxidoredutase; Atividade de transferência de elétrons e Ligação de íons ferro. Em contraste, o Grupo 4 está envolvido apenas com: Atividade catalítica e atividade remodeladora da cromatina dependente de ATP.

Figura 10 - Clusters que evidenciam as funções moleculares (GO).



Fonte: Elaborada pela autora a partir dos dados do STRING (2023).

A seta evidencia a proteína L-amino-acid oxidase.

As vias metabólicas e de sinalização presentes no banco de dados KEGG, conhecidas como vias KEGG, resultaram nas seguintes associações (taxa de descoberta falsa): Metabolismo do ácido araquidônico (0,0475); Metabolismo da biotina (0,0048); Metabolismo de ácidos graxos (0,00034); Biossíntese de ácidos graxos (0,00031); Vias metabólicas ($9.58e-12$) e Fosforilação oxidativa ($2.86e-16$).

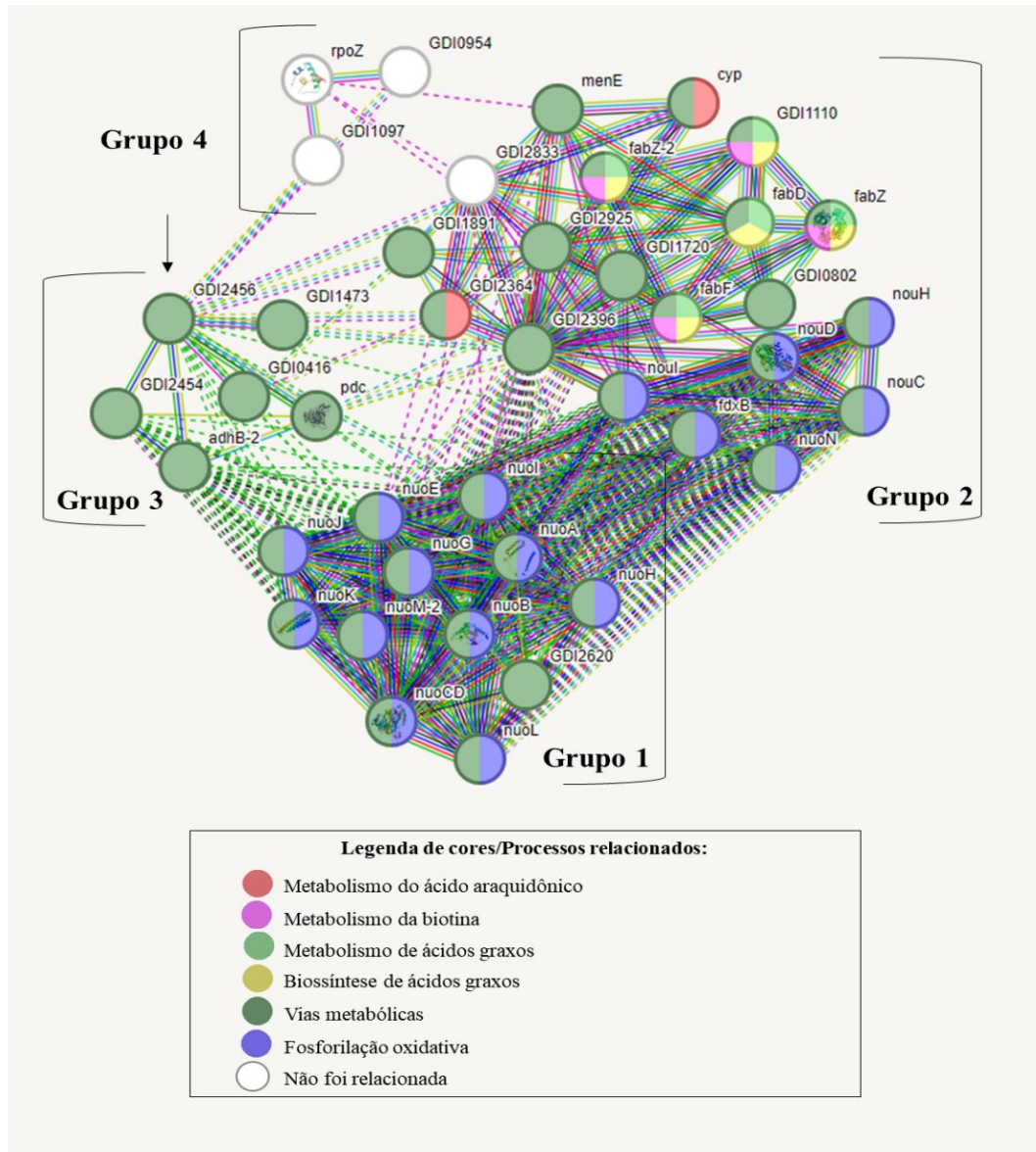
A análise da Figura 11 revela aspectos distintos dos grupos em relação à sua participação nas vias KEGG. O Grupo 4, notavelmente, não está associado a qualquer via metabólica

identificada no KEGG. Em contraste, o Grupo 3 está vinculado exclusivamente a vias metabólicas gerais.

O Grupo 1 se destaca por sua envolvimento tanto em vias metabólicas quanto na fosforilação oxidativa, uma característica que é compartilhada com o Grupo 2. Este último grupo é particularmente notável, pois inclui as proteínas fdxB, noul, nuoN, nouC, nouH e nouD. Estas proteínas estão implicadas tanto na fosforilação oxidativa quanto em outras vias metabólicas, conforme demonstrado na Figura 11.

Além disso, uma análise mais detalhada do Grupo 2 revela uma maior diversidade funcional. Este grupo abrange proteínas envolvidas em outras vias no KEGG, como o metabolismo do ácido araquidônico, representado pelas proteínas CGI2464 e cyp. Ademais, estão presentes também proteínas relacionadas às vias de biossíntese e metabolismo de ácidos graxos, assim como ao metabolismo da biotina, reforçando a complexidade e a variedade metabólica do Grupo 2.

