UNIVERSIDADE FEDERAL DO CEARÁ

CENTRO DE CIÊNCIAS

DEPARTAMENTO DE BIOQUÍMICA E BIOLOGIA MOLECULAR

PROGRAMA DE PÓS-GRADUAÇÃO EM BIOQUÍMICA

ERICA MONIK SILVA ROQUE

ANÁLISE GENÔMICA COMPARATIVA DA FAMÍLIA DE GENES *GLICOLATO OXIDASE* (*GOX*) EM PLANTAS E PERFIL DE EXPRESSÃO DE *VuGOX1* DURANTE ESTRESSE BIÓTICO EM *Vigna unguiculata* [L.] Walp.

FORTALEZA
2022

ERICA MONIK SILVA ROQUE

ANÁLISE GENÔMICA COMPARATIVA DA FAMÍLIA DE GENES *GLICOLATO OXIDASE* (*GOX*) EM PLANTAS E PERFIL DE EXPRESSÃO DE *VuGOX1* DURANTE ESTRESSE BIÓTICO EM *Vigna unguiculata* [L.] Walp.

<div style="text-align: right">

Dissertação apresentada ao Programa de Pós-Graduação em Bioquímica da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Bioquímica. Área de concentração: Bioquímica

Orientador: Prof. Dr. Murilo Siqueira Alves

</div>

FORTALEZA

2022

ERICA MONIK SILVA ROQUE

ANÁLISE GENÔMICA COMPARATIVA DA FAMÍLIA DE GENES *GLICOLATO OXIDASE* (*GOX*) EM PLANTAS E PERFIL DE EXPRESSÃO DE *VuGOX1* DURANTE ESTRESSE BIÓTICO EM *Vigna unguiculata* [L.] Walp.

Dissertação apresentada ao Programa de Pós-Graduação em Bioquímica da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Bioquímica. Área de concentração: Bioquímica

Aprovada em: 25/02/2022

BANCA EXAMINADORA

_____
Prof. Dr. Murilo Siqueira Alves
Universidade Federal do Ceará (UFC)

_____
Prof. Dr. Humberto Henrique de Carvalho
Universidade Federal do Ceará (UFC)

_____
Dr. Lucas Pinheiro Dias
Universidade Federal de São Paulo (Unifesp) - Escola Paulista de Medicina (EPM)

A Deus.

Aos meus pais, José Milton e Maria de Fátima
e aos meus amigos que me ajudaram muito
nessa jornada, dedico esse trabalho

AGRADECIMENTOS

A Deus, pelo dom da vida

A minha família, por me apoiar em tudo que me disponho a fazer

Ao Prof. Dr. Murilo Siqueira Alves, pela excelente orientação, paciência e ensinamentos durante o mestrado.

A minha amiga Joseli, por todo apoio que me dá desde a graduação e aos amigos que ganhei no mestrado, Felipe e Fernanda. Sem vocês três essa jornada seria mais difícil.

Aos professores do curso de pós-graduação em Bioquímica, pelas contribuições para a minha formação.

Aos colegas de laboratório, que estavam sempre dispostos a ajudar

Aos membros da banca examinadora Prof. Dr. Humberto Henrique de Carvalho e Dr. Lucas Pinheiro Dias por ter aceitado participar dessa avaliação

À Universidade Federal do Ceará, ao programa de Pós-graduação em Bioquímica e ao Laboratório de Proteínas Vegetais de Defesa (LPVD), pelos recursos que também contribuíram com a realização deste trabalho.

"Não é o mais forte que sobrevive, nem o mais inteligente, mas o que melhor se adapta às mudanças." (MEGGINSON, 1963).

RESUMO

A Glicolato Oxidase (GOX) é uma enzima chave da fotorrespiração, uma via metabólica complexa de plantas que afeta a eficiência da fotossíntese e apresenta como um de seus produtos mais proeminentes o peróxido de hidrogênio (H_2O_2). As vias fotossintéticas e a produção de H_2O_2 podem diferir drasticamente entre plantas C3 e C4, onde várias distinções dependem da maquinaria de fotorrespiração. Tais contrastes impactam em processos fisiológicos cruciais nas plantas, tais como desenvolvimento e respostas a estresses. No entanto, poucos estudos trazem luz aos aspectos evolutivos e estruturais dos componentes da fotorrespiração, bem como faltam análises comparativas de famílias de genes relacionados com a fotorrespiração em plantas C3 e C4. No presente estudo, apresentamos a primeira análise genômica comparativa da família de genes *GOX* em plantas, comparando características evolutivas e estruturais relevantes de distintos ortólogos de *GOX* em famílias vegetais. As relações evolutivas, estrutura genética, motivos conservados, predição de *cis*-elementos de promotores, localização cromossômica e colinearidade interespecífica foram analisadas para obter uma melhor compreensão da família de genes *GOX* em plantas. O perfil de expressão de um ortologo de *GOX* obtido após estímulos de estresse biótico foi investigado em plantas de *Vigna unguiculata* apresentando diferentes perfis genéticos, tendo em vista estudos recentes mostrando a expressão diferencial de proteínas GOX durante o estresse biótico em cultivares resistentes desta espécie. Foram observadas divergências evolutivas e estruturais dependentes da família vegetal entre os distintos genes *GOX*, com a maior conservação de genes sendo verificada entre os membros da família *Fabaceae*. *VuGOX1* é modulado precocemente durante interações incompatíveis com CPSMV em *Vigna unguiculata*, mostrando sua relevância em respostas rápidas de defesa de plantas. A elevada divergência verificada entre os ortólogos de *GOX* de *Fabaceae* e *Poaceae* pode ter impacto na divergência funcional entre estas famílias de genes. Este estudo comparativo fornece um quadro abrangente dos aspectos evolutivos e estruturais da família de genes *GOX* em plantas, bem como enfatiza o envolvimento de ortólogos de *GOX* nas respostas a estresses em plantas.

Palavras-chave: análise genômica comparativa; filogenética molecular; análise de sintenia; metabolismo do glicolato; estresse biótico.

ABSTRACT

Glycolate Oxidase (GOX) is a key enzyme in photorespiration, a complex metabolic pathway in plants that affects photosynthesis efficiency and presents as one of its most prominent products hydrogen peroxide (H_2O_2). Photosynthetic pathways and H_2O_2 production can drastically differ between C3 and C4 plants, where several distinctions rely on the photorespiration machinery. Such contrasts impact in crucial physiological processes in plants, such as development and stress responses. However, few studies bring light to evolutionary and structural aspects of the photorespiration components, as well as comparative analyses of gene families related to photorespiration in C3 and C4 plants are lacking. In the present study, we present the first genome-wide comparative analysis of *GOX* gene family in plants, comparing relevant evolutionary and structural aspects of distinct *GOX* orthologs in plant families. The evolutionary relationships, gene structure, conserved motifs, promoter *cis*-elements prediction, chromosome location and interspecific collinearity were analyzed in order to gain better understanding of *GOX* gene family in plants. The expression profile of a *GOX* ortholog following biotic stress stimuli was investigated in *Vigna unguiculata* plants exhibiting different genetic backgrounds, in view of recent studies showing differential expression of GOX proteins during biotic stress in resistant cultivars of this species. Family-dependent evolutionary and structural divergence were observed among distinct *GOX* genes, with higher gene conservation verified among members of *Fabaceae* family. *VuGOX1* is early modulated during incompatible interactions with CPSMV in *Vigna unguiculata*, showing relevance in rapid plant defense responses. High divergence verified among *Fabaceae* and *Poaceae GOX* orthologs might impact in functional divergence between these gene families. This comparative study provides a comprehensive picture of evolutionary and structural aspects of the *GOX* gene family in plants, as well as emphasizes the involvement of *GOX* orthologs in plant stress responses.

Keywords: genome-wide comparative analysis; molecular phylogenetics; synteny analysis; glycolate metabolism; biotic stress.

SUMÁRIO

1. INTRODUÇÃO

1.1. Vigna unguiculata

O feijão-caupí (*Vigna unguiculata* [L.]Walp.), pertencente à família *Fabaceae*, é uma planta anual geralmente ereta, que apresenta caule com nervuras, folhas trifolioladas lisas. e sementes que apresentam cores variadas, sendo estas consumidas cozidas, e a planta adulta apresentando potencial forrageiro (SAMAILA et. al., 2019). A subespécie *unguiculata* está dividida em cinco grupos: *Unguiculata, Sesquipedalis, Textilis, Biflora* e *Melanophthalmus*, sendo o grupo *unguiculata* o mais cultivado (BOUKAR et al, 2020). Apesar da espécie ser originária da África, não se sabe exatamente onde foi domesticada (LO et al, 2018).

O cultivo do feijão caupí contribui de maneira significativa com a agricultura camponesa, visto que a cultura tem poucas necessidades de insumos (GONÇALVES et al, 2016). A espécie tem uma considerável tolerância a seca e cresce em solos de baixo teor nutricional (JIE JI, 2019). É cultivada no sudeste asiático, África, sul dos Estados Unidos, América latina e algumas regiões do sul da Europa (JAYATHILAKE et al, 2018; KARAPANOS et al, 2017).

Em 2019, foram plantados mundialmente mais de 14 milhões de hectares de feijão caupí, obtendo como resultado uma produção de quase 9 milhões de toneladas de grãos secos. O continente Africano se destaca como o maior produtor dessa cultura, onde a Nigéria produziu um pouco mais de 3,5 milhões de toneladas, seguido de Níger e Bukina Faso. Ásia e Europa possuem uma produção menor comparada ao continente africano (FAO, 2021).

No Brasil, o feijão-caupí é produzido nas regiões Norte, Nordeste, Centro-Oeste e Sudeste, com destaque no Nordeste e Centro-Oeste. No período de 2018/2019 houve uma produção de mais de 600 mil toneladas de grãos, onde os maiores produtores foram o Ceará com mais de 105,7 mil toneladas e Mato Grosso com 151,5 mil toneladas. O Sudeste e o Norte não apresentaram uma produção significativa, quando comparada com as regiões Nordeste e Centro-Oeste (CONAB, 2019).

O feijão-caupí é uma cultura resiliente em determinadas condições edafoclimáticas adversas, porém alguns fatores podem prejudicar a produção, tais como estresses abióticos, como a seca, e estresses bióticos, que englobam doenças causadas por bactérias, fungos, nematoides, insetos e vírus (VARELA et al., 2017; SILVA et al., 2016). Muitas espécies de vírus podem infectar a cultura, dentre elas se destaca o Vírus do Mosaico Severo do Caupí (*Cowpea Severe Mosaic Virus*, CPSMV), responsável por causar grandes prejuízos ao cultivo (PAIVA et al., 2016; NORONHA SOUZA et al., 2019).

O vírus do CPSMV prejudica principalmente a fotossíntese (SOUZA et al., 2016), onde a planta acometida por essa doença apresenta sintomas severos como: encrespamento na lâmina foliar causado por bolhosidades, presença de mosqueado semelhante a um mosaico de cor verde claro ou escuro, e atraso no desenvolvimento das nervuras principais que resulta em deformação no limbo foliar. Quando infectada no início do ciclo de desenvolvimento a planta apresenta intenso nanismo, causando grandes prejuízos, com perdas podendo chegar a mais de 80%. Atualmente, as principais estratégias para manejo do mosaico severo do caupi se baseiam no uso de cultivares resistentes (SOBRINHO, 2016).

1.2. Mecanismos de defesa de plantas contra patógenos

A primeira linha de defesa é representada pelo reconhecimento de moléculas presentes nos organismos exógenos, conhecidos como MAMPs ou PAMPs (Microbe- or Pathogen-associated Molecular Patterns). Estes padrões moleculares são reconhecidos por proteínas localizadas na membrana celular da planta denominadas PRRs (Pattern Recognition Receptors), induzindo uma camada de defesa chamada de MTI (MAMP-triggered Immunity) e PTI (Pattern-triggered Immunity) (GRANDELIS et al, 2019; NEWMAN et al, 2013).
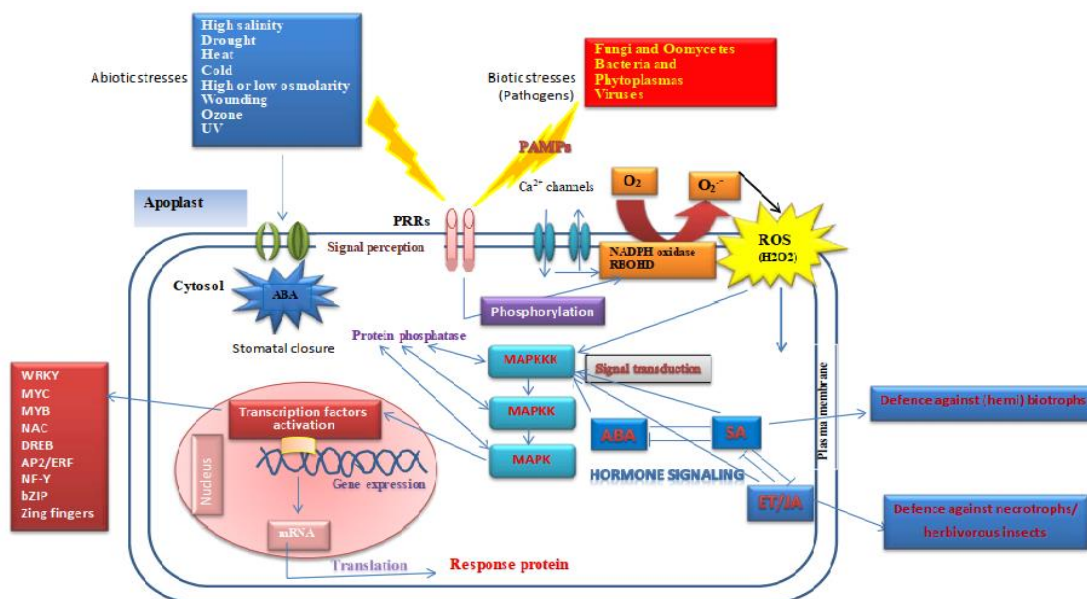
Essas PRRs encontradas em plantas são receptoras RLKs e RLPs (Receptor-like Kinases e Receptor-like Proteins). RLKs possuem um domínio extra celular, um domínio transmembrana de passagem única e um domínio cinase. RLPs são similares, porém não apresentam o domínio cinase. A partir da fosforilação de moléculas ligadas a essas proteínas a resposta de defesa é iniciada. Tais respostas incluem deposição de calose, geração de óxido nítrico (NO), produção de ROS (Reactive Oxygen Species), assim como a expressão de genes relacionados a imunidade (GRANDELIS et al., 2018; HAN, 2019).

A segunda linha de defesa consiste no reconhecimento de efetores de patógenos por produtos de genes de resistência (R), desencadeando a camada de defesa denominada ETI (Effector-triggered Immunity). Genes R podem atuar na regulação da expressão de genes envolvidos em interações planta-patógeno. A ETI é uma resposta de defesa eficaz que pode desencadear a denominada resposta hipersensitiva (Hypersensitive Response, HR), um tipo de morte celular programada associada a imunidade de plantas (HAN, 2018; MUTHAMILARASAN; PRASAD, 2013).

