IDENTIFICAÇÃO E CARACTERIZAÇÃO COMPUTACIONAL DE lncRNAs EM RAIZ DE MILHO SUBMETIDO A BIOESTIMULANTES VEGETAIS

LEANDRO DE OLIVEIRA SILVA

UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE
DARCY RIBEIRO – UENF

CAMPOS DOS GOYTACAZES - RJ
MAIO - 2022

IDENTIFICAÇÃO E CARACTERIZAÇÃO COMPUTACIONAL DE lncRNAs EM RAIZ DE MILHO SUBMETIDO A BIOESTIMULANTES VEGETAIS

LEANDRO DE OLIVEIRA SILVA

"Tese apresentada ao Centro de Biociências e Biotecnologia da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Doutor em Biotecnologia Vegetal."

Orientadora: Prof.ª Clícia Grativol Gaspar de Matos

CAMPOS DOS GOYTACAZES - RJ
MAIO - 2022

FICHA CATALOGRÁFICA
UENF - Bibliotecas
Elaborada com os dados fornecidos pelo autor.

S586 Silva, Leandro de Oliveira.

Identificação e Caracterização computacional de lncRNAs em raiz de milho submetido a Bioestimulantes Vegetais / Leandro de Oliveira Silva. - Campos dos Goytacazes, RJ, 2022.

111 f. : il.
Inclui bibliografia.

Tese (Doutorado em Biotecnologia Vegetal) - Universidade Estadual do Norte Fluminense Darcy Ribeiro, Centro de Biociências e Biotecnologia, 2022.
Orientadora: Clícia Grativol Gaspar.

1. *Zea mays*. 2. microorganismos benéficos. 3. substâncias húmicas. 4. ncRNAs. 5. bactéria endofítica diazotrófica. I. Universidade Estadual do Norte Fluminense Darcy Ribeiro. II. Título.

CDD - 660.6

IDENTIFICAÇÃO E CARACTERIZAÇÃO COMPUTACIONAL DE lncRNAs EM RAIZ DE MILHO SUBMETIDO A BIOESTIMULANTES VEGETAIS

LEANDRO DE OLIVEIRA SILVA

"Tese apresentada ao Centro de Biociências e Biotecnologia da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Doutor em Biotecnologia Vegetal."

Aprovada em 31 de maio de 2022.

Comissão Examinadora:

Mariana Ramos Leandro

Dr.ª Mariana Ramos Leandro (D.Sc., Biotecnologia Vegetal) – UNICAMP

Kátia Valevski Sales Fernandes

Dr.ª Kátia Valevski Sales Fernandes (D.Sc., Bioquímica e Biologia Molecular) – UENF

Dr. Thiago Motta Venancio (D.Sc., Bioinformática) – UENF

Clícia Grativol Gaspar de Matos

Dr.ª Clícia Grativol Gaspar de Matos (D.Sc., Química Biológica) – UENF (Orientadora)

DEDICATÓRIA

Aos meus pais, Fátima e Luiz, que me deram a vida,
dedico.

"[…] To try when your arms are too weary
To reach the unreachable star […]"
("The Impossible Dream", Joe Darion and Mitchell Leigh)

AGRADECIMENTOS

Agradeço, primeiramente, a Deus, com quem “briguei” e “fiz as pazes” (centenas de vezes) ao longo dos anos (sobretudo os últimos). Obrigado por não me deixar desistir de mim.

À Prof[a]. Clícia Grativol, por ter acreditado em meu potencial e aberto as portas de seu laboratório para que eu realizasse este sonho (quase) impossível.

Ao Prof. Fábio Olivares, por ceder os transcriptomas utilizados neste estudo, essenciais para o desenvolvimento deste trabalho.

À Prof[a]. Maria Emilia M. T. Walter e seu doutorando, Lucas Maciel Vieira, da Universidade de Brasília, pela colaboração com o desenvolvimento do modelo de predição de lncRNAs de milho.

Ao Me. Fabrício Almeida, pela ajuda em algumas análises.

A todos os amigos do LQFPP, os que decidiram ficar e os que, por alguma razão, precisaram ir: Deise, Jacymara, Juliana, Maria Luíza e Wálaci. Em especial, agradeço à doutoranda Paula Machado por ouvir minhas reclamações e sempre tentar me ajudar (com comandos ou simplesmente destravando o computador); à Dr.[a] Fernanda Coelho, obrigado pela ajuda com o PC e por ouvir minhas reclamações; ao Me. Thiago Cézar de Pádua Rosa, obrigado por sempre estar disposto a me ouvir.

Aos professores: Dr.[a] Mariana Ramos Leandro, Dr.[a] Kátia Valevski e Dr. Thiago Venancio, por se disporem a avaliar este trabalho.

À minha segunda família, em Campos dos Goytacazes: Sara Sangi, Mariana Ramos, Júlia Moreira e Carolina Isidório, por dividirmos não apenas o espaço físico de uma casa, mas o espaço fraterno de um lar. Elis Nymeria é o pet que completa esse time.

À Alexandra Elbakyan, por lutar por uma ciência acessível a todos.

À Universidade Estadual do Norte Fluminense Darcy Ribeiro, pelo custeio da bolsa de doutorado e financiamento deste trabalho.

SUMÁRIO

LISTA DE FIGURAS

LISTA DE QUADROS

SIGLAS E ABREVIATURAS

A.E.C Antes da Era Comum

BPCV Bactérias promotoras do crescimento de plantas

circRNAs RNAs circulares

GO Gene Ontolgy

HC-lncRNAs LncRNAs conhecidos de alta confiança

iRF Iterative Random Forests

lincRNA Long intergenic non-coding RNA

lncRNA RNA longo não-codificante

miRNA MicroRNA

NC-lncRNA Novos lncRNAs candidatos

ncRNA RNA não-codificante

PCR Polymerase chain reaction

PH Protein hydrolysate

piRNA RNA que interage com PIWI

rRNA RNA ribossomal

RF Random forest

SVM Support Vector Machine

siRNA Pequeno RNA de interferência

snoRNA pequeno RNA nucleolar

SWE Seaweed extracts

tRNA RNA transportador

RESUMO

O milho é uma das mais importantes culturas, empregado principalmente na alimentação humana e animal e, em países como os Estados Unidos, largamente utilizado na produção de etanol. A despeito dos diversos estudos realizados com milho ao longo dos anos, seu transcriptoma complexo ainda representa um desafio ao entendimento das redes de regulação associadas à interação com organismos benéficos ou compostos orgânicos, como as substâncias húmicas. Nos últimos anos, RNAs não-codificantes têm sido apontados como elementos-chave na regulação da expressão gênica, incluindo os chamados RNAs longos não-codificantes (lncRNAs). Somente um trabalho descreveu a participação de lncRNAs em milho submetido a bioestimulante microbiano (fungo arbuscular micorrízico). Não foram identificados estudos na literatura descrevendo a participação de lncRNAs em milho colonizado por organismos benéficos, como a bactéria endofítica diazotrófica - *Herbaspirillum seropedicae* -, ou na adição de ácido húmico. Este trabalho teve como objetivo caracterizar, por meio de análises de bioinformática, lncRNAs expressos em milho submetido a bioestimulantes como bactéria promotora do crescimento vegetal - *Herbaspirillum seropedicae* - e ácido húmico. Para identificar os lncRNAs conhecidos, foi realizada uma busca em bases de dados e artigos. Para a identificação de novos lncRNAs, um transcriptoma foi montado a partir de bibliotecas de RNA-seq de milho colonizado pela bactéria e também com a adição de ácido húmico. Através de um pipeline computacional, identificamos um total de 32.501 lncRNAs conhecidos de alta confiança (HC-lncRNAs) e 6.418 novos lncRNAs candidatos (NC-lncRNA) nos transcriptomas. O perfil de expressão de lncRNAs revelou 3.263 HC-lncRNAs conhecidos e 1,050 novos lncRNAs diferencialmente expressos (DE) nas seis diferentes comparações. Também identificamos as redes regulatórias de DE-lncRNAs por meio da previsão de genes-alvo e miRNAs. Tanto os lncRNAs novos quanto os conhecidos foram investigados quanto a atuarem como alvos/*decoys* de microRNAs. Identificamos 20 lncRNAs atuando como alvos de 41 diferentes miRNAs, e 3 lncRNAs atuando como *decoys* de 11 miRNAs. Os resultados mostraram que os lncRNAs têm como alvo genes associados a diferentes processos, como transdução de sinal hormonal, resposta ao ácido abscísico (ABA) e ao etileno, ligação ao Ca^{2+}, resposta à auxina; ubiquitinação; composição da parede celular externa e sinalização transmembrana, entre outros. A obtenção de redes de interação lncRNA-mRNA-miRNA nos ajudou a esclarecer as funções dos lncRNAs a partir das funções de seus alvos, sugerindo o envolvimento desses importantes transcritos não-codificantes com a regulação de hormônios vegetais. Esperamos, com os resultados obtidos, esclarecer o papel dos lncRNAs no processo de interação do milho com organismos benéficos e substâncias orgânicas.

Palavras-chave: *Zea mays*; microorganismos benéficos; substâncias húmicas; ncRNAs; bactéria endofítica diazotrófica.

ABSTRACT

Maize is one of the most important crops, used mainly in human and animal food and, in countries like the United States, widely used in the production of ethanol. Despite the many studies carried out with maize over the years, its complex transcriptome still represents a challenge to the understanding of the regulatory networks associated with the beneficial organisms or organic compounds, such as humic substances. In recent years, non-coding RNAs have been identified as key elements in the regulation of gene expression, including the so-called long non-coding RNAs (lncRNAs). Only one study described the participation of lncRNAs in maize submitted to microbial biostimulant (arbuscular mycorrhizal fungus). Thus, this work aimed to characterize, through bioinformatics analysis, lncRNAs expressed in maize subjected to biostimulants such as plant growth-promoting bacterium - *Herbaspirillum seropedicae* - and humic acid. To identify and characterize maize lncRNAs, we used RNA-seq libraries constructed from inoculated and uninoculated seedling roots with *H. seropedicae* and humic acid. Through a computational pipeline, we identified a total of 32,501 known high-confidence lncRNAs (HC-lncRNAs) and 6,418 new candidate lncRNAs (NC-lncRNAs) in the transcriptomes. The lncRNA expression profile revealed 3,263 known HC-lncRNAs and 1,050 new lncRNAs differentially expressed (DE) in the six different comparisons. We also identify the regulatory networks of DE-lncRNAs by predicting target genes and miRNAs. Both new and known lncRNAs were investigated as to whether they act as targets/decoys of microRNAs. We identified 20 lncRNAs acting as targets of 41 different miRNAs, and 3 lncRNAs acting as decoys of 11 miRNAs. The results showed that lncRNAs target genes associated with different processes, such as hormonal signal transduction, abscisic acid (ABA) and ethylene response, Ca^{2+} binding, auxin response; ubiquitination; composition of the outer cell wall and transmembrane signaling, among others. Furthermore, lncRNAs-mRNAs-miRNAs interaction networks were constructed and showed the potential of lncRNAs as a key regulatory element. Our results showed that lncRNAs are involved in different processes during the interaction of maize with the bacteria *H. seropedicae* and in the presence of humic acid, including the participation of elements associated with different plant hormones. We hope, with the results obtained, to clarify the role of lncRNAs in the process of maize interaction with beneficial organisms and organic substances.

Keywords: *Zea mays*; beneficial microorganisms; humic substances; ncRNAs; diazotrophic endophytic bacteria.

1. INTRODUÇÃO

1.1 Milho: taxonomia, aspectos gerais, genéticos e econômicos

O milho (*Zea mays* L.) pertence à classe Liliopsida (Monocotyledonae); à família Poaceae (Gramineae) (ILTIS; DOEBLEY, 1980; STRABLE; SCANLON, 2009) e à tribo Andropogoneae. A família Poaceae inclui também a cana-de-açúcar (*Saccharum* spp.*)* e importantes cereais, como o sorgo (*Sorghum bicolor*), o trigo (*Triticum aestivum*), a cevada (*Hordeum vulgare*), a aveia (*Avena sativa*) e o arroz (*Oryza sativa*), como podemos observar na **Figura 1**.

As gramíneas se originaram entre 55 e 70 milhões de anos, e englobam todas as principais culturas de cereais, além de aproximadamente 10.000 outras espécies relacionadas não domesticadas (BOLOT et al., 2009; KELLOGG, 2001).

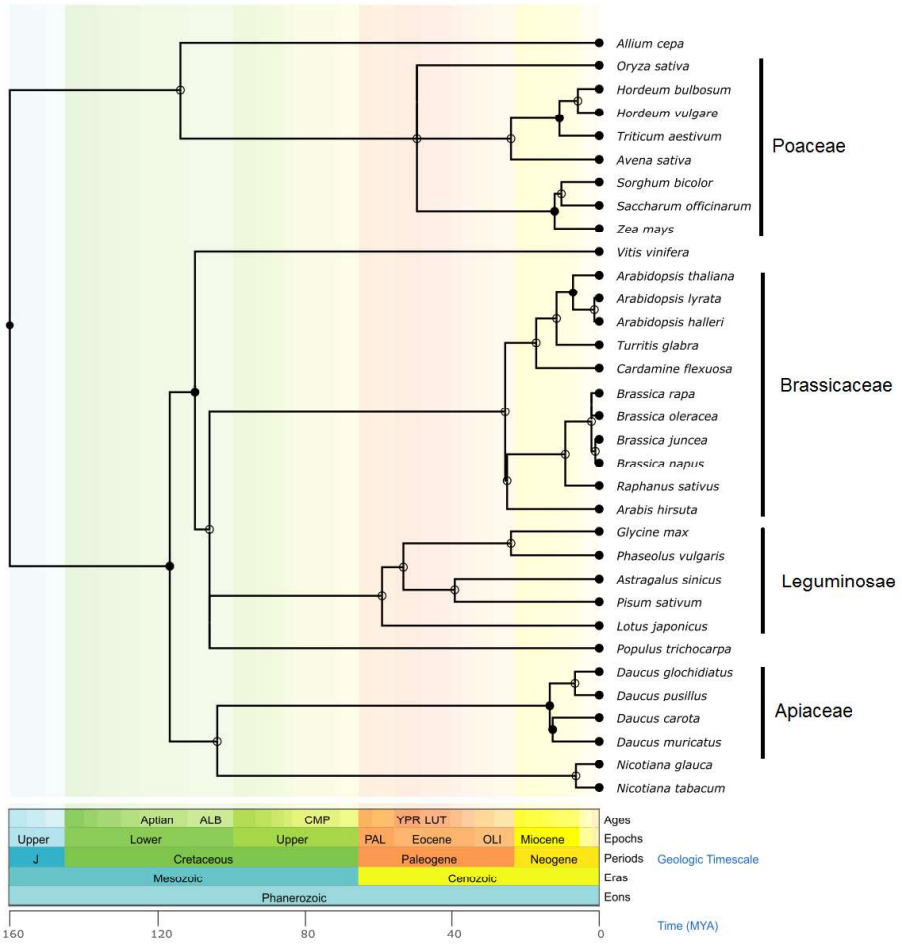


Figura 1 Filogenia mostrando o milho e algumas espécies próximas. **Fonte**: http://www.timetree.org/.

Para compreender a evolução genômica do milho é importante estudar sua domesticação (DOEBLEY, 2004). O milho tem como centro de origem a região Mesoamericana, e acredita-se que ele tenha se originado no México. Os registros arqueológicos e as análises filogenéticas mostram que sua domesticação aconteceu entre 6.000 e 10.000 anos A.E.C., a partir do teosinto, a espécie silvestre mais próxima do milho domesticado (HUFFORD et al., 2012; MATSUOKA et al., 2002; PIPERNO; FLANNERY, 2001).

O teosinto (*Zea mays* subsp. mexicana) é uma gramínea perene e anual do gênero *Zea*, nativo do México e da América Central (FUKUNAGA et al., 2005). O gênero *Zea* compreende quatro espécies: *Z. luxurians*, *Z. diploperennis*, *Z. perennis* e *Z. mays*, esta última compreendendo quatro subespécies: ssp. *mays*, ssp. *mexicana*, ssp. *parviglumis* e ssp. *huehuetenangensis* (FUKUNAGA et al., 2005; VOLLBRECHT; SIGMON, 2005). Postula-se que um único evento de domesticação aconteceu na região central do Rio Balsas, sul do México, com posterior diversificação do milho nas terras altas entre as províncias de Oaxaca e Jalisco (MATSUOKA et al., 2002) e, após a descoberta das Américas no século XV, o milho rapidamente se espalhou pelo mundo (GALINAT, 1988).

Um estudo mais recente sobre a domesticação e dispersão do milho sugere que o processo se iniciou no sudoeste do México, há 9.000 A.E.C., e sua dispersão pela América do Sul teria ocorrido há 7.000 A.E.C. (KISTLER et al., 2020). A América do Sul, por essa perspectiva, constituiu um centro secundário de domesticação do milho. Mais tarde, teria ocorrido um influxo do milho da América do Sul para a América Central.

O milho ancestral surgiu há 4,8 milhões de anos, através de um evento de alotetraploidização segmentar de dois genomas progenitores que divergiram do sorgo (GAUT; DOEBLEY, 1997; SWIGONˇOVA et al., 2004). Assim como em outras plantas, a poliploidização desempenhou um importante papel em sua evolução (MESSING, 2009; WENDEL, 2000). O genoma do milho moderno (N=10 cromossomos) tem 2,4 Gb e é 18 vezes o tamanho do genoma de *Arabidopsis thaliana*, 5 vezes o genoma do arroz (*O. sativa*) e 3 vezes o genoma do sorgo (*S. bicolor)* (GOFF et al., 2005; KAUL et al., 2000; PATERSON et al., 2009; SCHNABLE et al., 2009). Consiste, principalmente, de frações repetitivas não gênicas, pontuadas por ilhas com genes únicos ou

pequenos grupos de genes (LLACA; CAMPBELL; DESCHAMPS, 2011). Em sua versão 4 (APGv4), o genoma de milho possui 39.498 genes e 131.496 transcritos codificantes de proteínas anotados (JIAO et al., 2017). Na porção repetitiva do genoma foram identificados 130.000 elementos móveis.

O milho é usado como alimento humano, ração animal e matéria prima industrial em países desenvolvidos (TESTER; LANGRIDGE, 2010). Nos Estados Unidos, uma importante parcela do milho é utilizada para a produção de etanol (álcool etílico), empregado como combustível para motores e aditivo para a gasolina (ROSENTRATER, 2011). A produção global de milho alcançou 1,2 bilhões de toneladas, com os Estados Unidos ocupando a primeira posição, seguidos por China e Brasil. Estima-se um aumento de 7% da safra dos Estados Unidos (384 milhões de toneladas); 22,8% da China (273 milhões de toneladas) e aumento de 31% da safra brasileira (114 milhões de toneladas) (Global Trade, 2022).

1.2 LncRNAs: quem são, onde estão e o que fazem?

Análises de transcriptoma de eucariotos mostraram que mais de 90% do genoma é transcrito em RNAs não codificadores (ncRNAs, do inglês *non-coding RNA*) (CHEKANOVA et al., 2007; KAPRANOV et al., 2007). Os ncRNAS são transcritos de sequências de DNA, que não possuem ORFs funcionais (WILUSZ; SUNWOO; SPECTOR, 2009) e, durante muito tempo, foram ignorados pelos pesquisadores.

O desenvolvimento das novas tecnologias, aliado ao crescente número de projetos de sequenciamento, proporcionou um acúmulo de dados biológicos (BERNAL; EAR; KYRPIDES, 2001), possibilitando a descoberta de um número cada vez maior de ncRNAs de eucariotos (SABIN; DELÁS; HANNON, 2014), sendo alguns destes diretamente envolvidos com alterações epigenéticas e modificações celulares, dentre diversas funções. Em plantas, ncRNAs são classificados em ncRNAs lineares e RNAs circulares (circRNAs), conforme podemos observar na **Figura 2**. O subconjunto de ncRNAs lineares é constituído pelo RNA transportador (tRNA), RNA ribossomal (rRNA), pequeno RNA nucleolar (snoRNAs), microRNAs, pequenos RNAs de interferência

(siRNAs), RNAs que interagem com PIWI (piRNAs) e RNAs longos não codificadores (lncRNA) (SANTOSH; VARSHNEY; YADAVA, 2015).

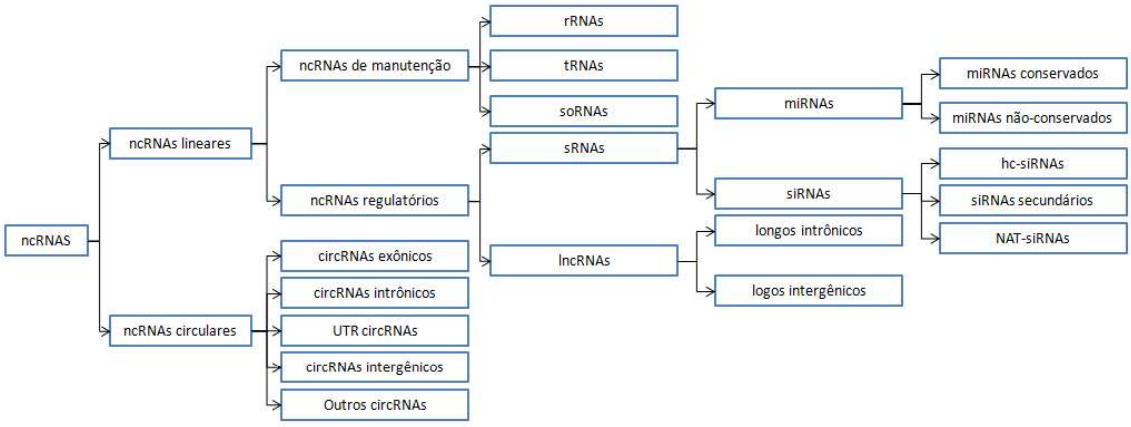


Figura 2 Classificação dos ncRNAs de plantas. **Fonte**: Adaptada de Liu et al. (2017).

Com base na função molecular, os ncRNAs lineares podem ser divididos em duas classes: ncRNAs de manutenção (rRNAs, tRNAs e snoRNAs) e ncRNAs reguladores, estes últimos classificados em duas subcategorias: pequenos RNAs (sRNAs) (que incluem os miRNAs e os siRNAs) e os lncRNAs, que incluem os lncRNAs intrônicos e os lncRNAs intergênicos (LIU et al., 2017).

Em plantas, uma classe tardiamente explorada de ncRNAs (quando comparada com animais) tem chamado a atenção dos pesquisadores: os lncRNAs. Os lncRNAs são transcritos com tamanho superior a 200 nucleotídeos e destituídos de potencial codificante (ou com potencial para codificar peptídeos com menos de 100 aminoácidos) (JIN et al., 2013; WANG et al., 2014a; ZHANG et al., 2014b) e podem ser classificados em seis principais tipos, como podemos observar na **Figura 3**.

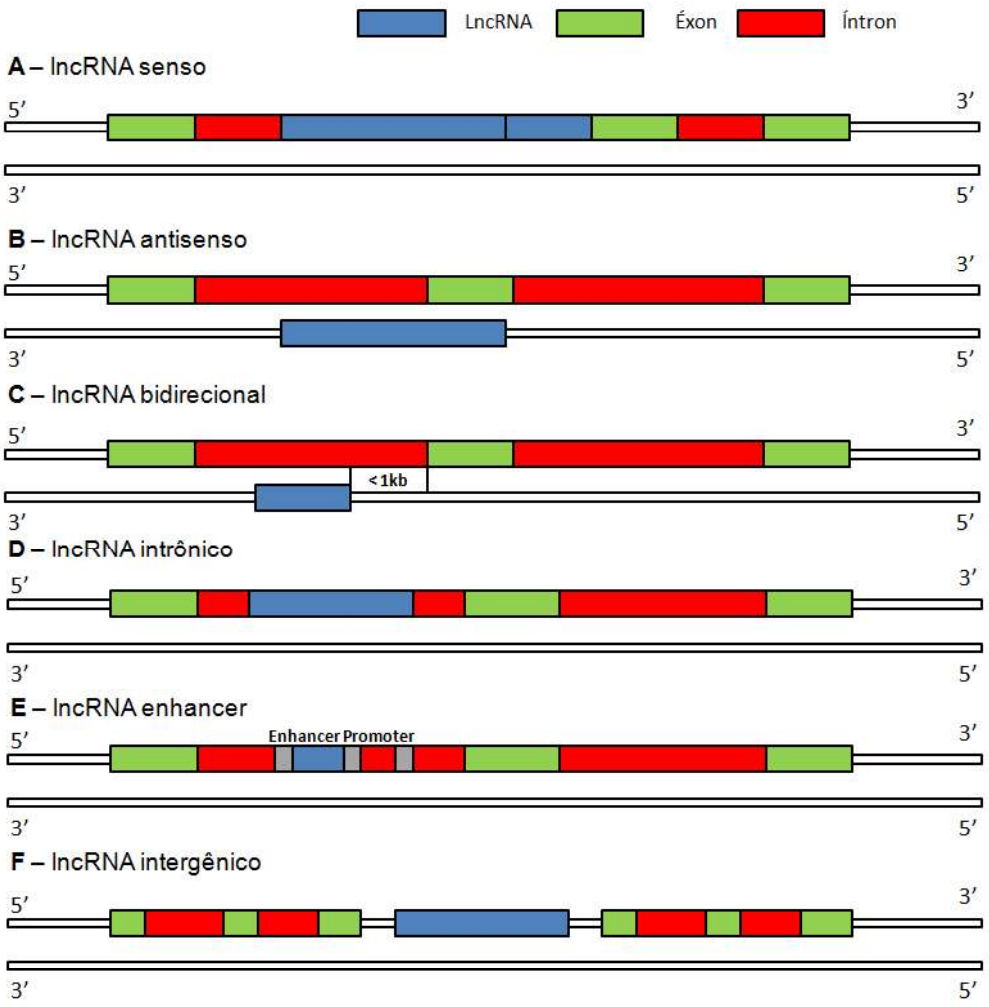


Figura 3 Classificação dos lncRNAs de acordo com a localização genômica. **Fonte**: Adaptada de Vieira et al. (2017).

Chama-se de lncRNA senso (A) quando o lncRNA se sobrepõe à região de transcrição de um ou mais éxons de outro gene na mesma cadeia e antisenso (B), quando se sobrepõe a um ou mais éxons , na cadeia oposta; bidirecional (C), quando o início da transcrição do lncRNA e outro gene na cadeia oposta estão próximos; intrônicos (D), quando se originam inteiramente de íntrons; *enhancer* (E), quando o lncRNA está localizado em regiões de *enhancer* e intergênico (ou lincRNA, do inglês *long intergenic non-coding RNA*) (F), quando está no intervalo entre dois genes (DEVAUX et al., 2015; PONTING; OLIVER; REIK, 2009).

Identificar novos lncRNAs em plantas é um grande empreitada, cuja realização tem sido facilitada pelos avanços das ferramentas de bioinformática, que proporcionam diversos tipos de análises por meio de programas e

ferramentas computacionais. Consultando os principais bancos de dados que possuem lncRNAs de plantas, fica claro como o avanço das pesquisas proporcionou a identificação de um grande número de sequências (**Quadro 1**). Esse acúmulo foi subsidiado pelo advento de novas técnicas de sequenciamento atreladas ao desenvolvimento de abordagens computacionais, uma vez que a maioria destes lncRNAs foi identificada por meio de abordagens *in silico*.

Quadro 1 Bancos de dados que contém de lncRNAs de plantas (consulta em 09 de março de 2022[1]).

Bases de dados	Total de espécies de plantas	Total de lncRNAs	Endereço
AlnC	682	10.855.598	http://14.139.61.8/AlnC/index.html
CANTATAdb	39	239.631	http://cantata.amu.edu.pl/
DsTRD	1	27.687	http://bi.sky.zstu.edu.cn/DsTRD/home.php
EVLncRNAs	56	506	https://www.sdklab-biophysics-dzu.net/EVLncRNAs2/#
GreeNC v. 2.0	94	491.595	http://greenc.sequentiabiotech.com/wiki2/Main_Page
MAPslnc 1.0	21	88.459	https://lncrnapipe.cimap.res.in/Default.aspx#templatemo-blog
NONCODE	23	94.697	http://www.noncode.org/index.php
PLncDB V2.0	80	1.246.372	https://www.tobaccodb.org/plncdb/
PlncRNAdb	4	5.107	http://bis.zju.edu.cn/PlncRNADB/index.php
RiceLncPedia	1	6.925	http://3dgenome.hzau.edu.cn/RiceLncPedia#/
RLncRNABase	1	6.064	http://bioinfo.tnau.ac.in/RLncRNABase/index.html

Inicialmente, uma limitação dos bancos de dados de lncRNAs de plantas era o fato de serem pouco compreensíveis em comparação aos bancos de lncRNAs de animais, servindo, muitas vezes, apenas de meros depósitos de onde as sequências podiam ser baixadas. Com o tempo, bancos de dados mais compreensíveis foram desenvolvidos. CANTATAdb (SZCZEŚNIAK; ROSIKIEWICZ; MAKAŁOWSKA, 2016) é um dos bancos de dados de lncRNAs de plantas mais compreensíveis, possibilitando, inclusive, a verificação da expressão dos lncRNAs. AlnC (SINGH; VIVEK; KUMAR, 2021) é um banco de dados de lncRNAs de angiospermas, contando com o maior número de

[1]Alguns bancos de dados de nosso conhecimento não puderam ser consultados pois os links de acesso não estavam ativos na data da consulta: PNRD; PlantNATsDB; lncRNAdb e PLNlncRbase.

espécies (682) dentre os bancos consultados. EVLncRNAs (ZHOU et al., 2021) é um banco de dados curado manualmente e validado por abordagens experimentais. GreeNC v. 2.0 (DI MARSICO et al., 2022) é um coleção de lncRNAs de plantas identificados computacionalmente. MAPslnc 1.0 (ZHAO et al., 2018) é um repositório de lncRNAs de plantas medicinais e aromáticas. NONCODE (ZHAO et al., 2016) é um banco de dados dedicado a RNAs não-codificantes (com a exceção de rRNAs e tRNAs), abrigando milhares de transcritos de diferentes espécies de animais e plantas. PLncDB V2.0 (JIN et al., 2021) é um banco de dados de lncRNAs de plantas, incluindo dados de alvos dos lncRNAs e ferramentas para explorar a regulação epigenética dos mesmos. PlncRNAdb (BAI et al., 2019b) também é um banco de dados dedicado a lncRNAs de plantas. Alguns bancos de dados abrigam sequências de apenas uma espécie: DsTRD (SHAO et al., 2016), um banco de lncRNAs da planta medicinal modelo *Salvia miltiorrhiza*; RiceLncPedia (ZHANG et al., 2021), um banco curado de lncRNAs de arroz, e RLncRNABase, constituído por lncRNAs de arroz compilados de outras três bases de dados: GreeNC; PNRD e PLNlncRBase.

Os bancos de dados são importantes por facilitarem o acesso dos pesquisadores às sequências, e são alimentados pela descoberta de novos lncRNAs, pelo uso de abordagens *in silico*. A qualidade/confiabilidade dos lncRNAs identificados depende da ferramenta de bioinformática empregada. Diferentes ferramentas de bioinformática foram desenvolvidas para a identificação de lncRNAs, como podemos ver no **Quadro 2**.

Quadro 2 Algumas ferramentas para a identificação de lncRNAs.

