UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E DO DESENVOLVIMENTO

Vitória Hirdes Glenzel

Análise da Ascorbato Oxidase em Fabaceae: uma abordagem evolutiva e funcional

Florianópolis
2023

Vitória Hirdes Glenzel

Análise da Ascorbato Oxidase em Fabaceae: uma abordagem evolutiva e funcional

Dissertação submetida ao Programa de Pós-Graduação em Biologia Celular e do Desenvolvimento da Universidade Federal de Santa Catarina como requisito parcial para a obtenção do título de Mestra em Estudos Celulares e Moleculares do Desenvolvimento

Orientadora: Profa. Dra. Franceli Rodrigues Kulcheski
Coorientadora: Profa. Dra. Andreia Carina Turchetto Zolet

Florianópolis
2023

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Glenzel, Vitória Hirdes
Análise da Ascorbato Oxidase em Fabaceae: uma abordagem evolutiva e funcional / Vitória Hirdes Glenzel ; orientadora, Franceli Rodrigues Kulcheski, coorientadora, Andréia Carina Turchetto Zolet, 2023.
111 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro de Ciências Biológicas, Programa de Pós Graduação em Biologia Celular e do Desenvolvimento, Florianópolis, 2023.

Inclui referências.

1. Biologia Celular e do Desenvolvimento. 2. Evolução. I. Kulcheski, Franceli Rodrigues. II. Zolet, Andréia Carina Turchetto. III. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Biologia Celular e do Desenvolvimento. IV. Título.

Vitória Hirdes Glenzel

Ascorbato Oxidase em Fabaceae: evolução e caracterização

O presente trabalho em nível de Mestrado foi avaliado e aprovado, em 21 de novembro de 2023, pela banca examinadora composta pelos seguintes membros:

Profa. Dra. Franceli Rodrigues Kulcheski
Universidade Federal de Santa Catarina

Profa. Dra. Norma Machado
Universidade Federal de Santa Catarina

Profa. Dra. Ana Lucia Segatto
Universidade Federal de Santa Maria

Certificamos que esta é a versão original e final do trabalho de conclusão que foi julgado adequado para obtenção do título de Mestre em Biologia Celular e do Desenvolvimento

ICP-Edu
Documento assinado digitalmente
Ricardo Castilho Garcez
Data: 21/02/2024 11:44:22-0300
CPF: ***.030.799-**
Verifique as assinaturas em https://v.ufsc.br

Coordenação do Programa de Pós-Graduação

ICP-Edu
Documento assinado digitalmente
Franceli Rodrigues Kulcheski
Data: 21/02/2024 16:30:40-0300
CPF: ***.987.540-**
Verifique as assinaturas em https://v.ufsc.br

Profa. Dra. Franceli Rodrigues Kulcheski
Orientadora

Florianópolis, 2023.

AGRADECIMENTOS

O mestrado definitivamente não é uma tarefa fácil - exige uma doação de corpo e mente, e ter pessoas que nos apoiam, faz total diferença! Por isso, não poderia deixar essa página em branco.

Quero começar agradecendo aos meus pais, Edson, e Jane, e minha irmã Iohana, que me deram o apoio e suporte necessários tanto na graduação, quanto no mestrado. Minha mãe que me ensinou a ser uma mulher forte e independente, e meu pai que me ensinou a ser uma pessoa justa e honesta, e minha irmã, que sempre acreditou em mim e me encheu de elogios mesmo nos momentos que eu não via nada de positivo em mim ou em meu trabalho. Obrigada, eu amo muito vocês!

Quero agradecer ao meu amor, Regis, por ser meu alicerce em tantos momentos difíceis, por me dar forças e demonstrar carinho nas pequenas (grandes) coisas do dia a dia. Sem você e a Cali, eu não teria conseguido, obrigada por absolutamente tudo, eu te amo muito!

Quero agradecer aos amigos que a vida me deu, tanto em Floripa, quanto àqueles que a vida permitiu que o vínculo fosse mantido, mesmo com a distância. E em especial amigas que de uma forma ou de outra, me apoiaram, me deram dicas, conselhos, e foram sol em dias de chuva: Bruna e Priscila, vocês são mulheres incríveis, agradeço demais pelo apoio que me deram, eu amo vocês! E não poderia deixar de agradecer minha parceira de trabalho, e amiga, Nicole, por todos os ensinamentos, risadas e conselhos! Te admiro muito e sou feliz de ter sua amizade!

Agradecer em especial ao Gabriel e a Monique, que foram tão solícitos e parceiros durante o mestrado, minhas colegas de laboratório, Sara e Rafaela, que alegraram meus dias, e a todos os colegas que conheci na UFSC, todos tão gentis e amigáveis. Em especial ao pessoal do grupo RNEC/UFSC, que fazem a diferença na divulgação da ciência. Agradecer também aos professores do PPGBCD, por todos os ensinamentos que levarei para a vida profissional e pessoal.

Quero agradecer a UFSC, pela oportunidade de concluir esse estudo e à FAPESC pelo apoio financeiro.

E por último, mas não menos importante, quero agradecer a minha orientadora, Profa. Dra. Franceli, por todos os ensinamentos, pela paciência e apoio! Te admiro muito, Profe, você é incrível! Agradecer a minha coorientadora, Profa. Dra. Andreia, pelas dicas que enriqueceram o trabalho, e pela parceria formada. Agradecer ao Edgar do PPGBM/UFRGS pelo auxílio nas análises, e a João Pedro, também do PPGBM, por toda parceria, apoio, pela sua contribuição nesse trabalho, e por ser um profissional tão gentil e exemplar. Aguardo sua visita em Floripa, João! Obrigada por tudo!

Finalizo afirmando que não conquistamos nada sozinhos, e que sou grata pelas pessoas que a vida colocou no meu caminho!

“Happiness is only real when shared.”

Christopher McCandless

RESUMO

A Ascorbato oxidase (AAO) é uma enzima da família das multicobre oxidases, que catalisa a oxidação do ácido ascórbico (AA), o principal antioxidante no espaço apoplástico das plantas, em dehidroascorbato (DHA). O AA é a principal molécula antioxidante presente no apoplasto, que por sua vez, é constituído pelo espaço intercelular, paredes celulares e xilema, e atua como barreira primária entre as células e o ambiente externo. Devido à regulação da oxidação do AA, a AAO é essencial para manter o estado redox do apoplasto e sua relação com várias situações fisiológicas e de estresses já foi demonstrada. No entanto, estudos evolutivos e funcionais sobre a enzima AAO na família de plantas Fabaceae ainda são escassos. A família de plantas Fabaceae inclui plantas com alto valor nutricional e econômico, como a soja, o feijão e a ervilha. Portanto, este trabalho teve como objetivo identificar e caracterizar os genes que codificam as proteínas AAO em espécies de Fabaceae, visando contribuir para a compreensão da evolução e função dessa família gênica. Para isso, foi realizada a análise de filogenia dos genes *AAO* em Fabaceae, a análise da estrutura gênica, bem como análise de motivos proteicos conservados. A região promotora dos genes *AAO* de *Glycine max* foi investigada a fim de identificar elementos *cis*-reguladores; e dados de expressão gênica *in silico* foram analisados durante situações de estresses bióticos e abióticos. Nossos resultados demonstraram uma conservação significativa da família de genes de *AAO* em todas as 21 espécies de Fabaceae que foram utilizadas nas análises. A análise filogenética dividiu os genes *AAO* de Fabaceae em dois grupos bem suportados, indicando que esses genes se originaram a partir de uma duplicação gênica que ocorreu antes da origem dessa família. A análise da estrutura gênica revelou a alta conservação dos genes *AAO* em Fabaceae e categorizou as sequências em dois grupos distintos. Além disso, a análise *de novo* de motivos revelou dez motivos conservados em quase todas as sequências de AAO. Os dados de localização cromossômica para os genes *AAO* de *G. max* e *Glycine soja* revelaram uma distribuição gênica altamente semelhante em todo o genoma. Na análise da região promotora dos genes *AAO* de *G. max,* identificamos motivos de ligação para Fatores de Transcrição (TFs) associados a vários cenários biológicos, incluindo desenvolvimento, crescimento e respostas a estresses bióticos e abióticos. Além disso, as análises de expressão gênica *in silico* revelaram variabilidade substancial nos perfis de expressão gênica de *AAO* sob diferentes estressores ambientais, destacando o papel funcional dinâmico da AAO durante as respostas a estresses bióticos e abióticos. Nossas investigações contribuíram para a identificação de múltiplos homólogos da AAO em diversas espécies pertencentes à família Fabaceae, proporcionando, assim, um aprimoramento significativo em nossa compreensão dos papéis funcionais desempenhados por essa família gênica.

Palavras-chave: Ascorbato oxidase; Fabaceae; Evolução; Estresses bióticos; Estresses abióticos; *Glycine max*.

ABSTRACT

Ascorbate oxidase (AAO) is an enzyme from the multicopper oxidase family, which catalyzes the oxidation of ascorbic acid (AA), the main antioxidant in the apoplastic space of plants, into dehydroascorbate (DHA). AA is the main antioxidant molecule present in the apoplast, which in turn is made up of the intercellular space, cell walls and xylem, and acts as a primary barrier between cells and the external environment. Due to the regulation of AA oxidation, AAO is essential for maintaining the redox state of the apoplast and its relationship with various physiological and stress situations has already been demonstrated. However, evolutionary, and functional studies on the AAO enzyme in the Fabaceae plant family are still scarce. The Fabaceae plant family includes plants with high nutritional and economic value, such as soybeans, beans and peas. Therefore, this work aimed to identify and characterize the genes that encode AAO proteins in Fabaceae species, aiming to contribute to the understanding of the evolution and function of this gene family. To this end, a phylogeny analysis of the *AAO* genes in Fabaceae, an analysis of the gene structure, as well as an analysis of conserved protein motifs were carried out. The promoter region of the *AAO* genes of *Glycine max* was investigated in order to identify *cis*-regulatory elements; and *in silico* gene expression data were analyzed during biotic and abiotic stress situations. Our results demonstrated significant conservation of the *AAO* gene family across all 21 Fabaceae species that were used in the analyses. Phylogenetic analysis divided the *AAO* genes of Fabaceae into two well-supported groups, indicating that these genes originated from a gene duplication that occurred before the origin of this family. Gene structure analysis revealed the high conservation of *AAO* genes in Fabaceae and categorized the sequences into two distinct groups. Furthermore, *de novo* motif analysis revealed ten conserved motifs in almost all AAO sequences. Chromosomal location data for the *AAO* genes of *G. max* and *Glycine soja* revealed a highly similar gene distribution across the genome. In the analysis of the promoter region of the *G. max AAO* genes, we identified binding motifs for Transcription Factors (TFs) associated with various biological scenarios, including development, growth, and responses to biotic and abiotic stresses. Furthermore, *in silico* gene expression analyzes revealed substantial variability in *AAO* gene expression profiles under different environmental stressors, highlighting the dynamic functional role of AAO during responses to biotic and abiotic stresses. Our investigations contributed to the identification of multiple AAO homologues in several species belonging to the Fabaceae family, thus providing a significant improvement in our understanding of the functional roles played by this gene family.

Keywords: Ascorbate oxidase; Fabaceae; Evolution; Biotic stress; Abiotic stress; *Glycine max*.

LISTA DE FIGURAS

Da introdução:

Do artigo científico:

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

AA: ..... Ácido ascórbico
AA-GSH: ..... Ascorbato-Glutationa
AAO:..... Ascorbato Oxidase
ABA: ..... Ácido abscísico
AP2/ERF: ..... APETALA2/Fator Responsivo ao Etileno (do inglês, *APETALA2/Ethylene Responsive Factor*)
APX: ..... Ascorbato Peroxidase
ARF: ..... Fator Responsivo a Auxina (do inglês, *Auxin Responsive Factor)*
BBR/BPC:..... Cevada B-Recombinante/PENTACISTEÍNA Básica (do inglês, *Barley B-Recombinant/Basic PENTACYSTEINE*)
BHLH:..... Hélice-Loop-Hélice Básica (do inglês, *Basic Helix-Loop-Helix*)
BZip:..... Zíper básico de Leucina (do inglês, *Basic Leucine zipper*)
C2H2: ..... Cisteína 2 Histidina 2 (do inglês, *Cysteine 2 Histidine 2*)
CDS:..... Sequência codificadora (do inglês, *Coding sequence*)
CMV: ..... Vírus do mosaico do pepino (do inglês, *Cucumber mosaic virus*)
CNV:..... (Variação de número de cópias (do inglês, *Copy number variation*)
DHA:..... Ácido dehidroascórbico
D-man/L-gal:..... D-manose/L-galactose
DOF:..... Ligação de DNA com um dedo (do inglês, *DNA-binding with One Finger*)
Dpi:..... Dias após infecção
GSDS: ..... Servidor de exibição de estrutura gênica (do inglês, *Gene Structure Display Server*)
H_2O_2:..... Peróxido de Hidrogênio
HD-Zip:..... Homeodomínio de Zíper de Leucina (do inglês, *Homeodomain Leucine Zipper*)
Hpi:..... Horas após inoculação
MCO:..... Multicobre Oxidases
MDHA:..... Monodehidroascorbato
MDHAR: ..... Monodehidroascorbato redutase
MEME: ..... Maximização de Múltiplas Expectativas para Elicitação de Motivos (do inglês, *Multiple Expectation Maximization for Motif Elicitation*)
MICK-MADS:..... Caixa MADS/Intervenção/Domínios semelhantes a queratina e C (do inglês, *MADS-box/Intervening/Keratin-like-and C-domains*)
MiRNAs: ..... MicroRNAs
MYB-Related:..... Relacionado a Mieloblastose (do inglês, *Myeloblastosis-Related*)
NAC:..... Domínios NAM-ATAF-CUC

Nin-Like: ..... Semelhante a Incepção do Nódulo (do inglês, *NODULE INCEPTION-Like*)
O_2: ..... Oxigênio
CQ: ..... Centro quiescente (do inglês, *quiescent center*)
ROS: ..... Espécie reativa de oxigênio (do inglês, *reactive oxygen species*)
SA ..... Ácido salicílico (do inglês, *Salicylic acid*)
SNV: ..... Variante de nucleotídeo único (do inglês, *Single nucleotide variant*)
TALE: ..... Extensão de loop de três aminoácidos (do inglês, *Three Amino acid Loop Extension*)
TCP: ..... TEOSINTE RAMIFICADO 1- CICLOIDEA- PCF1 (do inglês, *TEOSINTE BRANCHED 1- CYCLOIDEA- PCF1*)
TF: ..... Fator Transcricional (do inglês, *Transcription Factor*)
TNC: ..... Centro Trinuclear de Cobre
WOX: ..... Homeobox relacionado a WUSCHEL (do inglês, *WUSCHEL-related homeobox*)

SUMÁRIO

1. INTRODUÇÃO

1.1 ASCORBATO OXIDASE

A enzima Ascorbato Oxidase (AAO; EC 1.10.3.3), pertence à família de proteínas Multicobre Oxidases (MCO), que engloba lacases, ferroxidases e ceruloplasminas (MESSERSCHMIDT, 1990; HOEGGER *et al.*, 2006). Os membros desta família de proteínas possuem a capacidade de catalisar a oxidação do substrato e, ao mesmo tempo, reduzir o oxigênio molecular à água, sem gerar espécies reativas de oxigênio (ROS) (MESSERSCHMIDT; HUBER, 1990; GÓRALCZYK-BIŃKOWSKA; JASIŃSKA; DŁUGOŃSKI, 2019). A AAO está localizada no apoplasto das células vegetais, e tem como função catalisar a oxidação do ácido ascórbico (AA) em monodehidroascorbato (MDHA), que rapidamente se transforma em ácido dehidroascórbico (DHA) (MESSERSCHMIDT, A., 1997; SMIRNOFF, 2000). Entretanto, no citoplasma das células, a enzima Ascorbato Peroxidase (APX) é responsável por converter o peróxido de hidrogênio (H_2O_2) em água, e compartilha com a AAO, a afinidade pelo AA como doador de elétrons (SMIRNOFF, 2000).O AA é a principal molécula antioxidante presente no apoplasto, que por sua vez, é constituído pelo espaço intercelular, paredes celulares e xilema (FOYER; NOCTOR, 2005; FARVARDIN *et al.*, 2020). O apoplasto desempenha um papel fundamental ao permitir que as plantas detectem e respondam às mudanças ambientais, servindo como barreira primária entre as células e o ambiente externo (FARVARDIN *et al.*, 2020).

A rota de síntese mais amplamente distribuída e importante do AA em plantas superiores é a via D-manose/L-galactose (D-man/L-gal), inicialmente descrita em 1998 (WHEELER; JONES; SMIRNOFF, 1998; FOYER; KYNDT; HANCOCK, 2020). Quanto a via de regeneração do AA, o mesmo pode ser regenerado de DHA a AA por meio do ciclo Ascorbato-Glutationa (AA-GSH), e o radical MDHA também pode ser revertido em AA pela ação da enzima monodehidroascorbato redutase (MDHAR), como exemplificado na Figura 1 (SMIRNOFF, 2000).

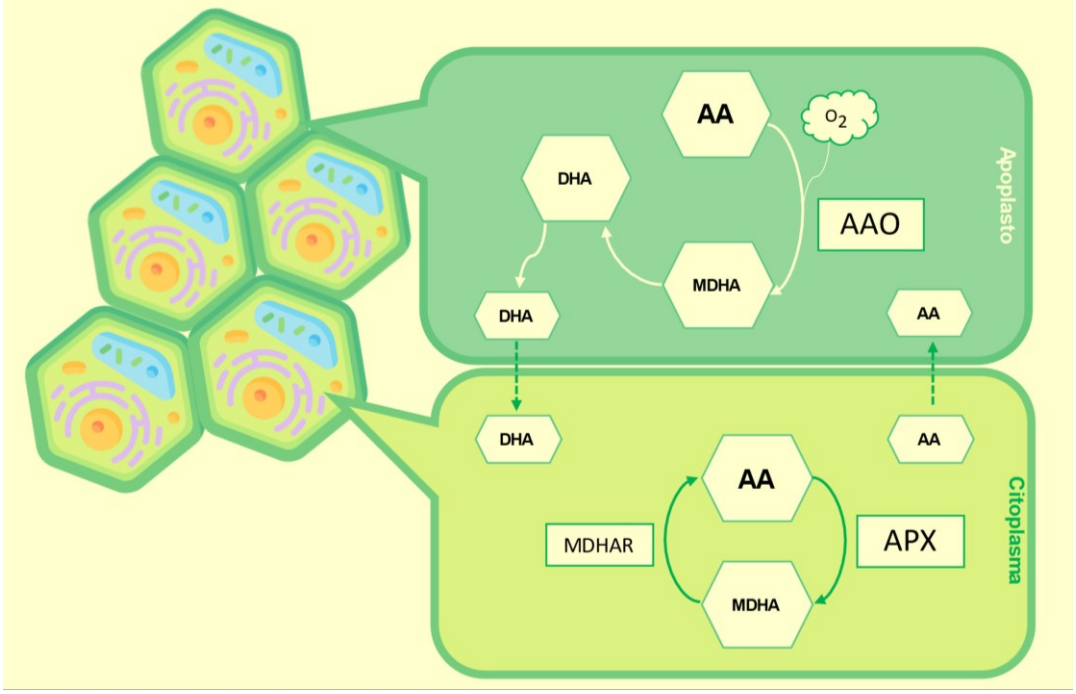


Figura 1. Ciclo de reciclagem do ácido ascórbico (AA). Os retângulos representam um aumento na região do citoplasma e do apoplasto da célula, demonstrando o ciclo de reciclagem do AA. Ascorbato Oxidase (AAO); Monodehidroascorbato (MDHA); Dehidroascorbato (DHA); Oxigênio (O2); Ascorbato Peroxidase (APX); Monodehidroascorbato Redutase (MDHAR). Fonte: da autora, baseada na figura de Smirnoff, 2000.

Todo esse processo enzimático que tem como chave a oxidação do AA pela AAO, é essencial para regular o estado redox do AA dentro do apoplasto, eliminando as ROS (SMIRNOFF, 2000). Experimentos realizados com tabaco evidenciaram que o aumento na atividade da enzima AAO tem um impacto significativo no nível de AA presente no apoplasto, influenciando, consequentemente, o equilíbrio redox nesse espaço extracelular (PIGNOCCHI *et al.*, 2003; DE TULLIO; GUETHER; BALESTRINI, 2013).

Quanto a sua estrutura, AAO é uma proteína dimérica que pode se dissociar em monômeros quando o pH do meio se encontra neutro ou alcalino (Di Venere et al. 2011). Cada monômero possui uma subunidade formada por 3 domínios distintos, que contém quatro íons de cobre, onde três formam um centro trinuclear próximo à interface entre os domínios (MESSERSCHMIDT et al., 1989). Esses íons de cobre estão organizados em um centro catalítico, que é subdividido em tipos T1 (ou azul) e T2, ambos contendo um átomo de cobre, e o tipo T3 contendo dois átomos de cobre (MESSERSCHMIDT et al., 1992). O local onde o oxigênio molecular é reduzido em água é chamado de Centro Trinuclear de Cobre

(TNC), e é formado pelos átomos de cobre de T2 e pelos cobres diatômicos de T3 (Figura 2). O mecanismo catalítico da AAO e outras MCOs inclui a redução do átomo de cobre T1 pelo elétron capturado do substrato, seguido pela transferência do elétron do sítio T1 para o TNC, resultando na redução do oxigênio com a formação de duas moléculas de água (FARVER; WHERLAND; PECHT, 1994; GÓRALCZYK-BIŃKOWSKA; JASIŃSKA; DŁUGOŃSKI, 2019).

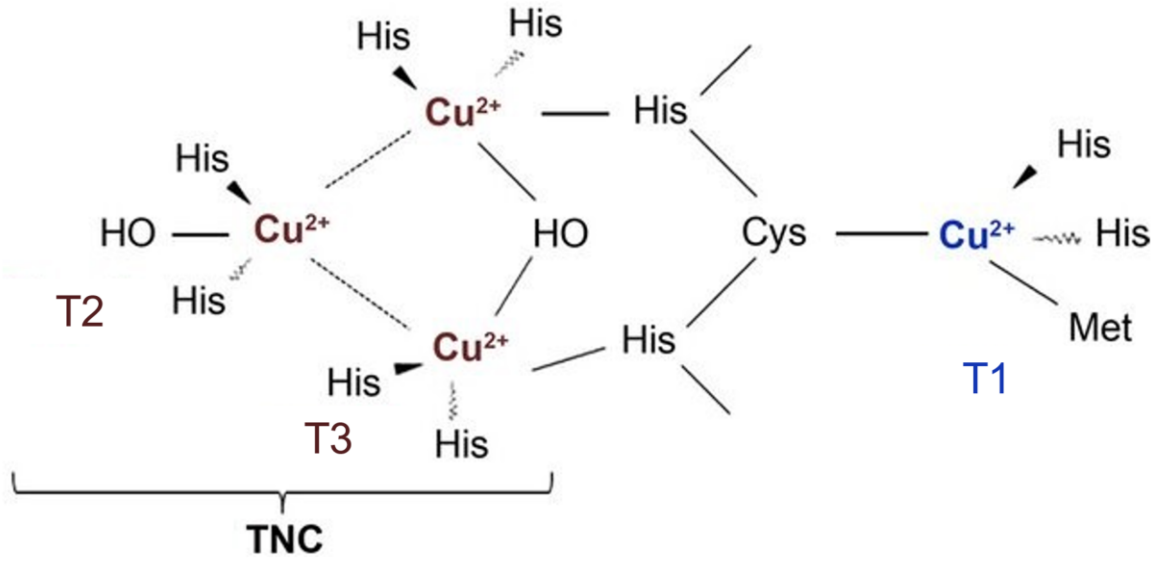


Figura 2: Modelo de Centro Trinuclear de Cobre (TNC). T1, T2, e T3: centro nucleares de cobre 1, 2 e 3. Cu: cobre; His: histidina; Met: metionina; Cys: cistetína. Fonte: Adaptada de Góralczyk-Bińkowska, Jasińska e Długoński, 2019.

1.2 FUNÇÕES E VIAS DE ATUAÇÃO DA ENZIMA AAO

Devido ao seu papel na regulação redox no espaço apoplástico, a AAO está intrinsecamente ligada a uma ampla gama de processos que incluem crescimento/alongamento e desenvolvimento celular, vias hormonais associadas à auxina, ácido abscísico (ABA), ácido salicílico e jasmonatos, formação de nódulos, percepção de luz, senescência, e respostas a estresses bióticos e abióticos (DIALLINAS *et al.*, 1997; KATO; ESAKA, 2000; PIGNOCCHI *et al.*, 2003; SANMARTIN *et al.*, 2007; BALESTRINI *et al.*, 2012; LI *et al.*, 2017; PAN *et al.*, 2019; SINGH *et al.*, 2021; ZHU *et al.*, 2023). Um dos primeiros artigos sobre AAO intitulado "*Ascorbate oxidase*", publicado em 1973 no The Journal of Biological

Biochemistry, demonstrou uma abordagem química da AAO, como suas propriedades espectrais de absorção, fluorescência e ressonância paramagnética eletrônica (LEE, DAWSON 1973). Além disso, a estrutura primária da AAO do pepino (*Cucumis sativus*) foi predita a partir da sequência de cDNA (OHKAWA *et al*., 1989). Esses estudos abriram portas para novas pesquisas acerca das funções dessa proteína. Nesse contexto, a relação entre a AAO e o crescimento/alongamento e desenvolvimento celular se destaca (DIALLINAS *et al*., 1997; KATO; ESAKA, 2000; SANMARTIN *et al*., 2007; LI *et al*., 2017). A interação entre o hormônio auxina e a AAO poderia impactar potencialmente o desenvolvimento da raiz, governando a regulação da formação do centro quiescente (QC) (KERK *et al*., 2000). A expressão e atividade da AAO no melão (*Cucumis melo*) aumentaram significativamente durante um período específico de desenvolvimento do fruto (DIALLINAS *et al*. 1997). No melão transgênico, o silenciamento de gene de *AAO* resulta em uma redução de crescimento, alteração dos níveis de AA e padrão de maturação diferentes (CHATZOPOULOU *et al*., 2020). Além disso, o crescimento celular foi significativamente aumentado em células BY-2 de tabaco cultivadas com superexpressão da *AAO*1 de algodão (*Gossypium hirsutum*) (LI *et al*., 2017). Este possível papel da AAO no crescimento e desenvolvimento foi explorado por Li e colaboradores (2017), que levantaram a hipótese de que o AAO regula a arquitetura da parede celular em parte pela oxidação do AA em DHA, causando o afrouxamento da parede celular e, portanto, o crescimento celular.

Alguns estudos demonstraram que a expressão ou atividade da AAO pode ser induzida/reprimida, por diferentes situações ou vias hormonais. Em frutos de abóbora, a auxina aumentou a abundância de transcritos do gene de *AAO*, e algo semelhante foi visto em tabaco (ESAKA *et al*., 1992; PIGNOCCHI *et al*., 2003). A atividade da AAO de melancia selvagem (*Citrullus lanatus* sp.) aumentou sob intensidade luminosa, indicando que a luz pode induzir a expressão gênica (NANASATO *et al*., 2005). Igualmente em *C. melo*, a abundância de mRNA da Cme*AAO4* aumentou progressivamente do sexto até o décimo dia em plântulas cultivadas na luz, em comparação com plantas cultivadas no escuro (SANMARTIN *et al*., 2007b). O mesmo estudo avaliou a abundância de mRNA sob tratamentos hormonais, e observaram que tratamentos com ácido jasmônico, resultaram em aumento nos níveis de transcritos de Cme*AAO4*,

enquanto ABA, ácido indol acético, e ácido salicílico provocaram diminuição na transcrição gênica de *AAO* (SANMARTIN *et al.*, 2007b). A aplicação de ABA em folhas de *Camellia sinensis* resultou na diminuição da expressão de todos os Cs*AAO*s. Além disso, o tratamento com ácido giberélico reprimiu Cs*AAO1* e Cs*AAO4*, mas Cs*AAO2* e Cs*AAO3* mostraram regulação positiva significativa, respectivamente, às 4 e 12 horas após os tratamentos (TIAN *et al*., 2019).

Estudos funcionais mostraram que a superexpressão ou repressão do gene *AAO* interfere nas respostas a estresses abióticos. A repressão da expressão de *AAO* aumentou a tolerância ao sal em tabaco, *Arabidopsis thaliana*, e tomate-cereja (YAMAMOTO *et al.*, 2005; ABDELGAWAD *et al.*, 2019). Além disso, a produção de frutos de tomate aumentou sob situações de deficiência hídrica em plantas com o gene de *AAO* silenciado (GARCHERY *et al*. 2013). Em *C. sinensis*, os níveis de expressão dos genes Cs*AAO* foram induzidos após 24h em situação de estresse salino (TIAN *et al.*, 2019). Entretanto em *Beta marítima* e *Beta vulgaris*, se observou uma repressão do gene *AAO* sob estresse salino ou seca (SKORUPA *et al*., 2022). De maneira análoga, observou-se um fenômeno correspondente durante o estresse de desidratação em sementes de *C. sinensis*, no qual os genes Cs*AAO1*, Cs*AAO3* e Cs*AAO4* apresentaram uma notável repressão (TIAN *et al.*, 2019). Em resposta a estresses de baixa ou alta temperatura, os níveis de expressão de todos os quatro genes *AAO* de *C. sinensis* também foram significativamente reduzidos. Por outro lado, a aclimatação ao frio aumentou a atividade de AAO no apoplasto de *A. nanus* (ZHU *et al*., 2022). Além disso, o controle da expressão da *AAO* a nível pós-transcricional é mediado por microRNAs (miRNAs). Em *O. sativa*, o miR528 regulou negativamente os transcritos de AAO, levando a uma tolerância ao estresse salino (WANG *et al*., 2021). No caso do *Ammopiptanthus nanus*, um arbusto da família Fabaceae, o miR4415 foi identificado como fator limitante da expressão de *AAO*, e se mostrou envolvido na aclimatação ao frio (ZHU *et al*., 2022)

No contexto de estresses bióticos, bioensaios em plantas de *B. vulgaris* revelaram que a pulverização das plantas com uma solução contendo AAO reduziu significativamente o número de fêmeas e cistos do nematóide *Heterodera schachtii* que infectam as raízes da planta, sugerindo que AAO induz resistência sistêmica contra esse patógeno (SINGH *et al.,* 2020a). Algo semelhante foi

observado na infecção do nematóide das galhas (*Meloidogyne graminicola)* em *O. sativa* (SINGH, Richard Raj *et al.*, 2020b). Foi observado que a oxidação do AA melhora a resposta de defesa à infecção por nematóides em plantas de arroz. Essa resistência foi baseada na via de fenilpropanóides, resultando em um fenótipo de tolerância (SINGH, Richard Raj et al., 2021). Além disso, a atividade elevada de AAO tem sido associada ao aumento do acúmulo de ROS, aumentando a capacidade de defesa de *O. sativa* em infecções pelo fungo *Magnaporthe oryzae* (HU *et al*. 2022b).

Os estudos mencionados evidenciam a ausência de um padrão uniforme no aumento ou diminuição da expressão gênica das AAO, dependendo da situação ou estresse. Esse fato pode indicar que as AAO são enzimas multifacetadas, e que por conta de seu caráter de regulação de ROS, estejam envolvidas nas mais variadas situações. Nesse sentido, torna-se interessante compreender os aspectos evolutivos e estruturais desses genes.