Figura 11 - Clusters evidencia as vias KEGG.



Fonte: Elaborada pela autora a partir dos dados do STRING (2023).

Seta evidencia a proteína L-amino-acid oxidase.

Seguindo, na parte de análise do cluster é encontrado a localização sub-celular: Complexo NADH desidrogenase ($3.43e-08$); Complexo I da cadeia respiratória da membrana plasmática ($1.14e-06$); Complexo oxidorreductase ($1.72e-06$); Respiratório ($1.96e-06$); Membrana ($0,00012$); Complexo catalítico ($0,0019$); Entidade anatômica celular ($0,0023$); Membrana de plasma ($0,0229$); Periferia celular ($0,0317$); Membrana cromófora derivada da membrana plasmática ($0,0323$).

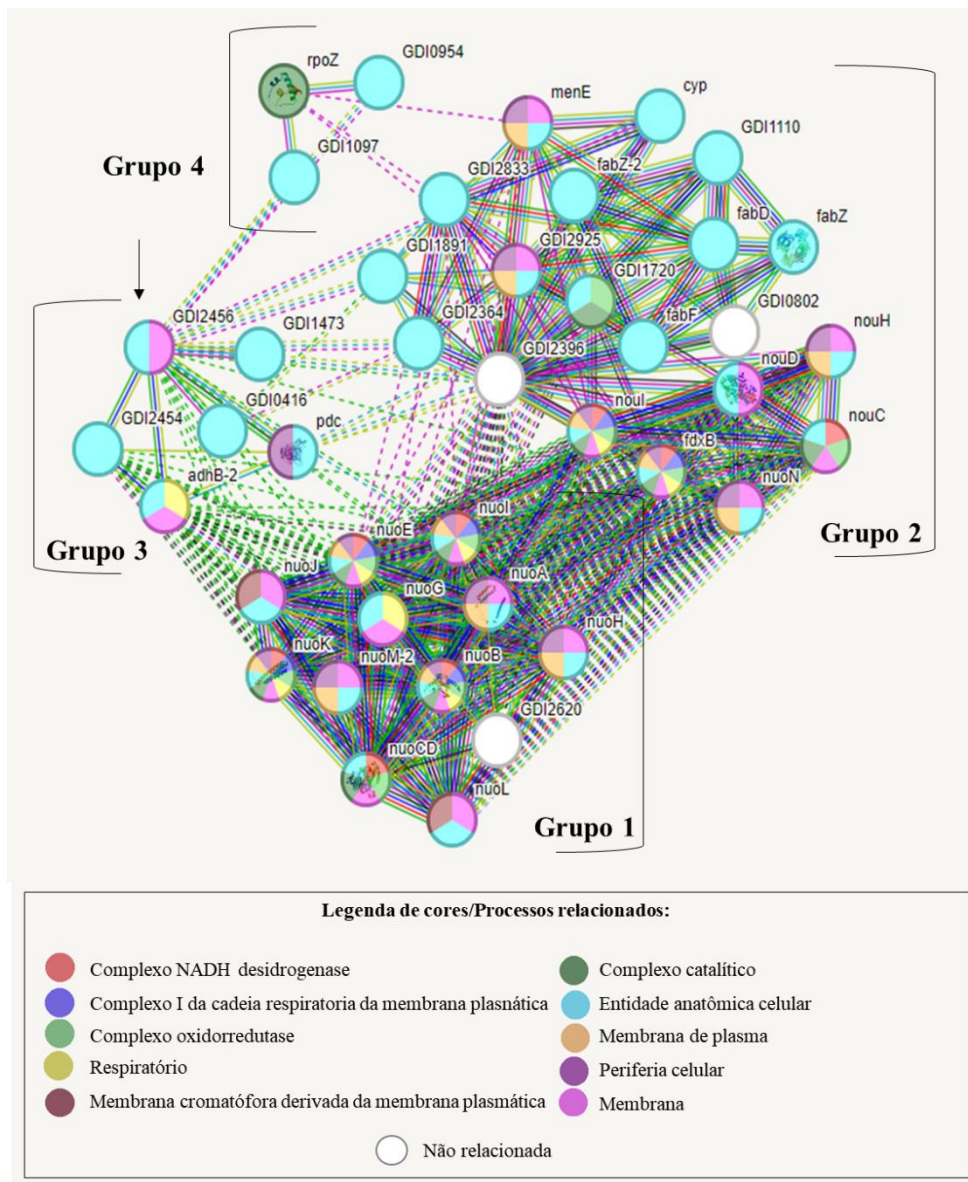
Nesta seção, observa-se que os Grupos 1 e 2 apresentam enzimas detectáveis em todas as áreas subcelulares abordadas no cluster. Dentro deste contexto, merecem destaque as

proteínas nuoI e fdxB do Grupo 2, assim como nuoI e nuoA do Grupo 1. Estas proteínas exibem todas as características previstas nesta análise.

Adicionalmente, identifica-se uma interconexão entre os grupos, evidenciada pelo fato de que a maioria das enzimas é localizada na entidade anatômica celular. No entanto, também são observadas proteínas associadas à membrana, ao sistema respiratório e à membrana plasmática. Estas últimas estiveram anteriormente vinculadas aos processos de fosforilação oxidativa, tanto no Grupo 1 quanto no Grupo 2, conforme ilustrado na Figura 12.

É importante ressaltar que, no Grupo 4, foi identificada uma proteína relacionada ao complexo oxidorreductase, destacando uma peculiaridade deste grupo em relação aos demais.

Figura 12 - Clusters evidencia a localização sub-celular.



Fonte: Elaborada pela autora a partir dos dados do STRING (2023).

A seta evidencia a proteína L-amino-acid oxidase.