Quanto à produção de ROS em respostas de defesa de plantas, um percentual de 70% do peróxido de hidrogenio (H_2O_2) liberado por plantas com metabolismo fotossintético C3 é oriundo da catálise da oxidação do composto glicolato pela enzima peroxissomal Glicolato

Oxidase durante a fotorrespiração, onde este percentual aumenta durante estresses ambientais (ZHANG et al, 2017). O H_2O_2 é uma espécie reativa de oxigênio que apresenta um importante papel na sinalização de processos moleculares ligados a defesa de plantas. Estudos demonstram que o H_2O_2 peroxissomal representa uma fonte fundamental para a resposta a estresses bióticos e no desencadeamento de processos de morte celular programada em plantas (CUI et al, 2016). Estresses bióticos e abióticos também desencadeiam respostas hormonais, como a ação do ácido salicílico (SA), ácido abscísico (ABA), etileno (ET) e ácido jasmônico (JA). Todos estes fitohormônios desempenham um papel crucial na modulação da defesa de plantas a estresses ambientais, integrando diferentes vias de sinalização celular (Figura A) (NEJAT; MANTRI, 2017; ABUQAMAR; MOUSTAFA; TRAN 2017). O H_2O_2 produzido durante a fotorrespiração apresenta papel fundamental neste sistema integrado de resposta.

Figura A. O sistema integrado de defesa de plantas em resposta a estresses bióticos e abióticos.

Fonte: (NEJAT; MANTRI, 2017)

1.3. Fotorrespiração

A fotorrespiração em plantas com metabolismo fotossintético C3 (Via glicolato) inicia no cloroplasto, com a atividade de oxigenase da enzima ribulose 1,5-bifosfato carboxilase-oxigenase (RuBisCO), produzindo fosfoglicolato (2-PG), que será transformado em glicolato pela 2-PG fosfatase (PGP). Na sequência, o glicolato é transferido para o peroxissomo, onde será oxidado a glioxilato por GOX, resultando na liberação de H_2O_2 (FOYER et al., 2009).

Em plantas com metabolismo fotossintético C4 (Via de Hatch-Slack), a fotorrespiração é reduzida, no entanto a via não é completamente inativa, pois é possível encontrar enzimas envolvidas na via glicolato, tais como GOX, hidroxipiruvato redutase (HPR) e catalase (ZELITCH et al, 2009; UENO 2005). Em plantas C4, ocorre a acumulação de CO_2 nas proximidades da RuBisCO, impossibilitando o oxigênio molecular de disputar o sitio ativo da enzima. O processo de fixação de carbono em plantas C4 acontece nas células da bainha de feixes e nas células do mesófilo. Essas células envolvem os vasos foliares em círculos semelhantes a grinaldas, que recebe o nome de anatomia Kranz. O CO_2 é fixado em compostos de 4 carbonos [ácido málico (MAL) ou ácido aspártico (ASP)], pela fosfoenolpiruvato (PEP) carboxilase, que na sequência é descarboxilado na bainha de feixes e fixado novamente como compostos de 3 carbonos (LANGDALE, 2011; FURBANK, 2011). Apesar de também ser encontrada nas células do mesófilo, é na bainha de feixes que GOX apresenta sua maior atividade (POPOV,2003).

Um estudo utilizando várias espécies C4 mostrou que a atividade da enzima GOX é maior naquelas que possuem o índice granal bem desenvolvido (UENO,2005). Quando comparados os parâmetros cinéticos de GOX1 e GOX2 de *Arabidopsis thaliana* com GOX1 de milho (*Zea mays*), ambos se mostraram similares (DELLERO et al., 2015). Em um outro trabalho realizada com plantas de milho mutantes com o gene *GOX* afetado, a deficiência de GOX prejudicou a eficiência fotossintética mesmo em um ambiente rico em CO_2, mostrando que o glicolato é um substrato processado exclusivamente por GOX, e sua acumulação inibe processos metabólicos ligados a via fotossintética C4 (ZELITCH et al, 2009).

Em estudos genômicos em plantas, é possível identificar diversos genes da família *GOX*, no entanto, alguns membros específicos apresentam uma maior atividade de oxidação do glicolato, atuando como pontos centrais de modulação da fotorrespiração. Em arroz, Arabidopsis e milho, *OsGLO*1 e *OsGLO4*, *AtGOX1* e *AtGOX2*, e *ZmGOX1* são os principais

homólogos responsáveis pela fotorrespiração, respectivamente (ROJAS et al., 2012; ZELITCH et al., 2009; ZHANG et al., 2017).

1.4. Glicolato oxidase

A Glicolato Oxidase (GOX) é uma enzima alfa-hidroxi-ácido oxidase dependente de Flavina Mononucleotídeo (FMN), que apresenta uma participação chave no processo de fotorrespiração. A catálise de GOX ocorre com a oxidação do glicolato a glioxilato pela coenzima FMN. Em seguida, FMN reduzido é reoxidado pelo oxigênio molecular, liberando peróxido de hidrogênio (H_2O_2). FMN é um cofator fundamental para a catálise de GOX (LIU; WU; CHEN, 2018; XU; YANG; CAI, 2018). Os resíduos de aminoácidos Tirosina, Serina, Treonina, Lisina, Arginina e Histidina são os principais envolvidos na regulação de FMN (Lindqvist, 1989; Dellero, 2015). Estudos com *A. thaliana* (*GOX1* e *GOX2*) e *Z. mays* (*GOX1*) realizaram mudanças estruturais em sítios de fosforilação de *GOX* recombinante, afetando a atividade da enzima devido a modificação dos resíduos de regulação, com consequente perda do cofator FMN ((JOSSIER et al., 2019)). Em arroz, o cofator FMN está envolvido na regulação da biossíntese de auxinas (LI et al., 2021b)

Atualmente, não existem trabalhos comparativos a respeito da família de genes *GOX* em plantas C3 e C4 utilizando dados genômicos. No entanto, trabalhos mais recentes mostram que em *Nicotiana benthamiana* e *A. thaliana* o H_2O_2 é indispensável no desencadeamento da resposta hipersensitiva e na imunidade não hospedeira (LI et al., 2015; ROJAS et al., 2012). Um estudo genômico realizado com *N. benthamiana* identificou 16 genes GOX, sendo que três deles mostraram relacionamento com vários tipos de resistência a doenças(XU; YANG; CAI, 2018) .

Varela e colaboradores (2019), identificaram várias proteínas diferencialmente expressas durante a infecção por CPSMV em uma cultivar resistente de *V. unguiculata* submetida a estresse salino, dentre elas uma proteína homóloga à glicolato oxidase peroxisomal (GOX-like), que se mostrou fortemente expressa no sexto dia após a inoculação do vírus. Em trabalho anterior realizado com a mesma cultivar resistente, e utilizando apenas o vírus CPSMV, a expressão induzida de GOX ocorreu no segundo dia após inoculação, e não foi detectada expressão diferencial no sexto dia (Varela et al, 2017).

Para entender melhor aspectos da evolução e estrutura de GOX em plantas com diferentes perfis fotossintéticos (C3 e C4), neste trabalho foi realizada uma caracterização estrutural do gene *GOX* em espécies vegetais com metabolismo fotossintético C3 e C4, bem como a

reconstrução das relações evolutivas entre os genes destas diferentes espécies, visando compreender a evolução da família de genes *GOX*. Também foi analisado o perfil de expressão do gene ortólogo *VuGOX1* em *V. unguiculata* sob tratamento com ácido salicílico e após inoculação com *Cowpea Severe Mosaic Vírus* (CPSMV) em plantas resistentes e suscetíveis ao vírus. Além de representar a primeira análise em escala genômica desta família gênica em plantas, este estudo visa verificar o padrão de expressão em resposta a estresse biótico de um representante da família *GOX* em *V. unguiculata*, organismo diplóide pertencente à família das *Fabaceae*.

2. OBJETIVO GERAL

Realizar a análise genômica comparativa da família de genes *GOX* em plantas com metabolismo fotossintético C3 e C4, visando compreender a relação entre evolução, estrutura e função da enzima em espécies com perfis fotossintéticos distintos, bem como avaliar o perfil de expressão de um gene ortólogo *GOX* em *Vigna unguiculata* sob estresse biótico.

3. ARTIGO

Genome-wide comparative analysis of *Glycolate oxidase* (*GOX*) gene family in plants and expression profile of *VuGOX1* during biotic stress in *Vigna unguiculata*

Erica Monik Silva Roque[1], Felipe de Castro Teixeira[1], Alex Martins de Aguiar[1], Sâmia Alves Silva[1], Jose Tadeu Abreu de Oliveira[1], Murilo Siqueira Alves[1*]

[1] Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará, Avenida Humberto Monte S/N, Campus Pici, Fortaleza, CE 60440-900, Brazil

*Corresponding author: murilo.alves@ufc.br

ABSTRACT

Glycolate Oxidase (GOX) is a key enzyme in photorespiration, a complex metabolic pathway in plants that affects photosynthesis efficiency and presents as one of its most prominent products hydrogen peroxide (H_2O_2). Photosynthetic pathways and H_2O_2 production can drastically differ between C3 and C4 plants, where several distinctions rely on the photorespiration machinery. Such contrasts impact in crucial physiological processes in plants, such as development and stress responses. However, few studies bring light to evolutionary and structural aspects of the photorespiration components, as well as comparative analyses of gene families related to photorespiration in C3 and C4 plants are lacking. In the present study, we present the first genome-wide comparative analysis of *GOX* gene family in plants, comparing relevant evolutionary and structural aspects of distinct *GOX* orthologs in plant families. The evolutionary relationships, gene structure, conserved motifs, promoter *cis*-elements prediction, chromosome location and interspecific collinearity were analyzed in order to gain better understanding of *GOX* gene family in plants. The expression profile of a *GOX* ortholog following biotic stress stimuli was investigated in *Vigna unguiculata* plants exhibiting different genetic backgrounds, in view of recent studies showing differential expression of GOX proteins during biotic stress in resistant cultivars of this species. Family-dependent evolutionary and structural divergence were observed among distinct *GOX* genes, with higher gene conservation verified among members of *Fabaceae* family. *VuGOX1* is early modulated during incompatible interactions with CPSMV in *Vigna unguiculata*, showing relevance in rapid plant defense responses. High divergence verified among *Fabaceae* and *Poaceae GOX* orthologs might impact in functional divergence between these gene families. This comparative study provides a comprehensive picture of evolutionary and structural aspects of the *GOX* gene family in plants, as well as emphasizes the involvement of *GOX* orthologs in plant stress responses.

Keywords: Genome-wide comparative analysis; Molecular phylogenetics; Synteny analysis; Biotic stress; Differential expression; Glycolate metabolism

Introduction

Glycolate Oxidase (GOX, GLO, GO, HAOX) is a Flavin Mononucleotide (FMN)-dependent alpha-hydroxyacid oxidase, which exhibits a key role in the process called photorespiration. Photorespiration is a complex metabolic pathway that directly impact photosynthesis efficiency, and is parametrical for differentiation of C3 and C4 plants (LIU; WU; CHEN, 2018b; XU; YANG; CAI, 2018).

In plants presenting C3 photosynthetic pathway (glycolate pathway), photorespiration starts in chloroplasts with the oxygenase activity of the enzyme ribulose 1,5-bisphosphate carboxylase-oxygenase (RuBisCO), producing phosphoglycolate (2-PG), which will be transformed in glycolate by 2-PG phosphatase (PGP). Following this step, glycolate is transferred to peroxisomes, where it will be oxidised to glyoxylate by GOX, resulting in hydrogen peroxide (H_2O_2) release (FOYER et al., 2009)(Fig. 1).

In plants with C4 photosynthetic pathway (Hatch-Slack pathway), carbon dioxide (CO_2) accumulation occurs in the vicinity of RuBisCO, impairing molecular oxygen (O_2) to dispute the active site of RuBisCO. The process of carbon fixation in C4 plants occurs in the bundle sheath cells and in the mesophyll cells, where it is fixed in 4-carbon compounds [malic acid (MAL) or aspartic acid (ASP)] by PEP carboxylase, after which it is decarboxylated in bundle sheath cells and fixed again as 3-carbon compounds (FURBANK, 2011; LANGDALE, 2011). The major GOX activity occurs in bundle sheath cells, but it is also found in mesophyll cells (POPOV et al., 2003)(Fig. 1). In grasses (*Poaceae*), this activity is different among species, and the feature that determines GOX activity level is the granule index in bundle sheath cells (UENO; YOSHIMURA; SENTOKU, 2005).

GOX enzymatic reaction occurs when glycolate is oxidized to glyoxylate by FMN. Following this step, reduced FMN is reoxidised by O_2, releasing H_2O_2 (LIU; WU; CHEN, 2018b). H_2O_2 is a reactive oxygen species (ROS) that plays a key role in signalling linked to plant defence. In plants, a plethora of biochemical sources of H_2O_2 occur, but studies demonstrate peroxisomal H_2O_2 represents a crucial source of H_2O_2 for biotic stresses responses and programmed cell death (PCD) triggering in plants (CUI et al., 2016). Photorespiration is the source of 70% of H_2O_2 in C3 plants, and this percentage increases if a plant is subjected to environmental stresses (ZHANG et al., 2017c). For instance, studies where *GOX* genes were overexpressed in rice showed there was an increase in photosynthesis when subjected to high temperatures and high light exposition (CUI et al., 2016). On the other hand, although plants with C4 photosynthetic pathway have as main advantage the reduction of photorespiration, optimizing carbon assimilation by photosynthesis, but the pathway is not completely inactive
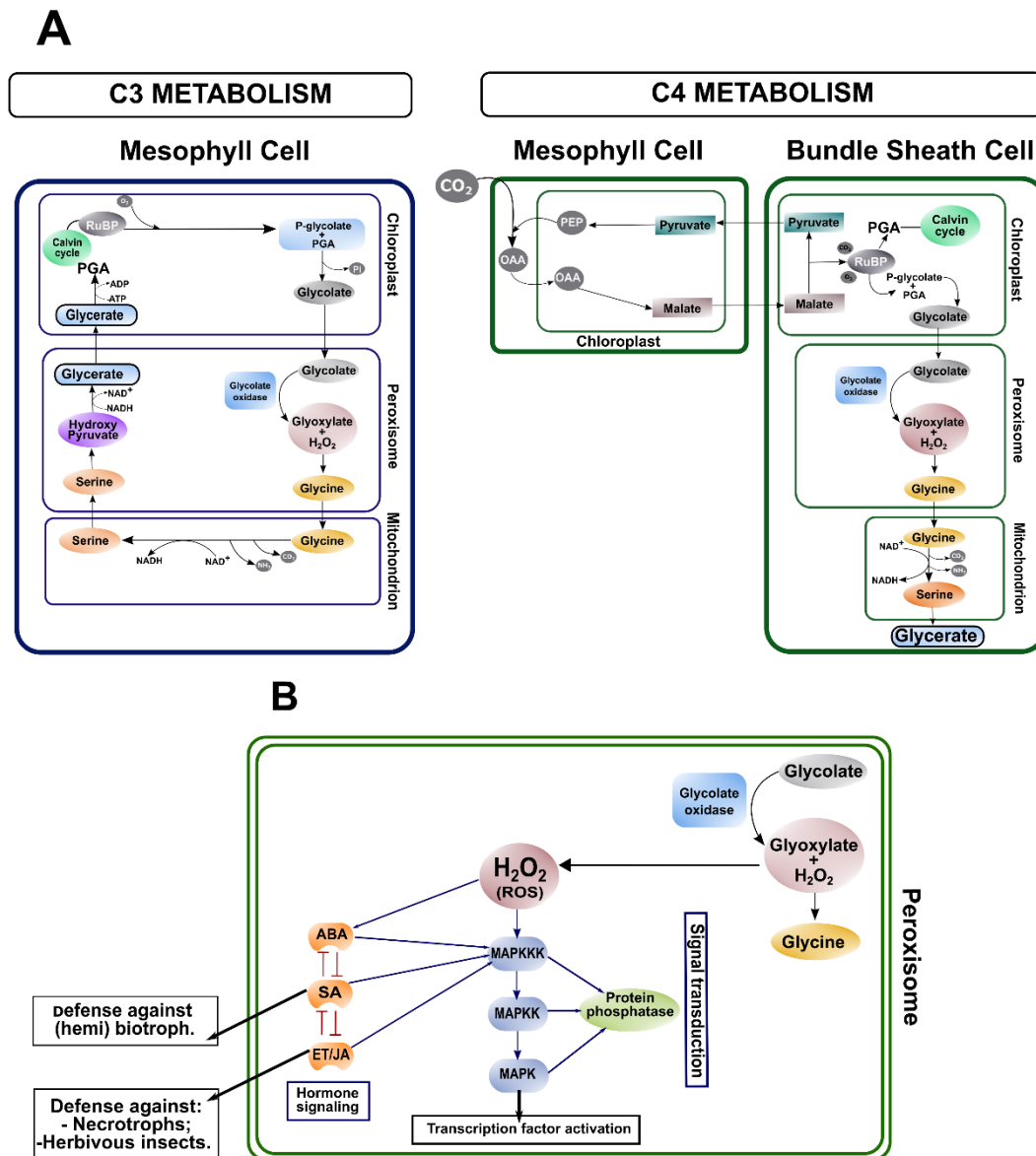
(UENO; YOSHIMURA; SENTOKU, 2005; ZELITCH et al., 2009) . Research carried out using mutant maize (*Zea mays*) showed glycolate oxidation is impaired in a *GOX-deficient* plant even in a CO$_2$-rich environment, since glycolate can only be oxidazed by *GOX,* and its accumulation can inhibit metabolic processes linked to C4 pathway (ZELITCH et al., 2009).