1.3 EVOLUÇÃO DA AAO

Apesar de haver conhecimento a respeito das funções e vias de atuação da AAO, ainda se sabe pouco sobre a história evolutiva dessa proteína. Estudos filogenéticos hipotetizam que a AAO tenha evoluído a partir de uma proteína de dois domínios, pela duplicação de um desses domínios (NAKAMURA *et al.*, 2003). A história evolutiva dos genes *AAO* tem sido explorada em espécies de plantas como *C. melo*, *Oryza sativa*, *A. thaliana*, *Glycine max*, *Sorghum bicolor*, *C. sinensis*, *G. hirsutum*, *Zea mays*, *B. vulgaris* e *A. nanus* (SANMARTIN *et al*., 2007b; BATTH *et al*., 2017; TIAN *et al*., 2019; PAN *et al*., 2019; WU *et al*., 2021; SKORUPA *et al*., 2022; ZHU *et al*., 2023). Em alguns desses estudos, os genes *AAO* se agruparam em três grupos, alguns contendo espécies dicotiledôneas e monocotiledôneas, no entanto, somente um dos genes de *O. sativa* se apresentou em um grupo separado (SANMARTIN *et al*., 2007b; BATTH *et al.*, 2017; TIAN *et al*., 2019; PAN *et al*., 2019). Em uma análise que abrangeu várias espécies, as sequências de Fabaceae formaram um grupo distinto (DE TULLIO; GUETHER; BALESTRINI, 2013). O mesmo estudo também encontrou sequências putativas de AAO em espécies ancestrais, como a *Selaginella moellendorffii*, o musgo *Physcomitrella patens* e as algas verdes *Chlorella* e *Chlamydomonas* (DE

TULLIO; GUETHER; BALESTRINI, 2013). Diferentemente, um estudo que focou em sequências de *Z. mays*, os genes *AAO* das espécies analisadas foram agrupados em quatro grupos distintos, com putativas sequências de *Z. mays* presentes em todos os grupos (WU *et al*. 2021).

Entretanto, ainda não foram realizados estudos que enfoquem na evolução das *AAO*s pertencentes a uma família de plantas, como as Fabaceae.

1.4 FAMÍLIA DE PLANTAS FABACEAE

A família Fabaceae, também conhecida como Legumes ou Leguminosae, é a terceira maior família de plantas terrestres e a segunda cultura alimentar mais importante depois dos cereais, servindo como um componente essencial de uma dieta equilibrada (MAGALLÓN, SANDERSON 2001; WANG *et al*., 2017) MAPHOSA *et al*., 2017; SINGH *et al*. 2021a). Esta família botânica abrange não apenas espécies economicamente importantes como *G. max* (soja), *Phaseolus vulgaris* (feijão-comum), *Pigeon hypogea* (amendoim), *Pisum sativum* (ervilha) e *Cicer arietinum* (grão de bico), mas também plantas modelo importantes como *Medicago truncatula*, juntamente com várias árvores, arbustos e plantas herbáceas (GOEBEL, 1969; WANG *et al*., 2017). As espécies de Fabaceae possuem valor nutricional significativo, fornecendo cerca de 20-45% de seu peso em proteínas com aminoácidos essenciais, 60% em carboidratos complexos e 5-37% em fibra alimentar (MAPHOSA *et al*., 2017; SINGH, D. *et al*., 2021). No entanto, a sua importância vai além do seu valor econômico e nutricional, uma vez que os legumes têm a notável capacidade de estabelecer relações simbióticas que facilitam a fixação de nitrogênio (ROY *et al.*, 2020). Esta associação simbiótica não só reduz a dependência de fertilizantes à base de nitrogênio, mas também enriquece o solo para futuras plantações (LINDSTRÖM, MOUSAVI 2020).

1.5 JUSTIFICATIVA

Dado o papel fundamental da enzima AAO em diversas situações nas células vegetais, bem como a significativa importância econômica e nutricional das espécies de Fabaceae, explorar mais detalhadamente a caracterização dessa família de genes dentro desse grupo de plantas pode proporcionar percepções sobre mecanismos que têm o potencial de aprimorar tanto a funcionalidade quanto às características dessas plantas. Além disso, a região promotora dos genes *AAO* de *G. max*, bem como os dados de expressão gênica desses genes sob condições de estresse biótico e abiótico, carecem de investigações

abrangentes. Portanto, a investigação dos aspectos funcionais desses genes poderá revelar potenciais alvos para análises subsequentes.

OBJETIVOS

1.6 OBJETIVO GERAL

Identificar e caracterizar os genes que codificam as proteínas AAO em espécies de Fabaceae visando contribuir com a compreensão sobre a evolução e função dessa família gênica.

1.7 OBJETIVOS ESPECÍFICOS

- Identificar as sequências de AAO pertencentes às espécies da família de plantas Fabaceae;
- Compreender as relações evolutivas entre os genes de *AAO* na família de plantas Fabaceae, por meio de análises filogenéticas;
- Caracterizar as proteínas AAO na família de plantas Fabaceae;
- Compreender a estrutura e a distribuição dos genes de *AAO* nas espécies de *G. max* e *G. soja*;
- Investigar os elementos regulatórios situados na região promotora dos genes de *AAO* de *G. max*;
- Elucidar a rede de processos biológicos associados aos genes *AAO* em *G. max* por meio de análises da expressão gênica frente a estresses bióticos e abióticos.

2 CONTEXTUALIZAÇÃO DO ARTIGO CIENTÍFICO

A enzima ascorbato oxidase (AAO) catalisa a reação de oxidação do ácido ascórbico em monodehidroascorbato (MDHA) com rápida transformação em dehidroascorbato (DHA). Essa enzima está localizada no apoplasto das células vegetais, que por sua vez, atua como primeira linha de defesa contra estresses e outros sinais externos. Por conta disso, a AAO tem se mostrado uma enzima crucial em vários cenários. Nesse contexto, alguns estudos já foram realizados mostrando a função, caracterização e evolução dessa família gênica, em espécies de plantas como *Z. mays*, *G. hirsutum*, *Amopitantus nanus*, entre outras. No entanto, até o momento não houve estudos evolutivos dentro da família de plantas Fabaceae. Essa família é também conhecida como Legumes ou Leguminosae, e é a terceira maior família de plantas terrestres e a segunda cultura alimentar mais importante depois dos cereais, servindo como um componente essencial de uma dieta equilibrada. Sendo assim, aprofundar a caracterização desta família de genes dentro das Fabaceae pode oferecer percepções sobre mecanismos que visam melhorar tanto a funcionalidade como os atributos dessas plantas. Portanto, para identificar e caracterizar as proteínas e genes que codificam AAO em diversas plantas Fabaceae, realizamos uma série de análises, incluindo análise filogenética, identificação de motivos conservados e análise de estrutura gênica. Além disso, conduzimos uma exploração de *cis*-elementos reguladores na região promotora dos genes *AAO* de *G. max,* e realizamos uma análise de expressão usando dados de mRNA-seq de *G. max* disponíveis publicamente. Nossos resultados contribuíram para a identificação de múltiplos homólogos de AAO em espécies de Fabaceae e melhoraram nossa compreensão dos papéis funcionais desta família de genes.

3 ARTIGO CIENTÍFICO

A ser submetido em revista da área

Caracterização e evolução dos genes de Ascorbato Oxidase (AAO) em plantas da família Fabaceae: um estudo filogenético e funcional*

* título mantido propositalmente em português para a dissertação

ABSTRACT

Ascorbate oxidase is a multicopper oxidase protein that catalyzes the oxidation of ascorbic acid (AA), the paramount antioxidant in the apoplastic space. Although the role of AAO has been widely studied, there are just a few researches into its evolution and function of this enzyme in the Fabaceae plant family, which includes plants with high nutritional and commercial value. Our investigation revealed significant conservation of the *AAO* gene family across all 21 studied Fabaceae species. Phylogenetic analysis consistently grouped Fabaceae *AAO* genes into two well-supported clusters, indicating their similar and conserved origin. This conservation extended to gene structure analyses, categorizing sequences into two groups based on intronic sizes. Moreover, a *de novo* motif discovery unveiled ten conserved motifs in almost all AAO sequences. Notably, the chromosomal localization data for *Glycine max* and *Glycine soja AAO* genes revealed a highly similar gene distribution across the genome. In our comprehensive *cis*-regulatory analysis of *G. max AAO* genes, we identified binding motifs for Transcription Factors (TFs) associated with various biological scenarios, including development, growth, and responses to both biotic and abiotic stresses. Further, *in silico* gene expression analyses revealed substantial variability in *AAO* gene expression profiles under different environmental stressors, emphasizing AAO's dynamic functional role during responses to biotic and abiotic stresses. Our findings helped us identify multiple *AAO* homologs in Fabaceae species and improved our understanding of the gene family's functional roles.

Keywords: Ascorbate oxidase; Fabaceae; Evolution; Biotic stress; Abiotic stress; *Glycine max*.

3.1 INTRODUCTION

The apoplast, comprising the intercellular space, cell walls, and xylem, serves as the primary barrier between plant cells and the external environment (Farvardin et al. 2020). It plays a pivotal role in enabling plants to detect and respond to environmental changes. In this compartment, one of the key enzymatic players is the ascorbate oxidase (AAO; EC 1.10.3.3). AAO catalyzes the oxidation of ascorbic acid (AA), the paramount antioxidant in the apoplastic space, into monodehydroascorbate (MDHA), which subsequently transforms into dehydroascorbic acid (DHA) at a rapid rate (Messerschmidt, A., 1997; Smirnoff, 2000). This enzymatic process is essential for regulating the redox state of AA within the apoplast by effectively scavenging and detoxifying reactive oxygen species (ROS) (Smirnoff, 2000). Experimental evidence from transgenic tobacco plants demonstrated that heightened AAO activity significantly influences the ascorbic acid pool within the apoplast, thereby impacting the redox equilibrium of this extracellular space (Pignocchi et al. 2003; De Tullio; Guether; Balestrini, 2013). The AAO enzyme belongs to the Multicopper Oxidase Family (MCO), which encompasses laccases, ferroxidases, and ceruloplasmin (Hoegger et al. 2006). Members of this protein family possess the unique ability to catalyze substrate oxidation while concomitantly reducing molecular oxygen to water, all without generating ROS (MESSERSCHMIDT; HUBER, 1990; GÓRALCZYK-BIŃKOWSKA; JASIŃSKA; DŁUGOŃSKI, 2019). Regarding its structural composition, AAO comprises three distinct functional domains, each containing four copper ions (Messerschmidt et al. 1989). These copper ions are organized into a catalytic center, which is further subdivided into T1 (or blue) and T2 types, both containing one copper atom, and the T3 type containing two copper atoms (Messerschmidt et al. 1992). This region serves as the site where molecular oxygen undergoes reduction into water (Farver; Wherland; Pecht, 1994; Góralczyk-Bińkowska; Jasińska; Długoński, 2019).

The initial characterization of the AAO enzyme involved a chemical approach, which unveiled its absorption, fluorescence, and electron paramagnetic resonance spectral properties (Lee and Dawson 1973). Later, the primary structure of cucumber (*Cucumis sativus*) *AAO* was elucidated through cDNA sequence analysis (Ohkawa et al. 1989). Since then, numerous investigations

have aimed to unravel the multifaceted functions of AAO. Due to its role in redox regulation in the apoplastic space, it has become evident that this enzyme is intricately linked to a wide array of processes in plants. These include cell growth/elongation and development, hormone pathways such as auxin, abscisic acid, salicylic acid, and jasmonates, nodule formation, light perception, senescence, as well as responses to biotic and abiotic stresses (Diallinas et al. 1997; Kato; Esaka, 2000; Pignocchi et al. 2003; Sanmartin et al. 2007; Balestrini et al. 2012; Li et al. 2017; Pan et al. 2019; SINGH, Richard Raj et al. 2021; Zhu *et* al. 2023).

The involvement of AAO in plant development was demonstrated in some experiments conducted with melon (Diallinas et al. 1997; Al-Madhoun; SANMARTIN; Kanellis, 2003; Sanmartin et al. 2007). During specific stages of melon fruit development, a notable and significant increase in *AAO* gene expression as well as enzyme activity was observed (Sanmartin et al., 2007). Conversely, the *AAO* gene silencing in transgenic melon reduced growth, altered ascorbic acid levels, and caused a different fruit ripening pattern (Chatzopoulou et al. 2020). Furthermore, cell growth was also significantly boosted in *Gossypium hirsutum AAO1*-overexpressing cultured tobacco BY-2 cells (Li et al. 2017). Interesting, while many studies emphasize the role of AAO in growth and development, other investigations shed light on its involvement in response to biotic and abiotic stresses. Suppressed *AAO* expression in tobacco and cherry tomato, improved tolerance to salt stress (Yamamoto et al. 2005; Abdelgawad et al. 2019). The AAO activity assay revealed that cold acclimation increased AAO activity in the apoplast of *Ammopiptanthus nanus* (Zhu et al. 2022). In the context of biotic stresses, plant bioassays have revealed that AAO application significantly reduces the number of females and cysts infesting sugar beet roots, suggesting that AAO induces systemic resistance against the cyst nematode *Heterodera schachtii* (Singh, et al. 2020a). Something similar was seen in the infection of root-knot nematode *Meloidogyne graminicola* in *Oryza sativa* (Singh, et al., 2020b). Ascorbate oxidation improves the phenylpropanoid-based response to nematode infection in treated rice plants, resulting in a tolerance phenotype (Singh, et al. 2021). Furthermore, elevated AAO activity has been associated with increased ROS accumulation, enhancing the plant's capacity to defend against a fungus in rice (Hu et al. 2022b).

The evolutionary history of the *AAO* genes has been explored in various plant species, including *Cucumis melo*, *O. sativa*, *Arabidopsis thaliana*, *Glycine max*, *Sorghum bicolor*, *Camellia sinensis*, *G. hirsutum*, *Zea mays*, *Beta vulgaris*, and *A. nanus* (Sanmartin et al. 2007; Batth et al. 2017; Tian et al. 2019; Pan et al. 2019; Wu et al. 2021; Skorupa et al. 2022; Zhu et al. 2023). However, there have been no prior evolutionary studies within the Fabaceae plant family. The Fabaceae family, alternatively known as Legume or Leguminosae, ranks as the third-largest land plant family and the second most important food crop after cereals, serving as an essential component of a balanced diet (Magallón; Sanderson, 2001; Wang, J. et al. 2017; Maphosa et al. 2017; Singh, D. et al. 2021). This botanical family encompasses not only valuable crops such as *G. max*, *Phaseolus vulgaris*, *Pigeon hypogea*, *Pisum sativum*, and *Cicer arietinum*, but also important model plants like *Medicago truncatula*, along with various trees, shrubs, and herbaceous plants (Goebel, 1969; Wang et al. 2017). Fabaceae species hold significant nutritional value, providing proteins (20-45%) with essential amino acids, complex carbohydrates (60%), and dietary fiber (5-37%) (Maphosa et al. 2017; Singh, D. et al. 2021). However, their importance extends beyond their economic value once the leguminous plants can establish symbiotic relationships that facilitate nitrogen fixation (Roy et al. 2020). This symbiotic association reduces dependence on nitrogen-based fertilizers and enriches the soil for future plantations (Lindström and Mousavi 2020).

Given the pivotal role of AAO as a crucial enzyme in various plant scenarios and the substantial global economic and nutritional relevance of Fabaceae species, delving into the characterization of this gene family within Fabaceae can offer insights into mechanisms aimed at enhancing both the functional and attributes of these plants. Therefore, in order to identify and characterize the proteins and genes that code for AAO in several Fabaceae plants, we conducted a series of analyses, including phylogenetic analysis, the identification of conserved motifs, and an examination of the genetic structure. Additionally, we explored *cis*-regulatory elements in the *G. max AAO*s promoter region and performed an *in silico* expression analysis using publicly available *G. max* mRNA-seq data. Our results contributed to identifying multiple *AAO* homologs in Fabaceae species and enhanced our understanding of the functional roles of this gene family.

3.2 METHODOLOGY

3.2.1 Database search and sequence acquisition

In order to elucidate the evolutionary patterns of the Fabaceae *AAO* genes, we conducted an extensive search in public genome databases. Our approach involved performing a BLASTp search against 21 complete genomes from the Fabaceae species, which were sourced from Phytozome v.13 (phytozome-next.jgi.doe.gov), Ensembl Plants v.57 (plants.ensembl.org), and NCBI (www.ncbi.nlm.nih.gov/) databases. We used the AAO3 protein sequence from *G. max* (Glyma.13G076900) as the query sequence in BLAST searches. The AAO sequences from *Selaginella moellendorffii*, *O. sativa*, and *A. thaliana* were used as outgroups in phylogenetic analysis. Our search in the Phytozome database yielded a total of sequences from 11 Fabaceae species, including *G. max*, *Arachis hypogaea*, *Phaseolus lunatus*, *M. truncatula*, *Phaseolus acutifolius*, *Vigna unguiculata*, *Lupinus albus*, *C. arietinum*, *Lotus japonicus*, *P. vulgaris*, and *Glycine soja*. Furthermore, we identified sequences from five species from the Ensembl Plants database: *Lupinus angustifolius*, *P. sativum*, *Vigna angularis*, *Vigna radiata*, and *Trifolium pratense*. Meanwhile, the NCBI database contained sequences from five Fabaceae species, namely *Abrus precatorius*, *Arachis duranensis*, *Arachis ipaensis*, *Cajanus cajan*, *Vigna umbellata*. For clarity and brevity, we abbreviated the plant names using the first letter of the genus and two letters of the species, followed by the identification code from the respective database. For example, Gma corresponds to *G. max*. All retrieved sequences were subjected to further analyses.

3.2.2 Multiple sequence alignment and Phylogenetic analysis

To elucidate the phylogenetic relationship among the Fabaceae AAO proteins, we performed multiple sequence alignments followed by phylogenetic analysis. The alignments were carried out using MAFFT software v.7 (mafft.cbrc.jp) with the FFT-NS-2 algorithm (Kuraku et al. 2013; (Katoh et al. 2019). The 70 resulting protein alignments contained 529 aminoacid sites. For the subsequent phylogenetic analysis, we employed the Bayesian information criterion (BIC) to identify the best substitution models for the protein sequences dataset (Kalyaanamoorthy et al. 2017). Model Finder suggested the WAG+I+G4 and

LG+I+G4 models as the most suitable ones. The WAG+I+G4 model was utilized to generate the phylogenetic tree encompassing all retrieved species, while the LG+I+G4 model was employed for the phylogeny used in gene structure analysis. Phylogenetic trees were estimated using IQ-TREE v2.2.2.5 web server (iqtree.cibiv.univie.ac.at) with 1,000 ultrafast bootstrap iterations based on the maximum likelihood method (Trifinopoulos et al. 2016). The resulting trees were visualized using FigTree v1.4.4 (tree.bio.ed.ac.uk/software/figtree/). Furthermore, to establish the evolutionary timeline of the Fabaceae species presented in this study, a Tree of Life was generated using the Time Tree v.5 online tool (http://www.timetree.org/).

3.2.3 Gene structure analysis and motif characterization of AAO proteins

We conducted a gene structure analysis to gain insights into the structural diversity of *AAO* genes. For this purpose, we utilized the Gene Structure Display Server (GSDS) v.2.0 by aligning the genomic sequences with the coding sequences (CDS) (Hu et al. 2015). Additionally, we integrated this analysis with a phylogeny tree in Newick format to better understand the evolutionary relationships of the *AAO* genes. For this analysis, we selected the following Fabaceae species: *G. max*, *A. hypogaea*, *P. vulgaris*, *M. truncatula*, *C. arietinum*, and *P. sativum*. To identify conserved motifs within the AAO proteins and their respective amino acid logos, we performed a *de novo* motif analysis using the Multiple Expectation Maximization for Motif Elicitation (MEME) v.5.5.3 tool (Bailey et al. 2015). The classic motif discovery mode and the zero or one occurrence per sequence motif distribution parameter were used, with ten maximum number of motifs. Furthermore, every motif was characterized utilizing a motif search site (www.genome.jp/tools/motif/) with the data from the Pfam database (www.ebi.ac.uk/interpro/entry/pfam).

3.2.4 *Cis*-acting Elements Analyses and Genic Distribution of *AAO*s in *Glycine* spp.

To comprehend the possible pathways that *AAO* genes are involved in, we investigated the regulatory regions of these genes. For this, the Plant Transcriptional Regulatory Map database (http://plantregmap.gao-lab.org) was

used to identify selected *cis*-regulatory elements in the 2000 pb upstream *G. max AAO*s gene sequence. A threshold p-value of 1e-5 was used. Furthermore, a comprehensive genic distribution analysis was performed to elucidate the organization of *Glycine* spp. *AAO* genes at the chromosomal level. The chromosome location data was obtained from the Genome Data Viewer of NCBI (www.ncbi.nlm.nih.gov/genome/gdv), and the Mapgene2chrom online tool (MG2C) v2.1 (http://mg2c.iask.in/mg2c_v2.1/) was employed to create a chromosome diagram.

3.2.5 *In Silico* Gene Expression analysis of *AAO* genes

An *in silico* gene expression analysis of the *G. max AAO*s was performed using mRNA-Seq data retrieved from Genevestigator Plants software (https://genevestigator.com/). A heatmap was generated with the HeatMaper online tool (http://heatmapper.ca/) using Log2-ratio, and *p* values to visualize *AAO* expression patterns under biotic and abiotic stresses. The Log2-ratio values were considered significant above -2.5 and upper 2.5, and the considered *p*-value was ≤ 0,005. The biotic stresses included infections caused by several pathogens, such as *Aphis glycines*, *Calonectria ilicicola*, *Phakopsora pachyrhizi*, *Fusarium oxysporum*, *Phytophthora sojae*, and soybean mosaic virus. Additionally, we examined information related to endosymbiotic relationships involving *Bradyrhizobium diazoefficiens* and *B. japonica*. Furthermore, we collected data on various abiotic stresses, including dehydration, salinity, extreme heat, drought, and submergence.

3.3 RESULTS

3.3.1 AAO Proteins Identification in Fabaceae

To know more about the AAO protein in the Fabaceae family and provide information about its molecular phylogeny and its gene and protein structure, we performed a wide search in the main databases. We retained only those sequences that possessed the three multicopper domains and exhibited characteristic residues consistent with well-characterized AAO (Messerschmidt, A. et al. 1992; Nakamura et al. 2003). We found 61 sequences in 21 Fabaceae species, three in *A. thaliana*, four in *O. sativa* and two in *S. moellendorffii,* totaling

70 sequences. Among them, 49 novel AAO sequences across 16 distinct Fabaceae species were identified, which were from *G. soja*, *A. hypogea*, *A. duranensis*, *A. ipaensis*, *P. lunatus, P. acutifolius*, *V. unguiculata*, *V. umbellate*, *V. angularis*, *V. radiata, L. albus*, *L. angustifolius*, *C. arietnum*, *T. pratense*, *A. precatorius*, and *C. cajan*. The information about these sequences are in Supplementary Table 1.

The number of *AAO* genes varied between Fabaceae species. We found five *AAO* genes in *G. max*, six in *G. soja,* and one in *V. radiata*. In *V. umbellate*, *M. truncatula*, *C. arietinum*, *P. sativum*, and *T. pratense* we found two *AAO* genes. *A. hypogaea*, *A. duranensis*, and *V. angularis* present three *AAO* genes. *P. acutifolius*, *P. vulgaris*, *V. unguiculata*, *A. precatorius*, and *C. cajan* have four *AAO* genes. The protein sizes of the AAOs varied from 541 to 1760 amino acids.

3.3.2 Phylogenetic analysis of Fabaceae AAOs

To understand the relationship among *AAO* genes from the Fabaceae species, a phylogenetic tree was constructed using the protein-aligned sequences, supported by UFbootstraps values above 90 in most (Fig. 1). The tree topology revealed two main clusters, Group I and Group II, which comprehend the AAOs from eudicots and monocot species, indicating that the duplication event that originated these groups occurred in the ancestral of these two angiosperms lineages. The two *S. moellendoffii* sequences grouped as an outgroup, suggesting that groups I and II diversified after the origin of lycophytes.

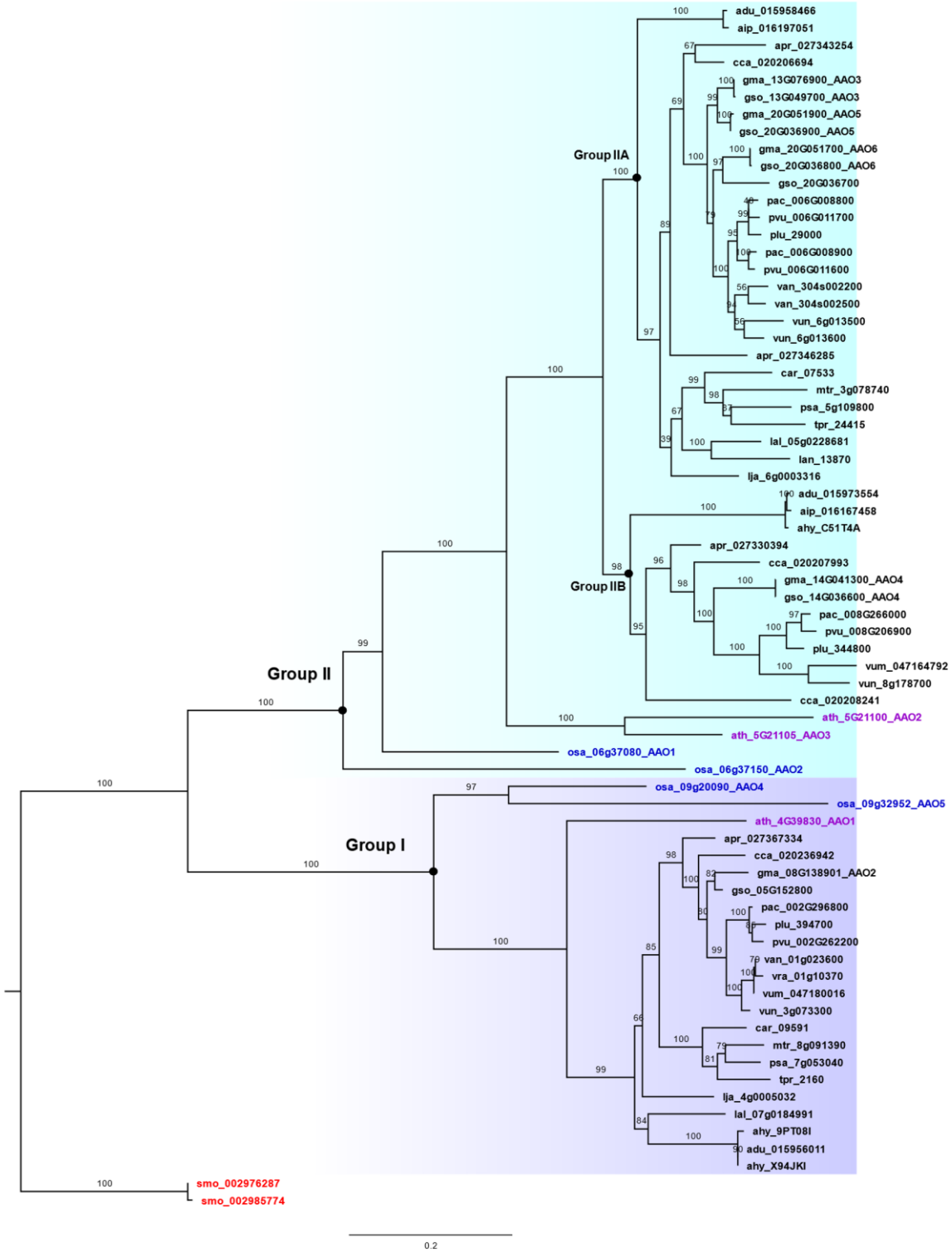


Fig. 1: Phylogenetic relationship of *Ascorbate oxidase* (*AAO*) genes in Fabaceae using the Maximum likelihood method.The tree was constructed by the maximum likelihood method using IQ-TREE web server with 1000 UFbootstrap replications, the WAG+I+G4 substitution model, and edited with FigTree software. Numbers on branches indicate UFbootstrap support values. The phylogenetic tree shows two well-supported groups (Group I and Group II), indicated with purple and blue background colors, respectively. Group II is subdivided into Group IIA and IIB. The Fabaceae sequences are represented in black, while *O. sativa* sequences are in blue, *A. thaliana*

sequences are in purple and *S. moellendorffii* AAOs are the outgroup in red. The protein sequences encoded by the AAO genes belong to the Fabaceae specie *Glycine max* (gma), *Glycine soja* (gso), *Arachis hypogaea* (ahy), *Phaseolus lunatus* (plu), *Medicago truncatula* (mtr), *Phaseolus acutifolius* (pac), *Vigna unguiculata* (vun), *Lupinus albus* (lal), *Cicer arietinum* (car), *Lotus japonicus* (lja), *Phaseolus vulgaris* (pvu), *Lupinus angustifolius* (lan), *Pisum sativum* (psa), *Vigna angularis* (van), *Vigna radiata* (vra), *Trifolium pratense* (tpr), *Abrus precatorius* (apr), *Arachis duranensis* (adu), *Arachis ipaensis* (aip), *Cajanus cajan* (cca), *Vigna umbellate* (vum) and from *Arabidopsis thaliana* (ath), *Selaginella moellendorffii* (smo) and *Oryza sativa* (osa) and are listed in Table S1.

Group I comprehend sequences from *A. thaliana*, O. sativa and the Fabaceae sequences grouped closely. Group II underwent a division, leading to the formation of two distinct clusters labeled Groups IIA and IIB. Almost all species were found in both Groups I and II, except for *L. angustifolius* AAO (Lan13870), represented by a unique sequence in Group IIA, and the *A. ipaensis* sequences (Aip016197051 and Aip016167458) present in Groups IIA and IIB, respectively. The phylogeny provides a glimpse into orthologous relationships, particularly between *Glycine* spp. AAOs. Within Group IIA, the orthology could be observed in *G. max* and *G. soja*, where AAO3, AAO5, and AAO6 are clustered together. The same occurs in Group IIB, with *G. max* and *G. soja* AAO4. In Group IIA, *P. acutifolius* and *P. vulgaris* AAOs also display orthologues such as Pac006g008800 and Pvu006g011700, Pac006G008900 and Pvu006g011600, and in Group IIB, Pac008G266000 and Pvu008G266000. The orthologous relationship is also found in Group IA, within the AAOs of *V. angularis* and *V. radiata*, in Van01g023600 and Vra01g10370. Furthermore, there is a notable presence of paralogous relationships in Groups I and II. However, the paralogy is more prevalent in Group IIA, with *V. angularis* AAOs (Van304s002200 and Van304s002500), as well as in *V. unguiculata* AAOs (Vun06g013500 and Vun06g013600).

3.3.3 Protein and Gene Characterization of AAO in Fabaceae, Using a *De Novo* Motif Discovery and a Gene Structure Analysis

To better understand and characterize the AAO proteins, especially in Fabaceae, we performed a *de novo* motif in the same species used in the phylogenetic analyzes. This analysis allowed us to find ten conserved motifs in 44 of 70 AAOs sequences. However, 23 sequences do not have the motif 8: AhyX94JKI, Ahy9PT081, GmaAAO2, Mtr8g091390, Pvu002G262200, Plu394700, Vum3g073300, Pac002G296800, Lal07g0184991, Car09591, Lja4g0005032, Gso05G152800, Psa7g053040, Van01g023600, Vra01g10370, Tpr2160, Apr027367334, Adu015956011, Cca020236942, Vum047180016, AthAAO2, OsaAAO2, and OsaAAO4. The sequence OsaAAO5 does not have motif 3, and two sequences do not have motifs 6 and 10, which were Pvu008G206900, and Gso20G036700. An amino acid logo analysis revealed high similarity between Fabaceae AAOs proteins (Fig. 2B). We also performed a domain characterization using the Pfam database in the ten discovered motifs (Fig. 2C). In this analysis, we found that motifs 4, 5, and 7 are comprehended on the Cu-oxidase domain, while motifs 2 and 3 are present in the Cu-oxidase_2 domain, and the motifs 1, 6, and 8 in the Cu-oxidase_3 domain.