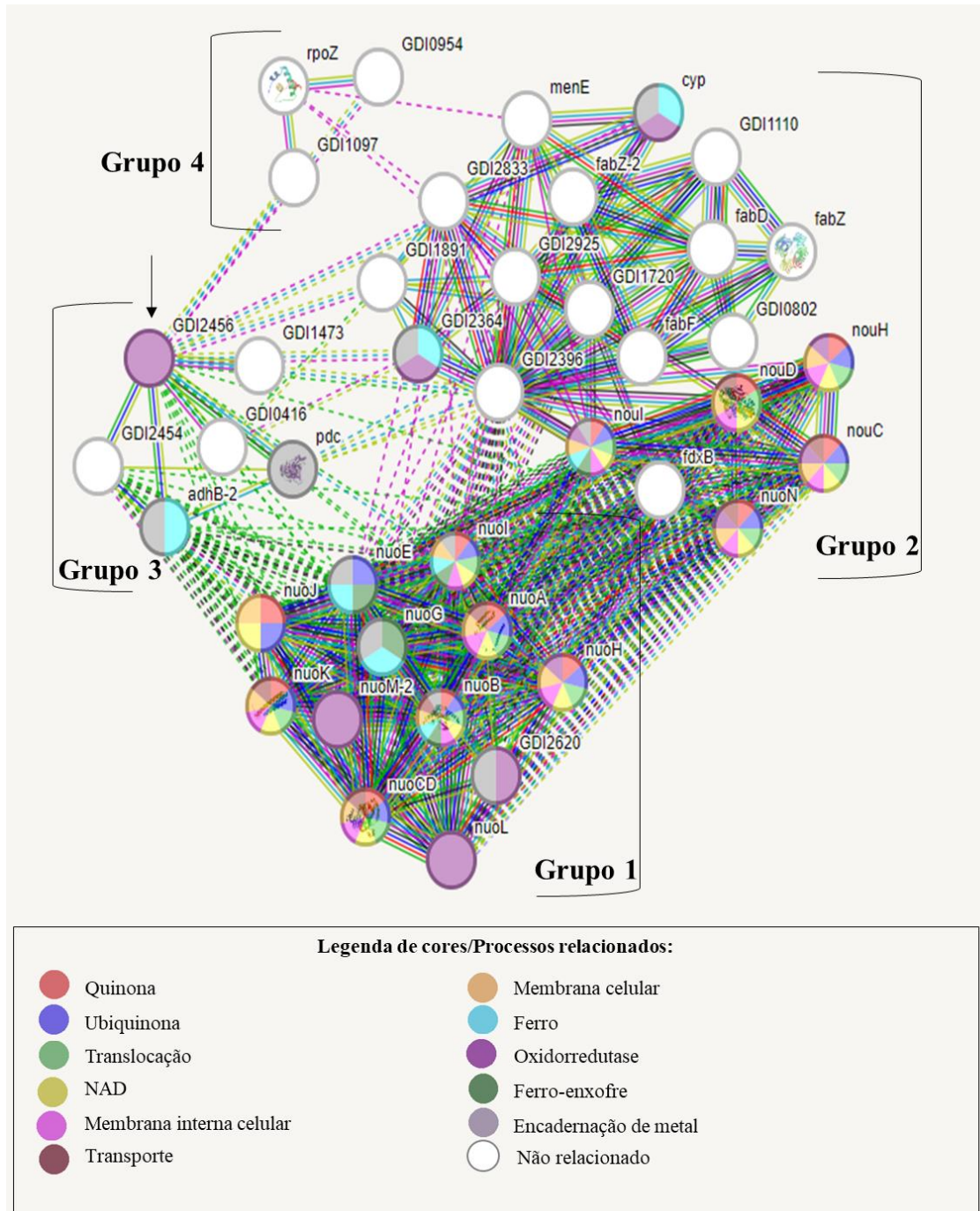
As anotações das palavras-chave disponíveis na plataforma UniProt, desempenhando um papel importante na classificação e descrição das proteínas. Essas anotações são atribuídas com base em características específicas das proteínas e em informações experimentais disponíveis. No contexto da análise realizada no STRING, as anotações das palavras-chave (UniProt) revelaram os seguintes resultados, acompanhados dos respectivos valores estatísticos de taxa de descoberta: Transporte (0,0391); Encadernação de metal (0,0090); Ferro-enxofre (0,0040); Oxidorredutase (0,00021) e Ferro (0,00016); Membrana celular (1.16e-05); Membrana interna celular (5.02e-07); NAD (4.93e-11); Translocação (5.03e-12); Ubiquinona (9.77e-14) e quinona (1.07e-14).

No que diz respeito às palavras-chave anotadas conforme a UniProt, observa-se uma diferenciação notável entre os grupos. O Grupo 4, por exemplo, não está associado a nenhuma das palavras-chave mencionadas no cluster. Em contraste, o Grupo 3 está vinculado a palavras-chave específicas como "ferro-enxofre", "ferro" e "oxidoreductase", conforme ilustrado na Figura 13.

O Grupo 1 se destaca por apresentar todas as palavras-chave previstas nesta análise. No entanto, essa característica é predominantemente concentrada em algumas proteínas específicas do cluster, tais como nuol, nuoA e nuoB. Existem outras proteínas neste grupo associadas a um número menor de palavras-chave; por exemplo, nuoL e nuoM-2 estão vinculados principalmente à "oxidoreductase", enquanto outras estão envolvidas na "ligação a metal". No contexto deste grupo, foi identificada uma associação específica com "ferro", conforme anotação da UniProt, além de uma ligação com NAD.

Por outro lado, no Grupo 2, uma proteína em particular, a nuol, parece englobar todas as anotações previstas na análise. Outras proteínas, como nouh, nouD e nouN, também possuem várias anotações. Além disso, dentro deste grupo, duas proteínas são notáveis pelas anotações "ligação a metal", "ferro" e "oxirredutase". Esses achados indicam uma diversidade funcional e bioquímica significativa entre os grupos, especialmente em relação às suas características e funções moleculares.

