



UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE BIOLOGIA

MAIARA CURTOLO

**"ESTRATÉGIAS DE IDENTIFICAÇÃO DE GENES ALVOS E
EDIÇÃO DE GENOMA DE LARANJA DOCE (*CITRUS
SINENSIS* L. OSB.) PARA TOLERÂNCIA
AO HUANGLONGBING"**

“TARGET IDENTIFICATION STRATEGIES AND SWEET
ORANGE (*CITRUS SINENSIS* L. OSB.) GENOME EDITING
FOR TOLERANCE TO HUANGLONGBING”

CAMPINAS

2021

MAIARA CURTOLO

**"ESTRATÉGIAS DE IDENTIFICAÇÃO DE GENES ALVOS E
EDIÇÃO DE GENOMA DE LARANJA DOCE (*CITRUS SINENSIS* L.
OSB.) PARA TOLERÂNCIA AO *HUANGLONGBING*"**

**“TARGET IDENTIFICATION STRATEGIES AND SWEET ORANGE
(*CITRUS SINENSIS* L. OSB.) GENOME EDITING FOR TOLERANCE TO
HUANGLONGBING”**

*Tese apresentada ao Instituto de Biologia
da Universidade Estadual de Campinas
como parte dos requisitos exigidos para a
obtenção do Título de Doutora em
Genética e Biologia Molecular na área de
Genética Vegetal e Melhoramento.*

*Thesis presented to the Institute of
Biology of the University of Campinas in
partial fulfillment of the requirements for
the degree of PhD in Genetics and
Molecular Biology in the area of Plant
Genetics and Breeding.*

Orientador: Marcos Antonio Machado

ESTE ARQUIVO DIGITAL CORRESPONDE À
VERSÃO FINAL DA TESE DEFENDIDA PELA
ALUNA MAIARA CURTOLO E ORIENTADA PELO
MARCOS ANTONIO MACHADO.

**CAMPINAS
2021**

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Biologia
Mara Janaina de Oliveira - CRB 8/6972

C948t Curtolo, Maiara, 1991-
Target identification strategies and sweet orange (*Citrus sinensis* L. Osb.) genome editing for tolerance to Huanglongbing / Maiara Curtolo. – Campinas, SP : [s.n.], 2021.

Orientador: Marcos Antonio Machado.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. Transcriptoma. 2. Mapeamento genômico vegetal. 3. RNA-seq. 4. *Huanglongbing*. I. Machado, Marcos Antonio. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Estratégias de identificação de genes alvos e edição de genoma de laranja doce (*Citrus sinensis* L. Osb.) para tolerância ao *Huanglongbing*

Palavras-chave em inglês:

Transcriptome

Plant genome mapping

RNA-seq

Greening disease

Área de concentração: Genética Vegetal e Melhoramento

Titulação: Doutora em Genética e Biologia Molecular

Banca examinadora:

Marcos Antonio Machado [Orientador]

Renato Vicentini dos Santos

Raquel Luciana Boscariol Camargo

Monalisa Sampaio Carneiro

Francisco de Assis Alves Mourão Filho

Data de defesa: 27-04-2021

Programa de Pós-Graduação: Genética e Biologia Molecular

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <http://orcid.org/0000-0003-3210-7614>

- Currículo Lattes do autor: <http://lattes.cnpq.br/3072353836156464>

Campinas, 27 de abril de 2021

COMISSÃO EXAMINADORA

Dr. Marcos Antonio Machado

Prof. Dr. Renato Vicentini dos Santos

Dra. Raquel Luciana Boscariol Camargo

Dra. Monalisa Sampaio Carneiro

Dr. Francisco de Assis Alves Mourão Filho

Os membros da Comissão Examinadora acima assinaram a Ata de defesa, que se encontra no processo de vida acadêmica do aluno.

A Ata da defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa de Pós-Graduação em Genética e Biologia Molecular do Instituto de Biologia da Unicamp.

AGRADECIMENTOS

Ao Instituto de Biologia da Unicamp, em especial ao programa de pós-graduação em Genética e Biologia Molecular e ao Centro de Citricultura Sylvio Moreira pela oportunidade de realizar esse trabalho;

Ao Dr. Marcos Antônio Machado, pela orientação, oportunidade e confiança em meu trabalho;

A Dra. Mariângela Cristofany-Yaly, pelas valiosas contribuições para o desenvolvimento desse trabalho, paciência e dedicação, por quem tenho enorme carinho e respeito.

Aos professores Marco Aurélio Takita, Alessandra Alves de Souza, Raquel Boscarol, Rodrigo Latado e Rodrigo Gazaffi por todo o conhecimento que tive o privilégio de receber ao longo desses anos e pela imensurável contribuição para a concretização deste trabalho.

Aos funcionários do Centro de Citricultura, especialmente a Anita, Cícera e Kleber pelo companheirismo e suporte técnico-científico;

Aos amigos de pós-graduação: Eduardas, Nathalias, Inaiara, Reinaldo, Isis, Mariana, Thais, Lais, Diogo, Paula, Maju, Cesar e Marcela pelos conhecimentos compartilhados, boas risadas e companheirismo;

A todos aqueles que de alguma forma possibilitaram a execução dessa pesquisa, o meu agradecimento e carinho.

Aos membros da banca pelo aceite e contribuições para o trabalho;

À Fundação de Amparo à pesquisa do Estado de São Paulo (Fapesp), pela concessão da bolsa de doutorado processo: 2016/22133-0.

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

Resumo

Apesar dos avanços na citricultura brasileira as pragas e doenças ainda são consideradas como os principais entraves na cultura dos citros. O *Huanglongbing* (HLB) ou Greening vem sendo considerada a doença mais devastadora dos citros, uma vez que todas as variedades comerciais são extremamente suscetíveis a doença. Deste modo, a busca do conhecimento dos mecanismos genéticos que desencadeiam muitos processos biológicos envolvidos na resposta ao HLB é de extrema importância. Sendo assim, o presente trabalho teve como finalidade: a) Utilizar diferentes técnicas biotecnológicas (RNA-seq e Mapeamento de eQTL) para identificação de genes alvos relacionados com suscetibilidade, tolerância e resistência ao HLB; b) Desenvolver uma plataforma de edição de genoma via CRISPR/Cas9 que, futuramente poderá permitir o desenvolvimento de plantas de citros mais tolerantes ao HLB. Neste trabalho, a análise de seis transcriptomas de diferentes genótipos infectados com *Candidatus Liberibacter asiaticus* (CLAs) permitiu a construção de um modelo hipotético para entender os mecanismos genéticos envolvidos na tolerância ao HLB. A deposição de grandes quantidades de calose e de proteínas P (proteínas de floema) nas placas crivadas do floema parece ser a principal alteração que determina os sintomas característicos do HLB. Os transcriptomas analisados neste trabalho também indicaram que genes relacionados a síntese de calose e proteínas P tiveram o nível expressão alterado pela infecção por CLAs. Por este motivo, a expressão de oito genes envolvidos na síntese de calose foi avaliada em híbridos de *Citrus sunki* e *Poncirus trifoliata* infectados com CLAs. Os dados de expressão foram associados a polimorfismo de nucleotídeo único (SNP) de *C. sunki* e *P. trifoliata*, permitindo a identificação de regiões genômicas potenciais para futuros estudos de programas de melhoramento de citros. As análises transcriptômicas indicaram que o gene *Sieve Element Occlusion c* (*SEOc*) pode estar relacionado à suscetibilidade ao HLB. Além disso, estudos anteriores também demonstraram que os membros da família *SEO* codificam as subunidades de proteínas P. Sendo assim, *SEOc* foi o gene alvo escolhido para ser trabalhado durante o desenvolvimento da plataforma de edição de genomas. O sistema CRISPR / Cas9 foi utilizado para modificar o genoma de tabaco e citros. Os resultados gerados por este trabalho mostram que otimizações da técnica de edição devem ser adotadas visando aumentar a taxa de edição em citros.

Palavras chave: transcriptoma, edição de genoma, mapeamento, Greening

Abstract

Despite the advances in the Brazilian citrus industry, pests and diseases are still considered the main obstacles in citrus groves. Huanglongbing (HLB) or Greening is currently considered the most devastating citrus disease, since all commercial varieties are highly susceptible to this disease. Thus, the pursuit of knowledge of genetic mechanisms that trigger many biological processes involved in the response to the HLB is of utmost importance for the citrus industry. The present work aimed: a) To use different biotechnological techniques (RNA-seq and eQTL Mapping) aiming the identification of target genes related to susceptibility, tolerance and resistance to HLB. b) To develop a genome editing platform based at CRISPR / Cas9 system. Posteriorly, that platform will allow the development of citrus plants more tolerant to HLB. The transcriptomic analysis of six different genotypes led us to build a hypothetical model to understand the genetic mechanisms involved in HLB tolerance. The deposition of large amounts of P protein and callose on the phloem sieve plates seems to be the main alteration that determines the typical HLB symptoms. Transcriptomic analysis also indicated that *calloses synthases* and *P proteins* genes also had their expression level altered by *Candidatus Liberibacter asiaticus* (CLAs) infection. The expression of eight *callose synthase* genes was evaluated in hybrids between *Citrus sunki* (HLB susceptible) and *Poncirus trifoliata* (HLB tolerant) infected with HLB. The expression data were associated with single nucleotide polymorphism (SNP) of *C. sunki* and *P. trifoliata*, allowing the identification of interesting genomic regions for citrus breeding programs. Transcriptomes indicated that *Sieve Element Occlusion c (SEOc)* may be related to HLB susceptibility. In addition, previous studies also demonstrated that members of the SEO family encode P-protein subunits. Thus, *SEOc* was the target gene chosen to be worked during the development of the genome editing platform. The CRISPR / Cas9 system was used to modify the tobacco and citrus genome. The results generated by this work indicated that optimization of the genome editing technique should be adopted in order to increase the edition rate in citrus.

Key words: transcriptome, genome editing, mapping, Greening

Summary

1. General introduction -----	10
1.1. Citriculture and Huanglongbing -----	10
1.2. The state of the art of transcriptomic analysis in Citrus-CLas interaction-----	14
1.3. Linkage Map Construction, a Brief Introduction-----	15
1.3.1. Developed Linkage Maps in Citrus-----	15
1.4. QTL Mapping for Particular Traits-----	18
1.4.1. eQTL studies in citrus -----	19
1.5. CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) -----	21
1.5.1. CRISPR as immunity system and technology-----	21
1.5.2. Citrus genome editing -----	23
2. Objectives-----	24
3. Results-----	25
3.1. Chapter 1: Expression Quantitative Trait Loci (eQTL) mapping for callose synthases in intergeneric hybrids of Citrus challenged with the bacteria <i>Candidatus Liberibacter asiaticus</i> . -----	25
3.1.1. Abstract -----	26
3.1.2. Introduction -----	27
3.1.3. Materials and Methods -----	28
3.1.3.1. Plant material -----	28
3.1.3.2. DNA extraction and molecular marker analysis -----	29
3.1.3.3. Linkage Maps -----	29
3.1.3.4. RNA extraction and cDNA synthesis -----	30
3.1.3.5. Real-time Quantitative PCR (RT-qPCR) -----	30
3.1.3.6. Gene expression profile and genetic parameter analyses -----	31
3.1.3.7. eQTL mapping-----	31
3.1.4. Results-----	32
3.1.4.1. <i>C. sunki</i> and <i>P. trifoliata</i> linkage maps-----	32
3.1.4.2. Gene expression profile-----	38
3.1.4.3. eQTL mapping-----	41
3.1.5. Discussion-----	46
3.1.6. Conclusion-----	51
3.1.7. Supplementary Information-----	52
3.1.8. References -----	53
3.1.9. Internet Resources-----	58
3.1.10. Supplementary material -----	59
3.2. Chapter 2: Wide-ranging transcriptomic analysis of <i>Poncirus trifoliata</i> , <i>Citrus sunki</i> , <i>Citrus sinensis</i> and contrasting hybrids reveals HLB tolerance mechanisms -----	62
3.2.1. Abstract -----	62
3.2.2. Introduction -----	63
3.2.3. Results -----	64
3.2.3.1. CLas quantification-----	64
3.2.3.2. Phenotypic analysis -----	66
3.2.3.3. Transcriptome assembly -----	68
3.2.3.4. Differential gene expression analysis-----	69
3.2.3.5. Main processes affected by CLas infection-----	70

3.2.3.6. Differentially expressed genes (DEGs) associated with a specific biological pathway. -----	71
3.2.3.6.1. Signaling receptor -----	71
3.2.3.6.2. Hormones -----	72
3.2.3.6.3. Transcription factors -----	72
3.2.3.6.4. Defense-related genes -----	72
3.2.3.6.5. Secondary metabolism and cell wall composition -----	73
3.2.3.6.6. Phloem-related genes -----	74
3.2.3.6.7. Carbohydrate metabolism -----	74
3.2.3.6.8. Transporters -----	75
3.2.4. Discussion -----	75
3.2.5. Materials and Methods -----	80
3.2.5.1. Plant material -----	80
3.2.5.2. CLas quantification -----	81
3.2.5.3. Phenotypic analysis -----	81
3.2.5.4. RNA extraction and sequencing (RNA-seq) -----	82
3.2.5.5. Data analysis -----	82
3.2.5.6. Real time PCR (RT-qPCR) validation -----	83
3.2.6. Supplementary Information -----	83
3.2.7. References -----	84
3.2.8. Supplementary Material -----	88
3.3. Chapter 3: CRISPR/Cas system targeting Sieve Element Occlusion gene to improve HLB tolerance in sweet orange trees -----	96
3.3.1. Abstract -----	96
3.3.2. Introduction -----	96
3.3.3. Material and Methods -----	98
3.3.3.1. Molecular Cloning -----	98
3.3.3.2. Plant transformation -----	99
3.3.3.3. Screening of transgenic and edited plants -----	99
3.3.3.4. Immunoblotting -----	100
3.3.4. Results -----	100
3.3.4.1. Molecular cloning sgRNAs for SEO gene in pDirect22c and genetic transformation of citrus and tobacco -----	100
3.3.4.2. Immunoblotting of citrus transgenic leaves -----	104
3.3.4.3. Identification of edited SEOc gene in citrus genotype and NtSEO1 in tobacco	105
3.3.5. Discussion -----	108
3.3.6. Conclusion -----	111
3.3.7. References -----	111
4. Final consideration -----	115
5. References -----	116
6. Anexos -----	125
6.1. Declaração de Bioética -----	125
6.2. Declaração sobre direitos autorais -----	126

1. General introduction

Part of the general introduction (Linkage Map Construction, a Brief Introduction, Developed Linkage Maps in Citrus and QTL Mapping for Particular Traits) was included in the book chapter “Markers, Maps, and Markers-Assisted Selection”; authors: Shimizu T, Kacar YA, Cristofani-Yaly M, Curtolo M, Machado MA. Chapter published in “The Citrus Genome” (2020). Compendium of Plant Genomes. Springer, Cham. ISBN 978-3-030-10799-4.

1.1. Citriculture and Huanglongbing

Brazil is one of the most important countries for citrus production worldwide. Our country is the second largest citrus producer in the world, being the first largest producer of sweet orange, sixth largest producer of tangerines and fifth largest producer of limes and lemons (Fao, 2020). Currently, São Paulo and Minas Gerais citrus belt are responsible for most of the national production (Sampaio Passos et al., 2019). The citrus industry is one the main commodities for the Brazilian economy, generating billions of dollars in exported orange juice (Fundecitrus, 2020). This wealth is distributed among hundreds of companies linked directly to the sector, including thousands of rural properties, generating thousands of direct and indirect jobs, collecting taxes and leveraging the economy.

Apart from the commercial and economic aspects, the biggest challenge of citrus farming are pests, diseases and abiotic factors which are the main limiting elements of the Brazilian citrus industry, representing a large part of the production cost (Donkersley et al., 2018). Among its limitations, a disease caused by a Gram-negative bacteria *Candidatus Liberibacter* species has caused huge losses in commercial production independently of the region where it occurs. In Brazil, specifically in the citrus belt, 18.5% of citrus trees were infected with HLB. Added to other recurrent phytosanitary problems a reduction of 28% in Brazilian citrus production was reported (Fundecitrus, 2018).

The following three *Candidatus Liberibacter* (*Ca. Liberibacter*) species have been associated with HLB: *Ca. Liberibacter asiaticus*, (CLas), *Ca. Liberibacter americanus* (CLam), and *Ca. Liberibacter africanus* (CLaf). CLas is the most widespread and responsible for large economic losses worldwide. It is worthy to mention that *Ca. Liberibacter* species have not proved possible to maintain under axenic culture conditions, making impossible to develop key studies to shed light on the pathogen biology and host-interaction aspects (Davis et al., 2008).

HLB can be transmitted by both the citrus vector *Diaphorina citri* (psyllid) and experimentally through grafting with CLas-contaminated buds (Hilf and Lewis, 2016). The disease leads to the development of several symptoms including blotchy chlorosis, mottling of leaves, shoot yellowing, corking veins, stunted growth and small, green, and lopsided fruits with aborted seeds (Johnson et al., 2014). HLB symptoms are considered as a consequence of a series of molecular, cellular and physiological disorders in the host plant. The most expressive modifications caused by CLAs in citrus are alterations in sucrose and starch metabolism, changes in hormones production, biosynthesis of secondary metabolites, phloem function disorders and source-sink communication (Balan et al., 2018).

Citrus plants recognize pathogen-associated molecular patterns (PAMPs) of CLAs, triggering callose deposition in the phloem sieve plates (Luna et al., 2011). The deposition of high amounts of callose and phloem proteins (P protein) on the phloem sieve plates interferes with the transport of photo assimilates from source leaves to the sink organs (Wang et al., 2017). This disruption results in excessive starch accumulation in chloroplasts (Wang and Trivedi, 2013; Boava et al., 2017). Starch accumulation causes the disintegration of the thylakoid system, resulting in the yellowing leaf mottle symptom (Etxeberria et al., 2009), consequently, other typical HLB symptoms occur. The roots of susceptible citrus trees after CLAs infection attest the evidence of partial obstruction of the host phloem nutrition system (Figure 1).

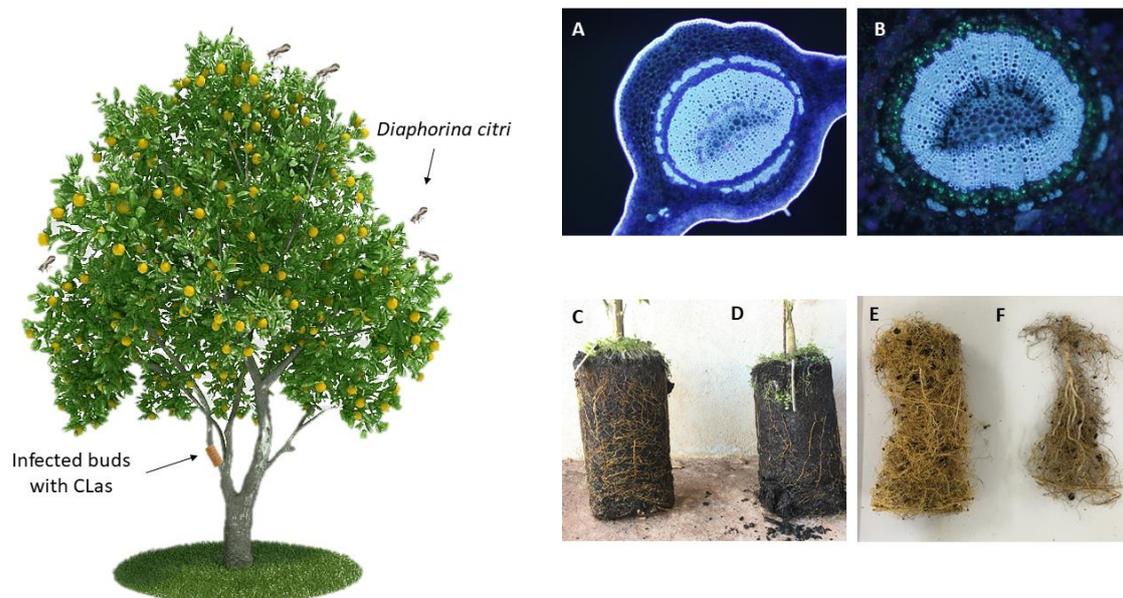


Figure 1: Representation of *Ca. Liberibacter asiaticus* acquisition and symptomatology. Grafting with contaminated buds and insect vector (*Diaphorina citri*) as inoculum sources. (A) Absence of callose accumulation in phloem in mock-inoculated sweet orange plants and normal

root development (C and E); (B) Expressive accumulation of callose in phloem and impaired root system in CLAs-inoculated sweet orange plants (D and F). Photos A-F: Maiara Curtolo.

Most of the HLB symptoms may be confused with nutrient deficiencies and other diseases. In addition to the long period of latency of the bacteria in infected plants, HLB diagnosis cannot be done exclusively through the observation of the visual symptoms.

Recently, effectors of CLAs have been predicted and some of them seem to be directly related to the HLB symptoms (Pitino et al., 2016). *Las5315mp* effector induced starch accumulation, callose deposition and cell death in *Nicotiana benthamiana* (Pitino et al., 2016, 2018). *Sec-delivered effector 1 (SDE1)* was previously characterized as an inhibitor of defense-related genes (Clark et al., 2018). Further, *SDE1* expression in *Arabidopsis thaliana* resulted in severe yellowing in mature leaves, similar to CLAs infection symptoms and accelerated senescence (Clark et al., 2020).

All commercial Citrus varieties are susceptible to CLAs infection and no effective source of resistance to HLB is known so far. Thus, the identification of tolerant genotypes is essential to the maintenance of citrus production.

Poncirus trifoliata, a close relative and sexually compatible species with the *Citrus* genus has been highlighted in numerous studies as a tolerant genotype. It does not present typical HLB symptoms and CLAs titer remains low or nonexistent (Albrecht and Bowman, 2012; Boava et al., 2017). Some Citrus x *P. trifoliata* hybrids have also been reported to have a significant tolerance to HLB (Figure 2) (Boava et al., 2017). Additionally, it is an important rootstock source for citriculture because of its tolerance/resistance to *Phytophthora*, citrus tristeza virus (CTV) and nematodes (Pang et al., 2007). Due to those traits, *P. trifoliata* and its hybrids have been reported as a possible source of tolerance/resistance to HLB. However, it remains unclear which mechanisms are involved in this tolerance/resistance.

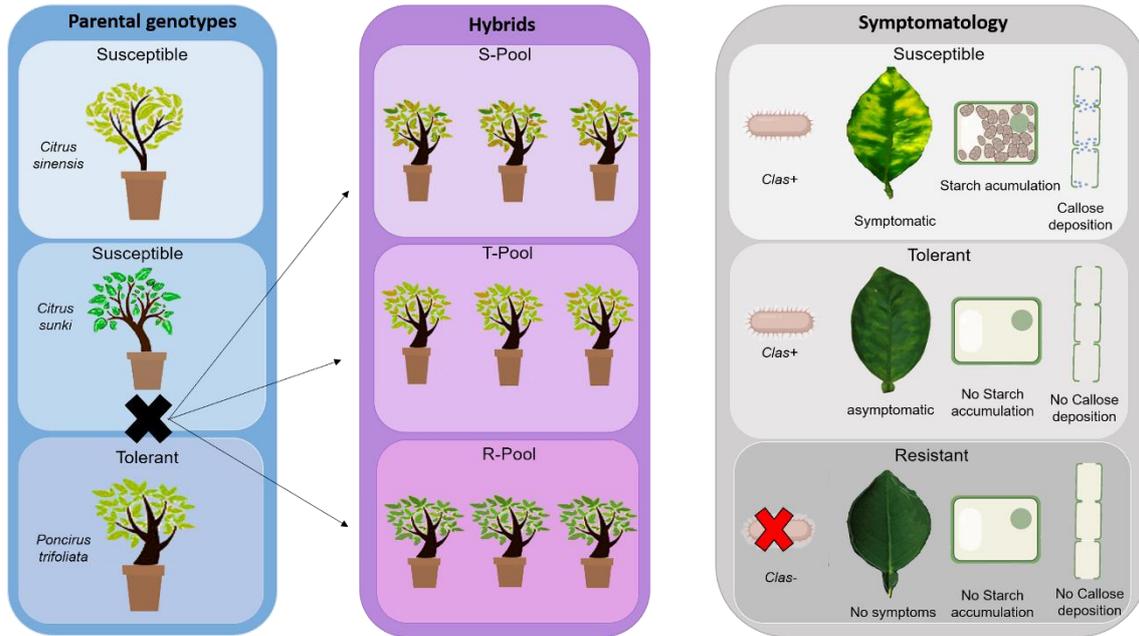


Figure 2: The crosses between the susceptible (*C. sunki*) and the tolerant genotypes (*P. trifoliata*) generated hybrids with different responses to HLB. Some hybrids were susceptible, such as plants that showed both CLAs titer and HLB typical symptoms, including mottle leaves and high accumulation of starch and callose. The plants that showed CLAs titer, non-visible HLB symptoms and no starch and callose accumulation were considered tolerant part of the population. And few hybrids were resistant, such as plants which presented neither detectable CLAs titer nor symptoms or starch and callose accumulation. Figure extracted from Curtolo et al., (2020a).

The study of HLB-tolerant / resistant genetic materials has provided the understanding of the virulence mechanisms and the development of many approaches to HLB control. Efforts have been made in order to test a series of anti-microbial peptides in Gram-negative bacteria, aiming the identification of molecules with functions of treating and preventing citrus Huanglongbing (Velasquez Guzman et al., 2018). A heat-stable antimicrobial peptide MaSAMP (*Microcitrus australasica* Stable Antimicrobial Peptide) was firstly identified from HLB-tolerant Australian finger lime (*Microcitrus australasica*). MaSAMP has triggered interest due to its capacity of both killing CLAs and inducing plant immunity (Huang et al., 2021). The anatomical characteristics of leaf lamina transverse sections in HLB tolerant trees, such as “Bearss” lemon and “LB8-9” Sugar Belle® mandarin demonstrated substantial phloem regeneration. That mechanism seems to compensate the dysfunctional phloem caused by CLAs (Deng et al., 2019). Previously, Granato et al., (2020) have already reported that anatomical divergences between tolerant and

susceptible citrus trees could represent an important feature to avoid sieve tube elements collapse.

1.2. The state of the art of transcriptomic analysis in Citrus-CLAs interaction

Different genotypes react differentially upon CLAs infection and this response can be extremely complex since it encompasses numerous physiological and metabolic reprogramming. Many genes and mechanisms are involved in the susceptibility, tolerance and resistance responses. Several studies using microarray and high-throughput RNA sequencing technology (RNA-seq) have been used to understand the global gene expression patterns in different genotypes infected with CLAs (Martinelli et al., 2012; Mafra et al., 2014; Wang et al., 2016). The transcriptomic analysis from both microarray and RNA-seq techniques together provide a complete set of transcripts and their quantity (Wang et al., 2009). The first transcriptomic analysis in Citrus genotypes infected with CLAs focused on the evaluation of a single genotype (Mafra et al., 2014). The information obtained by analyzing only one response/interaction provided a wide knowledge about the CLAs-citrus interaction. However, due to the complexity of the HLB disease, further analysis in contrasting genotypes became necessary to shed some light on the mechanisms underlying Citrus resistance to CLAs.

Global gene expression changes in leaves assessed by microarray indicated a large number of physiological processes impacted by HLB (Albrecht and Bowman, 2012; Mafra et al., 2014). The authors reported that HLB symptoms are caused by changes in metabolism, especially the ones related to sucrose and starch related processes, stimulation of hormone production, biosynthesis of secondary metabolites, phloem function disorders and source-drain communication. These observations corroborate with what was previously reported by physiological studies (Martinelli et al., 2012; Mafra et al., 2014).

Some key pathways and genes were linked with the susceptibility, tolerance and resistance responses, since they were significantly altered by CLAs infection. However, the results among the transcriptomic studies are somewhat divergent. Previous studies point out that hormone modulations, cell wall strengthening, signaling, transcription factors, secondary metabolites and Pathogenesis Related genes (PR-genes) are important functional categories related to tolerance responses in different genotypes. But there is no defined mechanism or sets of genes that have been widely reported to be responsible for different responses so far. For instance, the induction of *constitutive disease resistance (CDR)* genes (Rawat et al., 2017), downregulation of *beta glucanases*, *DRM6-like*, *expansin* and *DET2* and the induction of *NPR1* were pointed as genes responsible for inducing the tolerance and resistance (Wang et al., 2016). So, despite the recent

studies and their findings, there is no definitive solution, gene or manner to control HLB. But most certainly all those information supply valuable resource to genetic breeding.

1.3.Linkage Map Construction, a Brief Introduction

Techniques involving molecular markers have solved some of the limiting problems associated with classical breeding in citrus and other perennial species. Citrus breeding is affected by long juvenility, heterozygosis, gametophytic cross-incompatibility, male sterility, apomixis, seedlessness, and trait stability under different environmental conditions (Ollitrault et al., 2012; Xu et al., 2013; Curtolo et al., 2017). Since the beginning of the 1990s, molecular markers have been used for evaluation or characterization of active germplasm collections, identification of nucellar hybrid seedlings in progenies of controlled crosses, the study of phylogenetic relationships and genetic divergence, and genetic mapping (Machado et al., 2011). When using a polyembryonic female parent, the identification of zygotic embryos at the seedling stage makes it possible to obtain many progenies that can be used for genetic mapping and to study the heritability of traits. In the past, studies focused on the progenies obtained from crosses that included *Poncirus trifoliata* as one of the parents, because of its importance for rootstock breeding. *Poncirus* was also used because the trifoliolate leaf, which is a characteristic with monogenic and dominant inheritance, allowed the selection of zygotic plants in the progeny. The monoembryonic cultivars, like Fortune and Clementine mandarins, were also widely used as the female parent in breeding programs in Spain, France, and Italy (Cuenca et al., 2013). With the advent of molecular markers and the ease of genotyping, many progenies could be obtained even using polyembryonic cultivars as the female parent. Genetic mapping, which is of fundamental importance in breeding programs, has benefited from the improvement of genotyping techniques. A high-density linkage map is fundamental for QTL (Quantitative Trait Loci) mapping, marker-assisted selection (MAS), and candidate gene identification within the QTL intervals and gene cloning. Compared with other crops, genetic mapping in citrus is relatively less well developed, but this scenario is changing.