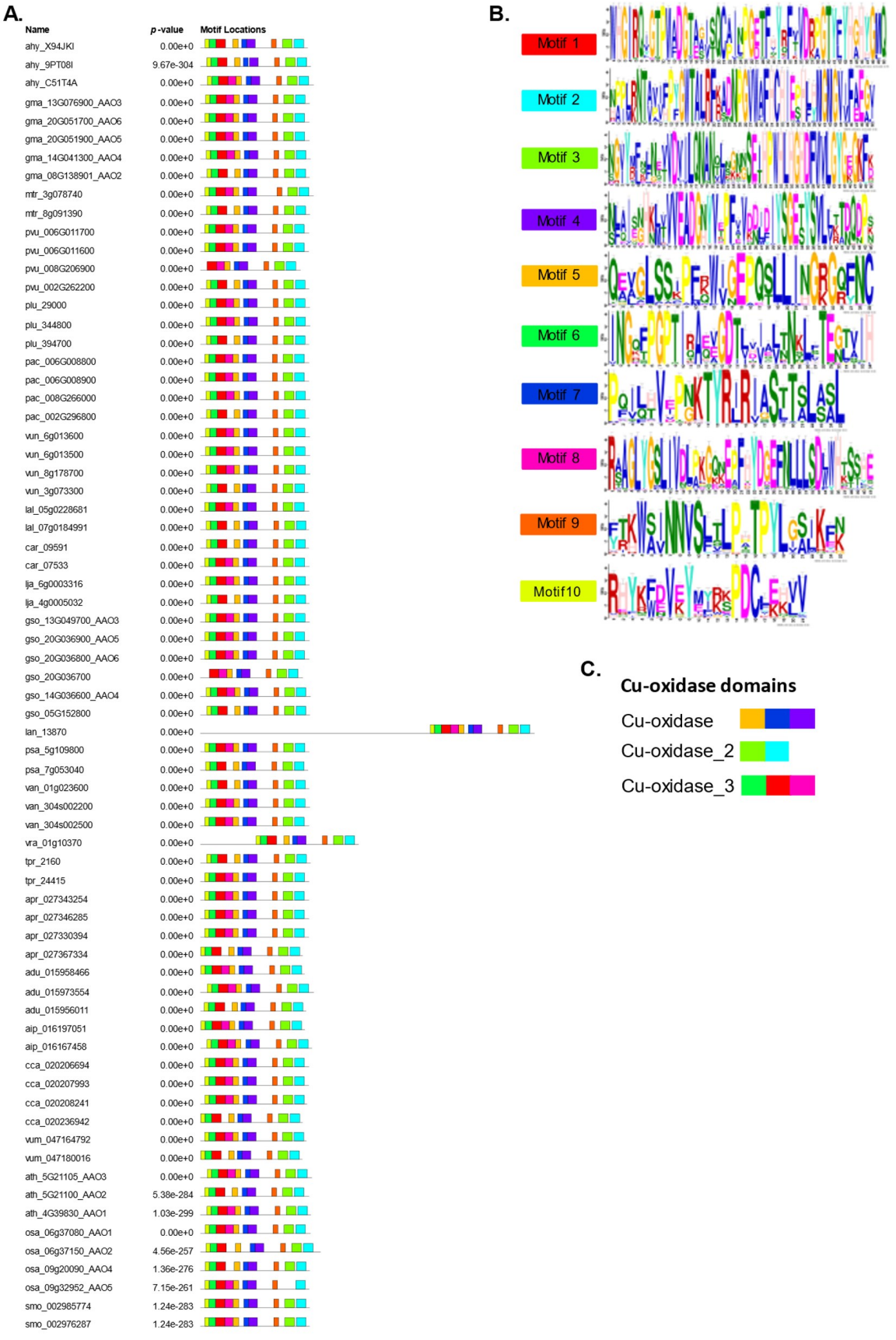


Fig. 2: *De novo* motif analysis of Ascorbate oxidase sequences from Fabaceae.The Fabaceae species were *Glycine max* (gma), *Glycine soja* (gso), *Arachis hypogaea* (ahy), *Phaseolus lunatus*

(plu), *Medicago truncatula* (mtr), *Phaseolus acutifolius* (pac), *Vigna unguiculata* (vun), *Lupinus albus* (lal), *Cicer arietinum* (car), *Lotus japonicus* (lja), *Phaseolus vulgaris* (pvu), *Lupinus angustifolius* (lan), *Pisum sativum* (psa), *Vigna angularis* (van), *Vigna radiata* (vra), *Trifolium pratense* (tpr), *Abrus precatorius* (apr), *Arachis duranensis* (adu), *Arachis ipaensis* (aip), *Cajanus cajan* (cca), *Vigna umbellate* (vum) and from *Arabidopsis thaliana* (ath), *Selaginella moellendorffii* (smo) and *Oryza sativa* (osa). **A)** Were found 10 motifs in the sequences analyzed. The MEME online tool was used to generate these results. **B)** The conserved motif are displayed in different coloured boxes, and the sequence information for each motif is displayed in the form of a seqlog. **C)** The Multicopper Oxidase Type 1 (Cu-oxidase), Multicopper Oxidase Type 2 (Cu-oxidase_2) and Multicopper Oxidase Type 3 (Cu-oxidase_3) domains are represented in the bars colored by the color of motifs that are present in each motif. Motifs 4, 5 and 7 are present in the Cu-oxidase domain, motifs 2 and 3 are present in Cu-oxidase_2 domain and motifs 6, 1 and 8, are present in the Cu-oxidase_3 domain. Data about the motifs 8, 9 and 10 are inconclusive about the domains.

To better understand the gene organization of *AAO*, especially from Fabaceae species, we performed a gene structure analysis (Fig. 3B), displayed in a phylogeny to see if the intron-exon organization is related to the phylogenetic position (Fig. 3A). The genes with bigger intronic sizes were in Group II, while the others were in Group I (Fig. 3A). The gene structure of *AAO*s within the Fabaceae family exhibits remarkable conservation, characterized by minimal variation among the number and size of exons and introns. AAO genes generally have four introns, except for Gma-*AAO4* and Pvu*006G01700*, which have three introns, and Ahy*9PT081,* which have five introns. The exon numbers are also conserved. Almost all analyzed Fabaceaes have five exons, except for Pvu*006G01700*, Gma-*AAO4* and Psa*7G053040*, which have four exons each, and Ahy*9PT081* with six exons (Fig. 3B). The evaluation of non-coding regions showed that up- and downstream sequences are predominantly smaller than 1 kilobase (kb) in size. However, it's noteworthy that only two genes displayed regulatory regions exceeding 1 kb in size, while three genes lacked discernible regulatory regions entirely (see Supplementary Fig. 2). Furthermore, a broader comparative analysis encompassing all AAO genes within the Fabaceae family revealed a similar trend, as depicted in Supplementary Fig. 3.

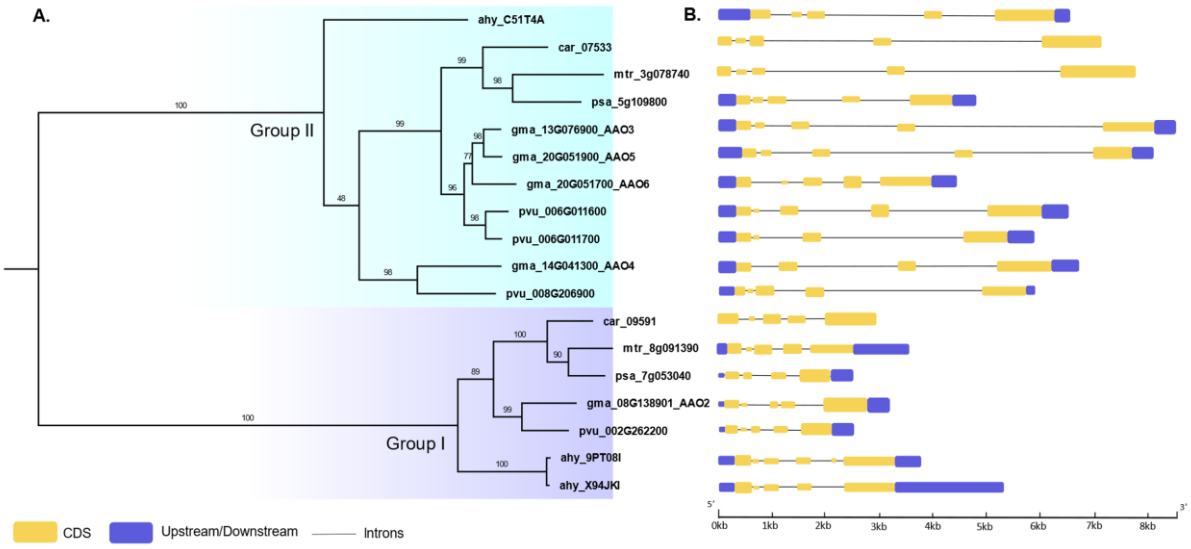


Fig. 3: The comparison of the gene structure of Fabaceae *AAO* in groups I and II.A) Phylogenetic tree of Fabaceae *AAO* genes. The phylogeny was generated using IQ-TREE web server with the LG+I+G4 substitution model and edited using FigTree software. **B)** Exon-intron structures of Fabaceae *AAO* genes. The black line represents introns, the CDS regions are in yellow, and the up/downstream regions are in blue. The gene size varied from 2.8kb to 8.9kb between analyzed *AAO* genes. The CDS regions are the most conserved, while the upstream/downstream and the intronic regions varied in size and number. The Gene Structure Display Server was used to infer the gene structure (gsds.gao-lab.org/). *G. max* (gma), *A. hypogaea* (ahy), *P. vulgaris* (pvu), *M. truncatula* (mtr), *C. arietinum* (car), *P. sativum* (psa).

3.3.4 Physical location of *AAO* genes on *Glycine* spp. chromosomes

In this study, we analyzed the genic distribution of *AAO* genes within the genomes of *Glycine* spp. Notably, *Glycine soja*, often called "wild soybean," is the ancestral precursor to *G. max*. Both species share an identical chromosome number, and a compelling observation emerges from our analysis: the vast majority of *AAO* genes in these species are localized to the same chromosomes. Gma-*AAO3* (Chr13: 17224503-17232484) and Gso-*AAO3* (Chr13: 17529729-17537967) are both positioned on chromosome 13. Furthermore, Gma-*AAO4* (Chr14: 3112468-3119143) and Gso-*AAO4* (Chr14: 3233536-3240107) are co-located on chromosome 14, while Gma-*AAO5* (Chr20: 11752947-11760913) Gso-*AAO5* (Chr20: 11774888-11782436), Gma-*AAO6* (Chr20: 11669323-11673938), and Gso-*AAO6* (Chr20: 11663108-11667377) are found on chromosome 20 (Fig. 4). However, Gma-*AAO2* (Chr08: 10718777-10721931) is situated on

chromosome 8, while *G. soja* have a gene on chromosome 5 (Chr05: 11774888-11782436) It's worth mentioning that *G. soja* exhibits an additional *AAO* gene, on chromosome 20 (Chr20: 11597536-11602725). This striking similarity in the chromosomal distribution of *AAO* genes between these closely related *Glycine* species underscores their genomic conservation in this context.

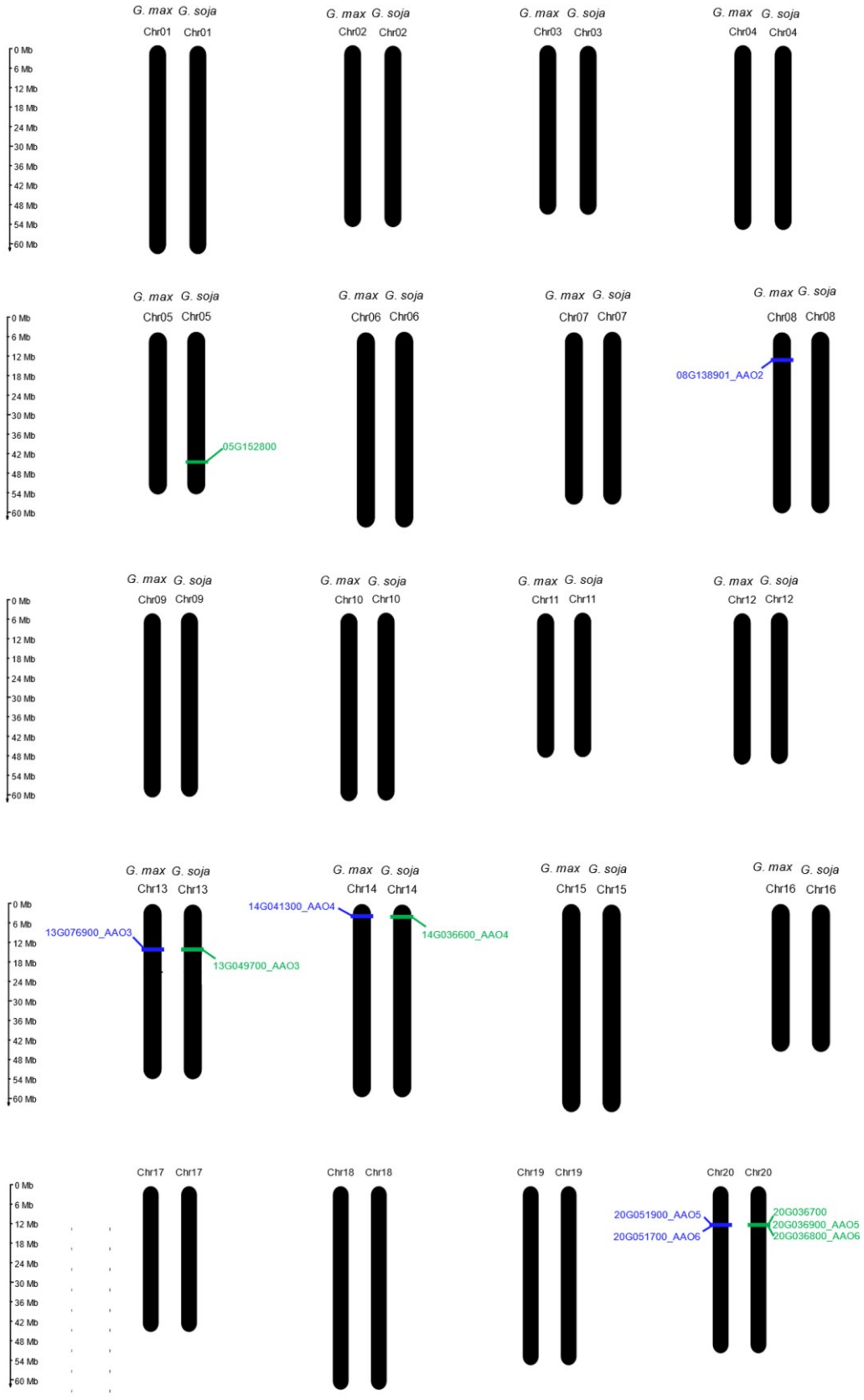


Fig. 4: Chromosome localization of *AAO* gene in *Glycine max* (Gma*)* and *Glycine Soja* (Gso). The five *G. max AAO*s are represented in blue and the six *G. soja* sequences are in green. *G. max*

and *G. soja AAO*s genes are distributed on chromosomes 13, 14, and 20, while *Gma_AAO2* is in chromosome 8, and *Gso_05G152800* is in chromosome 5. The figure was made with MG2C online tool with information from NCBI Genome Data Viewer.

3.3.5 *Cis*-acting Elements of *G. max AAO* Genes

An analysis of *cis*-acting elements in the promoter region was performed to understand better the functionalization of the *G. max AAO* gene family. In this analysis, we found binding motifs for 19 transcription factors in the forward strand of the five *G. max AAO* genes in a 2000 pb upstream region (Figura 5). The transcription factors are Barley B-Recombinant/Basic PENTACYSTEINE (BBR/BPC), Three Amino acid Loop Extension (TALE), MADS-box/ Intervening/Keratin-like-and C-domains (MICK-MADS), Cysteine 2 Histidine 2 (C2H2), DNA-binding with One Finger (DOF), Homeodomain Leucine Zipper (HD-Zip), Basic Leucine zipper (bZip), Basic Helix-Loop-Helix (BHLH), APETALA2/Ethylene Responsive Factor (AP2/ERF), NODULE INCEPTION-Like (Nin-Like), Auxin Response Factor (ARF), Myeloblastosis-Related (MYB-Related), Trihelix, WRKY, Lateral Organ Boundaries Domain Protein (LBD), (G/A/T/A) GATA, TEOSINTE BRANCHED 1- CYCLOIDEA- PCF1 (TCP), WUSCHEL-related homeobox (WOX), and NAM-ATAF-CUC-domain proteins (NAC). The transcription factor most encountered across *AAO* genes is BBR-BPC, not present just in Gma-*AAO4*. Otherwise, some TFs are present only in a few *AAO* genes. LBD, GATA, NAC, and TCP were present in Gma-*AAO2*; C2H2, and HD-Zip were present only in Gma-*AAO3*; MYB-related, WRKY, and Trihelix in Gma-*AAO4*; ARF in Gma-*AAO5*. The AP2/ERF motif are present in Gma-*AAO6* and Gma-*AAO2*, while the bZIP motif is present in Gma-*AAO5* and Gma-*AAO2*, and DOF in Gma-*AAO2*, and Gma-*AAO6* (Fig. 5; Supplementary Table 2). Among these, Gma-*AAO2* exhibited the highest abundance of TF binding motifs with eleven distinct motifs, and Gma-*AAO4* displayed the lowest abundance, harboring three different TF binding motifs. We found a total of 77 motifs, being the motifs for BBR/BPC the most numerous with 19 motifs, followed by 11 motifs for TALE, and 10 motifs to bZIP. The TFs that presented a smaller number of motifs were Nin-Like, ARF, Trihelix, WRKY, GATA, and NAC, with just one motif each (Supplementary Table 3).

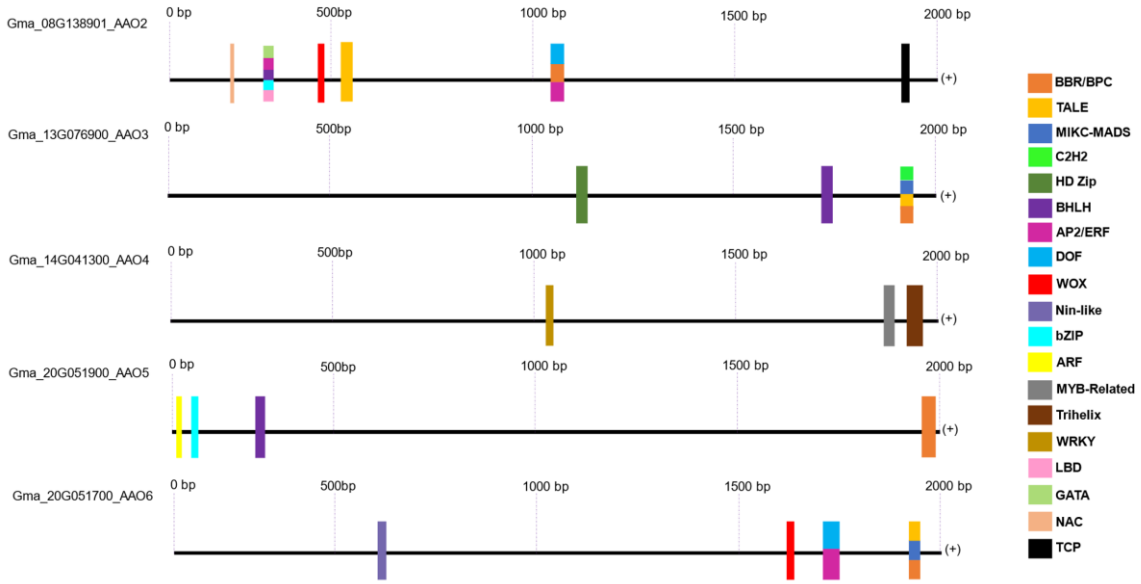


Fig. 5: *Glycine max* (Gma) *Ascorbate oxidase* (AAO) *cis*-regulatory motifs. Every motif is represented by a different color on the left legend, and the width of the columns represents approximately the transcription factors (TFs) motif sizes in nucleotides and their position on the upstream *AAO* region. The motifs have from 10 to 24 nucleotides per sequence. The information about *Glycine max Ascorbate oxidase cis*-regulatory motifs was provided by PlantRegMap online tool. Barley B-Recombinant/Basic PENTACYSTEINE (BBR/BPC), Three Amino acid Loop Extension (TALE), MADS-box/ Intervening/Keratin-like-and C-domains (MICK-MADS), Cysteine 2 Histidine 2 (C2H2), DNA-binding with One Finger (DOF), Homeodomain Leucine Zipper (HD-Zip), Basic Leucine zipper (bZip), Basic Helix-Loop-Helix (BHLH), APETALA2/Ethylene Responsive Factor (AP2/ERF), NODULE INCEPTION-Like (Nin-Like), Auxin Response Factor (ARF), Myeloblastosis-Related (MYB-Related), Trihelix, WRKY, Lateral Organ Boundaries Domain Protein (LBD), (G/A/T/A) GATA, TEOSINTE BRANCHED 1- CYCLOIDEA- PCF1 (TCP), WUSCHEL-related homeobox (WOX), and NAM-ATAF-CUC-domain proteins (NAC).

The identified transcription factors (TFs) are related in the literature with different plant biological functions (Fig. 6). In the case of *G. max*, nearly all transcription factors have already been associated with various processes, except for BBR/BPC, which has yet to be linked to *G. max*. TALE TF was identified in *G. max* in response to salt and drought stress, as well as DOF, and MYB-related. The HD-Zip TF is also related to salt and drought stress and nodule formation. Similarly, C2H2 is also associated with nodule formation, cold and drought stress, TCP with salt and biotic stresses, and nodule formation. Some TFs are related to seed and flower development, like MICK-MADS, and others just with nodule formation, like NIN-Like. Some TF are related to different stress and physiological situations. An example of this is the bZip TF, which is coupled with salt, drought,

and cold stress, seed and flower development, biotic stresses, nodule formation, light signal, abscisic acid (ABA), shade avoidance syndrome, root growing, and heavy metal stresses. Another TF related to various processes is the WOX TF, associated with heat, cold, salt and drought stresses and regeneration. Studies linked the bHLH TF with nodule growth, ammonium transport, biotic stress, ROS control, and iron homeostasis.

Prior studies have already established the connection between G. max's AP2/ERF genes and growth, auxin synthesis, seed development, and water stress. Trihelix are related to a variety of stresses in *G. max*, like flooding, salt, biotic, cold, and drought stresses. NAC genes in *G. max* are associated with salt and drought stresses, senescence, and lateral root formation. Besides that, this TF has the most extensive body of literature regarding their relationship with *G. max*. The literature about the WRKY TF and *G. max* show that this gene is related to heavy metal, drought, salt, and biotic stresses, as well as with development and flowering. The LBD TF is related to hormones and stresses for salt, drought, cold, and biotic stress in *G. max*. Other transcription factors are correlated to just a few processes, like ARF related to auxin and drought stress, GATA related to low nitrogen stress, chlorophyll synthesis, and drought stress.

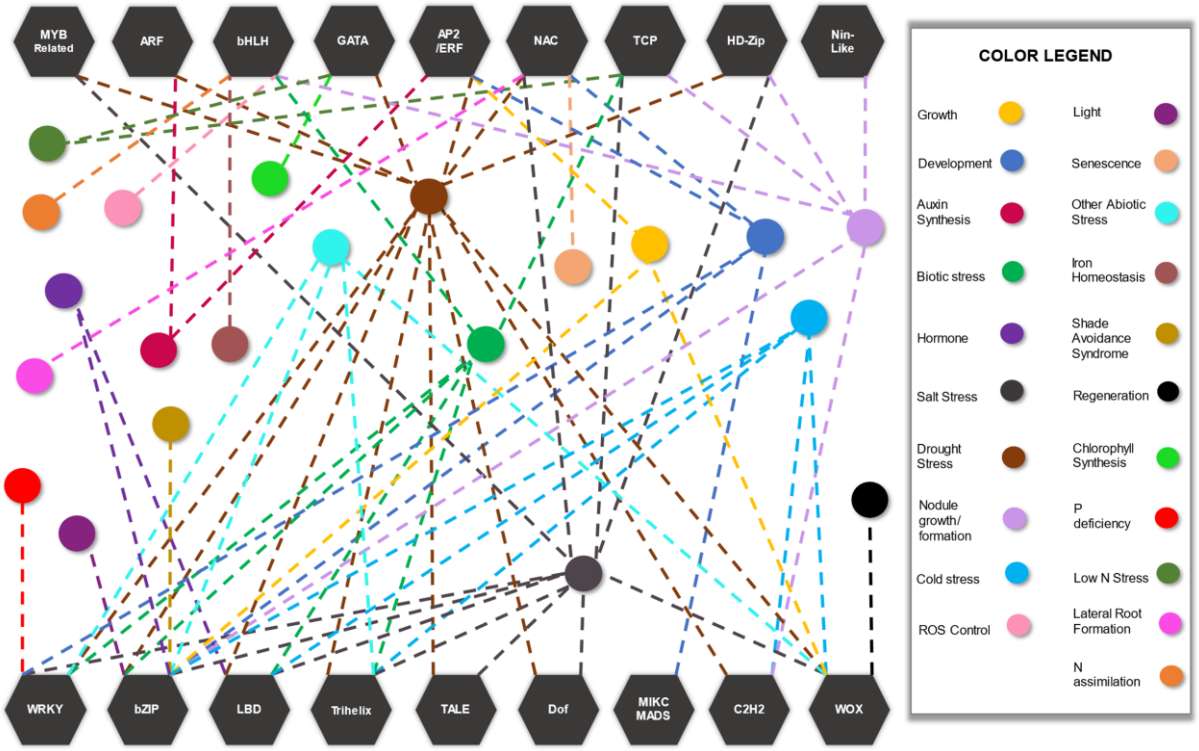


Fig. 6: Transcription factors (TFs) found in *Glycine max Ascorbate Oxidase* (Gma_*AAO*) upstream region, and their described functions in G. max. The left legend informs the representative color for each responsive situation, and the TFs are represented on the hexagons. Were found 18 TFs on Gma_*AAO*s with different functions, but the most common function found in TFs are responsive to drought and salt stress. The "other abiotic stress" legend referred to flooding stress, heavy metal stress and heat stress. The information about *Glycine max Ascorbate oxidase cis*-regulatory motifs was provided by PlantRegMap online tool and their respective functions were found in the literature. Three Amino acid Loop Extension (TALE); MADS-box; Intervening-, Keratin-like- and C-domains (MICK-MADS); Cysteine 2 Histidine 2 (C2H2); Cysteine 3 Histidine (C3H); DNA-binding with One Finger (Dof); Homeodomain Leucine Zipper (DH-Zip); Basic Leucine zipper (Bzip); Basic Helix-Loop-Helix (BHLH); APETALA2/Ethylene Responsive Factor (AP2/ERF); NODULE INCEPTION-Like (Nin-Like); Auxin Response Factor (ARF); Myeloblastosis-Related (MYB-Related); Lateral Organ Boundaries Domain Protein (LBD); TEOSINTE BRANCHED 1, CYCLOIDEA, PCF1 (TCP); WUSCHEL-related homeobox (WOX); NAM, ATAF, and CUC - domain proteins (NAC).

3.3.6 *In Silico* Expression Patterns of *G. max AAO* Genes

In silico expression analyses was performed and showed different expression patterns in *G. max AAO* genes under various biotic and abiotic stresses (Fig. 7). Differential expression patterns were discerned in response to various infection scenarios, yielding distinct regulatory outcomes across *AAO*s genes within *G. max* (Supplementary Table 4). Infection by *A. glycines* led to the downregulation of Gma-*AAO3* at 48 hours post inoculation (hpi) and Gma-*AAO4*

at 24 hpi. In contrast, infection by *B. diazoefficiens* resulted in the upregulation of Gma-*AAO5* expression at 5-7 days post-infection (dpi). Notably, *B. diazoefficiens* infection displayed dual effects: upregulating Gma-*AAO4* in nodules at 14-16 dpi while downregulating it in roots within the same timeframe. However, Gma-*AAO4* was upregulated in leaves inoculated with *B. japonicum*. Furthermore, upon infection by *Colletotrichum ilicicola*, Gma-*AAO6* expression was upregulated in hypocotyls, while infection of the radicle led to the downregulation of Gma-*AAO2*. *Fusarium oxysporum* elicited distinct expression patterns, with Gma-*AAO4* being upregulated at 72 hpi and Gma-*AAO2* upregulated at 96 hpi. Notably, during infection with one of the most prominent *G. max* pathogen, *Phakopsora pachyrhizi*, varying regulatory effects were evident across *AAO* genes. Gma-*AAO3* expression was downregulated, while Gma-*AAO5* and Gma-*AAO2* exhibited contrasting patterns of upregulation, particularly within the context of resistant cultivars (NIL-Rpp3). Distinctive patterns were also observed upon infection with *Phytophtora sojae*, whereby Gma-*AAO6* was upregulated between four and 14 dpi in both resistant (PI 449459) and non-resistant cultivars. In comparison, Gma-*AAO2* was downregulated in the resistant cultivar at 21 dpi. In response to Soybean Mosaic Virus infection, Gma-*AAO3* and Gma-*AAO4* expression was upregulated at specific time points after infection.

A spectrum of distinct expression profiles was observed among *G. max* AAO genes in abiotic stress responses. Under conditions of dehydration stress, a notable upregulation of Gma-*AAO2* is evident at 6-12- and 24-hours post-stress initiation. Intriguingly, this stress stimulus leads to a downregulation of Gma-*AAO3* expression at 12- and 24-hours. Meanwhile, extended periods of drought stress for six- and five-days yield downregulation of Gma-*AAO3* and Gma-*AAO5* in leaf and root tissue samples. Exposing the soybean plants to eight hours of heat stress resulted in a downregulation of Gma-*AAO6* expression. In salinity stress, durations of one- and two-hours prompt downregulation of Gma-*AAO4* expression, accompanied by an upregulation of Gma-*AAO3* in leaf samples. A contrasting expression pattern is observed under submergence stress, where Gma-*AAO3* experiences downregulation in root tissue samples, and Gma-*AAO2* displays an upregulated expression in leaves. These findings underscore the intricate and varied regulatory responses of *AAO* genes in the face of diverse biotic and abiotic stressors in *G. max*.

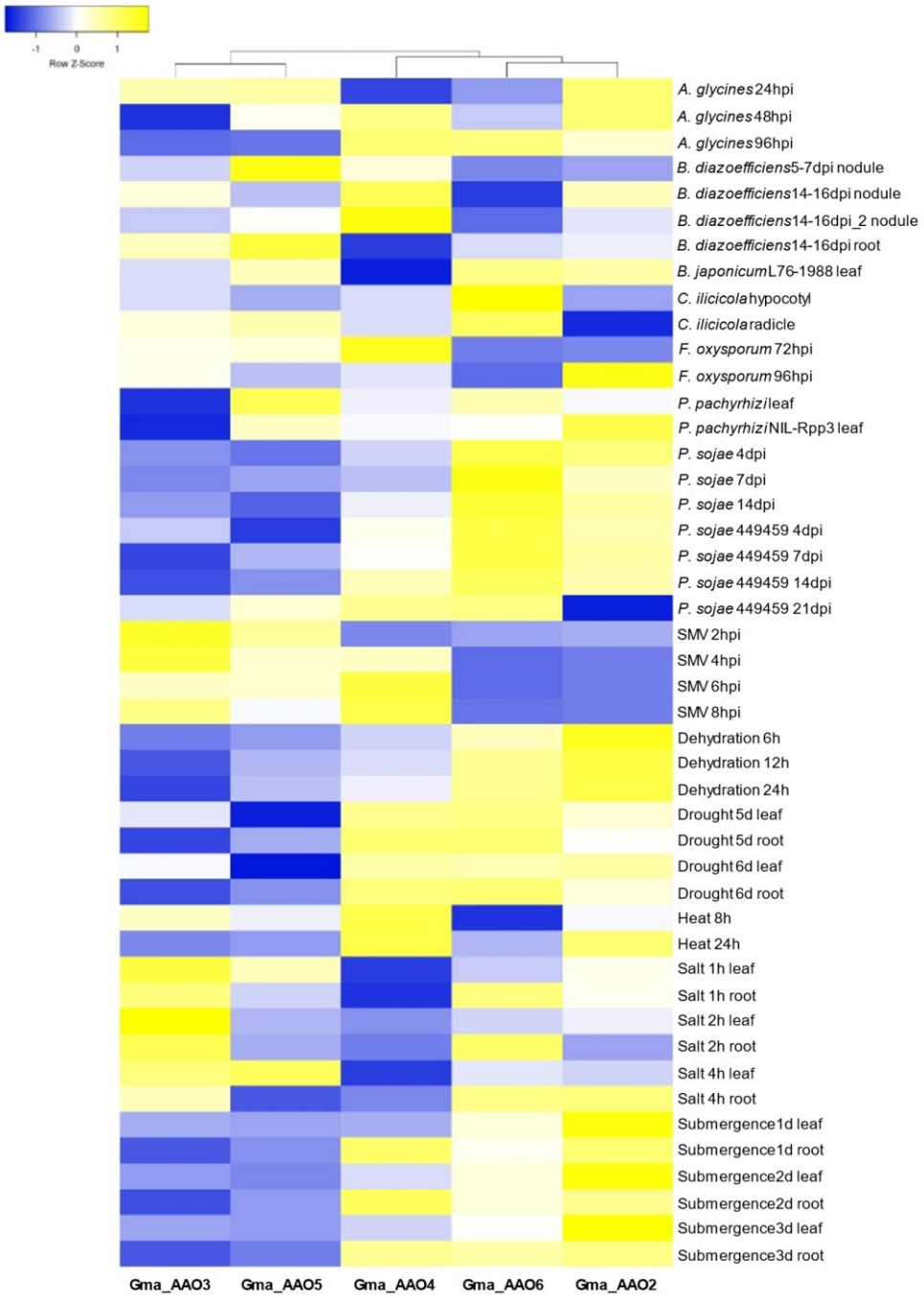


Fig. 7: The *in silico* expression patterns of *Glycine max AAO* in biotic and abiotic.

