THALLITON LUIZ CARVALHO DA SILVA

FENÔMICA E INTEGRAÇÃO DE TRANSCRITÔMICA E METABOLÔMICA NA ANÁLISE DAS RESPOSTAS DE *Gliricidia sepium* (JACQ.) STEUD. E *Portulaca oleracea* L. AO ESTRESSE SALINO

LAVRAS – MG
2021

THALLITON LUIZ CARVALHO DA SILVA

FENÔMICA E INTEGRAÇÃO DE TRANSCRITÔMICA E METABOLÔMICA NA ANÁLISE DAS RESPOSTAS DE G*liricidia sepium* (JACQ.) STEUD. E *Portulaca oleracea* L. AO ESTRESSE SALINO

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Biotecnologia Vegetal, área de concentração em Biotecnologia Vegetal, para a obtenção do título de Mestre.

Prof. Dr. Manoel Teixeira Souza Junior
Orientador

Dr. Leonardo Fonseca Valadares
Coorientador

LAVRAS – MG
2021

Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).

Silva, Thalliton Luiz Carvalho da.
Fenômica e integração de transcritômica e metabolômica na análise das respostas de *Gliricidia sepium* (Jacq.) Steud. E *Portulaca oleracea* L. ao estresse salino / Thalliton Luiz Carvalho da Silva. - 2021.
143 p. : il.

Orientador(a): Manoel Teixeira Souza Junior.
Coorientador(a): Leonardo Fonseca Valadares.

Dissertação (mestrado acadêmico) - Universidade Federal de Lavras, 2021.
Bibliografia.

1. Multi-ômica. 2. Salinidade. 3. Estresse Abiótico. I. Junior, Manoel Teixeira Souza. II. Valadares, Leonardo Fonseca.

THALLITON LUIZ CARVALHO DA SILVA

FENÔMICA E INTEGRAÇÃO DE TRANSCRITÔMICA E METABOLÔMICA NA ANÁLISE DAS RESPOSTAS DE *Gliricidia sepium* (JACQ.) STEUD. E P*ortulaca oleracea* L. AO ESTRESSE SALINO

PHENOMICS AND INTEGRATION OF TRANSCRIPTOMICS AND METABOLOMICS FOR ANALYSIS OF THE RESPONSES OF *Gliricidia sepium* (JACQ.) STEUD. AND P*ortulaca oleracea* L. TO SALINITY STRESS

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Biotecnologia Vegetal, área de concentração em Biotecnologia Vegetal, para a obtenção do título de Mestre.

APROVADA em 04 de agosto de 2021.

Dr. Manoel Teixeira Souza Júnior	EMBRAPA - Agroenergia
Dr. Leonardo Fonseca Valadares	EMBRAPA - Agroenergia
Dr. Carlos Antônio Ferreira de Sousa	EMBRAPA - Meio-Norte
Dra. Vivianny Nayse Belo Silva	EMBRAPA - Agroenergia

Prof. Dr. Manoel Teixeira Souza Junior
Orientador

Dr. Leonardo Fonseca Valadares
Coorientador

LAVRAS – MG
2021

Dedico este a todos que, direta ou indiretamente, participaram de minha vida e trouxeram consigo confiança, apoio e auxílio.

AGRADECIMENTOS

Agradeço primeiramente a Deus, pela vida, saúde e capacidade que tem me dado dia após dia para seguir meus caminhos e meu sonho.

Aos meus pais, Luiz e Adriana, que tanto tem me auxiliado, me apoiado e dado minha base de vida, minha educação e meus princípios.

A minha irmã e meu cunhado, Samille e Dailson, que em todos os momentos se dispõem prontamente para auxiliar no que for preciso.

Ao Manoel, por me orientar e por toda a paciência que teve comigo ao longo desses anos. Por me ensinar e me treinar em tudo o que fosse preciso.

Ao Leonardo, por todo o treinamento, paciência e por ter me dado oportunidades únicas e inesquecíveis (como trabalhar com a impressora 3D).

A toda equipe do grupo "Sal da Terra" pelos ensinamentos, risadas e ajuda em todos os momentos.

A todos os meus amigos, e todas as demais pessoas, que por descuido não lembrei no momento, mas que estão e estiveram presentes em minha vida e me ajudaram em algum momento.

A todos estes acima por toda a paciência no qual tiveram comigo, por todos os conselhos, palavras de carinho e, também, pelos "puxões de orelha" quando precisei.

A Universidade Federal de Lavras (UFLA) e a EMBRAPA Agroenergia pela oportunidade de realização deste mestrado.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001

A todos citados, meu mais sincero, muito obrigado!

“Acredite que você pode, e já terá percorrido metade do caminho!”
Theodore Roosevelt

RESUMO GERAL

A salinidade do solo é um dos estresses abióticos que mais ameaçam a agricultura. Este estresse está presente em mais de 100 países ao redor do mundo. Devido a estimativa de um aumento populacional mundial para cerca de 9 bilhões de pessoas em 2050 e, consequentemente, um aumento da demanda por produtos agrícolas, a pressão para a utilização dessas áreas tem aumentado. O objetivo geral do presente estudo foi aplicar estratégias de análise individual e integrada de dados ômicos provenientes de transcritômica e metabolômica visando ganhar conhecimento sobre os mecanismos moleculares responsáveis pela tolerância à salinidade observada em *Gliricidia sepium* e *Portulaca oleracea*. Para tal, foram utilizados dados do banco de dados "Sal da Terra", pertencentes ao programa de PD&I de mesmo nome desenvolvido na Embrapa Agroenergia, que contempla dados de fenômica, ionômica, genômica, transcritômica (mRNA e microRNA), metabolômica e proteômica caracterizando a resposta de dendê (*Elaeis guineensis*), beldroega (*Portulaca oleracea*) e gliricídia (*Gliricidia sepium*) ao estresse salino. As amostras do transcritoma foram submetidas a uma análise de RNA-Seq usando uma plataforma Illumina HiSeq e a estratégia "*paired-end*", a análise dos dados foi feita com o *software* OmicsBox versão 1.3. As amostras de metaboloma foram analisadas em um sistema UHPLC equipado com uma coluna de fase reversa. A espectrometria de massa de alta resolução (HRMS) foi realizada em um analisador Q-TOF usando fonte de eletrospray em ESI (+) - MS e ESI (-) - MS. Os dados adquiridos foram pré-processados usando o XCMS Online e posteriormente exportados para o MetaboAnalyst para análises estatísticas, anotação e observação das vias metabólicas. A plataforma Omics Fusion, foi utilizada para realizar a análise integrativa entre transcritos e metabólitos. Os resultados alcançados permitiram correlacionar e diferenciar grupos de plantas submetidas ao estresse salino, revelando genes / transcritos, metabólitos e vias responsivas a este estresse tanto em gliricídia, quanto em beldroega.

Palavras-chave: Multi-ômica. Salinidade. Estresse Abiótico

GENERAL ABSTRACT

Soil salinity is one of the abiotic stresses that most threaten agriculture. This stress is present in over 100 countries around the world. Due to an estimated global population increase to around 9 billion people in 2050, and the consequent increase in the demand for agricultural products, the pressure to use these areas has increased. The general objective of the present study was to apply single and integrated analysis strategies of omics data from transcriptomics and metabolomics to gain knowledge about the molecular mechanisms responsible for the salinity tolerance observed in *Gliricidia sepium* and *Portulaca oleracea*. To this end, data from the "Sal da Terra" database, belonging to the RD&I program of the same name developed at Embrapa Agroenergia, was used, which includes phenomic, ionomic, genomic, transcriptomic (mRNA and microRNA), metabolomic and proteomic data featuring the response of oil palm (*Elaeis guineensis*), purslane (*Portulaca oleracea*) and gliricidia (*Gliricidia sepium*) to salt stress. The transcriptome samples were submitted to an RNA-Seq analysis using an Illumina HiSeq platform using the paired-end strategy and the data analysis with the OmicsBox software version 1.3. Metabolome samples were analyzed on a UHPLC system equipped with a reversed-phase column. High-resolution mass spectrometry (HRMS) was performed on a Q-TOF analyzer using an electrospray source in ESI (+) - MS and ESI (-) - MS. The acquired data was pre-processed using XCMS Online and later exported to MetaboAnalyst for statistical analysis, annotation, and observation of metabolic pathways. The Omics Fusion platform was used to perform the integrative analysis between transcripts and metabolites. The results have allowed us to correlate and differentiate groups of plants subjected to salt stress, revealing genes/transcripts, metabolites, and responsive pathways to this stress, both in gliricidia and purslane.

General Keywords: Multi-omics. Salinity. Abiotic Stress

SUMÁRIO

CAPÍTULO 1

1 INTRODUÇÃO GERAL

Um dos problemas que mais afeta a atividade agrícola é a presença de sal nos solos, situação que aflige diversos países ao redor do mundo. Quando consideramos um contexto geral, cerca de 20% das terras agriculturáveis no mundo apresentam solos salinos e/ou sódicos. Olhando especificamente para as produções irrigadas, entre 25% e 30% dessas terras são afetadas pelo sal, não sendo produtivas em nível comercial (SHAHID et al., 2018).

Solos salinos, do ponto de vista agrícola, são descritos como aqueles que contêm sais solúveis neutros em quantidade suficiente para afetar negativamente o crescimento da maioria das plantas cultivadas. A priori, são considerados salinos aqueles solos que apresentam condutividade elétrica (CE) do extrato de saturação do solo >4 dS/m a 25 °C. Porém, devido ao fato de muitas espécies frutíferas, olerícolas e ornamentais sofrerem com os efeitos adversos da salinidade já em um intervalo de 2 dS/m a 4 dS/m, os solos com CE >2 dS/m a 25 °C passaram a também ser considerados salinos (BRESLER et al., 1982; VARGAS et al., 2018).

Existem dois grandes grupos de plantas, divididos com base em sua tolerância à salinidade: glicófitas e halófitas. Sendo que quase a totalidade (aproximadamente 99%) são glicófitas, plantas sensíveis ao sal, inclusive todas as principais culturas agrícolas. As halófitas são minoria (menos de 1%) e são capazes de completar seu ciclo de vida em ambiente onde a condutividade elétrica é maior ou igual a 20 dS/m (FLOWERS et al., 1986; FLOWERS; COLMER, 2008).

Apesar de ser frequentemente vista como um problema para o setor agrícola, suscitando ações voltadas à prevenção ou à remediação nas áreas afetadas, a salinidade pode ser vista como uma oportunidade. No âmbito da agricultura biossalina, a produção de alimentos, de fibras e de bioenergia é feita através de plantas tolerantes ao estresse salino, utilizando áreas e águas marginais para o seu cultivo (FAO, 2009; BORSAI et al., 2018).

No que diz respeito às espécies vegetais a serem utilizadas em um sistema de agricultura biossalina, existem duas possibilidades a serem exploradas: a) o uso de espécies glicófitas tolerantes à salinidade; e b) o uso de espécies halófitas. Essas duas possibilidades não são excludentes.

O objetivo do presente trabalho foi ganhar conhecimento sobre os mecanismos moleculares que conferem tolerância a salinidade em duas espécies previamente estudadas, beldroega (Portulaca oleracea L.) e gliricídia (Gliricidia sepium (Jacq.) Steud), utilizando estratégias de análise individuais e integradas (multi-ômica) de transcritômica e metabolômica. As análises foram feitas com base no banco de dados "Sal da Terra", um banco de dados desenvolvido pelo PD&I de mesmo nome, que reúne informações de diversas ômicas.

2 REVISÃO DE LITERATURA

2.1 Uso de estratégias de "*Multi-omics Integration*" (MOI) para caracterizar as respostas de plantas ao estresse salino

O projeto Genoma Humano (SCHMUTZ et al. 2004; NURK et al., 2021), concluído em 2003, pode ser considerado o marco que abriu as portas para o desenvolvimento da Biologia de Sistemas (IDEKER, 2004; VEENSTRA, 2021) e da Integração de Multi-ômicas (CAVILL et al., 2016, RAI et al., 2017). Foi a partir deste projeto que as ciências ômicas experimentaram um salto de magnitude na redução de custos e alavancagem operacional que contribuiu significativamente para sua popularização e consequente refinamento (GREEN et al., 2015).

Para bem conceituar o termo "-ômica", é necessário entender, primeiramente, o significado do sufixo "-oma", do qual este se deriva. O sufixo "-oma" pode ser definido como "conjunto de". Portanto, o termo genoma tem como significado o conjunto de genes. Tendo isso em mente, podemos compreender o termo "ômica" como "estudo do" (LEDERBERG; MCCRAY, 2001). Além da genômica, as principais ômicas são a transcritômica, proteômica e metabolômica, as quais podem ser definidas, respectivamente, como estudo do transcritoma, do conjunto de RNAs (mRNAs, miRNAs, lncRNAs, etc.) produzidos no organismo; estudo do proteoma, do conjunto de proteínas formadas no organismo; e estudo do metaboloma, do conjunto de metabólitos sintetizados no organismo (FIOCCHI, 2014).

Os investimentos e esforços em massa feitos de forma global no final do século XX para alcançar a elucidação do genoma humano permitiram que diversas ômicas emergissem e se aprimorassem cada vez mais, de forma que a partir desse momento histórico a tecnologia e a biologia começaram a andar lado a lado. Avanços tecnológicos permitiram novas descobertas biológicas e limitações biológicas instigaram o aprimoramento da tecnologia, permitindo que cada vez mais houvesse uma redução nos custos de aquisição dos diferentes dados ômicos e que estes fossem robustos e de alto rendimento (VEENSTRA, 2021). Além de que, junto ao elo "tecnologia-biologia" formando as ômicas, surgiu também o elo "tecnologia da informação-

ômicas", pois os grandes conjuntos de dados gerados não eram mais passiveis de serem analisados manualmente e, portanto, o processamento de computadores e o auxílio de softwares se tornaram parte fundamental dos estudos subsequentes, proporcionando o surgimento de um novo campo de estudo denominado bioinformática (BINNECK, 2004).

As abordagens ômicas individuais são utilizadas para avaliar as respostas biológicas a um amplo espectro de estímulos, incluindo a salinidade (KUMAR et al., 2019). Porém, nesse tipo de estudo é isolado apenas um nível, de toda a complexidade biológica existente, para verificar sua resposta. Dessa maneira, a genômica pode identificar diversos genes que não necessariamente estão sendo expressos. Ao passo que a transcritômica pode identificar múltiplos transcritos expressos, mas não nos dá a certeza de quais desses verdadeiramente se traduzem em proteínas, devido a diversos fatores de silenciamento, modificações pós-transcricionais e pós-traducionais.

Tendo em vista todas estas questões, desde a redução dos custos de aquisição dos dados junto ao alto rendimento dos mesmos até o advento da bioinformática para auxílio nas análises, a ciência biológica entrou na era da "Biologia de Big Data" (JAMIL et al., 2020). Isso levou a uma mudança de paradigma onde ocorreu uma transição da análise individual (*single*) para uma análise integrada correlacionando diferentes ômicas, como também a uma visão mais abrangente e robusta do sistema biológico (CAVILL et al., 2016).

Nessa nova era, um conceito nada novo amadurece e se expande, a "multi-ômica". Essa abordagem consiste na combinação de dois ou mais dados ômicos durante a análise, com a proposta de correlacionar os diversos dados e conseguir visualizar a resposta a um determinado estímulo de vários ângulos diferentes, tendo a bioinformática e trabalhos computacionais como principais coadjuvantes (CAVILL et al., 2016; JAMIL et al., 2020). Dessa maneira os cientistas conseguem encontrar novas associações entre os níveis biológicos.

Há diversas variações relacionadas ao termo "multi-ômica", tais como "poli-ômicas", "integração de ômicas", "trans-ômicas" e mais recentemente o surgimento do termo "Panômica" ou "Pan-ômica" para classificar todas as ômicas em uma mesma categoria (MISRA et al., 2019). Porém, tendo como base os artigos de revisão publicados nos últimos anos, o termo "multi-ômica" parece ser o mais correto e disseminado (CAVILL et al., 2016; JAMIL et al., 2020; MISRA et al., 2019; RAI et al., 2017; VEENSTRA, 2021).

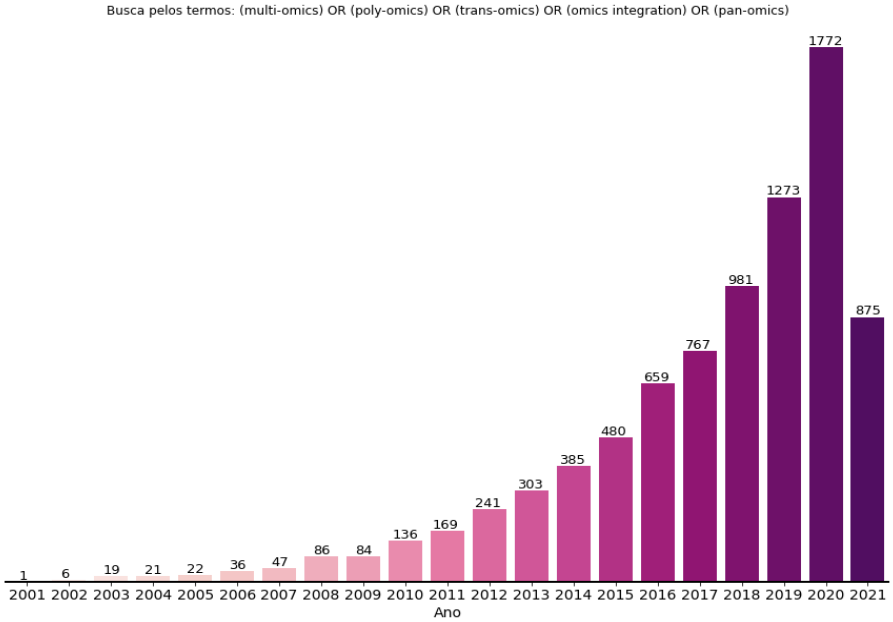
Quando realizamos uma busca pelo indexador de artigos PubMed, com os termos "multi-omics" e suas variações, são retornados cerca de 8.360 artigos, sendo os primeiros

publicados por volta de 2001 e o ano de 2020 sendo o que mais acumula artigos publicados com este tema, somando 1772 artigos (Figura 1).

Quando adicionamos o termo “plant” à busca, o montante total cai para aproximadamente 13% do seu valor, somando 1.071 artigos, com o primeiro sendo publicado em 2002. Da mesma maneira, o ano de 2020 conta com a maior quantidade de artigos publicados sobre esse tema em plantas, acumulando 217 artigos (Figura 2).

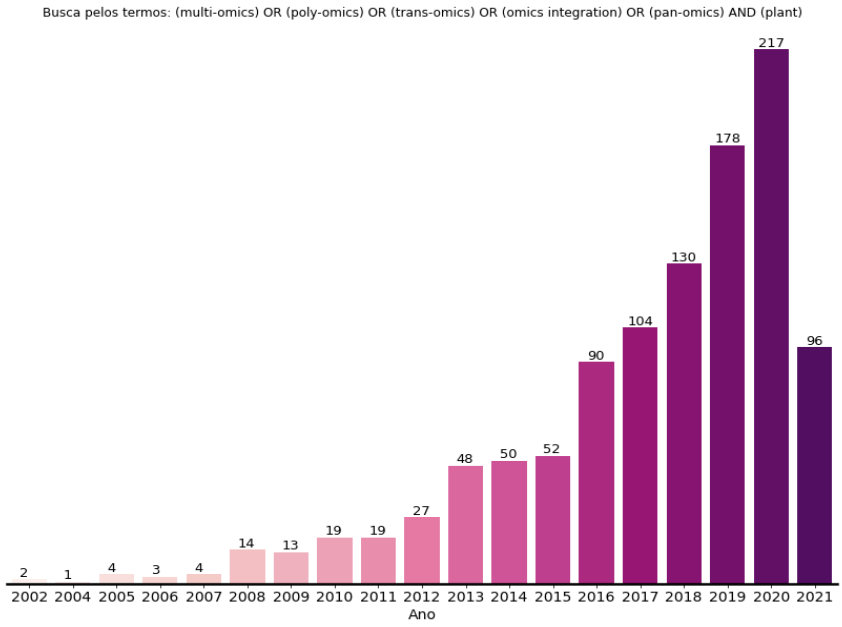
Isto mostra que a ideia de empregar a análise conjunta de diferentes ômicas para estudar um determinado fenômeno em plantas, ou em outros organismos, nasceu durante a execução do Projeto Genoma Humano, e não deixou de crescer desde então. No caso das plantas, uma das primeiras tentativas bem-sucedidas de integração de diferentes dados ômicos datam de 2003 (URBANCZYK-WOCHNIAK et al., 2003).

Figura 1 – Número de artigos científicos publicados no tema Multi-ômica até 2021. Busca pelo termo“multi-omics” e suas derivações no indexador de artigos PubMed.



Fonte: Do autor, 2021

Figura 2 – Número de artigos científicos publicados no tema Multi-ômica em plantas. Busca pelo termo “multi-omics” e suas derivações no indexador de artigos PubMed.



Fonte: Do autor, 2021

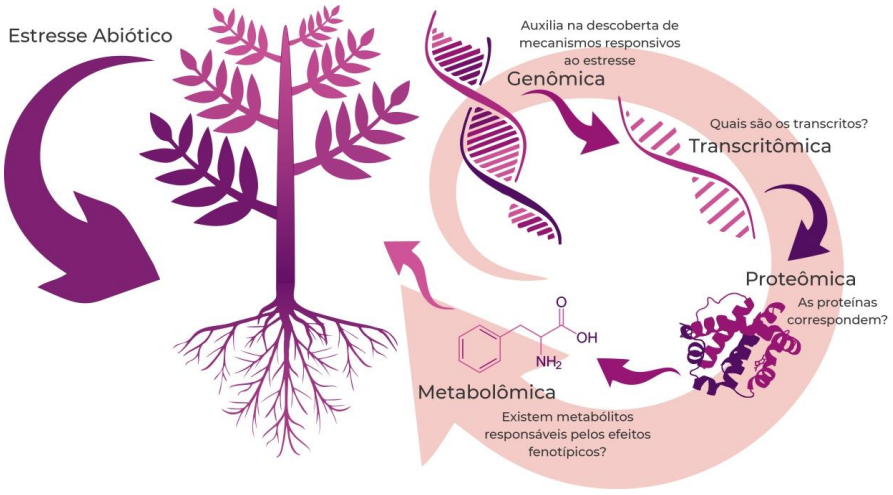
2.2 Fluxo da Informação Genética e Biologia de Sistemas

Na biologia molecular, o alicerce clássico que explica o fluxo da informação, desde o DNA até as proteínas, é o dogma central, proposto pela primeira vez por Francis Crick (CRICK, 1970). Este dogma descreve a transferência sequencial de informações das células desde a replicação do DNA, a transcrição em RNA e a tradução em cadeias de aminoácidos que posteriormente formarão proteínas (Figura 3); ao passo que afirma, também, que essa informação não pode fluir a partir da proteína para os outros níveis ômicos anteriores.

O aspecto geral dessas etapas descritas por Crick, não informando detalhes regulatórios complexos em etapas intermediárias entre os níveis ômicos, têm sido questionado e analisado por diversos autores (BUSTAMANTE et al., 2011; COSTA DOS SANTOS et al., 2021; PIRAS et al., 2012). Características regulatórias, como silenciamentos e modificações pós-transcricionais (*splicing* alternativo) e/ou pós-traducionais, eventos envolvendo miRNAs e modificações epigenéticas, possivelmente alteram o fluxo dessa informação (LUCO et al., 2011; KOONIN, 2012; PIRAS et al., 2012).

Mesmo com essas questões, o caráter simplista e macroscópico que o dogma central traz, em um nível amplo das diferentes ômicas, tende a continuar sendo um alicerce teórico extremamente influente dos sistemas vivos (PIRAS et al., 2012).

Figura 3 – Fluxo da informação genética sob o aspecto do estresse abiótico.



Fonte: Traduzido e adaptado de Raza et al. (2021).

Em diversos artigos que dissertam sobre multi-ômica e suas estratégias, podemos também observar a utilização do termo “Biologia de Sistemas” em conjunto com o termo “Multi-ômica” (CAVILL et al., 2016; FONDI; LIÒ, 2015; JAMIL et al., 2020; PINU et al., 2019; RAI et al., 2017; RAI et al., 2019; VEENSTRA et al., 2021). Embora ocorram divergências entre os próprios pesquisadores que contribuem ativamente para o avanço em pesquisas no âmbito da biologia de sistemas, devido principalmente à juvenilidade do campo e ao seu caráter interdisciplinar (BREITLING, 2010; VEENSTRA, 2021), o conceito mais simples e purista do termo pode ser atribuído a Dr. Trey Ideker (2004), que considera a biologia de sistemas um ramo no qual utiliza informações e dados adquiridos sistematicamente, a partir de diversas e diferentes ômicas, de forma a construir modelos preditivos para doenças e sistemas biológicos complexos.

Dessa maneira, a Biologia de Sistemas tem como objetivo a construção de modelos matemáticos bem projetados que prevejam, *in silico*, a mudança de um determinado organismo, no nível celular e molecular, quando este é perturbado ou está em um determinado meio (PINU et al., 2019). Já a multi-ômica tem um escopo mais extensivo, no qual o foco é compreender e correlacionar os diferentes níveis ômicos de forma a proporcionar à comunidade científica avanços no entendimento das regulações ômicas.

Conforme veremos adiante, podemos dizer que o ramo da Biologia de Sistemas está representado e incluso na integração de ômicas (MOI - “*Multi-omics Integration*”) nível 3, proposto por Jamil et al., (2020), enquanto a multi-ômica em si é um campo mais amplo e engloba não somente a modelagem do sistema biológico, mas os *insights* e descobertas promovidas pela análise e combinação de análises de diversas ômicas de forma integrada.

A multi-ômica compreende, portanto, uma análise global dos sistemas biológicos visando caracterizar grupos de moléculas em múltiplos níveis, sendo que as quatro grandes ômicas que ancoram estes diversos estudos são a genômica, transcritômica, proteômica e metabolômica (CAVILL et al., 2016; JAMIL et al., 2020).

A partir destas, diversos outros campos de estudo surgiram, carregando consigo seus próprios termos ômicos, alguns exemplos são: epigenômica, focada nos estudos das alterações epigenéticas resultantes da metilação do genoma (MALDONADO et al., 2021); peptidômica, caracterizada por estudar particularmente pequenos peptídeos como venenos e toxinas (AMADO et al., 2010); e a interatômica, que visa estudar as redes de interação proteína-proteína (SEATH et al., 2021).

2.2.1 Genômica

A genômica é a primeira grande ômica que ancora as principais ômicas estudadas, um campo que estuda a sequência completa de DNA, incluindo tanto os genes quanto as sequências intergênicas (BROWN, 2002). Embora o termo genoma remeta a "conjunto de genes", a definição mais correta seria "toda a informação que é herdável codificada no DNA de um organismo", dessa maneira é incluído tanto os genes quanto às regiões regulatórias e não-codificantes presentes na sequência de DNA (AIZAT et al., 2018).

As plantas precisam se adaptar e tolerar as distintas mudanças no ambiente, que geram estresses bióticos e abióticos, para garantir sua sobrevivência e perpetuação. Essa capacidade de se moldar a diferentes condições é denominada plasticidade fenotípica e está intimamente relacionada ao genoma do organismo, que através da ativação de genes específicos permite a regulação fisiológica e adaptação às diversas condições atípicas que sobrevêm (STOTZ et al., 2021). Dessa maneira, o genoma do organismo é quem dita a resposta aos diferentes tratamentos e estresses.

Portanto, a genômica visa não só verificar quais genes ou inferir quais proteínas estão presentes no organismo, mas verificar suas inter-relações e a influência no organismo, bem como descobrir e explorar a estrutura, função e a evolução dos diferentes genomas já sequenciados, além de realizar o sequenciamento de novas espécies (GUPPY et al., 2018; MISRA et al., 2019; SHENDURE et al., 2017).

2.2.2 Transcritômica

A segunda grande ômica é a transcritômica. O termo transcritoma pode ser entendido como o conjunto completo de todas as moléculas de RNA expressas em um organismo (WOLF, 2013). A transcritômica se caracteriza, então, pelo estudo tanto qualitativo, quanto quantitativo, dos diversos transcritos de um organismo (MILWARD et al., 2016; LIANG, 2013). Os mais conhecidos são os mRNAs, tRNAs e rRNAs. Porém, diversos outros transcritos já foram identificados e estão sendo cada vez mais estudados, alguns exemplos são os microRNAs e lncRNAs (*long non-coding* RNAs) (NAGANO; FRASER, 2011).

A transcritômica é fundamental devido ao seu papel como intermediário entre as informações contidas no DNA do organismo (genoma) e o proteoma, além das diversas funções regulatórias que os ncRNAs promovem (URANO et al., 2010). Da quantidade total de RNAs presentes em um organismo, cerca de 4% são traduzidos em proteínas, reafirmando a

importância desses RNAs não codantes na regulação dos processos fisiológicos do organismo (BROWN, 2002).

Outro importante aspecto da transcritômica é o *splicing* alternativo, onde um mesmo gene pode dar origem a mRNAs diferentes dependendo da forma com que seus éxons são processados (PUCKER; BROCKINGTON, 2018). Em uma ordem padrão, os genes contêm, em sua sequência, partes denominadas íntrons (que não são codificantes) e partes denominadas éxons (codificantes). Primeiramente, toda a sequência do gene é transcrita em um pré-mRNA e após isso ocorre o *splicing*, em que as regiões contendo íntrons são removidas e os éxons são unidos de forma sequencial. O *splicing* alternativo é o evento em que diferentes íntrons e éxons (ou parte destes) são alternativamente incluídos ou removidos durante o processamento do mRNA formando, dessa maneira, diferentes mRNAs a partir de um mesmo gene (PUCKER; BROCKINGTON, 2018; SIBLEY et al., 2016).

De acordo com Wang et al., (2009), podemos definir como principais objetivos da transcritômica: i) identificar e catalogar todos os tipos de transcritos; ii) determinar a estrutura da transcrição dos genes, bem como identificar seus padrões de *splicing* e outras modificações pós-transcricionais; iii) quantificar os níveis de expressão dos transcritos sob diferentes tratamentos e condições de crescimento, estádios de desenvolvimento e interferência de fatores bióticos e abióticos.

2.2.3 Proteômica

De forma a compreender, ao todo, um organismo, não basta somente saber quais são as sequências de nucleotídeos do seu genoma, nem quais são os transcritos expressos e seus níveis de expressão, em um determinado momento. É necessário, além de tudo isso, saber quais são os produtos dessa expressão.

A proteômica se caracteriza pelo estudo das diversas proteínas, incluindo sua identificação em larga escala, localização e compartimentalização, em um organismo (AEBERSOLD; MANN, 2003). As proteínas são moléculas orgânicas, de massa molecular elevada e estrutura complexa, formadas a partir de ligações covalentes entre os aminoácidos e têm diversas funções, como por exemplo, transporte de substâncias, catálise de reações, controle do metabolismo e componentes estruturais (ROBERTS, 2002). Diversas proteínas têm modificações pós-traducionais, como fosforilação, acetilação e glicosilação. Essas modificações regulam e realizam a manutenção da estrutura e função das proteínas (AEBERSOLD; MANN, 2016).

Tendo em vista esses aspectos, a proteômica nos permite visualizar o que ocorre no organismo provendo informações de eventos pós-transcricionais e pós-traducionais, além de que é o proteoma que especifica a natureza das reações bioquímicas que um organismo está capacitado a realizar. Dessa maneira, oferece a oportunidade de examinar as mudanças que ocorrem na produção e acúmulo das proteínas em processos complexos de desenvolvimento (BROWN, 2002).

2.2.4 Metabolômica

Visando compreender ao máximo as respostas de um organismo a uma determinada condição, precisamos chegar o mais próximo possível da avaliação do fenótipo daquele organismo. De todas as ômicas moleculares, a metabolômica é o elo mais próximo ao fenótipo do organismo (COSTA DOS SANTOS et al., 2021). Esta compreende o produto final da expressão de um gene e dos processos fisiológicos; e as mudanças em suas concentrações podem descrever melhor o estado bioquímico do organismo do que alterações visualizadas em níveis transcritômicos ou proteômicos (PALSSON, 2009).

Como último nível, das quatro grandes ômicas, temos a metabolômica, que consiste no estudo quantitativo e qualitativo de todos os metabólitos presentes em um organismo, em um determinado tempo e sob uma condição específica (FIEHN, 2001). Os metabólitos consistem em pequenas moléculas, com menos de 1.500 Da (DUNN et al., 2011).

Estes metabólitos podem ser classificados em dois tipos principais: os metabólitos primários e secundários (KABERA et al., 2014). Os metabólitos primários se caracterizam por moléculas envolvidas nos processos e funções básicas de uma célula para sua sobrevivência, sendo estes compartilhados por basicamente todos os organismos vivos. Estes metabólitos estão envolvidos nas principais vias metabólicas de uma célula, desempenhando funções como respiração celular e biossíntese de aminoácidos (KABERA et al., 2014).

Outro grupo bastante importante de metabólitos são os metabólitos secundários, estes são específicos para cada espécie (ou grupos próximos) e desempenham funções não vitais para a célula, mais ainda extremamente importantes para o organismo, como atrair polinizadores ou se defender contra pragas e doenças. De forma geral, nas plantas os metabólitos primários estão ligados ao seu crescimento e produção, enquanto os secundários estão ligados a características organolépticas (como sabor e cor) e de resistência a danos bióticos e abióticos (KABERA et al., 2014).

Os métodos de análise dos metabólitos (sejam eles primários ou secundários) se diferenciam em dois grupos: a metabolômica direcionada (do inglês, *targeted metabolomics*) (DUDLEY et al., 2010) e a metabolômica não-direcionada (do inglês, *untargeted metabolomics*) (DE VOS et al., 2007). A metabolômica direcionada se concentra na seleção *a priori* dos metabólitos a serem estudados e posterior aquisição e análise desses dados, com o objetivo principal de quantificação dos metabólitos de interesse, podendo ser selecionados alguns metabólitos específicos ou uma via metabólica alvo (DUDLEY et al., 2010).

Já a metabolômica não-direcionada consiste na aquisição de dados globais do perfil metabolômico, isto é, na aquisição da maior quantidade de dados possível referente ao metaboloma daquele organismo, realizando posteriormente a análise desses dados visando a classificação de amostras e a determinação de cada metabólito, não sendo necessário o prévio conhecimento dos compostos analisados (DE VOS et al., 2007). A escolha entre esses dois tipos de técnicas é determinada majoritariamente pelo foco do estudo, de forma que a metabolômica não-direcionada é utilizada normalmente para novas descobertas e geração de novas hipóteses e a metabolômica direcionada foca em testar estas hipóteses (DUNN; ELLIS, 2005).

2.3 Integração Multi-ômica (MOI)

Diversas revisões sobre o tema “Multi-ômica” foram escritas nos últimos anos, principalmente devido ao seu alto potencial de produzir novas ideias e observações sobre aspectos antes analisados somente sob uma perspectiva ômica (CAVILL et al., 2016; JAMIL et al., 2020; MISRA et al., 2019; RAI et al., 2017; VEENSTRA, 2021). Dentre elas, podemos destacar três revisões que fundamentam a pesquisa multi-ômica.

Cavill et al. (2016) trouxeram uma discussão sobre os diferentes aspectos da integração de dados entre metabolômica e transcritômica e os métodos de integração existentes. Mas o real impacto dessa revisão foi deixar bem elucidado a importância que o desenho experimental tem sob o aspecto de uma análise multi-ômica, descrevendo e exemplificando a diferença entre os desenhos experimentais e os vieses que cada desenho experimental pode causar durante as análises e no tratamento e processamento dos dados obtidos.

Um pouco mais à frente, Jamil et al. (2020) desenvolveram uma trilha de métodos para a integração de dados multi-ômicos, definindo essa integração em diferentes níveis, e guiando os pesquisadores, principalmente os novos nessa área, a como realizar suas análises, indicando ferramentas, softwares e fluxos de trabalho para uma integração bem-sucedida e precisa.

Com a importância que esse tema tem nos diversos campos da biologia e com a ampla adesão dos cientistas em embarcar nessa “nova” jornada, a revista *PROTEOMICS* fez uma edição especial, em fevereiro de 2021, com o tema “System Biology and Multi-omics”. Nessa edição, Veenstra (2021) trouxe uma revisão em que ele não só detalha a ascensão das ômicas e consequentemente da multi-ômica, mas discute, sobretudo, uma questão histórica na ciência: a pesquisa definida por hipóteses.

Veenstra (2021) discorreu sobre a pesquisa tradicional, relatando que esta é baseada em hipóteses e para estas hipóteses serem validadas ou rejeitadas, estudos são cuidadosamente desenhados e executados. Porém, com o advento da multi-ômica, as novas pesquisas que tenham como escopo a utilização dessas técnicas multi-ômicas e de biologia de sistemas, não necessariamente seguem a tradicional pesquisa baseada em hipóteses. Essa ordem se altera, e as hipóteses passam a ser geradas após as análises dos dados sob a ótica da multi-ômica. Isso muda a perspectiva científica de “pesquisas definidas por hipóteses” para uma “pesquisa dirigida por dados”. Reafirmando ainda mais a necessidade da realização cuidadosa e bem desenhada dos estudos, para que as novas hipóteses geradas possam ser robustas e bem definidas.