Ferramenta	Organismo	Web server?	Referência
PlantRNA_Sniffer	Planta	Não	(VIEIRA et al., 2017)
LncFinder	Humano Rato	https://bmbl.bmi.osumc.edu/lncfinder/	(HAN et al., 2019)
PLEK	Vertebrado Planta	Não	(LI; ZHANG; ZHOU, 2014)
LncRNA-ID	Humano Rato	Não	(ACHAWANANTAKUN et al., 2015)
lncRScan-SVM	Humano Rato	Não	(SUN et al., 2015)
DeepLNC	Humano	https://bioserver.iiita.ac.in/deeplnc/	(TRIPATHI et al., 2016)

(continuação do Quadro 2)

Ferramenta	Organismo	Web server?	Referência
COME	Humano Rato Mosca Verme Planta	Não	(HU et al., 2017)
SCORE	Vários	Não	(LERTAMPAIPORN et al., 2014)
Lncident	Humano Verme Rato	http://csbl.bmb.uga.edu/mirrors/JLU/Lncident/index.php	(HAN et al., 2016)
PLIT	Planta	Não	(DESHPANDE et al., 2019)
iSeeRNA	Humano Rato	https://sunlab.cpy.cuhk.edu.hk/iSeeRNA/webserver.html	(SUN et al., 2013a)
Linc-SF	Humano	Não	(WANG et al., 2014c)
lncRNA-MFDL	Humano	Não	(FAN; ZHANG, 2015)
lncScore	Animal	Não	(ZHAO; SONG; WANG, 2016)
LncRNApred	Humano	http://mm20132014.wicp.net:57203/LncRNApred/home.jsp	(PIAN et al., 2016)
longdist	Humano Rato Peixe	Não	(SCHNEIDER et al., 2017)
FEELnc	Animal	Não	(WUCHER et al., 2017)
PLncPRO	Planta	Não	(SINGH et al., 2017)
TLCLnc	Animal	Não	(HU; ANDREWS, 2017)
LncADeep	Humano	Não	(YANG et al., 2018)
CREMA	Planta	Não	(SIMOPOULOS; WERETILNYK; GOLDING, 2018)
TERIUS	Humano Rato	Não	(CHOI; NAM, 2018)
LncRNAnet	Humano Rato	Não	(BAEK et al., 2018)
LGC	Várias	https://ngdc.cncb.ac.cn/lgc/calculator	(WANG et al., 2019)
lncRNA-LSTM	Planta	Não	(MENG et al., 2019)
PLncWX	Planta	Não	(GUO et al., 2021)
DeepPlnc	Planta	https://scbb.ihbt.res.in/DeepPlnc/index.php	(GUPTA et al., 2021)
LncPred-IEL	Animal	Não	(XU et al., 2020)
PredLnc-GFStack	Humano Rato	Não	(LIU et al., 2019)
lncRNADetector	Planta	https://lncrnapipe.cimap.res.in/	(SHUKLA et al., 2021)
RNAplonc V1.1	Planta	Não	(NEGRI et al., 2019)

(final do Quadro 2)

Ferramenta	Organismo	Web server?	Referência
PlncRNA-HDeep	Planta	Não	(MENG et al., 2021)
ItLnc-BXE	Planta	Não	(ZHANG et al., 2020)
LncPipe	Humano	Não	(ZHAO et al., 2018)

Dentre as ferramentas que possuem *web server* e são aplicáveis ao estudo de lncRNAs em plantas, DeepPlnc obteve precisão superior a 95% (GUPTA et al., 2021). lncRNADetector foi desenvolvido para a identificação de lncRNAs em plantas, e testado com espécies medicinais e aromáticas (SHUKLA et al., 2021). Dentre as ferramentas que não possuem *web server*, PlantRNA_Sniffer é interessante porque foi desenvolvida para a predição de lincRNAs, e testada com cana-de-açúcar e milho (VIEIRA et al., 2017), integrando ferramentas conhecidas de bioinformática com técnicas de aprendizado de máquina (*machine learning*). COME (HU et al., 2017) foi desenvolvido para a identificação de novos lncRNAs em bibliotecas de RNA-seq por meio de aprendizado de máquina supervisionado, treinado com mRNAs e lncRNAs, e foi testado com *A. thaliana*. SCORE (LERTAMPAIPORN et al., 2014) foi testado com genomas de várias espécies diferentes, incluindo *O. sativa* e é capaz de identificar vários ncRNAs, combinando um modelo de regressão logística com um classificador híbrido baseado em RF (em inglês, *random forest*). PLIT (DESHPANDE et al., 2019) foi desenvolvido para a identificação de ncRNAs em transcriptomas de plantas, e utiliza um modelo baseado em *iterative Random Forests* (iRF). PLncPRO foi testado na identificação de lncRNAs responsivos a estresse abiótico (seca e salinidade) em arroz e grão-de-bico, e é baseado em aprendizagem de máquina e usa o algoritmo *random forest* (SINGH et al., 2017). A intenção dos pesquisadores era desenvolver um modelo de predição aplicável a plantas modelo e não-modelo, monocotiledôneas e eudicotiledôneas. CREMA é baseado em aprendizagem de máquina, e em seu treinamento foram utilizados apenas lncRNAs validados experimentalmente, e resultados superiores foram obtidos com lncRNAs de plantas (SIMOPOULOS; WERETILNYK; GOLDING, 2018). LncRNA-LSTM utiliza uma rede neural para a predição dos lncRNAs (MENG et al., 2019). PLncWX foi testado com algas, musgo, samambaias, monocotiledôneas e dicotiledôneas, e obteve desempenho superior a 90% (GUO et al., 2021). RNAplonc foi treinado com agrião, pepino, soja, álamo e

arroz, e apresentou resultados bastante positivos (NEGRI et al., 2019). PlncRNA-HDeep, um modelo híbrido de predição e que não necessita de conhecimentos prévios para a predição de lncRNAs, apresentou sensitividade, precisão e acurácia superiores a 90% com milho (MENG et al., 2021). ItLnc-BXE foi testado com seis espécies de plantas, e obteve desempenho superior a outras ferramentas disponíveis, com mais de 96% de eficiência (ZHANG et al., 2020).

Uma importante característica dos lncRNAs é a ausência de potencial de codificação de peptídeos, o que pode ser verificado por meio de diferentes ferramentas, algumas delas com *web server*, como podemos observar no **Quadro 3**.

Quadro 3 Ferramentas para cálculo de potencial codificante de sequências.

Ferramenta	Web server	Referência
RNASamba	https://rnasamba.lge.ibi.unicamp.br/	(CAMARGO et al., 2020)
CPC	http://cpc.cbi.pku.edu.cn/	(KONG et al., 2007)
CPC2	http://cpc2.gao-lab.org/	(KANG et al., 2017)
LGC	https://ngdc.cncb.ac.cn/lgc/	(WANG et al., 2019)
CNIT	http://cnit.noncode.org/CNIT/	(GUO et al., 2019)
PORTRAIT	http://bioinformatics.cenargen.embrapa.br/portrait/	(ARRIAL; TOGAWA; MARCELO, 2009)
CPPred	Não	(TONG; LIU, 2019)
CONC	Não	(LIU; GOUGH; ROST, 2006)
PhyloCSF	Não	(LIN; JUNGREIS; KELLIS, 2011)
CNCI	Não	(SUN et al., 2013b)
CodAn	Não	(NACHTIGALL; KASHIWABARA; DURHAM, 2021)

As diferentes ferramentas para cálculo do potencial codificante dos transcritos utilizam abordagens igualmente diversas. CPC (KONG et al., 2007) e CPC2 (KANG et al., 2017), por exemplo, dependem de bancos de dados de proteínas a fim de classificar as sequências. Transcritos que não apresentam *hits* com sequências dos bancos de proteínas, nesse caso, tendem a ser classificados como não-codificantes. RNASamba (CAMARGO et al., 2020), por

outro lado, busca superar as limitações impostas pela dependência de bancos de dados de sequências codificantes, procurando características intrínsecas das sequências (obtidas a partir de *k-mers*) que permitam identificá-las como não-codificantes, por meio do uso de uma rede neural. CNIT (*Coding-Non Coding Identification Tool*) é uma atualização de outra calculadora de potencial codificante (CNCI, *Coding-Non-Coding Index*), buscando maior precisão e rapidez, sendo 200 vezes mais rápido do que seu antecessor (GUO et al., 2019). PORTRAIT é baseado em SVM, e recomendado para espécies pouco estudadas, utilizando dois modelos: um dependente e outro independente de proteínas preditas a partir dos próprios transcritos (ARRIAL; TOGAWA; MARCELO, 2009). CPPred é uma ferramenta baseada em SVM e em múltiplas características extraídas dos lncRNAs, sendo mais eficiente com ncRNAs curtos (TONG; LIU, 2019). CONC é baseado em SVM, e utiliza múltiplas características das sequências caso elas estivessem codificando, como comprimento do peptídeo e presença de homólogos (LIU; GOUGH; ROST, 2006). PhyloCSF é um método de genômica comparativa, comparando alinhamentos multiespécie para determinar se uma dada sequência é ou não codificante (LIN; JUNGREIS; KELLIS, 2011). CodAn (*Coding sequence Annotator*) prevê regiões CDS e UTR de transcriptomas de eucariotos (NACHTIGALL; KASHIWABARA; DURHAM, 2021).

Além de todos os aspectos listados anteriormente, diversos outros podem ser explorados por meio de abordagens *in silico*. LncRNAs podem interagir com DNA, RNA, proteínas ou com a cromatina (PINKNEY; WRIGHT; DIERMEIER, 2020), e diferentes ferramentas têm sido desenvolvidas com o objetivo de predizer, dentre outros aspectos, a interação dos lncRNAs (**Quadro 4**).

Quadro 4 Ferramentas para diferentes estudos *in silico* envolvendo lncRNAs.

Ferramenta	Web server?	Referência
Predição da estrutura secundária		
RNAfold web server	http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi	(GRUBER et al., 2008)
Predição da localização subcelular		
DeepLncLoc	http://bioinformatics.csu.edu.cn/DeepLncLoc/	(ZENG et al., 2022)

(final do Quadro 4)

Ferramenta	Web server?	Referência
Interação RNA-RNA		
LncTar	http://www.cuilab.cn/lnctar	(LI et al., 2014)
IntaRNA	http://rna.informatik.uni-freiburg.de/IntaRNA/Input.jsp	(MANN; WRIGHT; BACKOFEN, 2017)
Interação lncRNA-proteína		
Global Score	Não	(CIRILLO et al., 2016)
PLRPIM		(WEKESA et al., 2019)
Vias funcionais		
LncRNAs2Pathways	Não	(HAN et al., 2017)
Ligação microRNA lncRNA		
spongeScan	Não	(FURIÓ-TARÍ et al., 2016)
Identificação de sítios de edição em lncRNAs		
REDItools	Não	(PICARDI et al., 2014)

RNAfold web server faz parte do Vienna RNA Websuite, um conjunto de ferramentas para analisar RNA (GRUBER et al., 2008), dotadas de uma interface de internet que torna o uso mais amigável. RNAfold web server é utilizado para predizer a estrutura secundária de energia livre mínima de sequências. DeepLncLoc é uma ferramenta que prediz a localização subcelular de lncRNAs (ZENG et al., 2022). LncTar é uma ferramenta de predição de interação lncRNA-RNA, que utiliza a minimização de energia livre para predizer os pares (LI et al., 2014). IntaRNA é uma ferramenta de predição da interação RNA-RNA, e se utiliza de acessibilidade ao sítio de interação e de predição de *seeds* para determinar os pares que interagem (MANN; WRIGHT; BACKOFEN, 2017). O algoritmo computacional Global Score foi o primeiro capaz de predizer interações globais e locais de RNAs e proteínas (CIRILLO et al., 2016). PLRPIM é um método desenvolvido para predizer a interação entre lncRNAs de plantas e proteínas baseado, dentre outros aspectos, em *k-mer* (WEKESA et al., 2019). LncRNAs2Pathways é capaz de identificar vias funcionais influenciadas por um conjunto de lncRNAs selecionados pelo usuário (HAN et al., 2017). spongeScan é uma ferramenta capaz de predizer lncRNAs atuando como ceRNAs (esponjas de miRNAs), com uma interface de internet (FURIÓ-

TARÍ et al., 2016). REDItools é uma ferramenta que se propõe a detectar eventos de edição de RNA (PICARDI et al., 2014).

A utilização de bibliotecas de RNA-seq para a identificação de lncRNAs tem se tornado rotina na pesquisa com essas moléculas, inclusive em plantas (WANG; CHEKANOVA, 2019). As próprias bibliotecas de RNA-seq podem ser utilizadas para o cálculo da expressão dos lncRNAs. Se as sequências forem conhecidas, a técnica de *microarrays* pode ser empregada para a obtenção da expressão dos lncRNAs (RAI et al., 2019).

As abordagens experimentais, que se seguem como verificação das sequências preditas computacionalmente, constituem uma importante etapa da caracterização funcional dos lncRNAs. Diante da vastidão de lncRNAs identificados computacionalmente, apenas uma pequena parcela foi validada e teve sua funcionalidade confirmada experimentalmente (**Figura 4**). Plantas modelo (*A. thaliana*, *O. sativa* e *Z. mays*) possuem o maior número de lncRNAs verificados experimentalmente.

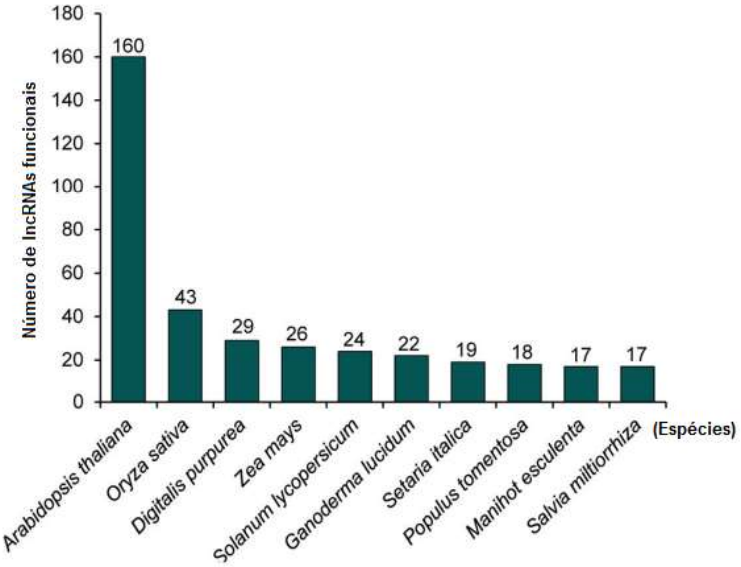


Figura 4 Número de lncRNAs funcionais experimentalmente validados das 10 primeiras plantas listadas no EVLncRNA2.0. Adaptada de (CHEN et al., 2021).

Em plantas, lncRNAs já foram caracterizados exercendo diferentes funções (**Figura 5**), associadas, por exemplo, a: resposta ao frio; resposta ao calor; amadurecimento de frutos; metabolismo de nitrogênio e de fosfato; resposta a hormônios vegetais, como auxinas e etileno; processos específicos

de raiz; nodulação; gametogênese; defesa contra organismos patogênicos; produção de antocianinas; florescimento; desenvolvimento da plântula e diversos outros.

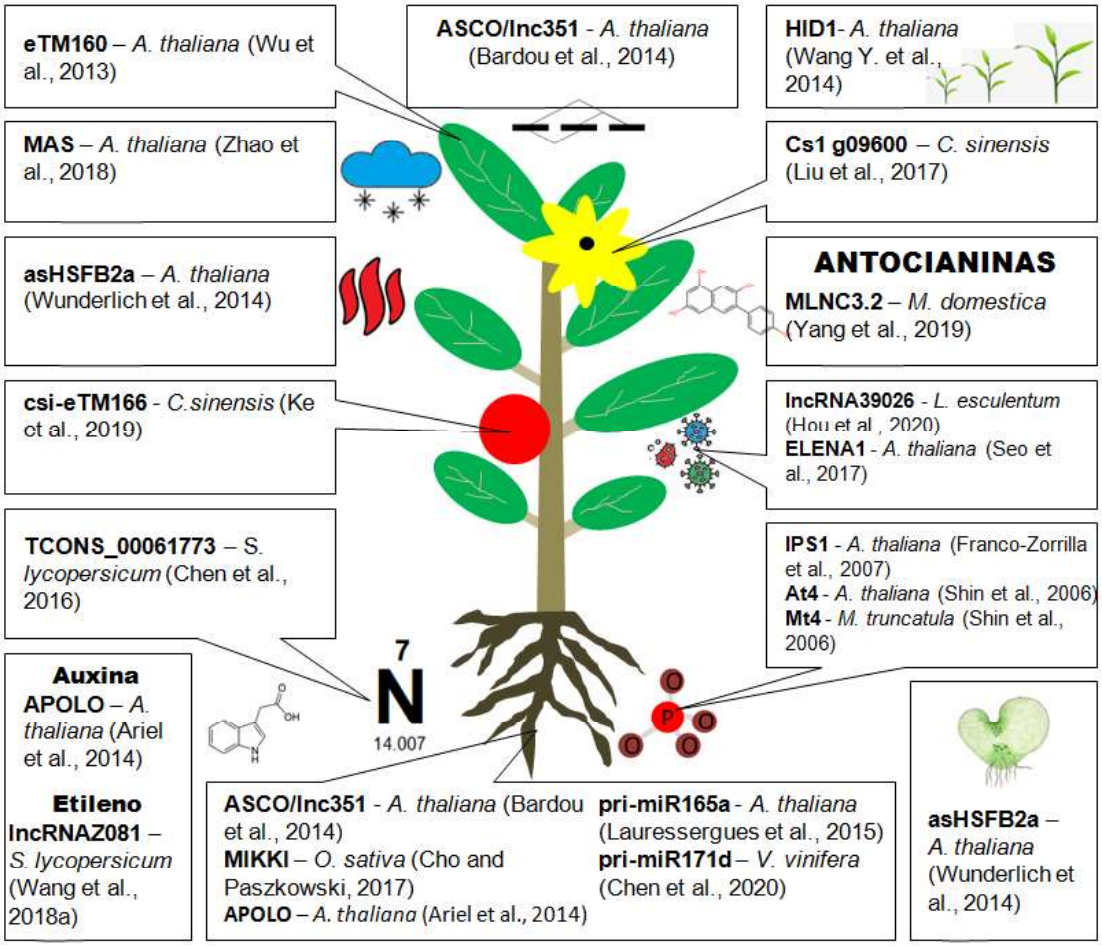


Figura 5 LncRNAs funcionais caracterizados em diferentes plantas.

Plantas modelo, como *A. thaliana*, *Medicago truncatula*, *O. sativa* e *Nicotiana tabacum* já tiveram lncRNAs identificados sob diferentes tipos de estresse biótico e abiótico. O primeiro lncRNA de planta, ENOD40, foi isolado de nódulos de *M. truncatula* (CRESPI et al., 1994). Em *M. truncatula*, 582 lncRNAs responsivos a deficiência de fosfato foram identificados, e três deles foram caracterizados funcionalmente (WANG et al., 2017). Na mesma espécie, foram identificados lncRNAs responsivos a frio (ZHAO et al., 2020b) e a estresse osmótico e salinidade (WANG et al., 2015b). Em *A. thaliana*, foram identificados lncRNAs responsivos a: luz azul (SUN et al., 2020b); temperatura (SEVERING et al., 2018); limitação de fosfato (YUAN et al., 2016). Em arroz,

foram identificados lncRNAs: responsivos ao frio (LENG et al., 2020); responsivos à seca (WEIDONG et al., 2020); desenvolvimento da semente (ZHAO et al., 2020a); resistência a patógeno (JAIN et al., 2017); fertilidade e fotomorfogênese (WANG et al., 2021); senescência (HUANG et al., 2021). Em *N. tabacum* foram identificados lncRNAs: associados a raiz (CHEN et al., 2019); desenvolvimento de brotos axilares (WANG et al., 2022), e em resposta a nematódeo (LI et al., 2018). A identificação de lncRNAs envolvidos em respostas de defesa de plantas descortina um interessante cenário, colocando-os com potencial de atuação no controle de pragas. *Magnaporthae oryzae* é um fungo que causa sérios danos à cultura de arroz, e o trabalho de Jain et al. (2017) comparou os lncRNAs de uma linhagem de arroz resistente ao fungo, com e sem a inoculação de *M. oryzae*, por meio de análise de transcriptoma. Foram identificados diversos lincRNAs passíveis de participarem da resposta da planta ao patógeno. Em tabaco, foram identificados 565 lncRNAs diferencialmente expressos em plantas colonizadas pelo nematóide-das-galhas (*Meloidogyne* sp.). A análise de *Gene Ontology* (GO) revelou que os alvos desses lncRNAs estão envolvidos na resposta da planta a: estresse biótico; indução da resistência sistêmica; transdução de sinal hormonal e resposta hipersensível, associando esses lncRNAs à resposta ao patógeno (LI et al., 2018).

Diversas plantas não-modelo também tiveram lncRNAs identificados, como: *Cajanus cajan* (NITHIN et al., 2017); *Populus euphratica* (MA et al., 2019); *Solanum lycopersicum* (WANG et al., 2015a); *S. bicolor* (SUN et al., 2020a); *Ananas comosus* (BAI et al., 2019a); *Musa itinerans* (LIU et al., 2018), dentre muitas outras espécies.

Alguns lncRNAs receberam mais atenção por parte dos pesquisadores ao longo do tempo, especialmente por causa de suas funções em eventos ainda não completamente compreendidos, como, por exemplo, a passagem por vernalização, nodulação em raiz e a resposta da planta à privação de nutrientes, como o fosfato (Pi). Nesse caso, destacamos quatro exemplos cujos mecanismos de ação têm sido bastante explorados: COOLDAIR/COOLAIR; ENOD40 e IPS1 (**Figura 6**).

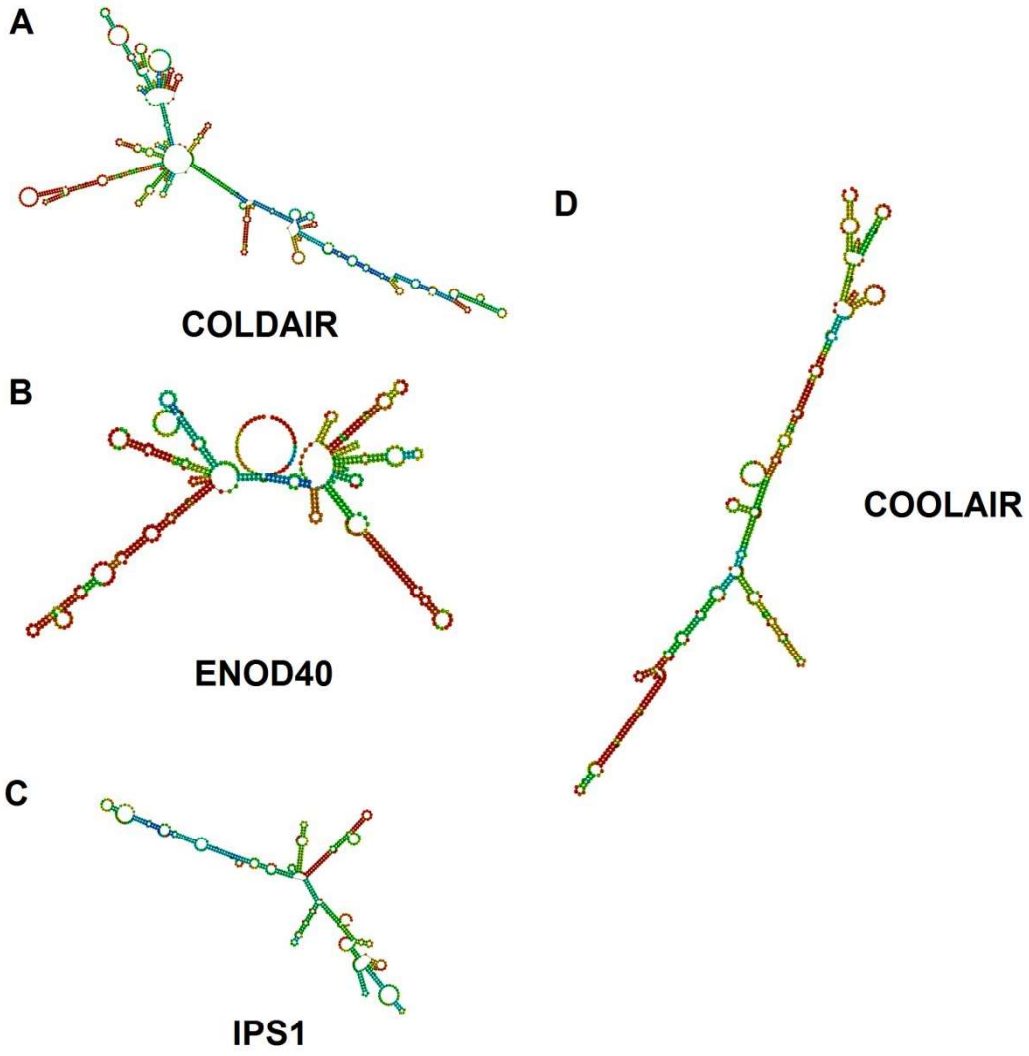


Figura 6 Estruturas secundárias de lncRNAs de plantas bem estudados obtidas pelo RNAfold web server. **A –** *A. thaliana* COLDAIR (GenBank: HG975388.1); **B** - *M.sativa* ENOD40 (GenBank: X80263.1); **C –** *A. thaliana* IPS1 (GenBank: NM_180219.1); **D –** *A. thaliana* COOLAIR (GenBank: GQ352646.1).

COLDAIR e COOLAIR são dois lncRNAs produzidos a partir do gene FLOWERING LOCUS C (FLC). FLC é um repressor floral, reprimido epigeneticamente e, devido ao interesse em elucidar os mecanismos de floração, FLC é um dos genes de plantas mais estudado (MICHAELS; AMASINO, 1999; WHITTAKER; DEAN, 2017). O FLC é reprimido por duas vias: via autônoma e via de vernalização (JIAO et al., 2019). Vernalização é um fenômeno bem estudado, encontrado em plantas de clima frio, por meio da qual há um impedimento do florescimento durante o inverno, com a permissão da floração apenas com a chegada da primavera (KIM et al., 2009). Em *A. thaliana*, COLDAIR é um lncRNA que reprime o gene FLC, e só se associa à

cromatina no lócus de FLC (KIM; XI; SUNG, 2017). Em *A. thaliana*, COOLAIR é um lncRNA antisenso expresso do lócus de FLC, e responde ao frio (HAWKES et al., 2019). A despeito da baixa conservação a nível de sequência primária, seu padrão de expressão e estrutura secundária *in vitro* são altamente conservados em Brassicaceae (JIAO et al., 2019).

ENOD40 (do inglês *early nodulin gene*) é um lncRNA envolvido em processos de regulação simbiótica entre leguminosas e bactérias (FUJITA et al., 1993). Interessantemente, ENOD40 é necessário para a simbiose do nódulo radicular, mas não é necessário o nódulo de simbiose radicular actinorrízico (GANGULY et al., 2021). Foi visto em *M. truncatula* que ENOD40 produz um ncRNA funcional e dois peptídeos curtos, cuja perturbação de ENOD40 provocou diminuição da nodulação. Além disso, ENOD40 atua alterando a localização da proteína MtRBP1, provocando uma mudança de MtRBP1 para o citoplasma e, na ausência do lncRNA, a proteína permanece no núcleo (CHEN; CHEN; ZHANG, 2015; SOUSA et al., 2001).

IPS1 (do inglês *Induced by Phosphate Starvation 1*) é um lncRNA identificado primeiramente em *A. thaliana* e que funciona como *target mimic* de mir399 (FRANCO-ZORRILLA et al., 2007; WU et al., 2013). IPS1 age como *decoy*, impedindo que mir399 se ligue ao seu alvo original (PHO2, *Phosphate 2*).

A despeito do interesse na identificação de novos lncRNAs em diversas espécies de plantas, a participação destes importantes reguladores da expressão gênica em eventos de interação com organismos endossimbiontes permanece pouco compreendida, necessitando de maiores investigações.

1.3 LncRNAs envolvidos em diferentes processos em raizes

As raízes das plantas possuem diferentes funções. São responsáveis por fixar os vegetais ao solo, resistindo à força do vento ou às águas das chuvas; captam água e nutrientes do solo, possibilitando que sejam conduzidos até as partes superiores e, não menos importante, interagem com microorganismos presentes no solo, os quais podem beneficiar ou prejudicar o desenvolvimento da planta. Não é de se espantar que lncRNAs tenham sido descritos como importantes elementos reguladores de processos em raiz.

Em *A. thaliana* submetida a seca, estresse salino ou limitação de fosfato, 42 lncRNAs com aumento de expressão em raiz e responsivos a pelo menos um desses estresses foram identificados (BEN AMOR et al., 2009). Os mesmos autores verificaram que a superexpressão de dois desses lncRNAs interferiram no crescimento das raízes de *Arabidopsis*.

Em células de raiz, APOLO (*AUXIN REGULATED PROMOTER LOOP RNA*) é um lncRNA que participa da regulação do crescimento primário e da resposta ao gravitropismo (BAZIN; BAILEY-SERRES, 2015). APOLO interfere na estrutura da cromatina próxima ao seu gene alvo, PINOID (PID) (ARIEL et al., 2014). PINOID codifica uma proteína quinase responsável pela localização do transportador de auxina PIN-FORMED 2 (PIN2) (HUANG et al., 2010). APOLO e PINOID são corregulados, e ambos respondem a auxina exógena.

Considerando a homeostase de fosfato (Pi), foi verificado em *A. thaliana* que dois lncRNAs (At4 and IPS1) têm sua expressão fortemente induzida em plantas submetidas a privação de fosfato. Mutantes de At4 exibem problemas de distribuição de Pi entre raiz e parte aérea quando em privação de fosfato (BURLEIGH; HARRISON, 1999; SHIN et al., 2006). At4 e IPS1 agem como *decoys* de mir399, prevenindo que esse miRNA se ligue ao seu transcrito alvo (PHO2), que está envolvido com a regulação do carregamento de Pi no xilema (HUANG et al., 2013; LIU et al., 2012). Em arroz, o lncRNA cis-NATPHO$_{1;2}$ está envolvido com a homeostase de Pi (JABNOUNE et al., 2013). Cis-NATPHO$_{1;2}$ atua potencializando a tradução de OsPHO1;2, que codifica uma proteína envolvida com o transporte de Pi do xilema da raiz para a parte superior da planta (SECCO; BAUMANN; POIRIER, 2010).

ASCO (*Alternative Splicing COmpetitor*) é um lncRNA que atua regulando o *splicing* alternativo de proteínas NSR (*Nuclear Speckle RNA-binding*), modulando diretamente o crescimento e desenvolvimento da raiz com regulação por auxina (BARDOU et al., 2014). A auxina é um regulador positivo do crescimento da raiz (HU et al., 2021), e a superexpressão de ASCO produz plantas extremamente sensíveis a esse hormônio. Diversos outros lncRNAs, além de ASCO, interagem com NSR, indicando que eles também direcionam eventos de *splicing* alternativo dependente de NSR (BAZIN et al., 2018).

Em Petunia x hybrida, o gene Sho, envolvido na produção de citocininas, produz um transcrito não-codificante de sua fita antisenso (ZUBKO; MEYER,

2007). Citocininas são hormônios vegetais inibidores do crescimento da raiz e essenciais para senescência e desenvolvimento de brotos e folhas (AMASINO, 2005; MOK; MARTIN; MOK, 2000). Produzidas na ponta da raiz e em locais específicos da parte aérea (NORDSTRÖM et al., 2004), citocininas são empregadas para regeneração de plantas (BISHOPP; HELP; HELARIUTTA, 2009). Parece haver uma correlação entre a expressão de Sho e seu transcrito antisenso o que leva à produção de sRNAs de 24 nucleotídeos em vários tecidos, menos na raiz. Isso sugere que a manutenção da produção de citocininas em raiz se deve à não degradação do lncRNA Sho, o qual parece estar diretamente relacionado com a biogênese de citocininas em Petunia x hybrida (ZUBKO; MEYER, 2007).

LncRNAs de raiz têm sido descritos como reguladores em resposta a diferentes fatores abióticos. Em *A. thaliana*, DRIR (*DRought-Induced lncRNA*) está intimamente relacionado à resistência da planta ao estresse hídrico e salino, controlando genes relacionados à translocação hídrica e resposta a ABA (QIN et al., 2017). Em algodão (*Gossypium hirsutum*), o lncRNA973 é responsável por regular genes envolvidos na resposta a estresse salino (ZHANG et al., 2019). Em mandioca (*Manihot esculenta*), o lncRNA340 atua como *mimic* do miR169, regulando a resposta a seca e ao frio (LI et al., 2017).

A disponibilidade de nutrientes, como o nitrogênio (N) é capaz de modular a arquitetura radicular em diferentes espécies (FORDE, 2014; FORDE; WALCH-LIU, 2009). Leguminosas possuem uma característica única: são capazes de se associar com bactérias fixadoras de nitrogênio (KISTNER; PARNISKE, 2002), o que as coloca como componentes importantes de práticas agrícolas menos dependentes de insumos químicos (ZHAO et al., 2020c). ENOD40 é um lncRNA envolvido no processo de nodulação, e seu gene possui conservação de certas regiões entre plantas não-leguminosas (GULTYAEV; ROUSSIS, 2007). O silenciamento de ENOD40 em *L. japonicus* provoca atenuação da formação de nódulos (KUMAGAI et al., 2006). É interessante ressaltar que outras funções são atribuídas a ENOD40, como elemento-chave na regulação da expressão gênica ou pela produção de dois peptídeos que interagem com a sacarose sintase, modulando o fluxo de sacarose na nodulação (BARDOU et al., 2011; CAMPALANS; KONDOROSI; CRESPI, 2004).

1.4 LncRNAs de milho associados a diferentes processos biológicos

Em um tempo relativamente curto, o milho se diferenciou grandemente de seu ancestral selvagem, e foi proposto que lncRNAs tiveram participação neste processo como intermediários de novos genes (HEZRONI et al., 2015). Um estudo recente identificou 18.165 lncRNAs em milho, por meio da análise de 749 experimentos de RNA-seq (HAN et al., 2018), e concluiu que o milho possui um transcriptoma extremamente complexo. Em milho, centenas de lncRNAs foram identificados em resposta a diferentes tipos de fatores bióticos e abióticos, como: calor; frio; seca; salinidade; boro; alagamento; colonização por fungo benéfico; giberelina e deficiência de nitrogênio (LV et al., 2016, 2019; PANG et al., 2019; WANG et al., 2018; YU et al., 2020; ZHANG et al., 2014a).

Em milho, a literatura relata que a identificação de lncRNAs tem sido feita, majoritariamente, por meio de bioinformática, com posterior verificação da existência do lncRNA, principalmente, por PCR (do inglês *Polymerase Chain Reaction*). Considerando bancos de dados que abrigam lncRNAs de milho e artigos: CANTATAdb (10761); GreeNC (7129); EVLncRNAs (26); NONCODE (4717); Lv et al. 2016 (7245); Li et al. (2014) (20163); Zhu et al. (2017) (753); Huanca-Mamani et al. (2018) (48345); Pang et al. (2019) (3488); Yu et al. (2020) (6099); Han et al. (2019a) (5941); Han et al. (2019b) (18165); Boerner e McGuinis (2492); Zhang et al. (2014) (664); Wang et al. (2015) (11105); Wang et al. (2018) (9933), totalizando 157026 lncRNAs (valores redundantes, autores podem ter depositado as sequências nos bancos de dados; alguns autores não disponibilizaram arquivos com as sequências).