1.3.1. Developed Linkage Maps in Citrus

There are a reasonable number of linkage maps for citrus so far. However, with the accumulation of knowledge and advancement of technologies for obtaining molecular markers, the maps are being continually updated. These maps are becoming increasingly representative, making it possible to compare them with reference genomes or even use them to assist in the assembly or updating of sequenced genomes (Ollitrault et al., 2012; Xu et al., 2013; Curtolo et

al., 2017). Initially, RFLP (Restriction Fragment Length Polymorphism), RAPD (Random Amplified Polymorphic DNA), AFLP (Amplified Fragment Length Polymorphism), and isoenzyme markers were the most commonly used markers for linkage mapping. From the twenty-first century onward, ISSR (Inter-Simple Sequence Repeat), IRAP (amplification of repeated sequences from sites of retrotransposons), SSR (Simple-sequence repeats), and those markers obtained through high throughput sequencing were also used to generate the linkage maps. Currently, the DNA markers derived using high-throughput technology are the most commonly used, allowing several genotypes to be analyzed at the same time and thousands of markers to be generated at once. Citrus plants are perennial species, so F₁ populations are typically used for the construction of citrus maps although backcrossing and F₂ progeny are also sometimes used (Chen et al., 2008; Raga et al., 2012). In these cases, linkage mapping can be performed for each heterozygous parental individual separately using single-dose DNA polymorphisms segregating 1:1. Such mating configurations are displayed where the marker is present in one parent, absent in the other, and segregating in the progeny. Grattapaglia and Sederoff (1994) called this mapping strategy ‘*pseudo-testcross*’ because the testcross mating configuration of the markers is not known a priori, as in a conventional testcross in which the tester is homozygous recessive for the locus of interest. A two-way pseudo-testcross has been conducted in the F₁ population in citrus (Cristofani et al., 1999; Weber et al., 2003). According to Weber et al., (2003) in this design, a map of each parent is constructed by grouping marker alleles originating from each parent for analysis but the collinearity between the maps cannot be determined without an intermediary map or codominant markers in both parents. For F₁ populations, with markers that segregate 3:1 (dominant), 1:2:1 (codominant), and 1:1:1:1 (codominant), an integrated map can be built (Curtolo et al., 2017). Durham et al., (1992), Cai et al., (1994), Gmitter et al., (1996) and Deng et al., (1997) constructed integrated maps considering *P. trifoliata* as one of the parents in population formation. Integration was possible because of the type of cross once F₂ and backcrossing were adopted.

The number of markers anchored in the maps reflects the evolution of the technologies for marker production and analysis. Durham et al., (1992), Liou et al., (1996) and Gulsen et al., (2010) associated at least two types of markers in the linkage analysis. Among them, Ling et al., (1999) gathered information from polymorphisms generated by AFLP, RFLP, and isoenzyme markers, then obtained the map with the highest number of markers (337). With high throughput genotyping technology, high-density linkage maps were developed for Clementine (Ollitrault et al., 2012), *P. trifoliata*, and *C. sunki* (Curtolo et al., 2018), for instance. Curtolo et al., (2018) using only NGS (Next-Generation Sequencing) combined

diversity arrays technology (DArTseq) markers, obtained the map with the highest number of markers so far. As the number of markers has increased, the genomic coverage of the maps has consequently increased as well. However, attention must be taken when associating genomic coverage with map saturation. In addition to genomic coverage, one should analyze the degree of density of the linkage groups or the number of markers per unit of recombination (cM – centimorgan). The presence of large gaps between markers can often give the false impression of high genomic coverage. For example, Luro et al., (1996) using RAPD, RFLP, and isoenzymes built a map with 95 markers for *Poncirus*, distributed in nine linkage groups, with genomic coverage of 1,503 cM representing, on average, one marker every 15 cM. With the improvement of genotyping methods, the density of markers in the maps increased. Guo et al., (2015) and Xu et al., (2013) published a dense map for citrus, with around one marker per cM. Curtolo et al., (2018) built the map with the largest number of anchored markers (3,084 for *P. trifoliata*); however, it is not the most saturated map, because some of the markers were positioned at the same loci. While these markers, with recombination frequency equal to zero ($Fr = 0$), are not genetically informative, this map provides genomic information for candidate gene identification within the QTL intervals. The number of genotyped individuals in the progeny establishes the maximum level of resolution that can be reached with a saturated number of markers in the genetic map. Curtolo et al., (2018, 2017) used a relatively large population (276 individuals) when compared with the other previous studies. Nevertheless, the number of individuals used was not large enough to require the use of a high-throughput genotyping system. According to Omura & Shimada, (2016) chromosome transmission to progeny in citrus tends to result in the inheritance of large linkage blocks and the frequency of recombination in a chromosome is low. To reach higher levels of polymorphism, it is necessary to advance generations through crosses or to work with large population sizes. Both approaches are difficult to apply in citrus because of the biological characteristics of the species. In the map of *P. trifoliata*, 1,782 loci were built using 276 hybrids with 3,084 DArTseq markers (Curtolo et al., 2018). Compared with the map of Zhou et al., (2018) the number of individuals in the F_1 population was three times greater. This demonstrates that increasing the amount of recombination allows the use of fewer individuals on the map. As codominant markers are the most informative, the use of this type of marker can help to minimize the difficulties associated with increasing the population size needed for citrus mapping (Curtolo et al., 2017). In conclusion, the genetic density of the markers in the map is defined not only by the number of markers obtained but also by the number of recombination events occurring in meiosis, the size

of the population, population types, the nature of the markers involved, and the required statistical confidence (Ferreira et al., 2006).

1.4.QTL Mapping for Particular Traits

Several approaches, including bulked segregation analysis, linkage mapping, and QTL analysis have been used for the development of DNA markers linked to specific traits (Imai et al., 2018). The bulked segregation analysis (BSA) approach divides a hybrid population into two groups according to their distinguishable phenotype, and then mines a DNA marker allele that is specific to either one of those groups. BSA is simple and effective for developing a DNA marker for MAS and can be applied to a population too small for linkage mapping. Conversely, application of BSA is limited to simple qualitative traits regulated by a single gene of higher genetic effect. It is also difficult to identify a selectable DNA marker for quantitative loci with minor effects. Furthermore, this method provides no information on the genetic distance between the loci for the trait and mapped position of the selected DNA marker. Linkage mapping analysis and QTL analysis are conventional approaches for the identification of loci linked to a trait of interest. These analyses predict the distance between the trait and mapped DNA marker position and estimate the genetic contribution of the trait. Using a large-sized population or dense DNA markers for the analyses will help to improve the resolution. A quantitative trait loci (QTL) mapping using bi-parental populations is a key approach to dissect complex traits and identify genomic regions underlying quantitative traits for breeding purposes. In citrus, efforts have been made over the last two decades to dissect complex traits using a QTL mapping approach. Most characteristics of agronomic interest are controlled by quantitative loci and study of their QTL allows the identification, mapping, and quantification of their effects. Several factors influence the detection of these regions such as number and frequency of recombination of QTL, the magnitude of their effects, heritability characteristics, interaction between genes and types of markers, and degree of saturation of the genetic map. The mapping of QTL has favored breeding programs of several perennial species; in citrus, it was possible to map several characteristics with qualitative and quantitative inheritance. The identification of QTL in citrus focused on morphological traits as well as resistance to abiotic or biotic factors (Curtolo et al., 2017; Lima et al., 2018; Soratto et al., 2020). The association of molecular markers with citrus characteristics has been previously studied since 1994 with cold acclimation (Cai et al., 1994). Thereafter, genetic maps have been extended to localize important traits such as citrus tristeza virus (CTV) resistance (Gmitter et al., 1996; Cristofani et al., 1999), fruit acidity (Fang et al., 1997), apomixes (García et al., 1999), nematode

resistance (Ling et al., 2000) (Ling et al., 2000), *Phytophthora* gummosis resistance (Lima et al., 2018) and HLB tolerance (Huang et al., 2018; Soratto et al., 2020).

In some citrus maps, both qualitative and quantitative identification of loci is available. For example, using the same map, CTV and gummosis of *Phytophthora* resistance loci were mapped. Reviews of QTL mapping efforts in fruit trees and citrus were published (Iwata et al., 2016; Omura and Shimada, 2016). Recent work on populations of citrus scion varieties, especially mandarins, has focused on fruit characteristics. In these studies, the availability of maps with marker sequences enables the identification of candidate genes within the QTL intervals. The number of studies examining fruit-quality QTL in citrus is increasing. Curtolo et al., (2017) identified 19 QTL regions for 12 fruit characteristics, including fruit diameter using 278 F₁ hybrids from a cross between Murcott tangor and Pera sweet orange. Yu et al., (2016) reported the identification of molecular markers and candidate genes linked to mandarin fruit-quality traits in maps built using data generated from a 1536-SNP Illumina Golden Gate assay in two mandarin parents (Fortune and Murcott) and their 116 F₁ progeny. Accordingly, they identified 48 QTL regions for eight important fruit-quality traits, including fruit size or weight and flavedo color (Yu et al., 2016). They also used the same population and maps to identify single nucleotide polymorphism (SNP) markers associated with volatile traits and detected a total of 206 quantitative trait loci (QTL) for 94 volatile compounds. Some fruit aroma QTL were identified and the candidate genes in the terpenoid biosynthetic pathway were found within the QTL intervals. According to the authors, these QTLs could lead to an efficient and feasible MAS approach to mandarin fruit quality improvement (Yu et al., 2016, 2017). In these studies, the sequences that flank the QTL regions were available, allowing comparison between the results of mapping of different studies. For example, Imai et al., (2018) used association mapping and a genome-wide association study (GWAS) of fruit-quality traits in citrus using SNPs obtained by GBS (Genotyping by Sequencing). They found two regions for fruit weight that were common with QTLs in the maps reported by Imai et al., (2017) and one region for pulp firmness that was common with that reported by Minamikawa et al., (2017).

1.4.1. eQTL studies in citrus

Currently, studies using differential profile of gene expression approaches, such as microarray, RNA-Seq and RT-qPCR (Real Time Quantitative polymerase chain reaction) have been used to determine levels of gene expression in a segregating population. The expression data from population mapping can be associated with genotyping data from molecular markers, being analyzed as quantitative traits. That strategy allows the identification of genomic regions (eQTL

– expression quantitative trait loci) which can be related with variation of transcripts in co-regulated genes. In other words, eQTL studies involve a direct association between genomic locations with gene expression levels (Nica and Dermitzakis, 2013). Jansen and Nap, (2001) proposed genomic genetics as a technique that encompasses the quantitative locus mapping and analysis of gene expression to identify the association between the allelic state of a genome region and the quantification of gene transcripts. Schadt et al., (2003) referred to such genomic regions as expression QTLs (eQTL). The identification of eQTL and the genes whose expression they regulate is of great interest in revealing the key components of genetic architecture that trigger many biological processes. The use of that strategy would be specially interesting for understanding the processes involved in resistance to HLB.

Few expression quantitative trait loci mapping studies have already been performed in citrus populations. Those previous eQTL researches led to studies related to carotenoid metabolism and resistance to *Phytophthora* (Sugiyama et al., 2014; Lima et al., 2018). QTL involved in the citrus - HLB interaction using CLAs and starch quantification data were most recently identified by Soratto et al., (2020). The identification of eQTL using the gene expression values was also employed in that previous study. Fourteen genes had the expression profile mapped out in a population of hybrids between *P. trifoliata* and *C. sunki*. In addition, expression results and eQTL were compared with the starch and CLAs quantification. Some genes which had the expression data using in the eQTL study were involved in trehalose biosynthesis, starch degradation, metabolism of carbohydrates, phloem functionality, cell wall, metal ion transport, glucose metabolic processes and transcription factors. The eQTL and expression results demonstrated that all fourteen genes were affected by HLB disease and all were responsive to CLAs infection. Moreover, some of them were related to CLAs and starch quantification (Soratto et al., 2020).

Only fourteen genes related to HLB have already had their behavior investigated in mapping populations (Soratto et al., 2020). However, due to the fact that the genetic response to HLB is an extremely complex and polygenic trait, the study of more candidate genes would help further clarify the different responses to CLAs by different citrus genotypes, since the question why some citrus plants are tolerant or even resistant is still left unanswered.

As exposed by Soratto et al., (2020), the mapping eQTL related to HLB disease in robust linkage maps was successful applied. More information about HLB disease was obtained. But one of the main challenges in genetic mapping is to establish an association between the mapped information and the location of the gene (s) in the genome and whether and how the regions (eQTL) can affect in molecular and phenotypic response.

1.5. CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)

1.5.1. CRISPR as immunity system and technology

Genome engineering using CRISPR/Cas technology has revolutionized science, including plant, animal, and human research. The technology has been widely used for gene function discovery and biological processes studies, as well as for genetic breeding.

CRISPR technology was developed from immune system of prokaryotes (Barrangou and Marraffini, 2014). Firstly, it was observed that some bacteria had the capacity to degrade exogenous sequences from an invading phage or plasmid (Ishino et al., 1987). The bacteria with this ability present a specific signature pattern with approximately 32 nucleotides (nt) of non-repetitive sequences and “tandem repeats”. The loci have been called “Clustered Regularly Interspaced Short Palindromic Repeats” (CRISPR). A sequence with exogenous origin derived from plasmids or viruses was identified in sequence spacer of CRISPR locus. From this evidence, this unique bacterial immune system was elucidated. Part of the invading DNA is incorporated into the host's CRISPR locus as a spacer from the cleavage of its nucleic acid into small pieces by the Cas1 and Cas2 proteins. Once incorporated into the bacterium genome, in the case of a second invasion, those spacers begin to be transcribed and posteriorly processed in several smaller RNAs, called as crRNA (CRISPR-derived RNA). The crRNAs form a complex with the Cas9 protein, which is able to recognize and destroy the exogenous sequence (Ratner et al., 2016).

The understanding of CRISPR as a defense-related system led to the development of a new technology which made it possible to modify genomes in a fast, targeted and effective manners (Chen et al., 2019). That strategy has the specificity of modification, avoiding the appearance of undesirable mutations in other genomic regions (*off targets*), and the absence of exogenous gene insertions in the host genome. In addition to the potential of this methodology in generating new alleles, genomic editing performs an essential role in breeding strategies aiming plant design, including control of gene expression and metabolic reprogramming (Chen et al., 2019). Moreover, CRISPR system has other numerous applications such as: alterations in gene expression through silencing, repression, induction, and gain of function; modulation and alteration in protein activity, introduction of exogenous genes and gene location.

CRISPR as genome editing technology has two mainly active components: a guide RNA (sgRNA) and an endonuclease (Cas protein). The sgRNA contains the target genome sequence for editing and should be positioned next to a protospacer adjacent motif (PAM). The endonuclease Cas is conjugated to the sgRNA and it is responsible for the enzymatic reactions

of cutting, editing and binding the DNA (Ratner et al., 2016). Currently, Cas9 protein is the most widely used as a genome editing tool. Alternatively, Cas9 variants have been designed together with the identification of other Cas proteins, such as Cas12a formerly Cpf1 and Cas3. Each Cas protein has its peculiarities, for instance: Cas9 from *Streptococcus pyogenes* (*SpCas9*) typically generates a blunt double-strand break (DSB) and the DNA targeted by *SpCas9* relies on the 20-nucleotide-long spacer and on the PAM 5'. *Streptococcus thermophilus* Cas9 (*StCas9*) recognizes the PAM 5'-NGG, however, *SpCas9* variant can recognize different motifs (5'-NG, 5'-GAA and 5'-GAT) and can also cleave only one strand as a nickase (Pickar-Oliver and Gersbach, 2019). Cas12a in contrast to Cas9, performs a staggered cut with a 5' overhang at DNA target sites and does not use a transactivating RNA. It also has intrinsic RNase activity that allows Cas12a to cleave crRNA arrays to generate its own crRNAs. This ability enables multigene editing from a single RNA transcript (Pickar-Oliver and Gersbach, 2019; Gier et al., 2020). Cas3 presents both the nickase activity and the helicase activity. This specific protein can separate DNA duplex strands and at the same time displace other DNA binding proteins during translocation in order to generate a single strand break in DNA by targeting DNA degradation through 3' to 5' due to its exonuclease activity (He et al., 2020). Cas9, Cas12a, and Cas3 exhibit different structural architectures and consequently act as distinct molecular mechanisms, Therefore, Cas proteins are categorically divided in different classes. There is the class 1, which include the types I, III and IV and the types II and V are clustered into the class 2. Briefly, class 2 system is simpler, since the functions of the effector complex are performed by the action of a single protein, such as Cas9 and Cas12a. Meanwhile, the Cas proteins from class 1 require multi-subunit crRNA–effector complexes, as well as Cas3 (Weiss and Clark, 2002; Ratner et al., 2016; Swarts and Jinek, 2018).

All genome editing technologies (ZFN - zinc finger nucleases, TALEN - transcription activator-like effector nucleases or CRISPR) result in a double or single strand DNA breaks. The strand break is seen by the cell as a damage, where an endogenous repair mechanism is activated. Two major repair mechanisms can take place: nonhomologous end-joining (NHEJ) and homologous recombination (HR). Typically, knock-outs resulting from short insertions or deletions (*indels*) are due to DNA repair by NHEJ, since this mechanism is inherently error-prone. When a homologous DNA donor template is available or it is provided, the HR can occur, and the sequence can be either perfectly corrected or exogenous DNA sequences can be inserted (Pickar-Oliver and Gersbach, 2019). Those repair mechanisms have allowed the development of genome editing strategies with several specific applications.

1.5.2. Citrus genome editing

CRISPR has showed itself as an important technology with countless variations and applications in animals, humans and plants researches. In plants, several studies using the Cas proteins have emerged and some of them are related to proof of concept using target genes like *GFP* (Green fluorescent protein) (Permyakova et al., 2019), *GUS* (*Beta-glucuronidase*) (Michno et al., 2015) and *PDS* (*Phytoene Desaturase*) (Dutt et al., 2020). CRISPR/Cas technology has also been successfully used to target specific genomic sequences of interest for the development of edited plants in unimaginable ways (Song et al., 2016). And it has been highlighted as the promise of solution for all crop challenges, particularly for HLB in citrus.

Currently, there are relatively few citrus genome editing works and different strategies have been used in those studies. Transient transformation in sweet orange leaves is widely used to initially identify the efficiency and functionality of the adopted CRISPR system (Jia and Wang, 2014). However, in order to fully achieve the potential of the CRISPR/Cas technology, it is necessary to obtain stable genetically modified plants. Different explant sources have been used in citrus transformation, among them epicotyl tissues, embryogenic cell cultures and protoplasts (Dutt et al., 2020).

To date, there are studies using CRISPR/Cas system in citrus, but most of them are from a restricted group of researchers. CRISPR/Cas9 system was firstly used to target the *CsPDS* gene in sweet orange via Xcc-facilitated agro infiltration (Jia and Wang, 2014). Recently Dutt et al., (2020) have successfully edited the *CsPDS* gene using citrus embryogenic cell cultures. CRISPR/Cas9 technology also has been applied to increase citrus canker resistance mediated modification the *CsLOB1* (*Lateral Organ Boundaries 1*) gene in Duncan grapefruit (Jia et al., 2016, 2017). *CsLOB1* gene was related with citrus canker susceptibility (Hu et al., 2014). The used strategy mutated only one allele, but it was enough to alleviate the canker symptoms. Later, the edition of both alleles of *CsLOB1* promoters showed a high degree of resistance to citrus canker (Peng et al., 2017). The transformation efficiency and editing rate is extremely variable among the previous citrus genome editing studies.

The CRISPR editing system is restricted to the NHEJ repair system so far, which primarily promotes the knock-out of genes (Jia and Wang, 2014; Jia et al., 2019; Dutt et al., 2020).

Genome editing systems can be easily applied to citrus, at first glance. However, biological and practical evidence make it difficult to establish genome editing efficiently. Genetic transformation and regeneration of plants through the juvenile epicotyl or mature stem tissues produce mostly chimeric shoots (Domínguez et al., 2004; Dutt et al., 2020). In this case, non-

edited and edited cells are mixed, composing the tissue. This fact can drastically damage the identification of mutants since the editing rate can be very low due to its dilution. In addition, chimeric plants are also undesirable in the plant breeding process since they do not reach their full potential. Protoplast transformation can be an option to enhance genetic transformation efficiency and avoid chimeric plants, however, citrus protoplast regeneration is not a simple and easy process. Although the Clementine (*C. clementine*) and sweet orange (*Citrus sinensis*) genomes are already sequenced and assembled, the sequence of the target gene may be different among oranges, since the citrus genome is highly polymorphic with several SNPs (Curtolo et al., 2020b). There are other challenges and limitations to the application of the technology in citrus.

Those difficulties combined with a particularly complicated disease such as HLB represent a great challenge for citriculture.

2. Objectives

- To identify possible tolerance loci combining the expression quantitative trait loci (eQTL) of different *callose synthases* and genetic Single-Nucleotide Polymorphism (SNP) maps of *C. sunki* and *P. trifoliata*.
- To perform a wide-ranging transcriptomic analysis using contrasting genotypes regarding HLB severity to identify the genetic mechanism associated with tolerance to HLB
- To establish a genome editing (CRISPR/Cas) platform which can be used in the development of HLB tolerant plants

The thesis was divided into three chapters in order to better present the results of each objective.

3. Results

3.1. Chapter 1: Curtolo M, Moreira Granato L, Aparecida T, et al (2020) Expression Quantitative Trait Loci (eQTL) mapping for *callose synthases* in intergeneric hybrids of Citrus challenged with the bacteria *Candidatus Liberibacter asiaticus*. **Genet. Mol. Biol.** [online]. vol.43, n.2, e20190133. Epub June 15, 2020. ISSN 1678-4685. <https://doi.org/10.1590/1678-4685-GMB-2019-0133>

Expression Quantitative Trait Loci (eQTL) mapping for *Callose Synthases* in intergeneric hybrids of *Citrus* challenged with the bacteria *Candidatus Liberibacter asiaticus*

eQTL mapping for *Callose Synthases*

Maiara Curtolo^{1,2*}, Laís Moreira Granato¹, Tatiany Aparecida Teixeira Soratto¹, Maisa Curtolo³, Rodrigo Gazaffi⁴, Marco Aurélio Takita¹, Mariângela Cristofani-Yaly¹, Marcos Antonio Machado¹

¹Centro APTA Citros Sylvio Moreira, Instituto Agronômico de Campinas, Cordeirópolis/SP, Brazil;

²Programa de Pós-Graduação em Genética e Biologia Molecular, Universidade Estadual de Campinas, Campinas/SP, Brazil;

³Programa de Pós-Graduação em Genética e Melhoramento de Plantas, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba/SP, Brazil;

⁴Departamento de Biotecnologia e Produção Vegetal e Animal, Centro de Ciências Agrárias, Universidade Federal de São Carlos, Araras/SP, Brazil.

*Corresponding author: maiaramc@hotmail.com

Maiara Curtolo: 0000-0003-3210-7614

Laís Moreira Granato: 0000-0002-1599-8729

Rodrigo Gazaffi: 0000-0001-6549-4222

Marco Aurélio Takita: 0000-0003-3935-1361

Mariângela Cristofani-Yaly: 0000-0003-0283-989X

Marcos Antonio Machado: 0000-0001-9780-3990

3.1.1. Abstract

Citrus plants have been extremely affected by Huanglongbing (HLB) worldwide, causing economic losses. HLB disease causes disorders in citrus plants, leading to callose deposition in the phloem vessel sieve plates. Callose is synthesized by callose synthases, which are encoded by 12 genes (*calS1– calS12*) in *Arabidopsis thaliana*. We evaluated the expression of eight callose synthase genes from *Citrus* in hybrids between *Citrus sunki* and *Poncirus trifoliata* infected with HLB. The objective of this work was to identify possible tolerance loci combining the expression quantitative trait loci (eQTL) of different callose synthases and genetic Single-Nucleotide Polymorphism (SNP) maps of *C. sunki* and *P. trifoliata*. The expression data from all *Cscals* ranged widely among the hybrids. Furthermore, the data allowed the detection of 18 eQTL in the *C. sunki* map and 34 eQTL in the *P. trifoliata* map. In both maps, some eQTL for different *Cscals* were overlapped; thus, a single region could be associated with the regulation of more than one *Cscals*. The regions identified in this work can be interesting targets for future studies of *Citrus* breeding programs to manipulate callose synthesis during HLB infection.

Keywords: gene expression, molecular markers, polymorphism

3.1.2. Introduction

The citrus industry plays an important role in the productivity chain in Brazilian agribusiness. Brazil is the largest sweet orange producer, and, during the period 2017/18, its yield was approximately 397 million of boxes of 40.8 kg each (Fundecitrus, 2018). Nevertheless, this important economic area has been challenged by Huanglongbing (HLB) (Colleta-Filho et al., 2004), which has caused great economic losses because of the fast dissemination and severity. In 2008, 0.61% of the crop trees were symptomatic; in 2016, this number increased to 16.92%. In four years of evaluation, 50% of the scion trees showed disease symptoms, with an approximately 60% decrease in production (Fundecitrus, 2018).

HLB is caused by the gram-negative bacterium *Candidatus Liberibacter asiaticus* (CLas) (Colleta-Filho et al., 2004), which is restricted to the phloem sieve tubes (Jagoueix et al., 1994), and is transmitted by the vector citrus psyllid (*Diaphorina citri*) (Gottwald, 2010). Citrus plants recognize pathogen-associated molecular patterns (PAMPs) of CLas, triggering callose deposition in the phloem sieve plates (Gómez-Gómez et al., 1999; Luna et al., 2011). The deposition of high amounts of callose and phloem proteins (PP2) on the phloem sieve plates interferes with the transport of photoassimilates of source leaves to the sink organs (Koh et al., 2012; Boava et al., 2017; Wang et al., 2017), resulting in excessive starch accumulation in leaf chloroplasts (Wang and Trivedi, 2013; Boava et al., 2017). Starch accumulation causes the disintegration of the chloroplast thylakoid system, producing the yellowing leaf mottle symptom (Schneider, 1968; Etxeberria et al., 2009). Consequently, other typical HLB symptoms occur, such as yellow shoots, hardened and small leaves, leaves showing zinc deficiency and corky veins, twig dieback, stunted growth, and tree decline (Bové, 2006; Wang and Trivedi, 2013).

Thus far, no source of resistance to HLB is known. However, the relative *Citrus* species *Poncirus trifoliata* does not present typical HLB symptoms, and multiplication of CLas remains low or nonexistent (Folimonova et al., 2009; Albrecht et al., 2012; Boava et al., 2015, 2017). Additionally, it is an important rootstock for the Citriculture because of its tolerance/resistance to *Phytophthora*, citrus tristeza virus and nematodes (Pang et al., 2007). Due to these characteristics, *P. trifoliata* and its hybrids have been highlighted as a possible source of tolerance/resistance to HLB. The hybrid population between *P. trifoliata* and *Citrus sunki* showed variability in response to CLas infection. Some hybrids were considered susceptible (CLas-positive and significant difference in starch levels), tolerant (CLas-positive, but no

significant difference in starch levels) and resistant (CLas-negative and no difference in starch levels) (Boava et al., 2017).

We mapped the genomic regions associated with the expression analyses (eQTL) of *Citrus* callose synthase genes (*Cscals*) in the linkage groups of *C. sunki* and *P. trifoliata* genetic maps. Callose synthase genes encode the enzymes callose synthases (*CalS*), which are key elements for callose synthesis in different plant locations (Verma and Hong, 2001). In *Arabidopsis thaliana* (*At*), 12 *calS* genes were identified and designated as *calS1–calS12* (Chen and Kim, 2009). In the *Citrus* genome, nine putative callose synthase (*calS*) genes could be found based on their amino-acid and DNA sequence similarities to *AtcalS* and they were named *Cscals2*, *Cscals3*, *Cscals5*, *Cscals7*, *Cscals8*, *Cscals9*, *Cscals10*, *Cscals11* and *Cscals12* (Granato et al., 2019).

Each *CalS* has a tissue-specific function (Ellinger and Voigt, 2014), and most are required for callose biosynthesis during pollen development (Jacobs et al., 2003; Enns et al., 2005; Töller et al., 2008). However, some callose synthases play important roles in response to pathogen infection (Dong et al., 2008; Luna et al., 2011). Particularly, *CalS7* has been demonstrated to be responsible for the synthesis of callose in sieve plates in *Arabidopsis* (Barratt et al., 2011; Wang et al., 2011).

Expression quantitative trait loci (eQTL) studies involve a direct association between genomic locations with gene expression levels (Nica and Emmanouil, 2013). eQTL evaluations using the *C. sunki* and *P. trifoliata* hybrids can be very important to understand the mechanisms involved in the development of HLB symptoms. Some regions associated with *Cscals* expression and, consequently, with callose deposition identified in this study can be considered potential targets for future citrus breeding programs aiming to obtain tolerance to HLB.

3.1.3. Materials and Methods

3.1.3.1. Plant material

The mapping population comprised 272 F₁ hybrids resulting from crosses between *C. sunki* ex Tan (female parent) and *P. trifoliata* Raf. cv. Rubidoux (male parent). All the plants were propagated using buds grafted onto six-month-old Rangpur lime (*C. limonia* Osbeck). After six months, the plant scions were grafted on the opposite side of the primary stem, with two CLas-infected budwoods obtained from *C. sinensis* (L.) Osbeck cv. Pera plants, the identification of which was confirmed by qPCR. Infected budwoods were left on the plants, but shoots from these budwoods were eliminated upon sprouting. All the plants were kept in a greenhouse at

Centro de Citricultura Sylvio Moreira of the Instituto Agronomico (IAC), Cordeiropolis/SP at an average temperature of 25 °C. The experiment comprised three biological replicates for each inoculated (CLas-infected budwood) and mock-inoculated (healthy budwood) genotypes.

For the gene expression assay and eQTL mapping, the leaves were collected from parental plants (*C. sunki* and *P. trifoliata*) and 72 hybrids from the F₁ population, randomly selected, at 24 months after CLas inoculation.

3.1.3.2. DNA extraction and molecular marker analysis

The leaves of 272 hybrids and the parental plants were collected at a similar age from four sides of the plants for DNA extraction. Five leaves were combined, and 200-mg subsamples were lysed by grinding with two beads (3-mm diameter) in 2-mL microtubes at 30 Hz for 120 s in a TissueLyser II (Qiagen). DNA extraction was performed using the CTAB method (Murray and Thompson, 1980), and DNA quality and concentration were checked using a NanoDrop™ 8000 spectrophotometer (Thermo Scientific, Waltham, Massachusetts, USA).