Como vimos anteriormente (Tópico 1.2), cada ômica individual tem seu próprio universo extremamente amplo de estudos. Porém, um nível ômico, por si só, não é capaz de responder e elucidar todas as questões referentes à resposta de um organismo a uma determinada perturbação. Cada ômica é uma peça primordial, fundamental e indispensável de um grande quebra-cabeça biológico, mas a visão global e sistemática desse quebra-cabeça só é possível quando juntamos essas peças (VEENSTRA et al., 2021).

Quando analisamos a proteômica e a metabolômica em conjunto, podemos ter uma visão ampla das reações presentes sob uma determinada condição. Podemos inferir, por meio da proteômica, quais são as vias metabólicas que estão sendo expressas naquele determinado momento e comparar essa inferência com as concentrações observadas dos metabólitos presentes, dando uma visão da regulação fisiológica do organismo e permitindo novas descobertas de supressão das atividades proteicas (CRAMER et al., 2011).

O mesmo vale para as análises conjuntas de transcritômica e proteômica, permitindo entender quais são as modificações e regulações pós-transcricionais (silenciamento, degradação, entre outros) que não seriam visíveis apenas no nível de transcritômica, bem como elucidar qual o resultado de uma superexpressão gênica na ampla gama de proteínas de um organismo (DALDOUL et al., 2014). Se associarmos a metabolômica junto às análises, conseguimos visualizar a resposta do organismo desde o nível de expressão gênica até o fenótipo, assimilando o quão complexo é o sistema biológico e verificando se uma observação ao nível transcritômico é realmente corroborada pelo nível metabolômico.

Integrando a genômica nesse complexo sistema, temos uma compreensão que parte desde o nível das sequências de nucleotídeos (mutações, modificações epigenéticas), quais as consequências dessas modificações a nível dos transcritos, as regulações sofridas e os produtos proteicos gerados e, por fim, o desfecho dessa complexidade nas moléculas constituídas, as vias metabólicas alteradas e a resposta final do organismo, no nível fenotípico, a um determinado estresse ou condição (DALDOUL et al., 2014).

Dessa maneira, temos claramente a compreensão de que uma estratégia multi-ômica permite avanços na elucidação da complexidade biológica e dos sistemas biológicos inter-relacionados, permitindo novas descobertas e a utilização desse conhecimento no melhoramento de culturas frente a diferentes pragas e estresses abióticos (DAS et al., 2015).

Esforços para propor maneiras de agrupar e dividir as análises de integração de dados ômicos de maneira a ficar mais compreensível para os novos pesquisadores que estão entrando nesse “mundo” foram feitas e alguns exemplos são expostos a seguir.

Ebbels e Cavill (2009) sugeriram três níveis de integração de dados: integração conceitual, integração estatística e integração baseada em modelo. A integração conceitual remete a análise separada de cada conjunto de dados ômicos e, posteriormente, os resultados e conclusões provenientes dessa análise são comparadas e sintetizadas pelo próprio autor. A integração estatística, como o próprio nome sugere, se caracteriza por encontrar associações estatísticas entre os dados. A integração baseada em modelo propõe uma descrição matemática do sistema, que pode modelar e prever cada nível de organização biológica separadamente, por exemplo, uma via metabólica parametrizada para um determinado organismo.

Wanichthanarak, Fahrmann e Grapov (2015) classificaram a integração entre diferentes ômicas em três grandes grupos: Integração baseada em vias metabólicas ou ontologia bioquímica, integração baseada em redes e integração baseada em correlação. O primeiro grupo se baseia em classificar os diferentes dados ômicos nas vias metabólicas já conhecidas. O segundo grupo visa construir uma rede de interação entre os diferentes tipos de dados ômicos, para possivelmente observar interações entre dados ômicos que não estão presentes em uma mesma via metabólica. O terceiro grupo tem como objetivo correlacionar os diferentes dados ômicos estatisticamente, principalmente em dados que possuem uma lacuna de conhecimento bioquímico prévio.

Bersanelli et al. (2016), classificaram e organizaram os métodos de integração de ômicas em quatro grandes classes: não bayesiano livre de rede (NF-NBY), bayesiano livre de rede (NF-BY), não bayesiano baseado em rede (NB-NBY) e bayesiano baseado em rede (NB-BY). Os autores explicaram e descreveram os fundamentos matemáticos das análises feitas sob o aspecto da multi-ômica, sendo importante para o desenvolvimento de novas ferramentas.

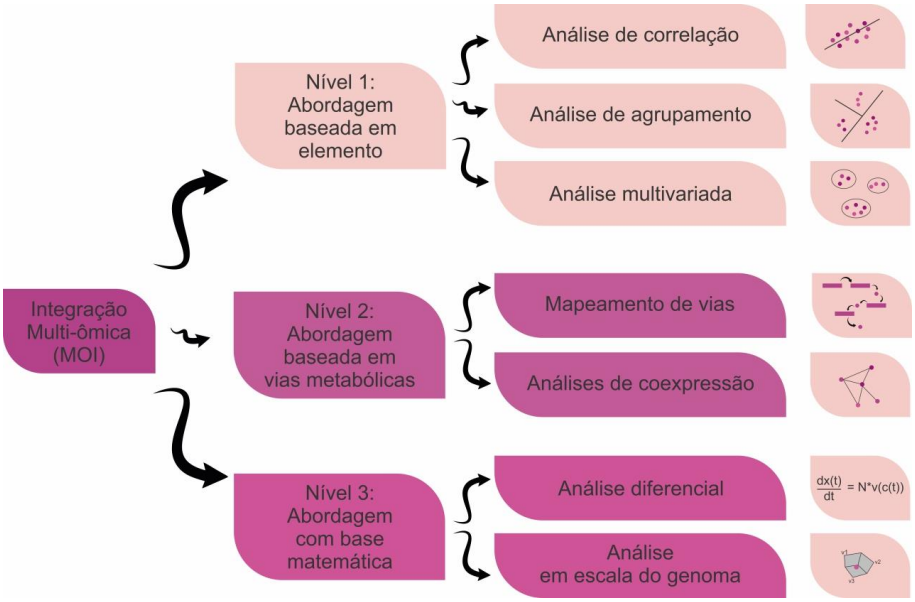
Cavill et al. (2016) além de explicar a importância do desenho experimental em uma análise multi-ômica, como já dito anteriormente, também descreveram diferentes formas de analisar os dados a partir de uma perspectiva multi-ômica. Eles separaram os dados em três grandes grupos, seguindo a linha de raciocínio do primeiro artigo publicado (EBBELS; CAVILL, 2009): integração conceitual, integração estatística e integração baseada em modelo. A integração conceitual e a integração baseada em modelo foram descritas anteriormente e mantêm o significado, mas a integração estatística foi subdividida em quatro grupos: integração

baseada em correlação, integração baseada em concatenação de dados, integração baseada em análises multivariadas e, por fim, integração baseada em vias metabólicas.

A integração baseada em correlação visa encontrar correlações entre dois grupos de dados ômicos distintos. Os métodos baseados na concatenação dos dados têm como objetivo agrupar as medições provenientes de diferentes ômicas em uma única tabela e, posteriormente, realizar uma análise integrada. A integração baseada em análises multivariadas utiliza técnicas padrão, como mínimos quadrados parciais (PLS) e análise de componentes principais (PCA) para encontrar relações entre variáveis e/ou amostras. Por fim, o último grupo consiste na utilização de conhecimento biológico para mapear os dados ômicos, com uma mudança estatística observada, em vias metabólicas conhecidas e presentes em banco de dados como KEGG e Wikipathways.

Uma das mais recentes publicações que visa classificar e direcionar as análises multi-ômicas foi redigida por Jamil et al. (2020), que teve como objetivo proporcionar diretrizes construtivas e metodológicas para uma realização bem-sucedida das análises multi-ômicas. Dessa maneira, os autores propuseram que um esquema metodológico bem definido, que permita a extração, combinação e associação crítica entre os diferentes dados ômicos, é necessário. De forma a garantir tudo que foi proposto, o fluxo de trabalho para estratégias de integração multi-ômica (MOI) foi redefinido em três níveis (Figura 4), com base na classificação anterior feita por Cavill et al. (2016), visando tornar a integração multi-ômica acessível a todos os pesquisadores, independente se estes são novos e não-treinados ou experientes.

Figura 4 – Níveis do fluxo de trabalho da integração multi-ômica (MOI).



Fonte: Traduzido e adaptado de Jamil et al. (2020)

2.3.1 Estratégia legado: A integração conceitual

A integração conceitual, como já descrita anteriormente, visa a análise de diferentes conjuntos de dados ômicos separadamente e, ao final das análises, os resultados são correlacionados pelo próprio autor de forma descritiva.

Cavill et al. (2016) chamaram a atenção para o fato de que essa abordagem pode produzir conhecimentos importantes e valiosos, mas também é uma abordagem que pode, muitas vezes, perder associações entre os dados ômicos que só poderiam ser observadas quando esses dados fossem analisados em conjunto, sob uma perspectiva estatística.

Jamil et al. (2020) concluíram, então, que esta análise quando não é feita de forma adequada se torna uma análise arbitrária. Dessa maneira, para a classificação proposta por estes autores, a integração conceitual não é inserida como uma abordagem MOI.

Seguindo as ideias propostas por Cavill et al. (2016), a integração estatística foi então reclassificada, a abordagem de integração por vias metabólicas foi separada em um novo grupo, para distinguir a integração imparcial da integração baseada em conhecimento prévio. A integração baseada em modelo também foi reformulada para separar a reconstrução de vias

metabólicas das abordagens puramente matemáticas. Os novos níveis de integração são descritos a seguir.

2.3.2 Integração Multi-ômica nível 1 – Baseada em elemento

Os níveis de MOI propostos por Jamil et al., (2020) tem como um dos objetivos serem complementares e com dificuldade e complexidade crescentes. Dessa maneira, o nível 1 engloba análises puramente estatísticas e imparciais, tendo como objetivo ser uma abordagem fácil e intuitiva.

Esse nível é dividido em três subclasses: correlação, agrupamento e análises multivariadas. A correlação é uma análise estatística que utiliza de coeficientes de correlação (Pearson, Spearman ou Kendall) para verificar o grau de correlação entre dois ou mais conjuntos de dados ômicos, sejam estas correlações diretas ou inversas.

O agrupamento consiste em deduzir associações e padrões entre os diferentes dados ômicos com base em atributos semelhantes, como seus níveis de expressão. O agrupamento é feito principalmente por meio de técnicas de aprendizado de máquina, como o agrupamento k-means e a análise por floresta aleatória, que permitem uma diferenciação por padrões de expressão e uma classificação para uma determinada característica, respectivamente.

A análise multivariada permite que o pesquisador consiga observar diferentes tendências nos conjuntos de dados ômicos, bem como investigar as relações entre esses dados. As técnicas mais comuns são o PCA, PLS e OPLS-DA (do inglês, *Orthogonal Partial Least Squares Discriminant Analysis*), bem como as variações dessas técnicas, como OnPLS (do inglês, *Orthogonal Projections to Latent Structures in Multiblock*). A análise multivariada é um pouco mais complexa e requer um estudo mais profundo para sua aprendizagem.

2.3.3 Integração Multi-ômica nível 2 – Baseada em vias metabólicas

A MOI nível 2 se baseia no conhecimento biológico prévio já estabelecido. Para pesquisadores com uma base biológica, tende a ser o modo de integração mais intuitivo. Esse nível é dividido em duas subclasses: mapeamento de via e análise de coexpressão.

O mapeamento de via consiste basicamente em mapear os diferentes conjuntos de dados ômicos, em banco de dados de vias metabólicas já existentes. O banco de dados mais comum e disseminado para este fim é o KEGG (Enciclopédia de Genes e Genomas de Kyoto) que engloba diversos organismos em todos os reinos. Porém, diversos bancos de dados com foco em organismos específicos já existem. Alguns exemplos são: Solcyc, com foco em espécies de

Solanaceae; AraCyc, com foco em Arabidopsis e CitrusCyc focado em diversas espécies de Citrus.

As análises de coexpressão tem como foco utilizar o resultado da correlação do MOI nível 1 para produzir redes de interação e avaliar a força das relações entre diferentes moléculas expressas. Essa análise permite revelar agrupamentos e módulos de interação importantes que contribuem para o avanço do conhecimento biológico.

2.3.4 Integração Multi-ômica nível 3 – Com base matemática

O último nível MOI consiste na aplicação matemática para produzir, com base nos dados ômicos, uma equação diferencial e um modelo bem definido de um determinado sistema biológico ou organismo. É a integração mais complexa e requer uma ampla cobertura de diferentes ômicas, bem como um organismo alvo bem caracterizado. Este nível é, também, dividido em duas subclasses: análise diferencial e análise em escala do genoma.

A análise diferencial consiste na aquisição de dados ômicos em diferentes tempos, para prever, por meio de uma equação estequiométrica, algum fator do organismo, como a taxa de tradução de um determinado mRNA ou o fluxo metabólico em uma determinada via metabólica já conhecida e bem caracterizada.

A análise em escala do genoma difere no fato de que o modelo matemático é construído primeiramente com base no genoma do organismo, considerando toda e qualquer reação que seja possível, para posteriormente, validar de maneira experimental os dados. É um processo complexo, principalmente para plantas e outros organismos eucariotos devido à alta compartimentalização e diversas vias metabólicas secundárias, bem como a poliploidia e o tamanho extenso de seus genomas.

Esse nível de integração permite que perturbações possam ser prevista *in silico*, porém o nível de conhecimento prévio exigido, tanto na questão biológica quanto no nível de programação e matemática, torna esta estratégia quase que uma utopia. É possível atualmente modelar amostras homogêneas e com um estado metabólico estacionário por um longo período de tempo.

2.4 Importância da salinidade e seus efeitos nas plantas

A salinidade do solo é um problema presente em mais de 100 países, espalhados em todos os continentes. Trata-se de um dos estresses abióticos que impõe as maiores limitações ao setor agrícola. Aproximadamente 20% das terras agriculturáveis no mundo apresentam solos salinos e/ou sódicos, entre 25% e 30% das terras irrigadas são afetadas pelo sal, sendo essencialmente improdutivas comercialmente (SHAHID et al., 2018).

Normalmente, a salinidade é vista como um problema para o setor agrícola, sendo constantemente realizadas ações voltadas para a prevenção ou à remediação nas áreas afetadas. Mas, sob a ótica da agricultura biossalina, os solos salinos são vistos como uma oportunidade para a produção de alimentos, de fibras, de bioenergia, como também para a recuperação de áreas degradadas e uso de áreas marginais, utilizando espécies tolerantes a essa condição (FAO, 2009; BORSAI et al., 2018).

Em geral, as espécies vegetais terrestres são divididas em dois grupos, de acordo com sua resposta ao estresse salino: glicófitas e halófitas. Aproximadamente 99% das plantas são glicófitas, plantas que são sensíveis ao sal e não conseguem completar seu ciclo de vida em um ambiente salino, estando neste grupo todas as principais culturas agrícolas. As halófitas correspondem a cerca de 1% das espécies vegetais terrestres. São plantas capazes de completar seu ciclo de vida em ambientes onde a concentração salina supera os 200 mM de NaCl – aproximadamente 20 dS/m (FLOWERS; COLMER, 2008; SCHOSSLER et al., 2012).

A salinidade causa estresses nas plantas de três maneiras principais: estresse osmótico; estresse iônico e estresse oxidativo. O estresse osmótico se caracteriza por um atraso no crescimento da planta, principalmente por efeito de estresse hídrico. O iônico se caracteriza por um processo dependente de íons, de forma que o acúmulo excessivo de íons na célula atinge níveis tóxicos, levando à atenuação dos processos metabólicos e, em alguns casos, à morte celular. Por fim, o estresse oxidativo se caracteriza pela formação das espécies reativas de oxigênio (*ROS* – do inglês*, Reactive Oxygen Species*), que em concentrações elevadas causam danos a todas as macromoléculas biológicas da célula (IBRAHIMOVA et al., 2021). Dessa maneira, o estresse salino afeta todos os principais processos vegetais, como a germinação e crescimento, fotossíntese, absorção de água, desequilíbrio de nutrientes e, portanto, o rendimento (PARIHAR et al., 2015).

Para lidar com as condições adversas, as plantas halófitas possuem mecanismos de adaptação aos íons e sais. Três mecanismos principais são conhecidos: absorção de íons de alta

concentração e seu acúmulo em vacúolos; liberação de sais absorvidos por células especiais nas folhas e restrição da absorção de sal por células da raiz (IBRAHIMOVA et al., 2021).

2.4.1 O uso de MOI visando entender as respostas das plantas ao estresse salino

Diversos estudos utilizando abordagens multi-ômicas de diferentes níveis e em diferentes organismos vêm sendo realizados nos últimos anos (Figura 1 e 2). Tais estudos visam melhor caracterizar a resposta dos organismos a uma determinada condição e, dessa forma, auxiliar no avanço do conhecimento científico (CAVILL et al., 2016). No que concerne à salinidade, as pesquisas em sua grande maioria visam descobrir novas informações que permitam auxiliar no aumento da tolerância ao estresse salino de espécies de interesse econômico, tendo em vista a necessidade de garantir a segurança alimentar em todo o mundo (DALDOUL et al., 2014; DAS et al., 2015; HO et al., 2020).

Esse auxílio no aumento da tolerância pode se dar por meio de técnicas de transgenia, inserindo genes já conhecidos sob o aspecto de tolerância à salinidade em plantas não tolerantes, visando caracterizar a resposta desse gene na planta de interesse, ou por meio de análises e comparações de cultivares tolerantes em espécies naturalmente não tolerantes, para descobrir novas regulações gênicas e novos *insights* para posterior utilização de técnicas de silenciamento ou outras alternativas para conferir tolerância (DAS et al., 2015).

Shen et al. (2016) estudaram por meio da multi-ômica dois acessos de cevada que diferiam na tolerância ao sal, o acesso XZ26 e XZ169. A integração utilizada foi a conceitual, comparando dados de metabolômica, proteômica e ionômica. Foi visto que o acesso XZ26 apresentou um maior crescimento e um menor acúmulo de sódio após 7 dias de tratamento salino quando comparado com o cultivar XZ169. Já o cultivar XZ169 apresentou uma redução significativa em concentrações de sacarose e metabólitos que estão envolvidos na via da glicólise, além de um elevado acúmulo de ácido cítrico, ácido aconítico e ácido succínico, resultando em um elevado nível do ciclo do ácido tricarboxílico (TCA). A análise proteômica corroborou os resultados obtidos pela metabolômica. O acesso XZ26 apresentou proteínas menos afetadas nos processos metabólicos e atividades catalíticas, além de uma fotossíntese mais estável, mostrando uma otimização dos processos que consomem energia.

Wanichthanarak et al. (2020) utilizaram uma abordagem integrativa entre ômicas, utilizando dados de transcritômica, metabolômica e fenômica, para observar as vias metabólicas perturbadas, os metabólitos alterados e os módulos mais importantes das redes metabólicas de arroz sob condições de estresse salino comparados com o controle. Foi verificado uma

reprogramação em vias metabólicas primárias, respiração celular, vias biossintéticas de antioxidantes e vias biossintéticas de fito-hormônios. Além dessa análise MOI nível 2, os autores, também, realizaram uma análise MOI nível 3, utilizando a abordagem em escala do genoma para modelar as respostas das vias metabólicas quando a planta está sob estresse salino. Os autores concluíram que a modelagem foi bem-sucedida, prevendo estados metabólicos que corroboram com os resultados da transcritômica e metabolômica, bem como das análises de fenômica, para algumas vias metabólicas.

Ho et al. (2020) utilizaram diversas abordagens MOI para estudar as respostas das raízes de dois cultivares de cevada (Clipper e Sahara) sob estresse salino. O estudo englobou dados transcritômicos, metabolômicos, lipidômicos e de microscopia, utilizando a estratégia MOI nível 1. Para correlação dos dados ômicos, as duas estratégias MOI nível 2 foram utilizadas, mapeamento de vias e análise de co-expressão, bem como a utilização dos dados de microscopia para corroborar com os resultados provenientes das ômicas. A via metabólica mais perturbada foi a via dos fenilpropanóides entre todas as respostas salinas observadas. Foi descrita uma intensa impregnação de lignina na parede celular da zona de alongamento Z2 do cultivar Clipper, em contraste com uma deposição de suberina na mesma zona Z2 do cultivar Sahara. Foi observado também que o fluxo simplástico que potencialmente ajusta a deposição de calose, no cultivar Clipper era praticamente constitutivo, independente do estresse por sal, enquanto esse fluxo diminuiu acentuadamente no cultivar Sahara quando exposta a salinidade.

Moreno et al. (2021) utilizaram uma abordagem multi-ômica, estudando a transcritômica, proteômica e metabolômica, para verificar quais alterações eram produzidas pela inserção do gene DcLCYB1 de cenoura (*Daucus carota*) em tabaco (*Nicotiana tabacum* cultivar Xanthi NN). Em contraste com o que se imaginava, a inserção de um gene que codifica uma enzima conversora do licopeno em beta-caroteno não somente alterou a quantidade de beta-caroteno produzido nas plantas transgênicas de tabaco, mas também, resultou em uma remodelagem nos níveis de transcritoma, proteoma e metaboloma na planta. Isso permitiu com que essa planta não somente fosse tolerante a estresses abióticos (como o sal), mas que o rendimento em termos de biomassa nessas condições adversas fosse maior do que o tipo selvagem, melhorando o crescimento e o desenvolvimento dessas plantas. A análise integrada dessas diferentes ômicas permitiu que os autores sugerissem novos processos e vias envolvidos nesse fenômeno de alta tolerância.

3 Uso da estratégia de moi para caracterização das respostas de *Gliricidia sepium* (Jacq.) Steud. E *Portulaca oleracea* L. ao estresse salino

Esta dissertação de Mestrado foi desenvolvida no âmbito do Programa de PD&I "Sal da Terra", desenvolvido na Embrapa Agroenergia. Este programa desenvolveu o Banco de Dados "Sal da Terra", que é constituído de dados de fenômica, ionômica, genômica, transcritômica (mRNA e microRNA), metabolômica e proteômica caracterizando a resposta de dendê (*Elaeis guineenses* Jacq.), beldroega (*Portulaca oleracea* L.) e gliricídia (*Gliricidia sepium* (Jacq.) Steud.) ao estresse salino.

Estudos visando a caracterização morfofisiológica da resposta destas espécies vegetais ao estresse salino, desenvolvidos no escopo deste programa, mostraram que tanto a beldroega quanto a gliricídia são altamente tolerantes a este estresse.

4 OBJETIVOS

O objetivo geral deste estudo é ganhar conhecimento sobre os mecanismos moleculares responsáveis pela tolerância a salinidade observada em *Gliricidia sepium* (Jacq.) Steud. e *Portulaca oleracea* L. através de estratégias de análise individual e integrativa de transcritômica e metabolômica.

5 ORGANIZAÇÃO DA DISSERTAÇÃO

A dissertação está organizada em quatro partes:

- Parte 1: Capítulo 1 - Revisão sobre o tema Multi-Omics Integration (MOI)
- Parte 2: Capítulo 2 – Artigo: "Integration of metabolomics and transcriptomics data to futher characterize *Gliricidia sepium* (Jacq.) Steud. under high salinity stress"
- Parte 3: Capítulo 3 – Artigo: "Multi-Omics analysis of young *Portulaca oleracea* L. plants' responses to high NaCl doses reveal insights on pathways and genes responsives to salinity stress in this halophyte species"
- Parte 4: Considerações finais

REFERÊNCIAS

AEBERSOLD R, MANN M. **Mass spectrometry-based proteomics**. Nature. Mar 13;422(6928):198-207, 2003. DOI: 10.1038/nature01511

AEBERSOLD, R.; MANN, M. **Mass-spectrometric exploration of proteome structure and function**. Nature, 537(7620), 347–355, 2016. DOI:10.1038/nature19949

AIZAT, W. M.; GOH, H.-H.; BAHARUM. **Omics Applications for Systems Biology**. Advances in Experimental Medicine and Biology. S. N. 2018. DOI:10.1007/978-3-319-98758-3

ALON, U. **An Introduction to Systems Biology:** Design Principles of Biological Circuits. Second Edition, Chapman and Hall/CRC, 2015.

AMADO, F., et al. **Salivary peptidomics**. Expert Review of Proteomics, 7(5), 709–721, 2010. DOI:10.1586/epr.10.48

BELHAJ, M. R., et al. **Metabolomics and Lipidomics:** Expanding the Molecular Landscape of Exercise Biology. Metabolites 11, no. 3: 151, 2021. DOI: https://doi.org/10.3390/metabo11030151

BERSANELLI, M.. et al. **Methods for the integration of multi-omics data:** mathematical aspects. BMC Bioinformatics 17, S15, 2016. DOI: https://doi.org/10.1186/s12859-015-0857-9

BINNEK E., **As ômicas:** integrando a bioinformação. Biotecnologia Ciência & Desenvolvimento. N 32 – janeiro/junho 2004. Disponível em: https://edisciplinas.usp.br/pluginfile.php/4119117/mod_resource/content/1/Estudos%20das%20%C3%B4micas.pdf

BORSAI, O.; et al. **The genus Portulaca as a suitable model to study the mechanisms of plant tolerance to drought and salinity**. The EuroBiotech Journal, v. 2, n. 2, p. 104-113, 2018. DOI: 10.2478/ebtj-2018-0014.

BREITLING, R. **What is systems biology?** Front. Physiol.,1, 9, 2010.

BRESLER, E. et al. **Saline and sodics soils: principles-dynamics- -modeling**. Berlin: Springer-Verlag, 1982. 236 p. (Advanced series in agricultural Sciences). DOI: 10.1007/978-3-642-68324-4

BROWN, T. A. **Genomes**. 2nd edition. Oxford: Wiley-Liss; 2002. Disponível em: https://www.ncbi.nlm.nih.gov/books/NBK21128/

BUSTAMANTE, C.; CHENG, W.; MEJIA, Y. X. **Revisiting the Central Dogma One Molecule at a Time**. Cell, v 144, issue 4, 480-497, 2011. DOI: https://doi.org/10.1016/j.cell.2011.01.033.

CASSAGO, A. L. L., et al. **Metabolomics as a marketing tool for geographical indication products**: a literature review. Eur Food Res Technol (2021). DOI: https://doi.org/10.1007/s00217-021-03782-2

CAVILL, R., et al. **Transcriptomic and metabolomic data integration**. Briefings in Bioinformatics, 17(5), 891–901, 2016. DOI:10.1093/bib/bbv090

COSTA DOS SANTOS, G. et al. **The remodel of the "central dogma":** a metabolomics interaction perspective. Metabolomics 17, 48, 2021. DOI: https://doi.org/10.1007/s11306-021-01800-8

CRAMER, G.R., et al. **Effects of abiotic stress on plants:** a systems biology perspective. BMC Plant Biol., 11, 163, 2011.

CRICK, F. **Central dogma of molecular biology**. Nature 227.5258, 561-563, 1970.

DALDOUL, S., et al. **Integration of omics and system biology approaches to study grapevine (Vitis vinifera L.) response to salt stress:** a perspective for functional genomics - A review. OENO One, 48(3), 189–200, 2014. DOI: https://doi.org/10.20870/oeno-one.2014.48.3.1573

DAS, P., et al. **Understanding salinity responses and adopting "omics-based" approaches to generate salinity tolerant cultivars of rice**. Frontiers in Plant Science, 6, 2015. DOI:10.3389/fpls.2015.00712

DE VOS, R. C., et al. **Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry.** Nature Protocols, 2(4), 778–791, 2007. DOI: 10.1038/nprot.2007.95

DUDLEY, E. et al. **Targeted metabolomics and mass spectrometry**. Advances in Protein Chemistry and Structural Biology, 45–83, 2010. DOI:10.1016/b978-0-12-381264-3.00002-3

DUNN, W. B., et al. **Systems level studies of mammalian metabolomes:** the roles of mass spectrometry and nuclear magnetic resonance spectroscopy. Chem. Soc. Rev., 40(1), 387–426, 2011. DOI:10.1039/b906712b

DUNN, W. B.; ELLIS, D. I. **Metabolomics:** Current analytical platforms and methodologies. TrAC Trends in Analytical Chemistry, 24(4), 285–294, 2005. DOI:10.1016/j.trac.2004.11.021

EBBELS, T. M. D.; CAVILL R . **Bioinformatic methods in NMR-based metabolic profiling**. Prog Nucl Magn Reson Spectrosc; 55 : 361 – 74, 2009. DOI:10.1016/j.pnmrs.2009.07.003

FAO. **Advances in the assessment and monitoring of salinization and status of biosaline agriculture:** report of an expert consultation held in Dubai, United Arab Emirates, 26–29. Rome, 2009. Disponível em: www.fao.org/3/i1220e/i1220.pdf.

FIEHN, O. C**ombining Genomics, Metabolome Analysis, and Biochemical Modelling to Understand Metabolic Networks**. Comparative and Functional Genomics, 2(3), 155–168, 2001. DOI:10.1002/cfg.82

FIOCCHI C: **Integrating Omics:** The Future of IBD? Dig Dis. 32(suppl 1):96-102, 2014. DOI: 10.1159/000367836

FLOWERS, T. et al. **Halophytes**. The Quarterly Review of Biology, v. 61, n. 3, p. 313-337, 1986

FLOWERS, T.; COLMER, T. **Salinity tolerance in halophytes**. New Phytologist, v. 179, n. 4, p. 945-963, 2008.

FONDI, M.; LIÒ, P. **Multi-omics and metabolic modelling pipelines**: challenges and tools for systems microbiology. Microbiol. Res. 171, 52–64, 2015. DOI: 10.1016/j.micres.2015.01.003

GREEN, E.; WATSON, J.; COLLINS, F. **Human Genome Project**: Twenty-five years of big biology. Nature 526, 29–31, 2015. DOI: https://doi.org/10.1038/526029a

GUPPY, J. L., et al. **The State of "Omics" Research for Farmed Penaeids**: Advances in Research and Impediments to Industry Utilization. Frontiers in Genetics. 2018. DOI: 10.3389/fgene.2018.00282

HO, W. W. H., et al. **Integrative Multi-omics Analyses of Barley Rootzones under Salinity Stress Reveal Two Distinctive Salt Tolerance Mechanisms.** Plant Communications, Volume 1, Issue 3, 2020. DOI: https://doi.org/10.1016/j.xplc.2020.100031.

IBRAHIMOVA U., et al. **Progress in understanding salt stress response in plants using biotechnological tools.** Journal of Biotechnology, 329 , pp. 180-191, 2021. DOI: https://doi.org/10.1016/j.jbiotec.2021.02.007.

IDEKER, T. **Systems biology 101—what you need to know.** Nature Biotechnology, 22(4), 473–475, 2004. DOI:10.1038/nbt0404-473

JAMIL, I. N., et al. **Systematic Multi-Omics Integration (MOI) Approach in Plant Systems Biology**. Frontiers in Plant Science, 11, 2020. DOI:10.3389/fpls.2020.00944

KABERA, J. N. et al. **Plant Secondary Metabolites:** Biosynthesis, Classification, Function and Pharmacological Properties. Journal of Pharmacy and Pharmacology, v. 2, p. 377-392. 2014.

KLIPP, E. et al. **Systems biology in practice:** concepts, implementation and application. Wiley VCH; 2nd edition, 2014.

KOONIN, E.V. **Does the central dogma still stand?**. Biol Direct 7, 27, 2012. DOI: https://doi.org/10.1186/1745-6150-7-27

KUMAR, A., et al. **Salinity-induced Physiological and Molecular Responses of Halophytes.** Research Developments in Saline Agriculture, 331–356, 2019. DOI:10.1007/978-981-13-5832-6_10

LEDERBERG J.; MCCRAY A. T. **'Ome Sweet 'Omics - A Genealogical Treasury of Words.** Genealogical Treasury of Words. Scientist.; 15(7):8, 2001.

LIANG, K.-H. **Transcriptomics**. Bioinformatics for Biomedical Science and Clinical Applications, 49–82, 2013. DOI:10.1533/9781908818232.49

LUCO, R. F. Et al. **Epigenética em splicing alternativo de pré-mRNA**. Cell 144, 16–26, 2011.

MALDONADO, R. et al. **Genomics and epigenomics of addiction**. Am J Med Genet Part B. 186B: 128– 139, 2021. DOI: https://doi.org/10.1002/ajmg.b.32843

Milward, E. A.,et al. **Transcriptomics**. Encyclopedia of Cell Biology, 160–165, 2016. DOI:10.1016/b978-0-12-394447-4.40029-5

MISRA, B. B., et al. **Integrated omics:** tools, advances and future approaches. Journal of Molecular Endocrinology 62, 1, R21-R45, 2019. DOI: https://doi.org/10.1530/JME-18-0055

MORENO, J. C. et al. **A Multi-OMICs Approach Sheds Light on the Higher Yield Phenotype and Enhanced Abiotic Stress Tolerance in Tobacco Lines Expressing the Carrot lycopene β-cyclase1 Gene**. Frontiers in plant science vol. 12 624365, 2021. DOI:10.3389/fpls.2021.624365

NAGANO, T.; FRASER, P. **No-Nonsense Functions for Long Noncoding RNAs**. Cell, 145(2), 178–181, 2011. DOI:10.1016/j.cell.2011.03.014

NURK, S. et al. The complete sequence of a human genome. 2021. DOI: https://doi.org/10.1101/2021.05.26.445798

OLIVER, S. (1998). **Systematic functional analysis of the yeast genome.** Trends in Biotechnology, 16(9), 373–378. DOI:10.1016/s0167-7799(98)01214-1

PALSSON, B. **Metabolic systems biology**. FEBS Letters, 583(24), 3900–3904, 2009. DOI:10.1016/j.febslet.2009.09.031

PARIHAR, P. et al. **Effect of salinity stress on plants and its tolerance strategies:** a review. Environ Sci Pollut Res., v. 22, n. 6, p. 4056-4075, 2015. DOI: 10.1007/s11356-014-3739-1

PINU, F. R. et al. **Systems biology and multi-omics integration:** Viewpoints from the metabolomics research community. Metabolites 9 (4), 76, 2019. DOI: 10.3390/metabo9040076

PIRAS, V. et al. **Is central dogma a global property of cellular information flow?** Frontiers in Physiology, 3, 2012. DOI:10.3389/fphys.2012.00439

PUCKER, B.; BROCKINGTON, S.F. **Genome-wide analyses supported by RNA-Seq reveal non-canonical splice sites in plant genomes**. BMC Genomics 19, 980, 2018. DOI: https://doi.org/10.1186/s12864-018-5360-z

RAI, A. et al. **A new era in plant functional genomics**. Curr. Opin. Syst. Biol. 15, 58–67, 2019. DOI: 10.1016/j.coisb.2019.03.005

RAI, A. et al. **Integrated omics analysis of specialized metabolism in medicinal plants**. Plant J. 90 (4), 764–787, 2017. DOI: 10.1111/tpj.13485

ROBERTS, J. K. M. **Plant Molecular Biology**, 48(1/2), 143–154. 2002 DOI:10.1023/a:1013736322130

SCHMUTZ, J., et al. **Quality assessment of the human genome sequence.** Nature 429, 365–368, 2004. DOI: https://doi.org/10.1038/nature02390

SCHOSSLER, T. R. et al. **Salinidade:** Efeitos na fisiologia e na nutrição mineral de plantas. Enciclopédia Biosfera, Centro Científico Conhecer, Goiânia, v. 8, n. 15, p. 1563-1578, 2012

SEATH, C. P. et al. Reactive intermediates for interactome mapping. The Royal Society of Chemistry., 5, 2911-2926, 2021. DOI: 10.1039/D0CS01366H

SHAHID, S. A. et al. **Soil salinity:** historical perspectives and a world overview of the problem. In: SHAHID, S. A.; ZAMAN, M.; HENG, L. Guideline for Salinity Assessment, Mitigation and Adaptation Using Nuclear and Related Techniques. Cham: Springer, 2018. 164 p. https://doi.org/10.1007/978-3-319-96190-3

SHEN, Q. et al. **Multi-omics analysis reveals molecular mechanisms of shoot adaption to salt stress in Tibetan wild barley**. BMC Genomics 17, 889, 2016. DOI: https://doi.org/10.1186/s12864-016-3242-9

SHENDURE, J., et al. **DNA sequencing at 40: past, present and future**. Nature, 550(7676), 345–353, 2017. DOI:10.1038/nature24286

SIBLEY, C. R., BLAZQUEZ, L., & ULE, J. **Lessons from non-canonical splicing**. Nature Reviews Genetics, 17(7), 407–421, 2016. DOI:10.1038/nrg.2016.46

STOTZ, G.C., et al. **Global trends in phenotypic plasticity of plants**. Ecology Letters, 00, 1– 15, 2021. DOI: https://doi.org/10.1111/ele.13827

URANO, K., et al. **"Omics" analyses of regulatory networks in plant abiotic stress responses.** Current Opinion in Plant Biology, 13(2), 132–138, 2010. DOI:10.1016/j.pbi.2009.12.006

URBANCZYK-WOCHNIAK, E. et al. **Parallel analysis of transcript and metabolic profiles:** a new approach in systems biology. EMBO reports, 4(10), 989–993, 2003. DOI: https://doi.org/10.1038/sj.embor.embor944

VARGAS, R. et al. **Handbook for saline soil management.** [Rome]: FAO, 2018. Disponível em: www.fao. org/3/i7318en/I7318EN.pdf

VEENSTRA, T.D. **Omics in Systems Biology:** Current Progress and Future Outlook. Proteomics, 21: 2000235, 2021. DOI: https://doi.org/10.1002/pmic.202000235

WANG, Z., et al. **RNA-Seq: a revolutionary tool for transcriptomics**. Nat Rev Genet 10, 57–63, 2009. DOI: https://doi.org/10.1038/nrg2484

WANICHTHANARAK K, FAHRMANN JF, GRAPOV D. **Genomic, Proteomic, and Metabolomic Data Integration Strategies**. Biomarker Insights. 2015. DOI:10.4137/BMI.S29511

WANICHTHANARAK, K., et al. **Deciphering rice metabolic flux reprograming under salinity stress via in silico metabolic modeling**. Computational and Structural Biotechnology Journal, 18, 3555–3566, 2020. DOI:10.1016/j.csbj.2020.11.023

WOLF, J.B.W. **Principles of transcriptome analysis and gene expression quantification:** an RNA-seq tutorial. Mol Ecol Resour, 13: 559-572, 2013. DOI: https://doi.org/10.1111/1755-0998.12109

CAPÍTULO 2

Integração de dados metabolômicos e transcritômicos para melhor caracterizar *Gliricidia sepium* (Jacq.) Walp. sob estresse de alta salinidade

A versão apresentada do presente artigo foi submetida a revista "*The Plant Genome*", sendo uma versão preliminar e o conselho editorial do periódico poderá sugerir alterações.