A maioria absoluta dos lncRNAs identificados por abordagens computacionais permanece como sequências putativas e sem caracterização de suas funções. Alguns lncRNAs de milho, por outro lado, foram fortemente associados aos respectivos estresses, por meio de diferentes ensaios/abordagens, sugerindo seus papéis como elementos reguladores dos processos (**Figura 7**).

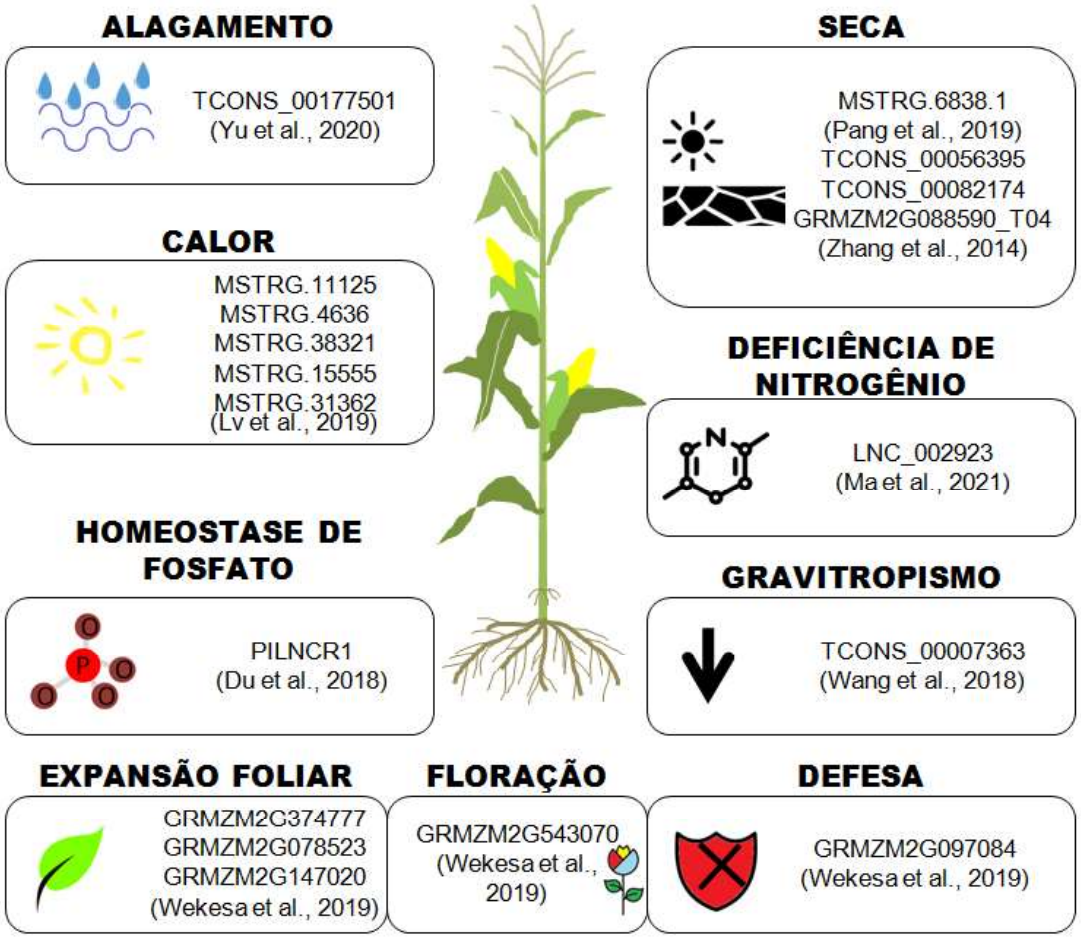


Figura 7 LncRNAs verificados experimentalmente em milho e suas funções.

Yu et al. (2020) associaram o lncRNA TCONS_00177501 à tolerância do milho ao alagamento, pois esse lncRNA faz parte de um módulo obtido a partir de uma rede de coexpressão que inclui transcritos codificantes de proteínas e lncRNAs de raiz. A expressão de TCONS_00177501 foi relacionada ao gene Zm00001d029280. Diferentes lncRNAs foram identificados como responsivos a calor (LV et al., 2019), e a tolerância a baixas concentrações de fosfato (Pi) em milho foi associada ao módulo formado pelo lncRNA PILNCR1 e miR399 (DU et al., 2018). Por meio de expressão transiente do lncRNA LNC_002923 em *Nicotiana benthamiana*, foi demonstrado que ele poderia guiar a clivagem de Zm00001d015521 com a participação de ZmmiR159 (MA et al., 2021). GRMZM2G088590_T04 é um lncRNA que se apresentou reprimido em raiz tratada por 5h, e induzido em raiz com 10h de tratamento (seca) (ZHANG et al., 2014a). Wekesa e colaboradores (2019) associaram diferentes lncRNAs de milho a defesa, floração e expansão foliar.

1.5 Bioestimulantes e plantas

Os bioestimulantes, como produtos comerciais, necessitam passar por uma regulamentação. O Regulamento de Produtos Fertilizantes da União Européia (EU) (*European Union (EU) Fertilizing Products Regulation*) propõe considerar como bioestimulante um produto capaz de estimular os processos nutricionais das plantas, melhorando um ou mais aspectos, como produtividade, tolerância a estresse abiótico ou uso mais eficiente de nutrientes (RICCI et al., 2019), e deverá constar no rótulo do produto as vantagens que ele proporciona com especificação do tipo de planta beneficiado.

Os bioestimulantes compreendem um conjunto heterogêneo de substâncias e microorganismos, caracterizadas por: natureza diversa; funções biológicas diversas; efeitos convergentes para as mesmas funções agrícolas (aumento da resistência a estresse abiótico; eficiência no uso de nutrientes; melhoria da cultura). É interessante lembrar que o aumento da resposta a estresse biótico é deixado de fora, mantendo bioestimulação e biocontrole separados. A definição de bioestimulante, logo, deve estar atrelada às funções agrícolas dos compostos e microorganismos utilizados (DU JARDIN, 2015).

Entre os bioestimulantes microbianos, estão os fungos microrrízicos arbusculares e bactérias promotoras do crescimento vegetal. Fungos frequentemente são encontrados associados a raízes de plantas, estabelecendo relações benéficas ou agindo como parasitas (BEHIE; BIDOCHKA, 2014). A partir do aparecimento das plantas terrestres, plantas e fungos coevoluíram, estabelecendo diversos tipos de relacionamento mutualístico (BONFANTE; GENRE, 2010; JOHNSON; GRAHAM, 2013). Fungos micorrízicos, que são encontrados em mais de 90% dos vegetais, são exemplos de organismos associados com plantas, sendo os fungos micorrízicos arbusculares bastante difundidos em plantas cultivadas (BONFANTE; GENRE, 2010; BEHIE; BIDOCHKA, 2014). Desta forma, as micorrizas têm constituído alvo de crescente interesse como um recurso para a ampliação das práticas agrícolas sustentáveis, pois promovem diversos benefícios às plantas, como resistência a estresse biótico e abiótico; aquisição de micro e macronutrientes e equilíbrio hídrico (AUGÉ, 2001; GIANINAZZI et

al., 2010; HARRIER; WATSON, 2004; VAN DER HEIJDEN et al., 2006). Alguns trabalhos têm sugerido a existência de uma verdadeira rede de conexão, permitindo a comunicação entre as hifas e as plantas de uma comunidade, o que implica em um potencial uso para a agricultura (JOHNSON; GILBERT, 2015; SIMARD et al., 2012). A utilização das micorrizas na agricultura exige uma adaptação das plantas aos microrganismos, assim como de mudanças no manejo agrícola (GIANINAZZI et al., 2010; PLENCHETTE et al., 2005). Apesar de promissor, o uso de produtos com fungos micorrízicos tem o desafio da propagação em larga escala (DALPÉ; MONREAL, 2004), assim como da compreensão das especificidades fungo-hospedeiro e da dinâmica populacional dos sistemas agrícolas (DU JARDIN, 2015). Diferentemente de micorrizas, fungos endofíticos (*Trichoderma* spp. e *Piriformospora indica*, por exemplo), conseguem viver parte do ciclo de vida fora da planta e, uma vez colonizando raízes, são capazes de transferir nutrientes para o hospedeiro (BEHIE; BIDOCHKA, 2014). As propriedades desses fungos como inoculantes têm sido estudadas, apontando para um potencial uso na agricultura sustentável, especialmente por serem mais fáceis de multiplicar *in vitro*. *Trichoderma spp*., inclusive, têm sido relatado como dotado de potencial biopesticida e de biocontrole (MUKHERJEE et al., 2013; NICOLÁS et al., 2014).

Bactérias interagem com plantas de diversas formas: de forma mutualística ou parasitária; interações na rizosfera, no rizoplano e no interior das células; interações perenes ou temporárias, até mesmo com transmissão de bactérias por meio das sementes; alterações na eficiência do uso de nutrientes; promovendo resistência a fatores abióticos e a patógenos; e regulação do crescimento da planta (AHMAD; PICHTEL; HAYAT, 2008). Dois tipos principais de bactérias são reconhecidos como bioestimulantes agrícolas: endossimbiontes mutualísticos (*Rhizobium*) e bactérias promotoras do crescimento de plantas (BPCV). *Rhizobium* já é comercializado como biofertilizante (DU JARDIN, 2015), e seu funcionamento e as contribuições para a prática agrícola já foram extensamente estudados. As BPCVs atuam sobre os mais diferentes processos, incluindo a nutrição e o crescimento vegetal e a defesa contra fatores bióticos e abióticos (AHMAD; PICHTEL; HAYAT, 2008; BABALOLA, 2010; BERENDSEN; PIETERSE; BAKKER, 2012; BERG et al.,

2014; BHATTACHARYYA; JHA, 2012; GAIERO et al., 2013; PHILIPPOT et al., 2013; (VACHERON et al., 2013). Os benefícios dos microorganismos variam enormemente a depender do genótipo da planta considerada, sendo alguns efeitos específicos das cepas utilizadas ou, ainda, do sinergismo entre diferentes microorganismos, o que pode limitar a utilização deste recurso. A formulação do bioinoculante também pode influenciar os resultados de campo (ARORA; KHARE; MAHESHWARI, 2011; BRAHMAPRAKASH; SAHU, 2012). Contudo, o uso de bactérias como bioestimulante de plantas vem encontrando espaço no mercado mundial (BERENDSEN; PIETERSE; BAKKER, 2012).

Existem, ainda, os chamados bioestimulantes de plantas não-microbianos, que compreendem principalmente hidrolisados de proteínas (*protein hydrolysates*, PHs) e extratos de algas marinhas (*seaweed extracts*, SWE) (ROUPHAEL; COLLA, 2020). Os hidrolisados de proteínas, que compreendem principalmente peptídeos sinalizadores e aminoácidos livres, são capazes de conferir maior resistência de plantas a estresse, assim como melhorar a germinação e o crescimento de plântulas, a qualidade e produtividade de frutos (COLLA et al., 2017). Os extratos de algas marinhas, obtidos a partir de algas vemelhas, verdes e marrons são os mais comumente utilizados na agricultura e na horticultura. As algas colhidas no mar dificultam a padronização dos produtos obtidos, o que torna necessária a procura por produção padronizada de algas para fins de pesquisa e uso industrial (ROUPHAEL; COLLA, 2020). As algas podem constituir fontes renováveis de bioestimulantes de plantas (CHIAIESE et al., 2018). Subprodutos derivados de atividades agrícolas e industriais, desde que atendam a determinados critérios, como estar livres de agrotóxicos e de metais pesados, assim como possuírem disponibilidade e facilidade de armazenamento, podem constituir fontes de bioestimulantes de plantas. Esses subprodutos incluem vermicomposto, lodo de esgoto e compostagem de lixo urbano (XU; GEELEN, 2018).

Ainda no grupo dos bioestimulantes não-microbianos temos as substâncias húmicas. Substâncias húmicas se referem ao ácido húmico e ácido fúlvico, e são pequenas moléculas anfifílicas com capacidade de formar agregados moleculares ou supramoleculares em solução e em superfícies minerais (PALUMBO et al., 2018; PICCOLO, 2001; SCHAUMANN, 2006). A atividade biológica do ácido húmico é conhecida desde 1917 (BOTTOMLEY,

1917). Os efeitos das substâncias húmicas sobre as plantas são dependentes da concentração utilizada e das propriedades físico-químicas, propriedades estas intimamente relacionadas ao material que deu origem às substâncias húmicas empregadas. Algumas pesquisas associaram os efeitos das substâncias húmicas à presença de compostos que possuem atividade semelhante à de hormônios (MORA et al., 2010; NARDI et al., 1994). Já foram demonstrados diferentes efeitos das substâncias húmicas sobre plantas, como: controle da disponibilidade de nutrientes em milho (EYHERAGUIBEL; SILVESTRE; MORARD, 2008); absorção e assimilação de nitrogênio (VACCARO et al., 2009); mudanças na arquitetura de raiz (CANELLAS et al., 2002; NARDI et al., 2002; ZANDONADI; CANELLAS; FAÇANHA, 2007); intensificação da atividade de H^+-ATPase (NARDI et al., 1991; ZANDONADI et al., 2010); alterações no metabolismo secundário (ciclo de Krebs, fotossíntese, glicólise) (LOTFI et al., 2018; ROOMI et al., 2018); aumento do número de compostos fenólicos (ERTANI et al., 2011); redução da infecção das plantas (OLIVARES et al., 2015) e incremento da proteção vegetal (HERNANDEZ et al., 2015).

Tanto os bioestimulantes de plantas não-microbianos quanto os microbianos são reconhecidos como fontes renováveis, capazes de contribuir para práticas agrícolas mais sustentáveis, seja agindo como biofertizantes em solos com baixa disponibilidade de N e P ou contribuindo para o aumento da resistência das culturas a fatores abióticos, principalmente seca e salinidade (BACKER et al., 2018; BITTERLICH et al., 2018; GRANADA et al., 2018). Já foi demonstrada uma ação complementar entre fungos micorrízicos e bactérias promotoras de crescimento, indicando os benefícios do uso de mais de um recurso combinado (BACKER et al., 2018; BITTERLICH et al., 2018; TURRINI et al., 2018). O desenvolvimento de produtos a partir de consórcios de microorganismos, como o consórcio Trichoderma-Azotobacter, pode constituir uma estratégia interessante para um manejo mais eficiente e práticas agrícolas menos prejudiciais ao ambiente (WOO; PEPE, 2018).

Uma BPCV que tem sido utilizada como modelo para estudos da interação planta-bactéria endofítica é a *H. seropedicae* (CHUBATSU et al., 2012; DIVAN BALDANI; BALDANI; DÖBEREINER, 2000; PEDROSA et al., 2011). Esta BPCV foi isolada de cana-de-açúcar, arroz, sorgo e milho

(FERRARI et al., 2014; MONTEIRO et al., 2012; OLIVARES et al., 1996) e tem auxiliado na compreensão do complexo processo de promoção de crescimento promovido pela bactéria (BODDEY et al., 1995; ELBELTAGY et al., 2001; GYANESHWAR et al., 2002; JAMES, 2000; JAMES et al., 2002; RONCATO-MACCARI et al., 2003).

Os mecanismos de interação das plantas com bioestimulantes, como BPCV e ácido húmico, são complexos e ainda necessitam de estudos relacionados ao entendimento da modulação de redes transcricionais. Considerando o papel regulatório que ncRNAs apresentam, alguns estudos identificaram que miRNAs podem estar associados à resposta da planta a microorganismos benéficos: em milho colonizado por bactérias endofíticas diazotróficas (*Azospirillum brasilense* e *H. seropedicae*), os miRNAs miR397, miR398, miR408, e miR528 se mostraram induzidos pela colonização por *H. seropedicae* (THIEBAUT et al., 2020). Esses miRNAs estão associados a homeostase de cobre, um micronutriente envolvido em vias de diferentes processos biológicos, como a fotossíntese e a eliminação de espécies reativas de oxigênio. Em feijão (*Phaseolus vulgaris*) inoculado com *Rhizobium tropici,* miR398 teve decréscimo de expressão (NAYA et al., 2014), com aumento de expressão de seus alvos (CSD1 e Nod19), sugerindo a participação de miR398 no estágio inicial de colonização por *R. tropici*. Em soja (*Glycine max*) inoculada com a bactéria fixadora de nitrogênio *Bradyrhizobium japonicum* miR172 apresentou aumento de expressão em nódulo. miR172 funciona promovendo a nodulação ao regular negativamente seu alvo, NNC1, um gene que possui papel central na repressão da nodulação (WANG et al., 2014). Ainda em soja com *B. japonicum*, miR167 também apresentou aumento de expressão em raiz, após 24 horas da inoculação, e está envolvido na sinalização por auxina (WANG et al., 2015). Também foram verificados altos níveis de miR167 em nódulos, sugerindo participação deste miRNA na nodulação, além da regulação hormonal (WANG et al., 2009; LUIS et al., 2012). A colonização de arroz pelo fungo *Piriformospora indica* provocou o aumento de expressão de miR159 e miR396 em folhas, durante estresse hídrico. Esses dois miRNAs têm como alvos os fatores de transcrição MYB e GRF, que estão envolvidos no controle do crescimento e na economia de água. Desta forma, mudanças na expressão de miR159 e miR396 levam ao aumento da tolerância

da planta ao estresse hídrico pelo controle de seus alvos (FARD et al., 2017). Em *M. truncatula* colonizado pelo fungo *Rhizophagus irregularis,* miR171 está associado à colonização pelo fungo, e mostrou aumento de expressão em raiz, enquanto seu alvo (NSP2) teve sua expressão diminuída. A superexpressão de miR171 provoca redução da colonização do fungo em raiz, o que sugere a participação desse miRNA em uma via de controle da colonização pela regulação de NSP2 (LAURESSERGUES et al., 2012). Com relação aos lncRNAs, Han e colaboradores (2019) reportaram que 63 lncRNAs foram diferencialmente expressos na associação de milho com *Rhizophagus irregulares* (fungo AM). Estes poderiam estar envolvidos na regulação da troca de nutrientes entre a planta e o fungo, mostrando o potencial regulatório de lncRNAs a ser explorado em plantas submetidas a bioestimulantes.

2. OBJETIVOS

2.1 OBJETIVO GERAL

Caracterizar, por meio de análises de bioinformática, lncRNAs expressos em milho submetido a colonização por *H. seropedicae* com adição de ácido húmico.

2.2 OBJETIVOS ESPECÍFICOS

A – Identificar os lncRNAs de milho descritos/caracterizados na literatura;
B – Identificar lncRNAs já caracterizados e novos lncRNAs em bibliotecas de transcriptoma de milho submetido a colonização por *H. seropedicae* com adição de ácido húmico;
C – Analisar a expressão dos lncRNAs em bibliotecas de transcriptoma de milho submetido a colonização por *H. seropedicae* com adição de ácido húmico;
D – Identificar lncRNAs que são alvos de miRNAs;
E – Identificar os mRNAs alvos de lncRNAs;
F – Construir redes de interação de miRNAs-lncRNAs-mRNAs.

3. ARTIGO – SUBMETIDO À REVISTA PLANTA – A1

Long non-coding RNAs from maize roots treated with plant biostimulants targets elements related to hormonal regulation pathways

Leandro de Oliveira Silva[1], Lucas Maciel Vieira[2], Maria Emilia M. T. Walter[2], Fábio Lopes Olivares[3] and Clicia Grativol[1*]

[1]Laboratory of Chemistry and Function of Proteins and Peptides (LQFPP), Center of Biosciences and Biotechnology, North Fluminense State University Darcy Ribeiro (UENF), Campos dos Goytacazes, RJ, Brazil; oliveiradasilvaleo@gmail.com (ORCID: 0000-0003-1651-7760.); cgrativol@uenf.br (ORCID: 0000-0002-8873-701X).
[2]Department of Computer Science, University of Brasilia, Brasília-DF 70910-900, Brazil; lucas.maciel.vieira@gmail.com (ORCID:0000-0001-6286-8209); memtwalter@gmail.com (ORCID: 0000-0001-6822-931X)
[3]Center for Development of Biological Inputs for Agriculture (NUDIBA), North Fluminense State University Darcy Ribeiro (UENF), Av. Alberto Lamego 2000, Campos dos Goytacazes, 28013-602, Rio de Janeiro, RJ, Brazil. fabioliv@uenf.br (ORCID:0000-0001-6541-0324)
*Corresponding author: cgrativol@uenf.br; Tel.: +55 22 2739-7107

Main conclusion The lncRNAs from maize roots target genes related to several biological processes when plants are subjected to colonization by *H. seropedicae* and humic acids, including regulation by hormones such as abscisic acid and auxins, and response to biotic and abiotic stresses.

Abstract: Long non-coding RNAs (lncRNAs) constitute a class of non-coding RNAs (ncRNAs) responsible for playing critical regulatory roles at different stages of development and in plants subjected to different types of biotic and abiotic stress. Despite the increase of studies on plant lncRNAs in recent years, the functions of lncRNAs in maize (*Zea mays* L.) subjected to colonization by beneficial microorganisms and the response to humic acids remain unknown. To identify and characterize maize lncRNAs, we used single-end RNA-seq libraries constructed from roots of non-inoculated seedlings and inoculated with diazotrophic endophytic bacteria *Herbaspirillum seropedicae* and humic acid. Through a computational pipeline, we identified a total of 32,501 known High-Confidence lncRNAs (HC-lncRNAs) and 6,418 novel candidate lncRNAs (NC-lncRNAs) in the transcriptomes. The expression profile of lncRNAs revealed 3,263 HC-lncRNAs and 1,050 NC-lncRNAs significantly differentially expressed (DE) in the six different comparisons. We also identified the regulatory networks of DE-lncRNAs by predicting cis-target genes and miRNAs. The results showed that lncRNAs target genes associated with different processes, such as

hormone signal transduction, abscisic acid (ABA) and ethylene response, Ca^{2+}-binding, auxin response; outer cell wall composition, and transmembrane signaling, among others. Furthermore, interaction networks of lncRNAs-mRNAs-miRNAs were constructed and showed the potential of lncRNAs as a key regulatory element. Our results showed that lncRNAs are involved in different processes during the interaction of maize with bacteria *H. seropedicae* and in the presence of humic acid, including the participation of elements associated with distinct plant hormones.

Keywords: *Zea mays* – ncRNAs – lncRNAs – miRNAs – endophytic bacteria - humic substances

Abbreviations

AA Amino acid
BNF Biological nitrogen fixation
CDS Coding sequence
CONAB Companhia Nacional de Abastecimento
CPC Coding potential calculator
DE Differentially expressed
dG deltaG
FC Fold change
HAs Humic acids
HC-lncRNAs High-confidence lncRNAs
lncRNA Long non-coding RNA
miRNA MicroRNA
mRNA Messenger RNA
NCBI National Center for Biotechnology Information
NC-lncRNAs Novel candidate lncRNAs
ncRNA Non-coding RNA
NT Nucleotide
ORF Open Reading Frame
PCT Protein-coding transcript
sRNA Small RNA
SVM Support Vector Machine
TPM Transcript Per Million

1. Introduction

Biostimulation of plants by biocontrol agents or natural or synthetic compounds has been growing enormously in agriculture in the past decade. The application of bioactive compounds or beneficial microorganisms to the plant enhances growth, tolerance to environmental stresses, boost nutrient uptake and crop quality (du Jardin 2015; Brown and Saa, 2015). Biostimulants are promising to face chemical fertilizers high costs and environmental damage once they benefit crops productivity with the activation of plants' metabolism (Bulgari et al., 2014).

Among the beneficial microorganisms used for plant biostimulation, the genus *Herbaspirillum* has been isolated and characterized in different plant

species, such as rice, sorghum, sugarcane, and wheat (Monteiro et al. 2012). The genus *Herbaspirillum* comprises gram-negative, vibroid bacteria that carry one to three polar flagella. Some members can fix atmospheric N_2, being called diazotrophic (Baldani et al. 1986, 1996). The association of the genus *Herbaspirillum* with several plant species, especially members of the Poaceae family, has been one of the best-studied models of plant-bacteria interaction (Monteiro et al. 2012).

Despite the different studies about the interaction of *H. seropedicae* with plants, several molecular aspects of this association have yet to be fully elucidated. In a field experiment, Dartora et al. (2016) found a 12% increase in the phosphorus (P) content of the corn leaf subjected to colonization by *A. brasilense* and *H. seropedicae*, supplemented by N fertilization. Azevedo et al. (2019) demonstrated biochemical and molecular changes in maize inoculated with *H. seropedicae* and with the addition of humic acids isolated from vermicompost (Azevedo et al. 2019). Under these conditions, maize showed overexpression of aquaporins, with an increase in the efflux of H^+ in the root. In addition to the bacteria, the use of humic substances in crops has shown positive results in productivity. Testing the combination of humic substances and *H. seropedicae* delivered as foliar spray application under field conditions, Canellas et al. (2013) obtained an increase of 65% of the maize grain production.

Humic substances have been successfully used as biostimulants (Jindo et al. 2020). In 1917, the biological activity of humic acids was described (Bottomley 1917). Subsequently, other properties of humic acids were described: hormone–like nature (Nardi et al. 2002; Canellas et al. 2008); promoting expression of genes encoding specific enzymes, such as membrane H^+ATPase (Elena et al. 2009) and increased nitrate absorption (Piccolo et al. 1992). It has even been shown that humic substances can affect the root growth and architecture, mainly inducinglateral roots formation of maize (Canellas et al. 2002).

Humic substances refer to humic acid (HA) and fulvic acid, and are small amphiphilic molecules capable of forming molecular or supramolecular aggregates in solution and on mineral surfaces (Palumbo et al. 2018; Piccolo 2002; Schaumann 2006). HAs are used in the pharmaceutical industry,

medicine, agriculture, and remediation of pollutants, among other technological applications (Melo et al. 2016). The biological activity of humic acids on maize was higher for low molecular weight fractions 35-175 kDa (Vlčková et al. 2009). Humic acids extracted from lignite increased different biological activities in maize seedlings (David et al. 2014). Stimulation of maize seedlings by humic acids positively modulated hormonal regulatory pathways and abiotic stress response genes (Canellas et al. 2020). Humic acids extracted from green compounds demonstrated positive effects on corn growth (Monda et al. 2018), as well as humic acids from coal and peat (Mindari et al. 2014).

Regardless of the strategy used for plant biostimulation, we must broaden our understanding of the molecular mechanisms involved in the plant's response to the treatment used. An important group of molecules that have been associated with different molecular mechanisms are the long non-coding RNAs (lncRNAs), belonging to the class of non-coding RNAs (ncRNAs). ncRNAs control gene expression by various mechanisms, including chromatin remodeling and regulation of expression at the transcriptional and post-transcriptional levels (Kaikkonen et al. 2011), and are divided into two classes: small ncRNAs (sRNAs, with <40 nucleotides) and long ncRNAs (lncRNAs, with >200 nucleotides).

LncRNAs comprise a class of ncRNAs, with the following characteristics: a total absence of open reading frame (ORF) or presence of an ORF with <100 amino acids (Kapranov et al. 2007); length >200 nt (Kim and Sung 2012); GC content lower in comparison with mRNAs (Niazi and Valadkhan 2012), as well as a smaller number of exons, and low expression when compared to protein-coding genes, which can even make their annotation more difficult (Derrien et al. 2012). The characterization of many lncRNAs revealed nuclear localization, the composition of polyadenylated and non-polyadenylated transcripts, low level of sequence conservation and expression (Kapranov et al. 2007; Wu et al. 2008). Moreover, lncRNAs can target genes or miRNAs and act as decoys of miRNAs (Statello et al. 2021).

Maize (*Zea mays* L.) is one of the most important plant crops. In addition to direct consumption in human food and animal feed production, maize is widely used as an industrial material in developed countries, as well as for the production of ethanol (Tester and Langridge 2010). In maize, different works

identified lncRNAs. Using ribosomal RNA depletion, Lv et al. (2016) identified 7,245 lncRNAs in plants grown under nitrogen-limited conditions, being 637 identified as nitrogen-responsive lncRNAs. Li et al. (2014b) identified 20,163 putative lncRNAs, with 1,704 lncRNAs classified as highly confidential. Zhu et al. (2017) identified 753 lncRNAs in maize seeds. These lncRNAs revealed specific tissue and stage specificity and differential expression, indicating participation in seed development. Huanca-Mamani et al. (2018), working with maize from the Atacama Desert, identified 48,345 lncRNAs in plants subjected to the combined stress of salinity and excess boron. Of these, 1,710 were shown to be responsive to combined stresses. Pang et al. (2019) identified 3,488 high-confidence lncRNAs in maize subjected to drought stress, being 1,535 identified as drought-responsive. Yu et al. (2020) identified 6,099 lncRNAs in maize plants subjected to waterlogging stress, being 3,190 differentially expressed. Han et al. (2019b) identified 18,165 high-confidence lncRNAs, with 6,872 of them conserved between maize and teosinte. Using full-length maize cDNA, Boerner and McGinnis (2012) identified 2,492 candidate lncRNAs. Zhang et al. (2014) identified 664 drought stress-responsive lncRNAs. Wang et al. (2015) identified 11,105 lincRNAs from the RNA-seq of different organs and developmental stages in maize. Using SNPs associated with agronomic traits of interest, the authors identified 951 maize lncRNAs genes associated with leaf development. Wang et al. (2018) identified 9,933 lncRNAs, 22 differentially expressed under gibberellin stimulation. Considering biostimulants, Han et al. (2019a) identified 5,941 lncRNAs in maize root submitted to colonization by the arbuscular mycorrhizal fungi *Rhizophagus irregularis*, being 63 differentially expressed. They found most of lncRNAs were involved in nutrient exchange between plant-benefic fungi.

Taking advantage of the significant amount of lncRNAs already identified in maize, we aimed to identify known and new candidate lncRNAs involved in maize root response to plant biostimulants - endophytic diazotrophic bacteria *H. seropedicae* and humic acids; conditions in which the participation of maize lncRNAs had not yet been investigated. We constructed a database of high-confidence lncRNAs (HC-lncRNAs) and built a support vector machine algorithm to identify novel candidate lncRNAs (NC-lncRNAs) responsive to the treatments. Computational analyses were applied to identify differentially

expressed lncRNAs and construct interaction networks of lncRNAs-miRNAs-mRNAs to understand the molecular mechanisms of plant-bacteria interaction and response to humic acid.

2. Materials and methods

2.1 Maize RNA-seq libraries

All procedures for germination, cultivation in pots, inoculation with *H. seropedicae*, and humic acid treatment were described in Nunes et al. (2019, 2021). Briefly, maize seeds var. Dekalb 7815 were surface sterilized in 0.5% NaClO solution for 30 min, followed by three times of 5 min washing in sterilized water and germinated for three days in filter paper. Then, seedlings were transferred to 2 L pots containing 2mM $CaCl_2$ and inoculated with a bacterial suspension of *H. seropedicae* strain HRC54 containing 10^8 cells. mL (B: bacteria treatment), treated with humic acids equivalent to 3.5 mM C. L^{-1} (HA: humic acid treatment) and its combination (B+HA) compared to non-treated plants (CTL: control). The experimental assay was carried out with four treatments (control, bacteria, humic acids and bacteria plus humic acids) and three biological replicates. After five days of treatment exposition, total RNA from control roots and five days treated plants was extracted and sequenced on Illumina HiSeq 2500 system to construct the RNA-seq libraries. Single-end libraries were generated, with RNA selection by the poly-A+ tail, with three biological replicas of each treatment and control, totaling 12 libraries (**Supplementary Table 1)**.

2.2 Transcript assembly

The quality of the libraries was evaluated using FASTQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to remove reads with average base quality lower than 20. Trimmomatic (v0.39) (Bolger et al., 2014) removed reads containing adapter sequences. The reads from each RNA-seq library were aligned against the B73 genome reference (RefGen_v4 AGPv4) by using HISAT2 version 2.1.0 (Kim et al. 2015) with default parameters, along with the maize gene annotation file containing exon-intron boundaries (in GFF3). Finally, StringTie (v. 2.1.4) (Pertea et al. 2015) was used to assemble and merge transcripts with default parameters.

2.3 Construction of high confidence dataset of lncRNAs

LncRNAs are available in two databases, CANTATA and GreeNC (Gallart et al. 2016; Szcześniak et al. 2016), and seven different works (Lv et al. 2016; Li et al. 2014b; Huanca-Mamani et al. 2018; Han et al. 2019; Boerner and McGinnis 2012; Zhang et al. 2014 and Pang et al. 2019) were used to construct a dataset of known lncRNAs. First, the sequences were concatenated into a single file and eliminated redundancies. Then, to increase data accuracy and decrease the number of false positives, the known lncRNAs sequences were aligned against maize CDS (RefGen_v4 AGPv4) and those with 100% of coverage and identity with CDS were discarded. These filtered sequences were identified as high confidence lncRNAs dataset (HC-lncRNAs). Finally, to determine the known HC-lncRNAs in our transcriptome, the dataset of lncRNAs was aligned against the transcriptome using BLASTN (cutoff threshold of an E-value less than 1e-4), and sequences with ≥70% alignment and 100% identity were considered known HC-lncRNAs.

2.4 Identification of NC-lncRNAs with support-vector machine learning

An identification pipeline was developed for the prediction of NC-lncRNAs. First, transcripts were size-filtered, and all with <200 nt were discarded. Then, the transcripts were submitted to the online RNASamba tool (Camargo et al. 2020) to calculate the potential coding and remove coding sequences. The transcripts were then submitted to BLAST against the Rfam database 14.7 (Griffiths-Jones et al. 2003). Next, we discard transcripts that presented hits with non-coding RNA and other structured RNA elements. Then, the remaining transcripts were submitted to the TransDecoder tool (https://github.com/TransDecoder/TransDecoder/releases) to discard transcripts with ORF >300 nt (>100 amino acids).