The colors represent gene expression levels that are either above (yellow), below (blue), or at the mean (white). Different stresses are mentioned in figure right, while the genes are mentioned below, and are clustered according to the similarity of the expression pattern. The heatmap was generated using the Heatmapper online tool and we manually selected the perturbations according to the Log2-ratio values from *Glycine max* mRNA seq data hosted in Genevestigator Plants software. The Log2-ratio values was considered significant in -2.5 to 2.5, and *p*-value ≤ 0,005. Soybean 449459 – resistant cultivar for *P. sojae.*

3.4 DISCUSSION

3.4.1 The search from Ascorbate Oxidase proteins

We did a massive search using BLASTp in the Phytozome, NCBI, and Ensembl Plants databases to identify the Ascorbate Oxidases from the Fabaceae species, an economically important family of plants. We used the well-characterized Gma-*AAO3* as the query for our searches. Despite the small size of the AAO protein family, the sequences are quite similar to the Multicopper Oxidases class of proteins (Nakamura et al. 2003; Batth et al. 2017). We collected more than 1000 sequences; after filtering analyses, 70 sequences were maintained for further analysis. We also successfully identified 49 novel AAO sequences across 16 distinct Fabaceae species. Until now, almost all the Fabaceae AAO putative sequences that we discovered had not been characterized in the literature.

The most recent investigation that performed the identification of AAO in *G. max*, *O. sativa*, and *A. thaliana* was conducted by Batth et al., 2017. Our findings exhibited partial agreement with this study. Specifically, in *A. thaliana*, we similarly identified three prospective *AAO* genes. It is worth noting that the acronyms employed by Batth et al. for these sequences has been maintained. Nevertheless, in the case of *G. max* and *O. sativa*, our analysis unveiled the presence of five and four *AAO* genes, respectively, which contradicts the findings of the previous work (Batth et al. 2017). In order to capture the *AAO* gene from *G. max*, we utilized the data accessible from the most recent *G. max* genome release (Wm82.a4.v1) within the Phytozome database. We could not find a correspondent sequence for Glyma_05G33470 (AAO1) and Glyma_20G12230 (AAO7) on this last release. Searching on Phytozome, we detected Gma-*AAO1* in the earlier Wm82.a2 release under the identifier Glyma_20G51600. However, upon investigating this particular identifier in the subsequent Wm82.a4.v1 release, a sequence labeled as Glyma.20G051602 was retrieved. Notably, this sequence lacked the principal attributes characteristic of AAO proteins, including the three Cu-oxidase domains and the conserved motifs (Nakamura et al. 2003). We hypothesize that the disparity between our findings and those of Batth, et al., 2017, could potentially arise from modifications in the latest versions of the genomes.

3.4.2 The evolutionary history of the Fabaceae AAO explains the preservation of genetic and protein structures.

Using the ML methodology, we provide a consistent phylogeny of the AAO protein family in 21 Fabaceae species. The AAOs clustered into groups designated as Group I and Group II. The Group II subsequently divide into Groups IIA, and IIB, characterized exclusively by the presence of sequences affiliated with the Fabaceae family. Even though a few recent studies on the evolution of AAO were published (Batth et al. 2017; Pan et al. 2019) they do not relate the evolutionary history of this enzyme in a plant family. Our phylogeny from Fabaceae AAOs provides a glimpse into the evolutionary story of these genes from the perspective of this plant family.

The first demonstration of the AAO protein's evolution was presented in the study by Sanmartin et al. in 2007. This study primarily centered on the AAOs of *C. melo*. Yet, it established an evolutionary linkage with three species that are integral to our investigation: *M. truncatula*, *A. thaliana*, and *O. sativa*. After that study, De Tullio et al., 2013, performed a phylogeny with a wider approach (De Tullio et al. 2013). They employed sequences extracted from a variety of plant species, with a notable focus on sequences sourced from *G. max*, *M. truncatula*, *P. sativum*, *L. japonicus*, *A. thaliana*, *O. sativa*, and *S. moellendorffii*, likewise utilized by us. Also, a few recent studies have been published showing AAO evolution from a single species perspective. Like in *C. sinensis* (Tian et al. 2019), *G. hirsutum* (Pan et al. 2019), *Z. mays* (Wu et al. 2021), *B. vulgaris* (Skorupa et al. 2022), and *A. nanus* (Zhu et al. 2023). *A. nanus* also belongs to the Fabaceae family; however, the absence of available sequences during our phylogeny construction precluded their inclusion in our analysis.

Our phylogeny was consistent with previous studies, which found that the analyzed genes were divided into two main groups (Batth et al. 2017; Pan et al. 2019). In those studies, Gma-AAO2, Ath-AAO1, and Osa-AAO4 and Osa-AAO5, were grouped in a cluster, whereas Gma-AAO3, Gma-AAO4, Gma-AAO5, and Gma-AAO6, Ath-AAO2 and Ath-AAO3, and Osa-AAO1 and Osa-AAO2, were grouped in another. Notably, an intriguing observation was the presence of all Fabaceae AAO sequences within Groups I, IIA, and IIB. This phenomenon can be explained by a duplication event that predates the emergence and evolutionary divergence of the Fabaceae plant family. Indeed, two ancestral whole genome

duplication events took place in the common ancestor of seed plants and angiosperms, respectively, and in the Fabaceae ancestor (Jiao et al. 2011; Zhao et al. 2020). These events may explain why almost all analyzed species had at least two AAO genes. A notable observation arises from the circumstance that the monocot sequences (from *O. sativa*) exhibited their presence in both Group I and Group II, associated with AAOs from the Fabaceae family. This intriguing finding suggests that the ancestral *AAO* genes underwent a division and subsequent independent evolution before the divergence of monocotyledons and dicotyledons (Tian et al. 2019).

Furthermore, our founds also agreed with the species tree (Supplementary Fig. 1), particularly when examining the relationships among *C. arietinum*, *M. truncatula*, *P. sativum*, and *T. pratense*. These findings not only support our tree topology but also offer evidence that the evolutionary trajectory of AAOs is consistent with the evolutionary pattern observed in Fabaceae species.

3.4.3 Conserved protein and gene structure of Fabaceae Ascorbate Oxidases

Understanding the sequence pattern of a protein within a plant family holds particular significance for characterizing, identifying, and assessing the level of conservation across sequences. AAOs are already thoroughly researched at the molecular level (Ohkawa et al. 1989; Batth et al. 2017; Wu et al. 2021; Zhu et al. 2023). Owing to this, the significance of the three Multicopper oxidase domains (Cu-oxidase, Cu-oxidase_2 and Cu-oxidase_3) for both the oxidation function of AAO proteins is already well established (Messerschmidt et al. 1992). Here, we show, through a *de novo* motif analysis, that almost all AAO from the Fabaceae family contains ten conserved motifs, some belonging to the three Cu-oxidase domains. Despite the high level of conservation, a few sequences had a lack of motifs. However, it is important to emphasize that the missing motifs occur mainly in the initial portion of the sequences, which comprise the Cu-oxidase 3 domain. These instances suggest that these sequences might have experienced the loss of specific nucleotides throughout evolution, consequently hindering the complete motif's identification. However, in our alignments, these sequences are still perfectly aligned with the other sequences, demonstrating discrete changes. A study which investigated the motif pattern between *O. sativa*, *A. thaliana*, *Z. mays*,

S. bicolor, and *G. max* AAOs demonstrated that the motif 3 showed resemblance to the Cu-oxidase domain, the motif 5 to the Cu-oxidase_2 domain, and the motifs 1 and 4 to the Cu-oxidase_3 domain, with the motif 6 showing resemblance with the Cu-oxidase domains 1 and 3 (Batth et al. 2017). Tian et al. (2019) found nine motifs from *C. sinensis* similar to the Cu-oxidase domains. Our findings corroborate with these authors who also found motifs comprehended in the Cu-oxidase domains, with a similar pattern of distribution in the motifs (Batth et al. 2017; Tian et al. 2019; Zhu *et al*. 2023). The exception was only in the Wu et al. study, which set and found 20 motifs in *Z. mays* (Wu et al. 2021).

Exon-intron organization comparisons are critical for understanding gene structure and organization rules, protein functionality, and evolutionary changes across species (Wang, Y. et al. 2013; Turchetto-Zolet et al. 2016). Our genic structure analyses also showed great conservation between *AAO* sequences from Fabaceae species. The genes were displayed in a phylogenetic tree, separating the sequences between two groups. Group II sequences present longer introns than the Group I sequences (Fig. 3). This phenomenon is most likely explained by intron reorganization in the sequences during evolution, which can result from processes such as transposable element insertion, nucleotide substitutions, or insertion/deletions, leading to the gain or loss of nucleotides (Roy and Penny 2007). Until now, few studies have presented information about the gene structure of *AAO* genes, and some of these studies exhibited sequences with notable disparities in terms of exon-intron counts, sizes, and the presence/absence of untranslated regions (UTR), in comparison to our analysis (Diallinas et al. 1997; Pan et al. 2019; Wu et al. 2021; Skorupa et al. 2022; Zhu et al. 2023). The exon-intron pattern observed in our analysis is similar to that found for the *AAO* genes in *G. hirsutum*, which typically exhibit four to five exons and four introns (Pan et al. 2019). In this study, they also conducted a phylogenetic analysis that revealed a clustering of these genes into two distinct groups, further distinguished by variations in intron sizes, resulting in a cluster divided into two groups differentiated by the intron sizes (Pan et al. 2019). Examining gene structure in *Z. mays AAO* genes reveals a consistent pattern of a single exon and one intron across all sequences, with the absence of a UTR region in some (Wu et al. 2021). Moreover, the gene size is notably smaller than the typical *AAO* size, measuring only 2kb (Wu et al. 2021). The analysis in *B. vulgaris* also highlights variations in

the count and size of exon-intron structures, along with differences in UTR' regions. However, certain genes, like Bvu-*AAO1*, Bvu-*AAO2*, Bvu-*AAO5*, and Bvu-*AAO9,* exhibit similarities with our analysis, featuring four to five exons and comparable UTR' regions (Skorupa et al. 2021). A comparable exon-intron pattern is also evident in *A. nanus AAO* genes, like Ana-*AAO3*, Ana-*AAO5*, Ana-*AAO11*, Ana-*AAO13*, and Ana-*AAO15* (Zhu et al. 2023).

Further investigation into comparative gene structures within *AAO* sequences is necessary. However, our studies helped shed light on this inquiry for being the only study to analyze the AAO gene structures across multiple species until now.

3.4.4 Chromosomal distribution between *Glycine* spp.

We presented an analysis of the chromosomal positioning of *Glycine* spp. *AAO* genes, highlighting the remarkable conservation in their gene locations. The wild soybean, *G. soja*, and its descendent, *G. max*, preserve a lot of similarities, including the chromosomal structure. However, we know from studies comparing the genomes of *G. soja* and *G. max* that, despite *G. max* domestication process preserving genetic diversity, many changes occurred, including duplications and deletions, in addition to identifying many variants of a single nucleotide (SNVs) and copy number (CNVs) (Sedivy et al. 2017). However, no study so far has reported on *AAO* in *G. soja*. We believe that in the case of the *AAO* genes, *G. max* has lost a copy compared to *G. soja*. The analysis offers insights into the chromosomal organization, shedding light on whether the *AAO* genes in *G. max* and *G. soja* originated from genome duplications or arose through tandem duplication events. Batth et al., 2019 suggest a tandem duplication between Gma-*AAO5* and Gma-*AAO6*, which could also be observed in our results of Gso-*AAO5*, Gso-*AAO6* and *Gso-20G036700*. Other studies also present *AAO* genes in tandem duplications in distinct species, like *G. hirsutum* (Pan et al. 2019), *Z. mays* (Wu et al. 2021), *B. vulgaris* (Skorupa et al. 2022), and *A. nanus* (Zhu et al. 2023). Comparative genomic analysis of closely related species revealed that tandem duplication is a primary mechanism for generating new genes, particularly those clustered within a gene family (Fan et al. 2008).

Nonetheless, duplicated genes may either acquire a novel function or retain their original function. Various degrees of functional redundancy are to be

expected following genetic duplication. However, redundant functions may exist not because they are useful, but because there was insufficient time for them to be lost (Panchy et al. 2016). It remains to be seen whether new functions have been acquired in the context of duplicated *AAO* genes.

3.4.5 Possible function for *G. max AAO*s showed by the *cis*-regulatory elements and gene expression analyses

Cis-regulatory elements are DNA sequence motifs targeted by sequence-specific TFs, which can regulate the transcriptional pattern of the genes. The analysis of the TF pattern could serve as a valuable tool for elucidating gene functions and discerning the specific cellular contexts in which a gene is active (Marand et al. 2023). In our study, we conducted a new investigation into the *cis*-regulatory elements within the five *AAO* genes of *G. max*. Additionally, we used the RNA-seq expression data of the *AAO*s from *G. max*, under biotic and abiotic stresses, to verify whether there was a correlation between the *cis*-regulatory elements and the expression pattern of these genes.

It is important to notice that BBR-BPC was the predominant TF predicted to bind to the promoter region of most *AAO* genes, except for Gmax-*AAO5*. Unfortunately, the function of this TF was not characterized in *G. max*, but it has established associations with growth and developmental processes in *A. thaliana* (Monfared et al. 2011) (Lee et al. 2022). However, other transcription factors were related to growth and development in *G. max*, like MICK-MADS (Fan et al. 2013), WKRY (Yang et al. 2016), NAC (Yang et al. 2019), WOX (Hao et al. 2019), bZIP (Hu et al. 2022a), and AP2/ERF (Xu et al. 2022).

A well-established connection between growth and developmental processes and *AAO* is corroborated by numerous prior investigations, in *C. melo*, *P. sativum*, *G. hirsutum*, *O. sativa*, *A. thaliana*, *Z. mays*, *S. bicolor* and in *G. max* (Diallinas et al. 1997; Al-Madhoun et al. 2003; Sanmartin et al. 2006; Xin; Tao; Li, 2016; Li et al. 2017; Batth et al. 2017). These findings helped to corroborate the relation of *G. max AAO*s with growth and development processes.

While there is a lack of research dedicated to analyzing the *AAO* regulatory region in *G. max*, analogous studies have explored this region in *AAO* genes among other plant species. Asao et al. (2013), found a wound-responsive *cis*-element in the promoter region of cucumber *AAO1* (Asao et al. 2003).

Balestrini et al., examined the promoter region of *L. japonicus AAO1* and identified motifs, including "AAAGAT" and "CTCTT," which have been previously recognized as nodulin consensus sequences (Balestrini et al. 2012). Intriguingly, Gma-*AAO3*, Gma-*AAO6*, Gma-*AAO5*, and Gma-*AAO2* were found to harbor transcription factor motifs associated with nodule formation, such as bHLH, HD-Zip, TCP, C2H2, Nin-Like and bZIP motifs (Chiasson et al. 2014; Yuan et al. 2018; Damodaran et al. 2019; Yue et al. 2023; Kim et al. 2023). The gene expression analyses of Gma-*AAO5*, during the beneficial interaction with the nitrogen-fixing symbiont *Bradyrhizobium diazoefficiens*, was upregulated in nodules, suggesting that it could be associated with nodule formation (Sandhu et al. 2023). Ascorbate oxidase association with nodule formation and nitrogen fixation have been already explored. The researchers identified and partially characterized a symbiotically induced *AAO* gene that was overexpressed in *L. japonicus* during interactions with either the nitrogen fixing *Mesorhizobium loti* or the fungus *Gigaspora margarita* (Balestrini et al. 2012). This relation probably occur because even low oxygen concentrations inactivate the nitrogen-fixing enzyme nitrogenase, and tight control over oxygen diffusion and the formation of ROS is required in plant-microbe interactions (Layzell; Hunt; Palmer, 1990; Balestrini et al. 2012). A recent study examining *AAO* genes in *B. vulgaris* identified binding sites for the TF bZIP, MYB, WRKY, NAC and bHLH in the promoter region, which are related to salinity and/or drought response (Skorupa et al. 2022). Interestingly, a significant proportion of the transcription factors associated with the six *G. max AAO*s identified were found to have connections with drought and salt stress responses in this plant. The TFs MYB-related (Du et al. 2018), HD-Zip (Chen et al. 2014; Belamkar et al. 2014), WKRY (Song et al. 2016; Shi et al. 2018), bZIP (Zhang et al. 2018; Yue et al. 2023), Trihelix, DOF (Wei et al. 2023) and WOX (Hao et al. 2019), are related with both stresses in *G. max*. Otherwise, some TFs are related only with drought stress, like ARF (Van Ha et al. 2013), GATA (Ouyang et al. 2022), and AP2/ERF (Wang et al. 2022), and TCP (Zhang et al. 2023), only with salt stress. In our study, gene expression data revealed that Gma-*AAO3* was upregulated in response to salinity stress, whereas Gma-*AAO4* showed the opposite regulation under the same stress conditions. Moreover, both Gma-*AAO5* and Gma-*AAO3* were downregulated during drought stress. This dynamic gene expression pattern aligns with findings from a study conducted by Batth et al., who employed

microarray and real-time PCR assays in *O. sativa* and *Z. mays* (Batth et al. 2017). More recently, promoter analyses of *A. nanus AAO*s showed the predicted *cis*-acting elements involved in low-temperature, drought stress, and hormones such as ABA, gibberellin, salicylic acid (SA), and auxin responses (Zhu et al. 2023). In our study, some TFs are also related to those hormones. *G. max* LBD and bZIP genes were induced with ABA, and SA treatments (Yang, H. et al. 2017; He et al. 2020; (Chai et al. 2022), and the ARF and AP2/ERF TFs, were related to auxin (Wang et al. 2015; Xu et al. 2022).

In our search, we also found transcription factors associated with biotic stresses. The WRKY (Zhang et al. 2012), bZIP (Alves et al. 2015), bHLH (Cheng et al. 2018), Trihelix (Liu et al. 2020), TCP (Fan et al. 2022), and LBD (Feng et al. 2022) transcription factors were found to be linked to biotic stress responses in *G. max*. Notably, *AAO*6 was the only *G. max* AAO that lacked the *cis* elements associated with these TFs linked to biotic stress. In contrast, the *AAO* most frequently associated with these biotic stress-related TF motifs was Gma-*AAO*2. Moreover, the TF with more evidence for different types of pathogens, like the Soybean Cyst Nematode, *P. sojae* and *P. pachyrhizi*, was the WRKY, which has corresponding cis-elements present only in *Gma-AAO4* (Bencke-malato et al., 2014; Yang, Y. et al., 2017; Fan, S. et al., 2017; DONG et al., 2019). Interestingly, the gene expression pattern of Gma-*AAO2* was upregulated in *F. oxysporum* at 96hpi, and *P. pachyrhizi* infection in the resistant cultivar. The *AAO4* from *C. sativus* interacts with the movement protein (MP) of the cucumber mosaic virus (CMV), facilitating the initiation of infection (Kumari et al. 2016). In our analyses, Gma-*AAO3* was upregulated in the SMV infection in 2hpi and 4hpi. This suggests that not only Gma-*AAO2* but also Gma-*AAO3* and Gma-*AAO4* could have an important role in biotic stress situations.

These findings illustrate the commonalities observed among *AAO* genes across various species and the wide array of their functions. Moreover, the *cis*-regulatory elements and gene expression analyses can provide valuable insights for defining stress conditions in functional and expression analyses for *G. max AAO* and those within the Fabaceae family.

4.5 CONCLUSION

Our analysis in this study led us to conclude that Fabaceae AAOs represent a compact family of proteins, characterized by a high degree of conservation and similarity in terms of both gene structure and protein motifs and domains. Phylogenetic analysis revealed that the AAO sequences could be categorized into two main groups, with high support UFbootstrap values, and further divided into groups specific to Fabaceae AAOs. Moreover, when considering data on *cis*-regulatory elements and conducting *in silico* expression analyses, we hypothesize that the Gma-*AAO2*, Gma-*AAO3*, and Gma-*AAO4* genes are implicated in biotic stress responses, while the Gma-*AAO5* gene may be linked to nodulation. Our work has successfully characterized Fabaceae AAOs and contributed to our understanding of the potential roles of certain *AAO* genes in *G. max*. As part of future perspectives, we have identified genes exhibiting differential expression, which can be selected as biomarkers or targets for further analyses, not only in *G. max* but also in other Fabaceae species. This endeavor aims to enhance our comprehension and subsequently improve the attributes of these plants.

4.6 REFERENCES

Abdelgawad KF, El-Mogy MM, Mohamed MIA, et al (2019) Increasing ascorbic acid content and salinity tolerance of cherry tomato plants by suppressed expression of the ascorbate oxidase gene. Agronomy 9:. https://doi.org/10.3390/agronomy9020051

Al-Madhoun AS, Sanmartin M, Kanellis AK (2003) Expression of ascorbate oxidase isoenzymes in cucurbits and during development and ripening of melon fruit. Postharvest Biol Technol 27:137–146. https://doi.org/10.1016/S0925-5214(02)00090-X

Alves MS, Soares ZG, Vidigal PMP, et al (2015) Differential expression of four soybean bZIP genes during Phakopsora pachyrhizi infection. Funct Integr Genomics 15:685–696. https://doi.org/10.1007/S10142-015-0445-0

Asao H, Yoshida K, Nishi Y, Shinmyo A (2003) Wound-responsive cis-element in the 5′-upstream region of cucumber ascorbate oxidase gene. Biosci Biotechnol Biochem 67:271–277. https://doi.org/10.1271/bbb.67.271

Bailey TL, Johnson J, Grant CE, Noble WS (2015) The MEME Suite. Nucleic Acids Res 43:W39–W49. https://doi.org/10.1093/NAR/GKV416

Balestrini R, Ott T, Güther M, et al (2012) Ascorbate oxidase: the unexpected involvement of a “wasteful enzyme” in the symbioses with nitrogen-fixing bacteria and arbuscular mycorrhizal fungi. Plant Physiol Biochem PPB 59:71–79. https://doi.org/10.1016/J.PLAPHY.2012.07.006

Batth R, Singh K, Kumari S, Mustafiz A (2017) Transcript profiling reveals the presence of abiotic stress and developmental stage specific ascorbate oxidase genes in plants. Front Plant Sci 8:198. https://doi.org/10.3389/FPLS.2017.00198/BIBTEX

Belamkar V, Weeks NT, Bharti AK, et al (2014) Comprehensive characterization and RNA-Seq profiling of the HD-Zip transcription factor family in soybean (Glycine max) during dehydration and salt stress. BMC Genomics 15:1–25. https://doi.org/10.1186/1471-2164-15-950/TABLES/4

Bencke-Malato M, Cabreira C, Wiebke-Strohm B, et al (2014) Genome-wide annotation of the soybean WRKY family and functional characterization of genes involved in response to Phakopsora pachyrhizi infection. BMC Plant Biol 14:1–18. https://doi.org/10.1186/S12870-014-0236-0

Bertioli DJ, Cannon SB, Froenicke L, et al (2016) The genome sequences of Arachis duranensis and Arachis ipaensis, the diploid ancestors of cultivated peanut. Nat Genet 2016 484 48:438–446. https://doi.org/10.1038/ng.3517

Bertioli DJ, Jenkins J, Clevenger J, et al (2019) The genome sequence of segmental allotetraploid peanut Arachis hypogaea. Nat Genet 2019 515 51:877–884. https://doi.org/10.1038/s41588-019-0405-z

Chai M, Fan R, Huang Y, et al (2022) GmbZIP152, a Soybean bZIP Transcription Factor, Confers Multiple Biotic and Abiotic Stress Responses in Plant. Int J Mol Sci 23:. https://doi.org/10.3390/IJMS231810935

Chatzopoulou F, Sanmartin M, Mellidou I, et al (2020) Silencing of ascorbate oxidase results in reduced growth, altered ascorbic acid levels and ripening pattern in melon fruit. Plant Physiol Biochem 156:291–303. https://doi.org/10.1016/j.plaphy.2020.08.040

Chen X, Chen Z, Zhao H, et al (2014) Genome-Wide Analysis of Soybean HD-Zip Gene Family and Expression Profiling under Salinity and Drought Treatments. PLoS One 9:e87156. https://doi.org/10.1371/JOURNAL.PONE.0087156

Cheng Q, Dong L, Gao T, et al (2018) The bHLH transcription factor GmPIB1 facilitates resistance to Phytophthora sojae in Glycine max. J Exp Bot 69:2527–2541. https://doi.org/10.1093/JXB/ERY103

Chiasson DM, Loughlin PC, Mazurkiewicz D, et al (2014) Soybean SAT1 (Symbiotic Ammonium Transporter 1) encodes a bHLH transcription factor involved in nodule growth and NH4+ transport. Proc Natl Acad Sci U S A 111:4814–4819. https://doi.org/10.1073/PNAS.1312801111

Damodaran S, Dubois A, Xie J, et al (2019) GmZPR3d Interacts with GmHD-ZIP III Proteins and Regulates Soybean Root and Nodule Vascular Development. Int J Mol Sci 2019, Vol 20, Page 827 20:827. https://doi.org/10.3390/IJMS20040827

De Tullio MC, Guether M, Balestrini R (2013) Ascorbate oxidase is the potential conductor of a symphony of signaling pathways. Plant Signal Behav 8:. https://doi.org/10.4161/PSB.23213

De Vega JJ, Ayling S, Hegarty M, et al (2015) Red clover (Trifolium pratense L.) draft genome provides a platform for trait improvement. Sci Reports 2015 51 5:1–10. https://doi.org/10.1038/srep17394

Di Venere A, Nicolai E, Rosato N, et al (2011) Characterization of monomeric

substates of ascorbate oxidase. FEBS J 278:1585–1593. https://doi.org/10.1111/J.1742-4658.2011.08084.X

Diallinas G, Pateraki I, Sanmartin M, et al (1997) Melon ascorbate oxidase: Cloning of a multigene family, induction during fruit development and repression by wounding. Plant Mol Biol 34:759–770. https://doi.org/10.1023/A:1005851527227

Dong H, Tan J, Li M, et al (2019) Transcriptome analysis of soybean WRKY TFs in response to Peronospora manshurica infection. Genomics 111:1412–1422. https://doi.org/10.1016/J.YGENO.2018.09.014

Du YT, Zhao MJ, Wang CT, et al (2018) Identification and characterization of GmMYB118 responses to drought and salt stress. BMC Plant Biol 18:1–18. https://doi.org/10.1186/S12870-018-1551-7/FIGURES/10

Esaka M, Fujisawa K, Goto M, Kisu Y (1992) Regulation of ascorbate oxidase expression in pumpkin by auxin and copper. Plant Physiol 100:231–237. https://doi.org/10.1104/pp.100.1.231

Fan C, Chen Y, Long M (2008) Recurrent tandem gene duplication gave rise to functionally divergent genes in Drosophila. Mol Biol Evol 25:1451–1458. https://doi.org/10.1093/molbev/msn089

Fan CM, Wang X, Wang YW, et al (2013) Genome-Wide Expression Analysis of Soybean MADS Genes Showing Potential Function in the Seed Development. PLoS One 8:e62288. https://doi.org/10.1371/JOURNAL.PONE.0062288

Fan S, Dong L, Han D, et al (2017) GmWRKY31 and GmHDL56 Enhances Resistance to Phytophthora sojae by Regulating Defense-Related Gene Expression in Soybean. Front Plant Sci 8:. https://doi.org/10.3389/FPLS.2017.00781

Fan S, Zhang Z, Song Y, et al (2022) CRISPR/Cas9-mediated targeted mutagenesis of GmTCP19L increasing susceptibility to Phytophthora sojae in soybean. PLoS One 17:e0267502. https://doi.org/10.1371/JOURNAL.PONE.0267502

Farvardin A, González-hernández AI, Llorens E, et al (2020) The Apoplast: A Key Player in Plant Survival. Antioxidants 2020, Vol 9, Page 604 9:604. https://doi.org/10.3390/ANTIOX9070604

Farver O, Wherland S, Pecht I (1994) Intramolecular electron transfer in ascorbate oxidase is enhanced in the presence of oxygen. J Biol Chem 269:22933–

22936. https://doi.org/10.1016/s0021-9258(17)31598-3

Feng S, Shi J, Hu Y, et al (2022) Genome-Wide Analysis of Soybean Lateral Organ Boundaries Domain Gene Family Reveals the Role in Phytophthora Root and Stem Rot. Front Plant Sci 13:865165. https://doi.org/10.3389/FPLS.2022.865165/BIBTEX

Foyer CH, Kyndt T, Hancock RD (2020) Vitamin C in Plants: Novel Concepts, New Perspectives, and Outstanding Issues. Antioxidants Redox Signal 32:463–485. https://doi.org/10.1089/ars.2019.7819

Foyer CH, Noctor G (2005) Oxidant and antioxidant signalling in plants: A re-evaluation of the concept of oxidative stress in a physiological context. Plant, Cell Environ 28:1056–1071. https://doi.org/10.1111/j.1365-3040.2005.01327.x

Garchery C, Gest N, Do PT, et al (2013) A diminution in ascorbate oxidase activity affects carbon allocation and improves yield in tomato under water deficit. Plant Cell Environ 36:159–175. https://doi.org/10.1111/J.1365-3040.2012.02564.X

Garcia T, Duitama J, Zullo SS, et al (2021) Comprehensive genomic resources related to domestication and crop improvement traits in Lima bean. Nat Commun 2021 121 12:1–17. https://doi.org/10.1038/s41467-021-20921-1

Góralczyk-Bińkowska A, Jasińska A, Długoński J (2019) Characteristics And Use Of Multicopper Oxidases Enzymes. Postępy Mikrobiol - Adv Microbiol 58:7–18. https://doi.org/10.21307/pm-2019.58.1.007

Hane JK, Ming Y, Kamphuis LG, et al (2017) A comprehensive draft genome sequence for lupin (Lupinus angustifolius), an emerging health food: insights into plant–microbe interactions and legume evolution. Plant Biotechnol J 15:318–330. https://doi.org/10.1111/PBI.12615

Hao Q, Zhang L, Yang Y, et al (2019) Genome-Wide Analysis of the WOX Gene Family and Function Exploration of GmWOX18 in Soybean. Plants 2019, Vol 8, Page 215 8:215. https://doi.org/10.3390/PLANTS8070215

He Q, Cai H, Bai M, et al (2020) A Soybean bZIP Transcription Factor GmbZIP19 Confers Multiple Biotic and Abiotic Stress Responses in Plant. Int J Mol Sci 21:1–19. https://doi.org/10.3390/IJMS21134701