RESUMO

Um dos estresses abióticos que mais ameaçam a agricultura é a salinidade do solo, um problema presente em mais de 100 países espalhados por todos os continentes. Devido ao aumento da demanda por produtos agrícolas, a pressão para o cultivo nesses solos tem aumentado. *Gliricidia sepium* (Jacq.) Walp. é uma árvore polivalente, cultivada para melhorar a fertilidade do solo, para fins medicinais, como madeira / lenha, como carvão e como sombra de plantações. Também é conhecido por sua capacidade de se adaptar a uma ampla variedade de solos, desde solos ácidos erodidos, solos arenosos, argila pesada, calcário e solos alcalinos. Os limites de tolerância à salinidade da gliricídia, bem como suas respostas ao estresse salino, ainda não são bem compreendidos. Os perfis de transcritoma e metaboloma da parte aérea de *G. sepium* foram realizados em plantas controle e com estresse salino em um delineamento inteiramente casualizado. As amostras do transcritoma foram submetidas ao RNA-Seq usando uma plataforma Illumina HiSeq e a estratégia "*paired end*", e a análise dos dados foi feita com o OmicsBox versão 1.3. As amostras de metaboloma foram analisadas em um sistema UHPLC equipado com uma coluna de fase reversa. A espectrometria de massa de alta resolução (HRMS) foi realizada em um analisador Q-TOF usando fonte de eletrospray em ESI (+) - MS e ESI (-) - MS. Os dados adquiridos foram pré-processados usando o XCMS Online e posteriormente exportados para o MetaboAnalyst 4.0 para análise multivariada, anotação metabólica e observação da atividade da via. Omics Fusion, uma plataforma web para análise integrativa de dados ômicos, foi empregada para realizar a análise integrativa de transcritos e metabólitos. A análise dos conjuntos de dados do transcritoma e do metaboloma caracterizou a resposta da planta em três cenários: Efeito idade (plantas controle aos 2 e 45 dias sob estresse - DAT), estresse de curto prazo (plantas controle e estressadas aos 2 DAT) e estresse de longo prazo (plantas estressadas aos 2 e 45 DAT). Um grupo de 5.672 transcritos e 107 metabólitos foram submetidos à análise integrativa para integrar transcritos e metabólitos diferencialmente expressos na parte aérea de gliricídia sob estresse salino; a biossíntese do fenilpropanóide apareceu em primeiro lugar entre as vias mais afetadas, com 15 metabólitos e cinco transcritos

(três genes) diferencialmente expressos. A análise única e integrada dos perfis de transcritoma e metaboloma gerados neste estudo foi eficiente para correlacionar e diferenciar grupos de plantas de *G. sepium* submetidas ao estresse salino, revelando genes / transcritos, metabólitos e vias responsivas a este estresse. A análise dos metabólitos e genes diferencialmente expressos na via de biossíntese dos fenilpropanóides revelou que ele desempenha um papel no estresse de curto prazo. A análise do transcritoma identificou dois genes que codificam proteínas que podem desempenhar um papel na resposta da gliricídia tanto no estresse salino de curto quanto no de longo prazo.

Palavras-chave: RNA-Seq, Quimiometria, Espectometria de Massa de Alta Resolução, Estresse Abiótico, Integratômica, Integração Multi-ômica.

ABSTRACT

One of the abiotic stresses that threaten agriculture the most is soil salinity, a problem present in more than 100 countries spread across all continents. Due to the increase in demand for agricultural products, the pressure for cultivation in these soils has increased. *Gliricidia sepium* (Jacq.) Walp. is a multipurpose tree, cultivated for improvement of soil fertility, for medicinal purposes, as wood/firewood, as charcoal, and as a shade of plantations. It is also known for its ability to adapt to a wide range of soils ranging from eroded acidic soils, sandy soils, heavy clay, limestone, and alkaline soils. Gliricidia salinity tolerance limits, alongside its responses to salt stress, are not yet well understood. The transcriptome and metabolome profiles of *G. sepium* shoots were performed on control and salt-stressed plants in a completely randomized design. Transcriptome samples were subjected to RNA-Seq using an Illumina HiSeq platform and the paired-end strategy, and data analysis was done with OmicsBox version 1.3. Metabolome samples were analyzed on a UHPLC system equipped with a reversed-phase column. High-resolution mass spectrometry (HRMS) was performed on a Q-TOF analyzer using electrospray source in ESI(+)-MS and ESI(-)-MS. Acquired data were pre-processed using XCMS Online and further exported to MetaboAnalyst 4.0 for multivariate analysis, metabolic annotation, and pathway activity observation. Omics Fusion, the web platform for integrative analysis of Omics data, was employed for carrying out the integrative analysis of transcripts and metabolites. The analysis on transcriptome and metabolome data sets characterized the plant response under three scenarios: Age effect (control plants at 2 and 45 days under stress - DAT), short-term (control and stressed plants at 2 DAT), and long-term stress (stressed plants at 2 and 45 DAT). A group of 5,672 transcripts and 107 metabolites were submitted to integrative analysis to integrate transcripts and metabolites differentially expressed in gliricidia shoots under salt stress; the phenylpropanoid biosynthesis came in first among the most affected pathways with 15 metabolites as well as five transcripts (three genes) differentially expressed. The single and integrated analysis of the transcriptome and the metabolome profiles generated in this study were efficient to correlate and differentiate groups of *G. sepium* plants submitted to salinity stress, revealing genes/transcripts, metabolites, and pathways responsive to this stress. The analysis of the metabolites and genes differentially expressed in the phenylpropanoid biosynthesis pathway revealed that it plays a role in short-term stress. The single analysis of the transcriptome identified two genes coding for proteins that might play a role in gliricidia response at both the short- and long-term salt stress.

Keywords: RNA-Seq, Chemometrics, High Resolution Mass Spectrometry, Abiotic Stress, Integratomics, Multi-Omics Integration.

Integration of metabolomics and transcriptomics data to futher characterize *Gliricidia sepium* (Jacq.) Walp. under high salinity stress

Thalliton Luiz Carvalho da Silva[1§]
Vivianny Nayse Belo Silva[1§]
Ítalo de Oliveira Braga[1]
Jorge Candido Rodrigues Neto[2]
André Pereira Leão[4]
José Antônio de Aquino Ribeiro[4]
Leonardo Fonseca Valadares[4]
Patrícia Verardi Abdelnur[2,4]
Carlos Antônio Ferreira de Sousa[3]
Manoel Teixeira Souza Júnior[1,4,*]

[1] – Graduate Program of Plant Biotechnology, Federal University of Lavras, CP 3037, Lavras, MG, Zip Code 37200-000, Brazil
[2] – Institute of Chemistry, Federal University of Goiás, Campus Samambaia, Goiânia, GO, Zip Code 74690-900, Brazil
[3] – Brazilian Agricultural Research Corporation, Embrapa Mid-North, Teresina, PI, Zip Code 64008-780, Brazil
[4] – Brazilian Agricultural Research Corporation, Embrapa Agroenergy, Brasília, DF, Zip Code 70770-901, Brazil
§ - These authors contributed equally to this study
* - Corresponding author

Keywords: RNA-Seq, Chemometrics, High Resolution Mass Spectrometry, Abiotic Stress, Integratomics, Multi-Omics Integration.

Abstract

Introduction: One of the abiotic stresses that threaten agriculture the most is soil salinity, a problem present in more than 100 countries spread across all continents. Due to the increase in demand for agricultural products, the pressure for cultivation in these soils has increased. *Gliricidia sepium* (Jacq.) Walp. is a multipurpose tree, cultivated for improvement of soil fertility, for medicinal purposes, as wood/firewood, as charcoal, and as a shade of plantations. It is also known for its ability to adapt to a wide range of soils ranging from eroded acidic soils, sandy soils, heavy clay, limestone, and alkaline soils. Gliricidia salinity tolerance limits, alongside its responses to salt stress, are not yet well understood.

Method: The transcriptome and metabolome profiles of *G. sepium* shoots were performed on control and salt-stressed plants in a completely randomized design. Transcriptome samples were subjected to RNA-Seq using an Illumina HiSeq platform and the paired-end strategy, and data analysis was done with OmicsBox version 1.3. Metabolome samples were analyzed on a UHPLC system equipped with a reversed-phase column. High-resolution mass spectrometry (HRMS) was performed on a Q-TOF analyzer using electrospray source in ESI(+)-MS and ESI(-)-MS. Acquired data were pre-processed using XCMS Online and further exported to MetaboAnalyst 4.0 for multivariate analysis, metabolic annotation, and pathway activity observation. Omics Fusion, the web platform for integrative analysis of Omics data, was employed for carrying out the integrative analysis of transcripts and metabolites.

Results: The analysis on transcriptome and metabolome data sets characterized the plant response under three scenarios: Age effect (control plants at 2 and 45 days under stress - DAT), short-term (control and stressed plants at 2 DAT), and long-term stress (stressed plants at 2 and 45 DAT). A group of 5,672 transcripts and 107 metabolites were submitted to integrative analysis to integrate transcripts and metabolites differentially expressed in gliricidia shoots under salt stress; the phenylpropanoid biosynthesis came in first among the most affected pathways with 15 metabolites as well as five transcripts (three genes) differentially expressed.

Conclusion: The single and integrated analysis of the transcriptome and the metabolome profiles generated in this study were efficient to correlate and differentiate groups of *G. sepium* plants submitted to salinity stress, revealing genes/transcripts, metabolites, and pathways responsive to this stress. The analysis of the metabolites and genes differentially expressed in the phenylpropanoid biosynthesis pathway revealed that it plays a role in short-term stress. The single analysis of the transcriptome identified two genes coding for proteins that might play a role in gliricidia response at both the short- and long-term salt stress.

1. Introduction:

The world population is on track to reach between nine and ten billion persons by 2050, resulting from an increase of more than three billion individuals in the first half of the 21st century. This scenario has challenged the biomass production system to produce more food, feed, fiber, bioenergy, and ornamentals, among other bioproducts derived from plants, in a sustainable way. The increase in biomass production must occur while plants are affected by several more intense abiotic and biotic stresses resulted from changes in climatic conditions (FAO, 2011). One of the abiotic stresses that threaten agriculture the most is soil salinity, a problem present in more than 100 countries spread across all continents. Approximately 20% of all agricultural land in the world has either saline or sodic soils, and between 25% and 30% of the irrigated land area is affected by salt (Shahid et al., 2018).

Gliricidia sepium (Jacq.) Walp., a medium-sized legume (10-15 m) that belongs to the Fabaceae family, is originated from Central America. It shows rapid growth and is one of the most well-known multipurpose trees. It is cultivated for improvement of soil fertility, for medicinal purposes, as wood/firewood, as charcoal, and as a shade of plantations (Rahman et al., 2019). At the economic level, the gliricidia role in improving water infiltration and increasing water retention capability of the soil, reducing soil erosion, and restoring and improving the soil quality, leading to a higher crop yield, is highlighted (Diouf et al., 2017). It is also known for its ability to adapt very well to a wide range of soils, from eroded acidic soils, sandy soils, heavy clay, limestone, and alkaline soils (Rahman et al., 2019).

Gliricidia salinity tolerance limits, alongside its responses to salt stress, are not yet well understood (Rahman et al., 2019). Rahman and colleagues showed that seawater-induced salinity negatively affected several growth-related attributes in one-month-old gliricidia seedlings; postulating that proline, which showed enhanced accumulation under salinity stress, might help gliricidia plants to adjust to water deficit conditions. Proline participates in metabolic signaling and is known to be metabolized by its own family of enzymes responding to stress (Phang et al., 2010).

Several studies are available about transcriptomics and metabolomics analysis in plants (Cavill et al., 2016; Jamil et al., 2020). Transcriptomics is a technology applied to characterize the transcriptome in a cell, tissue, or organism at any given time. Unlike the genome that tends to be static information, the transcriptome is variable; and is one of the links between the genome and the phenotype of an organism (Wang et al., 2009; Zhang et al., 2010). Metabolomics is a technology applied to characterize the complete set of small-molecule

chemicals found within a biological sample. Metabolites are functional products of metabolism, and their concentration levels vary according to genetic or physiological changes. Since it provides a better representation of an organism's phenotype than any other omic, metabolomics emerges as an efficient tool to fill the phenotype-genotype gap (Zampieri and Sauer, 2017).

Due to the rise in accessibility to high throughput biological data from different omics, efforts to analyze these data separately have given rise to a more comprehensive view and with a focus on integrating different omics to obtain more robust knowledge of biological systems (Cavill et al., 2016; Jamil et al., 2020). The first successful integrative attempts using these two omics in fungi and plants date almost two decades (Askenazi et al. 2003; Urbanczyk-Wochniaket al. 2003; Hoefgen & Nikiforova, 2008). Since then, many groups have used distinct integrative approaches to gain insights into many different plant traits. Transcript and metabolite are not directly associated; however, the process of integrating them provides information that allows us to base the phenotypic data and measures provided by the metabolomics on the genetic data from the transcriptome (Cavill et al., 2016; Jamil et al., 2020). Yan and colleagues identified new target genes and metabolites by integrating data from these two omics in *Tetrastigma hemsleyanum*. These molecules led to a gain of efficiency of the anthocyanin metabolic pathway (Yan et al., 2020). Rai et al. (2020) also did it to identify genes involved in the biosynthetic pathways of the dominant groups of bioactive metabolites in *Cornus officinalis*, an important medicinal plant.

In this study, we first carried out a morphophysiological characterization of the response of *Gliricidia sepium* to salinity stress, in both the short and long-term and at five different doses of NaCl. Then, used samples from the shoots of gliricidia plants to characterize the metabolomic profile in all these treatments. At third, generated the transcriptome profile in the short and long-term stress at 0.0 and 0.8 g of NaCl per 100 g of the substrate. At last, applied conceptual, element- and pathway-based strategies to integrate metabolome and transcriptome data.

2. Materials & Methods:

2.1. Plant material and growth conditions

The accession of gliricidia [*Gliricidia sepium* (Jacq.) Steud.] used in this study belongs to the Gliricidia Collection at Embrapa Tabuleiros Costeiros (www.embrapa.br/en/tabuleiros-costeiros). After soaking the seeds in 2% sodium hypochlorite and Tween® 20 for 5 min under slow agitation, we washed them with sterile water and dried them on sterilized filter paper.

Then they were placed in a Petri dish with filter paper moistened with sterilized water until the radicle emission. Subsequently, each germinated seed was transferred individually to a 5 L plastic pot containing 4 kg of substrate previously prepared by mixing sterile soil, vermiculite, and a commercial substrate (Bioplant®), in the ratio 2:1:1 (v:v:v); and kept in a greenhouse for three months.

2.2. Experimental design and Saline stress

Groups of three-month-old gliricidia plants were kept under control conditions or subjected to saline stress (0.4, 0.6, 0.8, and 1.0 g of NaCl per 100 g of substrate) for 2 (short-term stress) or 45 (long-term stress) days. The experimental design was completely randomized with 5 replicates (plants) per treatment.

The NaCl was dissolved in deionized water to salinize the substrate. The amount of deionized water used corresponded to the difference between the amount of water previously present in the substrate and the amount of water necessary for the substrate to reach field capacity. Applying the right amount of water – up to the substrate field capacity – was a means of ensuring no leakage of the solution out of the pot and no loss of Na^+ or Cl^-. For details about moisture content, field capacity, and electric conductivity in the substrate, which were determined preliminarily, details are in Silva (2019).

We replaced the water lost by evapotranspiration with deionized water in a daily basis, and monitored electric conductivity and water potential in the substrate solution at zero, 6, 35, and 45 days after imposing the treatments (DAT) for all replicates.

2.3. Biomass and mineral analysis

After taking fresh weight (FW), the root and shoot were dried in an oven for 72 hours at 65ºC to a constant weight (dry weight – DW). The mineral analysis of samples (roots, shoot, and soil), collected at the end of the experiment, was done by Soloquímica (www.soloquimica.com.br). The data from the mineral analysis was initially analyzed using bidirectional analysis of variance (ANOVA). To compare the treatments with significant differences, we used the Tukey test ($p < 0.05$).

2.4. Metabolomics analysis

Shoots (leaves and stem) for metabolomics analysis were collected from all replicates at 2 and 45 DAT, immediately immersed in liquid nitrogen, and then stored at -80 °C until extraction of metabolites. Based on the results of the morphophysiological characterization, we selected the following treatments for metabolomics analysis: control plants at 2 and 45 DAT, stressed plants (0.4 and 0.8 g of NaCl per 100 g of substrate) at 2 and 45 DAT.

2.4.1. Chemicals and metabolites extraction

Samples were grounded in liquid nitrogen before solvent extraction. The solvents methanol grade UHPLC, acetonitrile grade LC-MS, formic acid grade LC-MS and sodium hydroxide ACS grade LC-MS were from Sigma-Aldrich (St. Louis, MO, USA); and the water treated in a Milli-Q system (Millipore, Bedford, MA, USA).

We employed a protocol adapted from the Max Planck Institute (Vargas et al., 2016; Rodrigues-Neto et al., 2018), known as All-in-One Extraction, to extract the metabolites. After transferring aliquots of 50 mg of grounded sample to 2 mL microtubes, added 1 ml of 1:3 (v:v) methanol: methyl tert-butyl ether at -20 °C, and then left for homogenization at 4 °C on an orbital shaker for 10 min, followed by an ultrasound treatment in an ice bath for another 10 min. Next, added 500 µL of 1:3 (v:v) methanol: water mixture (1:3) to each microtube before centrifugation (12,000 rpm, 4 °C for 5 min). After centrifugation, it generated three phases: an upper nonpolar (green), a lower polar (brown), and a remaining protein pellet. The apolar and polar fractions were transferred separately to 1.5 mL microtubes and vacuum dried in a Speed vac (Centrivap, Labconco, Kansa, MO, USA).

2.4.2. UHPLC-MS and UHPLC-MS/MS

After resuspending the dry polar fraction by adding 500 µL of 1:3 (v:v) methanol: water mixture, it was transferred to a vial and analyzed by UHPLC-MS/MS. We used a UHPLC chromatographic system (Nexera X2, Shimadzu Corporation, Japan) equipped with an Acquity UPLC HSS T3 (1.8 µm, 2.1 x 150 mm) reverse phase column (Waters Technologies, Milford, MA), maintained at 35 °C. Solvent A was 0.1% (v/v) formic acid in water and solvent B was 0.1% (v/v) formic acid in acetonitrile / methanol (70/30, v/v). The gradient elution used, with a flow rate of 0.4 mL/min, was as follows: isocratic from 0 to 1 min (0% B), linear gradient from 1 to 3 min (5% B), from 3 to 10 min (50% B), and 10 to 13 min (100% B), isocratic from 13 to 15 min (100% B), followed by rebalancing in the initial conditions for 5 min. The rate of acquisition spectra was 3.00 Hz, monitoring a mass range from *m/z* 70-1200 (polar fraction) and *m/z* 300-1600 (lipidic fraction).

Detection was performed by high-resolution mass spectrometry (HRMS) (MaXis 4G Q-TOF MS, Bruker Daltonics, Germany) using electrospray source in positive (ESI (+) - MS) and negative (ESI (-) - MS). The settings of the MS instrument were: final plate offset, 500 V; capillary voltage, 3800 V; nebulizer pressure, 4 bar; dry gas flow, 9 L/min, dry temperature, 200 °C. The rate of acquisition spectra was 3.00 Hz, monitoring a mass range of 70 to 1200 m/z. A sodium formate solution (10 mM HCOONa solution in 50/50 v/v isopropanol/water containing 0.2% formic acid) was injected directly through a 6-way valve at the beginning of each chromatographic run for external calibration. Ampicillin ([M+H] + *m/z* 350.11867 and

[M-H] - *m/z* 348.10288) was added to each sample and was used as an internal standard for peak normalization.

Tandem mass spectrometry (MS/MS) parameters have been adjusted to improve mass fragmentation, with collision energy ranging from 20 to 50 eV, using a step method. Precursor ions were acquired using the 3.0 s cycle time. The general AutoMS settings were: mass range, *m/z* 70-1000 (polar fraction) and *m/z* 300-1600 (lipidic fraction); spectrum rate, 3 Hz; ionic, positive polarity; pre-pulse storage, 8 μs; funnel 1 RF, 250.0 Vpp. The UHPLC-MS and UHPLC-MS/MS data were acquired by HyStar Application version 3.2 (BrukerDatonics, Germany).

2.4.3. Metabolomics data analysis

The raw data from UHPLC-MS were exported as mzMXL files, using DataAnalysis 4.2 software (Bruker Daltonics, Germany) and pre-processed using XCMS Online (Gowda et al., 2014; Tautenhahn et al., 2012), for peak detection, retention time correction and alignment of the metabolites detected in the UHPLC-MS analysis. Peak detection was performed using centWave peak detection (Δm / z = 10 ppm; minimum peak width, 5 s; maximum peak width, 20 s) and mzwid = 0.015, minfrac = 0.5, bw = 5 for alignment of retention time. The unpaired parametric t-test (Welch t-test) was used for statistical analysis.

The processed data (csv file) were exported to MetaboAnalyst 4.0, and submitted to analysis in the Statistical Analysis module (Chong et al., 2019; Chong & Xia, 2020). Before the chemometric analysis, all data variables from the polar fraction were normalized by internal standard (ampicillin-rT = 7.9 min; [M+H], *m/z* = 350.11711, [M-H], *m/z* = 348.10212); and, all data variables from the lipidic fraction were normalized by internal standard (1,2-diheptadecanoyl-sn-glycero-3-phosphocholine = 4.85 min; [M+H] + *m/z* = 762.60063). All three sets of data were scaled using the pareto method.

The differentially expressed peaks (DEP) were selected according to the following criteria: Variable Importance in Projection - VIP values ≥ 1, obtained from the PLS-DA model; adjusted P-value (FDR) ≤ 0.05, of the Welch t-test; and Log_2 (Fold Change) ≠ 1. The selected DEPs were then submitted to analysis in the MS Peaks to Pathway module (Chong et al., 2019; Chong & Xia, 2020) and analyzed using the following parameters: molecular weight tolerance of 5 ppm; mixed ion mode; joint analysis using the mummichog algorithm (Li et al., 2013) with a P-value cutoff of 1.0 □ 10^{-5} and Gene Set Enrichment Analysis - GSEA (Subramanian et al., 2005) algorithms, and the latest KEGG version of the *Arabidopsis thaliana* pathway library.

In the case of a DEP with two or more matched forms (isotopes) and later a matched compound with two or more DEPs, the initial criterion of metabolite selection applied was the

mass difference comparing to the metabolite database – choosing the smallest one. The second criterion was the adduct study of each candidate back in its mass spectra. Then, we applied the formula and exact mass data from KEGG; and, finally, performed the putative annotation of the metabolites of interest, with one or two candidates on each detected ion.

The KEGG IDs of the matched compounds were then submitted to pathway analysis (integrating enrichment analysis and pathway topology analysis) and visualization in the Pathway Analysis module (Chong et al., 2019; Chong & Xia, 2020) and analyzed using the Hypergeometric Test and the latest KEGG version of the *Arabidopsis thaliana* pathway library.

2.5. Transcriptomics

Shoots (leaves and stem) for transcriptomics analysis were collected from all replicates at 2 and 45 DAT, immediately immersed in liquid nitrogen and then stored at -80 °C until RNA extraction. Based on the results of the morphophysiological and metabolomics characterization, we selected the following treatments for transcriptomics analysis: control plants at 2 and 45 DAT, stressed plants (0.8 g of NaCl per 100 g of substrate) at 2 and 45 DAT. Three replicates per treatment were randomly selected for total RNA extraction, library preparation, and sequencing.

2.5.1. Total RNA extraction and quality analysis, library preparation and sequencing

Total RNA was isolated from gliricidia shoots using the Qiagen RNeasy® Plant Mini kit (QIAGEN, CA, USA) following the manufacturer's protocol. RNA quantity was measured using a Nanodrop Qubit 2.0 Fluorometer (Life Technologies, CA, USA), and RNA quality was evaluated with an Agilent Bioanalyzer Model 2100 (Agilent Technologies, Palo Alto, CA). Samples were subjected to RNA-Seq using an Illumina HiSeq platform at the GenOne Company (Rio de Janeiro, Brazil), using the paired-end strategy. The raw sequence data (24 fastq files) have been uploaded in the Sequence Read Archive (SRA) database of the National Center for Biotechnology Information under the BioProject number of PRJNA634354.

2.5.2. Transcriptomics data analysis

All the transcriptomics analysis was performed with OmicsBox version 1.3 (OmicsBox, 2019). We used FastQC (Andrews, 2010) and Trimmomatic (Bolger et al., 2014) to perform the quality control, to filter reads and remove low-quality bases. The minimum average quality of reads kept was 30, and the minimum length of reads was 75. The default parameters from OmicsBox version 1.3 were used to create a "de novo" transcriptome without a reference genome through the softwares Trinity version 2.8.5 (Grabherr et al., 2011) and Bowtie2 version 2.3.5.1 (Langmead and Salzberg, 2012). The RNA-Seq data were aligned to the reference transcriptome using default parameters from OmicsBox version 1.3 through software STAR

(Dobin et al., 2013); and, to quantify expression at gene or transcript level we used the default parameters from OmicsBox version 1.3 through HTSeq version 0.9.0 (Anders et al., 2015).

The pairwise differential expression analysis between different experimental conditions was performed through edgeR version 3.28.0 (Robinson et al., 2010), applying a simple design and an exact statistical test, without the use of filter for low counts genes. The functional analysis of the differentially expressed genes was performed combining differential expression results with functional annotations from the high-throughput functional annotation and data mining pipeline in OmicsBox version 1.3 (Götz et al., 2008).

2.6. Integratomics analysis

Omics Fusion (Brink et al., 2016), the web platform for integrative analysis of Omics data (https://fusion.cebitec.uni-bielefeld.de/), was employed for carrying out the integrative analysis of transcripts and metabolites. The input data used was the Log_2 (Fold Change) data of differentially expressed transcripts and metabolites. First, to check the data distribution, we used the Data Overview module and then the Scatter Plot one for the correlation analysis between the two sets of data – a pairwise combination of the different scenarios evaluated.

For subsequent analysis, we used the modules KEGG feature distribution and Map data on the KEGG pathway. The former module employed to verify which metabolic pathways had more transcripts and metabolites differentially expressed, and the latter to map these data differentially expressed in the metabolic paths in question. For the KEGG feature distribution module, we applied the joint analysis of transcripts and metabolites with a threshold of 10, and for the Map data on the KEGG pathways, the organism code gmx (*Glycine max*) was used for mapping.

3. Results

3.1. *Gliricidia sepium* response to salt stress

As expected, the addition of increasing levels of NaCl to the substrate led to an increase in electrical conductivity (EC) and a reduction in the water potential (Ψ_w) of the saturation extract. The saline level of 0.0 g of NaCl, which did not receive the addition of NaCl, presented EC greater than zero dS/m and reduced Ψ_w; this was probably due to the ionic effect of the salts that are present in the chemical fertilizers that were added to the substrate. All saline levels used provided electrical conductivity higher than that of a soil considered as saline (Figure 1).

The doses of salt in the substrate and the weather and climatic conditions influenced the evapotranspiration rates (data not shown). It was observed that the evapotranspiration rate values showed a negative correlation to the NaCl dose applied to the plant substrate. Thus, the saline level that showed the highest value of real evapotranspiration rate was the control. As

the saline levels were increased, a proportional reduction in the actual evapotranspiration rate was observed. However, at 45 DAT, the evapotranspiration rate in all saline levels were practically equal to the control.

The methodology established for the application and monitoring of salinity stress in *G. sepium*, by means of a substrate salinization protocol, allowed the identification of two distinct responses of these gliricidia plants to salt stress, depending on the amount of NaCl used. First, plants grown for 45 days on a substrate with approximately 15 dS/m (0.4 g of NaCl per 100 g of substrate) or less, did not show any visual symptoms of stress on the aerial parts, such as leaf wilt, yellowing, burning or falling; although, they experienced a reduction in shoot and root biomass (Figure 2).

However, plants grown on a substrate with approximately 30 dS/m (0.6 g of NaCl per 100 g of substrate) or more lost all their leaves in the first week after the stress onset (Figure 3A-H). In the third DAT, the leaves in these plants started to show a strong wilt symptom (Figure 3B), and in the fourth DAT they started to fall (Figure 3C). Some plants even presented a symptom os burning in the leaves (Figure 3D). At the end of the first week of stress, the stressed plants – at ≥30 dS/m – lost almost all leaves (Figure 3E). However, about three weeks after the beginning of the stress, it was possible to see some new leaves starting to emerge from the lateral meristems (Figure 3F); and their growth continue throughout the rest of the experiment (Figure 3G and 3H). As the amount of salt on the substrate increased (1.0, 1.5, and 2.0 g of NaCl per 100 g of substrate), the emergency of new leaves took longer to happen, up to a point where it did not happen at all (data not shown, from additional experiments).

As expected, the addition of increasing levels of NaCl led to an increase in the concentration of ions Na^+ and Cl^- in the substrate, and a consequent increase in the index of saturation by sodium as well (Supplementary Table 1). Plants under salt stress accumulated more Na^+ in the shoots than in the roots; at the two highest doses of salt used it was close to twice the amount. Indeed, under salt stress, most halophytes accumulate more sodium in their shoots than in their roots (van Zelm et al., 2020). There was a four fold increase of ion Cl^- in the roots compared to shoots in the control plants; however, a different pattern was seen in the stress plants, where most of it accumulated in the shoots.

The addition of NaCl also led to an increase in the availability of potassium in the soil solution. Approximately 2/3 of the accumulated K^+ in the plant was present in the shoots (Supplementary Table 1). Halophyte plants maintain a higher level of potassium than glycophytes and a more optimal K^+/Na^+; the cellular balance between sodium and potassium is important for plant survival in saline soils (van Zelm et al., 2020). The K^+/Na^+ in the shoots

and roots of gliricidia plants decreased directly proportional to the increase of NaCl in the substrate; and, at 0.8 g of NaCl per 100 g of the substrate, it was approximately 0.4 in either case (Supplementary Table 1). The ion Ca^{2+} did not alter its availability in the solution due to the increasing level of NaCl added, and there was no significant difference in the amount of Ca^{2+} accumulated in the plant roots – independently of the NaCl dose applied. However, in the shoots, one can see a decrease in Ca^{2+} as the NaCl increases.

Plants commonly respond to salt stress with a reduction in N assimilation (Ashraf et al., 2018). In gliricidia, salt stress reduced the nitrogen content in the shoots and roots as the concentration of NaCl in the substrate increased (Supplementary Table 1). Mg^{2+} concentrations in the substrate and the plants do not change due to the increase in NaCl levels. The amount of boron in the shoots significantly reduced as the levels of NaCl increased. Zinc in the substrate did not change due to the increase in NaCl levels, except for the highest concentration. The increasing levels of NaCl added did not significantly alter the concentrations of zinc in the shoots; however, in the roots, one can see a drop of about 40% compared to the control.

Shoots from plants under short-term salt stress had approximately 12 times more Na^+ and four times more Cl^- than the control plants, while the roots showed close to eight times more Na^+ and about half the amount of Cl^- than the control plants. On the other hand, the shoots from plants under long-term salt stress had approximately 36 times more Na^+ and five more Cl^- than the control plants; while the roots showed close to eleven times more Na^+ and 70% the amount of Cl^- of the control plants (Supplementary Table 1).

Both shoot and root-dry biomass decreased directly proportional to the increase of NaCl in the substrate (Figure 2). Those plants under short-term salt stress, subjected to 0.4 g of NaCl per 100 g of substrate, had close to 58% less shoot biomass than the control plants. These plants did not lose the leaves during the 45 days of stress. As plants under stress at NaCl doses of ≥0.6 g per 100 g of substrate lost all leaves in the first week after the onset of salinity stress, regaining new leaves two weeks after that, one can not establish a comparison between shoot biomass from control and stressed plants. The addition of increasing levels of NaCl led to a linear decrease in the amount of root dry biomass, with the dose of 0.8 g per 100 g of the substrate leading to a 63% loss of biomass (Figure 2).

3.2. Gliricidia metabolome under salinity stress

To analyze the metabolome data, the following four data sets were used: all treatments – AT [control plants at 2 and 45 DAT, and stressed plants (0.4 and 0.8 g of NaCl per 100 g of substrate) at 2 and 45 DAT]; age effect – AE (samples from the control plants at two and 45 DAT); short-term stress – STS (samples from the control and the stressed plants (0.8 g of NaCl

per 100 g of substrate) at 2 DAT); and long-term stress – LTS (samples from the stressed plants (0.8 g of NaCl per 100 g of substrate) at 2 and 45 DAT); all with five biological replicates.

PLS-DA (partial least squares discriminant analysis) permutations tests were performed using the AT data set to validate the model, applying permutation number = 2,000. When evaluated by group separation distance, the probability that the model was created by chance was less than 0.0005%, independent of the fraction – polar-positive, polar-negative, and lipidic-positive (Figure 4); and the evaluation by prediction accuracy showed that it was less than 0.0065% for the polar-positive fraction, and less than 0.0005% for the other two fractions (Figure 4).

First, the AE data set was employed to measure the degree of a possible age effect in the changes of the metabolome profiles of gliricidia, considering that there is a 43 day gap between the two assessments. Then, the short-term stress data set was employed to evaluate how distinct are the metabolome profiles of the control and stressed plants at 2 DAT, just one day before the leaves started to wilt. At last, the long-term stress data set was used to evaluate how distinct are the metabolome profiles of the stressed plants at two different moments (2 and 45 DAT).

The accumulated variance explained by the first three out of five principal components tested using the supervised multivariate method PLS-DA in the polar-positive fraction was 53.1%, 72.5%, 69.8%, and 69.5% for AT, AE, STS, and LTS, respectively. In the polar-negative fraction, the accumulated variance was 64%, 73.4%, 74.2%, and 86.5%; and in the lipidic-positive fraction, they were 46%, 69.4%, 75.1%, and 76.1%, respectively (data not shown).

The cross-validation analysis determines the optimal number of components necessary to build the PLS-DA model by measuring three criteria. These criteria are R^2 (the sum of squares captured by the model), Q^2 (the predictive ability of the model estimated by cross-validation), and Accuracy (Chong et al., 2019). When choosing the three-component model based on Q^2, the Q^2 values obtained in the polar-positive fraction were: 80%, 95%, 87%, and 89% for AT, AE, STS, and LTS, respectively. In the polar-negative fraction, the Q^2 values obtained were 77%, 90%, 94%, and 86%; and in the lipidic positive one, they were 79%, 90%, 90%, and 81%, respectively (data not shown).

The samples applied to estimate the AE scenario contained 1,368, 1,798, and 4,190 peaks, respectively, in the polar-positive, polar-negative, and lipidic-positive fractions (Table 1). On average, 90.50% did not show a difference in expression between two and 45 DAT, while 1.90% up-regulated at 45 DAT, and 7.61% down-regulated. Three, 39, and 112 peaks up-regulated in the polar-negative, polar-positive, and lipidic-positive fractions, much less than the number of down-regulated, which were 123, 135, and 256, respectively (Table 1).

In the case of the samples applied to estimate the change in the STS scenario, they contained 1,380, 1,815, and 4,190 peaks, respectively, in the polar-positive, polar-negative, and lipidic-positive fractions (Table 1). On average, 93.03% did not show a difference in expression between control and stressed plants at 2 DAT, while 3.94% up-regulated in the stressed plants, and 3.04% down-regulated. A total of 37, 93, and 127 peaks up-regulated in the polar-negative, polar-positive, and lipidic-positive fractions, while 29, 46, and 175 down-regulated, respectively (Table 1).