Figura 13 - Clusters evidencia as palavras-chave.



Fonte: Elaborada pela autora a partir dos dados do STRING (2023).

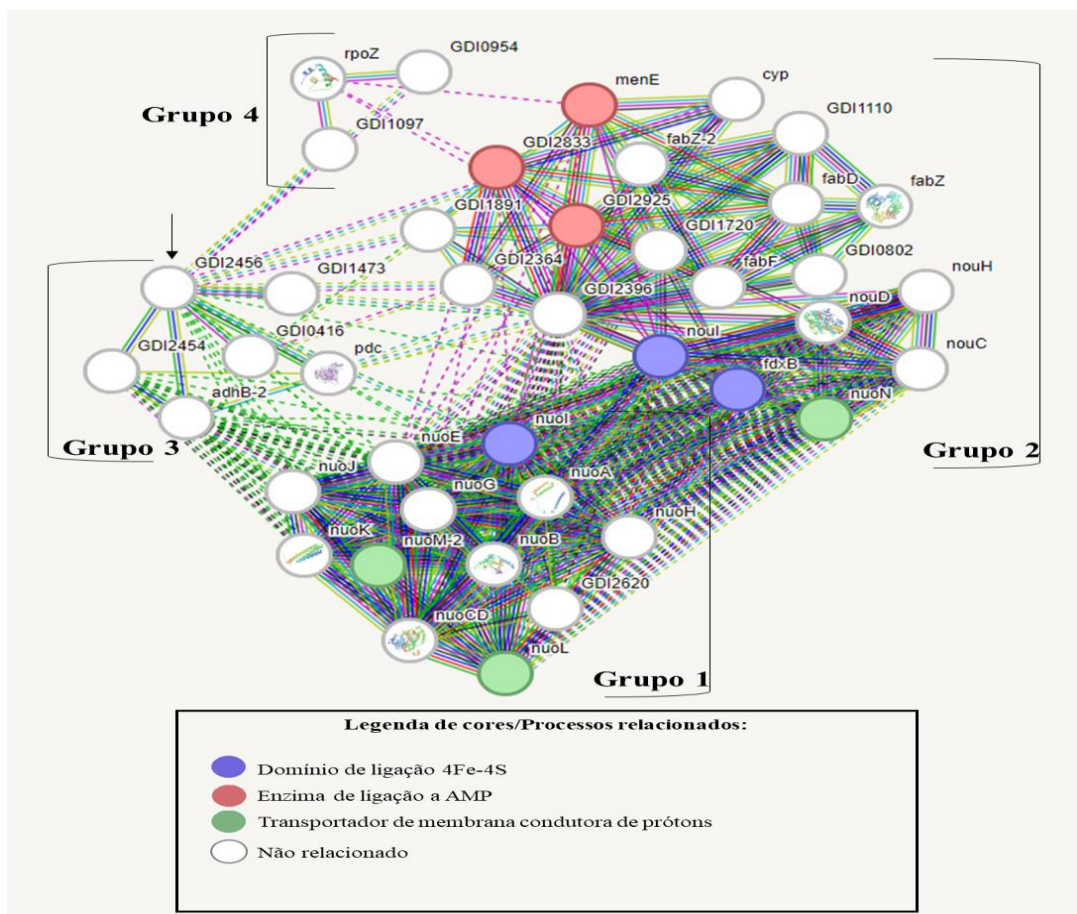
A seta evidencia a proteína L-amino-acid oxidase.

Os domínios das proteínas (Pfam) visualizados na plataforma STRING são componentes estruturais ou funcionais de uma proteína que representam um papel importante em sua atividade biológica. Sendo regiões conservadas que podem ser identificadas através de análises bioinformáticas. No contexto da análise realizada no STRING as anotações dos domínios das proteínas (Pfam) revelaram os seguintes resultados, acompanhados dos

respectivos valores estatísticos de taxa de descoberta: Domínio de ligação 4Fe-4s (0,0337); Enzima de ligação a AMP (0,0337); Transportador de membrana condutora de prótons (0,0337) (Figura 14).

Nesta seção observa-se que o Grupo 3 e Grupo 4 não estão relacionados a nenhum domínio das proteínas. Além disso, no que diz respeito ao Grupo 1 as enzimas estão apresentando o Domínio de ligação 4Fe-4s e Transportador de membrana condutora de prótons. Em contraste, o Grupo 2 está envolvido com os três domínios de proteínas, porém, é importante ressaltar que existem enzimas no Grupo 1 e 2 que não estão relacionadas com nenhum domínio, sendo visualizadas na Figura 14.

Figura 14 - Clusters evidenciam os Domínios de Proteínas (Pfam).



Fonte: Elaborada pela autora a partir dos dados do String.
A seta evidencia a proteína L-amino-acid oxidase.

A análise dos clusters gerados na plataforma STRING para uma rede de interação proteína-proteína revelou características distintas em quatro grupos distintos. Essas

características abrangem uma variedade de categorias, incluindo processos biológicos, funções moleculares, agrupamentos locais de rede, vias KEGG e palavras-chave anotadas (UniProt).

O Grupo 1 se destaca pela sua diversidade em processos biológicos. As proteínas deste grupo estão envolvidas em uma ampla gama de funções essenciais para a célula, como o metabolismo energético e processos biossintéticos. Esta diversidade é reforçada pela sua participação em vias metabólicas e na fosforilação oxidativa, destacando um papel vital na produção de energia e no metabolismo intermediário. A presença destas proteínas em todas as áreas subcelulares listadas indica uma função celular abrangente e multifacetada, potencialmente incluindo interações em diversos contextos celulares. Além disso, a abrangência das palavras-chave associadas a este grupo aponta para uma vasta gama de funções moleculares, incluindo atividades enzimáticas, de transporte e de ligação a metais.