Finally, the transcripts that survived the previous steps were submitted to our modelto increase the reliability of the lncRNAs prediction. The model uses the Support Vector Machine (SVM) method to predict maize lncRNAs and protein-coding transcripts (PCTs). The model's pipeline was based on three steps: building the database, choosing features, and constructing the model. Genome transcripts from B73 maize (RefGen_v4 AGPv4; Gramene) annotated

as non-coding or PCTs were obtained and used as raw data. To maintain data consistency, a filtering step was performed. Only transcripts with more than 200 bp were kept in the first step. In the second step, a BLAST between the groups (PCTs and lncRNAs) was performed to remove the outliers classified with the opposite class. The transcripts were filtered by standard deviation based on their characteristics in the third step. In the fourth stage, both classes were balanced, and the same number of transcripts from each were randomly selected. Finally, our database was assembled with 58,214 lncRNAs and 58,214 PCTs. The features used in the model were: set of nucleotide combinations, obtained using the method of Schneider et al. (2017): 'AA', 'AAA', 'AAAA', 'AC', 'ACA', 'ACT', 'AG', 'AGA', 'AGC', 'AGT', 'AT', 'ATA', 'ATC', 'ATG', 'CAC', 'CAG', 'CAT', 'CC', 'CCC', 'CG', 'CGA', 'CGC', 'CT', 'CTA', 'CTC', 'CTG', 'GAC', 'GAG', 'GAT', 'GCA', 'GCG', 'GCT', 'GG', 'GGG', 'GT', 'GTC', 'GTG', 'TAC', 'TAG', 'TAT', 'TCA', 'TCG', 'TCT', 'TG', 'TGA', 'TGC', 'TGT', 'TT', 'TTT', 'TTTT' (Schneider et al. 2017); the size of the ORFs; and the ratio of the size of the ORFs to the size of the transcript. Finally, the model was trained using the selected characteristics and the database as input, obtaining an accuracy of 92%.

2.5 Differential expression of mRNAs and known and new lncRNAs

Differential expression of known and new lncRNAs and mRNAs was done according to the general flowchart shown in **Supplementary Figure 1**. StringTie (v. 1.3.4) (Pertea et al. 2015) was used to assemble transcripts and estimate normalized gene expression. Multi-overlap reads or fragments were not considered in the count. Gffcompare (v0.10.5) (https://ccb.jhu.edu/software/stringtie/gffcompare.shtml) was used to compare assembled and reference transcripts. Further, *feature*Counts (subread-v1.6.2) (Liao et al. 2014) was used to count the number of reads per feature at transcript levels, while normalized expression was estimated in TPM. DESeq2 (Love et al. 2014) was used to calculate the differential gene expression. We used the log2 transformed TPM values for this analysis log2(TPM+1). For the calculation of fold change (FC), the treatments were compared as follows: (bacteria/control; humic_acid/control; bacteria+humic/control; bacteria+humic_acid/bacteria; bacteria+humic_acid/humic acid and

humic_acid/bacteria). We used log2 (fold-change) and adjusted p-value ≤ 0.05 (moderated t-statistic) to select significantly expressed genes. LncRNA expression heatmaps were generated using Morpheus (Ryan et al. 2020). Venn diagrams were generated using a webtool (http://bioinformatics.psb.ugent.be/webtools/Venn/).

2.6 Prediction of miRNA targets, functional target mimicries (miRNA decoys), and mRNA targets

The prediction of target mRNAs of lncRNAs was performed using the LncTar tool (Li et al. 2014a). The coding sequences of our transcriptome were used as targets (only the transcripts that were differentially expressed in different comparisons – 9,776 mRNAs). The mRNAs identified as targets were annotated using the Phytozome (v13) and/or NCBI.

A total of 324 maize mature miRNA sequences available in miRBase (version 22.1: May 2021, https://www.mirbase.org/) were downloaded to predict miRNA-lncRNA interactions. The interaction between miRNAs-target-lncRNAs was predicted using the psRNATarget tool (2017 Update) (Dai et al. 2018b). The parameters for the prediction of lncRNAs as potential decoys or miRNA targets were adopted as defined by Fan et al. 2015, as described: (1) for lncRNAs that are targeted by miRNAs,at most one mismatch or indel between the 9th and 12th position of the 5' end of the miRNA; the number of mismatches <4 nucleotides in other regions and absence of continuous mismatches; (2) for lncRNAs that are decoy of miRNAs, number of mismatches and indels >1 and <6 between the 9th and 12th positions of the 5' end of the miRNA; the perfect match between the 2nd and 8th positions of the 5' end of the miRNA, and number of mismatches and indels <4 in other regions.

2.8 Construction of miRNA-lncRNAs-mRNA interaction networks

The interaction networks of miRNAs-lncRNAs-mRNAs were constructed following the steps: *i*) lncRNAs identified as targets of miRNAs were connected; *ii*) differentially expressed mRNAs identified as targets of miRNAs were connected to the respective miRNAs; *iii*) differentially expressed mRNAs predicted as lncRNA targets, using the LncTar program were connected to the respective lncRNAs; *iv*) the expression of lncRNAs and mRNAs in each

treatment were signaled in the networks following a colorscheme (blue for down-regulated and red for up-regulated). After sorting the mRNAs predicted by LncTAr (smallest to largest dG, deltaG), the 10 mRNAs with the smallest dGs were selected for inclusion in the co-expression networks. The miRNA target genes were selected using the psRNATarget tool and the AGPv4 maize genome (maximum expectation value =3.0).

3. Results

3.1 Construction of a comprehensive set of HC-lncRNAs from maize and identification of NC-lncRNAs from twelve root transcriptomes

By searching the available lncRNAs already reported in two databases – CANTATA and GreeNC – and seven studies (Lv et al. 2016; Li et al. 2014b; Huanca-Mamani et al. 2018; Han et al. 2019; Boerner and McGinnis 2012; Zhang et al. 2014 and Pang et al. 2019), a total of 141,315 lncRNAs sequences were downloaded. The lncRNAs sequences were subjected to filtration of redundant sequences and maize CDS, which resulted in 124,739 unique lncRNAs (**Figure 1**). These sequences were identified as high confidence lncRNAs (HC-lncRNAs). Furthermore, we identified by BLASTN the HC-lncRNAs against twelve root transcriptomes from maize seedlings inoculated with *H. seropedicae*, treated with humic acid, and treated with its combination and non-treated plants. In total, 32,501 transcripts were identified as HC-lncRNAs.

The identification of NC-lncRNAs from maize root was made according to the general pipeline shown in **Figure 1**:

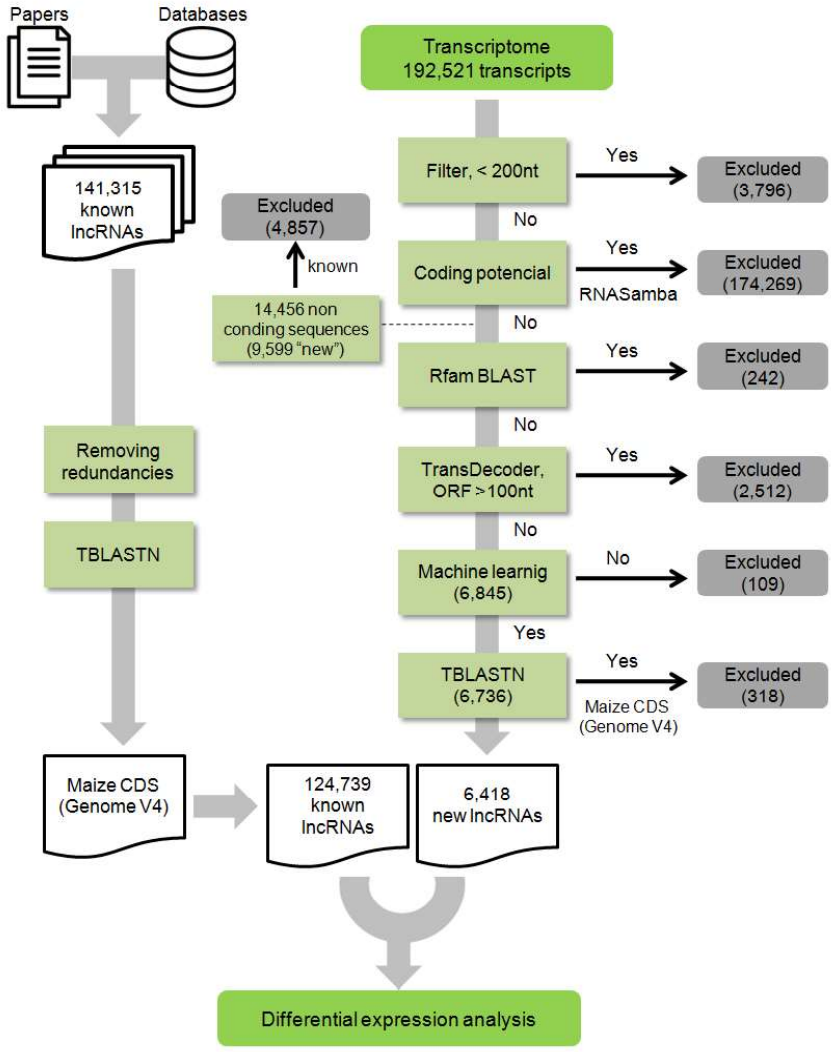


Figure 1 General flowchart of the pipeline used to identify the maize root lncRNAs.

To identify novel maize lncRNAs of twelve root transcriptomes, we used 432,144,184 cleaned reads from RNA-seq libraries, which were used to assemble 192,521 transcripts. After assembling, the transcripts were submitted to a series of steps to obtain novel candidate lncRNAs (NC-lncRNAs). First, the size filtering discarded 3,796 transcripts with a size smaller than 200 nt. Then, the coding transcripts were removed, leaving 14,456 sequences. Of these, 4,857 were similar to known HC-lncRNAs and were discarded. We also excluded 242 transcripts with a hit to Rfam and 2,512 that showed ORF >100 nt. The remaining 6,845 transcripts were submitted to our SVM algorithm, and 109 sequences are identified as coding and discarded. Finally, the 6,736 transcripts were aligned against maize CDS (B73 RefGen_v4) and those that showed 100% identity and plus orientation were discarded. Transcripts that aligned with minus orientation were classified as antisense; those who aligned

with less than 100% identity and plus orientation were classified as sense; transcripts that did not align to any CDS were classified as lincRNAs. Finally, 6,418 transcripts were identified as NC-lncRNAs.

The know HC-lncRNAs and novel lncRNAs showed similar characteristics from those already described in the literature for other plant species (**Figure 2A-B**), reinforcing the high quality of the lncRNAs identified in the present work.

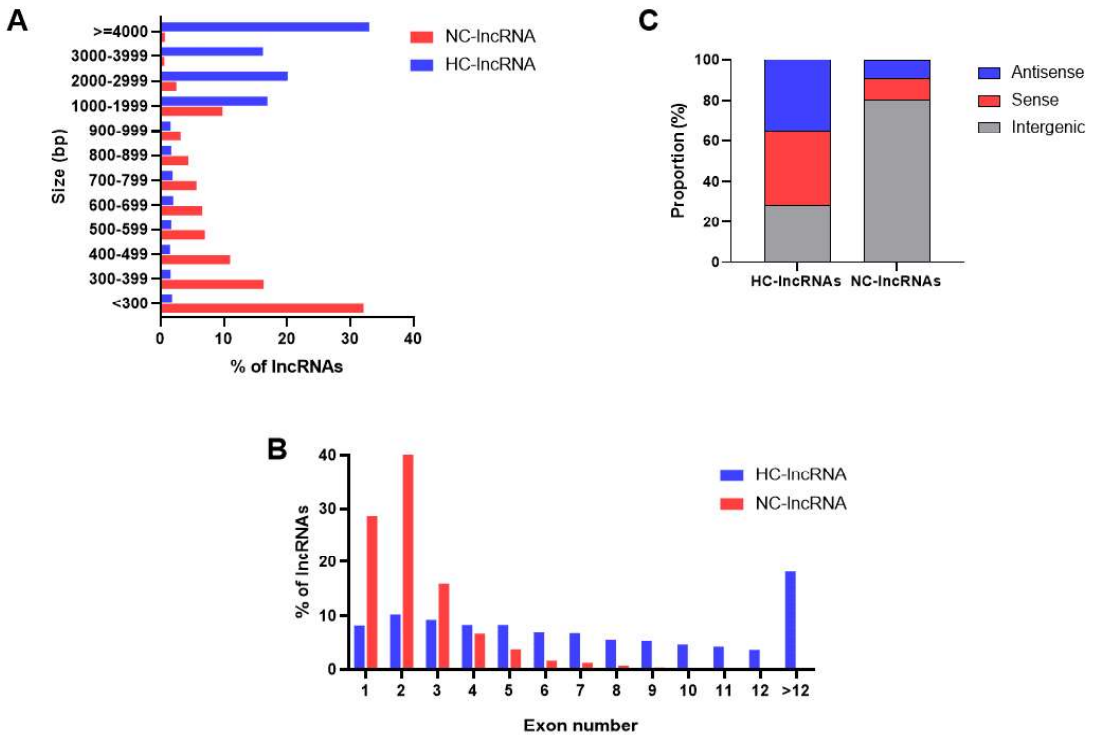


Figure 2 Characteristics of maize lncRNAs. **(A)** Size of lncRNAs. **(B)** The number of exons in maize lncRNAs. **(C)** Proportion of sense, antisense and intergenic lncRNAs.

We found that a significant part of the NC-lncRNAs (~32%) have a size <300 nt, while most of the HC-lncRNAs (33%) are greater than 4,000 nt (**Figure 2A**). Therefore, the NC-lncRNAs discovered in maize are smaller than the known ones. Regarding the number of exons, 18% of HC-lncRNAs showed number of exons greater than 12. While the majority of NC-lncRNAs (~40%) having two exons (**Figure 2B**). Moreover, a higher percentage of NC-lncRNAs have 1-3 exons when compared to HC-lncRNAs. The proportion of sense, antisense, and intergenic lncRNAs were variable among the two groups. A greater proportion of NC-lncRNAs showed an intergenic origin, while the HC-lncRNAs were sense lncRNAs (**Figure 2C**).

3.2 Expression profile of HC-lncRNAs and NC-lncRNAs in maize submitted to biostimulants

To explore the expression profile of HC- and NC-lncRNAs, we evaluate the amount of lncRNAs with exclusive expression in each of the three different treatments and control, and the lncRNAs with expression above one TPM (**Figure 3**). The proportion of exclusive HC-lncRNAs were greater than NC-lncRNAs. We found a similar amount of exclusive lncRNAs in all treatments, but in plants inoculated with the bacteria (**Figure 3A**). In contrast, the exclusive HC-lncRNAs from bacteria – B – treatment showed high expression, together with the treatment B+HA (**Figure 3B**). For NC-lncRNAs, the only B+HA showed high expressed lncRNAs (**Figure 3C**). Among the lncRNAs with >1 TPM, the three treatments induced the expression of lncRNAs with a gradual increase from CTL to B+HA treatment (**Figure 3D**). The expression of NC-lncRNAs followed the same pattern (**Figure 3F**), while a similar expression profile was found in HC-lncRNAs on the four conditions (CTL, B, HA, B+HA) (**Figure 3E**).

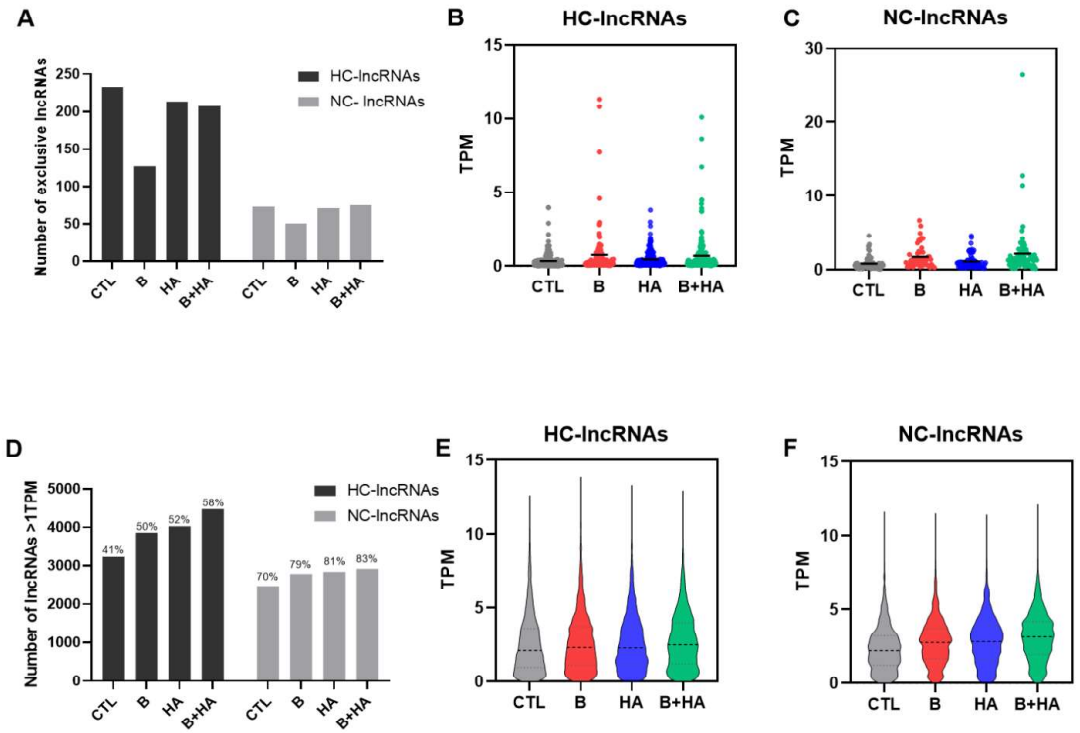


Figure 3 Proportion of differentially expressed lncRNAs identified after comparisons between different treatments. **(A)** Number of exclusive lncRNAs. **(B)** HC-lncRNAs expression. **(C)** NC-lncRNAs expression. **(D)** Number of lncRNAs >1 TPM. **(E)** HC-lncRNAs expression. **(F)** NC-lncRNAs expression.

We used StringTie and DESeq2 to identify the differentially expressed lncRNAs between CTL, B, HA, and B+HA. Among the differentially expressed lncRNAs, 3,263 were HC-lncRNAs and 1,050 were NC-lncRNAs. To investigate the regulation of differential expressed HC- and NC-lncRNAs in roots inoculated with the *H. seropedicae*, treated with humic acid, and submitted to both bacteria and humic acid, we constructed volcano plots with the fold change and p-value of HC- and NC-lncRNAs from DESeq2 analysis (**Figure 3**). We found a similar expression profile between HC- and NC-lncRNAs, with a more significant proportion of up-regulated transcripts in the B/CTL and B+HA/HA comparisons (**Figure 4A,F**). In general, the inoculation with the bacteria seems to induce lncRNAs in contrast with humic acid treatment **(Figure 4A, D, E, F)**, even when the roots were treated with both bacteria and humic acid (**Figure 4E-F**).

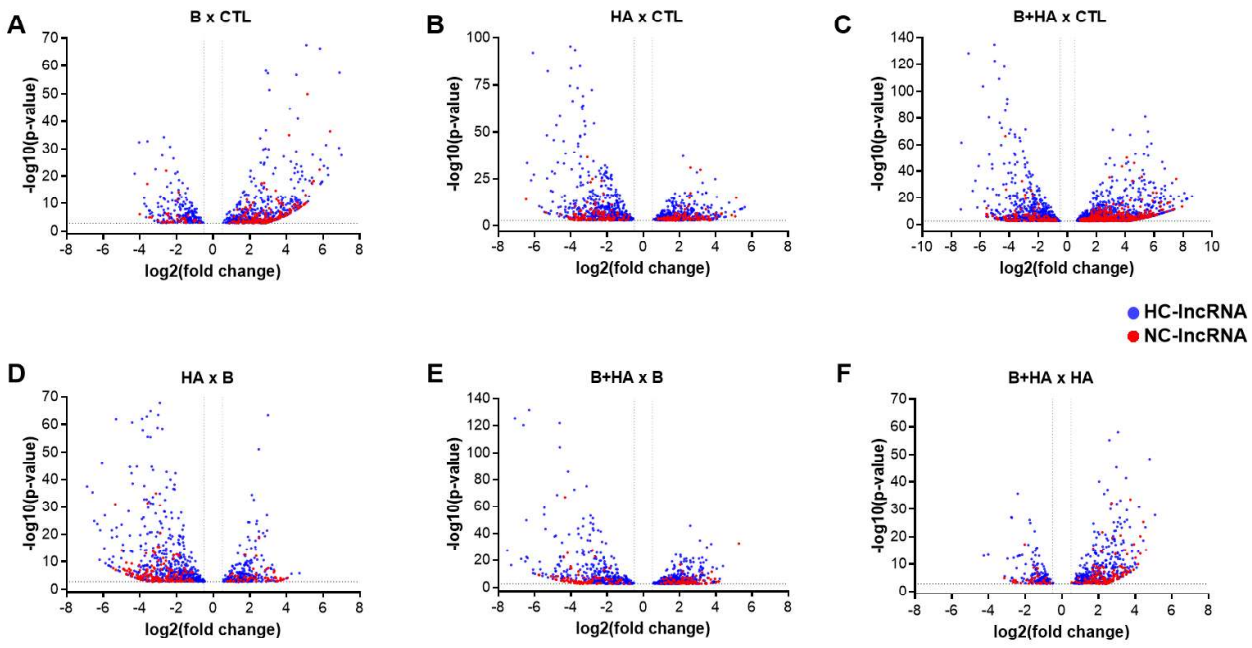


Figure 4 Proportion of differentially expressed lncRNAs identified after comparisons between different treatments. **(A)**

We compared the number of up- and down-regulated lncRNAs in the different conditions by constructed venn diagrams (**Figure 5A-B**). Considering the treatments versus control, the up- and down-regulated charts showed that few differentially expressed lncRNAs are shared between B/CTL and HA/CTL,

demonstrating the high specificity of lncRNAs related to the bacteria and humic acid treatment. Among the HC-lncRNAs, the highest number of shared lncRNAs were 188 and 262 in up- and down-regulated HC-lncRNAs respectively between HA/CTL and B+HA/CTL (**Figure 5 A**). In the NC-lncRNAs, the most shared lncRNAs were 34 and 50 in up- and down-regulated respectively between HA/CTL and B+HA/CTL (**Figure 5B**). Among the three treatments comparisons, the number of shared lncRNAs was less than 21 (**Figure 5**). The B+HA/B and B+HA/HA shared the greatest amount of up- and down-regulated lncRNAs. The B+HA/CTL showed to be the condition which highest numbers of HC- and NC-lncRNAs were specifically regulated (**Figure 5**).

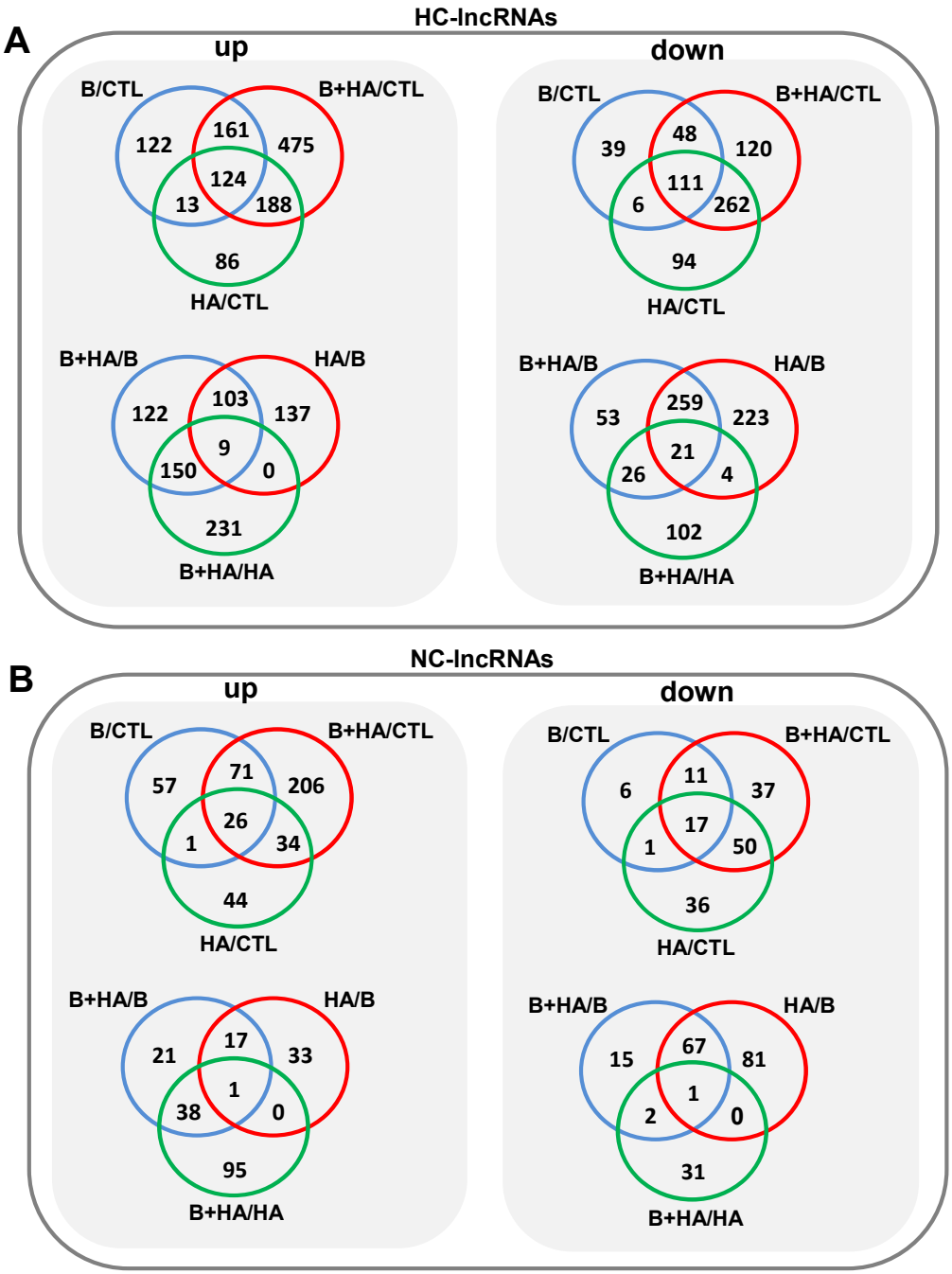


Figure 5 Venn diagrams showing the total differentially expressed lncRNAs identified when different treatments are compared. **(A)** HC-lncRNAs and **(B)** NC-lncRNAs.

In most of the comparisons, the up- and down-regulated lncRNAs were treatment specific; for example, 475 HC-lncRNAs were up-regulated only in the B+HA/CTL. These lncRNAs, which seem to respond only to a particular condition, were selected and, their expressions are shown in **Figure 6**.

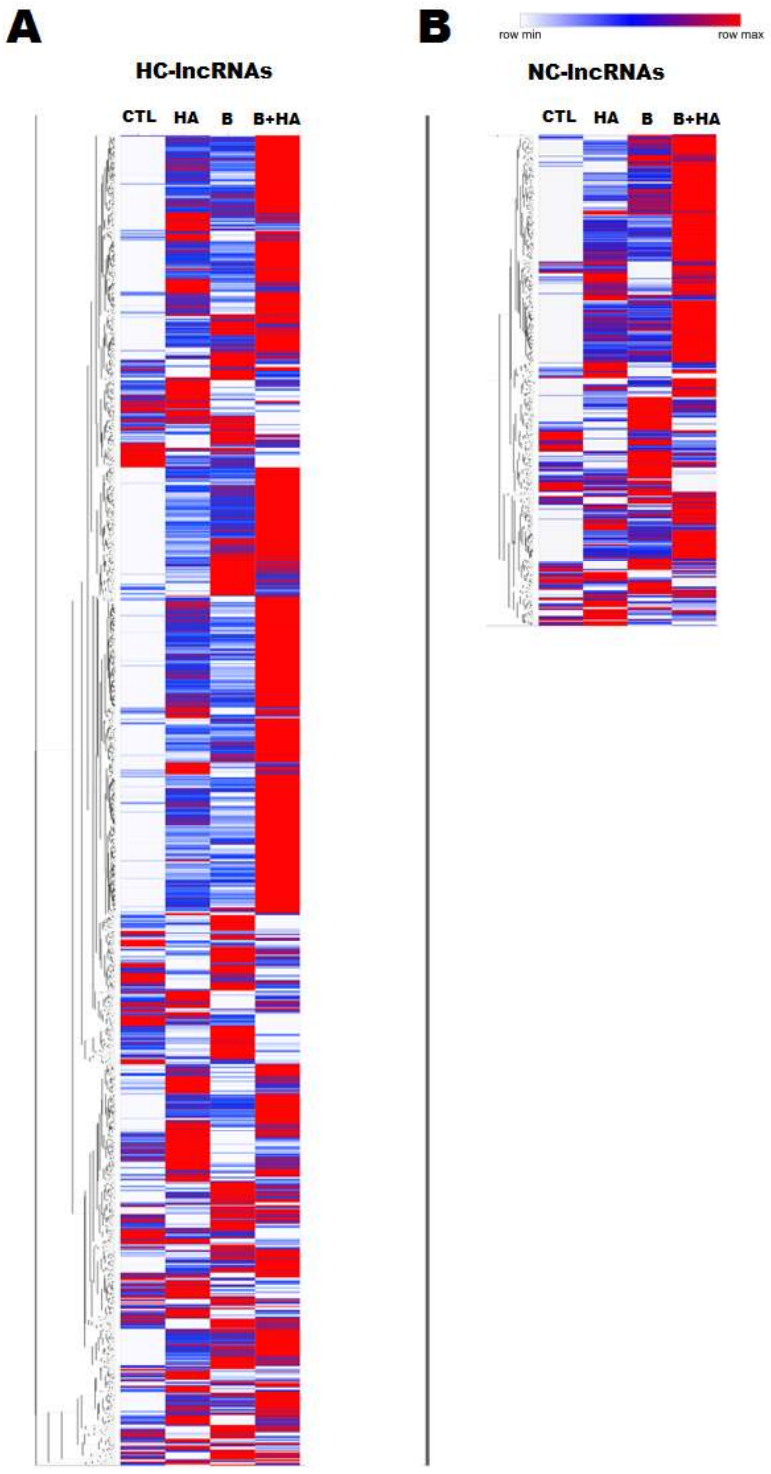


Figure 6 Expression profile of lncRNAs identified as differentially expressed in at least one specific condition. **(A)** 1,365 HC-lncRNAs and **(B)** 504 NC-lncRNAs.

Considering the differentially expressed HC-lncRNAs selected from Venn diagrams, we observed that 1,173 were upregulated and 631 were downregulated (**Figure 6A**). After removing the transcripts that were repeated between treatments, we obtained a set of 1,365 lncRNAs differentially

expressed in at least one comparison. For NC-lncRNAs, 456 lncRNAs were upregulated and 206 downregulated. A total of 504 differentially expressed lncRNAs were identified in at least one comparison (**Figure 6B**). The number of upregulated lncRNAs is practically double the number of downregulated ones for both HC-lncRNAs and NC-lncRNAs. Observing the HC-lncRNAs and the NC-lncRNAs, we noticed that a large part of them have their expression upregulated in the B+HA combination, suggesting a combinatorial effect of the treatments.

3.3 LncRNAs-mRNA-miRNA interaction networks

Several lncRNAs have been identified as precursors of microRNAs (miRNAs) and small interfering RNAs (siRNAs) (Reinhart et al. 2002; Hirsch et al. 2006), besides acting as decoys or targets of miRNAs. To explore the functional role of lncRNAs in response to *H. seropedicae* inoculation and humic acid treatment, we first identified lncRNAs that can interact with miRNA regulation. We have identified 20 lncRNAs that function as targets for 41 different miRNAs from 12 families (miR171, miR156, miR164, miR159, miR319, miR167, miR169, miR396, miR11970, miR394, miR395, and miR393). Of the 20 lncRNAs that function as targets, 17 are HC-lncRNAs, and three are NC-lncRNAs **(Table 1**). We found the most of the lncRNAs were up-regulated in B/CTL comparison and down-regulated in HA/B, suggesting again that the bacteria induces the expression of these lncRNAs.