At last, the samples applied to estimate the changes in the LTS scenario, in comparison with the STS one, contained 1,370, 1,817, and 4,190 peaks, respectively, in the polar-positive, polar-negative, and lipidic-positive fractions (Table 1). On average, 90,75% did not show any difference in expression between the stressed plants at two and 45 DAT, while 0.55% up-regulated in the stressed plants at 45 DAT, and 8.73% down-regulated. Zero, nine, and 41 peaks up-regulated in the polar-negative, polar-positive, and lipidic-positive fractions, while 139, 166, and 269 down-regulated, respectively (Table 1).

As already stated in the Materials & Methods section, a differentially expressed peak (DEP) is a peak with a VIP value ≥ 1, an adjusted P-value (FDR) ≤ 0.05, and Log_2 (Fold Change) > 1 (up-regulated) or Log_2 (Fold Change) < -1 (down-regulated). Once it was clear that the amount of DEPs was less than 10% in all scenarios tested (Table 1), a new question emerged: What would happens to the peaks differentially expressed in the STS scenario once the stressed plants reached the LTS scenario?

Out of the 257 DEPs up-regulated in STS, only two up-regulated again 45 DAT – in the LTS scenario (Table 2). On the other hand, 114 remained at the same level of expression as before, and 141 down-regulated. In the case of the 250 DEPs down-regulated in STS, five up-regulated, 238 maintained the same level of expression as before, and seven down-regulated further in the LTS scenario. At last, out of the 6,880 peaks not differently expressed in STS, 43 up-regulated, and 426 down-regulated in LTS, while 6,401 remained at the same level of expression as before. Ten polar-positive peaks that did not differentially expressed in STS, were deleted in the data integrity check stage due to a constant or single value across samples.

All 976 peaks differentially expressed in STS or in LTS, or in both (Table 2), were then submitted to functional interpretation via analysis in the MS Peaks to Pathway module (Chong et al., 2019; Chong & Xia, 2020), as described in the Materials and Methods section. The combined mummichog and GSEA pathway meta-analysis resulted in a list of 61 ranked pathways that are enriched in this group of DEPs. Five pathways were significantly perturbed in both algorithms – galactose metabolism, starch and sucrose metabolism, biosynthesis of

unsturated fatty acids, aminoacyl-tRNA biosynthesis, and arginine and proline metabolism pathways (Figure 5A).

After applying the initial criteria of metabolite selection, as described in the Materials and Methods section, 107 DEPs with a hit to just one known compound, 19 with hits to two compounds, two with hits to three, two with hits to four, one with hits to six, and one with hits to seven compounds were selected (Supplementary Table 2). The group of 107 DEPs with a hit to just one known compound was chosen to be submitted to the pathway topology analysis module, resulting in a list of 64 ranked pathways. The phenylpropanoid biosynthesis pathway, with a FDR (False Discovery Rate) of 0.00081712 and a pathway impact of 0.31153, came out in the top of this rank (Figure 5B). This pathway had 13 differentially expressed metabolites with the highest level of significance in the set of matched metabolites submitted to analysis.

3.3. Gliricidia transcriptome under salinity stress

The 24 fastq files (raw sequence data) generated using the 12 samples collected had 319,889,065 reads pairs or approximately 48 Gbases; 150 nucleotides per read (Supplementary Table 2). A total of 305,588,306 high-quality-pairs of reads - with an average quality of reads ≥30, and the minimum length of 75 nucleotides – remained after pre-processing the raw sequence data, or 95.53% of it. To assemble the Reference Transcriptome (RT), as well as to perform the mapping, counting, and differential expression, we used these high-quality sequences. The RT assembled has 53,735 features – longest ORFs per gene – ranging from 297 to 16,323 nucleotides in length. Previously to the submission to the functional annotation in OmicsBox, a BlastX analysis against the genome of *Glycine max* (L.) Merrill – available at NCBI (Glycine_max_v2.1, BioProject PRJNA19861, BioSample SAMN00002965) – was performed in May 2020 (data not shown).

The differential expression analysis was performed in order to measure the degree of a possible age effect using the AE data set, the differences in the short-term stress using the STS data set, and the differences in the long-term stress using the LTS data set. Differentially expressed transcripts (DET) are those with a FDR ≤ 0.05, and Log_2 (Fold Change) > 1 (up-regulated) or Log_2 (Fold Change) < -1 (down-regulated). Out of the 53,735 features in the RT, 1,347 (2.51%) up-regulated and 1,397 (2.60%) down-regulated in the control plants at 45 DAT comparing to 2 DAT (AE scenario). In the STS scenario, 824 (1.53%) up-regulated and 487 (0.91%) down-regulated in the stressed plants comparing to the control plants at 2 DAT. At last, in the LTS scenario, 1,920 (3.57%) up-regulated and 2,229 (4.15%) down-regulated (Table 1).

Once it was clear that the amount of DETs was also less than 10% – as it was seen before for DEPs (Table 1) – the same question made for DEPs emerged for DETs: What is the fate in LTS of the DETs from STS? Out of the 824 DETs up-regulated in STS, only two up-regulated again in LTS, while 158 remained at the same level of expression as before and 664 down-regulated (Table 2). In the case of the DETs down-regulated in STS, 264 up-regulated, 213 maintained the same level of expression as before, and ten down-regulated further in LTS.

A set of 12 DETs, being two that up-regulated twice – in STS and LTS as well – and ten that down-regulated twice was chosen to further characterization (Figure 6). First, the two DETs that up-regulated twice also up-regulated in AE scenario; however, in the STS+LTS cumulative scenario, one DET experienced a >260-fold increase, while in the AE scenario it had an approximately 51-fold increase. The second DET experienced a 40-fold increase in the STS+LTS cumulative scenario and an approximately 21-fold increase in the AE one. All ten DETs that down-regulated twice showed a level of expression lower than when in the AE scenario, showing that the stress did also contribute to the reduction in the expression level. For the rest of them, the decrease in expression was higher in the LTS scenario comparing to the AE (Figure 6).

Three out of these 12 DETs got positive hits in BlastX to uncharacterized proteins in soybean, including one of the two DETs that up-regulated twice and the one that down-regulated in STS but did not change the level of expression in the adaptation phase. The remaining DETs are homologs to genes coding for: a WAT1-related protein At4g28040, a bidirectional sugar transporter SWEET14-like protein, a G-type lectin S-receptor-like serine/threonine-protein kinase At4g27290, a protein NIM1-INTERACTING 1-like, a protein SAR DEFICIENT 1 isoform X1, a F-box protein At2g27310, a probable WRKY transcription factor 50, a VQ motif-containing protein 22, and a lysosomal Pro-X carboxypeptidase-like isoform X1 (Table 3).

3.4.) Integrating gliricidia metabolome and transcriptome data

A group of 5,672 DET (up and down-regulated from all three scenarios – AE, STS, and LTS) and 107 DEP (only those with a hit to just one known compound) were submitted to the Omics Fusion (Brink et al., 2016) platform for integrative analysis to integrate transcripts and metabolites differentially expressed in gliricidia under salt stress. The distribution of DETs and DEPs, in all three scenarios, obeyed a normal distribution (Figure 7). Due to the number of compounds employed in the analysis, the histograms for metabolites showed only one normal curve; however, in the case of transcripts, it is possible to see two well defined normal curves in the STS and LTS scenarios, representing both the down and up-regulated DETs (Figure 7).

The average Log_2 (Fold Change) of the down-regulated DETs in the STS scenario is bigger than the one in the LTS scenario, while the opposite is true for the up-regulated DETs.

The correlation analysis between the three different scenarios showed no correlation between AE and LTS, a negative correlation between STS and LTS, and a positive correlation between AE and STS for the metabolites (Figure 8). Regarding the transcripts, the correlation analysis showed no correlation between AE and STS, a negative correlation between STS and LTS, and a positive correlation between AE and LTS for the transcripts (Figure 8).

The phenylpropanoid pathway also came first in the rank of affected pathways when using the Omics Fusion (Brink et al., 2016) platform, with 15 metabolites and six transcripts (three genes). Eleven out of the 15 differentially expressed metabolites from this pathway were down-regulated in LTS when compared to their amounts in the STS (Figure 9). The four remaining metabolites were down-regulated in STS and slightly up-regulated in LTS compared to control plants and STS, respectively.

The expression profiles of these 15 differentially expressed metabolites were analyzed in the AT (all treatments) data set. A one-way parametric ANOVA – with an adjusted p-value (FDR) cutoff of 0.05 – and the Fisher's Least Significant Difference (Williams and Abdi, 2010) posthoc test was applied. All 15 metabolites had FDR $\leq$ 0.05 (data not shown).

Based on the expression profile in the AT data set, the metabolites were separated into different groups (Figure 10). The first group, up-regulated under salt stress (0.4 and 0.8 g of NaCl per 100 g of substrate) at 2 DAT, had L-phenylalanine (C00079), L-tyrosine (C00082), and spermidine (C00315) in it; however, independent of the NaCl level, they had at 45 DAT an average peak intensity statistically similar to the control plant at two and 45 DAT. Figure 10.A shows the expression profile of L-phenylalanine an example of this group.

Five metabolites from the phenylpropanoid biosynthesis pathway were down-regulated in LTS, being two coniferyl alcohol derivates (C02666 and C05619), two paracoumaryl alcohol derivates (C02646 and C05608), and one sinapate derivate (C00482). Three of them (C02646, C02666, and C05608) showed strong up-regulation under the highest level of salt (0.8 g of NaCl per 100 g of substrate) at 2 DAT, getting at 45 DAT to a level of expression similar or lower than the one in the control plants at two and 45 DAT. Figure 10.B shows the expression profile of coniferyl aldehyde as an example of the group. C00482 and C05619 did show a small up-regulation at 0.4 g of NaCl per 100 g of substrate at 2 DAT, dropping to a level similar or inferior to the control plants at 45 DAT (data not shown).

Another group of metabolites, composed of C05838 (2-Coumarinate), C05839 (*beta*-D-glucosyl-2-coumarinate), and C10434 (5-O-caffeoylshikimic acid), showed a reduction in

expression under salt stress (0.4 and 0.8 g of NaCl per 100 g of substrate) at two and 45 DAT; while C18069 (N1,N5,N10-tricoumaroyl spermidine) showed a reduction in expression under salt stress (0.4 and 0.8 g of NaCl per 100 g of substrate) only at 45 DAT. These four metabolites had expression levels under the highest NaCl dose at 45 DAT lower than the control plants. The expression profiles of N1,N5,N10-tricoumaroyl spermidine, and *cis*-beta-D-glucosyl-2-hydroxycinnamate are shown in Figure 10 (C and D), respectively, as examples of the group.

At last, C00852 (chlorogenate) showed a reduction of expression under salt stress (0.4 and 0.8 g of NaCl per 100 g of substrate) at two and 45 DAT; and C00761 (coniferin) and C02887 (sinapoyl malate) down-regulation at 45 DAT seems more a question of age effect (data not shown).

4. Discussion

4.1. *Gliricidia sepium* response to a long-term, and extremely high salt concentration, stress

According to their response to salt stress, plants are classified as glycophytes or halophytes. Halophytes can complete their life cycle in an environment where the salt concentration is equal or greater than 200 mM NaCl - approximately 20 dS/m at 25 °C, and glycophytes can not do so (Flowers et al., 1986; Flowers & Colmer, 2008). So far, there is no report showing that *G. sepium* is indeed a halophyte species. Although, based on the results of this study, it is probably the case. Here gliricidia plants grown under approximately 30 dS/m or more lost all their leaves in the first week after the stress onset; however, they started developing new leaves about three weeks after the beginning of the stress and continued to produce more new leaves for additional four weeks until the experiment ended. Once the salt stress level (≥30 dS/m) was the same throughout the entire experiment, this phenotype is not an example of recovery from the stress. It is likely a case of adaptation to the stress condition (Acosta-Motos et al., 2017) that started with a very drastic measure taken by the plant (loss of all leaves) and followed by a transition to a state that allowed the plant to resume developing new leaves.

Plants developed high phenotypic plasticity, such as rapid responses to aggressive environmental factors and adaptations to changing environments (Ashapkin et al., 2020). Plants tolerant to NaCl respond to this stress by implement a series of adaptations to acclimate to salinity, including morphological, physiological and biochemical changes; and increase in the root/canopy ratio is one of these changes (Acosta-Motos et al., 2017). In the present study, it was possible to see an increase in the root/canopy ratio on gliricidia plants grown for 45 days

under approximately 15 dS/m (Figure 2). Gliricidia plants grown under 30 dS/m or more for 45 days also showed an increase in the root/canopy ratio.

Adaptation to various abiotic stresses is critical for the survival and biomass accumulation of sessile plants, particularly those perennial tree species with a relatively long-life cycle (Liu et al., 2019), such as *G. sepium*. Acosta-Motos and colleagues also point out that changes in the leaf anatomy is one of these changes that ultimately lead to preventing leaf ion toxicity, thus maintaining the water status in order to limit water loss and protect the photosynthesis process (Acosta-Motos et al., 2017). In this study, the evapotranspiration rate measured in gliricidia plants at 45 DAT in all saline levels were practically equal to the control, and plants under stress from the two highest doses of salt used accumulated almost twice the amount of Na^+ in the shoots comparing to the roots.

Understanding plant responses and adaptation mechanisms to severe salt stress conditions, such as the one from gliricidia reported in this study, is the key to improve crops economically important that could then serve biosaline agriculture. As a relatively new way of dealing with salinity in agriculture, biosaline agriculture uses cultivation systems for saline environments developed taking advantage of the ability of halophytes and/or salt-tolerant glycophytes plants to grow under saline conditions in combination with the use of saline soils and water resources, and better soil and water management (Ventura et al., 2015; Duarte & Caçador, 2021). To further boost biosaline agriculture initiatives wherever possible, it is vital to understand the physiological, metabolic, and biochemical responses of plants to salt stress and to mine the salt tolerance-associated genetic resource in nature (Zhao et al., 2020).

Collecting, characterizing, and domesticating halophyte and/or salt-tolerant glycophytes species are the front runners to develop salt-tolerant crops. Besides that, there is the strategy to promote – whenever possible – the vertical or horizontal transference of this trait to crops economically important (e.g., soybean, wheat, corn, rice, sugarcane, to mention a few). The novel CRISPR–Cas genome editing system will be a factor key to achieve horizontal transfer (Zhu et al., 2020). Powered mainly by transcriptomics, proteomics, and metabolomics data sets, the era of multi-omics systems biology research (Jamil et al., 2020) will contribute significantly to expand the existing knowledge and facilitate either conventional and biotechnological breeding efforts to develop salt-tolerant crops. Besides reporting for the first time the ability of gliricidia plants to adapt to a severe salt stress condition, the present study is also the first one to integrate metabolomics and transcriptomics data to gain a better understanding of this species short- and long-term responses to such an abiotic condition.

4.2. The integration of the metabolome and transcriptome profiles of *Gliricidia sepium* suggests a role for the phenylpropanoid biosynthesis pathway in the response to salinity stress

In the present study, the phenylpropanoid pathway came as the first ranked pathway in two distinct analyses. In the first, using the integrating enrichment and pathway topology analysis – from the pathway analysis module in MetaboAnalyst 4.0 (Chong et al., 2019; Chong & Xia, 2020) - 13 out of the 46 metabolites of this pathway were among the metabolites in the list of 107 differentially expressed compounds submitted to analysis. In the second, using the Omics Fusion (Brink et al., 2016) platform for integrative analysis to integrate transcripts and metabolites differentially expressed in gliricidia under salt stress, this pathway also came first with 15 metabolites and five transcripts (three genes). These results put this pathway - and the metabolites and genes in it - in the center of interest to further learn about the role it plays in the response of *G. sepium* to salinity stress. The gliricidia adaptation phenomena to salinity stress described in this study is a model interesting for further understanding the flux control in this complex biosynthetic pathway, as well as to the identification of targets for biotechnological manipulation.

There are three principal kinds of secondary metabolites biosynthesized by plants: phenolic compounds, terpenoids/isoprenoids, and alkaloids and glucosinolates; and the former represents the largest group of secondary metabolites in plants (Sharma et al., 2019; Santos-Sánchez et al., 2019). Phenylpropanoids are phenolics compounds derived from phenylalanine and/or tyrosine; and are involved in plant defense, structural support, and survival (Deng and Lu, 2017; Sharma et al., 2019; Santos-Sánchez et al., 2019).

According to Sharma et al. (2019), abiotic stresses disturb the balance between ROS generation and scavenge and accelerate ROS propagation that damages nucleic acids, proteins, carbohydrates, and lipids; and eventually leads to cell death. Plants growing under stressful environments can biosynthesize more phenolic compounds in comparison to plants growing under normal conditions; and these compounds have antioxidative properties and are capable of scavenging free radicals, resulting in a reduction of cell membrane peroxidation, hence protecting plant cells from ill effects of oxidative stress (Sharma et al., 2019).

Phenylpropanoid metabolism is at the interface of primary and secondary metabolism, and the phenylpropanoid pathway, also known as the phenylalanine/hydroxycinnamate pathway, is likely the most studied secondary metabolism pathway in plants. Plants exhibiting increased polyphenols synthesis under abiotic stresses usually show better adaptability to

limiting environments (Kumar et al., 2019; Sharma et al., 2019). Based on chemical structures, there are five phenylpropanoid groups: flavonoids, monolignols, phenolic acids, stilbenes, and coumarins (Deng and Lu, 2017). Flavonoids, monolignols, coumarins, and stilbenes can act as defensive components in plants against various biotic and abiotic stresses, and salicylic acid is a phenolic phytohormone that acts as a signaling molecule in plant response to diverse biotic and abiotic stresses (Deng and Lu, 2017).

All 15 metabolites from the phenylpropanoid biosynthesis pathway differentially expressed in the shoots of gliricidia plants under salt stress (approximately 35 dS/m) at 45 DAT show an average peak intensity similar or lower than the one in the control plants. These results show that it is likely that these 15 metabolites have no role in maintaining the status of salt stress-adapted plants described above for *G. sepium*. However, as several of them were differentially expressed – either up or down-regulated - in the short-term stress phase (2 DAT), it signals a possible role of the metabolites from the phenylpropanoid biosynthesis pathway in this initial stage of salt stress. It is necessary to state that this present study has focused on the metabolome and transcriptome of the shoots of gliricidia plants under salt stress, and did so trying to understand what was different in the new leaves produced during the adaptation phenomena after losing all of them just one week after the stress onset. A similar or broader multi-omics study of the roots certainly will add additional insights into the process of understanding the adaptation phenomena seen in these gliricidia plants.

Phenylalanine ammonia-lyase (PAL), cinnamate-4-hydroxylase (C4H), and 4-coumarate—CoA ligase (4CL) are the enzymes that catalyze the first three steps in the reaction sequence of the phenylpropanoid pathway; and genetic inhibition of PAL, C4H, and 4CL genes significantly reduces the phenolic compounds content in several plant species (Feduraev et al., 2020). The Omics Fusion (Brink et al., 2016) platform showed via a transcriptome and metabolome integrative analysis that five transcripts (three genes) coding for proteins that are part of the phenylpropanoid pathway were differentially expressed in gliricidia shoots under salt stress. The transcriptomic analysis revealed that two of these genes (*PAL* and *4CL*) were differentially expressed in plants under salt stress at 45DAT. The homolog of the *PAL* gene in gliricidia had its normalized counts per million (CPM) reduced approximately six-folds due to the age effect (data not shown); however, it underwent an additional three folds reduction due to the salt stress. No significant reduction was seen for the *PAL* gene in gliricidia at 2 DAT. The *4CL* gene showed a four-fold fall in CPM due to the age effect and an additional two-fold reduction due to the salt stress (data not shown). These results show that the low amount of L-

phenylalanine in the shoots of salt stress-adapted *G. sepium* plants is not due to an overexpression of the *PAL* gene and the consequent production of *trans*-cinnamate (C00423).

The remaining three transcripts coded for homologs of peroxidase genes in *Glycine max,* the *PER1,* the *PRX2*, and the LOC100817540 gene. The *PER1* and the *PRX2* genes encode proteins with peroxidase activity that respond to oxidative stress. In plants, the cellular regulation through a complex network involving redox input elements, transmitters, targets, and sensory proteins, such as peroxiredoxins (Prx or Prdx), is part of the antioxidant defenses (Tovar-Méndez et al., 2011; Perkins et al., 2015; Rhee, 2016). Prx constitutes a large and highly conserved family of peroxidases that catalyze the reduction of H_2O_2, alkyl hydroperoxides, and peroxynitrite to water, alcohols, or nitrite, respectively; and contain one or two Cys residues at the active site and usually function as monomers or dimers (Perkins et al., 2015; Rhee, 2016). There are four types of Prxs in plants (1CPrx, 2CPrx, PrxII, and PrxQ), which protect the nuclei, plastids, cytosol, and the mitochondria against excess ROS in stressful conditions, besides being implicated in redox signaling (Tovar-Méndez et al., 2011).

In the present study, the homolog of the *PER1* gene in gliricidia had its normalized CPM increased approximately six-fold due to the age effect, and decreased approximately three-fold due to osmotic stress. At 45 DAT, its CPM in the stressed plants was about 75% of that in the control plants at the same age. On the other hand, the homolog of the *PRX2* gene in gliricidia had its CPM increased approximately four-fold due to the age effect, and decreased approximately four-fold due to osmotic stress. At 45 DAT, its CPM in the stressed plants was about 10% higher than the one in the control plants at the same age. At last, the homolog of the LOC100817540 gene in gliricidia had its CPM increased approximately 45% due to the age effect, and approximately five-fold due to osmotic stress. At 45 DAT, its CPM in the stressed plants was about the same as in the control plants at the same age. The *PRX2* and the LOC100817540 homolog genes were approximately 34 and 18-fold higher in the control gliricidia plants at 2 DAT, respectively, when compared to the *PER1* homolog gene (data not shown). These results show that whether any of these three peroxidases would play a role in the response of gliricidia plants to the salt stress, it would be probably the protein coded by the homolog of the LOC100817540 gene, and it would be in the STS scenario. In the LTS scenario, on the other hand, it seems that none of them has any role in maintaining the status of salt stress-adapted plants described above for *G. sepium*.

4.3. Genes up- or down-regulated twice – in the short- and long-term stress – give additional insights on the response of *Gliricidia sepium* to salinity stress

GENE_ORF_DN2854_c0_g2 and GENE_ORF_DN1810_c0_g1, the only two genes that up-regulated at the STS and at the LTS scenarios (Table 2), code for an uncharacterized protein and a bidirectional sugar transporter SWEET14-like protein, respectively. Sugars transporters perform an important role in development, metabolism, growth, and homeostasis in plants; and there are three distinct superfamilies of sugars transporters: the glucose transporters (GLUTs), the sodium solute symporter family, such as sodium glucose cotransporters (SGLTs) and Sugar Will Eventually be Exported Transporter (SWEET) proteins (Jeena et al., 2019). The proteins from the SWEET family contains seven predicted transmembrane domains with two internal triple-helix bundles; and plant genome normally contains about 20 SWEET paralogs which are found to be differentially expressed in tissues and are involved in the transport of different sugar molecules (Jeena et al., 2019). Salinity stress (150 mM of NaCl) up-regulated *SWEET14* gene in the stem of *Arabidopsis thaliana* plants and down-regulated in the leaves of rice (Chen et al., 2019; Sellami et al., 2019). The gene coding for the bidirectional sugar transporter SWEET14-like protein in gliricidia had its expression level increased of 51.21-fold at AE, 47.06-fold at STS, and 5.59-fold at LTS, resulting in an increase of 216 times in the level of expresion, already discounted the age effect. These results show that this protein might play a role in gliricidia response at both the short- and long-term salt stress. The same is true for the gene coding for an uncharacterized protein, which had its expression level increased almost 20 times only due to the salt stress, already discounted the age effect.

Among the ten genes that down-regulated twice, at the STS and at the LTS scenarios (Table 2), two code for uncharacterized proteins. The remaining eight genes code for a probable WRKY transcription factor 50, a WAT1-related protein, a G-type lectin S-receptor-like serine/threonine-protein kinase, a NIM1-INTERACTING 1-like protein, a SAR Deficient 1 protein, a F-box protein, a VQ motif-containing protein, and a lysosomal Pro-X carboxypeptidase-like. For all of them, most of the decrease in the level of expression at 45 DAT was a result of the age effect, with a minor contribution of the salt stress (data not shown).

5. Conclusion

Two distinct phenotypes were seen when applying salinity stress in *G. sepium* by means of a substrate salinization protocol, depending on the amount of NaCl used:

a) Plants grown under salinity stress up to 25 dS/m for 45 days did not show any visual symptoms of stress on the aerial parts, such as leaf wilt, yellowing, burning or falling; although, they experienced a reduction in shoot and root biomass yield; and

b) Plants grown on a substrate with ≥30 dS/m lost all their leaves in the first week after the stress onset; however, about three weeks after the beginning of the stress they started to develop new leaves from the lateral meristems that continue throughout the rest of the experiment.

The transcriptome and metabolome data sets were analyzed under three distinct scenarios: age effect, short-time stress, and long-time stress; and the integration of these two omics profiles pointed a central role of the phenylpropanoid biosynthesis pathway in the short-term response of *Gliricidia sepium* to salinity stress, but not in the long-term response.

The *de novo* transcriptomics analysis led to the identification of 5,672 differentially expressed transcripts (up and down-regulated), but only 12 of them were differentially expressed in both the short- and long-term stress scenarios. Among them, two code for proteins that might play a role in gliricidia response at both the short- and long-term salt stress.

6. References

Acosta-Motos, J.R., Ortuño, M.F., Bernal-Vicente, A., Diaz-Vivancos, P., Sanchez-Blanco, M.J., Hernandez, J.A. (2017). Plant Responses to Salt Stress: Adaptive Mechanisms. *Agronomy*, *7*, 18. https://doi.org/10.3390/agronomy7010018

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics (Oxford, England)*, *31*(2), 166–169. https://doi.org/10.1093/bioinformatics/btu638

Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Ashapkin, V. V., Kutueva, L. I., Aleksandrushkina, N. I., & Vanyushin, B. F. (2020). Epigenetic Mechanisms of Plant Adaptation to Biotic and Abiotic Stresses. *International journal of molecular sciences*, *21*(20), 7457. https://doi.org/10.3390/ijms21207457

Ashraf, M., Shahzad, S. H., Imtiaz, M., Rizwan, M. S. (2018). Salinity effects on nitrogen metabolism in plants – focusing on the activities of nitrogen metabolizing enzymes: A review, Journal of Plant Nutrition, 41:8, 1065-1081, DOI: 10.1080/01904167.2018.1431670

Askenazi, M., Driggers, E. M., Holtzman, D. A., Norman, T. C., Iverson, S., Zimmer, D. P., Boers, M. E., Blomquist, P. R., Martinez, E. J., Monreal, A. W., Feibelman, T. P., Mayorga, M. E., Maxon, M. E., Sykes, K., Tobin, J. V., Cordero, E., Salama, S. R., Trueheart, J., Royer, J. C., & Madden, K. T. (2003). Integrating transcriptional and metabolite profiles to direct the engineering of lovastatin-producing fungal strains. *Nature biotechnology*, *21*(2), 150–156. https://doi.org/10.1038/nbt781

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Brink, B. G., Seidel, A., Kleinbölting, N., Nattkemper, T. W., & Albaum, S. P. (2016). Omics Fusion - A Platform for Integrative Analysis of Omics Data. *Journal of integrative bioinformatics*, *13*(4), 296. https://doi.org/10.2390/biecoll-jib-2016-296

Cavill, R., Jennen, D., Kleinjans, J., & Briedé, J. J. (2016). Transcriptomic and metabolomic data integration. *Briefings in bioinformatics*, *17*(5), 891–901. https://doi.org/10.1093/bib/bbv090

Chen, G., Hu, J., Dong, L., Zeng, D., Guo, L., Zhang, G., Zhu, L., & Qian, Q. (2019). The Tolerance of Salinity in Rice Requires the Presence of a Functional Copy of FLN2. *Biomolecules*, *10*(1), 17. https://doi.org/10.3390/biom10010017

Chong, J., Wishart, D. S., & Xia, J. (2019). Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis. *Current protocols in bioinformatics*, *68*(1), e86. https://doi.org/10.1002/cpbi.86

Chong, J., & Xia, J. (2020). Using MetaboAnalyst 4.0 for Metabolomics Data Analysis, Interpretation, and Integration with Other Omics Data. *Methods in molecular biology (Clifton, N.J.)*, *2104*, 337–360. https://doi.org/10.1007/978-1-0716-0239-3_17

Diouf, A. , Ndiaye, M. , Fall-Ndiaye, M. and Diop, T. (2017) Maize Crop N Uptake from Organic Material of *Gliricidia sepium* Coinoculated with *Rhizobium* and Arbuscular

Mycorrhizal Fungus in Sub-Saharian Africa Sandy Soil. *American Journal of Plant Sciences*, 8, 428-440. doi: 10.4236/ajps.2017.83029

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Duarte, B.; Caçador, I. Iberian Halophytes as Agroecological Solutions for Degraded Lands and Biosaline Agriculture. *Sustainability,* v.13, p.1005, 2021. https://doi.org/10.3390/su13021005

FAO – Food and Agriculture Organization. (2011). FAO in the 21st century - ensuring food security in a changing world. www.fao.org/3/i2307e/i2307e.pdf

Feduraev, P., Skrypnik, L., Riabova, A., Pungin, A., Tokupova, E., Maslennikov, P., Chupakhina, G. (2020). Phenylalanine and Tyrosine as Exogenous Precursors of Wheat (*Triticum aestivum* L.) Secondary Metabolism through PAL-Associated Pathways. *Plants (Basel, Switzerland)*, *9*(4), 476. https://doi.org/10.3390/plants9040476

Flowers, T. J., & Colmer, T. D. (2008). Salinity tolerance in halophytes. *The New phytologist*, *179*(4), 945–963. https://doi.org/10.1111/j.1469-8137.2008.02531.x

Flowers, T. J., Hajibagheri, M. A., Clipson, N. J.W. (1986). Halophytes. The Quarterly Review of Biology, v. 61, n. 3, p. 313-337. https://doi.org/10.1086/415032

Götz, S., García-Gómez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., Robles, M., Talón, M., Dopazo, J., & Conesa, A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic acids research*, *36*(10), 3420–3435. https://doi.org/10.1093/nar/gkn176

Gowda, H., Ivanisevic, J., Johnson, C. H., Kurczy, M. E., Benton, H. P., Rinehart, D., Nguyen, T., Ray, J., Kuehl, J., Arevalo, B., Westenskow, P. D., Wang, J., Arkin, A. P., Deutschbauer, A. M., Patti, G. J., & Siuzdak, G. (2014). Interactive XCMS Online: simplifying advanced metabolomic data processing and subsequent statistical analyses. *Analytical chemistry*, *86*(14), 6931–6939. https://doi.org/10.1021/ac500734c

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology*, *29*(7), 644–652. https://doi.org/10.1038/nbt.1883

Hoefgen, R., & Nikiforova, V. J. (2008). Metabolomics integrated with transcriptomics: assessing systems response to sulfur-deficiency stress. *Physiologia plantarum*, *132*(2), 190–198. https://doi.org/10.1111/j.1399-3054.2007.01012.x

Jamil, I. N., Remali, J., Azizan, K. A., Nor Muhammad, N. A., Arita, M., Goh, H. H., & Aizat, W. M. (2020). Systematic Multi-Omics Integration (MOI) Approach in Plant Systems Biology. *Frontiers in plant science*, *11*, 944. https://doi.org/10.3389/fpls.2020.00944

Jeena, G. S., Kumar, S., Shukla, R. K. (2019). Structure, evolution and diverse physiological roles of SWEET sugar transporters in plants. Plant Molecular Biology. 100. https://doi.org/10.1007/s11103-019-00872-4

Kumar, S., Kumar, R., Pal, A., Chopra, D. S. (2019). Chapter 16 – Enzymes, Editor(s): Elhadi M. Yahia, Postharvest Physiology and Biochemistry of Fruits and Vegetables, Woodhead

Publishing, Pages 335-358, ISBN 9780128132784, https://doi.org/10.1016/B978-0-12-813278-4.00016-6

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Li, S., Park, Y., Duraisingham, S., Strobel, F. H., Khan, N., Soltow, Q. A., Jones, D. P., & Pulendran, B. (2013). Predicting network activity from high throughput metabolomics. *PLoS computational biology*, *9*(7), e1003123. https://doi.org/10.1371/journal.pcbi.1003123

Liu, J. G., Han, X., Yang, T., Cui, W. H., Wu, A. M., Wange, B. C., Liu, L. J. (2019). Genome-wide transcriptional adaptation to salt stress in *Populus*. *BMC Plant Biol* 19, 367. https://doi.org/10.1186/s12870-019-1952-2

OmicsBox – Bioinformatics Made Easy, BioBam Bioinformatics, March 3, 2019, https://www.biobam.com/omicsbox

Perkins, A., Nelson, K. J., Parsonage, D., Poole, L. B., & Karplus, P. A. (2015). Peroxiredoxins: guardians against oxidative stress and modulators of peroxide signaling. *Trends in biochemical sciences*, *40*(8), 435–445. https://doi.org/10.1016/j.tibs.2015.05.001

Phang, J. M., Liu, W., & Zabirnyk, O. (2010). Proline metabolism and microenvironmental stress. *Annual review of nutrition*, *30*, 441–463. https://doi.org/10.1146/annurev.nutr.012809.104638

Rahman, M., Das, A., Saha, S., Uddin, M., & Rahman, M. (2019). Morpho-physiological Response of Gliricidia sepium to Seawater-induced Salt Stress. *The Agriculturists*, *17*(1-2), 66-75. https://doi.org/10.3329/agric.v17i1-2.44697

Rai, A., Rai, M., Kamochi, H., Mori, T., Nakabayashi, R., Nakamura, M., Suzuki, H., Saito, K., & Yamazaki, M. (2020). Multiomics-based characterization of specialized metabolites biosynthesis in Cornus Officinalis. *DNA research : an international journal for rapid publication of reports on genes and genomes*, *27*(2), dsaa009. https://doi.org/10.1093/dnares/dsaa009

Rhee S. G. (2016). Overview on Peroxiredoxin. *Molecules and cells*, *39*(1), 1–5. https://doi.org/10.14348/molcells.2016.2368

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)*, *26*(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

Rodrigues-Neto, J. C., Correia, M. V., Souto, A. L., Ribeiro, J., Vieira, L. R., Souza Jr., M. T., Rodrigues, C. M., & Abdelnur, P. V. (2018). Metabolic fingerprinting analysis of oil palm reveals a set of differentially expressed metabolites in fatal yellowing symptomatic and non-symptomatic plants. *Metabolomics : Official journal of the Metabolomic Society*, *14*(10), 142. https://doi.org/10.1007/s11306-018-1436-7

Santos Sánchez, N. F., Salas-Coronado, R., Hernandez-Carlos, B., Villanueva, C. (2019). Shikimic Acid Pathway in Biosynthesis of Phenolic Compounds. https://doi.org/10.5772/intechopen.83815

Sellami, S., Le Hir, R., Thorpe, M. R., Vilaine, F., Wolff, N., Brini, F., & Dinant, S. (2019). Salinity Effects on Sugar Homeostasis and Vascular Anatomy in the Stem of the *Arabidopsis Thaliana* Inflorescence. *International journal of molecular sciences*, *20*(13), 3167. https://doi.org/10.3390/ijms20133167

Shahid S. A., Zaman M., Heng L. (2018) Soil Salinity: Historical Perspectives and a World Overview of the Problem. In: Guideline for Salinity Assessment, Mitigation and Adaptation Using Nuclear and Related Techniques. Springer, Cham. https://doi.org/10.1007/978-3-319-96190-3_2

Sharma, A., Shahzad, B., Rehman, A., Bhardwaj, R., Landi, M., & Zheng, B. (2019). Response of Phenylpropanoid Pathway and the Role of Polyphenols in Plants under Abiotic Stress. *Molecules (Basel, Switzerland)*, *24*(13), 2452. https://doi.org/10.3390/molecules24132452

Silva, V. N. B. (2019). Phenotypic characterization of the tolerance of *Portulaca oleracea* L. and *Gliricidia sepium* (Jacq.) Steud. to salinity. 326 p. Tese (Doutorado em Biotecnologia Vegetal) – Universidade Federal de Lavras, Lavras, 2019. http://repositorio.ufla.br/jspui/handle/1/38245

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(43), 15545–15550. https://doi.org/10.1073/pnas.0506580102

Tautenhahn, R., Patti, G. J., Rinehart, D., & Siuzdak, G. (2012). XCMS Online: a web-based platform to process untargeted metabolomic data. *Analytical chemistry*, *84*(11), 5035–5039. https://doi.org/10.1021/ac300698c

Tovar-Méndez, A., Matamoros, M. A., Bustos-Sanmamed, P., Dietz, K. J., Cejudo, F. J., Rouhier, N., Sato, S., Tabata, S., Becana, M. (2011). Peroxiredoxins and NADPH-dependent thioredoxin systems in the model legume Lotus japonicus. *Plant physiology*, *156*(3), 1535–1547. https://doi.org/10.1104/pp.111.177196

Urbanczyk-Wochniak, E., Luedemann, A., Kopka, J., Selbig, J., Roessner-Tunali, U., Willmitzer, L., & Fernie, A. R. (2003). Parallel analysis of transcript and metabolic profiles: a new approach in systems biology. *EMBO reports*, *4*(10), 989–993. https://doi.org/10.1038/sj.embor.embor944

van Zelm, E., Zhang, Y., & Testerink, C. (2020). Salt Tolerance Mechanisms of Plants. *Annual review of plant biology*, *71*, 403–433. https://doi.org/10.1146/annurev-arplant-050718-100005

Vargas, L. H. G., Neto, J. C. R., de Aquino Ribeiro, J. A., Ricci-Silva, M. E.; Souza Jr., M. T.; Rodrigues, C. M., Abdelnur, P. V. (2016). Metabolomics analysis of oil palm (*Elaeis guineensis*) leaf: evaluation of sample preparation steps using UHPLC–MS/MS. *Metabolomics : Official journal of the Metabolomic Society, 12,* 153. https://doi.org/10.1007/s11306-016-1100-z

Ventura, Y., Eshel, A., Pasternak, D., & Sagi, M. (2015). The development of halophyte-based agriculture: past and present. *Annals of botany*, *115*(3), 529–540. https://doi.org/10.1093/aob/mcu173

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews. Genetics*, *10*(1), 57–63. https://doi.org/10.1038/nrg2484

Williams, L. J. and Abdi, H. (2010). "Fisher's least significance difference (LSD) test," in Encyclopedia of Research Design. Thousand Oaks, pp. 491-494. http://dx.doi.org/10.4135/9781412961288.n154

Yan, J., Qian, L., Zhu, W., Qiu, J., Lu, Q., Wang, X., Wu, Q., Ruan, S., & Huang, Y. (2020). Integrated analysis of the transcriptome and metabolome of purple and green leaves of Tetrastigma hemsleyanum reveals gene expression patterns involved in anthocyanin biosynthesis. *PloS one*, *15*(3), e0230154. https://doi.org/10.1371/journal.pone.0230154

Deng, Y. and Lu, S. (2017). Biosynthesis and Regulation of Phenylpropanoids in Plants, Critical Reviews in Plant Sciences, 36:4, 257-290, DOI: 10.1080/07352689.2017.1402852

Zampieri, M., & Sauer, U. (2017). Metabolomics-driven understanding of genotype-phenotype relations in model organisms. Current Opinion in Systems Biology, 6, 28-36. https://doi.org/10.1016/j.coisb.2017.08.007

Zhang, G., Guo, G., Hu, X., Zhang, Y., Li, Q., Li, R., Zhuang, R., Lu, Z., He, Z., Fang, X., Chen, L., Tian, W., Tao, Y., Kristiansen, K., Zhang, X., Li, S., Yang, H., Wang, J., & Wang, J. (2010). Deep RNA sequencing at single base-pair resolution reveals high complexity of the rice transcriptome. *Genome research*, *20*(5), 646–654. https://doi.org/10.1101/gr.100677.109

Zhao, C, Zhang, H, Song, C, Zhu, J-K and Shabala, S (2020). Mechanisms of plant responses and adaptation to soil salinity. *The Innovation*, vol. 1, no. 1 , pp. 1-41 , doi: 10.1016/j.xinn.2020.100017.