Hoegger PJ, Kilaru S, James TY, et al (2006) Phylogenetic comparison and classification of laccase and related multicopper oxidase protein sequences. FEBS J 273:2308–2326. https://doi.org/10.1111/J.1742-4658.2006.05247.X

Hovde BT, Daligault HE, Hanschen ER, et al (2019) Detection of Abrin-Like and Prepropulchellin-Like Toxin Genes and Transcripts Using Whole Genome Sequencing and Full-Length Transcript Sequencing of Abrus precatorius. Toxins 2019, Vol 11, Page 691 11:691. https://doi.org/10.3390/TOXINS11120691

Hu B, Jin J, Guo AY, et al (2015) GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics 31:1296–1297. https://doi.org/10.1093/BIOINFORMATICS/BTU817

Hu D, Li X, Yang Z, et al (2022a) Downregulation of a gibberellin 3β-hydroxylase enhances photosynthesis and increases seed yield in soybean. New Phytol 235:502–517. https://doi.org/10.1111/NPH.18153

Hu J, Liu M, Zhang A, et al (2022b) Co-evolved plant and blast fungus ascorbate oxidases orchestrate the redox state of host apoplast to modulate rice immunity. Mol Plant 15:1347–1366. https://doi.org/10.1016/j.molp.2022.07.001

Hufnagel B, Marques A, Soriano A, et al (2020) High-quality genome sequence of white lupin provides insight into soil exploration and seed quality. Nat Commun 2020 111 11:1–12. https://doi.org/10.1038/s41467-019-14197-9

Jiao Y, Wickett NJ, Ayyampalayam S, et al (2011) Ancestral polyploidy in seed plants and angiosperms. Nat 2011 4737345 473:97–100. https://doi.org/10.1038/nature09916

Kalyaanamoorthy S, Minh BQ, Wong TKF, et al (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 2017 146 14:587–589. https://doi.org/10.1038/nmeth.4285

Kang YJ, Kim SK, Kim MY, et al (2014) Genome sequence of mungbean and insights into evolution within Vigna species. Nat Commun 5:5443–5443. https://doi.org/10.1038/NCOMMS6443

Kang YJ, Satyawan D, Shim S, et al (2015) Draft genome sequence of adzuki bean, Vigna angularis. Sci Reports 2015 51 5:1–8. https://doi.org/10.1038/srep08069

Kato N, Esaka M (2000) Expansion of transgenic tobacco protoplasts expressing pumpkin ascorbate oxidase is more rapid than that of wild-type protoplasts. Planta 210:1018–1022. https://doi.org/10.1007/s004250050712

Katoh K, Rozewicki J, Yamada KD (2019) MAFFT online service: multiple

sequence alignment, interactive sequence choice and visualization. Brief Bioinform 20:1160–1166. https://doi.org/10.1093/BIB/BBX108

Kaul S, Koo HL, Jenkins J, et al (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nat 2000 4086814 408:796–815. https://doi.org/10.1038/35048692

Kerk NM, Jiang K, Feldman LJ (2000) Auxin metabolism in the root apical meristem. Plant Physiol 122:925–932. https://doi.org/10.1104/pp.122.3.925

Kim Y, Wang J, Ma C, et al (2023) GmTCP and GmNLP Underlying Nodulation Character in Soybean Depending on Nitrogen. Int J Mol Sci 24:7750. https://doi.org/10.3390/IJMS24097750/S1

Kreplak J, Madoui MA, Cápal P, et al (2019) A reference genome for pea provides insight into legume genome evolution. Nat Genet 2019 519 51:1411–1422. https://doi.org/10.1038/s41588-019-0480-1

Kumari R, Kumar S, Singh L, Hallan V (2016) Movement Protein of Cucumber Mosaic Virus Associates with Apoplastic Ascorbate Oxidase. PLoS One 11:. https://doi.org/10.1371/JOURNAL.PONE.0163320

Kuraku S, Zmasek CM, Nishimura O, Katoh K (2013) aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. Nucleic Acids Res 41:W22–W28. https://doi.org/10.1093/NAR/GKT389

Layzell DB, Hunt S, Palmer GR (1990) Mechanism of nitrogenase inhibition in soybean nodules: Pulse-modulated spectroscopy indicates that nitrogenase activity is limited by o2. Plant Physiol 92:1101–1107. https://doi.org/10.1104/pp.92.4.1101

Lee MH, Dawson CR (1973) Ascorbate Oxidase. J Biol Chem 248:6603–6609. https://doi.org/10.1016/s0021-9258(19)43396-6

Lee YC, Tsai PT, Huang XX, Tsai HL (2022) Family Members Additively Repress the Ectopic Expression of BASIC PENTACYSTEINE3 to Prevent Disorders in Arabidopsis Circadian Vegetative Development. Front Plant Sci 13:919946. https://doi.org/10.3389/FPLS.2022.919946/BIBTEX

Li H, Jiang F, Wu P, et al (2020) A High-Quality Genome Sequence of Model Legume Lotus japonicus (MG-20) Provides Insights into the Evolution of Root Nodule Symbiosis. Genes 2020, Vol 11, Page 483 11:483. https://doi.org/10.3390/GENES11050483

Li R, Xin S, Tao C, et al (2017) Cotton ascorbate oxidase promotes cell growth in cultured tobacco bright yellow-2 cells through generation of apoplast oxidation. Int J Mol Sci 18:1–15. https://doi.org/10.3390/ijms18071346

Lindström K, Mousavi SA (2020) Effectiveness of nitrogen fixation in rhizobia. Microb Biotechnol 13:1314–1335. https://doi.org/10.1111/1751-7915.13517

Little DP, Moran RC, Brenner ED, Stevenson DW (2007) Nuclear genome size in Selaginella. https://doi.org/101139/G06-138 50:351–356. https://doi.org/10.1139/G06-138

Liu W, Zhang Y, Li W, et al (2020) Genome-wide characterization and expression analysis of soybean trihelix gene family. PeerJ 2020:e8753. https://doi.org/10.7717/PEERJ.8753/SUPP-15

Lonardi S, Muñoz-Amatriaín M, Liang Q, et al (2019) The genome of cowpea (Vigna unguiculata [L.] Walp.). Plant J 98:767–782. https://doi.org/10.1111/TPJ.14349

Magallón S, Sanderson MJ (2001) Absolute diversification rates in angiosperm clades. Evolution 55:1762–1780. https://doi.org/10.1111/J.0014-3820.2001.TB00826.X

Maphosa Y, Jideani VA, Maphosa Y, Jideani VA (2017) The Role of Legumes in Human Nutrition. Funct Food - Improv Heal through Adequate Food. https://doi.org/10.5772/INTECHOPEN.69127

Marand AP, Eveland AL, Kaufmann K, Springer NM (2023) cis-Regulatory Elements in Plant Development, Adaptation, and Evolution. https://doi.org/101146/annurev-arplant-070122-030236 74:111–137. https://doi.org/10.1146/ANNUREV-ARPLANT-070122-030236

Messerschmidt a (1990) Modelling and structural relationships. Enzyme 352:341–352

Messerschmidt A (1997) Multi-copper oxidases. World Scientific, Singapore

Messerschmidt A, Ladenstein R, Huber R, et al (1992) Refined crystal structure of ascorbate oxidase at 1.9 Å resolution. J Mol Biol 224:179–205. https://doi.org/10.1016/0022-2836(92)90583-6

Messerschmidt A, Rossi A, Ladenstein R, et al (1989) X-ray crystal structure of the blue oxidase ascorbate oxidase from Zucchini. Analysis of the polypeptide fold and a model of the copper sites and ligands. J Mol Biol 206:513–529. https://doi.org/10.1016/0022-2836(89)90498-1

Moghaddam SM, Oladzad A, Koh C, et al (2021) The tepary bean genome provides insight into evolution and domestication under heat stress. Nat Commun 2021 121 12:1–14. https://doi.org/10.1038/s41467-021-22858-x

Monfared MM, Simon MK, Meister RJ, et al (2011) Overlapping and antagonistic activities of BASIC PENTACYSTEINE genes affect a range of developmental processes in Arabidopsis. Plant J 66:1020–1031. https://doi.org/10.1111/J.1365-313X.2011.04562.X

Nakamura K, Kawabata T, Yura K, Go N (2003) Novel types of two-domain multi-copper oxidases: possible missing links in the evolution. FEBS Lett 553:239–244. https://doi.org/10.1016/S0014-5793(03)01000-7

Nanasato Y, Akashi K, Yokota A (2005) Co-expression of cytochrome b561 and ascorbate oxidase in leaves of wild watermelon under drought and high light conditions. Plant Cell Physiol 46:1515–1524. https://doi.org/10.1093/pcp/pci164

Ohkawa J, Okada N, Shinmyo A, Takano M (1989) Primary structure of cucumber (Cucumis sativus) ascorbate oxidase deduced from cDNA sequence: homology with blue copper proteins and tissue-specific expression. Proc Natl Acad Sci U S A 86:1239–1243. https://doi.org/10.1073/pnas.86.4.1239

Ouyang S, Zhu W, Hamilton J, et al (2007) The TIGR Rice Genome Annotation Resource: improvements and new features. Nucleic Acids Res 35:D883–D887. https://doi.org/10.1093/NAR/GKL976

Ouyang W, Chen L, Ma J, et al (2022) Identification of Quantitative Trait Locus and Candidate Genes for Drought Tolerance in a Soybean Recombinant Inbred Line Population. Int J Mol Sci 23:10828. https://doi.org/10.3390/IJMS231810828/S1

Pan Z, Chen L, Wang F, et al (2019) Genome-wide identification and expression analysis of the ascorbate oxidase gene family in gossypium hirsutum reveals the critical role of ghao1a in delaying dark-induced leaf senescence. Int J Mol Sci 20:. https://doi.org/10.3390/ijms20246167

Panchy N, Lehti-Shiu M, Shiu SH (2016) Evolution of gene duplication in plants. Plant Physiol 171:2294–2316. https://doi.org/10.1104/pp.16.00523

Pignocchi C, Fletcher JM, Wilkinson JE, et al (2003) The Function of Ascorbate Oxidase in Tobacco. Plant Physiol 132:1631–1641. https://doi.org/10.1104/PP.103.022798

Roy S, Liu W, Nandety RS, et al (2020) Celebrating 20 Years of Genetic Discoveries in Legume Nodulation and Symbiotic Nitrogen Fixation. Plant Cell 32:15–41. https://doi.org/10.1105/TPC.19.00279

Roy SW, Penny D (2007) Patterns of intron loss and gain in plants: Intron loss-dominated evolution and genome-wide comparison of O. sativa and A. thaliana. Mol Biol Evol 24:171–181. https://doi.org/10.1093/molbev/msl159

Sandhu AK, Brown MR, Subramanian S, Brözel VS (2023) Bradyrhizobium diazoefficiens USDA 110 displays plasticity in the attachment phenotype when grown in different soybean root exudate compounds. Front Microbiol 14:. https://doi.org/10.3389/fmicb.2023.1190396

Sanmartin M, Pateraki I, Chatzopoulou F, Kanellis AK (2007) Differential expression of the ascorbate oxidase multigene family during fruit development and in response to stress. Planta 225:873–885. https://doi.org/10.1007/S00425-006-0399-5/FIGURES/6

Schmutz J, Cannon SB, Schlueter J, et al (2010) Genome sequence of the palaeopolyploid soybean. Nat 2010 4637278 463:178–183. https://doi.org/10.1038/nature08670

Sedivy EJ, Wu F, Hanzawa Y (2017) Soybean domestication: the origin, genetic architecture and molecular bases. New Phytol 214:539–553. https://doi.org/10.1111/nph.14418

Shi WY, Du YT, Ma J, et al (2018) The WRKY Transcription Factor GmWRKY12 Confers Drought and Salt Tolerance in Soybean. Int J Mol Sci 19:. https://doi.org/10.3390/IJMS19124087

Singh D, Chaudhary P, Taunk J, et al (2021a) Fab Advances in Fabaceae for Abiotic Stress Resilience: From ‘Omics’ to Artificial Intelligence. Int J Mol Sci 2021, Vol 22, Page 10535 22:10535. https://doi.org/10.3390/IJMS221910535

Singh RR, Nobleza N, Demeestere K, Kyndt T (2020a) Ascorbate Oxidase Induces Systemic Resistance in Sugar Beet Against Cyst Nematode Heterodera schachtii. Front Plant Sci 11:1–15. https://doi.org/10.3389/fpls.2020.591715

Singh RR, Pajar JA, Audenaert K, Kyndt T (2021b) Induced Resistance by Ascorbate Oxidation Involves Potentiating of the Phenylpropanoid Pathway and Improved Rice Tolerance to Parasitic Nematodes. Front Plant Sci 12:. https://doi.org/10.3389/fpls.2021.713870

Singh RR, Verstraeten B, Siddique S, et al (2020b) Ascorbate oxidation activates systemic defence against root-knot nematode Meloidogyne graminicola in rice. J Exp Bot 71:4271–4284. https://doi.org/10.1093/jxb/eraa171

Skorupa M, Szczepanek J, Yolcu S, et al (2022) Characteristic of the Ascorbate Oxidase Gene Family in Beta vulgaris and Analysis of the Role of AAO in Response to Salinity and Drought in Beet. Int J Mol Sci 23:. https://doi.org/10.3390/ijms232112773

Smirnoff N (2000) Ascorbic acid: metabolism and functions of a multi-facetted molecule. Curr Opin Plant Biol 3:229–235. https://doi.org/10.1016/s1369-5266(00)80070-9

Song H, Wang P, Hou L, et al (2016) Global Analysis of WRKY Genes and Their Response to Dehydration and Salt Stress in Soybean. Front Plant Sci 7:. https://doi.org/10.3389/FPLS.2016.00009

Tang H, Krishnakumar V, Bidwell S, et al (2014) An improved genome release (version Mt4.0) for the model legume Medicago truncatula. BMC Genomics 15:1–14. https://doi.org/10.1186/1471-2164-15-312/FIGURES/7

Tian KH, Pan C, Yang YF, et al (2019) Differential expression of the ascorbate oxidase multigene family of Camellia sinensis in response to stress. J Hortic Sci Biotechnol 94:160–170. https://doi.org/10.1080/14620316.2018.1495581

Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res 44:W232–W235. https://doi.org/10.1093/NAR/GKW256

Turchetto-Zolet AC, Christoff AP, Kulcheski FR, et al (2016) Diversity and evolution of plant diacylglycerol acyltransferase (DGATs) unveiled by phylogenetic, gene structure and expression analyses. Genet Mol Biol 39:524–538. https://doi.org/10.1590/1678-4685-GMB-2016-0024

Valliyodan B, Cannon SB, Bayer PE, et al (2019) Construction and comparison of three reference-quality genome assemblies for soybean. Plant J 100:1066–1082. https://doi.org/10.1111/TPJ.14500

Varshney RK, Chen W, Li Y, et al (2011) Draft genome sequence of pigeonpea (Cajanus cajan), an orphan legume crop of resource-poor farmers. Nat Biotechnol 2011 301 30:83–89. https://doi.org/10.1038/nbt.2022

Varshney RK, Song C, Saxena RK, et al (2013) Draft genome sequence of chickpea (Cicer arietinum) provides a resource for trait improvement. Nat

Biotechnol 2013 313 31:240–246. https://doi.org/10.1038/nbt.2491

Wang H, Ni D, Shen J, et al (2022) Genome-Wide Identification of the AP2/ERF Gene Family and Functional Analysis of GmAP2/ERF144 for Drought Tolerance in Soybean. Front Plant Sci 13:848766. https://doi.org/10.3389/FPLS.2022.848766/BIBTEX

Wang J, Sun P, Li Y, et al (2017) Hierarchically Aligning 10 Legume Genomes Establishes a Family-Level Genomics Platform. Plant Physiol 174:284–300. https://doi.org/10.1104/PP.16.01981

Wang M, Guo W, Li J, et al (2021) The miR528- AO Module Confers Enhanced Salt Tolerance in Rice by Modulating the Ascorbic Acid and Abscisic Acid Metabolism and ROS Scavenging. J Agric Food Chem 69:8634–8648. https://doi.org/10.1021/acs.jafc.1c01096

Wang Y, Li K, Hen L, et al (2015) MicroRNA167-Directed Regulation of the Auxin Response Factors GmARF8a and GmARF8b Is Required for Soybean Nodulation and Lateral Root Development. Plant Physiol 168:984–999. https://doi.org/10.1104/PP.15.00265

Wang Y, You FM, Lazo GR, et al (2013) PIECE: A database for plant gene structure comparison and evolution. Nucleic Acids Res 41:1159–1166. https://doi.org/10.1093/nar/gks1109

Wei JT, Zhao SP, Zhang HY, et al (2023) GmDof41 regulated by the DREB1-type protein improves drought and salt tolerance by regulating the DREB2-type protein in soybean. Int J Biol Macromol 230:123255. https://doi.org/10.1016/J.IJBIOMAC.2023.123255

Wheeler GL, Jones MA, Smirnoff N (1998) The biosynthetic pathway of vitamin C in higher plants. Nat 1998 3936683 393:365–369. https://doi.org/10.1038/30728

Wu DG, Wang Y, Huang SC, et al (2021) Genome-wide identification and expression analysis of aao gene family in maize. Pakistan J Bot 53:181–190. https://doi.org/10.30848/PJB2021-1(11)

Xin S, Tao C, Li H (2016) Cloning and functional analysis of the promoter of an Ascorbate oxidase gene from Gossypium hirsutum. PLoS One 11:1–13. https://doi.org/10.1371/journal.pone.0161695

Xu Z, Wang R, Kong K, et al (2022) An APETALA2/ethylene responsive factor transcription factor GmCRF4a regulates plant height and auxin biosynthesis in

soybean. Front Plant Sci 13:983650. https://doi.org/10.3389/FPLS.2022.983650/BIBTEX

Yamamoto A, Bhuiyan MNH, Waditee R, et al (2005) Suppressed expression of the apoplastic ascorbate oxidase gene increases salt tolerance in tobacco and Arabidopsis plants. J Exp Bot 56:1785–1796. https://doi.org/10.1093/jxb/eri167

Yang H, Shi G, Du H, et al (2017a) Genome-Wide Analysis of Soybean LATERAL ORGAN BOUNDARIES Domain-Containing Genes: A Functional Investigation of GmLBD12. Plant Genome 10:plantgenome2016.07.0058. https://doi.org/10.3835/PLANTGENOME2016.07.0058

Yang X, Kim MY, Ha J, Lee SH (2019) Overexpression of the Soybean NAC Gene GmNAC109 Increases Lateral Root Formation and Abiotic Stress Tolerance in Transgenic Arabidopsis Plants. Front Plant Sci 10:462757. https://doi.org/10.3389/FPLS.2019.01036/BIBTEX

Yang Y, Chi Y, Wang Z, et al (2016) Functional analysis of structurally related soybean GmWRKY58 and GmWRKY76 in plant growth and development. J Exp Bot 67:4727–4742. https://doi.org/10.1093/JXB/ERW252

Yang Y, Zhou Y, Chi Y, et al (2017b) Characterization of Soybean WRKY Gene Family and Identification of Soybean WRKY Genes that Promote Resistance to Soybean Cyst Nematode. Sci Rep 7:1–13. https://doi.org/10.1038/s41598-017-18235-8

Yuan S, Li X, Li R, et al (2018) Genome-wide identification and classification of Soybean C2H2 Zinc Finger proteins and their expression analysis in legume-rhizobium symbiosis. Front Microbiol 9:287812. https://doi.org/10.3389/FMICB.2018.00126/BIBTEX

Yue L, Pei X, Kong F, et al (2023) Divergence of functions and expression patterns of soybean bZIP transcription factors. Front Plant Sci 14:1150363. https://doi.org/10.3389/FPLS.2023.1150363/BIBTEX

Zhang C, Grosic S, Whitham SA, Hill JH (2012) The requirement of multiple defense genes in soybean Rsv1-mediated extreme resistance to soybean mosaic virus. Mol Plant Microbe Interact 25:1307–1313. https://doi.org/10.1094/MPMI-02-12-0046-R

Zhang M, Liu Y, Shi H, et al (2018) Evolutionary and expression analyses of soybean basic Leucine zipper transcription factor family. BMC Genomics

19:1–14. https://doi.org/10.1186/S12864-018-4511-6/FIGURES/8

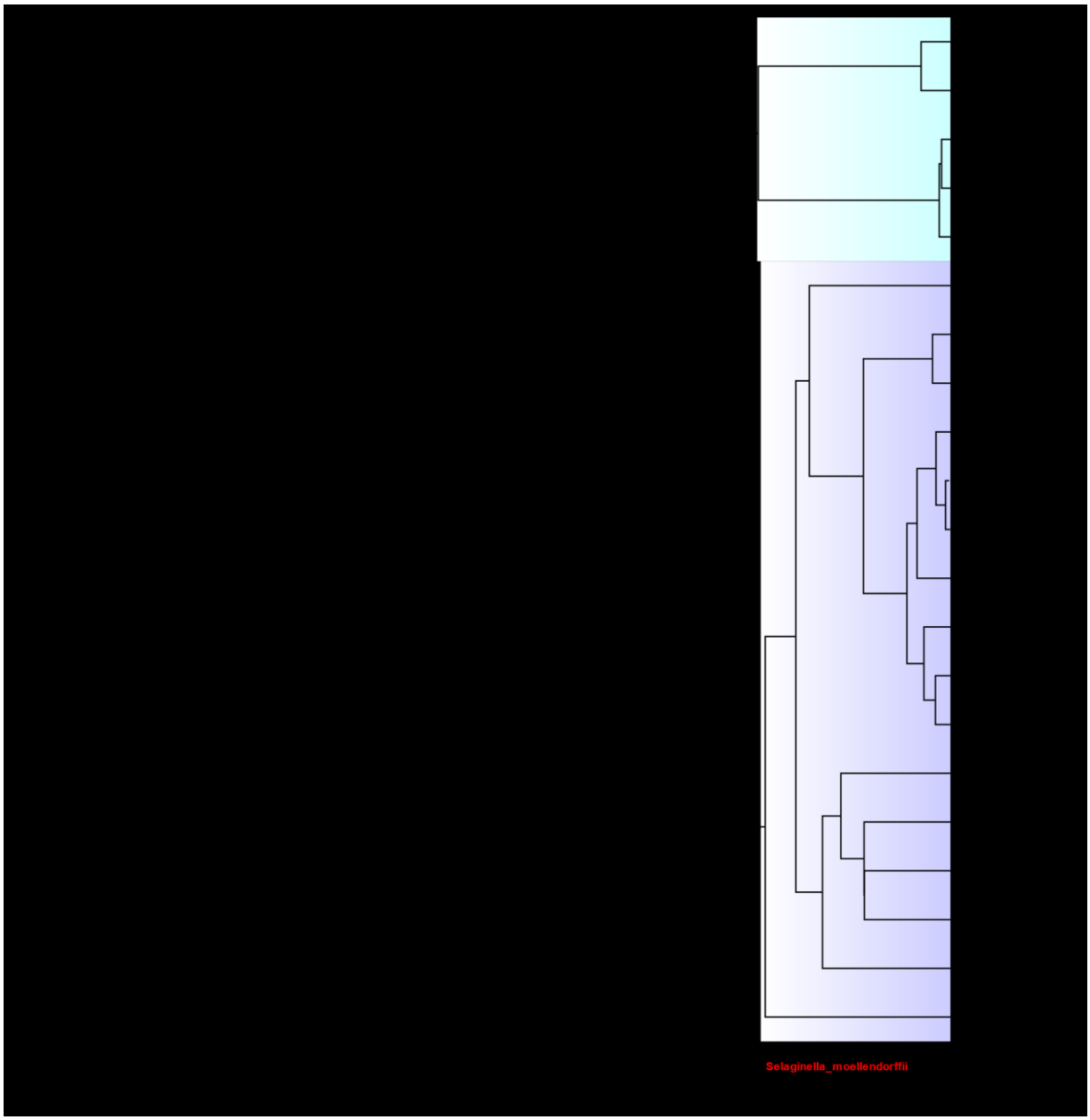
Zhang Z, Zhao Y, Chen Y, et al (2023) Overexpression of TCP9-like gene enhances salt tolerance in transgenic soybean. PLoS One 18:e0288985. https://doi.org/10.1371/JOURNAL.PONE.0288985

Zhao N, Ding X, Lian T, et al (2020) The Effects of Gene Duplication Modes on the Evolution of Regulatory Divergence in Wild and Cultivated Soybean. Front Genet 11:1–9. https://doi.org/10.3389/fgene.2020.601003

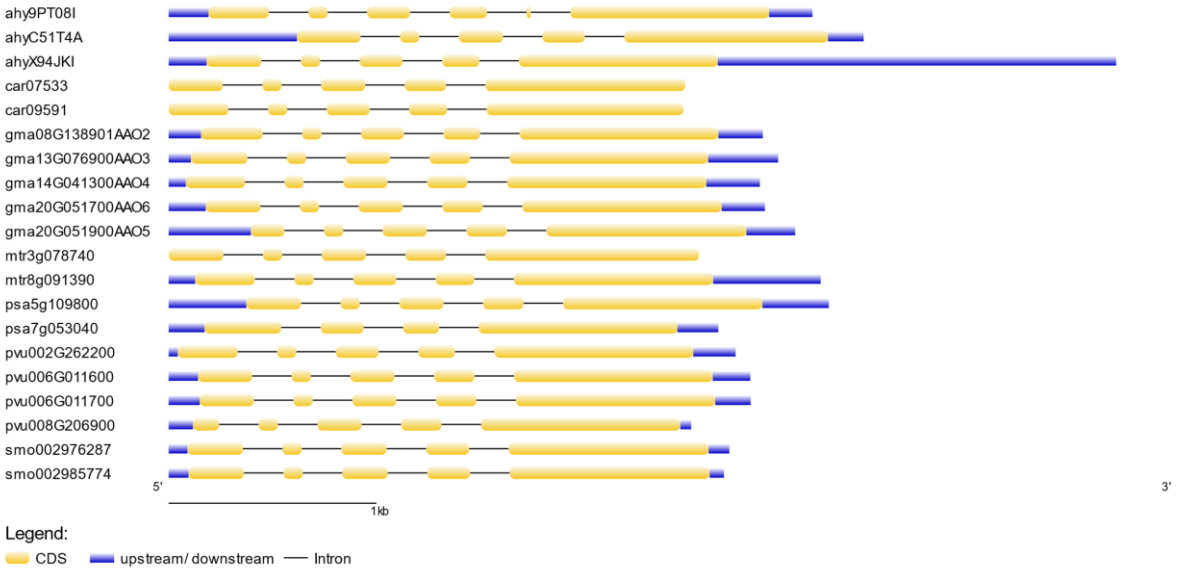
Zhu M, Liu Q, Liu F, et al (2023) Gene Profiling of the Ascorbate Oxidase Family Genes under Osmotic and Cold Stress Reveals the Role of AnAO5 in Cold Adaptation in Ammopiptanthus nanus. Plants 12:. https://doi.org/10.3390/plants12030677

Zhu M, Wang X, Zhou Y, et al (2022) Small RNA Sequencing Revealed that miR4415, a Legume-Specific miRNA, was Involved in the Cold Acclimation of Ammopiptanthus nanus by Targeting an L-Ascorbate Oxidase Gene and Regulating the Redox State of Apoplast. Front Genet 13:1–20. https://doi.org/10.3389/fgene.2022.870446

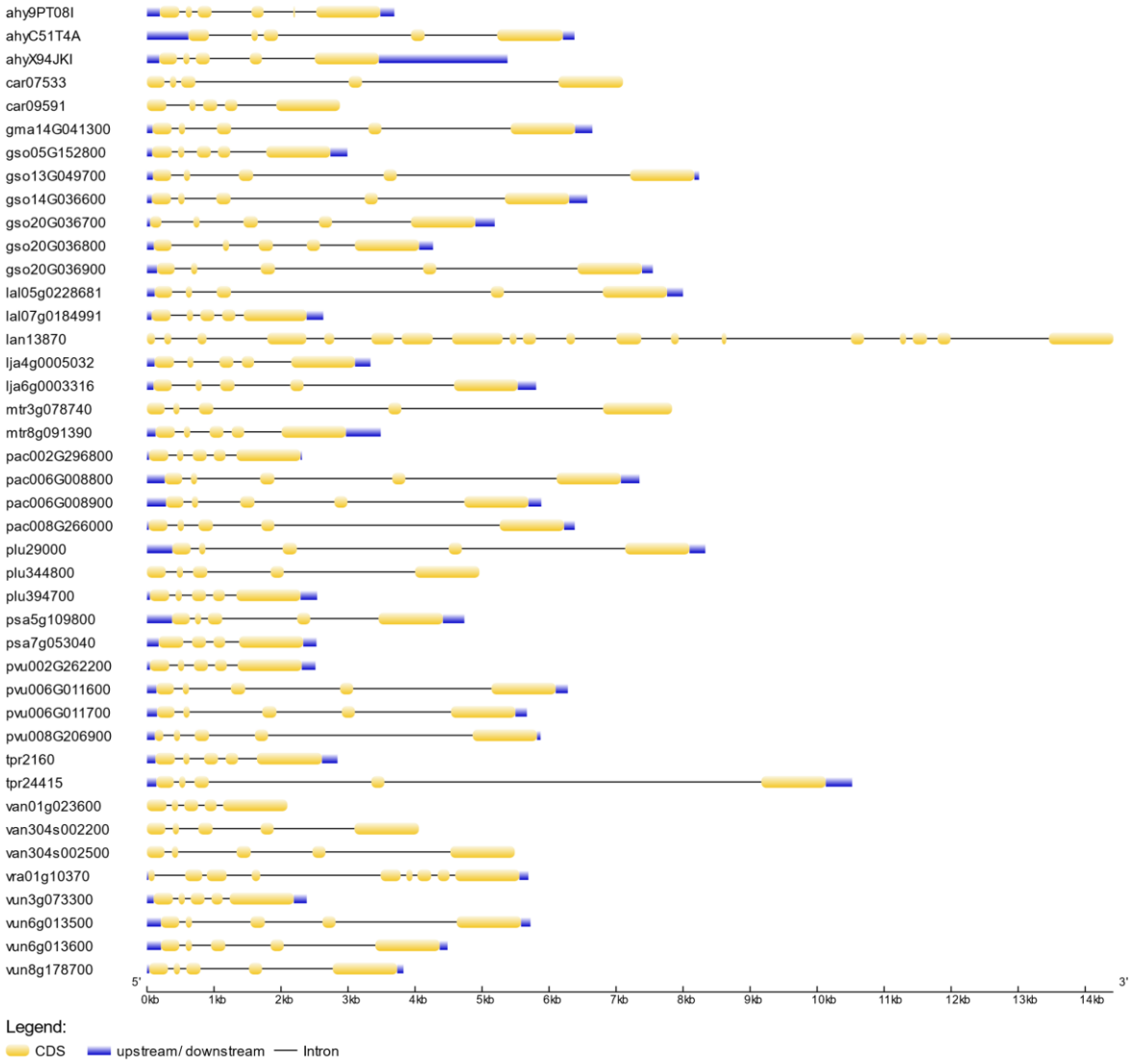
Supplementary material



Supplementary Fig. 1. Fabaceae species tree. The tree shows two groups which are indicated with purple and blue background colors. Fabaceae sequences are represented in black, while *S. moellendorffii* AAOs are the outgroup, in red. The tree was made on FigTree software with data from Time Tree online tool.