Por outro lado, o Grupo 2 se concentra mais especificamente no metabolismo do esteroide e no metabolismo lipídico, sugerindo uma especialização em aspectos chave do metabolismo desses compostos anteriormente mencionados, que são fundamentais para a estrutura celular e sinalização. A inclusão de proteínas específicas em diversas vias metabólicas ressalta a importância deste grupo na modulação e na execução de várias vias bioquímicas críticas.

Os Grupos 1 e 2 têm em comum proteínas que participam ativamente no processo de respiração celular, uma função compartilhada essencial para a bioenergética celular. Além disso, uma análise mais detalhada do Grupo 2 revela uma maior diversidade funcional, abrangendo proteínas envolvidas em outras vias no KEGG, como o metabolismo do ácido araquidônico, metabolismo de ácidos graxos e biotina, reforçando a complexidade e variedade metabólica deste grupo.

O Grupo 3 apresenta um envolvimento em processos metabólicos e celulares comuns, indicando um papel fundamental na célula com um foco particular em processos bioenergéticos. Este grupo desempenha funções essenciais, contribuindo para a eficiência e regulação desse metabolismo.

Finalmente, o Grupo 4 se distingue pela sua ausência de associação com as vias metabólicas conhecidas do KEGG e também do GO, sugerindo um papel diferenciado ou menos compreendido nas funções celulares. Esta característica pode envolver vias ou processos ainda não totalmente elucidados. A falta de associação com palavras-chave comuns e a identificação de proteínas relacionadas ao complexo oxidoredutase sugerem funções únicas ou especializadas, possivelmente envolvendo mecanismos celulares inovadores ou menos estudados.

Em resumo, a análise destaca a diversidade funcional e bioquímica entre os grupos, especialmente em relação às suas características e funções moleculares. A presença de proteínas com funções bioenergéticas em todos os grupos sugere uma sobreposição de atividades relacionadas ao metabolismo energético, essencial para a manutenção e eficiência da célula. Além disso, a participação em vias metabólicas chave, como a fosforilação oxidativa, indica um ponto de convergência na função celular, essencial para a bioenergética e metabolismo do organismo bacteriano.

Identificada no Grupo 3, a LAAO está diretamente envolvida na conversão de L-triptofano em IPyA, um precursor fundamental na biossíntese do fitohormônio IAA (ácido indolacético) (Rodrigues et al., 2016). Este processo é essencial para a promoção do crescimento vegetal, influenciando a elongação celular, a diferenciação de tecidos e as respostas ao estresse. Rodrigues et al. (2016) mostraram que a interrupção do gene *lao*, que codifica a LAAO, leva à redução da capacidade de biossíntese de IAA, resultando na supressão dos efeitos promotores de crescimento nas raízes das plantas. Este resultado confirma o papel significativo da LAAO na promoção do crescimento vegetal.

Embora os Grupos 1 e 2 não estejam diretamente envolvidos na produção de IPyA, eles desempenham funções vitais no metabolismo energético e em outras funções metabólicas celulares. Estes processos são fundamentais para manter a saúde e a eficiência metabólica da bactéria, influenciando indiretamente sua capacidade de produzir IPyA e outros compostos benéficos para as plantas. Além disso, a presença de proteínas relacionadas ao complexo oxidorreductase no Grupo 4 sugere um possível papel em processos de transferência de elétrons, essenciais para a respiração celular. Embora o papel exato desse grupo na promoção do crescimento vegetal não seja totalmente claro, suas funções podem ser importantes para a eficiência metabólica geral da bactéria.

Os resultados obtidos sugerem hipóteses importantes sobre a via de biossíntese de IPyA, essencial para a produção de IAA (ácido indolacético), e sua relação com o metabolismo energético da célula bacteriana. É plausível que a biossíntese de IPyA demande uma quantidade significativa de energia. Dentro desse contexto, o metabolismo energético, que abrange processos como a fosforilação oxidativa e a cadeia de transporte de elétrons, desempenha um papel crucial ao fornecer o ATP necessário para as reações enzimáticas, incluindo as catalisadas pela LAAO. A presença de proteínas relacionadas à respiração celular e à cadeia de transporte de elétrons nos clusters analisados indica uma conexão direta entre o metabolismo energético da célula e a eficiência da biossíntese de IPyA, facilitando assim a produção contínua e eficiente de IAA.

Além disso, o metabolismo de lipídeos tem um papel vital na manutenção da integridade da membrana celular e na execução de várias funções biológicas. Os lipídeos, servindo também como mensageiros secundários ou substratos para reações bioquímicas, podem influenciar a atividade de diversas enzimas, incluindo potencialmente a LAAO. Alterações no metabolismo de lipídeos podem impactar a disponibilidade de cofatores ou a estabilidade enzimática, afetando indiretamente a via de biossíntese de IPyA.

Portanto, a integração da via IPyA com os processos de metabolismo energético e lipídico na célula bacteriana parece ser uma estratégia adaptativa para assegurar a eficiência da biossíntese de IAA, um processo vital para a promoção do crescimento vegetal. A disponibilidade de energia e a regulação do metabolismo de lipídeos emergem como fatores chave influenciando a atividade da LAAO e, conseqüentemente, a habilidade da bactéria em auxiliar o crescimento das plantas. Esta interconexão complexa destaca a sofisticação do metabolismo bacteriano e sua adaptação para desempenhar um papel simbiótico na agricultura.

A interação sinérgica entre os diferentes clusters e a LAAO sublinha a complexidade do sistema bacteriano em promover o crescimento vegetal. Isso ressalta a importância da *Gluconacetobacter diazotrophicus* PA15 como um agente promotor de crescimento em plantas. Para entender mais profundamente essa correlação entre os clusters e a biossíntese de IAA, são necessários estudos adicionais que caracterizem todas as proteínas presentes na rede de interação. Essas investigações podem abrir novos caminhos para aplicações biotecnológicas desses conhecimentos na agricultura.