Zhu, H., Li, C., & Gao, C. (2020). Applications of CRISPR-Cas in agriculture and plant biotechnology. *Nature reviews. Molecular cell biology*, *21*(11), 661–677. https://doi.org/10.1038/s41580-020-00288-9

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

The authors acknowledge funding to T.L.C.S., V.N.B.S., I.O.B., and J.C.R.N. by the Coordination for the Improvement of Higher Education Personnel (CAPES), via the Graduate Program in Plant Biotechnology at the Federal University of Lavras (UFLA) and the Graduate Program in Chemistry at the Federal University of Goiás (UFG).

Funding

The authors disclose receipt of the following financial support for the research, authorship, and/or publication of this article: the grant (01.13.0315.00 - DendePalm Project) for this study was awarded by the Brazilian Ministry of Science, Technology, and Innovation (MCTI) via the Brazilian Research and Innovation Agency (FINEP). The authors confirm that the funder had no influence over the study design, the content of article, or selection of this journal.

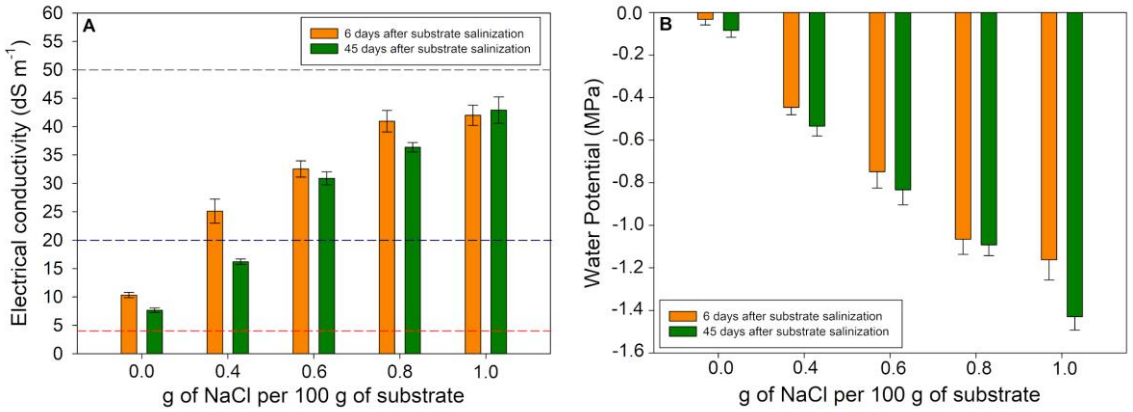


Figure 1. Electrical conductivity – EC (A) and water potential (B) of the substrate used for growing gliricidia plants to which different levels of NaCl have been added; at six and 45 days after substrate salinization. The values represent the average of five replicates, and the bars represent the standard error of the mean. red dashed line: EC = 4 dS/m (above that level is considered saline soil); blue dashed line: EC = 20 dS/m (plant completing its life cycle above this EC is considered as halophyte); and gray dashed line: EC = 50 dS/m (seawater salinity).

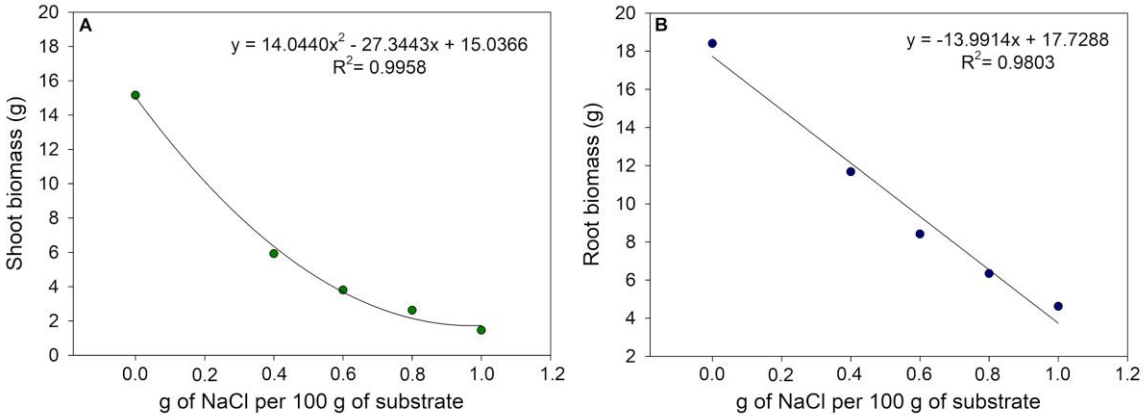


Figure 2. Biomass accumulation in shoot (A) and root (B) of gliricidia plants grown under different concentrations of NaCl for 45 days. The values used represent an average of five replicates.

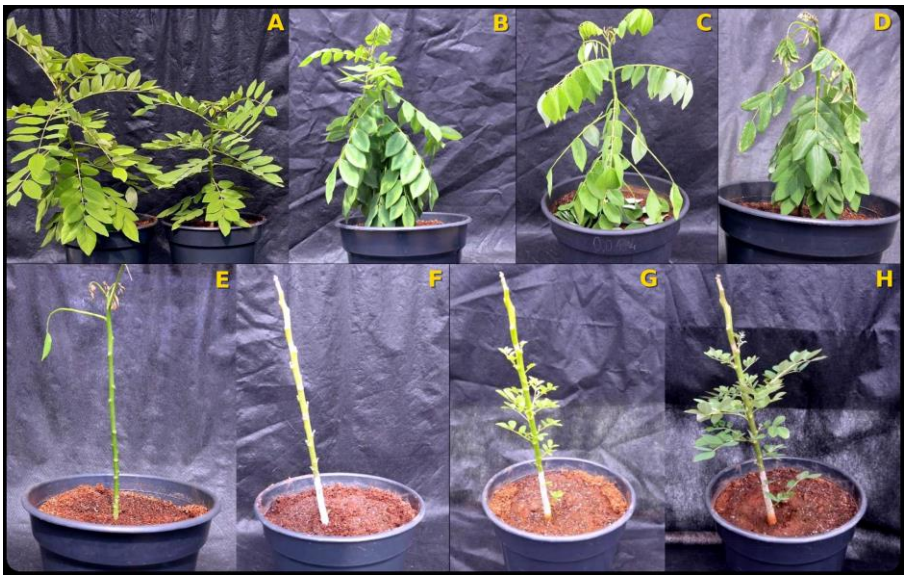


Figure 3. Symptoms of salt stress in gliricidia plants shoots. A – control plants, and B-H – stressed plants at 0.8 g of NaCl per 100 g of substrate. In the third days at treatment (DAT), the leaves started to show a strong wilt symptom (B), and in the fourth DAT they started to fall (C). Some plants presented a kind of burning symptoms in some leaves (D). At the end of the first week of stress, the stressed plants had lost almost all leaves (E), and about three weeks after the beginning of the stress, new leaves started to emerge (F), and kept growing continuously throughout the rest of the experiment (G and H).

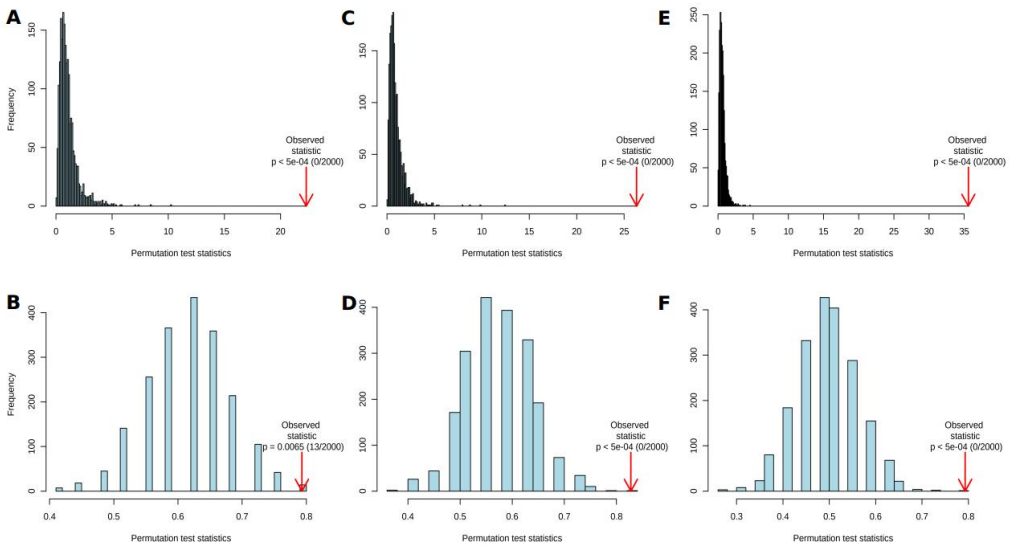


Figure 4. PLS-DA permutation validation evaluated by group separation distance (A, C, E) and by prediction accuracy (B, D, F), applying permutation number = 2,000. Polar-positive (A, B), polar-negative (C, D), and lipidic-positive (E, F) fractions.

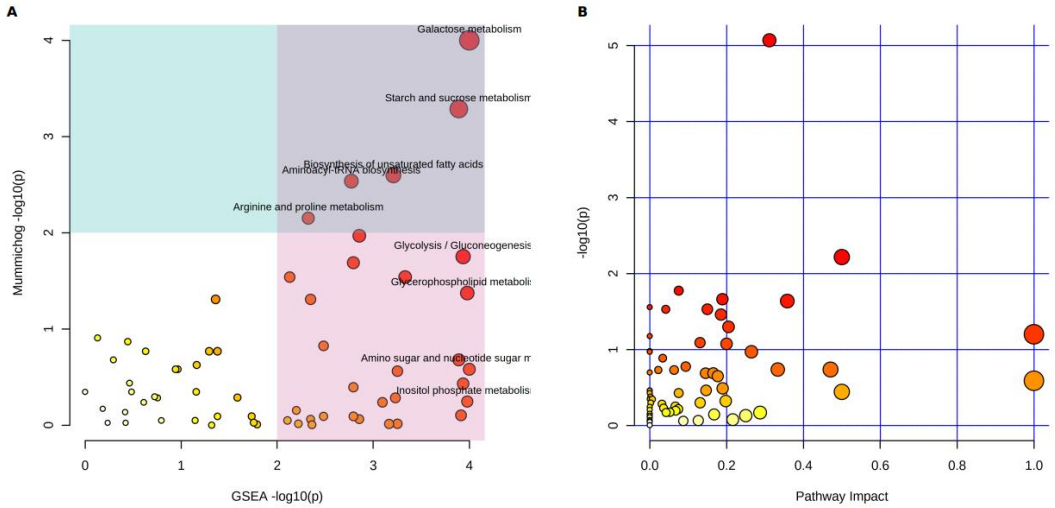


Figure 5. Summary of Pathway Analysis using the MS Peaks to Pathway and the Pathway Topology Analysis modules of MetaboAnalyst 4.0. The Integrated MS Peaks to Paths plot (A) summarizes the results of the Fisher's method for combining mummichog (y-axis) and Gene Set Enrichment Analysis - GSEA (x-axis) p-values. The metabolome view (B) resulted from the analysis in the Pathway Topology Analysis module using the Hypergeometric test, the relative betweenness centrality node importance measure, and the latest KEGG version of the *A. thaliana* pathway library.

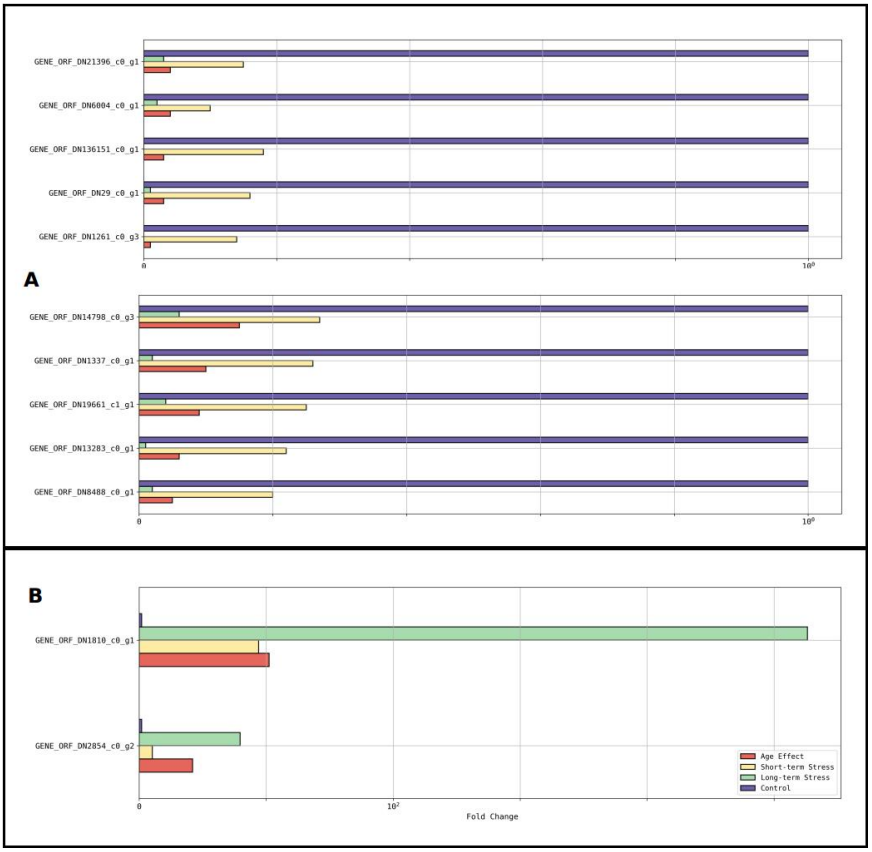


Figure 6. Expression profiles in percentage (x-axis) of the 12 differentially expressed transcripts, being two that up-regulated twice (A) - in short-term stress (control and the stress plants at two days under salinity stress – DAT) and long-term stress (tressed plants at two and 45 DAT) as well – and ten that down-regulated twice (B) resulted from submission of gliricidia plants to salinity stress, and in comparisson with the control treatment (FC = 1). AE – age effect (control plants at two and 45 DAT).

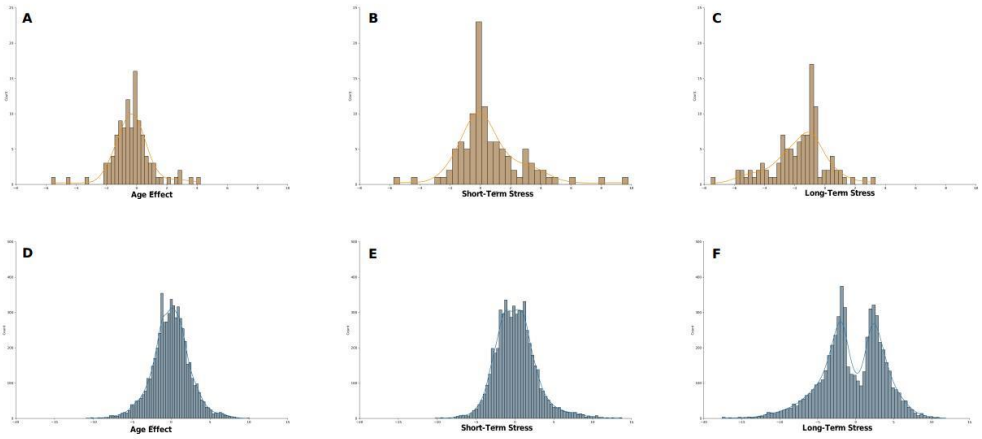


Figure 7. Histograms – the distribution in terms of number (y-axis) and Log_2 (Fold Change) (x-axis) – of 107 differentially expressed metabolites (A, B, C) and 5,672 differentially expressed transcripts (D, E, F) in three scenarios: age effect - AE (control plants at two and 45 days under salinity stress – DAT) – A and D; short-term stress – STS (control and the stress plants at 2 DAT) – B and E; and long-term stress – LTS (stressed plants at two and 45 DAT) – C and F.

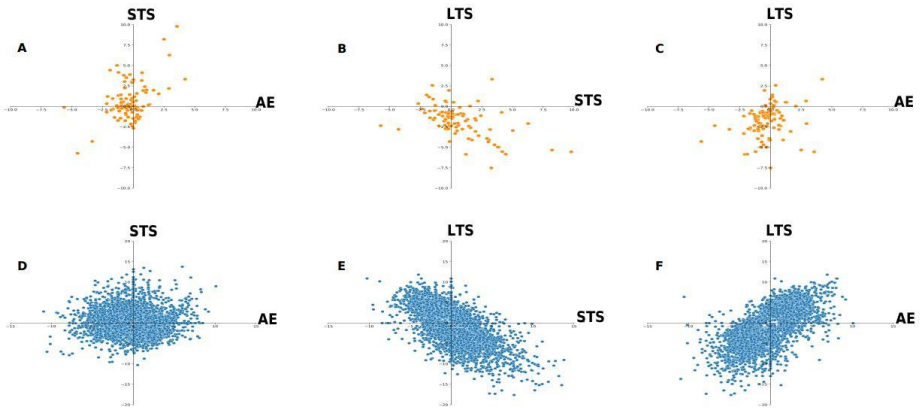


Figure 8. Correlation analysis of the Log_2 (Fold Change) of 107 differentially expressed metabolites (A, B, C) and 5,672 differentially expressed transcripts (D, E, F) by pairwise comparison of three scenarios: age effect - AE (control plants at two and 45 days under salinity stress – DAT); short-term stress – STS (control and the stress plants at 2 DAT); and long-term

stress – LTS (stressed plants at two and 45 DAT). STS x AE (A, D), STS x LTS (B, E), and AE x LTS (C, F).

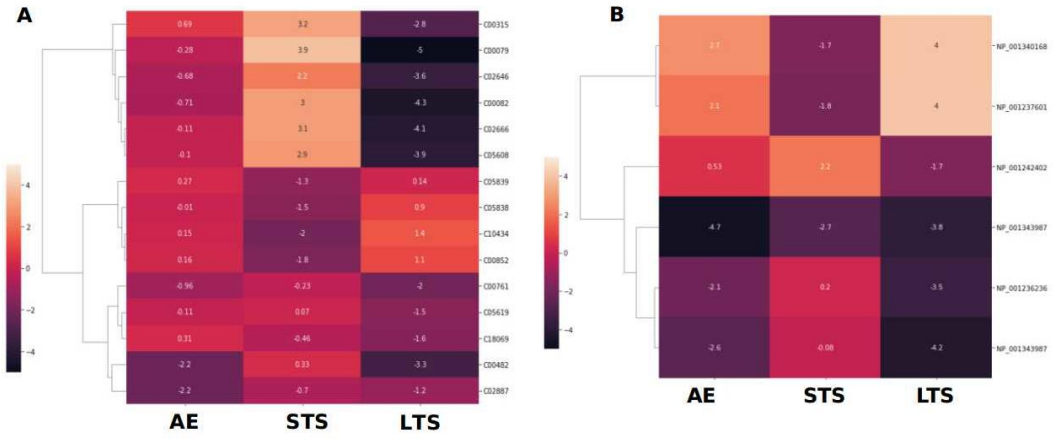


Figure 9. Cluster heat map of 15 metabolites (A) and six transcripts (B) from the phenylpropanoid pathway differentially expressed under salinity stress in the leaves of gliricidia plants. Hierarchical clustering of metabolites and transcripts with altered expression levels in three scenarios: age effect - AE (control plants at two and 45 days under salinity stress – DAT); short-term stress – STS (control and the stress plants at 2 DAT); and long-term stress – LTS (stressed plants at two and 45 DAT). Metabolites identified by the KEGG id, and transcripts by Protein id. Log_2 (Fold change) is presented in the center of each box.

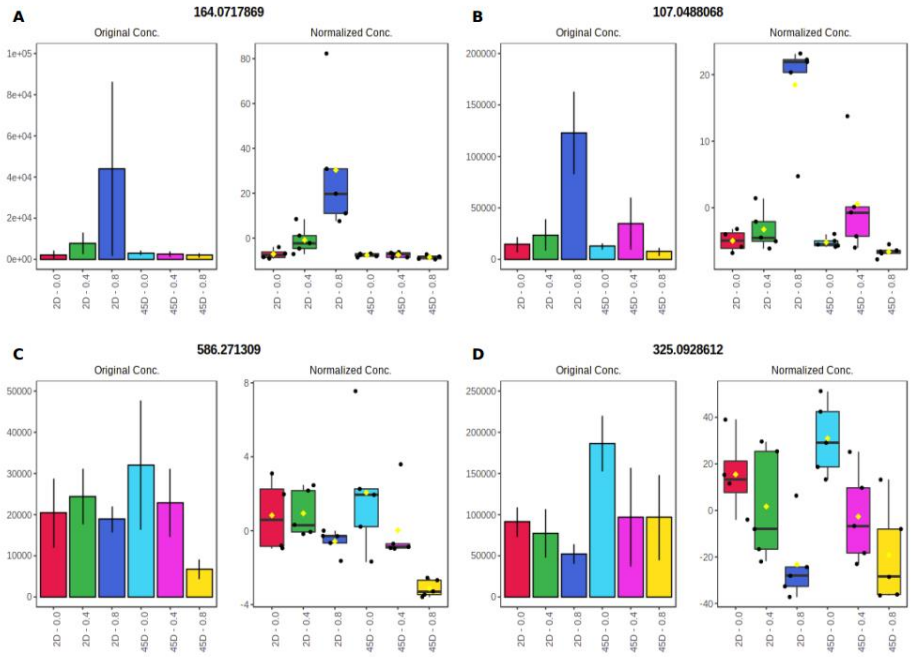
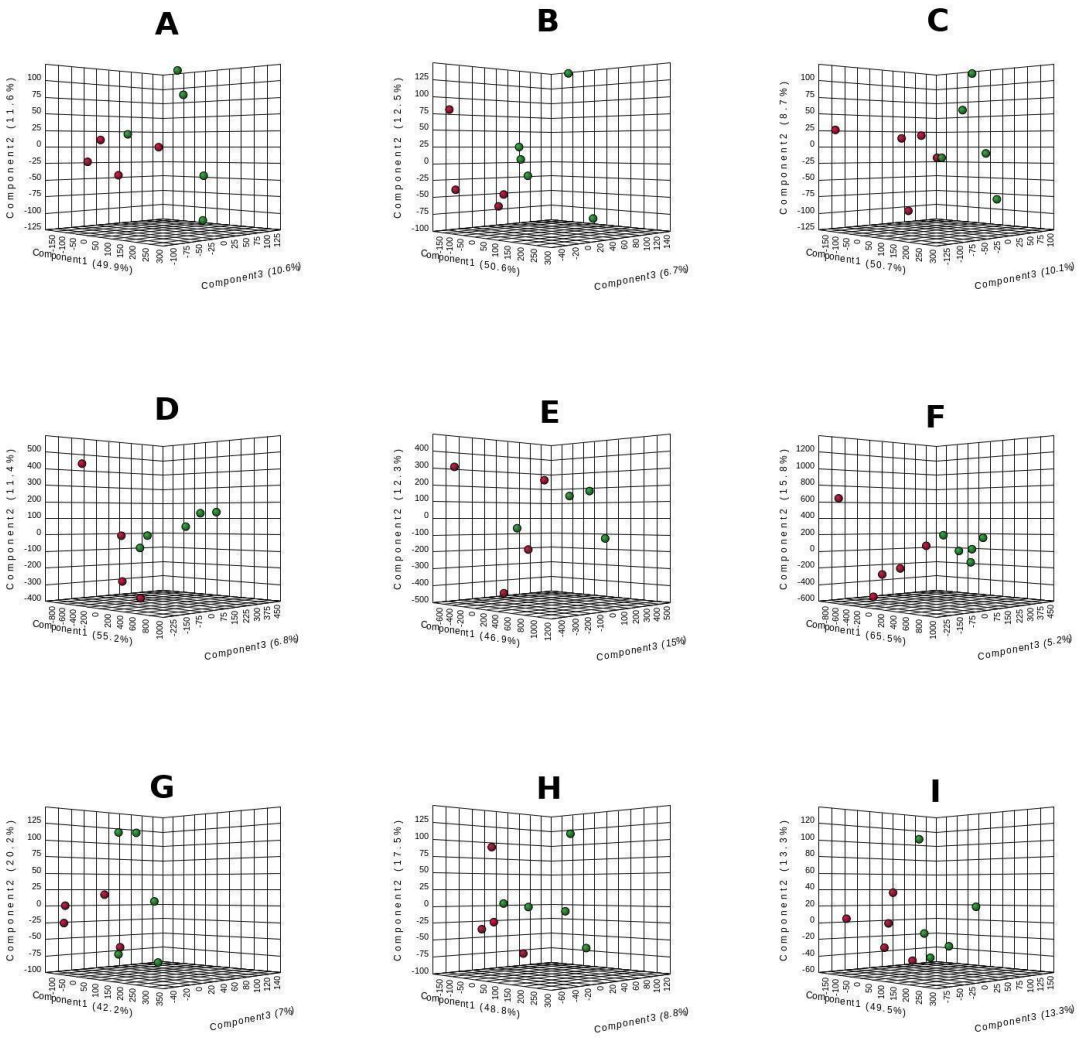


Figure 10. Box plot showing original and normalized concentration of metabolites from the phenylpropanoid pathway differentially expressed under salinity stress in the leaves of gliricidia plants. The expression profiles of L-phenylalanine (A), coniferyl aldehyde (B), N1,N5,N10-tricoumaroyl spermidine (C), and *cis*-beta-D-glucosyl-2-hydroxycinnamate (D). The values represent the average of five replicates, and the bars represent the standard error of the mean. Treatments (x-axis): 2D (2 days under salinity stress) or 45D (45 days under salinity stress), at 0, 04, or 0.8 g of NaCl per 100 g of substrate. Number on the top of the boxes: m.z.

Supplementary Figure 1. PLS-DA 3D score plots. PLS-DA 3D score plots in polar-positive mode (A, B, C), in polar-negative mode (D, E, F), and in lipidic-positive mode (G, H,



I), under three scenarios: age effect (control plants at two and 45 days under salinity stress – DAT) – A, D, and G; short-term stress (control and the stress plants at 2 DAT) – B, E, and H; and long-term stress (stressed plants at 2 and 45 DAT) – C, F, and I. Red dots: control plants at 2 DAT in A, B, C, D, E, and F; and stressed plants at 2 DAT in G, H, and I. Green dots: Control plants at 45 DAT in A, B, and C; stressed plants at 2 DAT in D, E, and F; and stressed plants at 45 DAT in G, H, and I.

List of Tables

showing the KEGG id of the matched compound, matched form, mass difference, name of the compound, Variable Importance in Projection (VIP), correlation, t.score, p.value, FDR (False Discovery Rate), fold change - FC, Log_2 (Fold Change), and profile, in each one of the three scenarios evaluated: age effect - AE (control plants at 2 and 45 days under salinity stress – DAT); short-term stress – STS (control and the stress plants at 2 DAT); and long-term stress – LTS (stressed plants at 2 and 45 DAT).

Supplementary Table 3. Statistics of RNA-Seq data from 12 samples of gliricidia plants submitted to two treatments (0 and 0.8 grams of NaCl per 100 g of the substrate); three replicates per treatment, used to generate the Reference Transcriptome and to perform differential expression analysis.

CAPÍTULO 3

Análise Multi-ômica das respostas de plantas jovens de *Portulaca oleracea* L. a altas doses de NaCl revelam percepções sobre as vias metabólicas e genes que respondem ao estresse salino nesta espécie halófita

A versão apresentada do presente artigo foi submetida a revista "*The Plant Genome*", sendo uma versão preliminar e o conselho editorial do periódico poderá sugerir alterações.

RESUMO

A salinidade do solo está entre os estresses abióticos que mais ameaçam a agricultura. Beldroega (*Portulaca oleracea* L.) é uma espécie de dicotiledônea adaptada ao deserto salino e habitats salinos e é uma planta hiperacumuladora de sal com alto potencial de fitorremediação. É considerada uma espécie modelo adequada para estudar os mecanismos de tolerância das plantas à seca e ao estresse salino. Este estudo aplicou diferentes técnicas ômicas para obter mais informações sobre a tolerância de plantas jovens de beldroegas ao estresse salino muito alto. Um protocolo robusto de estresse salino foi desenvolvido e usado para caracterizar as respostas morfofisiológicas de beldroegas jovens ao estresse salino. Uma análise, em larga escala, abrangente e integrada de metaboloma e transcritoma foi então empregada usando amostras de folhas. O protocolo de estresse salino gerou diferentes níveis de estresse por gradientes de condutividade elétrica e potencial hídrico no extrato de saturação do substrato, e as evidências mostraram que um mecanismo de exclusão de sal opera nas folhas desta espécie halófita. Três vias metabólicas e 20 metabólitos foram identificados por análises ômicas simples, além de milhares de genes expressos diferencialmente. A multi-ômica levou a um conjunto de dados de 51 vias metabólicas que tinham, pelo menos, uma enzima e um metabólito diferencialmente expressos devido ao estresse salino. Os conjuntos de dados gerados são valiosos para estudos futuros com o objetivo de aprofundar nosso conhecimento sobre os mecanismos por trás da alta tolerância de beldroegas jovens ao estresse salino.

Palavras-chave: Beldroega, RNA-Seq, Quimiometria, Espectometria de Massa de Alta Resolução, Estresse Abiótico, Integratômica, Integração Multi-ômica.

ABSTRACT

Soil salinity is among the abiotic stressors that threaten agriculture the most. Purslane (Portulaca oleracea L.) is a dicot species adapted to inland salt desert and saline habitats, and it is a salt hyperaccumulator plant with high phytoremediation potential. It is considered a suitable model species to study the mechanisms of plant tolerance to drought and salt stresses. This study applied different omics technics to gain further insights on young purslane plants' tolerance to very high salinity stress. A robust salinity stress protocol was developed and used to characterize the morphophysiological responses of young purslane plants to salinity stress. A comprehensive, large-scale metabolome and transcriptome single and integrated analysis was then employed using leaf samples. The salinity stress protocol did generate different levels of stress by gradients of electrical conductivity and water potential in the saturation extract of the substrate, and evidence showed that a mechanism of salt exclusion operates on the leaves of this halophyte species. Three metabolic pathways, and 20 metabolites, were identified by single-omics analyses, besides thousands of differentially expressed genes. The multi-omics led to a dataset of 51 metabolic pathways that had at least one enzyme and one metabolite differentially expressed due to salinity stress. The datasets generated are valuable for future studies aiming to deepen our knowledge on the mechanisms behind the high tolerance of young purslane plants to salinity stress.

Keywords: Purslane, RNA-Seq, Chemometrics, High Resolution Mass Spectrometry, Abiotic Stress, Integratomics, Multi-Omics Integration.

Multi-Omics analysis of young *Portulaca oleracea* L. plants' responses to high NaCl doses reveal insights on pathways and genes responsives to salinity stress in this halophyte species

Vivianny Nayse Belo Silva[1]

Thalliton Luiz Carvalho da Silva[1]

Thalita Massaro Malheiros Ferreira[1]

Jorge Candido Rodrigues Neto[2]

André Pereira Leão[4]

José Antônio de Aquino Ribeiro[4]

Patrícia Verardi Abdelnur[2,4]

Leonardo Fonseca Valadares[4]

Carlos Antônio Ferreira de Sousa[3]

Manoel Teixeira Souza Júnior[1,4,*]

[1] – Graduate Program of Plant Biotechnology, Federal University of Lavras, CP 3037, Lavras, MG, Zip Code 37200-000, Brazil

[2] – Institute of Chemistry, Federal University of Goiás, Campus Samambaia, Goiânia, GO, Zip Code 74690-900, Brazil

[3] – Brazilian Agricultural Research Corporation, Embrapa Mid-North, Teresina, PI, Zip Code 64008-780, Brazil

[4] – Brazilian Agricultural Research Corporation, Embrapa Agroenergy, Brasília, DF, Zip Code 70770-901, Brazil

* - Corresponding author

Keywords: Purslane, RNA-Seq, Chemometrics, High Resolution Mass Spectrometry, Abiotic Stress, Integratomics, Multi-Omics Integration.

Abstract

Soil salinity is among the abiotic stressors that threaten agriculture the most. Purslane (*Portulaca oleracea* L.) is a dicot species adapted to inland salt desert and saline habitats, and it is a salt hyperaccumulator plant with high phytoremediation potential. It is considered a suitable model species to study the mechanisms of plant tolerance to drought and salt stresses. This study applied different omics technics to gain further insights on young purslane plants' tolerance to very high salinity stress. A robust salinity stress protocol was developed and used to characterize the morphophysiological responses of young purslane plants to salinity stress. A comprehensive, large-scale metabolome and transcriptome single and integrated analysis was then employed using leaf samples. The salinity stress protocol did generate different levels of stress by gradients of electrical conductivity and water potential in the saturation extract of the substrate, and evidence showed that a mechanism of salt exclusion operates on the leaves of this halophyte species. Three metabolic pathways, and 20 metabolites, were identified by single-omics analyses, besides thousands of differentially expressed genes. The multi-omics led to a dataset of 51 metabolic pathways that had at least one enzyme and one metabolite differentially expressed due to salinity stress. The datasets generated are valuable for future studies aiming to deepen our knowledge on the mechanisms behind the high tolerance of young purslane plants to salinity stress.

1. Introduction

There are many abiotic stressors affecting the plant life cycle, and interfering with growth and productivity (Sunkar et al. 2007). Soil salinity is among the abiotic stressors that threaten agriculture the most, and it is a growing problem in several parts of the world, predominantly in arid and semi-arid regions (Allbed and Kumar, 2013), resulting in a considerable restriction on crop productivity (Mahajan and Tuteja, 2005). Approximately 20% of the agricultural land in the world has saline or sodic soils, and between 25% and 30% of the irrigated land area is affected by salt (Shahid et al., 2018). Soil salinity spreads over 100 countries (Kumari et al., 2015; Zhang et al., 2012), resulting in an annual cost of U$ 27.3 billion due to the loss of agricultural production (Qadir et al., 2014).