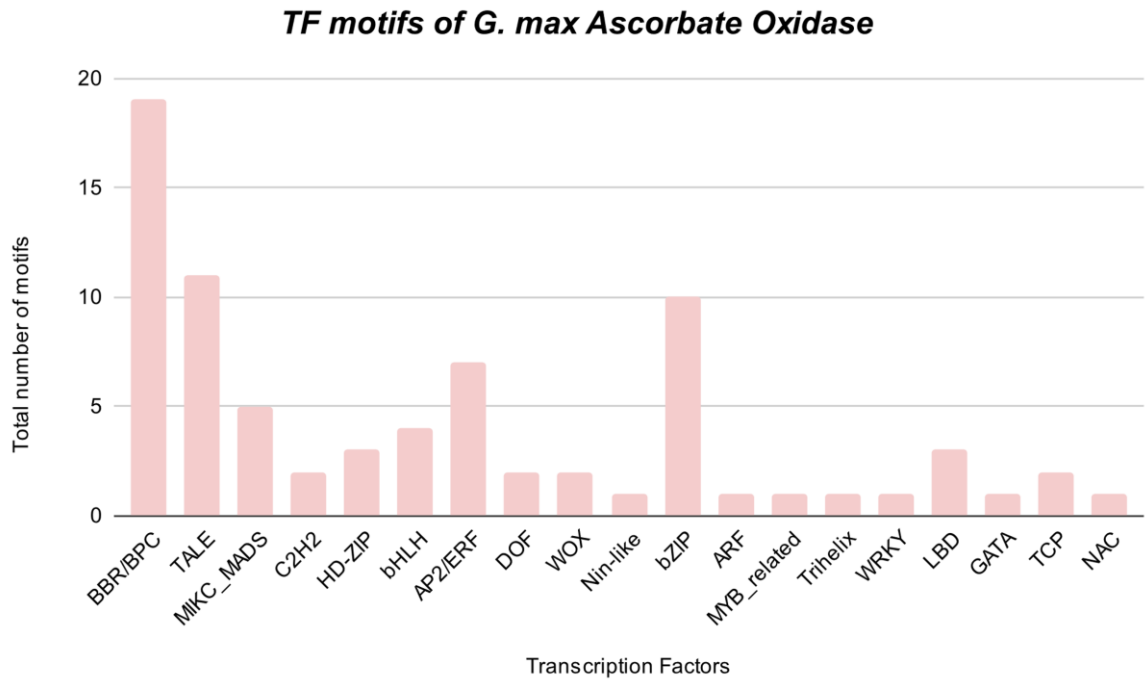


Supplementary Fig. 2. *Ascorbate Oxidase* gene structure with standardized intron sized. *Glycine max* (gma), *Arachis hypogaea* (ahy), *Phaseolus vulgaris* (pvu), *Medicago truncatula* (mtr), *Cicer arietinum* (car), *Pisum sativum* (psa) and *Selaginella moellendorffii* (smo) were selected for the analysis. The Gene Structure Display Server was used to infer the gene structure.



Supplementary Fig. 3. Gene structure analysis of *Ascorbate Oxidase* from Fabaceae family.

The analyses includes sequences from *Glycine max* (gma), *Glycine soja* (gso), *Arachis hypogaea* (ahy), *Phaseolus lunatus* (plu), *Medicago truncatula* (mtr), *Phaseolus acutifolius* (pac), *Vigna unguiculata* (vun), *Lupinus albus* (lal), *Cicer arietinum* (car), *Lotus japonicus* (lja), *Phaseolus vulgaris* (pvu), *Lupinus angustifolius* (lan), *Pisum sativum* (psa), *Vigna angularis* (van), *Vigna radiata* (vra), *Trifolium pratense* (tpr), *Abrus precatorius* (apr), *Arachis duranensis* (adu), *Arachis ipaensis* (aip), *Cajanus cajan* (cca) and *Vigna umbellate* (vum). The Gene Structure Display Server was used to infer the gene structure.



Supplementary Fig. 4: The total number of Transcription Factors (TF) motifs. The motifs were found in the upstream region of Ascorbate Oxidases (AAO) from *G. max*. Barley B-Recombinant/Basic PENTACYSTEINE (BBR-BPC); Three Amino acid Loop Extension (TALE); MADS-box; Intervening-, Keratin-like- and C-domains (MICK-MADS); Cysteine 2 Histidine 2 (C2H2); Cysteine 3 Histidine (C3H); DNA-binding with One Finger (Dof); Homeodomain Leucine Zipper (DH-Zip); Basic Leucine zipper (Bzip); Basic Helix-Loop-Helix (BHLH); APETALA2/Ethylene Responsive Factor (AP2/ERF); NODULE INCEPTION-Like (Nin-Like); Auxin Response Factor (ARF); Myeloblastosis-Related (MYB-Related); Lateral Organ Boundaries Domain Protein (LBD); TEOSINTE BRANCHED 1, CYCLOIDEA, PCF1 (TCP); WUSCHEL-related homeobox (WOX); NAM, ATAF, and CUC - domain proteins (NAC).

Supplementary Table 1: Main information of Ascorbate Oxidases sequences.

Specie	Acronim	Database	Access code	e-value	% identity	Align lenght	Bitscore	Size (aa)	Ploidy/Reference	
Arachis hypogaea	ahy_X94JKI	Phytozome	arahy.Tifrunner.gnm1.ann1.X94JKI.1	0	52	567	561.992	567	tetraploid (2n=4x=40)	Bertioll et al. 2019
Arachis hypogaea	ahy_9PT08I		arahy.Tifrunner.gnm1.ann1.9PT08I.1	0	51	574	554.288	583		
Arachis hypogaea	ahy_C51T4A		arahy.Tifrunner.gnm1.ann1.C51T4A.1	0	69	585	785.023	599		
Phaseolus lunatus	plu_29000		Pl06G0000029000.1.v1	0	91	575	1042.72	578	diploid (2n=22)	Garcia et al. 2021
Phaseolus lunatus	plu_344800		Pl08G0000344800.1.v1	0	74	575	873.233	580		
Phaseolus lunatus	plu_394700		Pl02G0000394700.1.v1	0	51	567	550.821	574		
Medicago truncatula	mtr_3g078740		Medtr3g078740.1	0	77	591	923.694	598	diploid (2n=16)	Tang et al. 2014
Medicago truncatula	mtr_8g091390		Medtr8g091390.1	0	50	553	547.74	578		
Phaseolus acutifolius	pac_006G008800		Phacu.CVR.006G008800.1	0	91	575	1034.63	574	diploid (2n=22)	Moghaddam et al. 2021
Phaseolus acutifolius	pac_006G008900		Phacu.CVR.006G008900.1	0	91	568	1031.94	573		
Phaseolus acutifolius	pac_008G266000		Phacu.CVR.008G266000.1	0	73	575	875.93	580		
Phaseolus acutifolius	pac_002G296800		Phacu.CVR.002G296800.1	0	51	566	560.836	574		
Vigna unguiculata	vun_6g013600		Vigun06g013600.1.p	0	90	571	1058.13	578	diploid (2n=22)	Lonardi et al. 2019
Vigna unguiculata	vun_6g013500		Vigun06g013500.1.p	0	89	571	1039.25	578		
Vigna unguiculata	vun_8g178700		Vigun08g178700.1.p	0	72	567	843.958	580		
Vigna unguiculata	vun_3g073300		Vigun03g073300.1.p	0	53	545	561.992	571		
Lupinus albus	lal_05g0228681		Lalb_Chr05g0228681	0	85	575	991.49	574	diploid (2n=50)	Hufnagel et al. 2020

Lupinus albus	lal_07g0184991		Lalb_Chr07g0184991	0	50	562	558.91	573		
Cicer arietinum	car_07533		Ca_07533	0	83	570	957.592	576	diploid (2n=16)	
Cicer arietinum	car_09591		Ca_09591	0	50	567	555.058	573		Varshney et al. 2013
Lotus japonicus	lja_6g0003316		Lj6g0003316.1	0	85	570	959.133	576	diploid (2n=12)	
Lotus japonicus	lja_4g0005032		Lj4g0005032.1	0	49	572	558.91	575		Li, H. et al. 2020
Phaseolus vulgaris	pvu_006G011700		Phvul.006G011700.1.p	0	91	575	1036.94	574		
Phaseolus vulgaris	pvu_006G011600		Phvul.006G011600.1.p	0	91	568	1034.25	573	diploid (2n=22)	
Phaseolus vulgaris	pvu_008G206900		Phvul.008G206900.1.p	0	76	527	829.321	529		
Phaseolus vulgaris	pvu_002G262200		Phvul.002G262200.1.p	0	51	566	552.362	574		Phytozome v13
Arabidopsis thaliana	ath_5G21105_AAO3		AT5G21105.1	0	67	566	790.03	589		
Arabidopsis thaliana	ath_5G21100_AAO2		AT5G21100.1	0	63	557	695.271	574	diploid (2n=10)	
Arabidopsis thaliana	ath_4G39830_AAO1		AT4G39830.1	0	48	582	558.14	583		Kaul et al. 2000
Oryza sativa	osa_06g37080_AAO1		LOC_Os06g37080.1	0	64	568	723.776	582		
Oryza sativa	osa_09g20090_AAO4		LOC_Os09g20090.1	0	51	552	541.576	578	diploid (2n=24)	
Oryza sativa	osa_09g32952_AAO5		LOC_Os09g32952.1	1.03e-173	50	551	504.597	575		
Oryza sativa	osa_06g37150_AAO2		LOC_Os06g37150.1	0	54	580	590.882	634		Ouyang, S. et al. 2007
Glycine max	gma_13G076900_AAO3		Glyma.13G076900.1.p	0	100	577	1163.67	577	tetraploid (n=4x=80)	
Glycine max	gma_20G051700_AAO6		Glyma.20G051700.1p	0	91	575	1064.29	575		Schmutz et al. 2010

Glycine max	gma_20G051900_AAO5		Glyma.20G051900.1p	0	95	576	1080.86	576		
Glycine max	gma_14G041300_AAO4		Glyma.14G041300.1p	0	76	574	882.093	582		
Glycine max	gma_08G138901_AAO2		Glyma.08G138901.1.p	0	49	570	555.444	577		
Glycine soja	gso_13G049700_AAO3		GlysoPI483463.13G049700.1.p	0	100	577	1161.36	577	diploid (2n=40)	Valliyodan et al. 2019
Glycine soja	gso_20G036900_AAO5		GlysoPI483463.20G036900.1.p	0	95	576	1105.12	576		
Glycine soja	gso_20G036800_AAO6		GlysoPI483463.20G036800.1.p	0	91	575	1063.14	575		
Glycine soja	gso_20G036700		GlysoPI483463.20G036700.1.p	0	91	571	982.245	541		
Glycine soja	gso_14G036600_AAO4		GlysoPI483463.14G036600.1.p	0	76	574	882.93	582		
Glycine soja	gso_05G152800		GlysoPI483463.05G152800.1.p	0	52	549	559.681	578		
Lupinus angustifolius	lan_13870	**Ensembl**	TanjilG_13870	0	79.7	311	1402	1760	diploid (2n=40)	Hane et al. 2017
Pisum sativum	psa_5g109800		Psat5g109800	1.6e-139	82.5	229	1057	574	diploid (2n=14)	Kreplak et al. 2019
Pisum sativum	psa_7g053040		Psat7g053040	1.5e-50	60.5	119	421	568		
Vigna angularis	van_01g023600		LR48_Vigan01g023600	1.4e-50	60.8	120	420	574	diploid (2n=22)	Kang et al. 2015
Vigna angularis	van_304s002200		LR48_Vigan304s002200	0	90.7	312	1576	577		
Vigna angularis	van_304s002500		LR48_Vigan304s002500	0	91.3	312	1583	573		
Vigna radiata	vra_01g10370		Vradi01g10370	4.1e-50	60.0	120	416	834	diploid (2n=22)	Kang et al. 2015
Trifolium pratense	tpr_2160		Tp57577_TGAC_v2_gene2160	5.7e-50	58.8	119	416	581	diploid (2n=14)	De Vega et al. 2015
Trifolium pratense	tpr_24415		Tp57577_TGAC_v2_gene24415	1.3e-142	81.1	238	1078	573		

Abrus precatorius	apr_027343254		XP_027343254.1	0.0	86.68	573	1.021	573		
Abrus precatorius	apr_027346285		XP_027346285.1	0.0	83.80	573	1004	573	diploid (2n=22)	
Abrus precatorius	apr_027330394		XP_027330394.1	0.0	80.31	573	971	573		
Abrus precatorius	apr_027367334		XP_027367334.1	0.0	51.73	565	583	565		Hovde et al. 2019
Arachis duranensis	adu_015958466		XP_015958466.1	0.0	83.27	574	962	574		
Arachis duranensis	adu_015973554		XP_015973554.1	0.0	69.57	598	811	598	diploid (2n=20)	
Arachis duranensis	adu_015956011		XP_015956011.1	0.0	51.68	566	578	566		Bertioli, et al. 2016
Arachis ipaensis	aip_016197051		XP_016197051.1	0.0	83.45	576	964	576	diploid (2n=20)	
Arachis ipaensis	aip_016167458	**NCBI**	XP_016167458.1	0.0	68.71	600	812	600		Bertioli, et al. 2016
Cajanus cajan	cca_020206694		XP_020206694.1	0.0	89.55	573	1049	573		
Cajanus cajan	cca_020207993		XP_020207993.2	0.0	78.92	579	950	579	diploid (2n=22)	
Cajanus cajan	cca_020208241		XP_020208241.1	0.0	72.34	562	863	562		
Cajanus cajan	cca_020236942		XP_020236942.1	0.0	52.82	576	588	576		Varshney et al. 2011
Vigna umbellata	vum_047164792		XP_047164792.1	0.0	70.69	570	846	570	diploid (2n=22)	
Vigna umbellata	vum_047180016		XP_047180016.1	0.0	52.66	622	585	622		NCBI
Selaginella moellendorffii	smo_002985774		XP_002985774.1	0.0	51.47	582	598	582	diploid (2n=20)	
Selaginella moellendorffii	smo_002976287		XP_002976287.1	0.0	51.30	582	592	582		Little et al. 2007

The current data was obtained from the three databases informed: Phytozome, Ensembl Plants, and NCBI.

Supplementary Table 2: **Transcription Factor motifs in the 2000bp upstream region of *Glycine max Ascorbate Oxidase* genes.**

Pattern name	TF Family name	TF abbreviation	Present on	start	stop	strand	score	p-value	q-value	matched sequence
Glyma.04G026300	Barley B-Recombinant/Basic PENTACYSTEINE	BBR/BPC	gma_13G076900_AAO3	1955	1975	+	209.167	1.23e-08	4.11e-05	CTCCCTCCCTCTCTCTATATA
Glyma.04G026300			gma_13G076900_AAO3	1953	1973	+	195.833	2.33e-08	4.11e-05	CCCTCCCTCCCTCTCTCTATA
Glyma.04G026300			gma_13G076900_AAO3	1951	1971	+	160.972	1.11e-07	1,31E+02	TTCCCTCCCTCCCTCTCTCTA
Glyma.04G026300			gma_13G076900_AAO3	1949	1969	+	136.111	3.13e-07	2,76E+02	TTTTCCCTCCCTCCCTCTCTC
Glyma.04G026300			gma_13G076900_AAO3	1945	1965	+	11.2778	7.84e-07	5,53E+02	CACCTTTTCCCTCCCTCCCTC
Glyma.04G026300			gma_13G076900_AAO3	1947	1967	+	877.778	1.99e-06	1,17E+02	CCTTTTCCCTCCCTCCCTCTC
Glyma.04G026300			gma_13G076900_AAO3	1957	1977	+	556.944	6.07e-06	3,06E+02	CCCTCCCTCTCTCTATATAAG
Glyma.04G026300			gma_20G051700_AAO6	1940	1960	+	353.472	1.99e-12	5.62e-09	CCCTCTCTCTCTCTCTCTCTT
Glyma.04G026300			gma_20G051700_AAO6	1938	1958	+	346.944	2.95e-12	5.62e-09	TTCCCTCTCTCTCTCTCTCTC
Glyma.04G026300			gma_20G051700_AAO6	1934	1954	+	276.806	3.44e-10	4.37e-07	CTTTTTCCCTCTCTCTCTCTC
Glyma.04G026300			gma_20G051700_AAO6	1936	1956	+	262.778	7.53e-10	5.93e-07	TTTTCCCTCTCTCTCTCTCTC
Glyma.04G026300			gma_20G051700_AAO6	1942	1962	+	262.222	7.78e-10	5.93e-07	CTCTCTCTCTCTCTCTCTTAT
Glyma.04G026300			gma_20G051700_AAO6	1932	1952	+	222.222	6.48e-09	4.12e-06	ACCTTTTTCCCTCTCTCTCTC
Glyma.04G026300			gma_20G051700_AAO6	1944	1964	+	147.917	1.93e-07	1,05E+02	CTCTCTCTCTCTCTCTTATAT
Glyma.04G026300			gma_20G051900_AAO5	1955	1975	+	776.389	2.85e-06	1,05E+02	CACCTTTTCCCTCTCTATATA
Glyma.05G137100			gma_08G138901_AAO2	1060	1083	+	166.825	7.45e-08	2,77E+02	ATGTAACCGAGAGAGAGAGAGAGG
Glyma.05G137100			gma_08G138901_AAO2	1070	1093	+	10.9365	8.75e-07	1,63E+02	GAGAGAGAGAGAGGTGTTTGAAGG
Glyma.05G137100			gma_08G138901_AAO2	1068	1091	+	7.7619	3.02e-06	3,74E+02	GAGAGAGAGAGAGAGGTGTTTGAA
Glyma.05G137100			gma_08G138901_AAO2	1062	1085	+	634.921	5.11e-06	4,75E+02	GTAACCGAGAGAGAGAGAGAGGTG

Glyma.14G091200	Three Amino acid Loop Extension	TALE	gma_13G076900_AAO3	1953	1972	+	210.541	4.79e-08	1,27E+02	CCCTCCCTCCCTCTCTCTAT
Glyma.14G091200			gma_13G076900_AAO3	1949	1968	+	206.892	6.42e-08	1,27E+02	TTTTCCCTCCCTCCCTCTCT
Glyma.14G091200			gma_13G076900_AAO3	1951	1970	+	195.541	1.53e-07	2,02E+02	TTCCCTCCCTCCCTCTCTCT
Glyma.14G091200			gma_13G076900_AAO3	1947	1966	+	176.486	5.77e-07	5,71E+02	CCTTTTCCCTCCCTCCCTCT
Glyma.14G091200			gma_13G076900_AAO3	1945	1964	+	130.135	8.19e-06	6,49E+02	CACCTTTTCCCTCCCTCCCT
Glyma.14G091200			gma_20G051700_AAO6	1940	1959	+	215.946	3.06e-08	1,21E+02	CCCTCTCTCTCTCTCTCTCT
Glyma.14G091200			gma_20G051700_AAO6	1938	1957	+	184.189	3.44e-07	6,79E+02	TTCCCTCTCTCTCTCTCTCT
Glyma.14G091200			gma_20G051700_AAO6	1934	1953	+	148.649	3.1e-06	3,06E+02	CTTTTTCCCTCTCTCTCTCT
Glyma.14G091200			gma_20G051700_AAO6	1936	1955	+	170.405	8.54e-07	1,12E+02	TTTTCCCTCTCTCTCTCTCT
Glyma.14G091200			gma_08G138901_AAO2	526	545	+	169.054	9.29e-07	3,45E+02	TCTCTGCTCTCCCTCTCTGC
Glyma.14G091200			gma_08G138901_AAO2	528	547	+	149.189	3.01e-06	5,58E+02	TCTGCTCTCCCTCTCTGCAC
Glyma.18G224500	MADS-box, Intervening-, Keratin-like- and C-domains.	MIKC_MADS	gma_13G076900_AAO3	1952	1972	+	159.275	1.84e-06	7,25E+02	TCCCTCCCTCCCTCTCTCTAT
Glyma.18G224500			gma_20G051700_AAO6	1937	1957	+	205.942	2.07e-10	4.31e-07	TTTCCCTCTCTCTCTCTCTCT
Glyma.18G224500			gma_20G051700_AAO6	1939	1959	+	205.652	2.23e-10	4.31e-07	TCCCTCTCTCTCTCTCTCTCT
Glyma.18G224500			gma_20G051700_AAO6	1935	1955	+	201.014	1.14e-09	1.47e-06	TTTTTCCCTCTCTCTCTCTCT
Glyma.18G224500			gma_20G051700_AAO6	1941	1961	+	19.058	1.81e-08	1.75e-05	CCTCTCTCTCTCTCTCTCTTA
Glyma.02G293300	Cysteine 2 Histidine 2	C2H2	gma_13G076900_AAO3	1947	1965	+	145.909	2.47e-06	9,80E+01	CCTTTTCCCTCCCTCCCTC
Glyma.02G293300			gma_13G076900_AAO3	1949	1967	+	132.727	5.23e-06	1,04E+02	TTTTCCCTCCCTCCCTCTC
Glyma.12G063800	DNA-binding with One Finger	Dof	gma_08G138901_AAO2	999	1019	+	149.219	3.03e-06	1,13E+02	GAAACTTTTTGTTTTTTTTTT
Glyma.17G089300			gma_20G051700_AAO6	1783	1803	+	13.619	8.66e-06	1,63E+02	AAAAAAAAAAAAAAGAAAACT
Glyma.01G044600	Homeodomain Leucine Zipper	HD-Zip	gma_13G076900_AAO3	1165	1175	+	130.735	4.32e-06	1,26E+02	CCAATAATTGT

Glyma.13G169900			gma_13G076900_AAO3	1165	1175	+	140.597	5.72e-06	1,58E+02	CCAATAATTGT
Glyma.18G258400			gma_13G076900_AAO3	1165	1175	+	137.639	9.62e-06	2,85E+02	CCAATAATTGT
Glyma.18G052500	Basic Leucine zipper	bZIP	gma_20G051900_AAO5	51	61	+	147.778	6.2e-06	2,47E+02	GCGGGCTGGCC
Glyma.02G058800			gma_08G138901_AAO2	350	359	+	163.763	2.63e-07	1,04E+02	CGACGTGGCG
Glyma.09G208500			gma_08G138901_AAO2	345	359	+	175.079	8.09e-07	3,21E+02	AAATACGACGTGGCG
Glyma.12G040600			gma_08G138901_AAO2	345	359	+	14.381	2.53e-06	0.01	AAATACGACGTGGCG
Glyma.15G232000			gma_08G138901_AAO2	353	367	+	154.921	2.59e-06	1,03E+02	CGTGGCGTCATCATC
Glyma.10G071700			gma_08G138901_AAO2	344	361	+	165.714	1.29e-06	5,13E+02	CAAATACGACGTGGCGTC
Glyma.13G193700			gma_08G138901_AAO2	353	367	+	154.921	2.59e-06	1,03E+02	CGTGGCGTCATCATC
Glyma.03G003400			gma_08G138901_AAO2	345	359	+	163.492	1.57e-06	6,21E+02	AAATACGACGTGGCG
Glyma.03G219300			gma_08G138901_AAO2	345	359	+	138.889	3.6e-06	1,43E+02	AAATACGACGTGGCG
Glyma.02G161100			gma_08G138901_AAO2	354	363	+	13.403	9.94e-06	1,98E+02	GTGGCGTCAT
Glyma.11G131200	Basic Helix-Loop-Helix	bHLH	gma_13G076900_AAO3	1733	1746	+	151.587	5.58e-06	2,22E+02	TTGTCAACTTGCTT
Glyma.02G175700			gma_13G076900_AAO3	1735	1745	+	149.722	7.07e-06	2,81E+02	GTCAACTTGCT
Glyma.08G239500			gma_20G051900_AAO5	1240	1250	+	149.524	7.33e-06	2,92E+02	TACAAGTTGCA
Glyma.14G089600			gma_08G138901_AAO2	342	362	+	13.75	7.61e-06	0.03	GGCAAATACGACGTGGCGTCA
Glyma.09G248200	APETALA2/Ethylene Responsive Factor	AP2/ERF	gma_20G051700_AAO6	1781	1800	+	174.359	4.46e-07	8,51E+02	TAAAAAAAAAAAAAAGAAA
Glyma.09G248200			gma_20G051700_AAO6	1789	1808	+	159.231	1.95e-06	1,45E+02	AAAAAAAAGAAAACTAGAAA
Glyma.09G248200			gma_20G051700_AAO6	1779	1798	+	147.564	5.15e-06	2,45E+02	TATAAAAAAAAAAAAAAGA
Glyma.09G248200			gma_08G138901_AAO2	1062	1081	+	154.103	3.04e-06	1,14E+02	GTAACCGAGAGAGAGAGAGA
Glyma.11G014200			gma_08G138901_AAO2	358	377	+	158.281	1.27e-06	5,00E+00	CGTCATCATCGCCA**CCC**GCG

Glyma.06G068800			gma_08G138901_AAO2	353	373	+	10.75	3.58e-06	1,40E+01	CGTGGCGTCATCATCGCCACC
Glyma.11G014200			gma_08G138901_AAO2	355	374	+	136.719	4.56e-06	8,97E+02	TGGCGTCATCATCGCCACCC
Glyma.06G017800	NODULE INCEPTION-Like	Nin-like	gma_20G051700_AAO6	608	622	+	148.507	4.9e-06	1,95E+02	TTAGCAGCATGAGCG
Glyma.05G200800	Auxin Response Factor	ARF	gma_20G051900_AAO5	30	39	+	131.746	7.37e-06	2,94E+02	CGCAGACAGG
Glyma.10G048500	Myeloblastosis-Related	MYB_related	gma_14G041300_AAO4	1840	1851	+	14.381	6.94e-06	2,73E+02	AGATATTTTTAG
Glyma.18G202400	Trihelix	Trihelix	gma_14G041300_AAO4	1910	1926	+	14.597	7.58e-06	2,96E+02	GCAACGTGTTAAATGTG
Glyma.03G109100	WRKY	WRKY	gma_14G041300_AAO4	1020	1029	+	137.217	8.96e-06	3,57E+02	CGGTTGACCA
Glyma.14G021100	Lateral Organ Boundaries Domain Protein	LBD	gma_08G138901_AAO2	356	370	+	16.619	1.47e-06	5,78E+02	GGCGTCATCATCGCC
Glyma.17G086300			gma_08G138901_AAO2	353	373	+	12.7143	3.96e-06	1,55E+02	CGTGGCGTCATCATCGCCACC
Glyma.05G040500			gma_08G138901_AAO2	353	373	+	12.7143	3.96e-06	1,54E+02	CGTGGCGTCATCATCGCCACC
Glyma.08G221800	GATA (G/A/T/A)	GATA	gma_08G138901_AAO2	358	376	+	12.3151	2.3e-06	9,06E+02	CGTCATCATCGCCACCCGC
Glyma.17G099100	TEOSINTE BRANCHED 1, CYCLOIDEA, PCF1	TCP	gma_08G138901_AAO2	1908	1927	+	471.765	3.15e-06	1,24E+02	TTTGAAAAGGTAGGGCCAAC
Glyma.10G285900			gma_08G138901_AAO2	1917	1927	+	140.794	3.5e-06	1,39E+02	GTAGGGCCAAC
Glyma.01G166800	WUSCHEL-related homeobox	WOX	gma_08G138901_AAO2	488	498	+	167.051	7.55e-07	2,94E+02	TCATTCATTCA
Glyma.10G289900			gma_20G051700_AAO6	1132	1141	+	132.595	4.9e-06	1,95E+02	TGTCAATCAA
Glyma.07G192900	NAM, ATAF and CUC - domain proteins	NAC	gma_08G138901_AAO2	263	270	+	138.925	7.84e-06	3,12E+02	CACGTAAC

Information was collected from the Plant Transcriptional Regulatory Map website.

Supplementary Table 3: Transcription factor motif count in the 2000bp upstream region of *Ascorbate Oxidase* genes from *Glycine max*.

Sequence	TF	Number of motifs
gma_AAO3	BBR-BPC	14
	TALE	5
	MIKC_MADS	1
	C2H2	2
	Dof	1
	AP2/ERF	1
	HD-ZIP	5
	bHLH	3
gma_AAO6	BBR-BPC	14
	MIKC_MADS	7
	AP2/ERF	9
	TALE	4
	Dof	3
	WOX	1
	Nin-like	1
	MYB_related	1
gma_AAO5	BBR-BPC	2
	C2H2	2
	bZIP	1
	bHLH	1
	ARF	1
	MIKC_MADS	1
gma_AAO4	bZIP	3
	MYB_related	2
	NAC	1
	Trihelix	1
	WRKY	1
gma_AAO2	BBR-BPC	8
	bZIP	13
	TALE	3

	LBD	5
	GATA	3
	AP2/ERF	3
	TCP	2
	Dof	1
	HD-ZIP	1
	AP2/ERF	1
	bHLH	1
	WOX	1
	NAC	1
	AP2/ERF	1

Information was collected from the Plant Transcriptional Regulatory Map website.

Supplementary Table 4: Gene expression information for the *G. max* AAOs.