6 CONCLUSÃO

Com base nos dados obtidos, a análise dos parceiros de interação da enzima Putative L-amino-acid oxidase (LAAO) em *Gluconacetobacter diazotrophicus* PAL 5 revela um panorama complexo de interações proteicas. Essas interações têm implicações significativas no metabolismo e na fisiologia bacteriana, podendo afetar a interação com organismos vegetais. Proteínas identificadas, como a D-isomer specific 2-hydroxyacid dehydrogenase NAD-binding e a Beta-ketoacyl synthase, destacam a multifuncionalidade do *G. diazotrophicus*, em que processos como a fixação de nitrogênio e a síntese de ácidos graxos são cruciais. Embora a relação direta entre a LAAO e estas proteínas não seja explicitamente clara, a natureza complementar de suas funções sugere um panorama metabólico interconectado, que pode potencializar o crescimento vegetal e a produtividade.

As 2-hidroxiácido desidrogenases têm implicações diretas tanto no metabolismo bacteriano quanto no vegetal, conectando as redes de interações entre diferentes organismos. A integração de proteínas como o Cytochrome C e NADH-quinona oxidoreductase subunidade I na rede de interações enfatiza a complexidade das vias metabólicas envolvidas na respiração celular e na geração de energia. A presença de enzimas como a Endoribonuclease L-PSP e a Pyridoxal-dependent decarboxylase, cujas funções ainda não são totalmente compreendidas, abre caminhos para investigações futuras que podem revelar aspectos desconhecidos da biologia de *G. diazotrophicus*.

Em relação à análise dos clusters, os Grupos 1 e 2 compartilham proteínas ativas no processo de respiração celular, uma função essencial para a bioenergética celular. Uma análise detalhada do Grupo 2 revela uma diversidade funcional maior, incluindo proteínas envolvidas em outras vias no KEGG, como o metabolismo do ácido araquidônico, de ácidos graxos e de biotina, reforçando a complexidade e variedade metabólica desse grupo.

O Grupo 3 é onde se encontra LAAO que está diretamente envolvida na conversão de L-triptofano em IPyA, um precursor essencial na biossíntese do fitormônio IAA (ácido indolacético). Esse grupo também está envolvido em processos metabólicos e celulares comuns, indicando um papel fundamental na célula com foco particular em processos bioenergéticos.

O Grupo 4 se distingue por sua ausência de associação com as vias metabólicas conhecidas do KEGG e do GO, sugerindo um papel diferenciado ou menos compreendido nas funções celulares. Esta característica pode indicar vias ou processos ainda não totalmente elucidados. A falta de associação com palavras-chave comuns e a identificação de proteínas

relacionadas ao complexo oxidorreductase sugerem funções únicas ou especializadas, possivelmente envolvendo mecanismos celulares inovadores ou menos estudados.

Estas múltiplas facetas não apenas reforçam a complexidade intrínseca dessas redes proteicas, mas também sublinham a necessidade de estudos futuros para decifrar as interações funcionais e os significados biológicos mais amplos. É importante reconhecer que este estudo se concentrou na bioinformática e que mais pesquisas são necessárias para validar experimentalmente as funções preditas e as interações proteína-proteína. Assim, é essencial que estudos futuros incluam análises *in vitro* e *in vivo* para elucidar as interações funcionais detalhadas e os significados biológicos mais amplos dessas redes proteicas. Isso permitirá uma compreensão mais profunda da relação simbiótica entre *G. diazotrophicus* e plantas hospedeiras, abrindo novos caminhos para aplicações inovadoras em biotecnologia agrícola.

REFERÊNCIAS

- Bacon, C.; Hinton, D. Bacterial endophytes: the endophytic niche, its occupants, and its utility. **Plant-associated bacteria**, p. 155-194, 2006.
- Bastián, F. et al. Production of indole-3-acetic acid and gibberellins A1 and A3 by *Acetobacter diazotrophicus* and *Herbaspirillum seropedicae* in chemically-defined culture media. **Plant Growth Regulation**, v.24, p.7-11, 1998.
- Beattie, G.A. Plant-associated bacteria: survey, molecular phylogeny, genomics and recent advances. In: S.S. Gnanamanickam (ed). **Plant-Associated Bacteria**. Springer, Netherlands, p. 1–56. 2006.
- Berleth, T.; Sachs, T. Plant morphogenesis: long-distance coordination and local patterning. **Current Opinion in Plant Biology**, v. 4, p. 57-62, 2001.
- Bertalan, M. *et al.* Complete genome sequence of the sugarcane nitrogen-fixing endophyte *Gluconacetobacter diazotrophicus* PAL5. **BMC Genomics**. v.10, p.1-17, 2009.
- Bertini, I. *et al.* Monomorphism of human cytochrome c. **Genomics**, v. 88, p. 669-672, 2006.
- Binenbaum, J.; Weinstain, R.; Shani, E. Gibberellin Localization and Transport in Plants. **Trends Plant Sci**, v. 23, p. 410-421, 2018.
- Bruckner, F. P. Aspectos da interação entre a proteína TCTP e o potyvírus PepYMV na infecção de tomateiro e *Nicotiana benthamiana*, 2012.
- Burman, J. D., Stevenson, C. E., Sawers, R. G., & Lawson, D. M. (2007). The crystal structure of *Escherichia coli* TdcF, a member of the highly conserved YjgF/YER057c/UK114 family. **BMC Struct Biol**, v. 7, p. 30, 2007.
- Caselli, D. M. A. **Importância do resíduo de lisina 79 no processo apoptótico induzido por citocromo c**. 2014. Dissertação (Mestrado) - Universidade de Mogi das Cruzes, 2014.