Approximately 99% of all plant species are glycophytes or salt-sensitive, including all major crops. Halophytes account for the remaining 1% and can complete their life cycle in an environment where the salt concentration is >200 mM NaCl or >20 dS m^{-1} (Flowers & Colmer, 2008; Shabala and Mackay, 2011). Purslane (*Portulaca oleracea* L.) is the most well-known and studied species of this single genus from the Portulacaceae family. It is a dicot species adapted to inland salt desert and saline habitats - xerohalophyte, and it is a salt hyperaccumulator plant with a high phytoremediation potential (Devi et al., 2016; Santos et al., 2016; Ozturk et al., 2020). It also produces many bioactive allelopathic compounds – such as growth regulators and natural herbicides, which makes it suitable to be used as an allelopathic plant (El-Shora and El-Gawad, 2015; Petropoulos et al., 2016).

Petropoulos et al. (2015), after studying the chemical composition and yield of six genotypes of common purslane, reported that the yield (fresh weight) was affected by genotype, with the highest yield of the tested genotypes being of 33 tons/hectare, and the lowest one being 11.5; with an average of about 22.5 among those genotypes. On the top of this high biomass productivity seen under a controlled agricultural environment, portulaca is know as a plant with

a nutritional quality higher than many other leafy vegetables, as it possesses a large spectrum of pharmacological properties, such as neuroprotective, antimicrobial, antidiabetic, antioxidant, anti-inflammatory, antiulcerogenic, and anticancer activities; and flavonoids, alkaloids, polysaccharides, fatty acids, terpenoids, sterols, proteins vitamins, and minerals (Zhou et al., 2015; Ozturk et al., 2020).

Due to its well-known tolerance to several abiotic stresses, as well as short growth cycle – completes its life cycle in 2 to 4 months, Borsai et al. (2018) presented the idea of using *P. oleracea* as a suitable model to study the mechanisms of plant tolerance to drought and salt stresses. Furthemore, it is a C4 plant that can develop the crassulacean acid metabolism (CAM) when subjected to water stress and short photoperiod, making the idea of using it as a model plant even more interesting (D'andrea et al., 2014; Koch and Kennedy, 1980).

Studies conducted by different researchers have shown that there are differences between the *P. oleracea* genotypes in terms of defense mechanisms against high salinity (Ozturk et al., 2020). Some studies have suggested an initial view of some of these mechanisms triggered in purslane plants under saline stress, such as the ability to avoid chlorine ion toxicity, the activation of the ethylene signaling pathway, the carriers' ability to discriminate cations, and increased antioxidant activity, the synthesis of osmoprotectant proline, and the synthesis of several other metabolites involved with many biochemical pathways (Sdouga et al., 2019; Xing et al., 2019; Xing et al., 2020; Zaman et al., 2020).

Metabolomics and transcriptomics are techniques intensively used to assist in the systemic understanding of plant responses to salinity, envisioning the use of genetic engineering in sensitive plants of economic importance. The transcriptome and metabolome are the complete set of RNA and metabolites (primary and secondary), respectively, produced under specific circumstances or in a cell, tissue, organ, or an entire organism, in a given moment of its development (Villate et al., 2021). In recent years, due to the technological advances and

cost reduction achieved with the RNA-seq technique, as well as Mass spectrometry and NMR spectroscopy, we have witnessed an explosion in the amount of transcriptome and metabolome data generated and made public (Lowe et al., 2017; Villate et al., 2021).

Salinity stress is known to alter the metabolic and transcriptomic profile of several plant species (Tada et al., 2019; Arif et al., 2020; Wang et al., 2021). There are many reports on the tolerance of purslane plants to salt stress, but not many reporting the use of the so-called multiomics integration (MOI) strategy (Cavill et al., 2016; Jamil et al., 2020). Xing et al. (2020), who performed an integrative analysis of the transcriptome and metabolome profiles of purslane, showed that resisting saline stress it reduces the expression levels of chloride channel proteins to avoid the toxicity of chloride, activates the signaling pathway of ethylene, and accumulates pyrophosphate and unsaturated fatty acid to regulate energy supply and minimize oxidative effects on cell structure.

A first step in pursuing a multiomics approach to gain additional insights on purslane's response to salinity stress is to set up and validate a bioassay to evaluate its morphophysiological changes due to this stress. This study reports on a robust salinity stress protocol and the characterization of the morphophysiological responses of young purslane plants to salinity stress using such protocol. Furthermore, it reports a comprehensive, large-scale metabolome and transcriptome single and integrated analyses to gain further insights on the salt-tolerance shown by this halophyte species.

2. Materials and Methods

2.1. Plant material, growth conditions, experimental design and saline stress

Seeds of the B1 accession of purslane, from the Purslane Collection at Embrapa Agroenergia, were disinfected by soaking in 2% sodium hypochlorite and Tween® 20 for 5 minutes, under slow agitation, then washed with sterile water and dried on sterile filter paper. After seeded on a culture medium (MS 1/2 strength, Phytagel 0.2%, and pH 5.8) (Murashige and Skoog, 1962), it was kept for germination in a growth chamber Conviron mod. Adaptis 1000TC (Controlled Environments Ltd, Winnipeg, Canada) at 150 µmol m^2 s^{-1} of light and 30°C. After 13 days, plantlets were individually transferred to 200 ml plastic cups containing 100g of sterilized substrate - clay soil, vermiculite, and a commercial substrate (Bioplant®), 2:1:1 (v:v:v) ratio – and transferred to another Conviron® growth chamber mod. PGW40 at 25±2°C, 500±20 µmol m^{-2} s^{-1} of light, 65±5% air relative humidity, and photoperiod of 16/8h (light/dark), and kept there until the end of the experiments. The plants were allowed to acclimatize for three days before starting the experiment.

Young purslane plants (13 days after sowing, plus three days of acclimatization in the growth chamber) were submitted to and kept under stress for five days. The treatments used consisted of control (no added salt) and four salinity levels (0.5, 1.0, 1.5, and 2.0 g of NaCl per 100 g of substrate), with ten replicates per treatment in a completely randomized design. NaCl was dissolved in deionized water to salinize the substrate. The amount of deionized water used corresponded to the difference between the amount of water previously present in the substrate and the amount of water necessary for the substrate to reach field capacity. Applying the right amount of water – up to the substrate field capacity – was a mean of ensuring no leakage of the solution out of the pot and no loss of Na^+ or Cl^-. The water lost by evapotranspiration was replaced with deionized water in a daily basis, and the electric conductivity and water potential in the substrate solution measured daily, for all replicates.

2.2. Phenomic data

2.2.1. Gas exchange measurements

The parameters of leaf gas exchange [net CO_2 assimilation rate (*A*), transpiration rate (*E*), stomatal conductance to water vapor (*gs*), and intercellular CO_2 concentration (*Ci*)] were measured by a portable infrared gas analyzer LI-COR Mod. 6400XT (LI-COR, Lincoln, NE, USA) equipped with a measuring chamber (2x3 cm) with artificial light system LI-COR Mod. 6400-02B. The extracted data was provided by the OPEN software version 6.3. Measurements were performed daily, between 9 am to 11 am, on the first fully expanded leaf. The block temperature was 25°C, PAR was 1500 μmol m^2 s^{-1}, the relative humidity of the air inside the measuring chamber was kept between 50 and 60%, the airflow index was 400 μmol s^{-1}, and the CO_2 concentration was 400 ppm in the reference cell, using the CO_2 mixer model 6400-01.

2.2.2. Chlorophyll fluorescence measurements

The parameters evaluated using the chlorophyll fluorescence technique (Saturation Pulse Method) via a Walz image fluorimeter model IMAGING-PAM Maxi version (Heinz Walz GmbH, Effeltrich, Germany), controlled by the ImaginWin software version 2.40b, were: Fm, Fo, Y(II), Fv/Fm, Y(NPQ), and Y(NO). The IMAG-MAX/L LED lighting head and the CCD camera were mounted on a 15 mm diameter metal bar on the optional support at a standard distance of 18.5 cm for all plants. We used the following settings: measurement light = 1, saturation pulse = 10 (2800 μmol m^2 s^{-1}), pulse amplification factor = 1, damping = 2, amplification factor in red = 25, Fm factor = 1.055, Factor = 0.999, and actinic light = 280 μmol m^{-2} s^{-1}. Plants were kept in the dark for 30 min before measurement, performed in the dark, and on the same leaf used for gas exchange measurements. The induction curve approach was used with a 40 s delay from the initial saturation pulse until the start of the actinic illumination, and, from then on, a saturation pulse was emitted every 20 s until completing 315 s. After

measuring the initial parameters, the software calculated all derivatives parameters. For this, an area of interest (AOI) that did not include the midrib was demarcated on the leaf surface.

2.2.3. Pigment content measurements

We used a hyperspectral camera Resonon Mod Pika XC (Resonon, Bozeman, Massachusetts, USA), controlled by software Spectronon version 2.1, to obtain the spectral signature of the samples and to estimate the levels of pigments. The calculated parameters were: chlorophyll index [CI = (R_{660}-R_{930}) x R_{930}] (Gitelson *et al.*, 2005), photochemical reflectance index [PRI = (R_{531}- R_{570}) / (R_{531} + R_{570})] (Gamon *et al.*, 1992), carotenoid index [CRI = (R_{510}) - (R_{550}) x R_{800}] (Gitelson *et al.*, 2002). The system consisted of a hyperspectral camera, a linear translation phase, and a fixed lighting system in the assembly tower. Hyperspectral images generated by maintaining the same distance for all plants and the configurations applied according to the manufacturer's recommendations. Five regions of interest for the average reflectance spectrum were marked on each plant.

2.3. Scanning electron microscopy analysis

We collected samples of purslane leaves and immediately immersed them in liquid nitrogen and then stored them at -80°C until lyophilization in a freeze dryer LIOTOP® model K120 for 48 hours. Lyophilized samples were stored in a desiccator until use. Leaf samples were coated with a gold layer, 12.3 mm thick, using a Quorum Technologies® model Q 150T ES with the QT GOLD program. High-resolution images of the regions of interest were obtained employing a scanning electron microscope (SEM) with detectors of energy dispersive spectroscopy (EDS). The composition of the sample was identified, qualitatively, in specific points of the image.

2.4. Metabolomics analysis

Leaves from all treatments – five replicates per treatment – were collected at one and four days after treatment (DAT), immediately immersed in liquid nitrogen, and then stored at -80 °C until extraction of metabolites.

2.4.1. Chemicals and metabolites extraction

The solvents methanol grade UHPLC, acetonitrile grade LC-MS, formic acid grade LC-MS and sodium hydroxide ACS grade LC-MS were from Sigma-Aldrich (St. Louis, MO, USA); and the water treated in a Milli-Q system (Millipore, Bedford, MA, USA).

Based on previous works, the analytical protocol employed was optimized for a fast and efficient metabolite extraction method. Aliquots of 50 mg of grounded sample were transferred to 2 mL microtubes, then 1 ml of a solvent mixture (1:3 (v:v) methanol: methyl tert-butyl ether) was added, and left for homogenization at 4 °C on an orbital shaker for 10 min, followed by an ultrasound treatment in an ice bath for another 10 min. Next, 500 µL of a 1:3 (v:v) methanol: water mixture (1:3) was added to each microtube before centrifugation (12,000 rpm, 4 °C for 5 min) in order to promote a phase separation. After centrifugation, three phases were generated: an upper nonpolar, a lower polar, and a protein pellet. The apolar and polar fractions were transferred separately to 1.5 mL microtubes and vacuum dried in a Speed vac overnight in room temperature (Centrivap, Labconco, Kansa, MO, USA).

2.4.2. UHPLC-MS and UHPLC-MS/MS

After resuspending the dry polar fraction by adding 500 µL of 1:3 (v:v) methanol: water mixture, samples were transferred to vials and analyzed by UHPLC-MS/MS. We used a UHPLC chromatographic system (Nexera X2, Shimadzu Corporation, Japan) equipped with an Acquity UPLC HSS T3 (1.8 µm, 2.1 x 150 mm) reverse phase column (Waters Technologies, Milford, MA), maintained at 35 °C. A polar mobile phase was used, where solvent A was 0.1% (v/v) formic acid in water and solvent B was 0.1% (v/v) formic acid in acetonitrile/methanol (70/30, v/v). The gradient elution used, with a flow rate of 0.4 mL min^{-1}, was the following:

isocratic from 0 to 1 min (0% B), linear gradient from 1 to 3 min (5% B), from 3 to 10 min (50% B), and 10 to 13 min (100% B), isocratic from 13 to 15 min (100% B), followed by re-balancing in the initial conditions for 5 min. The rate of acquisition spectra was 3.00 Hz, monitoring a mass range from *m/z* 70-1200 (polar fraction) and *m/z* 300-1600 (lipidic fraction).

A high resolution mass spectrometer was used after the UHPLC separation (MaXis 4G Q-TOF MS, Bruker Daltonics, Germany) using electrospray source in positive (ESI(+)-MS) and negative modes (ESI(-)-MS). The settings used on the MS method were: final plate offset, 500 V; capillary voltage, 3800 V; nebulizer pressure, 4 bar; dry gas flow, 9 L min^{-1}, dry temperature, 200 °C. The rate of acquisition spectra was 3.00 Hz. A sodium formate solution (10 mM HCOONa solution in 50/50 v/v isopropanol/water containing 0.2% formic acid) was injected directly through a 6-way valve at the beginning of each chromatographic run for external calibration. Ampicillin ([M+H] 350.11867 and [M-H] 348.10288) was added to each sample and was used as an internal standard for peak normalization.

Tandem mass spectrometry (MS/MS) parameters have been adjusted to improve mass fragmentation, with collision energy ranging from 20 to 50 eV, using a step method. Precursor ions were acquired using the 3.0 s cycle time. The general AutoMS settings were: mass range, *m/z* 70-1000 (polar fraction) and *m/z* 300-1600 (lipidic fraction); spectrum rate, 3 Hz; ionic, positive polarity; pre-pulse storage, 8 μs; funnel 1 RF, 250.0 Vpp. The UHPLC-MS and UHPLC-MS/MS data were acquired by HyStar Application version 3.2 (BrukerDatonics, Germany).

2.4.3. Metabolomics data analysis

The raw data from UHPLC-MS were exported as mzMXL files, using DataAnalysis 4.2 software (Bruker Daltonics, Germany) and pre-processed using XCMS Online (Gowda et al., 2014; Tautenhahn et al., 2012), for peak detection, retention time correction and alignment of the metabolites detected in the UHPLC-MS analysis. Peak detection was performed using

centWave peak detection (Δm/z = 10 ppm; minimum peak width, 5 s; maximum peak width, 20 s) and mzwid = 0.015, minfrac = 0.5, bw = 5 for alignment of retention time. The unpaired parametric t-test (Welch t-test) was used for statistical analysis.

The processed data (csv file) were exported to MetaboAnalyst 5.0, and submitted to analysis in the Statistical Analysis module (Chong et al., 2019; Chong & Xia, 2020). Before the chemometric analysis, all data variables from the polar fraction were normalized by internal standard (ampicillin-rT = 7.9 min; [M+H], *m/z* = 350.11867, [M-H], *m/z* = 348.10288); and, all data variables from the lipidic fraction were normalized by internal standard (1,2-diheptadecanoyl-sn-glycero-3-phosphocholine = 4.85 min; [M+H] + *m/z* = 762.60125). All three sets of data were scaled using the pareto method.

The differentially expressed peaks (DEP) were selected according to the following criteria: Variable Importance in Projection – VIP values ≥ 0.99, obtained from the PLS-DA model; adjusted p.value (FDR – False Discovery Rate) ≤ 0.05, of the Welch t-test; and Log_2 (FC – Fold Change) ≠ 1. The selected DEPs were then submitted to analysis in the MS Peaks to Pathway module (Chong et al., 2019; Chong & Xia, 2020) and analyzed using the following parameters: molecular weight tolerance of 5 ppm; mixed ion mode; joint analysis using the mummichog algorithm (Li et al., 2013) with a P-value cutoff of 1.0×10^{-5} and Gene Set Enrichment Analysis – GSEA (Subramanian et al., 2005) algorithms, and the latest KEGG version of the *Arabidopsis thaliana* pathway library.

In the case of a DEP with two or more matched forms (isotopes) and later a matched compound with two or more DEPs, the initial criterion of metabolite selection applied was the mass difference comparing to the metabolite database – choosing the smallest one. The second criterion was the adduct study of each candidate back in its mass spectra. Then, we applied the formula and exact mass data from KEGG; and, finally, performed the putative annotation of the metabolites of interest, with one or two candidates on each detected ion.

The KEGG IDs of the matched compounds were then submitted to pathway analysis (integrating enrichment analysis and pathway topology analysis) and visualization in the Pathway Analysis module (Chong et al., 2019; Chong & Xia, 2020) and analyzed using the Hypergeometric Test and the latest KEGG version of the *A. thaliana* pathway library.

2.5. Transcriptomics analysis

Based on the results of the morphophysiological characterization, we selected the following treatments for transcriptomics analysis: control and stressed plants – 0.0 and 1.5 g of NaCl per 100 g of substrate, respectively – at one and four days after treatment (DAT). Leaves for transcriptomics analysis were collected from five replicates, immediately immersed in liquid nitrogen, and then stored at -80 °C until RNA extraction. Three replicates per treatment were randomly selected for total RNA extraction, library preparation, and sequencing.

2.5.1. Total RNA extraction and quality analysis, library preparation and sequencing

Total RNA was isolated from purslane shoots using the Qiagen RNeasy® Plant Mini kit (QIAGEN, CA, USA) following the manufacturer's protocol. RNA quantity was measured using a Nanodrop Qubit 2.0 Fluorometer (Life Technologies, CA, USA), and RNA quality was evaluated with an Agilent Bioanalyzer Model 2100 (Agilent Technologies, Palo Alto, CA, USA). Samples were subjected to RNA-Seq using an Illumina HiSeq platform at the GenOne Company (Rio de Janeiro, Brazil), using the paired-end strategy. The raw sequence data (24 fastq files) have been uploaded in the Sequence Read Archive (SRA) database of the National Center for Biotechnology Information under the BioProject PRJNA575830 and BioSample SAMN12911623 (*Portulaca oleracea*_B1 – TaxID: 46147).

2.5.2. Transcriptomics data analysis

All the transcriptomics analysis was performed with OmicsBox version 1.3 (OmicsBox, 2019). We used FastQC (Andrews, 2010) and Trimmomatic (Bolger et al., 2014) to perform the quality control, to filter reads and remove low-quality bases. The minimum average quality of

reads kept was 30, and the minimum length of reads was 75. The default parameters from OmicsBox version 1.3 were used to create a "de novo" transcriptome without a reference genome through the softwares Trinity version 2.8.5 (Grabherr et al., 2011) and Bowtie2 version 2.3.5.1 (Langmead and Salzberg, 2012). The RNA-Seq data were aligned to the reference transcriptome using default parameters from OmicsBox version 1.3 through software STAR (Dobin et al., 2013); and, to quantify expression at gene or transcript level we used the default parameters from OmicsBox version 1.3 through HTSeq version 0.9.0 (Anders et al., 2015).

The pairwise differential expression analysis between different experimental conditions was performed through edgeR version 3.28.0 (Robinson et al., 2010), applying a simple design and an exact statistical test, without the use of filter for low counts genes. The functional analysis of the differentially expressed genes (DEGs) was performed combining differential expression results with functional annotations from the high-throughput functional annotation and data mining pipeline in OmicsBox version 1.3 (Götz et al., 2008).

2.6. Integratomics analysis

Previously to the integration of metabolome and transcriptome data, a fasta file containing all 97,613 ORFs from the reference transcriptome was submitted to analysis in the GhostKOALA annotation tool for K number assignment of KEGG genes (Kanehisa et al., 2016).

Omics Fusion (Brink et al., 2016), the web platform for integrative analysis of Omics data (https://fusion.cebitec.uni-bielefeld.de/), was employed for carrying out the integrative analysis of transcripts and metabolites. The input data used was the Log_2 (FC) data of differentially expressed transcripts and metabolites. First, to check the data distribution, we used the Data Overview module and then the Scatter Plot one for the correlation analysis between the two sets of data – a pairwise combination of the different scenarios evaluated.

For subsequent analysis, we used the modules KEGG feature distribution and Map data on the KEGG pathway. The former module employed to verify which metabolic pathways had more transcripts and metabolites differentially expressed, and the latter to map these data differentially expressed in the metabolic paths in question. For the KEGG feature distribution module, we applied the joint analysis of transcripts and metabolites with a threshold of 8, and for the Map data on the KEGG pathways, the organism code bvg (*Beta vulgaris* – sugar beet) was used for mapping.

3. Results

3.1. The addition of salt led to changes in the electrical conductivity and water potential of the substrate, and promoted changes in biomass production and visual aspect of the plants

As the amount of NaCl added to the substrate rose, it led to a dose-dependent rise in electrical conductivity – EC (Figure 1A) and a fall in water potential (Figure 1B) of the saturation extract. The control (no added salt) showed an EC lower than the minimal set to consider the soil to be saline (4 dS m^{-1}), as expected. The three highest amounts of NaCl used resulted in ECs higher than 20 dS m^{-1}; the 2.0 g of NaCl / 100 g of the substrate resulted in an EC a little over 50 dS m^{-1}. The water potential in the substrate decreased as the NaCl amount rose, and the treatment that received the highest amount of NaCl showed water potential of almost -1.8 MPa.

As the amount of NaCl in the substrate increased, the shoot (Figure 1C) and root (Figure 1D) biomass of young purslane plants decreased about 50% and 90%, respectively, in the stress treatments with ≥1.5 g of NaCl / 100 g of the substrate. The visual aspect of the purslane plants observed at the end of the experiment was recorded (Figure 2). When comparing to the control (Figure 2A), except for the smallest amounts of NaCl used (Figure 2B and 2C), changes in appearance and color of the leaves and stem were evident in the stress treatments with ≥1.5 g of NaCl / 100 g of the substrate (Figure 2D and 2E).

3.2. Salt stress affected gas exchange, and pigment content in purslane plants

All purslane plants showed similar values for all the gas exchange variables before adding NaCl – day zero (Figure 3). One day after adding NaCl to the substrate, one already can see a reduction in the rates of CO_2 assimilation (*A*), transpiration (*E*), and stomatal conductance to water vapor (*gs*), which correlated with the amount of NaCl used (Figure 3A, 3B, and 3C). Thereafter, those variables remained reasonably constant throughout the

experiment. On the other hand, there was an increase in intercellular CO_2 concentration (*Ci*) in the two treatments with the highest amount of NaCl (Figure 3D). The reductions in *A* and *E* were proportional to *gs* at the intermediate salt stress levels, indicating a stomatal limitation to gas exchange. The increase in *Ci* at the highest levels of salt stress pointed to a difficulty in CO_2 assimilation, which intensified over time.

The chlorophyll fluorescence variables remained relatively stable in control plants throughout the experiment (Figure 4). However, changes occurred as the saline concentration in the substrate increased. In general, from the addition of 1.0 g of NaCl to the substrate, the leaves of purslane plants began to show a reduction in the maximum fluorescence in the dark (Fm), the effective quantum yield of photosystem II [Y(II)], and maximum quantum yield of photosystem II (Fv/Fm). On the other hand, there was an increase in initial fluorescence in the dark (Fo), in the quantum yield of regulated [Y(NPQ)], and unregulated [Y(NO)] dissipation of the energy of the light. Such responses accentuate throughout the stress period, mainly in the treatments with ≥1.5 g of NaCl / 100 g of the substrate.

There were practically no differences between the pigment content before the onset of the saline stress (day zero). As the period of stress went on, there was a downward trend in the chlorophyll and carotenoid indices in a dose-dependent manner. At the same time, PRI values were also reduced in a dose-dependent way, indicating that the plants kept the xanthophylls cycle functioning, despite the degradation of photosynthetic pigments. In general, as higher was the NaCl level applied, the lower the values of chlorophyll, carotenoid, and photochemical reflectance indexes were on day 5 (Figure 5).

3.3. Appearance of salt crystals on the edges of purslane leaves

As some white crystals appeared on the margin of leaves in some young purslane plants under saline stress (data not shown), leaf samples were collected for microscopy analysis. Scanning electron microscopy (SEM) images showed the formation of wrinkles on the leaf

surface, probably due to dehydration (Supplementary Figure 1). In general, the stomata on the leaves of salt-stressed purslane plants started to close already in the lowest NaCl level used, and were completely closed on the highest levels of NaCl used. It was possible to see the white crystals in the surrounding and on top of the stomata (Supplementary Figure 1).

To identify the chemical composition of these crystals, we obtained SEM images with detectors of energy dispersive spectroscopy (EDS) (Figure 6A and 6B). The compositional map allowed identifying the crystals' main constituents as Na^+, Cl^-, and K^+ (Figure 6B). These results indicate that *P. oleracea* has a mechanism of salt exclusion operating on the leaves, which has its role in salt tolerance.

3.4. Purslane metabolome under salinity stress – single analysis

PLS-DA (partial least squares discriminant analysis) permutations tests were performed using the all treatments data set (control and stressed plants at 1 and 4 DAT) to validate the model, applying permutation number = 2,000. When evaluated by group separation distance, the probability that the model was created by chance was less than 0.0005, independent of the fraction – polar-positive, polar-negative, and lipidic-positive; and the evaluation by prediction accuracy also showed a probability less than 0.0005, independent of the fraction. The cross-validation analysis determines the optimal number of components necessary to build the PLS-DA model by measuring three criteria. These criteria are R^2 (the sum of squares captured by the model), Q^2 (the predictive ability of the model estimated by cross-validation), and Accuracy (Chong et al., 2019). When choosing the three-component model based on Q^2, the Q^2 value obtained for the all treatments data set was 89.98%, 80.55%, and 85.82%, for the polar-positive, polar-negative, and lipidic-positive fraction, respectively (data not shown).

Before submitting the data for analysis in the statistics module, they were organized as follows: age effect – AE (samples from the control plants at 1 and 4 days after treatment - DAT); short-term stress – STS (the control and the stressed plants at 1 DAT); long-term stress 1 – LTS1

(the control and the stressed plants at 4 DAT); and long-term stress 2 – LTS2 (the stressed plants at 1 and 4 DAT). Each data set had three biological replicates per treatment.

The AE data set contained 1,003, 2,008, and 1,113 peaks, respectively, in the polar-positive, polar-negative, and lipidic-positive fractions (Table 1). The accumulated variances explained by the first three principal components tested were 88.70%, 85.70%, and 50.80% for the fractions above-mentioned, respectively (data not shown). All but two peaks did not show a difference in expression between 1 and 4 DAT (Table 1). These results clearly show that there is no age effect modifying the metabolome profile of purslane plants between 1 and 4 DAT. As already stated in the Materials and Methods section, a differentially expressed peak (DEP) is a peak with a VIP $\geq$ 0.99, a FDR $\leq$ 0.05, and Log_2 (FC) > 0 (upregulated) or Log_2 (FC) < 0 (downregulated).

The short-term stress data set was employed to evaluate how distinct are the metabolome profiles of the control and stressed plants at 1 DAT. The samples applied to evaluate the STS scenario contained 1,034, 2,046, and 1,114 peaks, respectively (Table 1). The accumulated variances explained by the first three principal components tested were 96.10%, 95.10%, and 88.90% for the fractions above-mentioned, respectively (data not shown). On average, 90.38% of the peaks did not show a difference in expression between control and stressed plants at 1 DAT, while 141, 198, and 10 peaks upregulated, and 18, 52, and 4 downregulated, respectively (Table 1).

The long-term stress 1 data set was employed to evaluate how distinct are the metabolome profiles of the control and stressed plants at 4 DAT. The samples applied to evaluate the LTS1 scenario contained 1,034, 2,046, and 1,114 peaks, respectively (Table 1). The accumulated variances explained by the first three principal components tested were 98.40%, 91.80%, and 91.80% for the fractions above-mentioned, respectively (data not shown). On average, 92.95% of the peaks did not show a difference in expression between control and

stressed plants at 4 DAT, while 46, 182, and 5 peaks upregulated, and 33, 67, and 10 downregulated, respectively (Table 1).

At last, the long-term stress 2 data set was employed to evaluate how distinct are the metabolome profiles of the stressed plants at 1 and 4 DAT. The samples applied to evaluate the LTS2 scenario also contained 1,034, 2,046, and 1,114 peaks, respectively (Table 1). The accumulated variances explained by the first three principal components tested were 91.50%, 92.40%, and 90.80% for the fractions above-mentioned, respectively (data not shown). On average, 92.11% of the peaks did not show a difference in expression between the stressed plants at 1 and 4 DAT, while 2, 15, and 4 peaks upregulated, and 53, 127, and 123 downregulated, respectively (Table 1).

Taken together, the metabolome profiles obtained from the analysis of the STS, LTS1, and LTS2 data sets in the Statistical Analysis module of MetaboAnalyst 5.0 revealed 748 DEPs – those peaks differentially expressed in STS, plus those DEPs in LTS2 but not in STS, and plus those DEPs in LTS1 but not in STS or LTS2. This group of 748 DEPs was submitted to functional interpretation via analysis in the MS Peaks to Pathway module (Chong et al., 2019; Chong and Xia, 2020), as described in the Materials and Methods section. The combined mummichog and GSEA pathway meta-analysis resulted in a list of 68 ranked pathways enriched in this group of DEPs. Six had combined p.value ≤ 0.05 – Alanine, aspartate and glutamate metabolism, Monobactam biosynthesis, Tryptophan metabolism, Lysine biosynthesis, Aminoacyl-tRNA biosynthesis, and C5-Branched dibasic acid metabolism pathways (Figure 7A).

After applying the initial criteria of metabolite selection, as described in the Materials and Methods section, 109 DEPs (Supplementary Table 1) with a hit to just one known compound were submitted to the pathway topology analysis module, resulting in a list of 63 ranked pathways (Figure 7B). The Nicotinate and nicotinamide metabolism, C5-Branched

dibasic acid metabolism, and Phenylpropanoid biosynthesis pathways, with an FDR (False Discovery Rate) of 0.014831, 0.017755, and 0.041788, came out at the top of this rank, respectively; and an impact of 0.4, 1.0, and 0.2, respectively. These three pathways had 20 differentially expressed metabolites with the highest level of significance in the set of 109 matched metabolites submitted to analysis; being 6 out of its 13 metabolites in the Nicotinate and nicotinamide metabolism pathway, 4 out of its 6 in the C5-Branched dibasic acid metabolism, and 10 out of its 46 in Phenylpropanoid biosynthesis (Supplementary Table 1).

All the 109 DEPs with a hit to just one known compound were also submitted to correlation analysis by means of pairwise comparison of three out of the four scenarios tested (AE, STS, LTS1, and LTS2), using Log_2(FC) values. As there was not an age effect detected in the metabolomic analysis (Table 1), the AE scenario was not used for correlation analysis. The correlation analysis revealed strong positive correlations between STS and LTS1, and LTS1 and LTS2 (Figure 8A and 8C), and a weak positive correlation between STS and LTS2 (Figure 8B). As STS compares control and stressed plants at 1 DAT, and LTS2 compares control and stressed plants at 4 DAT, this weak positive correlation implies that the behavior seen for most of the 109 DEPs at short-term stress does not repeat at the long-term stress. Meanwhile, the behavior seen in most of the 109 DEPs in STS repeats itself in LTS1; the same is true when comparing LTS1 and LTS2. When comparing STS and LTS1, the Log_2(FC) value at LTS1 is already the result of the differential expressed changes seen at 1 and 4 DAT; but when comparing STS and LTS2, the Log_2(FC) value in LTS2 must be added to the value found in STS to reflect the final change in expression level. In the case of comparing LTS1 and LTS2, the Log_2(FC) value at LTS1 reflects the final change in expression level due to the stress.

Taken together, the results from this large-scale metabolome single analysis reveal three pathways affected by salt stress in the leaves of young purslane plants, as well as several

metabolites – from these and other pathways – that should be the focus of further characterization of the role of metabolites in the tolerance of this halophyte species to salt stress.

3.5. Purslane transcriptome under salinity stress – single analysis

The raw sequence data (26 fastq files) generated in this study have been uploaded in the Sequence Read Archive (SRA) database of the National Center for Biotechnology Information under *Portulaca oleracea*_B1 - BioProject number of PRJNA575830, BioSample SAMN12911623. A total of 296,493,969 high-quality-pairs of reads - with an average quality of reads ≥30, and the minimum length of 75 nucleotides – remained after pre-processing the raw sequence data (data not shown). These high-quality sequences were used to assemble the Reference Transcriptome (RT), as well as to perform the mapping, counting, and differential expression analysis. The RT assembled has 365,297,960 total assembled bases that resulted in 415,379 transcripts, 252,197 genes, and 49,412 complete ORFs (data not shown). The RT presented a 42.09% GC content, 879.43 bases as average size of transcripts, and a N50 equal to 1,493 bases per transcript.

The differential expression analysis was performed in order to measure the a possible age effect – AE (samples from the control plants at 1 and 4 days after treatment – DAT), as well as to measure the short-term stress – STS (the control and the stressed plants at 1 DAT) the long-term stress 1 – LTS1 (the control and the stressed plants at 4 DAT), and long-term stress 2 – LTS2 (the stressed plants at 1 and 4 DAT). Out of the 252,197 genes from the RT, 29,737 remained for differential expression analysis after applying the following criteria: CPM filter equals to 1.0, number of samples reaching CPM filter equals to 3, normalization method TMM (Trimmed mean of M values). Differentially expressed genes (DEGs) are those with a FDR < 0.05, and FC > 1 (upregulated) or FC < -1 (downregulated).

Differential expression analysis of the AE data set revealed 3,512 (11.81%) upregulated and 2,868 (9.64%) downregulated genes in the control plants at 4 DAT when comparing to 1

DAT (Table 1). In the STS scenario, a total of 8,430 (28.35%) upregulated and 8,280 (27.84%) downregulated in the stressed plants comparing to the control plants at 1 DAT. In the LTS1 scenario, a total of 11,005 (37.01%) upregulated and 11,550 (38.84%) downregulated in the stressed plants comparing to the control plants at 4 DAT. At last, in the LTS2 scenario, 8,693 (29.23%) upregulated and 8,994 (30.25%) downregulated in the stressed plants at 4 DAT comparing to the stressed ones at 1 DAT.

The 29,737 genes submitted to differential expression analysis in all four scenarios (AE, STS, LTS1, and LTS2) were separated accordingly to the combined profiles in 15 groups (Figure 9). A total of 442 genes (1,49%) were upregulated in all four scenarios (Figure 9A), while 2,755 (9,26%) did not differentially expressed in all scenarios (Figure 9B), and 524 (1,76%) were downregulated in all scenarios (Figure 9C). When considering those genes that up- or downregulated twice in STS and LTS2, but not in AE, a total of 3,053 (10,27%) and 2,864 (9,63%) genes were selected, respectively (Figure 9A and 9C). No genes were found up- or downregulated only in AE and STS, or in AE and LTS1, but not in the other scenarios (Figure 9A and 9C).

All 29,737 genes were also submitted to correlation analysis by means of pairwise comparison of the four scenarios tested (AE, STS, LTS1, and LTS2), using Log_2(FC) values. As there was an age effect detected in the transcriptomic analysis (Table 1), the AE scenario was also used for correlation analysis. The genes not differentially expressed in STS were not used in the correlation analysis against AE, LTS1 and LTS2; the same is true for LTS1 against AE and LTS2, and for LTS2 against AE (Figure 10). The correlation analysis revealed strong positive correlations between STS and LTS1, and LTS1 and LTS2 (Figure 10D and 10F), weak positive correlations between AE and STS, AE and LTS2, and STS and LTS2 (Figure 10A, 10C, and 10E), and weak negative correlation between AE and LTS1 (Figure 10B).

A weak positive or negative correlation implies that the behavior seen for most of the genes in one scenario does not repeat the second scenario evaluated. Even in such a case, it is possible to identify gene(s) with a behavior of interest for further studies. As an example, one can see in the fourth quadrant in Figure 10B the existence of a group of genes that the level of upregulation in AE is exactly the same of downregulation in LTS1; meaning that the saline stress affected the expression of these genes in a way that let their expression level to what it was in the control plants at 1 DAT. Another example is found in Figure 10C, on the line separating the first and the second quadrants, near the Log_2(FC) value of 10, where genes not differentially expressed in AE experienced an increase in expression of about 1,000 fold between 1 and 4 DAT.

In the case of the strong positive correlation seen when comparing STS and LTS1, and LTS1 and LTS2, the behavior seen for most of the genes in one scenario does repeat itself in the second scenario evaluated. When comparing STS and LTS1, the Log_2(FC) value at LTS1 is already the result of the differential expressed changes seen at 1 and 4 DAT; but when comparing LTS1 and LTS2, it is the Log_2(FC) value at LTS1 that reflects the final change in expression level due to the stress. It can be infer that genes upregulated twice - in STS and LTS2 – are among those most important to further characterization regarding their role in the response of this halophyte species to salinity stress. Even though there is a weak positive correlation when comparing STS and LTS2, it was possible to identify several genes that upregulated twice (Figure 10E).

Taken together, the results from this large-scale transcriptome single analysis reveal many opportunities for further characterization of salt-responsive genes aiming either the identification of candidate genes for salt tolerance or the prospection of promoter sequence for biotechnological application – salt stress-dependent expression of genes of interest.