The hybrid population and parental plants were genotyped using SNP (single-nucleotide polymorphism) markers. The method used to obtain the molecular markers for *Citrus* using the DArT-seq platform was previously reported (Curtolo et al., 2017). Briefly, all the samples (272 hybrids and parents) were genotyped using *Pst*I and *Taq*I digestion and were sequenced on a HiSeq2000 DArT-seq device (Illumina Inc., San Diego, California, USA) at Diversity Arrays Technology Ltd. (DArT P/L, Canberra, Australia). The resulting sequences were aligned to the Clementine tangerine reference genome (<https://phytozome.jgi.doe.gov/pz/portal.html>). The DArT-seq technology detects both SNPs (Single Nucleotide Polymorphisms) and DArT-seq markers, which are based only on presence–absence (Raman et al., 2014). The molecular markers were represented in a dataset matrix where columns were the genotypes and rows were the markers. Parameters for quality control such as the call rate and reproducibility over 90% were adopted to select SNP markers for genetic mapping construction.

3.1.3.3. Linkage Maps

The linkage maps were obtained as previously described by Curtolo et al. (2018). All SNP loci that showed no deviation from the expected segregation were included in the analysis. The SNP molecular markers were coded according to Wu et al. (2002) in OneMap software (Margarido et al., 2007). Because this technology provides biallelic markers, three possible segregation patterns were expected: marker segregation for only the female parent (*C. sunki*) [ab × aa]; only for the male parent (*P. trifoliata*) [aa × ab]; and for both parents simultaneously [ab × ab]. The

maps were constructed considering an LOD score = 8, and the maximum recombination fraction of 0.3. All the markers were aligned using BLASTn (Basic Local Alignment Search Tool) to the *C. sinensis* genome (<https://citrus.hzau.edu.cn/>) to establish the linkage groups because its assembly is based on pseudochromosomes while the Clementine genome is still based on scaffolds.

3.1.3.4. RNA extraction and cDNA synthesis

We sampled the leaves from 72 hybrids and parent plants (*C. sunki* and *P. trifoliata*) both CLAs and mock-inoculated (healthy plants). Leaves at a similar age were collected from four sides of the plants for RNA extraction. The samples were ground with liquid nitrogen, resulting in three microtubes with 100 mg for each genotype, consisting of three biological replicates per condition per genotype. Total RNA was extracted with lithium chloride (LiCl) using the protocol described by Chang et al. (1993) and adapted by Porto et al. (2010). The genomic DNA was eliminated using a DNase I, RNase-Free kit (Thermo Scientific, Waltham, Massachusetts, USA), according to the manufacturer's recommendations, followed by purification with phenol-chloroform and ethanol precipitation. RNA quality was verified by agarose gel electrophoresis, and the RNA concentration was determined using a NanoDrop™ ND-8000 spectrophotometer (Thermo Scientific, Waltham, Massachusetts, USA). cDNAs were synthesized from 1.0 µg of total RNA using Superscript III (200 U /µl) (Invitrogen, Carlsbad, California, USA) and oligo (dT) primers (dT12-18; Invitrogen) according to the manufacturer's instructions. The obtained cDNA from the biological replicates was diluted in RNase-free water at the ratio of 1:50 and mixed, forming a pool of samples for each genotype to be analyzed in gene expression and eQTL mapping assays.

3.1.3.5. Real-time Quantitative PCR (RT-qPCR)

The cDNA pool from each genotype was diluted in RNase-free water at the proportion of 1:25. The reaction comprised 6.0 µL of GoTaq qPCR Master Mix (Promega, São Paulo, Brazil), 2 µL of cDNA, 200 nM of each primer and water to a final volume of 10 µL. Amplifications were carried out using two replicates for each sample with appropriate negative controls in the 7500 Fast Real-Time PCR System (Applied Biosystems, Foster City, California, USA) thermal cycler with the following conditions: 50 °C for 2 min; 95 °C for 10 min; 40 cycles of 95 °C for 15 s and 60 °C for 1 min.

The *Cscals* primers were based on Granato et al. (2019), and the endogenous controls (FBOX and GAPC2) were based on Mafra et al. (2012) (Table S1). The primer specificities were

checked by melting curve analysis. Amplicons were sequenced using an ABI 3730 sequencer (Applied Biosystems, Foster City, CA, USA) and DyeTerminator chemistry to confirm their identities.

The amplification efficiency values (E) and Ct data were calculated for each RT-qPCR reaction using Real-time PCR Miner software (<http://ewindup.info/miner/>). The mean of the Ct values of the two technical replicates of each genotype was considered. Using these data, the relative quantification (fold change) was calculated using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). The fold change was calculated using CLas-inoculated plants compared with the respective mock-inoculated plants with FBOX and GAPC2 as reference genes.

During RT-qPCR, 74 genotypes (72 hybrids, *C. sunki* and *P. trifoliata*) were separated in four plates (incomplete blocks). In each one, 18 genotypes and the parents were evaluated under mock-inoculated (healthy plants) and CLas-inoculated conditions. The experimental design used to evaluate the samples was an incomplete block design. The model used was as follows: $Y_{ij} = \mu + B_j + G_i + e_{ij}$, where Y_{ij} corresponds to the gene expression of the i -th genotype evaluated in the j -th plate, μ is the model intercept, B_j is the fixed effect for plates, in which j varies from 1 to 4, G_i is the random effect of genotypes, in which i ranges from 1 to 74 and the genotypes 73 and 74 correspond to parents repeated along the four plates, and e_{ij} is the random residual effect. The function LME from package NLME of R software was used to analyze the mixed model and estimate the variance components.

3.1.3.6. Gene expression profile and genetic parameter analyses

Fold-change values adjusted by the mixed model were used as inputs to the MeV (MultiExperiment Viewer) program v. 4.9 (<http://sourceforge.net/projects/mev-tm4/>) to evaluate the gene expression profile. Evaluations were performed comparing the *Cscals* gene expression between the 72 hybrids and two parents (*C. sunki* and *P. trifoliata*) that were CLas inoculated and mock inoculated. The sets of genotypes with gene expression similarity were clustered using the hierarchical clustering method (HCL) and the Pearson correlation as the metric distance. The obtained values were graphically represented as a heatmap.

3.1.3.7. eQTL mapping

The genetic linkage maps obtained for *C. sunki* and *P. trifoliata* were used for eQTL mapping. Relative gene expression values were analyzed using the composite interval mapping (CIM) strategy (Zeng, 1994), adapted to a single fullsib cross and implemented in the *FullsibQTL* package (Gazaffi et al., 2014) of the R software. Cofactor selection was performed using

multiple linear regression analysis with a stepwise approach based on AIC (Akaike Information Criterion), similar to that performed by Souza et al. (2013) and Curtolo et al. (2018). The maximum number of selected cofactors was 20 with a window size of 1000 cM. The permutation test (Churchill and Doerge, 1994) was performed with 1000 replicates ($P < 0.05$) to obtain the threshold (LOD score) to declare eQTL. However, the modification proposed by Chen and Storey (2006) was used. All genetic markers flanking an eQTL interval for *Cscals* were aligned with the Citrus reference genome (<http://citrus.hzau.edu.cn/orange/>) to check the presence of cis/trans eQTL using the BLASTn tool (<https://blast.ncbi.nlm.nih.gov>).

3.1.4. Results

3.1.4.1. *C. sunki* and *P. trifoliata* linkage maps

The linkage maps constructed were generated by SNP markers using 272 F₁ hybrids from crosses between *C. sunki* and *P. trifoliata*. The F₁ hybrids sampled were genotyped using 17,482 SNP markers, but 16,337 were excluded. The exclusion criteria for SNP markers were as follows: 2,437 SNP markers had a call rate < 90 (percentage of successfully scored individuals for an allele); 1,338 SNP markers showed distorted segregation; 6,914 SNP markers were homozygous for both parents; and 455 and 5,193 SNP markers were missing calls for *C. sunki* and *P. trifoliata*, respectively. The distribution of SNP markers before and after the exclusion is observed in Figures 1 and 2.

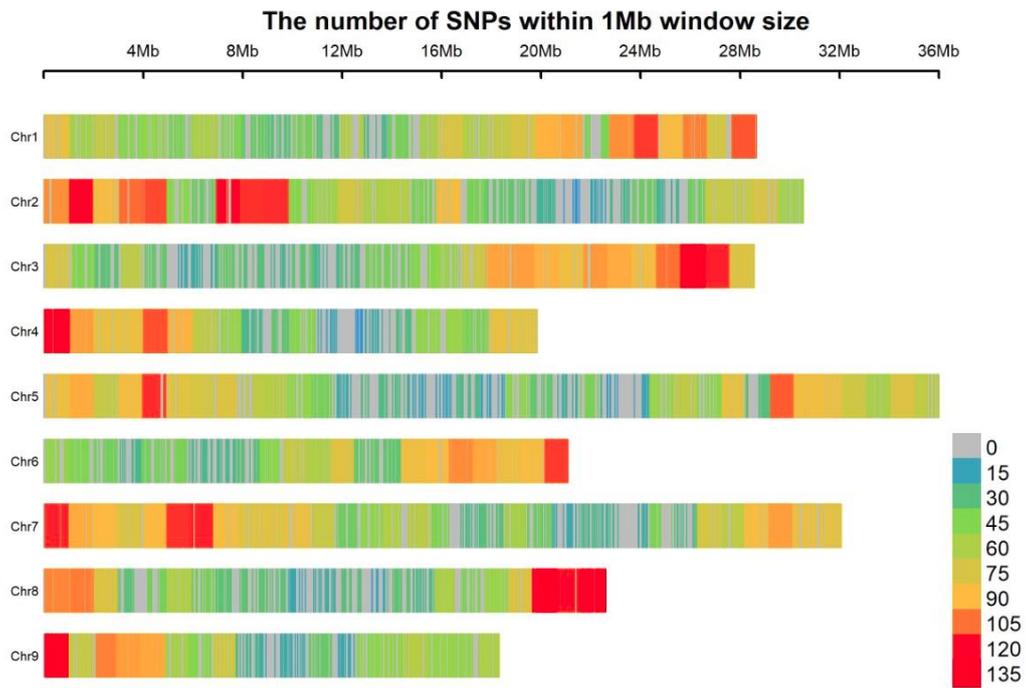


Figure 1. Density of markers in the chromosomes considering all markers resulting from the technology of SNP from DArT-seq

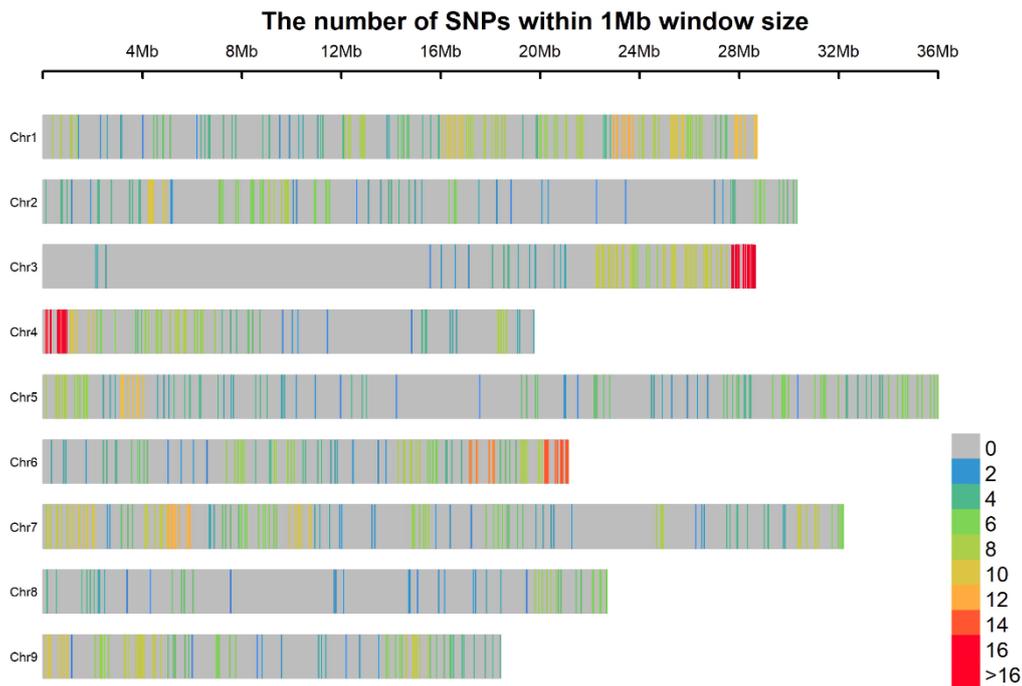


Figure 2. Density of markers in the chromosomes after considering a call rate < 90 , missing calls in the parent genotyping for *C. sunki* and *P. trifoliata* and distortion segregation.

Regarding the remaining 1,145 SNP markers that showed a segregation ratio of 1:1, 571 SNP markers were polymorphic for the parent *C. sunki* and 574 for *P. trifoliata*. Initially, only 109 markers were common and polymorphic for both parents. On the other hand, these markers showed segregation deviation and therefore they were excluded. This fact resulted in an impossible integration of the linkage groups of both maps. The original approach proposed by Wu et al. (2002) results in a single integrated genetic map modeling the linkage phases between markers. We applied this methodology but analyzed as two separated data sets derived for each parent, similar to the pseudo-testcross strategy (Grattapaglia and Sederoff, 1994) and resulting in two separated maps. The *C. sunki* linkage map exhibited 571 loci and genomic coverage of 2,855 cM, distributed in nine linkage groups (LG) (Figure 3). The groups ranged from 63.68 (LG8) to 530.91 (LG5) cM. LG3 had the highest density of markers (4.21 cM between markers), and LG4 had the lowest density of markers (6.48 cM between markers).

The *P. trifoliata* linkage map was constructed using 568 markers, and it had a genomic coverage of 3,334.1 cM, distributed in nine linkage groups (Table 1 and Figure 4). Only six SNP markers

were not positioned on the map. Some linkage groups (LG1, LG5 and LG6) exhibited some large gaps. To avoid an overestimation of genomic coverage, we divided the linkage groups in subgroups adding the letters “a” and “b”. Based on the genomic information, the linkage groups were identified as LG1 to LG9 and ranged from 143.55 (LG5b) to 439.51 (LG4) cM. LG6a had the highest density of markers (5.06 cM between markers), and LG5a had the lowest density of markers (7.07 cM between markers). However, the molecular markers were compared with the genomic information, and some further information could be obtained (Table 2) e.g., 87 molecular markers were assigned to LG1 for the *C. sunki* map, among which 71 were correctly aligned with chromosome one, 13 were referred with an unassigned chromosome, and two markers were not aligned with a reference genome. Only one marker was wrongly assigned with other linkage groups, but the genomic information was assigned as chromosome one.

Table 1. Distribution of mapped SNP marker numbers and sizes (cM) for each linkage group in the *C. sunki* and *P. trifoliata* linkage maps.

Linkage map <i>C. sunki</i>			Linkage map <i>P. trifoliata</i>		
	Number of markers	Size (cM)		Number of markers	Size (cM)
LG 1	87	398.78	LG 1a	57	291.84
			LG1b	42	238.78
LG 2	73	348.65	LG 2	49	269.44
LG 3	44	185.48	LG 3	46	246.80
LG 4	48	311.13	LG 4	72	439.51
LG 5	113	530.91	LG 5a	47	332.32
			LG 5b	23	143.55
LG 6	61	293.06	LG 6a	31	156.96
			LG 6b	30	153.72
LG 7	73	358.47	LG 7	63	399.75
LG 8	11	63.68	LG 8	46	304.76
LG 9	61	364.84	LG 9	62	356.67
Total	571	2855	Total	568	3334.1

Table 2. Number of markers not aligned to the reference genome, aligned on the unassigned chromosome (UnChr), in another chromosome (X) or in the corresponding chromosome (Chr).

<i>C. sunki</i>					<i>P. trifoliata</i>				
Linkage Groups	NotAlig	UnChr	X	Chr	Linkage Groups	NotAlig	UnChr	X	Chr
LG1	2	13	1	71	LG1a	1	13	1	42
					LG1b	0	4	0	38
LG2	1	14	2	56	LG2	0	8	3	38
LG3	0	0	0	44	LG3	0	2	2	42
LG4	1	3	8	36	LG4	0	4	14	54
LG5	0	30	5	78	LG5a	0	16	1	30
					LG5b	0	13	3	7
LG6	0	10	1	50	LG6a	0	5	3	23
					LG6b	0	0	0	30
LG7	1	3	0	69	LG7	0	8	0	55
LG8	0	0	0	11	LG8	0	5	7	34
LG9	1	15	4	41	LG9	0	15	6	41
Total	6	88	21	456	Total	1	93	40	434

* NotAlig represents all sequences that were not aligned to the reference genome; UnChr (unassigned chromosome) is a segment of the genome where none of the sequences are placed in pseudochromosomes; X represents all markers that were positioned in another chromosome which is not the one of the correspondences; Chr represents all markers that were aligned into corresponding chromosome.

A general view indicated that 456 (80%) of the markers from the *C. sunki* map and 434 (76%) of the markers from the *P. trifoliata* map were correctly grouped. Additionally, 88 (*C. sunki*) and 93 (*P. trifoliata*) molecular markers were assigned to an anonymous group (unassigned chromosome) in the reference genome i.e., they do not match any chromosome but the linkage approach provides extra information assigning along the genetic map. Only six markers of *C. sunki* and one marker of *P. trifoliata* were not assigned to the reference genome. Twenty-one markers of *C. sunki* and 40 markers of *P. trifoliata* were considered linked with groups that do not match genomic positions. In this case, the genomic position prevails to assign the markers to a specific group. Differences between genomic and map positions of markers may have resulted from false positives due to the multiple tests performed.

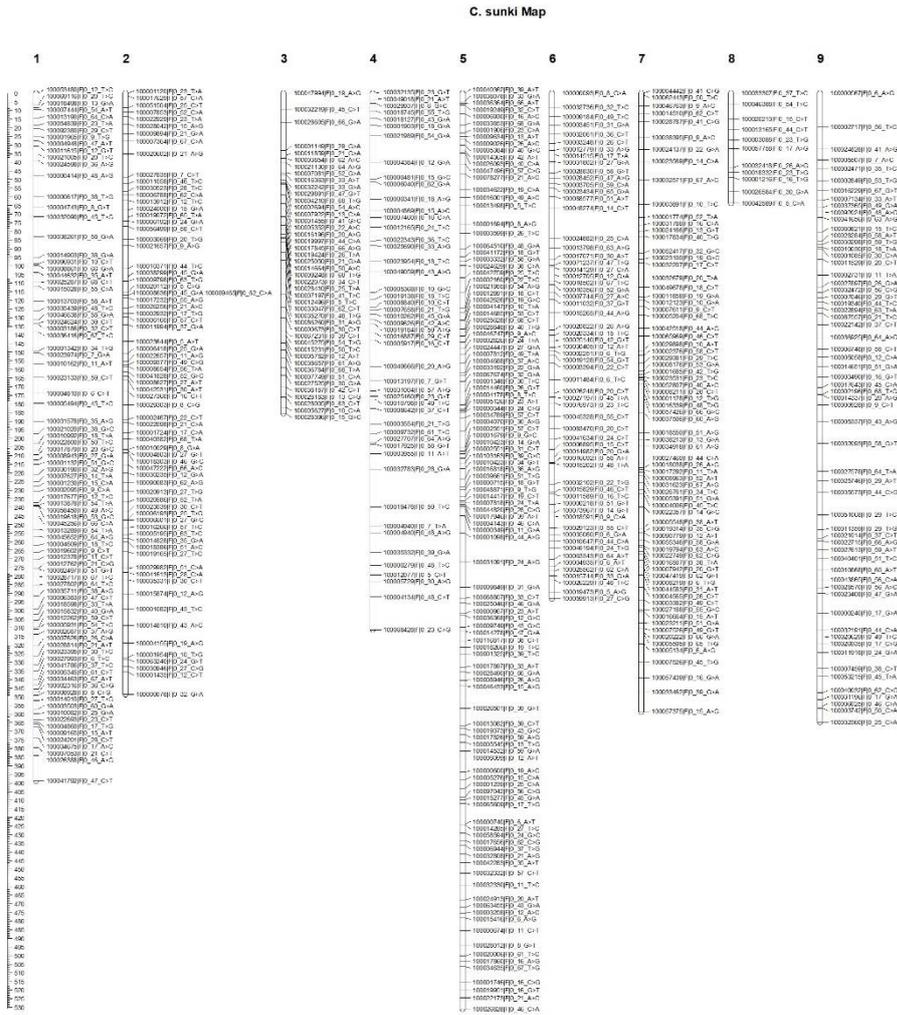


Figure 3. Linkage map of the *C. sunki* using the pseudo-testcross strategy. Distribution of the 571 SNP markers on nine linkage groups of the *C. sunki* linkage map. X-axis represents linkage groups, and Y-axis indicates the genetic location (cM).

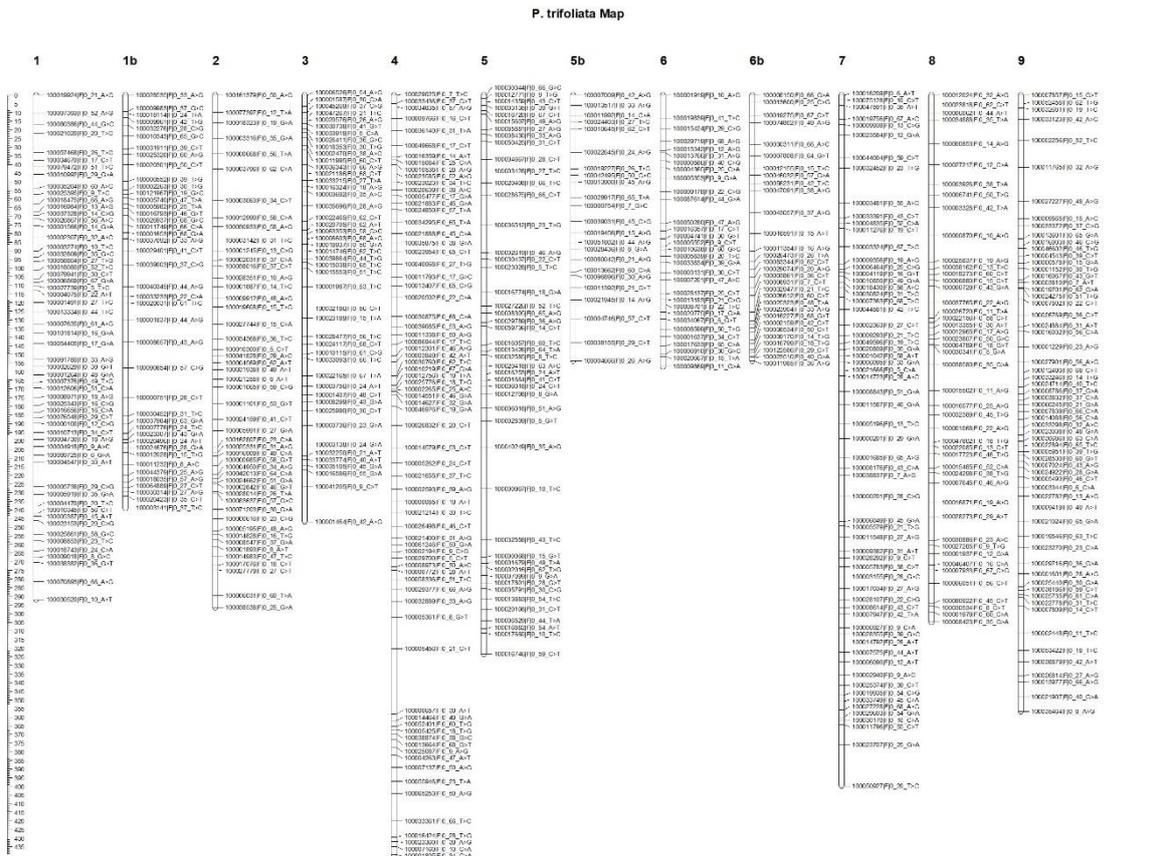


Figure 4. Linkage map of the *P. trifoliata* using the pseudo-testcross strategy. Distribution of the 568 SNP markers on the nine linkage groups of the *P. trifoliata* linkage map. X-axis represents linkage groups, and Y-axis indicates the genetic location (cM).

3.1.4.2. Gene expression profile

According to the heatmap (Figure 5), the parental *C. sunki* and 43% of hybrids plants showed a predominantly green overall expression pattern, indicating that genotypes 132, 130, 141, 146, 19, 99, 124, 166, 293, 163, 149, 187, 119, 134, 107, 109, 148, 217, 121, 70, 279, 143, 137, 31, 4, 129, 73, 136, 68, 49, 173, and the parental *C. sunki* showed upregulation of *CscalS* gene expression compared with the CLAs-infected plants and healthy controls. On the other hand, most of the genotypes (57%) i.e., hybrids 56, 126, 94, 24, 78, 125, 179, 154, 189, 111, 102, 26, 151, 101, 86, 66, 61, 23, 191, 54, 183, 90, 20, 42, 2, 96, 117, 150, 47, 14, 10, 35, 113, 16, 28, 110, 142, 1, 118, 184, 105, and the parental *P. trifoliata* exhibited downregulation in the expression of *CscalS* genes compared with that in the CLAs-infected plants and healthy controls.

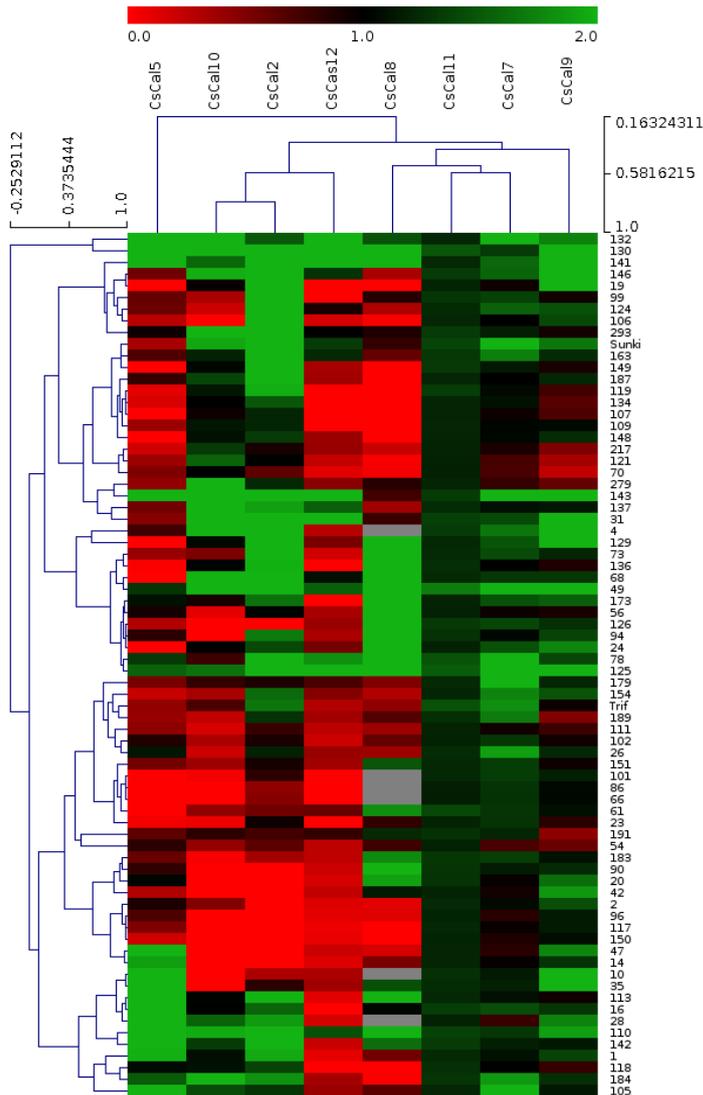


Figure 5. Heatmap of the gene expression profile by clustering analysis between the eight *Cscals* genes evaluated using the 74 genotypes (72 hybrids and the parent plants *P. trifoliata* and *C. sunki*). The heatmap was made using fold-change normalized data as inputs to the MeV (MultiExperiment Viewer) program v. 4.9 (<http://sourceforge.net/projects/mev-tm4/>). The names of genes and gene hierarchical clusters are shown at the top. Fold-change expression values ranged from green (highest expression) to red (lowest expression). The sample names (74 genotypes) are shown on the right side, while the sample hierarchical cluster is shown on the left side.

In the same analysis, the parental *P. trifoliata* showed upregulated expression of *Cscals2* and *Cscals7*, while *Cscals11* and the parental *C. sunki* displayed upregulated expression of *Cscals2*, *Cscals7*, *Cscals9*, *Cscals10*, *Cscals11* and *Cscals12*. Regarding the hybrids, it is possible to observe that regulation of the analyzed *Cscals* genes was very different among them. The expression of *Cscals2* and *Cscals7* was upregulated in most genotypes, including the

parental *C. sunki* and *P. trifoliata*. *Cscals9* and *Cscals10* also demonstrated upregulation in 53 genotypes. *Cscals5* and *Cscals12* were revealed to be largely downregulated in the genotypes. The expression of *Cscals11* presented upregulation in all the genotypes analyzed, and *Cscals8* was upregulated in 27 genotypes.

The heatmap (Figure 5), based on the comparative analysis performed by hierarchical clustering (HCL) of *Cscals* genes and the 72 hybrids plus their two parents (*C. sunki* and *P. trifoliata*) allowed the grouping of genes and related genotypes. Additionally, Pearson's correlation was used as a metric distance to obtain the best intra and intervariable grouping possible. The genotypes were separated into eight subgroups distributed into three main clusters. The parent *P. trifoliata* was internally clustered with the genotypes 154 and 189, while the parent *C. sunki* was clustered together with the genotypes 163 and 149. Both parent clusters were grouped with the remaining genotypes to form a larger main cluster.

The genes were separated into three clusters. The first cluster was formed by *Cscals2*, *Cscals10* and *Cscals12*, the second cluster was formed by *Cscals7*, *Cscals8*, *Cscals9* and *Cscals11*, and a third one was formed only by *Cscals5*.

The adjusted values of the *CsCals* relative gene expression from the F₁ hybrids were used to calculate the genetic parameters (heritability, variance, and coefficient of variation). The genotypic variance (V_g) ranged from 0.11 to 40.81, expressed as the genotypic variation coefficient (CV_g) that varied from 26.11 to 369.23% (Table 3). Phenotypic variance (V_f) estimates varied from 1.37 to 41.22, and the highest values were obtained for the genes *Cscals8* (41.22) and *Cscals12* (15.95). High values of heritability (h²) for the studied callose synthase genes were observed, with the exception of *Cscals11* (6.00), indicating that, for this gene, the genotypic variance was proportionally lower than the environmental variance.

Table 3. Estimates of genotypic and phenotypic variances, heritability and coefficients of variation for gene expression.