Stress	Log2-ratio	p-value	Abreviation
Gma_AAO3			
GM-00231 (1) *A. glycines* (DongNong 47; 24hpi) / mock treated leaf blade	0.74	0.056	*A. glycines* 24hpi
GM-00231 (2) *A. glycines* (DongNong 47; 48hpi) / mock treated leaf blade	-0.44	0.469	*A. glycines* 48hpi
GM-00231 (3) *A. glycines* (DongNong 47; 96hpi) / mock treated leaf blade	-1.84	0.026	*A. glycines* 96hpi
GM-00119 (1) *B. diazoefficiens* (5-7dpi; nodule) / B. diazoefficiens (5-7dpi; root)	-0.24	0.507	*B. diazoefficiens* 5-7dpi nodule
GM-00119 (2*) B. diazoefficiens* (14-16dpi; nodule) / B. diazoefficiens (5-7dpi; nodule)	-0.24	0.382	*B. diazoefficiens* 14-16dpi nodule
GM-00119 (3) *B. diazoefficiens* (14-16dpi; nodule) / B. diazoefficiens (14-16dpi.)	-1.16	0.001	*B. diazoefficiens* 14-16dpi nodule_2
GM-00119 (4) *B. diazoefficiens* (14-16dpi; root) / B. diazoefficiens (5-7dpi; root)	0.67	63	*B. diazoefficiens* 14-16dpi root
GM-00135 (1) *B. japonicum* / mock treated L76-1988 leaf samples	-0.62	0.052	*B. japonicum* L76-1988-leaf
GM-00311 (1) C. ilicicola (hypocotyl) / mock inoculated hypocotyl samples	0.00	0.990	*C. ilicicola* hypocotyl
GM-00311 (2*) C. ilicicola* (radicle) / mock inoculated radicle samples	0.35	0.037	*C. ilicicola* radicle
GM-00083 (1) *F. oxysporum* (FO36; 72hpi) / mock treated Forrest root samples	0.05	0.595	*F. oxysporum* 72hpi
GM-00083 (2) *F. oxysporum* (FO36; 96hpi) / mock treated Forrest root samples	0.04	485	*F. oxysporum* 96hpi
GM-00203 (1) *P. pachyrhizi* (BRS 184) / mock inoculated leaf samples (BRS 184)	-0.98	0.073	*P. pachyrhizi* leaf
GM-00203 (2) *P. pachyrhizi* (BRS 184-NIL-Rpp3) / mock inoculated leaf sample	-2.28	<0.001	*P. pachyrhizi* NIL-Rpp3 leaf

GM-00232 (1) *P. sojae* (Misty; 4dpi) / mock treated root samples (Misty; 4dpi)	-0.62	62	*P. sojae* 4dpi
GM-00232 (2) *P. sojae* (Misty; 7dpi) / mock treated root samples (Misty; 7dpi)	-0.47	1.000	*P. sojae* 7dpi
GM-00232 (3) *P. sojae* (Misty; 14dpi) / mock treated root samples (Misty; 14dpi)	-0.89	191	*P. sojae* 14dpi
GM-00232 (4) *P. sojae* (PI 449459; 4dpi) / mock treated root samples (PI 44945)	-0.47	0.222	*P. sojae* PI 449459 4dpi
GM-00232 (5) *P. sojae* (PI 449459; 7dpi) / mock treated root samples (PI 44945)	-1.38	0.008	*P. sojae* PI 449459 7dpi
GM-00232 (6) *P. sojae* (PI 449459; 14dpi) / mock treated root samples (PI 44945)	-3.13	<0.001	*P. sojae* PI 449459 14dpi
GM-00232 (7) *P. sojae* (PI 449459; 21dpi) / mock treated root samples (PI 44945)	-0.50	0.262	*P. sojae* PI 449459 21dpi
GM-00185 (5) SMV (Williams 82; 2hpi) / SMV (Williams 82; 0hpi)	2.87	0.002	SMV 2hpi
GM-00185 (6) SMV (Williams 82; 4hpi) / SMV (Williams 82; 0hpi	2.76	0.005	SMV 4hpi
GM-00185 (7) SMV (Williams 82; 6hpi) / SMV (Williams 82; 0hpi)	1.76	0.003	SMV 6hpi
GM-00185 (8) SMV (Williams 82; 8hpi) / SMV (Williams 82; 0hpi)	2.52	<0.001	SMV 8hpi
GM-00089 (1) dehydration (Benning; 6h) / mock treated Benning leaf samples	-3.25	<0.001	Dehydration 6h
GM-00089 (2) dehydration (Benning; 12h) / mock treated Benning leaf samples	-3.69	0.002	Dehydration 12h
GM-00089 (3) dehydration (Benning; 24h) / mock treated Benning leaf samples	-6.53	<0.001	Dehydration 24h
GM-00230 (1) drought (5d; leaf blade) / untreated leaf blade samples (5d)	-0.34	0.008	Drought 5d leaf
GM-00230 (2) drought (5d; root) / untreated root samples (5d)	-2.96	<0.001	Drought 5d root

GM-00230 (3) drought (6d; leaf blade) / untreated leaf blade samples (5d)	-0.34	0.019	Drought 6d leaf
GM-00230 (4) drought (6d; root) / untreated root samples (5d)	-2.51	<0.001	Drought 6d root
GM-00264 (1) heat (8h) / untreated trifoliolate leaf samples	2.68	<0.001	Heat 8h
GM-00264 (2) heat (24h) / untreated trifoliolate leaf samples	-0.15	0.161	Heat 24h
GM-00124 (1) salt (1h; leaf) / untreated unifoliolate leaf samples	2.38	<0.001	Salt 1h leaf
GM-00124 (2) salt (1h; root) / untreated root samples	4.50	<0.001	Salt 1h root
GM-00124 (3) salt (2h; leaf) / untreated unifoliolate leaf samples	1.54	<0.001	Salt 2h leaf
GM-00124 (4) salt (2h; root) / untreated root samples	3.36	<0.001	Salt 2h root
GM-00124 (5) salt (4h; leaf) / untreated unifoliolate leaf samples	0.46	0.049	Salt 4h leaf
GM-00124 (6) salt (4h; root) / untreated root samples	1.19	<0.001	Salt 4h root
GM-00230 (9) submergence (1d; leaf blade) / untreated leaf blade samples (0d)	-0.83	0.005	Submergence 1d leaf
GM-00230 (10) submergence (1d; root) / untreated root samples (0d)	-4.90	<0.001	Submergence 1d root
GM-00230 (11) submergence (2d; leaf blade) / untreated leaf blade samples (0d)	-1.58	<0.001	Submergence 2d leaf
GM-00230 (12) submergence (2d; root) / untreated root samples (0d)	-5.64	<0.001	Submergence 2d root
GM-00230 (13) submergence (3d; leaf blade) / untreated leaf blade samples (0d)	-1.37	<0.001	Submergence 3d leaf
GM-00230 (14) submergence (3d; root) / untreated root samples (0d)	-5.12	<0.001	Submergence 3d root

Gma_AAO6			Abreviation
GM-00231 (1) *A. glycines* (DongNong 47; 24hpi) / mock treated leaf blade	0.00	0.992	*A. glycines* 24hpi
GM-00231 (2) *A. glycines* (DongNong 47; 48hpi) / mock treated leaf blade	0.00	1.000	*A. glycines* 48hpi
GM-00231 (3) *A. glycines* (DongNong 47; 96hpi) / mock treated leaf blade	-0.03	0.806	*A. glycines* 96hpi
GM-00119 (1) *B. diazoefficiens* (5-7dpi; nodule) / B. d*iazoefficiens* (5-7dpi; root)	-0.44	0.169	*B. diazoefficiens* 5-7dpi nodule
GM-00119 (2) *B. diazoefficiens* (14-16dpi; nodule) / *B. diazoefficiens* (5-7dpi; nodule)	-0.69	0.005	*B. diazoefficiens* 14-16dpi nodule
GM-00119 (3) *B. diazoefficiens* (14-16dpi; nodule) / *B. diazoefficiens* (14-16dpi;	-1.53	<0.001	*B. diazoefficiens* 14-16dpi nodule_2
GM-00119 (4) *B. diazoefficiens* (14-16dpi; root) / *B. diazoefficiens* (5-7dpi; root)	0.40	0.204	*B. diazoefficiens* 14-16dpi root
GM-00135 (1) *B. japonicum* / mock treated L76-1988 leaf samples	0.10	0.216	*B. japonicum* L76-1988-leaf
GM-00311 (1) *C. ilicicola* (hypocotyl) / mock inoculated hypocotyl samples	0.80	0.008	*C. ilicicola* hypocotyl
GM-00311 (2) *C. ilicicola* (radicle) / mock inoculated radicle samples	0.86	0.007	*C. ilicicola* radicle
GM-00083 (1) *F. oxysporum* (FO36; 72hpi) / mock treated Forrest root samples	-0.05	0.437	*F. oxysporum* 72hpi
GM-00083 (2) *F. oxysporum* (FO36; 96hpi) / mock treated Forrest root samples	-0.05	0.453	*F. oxysporum* 96hpi
GM-00203 (1) *P. pachyrhizi* (BRS 184) / mock inoculated leaf samples (BRS 184)	0.57	0.044	*P. pachyrhizi* leaf
GM-00203 (2) *P. pachyrhizi* (BRS 184-NIL-Rpp3) / mock inoculated leaf sample	0.12	0.630	*P. pachyrhizi* NIL-Rpp3 leaf
GM-00232 (1) *P. sojae* (Misty; 4dpi) / mock treated root samples (Misty; 4dpi)	1.68	0.038	*P. sojae* 4dpi

GM-00232 (2) *P. sojae* (Misty; 7dpi) / mock treated root samples (Misty; 7dpi)	2.40	1.000	*P. sojae* 7dpi
GM-00232 (3) *P. sojae* (Misty; 14dpi) / mock treated root samples (Misty; 14dpi)	2.26	0.008	*P. sojae* 14dpi
GM-00232 (4) *P. sojae* (PI 449459; 4dpi) / mock treated root samples	01.05	0.015	*P. sojae* PI 449459 4dpi
GM-00232 (5) *P. sojae* (PI 449459; 7dpi) / mock treated root samples (PI 44945)	1.50	0.040	*P. sojae* PI 449459 7dpi
GM-00232 (6) *P. sojae* (PI 449459; 14dpi) / mock treated root samples (PI 44945)	1.27	0.003	*P. sojae* PI 449459 14dpi
GM-00232 (7) P. sojae (PI 449459; 21dpi) / mock treated root samples (PI 44945)	0.07	0.703	P. sojae PI 449459 21dpi
GM-00185 (5) SMV (Williams 82; 2hpi) / SMV (Williams 82; 0hpi)	0.23	0.187	SMV 2hpi
GM-00185 (6) SMV (Williams 82; 4hpi) / SMV (Williams 82; 0hpi)	0.02	0.895	SMV 4hpi
GM-00185 (7) SMV (Williams 82; 6hpi) / SMV (Williams 82; 0hpi)	0.03	0.850	SMV 6hpi
GM-00185 (8) SMV (Williams 82; 8hpi) / SMV (Williams 82; 0hpi)	0.03	0.867	SMV 8hpi
GM-00089 (1) dehydration (Benning; 6h) / mock treated Benning leaf samples	-0.46	0.192	Dehydration 6h
GM-00089 (2) dehydration (Benning; 12h) / mock treated Benning leaf samples	-0.61	0.053	Dehydration 12h
GM-00089 (3) dehydration (Benning; 24h) / mock treated Benning leaf samples	-0.56	0.008	Dehydration 24h
GM-00230 (1) drought (5d; leaf blade) / untreated leaf blade samples (5d)	0.01	0.672	Drought 5d leaf
GM-00230 (2) drought (5d; root) / untreated root samples (5d)	-0.04	0.757	Drought 5d root
GM-00230 (3) drought (6d; leaf blade) / untreated leaf blade samples (5d)	-0.02	0.750	Drought 6d leaf

GM-00230 (4) drought (6d; root) / untreated root samples (5d)	0.09	0.339	Drought 6d root
GM-00264 (1) heat (8h) / untreated trifoliolate leaf samples	0.09	0.267	Heat 8h
GM-00264 (2) heat (24h) / untreated trifoliolate leaf samples	0.03	0.594	Heat 24h
GM-00124 (1) salt (1h; leaf) / untreated unifoliolate leaf samples	0.46	0.016	Salt 1h leaf
GM-00124 (2) salt (1h; root) / untreated root samples	4.40	<0.001	Salt 1h root
GM-00124 (3) salt (2h; leaf) / untreated unifoliolate leaf samples	0.07	0.584	Salt 2h leaf
GM-00124 (4) salt (2h; root) / untreated root samples	3.15	<0.001	Salt 2h root
GM-00124 (5) salt (4h; leaf) / untreated unifoliolate leaf samples	0.00	0.989	Salt 4h leaf
GM-00124 (6) salt (4h; root) / untreated root samples	1.45	<0.001	Salt 4h root
GM-00230 (9) submergence (1d; leaf blade) / untreated leaf blade samples (0d)	-0.03	0.298	Submergence 1d leaf
GM-00230 (10) submergence (1d; root) / untreated root samples (0d)	-2.01	<0.001	Submergence 1d root
GM-00230 (11) submergence (2d; leaf blade) / untreated leaf blade samples (0d)	-0.04	0.669	Submergence 2d leaf
GM-00230 (12) submergence (2d; root) / untreated root samples (0d)	-2.06	<0.001	Submergence 2d root
GM-00230 (13) submergence (3d; leaf blade) / untreated leaf blade samples (0d)	-0.02	0.790	Submergence 3d leaf
GM-00230 (14) submergence (3d; root) / untreated root samples (0d)	-0.33	0.003	Submergence 3d root
Gma_AAO5			**Abreviation**

GM-00231 (1) *A. glycines* (DongNong 47; 24hpi) / mock treated leaf blade	0.77	0.017	*A. glycines* 24hpi
GM-00231 (2) *A. glycines* (DongNong 47; 48hpi) / mock treated leaf blade	0.18	0.682	*A. glycines* 48hpi
GM-00231 (3) *A. glycines* (DongNong 47; 96hpi) / mock treated leaf blade	-1.78	0.024	*A. glycines* 96hpi
GM-00119 (1) *B. diazoefficiens* (5-7dpi; nodule) / *B. diazoefficiens* (5-7dpi; root)	0.47	0.017	*B. diazoefficiens* 5-7dpi nodule
GM-00119 (2) *B. diazoefficiens* (14-16dpi; nodule) / *B. diazoefficiens* (5-7dpi; nodule)	-0.42	0.017	*B. diazoefficiens* 14-16dpi nodule
GM-00119 (3) *B. diazoefficiens* (14-16dpi; nodule) / *B. diazoefficiens* (14-16dpi)	-0.92	0.003	*B. diazoefficiens* 14-16dpi nodule_2
GM-00119 (4) *B. diazoefficiens* (14-16dpi; root) / *B. diazoefficiens* (5-7dpi; root)	0.96	0.003	*B. diazoefficiens* 14-16dpi root
GM-00135 (1) *B. japonicum* / mock treated L76-1988 leaf samples	-0.14	-1.10	*B. japonicum* L76-1988-leaf
GM-00311 (1) *C. ilicicola* (hypocotyl) / mock inoculated hypocotyl samples	-0.14	0.260	*C. ilicicola* hypocotyl
GM-00311 (2) *C. ilicicola* (radicle) / mock inoculated radicle samples	0.50	0.103	*C. ilicicola* radicle
GM-00083 (1) *F. oxysporum* (FO36; 72hpi) / mock treated Forrest root samples	0.06	0.384	*F. oxysporum* 72hpi
GM-00083 (2) *F. oxysporum* (FO36; 96hpi) / mock treated Forrest root samples	0.00	0.971	*F. oxysporum* 96hpi
GM-00203 (1) *P. pachyrhizi* (BRS 184) / mock inoculated leaf samples (BRS 184)	01.02	0.030	*P. pachyrhizi* leaf
GM-00203 (2) *P. pachyrhizi* (BRS 184-NIL-Rpp3) / mock inoculated leaf sample	0.66	0.056	*P. pachyrhizi* NIL-Rpp3 leaf
GM-00232 (1) *P. sojae* (Misty; 4dpi) / mock treated root samples (Misty; 4dpi)	-0.91	0.023	*P. sojae* 4dpi
GM-00232 (2) *P. sojae* (Misty; 7dpi) / mock treated root samples (Misty; 7dpi)	-0.22	1.000	*P. sojae* 7dpi

GM-00232 (3) *P. sojae* (Misty; 14dpi) / mock treated root samples (Misty; 14dpi)	-1.52	0.059	*P. sojae* 14dpi
GM-00232 (4) *P. sojae* (PI 449459; 4dpi) / mock treated root samples (PI 44945)	-1.45	0.001	*P. sojae* PI 449459 4dpi
GM-00232 (5) *P. sojae* (PI 449459; 7dpi) / mock treated root samples (PI 44945)	-0.48	0.071	*P. sojae* PI 449459 7dpi
GM-00232 (6) *P. sojae* (PI 449459; 14dpi) / mock treated root samples (PI 44945)	-2.16	0.017	*P. sojae* PI 449459 14dpi
GM-00232 (7) *P. sojae* (PI 449459; 21dpi) / mock treated root samples (PI 44945)	-0.21	0.485	*P. sojae* PI 449459 21dpi
GM-00185 (5) SMV (Williams 82; 2hpi) / SMV (Williams 82; 0hpi)	1.94	<0.001	SMV 2hpi
GM-00185 (6) SMV (Williams 82; 4hpi) / SMV (Williams 82; 0hpi)	1.67	0.005	SMV 4hpi
GM-00185 (7) SMV (Williams 82; 6hpi) / SMV (Williams 82; 0hpi)	1.64	0.003	SMV 6hpi
GM-00185 (8) SMV (Williams 82; 8hpi) / SMV (Williams 82; 0hpi)	1.39	0.004	SMV 8hpi
GM-00089 (1) dehydration (Benning; 6h) / mock treated Benning leaf samples	-2.90	0.001	Dehydration 6h
GM-00089 (2) dehydration (Benning; 12h) / mock treated Benning leaf samples	-2.66	0.001	Dehydration 12h
GM-00089 (3) dehydration (Benning; 24h) / mock treated Benning leaf samples	-3.96	<0.001	Dehydration 24h
GM-00230 (1) drought (5d; leaf blade) / untreated leaf blade samples (5d)	-0.88	<0.001	Drought 5d leaf
GM-00230 (2) drought (5d; root) / untreated root samples (5d)	-1.99	<0.001	Drought 5d root
GM-00230 (3) drought (6d; leaf blade) / untreated leaf blade samples (5d)	-1.32	<0.001	Drought 6d leaf
GM-00230 (4) drought (6d; root) / untreated root samples (5d)	-1.94	<0.001	Drought 6d root

GM-00264 (1) heat (8h) / untreated trifoliolate leaf samples	02.05	0.006	Heat 8h
GM-00264 (2) heat (24h) / untreated trifoliolate leaf samples	-0.06	0.769	Heat 24h
GM-00124 (1) salt (1h; leaf) / untreated unifoliolate leaf samples	1.41	<0.001	Salt 1h leaf
GM-00124 (2) salt (1h; root) / untreated root samples	2.39	<0.001	Salt 1h root
GM-00124 (3) salt (2h; leaf) / untreated unifoliolate leaf samples	-0.08	0.398	Salt 2h leaf
GM-00124 (4) salt (2h; root) / untreated root samples	0.98	0.002	Salt 2h root
GM-00124 (5) salt (4h; leaf) / untreated unifoliolate leaf samples	0.56	0.008	Salt 4h leaf
GM-00124 (6) salt (4h; root) / untreated root samples	0.00	0.955	Salt 4h root
GM-00230 (9) submergence (1d; leaf blade) / untreated leaf blade samples (0d)	-0.89	<0.001	Submergence 1d leaf
GM-00230 (10) submergence (1d; root) / untreated root samples (0d)	-4.05	-16.49	Submergence 1d root
GM-00230 (11) submergence (2d; leaf blade) / untreated leaf blade samples (0d)	-1.77	<0.001	Submergence 2d leaf
GM-00230 (12) submergence (2d; root) / untreated root samples (0d)	-4.30	<0.001	Submergence 2d root
GM-00230 (13) submergence (3d; leaf blade) / untreated leaf blade samples (0d)	-1.61	<0.001	Submergence 3d leaf
GM-00230 (14) submergence (3d; root) / untreated root samples (0d)	-4.37	<0.001	Submergence 3d root
Gma_AAO4			**Abreviation**
GM-00231 (1) *A. glycines* (DongNong 47; 24hpi) / mock treated leaf blade	-0.35	0.572	A. *glycines* 24hpi

GM-00231 (2) *A. glycines* (DongNong 47; 48hpi) / mock treated leaf blade	0.45	0.075	*A. glycines* 48hpi
GM-00231 (3) *A. glycines* (DongNong 47; 96hpi) / mock treated leaf blade	0.02	0.987	*A. glycines* 96hpi
GM-00119 (1) *B. diazoefficiens* (5-7dpi; nodule) / *B. diazoefficiens* (5-7dpi; root)	-0.02	0.463	*B. diazoefficiens* 5-7dpi nodule
GM-00119 (2) *B. diazoefficiens* (14-16dpi; nodule) / *B. diazoefficiens* (5-7dpi; nodule)	0.00	0.868	*B. diazoefficiens* 14-16dpi nodule
GM-00119 (3) *B. diazoefficiens* (14-16dpi; nodule) / *B. diazoefficiens* (14-16dpi)	-0.01	0.869	*B. diazoefficiens* 14-16dpi nodule_2
GM-00119 (4) *B. diazoefficiens* (14-16dpi; root) / *B. diazoefficiens* (5-7dpi; root)	-0.01	0.918	*B. diazoefficiens* 14-16dpi root
GM-00135 (1) *B. japonicum* / mock treated L76-1988 leaf samples	-1.58	0.090	*B. japonicum* L76-1988-leaf
GM-00311 (1) *C. ilicicola* (hypocotyl) / mock inoculated hypocotyl samples	0.00	1.000	*C. ilicicola* hypocotyl
GM-00311 (2) *C. ilicicola* (radicle) / mock inoculated radicle samples	0.00	0.951	*C. ilicicola* radicle
GM-00083 (1) *F. oxysporum* (FO36; 72hpi) / mock treated Forrest root samples	0.18	0.128	*F. oxysporum* 72hpi
GM-00083 (2) *F. oxysporum* (FO36; 96hpi) / mock treated Forrest root samples	0.02	0.849	*F. oxysporum* 96hpi
GM-00203 (1) *P. pachyrhizi* (BRS 184) / mock inoculated leaf samples (BRS 184)	0.08	0.587	*P. pachyrhizi* leaf
GM-00203 (2) *P. pachyrhizi* (BRS 184-NIL-Rpp3) / mock inoculated leaf sample	-0.01	0.964	*P. pachyrhizi* NIL-Rpp3 leaf
GM-00232 (1) *P. sojae* (Misty; 4dpi) / mock treated root samples (Misty; 4dpi)	-0.08	0.454	*P. sojae* 4dpi
GM-00232 (2) *P. sojae* (Misty; 7dpi) / mock treated root samples (Misty; 7dpi)	0.00	1.000	*P. sojae* 7dpi
GM-00232 (3) *P. sojae* (Misty; 14dpi) / mock treated root samples (Misty; 14dpi)	0.00	1.000	*P. sojae* 14dpi

GM-00232 (4) *P. sojae* (PI 449459; 4dpi) / mock treated root samples	0.02	0.886	*P. sojae* PI 449459 4dpi
GM-00232 (5) *P. sojae* (PI 449459; 7dpi) / mock treated root samples (PI 44945)	0.15	0.079	*P. sojae* PI 449459 7dpi
GM-00232 (6) *P. sojae* (PI 449459; 14dpi) / mock treated root samples (PI 44945)	0.05	0.500	*P. sojae* PI 449459 14dpi
GM-00232 (7) *P. sojae* (PI 449459; 21dpi) / mock treated root samples (PI 44945)	0.01	0.953	*P. sojae* PI 449459 21dpi
GM-00185 (5) SMV (Williams 82; 2hpi) / SMV (Williams 82; 0hpi)	-0.07	0.549	SMV 2hpi
GM-00185 (6) SMV (Williams 82; 4hpi) / SMV (Williams 82; 0hpi)	1.73	0.002	SMV 4hpi
GM-00185 (7) SMV (Williams 82; 6hpi) / SMV (Williams 82; 0hpi)	2.82	0.042	SMV 6hpi
GM-00185 (8) SMV (Williams 82; 8hpi) / SMV (Williams 82; 0hpi)	03.07	0.032	SMV 8hpi
GM-00089 (1) dehydration (Benning; 6h) / mock treated Benning leaf samples	-1.96	0.008	Dehydration 6h
GM-00089 (2) dehydration (Benning; 12h) / mock treated Benning leaf samples	-2.22	<0.001	Dehydration 12h
GM-00089 (3) dehydration (Benning; 24h) / mock treated Benning leaf samples	-2.99	<0.001	Dehydration 24h
GM-00230 (1) drought (5d; leaf blade) / untreated leaf blade samples (5d)	0.0	1.000	Drought 5d leaf
GM-00230 (2) drought (5d; root) / untreated root samples (5d)	0.00	1.000	Drought 5d root
GM-00230 (3) drought (6d; leaf blade) / untreated leaf blade samples (5d)	0.00	1.000	Drought 6d leaf
GM-00230 (4) drought (6d; root) / untreated root samples (5d)	0.00	1.000	Drought 6d root
GM-00264 (1) heat (8h) / untreated trifoliolate leaf samples	3.87	0.020	Heat 8h

GM-00264 (2) heat (24h) / untreated trifoliolate leaf samples	0.90	0.016	Heat 24h
GM-00124 (1) salt (1h; leaf) / untreated unifoliolate leaf samples	-0.68	0.003	Salt 1h leaf
GM-00124 (2) salt (1h; root) / untreated root samples	0.32	0.060	Salt 1h root
GM-00124 (3) salt (2h; leaf) / untreated unifoliolate leaf samples	-0.31	0.073	Salt 2h leaf
GM-00124 (4) salt (2h; root) / untreated root samples	0.55	0.018	Salt 2h root
GM-00124 (5) salt (4h; leaf) / untreated unifoliolate leaf samples	-0.54	0.006	Salt 4h leaf
GM-00124 (6) salt (4h; root) / untreated root samples	0.23	0.130	Salt 4h root
GM-00230 (9) submergence (1d; leaf blade) / untreated leaf blade samples (0d)	-0.81	0.007	Submergence 1d leaf
GM-00230 (10) submergence (1d; root) / untreated root samples (0d)	-0.03	0.844	Submergence 1d root
GM-00230 (11) submergence (2d; leaf blade) / untreated leaf blade samples (0d)	-0.81	0.008	Submergence 2d leaf
GM-00230 (12) submergence (2d; root) / untreated root samples (0d)	-0.03	0.852	Submergence 2d root
GM-00230 (13) submergence (3d; leaf blade) / untreated leaf blade samples (0d)	-0.77	0.005	Submergence 3d leaf
GM-00230 (14) submergence (3d; root) / untreated root samples (0d)	-0.03	0.843	Submergence 3d root
Gma_AAO2			**Abreviation**
GM-00231 (1) *A. glycines* (DongNong 47; 24hpi) / mock treated leaf blade	0.97	0.002	*A. glycines* 24hpi
GM-00231 (2) *A. glycines* (DongNong 47; 48hpi) / mock treated leaf blade	0.48	0.001	*A. glycines* 48hpi

GM-00231 (3) *A. glycines* (DongNong 47; 96hpi) / mock treated leaf blade	-0.55	0.091	*A. glycines* 96hpi
GM-00119 (1) *B. diazoefficiens* (5-7dpi; nodule) / *B. diazoefficiens* (5-7dpi; root)	-0.37	0.132	*B. diazoefficiens* 5-7dpi nodule
GM-00119 (2) *B. diazoefficiens* (14-16dpi; nodule) / *B. diazoefficiens* (5-7dpi; nodule)	-0.18	0.070	*B. diazoefficiens* 14-16dpi nodule
GM-00119 (3) *B. diazoefficiens* (14-16dpi; nodule) / *B. diazoefficiens* (14-16dpi)	-1.01	0.001	*B. diazoefficiens* 14-16dpi nodule_2
GM-00119 (4) *B. diazoefficiens* (14-16dpi; root) / *B. diazoefficiens* (5-7dpi; root)	0.46	0.105	*B. diazoefficiens* 14-16dpi root
GM-00135 (1) *B. japonicum* / mock treated L76-1988 leaf samples	-0.06	0.609	*B. japonicum* L76-1988-leaf
GM-00311 (1) *C. ilicicola* (hypocotyl) / mock inoculated hypocotyl samples	-0.17	0.401	*C. ilicicola* hypocotyl
GM-00311 (2) *C. ilicicola* (radicle) / mock inoculated radicle samples	-0.85	0.002	*C. ilicicola* radicle
GM-00083 (1) *F. oxysporum* (FO36; 72hpi) / mock treated Forrest root samples	-0.04	0.727	*F. oxysporum* 72hpi
GM-00083 (2) *F. oxysporum* (FO36; 96hpi) / mock treated Forrest root samples	0.15	0.239	*F. oxysporum* 96hpi
GM-00203 (1) *P. pachyrhizi* (BRS 184) / mock inoculated leaf samples (BRS 184)	0.12	0.794	*P. pachyrhizi* leaf
GM-00203 (2) P. pachyrhizi (BRS 184-NIL-Rpp3) / mock inoculated leaf sample	1.92	0.002	P. pachyrhizi NIL-Rpp3 leaf
GM-00232 (1) *P. sojae* (Misty; 4dpi) / mock treated root samples (Misty; 4dpi)	1.28	0.074	*P. sojae* 4dpi
GM-00232 (2) *P. sojae* (Misty; 7dpi) / mock treated root samples (Misty; 7dpi)	1.04	1.000	*P. sojae* 7dpi
GM-00232 (3) *P. sojae* (Misty; 14dpi) / mock treated root samples (Misty; 14dpi)	1.14	0.001	*P. sojae* 14dpi
GM-00232 (4) *P. sojae* (PI 449459; 4dpi) / mock treated root samples (PI 44945	0.43	0.362	*P. sojae* PI 449459 4dpi

GM-00232 (5) *P. sojae* (PI 449459; 7dpi) / mock treated root samples (PI 44945)	0.75	0.047	*P. sojae* PI 449459 7dpi
GM-00232 (6) *P. sojae* (PI 449459; 14dpi) / mock treated root samples (PI 44945)	0.27	0.430	*P. sojae* PI 449459 14dpi
GM-00232 (7) *P. sojae* (PI 449459; 21dpi) / mock treated root samples (PI 44945)	-1.24	0.011	*P. sojae* PI 449459 21dpi
GM-00185 (5) SMV (Williams 82; 2hpi) / SMV (Williams 82; 0hpi)	0.30	0.050	SMV 2hpi
GM-00185 (6) SMV (Williams 82; 4hpi) / SMV (Williams 82; 0hpi)	0.17	0.081	SMV 4hpi
GM-00185 (7) SMV (Williams 82; 6hpi) / SMV (Williams 82; 0hpi)	0.20	0.106	SMV 6hpi
GM-00185 (8) SMV (Williams 82; 8hpi) / SMV (Williams 82; 0hpi)	0.14	0.116	SMV 8hpi
GM-00089 (1) dehydration (Benning; 6h) / mock treated Benning leaf samples	1.68	0.004	Dehydration 6h
GM-00089 (2) dehydration (Benning; 12h) / mock treated Benning leaf samples	0.17	0.747	Dehydration 12h
GM-00089 (3) dehydration (Benning; 24h) / mock treated Benning leaf samples	0.86	0.021	Dehydration 24h
GM-00230 (1) drought (5d; leaf blade) / untreated leaf blade samples (5d)	-0.17	0.017	Drought 5d leaf
GM-00230 (2) drought (5d; root) / untreated root samples (5d)	-1.18	<0.001	Drought 5d root
GM-00230 (3) drought (6d; leaf blade) / untreated leaf blade samples (5d)	0.02	0.665	Drought 6d leaf
GM-00230 (4) drought (6d; root) / untreated root samples (5d)	-0.76	0.004	Drought 6d root
GM-00264 (1) heat (8h) / untreated trifoliolate leaf samples	2.14	0.002	Heat 8h
GM-00264 (2) heat (24h) / untreated trifoliolate leaf samples	0.75	0.018	Heat 24h

GM-00124 (1) salt (1h; leaf) / untreated unifoliolate leaf samples	1.04	0.001	Salt 1h leaf
GM-00124 (2) salt (1h; root) / untreated root samples	3.10	<0.001	Salt 1h root
GM-00124 (3) salt (2h; leaf) / untreated unifoliolate leaf samples	0.23	0.033	Salt 2h leaf
GM-00124 (4) salt (2h; root) / untreated root samples	0.93	0.001	Salt 2h root
GM-00124 (5) salt (4h; leaf) / untreated unifoliolate leaf samples	-0.06	0.511	Salt 4h leaf
GM-00124 (6) salt (4h; root) / untreated root samples	1.48	<0.001	Salt 4h root
GM-00230 (9) submergence (1d; leaf blade) / untreated leaf blade samples (0d)	1.18	<0.001	Submergence 1d leaf
GM-00230 (10) submergence (1d; root) / untreated root samples (0d)	-0.14	0.238	Submergence 1d root
GM-00230 (11) submergence (2d; leaf blade) / untreated leaf blade samples (0d)	2.25	<0.001	Submergence 2d leaf
GM-00230 (12) submergence (2d; root) / untreated root samples (0d)	-0.90	0.001	Submergence 2d root
GM-00230 (13) submergence (3d; leaf blade) / untreated leaf blade samples (0d)	3.34	<0.001	Submergence 3d leaf
GM-00230 (14) submergence (3d; root) / untreated root samples (0d)	0.12	0.285	Submergence 3d root

Information obtained using mRNA-Seq data from Genevestigator Plants software , March, 2023, (https://genevestigator.com/).

4 CONCLUSÃO

Através das análises realizadas nesse trabalho, concluímos que as AAOs de Fabaceaes, são uma pequena família de proteínas. Foram encontrados cinco genes de *AAO* em *G. max*, e uma média de dois a seis genes nas demais espécies, com a maioria das espécies apresentando quatro genes. Observamos também que as proteínas AAOs da família de plantas Fabaceae, apresentam alto grau de conservação e semelhança tanto em sua estrutura gênica, quanto em relação a motivos e domínios de função. A análise de filogenia nos mostrou que as sequências de AAO analisadas, foram distribuídas em dois grupos principais, com valores de *bootstrap* estatisticamente significativos, e posteriormente subdividas em grupos contendo somente AAO de Fabaceae. Além disso, sequências de mono e dicotiledônias estiveram presentem em ambos os grupos principais, demonstrando que a diversificação dessas AAOs ocorreu antes mesmo da divergência entre os dois grupos. Ainda, a presença de sequências de AAO em *S. moellendorffii*, uma planta ancestral, demonstram a antiga origem da AAO. A análise de distribuição gênica de *AAO* entre *G. max* e *G. soja* demonstrou que a maioria dos genes *AAO* dessas espécies estão presentes nos mesmos cromossomos. Analisando a região promotora dos cinco genes de *AAO* em *G. max*, encontramos diversos motivos de ligação a fatores de transcrição responsivos as mais diversas situações, como crescimento/desenvolvimento, nodulação, estresses bióticos e abióticos, demonstrando a diversidade de funções desses genes em *G. max*. Além disso, a análise de expressão gênica demonstrou um padrão de expressão diferencial nos genes *AAO* de *G. max* frente a estresses bióticos e abióticos. No entanto, em conjunto com os dados de elementos *cis*-reguladores, acreditamos que os genes Gma-*AAO2*, Gma-*AAO3* e Gma-*AAO4* estão envolvidas com situações de estresses bióticos, enquanto o gene Gma-*AAO5* pode estar relacionado a nodulação.