Table 1 The 20 lncRNAs identified as miRNAs targets and its regulations.

lncRNA	miRNA	Classification	log2 (FC) per comparison					
			B/CTL	HA/CTL	B+HA/CTL	B+HA/B	B+HA/HA	HA/B
MSTRG.28442.1	miR156	HC-lncRNA		2,13	5,24	3,26	2,87	
MSTRG.30248.1	miR159	HC-lncRNA	-2,71	-4,17	-3,98			
MSTRG.1701.4	miR159	HC-lncRNA	3,60		7,27	2,11	2,81	
MSTRG.1345.1	miR164	HC-lncRNA		-1,98	-2,76	-3,13	-0,78	-2,33
MSTRG.28681.1	miR164	HC-lncRNA	4,25	2,70	5,20	0,81	2,37	-1,59
MSTRG.6488.1	miR164	HC-lncRNA	2,10					-3,03
MSTRG.22772.1	miR167	HC-lncRNA	5,92		7,91		5,08	-4,51
MSTRG.32740.1	miR169	HC-lncRNA	1,57		-2,77	-4,17		-2,88
MSTRG.35167.2	miR169	HC-lncRNA	2,99			-1,25	2,17	-3,64
MSTRG.13870.1	miR169	HC-lncRNA	4,53			-2,45		-3,65
MSTRG.6199.1	miR171	HC-lncRNA	-2,99	-1,74	-3,60			
MSTRG.6278.1	miR171	HC-lncRNA			3,72	4,22	1,71	
MSTRG.27679.1	miR171	HC-lncRNA	3,98	2,88				
MSTRG.9813.3	miR393	HC-lncRNA	2,91			-4,05		-2,16
MSTRG.17640.2	miR394	HC-lncRNA	3,75		2,98			-2,31
MSTRG.17609.8	miR395	HC-lncRNA	4,23		6,18	1,46	3,24	-1,86
MSTRG.7192.1	miR396	HC-lncRNA	3,74		4,99		2,64	-2,43
MSTRG.10242.4	miR159	NC-lncRNA					-1,81	3,38
MSTRG.3891.1	miR167	NC-lncRNA	1,12	-1,67	-2,97	-3,98		-2,73
MSTRG.26611.2	miR11970	NC-lncRNA	3,06		2,63			-1,80

After predicting lncRNAs-miRNA targets, the lncRNAs acting as miRNAs decoys were identified (**Supplementary Table 2**). We have identified three HC-lncRNAs (MSTRG.13897.1, MSTRG.30197.3, and MSTRG.34402.1) acting as decoys for 11 miRNAs from three other families (miR399, miR169, and miR319, respectively). The lncRNAs were modeled using RNAfold web server (Gruber et al. 2008), and their secondary structures are shown in **Figure 7,** together with the alignments of the respective miRNAs identified as decoys.

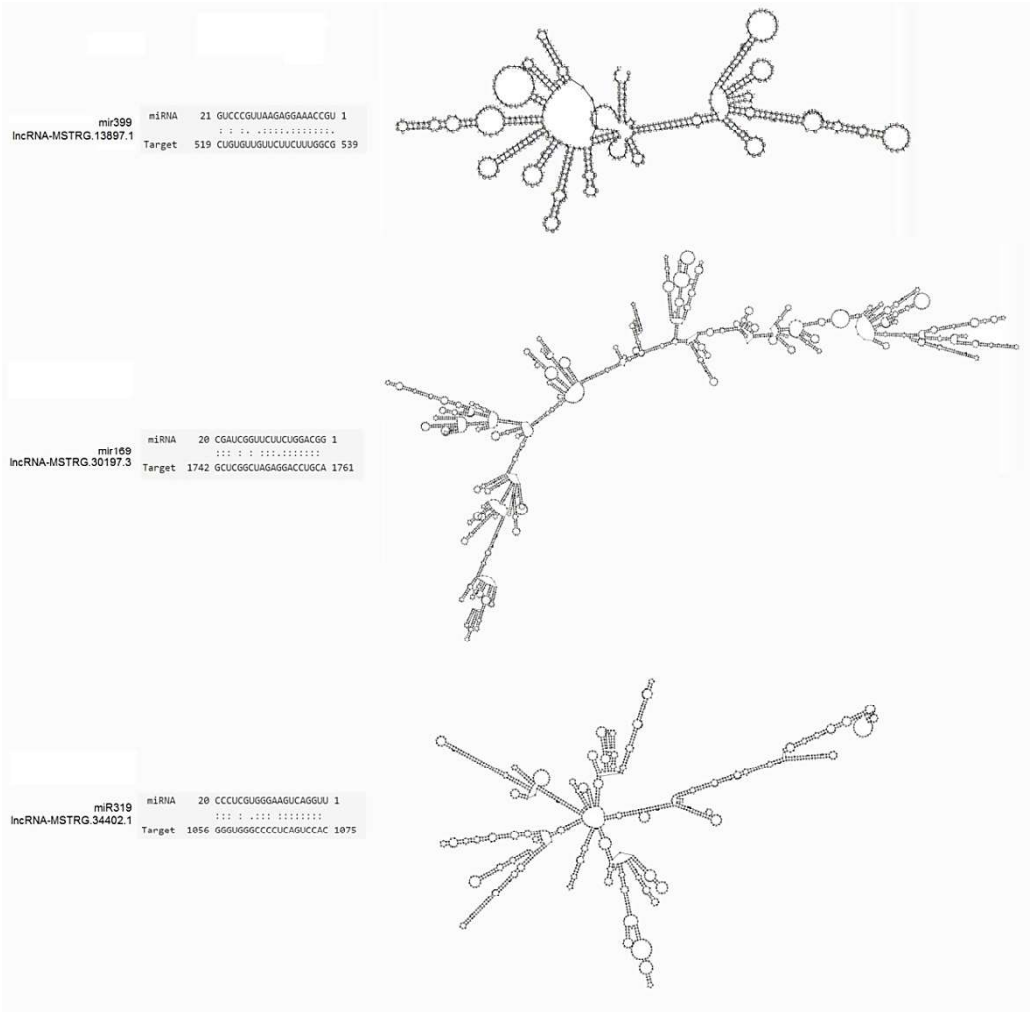


Figure 7 The three HC-lncRNAs identified as decoys of miRNAs in maize. Alignments showing gaps and mismatches.

The 20 lncRNAs predicted as targets of miRNAs were submitted to the LncTar program to predict their target mRNAs. Among the predicted mRNAs, the 10 with the lowest dG values for each of the analyzed lncRNAs were selected (**Supplementary Table 3**). The functional annotation of mRNAs predicted as lncRNA targets revealed an association with different processes, such as hormone signal transduction (MYB); abscisic acid (ABA) and ethylene response (ABA_WDS; AP2); Ca^{2+}-binding (EF-hand); auxin response (PIN9; SAUR); outer cell wall composition (GP1), and transmembrane signaling (TspO_MBR), among others.

For the 20 lncRNAs predicted as miRNA targets, miRNA-lncRNAs-mRNA interaction networks were constructed, which show the regulation mediated by miRNAs and lncRNAs in maize submitted to colonization by bacteria and/or humic acid treatment. **Figure 8** showed six up-regulated

lncRNAs related to hormonal regulation, targeting elements responsive to different hormones, such as ABA, auxins, and ethylene.

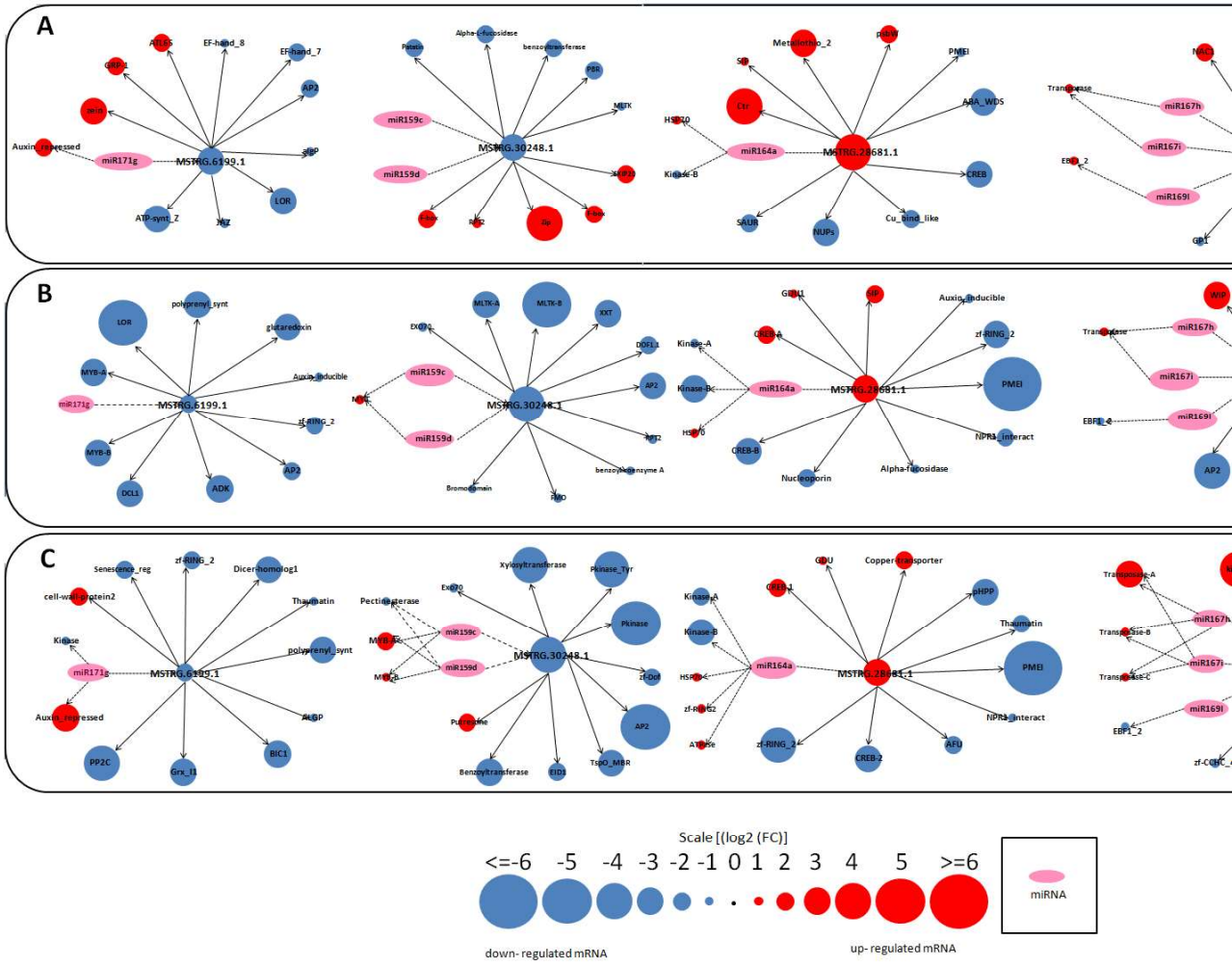


Figure 8 lncRNAs-mRNA-miRNA interaction networks of up- and down-regulated lncRNAs that appear to hormones. The networks were constructed with the lncRNA in the center, linked to the miRNA that cleave these networks, the regulation of lncRNAs was obtained by the comparison of treatments with the control co

Among the 20 lncRNAs identified as miRNA targets, 4 of them (MSTRG.6199.1, MSTRG.30248.1, MSTRG.28681.1 and MSTRG.3891.1) were differentially expressed in the three treatment *versus* control comparisons (B/CTL, HA/CTL and B +HA/CTL), allowing to study the contrasts of its regulation regarding these treatments.

MSTRG.6199.1 was downregulated in the three comparisons, having its expression suppressed by both bacteria and humic acid, and was predicted to be target of miR171g. This miRNA targets an auxin-related gene (named Auxin_repressed in the regulatory network), upregulated in the B/CTL comparison. Interestingly, this gene was not differentially expressed in the HA/CTL comparison. On the other hand, Auxin_repressed was upregulated in the B+HA/CTL comparison, in addition to another miR171g target (here called Kinase), downregulated. This gene is annotated as a non-specific serine/threonine protein kinase. Serine/threonine protein kinases are proteins responsible for phosphorylating the OH group of serine or threonine in other proteins. Auxin_repressed is annotated as DORMANCY/AUXIN ASSOCIATED FAMILY PROTEIN, and was upregulated in the combination of B+HA/CTL treatments. In *Arabidopsis*, this gene is a negative regulator of defense against virulent bacterial pathogens, with increased expression after inoculation with the pathogen *Pseudomonas syringae* (Roy et al. 2020).

MSTRG.30248.1 is also downregulated in the three comparisons (B/CTL, HA/CTL and B+HA/CTL), and was predicted to be target by miR159c/d, one of the targets of this miRNA being the MYB transcription factor gene (called MYB-A and MYB-B in the interaction network). In plants, MYB is related to several processes, including stress response and hormonal signal transduction. Two MYB genes were differentially expressed and upregulated in the B+HA/CTL comparison, and only one of them in the HA/CTL comparison. Another miR159c/d target gene that was differentially expressed only in the combination of B+HA/CTL treatments was the one noted as Pectinesterase/Pectin methylesterase (called Pectinesterase in the interaction network), downregulated. Pectin methylesterases are encoded by a multigene family, and are associated with cell wall modifications (Wu et al. 2018).

MSTRG.28681.1 was upregulated in the three comparisons (B/CTL, HA/CTL and B+HA/CTL), and was predicted to be a target of miR164a. A gene

encoding an Hsp70 protein (HSP70) was predicted to target miR164a. Hsp70 are heat shock proteins, ubiquitously expressed and highly conserved (Mayer and Bukau 2005). Hsp70 was upregulated in the three comparisons. In the combination of treatments, zf-RING_2, another miR164a target, was upregulated. Zinc finger proteins, like zf-RING_2, are proteins with a wide variety of molecular functions and can interact with DNA, RNA, and other proteins (Cassandri et al. 2017), suggesting a wide variety of possible functions within the cell. Still in the B+HA/CTL comparison, a gene annotated as Calcium-transporting ATPase (called ATPase in the regulatory network) was upregulated. Ca^{2+}-ATPases carry out the transport of ions across membranes, and are associated with the maintenance of intracellular calcium levels and the transmission of signals and, thus, modulating different metabolic pathways (Huda et al. 2013). Finally, two genes annotated as non-specific serine/threonine protein kinase (named Kinase-A and Kinase-B in the regulatory network) were downregulated in all treatments.

MSTRG.3891.1 was upregulated in the B/CTL comparison and downregulated in HA/CTL and B+HA/CTL, and was predicted to target miR167h/i and miR169l. miR167h/i targets genes annotated as Tam3-transposase (called Transposase-a/b/c in the interaction networks). Tam3, unlike the other transposable elements, is not epigenetically regulated in *Antirrhinum majus*, and its regulation is mediated by a BED-zinc finger domain of the gene itself (Zhou et al. 2017). The miR169l target, a gene annotated as EIN3-binding F-box protein (named EBF1_2 in the regulatory network) is related to ethylene regulation in *Arabidopsis*. This transcription factor senses ethylene and is related to different plant responses. In the absence of ethylene, different EIN proteins are ubiquitinated and degraded (Binder et al. 2007). In B/CTL, EBF1_2 was upregulated, while in HA/CTL and B+HA/CTL comparisons it was downregulated, suggesting that the miR169l is upregulated in these last two conditions.

The mRNAs predicted to be targets of the selected lncRNAs are associated with several different functions. In **Figure 8**, in the B/CTL comparison, we can see that 3 targets (zein, GRP-1 and ATL65) of MSTRG.6199.1 are upregulated, while another 7 are downregulated. The target annotated as Glycine-rich cell wall structural protein 2 (called GRP-1 in the

interaction network) belongs to a superfamily of genes associated with the response to different types of biotic and abiotic stress (Mangeon et al. 2010). The target annotated as RING-H2 finger protein ATL65 (named ATL65 in the interaction network) belongs to the group of ring finger proteins, associated with important roles in the response to abiotic stress. In tomato, a ring finger gene was constitutively expressed in several tissues, including roots (Song et al. 2016). Among the downregulated targets, EF-HAND CALCIUM-BINDING DOMAIN CONTAINING PROTEIN (called EF-hand-7/8 in the interaction network) is a domain present in proteins involved in calcium (Ca^{2+}) regulation. Calcium is related to the signaling of different hormones and the plant's response to environmental stress (Day et al. 2002). Another target, the gene annotated as jasmonate ZIM domain-containing protein (named JAZ) is related to programmed cell death in tomato and *Nicotiana benthamiana* and acts as negative regulators of jasmonate signaling in *Arabidopsis* (Ishiga et al. 2013).

In the HA/CTL comparison, all targets of MSTRG.6199.1 were downregulated, including Auxin_inducible, DCL1, and AP2. For example, DCL1 (Dicer homolog 1) is involved in RNA-mediated post-transcriptional gene silencing, acting in the miRNA biosynthesis. AP2 (AP2/ERF domain) responds to ethylene, an endogenous hormone that influences different plant growth and development aspects. The gene annotated as AUXIN-REGULATED PROTEIN-RELATED) (named Auxin_inducible in the interaction network) belongs to a family of auxin-response genes. Auxin is involved in most of the pathways responsible for plant growth (Hagen and Guilfoyle 2002).

In the B+HA/CTL comparison, all targets of MSTRG.6199.1 were downregulated, with the exception of the gene annotated as glycine-rich cell wall structural protein 2 (called cell-wall protein2), as mentioned before, belongs to a superfamily of genes associated with the response to different types of biotic and abiotic stress. Two targets related to plant defense are worth mentioning: Geranylgeranyl Diphosphate Synthase (called GGPP) and a thaumatin. The silencing of GGPP in tobacco decreases plant resistance to the pathogen *Manduca sexta* (Jassbi et al. 2008), and thaumatin comprises a family of proteins related to the development of plants, fungi and animals, as well as the defense response (de Jesús-Pires et al. 2019). Thaumatins even have their expression induced in maize subjected to abiotic stress (Frendo et al. 1992).

In the B/CTL comparison, MSTRG.30248.1 has half of the targets downregulated and half upregulated. The ROOT PHOTOTROPISM PROTEIN 2 gene (called RPT2) and the ZIP zinc/iron transport family (Zip) gene were upregulated. RPT2, in *Arabidopsis*, is related to hypocotyl photosensitivity (Haga et al. 2015) and Zip belong to a family of metal transport proteins (Guerinot 2000). In the HA/CTL comparison, all targets were downregulated, including RPT2, upregulated in the B/CTL comparison. AP2, responsive to ethylene, is downregulated in both HA/CTL and B+HA/CTL comparisons. These expression profiles suggest that the miR159c/d could be upregulated in B/CTL, which will target the MSTRG.30248.1, releasing the RPT2, Zip, among others.

In the combined treatment B+HA/CTL, we found a single upregulated MSTRG.30248.1 target, a Putrescine N-hydroxycinnamoyltransferase (here called Putrescine), an enzyme that has caffeoyl-CoA and putrescine as substrates, and which catalyzes the reaction that produces CoA and N-caffeoylputrescine. Other interesting downregulated targets are Xylosyltransferase (Xyloglucan 6-alpha-D-xylosyltransferase) and TspO_MBR (TspO/MBR-related protein). Xylosyltransferases act on xylose, a monosaccharide obtained from the hydrolysis of xylan, the main fraction of hemicellulose. TspO_MBR, on the other hand, are members of a superfamily of proteins involved in transmembrane signaling.

In the B/CTL comparison, MSTRG.28681.1 has the ABA/WDS induced protein (ABA_WDS) gene and a SAUR family protein (SAUR) as downregulated targets. SAUR (small auxin up-regulated RNA) proteins are a family of auxin-responsive genes (Stortenbeker and Bemer 2019). Proteins with the ABA-Water Deficit Stress Domain (ABA-WDS) domain belong to the family of ASR proteins, involved in abiotic stress tolerance (Yacoubi et al. 2021). Among the upregulated genes, we have Metallothionein (named Metallothio_2) and a copper transporter (named Ctr). In maize, metallothioneins are small proteins related to the plant's response to heavy metals, in addition to participating in plant growth and development (Gao et al. 2022). The copper transporter comprises a family of copper transporters characterized in rice (Yuan et al. 2011). In the HA/CTL comparison, two downregulated genes are worth mentioning: Auxin_inducible and NPR1 interacting (NPR1_interact). NPR1 is associated with plant response to salicylic acid in *Arabidopsis*, participating in

acquired systemic resistance (Kuai et al. 2015). In the B+HA/CTL comparison, the COPPER TRANSPORTER 1-RELATED gene (called Copper-transporter in the interaction network) is upregulated, as is PROTEIN GLUTAMINE DUMPER (GDU), a plant-specific membrane protein (Pratelli et al. 2010). Among the downregulated targets, we have: zf-RING_2; thaumatin and NPR1_interact.

The lncRNA MSTRG.3891.1 was upregulated in the B/CTL comparison and downregulated in the HA/CTL and B+HA/CTL comparisons. Two of its targets, the auxin efflux carrier PIN9 and the tubulin-folding cofactor D (named tubulin in the interaction network), associated with tubulin (Lopez-Fanarraga et al. 2001), were upregulated in both the B/CTL and HA/CTL comparisons. AP2, responsive to ethylene, was downregulated in the three comparisons (B/CTL, HA/CTL and B+HA/CTL). GP1 (vegetative cell wall protein gp1 precursor) was downregulated both in the B/CTL comparison and in the HA/CTL comparison. Finally, observing the B+HA/CTL comparison, only one target of MSTRG.3891.1 was upregulated, the gene annotated as cysteine-rich receptor-like protein kinase 19 (called kinase in the interaction network).

4. Discussion

This study identified and characterized maize lncRNAs, particularly those associated with plant response to plant bioestimulants (*H. seropedicae* bacteria and humic acid). Bioinformatics analyses allowed the identification of 124,739 HC-lcnRNAs and 6,418 candidate NC-lncRNAs. The lncRNAs associated with the plant's response to the treatments were characterized, mainly, in terms of the following aspects: *i)* their expression levels; *ii)* their miRNAs and mRNAs targets and, *iii)* their potential molecular functions, inferred through the construction of interaction networks.

For a long time, the non-coding sequences of genomes were ignored by researchers, who focused only on the coding regions. For example, despite the extensive transcription of the human genome, supposedly only 2% of transcripts encode functional proteins (Birney et al. 2007). The other unidentified transcripts constituted the called “dark matter” of the genome (Karlik et al. 2019), and the lncRNAs come from these regions (Kung et al. 2013). It has even been suggested that lncRNAs constitute 80% of the total ncRNAs in the genome (Zhang et al. 2013).

Maize lncRNAs respond to different biotic and abiotic stresses as identified in other studies (Lv et al. 2016; Huanca-Mamani et al. 2018; Wang et al. 2018; Pang et al. 2019; Han et al. 2019a). However, the regulation of lncRNAs in maize colonized by *H. seropedicae* with humic acid was not explored yet. Our study used RNA-seq libraries obtained from maize root submitted to the inoculation with *H. seropedicae* and/or humic acid treatment to explore the potential role of lncRNAs in response to these conditions. After assembling the transcriptome, 6,418 transcripts out of 192,521 were identified as new lncRNAs (3.3% of the transcriptome). We also constructed a comprehensive dataset of known HC-lncRNAs, representing 16.8% of the transcriptome.

The HC-lncRNAs and NC-lncRNAs showed characteristics described in the literature; most transcripts have between 200 nt and 300 nt (35%) and only one exon (80%). Furthermore, the expression of lncRNAs was lower than that of mRNAs. Comparison between HC-lncRNAs and NC-lncRNAs revealed important differences between them. Regarding size, 32% of NC-lncRNAs are between 200-300 nt in length. On the other hand, 33% of HC-lncRNAs are >4,000 nt. Most NC-lncRNAs have 1-2 exons, while 18% of HC-lncRNAs have more than 12 exons. These characteristics were observed by work with maize (Zhu et al. 2017) and are general characteristics of lncRNAs.

The analysis of the expression of lncRNAs with >1 TPM showed a gradual increase in the expression of the control for the combination of the bacteria with humic acid, which suggests that the combined use of the two treatments may have a synergistic effect on the expression of some lncRNAs. Regarding the treatments, the expression profile of HC-lncRNAs and NC-lncRNAs was similar, with a predominance of upregulated lncRNAs in B/CTL and B+HA/CTL. The bacteria appears to induce the expression of a greater number of lncRNAs when compared to humic acid. The same can be observed in the expression heatmap, where the combination of treatments (B+HA) presents a large number of lncRNAs with higher expression than in the other treatments or in the control. Wang et al. (2016) identified 867 novel lncRNAs in maize, many of which exhibited tissue-specific expression. The largest tissue-specific lncRNAs were identified in pollen and other smaller groups in the root, endosperm, tassel, and ear (Wang et al. 2016). In addition, our work identified

root-derived lncRNAs whose expression seems to be higher when maize is colonized by *H. seropedicae* and treated with humic acid.

Although thousands of lncRNAs have been identified to date, a significant portion of them has an unknown function (Sherstyuk et al. 2018). In this case, interaction networks can help understand the role interest-based on relationships with other regulatory elements. As suggested by Yang et al. (2019), lncRNAs can regulate the expression of mRNAs through competitive endogenous RNA. Similarly, maize lncRNAs may be controlling their target mRNAs through competition with miRNAs. In this context, it is possibly better to understand the functions of lncRNAs from their targets. Our results suggest that lncRNAs can act in signaling processes by regulating their mRNA targets or the targets of associated miRNAs. In the B/CTL comparison, the expression of MSTRG.6199.1 and their targets, suggested that the miR171g could be upregulated, cleaving this lncRNA, and releasing the Auxin_repressed (the miR171g target). This can induce the expression of GRP-1 e ATL65 in plants stimulated by the bacteria, which could impact their response to biotic and abiotic stresses. In work by Thiebaut et al. (2014), miR171 was shown to be up-regulated in maize colonized by endophytic beneficial diazotrophic bacteria. A similar profile was observed in B+HA/CTL, where the possible cleavage of MSTRG.6199.1 by miR171g, release the cell-wall protein2. The downregulation of Kinase by miR171g in B+HA/CTL reinforces the role of MSTRG.6199.1 in modulation of stresses responses when the plant is submitted to biostimulants analyzed here.

MiR159, which targets MSTRG.30248.1, was up-regulated in the presence of *Pseudomonas syringae* and also in maize colonized by the bacteria *H. seropedicae* (Thiebaut et al. 2014). miR159c/d targets a MYB transcription factor gene. MYB constitutes a large family of transcription factors related to the most diverse functions, such as response to biotic and abiotic stress, defense, metabolism, among others (Ambawat et al. 2013). Interestingly, our results showed that the MSTRG.30248.1 could be targeted by the miR159c/d in plants submitted to the bacteria, releasing the RPT2, Zip, among others. In contrast, plants submitted to B+HA the MSTRG.30248.1 and Pectinesterase seems to be preferentially cleaved by the miR159 c/d, increasing the expression of Putrescine, which is related to ABA responses and development. This suggests

that bacteria and humic acid in combination could modulate the cell wall structure and inducing plant development by regulate lncRNAs.

Genes related to hormone response were associated with different lncRNAs. MSTRG.28681.1 targets different genes related to plant response to ABA and auxins, suggesting its participation in these hormones' processes. Curá et al. (2017) observed that the inoculation of maize with *Azospirillum* sp. and *Herbaspirillum* sp. increased maize drought tolerance. In inoculated plants, the ZmVP14 gene involved in ABA synthesis showed low expression (Curá et al. 2017). MSTRG.28681.1 was upregulated in B, HA and B+HA treatments compared to control plants. However, the downregulation of ABA_WDS and SAUR was only found in B treatment. In HA and B+HA, the downregulated targets of MSTRG.28681.1 were related to biotic stress responses. In the MSTRG.28681.1 network, the miR164a targets were more diversified in B+HA treatment, although bacteria and humic acid treatments regulated genes related to abiotic and biotic stress (e.g., HSP70 and Kinase).

Other hormone-related genes were identified as targets for MSTRG.3891.1 - AP2, associated to ethylene and PIN9, related to auxin transport. The miR169l targets EBF1_2 and also MSTRG.3891.1. Interestingly, the induction of EBF1_2 was only found In B treatment, suggesting that the miR169l is downregulated by the bacteria. Thiebaut et al. (2014) found the miR169 downregulated in plants inoculated with *H. seropedicae*.

The effects of colonization by beneficial bacteria or humic acid on plants have been investigated for some time. This work shows that maize lncRNAs may be associated with several biological processes, emphasizing regulatory elements related to hormones and comparing the responses mobilized when a biostimulant is placed in contact with a plant. Finally, our results showed the participation of lncRNAs and miRNAs in the plant network response to bacteria and humic acid, with lncRNAs targeting target genes associated with different processes, such as resistance to biotic and abiotic stress, and plant hormones signaling.

5. Conclusions and perspectives

Considering climate change and the growing demand for food, it is necessary to characterize better the molecular mechanisms of plant responses

to a/biotic stress and the plant-beneficial microorganism interaction. In this scenario, the identification and characterization of lncRNAs, molecules whose roles in the regulation of critical biological processes has been evidenced by several studies, can help to understand the interaction with other coding and non-coding molecules, defining communication networks that can be manipulated to the improvement and obtaining of new cultivars.

We identified and characterized lncRNAs in maize submitted to colonization by *H. seropedicae* and humic acid addition; identified miRNAs related to these lncRNAs, and obtained interaction networks, associating these lncRNAs with different biological processes, including elements involved in the plant's response to hormones, such as ABA and auxins. The regulation of MSTRG.6199.1 targets (GRP-1; ATL65 and cell-wall protein2) suggests the participation of this lncRNA in the plant response to biotic and abiotic stresses. In the case of lncRNA MSTRG.30248.1, some of its targets suggest involvement in the response to ABA, while MSTRG.3891.1 has auxin-associated targets. The downregulated targets of MSTRG.28681.1 were related to biotic stress responses, suggesting the participation of this lncRNA in the plant response to the presence of the bacteria.

Finally, our results suggest that these lncRNAs are involved in the plant's response to biotic and abiotic stress, as well as in hormone-regulated processes. Furthermore, the participation of lncRNAs in the process of colonization by the bacteria *H. seropedicae* and with the addition of humic acid was evidenced by the regulation of different lncRNAs, which seem to respond to treatments. The interaction networks between lncRNAs-mRNAs-miRNAs also suggest an interesting conversation between gene expression regulatory elements.

We hope that the characterization of maize lncRNAs will contribute to a better understanding of the roles of these essential regulatory elements of gene expression, including plant responses to beneficial organisms. The molecular mechanisms behind plant colonization by endophytic microorganisms and response to humic acid can help to enhance their use to increase crop productivity.

Authors contribution statements CG and LOS conceived the project; CG and LOS analyzed the data and wrote the manuscript; LMC developed the script for identifying maize lncRNAs; FLO provided the maize RNA-seq libraries; MEMTW and FLO revised the manuscript. All authors approved the final version of the manuscript.

Declarations

Conflict of interests The authors declare no competing interests. The authors have no relevant financial or non-financial interests to disclose.