Castellano, F.; Molinier-Frenkel, V. An overview of l-amino acid oxidase functions from bacteria to mammals: focus on the immunoregulatory phenylalanine oxidase IL4I1.

Molecules, v. 22, p. 2151, 2017.

Cavalcante, V.A.; Dobereiner, J. A new acid-tolerant nitrogen-fixing bacterium associated with sugarcane. **Plant and Soil**. v.108, p.23-31, 1988.

Christopherson, M. R.; Schmitz, G. E.; Downs, D. M. YjgF is required for isoleucine biosynthesis when *Salmonella enterica* is grown on pyruvate medium. **J Bacteriol**, v. 190, p. 3057-3062, 2008.

Cortleven, A. *et al.* Cytokinin action in response to abiotic and biotic stresses in plants. **Plant, cell & environment**, v. 42, n. 3, p. 998-1018, 2019.

Costacurta, A.; Vanderleyden, J. Synthesis of phytohormones by plant-associated bacteria. **Critical Reviews in Microbiology**, v.21, p.1-18, 1995.

Davies, P.J. The plant hormones: their nature, occurrence and functions. In: Davies P.J, ed, **Plant Hormones: Physiology, Biochemistry and Molecular Biology**, Ed 5. Kluwer Academic Publishers, Dordrecht, The Netherlands, p. 1-12, 1995.

Döbereiner; V.L.D., B.; F., O.; Reis, V. M. Endophytic diazotrophs: the key to BNF in gramineous plants. In: HEGAZI, N. A.; FAYEZ, M.; MONIB, M. (Ed.). **Nitrogen Fixation with Non-Legumes**. Giza, Egypt: Faculty of Agriculture, Cairo University, p. 395-408, 1993.

Dzantor, E. K. Phytoremediation: the state of rhizosphere engineering for accelerated rhizodegradation. of xenobiotic contaminants, **Technol. Biotechnol**, v. 82, p. 228-232, 2007.

Firdous, J. et al. Endophytic bacteria and their potential application in agriculture: a review. *Indian Journal of Agricultural Research*, v. 53, n. 1, p. 1-7, 2019.

Galvão, P. G. **Análise proteômica de estirpes selvagem pal5t e mutante lao- de *gluconacetobacter diazotrophicus* na presença e ausência de triptofano e o efeito de sua**

inoculação em plantas micropropagadas de cana-de-açúcar. 2012. Tese (Doutorado) – Instituto de Agronomia, Universidade Federal Rural do Rio de Janeiro, 2012.

Gillis, M.; Kersters, K. *Acetobacter diazotrophicus* sp. nov., a nitrogen fixing acetic acid bacterium associated with sugarcane. **International Journal of Systematic Bacteriology**, v.39, n.3, p.361-364, 1989.

Gírio, L. A. et al. Bactérias promotoras de crescimento e adubação nitrogenada no crescimento inicial de cana-de-açúcar proveniente de mudas pré-brotadas. **Pesquisa Agropecuária Brasileira**, v. 50, p. 33–43 2015.

Glick, B. R. *et al.* Biochemical and genetic mechanisms used by plant growth promoting bacteria. **Imperial College Press**, 1999.

Glick, B. R. *et al.* Plant growth-promoting bacteria: mechanisms and applications. **Scientifica**, 2012. DOI: 10.6064/2012/963401

Gomes, G. L. B. **Caracterização interação molecular da proteína reprimida por auxina em resposta ao controle de desenvolvimento de tomateiro.** 2020. Tese (Doutorado) - Universidade Federal do Rio Grande do Norte, 2020.

Grant, G.A. A new family of 2-hydroxyacid dehydrogenases. **Biochem Biophys Res Commun**, v. 165, p. 1371-1374, 1989.

Hinsinger, P. et al. Rhizosphere geometry and heterogeneity arising from root-mediated physical and chemical processes. **New Phytol.** p. 293-303, 2005.

Hungria, M.; CAMPO, R. J; Mendes, I. de C. Fixação biológica do nitrogênio na cultura da soja. **Embrapa Soja-Circular Técnica (INFOTECA-E)**, 2001.

Huynh, T. N. *et al.* Cyclic di-AMP targets the cystathionine beta-synthase domain of the osmolyte transporter OpuC. **molecular microbiology**, v. 102, p. 233-243, 2016.

Ignoul, S.; Eggermont, J. CBS domains: structure, function, and pathology in human proteins. **Am J Physiol Cell Physiol**, v. **289**, p C1369-1378, 2005.

James, E.K.; Olivares, F.L. Further observation on the interaction between sugar cane and *Gluconacetobacter diazotrophicus* under laboratory and greenhouse conditions. **Journal of Experimental Botany**, v.52, n.357, p. 747-760, 2001.

Jensen, L. J. et al. STRING 8—a global view on proteins and their functional interactions in 630 organisms. *Nucleic acids research*, v. 37, n. suppl_1, p. D412-D416, 2009.

Kumar, R. Evolutionary Trails of Plant Group II Pyridoxal Phosphate-Dependent Decarboxylase Genes. **Front. Plant Sci**, v. 7, 2016.

Kumsab, J. *et al.* Characterization of a novel class of glyoxylate reductase belonging to the β -hydroxyacid dehydrogenase family in *Acetobacter acetii*, *Bioscience*. **Biotechnology, and Biochemistry**, v. 84, p. 2303-2310, 2020.

Lambrecht, J. A.; Browne, B. A.; Downs, D. M. Members of the YjgF/YER057c/UK114 family of proteins inhibit phosphoribosylamine synthesis in vitro. **J Biol Chem**, v. 285, p. 34401-34407, 2010.

Liu, H. et al. Inner plant values: diversity, colonization and benefits from endophytic bacteria. *Frontiers in microbiology*, v. 8, p. 2552, 2017.