Studies performed in *Arabidopsis thaliana*, *Nicotiana benthamiana*, and melon (*Cucumis melo*) demonstrate *GOX* is related to pathogen response, with function varying among different plant species, gene family members and types of response. In Arabidopsis, the GOX1 protein is primarily responsible for producing H$_2$O$_2$ directed to plant defense responses (Fig. 1) (DELLERO et al., 2016a; NOCTOR; MHAMDI, 2017; XU; YANG; CAI, 2018) . In *Fabaceae* family, Varela and collaborators identified several proteins differentially expressed during CPSMV infection in resistant cultivars of identified several proteins differentially expressed during CPSMV infection in a resistant cultiv*Vigna unguiculata* (VARELA et al., 2017)Among them, a protein homologous to peroxisomal glycolate oxidase (GOX-like), which was differentially expressed on the second day after virus inoculation. Another work performed using the same resistant cultivar during CPSMV infection and salt stress showed strong *GOX* expression on the sixth day post-infection, and differential expression was not detected on the second day (VARELA et al., 2017, 2019).

Despite the evident importance of *GOX* family during defense responses in both C3 and C4 plant species, there is no comparative study about *GOX* genes from C3 and C4 plants in scientific literature. Many plant species had the *GOX* gene identified in their genomes, where is generally found more than one member of the gene family in each species analysed. For instance, *Arabidopsis thaliana, Oryza sativa and Nicotiana benthamiana* present 5, 6 and 16 members of the *GOX* family, respectively (LIU; WU; CHEN, 2018b; XU; YANG; CAI, 2018). However, even with several plant organisms presenting genome available in public databases, the knowledge concerning structure, evolution and distribution of *GOX* genes on such genomes is lacking.

In this work, we present a comparative analysis of the *GOX* gene family in model and crop plant species, presenting C3 and C4 photosynthetic metabolism. Structural characterization, phylogenetic and synteny analyses of *GOX* genes were performed, aiming to better understand the evolution of this enzyme in plants presenting distinct photosynthetic pathways. Using a diploid species belonging to the *Fabaceae* family, an agronomically relevant plant family, the expression profile of a *V. unguiculata GOX* ortholog gene (*VuGOX1*) was obtained, under salicylic acid (SA) treatment, and after inoculation with Cowpea Severe Mosaic Virus (CPSMV), in resistant and susceptible plants. In addition to be the first genome-wide

analysis of the *GOX* family in plants, this study aims to shed light to important disparities among C3 and C4 plants that relies on aspects of *GOX* gene evolution, structure and function.

Source: Elaborated by author

Fig. 1 The role of GOX in C3 and C4 photosynthetic pathways. A) Inside mesophyll cell peroxisomes of C3 plants, GOX catalyzes the oxidation reaction transforming glycolate to glyoxylate, releasing hydrogen peroxide (H_2O_2). On the other hand, in C4 plants, the action of GOX occurs predominantly in the bundle sheath cells, also resulting in H_2O_2 release during the oxidation of glycolate to glyoxylate. B) During pathogen and insect attacks in *Arabidopsis*, H_2O_2 released from GOX reaction triggers MAPK cascades and hormone signaling pathways, which are responsible to activate a plethora of plant immunity mechanisms. Glycolate oxidase (GOX) is represented by a light blue box in the scheme. Enzymes not shown are implicit in the scheme

Materials and methods

Identification of GOX gene orthologs and phylogenetic analysis

All model and crop plant species chosen in this study have sequenced genomes deposited in public databases. Plant species from different families and photosynthetic systems (C3 and C4) were selected in our study. Firstly, using an amino acid sequence from a glycolate oxidase (GOX) ortholog of *Vigna unguiculata* (genome version v1.2) as query, searches were performed in the Phytozome comparative genomics database version 13 (https://phytozome.jgi.doe.gov/pz/portal.html)(GOODSTEIN et al., 2012), using the BLAST tool, to obtain nucleotide and amino acid sequences of members from *GOX* gene family from *Vigna unguiculata, Glycine max, Lotus japonicus, Medicago truncatula, Phaseolus vulgaris, Oryza sativa, Sorghum bicolor, Zea mays, Arabidopsis thaliana, Solanum tuberosum, Gossypium raimondii* and *Chlamydomonas reinhardtii* species. The NCBI database (https://www.ncbi.nlm.nih.gov/) was used to obtain the sequences of *Vigna radiata*, and the specialized database Vigna Genome Server (https://viggs.dna.affrc.go.jp/)(SAKAI et al., 2016) was used for the sequences of *Vigna angularis*. Secondly, using the predicted amino acid sequences obtained from previous BLAST searches, the probabilistic prediction method of Hidden Markov chains embedded the Pfam platform (http://pfam.xfam.org/)(MISTRY et al., 2021) was applied to detect and confirm the FMN (FMN-dh) domain. Finally, using peptide profile signatures in the InterPro database (https://www.ebi.ac.uk/interpro/)(BLUM et al., 2021), the experimentally curated UniProt database (https://www.uniprot.org/) and the conserved domains CD-NCBI database (https://www.ncbi.nlm.nih.gov/), the sequences coding the FMN domain were validated for further analysis.

Phylogenetic relationships were reconstructed by performing a multiple sequence alignment using the coding DNA sequence (CDS) retrieved from the 14 species selected. For multiple alignment, the MUSCLE tool embedded in MEGA X software suite (KUMAR et al., 2018)was used. All multiple sequence alignments were manually edited and curated. A Maximum Parsimony analysis in MEGA X was performed for phylogram generation, with bootstrapping of 1000 replications validating the reconstructed topology.

Chromosomal Location and Synteny Analysis

Information regarding *GOX* orthologs' locations in genomic sequences of *V. unguiculata*, *G. max*, *O. sativa* and *S. bicolor* was obtained from Phytozome database. The start and end positions of *GOX* genes and chromosome sizes were arranged in excel files for submission in

the chromosomal mapping. For mapping of gene position on chromosomes, MG2C tool v2.1 (http://mg2c.iask.in/mg2c_v2.1/)(CHAO et al., 2021) was used, using default parameters. To perform the synteny analysis, the MCScanX tool(WANG et al., 2012), embedded in TBtools software suite (version 1.098661) (CHEN et al., 2020)was used with default parameters.

Gene structure and conserved motif identification

To obtain the structure of *GOX* genes, the genomic and CDS sequences corresponding to the 12 plant species were retrieved from Phytozome database and submitted to GSDS 2.0 program (http://gsds.gao-lab.org/index.php)(HU et al., 2015), that performs sequence mapping using est2genome, a tool able to detect introns into a spliced sequence. To obtain the conserved motifs of the predicted GOX proteins, a file containing the CDS sequences of the *GOX* genes was submitted to MEME online tool (http://meme-suite.org)(BAILEY; ELKAN, [s.d.]). The map of the conserved motifs was visualized using the TBtools.

Cis-element identification

The 1500 bp sequence upstream the translation start codon of *GOX* genes, retrieved from Phytozome database, was used. The sequences were submitted to the online tool PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/)(LESCOT, 2002) to predict putative *cis*-elements. The map of *cis*-elements was visualized using TBtools.

Plant materials

For the expression analysis of *VuGOX1* in *Vigna unguiculata*, the cultivar Pitiúba (susceptible to CPSMV) and BR 14 Mulato (resistant to CPSMV) were used. Seeds from both cultivars were obtained from the Seed Analysis Laboratory (LAS) of the Universidade Federal do Ceará. The plants were maintained in a greenhouse at the Department of Biochemistry and Molecular Biology of the Universidade Federal do Ceará, under local climatic conditions (12-hour photoperiod, and temperature of 32 ± 2 °C during the day and 27 ± 2 °C at night).

The seeds were planted in plastic pots containing a substrate composed of 2 parts of soil, 1 part of sand and 1 part of peat, according to (CARVALHO et al., 2019). One seed per pot was watered until it reached the V3 developmental stage. Following this, experiments were conducted using CPSMV inoculation and Salicylic acid (SA) treatment for further RNA extraction.

CPSMV inoculation

Leaves from *V. unguiculata* plants, cultivar Pitiúba, previously infected with CPSMV and stored at -80°C, were used for inoculation of plants at V3 stage, according to the protocol described by (VARELA et al., 2017), with modifications. Briefly, leaves were triturated with 10 mM sodium phosphate buffer where carborundum abrasive powder was added. Wearing gloves, the suspension was gently rubbed onto leaves, which were collected 1 hour and 16 hours after inoculation, such as described in (FERREIRA-NETO et al., 2021). As control treatment, a false inoculation (MOCK) was performed using the same suspension without CPSMV. For each time, 3 biological replicates were used, containing one plant in each pot. After each incubation time, leaves were collected, and RNA extraction of the collected samples was immediately performed. Plants symptoms were monitored along the days after inoculation, in order to confirm CPSMV infection (Supplementary Fig. 1).

A. Pitiúba (CE-31)

B. Pitiúba (CE-31)

C. BR 14 Mulato

D. BR 14 Mulato

Source: Elaborated by author

Supplementary Fig. 1 CPSMV infection symptoms in two contrastant *V. unguiculata* cultivars, 12 days after inoculation. A) Pitiúba (susceptible) leaf after false-inoculation (mock), without infection symptoms. B) Pitiúba leaf after CPSMV inoculation, showing infection symptoms, such as leaf chlorosis, leaf lamina deformation and leaf growth restriction. C) and D) BR 14 Mulato (resistant) leaves after false-inoculation and CPSMV inoculation, respectively, both presenting absence of infection symptoms

Salicylic acid treatment

An alcohol solution containing salicylic acid (SA) at 1mM, with ethanol concentration of 0.1g/1ml in water, was used for application on plants at V3 stage, according to Razmi and collaborators (RAZMI et al., 2017). Leaves were collected 1 hour and 16 hours post application, and as control treatment alcohol solution without SA was applied. For each time point, 2 biological replicates were used, containing one plant in each pot, and after each application time, the leaves were collected and RNA extraction was performed immediately.

RNA extraction and gene expression analysis by quantitative real-time PCR (qPCR)

For RNA extraction, the plant materials obtained from the respective treatments were collected and triturated using autoclaved graal and pistil. Total RNA was extracted using the TRIzol reagent (Invitrogen), according to the manufacturer's instructions. Following the extraction, the RNA was quantified using a Microplate Spectrophotometer (Epoch), treated with RNAse-free DNAse (Promega), and then used for cDNA synthesis, using M-MLV reverse transcriptase (Promega), according to the manufacturer's guidelines.

Finally, gene expression analysis was performed using quantitative real-time PCR (qPCR), where specific primers for each target and reference gene (Supplementary Table 1), GoTaq Master Mix for qPCR (Promega), and the synthetized cDNA were used for qPCR reactions, according to the manufacturer's instructions (Promega), in a volume of 10 uL. The amplification parameters used were 10 min at 95 °C; 45 cycles of 94 °C for 15 s, and 60 °C for 1 min, in a QuantStudio 3 equipment (Thermofisher scientific). To quantify relative gene expression, Pfaffl's mathematical model was applied (PFAFFL, 2001)using F-BOX, UBIQUITIN and L23A3 as reference genes. The reference genes were validated using the BestKeeper tool, according to PFAFFL (PFAFFL et al., 2004). Statistical analysis was perform using ΔCq values, according to Yuan and collaborators (YUAN et al., 2006).

Compliance with biodiversity regulations

All organisms used in the present work were registered in SISGEN (Sistema Nacional de Gestão do Patrimônio Genético e do Conhecimento Tradicional Associado), under accession A0E5BE1.