References

Acosta-García G, Vielle-Calzada JP (2004) A classical arabinogalactan protein is essential for the initiation of female gametogenesis in *Arabidopsis*. Plant Cell 16:2614–2628. https://doi.org/10.1105/tpc.104.024588

Adesemoye AO, Kloepper JW (2009) Plant-microbes interactions in enhanced fertilizer-use efficiency. Appl Microbiol Biotechnol 85:1–12. https://doi.org/10.1007/s00253-009-2196-0

Aguiar NO, Medici LO, Olivares FL, et al (2016) Metabolic profile and antioxidant responses during drought stress recovery in sugarcane treated with humic acids and endophytic diazotrophic bacteria. Ann Appl Biol 168:203–213. https://doi.org/10.1111/aab.12256

Aguiar NO, Olivares FL, Novotny EH, Canellas LP (2018) Changes in metabolic profiling of sugarcane leaves induced by endophytic diazotrophic bacteria and humic acids. PeerJ 2018:1–28. https://doi.org/10.7717/peerj.5445

Ambawat S, Sharma P, Yadav NR, Yadav RC (2013) MYB transcription factor genes as regulators for plant responses: An overview. Physiol Mol Biol Plants 19:307–321. https://doi.org/10.1007/s12298-013-0179-1

Amor B Ben, Wirth S, Merchan F, et al (2009) Novel long non-protein coding RNAs involved in Arabidopsis differentiation and stress responses. Genome Res 19:57–69. https://doi.org/10.1101/gr.080275.108

Atilio B, Causin HUF (1996) The central role of amino acids on nitrogen utilization and plant growth. J Plant Physiol 149:358–362. https://doi.org/10.1016/s0176-1617(96)80134-9

Axtell MJ (2013) Classification and comparison of small RNAs from plants. Annu Rev Plant Biol 64:137–159. https://doi.org/10.1146/annurev-arplant-050312-120043

Azevedo IG, Olivares FL, Ramos AC, et al (2019) Humic acids and *Herbaspirillum seropedicae* change the extracellular H+ flux and gene expression in maize roots seedlings. Chem Biol Technol Agric 6:1–10. https://doi.org/10.1186/s40538-019-0149-0

Baldani JI, Baldani VLD, Seldin L, Dobereiner ANDJ (1986) Characterization of *Herbaspirillurn seropedicae* gen. nov. sp. nov. a Root- Associated Nitrogen-Fixing Bacterium. Int J Syst Bacteriol 36:86–93

Baldani JI, Pot B, Kirchhof G, et al (1996) Emended description of *Herbaspirillum*; inclusion of [Pseudomonas] rubrisubalbicans, a mild plant pathogen, as *Herbaspirillum rubrisubalbicans* comb. nov.; and classification of a group of clinical isolates (EF Group 1) as H. Int J Syst Bacteriol 46:802–810

Bartel DP (2009) MicroRNAs: Target Recognition and Regulatory Functions. Cell 136:215–233. https://doi.org/10.1016/j.cell.2009.01.002

Bhattacharjee RB, Singh A, Mukhopadhyay SN (2008) Use of nitrogen-fixing bacteria as biofertiliser for non-legumes: Prospects and challenges. Appl Microbiol Biotechnol 80:199–209. https://doi.org/10.1007/s00253-008-1567-2

Bhattacharya S, Bunick CG, Chazin WJ (2004) Target selectivity in EF-hand calcium binding proteins. Biochim Biophys Acta - Mol Cell Res 1742:69–79. https://doi.org/10.1016/j.bbamcr.2004.09.002

Binder BM, Walker JM, Gagne JM, et al (2007) The *Arabidopsis* EIN3 binding F-box proteins EBF1 and EBF2 have distinct but overlapping roles in ethylene signaling. Plant Cell 19:509–523. https://doi.org/10.1105/tpc.106.048140

Birney E, Stamatoyannopoulos JA, Dutta A, et al (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 447:799–816. https://doi.org/10.1038/nature05874

Boerner S, McGinnis KM (2012) Computational identification and functional predictions of long noncoding RNA in *Zea mays*. PLoS One 7:. https://doi.org/10.1371/journal.pone.0043047

Brown P, Saa S (2015) Biostimulants in agriculture. Front Plant Sci 6:671-. https://doi.org/10.1111/mec.12571

Bulgari R, Cocetta G, Trivellini A, et al (2014) Biostimulants and crop responses: A review. Biol Agric Hortic 31:1–17. https://doi.org/10.1080/01448765.2014.964649

Camargo AP, Sourkov V, Pereira GAG, Carazzolle MF (2020) RNAsamba: neural network-based assessment of the protein-coding potential of RNA sequences. NAR Genomics Bioinforma 2:1–12. https://doi.org/10.1093/nargab/lqz024

Canellas LP, Balmori DM, Médici LO, et al (2013) A combination of humic substances and *Herbaspirillum seropedicae* inoculation enhances the growth of maize (*Zea mays* L.). Plant Soil 366:119–132. https://doi.org/10.1007/s11104-012-1382-5

Canellas LP, da Silva SF, Olk DC, et al (2015) Foliar application of plant growth-promoting bacteria and humic acid increase maize yields. J Food, Agric Environ 13:131–138

Canellas LP, Olivares FL, Okorokova-Façanha AL, Façanha AR (2002) Humic acids isolated from earthworm compost enhance root elongation, lateral root emergence, and plasma membrane H+-ATPase activity in maize roots. Plant Physiol 130:1951–1957. https://doi.org/10.1104/pp.007088

Cao H, Xu D, Cai Y, et al (2021) Very long intergenic non-coding (vlinc) RNAs directly regulate multiple genes in cis and trans. BMC Biol 19:1–24. https://doi.org/10.1186/s12915-021-01044-x

Chen D, Shao Q, Yin L, et al (2019) Polyamine function in plants: Metabolism, regulation on development, and roles in abiotic stress responses. Front Plant Sci 9:1–13. https://doi.org/10.3389/fpls.2018.01945

Chen L, Zhu QH, Kaufmann K (2020) Long non-coding RNAs in plants: emerging modulators of gene activity in development and stress responses. Planta 252:1–14. https://doi.org/10.1007/s00425-020-03480-5

Cheng J, Kapranov P, Drenkow J, et al (2005) Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. Science (80-) 308:1149–1154. https://doi.org/10.1126/science.1108625

Cocking EC (2003) Endophytic colonization of plant roots by nitrogen-fixing bacteria. Plant Soil 252:169–175. https://doi.org/10.1023/A:1024106605806

Conab (2021) Perspectivas para a agropecuária - Safra 2021/22. In: Conab. https://www.conab.gov.br/institucional/publicacoes/perspectivas-para-aagropecuaria/%0Aitem/download/38908_7e9804e5a9e69f38adbdba8271f31918

Curá JA, Franz DR, Filosofía JE, et al (2017) Inoculation with *Azospirillum* sp. and *Herbaspirillum* sp. bacteria increases the tolerance of maize to drought stress. Microorganisms 5:. https://doi.org/10.3390/microorganisms5030041

Dai W, Chen X, Wang X, et al (2018a) The algicidal fungus *Trametes versicolor* F21a eliminating blue algae via genes encoding degradation enzymes and metabolic pathways revealed by transcriptomic analysis. Front Microbiol 9:1–11. https://doi.org/10.3389/fmicb.2018.00826

Dai X, Zhuang Z, Zhao PX (2018b) PsRNATarget: A plant small RNA target analysis server (2017 release). Nucleic Acids Res 46:W49–W54. https://doi.org/10.1093/nar/gky316

Dartora J, Marini D, Gonçalves EDV, Guimarães VF (2016) Co-inoculation of *Azospirillum brasilense* and *Herbaspirillum seropedicae* in maize. Rev Bras Eng Agric e Ambient 20:545–550. https://doi.org/10.1590/1807-1929/agriambi.v20n6p545-550

Day IS, Reddy VS, Shad Ali G, Reddy AS (2002) Analysis of EF-hand-containing proteins in *Arabidopsis*. Genome Biol 3:1–24

Datta R, Paul S (2019) Long non-coding RNAs: Fine-tuning the developmental responses in plants. J Biosci 44:1–11. https://doi.org/10.1007/s12038-019-9910-6

Derrien T, Johnson R, Bussotti G, et al (2012) The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. Genome Res 22:1775–1789. https://doi.org/10.1101/gr.132159.111

de Jesús-Pires C, Ferreira-Neto JRC, Pacifico Bezerra-Neto J, et al (2019) Plant Thaumatin-like Proteins: Function, Evolution and Biotechnological Applications. Curr Protein Pept Sci 21:36–51. https://doi.org/10.2174/1389203720666190318164905

du Jardin P (2015) Plant biostimulants: Definition, concept, main categories and regulation. Sci Hortic (Amsterdam) 196:3–14. https://doi.org/10.1016/j.scienta.2015.09.021

Duret L, Chureau C, Samain S, et al (2006) The Xist RNA gene evolved in eutherians by pseudogenization of a protein-coding gene. Science (80-) 312:1653–1655. https://doi.org/10.1126/science.1126316

Fageria NK, Baligar VC (2005) Enhancing Nitrogen Use Efficiency in Crop Plants. Adv Agron 88:97–185. https://doi.org/10.1016/S0065-2113(05)88004-6

Fan C, Hao Z, Yan J, Li G (2015) Genome-wide identification and functional analysis of lincRNAs acting as miRNA targets or decoys in maize. BMC Genomics 16:. https://doi.org/10.1186/s12864-015-2024-0

Fanucchi S, Fok ET, Dalla E, et al (2019) Immune genes are primed for robust transcription by proximal long noncoding RNAs located in nuclear compartments. Nat Genet 51:138–150. https://doi.org/10.1038/s41588-018-0298-2

Fok ET, Scholefield J, Fanucchi S, Mhlanga MM (2017) The emerging

molecular biology toolbox for the study of long noncoding RNA biology. Epigenomics 9:1317–1327. https://doi.org/10.2217/epi-2017-0062

Frendo P, Didierjean L, Passelegue E, Burkard G (1992) Abiotic stresses induce a thaumatin-like protein in maize; cDNA isolation and sequence analysis. Plant Sci 85:61–69. https://doi.org/10.1016/0168-9452(92)90094-3

Gallart AP, Pulido AH, De Lagrán IAM, et al (2016) GREENC: A Wiki-based database of plant lncRNAs. Nucleic Acids Res 44:D1161–D1166. https://doi.org/10.1093/nar/gkv1215

Gao C, Gao K, Yang H, et al (2022) Genome-wide analysis of metallothionein gene family in maize to reveal its role in development and stress resistance to heavy metal. Biol Res 55:1–13. https://doi.org/10.1186/s40659-021-00368-w

Gao J, Luo M, Peng H, et al (2019) Characterization of cadmium-responsive MicroRNAs and their target genes in maize (*Zea mays*) roots. BMC Mol Biol 20:1–9. https://doi.org/10.1186/s12867-019-0131-1

Gao X, Wang C, Dai W, et al (2017) Proteomic analysis reveals large amounts of decomposition enzymes and major metabolic pathways involved in algicidal process of *Trametes versicolor* F21a. Sci Rep 7:3–12. https://doi.org/10.1038/s41598-017-04251-1

German MA, Luo S, Schroth G, et al (2009) Construction of parallel analysis of rna ends (Pare) libraries for the study of cleaved mirna targets and the rna degradome. Nat Protoc 4:356–362. https://doi.org/10.1038/nprot.2009.8

German MA, Pillay M, Jeong DH, et al (2008) Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. Nat Biotechnol 26:941–946. https://doi.org/https://doi.org/10.1038/nbt1417

Griffiths-Jones S, Bateman A, Marshall M, et al (2003) Rfam: An RNA family database. Nucleic Acids Res 31:439–441. https://doi.org/10.1093/nar/gkg006

Guerinot M Lou (2000) The ZIP family of metal transporters. Biochim Biophys Acta 190–198

Gurjar AKS, Panwar AS, Gupta R, Mantri SS (2016) PmiRExAt: Plant miRNA expression atlas database and web applications. Database 2016:1–10. https://doi.org/10.1093/database/baw060

Haga K, Tsuchida-Mayama T, Yamada M, Sakai T (2015) *Arabidopsis* ROOT PHOTOTROPISM2 contributes to the adaptation to high-intensity light in phototropic responses. Plant Cell 27:1098–1112. https://doi.org/10.1105/tpc.15.00178

Han G, Cheng C, Zheng Y, et al (2019a) Identification of long non-coding RNAs and the regulatory network responsive to arbuscular mycorrhizal fungi colonization in maize roots. Int J Mol Sci 20:1–15. https://doi.org/10.3390/ijms20184491

Hagen G, Guilfoyle T (2002) Auxin-responsive gene expression: Genes, promoters and regulatory factors. Plant Mol Biol 49:373–385. https://doi.org/10.1023/A:1015207114117

Han L, Mu Z, Luo Z, et al (2019b) New lncRNA annotation reveals extensive functional divergence of the transcriptome in maize. J Integr Plant Biol 61:394–405. https://doi.org/10.1111/jipb.12708

Hezroni H, Koppstein D, Schwartz MG, et al (2015) Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17

species. Physiol Behav 11:139–148. https://doi.org/10.1016/j.celrep.2015.04.023.
Hirsch J, Lefort V, Vankersschaver M, et al (2006) Characterization of 43 non-protein-coding mRNA genes in arabidopsis, including the MIR162a-derived transcripts. Plant Physiol 140:1192–1204. https://doi.org/10.1104/pp.105.073817
Huanca-Mamani W, Arias-Carrasco R, Cárdenas-Ninasivincha S, et al (2018) Long non-coding RNAs responsive to salt and boron stress in the hyper-arid lluteño maize from atacama desert. Genes (Basel) 9:. https://doi.org/10.3390/genes9030170
Huang W, Peng S, Xian Z, et al (2017) Overexpression of a tomato miR171 target gene SlGRAS24 impacts multiple agronomical traits via regulating gibberellin and auxin homeostasis. Plant Biotechnol J 15:472–488. https://doi.org/10.1111/pbi.12646
Huda KMK, Banu MSA, Tuteja R, Tuteja N (2013) Global calcium transducer P-type Ca2+-ATPases open new avenues for agriculture by regulating stress signalling. J Exp Bot 64:3099–3109. https://doi.org/10.1093/jxb/ert182
Hurni S, Scheuermann D, Krattinger SG, et al (2015) The maize disease resistance gene Htn1 against northern corn leaf blight encodes a wall-associated receptor-like kinase. Proc Natl Acad Sci U S A 112:8780–8785. https://doi.org/10.1073/pnas.1502522112
Ihn T, Hannett NM, Harbison CT, et al (2002) Transcriptional Regulatory Networks in *Saccharomyces cerevisiae*. Science (80-) 298:799–804. https://doi.org/10.1126/science.1075090
Ishiga Y, Ishiga T, Uppalapati SR, Mysore KS (2013) Jasmonate ZIM-Domain (JAZ) Protein Regulates Host and Nonhost Pathogen-Induced Cell Death in Tomato and Nicotiana benthamiana. PLoS One 8:. https://doi.org/10.1371/journal.pone.0075728
Jassbi AR, Gase K, Hettenhausen C, et al (2008) Silencing geranylgeranyl diphosphate synthase in Nicotiana attenuata dramatically impairs resistance to tobacco hornworm. Plant Physiol 146:974–986. https://doi.org/10.1104/pp.107.108811
Jia H, Osak M, Bogu GK, et al (2010) Genome-wide computational identification and manual annotation of human long noncoding RNA genes. Rna 16:1478–1487. https://doi.org/10.1261/rna.1951310
Kaikkonen MU, Lam MTY, Glass CK (2011) Non-coding RNAs as regulators of gene expression and epigenetics. Cardiovasc Res 90:430–440. https://doi.org/10.1093/cvr/cvr097
Kakrana A, Hammond R, Patel P, et al (2014) SPARTA: A parallelized pipeline for integrated analysis of plant miRNA and cleaved mRNA data sets, including new miRNA target-identification software. Nucleic Acids Res 42:1–13. https://doi.org/10.1093/nar/gku693
Kapranov P, Cheng J, Dike S, et al (2007) RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science (80-) 316:1484–1488. https://doi.org/10.1126/science.1138341
Karlik E, Ari S, Gozukirmizi N (2019) LncRNAs: genetic and epigenetic effects in plants. Biotechnol Biotechnol Equip 33:429–439. https://doi.org/10.1080/13102818.2019.1581085
Kim D, Langmead B, Salzberg SL (2015) HISAT: A fast spliced aligner with low memory requirements. Nat Methods 12:357–360.

https://doi.org/10.1038/nmeth.3317
Kim ED, Sung S (2012) Long noncoding RNA: Unveiling hidden layer of gene regulatory networks. Trends Plant Sci 17:16–21. https://doi.org/10.1016/j.tplants.2011.10.008
Kuai X, Mac Leod BJ, Després C (2015) Integrating data on the *Arabidopsis* NPR1/NPR3/NPR4 salicylic acid receptors; a differentiating argument. Front Plant Sci 6:1–5. https://doi.org/10.3389/fpls.2015.00235
Kung JTY, Colognori D, Lee JT (2013) Long noncoding RNAs: Past, present, and future. Genetics 193:651–669. https://doi.org/10.1534/genetics.112.146704
Lam HM, Coschigano KT, Oliveira IC, et al (1996) The molecular-genetics of nitrogen assimilation into amino acids in higher plants. Annu Rev Plant Physiol Plant Mol Biol 47:569–593. https://doi.org/10.1146/annurev.arplant.47.1.569
Li J, Ma W, Zeng P, et al (2014a) LncTar: A tool for predicting the RNA targets of long noncoding RNAs. Brief Bioinform 16:806–812. https://doi.org/10.1093/bib/bbu048
Li L, Eichten SR, Shimizu R, et al (2014b) Genome-wide discovery and characterization of maize long non-coding RNAs. Genome Biol 2014 3:176–190. https://doi.org/https://doi.org/10.1186/gb-2014-15-2-r40
Li W, Xiang F, Zhong M, et al (2017) Transcriptome and metabolite analysis identifies nitrogen utilization genes in tea plant (*Camellia sinensis*). Sci Rep 7:1–12. https://doi.org/10.1038/s41598-017-01949-0
Liao Y, Smyth GK, Shi W (2014) FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30:923–930. https://doi.org/10.1093/bioinformatics/btt656
Liu J, Jung C, Xu J, et al (2012) Genome-wide analysis uncovers regulation of long intergenic noncoding RNAs in arabidopsisC W. Plant Cell 24:4333–4345. https://doi.org/10.1105/tpc.112.102855
Lopez-Fanarraga M, Avila J, Guasch A, et al (2001) Review: Postchaperonin tubulin folding cofactors and their role in microtubule dynamics. J Struct Biol 135:219–229. https://doi.org/10.1006/jsbi.2001.4386
Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:1–21. https://doi.org/10.1186/s13059-014-0550-8
Lv Y, Liang Z, Ge M, et al (2016) Genome-wide identification and functional prediction of nitrogen-responsive intergenic and intronic long non-coding RNAs in maize (*Zea mays* L.). BMC Genomics 17:1–15. https://doi.org/10.1186/s12864-016-2650-1
Ma X, Yin X, Tang Z, et al (2020) The RNA degradome: a precious resource for deciphering RNA processing and regulation codes in plants. RNA Biol 17:1223–1227. https://doi.org/10.1080/15476286.2020.1757898
Mangeon A, Junqueira RM, Sachetto-Martins G (2010) Functional diversity of the plant glycine-rich proteins superfamily. Plant Signal Behav 5:99–104. https://doi.org/10.4161/psb.5.2.10336
Masclaux-Daubresse C, Daniel-Vedele F, Dechorgnat J, et al (2010) Nitrogen uptake, assimilation and remobilization in plants: Challenges for sustainable and productive agriculture. Ann Bot 105:1141–1157. https://doi.org/10.1093/aob/mcq028
Mayer MP, Bukau B (2005) Hsp70 chaperones: Cellular functions and

molecular mechanism. Cell Mol Life Sci 62:670–684. https://doi.org/10.1007/s00018-004-4464-6
Melo BAG, Motta FL, Santana MHA (2016) Humic acids: Structural properties and multiple functionalities for novel technological developments. Mater Sci Eng C 62:967–974. https://doi.org/10.1016/j.msec.2015.12.001
Monteiro RA, Balsanelli E, Wassem R, et al (2012) Herbaspirillum-plant interactions: Microscopical, histological and molecular aspects. Plant Soil 356:175–196. https://doi.org/10.1007/s11104-012-1125-7
Morange M (2008) What history tells us XIV. Regulation of gene expression by non-coding RNAs: The early steps. J Biosci 33:327–331. https://doi.org/10.1007/s12038-008-0051-6
Niazi F, Valadkhan S (2012) Computational analysis of functional long noncoding RNAs reveals lack of peptide-coding capacity and parallels with 3′ UTRs. Rna 18:825–843. https://doi.org/10.1261/rna.029520.111
Novoa R, Loomis RS (1981) Nitrogen and plant production. In: Plant Production. pp 177–204
Olk DC, Dinnes DL, Rene Scoresby J, et al (2018) Humic products in agriculture: potential benefits and research challenges—a review. J Soils Sediments 18:2881–2891. https://doi.org/10.1007/s11368-018-1916-4
Paci P, Colombo T, Farina L (2014) Computational analysis identifies a sponge interaction network between long non-coding RNAs and messenger RNAs in human breast cancer. BMC Syst Biol 8:1–15. https://doi.org/10.1186/1752-0509-8-83
Palumbo G, Schiavon M, Nardi S, et al (2018) Biostimulant potential of humic acids extracted from an amendment obtained via combination of olive mill wastewaters (OMW) and a pre-treated organic material derived from municipal solid waste (MSW). Front Plant Sci 9:1–14. https://doi.org/10.3389/fpls.2018.01028
Pang J, Zhang X, Ma X, Zhao J (2019) Spatio-temporal transcriptional dynamics of maize long non-coding RNAs responsive to drought stress. Genes (Basel) 10:. https://doi.org/10.3390/genes10020138
Panni S, Lovering RC, Porras P, Orchard S (2019) Non-coding RNA regulatory networks. BBA - Gene Regul Mech 194417. https://doi.org/10.1016/j.bbagrm.2019.194417
Parameswari P, Dhandapani V (2019) Plant Noncoding RNAs Identification and Characterization - A Short Review. Ann Adv Biomed Sci 2:. https://doi.org/10.23880/aabsc-16000134
Pauli A, Valen E, Lin MF, et al (2012) Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. Genome Res 22:577–591. https://doi.org/10.1101/gr.133009.111
Pertea M, Pertea GM, Antonescu CM, et al (2015) StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol 33:290–295. https://doi.org/10.1038/nbt.3122
Pratelli R, Voll LM, Horst RJ, et al (2010) Stimulation of nonselective amino acid export by glutamine dumper proteins. Plant Physiol 152:762–773. https://doi.org/10.1104/pp.109.151746
Reinhart BJ, Weinstein EG, Rhoades MW, et al (2002) Revealing micro-RNAs in plants. Trends Plant Sci 16:1616–1626. https://doi.org/10.1016/S1360-1385(02)02361-0
Rinn JL, Kertesz M, Wang JK, et al (2007) Functional Demarcation of Active

and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. Cell 129:1311–1323. https://doi.org/10.1016/j.cell.2007.05.022

Rosentrater KA (2011) Overview of Corn-Based Fuel Ethanol Coproducts: Production and Use. In: Biofuel's Engineering Process Technology. pp 141–166

Roy S, Saxena S, Sinha A, Nandi AK (2020) DORMANCY/AUXIN ASSOCIATED FAMILY PROTEIN 2 of *Arabidopsis thaliana* is a negative regulator of local and systemic acquired resistance. J Plant Res 133:409–417. https://doi.org/10.1007/s10265-020-01183-2

Ryan MC, Stucky M, Wakefield C, et al (2020) Interactive Clustered Heat Map Builder: An easy web-based tool for creating sophisticated clustered heat maps. F1000Research 8:. https://doi.org/10.12688/f1000research.20590.2

Salmena L, Poliseno L, Tay Y, et al (2011) A ceRNA hypothesis: The rosetta stone of a hidden RNA language? Cell 146:353–358. https://doi.org/10.1016/j.cell.2011.07.014

Sanchita, Trivedi PK, Asif MH (2020) Updates on plant long non-coding RNAs (lncRNAs): the regulatory components. Plant Cell Tissue Organ Cult 140:259–269. https://doi.org/10.1007/s11240-019-01726-z

Schaumann GE (2006) Soil organic matter beyond molecular structure Part I: Macromolecular and supramolecular characteristics. J Plant Nutr Soil Sci 169:145–156. https://doi.org/10.1002/jpln.200521785

Schneider HW, Raiol T, Brigido MM, et al (2017) A Support Vector Machine based method to distinguish long non-coding RNAs from protein coding transcripts. BMC Genomics 18:1–14. https://doi.org/10.1186/s12864-017-4178-4

Shadchina TM, Dmitrieva V V. (1995) Leaf Chlorophyll Content as a Possible Diagnostic Mean for the Evaluation of Plant Nitrogen Uptake from the Soil. J Plant Nutr 18:1427–1437. https://doi.org/10.1080/01904169509364992

Sherstyuk V V., Medvedev SP, Zakian SM (2018) Noncoding RNAs in the Regulation of Pluripotency and Reprogramming. Stem Cell Rev Reports 14:58–70. https://doi.org/10.1007/s12015-017-9782-9

Shuai P, Liang D, Tang S, et al (2014) Genome-wide identification and functional prediction of novel and drought-responsive lincRNAs in *Populus trichocarpa*. J Exp Bot 65:4975–4983. https://doi.org/10.1093/jxb/eru256

Simopoulos CMA, Weretilnyk EA, Golding GB (2019) Molecular traits of long non-protein coding RNAs from diverse plant species show little evidence of phylogenetic relationships. G3 Genes, Genomes, Genet 9:2511–2520. https://doi.org/10.1534/g3.119.400201

https://doi.org/10.1590/S1415-475738420150053

Song J, Xing Y, Munir S, et al (2016) An ATL78-Like RING-H2 finger protein confers abiotic stress tolerance through interacting with RAV2 and CSN5B in tomato. Front Plant Sci 7:1–11. https://doi.org/10.3389/fpls.2016.01305

St Laurent G, Shtokalo D, Dong B, et al (2013) VlincRNAs controlled by retroviral elements are a hallmark of pluripotency and cancer. Genome Biol 14:R73. https://doi.org/10.1186/gb-2013-14-7-r73

Statello L, Guo CJ, Chen LL, Huarte M (2021) Gene regulation by long non-coding RNAs and its biological functions. Nat Rev Mol Cell Biol 22:96–118. https://doi.org/10.1038/s41580-020-00315-9

Stortenbeker N, Bemer M (2019) The SAUR gene family: The plant's toolbox for adaptation of growth and development. J Exp Bot 70:17–27.

https://doi.org/10.1093/jxb/ery332
Szcześniak MW, Rosikiewicz W, Makałowska I (2016) CANTATAdb: A collection of plant long non-coding RNAs. Plant Cell Physiol 57:e8(1-7). https://doi.org/10.1093/pcp/pcv201
Tester M, Langridge P (2010) Breeding Technologies to Increase. Science (80-) 818:818–822
Thiebaut F, Rojas CA, Grativol C, et al (2014) Genome-wide identification of microRNA and siRNA responsive to endophytic beneficial diazotrophic bacteria in maize. BMC Genomics 15:1–18. https://doi.org/10.1186/1471-2164-15-766
Thomas M, Lieberman J, Lal A (2010) Desperately seeking microRNA targets. Nat Struct Mol Biol 17:1169–1174. https://doi.org/10.1038/nsmb.1921
Tills AR, Alloway BJ (1981) The Effect Of Ammonium And Nitrate Nitrogen Sources On Copper Uptake And Amino Acid Status Of Cereals. Plant Soil 290:279–290
Tilman D, Balzer C, Hill J, Befort BL (2011) Global food demand and the sustainable intensification of agriculture. Proc Natl Acad Sci U S A 108:20260–20264. https://doi.org/10.1073/pnas.1116437108
Trapnell C, Williams BA, Pertea G, et al (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 28:511–515. https://doi.org/10.1038/nbt.1621
Ulitsky I, Shkumatava A, Jan CH, et al (2011) Conserved Function of lincRNAs in Vertebrate Embryonic Development Despite Rapid Sequence Evolution. Cell 7:1–7. https://doi.org/10.1016/j.cell.2011.11.055.Conserved
Vieira LM, Grativol C, Thiebaut F, et al (2017) PlantRNA_sniffer: A SVM-based workflow to predict long intergenic non-coding RNAs in plants. Non-coding RNA 3:. https://doi.org/10.3390/ncrna3010011
Wang H, Chung PJ, Liu J, et al (2014) Genome-wide identification of long noncoding natural antisense transcripts and their responses to light in Arabidopsis. Genome Res 24:444–453. https://doi.org/10.1101/gr.165555.113
Wang H, Niu QW, Wu HW, et al (2015) Analysis of non-coding transcriptome in rice and maize uncovers roles of conserved lncRNAs associated with agriculture traits. Plant J 84:404–416. https://doi.org/10.1111/tpj.13018
Wang H, Wang H (2015) The miR156/SPL module, a regulatory hub and versatile toolbox, gears up crops for enhanced agronomic traits. Mol Plant 8:677–688. https://doi.org/10.1016/j.molp.2015.01.008
Wang J, Meng X, Dobrovolskaya OB, et al (2017) Non-coding RNAs and Their Roles in Stress Response in Plants Wang J et al / miRNA and lncRNA in Plant Stress Response. Genomics, Proteomics Bioinforma 15:301–312. https://doi.org/10.1016/j.gpb.2017.01.007
Wang KC, Yang YW, Liu B, et al (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature 472:120–126. https://doi.org/10.1038/nature09819
Wang H-L V., Chekanova JA (2017) Long Noncoding RNAs in Plants. Adv Exp Med Biol 1008:133–154. https://doi.org/10.1007/978-981-10-5203-3_5
Wang Y, Wang Y, Zhao J, et al (2018) Unveiling gibberellin-responsive coding and long noncoding RNAs in maize. Plant Mol Biol 98:427–438. https://doi.org/10.1007/s11103-018-0788-8

Wu HC, Bulgakov VP, Jinn TL (2018) Pectin methylesterases: Cell wall remodeling proteins are required for plant response to heat stress. Front Plant Sci 871:1–21. https://doi.org/10.3389/fpls.2018.01612

Wu Q, Kim YC, Lu J, et al (2008) Poly A- transcripts expressed in Hela cells. PLoS One 3:e2803. https://doi.org/10.1371/journal.pone.0002803

Xu G, Fan X, Miller AJ (2012) Plant nitrogen assimilation and use efficiency. Annu Rev Plant Biol 63:153–182. https://doi.org/10.1146/annurev-arplant-042811-105532

Yacoubi I, Hamdi K, Fourquet P, et al (2021) Structural and Functional Characterization of the ABA-Water Deficit Stress Domain from Wheat and Barley: An Intrinsically Disordered Domain behind the Versatile Functions of the Plant Abscissic Acid, Stress and Ripening Protein Family. Int J Mol Sci 22:1–26. https://doi.org/10.3390/ijms22052314

Yang Q, Balint-Kurti P, Xu M (2017a) Quantitative Disease Resistance: Dissection and Adoption in Maize. Mol Plant 10:402–413. https://doi.org/10.1016/j.molp.2017.02.004

Yang Q, He Y, Kabahuma M, et al (2017b) A gene encoding maize caffeoyl-CoA O-methyltransferase confers quantitative resistance to multiple pathogens. Nat Genet 49:1364–1372. https://doi.org/10.1038/ng.3919

Yruela I (2005) Copper in plants. Brazilian J Plant Physiol 17:145–156. https://doi.org/10.1590/s1677-04202005000100012

Yuan M, Li X, Xiao J, Wang S (2011) Molecular and functional analyses of COPT/Ctr-type copper transporter-like gene family in rice. BMC Plant Biol 11:. https://doi.org/10.1186/1471-2229-11-69

Yu F, Tan Z, Fang T, et al (2020) A comprehensive transcriptomics analysis reveals long non-coding RNA to be involved in the key metabolic pathway in response to waterlogging stress in maize. Genes (Basel) 11:1–19. https://doi.org/10.3390/genes11030267

Zhai J, Arikit S, Simon SA, et al (2014) Rapid construction of parallel analysis of RNA end (PARE) libraries for Illumina sequencing. Methods 67:84–90. https://doi.org/10.1016/j.ymeth.2013.06.025

Zhang B, Pan X, Anderson TA (2006) Identification of 188 conserved maize microRNAs and their targets. FEBS Lett 580:3753–3762. https://doi.org/10.1016/j.febslet.2006.05.063

Zhang J, Mujahid H, Hou Y, et al (2013) Plant Long ncRNAs: A New Frontier for Gene Regulatory Control. Am J Plant Sci 04:1038–1045. https://doi.org/10.4236/ajps.2013.45128

Zhang W, Han Z, Guo Q, et al (2014) Identification of maize long non-coding RNAs responsive to drought stress. PLoS One 9:e98958. https://doi.org/10.1371/journal.pone.0098958

Zhou H, Hirata M, Osawa R, et al (2017) Detainment of tam3 transposase at plasmamembrane by its BED-zinc finger domain. Plant Physiol 173:1492–1501. https://doi.org/10.1104/pp.16.00996

Zhu M, Zhang M, Xing L, et al (2017) Transcriptomic analysis of long non-coding RNAs and coding genes uncovers a complex regulatory network that is involved in maize seed development. Genes (Basel) 8:1–14. https://doi.org/10.3390/genes8100274

4. CONCLUSÕES GERAIS

A partir dos resultados obtidos concluímos que:

- Foram identificados 6.418 novos lncRNAs e 32.501 lncRNAs conhecidos no transcriptoma;
- Os lncRNAs identificados possuem características semelhantes àquelas já observadas em outros trabalhos: grande parte possui entre 200 e 300 nucleotídeos e possuem entre 1 e 2 éxons;
- Os mRNAs preditos como alvos dos lncRNAs estão associados a diferentes processos, como: resposta a fatores bióticos a abióticos; ligação a cálcio; defesa; resposta a hormônios (como ABA e auxinas); transporte de cobre, dentre outros;
- 20 lncRNAs foram preditos como alvos de miRNAs, e 3 lncRNAs como *decoys*;
- As redes de interação podem auxiliar na identificação das funções dos lncRNAs a partir das funções de seus alvos e dos miRNAs associados;
- lncRNAs de milho submetido à bactéria *H. seropedicae* com adição de ácido húmico possuem como alvos genes relacionados a vias reguladas por hormônios.

REFERÊNCIAS

ACHAWANANTAKUN, R. et al. LncRNA-ID: Long non-coding RNA IDentification using balanced random forests. **Bioinformatics**, v. 31, n. 24, p. 3897-3905, 2015.

AHMAD, I.; PICHTEL, J.; HAYAT, S. **Plant-Bacteria Interactions**. 1. ed. Weinheim: Wiley&VCH Verlag GmbH & Co. KGaA, 2008.

AMASINO, R. 1955: Kinetin arrives. The 50th anniversary of a new plant hormone.**Plant Physiology**, v. 138, n. 3, p. 1177-1184, 2005.

ARIEL, F. et al. Noncoding transcription by alternative RNA polymerases dynamically regulates an auxin-driven chromatin loop. **Molecular Cell**, v. 55, n. 3, p. 383-396, 2014.

ARORA, N. K.; KHARE, E.; MAHESHWARI, D. K. Plant Growth Promoting Rhizobacteria: Constraints in Bioformulation, Commercialization, and Future Strategies. In: MAHESHWARI, D. K. (Ed.). **Plant Growth and Health Promoting Bacteria**. 1. ed. Berlin/Heidelberg: Springer-Verlag Berlin Heidelberg, v. 18, p. 251-272, 2011.

ARRIAL, R. T.; TOGAWA, R. C.; MARCELO, M. M. Screening non-coding RNAs in transcriptomes from neglected species using PORTRAIT: Case study of the pathogenic fungus *Paracoccidioides brasiliensis*. **BMC Bioinformatics**, v. 10, p. 1-9, 2009.

AUGÉ, R. M. Water relations mycorrhiza review. **Mycorrhiza**, v. 11, p. 3-42, 2001.

BABALOLA, O. O. Beneficial bacteria of agricultural importance. **Biotechnology Letters**, v. 32, n. 11, p. 1559-1570, 2010.

BAEK, J. et al. LncRNAnet: Long non-coding RNA identification using deep learning. **Bioinformatics**, v. 34, n. 22, p. 3889-3897, 2018.

BAI, Y. et al. Identification and characterization of pineapple leaf lncRNAs in crassulacean acid metabolism (CAM) photosynthesis pathway. **Scientific Reports**, v. 9, n. 1, p. 1-11, 2019a.

BAI, Y. et al. PlncRNADB: A Repository of Plant lncRNAs and lncRNA-RBP Protein Interactions. **Current Bioinformatics**, v. 14, n. 7, p. 621-627, 2019b.

BACKER, R. et al. Plant growth-promoting rhizobacteria: Context, mechanisms of action, and roadmap to commercialization of biostimulants for sustainable agriculture. **Frontiers in Plant Science**, v. 871, n. October, p. 1-17, 2018.

BARDOU, F. et al. Dual RNAs in plants. **Biochimie**, v. 93, n. 11, p. 1950-1954, 2011.