Genes	V _g	V _f	h ² (%)	CV _r (%)	CV _g (%)
<i>Cscals2</i>	7.94	8.44	94.07	33.11	137.04
<i>Cscals5</i>	11.33	11.83	95.77	44.75	213.03
<i>Cscals7</i>	0.80	1.55	51.61	61.48	64.81
<i>Cscals8</i>	15.48	15.95	97.05	32.18	184.71
<i>Cscals9</i>	1.23	1.37	89.78	24.94	73.93
<i>Cscals10</i>	40.81	41.22	99.00	37.01	369.26

<i>CscalS11</i>	0.11	1.69	6.00	98.97	26.11
<i>CscalS12</i>	1.42	1.62	87.65	72.13	192.19

V_g = genotypic variance; V_f = phenotypic variance; h² = heritability; CV_r = coefficient of variation of the residue; CV_g = coefficient of variation of the genotype.

3.1.4.3. eQTL mapping

It was possible to detect eQTL in response to infection caused by CLAs using the *C. sunki* and *P. trifoliata* linkage maps and gene expression profiles from the relative expression values (fold change) of *CscalS* genes evaluated in the 72 hybrids.

Considering the *CscalS* expression profile, 18 eQTL were mapped in the *C. sunki* linkage map, and the LOD scores of the eQTL ranged from 3.22 to 17.87 (Figure 6 and Table 4).

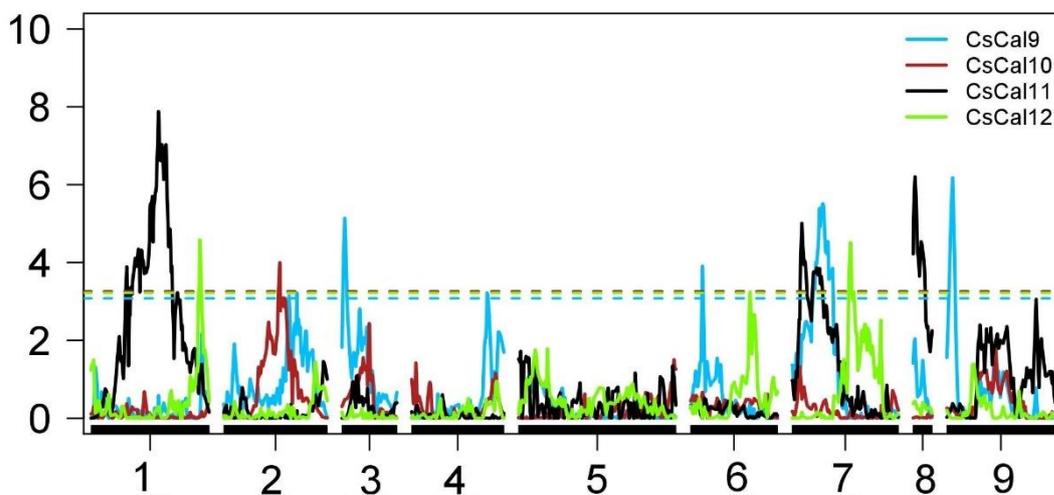
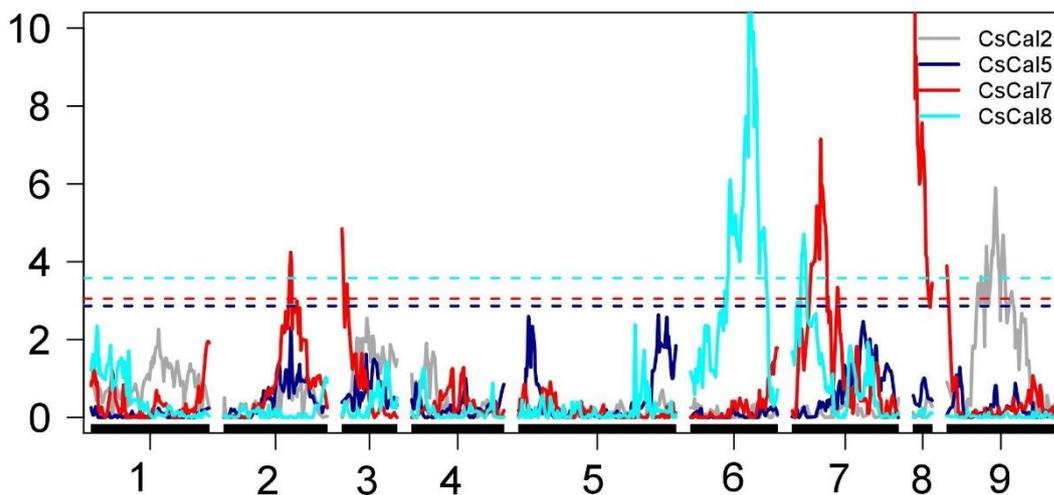


Figure 6. Detection of eQTL in the *C. sunki* linkage map related to the expression of the *Cscals* genes evaluated. Y-axis: LOD; X-axis: distance in centiMorgans; the dashed lines represent threshold values obtained using 1000 replicates.

Table 4. eQTL mapping for *Cscals2*, *Cscals7*, *Cscals8*, *Cscals9*, *Cscals10*, *Cscals12* in *C. sunki* linkage map

Genes	SNP Markers	Genome position	LG	cM	Lod-Score	Additive Effect	R ²
<i>Cscals2</i>	100003490 F 0_16_G>T	ChrUn,1142507	9	164.32	5.92	0.78	12.31
* <i>Cscals7</i>	100090083 F 0_62_A>G	Chr2,7755160	2	225.47	4.25	-1.01	0.82
* <i>Cscals7</i>	100047994 F 0_19_A>G	Chr3,19075229	3	0.00	5.06	1.79	7.42
* <i>Cscals7</i>	100023100 F 0_19_G>C	N/D	7	96.17	7.19	2.08	17.99
<i>Cscals7</i>	100033307 F 0_37_T>C	Chr8,19898080	8	0.00	17.87	3.05	20.18
* <i>Cscals7</i>	100000567 F 0_6_A>G	Chr9,17314839	9	0.00	3.90	-1.16	6.71
<i>Cscals8</i>	100041634 F 0_24_C>T- 100006895 F 0_15_C>T	Chr6,15796184- 15817077	6	203.00	11.50	-0.30	10.91
<i>Cscals8</i>	100023569 F 0_14_C>A	Chr7,1786000	7	39.42	4.71	0.21	5.29
* <i>Cscals9</i>	100006193 F 0_25_T>G	Chr2,7224068	2	246.57	3.22	-0.33	7.11
* <i>Cscals9</i>	100032219 F 0_45_C>T	Chr3,19755543	3	9.20	5.17	0.36	3.34
<i>Cscals9</i>	100004940 F 0_48_A>G	Chr7,3129395	4	254.71	3.25	-0.27	3.31
<i>Cscals9</i>	100031802 F 0_27_G>A	Chr6,5552031	6	39.72	3.91	-0.30	1.23
* <i>Cscals9</i>	100032207 F 0_17_C>T- 100032679 F 0_20_T>A	Chr7,6721626- 7216583	7	103.00	5.51	0.38	6.04
* <i>Cscals9</i>	100002717 F 0_56_T>C	ChrUn,50210454	9	19.40	6.18	-0.39	9.35
<i>Cscals10</i>	100002467 F 0_22_C>T	Chr2,13556907	2	189.02	4.01	-0.52	0.49
<i>Cscals12</i>	100001230 F 0_15_C>A	Chr1,16786655	1	367.38	4.59	0.42	7.57
<i>Cscals12</i>	100024137 F 0_22_G>A	Chr7,1434034	6	200.00	3.26	-0.27	11.46
<i>Cscals12</i>	100046388 F 0_54_T>C	Chr8,20056662	7	196.59	4.51	-0.43	11.43

SNP markers = flanking markers; LG = Linkage Group; cM = position; R² = explained phenotypic variation; * = hot spot

All eQTL detected showed a 1:1 segregation pattern, and they were mapped in all linkage groups, except LG5. One eQTL was detected for *Cscals2* on LG9; five eQTL for *Cscals7* were detected on LG2, LG3, LG7, LG8 and LG9; two eQTL for *Cscals8* were detected on LG6 and LG7; six eQTL for *Cscals9* were detected on LG2, LG3, LG4, LG6, LG7 and LG9; one eQTL for *Cscals10* was detected on LG2; and three eQTL for *Cscals12* were detected on LG1, LG6 and LG7. It was not possible to detect eQTL for *Cscals5* and *Cscals11*. The phenotypic variance values (R²) explained by the eQTL mapped varied from 0.49% to 20.18%. The eQTL detected for *Cscals7* on LG8 exhibited the highest R² using the *C. sunki* map (20.18%). Together, the five eQTL for *Cscals7* explained 53.12% of the phenotypic variation; thus, *Cscals7* had the highest percentage of the phenotypic variation explained by the eQTL

mapping. The highest number of eQTL was detected for *Cscals9* (six eQTL), and, overall, they represented 30.38% of the phenotypic variation. The three eQTL were identified for *Cscals12*, explaining 30.46% of the phenotypic variation.

The colocalization of eQTL may suggest the existence of hot spots. eQTL for *Cscals7* and *Cscals9* could be observed on LG2, LG3, LG7, and LG9 separated by 21.00, 9.20, 6.83, and 19.40 cM, respectively. Considering the 18 eQTL identified in the *C. sunki* map, eight were clustered in four different hot spots.

In the *P. trifoliata* linkage map, it was possible to map 34 eQTL (Figure 7 and Table 5): eight eQTL for *Cscals2* were distributed on LG2, LG4, LG5, LG6, LG7, and LG8; seven eQTL for *Cscals5* were distributed on LG1b, LG2, LG5, LG7, LG9; seven eQTL for *Cscals7* were distributed on LG2, LG4, LG5, LG8, LG9; two eQTL for *Cscals8* were distributed on LG4 and LG8; five eQTL for *Cscals9* were distributed on the LG1, LG1b, LG2, LG5b, LG7; and five eQTL for *Cscals12* were distributed on LG2, LG5, LG5b, LG7, LG8. No eQTL was identified for either *Cscals10* or *Cscals11*.

Overall, R^2 varied from 0.4 to 22.63%, the LOD score ranged from 3.21 to 9.56 and all segregated in a 1:1 fashion. Considering the eQTL mapping for *P. trifoliata*, eQTL for *Cscals7* had the highest R^2 (22.63%) and, when the seven eQTL were considered together, they summed the highest R^2 (55.61%). The region with the lowest R^2 was identified for *Cscals2*, explaining only 0.4% of the phenotypic variation.

Cscals2 had the highest number of regions detected in this study. Thirty-nine percent of the phenotypic variation were explained by the eight eQTL detected for *Cscals2*. Five other markers were associated with *Cscals8*, and, overall, they summed an R^2 of 39.62%. Two eQTL detected for *Cscals2* and *Cscals12* were overlapped. They were located on LG2 approximately 203-206 cM and further on two eQTL that were overlapped for *Cscals5* and *Cscals12* (230 cM). Another overlap eQTL for *Cscals5* and *Cscals9* was found on LG1b. The co-location of eQTL was detected for *Cscals2* and *Cscals12* on LG8, separated by 2.42 cM. Three overlap loci were identified between *Cscals2* and *Cscals7*: the first on LG4, the second separated by 14 cM on LG5 and the last on LG9 distant by 14 cM.

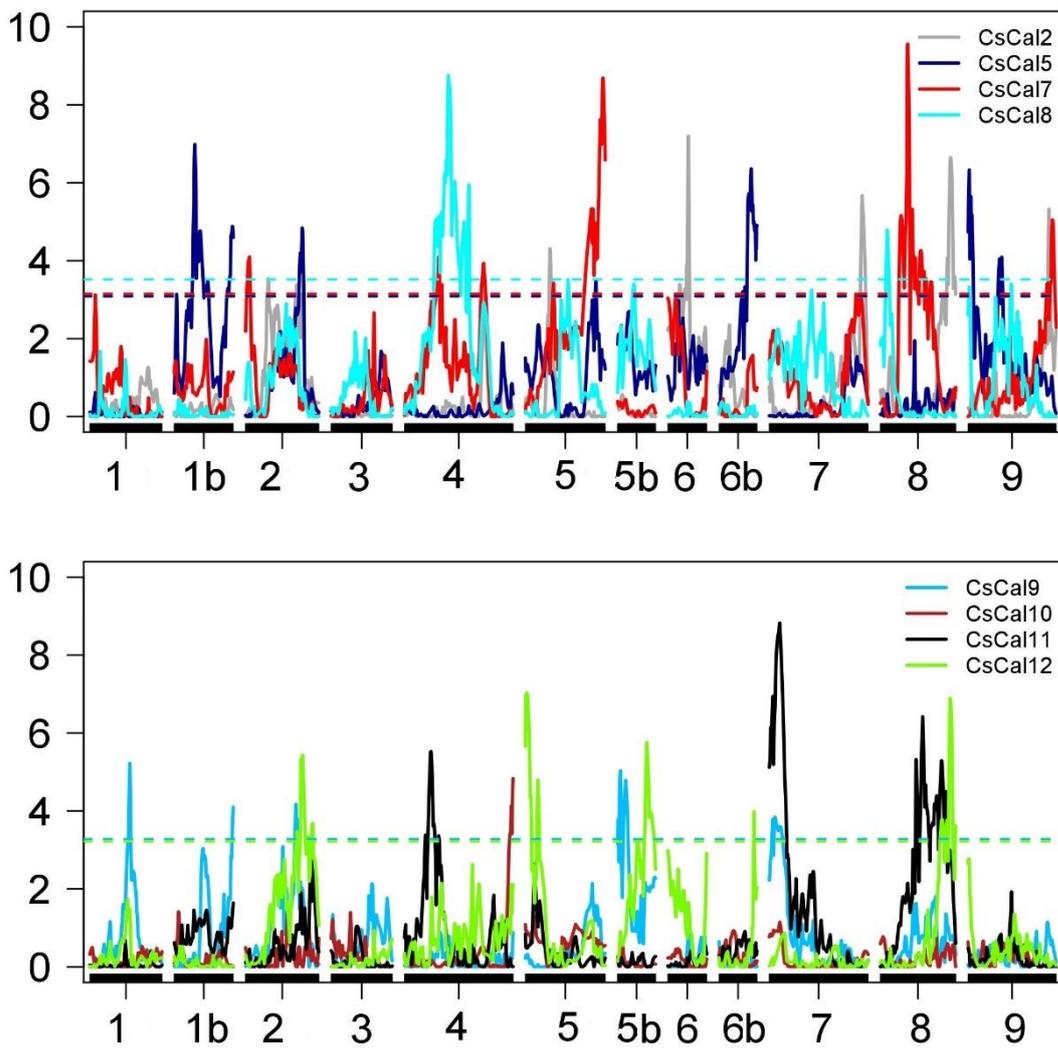


Figure 7. Detection of eQTL in the *P. trifoliata* linkage map related to the expression of the *CscaIS* genes evaluated. Y-axis: LOD; X-axis: distance in centiMorgans; the dashed lines represent threshold values obtained with 1000 replicates.

Table 5. eQTL mapped for *Cscals2*, *Cscals5*, *Cscals7*, *Cscals8*, *Cscals9*, *Cscals12* in *P. trifoliata* linkage map.

Genes	SNP Markers	Genome Position	LG	cM	Lod-score	Additive Effect	R ²
<i>Cscals2</i>	100001245 F 0_13_C>G-100002031 F 0_37_C>A	Chr2,11496268-12594168	2	92.00	3.54	0.70	3.68
* <i>Cscals2</i>	100025331 F 0_31_A>G	Chr2,9722118	2	206.42	3.40	0.69	2.13
* <i>Cscals2</i>	100005456 F 0_21_C>T	Chr7,11995806	4	320.42	3.20	-0.59	0.4
<i>Cscals2</i>	100023028 F 0_5_T>C	Chr5,6320268	5	99.56	4.30	0.81	8.32
<i>Cscals2</i>	100004741 F 0_30_G>T	Chr6,7357918	6	83.01	7.19	1.00	12.47
<i>Cscals2</i>	100023707 F 0_25_G>A	Chr7,1472171	7	375.76	5.67	-0.80	4.22
* <i>Cscals2</i>	100006051 F 0_56_C>T-100080922 F 0_45_C>T	Chr8,158039	8	284.00	6.64	-0.87	9.27
* <i>Cscals2</i>	100038879 F 0_42_A>T	Chr9,168999717	9	327.91	5.32	1.00	9.14
<i>Cscals5</i>	100037092 F 0_33_A>G	Chr1,24513911	1b	83.13	6.98	-0.89	2.29
* <i>Cscals5</i>	100020423 F 0_35_C>T-100003141 F 0_37_T>C	ChrUn,62887483-62915479	1b	235.00	4.88	-0.75	2.55
* <i>Cscals5</i>	100028014 F 0_26_T>A	Chr2,8399713	2	229.25	4.83	1.15	8.67
<i>Cscals5</i>	100005791 F 0_30_C>G	ChrUn,38031312	5	285.46	3.49	-0.79	3.33
<i>Cscals5</i>	100026612 F 0_60_C>T	Chr6,19905462	7	130.47	6.35	0.79	8.18
<i>Cscals5</i>	100052458 F 0_62_T>G	Chr9,752864	9	5.10	6.32	0.95	17.15
<i>Cscals5</i>	100016032 F 0_56_C>A	Chr9,7003215	9	136.41	4.15	0.78	3.40
<i>Cscals7</i>	100018323 F 0_19_G>A	ChrUn,32178022	2	15.96	4.09	-0.88	4.69
<i>Cscals7</i>	100011338 F 0_50_A>G	Chr4,6197839	4	138.66	4.09	-0.92	4.13
* <i>Cscals7</i>	100005456 F 0_21_C>T	Chr7,11995806	4	320.42	3.96	-0.90	3.24
* <i>Cscals7</i>	100016774 F 0_18_G>A	Chr5,7632775	5	114.08	3.43	-0.92	4.37
<i>Cscals7</i>	100017660 F 0_10_T>C-100016746 F 0_59_C>T	Chr5,27887080-29928945	5	314.00	8.68	-1.69	10.85
<i>Cscals7</i>	100000729 F 0_43_G>A	Chr6,13766295	8	110.29	9.55	-2.09	22.63
* <i>Cscals7</i>	100013977 F 0_66_A>G-100021907 F 0_40_G>A	Chr9,18067045	9	342.00	5.04	-1.09	5.7
<i>Cscals8</i>	100014627 F 0_32_G>A-100046976 F 0_19_G>A	Chr4,7777178	4	178.00	8.75	-0.26	8.88
<i>Cscals8</i>	100000853 F 0_14_A>G	ChrUn,88833722	8	27.94	4.78	-0.17	4.09
<i>Cscals9</i>	100001264 F 0_48_G>A	ChrUn,22371945	1	161.86	5.21	-0.36	7.64
* <i>Cscals9</i>	100003141 F 0_37_T>C	ChrUn,62915479	1b	238.77	4.09	0.29	4.87
* <i>Cscals9</i>	100162807 F 0_23_C>A	Chr2,9832235	2	203.43	4.17	0.24	0.8
<i>Cscals9</i>	100011992 F 0_14_C>A	ChrUn,4717070	5b	11.66	5.03	-0.30	4.83
<i>Cscals9</i>	100023584 F 0_12_G>A	Chr7,31022976	7	22.95	3.84	0.30	5.49
* <i>Cscals12</i>	100083637 F 0_57_G>C	Chr2,8444059	2	230.56	5.42	0.31	5.48
<i>Cscals12</i>	100003135 F 0_39_G>T	Chr5,8444059	5	5.83	7.03	-0.35	5.98
<i>Cscals12</i>	100021945 F 0_14_A>G	Chr5,33708464	5b	118.36	5.76	0.30	3.72
<i>Cscals12</i>	100002159 F 0_42_C>T	Chr6,21087431	7	141.72	3.99	-0.33	7.15
* <i>Cscals12</i>	100006051 F 0_56_C>T	Chr8,2038979	8	282.42	6.89	-0.33	7.6

SNP markers = flanking markers; LG = Linkage Group; cM = position; R^2 = explained phenotypic variation; * = hot spot

The existence of eQTL was noticed for the same *CsCalS* and LG in *C. sunki* and *P. trifoliata* maps. In both maps, eQTL were detected for *Cscals2* on LG9, *Cscals7* on LG2, *Cscals7* on LG8 and LG9, *Cscals9* on LG2 and LG7 and *Cscals12* on LG7. It is worth highlighting that the major eQTL identified in the *C. sunki* and *P. trifoliata* maps was positioned in the same linkage group (LG8).

Genomic information, such as the physical position, is not always accessible for *Cscals*; thus, inferring whether *cis* or *trans* eQTL exist becomes a challenge. Only the physical position is available for *Cscals2* (Chr 7), *Cscals5* (Chr 1), *Cscals7* (Chr 7), *Cscals8* (Chr 5) *Cscals10* (Chr 5), and *Cscals11* (Chr 2) (Granato et al., 2019). However, there is no eQTL close to the genes, suggesting the presence of epistatic eQTL or *trans* eQTL. In the cases of *Cscals9* and *Cscals12*, for which the physical locations are not described, an inference between *cis* and *trans* is not feasible.

3.1.5. Discussion

The hybrid population obtained from *C. sunki* and *P. trifoliata* crossing was genotyped using 17,482 SNP markers. However, the *C. sunki* and *P. trifoliata* genetic linkage maps were constructed using 571 and 568 representative SNP markers, respectively. Although a high number of SNP markers has been generated by genotyping using sequencing technology, many markers were excluded from the analysis due to the drawback of many lines being multiplexed during sequencing. Moreover, 1,338 SNP markers did not show the expected segregation. Deviations from the segregation can be the result of crosses among different genera (*Citrus* and *Poncirus*), as previously reported (Curtolo et al., 2018). The SNP marker exclusion resulted in a low number of polymorphic markers. We believe that monomorphic markers are often generated by technical and biological reasons. Genotyping technology with library construction, read depth, and data handling are possible causes of the presence of noninformative markers. Additionally, we should consider the limited population size as a possible explanation of monomorphic marker presence because the number of genotyped individuals determines the chance to detect recombinant loci. A large ratio of monomorphic markers has been reported as a disadvantage of high-throughput genotyping (Shimada et al., 2014; Guo et al., 2015; Yu et al., 2016a; Imai et al., 2017; Curtolo et al., 2017). It should be noted that the crossing between two parents from different genera contributes to few

polymorphic markers at the same time for both parents i.e., SNPs are not as old as that required for being shared by *C. sunki* and *P. trifoliata* because SNPs are conservative markers. This corroborates the idea that both parents are not genetically related and explains why two maps were obtained, one for each parent. Previously, Curtolo et al. (2018) used dominant markers such as DArTseq and obtained loci shared by *C. sunki* and *P. trifoliata*; however, the number of markers was not sufficient to enable information integration from both parents.

SNPs have been considered the most attractive markers to obtain genetic mapping, and they can be genotyped in parallel assays at low costs in marker-assisted breeding (Bertioli et al., 2014). There are six genetic maps for *Citrus* using SNP markers (Ollitrault et al., 2012; Xu et al., 2012; Guo et al., 2015; Yu et al., 2016a; Imai et al., 2017; Huang et al., 2018). However, this study is the first to demonstrate a linkage map for *Citrus* using SNP markers obtained from DArT-seq technology.

C. sunki and *P. trifoliata* linkage maps showed SNP markers distributed in nine linkage groups, corresponding to the haploid number of chromosomes of citrus. In both maps, few SNP markers were positioned in a different chromosome where most of the markers were located (Table 2). The difference in the marker position can be caused by the assembled difference between the species used in the reference genome and constructed linkage maps. The establishment of the marker position that has been grouped in the unassigned chromosome (UnChr) is a contribution of the present work. Furthermore, it could help update the *Citrus sinensis* genome, as previously reported by Curtolo et al. (2017). In the *P. trifoliata* map, some linkage groups were separated into “a” and “b” groups to avoid an overestimation of the genomic coverage. Nevertheless, the map and some groups of *P. trifoliata* are larger than those designed for *C. sunki*. Other authors also showed difference among the linkage group sizes (Chen et al., 2008; Huang et al., 2018). The recombination rate, which is used to obtain the maps, is distinct between females and males, both in plants and animals (Lorch, 2005). Ollitrault et al. (2012) and Huang et al. (2018) noticed that the size of male genetic maps is usually larger than that of female genetic maps. It corroborates the linkage maps obtained in this study because *C. sunki* was the female parent and *P. trifoliata* was the male parent of the crossing, generating the studied hybrid population. The presented linkage maps are a substantial resource for future studies of *Citrus*. The parents and hybrids used for the analyses revealed many important characteristics for citriculture. For example, both parents are important rootstocks, and *C. sunki* has high vigor and good fruit yield, as well as tolerance to Tristeza, citrus blight disease and salinity (Castle et al., 1993). *P. trifoliata* is immune to citrus tristeza virus and resistant to nematodes, although it has low tolerance to drought (Passos et al., 2006). *P. trifoliata* was also reported to be more tolerant to

HLB because it does not show starch accumulation in leaf chloroplasts and does not show typical HLB symptoms, unlike *C. sunki* (Boava et al., 2017).

The excessive accumulation of starch in *Citrus* leaves during CLAs infection has often been associated with photoassimilate transport disturbance (Koh et al., 2012; Boava et al., 2017; Wang et al., 2017). The reduction of photoassimilate transport of leaf sources to the sink organs results from deposition of callose and phloem proteins (PP2) in the phloem of infected plants (Koh et al., 2012; Wang and Trivedi, 2013; Boava et al., 2017). Callose is synthesized by the callose synthase enzymes (*CalS*), whose activity is highly regulated by pathogen infection (Yu et al., 2016b; Granato et al., 2019). In this study, the expression of all evaluated *Cscals* was regulated in CLAs-infected citrus leaves, demonstrating that multiple callose synthase genes can be expressed in the same organ (Dong et al., 2008; Granato et al., 2019). Most of the genotypes analyzed (57%), including the parental *P. trifoliata*, showed *Cscals* gene expression downregulation comparing the CLAs-infected plants and healthy controls. On the other hand, the parental *C. sunki* and 43% of the genotypes showed upregulation of *Cscals* gene expression after CLAs infection.

The *Cscals2* gene was upregulated in many genotypes, including the parental *C. sunki*. *CalS2* has not been characterized yet. However, in *Arabidopsis*, it shares high homology (92% identity) with *CalS1*, suggesting that a gene duplication event may have occurred, and it is possible that the two genes encoding both enzymes are functionally redundant (Hong et al., 2001). *Cscals2* upregulated expression in *C. sunki* and hybrids may indicate that this gene plays an important role in callose accumulation, as a strategy to alter plasmodesma permeability under CLAs infection because it occurs in *Arabidopsis* rosette leaves after salicylic acid (SA) and *Hyaloperonospora arabidopsis* infection (Cui and Lee, 2006; Dong et al., 2008).

Cscals7 has been demonstrated to be responsible for callose deposition specifically in the phloem sieve tubes (Barratt et al., 2011; Xie et al., 2011). *Cscals7* was upregulated in *P. trifoliata* in CLAs-infected plants. However, upregulation was lower than that observed for *C. sunki* (Table S2). The *Cscals7* gene was also upregulated in 49 other genotypes. The lower expression value of *P. trifoliata* can be due to its tolerance to HLB, or callose deposition in *P. trifoliata* does not cause hypertrophy of the phloem parenchyma cells and collapse of the sieve tube elements (STE) because it occurs in *C. sunki* (Folimonova et al., 2009; Koh et al., 2012). As previously shown for the HLB pathosystem (Granato et al., 2019) and grapevine-resistant cultivar *Vitis amurensis* ‘Shuanghong’ infected with *Plasmopara viticola* (Yu et al., 2016b), *calS7* upregulation after infection indicates that callose deposition specifically at phloem sieve

tubes occurs to block the flow of the pathogens, which probably occurred in *C. sunki*, *P. trifoliata* and their hybrids.

Other *Cscals* also presented upregulation in the analyzed genotypes, such as *Cscals9*, *Cscals10*, and *Cscals12*. CalS9 and CalS10 functions have been more related to gametophyte development (Töller et al., 2008) than the plant defense response. Nevertheless, the biological role of *calS12* has been well studied in the stress and pathogen response (Nishimura et al., 2003; Dong et al., 2008; Luna et al., 2011; Ellinger and Voigt, 2014). For example, *calS12* is required for callose deposition in cell wall thickenings at the sites of fungal pathogen attack during powdery mildew infection (Dong et al., 2008). Additionally, Granato et al., (2019) also demonstrated that, in *C. sinensis*, at 360 days after infection, *Cscals12* was significantly upregulated in HLB-positive plants. These results indicate that *Cscals12* is also likely involved in callose deposition after CLAs infection. Because all callose synthase genes showed regulation of expression after CLAs infection, it is possible that multiple *Cscals* work like a complex in the phloem sieve tubes, causing callose accumulation after pathogen attack (Granato et al., 2019).

Some genotypes studied in this work were classified by Boava et al. (2015) as tolerant or susceptible, based on the starch accumulation and titer of CLAs. Genotypes 19, 119, 124, 217 and *C. sunki* were previously classified as susceptible, and our results showed upregulation of *Cscals2*, *Cscals7* and *Cscals11* expression and downregulation of *Cscals5* and *Cscals8* expression after CLAs infection. Additionally, genotypes 66, 102 and *P. trifoliata*, classified by Boava et al., (2015) as tolerant, presented the same expression pattern of susceptible plants (19, 119, 124 and 217), except for *Cscals2*. Thus, making a connection between the expression values and level of tolerance or susceptibility is unlikely.

To find an association between the quantification of *Cscals* transcripts and allelic status of a genome region, we mapped the genomic regions associated with *Cscals* expression analysis in the linkage groups of *C. sunki* and *P. trifoliata* genetic maps. These genomic regions, referred to as eQTL, are important to understand the CLAs-host plant interaction and mechanisms of tolerance and response to HLB.

It was possible to identify eQTL for *Cscals2*, *Cscals7*, *Cscals8*, *Cscals9*, and *Cscals12* for both parents, although *P. trifoliata* is tolerant and does not exhibit callose deposition or starch accumulation after CLAs infection (Boava et al., 2017). Instead, no eQTL was found for *Cscals11* due to the low variation of expression data among CLAs-infected and healthy plants. Based on the estimation of the genetic parameters, *Cscals11* presented low heritability, indicating that the environment has great influence on this gene. Presumably, the regions that

control the genetic variability for *Cscals11* were not segregated in the study population, making it impossible to detect eQTL. The presence of important loci in homozygosity in both parents is a likely explanation for the absence of segregation for *Cscals11*.

Considering all eQTL mapped for the *Cscals7* gene, they explained the highest percentage of the phenotype variation between CLAs-infected and healthy plants. Thus, it is possible to state that *Cscals7* is the most affected evaluated gene after CLAs infection and is the most responsible for callose synthesis in the CLAs-infected plants.