Supplementary Table 1 Primer sequences of *Vigna unguiculata* reference and target genes, gene product description, amplicon size, accession number at Phytozome database, and optimal melting temperature (Tm) of each oligonucleotide

Gene	Gene product	Primer sequence (5′—3′)	Amplicon (bp)	Fw/ Rv Tm (°C)	Accession number
VuGOX1	(S)-2-hydroxy-acid oxidase / Hydroxy-acid oxidase B (GOX)	Fw: GCTGGATTCAAAGCCATTGC Rv: CGGTGGCAGTGTGAATCTGT	90	59/58	Vigun01g052400.1
VuPR1	Pathogenesis-related protein 1 (PR1)	Fw: CAAGGTCGCAGGTTGGTGTT Rv: AGTCACCTCTGCGTTGGTTTG	90	60/59	Vigun06g213100.1
VuMC1	Metacaspase-1 (MC1)	Fw: CCCACGGCCGCAAGA Rv: GTCGTTGATGCATCCCTTGA	80	60/58	Vigun05g144300.1
VuVSP2	Vegetative storage protein 2 (VSP2)	Fw: TGCAAACTGGGTTGCTGAAG Rv: TGCCGAGAGACAGGAGTTTGT	80	59/59	Vigun10g178100.1
VuFBOX3	F-box protein 3 (F-BOX3)	Fw:AAATGAATATGGCCGAAGCATT Rv: AATGCAGACGAGCGAACCTT	96	59/59	Vigun11g167000.1
VuUBQ3	Ubiquitin 3 (UBQ3)	Fw: TCTTGTCTTGCGACTCCGTG Rv: TCGTGTCTGAACTCTCGACC	90	60/60	Vigun03g105700.2
VuL23a3*	60S ribosomal protein L23a (L23a3)	Fw: CAGGGCATCATAGTCAGGT Rv: AGGCTTAACATCGAGTAGG	125	57.5/ 57.5	XM_014641511.1

Source: Elaborated by author

* Retrieved from Martins et al., (2020)(MARTINS et al., 2020)

Results

Identification and evolutionary relationships among plant Glycolate oxidase (GOX) orthologs

Information regarding gene identity (ID), gene description (when available on databases), gene species, chromosome location, main transcript, CDS length and predicted peptide length of all *GOX* orthologs identified in 14 plant species (*Vigna unguiculata, Vigna radiata, Vigna angulares, Glycine max, Lotus japonicus, Medicago truncatula, Phaseolus vulgaris, Oryza sativa, Sorghum bicolor, Zea mays, Arabidopsis thaliana, Solanum tuberosum, Gossypium raimondii* and *Chlamydomonas reinhardtii*) are displayed at Table 1, totaling 66 *GOX* gene accessions. The presence of FMN domain in the predicted amino acid sequences represented one of the main cut-off points for identification of *GOX* orthologs, along with other search criteria used (Table 1)

Table 1 List of *Glycolate oxidase* orthologs from 14 plant species distributed among 6 plant families. Gen

e information was obtained from the following databases: NCBI, Phytozome and Vigna Genome Server (VigGS)
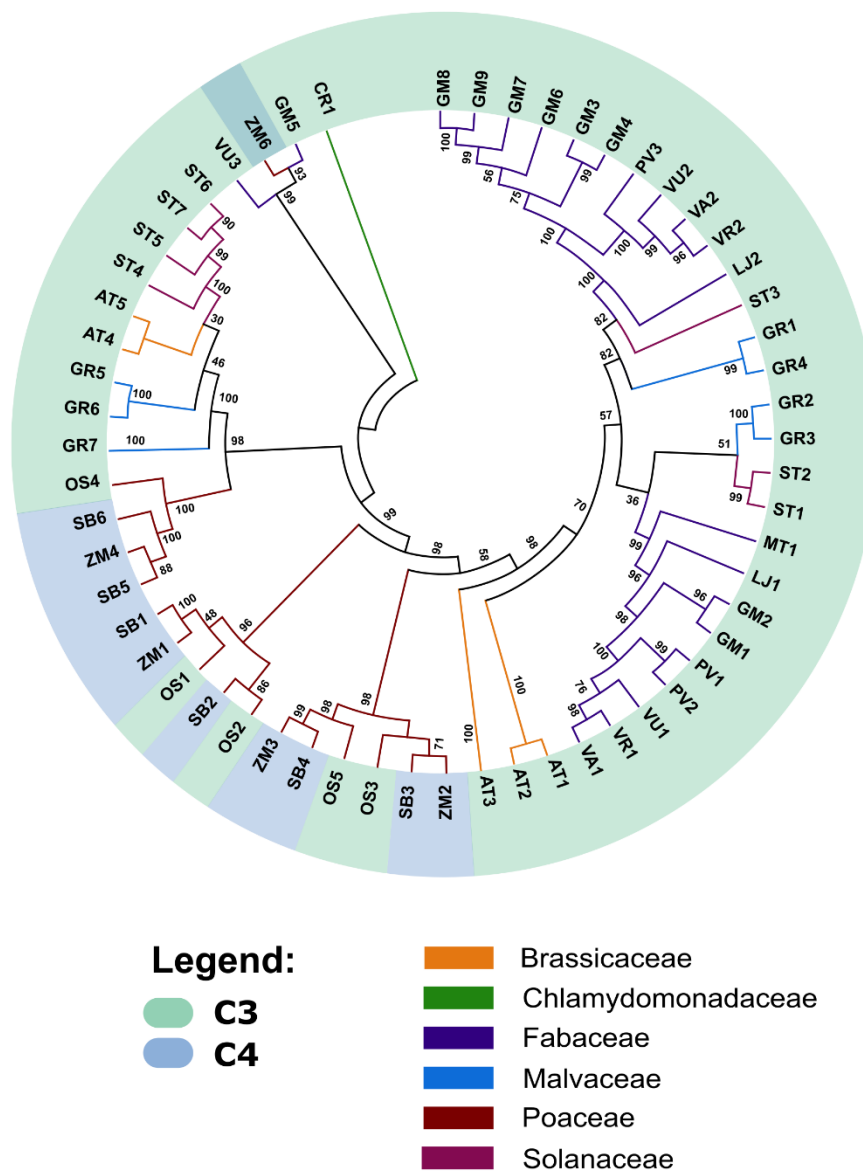
Especies	Gene	Gene ID	Description	Chromosome Location	Transcript sequence ID	CDS Length(bp)	Protein Length(aa)
Arabidopsis thaliana	AT1	AT3G14420	GOX1	Chr03: 4821522..4824178	AT3G14420.2	1104	368
	AT2	AT3G14415	GOX2	Chr03: 4818297..4821006	AT3G14415.1	1104	368
	AT3	AT4G18360	GOX3	Chr04: 10145918..10148682	AT4G18360.1	1107	369
	AT4	AT3G14130	HAOX1	Chr03: 4685645..4688309	AT3G14130.1	1092	364
	AT5	AT3G14150	HAOX2	Chr03: 4690449..4693001	AT3G14150.1	1092	364
Chlamydomonas reinhardtii	CR1	Cre03.g171300		Chr03: 3908304..3911636	Cre03.g171300.t1.2	1155	385
Glycine max	GM1	Glyma.08G175800	GLO1-like	Chr08: 14037038..14042017	Glyma.08G175800.3	1116	372
	GM2	Glyma.15G250800		Chr15: 47892520..47898180	Glyma.15G250800.1	1116	372
	GM3	Glyma.08G096800		Chr08: 7418970..7424007	Glyma.08G096800.1	1113	371
	GM4	Glyma.08G097300	GLO1-like	Chr08: 7438009..7444035	Glyma.08G097300.1	1101	367
	GM5	Glyma.08G096600		Chr08: 7404774..7409459	Glyma.08G096600.1	1107	369
	GM6	Glyma.05G141400	GLO5-like	Chr05: 33475993..33481973	Glyma.05G141400.3	1107	369
	GM7	Glyma.05G141300		Chr05: 33467152..33473778	Glyma.05G141300.10	885	295
	GM8	Glyma.08G096700	GLO1	Chr08: 7412769..7417800	Glyma.08G096700.1	912	304
	GM9	Glyma.08G097100	GLO1	Chr08: 7432417..7434453	Glyma.08G097100.1	618	206
	GM10	Glyma.12G125600		Chr12: 13776445..13776852	Glyma.12G125600.1	225	85
Gossypium raimondii	GR1	Gorai.009G154200		Chr09: 11786500..11790172	Gorai.009G154200.5	1104	368
	GR2	Gorai.002G037900	GLO1-like	Chr02: 2992581..2996498	Gorai.002G037900.1	1110	370
	GR3	Gorai.002G038000	GLO1-like	Chr02: 3012783..3016729	Gorai.002G038000.7	1110	370
	GR4	Gorai.006G092300		Chr06: 32925976..32931171	Gorai.006G092300.6	1107	369
	GR5	Gorai.011G062200	GLO4-like	Chr11: 5142108..5146554	Gorai.011G062200.2	1101	367
	GR6	Gorai.011G062400	GLO4-like	Chr11: 5150947..5154046	Gorai.011G062400.6	1098	366
	GR7	Gorai.010G174000	GLO4-like	Chr10: 50713340..50716489	Gorai.010G174000.7	879	293
	GR8	Gorai.010G173900	GLO4-like	Chr10: 50658236..50659145	Gorai.010G173900.1	564	188

Species							
	GR9	Gorai.010G174100		Chr10: 50733219..50733991	Gorai.010G174100.1	354	118
	GR10	Gorai.003G061400	GLO1-like	Chr03: 10989574..10991024	Gorai.003G061400.1	318	106
Lotus japonicus	LJ1	Lj3g0019836		Chr03: 76246792..76252075	Lj3g0019836.1	1116	372
	LJ2	Lj4g0019266		Chr04: 9712297..9718486	Lj4g0019266.2	1110	370
Medicago truncatula	MT1	Medtr2g070420		Chr02: 29641870..29646961	Medtr2g070420.1	1119	373
	MT2	Medtr1g072440		Chr01: 32146602..32146984	Medtr1g072440.1	267	89
	MT3	Medtr3g076640		Chr03: 34447455..34448279	Medtr3g076640.1	333	111
Oryza sativa	OS1	LOC_Os07g05820	GLO5-like	Chr07: 2797690..2801364	LOC_Os07g05820.2	1110	370
	OS2	LOC_Os03g57220	GLO1-like	Chr03: 32628527..32632535	LOC_Os03g57220.2	1110	370
	OS3	LOC_Os04g53210	GLO3-like	Chr04: 31688716..31692592	LOC_Os04g53210.1	1104	368
	OS4	LOC_Os07g42440	GLO4-like	Chr07: 25408331..25413871	LOC_Os07g42440.1	1101	367
	OS5	LOC_Os04g53214	GLO2-like	Chr04: 31693115..31696647	LOC_Os04g53214.2	831	277
	OS6	LOC_Os08g09860		Chr08: 5693125..5695423	LOC_Os08g09860.1	645	215
Phaseolus vulgaris	PV1	Phvul.005G051800		Chr05: 6669490..6674324	Phvul.005G051800.1	1116	372
	PV2	Phvul.005G052101		Chr05: 6729269..6747460	Phvul.005G052101.1	1116	372
	PV3	Phvul.002G180300		Chr02: 33972895..33980902	Phvul.002G180300.1	1107	369
Sorghum bicolor	SB1	Sobic.002G036000		Chr02: 3388484..3392015	Sobic.002G036000.3	1110	370
	SB2	Sobic.001G065600	GLO1	Chr01: 4988889..4994406	Sobic.001G065600.1	1107	369
	SB3	Sobic.006G220500	GLO3	Chr06: 56741900..56746454	Sobic.006G220500.2	1104	368
	SB4	Sobic.006G220600	GLO2	Chr06: 56748069..56751113	Sobic.006G220600.3	1104	368
	SB5	Sobic.002G374650	GLO4	Chr02: 73247928..73253132	Sobic.002G374650.2	1104	368
	SB6	Sobic.002G374700		Chr02: 73253683..73259983	Sobic.002G374700.1	1107	369
Solanum tuberosum	ST1	PGSC0003DMG400027654		Chr07: 51748835..51752517	S.tub_PGSC0003DMT400071115	1116	372
	ST2	PGSC0003DMG400021248		Chr10: 2622996..2627053	S.tub_PGSC0003DMT400054731	1116	372
	ST3	PGSC0003DMG400007514		Chr08: 14435561..14440383	S.tub_PGSC0003DMT400019434	1107	369
	ST4	PGSC0003DMG403022901		Chr03: 61422525..61426628	S.tub_PGSC0003DMT400058959	1095	365
	ST5	PGSC0003DMG401022901		Chr03: 61409112..61412786	S.tub_PGSC0003DMT400058960	1095	365

	ST6	PGSC0003DMG400022890		Chr03: 61413622..61416259	S.tub_PGSC0003DMT400058920	1092	364
	ST7	PGSC0003DMG402022901		Chr03: 61416805..61421122	S.tub_PGSC0003DMT400058950	1134	378
Vigna angularis	VA1	Vigan07G209000		Chr07: 28955992..28961458	Vigan.07G209000.01	1116	372
	VA2	Vigan01G193100		Chr01: 23724224..23733494	Vigan.01G193100.01	1107	369
Vigna radiata	VR1	LOC106762192	GLO1	Chr05: 29325844..29331164	XP_014501443.1	1116	372
	VR2	LOC106769099	GLO1	Chr07: 21822060..21830187	XP_014510058.1	1107	369
Vigna unguiculata	VU1	Vigun01g052400	GLO1	Chr01: 8835016..8840764	Vigun01g052400.1	1119	373
	VU2	Vigun03g162000	GLO1-like	Chr03: 18056093..18065564	Vigun03g162000.5	1107	369
	VU3	VigunL052501		contig_354: 87862..89020	VigunL052501.1	1158	386
Zea Mays	ZM1	Zm00001d018810	GLO1	Chr07: 5842586..5845546	Z.mays_Zm00001d018810_T002	1257	419
	ZM2	Zm00001d002261		Chr02: 9094316..9098515	Z.mays_Zm00001d002261_T003	1200	400
	ZM3	Zm00001d002260		Chr02: 9089886..9093334	Z.mays_Zm00001d002260_T013	1107	369
	ZM4	Zm00001d006868		Chr02: 218291717..218297633	Z.mays_Zm00001d006868_T006	1164	388
	ZM5	Zm00001d024312		Chr10: 63550608..63551472	Z.mays_Zm00001d024312_T001	360	120
	ZM6	Zm00001d023759		Chr10: 19056768..19063074	Z.mays_Zm00001d023759_T001	783	261