3.6. Purslane transcriptome and metabolome under salinity stress – integrative analysis

The annotation of the fasta file containing all 97,613 ORFs from the purslane RT in GhostKOALA, against the genus_prokaryotes + family_eukaryotes databases, resulted in 38,700 entries annotated under functional categories (data not shown). The KEGG Mapper reconstruction analysis allowed the identification of 1,802 enzymes from eudicots (data not shown).

Out of the 415,379 isoform features from the purslane RT, 72,830 remained for differential expression analysis after applying the following criteria: CPM filter equals to 1.0, number of samples reaching CPM filter equals to 3, normalization method TMM (Trimmed mean of M values). Out of these 72,830 isoform features, 23,834 did not differentially expressed in any of the four scenarios tested (AE, STS, LTS1, and LTS2), and 48,996 did in at least one of these scenarios (data not shown). Differentially expressed isoforms (DEIs) are those with a FDR < 0.05, and FC > 1 (upregulated) or FC < -1 (downregulated). A search for the 1,802 enzymes among the 48,996 DEIs led to a list of 5,883 DEIs with KO number and E.C. number.

Only two out the four scenarios studied underwent transcriptome and metabolome integration analysis; short-term stress - STS (control vs stressed plants at 1 DAT), and long-term stress 1 – LTS1 (control and stressed plants at 4 DAT). In order to do that, the list of 109 DEPs was first filtered to select only those differentially expressed in STS or LTS1; and the same was done with the list of 5,883 DEIs. The filtered lists from DETs and DEIs were then combined for the integration analysis. In the STS scenario, the Glycerophospholipid metabolism pathway came first in the rank of combined enzymes and compounds occurrency, with 6 enzymes and 5 compounds; while in the LTS1 scenario the Purine metabolism pathway came first with 4 enzymes and 7 compounds (Table 2). When considering both scenarios

studied, the Glycerophospholipid metabolism pathway comes first as the most affected pathway in young purslane plants under high salinity stress, followed by Cysteine and methionine metabolism (Table 2).

4. Discussion

4.1. Morphophysiological responses of young purslane plants to saline stress

When adding NaCl to a solution or wet substrate, this salt dissociates into its component ions - Na^+ and Cl^-. These ions enable the conduction of electric current (Visconti and de Paz, 2016). The higher the concentration of such ions in the solution, the greater its electrical conductivity (EC) (Polle and Chen, 2015; Rhoades et al., 1999), as as seen in the Figure 1A. On the other hand, the dissociation of the component ions of the salt in solution reduces the osmotic potential (Figure 1B), as water molecules are needed to dissolve the ions (Cordeiro, 2001; Ramoliya et al., 2004). Both increase in EC and reduction in water potential affect plant metabolism (Ali et al., 2019; Maksimovic and Ilin, 2012; Munns, 2002; Parida and Das, 2005; Duarte and Souza, 2016; Mane et al., 2011).

The drop in leaf gas exchange rates is a plant common response to salt and its intensity depends on the level of plant tolerance to salinity stress (Everard et al., 1994; Gale, 1975; Koyro, 2006; Parihar et al., 2015; Alam et al., 2015; Xing et al., 2019). At the lowest doses of salt, that is, up to 1.0 g of NaCl / 100 g of soil, the drop in gas exchange rates for young purslane plants seems to be due to stomatal closure, as net CO_2 assimilation, transpiration rate and stomatal conductance fell in the same proportion (Figure 3). This is basically due to the reduction of water potential, which makes difficult for the plants to absorb water and results in partial or total closure of stomata, depending on the severity of the salt stress (Flowers et al., 2015; Gale, 1975; Qiu et al. , 2003). At higher doses, probably, the activities of CO_2-fixing enzymes decreased during stress, as the intercellular CO_2 concentration increased. Such increase was intensified over time (Figure 3D). Similar results were previously observed in purslane (Alam et al., 2015; Xing et al., 2019). According to Van Zelm et al. (2020), this finding suggests that there may also be an ionic effect, or at least a stomatal closure-independent effect, of sodium on photosynthesis.

In purslane plants under saline stress of 0.5 g of NaCl, or approximately 8 dS m^{-1}, gas exchange rates did not differ much from the control plants, mainly at the end of the stress period (Figure 3). These plants managed to lose only a small percentage (13.27%) of shoot biomass (Figure 1C), while they lost almost 50% of the root biomass (Figure 1D). Under saline stress of 20 dS m^{-1} or more, stomatal closure was more pronounced. As for the intercellular CO_2 concentration, the results suggest that the mesophyll's resistance to carboxylation occurred from the 1.5 g NaCl level (Figure 3D). In summary, as a consequence of the changes in gas exchange rates at salt stress up to 50 dS m^{-1}, purslane plants got smaller (shoots and roots), but did not die (Figure 2).

In general, the aerial part of the plant is more affected than roots by salt stress (Acosta-Motos et al., 2017; Munns and Tester, 2008). In some cases, biomass production is stimulated by salinity depending on the salt concentration, as in *Atriplex nummularia* L. that is stimulated at 300 mM but inhibited at 600 mM NaCl (de Araújo et al., 2006). In *Chloris gayana* and *Salvadora persica*, root growth was more affected than shoot (Céccoli et al., 2011; Rao et al., 2004), as observed in young purslane plants (Figure 1). Due to their direct contact with the soil, roots are considered the first damage sites (Rewald et al., 2013). In another study, purslane plants showed reductions in the aerial part and roots as a function of salinity (150 mM and 200 mM NaCl) (Xing et al., 2019).

Recently, Xing et al. (2019) reported that, under saline stress between 100 and 200 mM NaCl, purslane plants significantly decreased the net photosynthetic rate, increased the intercellular concentration of CO_2, the content of malondialdehyde, the production rate of O_2^-, and the activity of the enzymes SOD, POD, and CAT. Such results suggest the inhibition of photosynthesis and the occurrence of oxidative stress. Taking into consideration the changes in the net photosynthetic rate and the intercellular concentration of CO_2, our results correlate to

the one found by Xing et al. (2019), even though the NaCl concentrations and plant ages used were different.

Chlorophyll fluorescence variables showed a differential effect of the level of saline stress on the photochemical apparatus (Figure 4). The addition of salt to the substrate, regardless of the level, caused damage to the chloroplast membrane system, which became more pronounced throughout the stress period, as can be inferred from the increase in minimum fluorescence yield on dark-adapted plants (Fo), the drop in maximum fluorescence yield on the dark-adapted plants (Fm), and, consequently, in maximum PSII quantum yield (Fv/Fm). The electron flow in the Z scheme of photosynthesis fell over time. After five days, such drop was directly proportional to the level of NaCl added to the substrate, as can be inferred from the PSII effective quantum yield [Y(II)] data. The light energy not used to flow the electrons in the Z scheme was directed towards the generation of heat, as seen in the regulated energy dissipation quantum yield [Y(NPQ)], but up to a certain limit of saline stress. This limit was up to 1.0 g of NaCl in the substrate, whose fluorescence emission in the light, despite the initial increase, was practically the same as the control at the end of the stress period (Figure 4). Purslane plants continued to increase the emission of fluorescence in the light with the increase of the level of salt in the substrate, as seen in the unregulated energy dissipation quantum yield [Y(NO)]; meaning that plants no longer had the means to regulate the extinction of fluorescence by their photochemical apparatus.

The drop in the levels of CI and CRI (Figure 5) was an expected result for purslane plants under salt stress, considering that it is recurrently reported for several plant species, such as *Salicornia persica* and *S. europaea* (Aghaleh et al., 2009), *Ricinus communis* L. (Li et al., 2010) and *Cakile maritime* (Megdiche et al., 2008). In the case of purslane, the effect was observed especially from 1.0 g of NaCl in the substrate, and more sharply at higher levels of NaCl. It is interesting to note that at the end of the stress period, the decrease in PRI was

proportional to the salt stress (Figure 5C), which indicates that the plants probably kept the xanthophyll cycle in full operation at all salt levels. The activation of the xanthophyll cycle is a response observed in some halophyte species under salt stress (Qiu et al., 2003; Rabhi et al., 2012) and is related to the protection of the photosynthetic apparatus against photoinhibition damage (Qiu et al., 2003).

The effect of salt stress on biomass accumulation in purslane plants was most likely related to the osmotic effect, which resulted in stomatal closure and restricted the entry of CO_2 into the leaf mesophyll. So much so that the drop in biomass accumulation was proportional to the drop in net CO_2 assimilation rates. Evidently, throughout stress, additional mechanisms contributed to the effect of saline stress being more pronounced in the roots than in the aerial part. The reduction of biomass due to high salinity is a common response from several species such as *Atriplex griffithii* var. *stocksii* (Khan et al., 2000b), *Suaeda fruticosa* L. (Khan et al., 2000a), *Mentha piperita* L. (Khorasaninejad et al., 2010), *Vetiveria zizanioides* (L.) Nash (Mane et al., 2011), and even for *Portulaca oleracea* (Alam et al., 2015; Alam et al., 2014).

Some purslane plants under salt stress showed structures resembling salt crystals on the leaf surface, which was confirmed to be constituted mainly of Na^+, Cl^- and K^+ (Figure 6). This salt found there could only have been excluded by the leaves. The ability to exclude salt thought roots and leaves is a characteristic observed in several halophyte species, such as *Limonium bicolor*, *Reaumuria hirtella*, *Spartina anglica*, *Limonium vulgare*, *Armeria maritima*, *Glaux maritime*, and others (Rozema et al., 1981; Ramadan, 1998; Flowers et al., 2010; Yuan et al., 2016). Such ability represents a self-regulating behavior, and the secretion can occur through epidermal pores and glands located in roots, shoots, and leaves. The intracellular transport mechanisms are responsible for moving excess salt from the surface cells to the outside of leaves or stem, and as the water evaporates, it is possible to observe salt crystals (Arora and Rao, 2017). In the present study, the salt crystal-like structures seen on and around closed

stomata on the leaves of young purslane plants had its composition mapped, indicating that sodium, chlorine, and potassium ions were excluded probably though salt glands, and accumulated on the leaf surface.

4.2. Single and integrated changes in the transcriptome and metabolome profiles of young purslane plants in response to saline stress

4.2.1. Insights from the transcriptome and metabolome profiles as a result of single-omics analyses

The conceptual integration strategy analyses the different omics data sets separately (single analysis), and the connection of the resulting conclusions results arbitrarily without further analysis of the data sets (Cavil et al., 2016; Rai et al., 2017; Jamil et al., 2020). This approach to connect conclusions from the single-omic analysis can produce valuable insights; however, it may miss reproducible associations when multiple omics data sets are analyzed together (Cavil et al., 2016; Rai et al., 2017). Jamil et al. (2020) state that this approach should not be considered a part of the MOI approach.

The single-metabolomic analysis done in the present study led to the identification of three pathways - Nicotinate and nicotinamide metabolism, C5-Branched dibasic acid metabolism, and Phenylpropanoid biosynthesis - at the top of a rank of 63 generated when using the integrating enrichment and pathway topology analysis protocol from the pathway analysis module in MetaboAnalyst 5.0 (Chong et al., 2019; Chong and Xia, 2020). Together, these three pathways had 20 differentially expressed metabolites out of 109 identified in the MS Peaks to Pathway module (Supplementary Table 1).

The (single) analysis of the changes in the transcriptome profile of young purslane plants in response to saline stress led to the identification of several different groups of genes, accordingly to the response seen in four distinct scenarios analyzed – AE, STS, LTS1, and LTS2 (Figure 9). A group of 5,917 genes – out of the 29,737 genes from the purslane reference transcriptome – were up or downregulated twice, respectively, in STS and LTS2 but not in AE. The set of genes that upregulated in the STS scenario, as a consequence of the osmotic stress, and again in the LTS1, as a consequence of the ionic and oxidative stresses, is a database

valuable for further studies aiming the selection and deep structural/functional characterization of genes that may play a role in purslane high tolerance to salinity stress.

This present study did not attempt to connect conclusions from these two single-omic analyses. It has focused on getting insights on candidate genes and metabolites for further characterization – via structural/functional characterization of genes and proteins or targetted metabolomics. The strategy applied in this study to perform this element-based correlation analysis, comparing different scenarios in a pairwise manner, turned out to be a powerful tool to select candidate salt-responsive genes for future work aiming for functional genomics studies.

4.2.2. Insights from the transcriptome and metabolome profiles as a result of MOI analyses

The Multi-Omics Integration (MOI) System presented by Jamill et al. (2020) has classified the integration strategies in three levels with increasing degrees of complexity: element- (level 1), pathway- (level 2), and mathematical-based approach (level 3). Here, a level 2 approach integrated transcriptome and metabolome data sets resulted from single–omics analysis. The differently expressed enzymes selected from the transcriptomics dataset are the central factor needed to perform the level 2 integration strategy in the Omics Fusion (Brink et al., 2016). They allow to link genes and metabolites from a metabolic pathway and consequently point out the ones most affected by the salt stress.

The transcriptome-single analysis using the Omics Fusion platform (Brink et al., 2016) led to 75 pathways in the STS scenario, and the same was true to metabolome-single one (data not shown). The MOI (transcriptome and metabolome) analysis of the STS scenario led to 44 pathways having differentially expressed transcripts and metabolites. Meanwhile, the single analysis revealed 76 and 63 affected ones in the LTS1 scenario. The MOI analysis led to 42 pathways having differentially expressed transcripts and metabolites. When combining the

results of the MOI analysis – STS and LTS1 – it showed that 35 were common to both scenarios, while nine were present only in STS, and seven only in LTS1 (Table 2).

The glycerophospholipid metabolism pathway was the one with the highest number of combined enzymes and metabolites differentially expressed in STS, followed by the cysteine and methionine metabolism (Table 2). Previously, these pathways were linked to plant response to abiotic stress, including salt stress (Hou et al., 2016; Zhang et al., 2017; Sui et al., 2017; Xia et al., 2019; He and Ding, 2020), and the overexpression of genes from these pathways have shown to be effective in conferring tolerance to salt stress (Sui et al., 2017; Ma et al., 2017).

Taken together, the results from the MOI analysis reveals a set of pathways with enzymes and metabolites differentially expressed due to salinity stress at 1 DAT (STS) and 4 DAT (LTS), which is also a database valuable for new studies aiming to further characterize the mechanisms behind the strong tolerance seen in young purslane plants to saline stress.

This present study did not explore the changes in the metabolome and transcriptome profiles in the roots of young purslane plants. The main reason for that was the small amount of root tissue available at the end of the experiments (Figure 1D). We performed new studies where new omics data - transcriptomic, metabolomic, ionomic, and proteomic - were gathered from adult purslane plants under salinity stress to circumvent this problem seen in young purslane plants (Salgado et al., unpublished).

5. Conclusion

- The protocol developed in this study for assessing purslane (*Portulaca oleracea*) responses to saline stress is successful in generating different levels of stress by gradients of electrical conductivity and water potential in the saturation extract of the substrate, according to the added NaCl;

- As expected from a halophyte species, young purslane plants remained alive under very high levels of salinity stress (> 20 dS m^{-1}). The salt crystal-like structures seen on and around closed stomata on the leaves of these plants are constituted mainly by Na^+, Cl^-, and K^+, indicating that *P. oleracea* has a mechanism of salt exclusion operating on the leaves, which has its role in salt tolerance;

- The correlation analysis strategy applied in this study produced a list of salt-responsive metabolites and genes valuable for future studies aiming at prospecting genes conferring high tolerance to salinity stress; and

- The MOI strategy applied in this study led to a group of 51 pathways that had at least one enzyme and one metabolite differentially expressed due to salinity stress. This pathways list is a valuable tool for future targetted-metabolomics and transcriptomics studies aiming to deepen our knowledge on the mechanisms behind the high tolerance of young purslane plants to salinity stress.

6. References

Acosta-Motos, J., Ortuño, M., Bernal-Vicente, A., Diaz-Vivancos, P., Sanchez-Blanco, M., & Hernandez, J. (2017). Plant responses to salt stress: Adaptive mechanisms. *Agronomy*, *7*(1), 18. https://doi.org/10.3390/agronomy7010018

Aghaleh, M., Niknam, V., Ebrahimzadeh, H., & Razavi, K. (2009). Salt stress effects on growth, pigments, proteins and lipid peroxidation in Salicornia persica and S. europaea. *Biologia plantarum*, *53*(2), 243–248. https://doi.org/10.1007/s10535-009-0046-7

Alam, Md. A., Juraimi, A. S., Rafii, M. Y., & Abdul Hamid, A. (2015). Effect of salinity on biomass yield and physiological and stem-root anatomical characteristics of purslane (Portulaca oleracea L.) accessions. *BioMed Research International*, *2015*, 1–15. https://doi.org/10.1155/2015/105695

Alam, Md. A., Juraimi, A. S., Rafii, M. Y., Abdul Hamid, A., & Aslani, F. (2014). Screening of purslane (Portulaca oleracea L.) accessions for high salt tolerance. *The Scientific World Journal*, *2014*, 1–12. https://doi.org/10.1155/2014/627916

Ali, A., Maggio, A., Bressan, R., & Yun, D.-J. (2019). Role and functional differences of HKT1-type transporters in plants under salt stress. *International Journal of Molecular Sciences*, *20*(5), 1059. https://doi.org/10.3390/ijms20051059

Allbed, A., & Kumar, L. (2013). Soil salinity mapping and monitoring in arid and semi-arid regions using remote sensing technology: A review. *Advances in Remote Sensing*, *02*(04), 373–385. https://doi.org/10.4236/ars.2013.24040

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics*, *31*(2), 166–169. https://doi.org/10.1093/bioinformatics/btu638

Andrews, S. (2010). *FastQC: a quality control tool for high throughput sequence data*. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Arif, Y., Singh, P., Siddiqui, H., Bajguz, A., & Hayat, S. (2020). Salinity induced physiological and biochemical changes in plants: An omic approach towards salt stress tolerance. *Plant Physiology and Biochemistry*, *156*, 64–77. https://doi.org/10.1016/j.plaphy.2020.08.042

Arora, S., & Rao, G. G. (2017). Bio-amelioration of salt-affected soils through halophyte plant species. In S. Arora, A. K. Singh, & Y. P. Singh (Orgs.), *Bioremediation of Salt Affected Soils: An Indian Perspective* (p. 71–85). Springer International Publishing. https://doi.org/10.1007/978-3-319-48257-6_4

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Borsai, O., Hassan, M. A., Boscaiu, M., Sestras, R. E., & Vicente, O. (2018). The genus *Portulaca* as a suitable model to study the mechanisms of plant tolerance to drought and salinity. *The EuroBiotech Journal*, *2*(2), 104–113. https://doi.org/10.2478/ebtj-2018-0014

Brink, B. G., Seidel, A., Kleinbölting, N., Nattkemper, T. W., & Albaum, S. P. (2016). Omics fusion—A platform for integrative analysis of omics data. *Journal of Integrative Bioinformatics - JIB*. https://doi.org/10.2390/BIECOLL-JIB-2016-296

Cavill, R., Jennen, D., Kleinjans, J., & Briedé, J. J. (2016). Transcriptomic and metabolomic data integration. *Briefings in Bioinformatics*, *17*(5), 891–901. https://doi.org/10.1093/bib/bbv090

Céccoli, G., C. Ramos, J., I. Ortega, L., M. Acosta, J., & G. Perreta, M. (2011). Salinity induced anatomical and morphological changes in Chloris gayana Kunth roots. *BIOCELL*, *35*(1), 9–17. https://doi.org/10.32604/biocell.2011.35.009

Chong, J., Wishart, D. S., & Xia, J. (2019). Using metaboanalyst 4. 0 for comprehensive and integrative metabolomics data analysis. *Current Protocols in Bioinformatics*, *68*(1). https://doi.org/10.1002/cpbi.86

Chong, J., & Xia, J. (2020). Using metaboanalyst 4. 0 for metabolomics data analysis, interpretation, and integration with other omics data. In S. Li (Org.), *Computational Methods and Data Analysis for Metabolomics* (Vol. 2104, p. 337–360). Springer US. https://doi.org/10.1007/978-1-0716-0239-3_17

Cordeiro, G. G. (2001). Salinidade em areas irrigadas. Embrapa Semiárido-Artigo em periódico indexado. *Embrapa Semiárido-Artigo em periódico indexado (ALICE).*

D'Andrea, R. M., Andreo, C. S., & Lara, M. V. (2014). Deciphering the mechanisms involved in Portulaca oleracea (C 4) response to drought: Metabolic changes including crassulacean acid-

like metabolism induction and reversal upon re-watering. *Physiologia Plantarum*, *152*(3), 414–430. https://doi.org/10.1111/ppl.12194

de Araújo, S. A. M., Silveira, J. A. G., Almeida, T. D., Rocha, I. M. A., Morais, D. L., & Viégas, R. A. (2006). Salinity tolerance of halophyte Atriplex nummularia L. grown under increasing NaCl levels. *Revista Brasileira de Engenharia Agrícola e Ambiental*, *10*(4), 848–854. https://doi.org/10.1590/S1415-43662006000400010

Devi, S., Nandwal, A. S., Angrish, R., Arya, S. S., Kumar, N., & Sharma, S. K. (2016). Phytoremediation potential of some halophytic species for soil salinity. *International Journal of Phytoremediation*, *18*(7), 693–696. https://doi.org/10.1080/15226514.2015.1131229

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Duarte, H. H. F., & Souza, E. R. de. (2016). Soil water potentials and Capsicum annuum L. under salinity. *Revista Brasileira de Ciência do Solo*, *40*(0). https://doi.org/10.1590/18069657rbcs20150220

El-Shora, H. M., & El-Gawad, A. M. (2015). Response of *Cicer arietinum* to Allelopathic Effect of *Portulaca oleracea* Root Extract shora. Phyton, 55(2), 215-232. https://doi.org/10.12905/0380.phyton55(2)2015-0215

Everard, J. D., Gucci, R., Kann, S. C., Flore, J. A., & Loescher, W. H. (1994). Gas exchange and carbon partitioning in the leaves of celery (Apium graveolens L.) at various levels of root zone salinity. *Plant Physiology*, *106*(1), 281–292. https://doi.org/10.1104/pp.106.1.281

Flowers, T. J., & Colmer, T. D. (2008). Salinity tolerance in halophytes. *New Phytologist*, *179*(4), 945–963. https://doi.org/10.1111/j.1469-8137.2008.02531.x

Flowers, T. J., Galal, H. K., & Bromham, L. (2010). Evolution of halophytes: Multiple origins of salt tolerance in land plants. *Functional Plant Biology*, *37*(7), 604. https://doi.org/10.1071/FP09269

Flowers, T. J., Munns, R., & Colmer, T. D. (2015). Sodium chloride toxicity and the cellular basis of salt tolerance in halophytes. *Annals of Botany*, *115*(3), 419–431. https://doi.org/10.1093/aob/mcu217

Gale, J. (1975). Water balance and gas exchange of plants under saline conditions. In A. Poljakoff-Mayber & J. Gale (Orgs.), *Plants in Saline Environments* (Vol. 15, p. 168–185). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-80929-3_11

Gotz, S., Garcia-Gomez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., Robles, M., Talon, M., Dopazo, J., & Conesa, A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research*, *36*(10), 3420–3435. https://doi.org/10.1093/nar/gkn176

Gowda, H., Ivanisevic, J., Johnson, C. H., Kurczy, M. E., Benton, H. P., Rinehart, D., Nguyen, T., Ray, J., Kuehl, J., Arevalo, B., Westenskow, P. D., Wang, J., Arkin, A. P., Deutschbauer, A. M., Patti, G. J., & Siuzdak, G. (2014). Interactive xcms online: Simplifying advanced metabolomic data processing and subsequent statistical analyses. *Analytical Chemistry*, *86*(14), 6931–6939. https://doi.org/10.1021/ac500734c

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652. https://doi.org/10.1038/nbt.1883

He, M., & Ding, N. Z. (2020). Plant Unsaturated Fatty Acids: Multiple Roles in Stress Response. Frontiers in plant science, 11, 562785. https://doi.org/10.3389/fpls.2020.562785

Hou, Q., Ufer, G., & Bartels, D. (2016). Lipid signalling in plant responses to abiotic stress. Plant, cell & environment, 39(5), 1029–1048. https://doi.org/10.1111/pce.12666

Jamil, I. N., Remali, J., Azizan, K. A., Nor Muhammad, N. A., Arita, M., Goh, H.-H., & Aizat, W. M. (2020). Systematic multi-omics integration (MOI) approach in plant systems biology. *Frontiers in Plant Science*, *11*, 944. https://doi.org/10.3389/fpls.2020.00944

Kanehisa, M., Sato, Y., & Morishima, K. (2016). BlastKOALA and ghostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of Molecular Biology*, *428*(4), 726–731. https://doi.org/10.1016/j.jmb.2015.11.006

Khan, M. A., Ungar, I. A., & Showalter, A. M. (2000a). The effect of salinity on the growth, water status, and ion content of a leaf succulent perennial halophyte, Suaeda fruticosa (L.) Forssk. *Journal of Arid Environments*, *45*(1), 73–84. https://doi.org/10.1006/jare.1999.0617

Khan, M., Ungar, I. A., & Showalter, A. M. (2000b). Effects of salinity on growth, water relations and ion accumulation of the subtropical perennial halophyte, atriplex griffithii var. Stocksii. *Annals of Botany*, *85*(2), 225–232. https://doi.org/10.1006/anbo.1999.1022

Khorasaninejad, S., Mousavi, A., Soltanloo, H., Hemmati, K., & Khalighi, A. (2010). The effect of drought stress on growth parameters, essential oil yield and constituent of peppermint (Mentha piperita L.). *World Applied Sciences Journal 11 (11): 1403-1407*.

Koch, K., & Kennedy, R. A. (1980). Characteristics of crassulacean acid metabolism in the succulent C4 dicot, Portulaca oleracea L. *Plant Physiology*, *65*(2), 193–197. https://doi.org/10.1104/pp.65.2.193

Koyro, H.-W. (2006). Effect of salinity on growth, photosynthesis, water relations and solute composition of the potential cash crop halophyte Plantago coronopus (L.). *Environmental and Experimental Botany*, *56*(2), 136–146. https://doi.org/10.1016/j.envexpbot.2005.02.001

Kumari, A., Das, P., Parida, A. K., & Agarwal, P. K. (2015). Proteomics, metabolomics, and ionomics perspectives of salinity tolerance in halophytes. *Frontiers in Plant Science*, *6*. https://doi.org/10.3389/fpls.2015.00537

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Li, G., Wan, S., Zhou, J., Yang, Z., & Qin, P. (2010). Leaf chlorophyll fluorescence, hyperspectral reflectance, pigments content, malondialdehyde and proline accumulation responses of castor bean (Ricinus communis L.) seedlings to salt stress levels. *Industrial Crops and Products*, *31*(1), 13–19. https://doi.org/10.1016/j.indcrop.2009.07.015

Li, S., Park, Y., Duraisingham, S., Strobel, F. H., Khan, N., Soltow, Q. A., Jones, D. P., & Pulendran, B. (2013). Predicting network activity from high throughput metabolomics. *PLoS Computational Biology*, *9*(7), e1003123. https://doi.org/10.1371/journal.pcbi.1003123

Lowe, R., Shirley, N., Bleackley, M., Dolan, S., & Shafee, T. (2017). Transcriptomics technologies. PLoS computational biology, 13(5), e1005457. https://doi.org/10.1371/journal.pcbi.1005457

Ma, C., Wang, Y., Gu, D., Nan, J., Chen, S., & Li, H. (2017). Overexpression of S-Adenosyl-l-Methionine Synthetase 2 from Sugar Beet M14 Increased Arabidopsis Tolerance to Salt and Oxidative Stress. International journal of molecular sciences, 18(4), 847. https://doi.org/10.3390/ijms18040847

Mahajan, S., & Tuteja, N. (2005). Cold, salinity and drought stresses: An overview. *Archives of Biochemistry and Biophysics*, *444*(2), 139–158. https://doi.org/10.1016/j.abb.2005.10.018

Maksimovic, I., & Ilin, Z. (2012). Effects of salinity on vegetable growth and nutrients uptake. *Irrigation Systems and Practices in Challenging Environments, 9.*

Mane, A., Saratale, G., Karadge, B., & Samant, J. (2011). Studies on the effects of salinity on growth, polyphenol content andphotosynthetic response in Vetiveria zizanioides (L.) Nash. *Emirates Journal of Food and Agriculture*, *23*(1), 59. https://doi.org/10.9755/ejfa.v23i1.5313

Megdiche, W., Hessini, K., Gharbi, F., Jaleel, C. A., Ksouri, R., & Abdelly, C. (2008). Photosynthesis and photosystem 2 efficiency of two salt-adapted halophytic seashore Cakile maritima ecotypes. *Photosynthetica*, *46*(3), 410–419. https://doi.org/10.1007/s11099-008-0073-1

Munns, R. (2002). Comparative physiology of salt and water stress. *Plant, Cell & Environment*, *25*(2), 239–250. https://doi.org/10.1046/j.0016-8025.2001.00808.x

Munns, R., & Tester, M. (2008). Mechanisms of salinity tolerance. *Annual Review of Plant Biology*, *59*(1), 651–681. https://doi.org/10.1146/annurev.arplant.59.032607.092911

Murashige, T., & Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiologia Plantarum*, *15*(3), 473–497. https://doi.org/10.1111/j.1399-3054.1962.tb08052.x

OmicsBox—Bioinformatics made easy. (2019). BioBam. https://www.biobam.com/omicsbox/

Parida, A. K., & Das, A. B. (2005). Salt tolerance and salinity effects on plants: A review. *Ecotoxicology and Environmental Safety*, *60*(3), 324–349. https://doi.org/10.1016/j.ecoenv.2004.06.010

Parihar, P., Singh, S., Singh, R., Singh, V. P., & Prasad, S. M. (2015). Effect of salinity stress on plants and its tolerance strategies: A review. *Environmental Science and Pollution Research*, *22*(6), 4056–4075. https://doi.org/10.1007/s11356-014-3739-1

Petropoulos, S. A., Karkanis, A., Fernandes, Â., Barros, L., Ferreira, I. C. F. R., Ntatsi, G., Petrotos, K., Lykas, C., & Khah, E. (2015). Chemical Composition and Yield of Six Genotypes of Common Purslane (*Portulaca oleracea* L.): An Alternative Source of Omega-3 Fatty Acids. *Plant Foods Hum Nutr* **70,** 420–426. https://doi.org/10.1007/s11130-015-0511-8

Petropoulos, S. A., Karkanis, A., Martins, N., & Ferreira, I. C. F. R. (2016). Phytochemical composition and bioactive compounds of common purslane (*Portulaca oleracea* L.) as affected by crop management practices. *Trends in Food Science & Technology,* 55, 1-10. https://doi.org/10.1016/j.tifs.2016.06.010

Polle, A., & Chen, S. (2015). On the salty side of life: Molecular, physiological and anatomical adaptation and acclimation of trees to extreme habitats: Salt tolerance in trees. *Plant, Cell & Environment*, *38*(9), 1794–1816. https://doi.org/10.1111/pce.12440

Qadir, M., Quillérou, E., Nangia, V., Murtaza, G., Singh, M., Thomas, R. J., Drechsel, P., & Noble, A. D. (2014). Economics of salt-induced land degradation and restoration. *Natural Resources Forum*, *38*(4), 282–295. https://doi.org/10.1111/1477-8947.12054

Qiu, N., Lu, Q., & Lu, C. (2003). Photosynthesis, photosystem II efficiency and the xanthophyll cycle in the salt-adapted halophyte *Atriplex centralasiatica*. *New Phytologist*, *159*(2), 479–486. https://doi.org/10.1046/j.1469-8137.2003.00825.x

Rabhi, M., Castagna, A., Remorini, D., Scattino, C., Smaoui, A., Ranieri, A., & Abdelly, C. (2012). Photosynthetic responses to salinity in two obligate halophytes: Sesuvium portulacastrum and Tecticornia indica. *South African Journal of Botany*, *79*, 39–47. https://doi.org/10.1016/j.sajb.2011.11.007

Ramadan, T. (1998). Ecophysiology of salt excretion in the xero-halophyte Reaumuria hirtella. *New Phytologist*, *139*(2), 273–281. https://doi.org/10.1046/j.1469-8137.1998.00159.x

Ramoliya, P. J., Patel, H. M., & Pandey, A. N. (2004). Effect of salinization of soil on growth and macro- and micro-nutrient accumulation in seedlings of Salvadora persica (Salvadoraceae). *Forest Ecology and Management*, *202*(1–3), 181–193. https://doi.org/10.1016/j.foreco.2004.07.020

Rao, G. G., Nayak, A. K., Chinchmalatpure, A. R., Nath, A., & Babu, V. R. (2004). Growth and yield of Salvadora persica, a facultative halophyte grown on saline black soil (Vertic haplustept). *Arid Land Research and Management*, *18*(1), 51–61. https://doi.org/10.1080/15324980490245013

Rewald, B., Shelef, O., Ephrath, J. E., & Rachmilevitch, S. (2013). Adaptive plasticity of salt-stressed root systems. In P. Ahmad, M. M. Azooz, & M. N. V. Prasad (Orgs.), *Ecophysiology and Responses of Plants under Salt Stress* (p. 169–201). Springer New York. https://doi.org/10.1007/978-1-4614-4747-4_6

Rhoades, J.D., Chanduvi, F., Lesch, S., 1999. Soil salinity assessment: methods and interpretation of electrical conductivity measurements. FAO Irrigation and Drainage Paper #57. Food and Agriculture Organization of the United Nations, Rome, Italy, pp. 1–150

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

Rozema, J., Gude, H., & Pollak, G. (1981). An ecophysiological study of the salt secretion of four halophytes. *New Phytologist*, *89*(2), 201–217. https://doi.org/10.1111/j.1469-8137.1981.tb07483.x

Santos, J., Al-Azzawi, M., Aronson, J., & Flowers, T. J. (2016). eHALOPH a database of salt-tolerant plants: Helping put halophytes to work. *Plant and Cell Physiology*, *57*(1), e10–e10. https://doi.org/10.1093/pcp/pcv155

Sdouga, D., Ben Amor, F., Ghribi, S., Kabtni, S., Tebini, M., Branca, F., Trifi-Farah, N., & Marghali, S. (2019). An insight from tolerance to salinity stress in halophyte Portulaca oleracea L.: Physio-morphological, biochemical and molecular responses. *Ecotoxicology and Environmental Safety*, *172*, 45–52. https://doi.org/10.1016/j.ecoenv.2018.12.082

Shabala, S., & Mackay, A. (2011). Ion transport in halophytes. In *Advances in Botanical Research* (Vol. 57, p. 151–199). Elsevier. https://doi.org/10.1016/B978-0-12-387692-8.00005-9

Shahid, S. A., Zaman, M., & Heng, L. (2018). Soil salinity: Historical perspectives and a world overview of the problem. In M. Zaman, S. A. Shahid, & L. Heng, *Guideline for Salinity Assessment, Mitigation and Adaptation Using Nuclear and Related Techniques* (p. 43–53). Springer International Publishing. https://doi.org/10.1007/978-3-319-96190-3_2

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, *102*(43), 15545–15550. https://doi.org/10.1073/pnas.0506580102

Sui, N., Tian, S., Wang, W., Wang, M., & Fan, H. (2017). Overexpression of Glycerol-3-Phosphate Acyltransferase from *Suaeda salsa* Improves Salt Tolerance in Arabidopsis. Frontiers in plant science, 8, 1337. https://doi.org/10.3389/fpls.2017.01337

Sunkar, R., Chinnusamy, V., Zhu, J., & Zhu, J. K. (2007). Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. Trends in plant science, 12(7), 301–309. https://doi.org/10.1016/j.tplants.2007.05.001

Tada, Y., Kawano, R., Komatsubara, S., Nishimura, H., Katsuhara, M., Ozaki, S., Terashima, S., Yano, K., Endo, C., Sato, M., Okamoto, M., Sawada, Y., Hirai, M. Y., & Kurusu, T. (2019). Functional screening of salt tolerance genes from a halophyte Sporobolus virginicus and transcriptomic and metabolomic analysis of salt tolerant plants expressing glycine-rich RNA-binding protein. *Plant Science*, *278*, 54–63. https://doi.org/10.1016/j.plantsci.2018.10.019

Tautenhahn, R., Patti, G. J., Rinehart, D., & Siuzdak, G. (2012). XCMS Online: A web-based platform to process untargeted metabolomic data. *Analytical Chemistry*, *84*(11), 5035–5039. https://doi.org/10.1021/ac300698c

van Zelm, E., Zhang, Y., & Testerink, C. (2020). Salt tolerance mechanisms of plants. *Annual Review of Plant Biology*, *71*(1), 403–433. https://doi.org/10.1146/annurev-arplant-050718-100005

Villate, A., San Nicolas, M., Gallastegi, M., Aulas, P. A., Olivares, M., Usobiaga, A., Etxebarria, N., & Aizpurua-Olaizola, O. (2021). Review: Metabolomics as a prediction tool for plants performance under environmental stress. Plant science : an international journal of experimental plant biology, 303, 110789. https://doi.org/10.1016/j.plantsci.2020.110789

Visconti, F., & de Paz, J. M. (2016). *Electrical conductivity measurements in agriculture: The assessment of soil salinity*. IntechOpen. https://doi.org/10.5772/62741

Wang, Y., Huang, L., Du, F., Wang, J., Zhao, X., Li, Z., Wang, W., Xu, J., & Fu, B. (2021). Comparative transcriptome and metabolome profiling reveal molecular mechanisms underlying OsDRAP1-mediated salt tolerance in rice. *Scientific Reports*, *11*(1), 5166. https://doi.org/10.1038/s41598-021-84638-3

Xia, Z., Zhou, X., Li, J., Li, L., Ma, Y., Wu, Y., Huang, Z., Li, X., Xu, P., & Xue, M. (2019). Multiple-Omics Techniques Reveal the Role of Glycerophospholipid Metabolic Pathway in the Response of *Saccharomyces cerevisiae* Against Hypoxic Stress. Frontiers in microbiology, 10, 1398. https://doi.org/10.3389/fmicb.2019.01398

Xing, J. C., Zhao, B. Q., Dong, J., Liu, C., Wen, Z. G., Zhu, X. M., Ding, H. R., He, T. T., Yang, H., Wang, M. W., & Hong, L. Z. (2020). Transcriptome and metabolome profiles revealed

molecular mechanisms underlying tolerance of Portulaca oleracea to saline stress. *Russian Journal of Plant Physiology*, *67*(1), 146–152. https://doi.org/10.1134/S1021443720010240

Xing, J.-C., Dong, J., Wang, M.-W., Liu, C., Zhao, B.-Q., Wen, Z.-G., Zhu, X.-M., Ding, H.-R., Zhao, X.-H., & Hong, L.-Z. (2019). Effects of NaCl stress on growth of Portulaca oleracea and underlying mechanisms. *Brazilian Journal of Botany*, *42*(2), 217–226. https://doi.org/10.1007/s40415-019-00526-1

Yuan, F., Lyu, M.-J. A., Leng, B.-Y., Zhu, X.-G., & Wang, B.-S. (2016). The transcriptome of NaCl-treated Limonium bicolor leaves reveals the genes controlling salt secretion of salt gland. *Plant Molecular Biology*, *91*(3), 241–256. https://doi.org/10.1007/s11103-016-0460-0

Zaman, S., Bilal, M., Du, H., & Che, S. (2020). Morphophysiological and comparative metabolic profiling of purslane genotypes (Portulaca oleracea L.) under salt stress. *BioMed Research International*, *2020*, 1–17. https://doi.org/10.1155/2020/4827045

Zhang, H., Han, B., Wang, T., Chen, S., Li, H., Zhang, Y., & Dai, S. (2012). Mechanisms of plant salt response: Insights from proteomics. *Journal of Proteome Research*, *11*(1), 49–67. https://doi.org/10.1021/pr200861w

Zhang, Z., Mao, C., Shi, Z., & Kou, X. (2017). The Amino Acid Metabolic and Carbohydrate Metabolic Pathway Play Important Roles during Salt-Stress Response in Tomato. Frontiers in plant science, 8, 1231. https://doi.org/10.3389/fpls.2017.01231

Zhou, Y.-X., Xin, H.-L., Rahman, K., Wang, S.-J., Peng, C., & Zhang, H. (2015). *Portulaca oleracea* L.: A review of phytochemistry and pharmacological effects. *BioMed Research International*, *2015*, 1–11. https://doi.org/10.1155/2015/925631

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

The authors acknowledge funding to V.N.B.S., T.L.C.S., and J.C.R.N. by the Coordination for the Improvement of Higher Education Personnel (CAPES), via the Graduate Program in Plant Biotechnology at the Federal University of Lavras (UFLA) and the Graduate Program in Chemistry at the Federal University of Goiás (UFG).