Other evaluated genes were also affected by CLAs infection. eQTL were mapped for *Cscals2*, *Cscals7*, *Cscals8*, *Cscals9*, *Cscals10*, and *Cscals12* in the *C. sunki* map and for *Cscals2*, *Cscals5*, *Cscals7*, *Cscals8*, *Cscals9*, and *Cscals12* in the *P. trifoliata* map. In *C. sunki*, more than 44% of the eQTL observed were overlapped, characterizing hot spots. Thus, there are genomic regions that regulate the expression of more than one *Cscals* gene e.g., the main region on LG6 (200-203 cM) probably modulates *Cscals8* and *Cscals12* expression. In the *P. trifoliata* map, seven regions were considered hot spots and another 20 regions were mapped. Almost half of eQTL detected for *Cscals2* and *Cscals7* were overlapped. These regions and the other hot spots detected could probably be related to callose synthesis after CLAs infection.

Apparently, both parents contribute to the response of the callose synthase gene expression because many eQTL were observed in the same chromosome for *Cscals* in both maps. Based only on the SNP markers, it is hard to establish a direct correlation between the maps. However, comparing the eQTL for *Cscals*, an important region was verified for *P. trifoliata* on chromosome 8 that could influence the expression of *Cscals7* in plants affected by HLB.

The data sets obtained in this study revealed that it is not possible to determine whether the eQTL detected for *Cscals* in both maps represent the same genomic regions. Future studies should be considered to integrate the information from different materials.

Some eQTL can alter the expression of other genes located near them (cis-eQTL), explaining the variation of gene expression in the chromosomal region where the gene was found. On the other hand, other eQTL can regulate the expression of genes located distant from them (trans-eQTL), representing an effect of genetic polymorphisms that are located in other regions of the genome (Lima et al., 2018). The position of *calS* was confirmed to be in the *Citrus sinensis* genome (<http://citrus.hzau.edu.cn/orange/>); however, some genes did not have a defined position on pseudochromosomes because *Cscals9* and *Cscals12* were grouped on UnChr. Thus, for some cases, it was appropriate to determine whether the eQTL identified altered expression of nearby transcripts (cis-eQTL) or remote transcripts (trans-eQTL), usually on different chromosomes. Four SNP markers from the *P. trifoliata* map associated with *Cscals2*, *Cscals5*

and *Cscals7* were exclusively on the same chromosome as the genes, although they have been classified as trans-eQTL, because they are separated by more than 1 kb. Based on this investigation, we concluded that it is necessary to allocate *Cscals9* and *Cscals12* on the nine *Citrus* pseudochromosomes to make it possible to identify cis-eQTL. None of the SNP markers associated with *Cscals* expression was located on the region where the gene was found; therefore, probably all of the eQTL described in this study have an epistatic effect. The nonidentification of cis-eQTL could be due to two reasons for *Cscals* that has a physical position in the genome. First, the effect of some eQTL could be relatively low, hindering its mapping. Second, the polymorphism could be homozygous, causing possible variation in cis, such as promoters or enhancers (or other gene regulatory agents), with no segregation of the loci in the progeny.

Considering that *Cscals9* and *Cscals12* do not have known physical positions, this work warrants suggestions for future studies. Regions with eQTL can be considered as targets for other studies searching for regions where the *Cscals* genes can be located. Equally important, there is the possibility of identifying other genes that are related to *Cscals* functions. The identification of hot spots reinforces the idea that the eQTL detected in this study may be influencing the expression of *CsCalS*. Additionally, any gene physically located in a hotspot is a candidate, possibly explaining the studied process.

The gene expression and eQTL mapping results revealed that reprogramming occurs in callose synthesis in *P. trifoliata* as well as in *C. sunki*. However, there is evidence that *P. trifoliata* does not accumulate or accumulates much less callose than *C. sunki* (Boava et al., 2017). Thus, we believe that *P. trifoliata* has mechanisms that prevent callose deposition.

3.1.6. Conclusion

Despite the importance of eQTL mapping to provide a better understanding of the phenotypic variation (including those occurring during HLB), few related works exist in the literature. This study is the first to detect genomic regions associated with *Cscals* expression in plants infected with the causal agent of HLB disease.

The expression of all callose synthase genes was affected after CLas infection in the hybrid population studied. Thus, eQTL for *Cscals2*, *Cscals7*, *Cscals8*, *Cscals9*, *Cscals10*, and *Cscals12* were mapped in the *C. sunki* map and eQTL for *Cscals2*, *Cscals5*, *Cscals7*, *Cscals8*, *Cscals9* and *Cscals12* were mapped in the *P. trifoliata* map. eQTL analysis indicated that multiple regions can contribute to *Cscals* expression regulation and some eQTL have an

epistatic effect for more than one *Cscals* gene. An important region was also verified on linkage group 8 that could influence the expression of *Cscals7* in plants affected by HLB.

The identification of hot spots reinforces the idea that eQTL identified in this study may influence the expression of *Cscals*. Additionally, any gene physically located in a hotspot is a candidate that can explain the studied process. This work suggests eQTL for *Cscals* related to HLB.

3.1.7. Supplementary Information

Acknowledgments

Funding was provided by the Instituto Nacional de Ciência e Tecnologia (INCT) de Genômica para Melhoramento de Citros (CNPq 88887136353/2017-0 and FAPESP 2014/50880-0), FAPESP 2018/00133-4. The authors Maiara Curtolo and Laís Moreira Granato are recipients of research fellowships FAPESP 2016/22133-0 and FAPESP 2019/01901-8, respectively.

Conflict of Interest

The authors declare that there is no conflict of interest.

Authors Contributions

MAM and MCY planned and supervised the study. MC, LMG and TAST conducted the experiments. MC and RG analyzed the data. MC, LMG, TAST and MC wrote the manuscript. MAM, MCY and MAT revised the manuscript. All authors have read and approved the final manuscript.

3.1.8. References

- Albrecht U and Bowman KD (2012) Transcriptional response of susceptible and tolerant citrus to infection with *Candidatus Liberibacter asiaticus*. *Plant Sci.* 185–186: 118–130.
- Barratt DH, Kolling K, Graf A, Pike M, Calder G, Findlay K, Zeeman SC and Smith AM (2011) Callose synthase *GSL7* is necessary for normal phloem transport and inflorescence growth in *Arabidopsis*. *Plant Physiol* 155:328–341.
- Bertioli DJ, Ozias-Akins P, Chu Y, Dantas KM, Santos SP, Gouvea E, Guimarães PM, Leal-Bertioli SC, Knapp SJ and Moretzsohn MC (2014) The Use of SNP Markers for Linkage Mapping in Diploid and Tetraploid Peanuts. *G3 (Bethesda)* 4: 89-96.
- Boava LP, Cristofani-Yaly M and Machado MA (2017) Physiologic anatomic and gene expression changes in *Citrus sunki*, *Poncirus trifoliata* and their hybrids after *Liberibacter asiaticus* infection. *Phytopathology* 107(5):590-599
- Boava LP, Sagawa CH, Cristofani-Yaly M and Machado MA (2015) Incidence of “*Candidatus Liberibacter asiaticus*”-infected plants among citrandarins as rootstock and scion under field conditions. *Phytopathology* 105(4): 518–524.
- Bové JM (2006) Huanglongbing: a destructive newly-emerging century-old disease of citrus. *J Plant Physiol Pathol* 88 (1): 7-37.
- Castle WS, Tucker DPH, Krezdorn AH and Youtsey CO (1993) Rootstocks for Florida Citrus; rootstock selection - the first step to success. 2.ed. Gainesville University of Florida.
- Chang S, Puryear J and Cairney J (1993) A simple and efficient method for isolating RNA from pine trees. *Plant Mol Biol R.* 11(2): 113–116.
- Chen CX, Bowman KD, Choi YA, Dang PM, Rao MN, Huang S, Soneji JR, McCollum TG and Gmitter FG (2008) EST-SSR genetic maps for *Citrus sinensis* and *Poncirus trifoliata*. *Tree Genet Genomes* 4: 1-10.
- Chen L and Storey JD (2006) Relaxed significance criteria for linkage analysis. *Genetics* 173: 2371-2381.

- Chen XY and Kim JY (2009) Callose synthesis in higher plants. *Plant Signal Behav* 4: 489–492.
- Churchill GA and Doerge RW (1994) Empirical threshold values for quantitative trait mapping. *Genetics* 138: 963–971.
- Colleta-Filho HD, Tagon MLPN, Takita MA, De Negri JD, Pompeu Júnior J, Carvalho AS and Machado MA (2004) First report of the causal agent of Huanglongbing (“*Candidatus Liberibacter asiaticus*”) in Brazil. *Plant Dis* 88:1382.
- Cui W and Lee JY (2006) Arabidopsis callose synthases CalS1/8 regulate plasmodesmal permeability during stress. *Nat. Plants* 2: 16034.
- Curtolo M, Cristofani-Yaly M, Gazaffi R, Takita MA, Figueira A and Machado MA (2017) QTL mapping for fruit quality in *Citrus* using DArTseq markers. *BMC Genomics* 18: 289.
- Curtolo M, Soratto T, Gazaffi R, Takita MA, Cristofani-Yaly M and Machado MA (2018) High-density linkage maps for *Citrus sunki* and *Poncirus trifoliata* using DArTseq markers. *Tree Genet Genomes*.14:1.
- Dong X, Hong Z, Chatterjee J, Kim S and Verma DP (2008) Expression of callose synthase genes and its connection with Npr1 signaling pathway during pathogen infection. *Planta* 229: 87–98
- Ellinger D and Voigt CA (2014) Callose biosynthesis in *Arabidopsis* with a focus on pathogen response: what we have learned within the last decade. *Ann. Bot.* 114: 1349–1358.
- Enns LC, Kanaoka MM, Torii KU, Comai L, Okada K and Cleland RE (2005) Two callose synthases GSL1 and GSL5 play an essential and redundant role in plant and pollen development and in fertility. *Plant Mol Biol* 58: 333–349.
- Etxeberria E, Gonzalez P, Achor D and Albrigo G (2009) Anatomical distribution of abnormally high levels of starch in HLB-affected Valencia orange trees. *Physiol Mol Plant Pathol* 74: 76–83.
- Folimonova SY, Robertson CJ, Garnsey SM, Gowda S, Dawson WO (2009) Examination of responses of different genotypes of citrus to Huanglongbing (citrus greening) under different conditions. *Phytopathology* 99: 1346-54.

- Gazaffi R, Margarido GRA, Pastina MM, Mollinari M and Garcia AAF (2014) A model for quantitative trait loci mapping linkage phase and segregation pattern estimation for a full-sib progeny. *Tree Genet Genomes* 10: 791–801.
- Granato LM, Galdeano DM, Alessandre NR, Breton MC and Machado MA (2019) Callose Synthase Family Genes Play an Important Role in the *Citrus* Defense Response to *Candidatus Liberibacter asiaticus*. *Eur J Plant Pathol*, 1-14.
- Gomez-Gomez L, Felix G and Boller T (1999) A single locus determines sensitivity to bacterial flagellin in *Arabidopsis thaliana*. *Plant J* 19: 277-284.
- Gottwald T (2010) Current Epidemiological Understanding of *Citrus* Huanglongbing. *Annu Rev Phytopathol* 48:119-139.
- Grattapaglia D and Sederoff R (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. *Genetics* 137(4):1121-37.
- Guo F, Yu HW, Zheng T, Jiang XL, Wang L, Wang X, Xu Q and Deng X (2015) Construction of a SNP-based high-density genetic map for pumelo using RAD sequencing. *Tree Genet Genomes*. 11: 2.
- Hong Z, Delauney AJ and Verma DPS (2001) A Cell Plate–Specific Callose Synthase and Its Interaction with Phragmoplastin. *Plant Cell* 13(4): 755–768.
- Huang M, Roose ML, Yu Q, Du D, Yu Y, Zhang Y, Deng Z, Stover E and Gmitter FG (2018) Construction of high-density genetic maps and detection of QTLs associated with Huanglongbing tolerance in citrus. *Front Plant Sci*. 9:1694.
- Imai A, Yoshioka T and Hayashi T (2017) Quantitative trait locus (QTL) analysis of fruit-quality traits for mandarin breeding in Japan. *Tree Genet Genomes* 13:79.
- Jacobs AK, Lipka V, Burton RA, Panstruga R, Strizhov N, Schulze-Lefert P and Fincher GB (2003) An *Arabidopsis* callose synthase *GSL5* is required for wound and papillary callose formation. *Plant Cell* 15:2503–2513.

- Jagoueix S, Bove JM and Garnier M (1994) The phloem-limited bacterium of greening disease of the proteobacteria is a member of the alpha subdivision of the Proteobacteria. *Intl J Syst Bacteriol* 44:379–386.
- Koh EJ, Zhou L, Williams DS, Park J, Ding N, Duan YP and Kang BH (2012) Callose deposition in the phloem plasmodesmata and inhibition of phloem transport in citrus leaves infected with “*Candidatus Liberibacter asiaticus*”. *Protoplasma* 249(3): 687–697.
- Lima RPM, Curtolo M, Merfa MV, Cristofani-Yaly M and Machado M (2018) QTLs and eQTLs mapping related to citrandarins resistance to citrus gummosis disease. *BMC Genomics* 19: 516.
- Livak KJ and Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method 25: 402–408.
- Lorch PD (2005) Sex differences in recombination and mapping adaptations. *Genetica* 123: 39-47.
- Luna E, Pastor V, Robert J, Flors V, Mauch-Mani B and Ton J (2011) Callose deposition: a multifaceted plant defense response. *Mol Plant Microbe Interact* 24(2): 183-193.
- Mafra V, Kubo KS, Alves-Ferreira M, Ribeiro-Alves M, Stuart RM, Boava LP, Rodrigues CM and Machado (2012) Reference genes for accurate transcript normalization in citrus genotypes under different experimental conditions. *PLoS One* 7(2):e31263.
- Margarido GRA, Souza AP and Garcia AAF (2007) OneMap software for genetic mapping in outcrossing species. *Hereditas* 144: 78-79.
- Murray MG and Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* 8: 4321-4325.
- Nica AC and Emmanouil TD (2013) Expression quantitative trait loci: present and future. *Philos Trans R Soc Lond B Biol Sci* 368(1620): 20120362.
- Nishimura MT, Stein M, Hou BH, Vogel JP, Edwards H and Somerville SC (2003) Loss of a callose synthase results in salicylic acid-dependent disease resistance. *Sci* 301: 969–972.

Ollitrault P, Terol J, Garcia-Lor A, Bérard A, Chauveau A, Froelicher Y, Belzile C, Morillon R, Navarro L, Brunel D et al. (2012) SNP mining in *C. clementina* BAC end sequences; transferability in the *Citrus* genus (Rutaceae) phylogenetic inferences and perspectives for genetic mapping. *BMC Genomics* 13: 13-16.

Pang XM, Hu CG and Deng XX (2007) Phylogenetic relationships within *Citrus* and its related genera as inferred from AFLP markers. *Genet Resour Crop Evol* 54: 429-436.

Passos OS, Peixoto LS, Santos LC, Caldas RC and Soares Filho WS (2006) Caracterização de híbridos de *Poncirus trifoliata* e de outros porta-enxertos de citros no Estado da Bahia. *Ver. Bras. Frutic.* 28: 410-413.

Porto BN, Magalhães PC, Campos NA, Alves JD and Magalhães MM (2010) Otimização de protocolos de extração de RNA em diferentes tecidos de milho. *Revista Brasileira de Milho e Sorgo* 9(2): 189–200.

Raman H, Raman R, Kilian A, Detering F, Carling J, Coombes N, Diffey S, Kadkol G, Edwards D, McCully M, Ruperap P, Parkin IAP, Batley J, Luckett DJ, Wratten N (2014). Genome-wide delineation of natural variation for pod shatter resistance in *Brassica napus*. *PLoS One* 9: e101673.

Schneider H (1968) Anatomy of greening-disease sweet orange shoots. *Phytopathology* 58: 1155–1160.

Shimada T, Fujii H, Endo T, Ueda T, Sugiyama A, Nakano M, Kita M, Yoshioka T, Shimizu T and Nesumi H (2014) Construction of a citrus framework genetic map anchored by 708 gene-based markers. *Tree Genet Genomes* 10: 1001-1013.

Souza LM, Gazaffi R, Mantello CC, Silva CC, Garcia D, Le Guen V, Cardoso SEA, Garcia AAF and Souza AP (2013) QTL Mapping of Growth-Related Traits in a Full-Sib Family of Rubber Tree (*Hevea brasiliensis*) evaluated in a Sub-Tropical Climate. *PLoS ONE* 8(4): e61238.

Töller A, Brownfield L, Neu C, Twell D and Schulze-Lefert P (2008) Dual function of *Arabidopsis* glucan synthase-like genes GSL8 and GSL10 in male gametophyte development and plant growth. *Plant J* 54: 911–923.

- Verma DPS and Hong Z (2001) Plant callose synthase complexes. *Plant Mol Biol* 47:693–701.
- Wang N and Trivedi P (2013) Citrus Huanglongbing: a newly relevant disease presents unprecedented challenges. *Phytopathology* 103(7):652-65.
- Wang N, Pierson EA, Setubal JC, Xu J, Levy JG, Zhang Y, Li J, Rangel LT and Martins J Jr (2017) Host Interface: Insights into Pathogenesis Mechanisms and Disease Control. *Annu Rev Phytopathol* 55:451-482.
- Wang X, Zhu M, Zhang Z and Hong Z (2011) CalS7 encodes a callose synthase responsible for callose deposition in the phloem. *Plant J* 65:1–14.
- Wu R, Ma CX, Painter I and Zeng ZB (2002) Simultaneous maximum likelihood estimation of linkage and linkage phases in outcrossing species. *Theor Popul Biol* 61: 349-363.
- Xie B, Wang X, Zhu M, Zhang Z and Hong Z (2011) CalS7 encodes a callose synthase responsible for callose deposition in the phloem. *Plant J* 65: 1-14.
- Xu Q, Chen LL, Ruan X, Chen D, Zhu A, Chen C, Bertrand D, Jiao B, Hao B, Lyon MP et al. (2012) The draft genome of sweet orange (*Citrus sinensis*). *Nature Genetics* 45: 59-66.
- Yu Y, Chen C and Gmitter FG (2016a) QTL mapping of mandarin (*Citrus reticulata*) fruit characters using high-throughput SNP markers. *Tree Genet Genome* 12(4): 77.
- Yu Y, Jiao L, Fu S, Yin L, Zhang Y and Lu J (2016b) Callose synthase family genes involved in the grapevine defense response to downy mildew disease. *Phytopathology* 106:56-64.
- Zeng Z (1994) Precision mapping of quantitative trait loci. *Genetics* 136: 1457–1468.

3.1.9. Internet Resources

- BLASTn tool*, <https://blast.ncbi.nlm.nih.gov> (May 28, 2018)
- Citrus sinensis* reference genome, <http://citrus.hzau.edu.cn/> (May 28, 2018)
- Clementine tangerine reference genome, <https://www.phytozome.org> (September 10, 2017)
- Fundecitrus, <https://www.fundecitrus.com.br/> (Apr 30, 2018)
- MeV (MultiExperiment Viewer) program v. 4.9, <http://sourceforge.net/projects/mev-tm4/> (Jun 25, 2018)

Miner software, <http://ewindup.info/miner> (May 15, 2018)

R software, <https://www.r-project.org> (May 15, 2018)

3.1.10. Supplementary material

Table S1: Sequences of primer pairs used for qPCR analysis

Gene	Localization	Primers sequences (5'-3')	Reference
<i>Cscals2</i>	LOC102624514	F, ATCTCTGCCGGTTCTATGCG R, CGGGCATCACTCTTTGACCT	Granato et al., 2019
<i>Cscals5</i>	LOC102618167	F, GTGTGATTGAAACGGAAGCCA R, CCATCATCACGCATAGGCCA	Granato et al., 2019
<i>Cscals7</i>	LOC102612996	F, GACGCCTAACCGAGTACCTGC R, GTGCAGCTGGTGATCCATCA	Granato et al., 2019
<i>Cscals8</i>	LOC102631245	F, AGGATGTTTTTCGCCGGTACA R, ATCACGACCTTTGCCCACTT	Granato et al., 2019
<i>Cscals9</i>	LOC102612131	F, TCCTTTCTCGAATTGGCCGT R, TGTCTGTGCGCGATATGAGG	Granato et al., 2019
<i>Cscals10</i>	LOC102616583	F, GGCTCGACTTGGCATACTTG R, AACTGTTCCAAGCAAGGCGT	Granato et al., 2019
<i>Cscals11</i>	LOC102627313	F, GATGTGTACCGCTTGGGTCA R, AGCAAGATAAAGACGCCCC	Granato et al., 2019
<i>Cscals12</i>	LOC102610237	F, CTTGGGTCAGCGTGTTTTGG R, CTCCTCGCAGTGTGCAGTTA	Granato et al., 2019
<i>GAPDH</i>	At1g13440	F, GGAAGGTCAAGATCGGAATCAA R, CGTCCCTCTGCAAGATGACTCT	Mafra et al., 2012
<i>FBOX</i>	At5g15710	F, GGCTGAGAGGTTTCGAGTGTT R, GGCTGTTGCATGACTGAAGA	Mafra et al., 2012

Table S2: Adjusted values of the expression of *CsCals* 2, 5, 7, 8, 9, 10, 11 and 12.

Genotypes	Genes							
	<i>CsCals2</i>	<i>CsCals5</i>	<i>CsCals7</i>	<i>CsCals8</i>	<i>CsCals9</i>	<i>CsCals10</i>	<i>CsCal11</i>	<i>CsCals12</i>
1	1.95	1.99	1.10	0.56	1.38	1.08	1.24	0.09
10	0.32	5.19	1.15	#N/D	2.16	-0.19	1.27	0.33
101	0.82	-0.74	1.34	#N/D	1.18	0.04	1.22	-0.22
102	0.90	0.87	1.31	0.61	0.93	0.33	1.19	0.20
105	1.30	2.67	2.63	0.64	1.13	1.42	1.22	0.40
106	2.28	0.28	1.00	0.03	1.40	-0.16	1.22	0.16
107	1.22	-0.15	0.94	-0.74	0.69	0.94	1.21	-0.25
109	1.22	0.40	1.04	-0.62	1.05	1.14	1.22	-0.37
110	6.39	8.98	1.31	4.30	1.90	1.96	1.37	1.45
111	0.78	0.45	0.93	0.38	0.77	0.15	1.20	0.26
113	2.24	20.18	1.09	4.35	0.94	1.03	1.22	0.11
117	-0.56	0.50	0.96	-0.11	1.15	-0.16	1.22	0.07
118	1.38	1.06	1.01	-0.38	0.80	1.10	1.27	-0.27
119	1.98	0.11	1.06	-0.46	0.74	1.13	1.35	-0.10
121	1.02	0.38	0.73	0.02	0.35	1.56	1.20	0.23

124	3.32	0.58	1.55	0.32	1.46	0.20	1.24	0.90
125	3.21	1.56	2.74	10.43	2.53	1.65	1.52	3.93
126	-0.84	0.30	1.40	12.68	1.26	-0.32	1.32	0.39
129	3.20	-0.47	1.48	2.58	2.28	1.05	1.22	0.52
130	3.44	2.94	1.33	2.28	3.17	4.10	1.49	5.00
132	1.50	2.68	2.40	1.46	1.73	2.51	1.22	3.61
134	1.49	0.15	1.10	-0.63	0.66	1.00	1.22	-0.35
136	2.00	0.02	1.02	2.93	0.87	1.03	1.26	-0.07
137	1.90	0.56	1.10	0.37	1.13	2.26	1.25	1.54
14	-0.74	1.89	0.98	0.52	1.29	-0.31	1.21	0.13
141	8.00	3.54	1.59	2.23	3.69	1.58	1.22	3.50
142	3.47	3.08	1.18	1.62	1.13	1.33	1.33	0.21
143	18.08	9.77	3.10	0.74	2.07	54.37	1.34	4.86
146	7.34	0.56	1.57	0.34	4.94	1.98	1.33	1.29
148	1.33	-0.03	1.04	-0.41	1.25	1.09	1.21	0.39
149	2.29	-0.24	1.14	-0.23	0.90	1.04	1.32	0.34
150	-0.78	0.18	0.88	-0.23	1.08	-0.35	1.20	0.08
151	0.91	0.55	1.34	1.48	0.94	0.38	1.23	0.37
154	1.60	0.22	1.75	0.30	1.47	0.34	1.21	0.48
16	1.56	14.60	1.44	1.04	1.30	1.02	1.34	-0.37
163	2.05	0.69	1.72	0.61	1.24	1.19	1.32	1.22
173	1.64	1.10	1.45	5.68	1.53	0.91	1.20	-0.12
179	0.89	0.53	2.13	0.50	1.21	0.80	1.23	0.71
183	0.35	0.59	1.35	1.78	1.10	0.03	1.31	0.25
184	1.83	1.51	1.86	0.01	1.26	2.09	1.30	0.35
187	2.14	0.79	1.00	-0.41	1.22	1.39	1.22	0.37
189	1.29	0.41	1.70	0.68	0.48	0.24	1.26	0.36
19	6.92	-0.11	0.94	-0.06	3.23	0.97	1.21	-0.30
191	0.74	0.63	1.21	1.24	0.45	0.83	1.29	0.80
2	-0.09	0.89	1.06	0.11	1.44	0.49	1.24	0.13
20	-0.19	1.03	0.97	1.90	1.62	-0.24	1.29	0.16
217	0.91	0.19	0.89	0.22	0.50	1.33	1.20	0.40
23	0.95	0.04	1.29	0.80	0.84	0.06	1.20	-0.25
24	1.41	-0.06	1.46	7.91	1.75	1.01	1.20	0.54
26	1.18	1.11	1.89	0.38	1.22	0.21	1.25	0.40
279	1.22	0.42	0.79	0.85	0.60	2.28	1.21	0.46
28	1.88	4.71	0.77	#N/D	1.74	1.55	1.19	0.16
293	4.97	0.94	1.17	0.88	0.92	2.42	1.33	1.00
31	3.39	0.47	1.42	0.77	3.80	4.55	1.36	2.28
35	0.83	6.42	1.19	1.45	3.14	-0.28	1.25	0.37
4	2.52	0.75	1.67	#N/D	4.46	3.50	1.33	0.28
42	-0.66	0.31	0.92	1.15	1.85	-0.38	1.21	0.24
47	-0.51	2.02	0.84	0.16	1.76	-0.35	1.21	0.20
49	9.05	1.30	2.95	19.66	6.01	8.51	1.71	1.53
54	0.64	0.83	0.71	0.75	0.59	0.42	1.19	0.26
56	1.02	0.92	0.95	3.91	0.90	0.10	1.20	0.34

61	0.56	-0.01	1.30	1.80	1.09	0.43	1.40	0.59
66	0.47	-0.64	1.29	#N/D	1.04	-0.14	1.17	-0.26
68	3.64	-0.21	1.31	6.37	1.31	2.12	1.24	1.10
70	0.63	0.50	0.71	0.04	0.24	1.01	1.19	0.11
73	3.06	0.41	1.42	3.83	1.21	0.52	1.24	0.18
78	2.94	1.32	4.81	13.67	1.40	0.76	1.47	1.81
86	0.44	-0.78	1.29	#N/D	1.05	-0.15	1.18	-0.31
90	-0.19	0.81	1.16	2.60	1.24	-0.08	1.30	0.20
94	1.69	0.83	1.04	15.20	1.38	-0.06	1.28	0.33
96	-0.32	0.70	0.84	0.12	1.15	-0.15	1.20	0.11
99	3.78	0.60	1.37	0.86	0.94	0.33	1.30	-0.06
<i>C. Sunki</i>	2.79	0.34	2.11	0.80	1.66	1.94	1.39	1.33
<i>P. trif</i>	1.67	0.43	1.80	0.43	0.95	0.70	1.46	0.30

3.2. Chapter 2: Curtolo M, de Souza Pacheco I, Boava LP, et al (2020) Wide-ranging transcriptomic analysis of *Poncirus trifoliata*, *Citrus sunki*, *Citrus sinensis* and contrasting hybrids reveals HLB tolerance mechanisms. *Sci Rep* 10:1–14. <https://doi.org/10.1038/s41598-020-77840-2>.

Wide-ranging transcriptomic analysis of *Poncirus trifoliata*, *Citrus sunki*, *Citrus sinensis* and contrasting hybrids reveals HLB tolerance mechanisms

Maiara Curtolo*^{1,2}; Inaiara de Souza Pacheco^{1,2}; Leonardo Pires Boava¹; Marco Aurélio Takita¹; Laís Moreira Granato¹; Diogo Manzano Galdeano¹; Alessandra Alves de Souza¹; Mariângela Cristofani-Yaly¹ and Marcos Antonio Machado¹

¹Centro de Citricultura Sylvio Moreira, Instituto Agronômico de Campinas, Cordeirópolis/SP, Brazil. ²Universidade Estadual de Campinas, Campinas/SP, Brazil.

*Corresponding author: Maiara Curtolo; e-mail: maiaramc@hotmail.com

3.2.1. Abstract

Huanglongbing (HLB), caused mainly by ‘*Candidatus Liberibacter asiaticus*’ (CLas), is the most devastating citrus disease because all commercial species are susceptible. HLB tolerance has been observed in *Poncirus trifoliata* and their hybrids. A wide-ranging transcriptomic analysis using contrasting genotypes regarding HLB severity was performed to identify the genetic mechanism associated with tolerance to HLB. The genotypes included *Citrus sinensis*, *Citrus sunki*, *Poncirus trifoliata* and three distinct groups of hybrids obtained from crosses between *C. sunki* and *P. trifoliata*. According to bacterial titer and symptomatology studies, the hybrids were clustered as susceptible, tolerant and resistant to HLB. In *P. trifoliata* and resistant hybrids, genes related to specific pathways were differentially expressed, in contrast to *C. sinensis*, *C. sunki* and susceptible hybrids, where several pathways were reprogrammed in response to CLas. Notably, a genetic tolerance mechanism was associated with the downregulation of gibberellin (GA) synthesis and the induction of cell wall strengthening. These defense mechanisms were triggered by a class of receptor-related genes and the induction of WRKY transcription factors. These results led us to build a hypothetical model to understand the genetic mechanisms involved in HLB tolerance that can be used as target guidance to develop citrus varieties or rootstocks with potential resistance to HLB.