Source: Elaborated by author

The evolutionary relationships of *GOX* orthologs in plants were reconstructed using 58 sequences distributed among 14 species from 6 plant families (Fig. 2). From the species analyzed, 12 are from plants presenting C3 photosynthetic metabolism and 2 from the C4 type. Two very distinct clusters can be evidenced, one of these being formed by only 3 sequences (*Glycine max* and *Vigna unguiculata,* C3; *Zea mays,* C4), while the other is formed by 54 sequences distributed in different subclusters (Fig. 2). Interestingly, two major subclusters can be verified inside the biggest cluster of the phylogram, *both showing almost the same species* (exception is *M. truncatula,* present in just one subcluster) and being predominantly composed of *Fabaceae* members, indicating evolutionary divergence between *GOX* genes inside species. Species of *Poaceae* family are grouped in three small clades that include *O. sativa, Z. mays* and *S. bicolor*. Although sequences from C3 plants are mostly grouped in the same subclusters, as well as sequences from C4 plants, *O. sativa* (C3) presents sequences inserted in subclusters majorly composed of C4 plants, and the sequence ZM6 of *Z. mays* (C4) is grouped in a small cluster containing *V. unguiculata* and G. *max* sequences (Fig. 2). Such exceptions indicate a significant structural and functional divergence degree among certain C3 and C4 *GOX* orthologs.
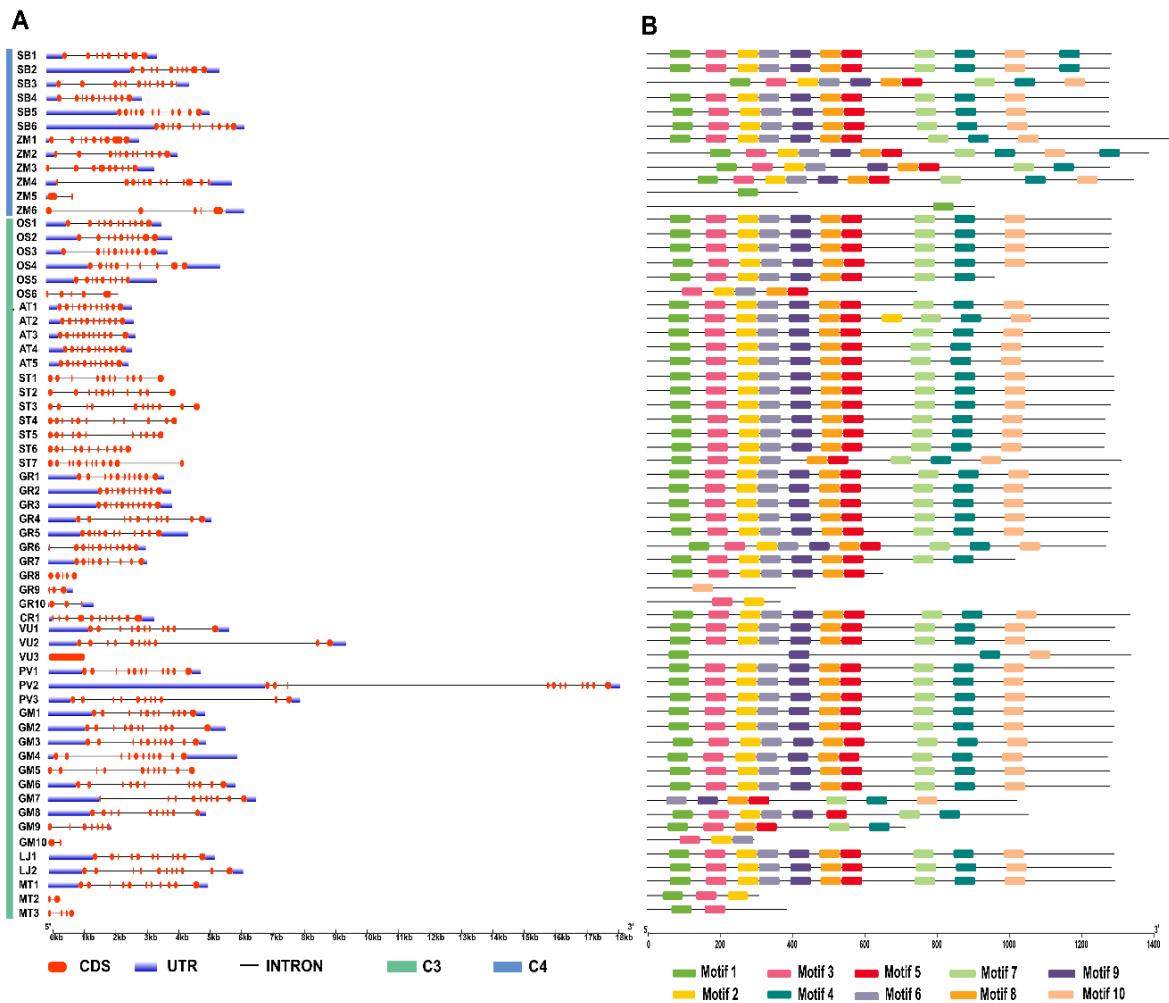
Source: Elaborated by author

Fig. 2 Phylogenetic relationships among *Glycolate oxidase* orthologs in plants, using the maximum parsimony method. Each sequence is identified by a code, described at Table 1. Numbers in nodes represent bootstrap values, obtained from 1000 replications. C3 and C4 plants are highlighted in the tree by the colors green and blue, respectively. The different plant families are evidenced by different colors in branches

Gene structure and conserved motifs organization of plant GOX orthologs

The gene structure and conserved motif organization of *GOX* orthologs in plants was analysed using 62 sequences from 12 species (Fig. 3), where 50 sequences were derived from 10 C3 plant species, and 12 sequences were derived from 2 C4 plant species. We observed the gene structures of the two C4 species showed variation from 9 to 12 exons. On the other hand, most C3 species presented 11 exons in their structure. However, *O. sativa* showed a variation from 6 to 11 exons among its representatives, and *G. max*, *G. raimondii*, *S. tuberosum* and *M. trucatula* presented a high variation in number of exons among their sequences (Fig. 3a), which corroborates with the evolutionary divergence among sequences from same species verified in Figure 2. All sequences, including the smallest ones, presented íntrons.

Source: Elaborated by author

Fig. 3 Gene structure and conserved motifs in plant *GOX* orthologs. A) Structure of *GOX* genes in the 12 plant species analyzed. In blue: 5' and 3' UTR; in red: exons. Introns are represented by black lines. C3 sequences are indicated by a green bar, and C4 sequences are indicated by a blue bar. B) Composition of conserved motifs in predicted amino acid sequences from plant GOX proteins. Motifs are distributed in 10 different colors, indicated below the scheme. Bars below schemes indicate number of residues

We identified 10 conserved motifs distributed among the predicted protein sequences. In C4 species, motifs 2, 3, 4, 5, 6, 7, 8 and 9 (Supplementary Table 2) were shown to be conserved in 10 sequences. Motif 1 was present in all 12 sequences of C4 species, while motif 10 was verified in 9 sequences. Interestingly, motif 4 is duplicated in sequences of C4 organisms (SB1, SB2 and ZM2). In C3 species, the 10 conserved motifs are present in the vast majority of the sequences, with fewer motifs only among shorter sequences. Despite that, the AT2 sequence presents motif 2 duplicated (Fig. 3b). The higher variation in conserved motif distribution in GOX orthologs sequences of C4 plants corroborates with the higher evolutionary and structural divergence verified for such sequences in Figure 2 and Figure 3a.
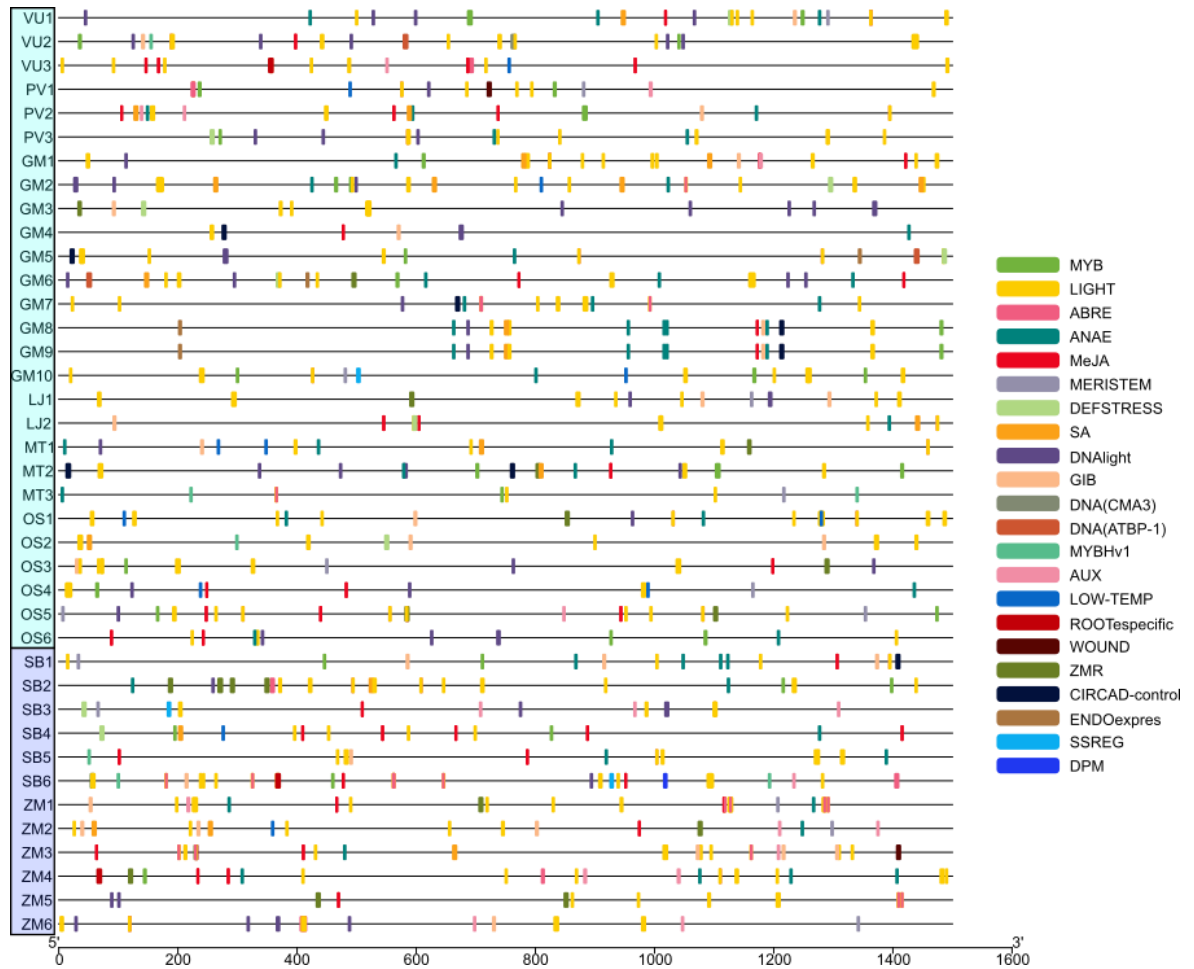
Supplementary Table 2. List of conserved motifs in predicted amino acid sequences of GOX proteins

Motif	Amino Acid Sequence	E-value
Motif 1	CCAAAGATGGTBTATGACTACTAYGCWTCWGGDGCAGAGGAYCAGTGGAC	$4.1 \, e^{-823}$
Motif 2	CCHATCATGATTGCYCCAACWGCYATGCAGAARATGGCTCAYCCTGAAGG	$1.4 \, e^{-827}$
Motif 3	GTTYCGGCCDCGNATTCTTRTTGATGTRAGCAAGATAGACWTGWCHACAA	$3.7 \, e^{-767}$
Motifi4	GGHGCDGCTGGRATCATYGTSTCCAAYCAYGGAGCTCGSCARCTDGAYTA	$8.6 \, e^{-819}$
Motif 5	ACTGTKGAYACTCCAAKRCTTGGYCGYAGGGARGCTGAYATCAARAACAG	$1.1 \, e^{-798}$
Motif 6	GARTWTGCHACWGCHAGAGCWGCAKCWGCWGCTGGMACWATMATGACA`	$2.2 \, e^{-712}$
Motif 7	TGGAAGGATRTSAARTGGCTNCARWCAATCACCWMVYTGCCAATYCTRGT	$2.2 \, e^{-729}$
Motif 8	GCWCAGCTKGTGAGAAGAGCTGAAAGRGCTGGDTTCAAGGCNATTGCYCT	$1.2 \, e^{-707}$
Motif 9	GAGGAGGTTGCTTCAACHGGMCCTGGCATTCGYTTYTTCCARCTHTATGT	$2.2 \, e^{-683}$
Motif10	GAYGGHGGMGTYCGSCGWGGMACHGAYGTYTTYAARGCAYTRGCDCTBGG	$3.5 \, e^{-723}$

Source: Elaborated by author

Cis-element analysis in plant GOX orthologs promoters

A plethora of elements involved in biotic and abiotic stresses responses as well as hormonal control are shown in *cis*-element analysis using promoter regions of plant *GOX* orthologs from *Fabaceae* and *Poaceae* families: low temperature response (LOW-TEMP), abscisic acid (ABRE), salicylic acid (SA), methyl jasmonate (MeJA), gibberellin (GIB), anaerobic induction (ANAE), Auxin (AUX), stress response (DEFSTRESS) and elements associated to MYB transcription factors (MYB) (Fig. 4) (GENG et al., 2021; LI et al., 2021a; SHARMA; TAGANNA, 2020). The elements LIGHT, DNAlight, ABRE, MeJA, SA, GIB, MERISTEM and ANAE are present in all promoter sequences analysed, where *cis*-elements linked to light response (LIGHT and DNAlight) represent the absolute majority in all plant *GOX* orthologs (Fig. 4). One element that appeared in almost all sequences was zein metabolism regulation (ZMR), which is involved in endosperm development. Other development-related elements were also identified, but in smaller frequencies, such as ENDOexpress, CIRCARD-control, ROOTspecific, SSREG and DPM (ZHANG; YANG; WU, 2015). The presence of such *cis*-elements indicates the *GOX* gene transcription is modulated by both biotic and abiotic stress responses and development stimuli in some plant species. Other elements were also identified in lower frequency (Supplementary Table 3).

Source: Elaborated by author

Fig. 4 *Cis*-elements from promoter region of *GOX* genes in plants of *Fabaceae* and *Poaceae* families. Colored blocks indicate distinct *cis*-elements present in the promoter region. C3 and C4 in green and blue boxes respectively

Supplementary Table 3 List of all *cis*-elements found in the promoter region of *GOX* orthologs in plants, showing *cis*-element identification, class, and description

ID	Cis-element class	Description
ABRE	ABRE	Cis-acting element involved in the abscisic acid responsiveness
ANAE	ARE	Cis-acting regulatory element essential for the anaerobic induction
AUX	AuxRR-core	Cis-acting regulatory element involved in auxin responsiveness
	TGA-box	Part of an auxin-responsive element
	TGA-element	Auxin-responsive element
CIRCAD-control	circadian	Cis-acting regulatory element involved in circadian control
DEFSTRESS	TC-rich repeats	Cis-acting element involved in defense and stress responsiveness
DNA(ATBP-1)	AT-rich element	Binding site of AT-rich DNA binding protein (ATBP-1)
DNA(CMA3)	3-AF3 binding site	Part of a conserved DNA module array (CMA3)
DNAlight	ATC-motif	Part of a conserved DNA module involved in light responsiveness
	ATCT-motif	Part of a conserved DNA module involved in light responsiveness
	Box 4	Part of a conserved DNA module involved in light responsiveness
DPM	HD-Zip 1	Element involved in differentiation of the palisade mesophyll cells
ENDOexpres	GCN4_motif	Cis-regulatory element involved in endosperm expression
GIB	GARE-motif	Gibberellin-responsive element
	P-box	Gibberellin-responsive element
	TATC-box	Cis-acting element involved in gibberellin-responsiveness
LIGHT	3-AF1 binding site	Light responsive element
	4cl-CMA1b	Light responsive element
	ACE	Cis-acting element involved in light responsiveness
	AE-box	Part of a module for light response
	AT1-motif	Part of a light responsive module
	Box II	Part of a light responsive module
	chs-CMA1a	Part of a light responsive element
	chs-CMA2a	Part of a light responsive element
	GA-motif	Part of a light responsive element
	Gap-box	Part of a light responsive element
	GATA-motif	Part of a light responsive element
	GATT-motif	Part of a light responsive element
	G-box	Cis-acting regulatory element involved in light responsiveness
	GT1-motif	Light responsive element
	I-box	Part of a light responsive element
	LAMP-element	Part of a light responsive element
	Sp1	Light responsive element
	TCCC-motif	Part of a light responsive element
	TCT-motif	part of a light responsive element