Ljung, K. et al. Biosynthesis, conjugation, catabolism and homeostasis of indole-3-acetic acid in *Arabidopsis thaliana*. **Plant Molecular Biology**, v. 49, p. 249-272, 2002.

Mariano, R. L. R. et al. Importância de bactérias promotoras de crescimento e de biocontrole de doenças de plantas para uma agricultura sustentável, vol. 1, p.89-111, 2004.

Marin, V. A. et al. Fixação biológica de nitrogênio: Bactérias fixadoras de nitrogênio de importância para a agricultura tropical. 1999.

Marreiros, B.C. **Exploring NADH:quinone oxidoreductases Complex I and alternative NADH dehydrogenase (NDH-2)**. 2016. Dissertação (Mestrado) - Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, 2016.

Matelska, D. *et al.* Classification, substrate specificity and structural features of D-2-hydroxyacid dehydrogenases: 2HADH knowledgebase. **BMC Evol Biol**, v. 18, p. 1-23, 2018.

Miguel, P. S. B. *et al.* Bactérias endofíticas: Colonização, benefícios e identificação. **Brazilian Journal of Development**, v. 7, n. 1, p. 8777-8791, 2021.

Mistinieni, E. *et al.* Oligomeric assembly and ligand binding of the members of protein family YER057c/YIL051c/YJGF. **Bioconjug Chem**, v. 14, p. 1243-1252, 2003.

Mou, S. *et al.* Over-Expression of Rice CBS Domain Containing Protein, OsCBSX3, Confers Rice Resistance to *Magnaporthe oryzae* Inoculation. **Int J Mol Sci**, v. 16, p. 15903-15917, 2015.

Munõz D.B. Bactérias promotoras del crecimiento vegetal: Mecanismos Y aplicaciones, - Facultad de Farmacia Universidad de Sevilla, 2017

Oliveira, Izamara Gesiele Bezerra de. Modelagem in silico do gene gumD e sua rede de interação em *Gluconacetobacter diazotrophicus* PAL5. 2023. 115f. Trabalho de Conclusão Dissertação (Programa de Pós-Graduação em Ciências Agrárias - PPGCA) - Universidade Estadual da Paraíba, Campina Grande - PB, 2023.

Patten, C.; Glick, B.R. Bacterial biosynthesis of indole-3-acetic acid. **Canadian Journal of Microbiology**, v.42, p.207-220, 1996.

Patterson, E. I. *et al.* Structural characterization of β -ketoacyl ACP synthase I bound to platencin and fragment screening molecules at two substrate binding sites. **Proteins**, v. 88, p. 47-56, 2020.

Pittard, J. Biosynthesis of the aromatic amino acids. In Neidhardt FC (ed) *Escherichia coli and salmonella: cellular and molecular biology*, vol 1 ASM, Washington, DC, pp 458-484, 1996.

Qiu, X.; Zhang, Y.; Hong, H. Classification of acetic acid bacteria and their acid resistant mechanism. **AMB Expr**, 2021.

Rashid, I. *et al.* Bacteria and fungi can contribute to nutrients bioavailability and aggregate formation in degraded soils. **Microbiological Research**, v.183, p. 26-41, 2016.

Reis, V.M.; Olivares, F.L. Improved methodology for isolation and identification of *Acetobacter diazotrophicus* and confirmation of its endophytic habitat. **World Journal of Microbiology and Biotechnology**, v.10, n.4, p.401-405, 1994.

Rodrigues, E. P. *et al.* Identification of genes involved in indole-3-acetic acid biosynthesis by *Gluconacetobacter diazotrophicus* PAL5 strain using transposon mutagenesis. **Front Microbiol**, v. 7, p. 1572, 2016.

Sabino, D. C. *et al.* Bactérias diazotróficas como promotoras do desenvolvimento inicial de plântulas de arroz. **Enciclopedia Biosfera**, v. 8, 2012.

Santos Dias, A.; Santos, C. C. Bactérias promotoras de crescimento de plantas: conceitos e potencial de uso. Nova Xavantina-MT: Pantanal, 2022. E-book: 100p.

Santos, I. B. *et al.* Selection and characterization of *Burholderia* spp. For Their Plant-Growth Promoting Effects and influence on Maize Seed Germination. **Frontiers in Soil Science**, v. 1, 2021.

Savino, S. *et al.* Kinetic and structural properties of a robust bacterial L-amino acid oxidase. **Catalysts**, v. 11, p. 1309, 2021.

Somers, E. *et al.* bacterial signaling: Alove parade beneath our feet. **Critical Reviews in Microbiology**, v.30, p.205-240, 2004.

Spero, M. A. *et al.* Phylogenomic Analysis and Predicted Physiological Role of the Proton-Translocating NADH:Quinone Oxidoreductase (Complex I) Across Bacteria. **mBio**, v. 6, 2015.

Subki, A. *et al.* Identification and characterisation of thiamine pyrophosphate (TPP) riboswitch in *Elaeis guineensis*. **PLoS ONE**, v. 15, 2020.

Szklarczyk, D. *et al.* STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic acids research*, v. 43, n. D1, p. D447-D452, 2015.

Taiz, L.; Zeiger, E. **Fisiologia vegetal**, 4.ed, Porto Alegre: ARTMED, p.820, 2009.

Timm, D. E. *et al.* Crystal structure of thiamin pyrophosphokinase. **J Mol Biol**, v. 310 p. 195-204, 2001.

Walport, L. J. *et al.* The characterization of protein interactions—what, how and how much?. **Chemical Society Reviews**, v. 50, n. 22, p. 12292-12307, 2021.

Yang, Z. *et al.* Overexpression of β -Ketoacyl-CoA Synthase From *Vitis vinifera* L. Improves Salt Tolerance in *Arabidopsis thaliana*. **Front Plant Sci.** v, 11. 2020.