3.2.2. Introduction

Huanglongbing (HLB) or Greening has been considered the most devastating citrus disease. HLB is caused by the gram-negative, phloem-limited, α -proteobacterium *Candidatus Liberibacter* species. The following three *Liberibacter* species have been associated with HLB: *Candidatus Liberibacter asiaticus* (CLas), *Candidatus Liberibacter americanus* (CLam) and *Candidatus Liberibacter africanus* (CLaf). CLas is the most widespread and is responsible for large economic losses worldwide ^{1,2}.

HLB symptoms include blotchy chlorosis, mottling of leaves, yellow shoots, vein corking, stunted growth and small, green, and lopsided fruits with aborted seeds ³. HLB symptom development is considered a consequence of a series of molecular, cellular, and physiological disorders in the plant host. The most expressive modifications caused by CLas in the citrus host are alterations in sucrose and starch metabolism, changes of hormone production, biosynthesis of secondary metabolites, phloem function disorders, and source-sink communication ^{4,5}.

Poncirus trifoliata is closely related and sexually compatible with the citrus genus, and it shows attenuated HLB symptoms and lower CLas titer, indicating that this genus possibly presents genetic defense mechanism against CLas ^{6,7}. Moreover, some citrus hybrids of *P. trifoliata* have also been reported to present a significant tolerance to HLB ^{7,8}; however, it remains unclear which mechanisms are involved in this tolerance. In contrast, all commercial *Citrus* species are susceptible to CLas infection, and the identification of tolerant genotypes is essential to the maintenance of citrus production ². Studies are still necessary to understand better the differences of genetic responses involved in the susceptibility, tolerance or resistance to such genotypes, aiming to obtain new citrus variety tolerant to HLB by conventional breeding or genetic engineering.

Our study provides a wide-ranging transcriptomic analysis of two CLas-susceptible citrus genotypes (*Citrus sinensis* and *C. sunki*), one CLas-tolerant genotype (*P. trifoliata*), and three pools of hybrids between *P. trifoliata* and *C. sunki*, which are classified as susceptible, tolerant, and resistant to HLB. Therefore, this work was the first to study transcriptional reprogramming and to compare the results of a large volume of transcriptomes, including individuals from a population of hybrids infected by CLas, which consequently inherited the susceptible and tolerance genetic mechanisms from their parents.

The results revealed that only a few genes associated with specific pathways were modulated in resistant genotypes to avoid CLas proliferation and plant disease severity. Using the transcriptomic analysis of the hybrid genotypes, we revalidated the mechanisms of

susceptibility and tolerance of their parents. Based on the analysis, we built a hypothetical model to explain the genetic mechanism involved in HLB tolerance conferred by *P. trifoliata* and inherited by its hybrids that could be further used in breeding or biotechnological approaches.

3.2.3. Results

3.2.3.1. CLas quantification

CLas quantification analysis showed that all plants from *C. sinensis*, *C. sunki*, and *P. trifoliata* were infected by CLas after 240 days of inoculation. From the analysis of the 21 hybrids, nine of them (H68, H106, H109, H113, H142, H156, H154, H161, and H165) were selected for the subsequent steps. The H109, H161, H165, H113, H154, and H146 hybrids were infected, but the H68, H106, and H142 hybrids were negative for the presence of CLas in all biological replicates (Table 1 and 2).

Table 1. Detection and quantification of the bacteria by quantitative PCR (qPCR) in *Citrus sinensis*, *C. sunki*, *Poncirus trifoliata* and nine hybrids from an F₁ population obtained from the cross between *C. sunki* and *P. trifoliata* Raf. cv Rubidoux. Each individual is represented by five repetitions.

Genotypes	HLB diagnosis (qPCR) days after inoculation				
	30	90	180	240	360
<i>C. sunki</i>	0/5	3/5	4/5	4/5	5/5
<i>C. sinensis</i>	0/5	3/5	5/5	5/5	5/5
<i>P. trifoliata</i>	0/5	0/5	0/5	3/5	3/5
H106	0/5	0/5	0/5	0/5	0/5
H109	1/5	2/5	4/5	5/5	5/5
H146	1/5	1/5	3/5	5/5	5/5
H68	0/5	0/5	0/5	0/5	0/5
H161	0/5	3/5	5/5	5/5	5/5
H142	0/5	0/5	0/5	0/5	0/5
H165	0/5	3/5	5/5	5/5	5/5
H154	-	-	4/5	5/5	5/5
H113	-	-	5/5	5/5	5/5

Table 2. *Candidatus Liberibacter asiaticus* (CLas) quantification obtained by comparing the standard curve of the HLB primers with the standard curve of the internal control gene (GAPDH) initiators. The value quantification refers to Log₁₀ of the number of copies of the

CLas fragment after 240 days from inoculation in each repetition per genotype included in RNAseq analysis.

Genotype	Ct value of GAPDH	Ct value of HLB	Quantification/Log₁₀ number of copies
<i>C. sinensis</i>	24.24	26.48	3.48
	22.07	24.30	4.09
	20.12	31.74	2.01
<i>C. sunki</i>	18.43	24.90	3.92
	18.18	20.51	5.15
	18.56	25.17	3.85
<i>P. trifoliata</i>	19.99	20.31	5.20
	19.20	21.85	4.78
	18.30	25.80	3.67
H109	19.16	17.95	5.86
	18.30	20.35	5.19
	19.13	20.38	5.19
H161	20.32	21.12	4.98
	19.28	18,87	5.61
	19.23	18.75	5.64
H165	18.98	18.85	5.61
	19.05	22.18	4.68
	18.52	18.87	5.61
H113	17.67	21.93	4.75
	19.14	26.25	3.55
	19.05	20.81	5.07
H154	19.47	30.79	2.28
	19.05	19.33	5.48
	19.05	20.55	5.14
H146	18.78	21.33	4.92
	18.35	22.16	4.69
	18.48	25.60	3.73
H68	20.28	Undetermined	0

	19.39	Undetermined	0
	19.50	Undetermined	0
H106	19.96	Undetermined	0
	19.86	Undetermined	0
	18.34	Undetermined	0
H142	20.93	Undetermined	0
	18.59	Undetermined	0
	18.99	Undetermined	0

3.2.3.2. Phenotypic analysis

A significant increase in callose deposition was observed for the CLas-infected *C. sinensis*, *C. sunki*, H109, H161, and H165 plants compared to the control (Fig. 1). Moreover, *P. trifoliata*, H113, H154, H146, H68, H106, and H142 showed no difference between the mock and CLas-inoculated plants (Fig. 1). Compared with inoculated and mock-inoculated plants, *C. sinensis*, *C. sunki*, and three infected hybrids (H109, H161 and H165) showed a significant difference ($p < 0.05$) in the amount of starch. In contrast, no significant difference in starch accumulation was observed in *P. trifoliata* and the other six hybrids (H113, H154, H146, H68, H106, and H142) (Fig. 1).

In general, the visual symptoms were more evident in the susceptible plants, while the visual HLB symptoms were undefined in *P. trifoliata* and its hybrids. However, according to CLas detection, starch and callose quantification between different treatments, the hybrids were clustered into three distinct groups as follows: Susceptible Pool (S Pool), composed of three different hybrids (H109, H161, and H165) that were diagnosed as HLB-positive and presented elevated starch and callose deposition, similar to that observed for susceptible parental genotypes (Fig. 1); Tolerant Pool (T Pool), composed of three different hybrids (H113, H154, and H146) that were diagnosed as HLB-positive but did not exhibit a significant starch and callose accumulation as observed in susceptible genotypes (Fig. 1); and Resistant Pool (R Pool), composed of three different hybrids (H68, H106, and H142) that were diagnosed as HLB-negative with starch quantification similar to healthy plants (mock-inoculated plants) (Fig. 1).

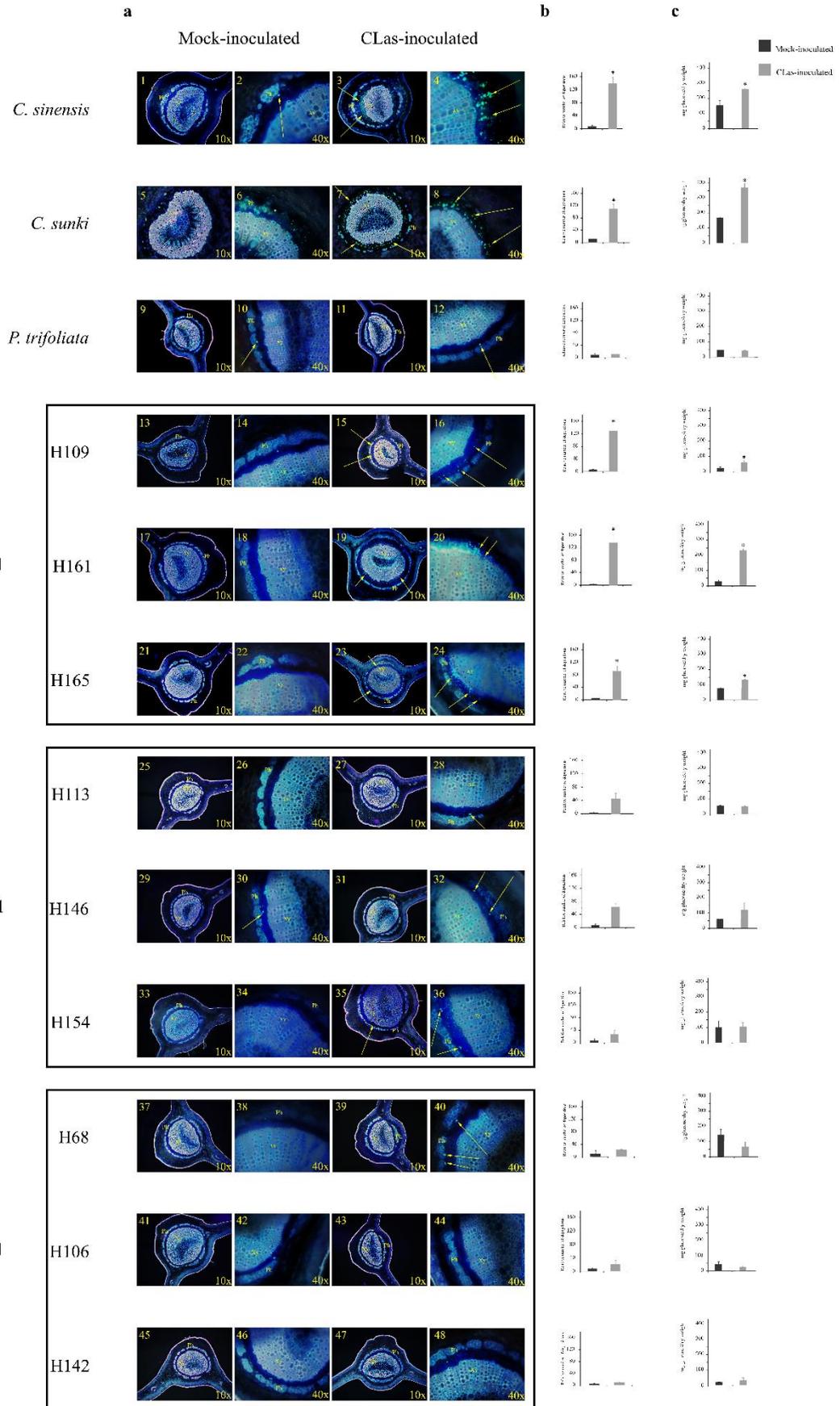


Fig. 1. Callose deposition. **a.** Cross sections of leaf petioles of *C. sinensis* mock-inoculated (1 and 2) and CLAs inoculated (3 and 4), *C. sunki* mock-inoculated (5 and 6) and CLAs inoculated (7 and 8), *P. trifoliata* mock-inoculated (9 and 10) and CLAs inoculated (11 and 12), H109 mock inoculated (13 and 14) and CLAs inoculated (15 and 16), H161 mock-inoculated (17 and 18) and CLAs inoculated (19 and 20), H165 mock-inoculated (21 and 22) and CLAs inoculated (23 and 24), H113 mock-inoculated (25 and 26) and CLAs inoculated (27 and 28), H146 mock-inoculated (28 and 30) and CLAs inoculated (31 and 32), H154 mock-inoculated (33 and 34) and CLAs inoculated (35 and 36), H68 mock-inoculated (37 and 38) and CLAs inoculated (39 and 40), H106 mock-inoculated (41 and 42) and CLAs inoculated (42 and 44), H142 mock-inoculated (45 and 46) and CLAs-inoculated (47 and 48). FL, phloem; Xi, xylem. **b.** The bar graph next to the microscopy plates show the callose quantification performed by counting fluorescent spots marked by aniline blue dye. Quantification was performed with three replicates per genotype, inoculated plants (positive or negative HLB) and mock-inoculated plants. **c.** Starch quantification. Individuals were inoculated with CLAs (CLAs-infected) or mock-inoculated (CLAs-free) and collection was performed after 240 days, and quantification was carried by the enzymatic method. Bars represent the standard deviation between 3 biological replicates. * $p_value < 0.05$ (Mock-inoculated x CLAs inoculated).

3.2.3.3. Transcriptome assembly

To elucidate the different responses to CLAs infection, we studied the changes in global transcriptional level in susceptible, tolerant, and resistant genotypes infected by CLAs. In this work, 36 cDNA libraries from six different genotypes of either CLAs-inoculated or mock-inoculated (control) samples were evaluated. After trimming, 487 million reads were obtained, and 95% of the total was assigned (see Supplementary Table S1). The reads were mapped in 133,976 transcripts on the *C. sinensis* genome available on <http://citrus.hzau.edu.cn/>.

HLB-susceptible genotypes, *C. sinensis* and *C. sunki*, showed a high number of differentially expressed genes (6,141 and 5,624 DEGs, respectively) compared with the tolerant parental, *P. trifoliata* (100 DEGs) (Table 3). A similar pattern was observed between the pool of hybrids. The S Pool showed 708 differentially expressed genes (DEGs), while the R Pool presented only 92 DGEs. The Tolerant Pool (T Pool) showed the highest number of DEGs (2,027) among the hybrid pools. Most of these genes were downregulated in HLB-infected plants compared with healthy ones (Table 3).

Table 3. Number of differentially expressed genes in *C. sinensis*, *C. sunki*, *P. trifoliata*, S Pool, T Pool and R Pool. CLAs-infected plants compared with healthy plants.

Genotypes	Up-regulated	Downregulated	Total
<i>C. sinensis</i>	3,175	2,966	6,141
<i>C. sunki</i>	3,288	2,336	5,624
<i>P. trifoliata</i>	70	30	100
S Pool	288	420	708
T Pool	939	1,088	2,027
R Pool	63	29	92
Total	5,812	5,331	14,692

The principal component analysis (PCA) using the Bioconductor package (see Supplementary Fig. S1) showed the replicates of the different genotypes in general grouped according to the analyzed condition for *C. sunki*, *C. sinensis*, and the susceptible and tolerant hybrids. The resistant groups in fact presented a mixed grouping, which is not surprising if we consider that these populations were the ones that showed the fewer number of DEGs. The genotype grouping indicated that the global expression landscape is related more to the different genotypes and not the analyzed condition (infection by CLAs). In this case, the identification of genes exclusively differentially expressed in the genotypes considered susceptible, tolerant, or resistant as well as genes that had antagonistic expression between the opposite phenotypes became important to increase our understanding of the different responses.

3.2.3.4. Differential gene expression analysis

The results are summarized in a Venn diagram (Fig. 2 and Table S2). The susceptible genotypes, *C. sinensis* and *C. sunki*, exhibited the highest number of overlapping DEGs (1,634), and 88% of these genes presented a similar expression pattern (Fig. 2, Supplementary Table S2), suggesting that a similar gene modulation is caused by CLAs infection. In *P. trifoliata*, 47% of the DEGs were exclusive of this genotype (Fig. 2 and Supplementary Table S2), and 26% of the DGEs were overlapped and showed antagonistic expression compared to susceptible genotypes. Five of the downregulated genes in *P. trifoliata* were upregulated in both *C. sinensis* and *C. sunki* genotype, and one gene was upregulated in the S Pool (see Supplementary Table S3 and S4).

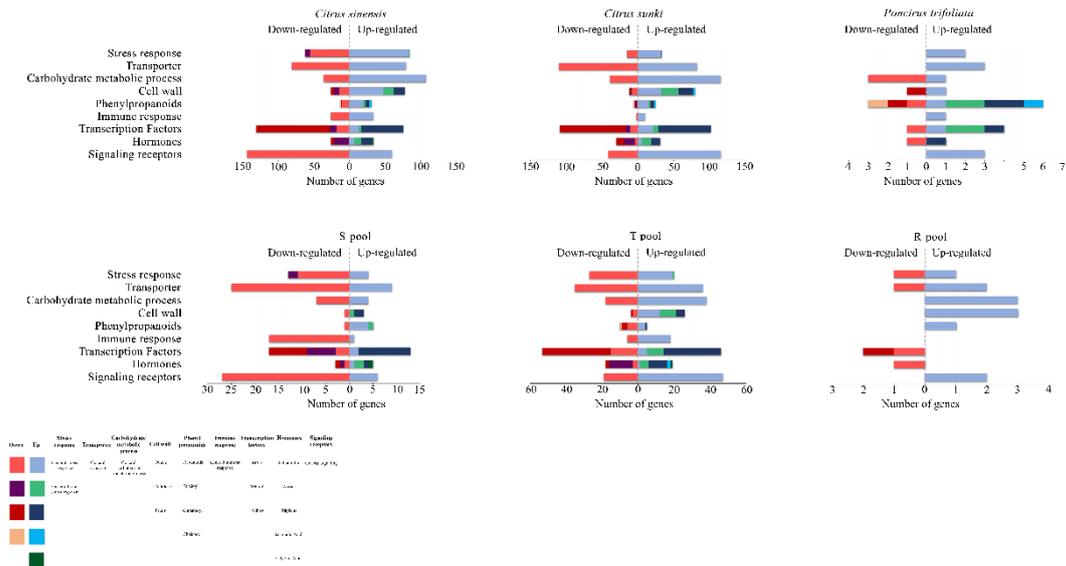


Fig. 3. *C. sinensis*, *C. sunki*, *P. trifoliata*, S Pool, T Pool and R Pool responses to 240 days of infection by CLAs. Genes are classified into nine groups (Stress response, Transporter, Carbohydrate metabolic process, Cell wall, Phenylpropanoids, Immune response, Transcription Factors, Hormones and Signaling receptors) according to Blast2GO analysis and based on their expression pattern. The number of downregulated genes in response to CLAs is represented by the bars in reddish tones and upregulated in blue tones. Some bars present subdivisions and the color legend for each pathway is indicating the specific, related gene or specific pathways, which were important to illustrate the proposed tolerance mechanism to HLB.

3.2.3.6. Differentially expressed genes (DEGs) associated with a specific biological pathway.

3.2.3.6.1. Signaling receptor

Plant receptors are responsible for the recognition of several external stimuli, including pathogen attack. These transmembrane proteins are directly associated with signaling pathways, which trigger a proper physiological response¹¹. Several types of receptors were regulated in *C. sinensis*, *C. sunki*, the S Pool, and the T Pool, and most of them were downregulated in those genotypes (Fig. 3). In *P. trifoliata* and the R Pool only a few receptors were differentially expressed, and most of them were induced (Fig. 3). These receptors included *G-type lectin S-receptor-like*, *cysteine-rich receptor kinase*, and *serine/threonine-protein kinase*, which were upregulated in *P. trifoliata*, and *leucine-rich repeat transmembrane kinase* and *leucine-rich repeat receptor-like protein kinase*, which were induced in the R Pool (see Supplementary Table S6). Therefore, our results suggested that downregulation of receptors may be associated with susceptible response to CLAs.

3.2.3.6.2. Hormones

Genes associated with auxin and ethylene pathways were barely or not affected in *P. trifoliata* and the R Pool, whereas many auxin and ethylene-related genes were differentially expressed in *C. sunki*, *C. sinensis*, the T Pool, and the S Pool under CLAs infection. Interestingly, no important changes in the transcriptional profiles of genes related to SA and JA biosynthesis were found (Fig. 3). In addition, CLAs induced key genes involved with gibberellin (GA) degradation in tolerant and resistant genotypes, while the related GA synthesis genes were downregulated. In *P. trifoliata*, the *gibberellin-induced* gene was one of the top three downregulated DEGs (log₂ fold change= -10) (see Supplementary Table S6). The opposite pattern was observed in CLAs-susceptible genotypes, in which an induction of genes involved with GA synthesis and downregulation of GA degradation was observed. Thus, these findings suggested that GA plays an important role in CLAs-citrus interactions, affecting plant physiology and consequently HLB symptoms.

3.2.3.6.3. Transcription factors

Plant responses to pathogen attack require large-scale transcriptional reprogramming. *P. trifoliata* showed only five transcription factor (TF)-related genes modulated by CLAs infection. Only the MYB TF was downregulated. The other four TFs were upregulated, including two WRKY TFs (Fig. 3). The resistant hybrids suppressed the expression of another class of transcription factor, the *SCL domain* (see Supplementary Table S6). In contrast, hundreds of TF genes showed changes at the transcription level in *C. sinensis*, *C. sunki*, the S Pool and the T Pool (Fig. 3). In this context, the large number of TFs affected in these genotypes may be directly related to the regulation of genes responsive to HLB infection. Of note, several WRKY TFs were identified in *C. sinensis* and *C. sunki*, and most of them were repressed in CLAs-infected plants (Fig. 3). Therefore, these results indicated that the increase in transcription of WRKY TFs in *P. trifoliata* is associated with the genetic defense mechanism involved with HLB tolerance.

3.2.3.6.4. Defense-related genes

Defense-related genes are directly related to processes or production of compounds able to inhibit pathogen reproduction or to make further infection more difficult¹². In particular, one defense-related gene, *endochitinase B*, was differentially expressed and highly upregulated in resistant hybrids (see Supplementary Table S6). Endochitinases have previously been reported

as important bactericides, and some of them have ability to cleave peptidoglycan chains, promoting bacterial cell lysis¹³. Other defense-related genes were differentially expressed in susceptible plants by CLAs. Among them, regions encoding lipid transfer, molecular factors that help the innate immune system of plants, and small lipid-transfer proteins can inhibit fungal growth and pathogenic bacteria¹⁴. Genes encoding these proteins were differentially expressed in *C. sinensis*, *C. sunki*, and the S Pool (see Supplementary Table S6 and Fig. S2). These results indicated the activation of defense pathways in response to CLAs infection in susceptible genotypes.

CDR1 also represents an important defense related gene in *Poncirus* and *Poncirus*-hybrids¹⁵. CDR1 showed high expression in all the *Poncirus* hybrids, including the S pool, but it was only induced in the R pool. Therefore, even though it could be associated with resistance, high CDR1 constitutive expression level seems not to be sufficient to lead to the resistance phenotype.

3.2.3.6.5. Secondary metabolism and cell wall composition

Secondary metabolites often play an important role in many physiological responses, such as growth, photosynthesis, reproduction, and plant defenses against pathogens¹⁶. The most upregulated genes in *P. trifoliata* included a variety of phenylpropanoids and lignin-related genes, such as *caffeic acid O-methyltransferase*, *chalcone synthase*, *feruloyl ortho-hydroxylase 1*, *hydroxycinnamoyl transferase* and *laccase precursor* (see Supplementary Fig. S2). In our study, the *laccase precursor* gene, whose protein catalyzes lignin and its derivatives¹⁷, was exclusive and highly induced in CLAs-infected *P. trifoliata* (see Supplementary Table S6).

Pectin hydrolysis occurs frequently in response to bacterial infection¹⁸. Just one pectin degradation-related gene was differentially expressed (downregulated) in *P. trifoliata* (Fig. 3 and Supplementary Table S6). Many genes involved in pectin synthesis and degradation were differentially expressed in *C. sinensis* and *C. sunki*. Pectin methyltransferases are enzymes that induce pectin modification. In *C. sinensis* and *C. sunki* under stress caused by CLAs infection, the *pectin methyltransferase 1* gene was upregulated (see Supplementary Fig. S2).

A larger number of DEGs involved in cellulose synthesis showed mRNA levels altered in susceptible genotypes; however, *P. trifoliata* and the R Pool did not exhibit differentially expressed regions encoding cellulose (see Supplementary Table S6).

These results demonstrated that the cell wall is highly affected in susceptible plants even at 240 days after CLAs inoculation. At the same time, genes involved in cell strengthening proved to be important in *P. trifoliata*.

3.2.3.6.6. Phloem-related genes

It is already known that callose deposition and phloem proteins (PP2) act as a physical barrier, attempting to block systemic spread of CLAs; however, they also likely cause phloem disorders¹⁹. The current study identified DEGs coding phloem proteins that had altered expression induced by CLAs in *C. sinensis*, *C. sunki*, the S Pool, and T Pool. Although *P. trifoliata* did not present callose-induced phloem blockage (Fig. 1), we observed modulation of PP2-B15 in response to CLAs with 9-fold higher expression than the control (see Supplementary Table S6). That result suggests that *P. trifoliata* modulates phloem genes in response to CLAs without over-deposition of callose, consequently not causing important phloem function disorders. Anatomical divergences between *P. trifoliata* and *Citrus* may represent an important feature to avoid collapse of the sieve tube elements²⁰.

As shown by our phenotypic data, only susceptible plants had affected callose deposition. Different callose synthases were differentially expressed in the susceptible plants, whereas those genes were absent in *P. trifoliata* and the R Pool inoculated with CLAs (see Supplementary Table S7).

Interestingly, genes encoding *sieve element occlusion c* (*SEOc*) and *d* (*SEOd*), which are part of a protein family that encodes specialized crystalloid phloem proteins²¹, were largely upregulated in all susceptible plants under study. Some of these genes were also upregulated in tolerant hybrids (see Supplementary Table S6 and S7).

3.2.3.6.7. Carbohydrate metabolism

Carbohydrate metabolism was the biological function most affected by HLB (Fig. 3). In the presence of CLAs, susceptible genotypes overexpressed genes involved with starch synthesis and suppressed genes that encode enzymes for starch degradation (see Supplementary Table S8 and S9). This phenomenon was not observed for the tolerant and resistant genotypes. Several DEGs involved in the metabolism of starch were identified in *C. sinensis*, *C. sunki*, and the T Pool, especially in the former two (see Supplementary Table S6). Genes encoding *ADP-glucose pyrophosphorylase* and *starch branching enzyme II*, which participate in the synthesis of starch and starch granules, were upregulated in *C. sinensis* and *C. sunki* (see Supplementary Table S6 and S8). Beta and alpha-amylase, important enzymes for normal degradation of the starch in plants²², also had their genes expression modulated in both susceptible plants (*C. sinensis* and *C. sunki*) and the T Pool (see Supplementary Table S9). Corroborating our phenotypic data (Fig. 1), resistant and tolerant genotypes did not exhibit altered expression of the main genes

involved in synthesis of starch (see Supplementary Table S6). While the R Pool had only *beta-amylase*-encoding gene upregulated, *P. trifoliata* did not have any DEGs related to synthesis and degradation of starch (see Supplementary Table S6).

3.2.3.6.8. Transporters

The transport of substances was also one of the main biological functions affected by CLAs. The transcription levels of genes related to transporters were overwhelmingly affected by CLAs infection in all genotypes and hybrids (Fig. 3). In general, susceptible plants had the greatest number of transport-related genes affected by CLAs (Fig. 3). The R Pool showed few DEGs related to transport function, including *ABC transporter family*, *phosphate transporter (PHO1-2)*, and *amino acid transmembrane transport* (Supplementary Table 2). Zinc transporter (*ZIP1* and *ZIP8*) genes were differentially expressed in *C. sinensis*, *C. sunki*, and the T Pool (see Supplementary Table S6). Most transport family genes affected by CLAs infection were involved with transport of sugars, amino acids, and ions (see Supplementary Table S6). When comparing the transporter-related DEGs in the tolerant genotypes, *P. trifoliata*, and the T Pool, we observed different responses among them. The T Pool exhibited 73 differentially expressed transporter-related genes. The parental *P. trifoliata* showed only 4 differentially expressed transporter-related genes, among which *potassium transporter* was exclusively differentially expressed in *P. trifoliata* (Fig. 3 and Supplementary Table S6).

3.2.4. Discussion

The hybrids evaluated in this work and the parents, *Citrus sunki* and *P. trifoliata*, were classified as susceptible, tolerant, or resistant according to bacterial presence, callose deposition, and starch accumulation (Fig. 1). RNAseq data indicated that the genotypes responded differently under CLAs infection, which was confirmed by RT-qPCR analysis. Overall, the genes showed similar patterns in the RNAseq and RT-qPCR data, but some divergent values were found, which was similar to other transcriptome studies when the results of different techniques were compared²³.

Our findings indicated that few genes were differentially expressed according to RNAseq analysis of the tolerant and resistant plants. In contrast, RNAseq analysis of susceptible plants showed transcription modulation of many genes. Resistant and tolerant plants have a tendency to respond more rapidly and vigorously to a pathogen than susceptible plants¹². It is possible that the resistant hybrids have an early response to CLAs presence because early molecular interactions are well-known mechanisms in plant-pathogen interactions^{24,25,26}. Nevertheless, to

verify that the genetic responses were due to CLas infection and to avoid false positives, the samples for transcriptomic analysis were collected eight months after CLas infection.

P. trifoliata showed upregulation of receptor-related genes, which presented an efficient recognition of CLas and possibly an effective signaling and activation of defense response against CLas. The reprogramming of defense signaling pathways has previously been reported as a critical element of the early response to CLas in tolerant genotypes²⁷, such as *P. trifoliata*. Previous studies have also highlighted the induction of phenylpropanoid-related genes as a molecular mechanism of HLB tolerance⁵. Lignin-related genes and several phenylpropanoids were strongly upregulated in *P. trifoliata* transcriptome (Supplementary Table S6). As reorganization of plant growth and development are critical to maximize plant survival under stress²⁸, cell wall reinforcement is a tolerance mechanism of *P. trifoliata* against CLas. When comparing *P. trifoliata* and resistant hybrids, we observed a distinct transcriptional response to CLas (Fig. 2). However, all replicates of the resistant hybrids did not present any detection of CLas, even after almost one year of the experiment (Table 1 and 2), and probably for this reason, they exhibited few DEGs in RNAseq. Interestingly, the exclusive DEGs of the R Pool, formed by the CLas-negative hybrids, may be linked with genes and mechanisms capable of eliminating the bacteria from the plant, such as *endochitinase B*. Plant endochitinases cleave peptidoglycan chains, thereby promoting bacterial cell lysis¹³.

CLas infection is erratic and unpredictable, and even susceptible plants can escape from infection. Until almost one year, all plant replicates classified as resistant did not present CLas titer (Table 1 and 2). Therefore, until that moment, we considered that those plants were resistant to CLas infection and that a mechanism was utilized to avoid spreading the disease.