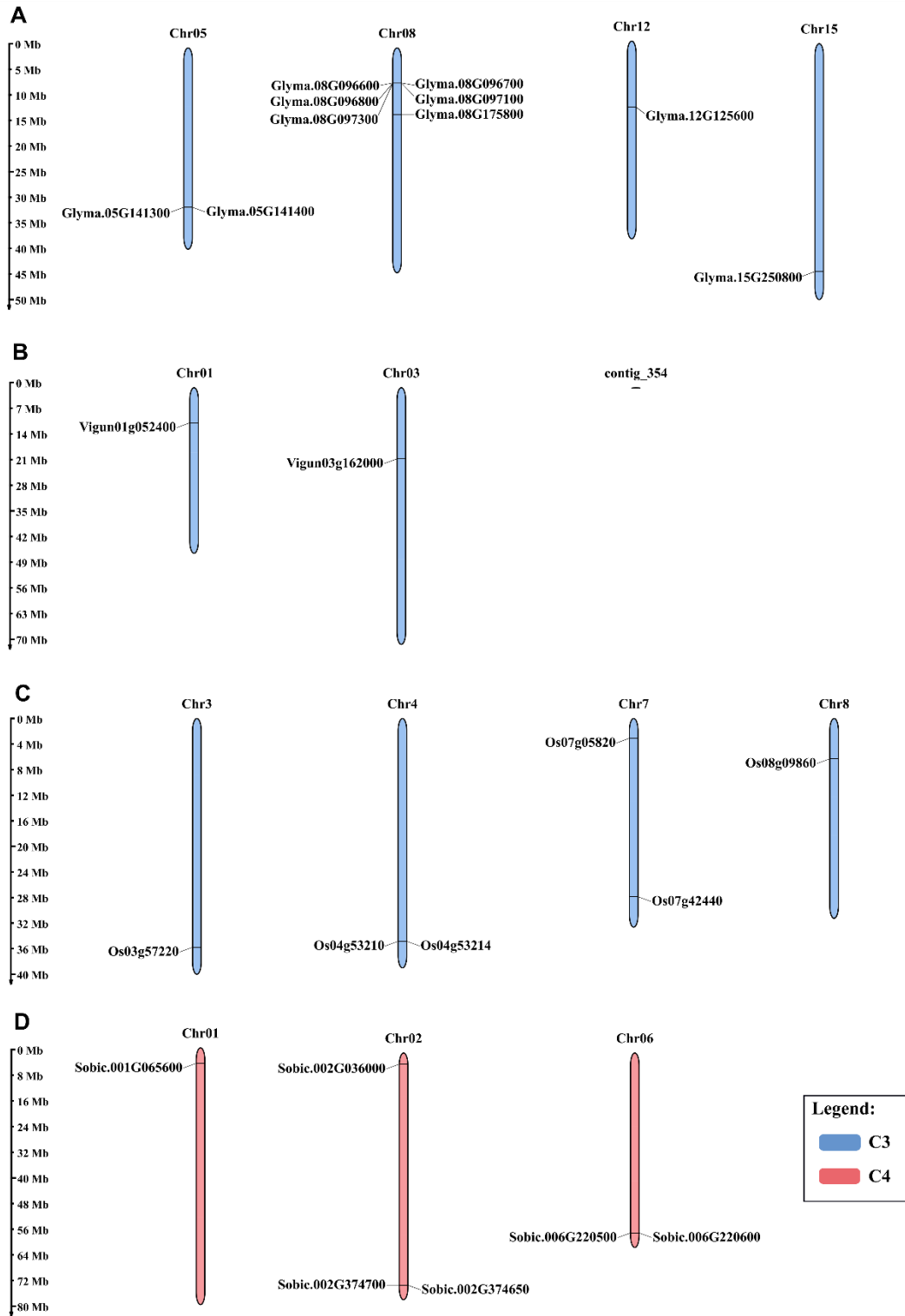|		
LOW-TEMP	LTR	Cis-acting element involved in low-temperature responsiveness
MeJA	TGACG-motif	Cis-acting regulatory element involved in the MeJA-responsiveness
	CGTCA-motif	Cis-acting regulatory element involved in the MeJA-responsiveness
MERISTEM	CAT-box	Cis-acting regulatory element related to meristem expression
MYB	MBS	MYB binding site involved in drought-inducibility
	MBSI	MYB binding site involved in flavonoid biosynthetic genes regulation
	MRE	MYB binding site involved in light responsiveness
MYBHv1	CCAAT-box	MYBHv1 binding site
ROOTespecific	motif I	Cis-acting regulatory element root specific
SA	TCA-element	Cis-acting element involved in salicylic acid responsiveness
SSREG	RY-element	Cis-acting regulatory element involved in seed-specific regulation
WOUND	WUN-motif	Wound-responsive element
ZMR	O2-site	Cis-acting regulatory element involved in zein metabolism regulation

Source: Elaborated by author

Chromosomal localization and synteny between C3 and C4 plant species of agronomic relevance

For chromosomal location and synteny analysis, we selected representatives of C3 and C4 species from *Fabaceae* and *Poaceae* plant families presenting distinct phylogenetic relationships (Fig. 5 and Fig. 6). *GOX* orthologs from *G. max* (C3)*, V. unguiculata* (C3)*, O.sativa* (C3) and *S. bicolor* (C4) were chosen, where there is phylogenetic proximity between *G. max* and *V. unguiculata* sequences, as well as between *O.sativa* and *S. bicolor* sequences (Fig. 2).

Among the representatives of C3 plants, in *Fabaceae* family, the *GOX* genes of *G. max* are distributed on chromosomes 5, 8, 12 and 15, and those of *V. unguiculata* are located on chromosomes 1 e 3, and on contig 354 (according to the latest version of the *Vigna unguiculata* genome). In *Poaceae* family, the *O.sativa* genes are on chromosomes 3, 4, 7 e 8. In the C4 plant representative, *S. bicolor* genes are present on chromosomes 1, 2 and 6 (Fig. 5).
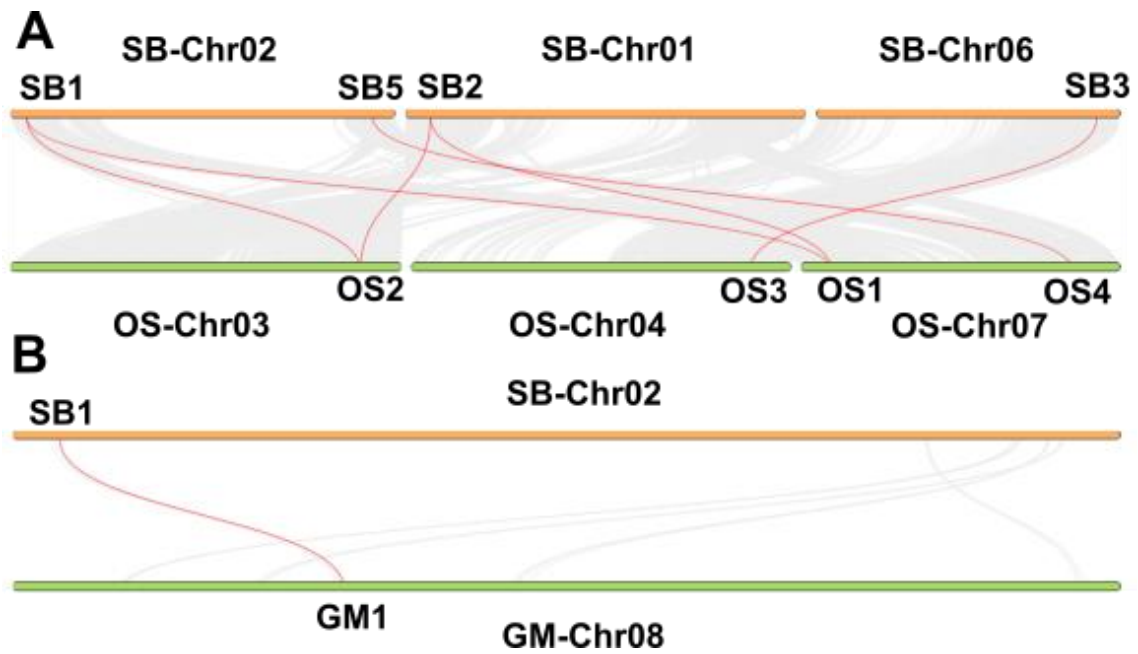
In the synteny analysis, colinearities were found between *S. bicolor* and *G. max* species and between *S. bicolor* and *O. sativa*. However, no colinearity was observed between *S. bicolor* and *V. unguiculata* (Fig. 6), indicating high evolutionary distance between *GOX* genes in such plant species, along with suggested functional divergence between certain C3 and C4 *GOX* orthologs.

O. sativa and *S. bicolor* showed the larger number of colinearities verified (Fig. 6). Both sequences SB1 and SB2 presented colinearity with sequences OS1 and OS2, SB3 presented colinearity with sequence OS3, and SB5 presented colinearity with sequence OS4. On the other hand, *G. max* and *S. bicolor* presented synteny only between sequence GM1 and SB1. This colinearity pattern among the studied species may indicate similar function of their *GOX* orthologs.

Source: Elaborated by author

Fig. 5 Distribution of *GOX* genes in the different chromosomes of *G. max* (A), *V. unguiculata* (B), *O. sativa* (C) and *S.bicolor* (D). Blue chromosomes depict C3 plant species, and red chromosomes depict C4 plant species

Source: Elaborated by author

Fig. 6 Synteny analysis of *GOX* genes between C3 and C4 species. Green and orange bars indicate distinct chromosomes, red lines indicate interspecific colinearities. A) Colinearities between *S. bicolor* (C4) and *G. max* (C3). B) Colinearities between *S. bicolor* and *O. sativa*
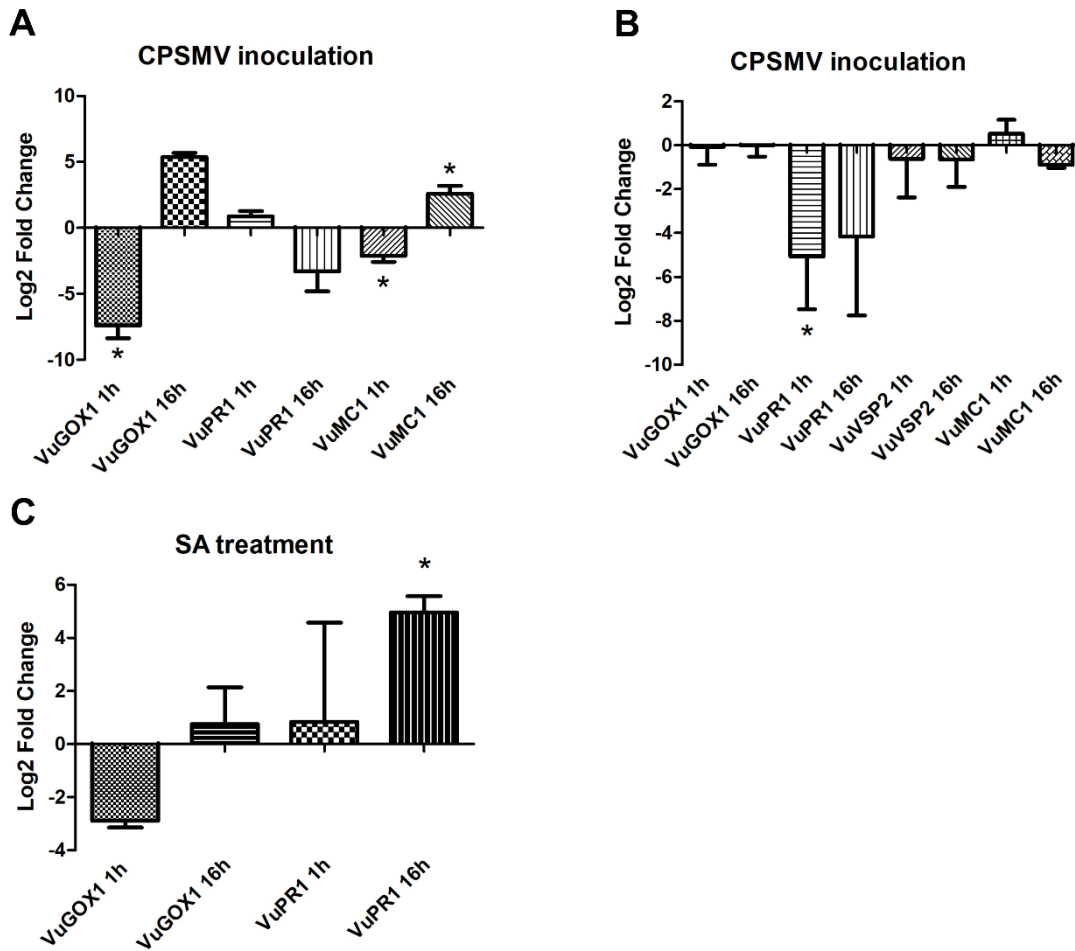
Expression profile of the GOX1 ortholog in Vigna unguiculata

In view of recent studies reporting differential expression of GOX proteins during biotic stress in *V. unguiculata* (VARELA et al., 2017, 2019), the expression profile of the *V. unguiculata GOX* ortholog with the highest phylogenetic proximity with *AtGOX1* (*VuGOX1),* which is involved in defense responses in *Arabidopsis* (DELLERO et al., 2016a; NOCTOR; MHAMDI, 2017; XU; YANG; CAI, 2018) , was verified during early times points after biotic stress stimuli (Fig. 7). Along with *VuGOX1,* gene orthologs associated with plant defense and pathogen response (*VuPR1, VuVSP2* and *VuMC1*) had their expression pattern evaluated. Susceptible (Pitiúba) and resistant (BR 14 Mulato) cultivars of *V.unguiculata* plants were used for CPSMV inoculation. Moreover, gene expression in response to Salicylic Acid (SA) was obtained in Pitiúba cultivar.

In BR 14 Mulato plants inoculated with CPSMV, *VuGOX1* was significantly repressed 1 hour after inoculation (hai), as well as *VuMC1*, whose orthologous gene in *Arabidopsis* is involved in environmental programmed cell death (ePCD) (COLL et al., 2010). On the other hand, at 16 hai, only *VuMC1* was significantly induced. *VuPR1* was not significantly expressed in both time points, indicating BR 14 Mulato resistance phenotype is not SA-dependent (Fig. 7).

In Pitiúba plants inoculated with CPSMV, another plant defense gene had its expression evaluated. *VSP2* orthologs in plants are responsive to wound, herbivory and Jasmonate (JA) (TAKI et al., 2005). However, *VuVSP2* was not significantly expressed in the time points analysed. Only *VuPR1* was significantly repressed at 1 hai. This result also indicates JA is not generated during CPSMV infection in susceptible plants (Fig. 7).

In view of the result in BR 14 Mulato plants, the experiment using SA was performed using Pitiúba plants (Fig. 7). *PR1* gene, whose orthologs are classic markers of SA-related response in plants (GONZÁLEZ-BOSCH, 2018), was used as a treatment control. *VuGOX1* did not present significant expression in any time point, whereas *VuPR1* was significantly induced at 16 hai, showing *VuGOX1* is not modulated by SA during plant defense responses (Fig. 7).