BARDOU, F. et al. Long Noncoding RNA Modulates Alternative Splicing Regulators in *Arabidopsis*. **Developmental Cell**, v. 30, n. 2, p. 166-176, 2014.

BAZIN, J. et al. Nuclear speckle rna binding proteins remodel alternative splicing and the non-coding *Arabidopsis* transcriptome to regulate a cross-talk between auxin and immune responses. **Frontiers in Plant Science**, v. 9, n. August, p. 1-13, 2018.

BAZIN, J.; BAILEY-SERRES, J. Emerging roles of long non-coding RNA in root developmental plasticity and regulation of phosphate homeostasis. **Frontiers in Plant**

Science, v. 6, n. June, p. 1-6, 2015.

BEHIE, S. W.; BIDOCHKA, M. J. Nutrient transfer in plant-fungal symbioses. **Trends in Plant Science**, v. 19, n. 11, p. 734-740, 2014.

BEN AMOR, B. et al. Novel long non-protein coding RNAs involved in *Arabidopsis* differentiation and stress responses. **Genome Research**, v. 19, n. 1, p. 57-69, 2009.

BERENDSEN, R. L.; PIETERSE, C. M. J.; BAKKER, P. A. H. M.The rhizosphere microbiome and plant health. **Trends in Plant Science**, v. 17, n. 8, p. 478-486, 2012.

BERG, G. et al. Unraveling the plant microbiome: Looking back and future perspectives. **Frontiers in Microbiology**, v. 5, n. JUN, p. 1-7, 2014.

BERNAL, A.; EAR, U.; KYRPIDES, N. Genomes Online Database (GOLD): A monitor of complete and ongoing genome projects world-wide. **Bioinformatics**, v. 29, p. 126-127, 2001.

BHATTACHARYYA, P. N.; JHA, D. K. Plant growth-promoting rhizobacteria (PGPR): Emergence in agriculture. **World Journal of Microbiology and Biotechnology**, v. 28, n. 4, p. 1327-1350, 2012.

BISHOPP, A.; HELP, H.; HELARIUTTA, Y. Chapter 1 Cytokinin Signaling During Root Development. In: **International Review of Cell and Molecular Biology**. 1. ed., Elsevier Inc., v. 276, p.1-48, 2009.

BITTERLICH, M. et al. Arbuscular mycorrhizas: a promising component of plant production systems provided favorable conditions for their growth. **Frontiers in Plant Science**, v. 9, n. September, p. 1-6, 2018.

BODDEY, R. M. et al. Biological nitrogen fixation associated with sugar cane and rice: Contributions and prospects for improvement. **Plant and Soil**, v. 174, n. 1-2, p. 195-209, 1995.

BOLOT, S. et al.The “inner circle” of the cereal genomes. **Current Opinion in Plant Biology**, v. 12, n. 2, p. 119-125, 2009.

BONFANTE, P.; GENRE, A. Mechanisms underlying beneficial plant - Fungus interactions in mycorrhizal symbiosis. **Nature Communications**, v. 1, n. 4, p. 1-11, 2010.

BOTTOMLEY, W. B. Some effects of organic growth-promotion substances (auximones) on the growth of *Lemna minor* in mineral cultural solutions. **Proceedings of the Royal Society B: Biological Sciences**, v. 88, p. 481-505, 1917.

BRAHMAPRAKASH, G. P.; SAHU, P. K. Biofertilizers for sustainability. **Journal of the Indian Institute of Science**, v. 92, n. 1, p. 37-62, 2012.

BURLEIGH, S. H.; HARRISON, M. J. The down-regulation of Mt4-like genes by phosphate fertilization occurs systemically and involves phosphate translocation to the shoots. **Plant Physiology**, v. 119, n. 1, p. 241-248, 1999.

CAMARGO, A. P. et al. RNAsamba: Neural network-based assessment of the protein-coding potential of RNA sequences. **NAR Genomics and Bioinformatics**, v. 2, n. 1, p. 1-12, 2020.

CAMPALANS, A.; KONDOROSI, A.; CRESPI, M. Enod40, a short open reading frame-containing mRNA, induces cytoplasmic localization of a nuclear RNA binding protein in *Medicago truncatula*. **Plant Cell**, v. 16, n. 4, p. 1047-1059, 2004.

CANELLAS, L. P. et al. Humic acids isolated from earthworm compost enhance root elongation, lateral root emergence, and plasma membrane H+-ATPase activity in maize roots. **Plant Physiology**, v. 130, n. 4, p. 1951-1957, 2002.

CANELLAS, L. P. et al. A combination of humic substances and *Herbaspirillum seropedicae* inoculation enhances the growth of maize (*Zea mays L.*). **Plant and Soil**, v. 366, n. 1–2, p. 119-132, 2013.

CHEKANOVA, J. A. et al. Genome-Wide High-Resolution Mapping of Exosome Substrates Reveals Hidden Features in the *Arabidopsis* Transcriptome. **Cell**, v. 131, n. 7, p. 1340–1353, 2007.

CHEN, M.; CHEN, J.; ZHANG, D. Exploring the secrets of long noncoding RNAs. **International Journal of Molecular Sciences**, v. 16, n. 3, p. 5467-5496, 2015.

CHEN, Q. et al. From “Dark Matter” to “Star”: Insight Into the Regulation Mechanisms of Plant Functional Long Non-Coding RNAs. **Frontiers in Plant Science**, v. 12, n. June, p. 1-13, 2021.

CHEN, X. et al. A transcriptomic profile of topping responsive non-coding RNAs in tobacco roots (*Nicotiana tabacum*). **BMC Genomics**, v. 20, n. 1, p. 1-14, 2019.

CHIAIESE, P. et al. Renewable sources of plant biostimulation: Microalgae as a sustainable means to improve crop performance. **Frontiers in Plant Science**, v. 871, n. December, p. 1-6, 2018.

CHOI, S. W.; NAM, J. W. TERIUS: Accurate prediction of lncRNA via high-throughput sequencing data representing RNA-binding protein association. **BMC Bioinformatics**, v. 19, n. Suppl 1, 2018.

CHUBATSU, L. S. et al. Nitrogen fixation control in *Herbaspirillum seropedicae*. **Plant and Soil**, v. 356, n. 1-2, p. 197-207, 2012.

CIRILLO, D. et al. Quantitative predictions of protein interactions with long noncoding RNAs: To the Editor. **Nature Methods**, v. 14, n. 1, p. 5-6, 2016.

COLLA, G. et al. Biostimulant action of protein hydrolysates: Unraveling their effects on plant physiology and microbiome. **Frontiers in Plant Science**, v. 8, n. December, p. 1-14, 2017.

CRESPI, M. D. et al. Enod40, a gene expressed during nodule organogenesis, codes for a non-translatable RNA involved in plant growth. **EMBO Journal**, v. 13, n. 21, p. 5099-5112, 1994.

DALPÉ, Y.; MONREAL, M. Arbuscular Mycorrhiza Inoculum to Support Sustainable Cropping Systems. **Crop Management**, v. 3, n. 1, p. 1-11, 2004.

DESHPANDE, S. et al. PLIT: An alignment-free computational tool for identification of long non-coding RNAs in plant transcriptomic datasets. **Computers in Biology and Medicine**, v. 105, p. 169-181, 2019.

DEVAUX, Y. et al. Long noncoding RNAs in cardiac development and ageing. **Nature Reviews Cardiology**, v. 12, n. 7, p. 415-425, 2015.

DI MARSICO, M. et al. GreeNC 2.0: a comprehensive database of plant long non-coding RNAs. **Nucleic acids research**, v. 50, n. D1, p. D1442-D1447, 2022.

DIVAN BALDANI, V. L.; BALDANI, J. I.; DÖBEREINER, J. Inoculation of rice plants with the endophytic diazotrophs *Herbaspirillum seropedicae* and Burkholderia spp. **Biology and Fertility of Soils**, v. 30, n. 5-6, p. 485-491, 2000.

DOEBLEY, J. The Genetics of Maize Evolution. **Annual Review of Genetics**, v. 38, n. 1, p. 37-59, 2004.

DU JARDIN, P. Plant biostimulants: Definition, concept, main categories and regulation. **Scientia Horticulturae**, v. 196, p. 3-14, 2015.

DU, Q. et al. The PILNCR1-miR399 regulatory module is important for low phosphate tolerance in maize. **Plant Physiology**, v. 177, n. 4, p. 1743-1753, 2018.

ELBELTAGY, A. et al. Endophytic Colonization and In Planta Nitrogen Fixation by a *Herbaspirillum*sp. **Applied and Environmental Microbiology**, v. 67, n. 11, p. 5285-5293, 2001.

ERTANI, A. et al. Effect of commercial lignosulfonate-humate on *Zea mays* L. metabolism. **Journal of Agricultural and Food Chemistry**, v. 59, n. 22, p. 11940-11948, 2011.

EYHERAGUIBEL, B.; SILVESTRE, J.; MORARD, P. Effects of humic substances derived from organic waste enhancement on the growth and mineral nutrition of maize. **Bioresource Technology**, v. 99, n. 10, p. 4206-4212, 2008.

FAN, X. N.; ZHANG, S. W. LncRNA-MFDL: Identification of human long non-coding RNAs by fusing multiple features and using deep learning. **Molecular BioSystems**, v. 11, n. 3, p. 892-897, 2015.

FARD, E. M. et al. Regulation of mir159 and mir396 Mediated by *Piriformospora indica* Confer Drought Tolerance in Rice. **Journal of Plant Molecular Breeding**, v. 5, n. 1, p. 10-18, 2017.

FERRARI, C. S. et al. Expressed Proteins of *Herbaspirillum seropedicae* in Maize (DKB240) Roots-Bacteria Interaction Revealed Using Proteomics. **Applied Biochemistry and Biotechnology**, v. 174, n. 6, p. 2267-2277, 2014.

FORDE, B. G. Nitrogen signalling pathways shaping root system architecture: An update. **Current Opinion in Plant Biology**, v. 21, p. 30-36, 2014.

FORDE, B. G.; WALCH-LIU, P. Nitrate and glutamate as environmental cues for behavioural responses in plant roots. **Plant, Cell and Environment**, v. 32, n. 6, p. 682-693, 2009.

FRANCO-ZORRILLA, J. M. et al. Target mimicry provides a new mechanism for regulation of microRNA activity. **Nature Genetics**, v. 39, n. 8, p. 1033-1037, 2007.

FUJITA, T. et al. Isolation and characterization of a cDNA that encodes a novel

proteinase inhibitor from a tobacco genetic tumor. **Plant and Cell Physiology**, v. 34, n. 1, p. 137-142, 1993.

FUKUNAGA, K. et al. Genetic diversity and population structure of teosinte. **Genetics**, v. 169, n. 4, p. 2241-2254, 2005.

FURIÓ-TARÍ, P. et al. spongeScan: A web for detecting microRNA binding elements in lncRNA sequences. **Nucleic acids research**, v. 44, n. W1, p. W176-W180, 2016.

GAIERO, J. R. et al. Inside the root microbiome: Bacterial root endophytes and plant growth promotion. **American Journal of Botany**, v. 100, n. 9, p. 1738-1750, 2013.

GALINAT, W. C. The history and evolution of maize. **Critical Reviews in Plant Sciences**, v. 7, n. 3, p. 197-220, 1988.

GANGULY, P. et al. The natural antisense transcript DONE40 derived from the lncRNA ENOD40 locus interacts with SET domain protein ASHR3 during inception of symbiosis in *Arachis hypogaea*. **Molecular Plant-Microbe Interactions**, v. 34, n. 9, p. 1057-1070, 2021.

GAUT, B. S.; DOEBLEY, J. F. DNA sequence evidence for the segmental allotetraploid origin of maize. **Proceedings of the National Academy of Sciences**, v. 94, n. 13, p. 6809-6814, 1997.

GIANINAZZI, S. et al. Agroecology: The key role of arbuscular mycorrhizas in ecosystem services. **Mycorrhiza**, v. 20, n. 8, p. 519-530, 2010.

GOFF, S. A. et al. A Draft Sequence of the Rice Genome (*Oryza sativa*L . ssp) . **Science**, v. 296, n. August, p. 92-101, 2005.

GRUBER, A. R. et al. The Vienna RNA websuite. **Nucleic acids research**, v. 36, n. Web Server issue, p. 70-74, 2008.

GRANADA, C. E. et al. Is phosphate solubilization the forgotten child of plant growth-promoting rhizobacteria? **Frontiers in Microbiology**, v. 9, n. Sep, p. 2016-2019, 2018.

GULTYAEV, A. P.; ROUSSIS, A. Identification of conserved secondary structures and expansion segments in enod40 RNAs reveals new enod40 homologues in plants. **Nucleic Acids Research**, v. 35, n. 9, p. 3144-3152, 2007.

GUO, F. et al. PLncWX: A Machine-Learning Algorithm for Plant lncRNA Identification Based on WOA-XGBoost. **Journal of Chemistry**, v. 2021, 2021.

GUO, J. C. et al. CNIT: a fast and accurate web tool for identifying protein-coding and long non-coding transcripts based on intrinsic sequence composition. **Nucleic Acids Research**, v. 47, n. W1, p. W516-W522, 2019.

GUPTA, S. et al. DeepPlnc: Discovering plant lncRNAs through multimodal deep learning on sequential data. **BioRxiv preprint**, 2021.

GYANESHWAR, P. et al. *Herbaspirillum colonization* increases growth and nitrogen accumulation in aluminium-tolerant rice varieties. **New Phytologist**, v. 154, n. 1, p. 131-145, 2002.

HAN, J. et al. LncRNAs2Pathways: Identifying the pathways influenced by a set of

lncRNAs of interest based on a global network propagation method. **Scientific Reports**, v. 7, n. March, p. 1-14, 2017.

HAN, L. et al. New lncRNA annotation reveals extensive functional divergence of the transcriptome in maize. **Journal of Integrative Plant Biology**, v. 61, n. 4, p. 394-405, 2019.

HAN, S. et al. Lncident: A Tool for Rapid Identification of Long Noncoding RNAs Utilizing Sequence Intrinsic Composition and Open Reading Frame Information. **International Journal of Genomics**, v. 2016, p. 1-11, 2016.

HAN, S. et al. LncFinder: An integrated platform for long non-coding RNA identification utilizing sequence intrinsic composition, structural information and physicochemical property. **Briefings in Bioinformatics**, v. 20, n. 6, p. 2009-2027, 2019.

HARRIER, L. A.; WATSON, C. A.The potential role of arbuscular mycorrhizal (AM) fungi in the bioprotection of plants against soil-borne pathogens in organic and/or other sustainable farming systems. **Pest Management Science**, v. 60, n. 2, p. 149-157, 2004.

HAWKES, E. J. et al. COOLAIR antisense RNAs form evolutionarily conserved elaborate secondary structures. **Cell Reports**, v. 16, n. 12, p. 3087-3096, 2019.

HERNANDEZ, O. L. et al. Humic substances from vermicompost enhance urban lettuce production. **Agronomy for Sustainable Development**, v. 35, n. 1, p. 225-232, 2015.

HEZRONI, H. et al. Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species. **Cell Reports**, v. 11, n. 7, p. 1110-1122, 2015.

HU, J.; ANDREWS, B. **Distinguishing long non-coding RNAs from mRNAs using a two-layer structured classifier**.IEEE International Conference on Computational Advances in Bio and Medical Sciences, ICCABS. **Anais**...2017

HU, L. et al. COME: A robust coding potential calculation tool for lncRNA identification and characterization based on multiple features. **Nucleic Acids Research**, v. 45, n. 1, p. 1-13, 2017.

HU, Y. et al. Cell kinetics of auxin transport and activity in *Arabidopsis* root growth and skewing. **Nature Communications**, v. 12, n. 1, p. 1-13, 2021.

HUANG, F. et al. Phosphorylation of conserved PIN motifs directs *Arabidopsis* PIN1 polarity and auxin transport. **Plant Cell**, v. 22, n. 4, p. 1129-1142, 2010.

HUANG, T. K. et al. Identification of downstream components of ubiquitin-conjugating enzyme PHOSPHATE2 by quantitative membrane proteomics in *Arabidopsis* roots. **Plant Cell**, v. 25, n. 10, p. 4044-4060, 2013.

HUANG, X. et al. Genome-wide identification and characterization of long non-coding RNAs involved in flag leaf senescence of rice. **Plant Molecular Biology**, v. 105, n. 6, p. 655-684, 2021.

HUFFORD, M. B. et al. Comparative population genomics of maize domestication and improvement. **Nature Genetics**, v. 44, n. 7, p. 808-811, 2012.

ILTIS, H. H.; DOEBLEY, J. F. Taxonomy of Zea (Gramineae). Ii. Subspecific Categories in The *Zea Mays* Complex and a Generic Synopsis. **America**, v. 67, n. 6, p. 994-1004, 1980.

JABNOUNE, M. et al. A Rice cis-Natural antisense RNA acts as a translational enhancer for its cognate mRNA and contributes to phosphate homeostasis and plant fitness. **Plant Cell**, v. 25, n. 10, p. 4166-4182, 2013.

JAIN, P. et al. Identification of long non-coding RNA in rice lines resistant to Rice blast pathogen *Maganaporthe oryzae*. **Bioinformation**, v. 13, n. 08, p. 249-255, 2017.

JAMES, E. K. Nitrogen fixation in endophytic and associative symbiosis. **Field Crops Research**, v. 65, n. 2-3, p. 197-209, 2000.

JAMES, E. K. et al. Infection and Colonization of Rice Seedlings by the Plant Growth-Promoting Bacterium *Herbaspirillum seropedicae* Z67. **Molecular Plant-Microbe Interactions**, v. 15, n. 9, p. 894-906, 2002.

JIAO, F. et al. Cold induced antisense transcription of FLOWERING LOCUS C in distant grasses. **Frontiers in Plant Science**, v. 10, n. February, p. 1-9, 2019.

JIAO, Y. et al. Improved maize reference genome with single-molecule technologies. **Nature**, v. 546, n. 7659, p. 524-527, 2017.

JIN, J. et al. PLncDB: Plant long non-coding RNA database. **Bioinformatics**, v. 29, n. 8, p. 1068-1071, 2013.

JIN, J. et al. PLncDB V2.0: A comprehensive encyclopedia of plant long noncoding RNAs. **Nucleic Acids Research**, v. 49, n. D1, p. D1489-D1495, 2021.

JINDO, K. et al. Root growth promotion by humic acids from composted and non-composted urban organic wastes. **Plant and Soil**, v. 353, n. 1-2, p. 209-220, 2012.

JOHNSON, D.; GILBERT, L. Interplant signalling through hyphal networks. **New Phytologist**, v. 205, n. 4, p. 1448-1453, 2015.

JOHNSON, N. C.; GRAHAM, J. H. The continuum concept remains a useful framework for studying mycorrhizal functioning. **Plant and Soil**, v. 363, n. 1-2, p. 411-419, 2013.

KANG, Y. J. et al. CPC2: A fast and accurate coding potential calculator based on sequence intrinsic features. **Nucleic Acids Research**, v. 45, n. W1, p. W12-W16, 2017.

KAPRANOV, P. et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. **Science**, v. 316, n. 5830, p. 1484-1488, 2007.

KAUL, S. et al. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. **Nature**, v. 408, n. December, p. 796-815, 2000.

KAWAGUCHI, M.; MINAMISAWA, K. Plant-microbe communications for symbiosis. **Plant and Cell Physiology**, v. 51, n. 9, p. 1377-1380, 2010.

KELLOGG, E. Update on Evolution Evolutionary History of the Grasses. **Plant physiology**, v. 125, p. 1198-1205, 2001.

KIM, D. H. et al. Vernalization: Winter and the timing of flowering in plants. **Annual Review of Cell and Developmental Biology**, v. 25, p. 277-299, 2009.

KIM, D. H.; XI, Y.; SUNG, S. Modular function of long noncoding RNA, COLDAIR, in the vernalization response. **PLoS Genetics**, v. 13, n. 7, p. 1-18, 2017.

KISTLER, L. et al. Archaeological Central American maize genomes suggest ancient gene flow from South America. **Proceedings of the National Academy of Sciences of the United States of America**, v. 117, n. 52, p. 33124-33129, 2020.

KISTNER, C.; PARNISKE, M. Evolution of signal transduction in intracellular symbiosis. **Plant Scien**, v. 7, n. 11, p. 511-518, 2002.

KONG, L. et al. CPC: Assess the protein-coding potential of transcripts using sequence features and support vector machine. **Nucleic Acids Research**, v. 35, n. SUPPL.2, p. 345-349, 2007.

KUMAGAI, H. et al. RNAi knock-down of ENOD40s leads to significant suppression of nodule formation in *Lotus japonicus*. **Plant and Cell Physiology**, v. 47, n. 8, p. 1102-1111, 2006.

LAURESSERGUES, D. et al. The microRNA miR171h modulates arbuscular mycorrhizal colonization of *Medicago truncatula* by targeting NSP2. **Plant Journal**, v. 72, n. 3, p. 512-522, 2012.

LENG, Y. et al. Genome-wide lncRNAs identification and association analysis for cold-responsive genes at the booting stage in rice (*Oryza sativa* L.). **Plant Genome**, v. 13, n. 2, p. 1-18, 2020.

LERTAMPAIPORN, S. et al. Identification of non-coding RNAs with a new composite feature in the Hybrid Random Forest Ensemble algorithm. **Nucleic Acids Research**, v. 42, n. 11, 2014.

LI, A.; ZHANG, J.; ZHOU, Z. PLEK: A tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. **BMC Bioinformatics**, v. 15, n. 1, p. 1-10, 2014.

LI, J. et al. LncTar: A tool for predicting the RNA targets of long noncoding RNAs. **Briefings in Bioinformatics**, v. 16, n. 5, p. 806-812, 2014.

LI, S. et al. Genome-wide identification and functional prediction of cold and/or drought-responsive lncRNAs in cassava. **Scientific Reports**, v. 7, n. December 2016, 2017.

LI, X. et al. Genome-wide identification and functional prediction of tobacco lncRNAs responsive to root-knot nematode stress. **PLoS ONE**, v. 13, n. 11, 2018.

LIN, M. F.; JUNGREIS, I.; KELLIS, M. PhyloCSF: A comparative genomics method to distinguish protein coding and non-coding regions. **Bioinformatics**, v. 27, n. 13, p. 1-8, 2011.

LIU, D. et al. New technologies accelerate the exploration of non-coding RNAs in horticultural plants. **Horticulture Research**, v. 4, p. 1-8, 2017.

LIU, J.; GOUGH, J.; ROST, B. Distinguishing protein-coding from non-coding RNAs

through support vector machines. **PLoS Genetics**, v. 2, n. 4, p. 529-536, 2006.

LIU, S. et al. Predlnc-gfstack: A global sequence feature based on a stacked ensemble learning method for predicting lncrnas from transcripts. **Genes**, v. 10, n. 9, 2019.

LIU, T. Y. et al. PHO2-dependent degradation of PHO1 modulates phosphate homeostasis in *Arabidopsis*. **Plant Cell**, v. 24, n. 5, p. 2168–2183, 2012.

LIU, W. et al. Genome-wide identification and characterization of the CKII gene family in the cultivated banana cultivar (*Musa* spp. cv Tianbaojiao) and the wild banana (*Musa itinerans*). **PLoS ONE**, v. 13, n. 7, p. 1-27, 2018.

LLACA, V.; CAMPBELL, M. A.; DESCHAMPS, S. Genome Diversity in Maize. **Journal of Botany**, v. 2011, p. 1-10, 2011.

LOTFI, R. et al. Effects of humic acid on photosynthetic efficiency of rapeseed plants growing under different watering conditions. **Photosynthetica**, v. 56, n. 3, p. 962-970, 2018.

LUIS, A. D. et al. Two MicroRNAs Linked to Nodule Infection and Nitrogen-Fixing Ability in the Legume *Lotus japonicus*. **Plant Physiology**, v. 160, n. 4, p. 2137-2154, 2012.

LV, Y. et al. Genome-wide identification and functional prediction of nitrogen-responsive intergenic and intronic long non-coding RNAs in maize (*Zea mays* L.). **BMC Genomics**, v. 17, n. 1, p. 1-15, 2016.

LV, Y. et al. Maize transposable elements contribute to long non-coding RNAs that are regulatory hubs for abiotic stress response. **BMC Genomics**, v. 20, n. 1, p. 1-17, 2019.

MA, J. et al. Genome-Wide Identification of Long Noncoding RNAs and Their Responses to Salt Stress in Two Closely Related Poplars. **Frontiers in Genetics**, v. 10, n. September, p. 1-13, 2019.

MA, P. et al. Transcriptomic and genome-wide association study reveal long noncoding RNAs responding to nitrogen deficiency in maize. **BMC Plant Biology**, v. 21, n. 1, p. 1-19, 2021.

MANN, M.; WRIGHT, P. R.; BACKOFEN, R. IntaRNA 2.0: Enhanced and customizable prediction of RNA-RNA interactions. **Nucleic Acids Research**, v. 45, n. W1, p. W435-W439, 2017.

MATSUOKA, Y. et al.A single domestication for maize shown by multilocus microsatellite genotyping. **Proceedings of the National Academy of Sciences**, v. 99, n. 9, p. 6080-6084, 2002.

MENG, J. et al. lncRNA-LSTM: Prediction of Plant Long Non-coding RNAs Using Long Short-Term Memory Based on p-nts Encoding. **Springer International Publishing**, v. 11645 LNAI, 2019.

MENG, J. et al. PlncRNA-HDeep: plant long noncoding RNA prediction using hybrid deep learning based on two encoding styles. **BMC Bioinformatics**, v. 22, p. 1-16, 2021.

MESSING, J. The polyploid origin of maize. **Handbook of Maize: Genetics and**

Genomics, v. II, p. 221-238, 2009.

MICHAELS, S. D.; AMASINO, R. M. FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. **Plant Cell**, v. 11, n. 5, p. 949-956, 1999.

MOK, M. C.; MARTIN, R. C.; MOK, D. W. S. Cytokinins: Biosynthesis, metabolism and perception. **In Vitro Cellular and Developmental Biology - Plant**, v. 36, n. 2, p. 102-107, 2000.

MONTEIRO, R. A. et al. *Herbaspirillum*-plant interactions: Microscopical, histological and molecular aspects. **Plant and Soil**, v. 356, n. 1-2, p. 175-196, 2012.

MORA, V. et al. Action of humic acid on promotion of cucumber shoot growth involves nitrate-related changes associated with the root-to-shoot distribution of cytokinins, polyamines and mineral nutrients. **Journal of Plant Physiology**, v. 167, n. 8, p. 633-642, 2010.

MUKHERJEE, P. K. et al. *Trichoderma* research in the genome era. **Annual Review of Phytopathology**, v. 51, p. 105-129, 2013.

MUSCOLO, A.; SIDARI, M.; NARDI, S. Humic substance: Relationship between structure and activity. Deeper information suggests univocal findings. **Journal of Geochemical Exploration**, v. 129, p. 57-63, 2013.

NACHTIGALL, P. G.; KASHIWABARA, A. Y.; DURHAM, A. M. CodAn: Predictive models for precise identification of coding regions in eukaryotic transcripts. **Briefings in Bioinformatics**, v. 22, n. 3, p. 1-11, 2021.

NARDI, S. et al. Nitrate uptake and ATPase activity in oat seedlings in the presence of two humic fractions. **Soil Biology and Biochemistry**, v. 23, n. 9, p. 833-836, 1991.

NARDI, S. et al. Auxin-like effect of humic substances extracted from faeces of *Allolobophora caliginosa* and *A. rosea*. **Soil Biology and Biochemistry**, v. 26, n. 10, p. 1341-1346, 1994.

NARDI, S. et al. Physiological effects of humic substances on higher plants. **Soil Biology and Biochemistry**, v. 34, n. 11, p. 1527-1536, 2002.

NARDI, S. et al. Plant biostimulants: Physiological responses induced by protein hydrolyzed-based products and humic substances in plant metabolism. **Scientia Agricola**, v. 73, n. 1, p. 18-23, 2016.

NARDI, S.; ERTANI, A.; FRANCIOSO, O. Soil–root cross-talking: The role of humic substances. **Journal of Plant Nutrition and Soil Science**, v. 180, n. 1, p. 5-13, 2017.

NAYA, L. et al. Regulation of copper homeostasis and biotic interactions by microRNA 398b in common bean. **PLoS ONE**, v. 9, n. 1, 2014.

NEGRI, T. D. C. et al. Pattern recognition analysis on long noncoding RNAs: A tool for prediction in plants. **Briefings in Bioinformatics**, v. 20, n. 2, p. 682-689, 2019.

NICOLÁS, C. et al. *Trichoderma* genes in plants for stress tolerance - status and prospects. **Plant Science**, v. 228, p. 71-78, 2014.

NITHIN, C. et al. Genome-wide identification of miRNAs and lncRNAs in *Cajanus cajan*. **BMC Genomics**, v. 18, n. 1, p. 1-14, 2017.

NORDSTRÖM, A. et al. Auxin regulation of cytokinin biosynthesis in *Arabidopsis thaliana*: A factor of potential importance for auxin-cytokinin-regulated development. **Proceedings of the National Academy of Sciences of the United States of America**, v. 101, n. 21, p. 8039–8044, 2004.

OLIVARES, F. L. et al. Occurrence of the endophytic diazotrophs *Herbaspirillum* spp. in roots, stems, and leaves, predominantly of Gramineae. **Biology and Fertility of Soils**, v. 21, p. 197-200, 1996.

OLIVARES, F. L. et al. Substrate biofortification in combination with foliar sprays of plant growth promoting bacteria and humic substances boosts production of organic tomatoes. **Scientia Horticulturae**, v. 183, n. 1, p. 100-108, 2015.

PALUMBO, G. et al. Biostimulant potential of humic acids extracted from an amendment obtained via combination of olive mill wastewaters (OMW) and a pre-treated organic material derived from municipal solid waste (MSW). **Frontiers in Plant Science**, v. 9, n. July, p. 1-14, 2018.

PANG, J. et al. Spatio-temporal transcriptional dynamics of maize long non-coding RNAs responsive to drought stress. **Genes**, v. 10, n. 2, 2019.

PATERSON, A. H. et al. The *Sorghum bicolor* genome and the diversification of grasses. **Nature**, v. 457, n. 7229, p. 551-556, 2009.

PEDROSA, F. O. et al. Genome of *Herbaspirillum seropedicae* strain SmR1, a specialized diazotrophic endophyte of tropical grasses. **PLoS Genetics**, v. 7, n. 5, 2011.

PHILIPPOT, L. et al. Going back to the roots: The microbial ecology of the rhizosphere. **Nature Reviews Microbiology**, v. 11, n. 11, p. 789-799, 2013.

PIAN, C. et al. LncRNApred: Classification of long non-coding RNAs and protein-coding transcripts by the ensemble algorithm with a new hybrid feature. **PLoS ONE**, v. 11, n. 5, p. 1-20, 2016.

PICARDI, E. et al. Uncovering RNA editing sites in long non-coding RNAs. **Frontiers in Bioengineering and Biotechnology**, v. 2, n. Dec, p. 1-6, 2014.

PICCOLO, A. The Supramolecular Structure of Humic Substances. **Soil Science**, v. 166, n. 11, p. 810-832, 2001.

PINKNEY, H. R.; WRIGHT, B. M.; DIERMEIER, S. D. The lncrna toolkit: Databases and in silico tools for lncrna analysis. **Non-coding RNA**, v. 6, n. 4, p. 1-25, 2020.

PIPERNO, D. R.; FLANNERY, K. V. The earliest archaeological maize (*Zea mays* L.) from highland Mexico: New accelerator mass spectrometry dates and their implications. **Proceedings of the National Academy of Sciences**, v. 98, n. 4, p. 2101-2103, 2001.

PLENCHETTE, C. et al. Managing arbuscular mycorrhizal fungi in cropping systems. **Canadian Journal of Plant Science**, v. 85, n. 1, p. 31-40, 2005.

PONTING, C. P.; OLIVER, P. L.; REIK, W. Evolution and Functions of Long Noncoding RNAs. **Cell**, v. 136, n. 4, p. 629-641, 2009.

QIN, T. et al. A nucleus-localized long non-coding rna enhances drought and salt stress tolerance. **Plant Physiology**, v. 175, n. 3, p. 1321-1336, 2017.

RAI, M. I. et al. Classification and experimental identification of plant long non-coding RNAs. **Genomics**, v. 111, n. 5, p. 997-1005, 2019.

RICCI, M. et al. General principles to justify plant biostimulant claims. **Frontiers in Plant Science**, v. 10, n. March, p. 1-8, 2019.

RONCATO-MACCARI, L. D. B. et al. Endophytic *Herbaspirillum seropedicae* expresses nif genes in gramineous plants. **FEMS Microbiology Ecology**, v. 45, n. 1, p. 39-47, 2003.

ROOMI, S. et al. Protein profiling of *Arabidopsis* roots treated with humic substances: Insights into the metabolic and interactome networks. **Frontiers in Plant Science**, v. 871, n. December, p. 1-19, 2018.

ROSE, M. T. et al. A meta-analysis and review of plant-growth response to humic substances: Practical implications for agriculture. In: **Advances in Agronomy**. 1. ed. Elsevier Inc., v. 124, p.37-89, 2014.

ROSENTRATER, K. A. Overview of Corn-Based Fuel Ethanol Coproducts: Production and Use. In: **Biofuel's Engineering Process Technology**, p. 141-166, 2011.

ROUPHAEL, Y.; COLLA, G. Editorial: Biostimulants in Agriculture. **Frontiers in Plant Science**, v. 11, n. February, p. 1-7, 2020.