Nosso trabalho conseguiu caracterizar as AAO de Fabaceae, e servirá como base para aprofundar a compreensão e possível papel de alguns genes *AAO* em *G. max*. Como perspectivas futuras, nosso trabalho indicou genes com expressão diferencial, que podem ser selecionados como biomarcadores ou alvos para análises posteriores, não só em *G. max* mas em outras espécies de Fabaceae, com a finalidade de compreender, e melhorar os atributos dessas plantas.

REFERÊNCIAS

Abdelgawad KF, El-Mogy MM, Mohamed MIA, et al (2019) Increasing ascorbic acid content and salinity tolerance of cherry tomato plants by suppressed expression of the ascorbate oxidase gene. Agronomy 9:. https://doi.org/10.3390/agronomy9020051

Al-Madhoun AS, Sanmartin M, Kanellis AK (2003) Expression of ascorbate oxidase isoenzymes in cucurbits and during development and ripening of melon fruit. Postharvest Biol Technol 27:137–146. https://doi.org/10.1016/S0925-5214(02)00090-X

Alves MS, Soares ZG, Vidigal PMP, et al (2015) Differential expression of four soybean bZIP genes during Phakopsora pachyrhizi infection. Funct Integr Genomics 15:685–696. https://doi.org/10.1007/S10142-015-0445-0

Asao H, Yoshida K, Nishi Y, Shinmyo A (2003) Wound-responsive cis-element in the 5′-upstream region of cucumber ascorbate oxidase gene. Biosci Biotechnol Biochem 67:271–277. https://doi.org/10.1271/bbb.67.271

Bailey TL, Johnson J, Grant CE, Noble WS (2015) The MEME Suite. Nucleic Acids Res 43:W39–W49. https://doi.org/10.1093/NAR/GKV416

Balestrini R, Ott T, Güther M, et al (2012) Ascorbate oxidase: the unexpected involvement of a “wasteful enzyme” in the symbioses with nitrogen-fixing bacteria and arbuscular mycorrhizal fungi. Plant Physiol Biochem PPB 59:71–79. https://doi.org/10.1016/J.PLAPHY.2012.07.006

Batth R, Singh K, Kumari S, Mustafiz A (2017) Transcript profiling reveals the presence of abiotic stress and developmental stage specific ascorbate oxidase genes in plants. Front Plant Sci 8:198. https://doi.org/10.3389/FPLS.2017.00198/BIBTEX

Belamkar V, Weeks NT, Bharti AK, et al (2014) Comprehensive characterization and RNA-Seq profiling of the HD-Zip transcription factor family in soybean (Glycine max) during dehydration and salt stress. BMC Genomics 15:1–25. https://doi.org/10.1186/1471-2164-15-950/TABLES/4

Bencke-Malato M, Cabreira C, Wiebke-Strohm B, et al (2014) Genome-wide annotation of the soybean WRKY family and functional characterization of genes involved in response to Phakopsora pachyrhizi infection. BMC Plant Biol 14:1–18. https://doi.org/10.1186/S12870-014-0236-0

Bertioli DJ, Cannon SB, Froenicke L, et al (2016) The genome sequences of Arachis duranensis and Arachis ipaensis, the diploid ancestors of cultivated peanut. Nat Genet 2016 484 48:438–446. https://doi.org/10.1038/ng.3517

Bertioli DJ, Jenkins J, Clevenger J, et al (2019) The genome sequence of segmental allotetraploid peanut Arachis hypogaea. Nat Genet 2019 515 51:877–884. https://doi.org/10.1038/s41588-019-0405-z

Chai M, Fan R, Huang Y, et al (2022) GmbZIP152, a Soybean bZIP Transcription Factor, Confers Multiple Biotic and Abiotic Stress Responses in Plant. Int J Mol Sci 23:. https://doi.org/10.3390/IJMS231810935

Chatzopoulou F, Sanmartin M, Mellidou I, et al (2020) Silencing of ascorbate oxidase results in reduced growth, altered ascorbic acid levels and ripening pattern in melon fruit. Plant Physiol Biochem 156:291–303. https://doi.org/10.1016/j.plaphy.2020.08.040

Chen X, Chen Z, Zhao H, et al (2014) Genome-Wide Analysis of Soybean HD-Zip Gene Family and Expression Profiling under Salinity and Drought Treatments. PLoS One 9:e87156. https://doi.org/10.1371/JOURNAL.PONE.0087156

Cheng Q, Dong L, Gao T, et al (2018) The bHLH transcription factor GmPIB1 facilitates resistance to Phytophthora sojae in Glycine max. J Exp Bot 69:2527–2541. https://doi.org/10.1093/JXB/ERY103

Chiasson DM, Loughlin PC, Mazurkiewicz D, et al (2014) Soybean SAT1 (Symbiotic Ammonium Transporter 1) encodes a bHLH transcription factor involved in nodule growth and NH4+ transport. Proc Natl Acad Sci U S A 111:4814–4819. https://doi.org/10.1073/PNAS.1312801111

Damodaran S, Dubois A, Xie J, et al (2019) GmZPR3d Interacts with GmHD-ZIP III Proteins and Regulates Soybean Root and Nodule Vascular Development. Int J Mol Sci 2019, Vol 20, Page 827 20:827. https://doi.org/10.3390/IJMS20040827

De Tullio MC, Guether M, Balestrini R (2013) Ascorbate oxidase is the potential conductor of a symphony of signaling pathways. Plant Signal Behav 8:. https://doi.org/10.4161/PSB.23213

De Vega JJ, Ayling S, Hegarty M, et al (2015) Red clover (Trifolium pratense L.) draft genome provides a platform for trait improvement. Sci Reports 2015 51 5:1–10. https://doi.org/10.1038/srep17394

Di Venere A, Nicolai E, Rosato N, et al (2011) Characterization of monomeric substates of ascorbate oxidase. FEBS J 278:1585–1593.

https://doi.org/10.1111/J.1742-4658.2011.08084.X

Diallinas G, Pateraki I, Sanmartin M, et al (1997) Melon ascorbate oxidase: Cloning of a multigene family, induction during fruit development and repression by wounding. Plant Mol Biol 34:759–770. https://doi.org/10.1023/A:1005851527227

Dong H, Tan J, Li M, et al (2019) Transcriptome analysis of soybean WRKY TFs in response to Peronospora manshurica infection. Genomics 111:1412–1422. https://doi.org/10.1016/J.YGENO.2018.09.014

Du YT, Zhao MJ, Wang CT, et al (2018) Identification and characterization of GmMYB118 responses to drought and salt stress. BMC Plant Biol 18:1–18. https://doi.org/10.1186/S12870-018-1551-7/FIGURES/10

Esaka M, Fujisawa K, Goto M, Kisu Y (1992) Regulation of ascorbate oxidase expression in pumpkin by auxin and copper. Plant Physiol 100:231–237. https://doi.org/10.1104/pp.100.1.231

Fan C, Chen Y, Long M (2008) Recurrent tandem gene duplication gave rise to functionally divergent genes in Drosophila. Mol Biol Evol 25:1451–1458. https://doi.org/10.1093/molbev/msn089

Fan CM, Wang X, Wang YW, et al (2013) Genome-Wide Expression Analysis of Soybean MADS Genes Showing Potential Function in the Seed Development. PLoS One 8:e62288. https://doi.org/10.1371/JOURNAL.PONE.0062288

Fan S, Dong L, Han D, et al (2017) GmWRKY31 and GmHDL56 Enhances Resistance to Phytophthora sojae by Regulating Defense-Related Gene Expression in Soybean. Front Plant Sci 8:. https://doi.org/10.3389/FPLS.2017.00781

Fan S, Zhang Z, Song Y, et al (2022) CRISPR/Cas9-mediated targeted mutagenesis of GmTCP19L increasing susceptibility to Phytophthora sojae in soybean. PLoS One 17:e0267502. https://doi.org/10.1371/JOURNAL.PONE.0267502

Farvardin A, González-hernández AI, Llorens E, et al (2020) The Apoplast: A Key Player in Plant Survival. Antioxidants 2020, Vol 9, Page 604 9:604. https://doi.org/10.3390/ANTIOX9070604

Farver O, Wherland S, Pecht I (1994) Intramolecular electron transfer in ascorbate oxidase is enhanced in the presence of oxygen. J Biol Chem 269:22933–22936. https://doi.org/10.1016/s0021-9258(17)31598-3

Feng S, Shi J, Hu Y, et al (2022) Genome-Wide Analysis of Soybean Lateral Organ Boundaries Domain Gene Family Reveals the Role in Phytophthora Root and

Stem Rot. Front Plant Sci 13:865165. https://doi.org/10.3389/FPLS.2022.865165/BIBTEX

Foyer CH, Kyndt T, Hancock RD (2020) Vitamin C in Plants: Novel Concepts, New Perspectives, and Outstanding Issues. Antioxidants Redox Signal 32:463–485. https://doi.org/10.1089/ars.2019.7819

Foyer CH, Noctor G (2005) Oxidant and antioxidant signalling in plants: A re-evaluation of the concept of oxidative stress in a physiological context. Plant, Cell Environ 28:1056–1071. https://doi.org/10.1111/j.1365-3040.2005.01327.x

Garchery C, Gest N, Do PT, et al (2013) A diminution in ascorbate oxidase activity affects carbon allocation and improves yield in tomato under water deficit. Plant Cell Environ 36:159–175. https://doi.org/10.1111/J.1365-3040.2012.02564.X

Garcia T, Duitama J, Zullo SS, et al (2021) Comprehensive genomic resources related to domestication and crop improvement traits in Lima bean. Nat Commun 2021 121 12:1–17. https://doi.org/10.1038/s41467-021-20921-1

Góralczyk-Bińkowska A, Jasińska A, Długoński J (2019) Characteristics And Use Of Multicopper Oxidases Enzymes. Postępy Mikrobiol - Adv Microbiol 58:7–18. https://doi.org/10.21307/pm-2019.58.1.007

Hane JK, Ming Y, Kamphuis LG, et al (2017) A comprehensive draft genome sequence for lupin (Lupinus angustifolius), an emerging health food: insights into plant–microbe interactions and legume evolution. Plant Biotechnol J 15:318–330. https://doi.org/10.1111/PBI.12615

Hao Q, Zhang L, Yang Y, et al (2019) Genome-Wide Analysis of the WOX Gene Family and Function Exploration of GmWOX18 in Soybean. Plants 2019, Vol 8, Page 215 8:215. https://doi.org/10.3390/PLANTS8070215

He Q, Cai H, Bai M, et al (2020) A Soybean bZIP Transcription Factor GmbZIP19 Confers Multiple Biotic and Abiotic Stress Responses in Plant. Int J Mol Sci 21:1–19. https://doi.org/10.3390/IJMS21134701

Hoegger PJ, Kilaru S, James TY, et al (2006) Phylogenetic comparison and classification of laccase and related multicopper oxidase protein sequences. FEBS J 273:2308–2326. https://doi.org/10.1111/J.1742-4658.2006.05247.X

Hovde BT, Daligault HE, Hanschen ER, et al (2019) Detection of Abrin-Like and Prepropulchellin-Like Toxin Genes and Transcripts Using Whole Genome Sequencing and Full-Length Transcript Sequencing of Abrus precatorius. Toxins 2019, Vol 11, Page 691 11:691. https://doi.org/10.3390/TOXINS11120691

Hu B, Jin J, Guo AY, et al (2015) GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics 31:1296–1297. https://doi.org/10.1093/BIOINFORMATICS/BTU817

Hu D, Li X, Yang Z, et al (2022a) Downregulation of a gibberellin 3β-hydroxylase enhances photosynthesis and increases seed yield in soybean. New Phytol 235:502–517. https://doi.org/10.1111/NPH.18153

Hu J, Liu M, Zhang A, et al (2022b) Co-evolved plant and blast fungus ascorbate oxidases orchestrate the redox state of host apoplast to modulate rice immunity. Mol Plant 15:1347–1366. https://doi.org/10.1016/j.molp.2022.07.001

Hufnagel B, Marques A, Soriano A, et al (2020) High-quality genome sequence of white lupin provides insight into soil exploration and seed quality. Nat Commun 2020 111 11:1–12. https://doi.org/10.1038/s41467-019-14197-9

Jiao Y, Wickett NJ, Ayyampalayam S, et al (2011) Ancestral polyploidy in seed plants and angiosperms. Nat 2011 4737345 473:97–100. https://doi.org/10.1038/nature09916

Kalyaanamoorthy S, Minh BQ, Wong TKF, et al (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 2017 146 14:587–589. https://doi.org/10.1038/nmeth.4285

Kang YJ, Kim SK, Kim MY, et al (2014) Genome sequence of mungbean and insights into evolution within Vigna species. Nat Commun 5:5443–5443. https://doi.org/10.1038/NCOMMS6443

Kang YJ, Satyawan D, Shim S, et al (2015) Draft genome sequence of adzuki bean, Vigna angularis. Sci Reports 2015 51 5:1–8. https://doi.org/10.1038/srep08069

Kato N, Esaka M (2000) Expansion of transgenic tobacco protoplasts expressing pumpkin ascorbate oxidase is more rapid than that of wild-type protoplasts. Planta 210:1018–1022. https://doi.org/10.1007/s004250050712

Katoh K, Rozewicki J, Yamada KD (2019) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform 20:1160–1166. https://doi.org/10.1093/BIB/BBX108

Kaul S, Koo HL, Jenkins J, et al (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nat 2000 4086814 408:796–815. https://doi.org/10.1038/35048692

Kerk NM, Jiang K, Feldman LJ (2000) Auxin metabolism in the root apical meristem. Plant Physiol 122:925–932. https://doi.org/10.1104/pp.122.3.925

Kim Y, Wang J, Ma C, et al (2023) GmTCP and GmNLP Underlying Nodulation Character in Soybean Depending on Nitrogen. Int J Mol Sci 24:7750. https://doi.org/10.3390/IJMS24097750/S1

Kreplak J, Madoui MA, Cápal P, et al (2019) A reference genome for pea provides insight into legume genome evolution. Nat Genet 2019 519 51:1411–1422. https://doi.org/10.1038/s41588-019-0480-1

Kumari R, Kumar S, Singh L, Hallan V (2016) Movement Protein of Cucumber Mosaic Virus Associates with Apoplastic Ascorbate Oxidase. PLoS One 11:. https://doi.org/10.1371/JOURNAL.PONE.0163320

Kuraku S, Zmasek CM, Nishimura O, Katoh K (2013) aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. Nucleic Acids Res 41:W22–W28. https://doi.org/10.1093/NAR/GKT389

Layzell DB, Hunt S, Palmer GR (1990) Mechanism of nitrogenase inhibition in soybean nodules: Pulse-modulated spectroscopy indicates that nitrogenase activity is limited by o2. Plant Physiol 92:1101–1107. https://doi.org/10.1104/pp.92.4.1101

Lee MH, Dawson CR (1973) Ascorbate Oxidase. J Biol Chem 248:6603–6609. https://doi.org/10.1016/s0021-9258(19)43396-6

Lee YC, Tsai PT, Huang XX, Tsai HL (2022) Family Members Additively Repress the Ectopic Expression of BASIC PENTACYSTEINE3 to Prevent Disorders in Arabidopsis Circadian Vegetative Development. Front Plant Sci 13:919946. https://doi.org/10.3389/FPLS.2022.919946/BIBTEX

Li H, Jiang F, Wu P, et al (2020) A High-Quality Genome Sequence of Model Legume Lotus japonicus (MG-20) Provides Insights into the Evolution of Root Nodule Symbiosis. Genes 2020, Vol 11, Page 483 11:483. https://doi.org/10.3390/GENES11050483

Li R, Xin S, Tao C, et al (2017) Cotton ascorbate oxidase promotes cell growth in cultured tobacco bright yellow-2 cells through generation of apoplast oxidation. Int J Mol Sci 18:1–15. https://doi.org/10.3390/ijms18071346

Lindström K, Mousavi SA (2020) Effectiveness of nitrogen fixation in rhizobia. Microb Biotechnol 13:1314–1335. https://doi.org/10.1111/1751-7915.13517

Little DP, Moran RC, Brenner ED, Stevenson DW (2007) Nuclear genome size in Selaginella. https://doi.org/101139/G06-138 50:351–356.

https://doi.org/10.1139/G06-138

Liu W, Zhang Y, Li W, et al (2020) Genome-wide characterization and expression analysis of soybean trihelix gene family. PeerJ 2020:e8753. https://doi.org/10.7717/PEERJ.8753/SUPP-15

Lonardi S, Muñoz-Amatriaín M, Liang Q, et al (2019) The genome of cowpea (Vigna unguiculata [L.] Walp.). Plant J 98:767–782. https://doi.org/10.1111/TPJ.14349

Magallón S, Sanderson MJ (2001) Absolute diversification rates in angiosperm clades. Evolution 55:1762–1780. https://doi.org/10.1111/J.0014-3820.2001.TB00826.X

Maphosa Y, Jideani VA, Maphosa Y, Jideani VA (2017) The Role of Legumes in Human Nutrition. Funct Food - Improv Heal through Adequate Food. https://doi.org/10.5772/INTECHOPEN.69127

Marand AP, Eveland AL, Kaufmann K, Springer NM (2023) cis-Regulatory Elements in Plant Development, Adaptation, and Evolution. https://doi.org/101146/annurev-arplant-070122-030236 74:111–137. https://doi.org/10.1146/ANNUREV-ARPLANT-070122-030236

Messerschmidt a (1990) Modelling and structural relationships. Enzyme 352:341–352

Messerschmidt A (1997) Multi-copper oxidases. World Scientific, Singapore

Messerschmidt A, Ladenstein R, Huber R, et al (1992) Refined crystal structure of ascorbate oxidase at 1.9 Å resolution. J Mol Biol 224:179–205. https://doi.org/10.1016/0022-2836(92)90583-6

Messerschmidt A, Rossi A, Ladenstein R, et al (1989) X-ray crystal structure of the blue oxidase ascorbate oxidase from Zucchini. Analysis of the polypeptide fold and a model of the copper sites and ligands. J Mol Biol 206:513–529. https://doi.org/10.1016/0022-2836(89)90498-1

Moghaddam SM, Oladzad A, Koh C, et al (2021) The tepary bean genome provides insight into evolution and domestication under heat stress. Nat Commun 2021 121 12:1–14. https://doi.org/10.1038/s41467-021-22858-x

Monfared MM, Simon MK, Meister RJ, et al (2011) Overlapping and antagonistic activities of BASIC PENTACYSTEINE genes affect a range of developmental processes in Arabidopsis. Plant J 66:1020–1031. https://doi.org/10.1111/J.1365-313X.2011.04562.X

Nakamura K, Kawabata T, Yura K, Go N (2003) Novel types of two-domain multi-

copper oxidases: possible missing links in the evolution. FEBS Lett 553:239–244. https://doi.org/10.1016/S0014-5793(03)01000-7

Nanasato Y, Akashi K, Yokota A (2005) Co-expression of cytochrome b561 and ascorbate oxidase in leaves of wild watermelon under drought and high light conditions. Plant Cell Physiol 46:1515–1524. https://doi.org/10.1093/pcp/pci164

Ohkawa J, Okada N, Shinmyo A, Takano M (1989) Primary structure of cucumber (Cucumis sativus) ascorbate oxidase deduced from cDNA sequence: homology with blue copper proteins and tissue-specific expression. Proc Natl Acad Sci U S A 86:1239–1243. https://doi.org/10.1073/pnas.86.4.1239

Ouyang S, Zhu W, Hamilton J, et al (2007) The TIGR Rice Genome Annotation Resource: improvements and new features. Nucleic Acids Res 35:D883–D887. https://doi.org/10.1093/NAR/GKL976

Ouyang W, Chen L, Ma J, et al (2022) Identification of Quantitative Trait Locus and Candidate Genes for Drought Tolerance in a Soybean Recombinant Inbred Line Population. Int J Mol Sci 23:10828. https://doi.org/10.3390/IJMS231810828/S1

Pan Z, Chen L, Wang F, et al (2019) Genome-wide identification and expression analysis of the ascorbate oxidase gene family in gossypium hirsutum reveals the critical role of ghao1a in delaying dark-induced leaf senescence. Int J Mol Sci 20:. https://doi.org/10.3390/ijms20246167

Panchy N, Lehti-Shiu M, Shiu SH (2016) Evolution of gene duplication in plants. Plant Physiol 171:2294–2316. https://doi.org/10.1104/pp.16.00523

Pignocchi C, Fletcher JM, Wilkinson JE, et al (2003) The Function of Ascorbate Oxidase in Tobacco. Plant Physiol 132:1631–1641. https://doi.org/10.1104/PP.103.022798

Roy S, Liu W, Nandety RS, et al (2020) Celebrating 20 Years of Genetic Discoveries in Legume Nodulation and Symbiotic Nitrogen Fixation. Plant Cell 32:15–41. https://doi.org/10.1105/TPC.19.00279

Roy SW, Penny D (2007) Patterns of intron loss and gain in plants: Intron loss-dominated evolution and genome-wide comparison of O. sativa and A. thaliana. Mol Biol Evol 24:171–181. https://doi.org/10.1093/molbev/msl159

Sandhu AK, Brown MR, Subramanian S, Brözel VS (2023) Bradyrhizobium diazoefficiens USDA 110 displays plasticity in the attachment phenotype when grown in different soybean root exudate compounds. Front Microbiol 14:. https://doi.org/10.3389/fmicb.2023.1190396

Sanmartin M, Pateraki I, Chatzopoulou F, Kanellis AK (2007) Differential expression of the ascorbate oxidase multigene family during fruit development and in response to stress. Planta 225:873–885. https://doi.org/10.1007/S00425-006-0399-5/FIGURES/6

Schmutz J, Cannon SB, Schlueter J, et al (2010) Genome sequence of the palaeopolyploid soybean. Nat 2010 4637278 463:178–183. https://doi.org/10.1038/nature08670

Sedivy EJ, Wu F, Hanzawa Y (2017) Soybean domestication: the origin, genetic architecture and molecular bases. New Phytol 214:539–553. https://doi.org/10.1111/nph.14418

Shi WY, Du YT, Ma J, et al (2018) The WRKY Transcription Factor GmWRKY12 Confers Drought and Salt Tolerance in Soybean. Int J Mol Sci 19:. https://doi.org/10.3390/IJMS19124087

Singh D, Chaudhary P, Taunk J, et al (2021a) Fab Advances in Fabaceae for Abiotic Stress Resilience: From 'Omics' to Artificial Intelligence. Int J Mol Sci 2021, Vol 22, Page 10535 22:10535. https://doi.org/10.3390/IJMS221910535

Singh RR, Nobleza N, Demeestere K, Kyndt T (2020a) Ascorbate Oxidase Induces Systemic Resistance in Sugar Beet Against Cyst Nematode Heterodera schachtii. Front Plant Sci 11:1–15. https://doi.org/10.3389/fpls.2020.591715

Singh RR, Pajar JA, Audenaert K, Kyndt T (2021b) Induced Resistance by Ascorbate Oxidation Involves Potentiating of the Phenylpropanoid Pathway and Improved Rice Tolerance to Parasitic Nematodes. Front Plant Sci 12:. https://doi.org/10.3389/fpls.2021.713870

Singh RR, Verstraeten B, Siddique S, et al (2020b) Ascorbate oxidation activates systemic defence against root-knot nematode Meloidogyne graminicola in rice. J Exp Bot 71:4271–4284. https://doi.org/10.1093/jxb/eraa171

Skorupa M, Szczepanek J, Yolcu S, et al (2022) Characteristic of the Ascorbate Oxidase Gene Family in Beta vulgaris and Analysis of the Role of AAO in Response to Salinity and Drought in Beet. Int J Mol Sci 23:. https://doi.org/10.3390/ijms232112773

Smirnoff N (2000) Ascorbic acid: metabolism and functions of a multi-facetted molecule. Curr Opin Plant Biol 3:229–235. https://doi.org/10.1016/s1369-5266(00)80070-9

Song H, Wang P, Hou L, et al (2016) Global Analysis of WRKY Genes and Their

Response to Dehydration and Salt Stress in Soybean. Front Plant Sci 7:. https://doi.org/10.3389/FPLS.2016.00009

Tang H, Krishnakumar V, Bidwell S, et al (2014) An improved genome release (version Mt4.0) for the model legume Medicago truncatula. BMC Genomics 15:1–14. https://doi.org/10.1186/1471-2164-15-312/FIGURES/7

Tian KH, Pan C, Yang YF, et al (2019) Differential expression of the ascorbate oxidase multigene family of Camellia sinensis in response to stress. J Hortic Sci Biotechnol 94:160–170. https://doi.org/10.1080/14620316.2018.1495581

Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res 44:W232–W235. https://doi.org/10.1093/NAR/GKW256

Turchetto-Zolet AC, Christoff AP, Kulcheski FR, et al (2016) Diversity and evolution of plant diacylglycerol acyltransferase (DGATs) unveiled by phylogenetic, gene structure and expression analyses. Genet Mol Biol 39:524–538. https://doi.org/10.1590/1678-4685-GMB-2016-0024

Valliyodan B, Cannon SB, Bayer PE, et al (2019) Construction and comparison of three reference-quality genome assemblies for soybean. Plant J 100:1066–1082. https://doi.org/10.1111/TPJ.14500

Varshney RK, Chen W, Li Y, et al (2011) Draft genome sequence of pigeonpea (Cajanus cajan), an orphan legume crop of resource-poor farmers. Nat Biotechnol 2011 301 30:83–89. https://doi.org/10.1038/nbt.2022

Varshney RK, Song C, Saxena RK, et al (2013) Draft genome sequence of chickpea (Cicer arietinum) provides a resource for trait improvement. Nat Biotechnol 2013 313 31:240–246. https://doi.org/10.1038/nbt.2491

Wang H, Ni D, Shen J, et al (2022) Genome-Wide Identification of the AP2/ERF Gene Family and Functional Analysis of GmAP2/ERF144 for Drought Tolerance in Soybean. Front Plant Sci 13:848766. https://doi.org/10.3389/FPLS.2022.848766/BIBTEX

Wang J, Sun P, Li Y, et al (2017) Hierarchically Aligning 10 Legume Genomes Establishes a Family-Level Genomics Platform. Plant Physiol 174:284–300. https://doi.org/10.1104/PP.16.01981

Wang M, Guo W, Li J, et al (2021) The miR528- AO Module Confers Enhanced Salt Tolerance in Rice by Modulating the Ascorbic Acid and Abscisic Acid Metabolism and ROS Scavenging. J Agric Food Chem 69:8634–8648.

https://doi.org/10.1021/acs.jafc.1c01096

Wang Y, Li K, Hen L, et al (2015) MicroRNA167-Directed Regulation of the Auxin Response Factors GmARF8a and GmARF8b Is Required for Soybean Nodulation and Lateral Root Development. Plant Physiol 168:984–999. https://doi.org/10.1104/PP.15.00265

Wang Y, You FM, Lazo GR, et al (2013) PIECE: A database for plant gene structure comparison and evolution. Nucleic Acids Res 41:1159–1166. https://doi.org/10.1093/nar/gks1109

Wei JT, Zhao SP, Zhang HY, et al (2023) GmDof41 regulated by the DREB1-type protein improves drought and salt tolerance by regulating the DREB2-type protein in soybean. Int J Biol Macromol 230:123255. https://doi.org/10.1016/J.IJBIOMAC.2023.123255

Wheeler GL, Jones MA, Smirnoff N (1998) The biosynthetic pathway of vitamin C in higher plants. Nat 1998 3936683 393:365–369. https://doi.org/10.1038/30728

Wu DG, Wang Y, Huang SC, et al (2021) Genome-wide identification and expression analysis of aao gene family in maize. Pakistan J Bot 53:181–190. https://doi.org/10.30848/PJB2021-1(11)

Xin S, Tao C, Li H (2016) Cloning and functional analysis of the promoter of an Ascorbate oxidase gene from Gossypium hirsutum. PLoS One 11:1–13. https://doi.org/10.1371/journal.pone.0161695

Xu Z, Wang R, Kong K, et al (2022) An APETALA2/ethylene responsive factor transcription factor GmCRF4a regulates plant height and auxin biosynthesis in soybean. Front Plant Sci 13:983650. https://doi.org/10.3389/FPLS.2022.983650/BIBTEX

Yamamoto A, Bhuiyan MNH, Waditee R, et al (2005) Suppressed expression of the apoplastic ascorbate oxidase gene increases salt tolerance in tobacco and Arabidopsis plants. J Exp Bot 56:1785–1796. https://doi.org/10.1093/jxb/eri167

Yang H, Shi G, Du H, et al (2017a) Genome-Wide Analysis of Soybean LATERAL ORGAN BOUNDARIES Domain-Containing Genes: A Functional Investigation of GmLBD12. Plant Genome 10:plantgenome2016.07.0058. https://doi.org/10.3835/PLANTGENOME2016.07.0058

Yang X, Kim MY, Ha J, Lee SH (2019) Overexpression of the Soybean NAC Gene GmNAC109 Increases Lateral Root Formation and Abiotic Stress Tolerance in Transgenic Arabidopsis Plants. Front Plant Sci 10:462757.

https://doi.org/10.3389/FPLS.2019.01036/BIBTEX

Yang Y, Chi Y, Wang Z, et al (2016) Functional analysis of structurally related soybean GmWRKY58 and GmWRKY76 in plant growth and development. J Exp Bot 67:4727–4742. https://doi.org/10.1093/JXB/ERW252

Yang Y, Zhou Y, Chi Y, et al (2017b) Characterization of Soybean WRKY Gene Family and Identification of Soybean WRKY Genes that Promote Resistance to Soybean Cyst Nematode. Sci Rep 7:1–13. https://doi.org/10.1038/s41598-017-18235-8

Yuan S, Li X, Li R, et al (2018) Genome-wide identification and classification of Soybean C2H2 Zinc Finger proteins and their expression analysis in legume-rhizobium symbiosis. Front Microbiol 9:287812. https://doi.org/10.3389/FMICB.2018.00126/BIBTEX

Yue L, Pei X, Kong F, et al (2023) Divergence of functions and expression patterns of soybean bZIP transcription factors. Front Plant Sci 14:1150363. https://doi.org/10.3389/FPLS.2023.1150363/BIBTEX

Zhang C, Grosic S, Whitham SA, Hill JH (2012) The requirement of multiple defense genes in soybean Rsv1-mediated extreme resistance to soybean mosaic virus. Mol Plant Microbe Interact 25:1307–1313. https://doi.org/10.1094/MPMI-02-12-0046-R

Zhang M, Liu Y, Shi H, et al (2018) Evolutionary and expression analyses of soybean basic Leucine zipper transcription factor family. BMC Genomics 19:1–14. https://doi.org/10.1186/S12864-018-4511-6/FIGURES/8

Zhang Z, Zhao Y, Chen Y, et al (2023) Overexpression of TCP9-like gene enhances salt tolerance in transgenic soybean. PLoS One 18:e0288985. https://doi.org/10.1371/JOURNAL.PONE.0288985

Zhao N, Ding X, Lian T, et al (2020) The Effects of Gene Duplication Modes on the Evolution of Regulatory Divergence in Wild and Cultivated Soybean. Front Genet 11:1–9. https://doi.org/10.3389/fgene.2020.601003

Zhu M, Liu Q, Liu F, et al (2023) Gene Profiling of the Ascorbate Oxidase Family Genes under Osmotic and Cold Stress Reveals the Role of AnAO5 in Cold Adaptation in Ammopiptanthus nanus. Plants 12:. https://doi.org/10.3390/plants12030677

Zhu M, Wang X, Zhou Y, et al (2022) Small RNA Sequencing Revealed that miR4415, a Legume-Specific miRNA, was Involved in the Cold Acclimation of

Ammopiptanthus nanus by Targeting an L-Ascorbate Oxidase Gene and Regulating the Redox State of Apoplast. Front Genet 13:1–20. https://doi.org/10.3389/fgene.2022.870446