Source: Elaborated by author

Fig. 7 Expression analysis of *VuGOX1* and marker genes in susceptible and resistant plants during CPSMV infection and SA treatment. Relative expressions were calculated according to PFAFFL, 2001, and statistical significance was obtained according to YUAN, 2006, using nonparametric Mann-Whitney U tests. *VuF-BOX3* and *VuUBQ3* were used as reference genes for CPSMV inoculation and SA application treatments in Pitiúba plants, whereas *VuL23a3* was used as reference gene for CPSMV inoculation treatment in BR 14 Mulato plants. False CPSMV inoculation (MOCK) and alcoolic solution were used as control treatments for CPSMV inoculation and SA application, respectively. The cDNAs used were obtained from three biological replicates. Means indicated by an asterisk represent significant difference between treated and control samples at 5% significance level. Bars indicate standard error of the means.

Discussion

Genome-wide studies are crucial for understanding of how the evolution of gene families relates to their structure and function. Associated with comparative analyses among several species, one can obtain insights of fundamental differences between gene orthologs in distinct species that can impact species physiology. In the present work, using a genome-wide comparative analysis, the evolutionary and structural aspects of *GOX* orthologs in plant species with C3 and C4 photosynthetic pathways are dissected.

The evolution of the C4 photosynthetic pathway is strongly associated with changes in natural environments of C3 plant species over time (OSBORNE; FRECKLETON, 2009). Previous studies indicate grasses may have arisen from the common ancestor of BOP (*Bambusoideae*, *Oryzoideae* and *Pooideae*) and PACMAD (*Panicoideae*, *Aristidoideae*, *Chloridoideae*, *Micrairoideae*, *Arundinoideae* and *Danthonioideae*). Gradative alleviation in photorespiration process with concomitant improvement of photosynthesis efficiency was the main outcome of the rise of C4 plants (GALLAHER et al.; OSBORNE; FRECKLETON, 2009). With the impact on photorespiration along C4 plants evolution, genes coding molecular players underlying the process also changed significantly, generating relevant structural and functional divergence among distinct plant species. The gene coding the enzyme responsible for glycolate oxidation in peroxisomes, *Glycolate oxidase* (*GOX*), was one that generated most divergent orthologs over time (Fig.1), acting in a variety of molecular functions, such as plant defense responses (LIU; WU; CHEN, 2018; XU; YANG; CAI, 2018).

In our phylogenetic reconstruction using Maximum Parcimony, 3 distinct subclades containing only sequences of species from *Poaceae* family were observed (*O. sativa*, *S. bicolor* and *Z. mays*), suggesting *GOX* gene has remained highly conserved among these species, even with *O. sativa* (C3) presenting a distinct photosynthetic apparatus from the other species (C4) (Fig. 2). Apart of the large subcluster comprising *GOX* sequences from several C3 plant species, formed by *Fabaceae*, *Malvaceae* and *Solanaceae* plant families, all *A. thaliana* sequences formed separated subclusters, where AT1 and AT2 are the main *GOX* orthologs responsible for photorespiration in this species (DELLERO et al., 2016a)Other 2 *A. thaliana* orthologs clustered with *S. tuberosum* members (Fig. 2), corroborating with previous work showing high similarity between GOX2 of tomato (another species from *Solanaceae* family) and GOX2 of *A. thaliana* (AHAMMED et al., 2018). An interesting finding regarding *GOX* orthologs from *Fabaceae* family is the small cluster of 3 sequences [*V.unguiculata* (VU3), *G. max* (GM5) and *Z. mays* (ZM6)], where 2 close C3 organisms group with one evolutionary distant C4 species (Fig. 2), indicating phylogenetic proximity among the 3 *GOX* orthologs inside this cluster may

represent functional conservation, a hypothesis to be further confirmed.

Previous research has shown the intron-exon ratio is an relevant indicator of gene evolution (GORLOVA et al., 2014).The 2 C4 species analyzed exhibited the highest variation of intron-exon ratio among their *GOX* orthologs, indicating fewer gene conservation when compared with C3 sequences (Fig. 3a). In contrast, C3 species showed a smaller variation among their *GOX* orthologs, where 10 introns and 11 exons were observed in most C3 *GOX* sequences (Fig. 3a). In addition to a higher gene conservation among C3 species, the large number of introns suggests frequent occurrence of alternative splicing in C3 *GOX* sequences, which might impact transcript abundance during cell signalling, as indicated during plant stress responses (LALOUM; MARTÍN; DUQUE, 2018).

Despite the FMN domain be highly conserved among plant GOX proteins, as confirmed in our analyses using Hidden Markov chains in Pfam and multiple sequence alignments (Supplementary Figure 2), the number of motifs present considerable variation between C3 and C4 predicted amino acid sequences (Fig. 3b). Larger *GOX* orthologs presenting higher intron-exon ratio exhibited all the 10 motifs detected using MEME, whereas smaller *GOX* orthologs with lower intron-exon ratio lacked some motifs, impacting in FMN domain conservation (Fig. 3b and Supplementary Figure 2). Such findings corroborate with the phylogenetic relationships at Figure 2, where most C3 *GOX* sequences which forms large subclusters code predicted amino acid sequences with all 10 conserved motifs.

Source: Elaborated by author

Supplementary Fig. 2 Multiple alignment of amino acid sequences from C3 and C4 GOX proteins. C3 species are identified in cyan and C4 species are identified in purple. Dashes indicate gaps. Amino acid residues involved in FMN cofactor stabilization [in red, Tyrosine (Y), Serine (S), Threonine (T), Lysine (K), Histidine (H) and Arginine (R)] are highlighted in pink boxes. Bioedit tool was used for multiple alignment visualization

Analyzing synteny between genomes is a powerful resource to detect important genomic processes occurring along evolution, like gene duplication and loss, as well as to verify high similarity regions common to compared genomes (SCHMUTZ et al., 2010). Polyploidy and aneuploidy occur frequently in several plant families, giving the synteny analysis a crucial relevance in detection of genomic processes in plants (SCHMUTZ et al., 2010). Among the 6 *GOX* orthologs present in both *S. bicolor* and *O. sativa* genomes (Fig. 5)*,* synteny analysis showed colinearity between 4 genes from *S. bicolor* and 4 from *O. sativa* (Fig. 6), emphasizing the homology and high similarity among *GOX* orthologs in *Poaceae* family, even with distinct photosynthetic apparatus, corroborating with the phylogenetic proximity between these two species (Fig. 2), as also reported in a recent study (GENG et al., 2021). When comparing *S. bicolor* and *G. max* genomes, only one colinearity is detected, between GM1 (the soybean GOX1 gene) from *G. max* and SB1 (without detailed information about this gene) from *S. Bicolor* (Fig. 5), suggesting a significant loss of genetic conservation in *GOX* genes along plant evolution, as observed in previous studies (SHARMA; TAGANNA, 2020). Interestingly, the synteny analysis between *S. bicolor* and *V. unguiculata* shows no colinearity (Fig. 5), suggesting the high divergence accumulated through long plant evolution eliminated fundamental similarities among *GOX* orthologs structures that may drastically impact gene function. Such hypothesis should be validated comparing molecular and physiological functions of *GOX* genes in *S. bicolor* and *V. unguiculata*.

In *cis*-element analysis using promoter regions of *GOX* orthologs, we observed 8 types of *cis*-elements common to all C3 and C4 plant species, distributed among *Fabaceae* and *Poaceae* plant families (Fig. 4). Many *cis*-elements detected are related to hormonal responses, such as abscisic acid (ABRE), methyl jasmonate (MeJA), salicylic acid (SA) and gibberellin (GIB), most are linked to light response (LIGHT and DNAlight), 1 element is related to meristem expression and regulation (MERISTEM) and another is related to anaerobic induction (ANAE). The *cis*-elements linked to light response (LIGHT and DNAlight) represents the majority in both *Fabaceae* and *Poaceae* orthologs (Fig. 4). Interestingly, in *V.unguiculata GOX* orthologs, following the light response *cis*-elements, the most prominent are abscisic acid, methyl jasmonate and MYB-related *cis*-elements, which are directly linked to plant stress responses (Fig. 4).

Some *cis*-elements were observed in lower frequency in specific *GOX* orthologs (Fig. 4). The root specific element (ROOT SPECIFIC) appears in both C4 species and in *V. unguiculata*; the wound responsive element (WOUND) is present in *Z.mays* and *P.vulgaris*; the element

involved in circadian control (CIRCARD CONTROL) is present in *S. bicolor*, *G. max* and *M. truncatula*; the MYBhv1 binding site (MYBhv1) is present in *S. bicolor*, *M. truncatula*, *V. unguiculata* and *O. sativa*; the element involved in seed-specific regulation (SSreg) was found in *S. bicolor* and *G. max*; the element involved in differentiation of palisade parenchyma cells (DPM) was found only in *S. bicolor*; the element related to meristem expression (ENDO-EXPRESS) is present only in *G. max*; the ATBP-1 protein binding site is present in *G. max* and *V. unguiculata*; and an element called part of a set of conserved DNA modules (CMA3) is present only in *V. unguiculata*. The presence of particular *cis*-elements in some *GOX* orthologs suggests transcription activation under specific developmental and stress signaling pathways.

Recently, a genome-wide study presented the first differential expression analysis of all kinase gene families at early time points during CPSMV infection in *V. unguiculata* plants, using the BR 14 Mulato cultivar (FERREIRA-NETO et al., 2021)For the first time, a *V. unguiculata GOX* ortholog had its expression evaluated in early time points following CPSMV infection and SA treatment (Fig. 7). The analysis of *VuGOX1* expression following CPSMV inoculation shows this gene ortholog is early repressed during incompatible interaction with the resistant cultivar BR 14 Mulato (Fig. 7). No significant differences in *VuGOX1* expression is observed in susceptible cultivar Pitiúba following both CPSMV inoculation and SA treatment (Fig. 7), showing *VuGOX1* is not regulated by SA during plant defenses and does not present modulation by CPMSV in susceptible plants. Previous works showed a significant increase of GOX proteins abundance during incompatible interactions between CPSMV and the resistant *V. unguiculata* cultivar Marataoã, days after inoculation (VARELA et al., 2017, 2019). Our findings suggest *VuGOX1* is repressed in early time points following CPSMV infection in resistant plants due to rapid activation of defense signaling responses that impact negatively on photosynthetic efficiency, as reported previously in other plant species (SOUZA; GARCIA-RUIZ; CARVALHO, 2019; ZANARDO; DE SOUZA; ALVES, 2019)After defense responses remission, *VuGOX1* can be induced following photosynthesis activation, a hypothesis to be experimentally confirmed.

Conclusion

In this first genome-wide comparative study of *Glycolate oxidase* in plant families, we conclude *GOX* gene family presents distinct levels of evolutionary and structural divergence, strongly dependent on plant family. *GOX* orthologs from *Fabaceae* members present the higher conservation among the plant families analyzed, with similar structures among *GOX* orthologs from different species. High divergence is dominant among *Fabaceae* and *Poaceae* members, even when species wih duplicated genomes such as *G. max* are compared, which might strongly impact gene function conservation among orthologs. As observed in other studies, *GOX* orthologs are modulated during incompatible interactions with CPSMV in *Vigna unguiculata*, a diploid member from *Fabaceae* family. The high divergence degree present between *GOX* orthologs from *Poaceae* C4 plants and *Fabaceae* C3 plants may result in *VuGOX1* presenting a distinct function from all *S. bicolor GOX* orthologs. Reverse genetics assays modulating *V. unguiculata* and *S. bicolor GOX* orthologs expression, such as Virus Induced Gene Silencing (VIGS), may prove invaluable to compare and dissect their molecular role in C3 and C4 plants, as well as to evaluate their relevance during plant defense responses. Such studies are fundamental to design new molecular strategies aiming to obtain tolerance to environmental stresses in plants presenting contrasting photosynthetic machineries.

Authors contributions and Acknowledgements

Statements and Declarations

Authors declare don't have competing interests related to the work submitted for publication

References

AHAMMED, G. J. et al. Tomato photorespiratory glycolate-oxidase-derived H_2O_2 production contributes to basal defence against *Pseudomonas syringae*: Photorespiratory GOX in tomato basal defence. **Plant, Cell & Environment**, v. 41, n. 5, p. 1126–1138, maio 2018.

BAILEY, T. L.; ELKAN, C. Fitting a Mixture Model By Expectation Maximization To Discover Motifs In Biopolymer. p. 9, [s.d.].

BLUM, M. et al. The InterPro protein families and domains database: 20 years on. **Nucleic Acids Research**, v. 49, n. D1, p. D344–D354, 8 jan. 2021.

CARVALHO, M. et al. Evaluating stress responses in cowpea under drought stress. **Journal of Plant Physiology**, v. 241, p. 153001, out. 2019.

CHAO, J. et al. MG2C: a user-friendly online tool for drawing genetic maps. **Molecular Horticulture**, v. 1, n. 1, p. 16, dez. 2021.

CHEN, C. et al. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. **Molecular Plant**, v. 13, n. 8, p. 1194–1202, ago. 2020.

COLL, N. S. et al. *Arabidopsis* Type I Metacaspases Control Cell Death. **Science**, v. 330, n. 6009, p. 1393–1397, 3 dez. 2010.

CUI, L.-L. et al. Overexpression of Glycolate Oxidase Confers Improved Photosynthesis under High Light and High Temperature in Rice. **Frontiers in Plant Science**, v. 7, 4 ago. 2016.

DELLERO, Y. et al. Decreased glycolate oxidase activity leads to altered carbon allocation and leaf senescence after a transfer from high CO_2 to ambient air in *Arabidopsis thaliana*. **Journal of Experimental Botany**, v. 67, n. 10, p. 3149–3163, maio 2016a.

DELLERO, Y. et al. Photorespiratory glycolate–glyoxylate metabolism. **Journal of Experimental Botany**, v. 67, n. 10, p. 3041–3052, maio 2016b.

FERREIRA-NETO, J. R. C. et al. The Cowpea Kinome: Genomic and Transcriptomic Analysis Under Biotic and Abiotic Stresses. **Frontiers in Plant Science**, v. 12, p. 667013, 14 jun. 2021.

FOYER, C. H. et al. Photorespiratory Metabolism: Genes, Mutants, Energetics, and Redox Signaling. **Annual Review of Plant Biology**, v. 60, n. 1, p. 455–484, 1 jun. 2009.

FURBANK, R. T. Evolution of the C4 photosynthetic mechanism: are there really three C4 acid decarboxylation types? **Journal of Experimental Botany**, v. 62, n. 9, p. 3103–3108, 1 maio 2011.

GENG, L. et al. **Genome-Wide Analysis of NAC Proteins and Characterization of the ATAF Member RcNAC091 in Rose (Rosa Chinensis)**. [s.l.] In Review, 24 ago. 2021. Disponível em: <https://www.researchsquare.com/article/rs-803598/v1>. Acesso em: 24 jan. 2022.

GONZÁLEZ-BOSCH, C. Priming plant resistance by activation of redox-sensitive genes. **Free Radical Biology and Medicine**, v. 122, p. 171–180, jul. 2018.

GOODSTEIN, D. M. et al. Phytozome: a comparative platform for green plant genomics. **Nucleic Acids Research**, v. 40, n. D1, p. D1178–D1186, jan. 2012.

GORLOVA, O. et al. Genes with a large intronic burden show greater evolutionary conservation on the protein level. **BMC Evolutionary Biology**, v. 14, n. 1, p. 50, 2014.