Funding

The authors disclose receipt of the following financial support for the research, authorship, and/or publication of this article: the grant (01.13.0315.00 - DendePalm Project) for this study was awarded by the Brazilian Ministry of Science, Technology, and Innovation (MCTI) via the Brazilian Research and Innovation Agency (FINEP). The authors confirm that the funder had no influence over the study design, the content of article, or selection of this journal.

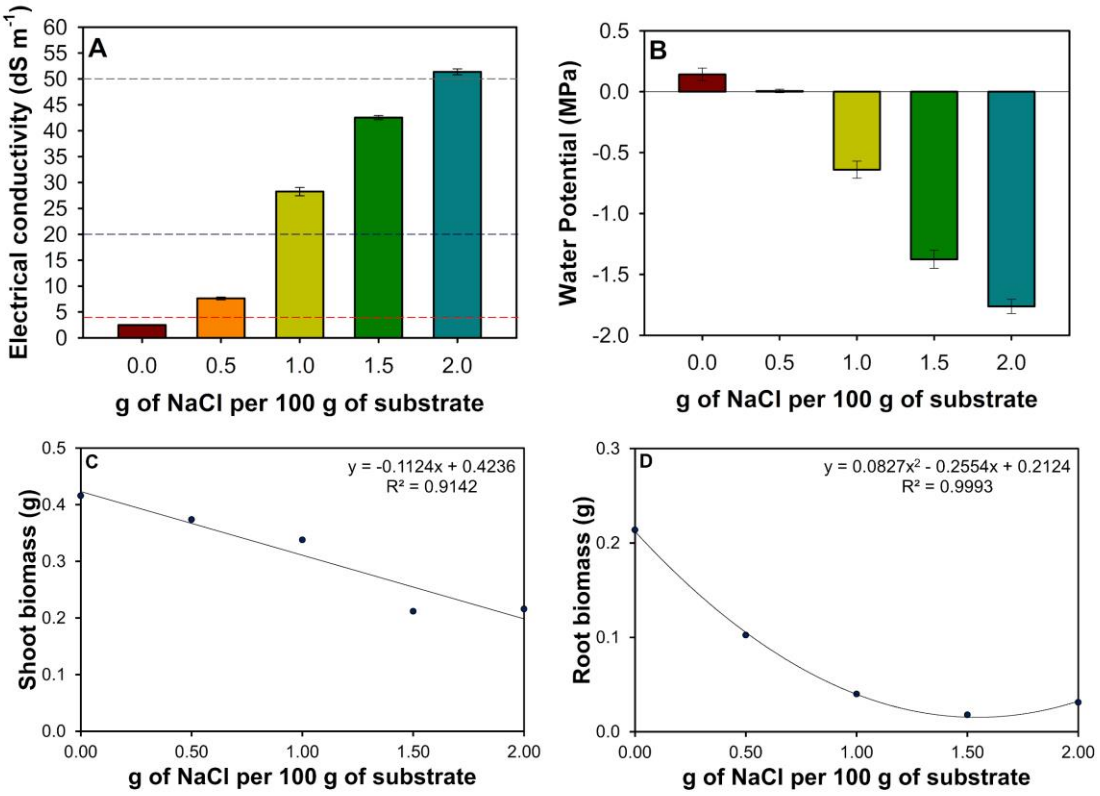


Figure 1. Electrical conductivity (A) and water potential (B) of the substrate used for growing purslane plants to which different levels of NaCl have been added.. Biomass accumulation in shoots (C) and roots (D) of young purslane plants grown for five days under different concentrations of NaCl. The values represent the average of five replicates, and the bars represent the standard error of the mean. Red dashed line: EC = 4 dS/m (above that level is considered saline soil); Blue dashed line: EC = 20 dS/m (plant completing its life cycle above this EC is considered as halophyte); and Gray dashed line: EC = 50 dS/m (seawater salinity).

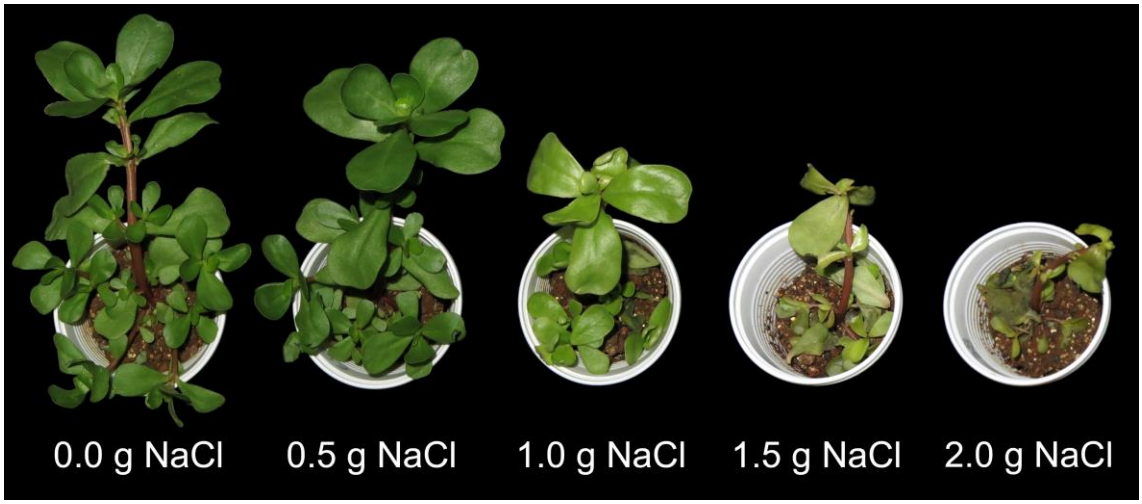


Figure 2. Young purslane plants grown for five days under different concentrations of NaCl.

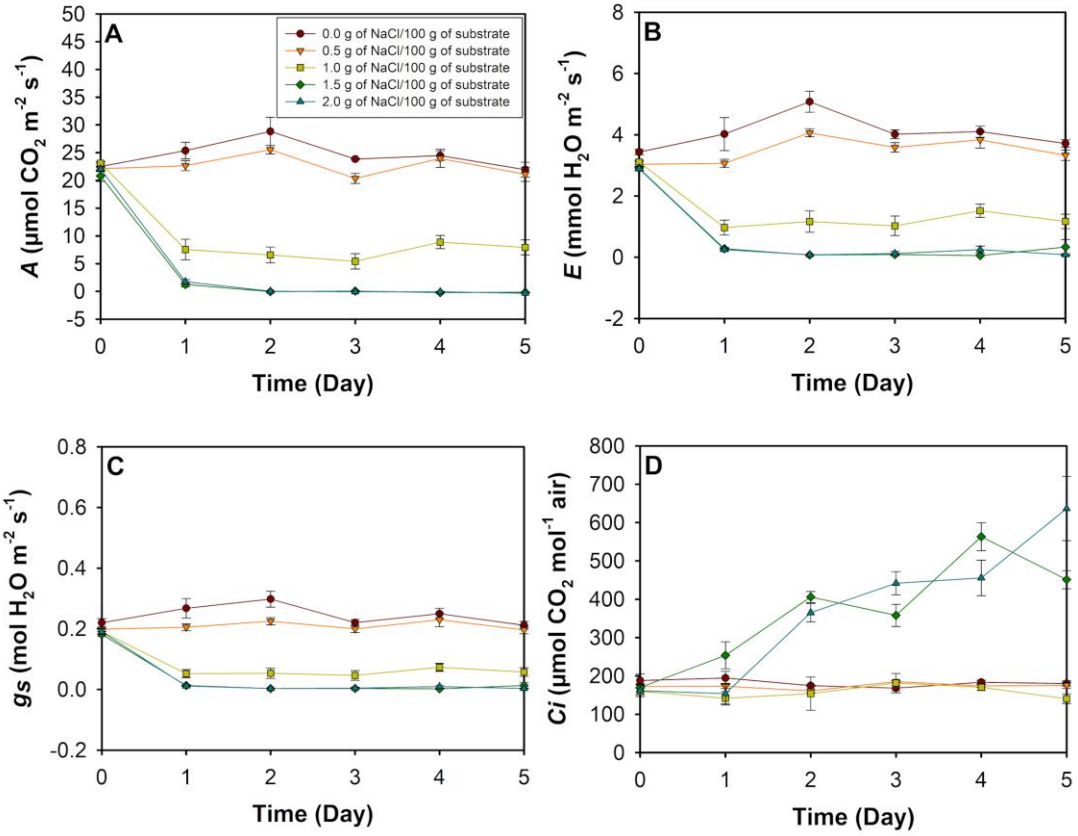


Figure 3. Leaf gas exchange in young purslane plants grown under increasing concentrations of NaCl in the substrate: (A) Net CO_2 assimilation rate (*A*); (B) stomatal conductance rate to water vapor (*gs*); (C) transpiration rate (*E*); (D) intercellular CO_2 concentration (*Ci*). The values represent the average of five replicates, and the bars represent the standard error of the mean.

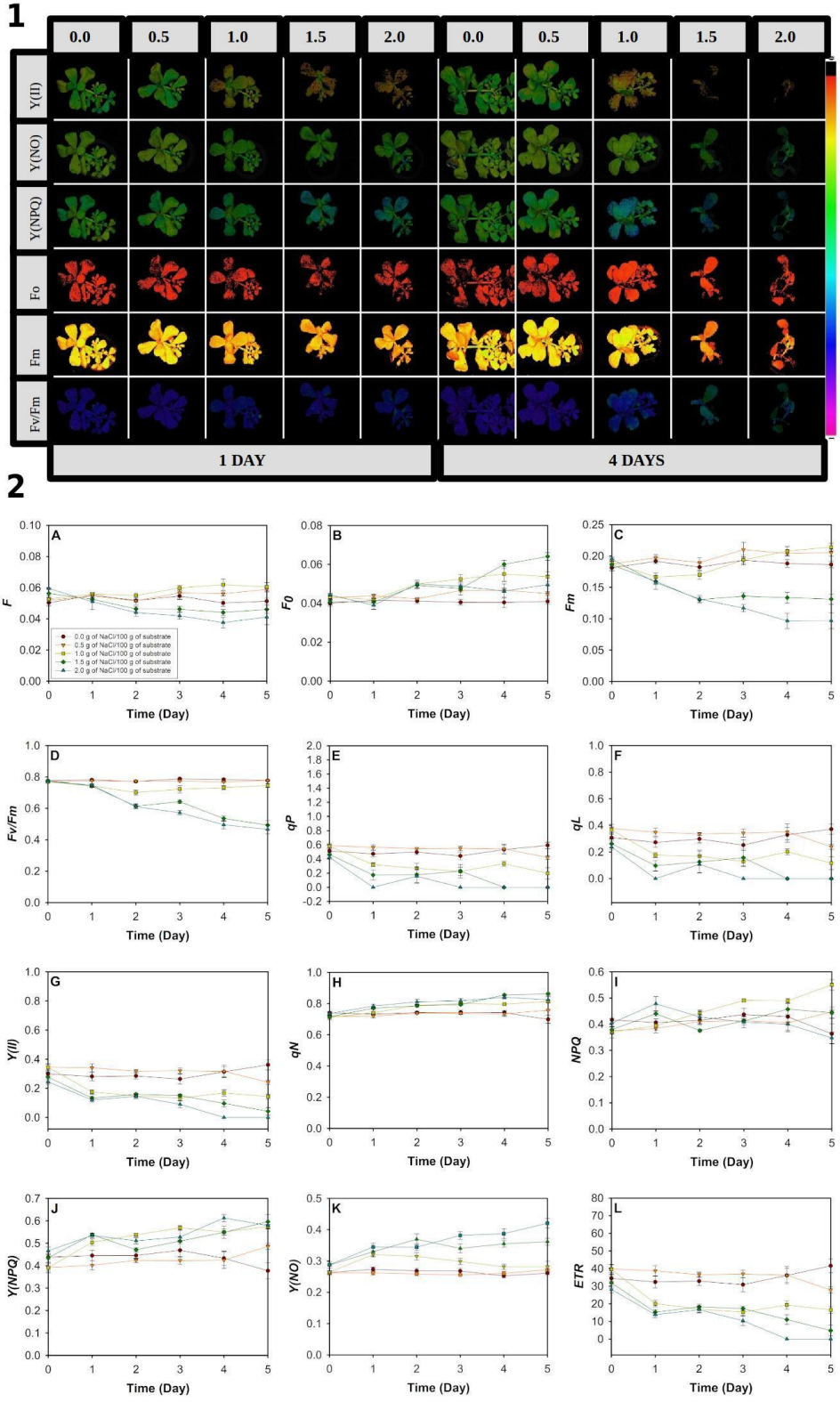

1
0.0
0.5
1.0
1.5
2.0
0.0
0.5
1.0
1.5
2.0
Y(II)
Y(NO)
Y(NPQ)
Fo
Fm
Fv/Fm
1 DAY
4 DAYS
2
A
B
C
D
E
F
G
H
I
J
K
L
F
Fo
Fm
Fv/Fm
qP
qL
Y(II)
qN
NPQ
Y(NPQ)
Y(NO)
ETR
Time (Day)
0.0 g of NaCl/100 g of substrate
0.5 g of NaCl/100 g of substrate
1.0 g of NaCl/100 g of substrate
1.5 g of NaCl/100 g of substrate
2.0 g of NaCl/100 g of substrate

Figure 4. Representative images (1) of the mean of variables derived from the chlorophyll fluorescence technique (saturation pulse method) in purslane plants after 1 and 4 days of submission to increasing levels of NaCl in the growing substrate. The images of Fo, Fv/Fm, and Fm' were obtained in plants adapted to the dark for 30 minutes, while the images of Y(II), Y(NO) and Y(NPQ) were captured after 5 minutes of actinic illumination at 280 μmol of light m^{-2} s^{-1}. The values of the chlorophyll fluorescence parameters in the images can be compared with the color scale in the right bar. Changes over time in chlorophyll fluorescence parameters (2) for control and salinity stressed young purslane plants. The values represent an average of five replicates, and bars represent the standard error of the mean. Fo: minimum fluorescence yield on dark-adapted plants; Fm: maximum fluorescence yield on the dark-adapted plants; and Fv/Fm: maximum PSII quantum yield; Y(II): PSII effective quantum yield; Y(NPQ): regulated energy dissipation quantum yield; Y(NO): unregulated energy dissipation quantum yield. The images of Fo, Fv/Fm, and Fm were obtained in plants adapted to the dark, while the images of Y(II), Y(NO) and Y(NPQ) were captured after 5 minutes of actinic illumination at 280 μmol of light/m^2/s.

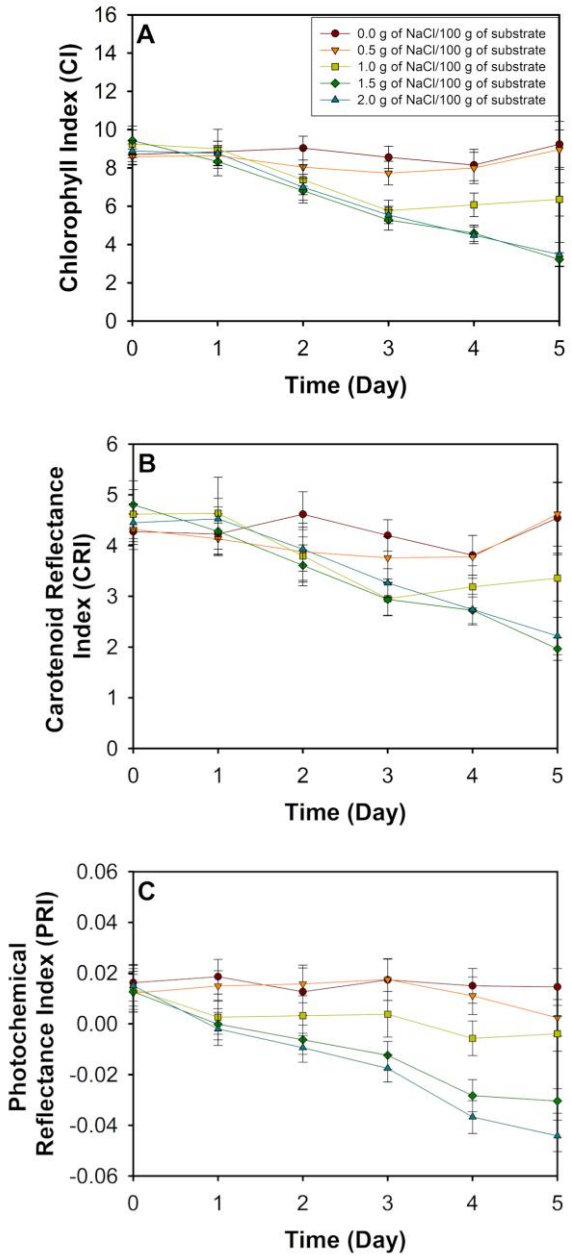


Figure 5. Chlorophyll index (CI) - A, carotenoid index (CRI) – B, and photochemical reflectance index (PRI) – C, for control and salinity stressed young purslane plants (0.0, 0.5, 1.0, 1.5, 2.0 g NaCl / 100 g of the substrate) throughout the period of 5 days of stress. The values represent an average of five replicates, and bars represent the standard error of the mean.

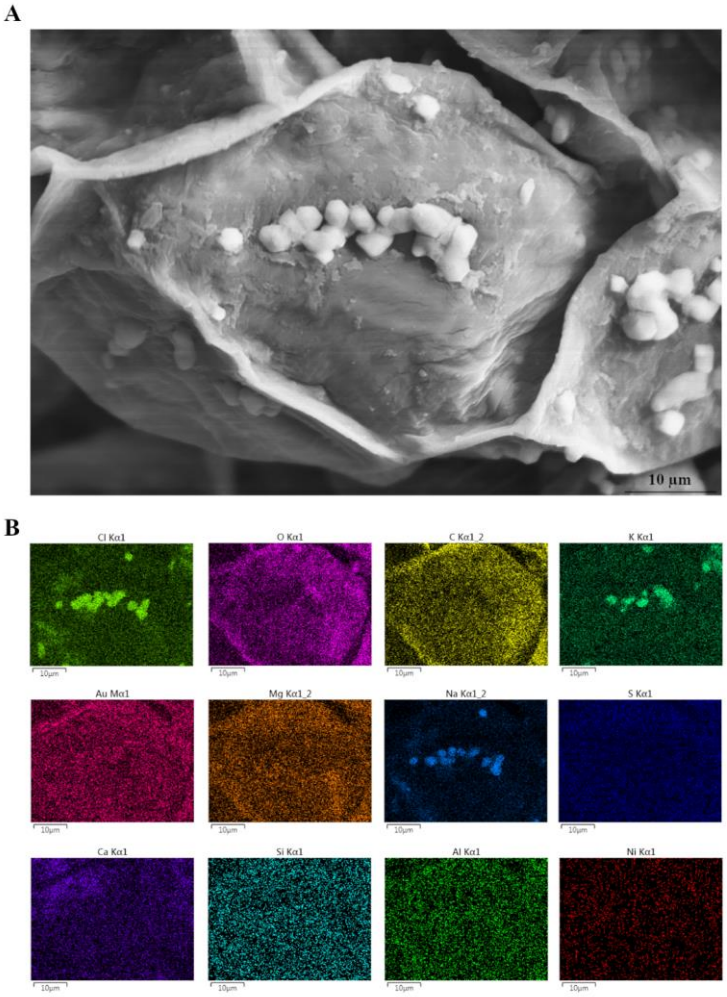


Figure 6. Compositional map of the elements and image obtained by scanning electron microscope (SEM) with detectors of energy dispersive spectroscopy (EDS) of a stoma on the leaf of a salt-stressed purslane plant. Formation of salt crystals above and around a closed stoma in salt-stressed plants (A). Compositional map of the elements showing that the crystals are made of Cl^-, K^+, and Na^+ (B).

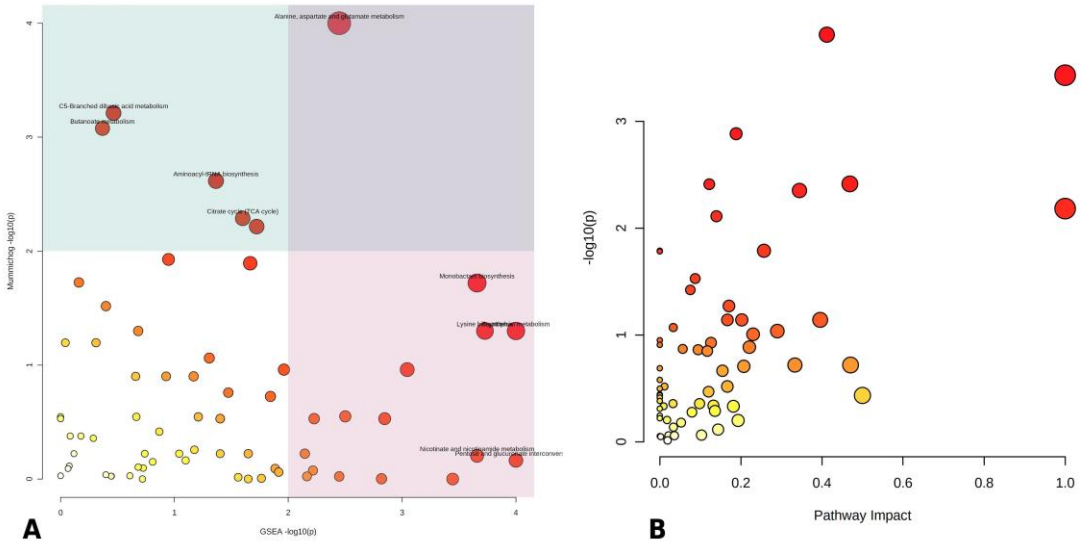


Figure 7. Summary of Pathway Analysis using the MS Peaks to Pathway and the Pathway Topology Analysis modules of MetaboAnalyst 5.0. The Integrated MS Peaks to Paths plot (A) summarizes the results of the Fisher's method for combining mummichog (y-axis) and Gene Set Enrichment Analysis - GSEA (x-axis) p-values. The metabolome view (B) resulted from the analysis in the Pathway Topology Analysis module using the Hypergeometric test, the relative betweenness centrality node importance measure, and the latest KEGG version of the *A. thaliana* pathway library.

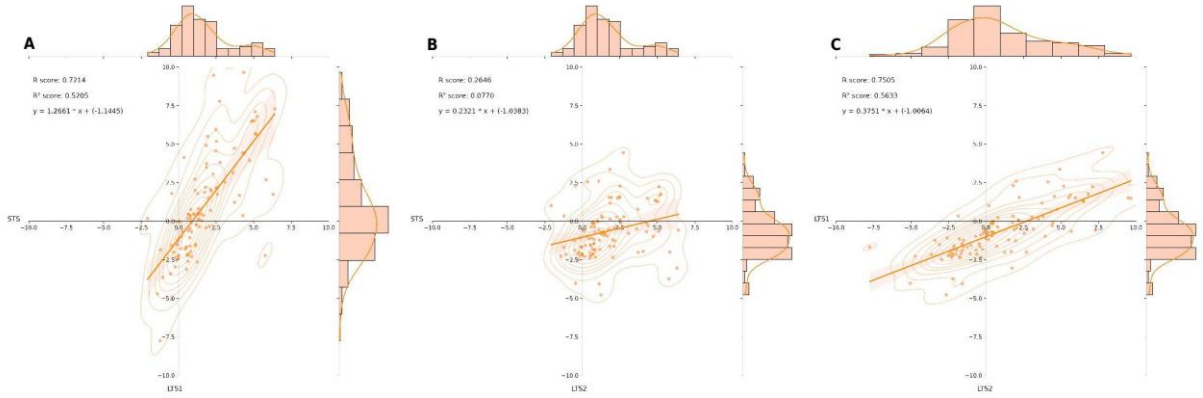


Figure 8. Histogram and correlation analysis of the Log_2 (Fold Change) of 109 differentially expressed metabolites by pairwise comparison of three scenarios: short-term stress – STS (control vs stress plants at 1 DAT); long-term stress 1 – LTS1 (control vs stress plants at 4 DAT); and long-term stress 2 – LTS2 (stressed plants at 1 and 4 DAT). A – STS vs LTS1, B – STS vs LTS2, and C – LTS1 vs LTS2.

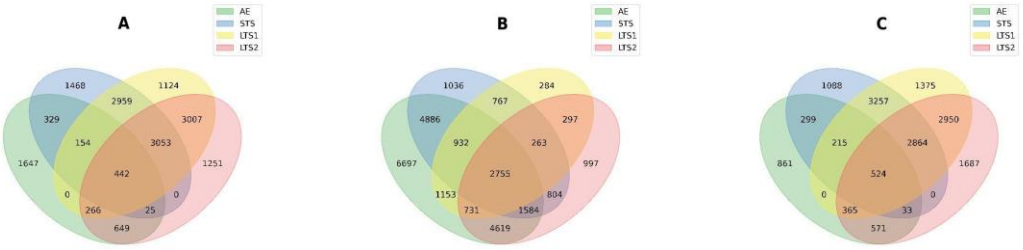


Figure 9. Venn Diagram of genes from *Portulaca oleracea* differentially expressed in the leaves of young purslane plants submitted to salt stress, under four distinct scenarios: age effect - AE (control plants at 1 and 4 days under salinity stress – DAT); short-term stress – STS (control vs stress plants at 1 DAT); long-term stress 1 – LTS1 (control vs stress plants at 4 DAT); and long-term stress 2 – LTS2 (stressed plants at 1 and 4 DAT). A – upregulated genes, B – non-differentially expressed genes, and C – downregulated genes.

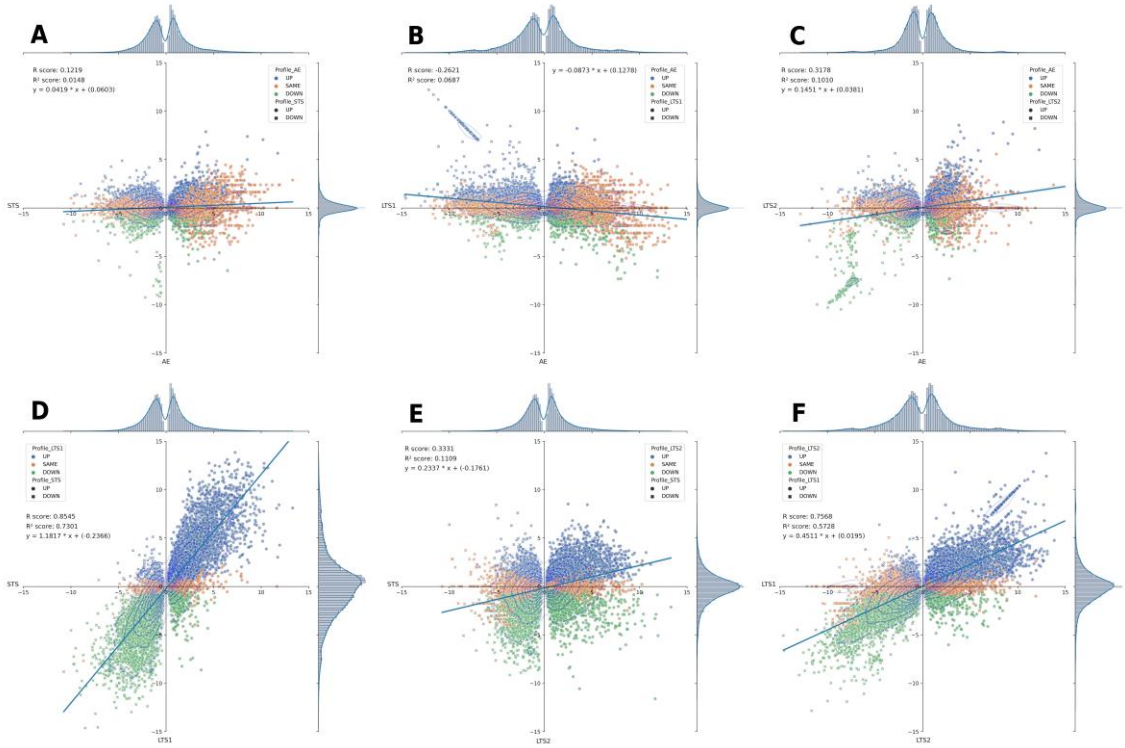
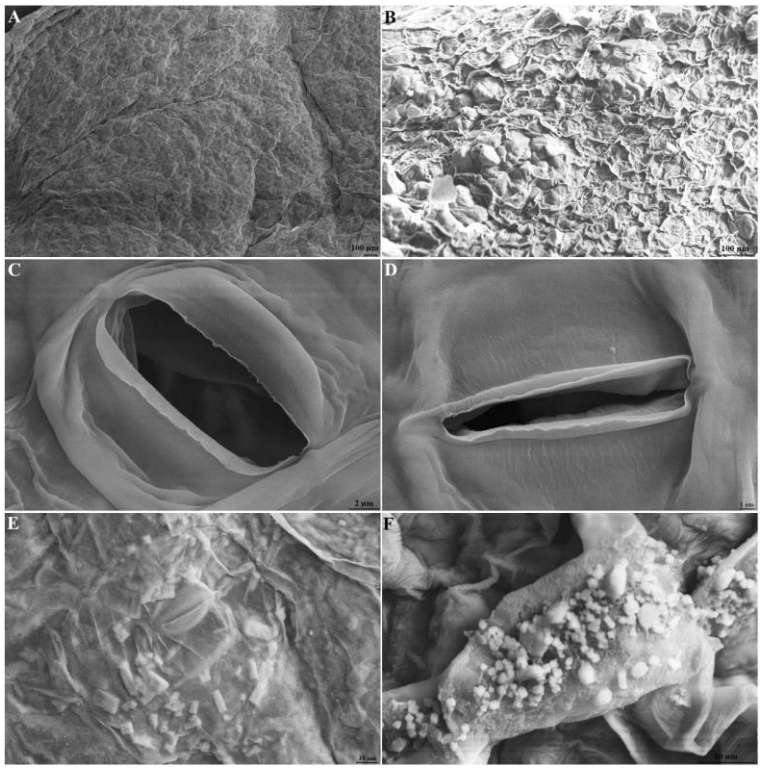


Figure 10. Histogram and correlation analysis of the Log_2 (Fold Change) of differentially expressed genes by pairwise comparison of four scenarios: age effect - AE (control plants at 1 and 4 days under salinity stress – DAT); short-term stress – STS (control vs stress plants at 1 DAT); long-term stress 1 – LTS1 (control vs stress plants at 4 DAT); and long-term stress 2 – LTS2 (stressed plants at 1 and 4 DAT). A – STS vs AE, B – LTS1 vs AE, C – LTS2 vs AE, D – STS vs LTS1, E – STS vs LTS2, F – LTS1 vs LTS2.



Supplementary Figure 1. Scanning electron microscope (SEM) images of the leaf surface of purslane plants. A) Control; B) stressed plant (2.0 g NaCl) showing wrinkle due to dehydration; (C) degree of stomata opening in control compared to (D) stressed (0.5 g NaCl). Formation of salt crystals around (E) and above (F) a closed stomata in stressed plants.

List of Tables

Considerações Finais

Com base nos estudos realizados no Programa de PD&I "Sal da Terra", foi observado que as espécies vegetais *G. sepium* e *P. oleracea* são altamente tolerantes à salinidade e que possuem mecanismos de tolerância distintos e notáveis (fenótipo de adaptação em gliricídia e liberação de cristais de sal nas folhas em beldroega). Dessa maneira, com base no robusto banco de dados gerado, uma análise mais profunda desses dados se mostrou crucial para tentar compreender os mecanismos que tornam a gliricídia e a beldroega tolerantes ao sal, sendo o objetivo principal deste trabalho.

Os resultados obtidos permitiram alcançar esse objetivo, revelando genes/transcritos, metabólitos e vias metabólicas diferencialmente expressas e com importância para a tolerância das espécies estudadas à condição salina. Neste estudo foi possível observar as diferenças entre analisar somente uma ômica e aplicar a estratégia MOI; novas vias metabólicas emergiram quando se aplicou a estratégia MOI e, também, foi possível verificar a diferença entre os cenários analisados (AE, STS e LTS).

Mas, não foi somente a estratégia de integração entre as ômicas que proporcionou conhecimentos interessantes; devido ao fato de que a estratégia de integração utilizada, visando mapear as vias metabólicas, leva em consideração apenas as enzimas e metabólitos, outros transcritos (não enzimáticos) que poderiam ter um papel importante no aspecto da tolerância destas espécies não são analisados. Dessa maneira, quando analisamos individualmente a transcritômica, foram observados transcritos que não codificavam enzimas e que tiveram sua expressão diferenciada, significativamente, duas vezes para cima (*up regulated*) ou duas vezes para baixo (*down regulated*); mostrando que a análise individual das ômicas continua sendo importante.

Durante o exercício de integração, viu-se que a estratégia MOI nível 2, de mapeamento em vias, se mostrou ser realmente mais fácil de realizar e analisar, conforme discutido anteriormente (Capítulo 1, Tópico 1.3.3), sendo uma porta de entrada para a aplicação de estratégias MOI neste grupo de pesquisa. Outro ponto importante foi que, conforme apresentado no tópico 1.3 (Capítulo 1), as estratégias MOI têm como características gerar hipóteses e não testá-las; sendo que, no presente estudo, as listas de transcritos e metabólitos que possuem relação com a tolerância à salinidade, geradas para ambas as espécies estudadas, são ferramentas valiosas para gerar hipóteses e aprofundar o conhecimento sobre os mecanismos de tolerância ao sal.

Por fim, durante a execução do trabalho, notou-se que a estratégia empregada no estudo da transcritômica de beldroega, utilizando a anotação GhostKOALA, parece ser mais apropriada para o estudo de espécies não-modelo, devido ao extenso grupo de plantas presentes na base de dados, caracterizando melhor os transcritos estudados.