SABIN, L. R.; DELÁS, M. J.; HANNON, G. J. Dogma Derailed: The Many Influences of RNA on the Genome. **Molecular Cell**, v. 49, n. 5, p. 783-794, 2014.

SANTOSH, B.; VARSHNEY, A.; YADAVA, P. K. Non-coding RNAs: Biological functions and applications. **Cell Biochemistry and Function**, v. 33, n. 1, p. 14-22, 2015.

SCHAUMANN, G. E. Soil organic matter beyond molecular structure Part I: Macromolecular and supramolecular characteristics. **Journal of Plant Nutrition and Soil Science**, v. 169, n. 2, p. 145-156, 2006.

SCHNABLE, P. S. et al. The B73 maize genome: Complexity, diversity, and dynamics. **Science**, v. 326, n. 5956, p. 1112-1115, 2009.

SCHNEIDER, H. W. et al. A Support Vector Machine based method to distinguish long non-coding RNAs from protein coding transcripts. **BMC Genomics**, v. 18, n. 1, p. 1-14, 2017.

SECCO, D.; BAUMANN, A.; POIRIER, Y. Characterization of the rice PHO1 gene family reveals a key role for OsPHO1;2 in phosphate homeostasis and the evolution of a distinct clade in dicotyledons. **Plant Physiology**, v. 152, n. 3, p. 1693-1704, 2010.

SEVERING, E. et al. *Arabidopsis thaliana* ambient temperature responsive lncRNAs. **BMC Plant Biology**, v. 18, n. 1, p. 1-10, 2018.

SHAO, Y. et al. DsTRD: Danshen transcriptional resource database. **PLoS ONE**, v. 11,

n. 2, p. 1-8, 2016.

SHIN, H. et al. Loss of At4 function impacts phosphate distribution between the roots and the shoots during phosphate starvation. **Plant Journal**, v. 45, n. 5, p. 712–726, 2006.

SHUKLA, B. et al. lncRNADetector: a bioinformatics pipeline for long non-coding RNA identification and MAPslnc: a repository of medicinal and aromatic plant lncRNAs. **RNA Biology**, v. 18, n. 12, p. 2290-2295, 2021.

SIMARD, S. W. et al. Mycorrhizal networks: Mechanisms, ecology and modelling. **Fungal Biology Reviews**, v. 26, n. 1, p. 39-60, 2012.

SIMOPOULOS, C. M. A.; WERETILNYK, E. A.; GOLDING, G. B. Prediction of plant lncRNA by ensemble machine learning classifiers. **BMC Genomics**, v. 19, n. 1, p. 1-11, 2018.

SINGH, A.; VIVEK, A. T.; KUMAR, S. AlnC: An extensive database of long noncoding RNAs in angiosperms. **PLoS ONE**, v. 16, n. 4 April, p. 1-11, 2021.

SINGH, U. et al. PLncPRO for prediction of long non-coding RNAs (lncRNAs) in plants and its application for discovery of abiotic stress-responsive lncRNAs in rice and chickpea. **Nucleic Acids Research**, v. 45, n. 22, 2017.

SOUSA, C. et al. Translational and Structural Requirements of the Early Nodulin Gene enod40, a Short-Open Reading Frame-Containing RNA, for Elicitation of a Cell-Specific Growth Response in the Alfalfa Root Cortex. **Molecular and Cellular Biology**, v. 21, n. 1, p. 354-366, 2001.

STRABLE, J.; SCANLON, M. J. Maize (*Zea mays*): A model organism for basic and applied research in plant biology. **Cold Spring Harbor Protocols**, v. 4, n. 10, p. 1-10, 2009.

SUN, K. et al. iSeeRNA: Identification of long intergenic non-coding RNA transcripts from transcriptome sequencing data. **BMC Genomics**, v. 14, n. Suppl 2, 2013a.

SUN, L. et al. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. **Nucleic Acids Research**, v. 41, n. 17, 2013b.

SUN, L. et al. lncRScan-SVM: A tool for predicting long non-coding RNAs using support vector machine. **PLoS ONE**, v. 10, n. 10, p. 1-16, 2015.

SUN, X. et al. Comparative Transcriptome Analysis Reveals New lncRNAs Responding to Salt Stress in Sweet *Sorghum*. **Frontiers in Bioengineering and Biotechnology**, v. 8, n. April, p. 1-14, 2020a.

SUN, Z. et al. Genome-wide identification of *Arabidopsis* long noncoding RNAs in response to the blue light. **Scientific Reports**, v. 10, n. 1, p. 1-10, 2020b.

SWIGONˇOVA, Z. et al. Close Split of *Sorghum* and Maize Genome Progenitors. **Genome Research**, v. 14, p. 1916-1923, 2004.

SZCZEŚNIAK, M. W.; ROSIKIEWICZ, W.; MAKAŁOWSKA, I. CANTATAdb: A collection of plant long non-coding RNAs. **Plant and Cell Physiology**, v. 57, n. 1, p. e8(1-7), 2016.

TESTER, M.; LANGRIDGE, P. Breeding technologies to increase crop production in a changing world. **Science**, v. 327, n. 5967, p. 818-822, 2010.

THIEBAUT, F. et al. Genome-wide identification of microRNA and siRNA responsive to endophytic beneficial diazotrophic bacteria in maize. **BMC Genomics**, v. 15, n. 1, p. 1-18, 2014.

THIEBAUT, F. et al. Role of small RNA in plant interaction with microbes. In: GULERIA, P.; KUMAR, V. (Eds.). **Plant Small RNA Biogenesis, Regulation and Application**. 1. ed. India: Elsevier, 2020. p. 299-319.

TONG, X.; LIU, S. CPPred: Coding potential prediction based on the global description of RNA sequence. **Nucleic Acids Research**, v. 47, n. 8, 2019.

TRIPATHI, R. et al. DeepLNC, a long non-coding RNA prediction tool using deep neural network. **Network Modeling Analysis in Health Informatics and Bioinformatics**, v. 5, n. 1, 2016.

TURRINI, A. et al. Functional complementarity of arbuscular mycorrhizal fungi and associated microbiota: the challenge of translational research. **Frontiers in Plant Science**, v. 9, n. September, p. 10-13, 2018.

VACCARO, S. et al. Effect of a compost and its water-soluble fractions on key enzymes of nitrogen metabolism in maize seedlings. **Journal of Agricultural and Food Chemistry**, v. 57, n. 23, p. 11267-11276, 2009.

VACHERON, J. et al. Plant growth-promoting rhizobacteria and root system functioning. **Frontiers in Plant Science**, v. 4, n. SEP, p. 1-19, 2013.

VAN DER HEIJDEN, M. G. A. et al. The mycorrhizal contribution to plant productivity, plant nutrition and soil structure in experimental grassland. **New Phytologist**, v. 172, n. 4, p. 739–752, 2006.

VIEIRA, L. M. et al. PlantRNA_sniffer: A SVM-based workflow to predict long intergenic non-coding RNAs in plants. **Non-coding RNA**, v. 3, n. 1, 2017.

VOLLBRECHT, E.; SIGMON, B. Amazing grass: developmental genetics of maize domestication. **Biochemical Society Transactions**, v. 33, n. 6, p. 1502, 2005.

WANG, G. et al. Characterization and identification of long non-coding RNAs based on feature relationship. **Bioinformatics**, v. 35, n. 17, p. 2949-2956, 2019.

WANG, H. et al. Genome-wide identification of long noncoding natural antisense transcripts and their responses to light in *Arabidopsis*. **Genome Research**, v. 24, n. 3, p. 444-453, 2014a.

WANG, J. et al. Genome-wide analysis of tomato long non-coding RNAs and identification as endogenous target mimic for microRNA in response to TYLCV infection. **Scientific Reports**, v. 5, n. December, 2015a.

WANG, L. et al. Comprehensive Analysis of Long Non-coding RNA Modulates Axillary Bud Development in Tobacco (*Nicotiana tabacum* L.). **Frontiers in Plant Science**, v. 13, n. February, p. 1-17, 2022.

WANG, T. et al. Novel phosphate deficiency-responsive long non-coding RNAs in the legume model plant *Medicago truncatula*. **Journal of Experimental Botany**, v. 68, n. 21–22, p. 5937–5948, 2017.

WANG, T. Z. et al. Identification and characterization of long non-coding RNAs involved in osmotic and salt stress in *Medicago truncatula* using genome-wide high-throughput sequencing. **BMC Plant Biology**, v. 15, n. 1, p. 1–13, 2015b.

WANG, H.-L.V.; CHEKANOVA, J. A. An Overview of Methodologies in Studying lncRNAs in the High-Throughput Era: When Acronyms ATTACK! In: **Plant Long Non-Coding RNAs: Methods and Protocols, Methods in Molecular Biology**. v. 1933, p.1-21.

WANG, Y. et al. Soybean miR172c Targets the Repressive AP2 Transcription Factor NNC1 to Activate *ENOD40* Expression and Regulate Nodule Initiation. **The Plant Cell Online**, v. 26, n. 12, p. 4782-4801, 2014b.

WANG, Y. et al. Soybean miR172c targets the repressive AP2 transcription factor NNC1 to activate ENOD40 expression and regulate nodule initiation. **Plant Cell**, v. 26, n. 12, p. 4782-4801, 2014.

WANG, Y. et al. Computational identification of human long intergenic non-coding RNAs using a GA-SVM algorithm. **Gene**, v. 533, n. 1, p. 94-99, 2014c.

WANG, Y. et al. Identification and expression analysis of miRNAs from nitrogen-fixing soybean nodules. **Biochemical and Biophysical Research Communications**, v. 378, n. 4, p. 799–803, 2009.

WANG, Y. et al. Microrna167-directed regulation of the auxin response factors GmARF8a and GmARF8b is required for soybean nodulation and lateral root development. **Plant Physiology**, v. 168, n. 3, p. 101-116, 2015.

WANG, Y. et al. Unveiling gibberellin-responsive coding and long noncoding RNAs in maize. **Plant Molecular Biology**, v. 98, n. 4-5, p. 427-438, 2018.

WANG, Y. et al. Genome-Wide Identification of lncRNAs Involved in Fertility Transition in the Photo-Thermosensitive Genic Male Sterile Rice Line Wuxiang S. **Frontiers in Plant Science**, v. 11, n. January, p. 1-17, 2021.

WEIDONG, Q. et al. Systematic Characterization of Long Non-Coding RNAs and Their Responses to Drought Stress in Dongxiang Wild Rice. **Rice Science**, v. 27, n. 1, p. 21-31, 2020.

WEKESA, J. S. et al. A Hybrid Prediction Method for Plant lncRNA-Protein Interaction. **Cells**, v. 8, n. 6, p. 521, 2019.

WENDEL, J. F. Genome evolution in polyploids. **Plant Molecular Biology**, v. 42, n. 1, p. 225-249, 2000.

WHITTAKER, C.; DEAN, C. The FLC locus: A platform for discoveries in epigenetics and adaptation. **Annual Review of Cell and Developmental Biology**, v. 33, n. June, p. 555-575, 2017.

WILUSZ, J. E.; SUNWOO, H.; SPECTOR, D. L. Long noncoding RNAs: Functional surprises from the RNA world. **Genes and Development**, v. 23, n. 13, p. 1494-1504,

2009.

WOO, S. L.; PEPE, O. Microbial consortia: Promising probiotics as plant biostimulants for sustainable agriculture. **Frontiers in Plant Science**, v. 9, n. 2003, p. 7–12, 2018.

WU, H. J. et al. Widespread long noncoding RNAs as endogenous target mimics for microRNAs in plants. **Plant Physiology**, v. 161, n. 4, p. 1875–1884, 2013.

WUCHER, V. et al. FEELnc: A tool for long non-coding RNA annotation and its application to the dog transcriptome. **Nucleic Acids Research**, v. 45, n. 8, p. 1-12, 2017.

XU, L.; GEELEN, D. Developing biostimulants from agro-food and industrial by-products. **Frontiers in Plant Science**, v. 871, n. October, p. 1-13, 2018.

XU, Y. et al. Predicting Long non-coding RNAs through feature ensemble learning. **BMC Genomics**, v. 21, n. Suppl 13, p. 1-12, 2020.

YANG, C. et al. LncADeep: An ab initio lncRNA identification and functional annotation tool based on deep learning. **Bioinformatics**, v. 34, n. 22, p. 3825-3834, 2018.

YU, F. et al. A comprehensive transcriptomics analysis reveals long non-coding RNA to be involved in the key metabolic pathway in response to waterlogging stress in maize. **Genes**, v. 11, n. 3, p. 1-19, 2020.

YUAN, J. et al. Systematic characterization of novel lncRNAs responding to phosphate starvation in *Arabidopsis thaliana*. **BMC Genomics**, v. 17, n. 1, p. 1-16, 2016.

ZANDONADI, D. B. et al. Nitric oxide mediates humic acids-induced root development and plasma membrane H^+-ATPase activation. **Planta**, v. 231, n. 5, p. 1025-1036, 2010.

ZANDONADI, D. B.; CANELLAS, L. P.; FAÇANHA, A. R. Indolacetic and humic acids induce lateral root development through a concerted plasmalemma and tonoplast H^+ pumps activation. **Planta**, v. 225, n. 6, p. 1583-1595, 2007.

ZENG, M. et al. DeepLncLoc: a deep learning framework for long non-coding RNA subcellular localization prediction based on subsequence embedding. **Briefings in Bioinformatics**, v. 23, n. 1, p. 1-10, 2022.

ZHANG, G. et al. ItLnc-BXE: A Bagging-XGBoost-Ensemble Method with Comprehensive Sequence Features for Identification of Plant lncRNAs. **IEEE Access**, v. 8, p. 68811-68819, 2020.

ZHANG, W. et al. Identification of maize long non-coding RNAs responsive to drought stress. **PLoS ONE**, v. 9, n. 6, p. e98958, 2014a.

ZHANG, X. et al. The long non-coding RNA lncRNA973 is involved in cotton response to salt stress. **BMC plant biology**, v. 19, n. 1, p. 459, 2019.

ZHANG, Y. C. et al. Genome-wide screening and functional analysis identify a large number of long noncoding RNAs involved in the sexual reproduction of rice. **Genome Biology**, v. 15, n. 12, p. 512, 2014b.

ZHANG, Z. et al. RiceLncPedia: a comprehensive database of rice long non-coding RNAs. **Plant Biotechnology Journal**, v. 19, n. 8, p. 1492-1494, 2021.

ZHAO, J. et al. Genome-wide identification of lncRNAs during rice seed development. **Genes**, v. 11, n. 3, 2020a.

ZHAO, J.; SONG, X.; WANG, K. LncScore: Alignment-free identification of long noncoding RNA from assembled novel transcripts. **Scientific Reports**, v. 6, n. July, p. 1-12, 2016.

ZHAO, M. et al. Identification of tissue-specific and cold-responsive lncRNAs in *Medicago truncatula* by high-throughput RNA sequencing. **BMC Plant Biology**, v. 20, n. 1, p. 1-16, 2020b.

ZHAO, Q. et al. LncPipe: A Nextflow-based pipeline for identification and analysis of long non-coding RNAs from RNA-Seq data. **Journal of Genetics and Genomics**, v. 45, n. 7, p. 399-401, 2018.

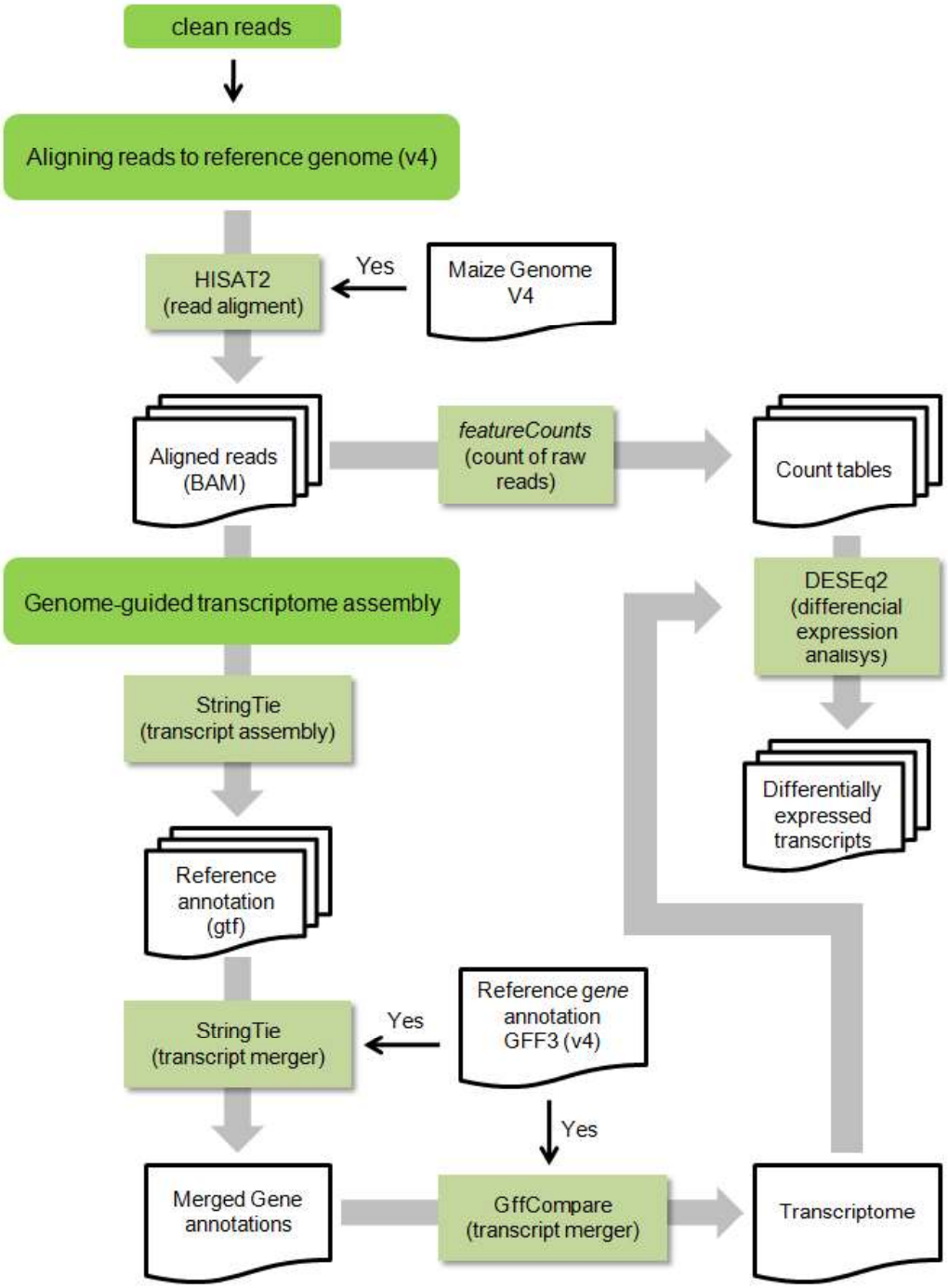
ZHAO, Y. et al. NONCODE 2016: An informative and valuable data source of long non-coding RNAs. **Nucleic Acids Research**, v. 44, n. D1, p. D203-D208, 2016.

ZHAO, Y. et al. Effect of root interaction on nodulation and nitrogen fixation ability of alfalfa in the simulated alfalfa/triticale intercropping in pots. **Scientific Reports**, v. 10, n. 1, p. 1-11, 2020c.

ZHOU, B. et al. EVLncRNAs 2.0: An updated database of manually curated functional long non-coding RNAs validated by low-throughput experiments. **Nucleic Acids Research**, v. 49, n. D1, p. D86-D91, 2021.

ZUBKO, E.; MEYER, P. A natural antisense transcript of the *Petunia* hybrida Sho gene suggests a role for an antisense mechanism in cytokinin regulation. **Plant Journal**, v. 52, n. 6, p. 1131-1139, 2007.

APÊNDICE – ARQUIVOS SUPLEMENTARES DO ARTIGO



Supplementary Figure 1 General flowchart of the pipeline used to calculate the expression of known and new lncRNAs from root maize RNA-seq libraries.

Supplementary Table 1 Statistics of RNA-seq libraries.

Sample	Run	Clean Reads	Aligned reads	Aligned read
Control	Sample_T1	78363202	66424680	84.77
	Sample_T2	27991757	24211511	86.50
	Sample_T3	52062618	45023166	86.48
Bacteria	Sample_T4	45585879	29945577	65.69
	Sample_T5	13358763	11249226	84.21
	Sample_T6	41012903	29574629	72.11
Humic acid	Sample_T7	39491808	33104418	83.83
	Sample_T8	21056212	17952633	85.26
	Sample_T9	32687190	26683791	81.63
Bacteria + humic acid	Sample_T10	14785114	12752339	86.25
	Sample_T11	31536296	26564291	84.23
	Sample_T12	34212442	26850061	78.48

Supplementary Table 2 LncRNAs acting as decoys of miRNAs.

miRNA_Acc.	Target_Acc.	Expectation	miRNA_start	miRNA_end	Target_start	Targe
zma-miR399a-3p	MSTRG.13897.1	4.0	1	21	519	5
zma-miR399c-3p		4.0	1	21	519	5
zma-miR399d-3p		4.0	1	21	519	5
zma-miR399h-3p		4.0	1	21	519	5
zma-miR399b-3p		4.5	1	21	519	5
zma-miR399e-3p		4.5	1	21	519	5
zma-miR169o-3p	MSTRG.30197.3	5.0	1	20	1742	17
zma-miR319a-3p	MSTRG.34402.1	5.0	1	20	1056	10
zma-miR319c-3p		5.0	1	20	1056	10
zma-miR319d-3p		5.0	1	20	1056	10
zma-miR319b-3p		5.0	1	20	1056	10

Supplementary Table 3 Annotations of mRNAs identified as lncRNA and miRNAs targets, using the LncTar

mRNAs bacteria|control		
		lncRNA MSTRG.6199.1
Targets	**Phytozome**	**NCBI**
MSTRG.34864.1	x	19-kDa-zein
Zm00001d051795_T001	EF-HAND CALCIUM-BINDING DOMAIN CONTAINING PROTEIN	x
Zm00001d017438_T001	x	glycine-rich cell wall structural protein 2
MSTRG.29272.1	x	transcriptional regulatory protein algP
Zm00001d003197_T001	EF-HAND CALCIUM-BINDING DOMAIN CONTAINING PROTEIN	x
MSTRG.13988.1	x	RING-H2 finger protein ATL65
Zm00001d011925_T001	LURP-one-related (LOR)	x
Zm00001d027457_T001	ATP-synthase-associated (ATP-synt_Z)	x
Zm00001d038216_T001	AP2 domain (AP2)	x
Zm00001d048268_T001	jasmonate ZIM domain-containing protein (JAZ)	x
		lncRNA MSTRG.30248.1
Targets	**Phytozome**	**NCBI**
Zm00001d038884_T001	Mitogen-activated protein kinase kinase kinase / MLTK	x
Zm00001d010856_T001	PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR	x
Zm00001d014748_T001	Benzyl alcohol O-benzoyltransferase	x
MSTRG.34452.1	x	F-box/LRR-repeat MAX2

Zm00001d005051_T001	ROOT PHOTOTROPISM PROTEIN 2	x
Zm00001d021197_T001	Alpha-L-fucoside fucohydrolase	x
Zm00001d033256_T001	Ca2+-independent phospholipase A2	x
Zm00001d033446_T001	ZIP zinc/iron transport family	x
Zm00001d036925_T001	F-box	x
Zm00001d007274_T001	KELCH-REPEAT PROTEIN SKIP20	x
		lncRNA 28681.1

Targets	Phytozome	NCBI
Zm00001d035021_T001	Pectin methylesterase inhibitor (PMEI)	x
Zm00001d040786_T001	ABA/WDS induced protein (ABA_WDS)	x
Zm00001d046664_T001	x	x
Zm00001d043323_T001	Ctr copper transporter	x
Zm00001d002773_T001	Cupredoxin	x
Zm00001d042633_T001	NUCLEOPORIN-RELATED	x
Zm00001d048959_T001	aquaporin SIP (SIP)	x
Zm00001d035659_T001	Metallothionein (Metallothio_2)	x
Zm00001d032475_T001	SAUR family protein (SAUR)	x
Zm00001d043299_T001	photosystem II PsbW protein (psbW)	x
		lncRNA 3891.1

Targets	Phytozome	NCBI
MSTRG.20520.1	x	NAC1 transcription factor-like
Zm00001d019734_T001	AP2 domain (AP2)	x
Zm00001d026028_T001	zinc-finger of the FCS-type, C2-C2	x

	(zf-FLZ)	
Zm00001d040292_T001	Late embryogenesis abundant protein (LEA_3)	x
MSTRG.5077.1	x	vegetative cell wall protein gp1
MSTRG.27054.1	x	aldose reductase
MSTRG.17763.1	x	Mu transposon
MSTRG.5071.1	x	vegetative cell wall protein gp1 precursor
MSTRG.3731.2	x	tubulin-folding cofactor D
MSTRG.11857.1	x	putative auxin efflux carrier PIN9

humic acid|control

		lncRNA MSTRG.6199.1
Targets	**Phytozome**	**NCBI**
Zm00001d008370_T001	Geranylgeranyl-PP synthetase	x
Zm00001d043784_T002	Zinc finger, RING-type	x
Zm00001d025947_T001	AUXIN-REGULATED PROTEIN-RELATED	x
Zm00001d017383_T001	MYB-LIKE DNA-BINDING PROTEIN MYB	x
Zm00001d011925_T001	LURP-one-related (LOR)	x
Zm00001d025477_T001	AP2 domain (AP2)	x
MSTRG.33057.1		Grx_I1 - glutaredoxin subgroup III
Zm00001d026584_T003	ADENYLATE KINASE 1	x
Zm00001d029172_T001	x	endoribonuclease Dicer homolog 1
Zm00001d014029_T001	MYB-LIKE DNA-BINDING PROTEIN MYB	x
		lncRNA MSTRG.30248.1

Targets	Phytozome	NCBI
Zm00001d014642_T001	EXOCYST COMPLEX PROTEIN EXO70	x
Zm00001d038884_T001	Mitogen-activated protein kinase kinase kinase / MLTK	x
Zm00001d011655_T001	Mitogen-activated protein kinase kinase kinase / MLTK	x
Zm00001d038676_T001	Xyloglucan 6-alpha-D-xylosyltransferase	x
Zm00001d017900_T001	DOF ZINC FINGER PROTEIN DOF1.1-RELATED	x
Zm00001d021205_T001	AP2 domain (AP2)	x
Zm00001d005051_T001	ROOT PHOTOTROPISM PROTEIN 2	x
Zm00001d014748_T001	Benzyl alcohol O-benzoyltransferase	x
Zm00001d029924_T001	DIMETHYLANILINE MONOOXYGENASE	x
Zm00001d031454_T001	BROMODOMAIN CONTAINING PROTEIN	x
		lncRNA 28681.1

Targets	Phytozome	NCBI
Zm00001d031194_T001	CAMP-RESPONSE ELEMENT BINDING PROTEIN-RELATED	x
Zm00001d046154_T001	RING FINGER DOMAIN-CONTAINING	x
Zm00001d035021_T001	Plant invertase/pectin methylesterase inhibitor (PMEI)	x
Zm00001d039997_T001	NPR1 interacting (NPR1_interact)	x
Zm00001d021961_T001	Alpha-L-fucosidase / Alpha-L-fucoside fucohydrolase	x
Zm00001d046927_T001	PROTEIN GLUTAMINE DUMPER 1-	x

	RELATED	
Zm00001d042633_T001	NUCLEOPORIN-RELATED	x
Zm00001d048959_T001	aquaporin SIP (SIP)	x
Zm00001d032088_T001	Auxin responsive protein (Auxin_inducible)	x
Zm00001d046664_T001	CAMP-RESPONSE ELEMENT BINDING PROTEIN-RELATED	x
		lncRNA 3891.1
Targets	**Phytozome**	**NCBI**
Zm00001d053908_T001	Zinc finger, CCHC-type	x
Zm00001d005663_T001	Succinic dehydrogenase	x
Zm00001d050462_T001	EXOCYST COMPLEX PROTEIN EXO70	x
MSTRG.20520.1	x	NAC1
Zm00001d053843_T001	Wound-induced protein	x
MSTRG.5077.1	vegetative cell wall protein gp1	x
Zm00001d019734_T001	AP2 domain (AP2)	x
MSTRG.3731.2	x	tubulin-folding cofactor D
MSTRG.11857.1	x	auxin efflux carrier PIN9
Zm00001d040751_T001	x	non-specific lipid-transfer protein

bac+humic acid|control

		lncRNA MSTRG.6199.1
Targets	**Phytozome**	**NCBI**
Zm00001d006489_T001	Senescence regulator S40	x
Zm00001d043784_T002	zinc ion binding	x
Zm00001d029172_T001	x	endoribonuclease Dicer homolog 1

Zm00001d017438_T001	x	glycine-rich cell wall structural protein 2
Zm00001d014563_T001	pathogenesis-related group 5 protein, thaumatin type	x
Zm00001d011495_T001	protein serine/threonine phosphatase activity	x
Zm00001d008370_T001	Geranylgeranyl-PP synthetase	x
MSTRG.29272.1	x	transcriptional regulatory protein algP
Zm00001d016698_T001	x	protein BIC1
MSTRG.33057.1	x	Grx_I1 - glutaredoxin subgroup III
		lncRNA MSTRG.30248.1

Targets	Phytozome	NCBI
Zm00001d014642_T001	EXO70-G1 PROTEIN	x
Zm00001d021329_T001	Putrescine N-hydroxycinnamoyltransferase	x
Zm00001d038884_T001	Tyrosine-protein kinase catalytic domain	x
Zm00001d011655_T001	Protein kinase domain	x
Zm00001d038676_T001	Xyloglucan 6-alpha-D-xylosyltransferase	x
Zm00001d017900_T001	DOF ZINC FINGER PROTEIN DOF1.1-RELATED	x
Zm00001d021205_T001	AP2 domain (AP2)	x
Zm00001d010856_T001	TspO/MBR-related protein	x
Zm00001d010652_T001	EID1-LIKE F-BOX PROTEIN 3	x
Zm00001d014748_T001	Benzyl alcohol O-benzoyltransferase	x
		lncRNA 28681.1

Targets	Phytozome	NCBI
Zm00001d019365_T001	P-hydroxyphenylpyruvate oxidase	x

Zm00001d031194_T001	CAMP-RESPONSE ELEMENT BINDING PROTEIN-RELATED	x
Zm00001d041553_T001	Thaumatin family (Thaumatin)	x
Zm00001d035021_T001	Plant invertase/pectin methylesterase inhibitor (PMEI)	x
Zm00001d039997_T001	NPR1 interacting (NPR1_interact)	x
Zm00001d021961_T001	ALPHA-L-FUCOSIDASE 3	x
Zm00001d046927_T001	PROTEIN GLUTAMINE DUMPER 1-RELATED	x
Zm00001d046664_T001	CAMP-RESPONSE ELEMENT BINDING PROTEIN-RELATED	x
Zm00001d046154_T001	Zinc finger, RING-type	x
Zm00001d043323_T001	COPPER TRANSPORTER 1-RELATED	x
		lncRNA 3891.1

Targets	Phytozome	NCBI
Zm00001d005663_T001	Succinic dehydrogenase	x
Zm00001d050462_T001	Exocyst complex protein Exo70	x
MSTRG.20520.1	NAC1 transcription factor-like	x
Zm00001d019734_T001	AP2/ERF domain	x
Zm00001d036044_T001	Exonuclease III	x
Zm00001d053908_T001	COLD SHOCK DOMAIN CONTAINING PROTEINS	x
Zm00001d026028_T001	611zinc-finger of the FCS-type	x
MSTRG.11727.1	x	classical arabinogalactan protein 4
MSTRG.11555.2	x	cysteine-rich receptor-like protein kinase 19
Zm00001d002394_T001	Methionyl-tRNA synthetase	x

miRNAs

Targets	miRNA	Phytozome
Zm00001d032081_T001	zma-miR171g-5p	DORMANCY/AUXIN ASSOCIATED FAMILY PROTE
Zm00001d025554_T001	zma-miR164a-3p	Hsp70 protein (HSP70) // Xylanase inhibitor N-ter (TAXi_N)
Zm00001d027463_T001	zma-miR171g-5p	Non-specific serine/threonine protein kinase
Zm00001d035467_T001	zma-miR164a-3p	Non-specific serine/threonine protein kinase / Thre specific protein kinase
Zm00001d010961_T001	zma-miR164a-3p	Non-specific serine/threonine protein kinase / Thre specific protein kinase
Zm00001d028277_T001	zma-miR167h-3p / zma-miR167i-3p	Tam3-transposase (Ac family)
Zm00001d030380_T001	zma-miR167h-3p / zma-miR167i-3p	Tam3-transposase (Ac family)
Zm00001d021994_T001	zma-miR167h-3p / zma-miR167i-3p	Tam3-transposase (Ac family)
Zm00001d015778_T001	zma-miR169l-3p	EIN3-binding F-box protein (EBF1_2)
Zm00001d046518_T001	zma-miR159c-3p / zma-miR159d-3p	MYB TRANSCRIPTION FACTOR
Zm00001d046519_T001	zma-miR159c-3p / zma-miR159d-3p	MYB TRANSCRIPTION FACTOR
Zm00001d022460_T001	zma-miR159c-3p / zma-miR159d-3p	Pectinesterase / Pectin methylesterase
Zm00001d021225_T004	zma-miR164a-3p	RING FINGER DOMAIN-CONTAINING // SUBFAMIL NAMED
Zm00001d028032_T001	zma-miR164a-3p	Calcium-transporting ATPase / Calcium-translocat type ATPase