HU, B. et al. GSDS 2.0: an upgraded gene feature visualization server. **Bioinformatics**, v. 31, n. 8, p. 1296–1297, 15 abr. 2015.

KUMAR, S. et al. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. **Molecular Biology and Evolution**, v. 35, n. 6, p. 1547–1549, 1 jun. 2018.

LALOUM, T.; MARTÍN, G.; DUQUE, P. Alternative Splicing Control of Abiotic Stress Responses. **Trends in Plant Science**, v. 23, n. 2, p. 140–150, fev. 2018.

LANGDALE, J. A. C $_4$ Cycles: Past, Present, and Future Research on C $_4$ Photosynthesis. **The Plant Cell**, v. 23, n. 11, p. 3879–3892, nov. 2011.

LESCOT, M. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. **Nucleic Acids Research**, v. 30, n. 1, p. 325–327, 1 jan. 2002.

LI, L. et al. Genome-Wide Analysis and Expression Profiling of the Phospholipase D Gene Family in Solanum tuberosum. **Biology**, v. 10, n. 8, p. 741, 2 ago. 2021.

LIU, Y.; WU, W.; CHEN, Z. Structures of glycolate oxidase from Nicotiana benthamiana reveal a conserved pH sensor affecting the binding of FMN. **Biochemical and Biophysical Research Communications**, v. 503, n. 4, p. 3050–3056, set. 2018

MARTINS, T. F. et al. Identification, characterization, and expression analysis of cowpea (Vigna unguiculata [L.] Walp.) miRNAs in response to cowpea severe mosaic virus (CPSMV) challenge. **Plant Cell Reports**, v. 39, n. 8, p. 1061–1078, ago. 2020.

MISTRY, J. et al. Pfam: The protein families database in 2021. **Nucleic Acids Research**, v. 49, n. D1, p. D412–D419, 8 jan. 2021.

NOCTOR, G.; MHAMDI, A. Climate Change, CO $_2$, and Defense: The Metabolic, Redox, and Signaling Perspectives. **Trends in Plant Science**, v. 22, n. 10, p. 857–870, out. 2017.

OSBORNE, C. P.; FRECKLETON, R. P. Ecological selection pressures for C $_4$ photosynthesis in the grasses. **Proceedings of the Royal Society B: Biological Sciences**, v. 276, n. 1663, p. 1753–1760, 22 maio 2009.

PAIVA, A. L. S. et al. Label-free Proteomic Reveals that Cowpea Severe Mosaic Virus Transiently Suppresses the Host Leaf Protein Accumulation During the Compatible Interaction with Cowpea (*Vigna unguiculata* [L.] Walp.). **Journal of Proteome Research**, v. 15, n. 12, p. 4208–4220, 2 dez. 2016.

PFAFFL, M. W. A new mathematical model for relative quantification in real-time RT–PCR. **Nucleic Acids Research**, v. 29, n. 9, p. e45–e45, 5 jan. 2001.

PFAFFL, M. W. et al. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper – Excel-based tool using pair-wise correlations. **Biotechnology Letters**, v. 26, n. 6, p. 509–515, mar. 2004.

POPOV, V. N. et al. Glycolate oxidase isoforms are distributed between the bundle sheath and mesophyll tissues of maize leaves. **Journal of Plant Physiology**, v. 160, n. 8, p. 851–857, jan. 2003.

RAZMI, N. et al. Salicylic acid induced changes on antioxidant capacity, pigments and grain yield of soybean genotypes in water deficit condition. **Journal of Plant Interactions**, v. 12, n. 1, p. 457–464, 1 jan. 2017.

SAKAI, H. et al. The *Vigna* Genome Server, ' *Vig* GS': A Genomic Knowledge Base of the Genus *Vigna* Based on High-Quality, Annotated Genome Sequence of the Azuki Bean, *Vigna angularis* (Willd.) Ohwi & Ohashi. **Plant and Cell Physiology**, v. 57, n. 1, p. e2–e2, jan. 2016.

SCHMUTZ, J. et al. Genome sequence of the palaeopolyploid soybean. **Nature**, v. 463, n. 7278, p. 178–183, jan. 2010.

SHARMA, B.; TAGANNA, J. Genome-wide analysis of the U-box E3 ubiquitin ligase enzyme gene family in tomato. **Scientific Reports**, v. 10, n. 1, p. 9581, dez. 2020.

SOUZA, P. F. N.; GARCIA-RUIZ, H.; CARVALHO, F. E. L. What proteomics can reveal about plant–virus interactions? Photosynthesis-related proteins on the spotlight. **Theoretical and experimental plant physiology**, v. 31, n. 1, p. 227–248, mar. 2019.

TAKI, N. et al. 12-Oxo-Phytodienoic Acid Triggers Expression of a Distinct Set of Genes and Plays a Role in Wound-Induced Gene Expression in Arabidopsis. **Plant Physiology**, v. 139, n. 3, p. 1268–1283, 1 nov. 2005.

UENO, O.; YOSHIMURA, Y.; SENTOKU, N. Variation in the Activity of Some Enzymes of Photorespiratory Metabolism in C4 Grasses. **Annals of Botany**, v. 96, n. 5, p. 863–869, 1 out. 2005.

VARELA, A. L. N. et al. Gel-free/label-free proteomic, photosynthetic, and biochemical analysis of cowpea (Vigna unguiculata [L.] Walp.) resistance against Cowpea severe mosaic virus (CPSMV). **Journal of Proteomics**, v. 163, p. 76–91, jun. 2017.

VARELA, A. L. N. et al. A resistant cowpea (Vigna unguiculata [L.] Walp.) genotype became susceptible to cowpea severe mosaic virus (CPSMV) after exposure to salt stress. **Journal of Proteomics**, v. 194, p. 200–217, mar. 2019.

WANG, Y. et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. **Nucleic Acids Research**, v. 40, n. 7, p. e49–e49, 1 abr. 2012.

XU, Y.-P.; YANG, J.; CAI, X.-Z. Glycolate oxidase gene family in Nicotiana benthamiana: genome-wide identification and functional analyses in disease resistance. **Scientific Reports**, v. 8, n. 1, p. 8615, dez. 2018.

YUAN, J. S. et al. Statistical analysis of real-time PCR data. **BMC Bioinformatics**, v. 7, n. 1, p. 85, dez. 2006.

ZANARDO, L. G.; DE SOUZA, G. B.; ALVES, M. S. Transcriptomics of plant–virus interactions: a review. **Theoretical and Experimental Plant Physiology**, v. 31, n. 1, p. 103–125, mar. 2019.

ZELITCH, I. et al. High Glycolate Oxidase Activity Is Required for Survival of Maize in Normal Air. **Plant Physiology**, v. 149, n. 1, p. 195–204, 6 jan. 2009.

ZHANG, Z. et al. Catalytic and functional aspects of different isozymes of glycolate oxidase in rice. **BMC Plant Biology**, v. 17, n. 1, p. 135, dez. 2017.

ZHANG, Z.; YANG, J.; WU, Y. Transcriptional Regulation of Zein Gene Expression in Maize through the Additive and Synergistic Action of opaque2, Prolamine-Box Binding Factor, and O2 Heterodimerizing Proteins. **The Plant Cell**, v. 27, n. 4, p. 1162–1172, 5 maio 2015.

REFERÊNCIAS

BOUKAR O, ABBERTON M, OYATOMI O, TOGOLA A, TRIPATHI L AND FATOKUN, C. Introgression Breeding in Cowpea [*Vigna unguiculata (*L.) Walp.]. **Frontiers in Plant Science**, v. 11, set. 2020.

Companhia Nacional de Abastecimento: Observatório Agrícola **- Acompanhamento da safra brasileira de grãos, V.6, safra 2018/2019**, N.12, Décimo segundo levantamento, set, 2019.

CUI, L.-L. et al. Overexpression of Glycolate Oxidase Confers Improved Photosynthesis under High Light and High Temperature in Rice. **Frontiers in Plant Science**, v. 7, 4 ago. 2016.

DELLERO, Y. et al. Experimental Evidence for a Hydride Transfer Mechanism in Plant Glycolate Oxidase Catalysis. **Journal of Biological Chemistry**, v. 290, n. 3, p. 1689–1698, jan. 2015

FAO, Food and Agriculture Organization of the United Nations.2021. Disponível em http://www.fao.org/faostat/en/#data/QC/visualize Acessado em 2 de março de 2021.

FOYER, C. H. et al. Photorespiratory Metabolism: Genes, Mutants, Energetics, and Redox Signaling. **Annual Review of Plant Biology**, v. 60, n. 1, p. 455–484, 1 jun. 2009.

FURBANK, R. T. Evolution of the C4 photosynthetic mechanism: are there really three C4 acid decarboxylation types? **Journal of Experimental Botany**, v. 62, n. 9, p. 3103–3108, 1 maio 2011.

GONÇALVES, A., GOUFO, P., BARROS, A., DOMÍNGUEZ-PERLES, R., TRINDADE, H., ROSA, E. A. S., … RODRIGUES, M. Cowpea (Vigna unguiculataL. Walp), a renewed multipurpose crop for a more sustainable agri-food system: nutritional advantages and constraints. **Journal of the Science of Food and Agriculture**, v.96(9), p.2941–2951. 2016.

GRANDELLIS, C. et al. DOTAP, a lipidic transfection reagent, triggers Arabidopsis plant defense responses. **Planta,** v. 249, n. 2, p. 469–480, fev. 2019.

HAN, G. Origin and evolution of the plant immune system. **New Phytologist**, v. 222, n. 1, p. 70–83, abr. 2019.

JI, J. et al. Genome Editing in Cowpea *Vigna unguiculata* Using CRISPR-Cas9. **International Journal of Molecular Sciences**, v. 20, n. 10, p. 2471, 19 maio 2019.

JOSSIER, M. et al. Enzymatic Properties of Recombinant Phospho-Mimetic Photorespiratory Glycolate Oxidases from Arabidopsis thaliana and Zea mays. **Plants**, v. 9, n. 1, p. 27, 24 dez. 2019.

KARAPANOS, I. et al. Cowpea fresh pods - a new legume for the market: assessment of their quality and dietary characteristics of 37 cowpea accessions grown in southern Europe: Quality and dietary characteristics of cowpea fresh pods. **Journal of the Science of Food and Agriculture**, v. 97, n. 13, p. 4343–4352, out. 2017

LANGDALE, J. A. C $_4$ Cycles: Past, Present, and Future Research on C $_4$ Photosynthesis. **The Plant Cell**, v. 23, n. 11, p. 3879–3892, nov. 2011.

LI, L. et al. Genome-Wide Analysis and Expression Profiling of the Phospholipase D Gene Family in Solanum tuberosum. **Biology**, v. 10, n. 8, p. 741, 2 ago. 2021a.

LIU, Y.; WU, W.; CHEN, Z. Structures of glycolate oxidase from Nicotiana benthamiana reveal a conserved pH sensor affecting the binding of FMN. **Biochemical and Biophysical Research Communications**, v. 503, n. 4, p. 3050–3056, set. 2018.

LO, S. et al. Identification of QTL controlling domestication-related traits in cowpea (Vigna unguiculata L. Walp). **Scientific Reports,** v. 8, n. 1, p. 6261, dez. 2018.

MUTHAMILARASAN, M.; PRASAD, M. Plant innate immunity: An updated insight into defense mechanism. **Journal of Biosciences**, v. 38, n. 2, p. 433–449, jun. 2013.

NEJAT, NAGHMEH AND MANTRI, NITIN. Plant Immune System: Crosstalk Between Responses to Biotic and Abiotic Stresses the Missing Link in Understanding Plant Defence **Curr. Issues Mol. Biol.** 23: 1-16. 2016

NEWMAN, M.-A. et al. MAMP (microbe-associated molecular pattern) triggered immunity in plants. **Frontiers in Plant Science**, v. 4, 2013.

NORONHA SOUZA, P. F. et al. H2O2Accumulation, Host Cell Death and Differential Levels of Proteins Related to Photosynthesis, Redox Homeostasis, and Required for Viral Replication Explain the Resistance of EMS-mutagenized Cowpea to Cowpea Severe Mosaic Virus. **Journal of Plant Physiology**, v. 245, p. 153110, fev. 2020.

OSBORNE, C. P.; FRECKLETON, R. P. Ecological selection pressures for C $_4$ photosynthesis in the grasses. **Proceedings of the Royal Society B: Biological Sciences**, v. 276, n. 1663, p. 1753–1760, 22 maio 2009.

PAIVA, A. L. S. et al. Label-free Proteomic Reveals that Cowpea Severe Mosaic Virus Transiently Suppresses the Host Leaf Protein Accumulation During the Compatible Interaction with Cowpea (*Vigna unguiculata* [L.] Walp.). **Journal of Proteome Research**, v. 15, n. 12, p. 4208–4220, 2 dez. 2016.

PFAFFL, M. W. A new mathematical model for relative quantification in real-time RT–PCR.

ROJAS, C. M. et al. Glycolate oxidase modulates reactive oxygen species-mediated signal transduction during nonhost resistance in Nicotiana benthamiana and Arabidopsis. **The Plant Cell**, v. 24, n. 1, p. 336–352, jan. 2012.

SAMAILA, A. E., DEGRI, M. M. AND MSHELIA, J. S. Assessment of Field Insect Pests Damage on Cowpea in Gombe State, Nigeria. **International Journal of Entomology and Nematology Research.** v.4, n.1, p.10-23, jun. 2019.

SOBRINHO, CANDIDO ATHAYDE principais doenças do feijão-caupi no Brasil. In:Bastos EA (ed) **A Cultura do feijão-caupi no Brasil**. Embrapa Meio-Norte, Teresina, p. 44–67. 2016

UENO, O.; YOSHIMURA, Y.; SENTOKU, N. Variation in the Activity of Some Enzymes of Photorespiratory Metabolism in C4 Grasses. **Annals of Botany**, v. 96, n. 5, p. 863–869, 1 out. 2005.

VARELA, A. L. N. et al. A resistant cowpea (Vigna unguiculata [L.] Walp.) genotype became susceptible to cowpea severe mosaic virus (CPSMV) after exposure to salt stress. **Journal of Proteomics**, v. 194, p. 200–217, mar. 2019.

VARELA, A. L. N. et al. Gel-free/label-free proteomic, photosynthetic, and biochemical analysis of cowpea (Vigna unguiculata [L.] Walp.) resistance against Cowpea severe mosaic virus (CPSMV). **Journal of Proteomics**, v. 163, p. 76–91, jun. 2017.

XU, Y.-P.; YANG, J.; CAI, X.-Z. Glycolate oxidase gene family in Nicotiana benthamiana: genome-wide identification and functional analyses in disease resistance. **Scientific Reports**, v. 8, n. 1, p. 8615, dez. 2018.

ZELITCH, I. et al. High Glycolate Oxidase Activity Is Required for Survival of Maize in Normal Air. **Plant Physiology**, v. 149, n. 1, p. 195–204, 6 jan. 2009.

ZHANG, Z. et al. Catalytic and functional aspects of different isozymes of glycolate oxidase in rice. **BMC Plant Biology**, v. 17, n. 1, p. 135, dez. 2017.