In the transcriptome of tolerant genotypes, downregulation of GA synthesis genes and upregulation of genes involved with GA degradation were observed, and the opposite behavior was observed in the susceptible genotypes (induction of GA synthesis and repression of GA degradation). In addition, we observed upregulation of several auxin-induced genes and repression of auxin responsive factors (Supplementary Table S6). It is known that the GA pathway presents cross-talk with auxin and ethylene hormones, which are plant growth regulators that also have been associated with plant defense and microbial pathogenesis^{29,30}. The present study showed that these regulators were strongly differentially expressed in the tolerant plants by CLas. It has been reported that auxin induces GA biosynthesis and suppresses GA degradation through modulation of several transcription factors and transporters^{31,32}. In citrus-pathogen interactions, crosstalk between auxin and GA has also been reported. Inhibition

of GA synthesis promotes inhibition of auxin-induced transcription, consequently reducing symptoms in the citrus-*Xanthomonas citri* interaction³³.

The plant tolerance mechanism is better explained by the interaction of GA and the salicylic acid (SA) hormone. The GA pathway is considered a hormone modulator of the SA signaling backbone during plant responses to pathogens^{34–36}. In *Arabidopsis thaliana*, Alonso-Ramírez et al. (2009)³⁶ showed that GAs and the overexpression of GA-responsive genes increase not only the endogenous levels of SA but also the expression of *ics1* and *npr1* genes involved in SA biosynthesis and action, respectively. However, SA-related genes were almost not modulated in the present study, which might be due to the high SA level in the evaluated stage, resulting in the expression of SA synthesis-related genes no longer being necessary as shown by Oliveira et al., 2019²⁰. Moreover, it is known that SA accumulation and downstream signaling events are important components of both pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI) and effector-triggered immunity (ETI)^{37,38} through increasing the expression of WRKY transcription factors. Many WRKY TFs were induced in the tolerant genotypes and affected in the susceptible plants (Fig. 5). WRKY TFs have been considered key regulators of plant defense against many pathogens, including CLas²⁷. The function of some WRKY genes remains unexplored, but in some crop species, specific WRKYs promote tolerance or even resistance to biotic and abiotic stresses²⁷. Thus, the induction of WRKY TFs may also be related to the activation of genes involved with the tolerance mechanism. For example, in *P. trifoliata*, the *WRKY transcription factor 14-1* was induced, and its orthologue in *Arabidopsis* (known as WRKY22) is an essential component of MAPK-mediated plant defense responses against pathogens. MAPKs are associated with one of the earliest signaling events after plant sensing of PAMPs and pathogen effectors.

Moreover, the tolerant and susceptible genotypes had changes in the level of transcription of many *callose synthases* and *phloem protein (PP2)* genes in response to CLas infection (Supplementary Table S6 and S7). Moreover, all susceptible plants showed induction of a class of genes that includes the *SEOc* gene (Supplementary Table S7). This class of genes has been reported to encode P-protein subunits²¹. Overexpression of these genes increases callose and PP2 protein synthesis in the citrus phloem sieve elements and leads to the callose and PP2 accumulation. Callose and PP2 accumulation is a crucial factor of phloem blockage in CLas-infected plants^{19,39,40}. Phloem blockage causes disturbance of photoassimilate flows from source organs (leaves) to sink organs (roots), resulting in starch accumulation in the leaves as observed in this work and in previous studies⁴¹.

Based on the knowledge of CLas-susceptible plant interaction that culminates in HLB symptoms, a zig-zag model as illustrated previously by Jones & Dang (2006)⁴² was adapted to explain such genetic molecular response to CLAs (Fig. 4). During the beginning of infection, receptors from citrus plants detect the CLAs PAMPs, which triggers a PTI response, resulting in the production of GA and SA as well as in the induction of several downstream genes (asymptomatic stage). In a second phase, CLAs delivers effectors, such as Las5315⁴³ and others⁴⁴, which interfere with PTI or enable pathogen nutrition and dispersal, resulting in effector-triggered susceptibility (ETS). In phase 3, effectors activate an ETI and an amplified version of PTI leading to induction of *callose synthases* and *pp2* gene expression that results in callose and PP2 accumulation. Therefore, callose and PP2 accumulation and the consequent anatomical alterations of the sieve pores may lead to hypersensitive cell death (HR) of the infected plants, which spatially isolate the CLAs to reduce their colonizing ability via the phloem^{19,40}.

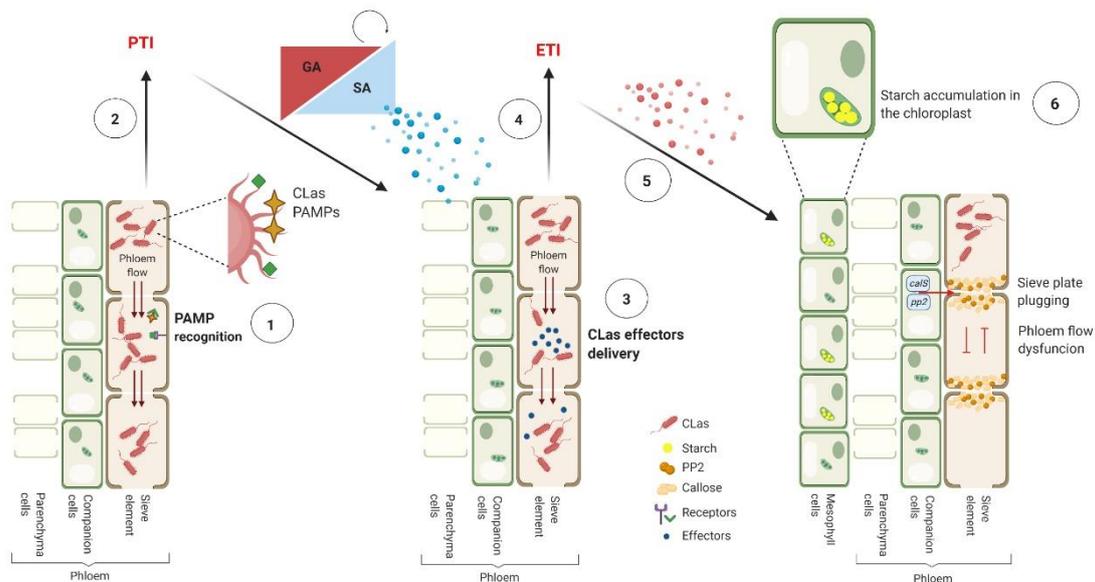


Fig.4. Defense response of susceptible genotypes against CLAs. In the phase 1 of this model, citrus plants receptors detect the CLAs PAMPs. In phase 2, a PAMP-triggered immunity (PTI) response is initiated, resulting in the production of gibberellic acid (GA), salicylic acid (SA) and the SA-dependent gene expression activation (in blue). In phase 3, CLas deliver effectors leading in effector-triggered susceptibility (ETS). In phase 4, effectors are recognized by plants proteins, activating effector-triggered immunity (ETI). In phase 5, ETI triggers a series of genetic events (in red), including the induction of *calloses synthases* and *pp2* expression. This exaggerated response could be considered as hypersensitive cell death (HR), since the attempt to isolate spatially the CLas leading to callose and PP2 accumulation, that cause phloem dysfunctions. The phase 6 represents the starch accumulation in the mesophyll chloroplasts (Created with BioRender.com).

To describe the genetic mechanisms potentially involved in a susceptible, tolerant, and resistant interaction with CLAs based on the data obtained in this study, we built a hypothetical model (Fig. 5). The model shows that in the susceptible plants (Fig. 5), auxin-related genes positively modulate GA synthesis, which activates response mechanisms to CLAs infection, such as callose deposition, PP2 deposition, phloem dysfunction, and impaired flow transport. The impaired flow results in starch accumulation on mesophyll chloroplasts, which promotes thylakoid rupture and chlorophyll degradation, culminating in HLB typical symptoms. In the tolerant plants, including *P. trifoliata* (Fig. 5), the induction of signaling receptors cause a fast and efficient defense response modulated by suppression of the auxin pathway and induction of GA degradation. The suppression of these pathways prevents the events that lead to phloem dysfunction (callose deposition, starch accumulation, and transport alteration), and it activates the defense response through the synthesis of phenylpropanoids and cell wall strengthened-related genes. This transcriptional reprogramming is efficient to impair the development of symptoms. In the resistant genotypes (Fig. 5), a potentially early and rapid defense may occur in response to CLAs because only a few genes were differentially expressed after 240 days after inoculation. However, this response is related to induction of signaling receptors and upregulation of *endochitinase B*, which is associated with bacterial cell lysis.

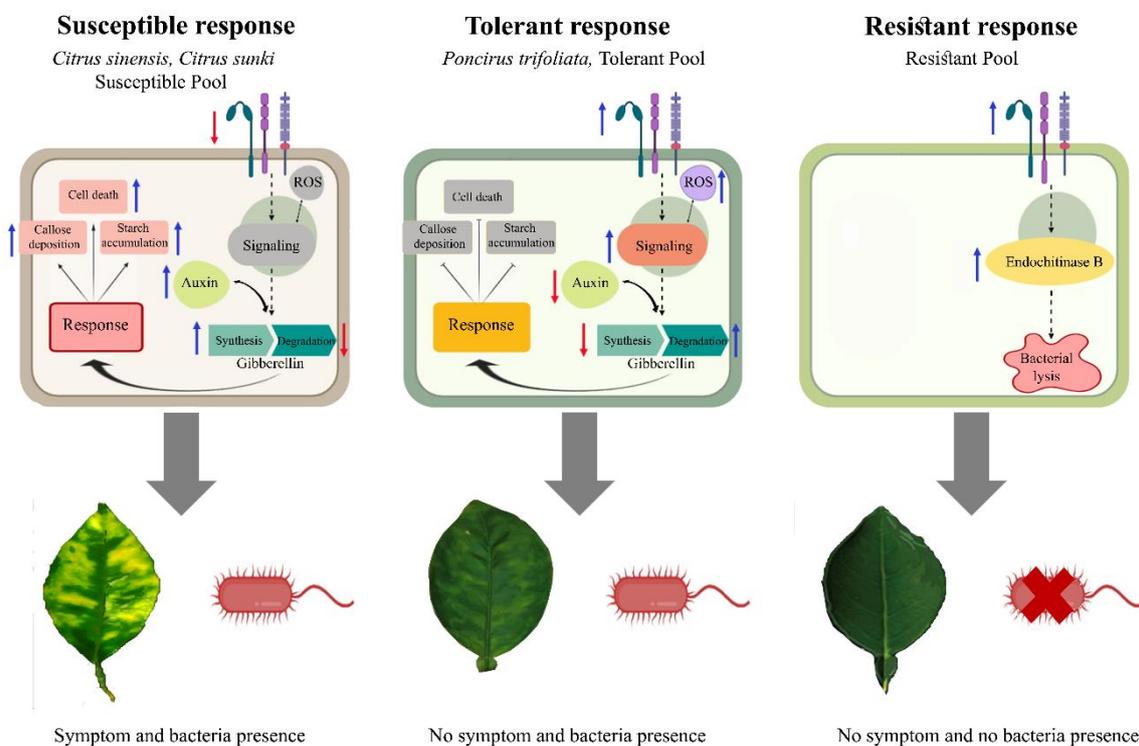


Fig. 5. Model of interaction between CLAs and Citrus plants. Susceptible plants, the downregulation of signaling receptors promotes a late recognition of CLAs infection, and,

consequently, no proper signaling is activated. Auxin-related genes positively modulate the gibberellin synthesis, which activates response mechanisms to CLAs infection, such as callose and PP2 deposition and impaired substances transport. Interference on substance transport along with callose deposition causes phloem dysfunction resulting in starch accumulation on photosynthetic tissues. Starch accumulation promotes thylakoid rupture and chlorophyll degradation culminating in HLB classical symptoms. Tolerant plants, the induction of signaling receptors causes a fast and efficient defense response modulated by suppression of auxin pathway and induction of GA degradation. The suppression of these pathways prevents the events that lead to the phloem dysfunction (callose deposition, starch accumulation and transport alteration) and activates defense response through the synthesis of phenylpropanoids and cell wall-strengthened related genes. This transcriptional reprogramming is efficient to impair the development of symptoms. Resistant genotypes, a possibly early and fast defense may occur in response to CLAs, since low numbers of the genes are modulated after 240 days post inoculation. Nonetheless, this response is related to induction of signaling receptors and upregulation of *Endochitinase B*, which might be associated with bacterial cell lysis (Created with BioRender.com).

Both hypothetical models showed that there are many pathways acting in citrus defense against CLAs infection. The data acquired in this study can help to generate citrus varieties of scions or rootstocks with potential resistance to HLB based on citrus conventional breeding programs or biotechnological approaches, including the development of transgenic or cisgenic lines as well as genome editing and host-induced gene silencing.

3.2.5. Materials and Methods

3.2.5.1. Plant material

C. sinensis, *C. sunki*, *P. trifoliata*, and 21 hybrids obtained from a controlled cross between *Citrus sunki* ex Tan (female parent and susceptible to HLB) and *Poncirus trifoliata* Raf. cv Rubidoux (male parent and tolerant to HLB) were used in the analysis. *C. sinensis* was included because it is one of most important citrus scions in the world, and it can also be considered an internal control of the experiment considering that *C. sinensis* is characterized as a species highly susceptible to HLB (Boava et al., 2017)³⁹. The experimental design was completely randomized and consisted of five biological replicates for each inoculated genotype (CLAs-infected budwoods) and mock-inoculated genotype (health budwoods). Plants were propagated

using buds that were grafted onto rootstocks of Rangpur lime (*C. limonia* Osb.). At the end of six months, the plant scions were grafted using two CLAs-infected buds obtained from *C. sinensis* (L.) Osbeck cv Pera. All plants were kept in a greenhouse at Centro de Citricultura Sylvio Moreira of the Agronomic Institute (IAC), SP with an average temperature of 25°C for 12 months. The starch content and callose deposition were estimated only in the genotypes selected for the further analysis (*C. sinensis*, *C. sunki*, *P. trifoliata*, and 15 hybrids obtained from crosses between *C. sunki* and *P. trifoliata*). Leaves from inoculated and mock-inoculated plants from all evaluated genotypes were collected with three biological replicates of each genotype after eight months of CLAs infection.

3.2.5.2. CLAs quantification

CLAs presence and HLB symptoms were evaluated according to previously described methodology⁴⁰. Briefly, 30, 90, 180, 240, and 360 days after inoculation, to confirm HLB infection, leaves above the inoculation point were collected and tested by qPCR using 16S ribosomal DNA primer sets and FAM/Iowa Black FQ label probe (IDT Inc., Coralville, IA) probes as described by Li et al. (2006)⁴⁵. Citrus GAPDH (glyceraldehyde 3- phosphate dehydrogenase F: GGAAGGTCAAGATCGGAATCAA; R: CGTCCCTCTGCAAGATGACTCT) was used as the reference gene. Values above 34 Ct were considered negative for CLAs infection⁷. After 240 days of CLAs inoculations, the bacterial titer was evaluated according to Boava et al. (2015)⁷ by qPCR using a standard curve with 10-fold serial dilutions of 16S ribosomal DNA (rDNA) cloning into pGEM-T vector (PROMEGA).

3.2.5.3. Phenotypic analysis

Starch and callose quantification of CLAs-inoculated and mock-inoculated plants was performed after 240 days of infection. Callose quantification was performed following the methodology reported previously⁴⁰. Leaf petioles were fixed in FAA solution (50 mL of formaldehyde, 50 mL of glacial acetic acid, and 900 mL of 70% ethanol) for 72 h and then kept in 70% ethanol. Transversal sections of 10 µm were generated using an automatic slide microtome (Leica SM2010R). The sections were stained with blue aniline, and the stained samples were examined on an Olympus BX61 fluorescence microscope using 355–375 nm excitation filter, 400-nm dichromatic mirror, and 435–490 nm emission filter. Callose quantification was performed by counting fluorescent spots in the total phloem area in 10 fields of view for each sample. The starch measurement was performed using leaves dried in an oven at 60°C for 48 h and ground. Starch content was estimated by enzymatic analysis using 10 mg

of dried leaves according to ⁴⁶. Absorbance was measured in 96-well microtiter plates using a Microplate Reader (Model 3550 – BIO-RAD) at 490 nm. A standard curve was performed using a glucose solution (SIGMA) at concentrations of 0, 2.5, 5.0, 7.5, and 10 µg/mL.

According to starch, callose, and CLas quantification, the genotypes were classified as susceptible, tolerant, and resistant (see supplementary Fig. S3).

3.2.5.4. RNA extraction and sequencing (RNA-seq)

Leaves from three biological replicates of the three genotypes (*C. sinensis*, *C. sunki*, and *P. trifoliata*) and the three hybrid pools (S Pool: H109, H161, and H165; T Pool: H113, H154, and H146; and R Pool: H68, H106, and H142), either CLas-infected inoculated or mock-inoculated plants, were collected for transcriptomic analysis after 240 days of infection. It is difficult to establish the ideal time for studying the first responses and stages of infection because it is difficult to confirm that the plant tissue is colonized by bacteria. Thus, to verify that the genetic responses were due to CLas infection, we performed RNA-seq analysis at eight months. Total RNA was isolated with the MasterPure Plant RNA Purification Kit (EPICENTRE Biotechnologies, Madison, WI, USA) according to the manufacturer's instructions. A total of 10 µg of RNA from each sample was sent for sequencing at the Centro de Genômica Funcional in Centro de Biotecnologia Agrícola in ESALQ/USP (<http://www.esalq.usp.br/genomicafuncional/>). RNA-seq was performed using the Illumina HiSeq 2500 platform. All procedures were performed according to Illumina's protocols. RNA-seq was performed in triplicate with a total of 36 samples.

3.2.5.5. Data analysis

The quality of obtained fragments from the sequencing was verified using CLC Genomics Workbench v.6 program (CLC BIO) software (<https://www.qiagenbioinformatics.com/products/clc-genomics-workbench/>). The sequences were trimmed using the trimmomatic tool ⁴⁷ and mapped on the v 2.0 *C. sinensis* genome (<http://citrus.hzau.edu.cn/>) using the STAR-2.5.2b program ⁴⁸. The R subread package was used for counting. DEGs between the control and CLas-infected plants were established using the DESeq in Bioconductor package ⁴⁹ using an adjusted *p-value* of 0.005 and FDR threshold of 0.05. Venn diagrams (<http://bioinformatics.psb.ugent.be/webtools/Venn/>) were used to identify common and unique DEGs among the analyzed genotypes. We used Blast2Go ⁹ for functional categorization, and the DEGs were annotated by Gene Ontology (GO) using default parameters

3.2.5.6. Real time PCR (RT-qPCR) validation

To ensure reproducibility of the biological phenomenon observed by transcriptomic analysis, we performed a second experiment with other plants following the same design used for RNA-seq. We sampled one hybrid of each pool to represent the susceptible, tolerant, and resistant pools. We used only one hybrid from each pool because it represents the hybrids that comprise each pool regarding CLas infection behavior. Total RNA was extracted using the protocol described by Chang et al. (1993)⁵⁰. Traces of genomic DNA were eliminated using the DNase RNase-Free Kit (QIAGEN, Valencia, CA, USA) according to the manufacturer's instructions. cDNAs were synthesized from 1.0 µg of total RNA using Superscript III (200 U/µL) (INVITROGEN) with an oligo (dT) primer (dT12-18, INVITROGEN) according to the manufacturer's instructions. cDNAs were treated with RNase H (1 U) for 20 min at 37°C to remove any contaminating RNA.

Ten genes that showed the opposite expression profile between the genotypes with different responses were selected, including *chalcone synthase*, *lipid transfer*, *cytochrome P450*, *gibberellin-regulated 9*, *sieve element occlusion c*, *cinnamoyl-reductase*, *pectin methylesterase 1*, *starch branching enzyme II*, *PRR response regulator*, and *choline transporter like-protein 2* (see Supplementary Table S10). Primers were designed using Primer3Plus⁵¹, and the Primer-BLAST tool⁵² was used to check the specificity of the primers. Two endogenous genes, GAPDH and FBOX, were used for normalization of the data. Relative gene expression was calculated with the $2^{-\Delta\Delta C_t}$ method⁵³.

3.2.6. Supplementary Information

Acknowledgments

This study was supported by INCT-Citros (FAPESP 2014/50880-0 and CNPq 465440/2014-2); PhD fellowships to ISP (FAPESP 2015/13971-0) and MC (FAPESP 2016/22133-0); and postdoc fellowships to DMG (CNPq 103228/2018-7) and LMG (FAPESP 2019/01901-8).

Authors' contributions

MCY and MAM planned and supervised the study. MC and ISP contributed to the design and execution of the experiments detailed. MC and MAT performed the functional genomic and bioinformatics data analyses. MC, LPB, LMG, and DMG conducted and evaluated plant growth. MC and ISP drafted the manuscript. MCY, LMG, DMG, MAT, AADS, and MAM provided intellectual input. All authors have read and approved the final manuscript.

Conflict of Interest Statement

The authors declare that they have no competing interests.

3.2.7. References

1. Bové, J. M. & Ayres, A. J. Etiology of three recent diseases of citrus in São Paulo State: sudden death, variegated chlorosis and huanglongbing. *IUBMB Life* **59**, 346–354 (2007).
2. Munir, S. *et al.* Huanglongbing Control: Perhaps the End of the Beginning. *Microb. Ecol.* **76**, 192–204 (2018).
3. Johnson, E. G., Wu, J., Bright, D. B. & Graham, J. H. Association of ‘*Candidatus Liberibacter asiaticus*’ root infection, but not phloem plugging with root loss on huanglongbing-affected trees prior to appearance of foliar symptoms. *Plant Pathol.* **63**, 290–298 (2014).
4. Albrecht, U. & Bowman, K. D. Transcriptional response of susceptible and tolerant citrus to infection with *Candidatus Liberibacter asiaticus*. *Plant Sci.* **185–186**, 118–130 (2012).
5. Balan, B., Ibáñez, A. M., Dandekar, A. M., Caruso, T. & Martinelli, F. Identifying host molecular features strongly linked with responses to huanglongbing disease in citrus leaves. *Front. Plant Sci.* **9**, 1–13 (2018).
6. Folimonova, S. Y., Robertson, C. J., Garnsey, S. M., Gowda, S. & Dawson, W. O. Examination of the responses of different genotypes of citrus to huanglongbing (Citrus Greening) under different conditions. *Phytopathology* **99**, 1346–1354 (2009).
7. Boava, L. P., Sagawa, C. H. D., Cristofani-Yaly, M. & Machado, M. A. Incidence of ‘*Candidatus Liberibacter asiaticus*’-infected plants among citrandarins as rootstock and scion under field conditions. *Phytopathology* **105**, 518–524 (2015).
8. Albrecht, U. & Bowman, K. D. Tolerance of the trifoliolate citrus hybrid US-897 (Citrus reticulata Blanco × poncirus trifoliata L. Raf.) to huanglongbing. *HortScience* **46**, 16–22 (2011).
9. Conesa, A. & Götz, S. Blast2GO: A comprehensive suite for functional analysis in plant genomics. *Int. J. Plant Genomics* **2008**, (2008).
10. Consortium, T. G. O. Gene Ontology : tool for the unification of biology. *Nature* **25**, 25–29 (2000).
11. Villena, J., Kitazawa, H., Van Wees, S. C. M., Pieterse, C. M. J. & Takahashi, H. Receptors and Signaling Pathways for Recognition of Bacteria in Livestock and Crops: Prospects for Beneficial Microbes in Healthy Growth Strategies. *Front. Immunol.* **9**, 2223 (2018).
12. Andersen, E. J., Ali, S., Byamukama, E., Yen, Y. & Nepal, M. P. Disease resistance

- mechanisms in plants. *Genes (Basel)*. **9**, (2018).
13. De Medeiros, S. C., Monteiro-Júnior, J. E., Passos Sales, G. W., Grangeiro, T. B. & Pinto Nogueira, N. A. Chitinases as antibacterial proteins: A systematic review. *J. Young Pharm.* **10**, 144–148 (2018).
 14. Finkina, E. I., Melnikova, D. N., Bogdanov, I. V. & Ovchinnikova, T. V. Lipid transfer proteins as components of the plant innate immune system: Structure, functions, and applications. *Acta Naturae* **8**, 47–61 (2016).
 15. Rawat, N. *et al.* Genome resequencing and transcriptome profiling reveal structural diversity and expression patterns of constitutive disease resistance genes in Huanglongbing-tolerant *Poncirus trifoliata* and its hybrids. *Hortic. Res.* **4**, 1–8 (2017).
 16. Bartwal, A., Mall, R., Lohani, P., Guru, S. K. & Arora, S. Role of Secondary Metabolites and Brassinosteroids in Plant Defense Against Environmental Stresses. *J. Plant Growth Regul.* **32**, 216–232 (2013).
 17. Bateman, A. UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515 (2019).
 18. Caffall, K. H. & Mohnen, D. The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydr. Res.* **344**, 1879–1900 (2009).
 19. Granato, L., Galdeano, D., D’Alessandre, N., Breton, M. & Machado, M. *Callose synthase* family genes plays an important role in the Citrus defense response to *Candidatus Liberibacter asiaticus*. *Eur. J. Plant Pathol.* **155**, 25–38 (2019).
 20. Oliveira, T. S. *et al.* Genetic analysis of salicylic acid-mediated defenses responses and histopathology in the huanglongbing pathosystem. *Citrus Res. Technol.* **40**, 1–13 (2019).
 21. Ernst, A. M. *et al.* Sieve element occlusion (SEO) genes encode structural phloem proteins involved in wound sealing of the phloem. *Proc. Natl. Acad. Sci. U. S. A.* **109**, (2012).
 22. Lloyd, J. R., Kossmann, J. & Ritte, G. Leaf starch degradation comes out of the shadows. *Trends Plant Sci.* **10**, 130–137 (2005).
 23. Wang, Y., Zhou, L., Yu, X., Stover, E. & Luo, F. Transcriptome Profiling of Huanglongbing (HLB) Tolerant and Susceptible Citrus Plants Reveals the Role of Basal Resistance in HLB Tolerance. *Front. Plant Sci.* **7**, 1–13 (2016).
 24. Liu, T. *et al.* Genome-wide identification, classification and expression analysis in fungal–plant interactions of cutinase gene family and functional analysis of a putative CICUT7 in *Curvularia lunata*. *Mol. Genet. Genomics* **291**, 1105–1115 (2016).
 25. Shen, Y. *et al.* The early response during the interaction of fungal phytopathogen and host plant. *Open Biol.* **7**, (2017).

26. Nirmala, J. *et al.* Concerted action of two avirulent spore effectors activates Reaction to *Puccinia graminis* 1 (rpg1)-mediated cereal stem rust resistance. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 14676–14681 (2011).
27. Yu, Q. *et al.* Reprogramming of a defense signaling pathway in rough lemon and sweet orange is a critical element of the early response to *Candidatus Liberibacter asiaticus*. *Hortic. Res.* **4**, 1–15 (2017).
28. Iglesias, M. J., Terrile, M. C., Bartoli, C. G., D’Ippólito, S. & Casalongué, C. A. Auxin signaling participates in the adaptative response against oxidative stress and salinity by interacting with redox metabolism in *Arabidopsis*. *Plant Mol. Biol.* **74**, 215–222 (2010).
29. Kazan, K. & Manners, J. M. Linking development to defense: auxin in plant-pathogen interactions. *Trends Plant Sci.* **14**, 373–382 (2009).
30. Robert-Seilaniantz, A., Grant, M. & Jones, J. D. G. Hormone Crosstalk in Plant Disease and Defense: More Than Just JASMONATE-SALICYLATE Antagonism. *Annu. Rev. Phytopathol.* **49**, 317–343 (2011).
31. Björklund, S., Antti, H., Uddestrand, I., Moritz, T. & Sundberg, B. Cross-talk between gibberellin and auxin in development of *Populus* wood: Gibberellin stimulates polar auxin transport and has a common transcriptome with auxin. *Plant J.* **52**, 499–511 (2007).
32. Richter, R., Behringer, C., Zourelidou, M. & Schwechheimer, C. Convergence of auxin and gibberellin signaling on the regulation of the GATA transcription factors GNC and GNL in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 13192–13197 (2013).
33. Cernadas, R. A. & Benedetti, C. E. Role of auxin and gibberellin in citrus canker development and in the transcriptional control of cell-wall remodeling genes modulated by *Xanthomonas axonopodis* pv. *citri*. *Plant Sci.* **177**, 190–195 (2009).
34. Pieterse, C. M. J., Van der Does, D., Zamioudis, C., Leon-Reyes, A. & Van Wees, S. C. M. Hormonal Modulation of Plant Immunity. *Annu. Rev. Cell Dev. Biol.* **28**, 489–521 (2012).
35. De Bruyne, L., Höfte, M. & De Vleeschauwer, D. Connecting growth and defense: The emerging roles of brassinosteroids and gibberellins in plant innate immunity. *Mol. Plant* **7**, 943–959 (2014).
36. Alonso-Ramírez, A. *et al.* Cross-talk between gibberellins and salicylic acid in early stress responses in *Arabidopsis thaliana* seeds. *Plant Signal. Behav.* **4**, 750–751 (2009).
37. Li, J. *et al.* ‘*Candidatus Liberibacter asiaticus*’ Encodes a Functional Salicylic Acid (SA) Hydroxylase That Degrades SA to Suppress Plant Defenses. *Mol. Plant-Microbe Interact.* **30**, 620–630 (2017).
38. An, C. & Mou, Z. Salicylic Acid and its Function in Plant Immunity. *J. Integr. Plant*

- Biol.* **53**, 412–428 (2011).
39. Xie, B. & Hong, Z. Unplugging the callose plug from sieve pores. *Plant Signal. Behav.* **6**, 491–493 (2011).
 40. Boava, L. P., Cristofani-Yaly, M. & Machado, M. A. Physiologic, anatomic, and gene expression changes in *citrus sunki*, *poncirus trifoliata*, and their hybrids after ‘*candidatus liberibacter asiaticus*’ infection. *Phytopathology* **107**, 590–599 (2017).
 41. Etxeberria, E., Gonzalez, P., Achor, D. & Albrigo, G. Anatomical distribution of abnormally high levels of starch in HLB-affected Valencia orange trees. *Physiol. Mol. Plant Pathol.* **74**, 76–83 (2009).
 42. Jones, J. D. G. & Dangl, J. L. The plant immune system. *Nature* **444**, 323–9 (2006).
 43. Pitino, M., Allen, V. & Duan, Y. Las Δ 5315 Effector Induces Extreme Starch Accumulation and Chlorosis as *Ca. Liberibacter asiaticus* Infection in *Nicotiana benthamiana*. *Front. Plant Sci.* **9**, 1–11 (2018).
 44. Granato, L. M. *et al.* ‘*Candidatus Liberibacter asiaticus*’ putative effectors : in silico analysis and gene expression in citrus leaves displaying distinct huanglongbing symptoms. *Trop. Plant Pathol.* (2020).
 45. Li, W., Hartung, J. S. & Levy, L. Quantitative real-time PCR for detection and identification of *Candidatus Liberibacter* species associated with citrus huanglongbing. *J. Microbiol. Methods* **66**, 104–15 (2006).
 46. Amaral, L., Gaspar, M., Costa, P., Aidar, M. & Buckeridge, M. Novo método enzimático rápido e sensível de extração e dosagem de amido em materiais vegetais. *Hoehnea* **34**, 425–431 (2007).
 47. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
 48. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
 49. Anders, S. & Wolfgang, H. Differential expression and sequence-specific interaction of karyopherin α with nuclear localization sequences. *Genome Biol.* **11**, (2010).
 50. Chang, S., Puryear, J. & Cairney, J. A Simple and Efficient Method for Isolating RNA from Pine Trees. *Plant Mol. Biol. Report.* **11**, 113–116 (1993).
 51. Untergasser, A. *et al.* Primer3-new capabilities and interfaces. *Nucleic Acids Res.* **40**, 1–12 (2012).
 52. Ye, J. *et al.* Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* **13**, 134 (2012).

53. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and. *Methods* **25**, 402–408 (2001).

3.2.8. Supplementary Material

Supplementary Table S1: RNA-seq reads and mapping information, C: mock inoculated samples, I: CLas-inoculated samples

Genotype	HLB	Treatment	Total reads	Unique mapped reads	% of Unmapped Reads
<i>C. sinensis</i>	Susceptible	C	39,463,048	36,194,913	2.71
		I	45,000,006	40,458,529	3.01
<i>C. sunki</i>	Susceptible	C	39,075,570	34,672,395	3.87
		I	45,647,733	39,644,055	4.23
<i>P. trifoliata</i>	Tolerant	C	46,269,608	39,685,082	3.86
		I	41,622,797	34,541,507	4.49
Pool S	Susceptible	C	36,266,233	31,762,851	4.04
		I	41,328,832	36,650,534	3.81
Pool T	Tolerant	C	38,072,190	34,012,956	3.72
		I	35,495,818	31,859,476	3.54
Pool R	Resistant	C	39,059,147	34,931,327	3.72
		I	39,707,545	34,582,737	4.43

Supplementary Table S2: Differentially expressed genes among the genotypes

This material is online available at: <https://www.nature.com/articles/s41598-020-77840-2>

Supplementary Table S3: Genes downregulated in *P. trifoliata*, which were upregulated in other genotypes

Gene description	Gene ID	Genotype
Putative uncharacterized protein Sb01g047790	orange1.1t00904	<i>C. sinensis</i> <i>C. sunki</i>
<i>flavonoid 3 -monooxygenase-like</i>	Cs3g05810	
<i>nucleic acid-binding</i>	Cs4g09300	
<i>lachrymatory-factor synthase-like</i>	orange1.1t03813	

<i>Chalcone synthase</i>	Cs2g14720	
<i>dihydrofolate reductase</i>	Cs6g16160	<i>C. sinensis</i>
<i>MYB transcription factor MYB128</i>	Cs4g13690	
<i>unnamed protein product</i>	Cs4g02690	
<i>Licodione synthase</i>	Cs5g18660	
<i>bifunctional 3-dehydroquinate dehydratase shikimate chloroplastic-like isoform XI</i>	Cs5g32370	
<i>kinase 2B</i>	Cs3g13410	<i>C. sunki</i>
<i>Serine carboxypeptidase-like 18</i>	Cs8g03880	
<i>glutathione S-transferase U8</i>	Cs6g07260	
<i>basic leucine zipper 61</i>	Cs2g15930	
<i>MLP-like protein 423</i>	Cs7g08260	S Pool

Supplementary Table S4: Genes up-regulated in *P. trifoliata* which were downregulated in other genotypes

Gene description	Gene ID	Genotype
<i>serine threonine kinase</i>	Cs5g14090	<i>C. sinensis</i> S Pool
uncharacterized protein LOC100777990	Cs3g24260	<i>C. sunki</i> S Pool
<i>class IV chitinase</i>	orange1.1t03118	<i>C. sunki</i>
<i>White-brown-complex ABC transporter family</i>	orange1.1t01993	
<i>Reticuline oxidase</i>	orange1.1t01957	
<i>basic chitinase</i>	Cs5g21860	
<i>Cysteine-rich receptor kinase</i>	Cs2g07450	S Pool
<i>Reticuline oxidase</i>	Cs2g10150	
<i>probable glutathione S-transferase</i>	orange1.1t03629	
<i>amino-acid permease BAT1 homolog isoform XI</i>	Cs3g20130	<i>C. sinensis</i>
<i>geraniol 8-hydroxylase</i>	Cs8g09390	

Supplementary Table S5: Genes upregulated in Pool R which were downregulated in other genotypes

Gene description	Gene ID	Genotype
<i>Quinone oxidoreductase</i>	orange1.1t00259	<i>C. sinensis</i>
<i>cycloartenol synthase</i>	Cs4g04730	

<i>hypothetical protein VITISV_011279</i>	Cs1g15880	
<i>hypothetical protein VITISV_000078</i>	orange1.1t03893	
<i>sterol regulatory element-binding site 2 protease</i>	Cs8g03710	
<i>uncharacterized protein LOC100795901 precursor</i>	Cs1g07510	
<i>cycloartenol synthase</i>	Cs4g04680	
<i>rubber peroxidase 1</i>	orange1.1t02045	T Pool
<i>class IV chitinase</i>	orange1.1t03118	<i>C. sunki</i>

Supplementary Table S6: Differentially expressed genes in *C. sinensis*, *C. sunki*, *P. trifoliata*, Susceptible Pool, Tolerant Pool and Resistant Pool

This material is online available at: <https://www.nature.com/articles/s41598-020-77840-2>

Supplementary Table S7. Differentially expressed phloem related in the *C. sinensis*, *C. sunki*, S Pool, T Pool and *P. trifoliata*. ID gene: access number on *C. sinensis* genome.

Genotype	DGEs	ID gene	log2FoldChange
<i>C. sinensis</i>	<i>Callose synthase 5</i>	Cs1g05830	-1.57
	<i>Plasmodesmata Callose-Binding Protein 3</i>	Cs5g11770	0.89
	<i>PP2-B1</i>	orange1.1t04174	-1.79
	<i>PP2-B10</i>	Cs2g10930	0.94
	<i>PP2-A13</i>	orange1.1t00304	1.47
	<i>PP2-B1</i>	Cs9g10910	1.73
	<i>PP2-A12</i>	Cs5g10330	2.31
	<i>PP2-B15</i>	Cs3g14740	3.38
	<i>PP2-B15</i>	Cs3g14720	3.40
	<i>PP2-B15</i>	Cs3g14680	7.49
	<i>PP2-like A1</i>	Cs7g16020	1.11
	<i>PP2-like A2</i>	Cs2g10920	3.33
	<i>PP2-like B13</i>	Cs3g14690	4.79
	<i>Sieve element occlusion c</i>	Cs5g11280	1.18
	<i>Sieve element occlusion c</i>	Cs5g06490	2.68
	<i>Sieve element occlusion d</i>	Cs7g09710	3.69
	<i>Sieve element occlusion d</i>	Cs2g26900	4.23
<i>C. sunki</i>	<i>Callose synthase 2</i>	Cs7g01200	-1.59
	<i>Callose synthase 3</i>	orange1.1t02029	1.24
	<i>PP2-B1</i>	orange1.1t04174	-2.60
	<i>PP2-A12</i>	Cs5g10330	0.82
	<i>PP2-B10</i>	Cs2g10930	0.91
	<i>PP2-A13</i>	Cs9g16920	1.37
	<i>PP2-A13</i>	orange1.1t00304	1.56

	<i>Sieve element occlusion c</i>	Cs5g11280	1.48
	<i>Sieve element occlusion c</i>	Cs5g06500	1.96
	<i>Sieve element occlusion c</i>	Cs5g06490	2.27
	<i>Sieve element occlusion d</i>	Cs7g09710	2.94
	<i>Sieve element occlusion d</i>	Cs2g26900	3.43
S Pool	<i>PP2-B1</i>	orange1.1t03219	-2.06
	<i>Sieve element occlusion c</i>	Cs5g11280	1.75
T Pool	<i>Callose synthase 12</i>	orange1.1t00806	1.77
	<i>PP2-A13</i>	Cs9g16920	-0.83
	<i>PP2-A12</i>	Cs5g10330	1.42
	<i>PP2-B15</i>	Cs3g14680	7.40
	<i>PP2-like B13</i>	Cs3g14690	5.52
	<i>Sieve element occlusion c</i>	Cs5g06500	1.82
	<i>Sieve element occlusion d</i>	Cs7g09710	2.13
	<i>Sieve element occlusion c</i>	Cs5g06490	2.19
<i>P. trifoliata</i>	<i>PP2-B15</i>	Cs3g14680	9.81

Supplementary Table. S8. Differentially expressed related with starch synthesis in *C. sinensis*, *C. sunki*. ID gene: access number on *C. sinensis* genome.

Genotype	DGEs	ID gene	log2Fold Change
<i>C. sinensis</i>	<i>ADP-glucose pyrophosphorylase small subunit</i>	Cs2g18800	1.02
	<i>Starch branching enzyme II</i>	Cs6g15320	1.8
<i>C. sunki</i>	<i>ADP-glucose pyrophosphorylase family</i>	Cs5g04870	1.17
	<i>Starch branching enzyme II</i>	Cs6g15320	0.5

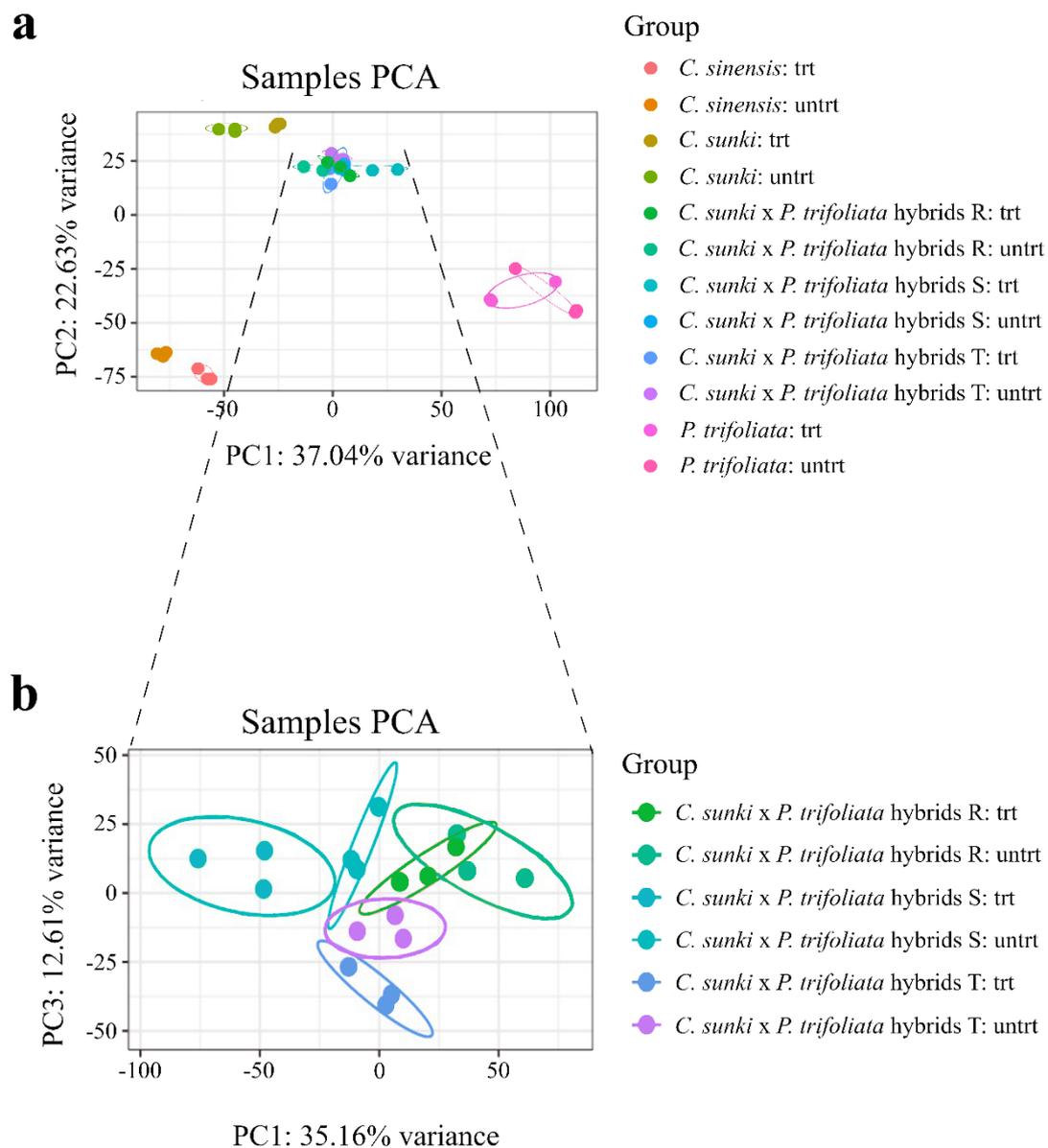
Supplementary Table. S9. Differentially expressed related with starch degradation in *C. sinensis*, *C. sunki*, S Pool, T Pool and R Pool. ID gene: access number on *C. sinensis* genome.

Genotype	DGEs	ID gene	log2Fold Change
<i>C. sinensis</i>	<i>Beta-amylase Family</i>	Cs5g07550	-3.9
	<i>Beta-amylase 7-like</i>	orange1.1t00361	-0.96
	<i>Inactive beta-amylase 9</i>	Cs9g04980	-1.29
	<i>Alpha-amylase 1 large isoform</i>	Cs3g26820	1.25
	<i>Alpha amylase domain</i>	Cs3g23560	2.67
	<i>Beta-amylase</i>	Cs2g22040	3.29
<i>C. sunki</i>	<i>Beta-amylase chloroplast-like</i>	orange1.1t03470	-2.57
	<i>Beta-amylase Family</i>	Cs5g07550	-2.37
	<i>Inactive beta-amylase 9</i>	Cs9g04980	-0.95
	<i>Alpha-amylase 1 large isoform</i>	Cs3g26820	-0.87
	<i>Alpha-amylase chloroplast-like</i>	Cs7g04310	-0.65
	<i>Alpha amylase domain</i>	Cs3g23560	2.17
	<i>Beta-amylase</i>	Cs2g22040	2.21

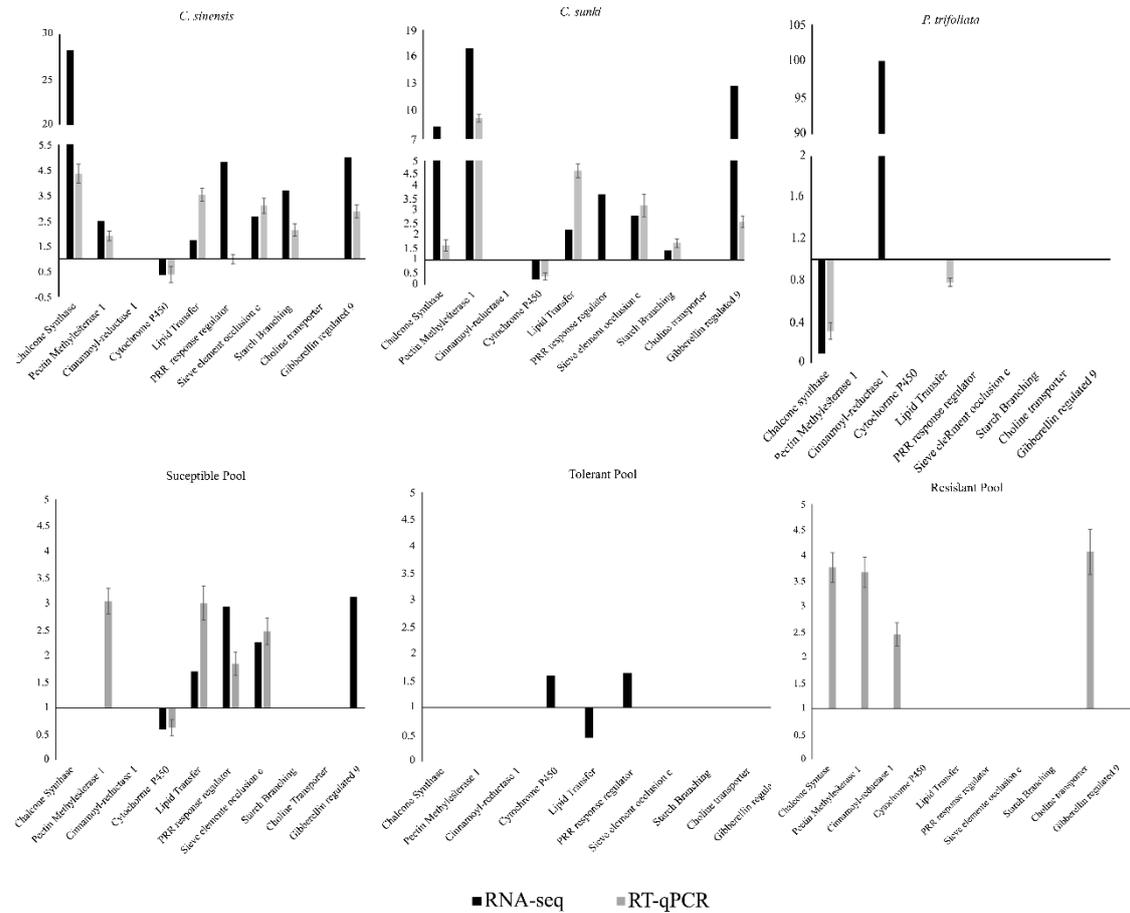
S Pool	<i>Beta-amylase activity gene</i>	Cs5g07550	-1.27
	<i>Beta-amylase Family</i>	Cs5g07550	-1.77
T Pool	<i>Beta-amylase chloroplastic-like</i>	orange1.1t03470	-0.90
	<i>Alpha amylase domain</i>	Cs3g23560	2.89
R Pool	<i>Beta-amylase</i>	Cs2g22040	2.61

Supplementary Table. S10. Primers designed and used for real-time PCR amplification

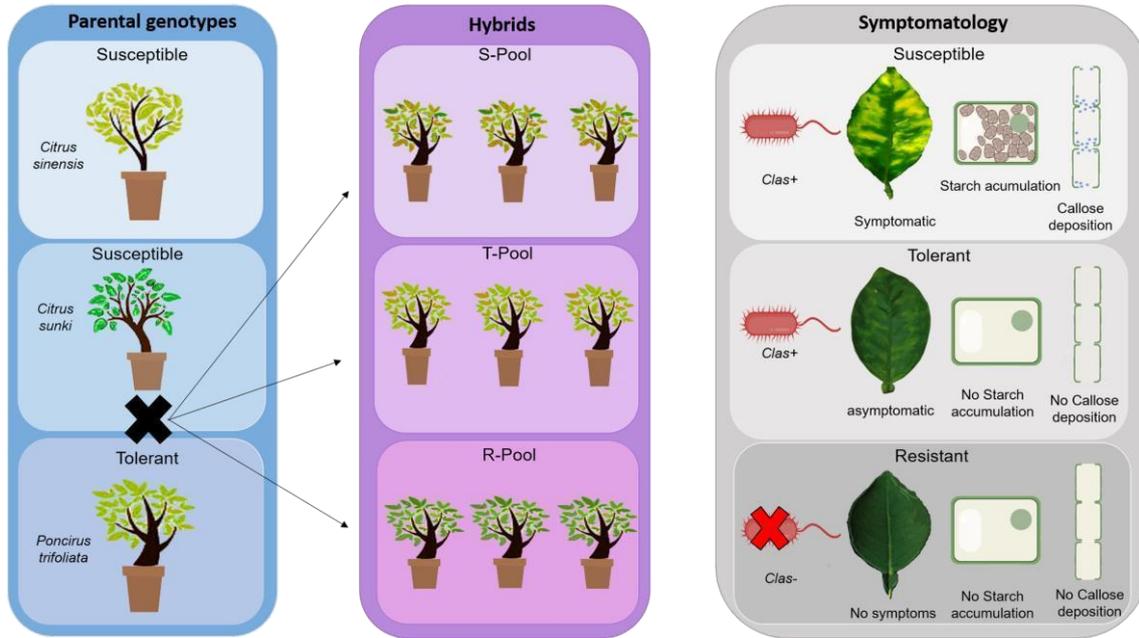
Primers			
<i>Chalcone Synthase</i>	F	TCGCCTCGCTAAAGACTTGG	
	R	ACCATCACCGAACAAAGCCT	
<i>Lipid Transfer</i>	F	AACCAAGCAAAAGCCTCCCT	
	R	AACGCCCTCCAGTTCTCAAG	
<i>Cytochrome P450 71A26-Like</i>	F	GATGATGGAGGCAGTGCAGA	
	R	GCAATGGAAGTGGTGGGTGA	
<i>Gibberellin Regulated 9</i>	F	CCTGCAGTTTCGATTCACAA	
	R	GTGCCTGCAGAAACAGGATT	
<i>Sieve Element Occlusion C</i>	F	GGCGATCCTAGTGTCAGTGG	
	R	TCAGCAGTGAAAGGGAAGGC	
<i>Cinnamoyl-Reductase</i>	F	GTGGATGTTAGGGATGTGGCA	
	R	GGGTTTTGCTCTTGGGCTCT	
<i>Pectin Methylesterase 1</i>	F	TCTCTCCCGAAAATCCGTGC	
	R	GGAAGTGCTGACAGGGAGTT	
<i>Starch Branching Enzyme II</i>	F	AGGTCACCGTCAGCATCTTG	
	R	TTATGCCTGTGTCACCTGCGT	
<i>PRR Response Regulator</i>	F	CACGGCAGCAATGGACAAAA	
	R	CACTATTTCCCTGCTGCCCA	
<i>Choline Transporter-Like Protein 2</i>	F	TGTGTCAGCCTCTCAAGTGC	
	R	ACCAAGGAACCAGCAAACCA	



Supplementary Figure S1: (a) Two dimensional PCA analysis with nonmetric multidimensional scaling using RNA-seq expression data from the 36 samples analyzed. (b) Two dimensional PCA analysis with nonmetric multidimensional scaling using the RNA-seq expression data from the nine hybrids samples, but on an expanded scale. *C. sunki* x *P. trifoliata* hybrids S: Susceptible Pool; *C. sunki* x *P. trifoliata* hybrids T: Tolerant Pool; *C. sunki* x *P. trifoliata* hybrids R: Resistant Pool; trt CLAs inoculated plants; untrt: mock-inoculated plants.



Supplementary Fig. S8. Validation of RNA-seq expression profiles by RT-qPCR targeting 10 genes in *C. sinensis*, *C. sunki*, *P. trifoliata*, S Pool, T Pool and R Pool. RTqPCR analyses verified differences in gene expression considering three biological replicates of mock-inoculated and CLas inoculated plants. qRT-PCR analyses were normalized using GAPDH and FBOX as an internal control genes. The fold change of each gene was calculated by the $2^{-\Delta\Delta Ct}$ method. Error bars on the black boxes indicate the standard error of three biological replicates of RT-qPCR analysis. Significant differences in comparison with treatments were verified by Tukey test ($\alpha=0,05$).



Supplementary Figure S3: Pool selection and characterization. Experimental design of hybrid development and symptomatologic standards applied for the determination of different levels of susceptibility, tolerance or resistance observed in parental genotypes and hybrid progeny. The crosses between the susceptible (*C. sunki*) and the tolerant genotype (*P. trifoliata*) generated hybrids with different responses to HLB. Susceptible are those plants that showed both CLAs titer and HLB typical symptoms, such as mottle leaves and high accumulation of starch and callose. Tolerant are the plants that showed CLAs titer and non-visible HLB symptoms, and no starch and callose accumulation. Resistant are the plants, which presented neither detectable CLAs titer nor symptoms or starch and callose accumulation (Created with BioRender.com).

3.3. Chapter 3: CRISPR/Cas system targeting *Sieve Element Occlusion* gene to improve HLB tolerance in sweet orange trees

3.3.1. Abstract

All commercial citrus varieties are highly susceptible to Huanglongbing (HLB), more commonly known as Greening. Currently it has been considered the most devastating citrus disease in the world. Thus, the development of tolerant commercial varieties has become the biggest challenge of citrus industry. A transcriptomic analysis indicated that *Sieve Element Occlusion c* (*SEOc*) could be related to HLB susceptibility. The use of genomic editing has been shown as a promising tool to generate HLB tolerant citrus varieties. Therefore, our objective was to create site-directed gene mutagenesis in sweet orange *SEOc* and its homologous (*NtSEOI*) in tobacco, using CRISPR/Cas9 technology. In this study, to create transformants via CRISPR/Cas9, the sgRNAs were cloned into pDirect22c vector (35S::Csy4-P2A-AtCas9, 35S::gRNA-array) which is composed by Csy4 system that allows the multiplexed editing. Both citrus and tobacco T0 transgenic events were verified by PCR analysis. DNA sequencing was used to confirm the *SEO* mutation at the target site in treated sweet orange and tobacco. Nine and 78 genetically transformed plants of citrus and tobacco were obtained, respectively. As expected, the efficiency of transformation process between the species was very divergent: 2.14 % for citrus and 96% for tobacco. Targeted sequencing of the nine citrus lines showed that one plant probably had mutations in only one sgRNA site and the TIDE analysis demonstrated the mutation rate was 10%. The other three target sites at *SEOc* presented a sequence similar to the wild-type. The amplification of the tobacco targets and sequencing indicated that two plants had a large-sequence deletion by CRISPR/Cas9 system. Our study showed that, despite using the same strategy for citrus and tobacco, different results were obtained. Thus, we consider the efficiency of the CRISPR/Cas technology is species-dependent. The use of different transformation and editing strategies can optimize the process and improve editing citrus rates.

Key words: genome editing, Huanglongbing, citrus, breeding

3.3.2. Introduction

Efforts have been made to develop citrus cultivars more resistant to abiotic and biotic stresses and at the same time more productive (Dutt et al., 2020). Currently, a disease caused by the negative Gram-bacteria *Candidatus Liberibacter asiaticus* has been causing huge losses in citrus

commercial production areas. All commercial citrus varieties are highly susceptible to HLB, thus the understanding and development of tolerant commercial plants have become the main challenge for citrus industry in the world.

Previous studies have reported tolerance to HLB in *P. trifoliata* and their related genotypes (Boava et al., 2017). The gene expression profile and transcriptomic analysis of susceptible, tolerant and resistant hybrids infected with CLAs revealed that there is a differential gene expression in several biological pathways (Boava et al., 2017; Curtolo et al., 2020a). The *Sieve Element Occlusion c (SEOc)* gene was up-regulated in all susceptible genotypes and it was not differently expressed in the tolerant plants, which could indicate that *SEOc* is involved with susceptibility (Curtolo et al., 2020a). Previous studies also demonstrated that members of the SEO family encode P-protein subunits that affect phloem translocation. The deposition of high amounts of callose and P-protein on the phloem sieve plates seems to be the major alteration that determines the typical HLB symptoms (Granato et al., 2019; Curtolo et al., 2020a). Therefore, *SEOc* represents a potential target to citrus genome editing aiming the development of HLB tolerant citrus trees.

The use of genomic editing in plants is a high promising strategy for the development of cultivars with specific characteristics. Several tools of genome editing have been developed such as: zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and the most recent clustered regularly interspaced short palindromic repeats (CRISPR). Variations and combinations of techniques also have been developed. For example, CRISPR system may have enhanced specificity fusing FokI (endonuclease originally from used in ZFNs) to catalytically inactive versions of Cas9 (dCas9 – dead Cas9) (Guha and Edgell, 2017).

CRISPR allows specific genetic modifications, in a fast, targeted, effective and moderate costs (Chen et al., 2019). That strategy avoids the appearance of undesirable mutations in other genomic regions. CRISPR technology is especially important for perennial and semi-perennial species since the backcrossing required to usually segregate away the host plasmid DNA is only feasible to short life cycles plants. Using CRISPR non-transgenic mutants can be generated applying a pre-assembled enzymatic ribonucleoprotein (RNP) Cas9-sgRNA complex. Non-transgenic mutants' approach allows developed plants to be used outside of the GMO (Genetically modified organism) regulatory framework (Dort et al., 2020).

In plants, there are several genome editing studies using model plants or crops. CRISPR technology can be more easily applicable to model plants, such as: *Arabidopsis thaliana* (Miki et al., 2018), *Nicotiana tabacum* (Huang et al., 2021; Tian et al., 2021) and *Nicotiana*

benthamiana (Ma et al., 2020). On the other hand, for other crops as citrus, natural traits can directly affect the establishment of CRISPR technology. The CRISPR/Cas9 system was firstly used to target the *CsPDS* (*Phytoene desaturase*) gene in sweet orange via *Xcc*-facilitated agro infiltration (Jia and Wang, 2014). Recently Dutt et al., (2020) has successfully edited the *CsPDS* gene using citrus embryogenic cell cultures. CRISPR/Cas9 technology also has been applied to increase citrus canker resistance mediated modification of *CsLOB1* (*Lateral organ boundaries 1*) gene in Duncan grapefruit (Jia et al., 2017, 2016). *CsLOB1* gene was related to citrus canker susceptibility (Hu et al., 2014). The main factors which can interfere with the development of editing citrus trees generally are the same observed when using other genetic breeding technologies.

Those difficulties combined with a particularly complicated disease such as HLB represent a great challenge to citriculture. So, in this work we aimed to study the efficiency of CRISPR/Cas9 technique in citrus through induction site-directed mutagenesis in sweet orange *SEOc*, since it can be a potential target to develop HLB tolerant citrus genotypes. Simultaneously, we also created mutations in *NtSEO1* in tobacco (homologous of *SEOc*) in order to validate the CRISPR system adopted. Tobacco is a model plant species and it has a short development cycle, high capacity for regeneration and transformation. Also, it can be used as an experimental host system to validate useful candidate genes to control plant pathogens, including *Candidatus Liberibacter*, the causal agent of HLB (Francischini et al., 2007).

3.3.3. Material and Methods

3.3.3.1. Molecular Cloning

The sgRNAs were cloned into pDirect22c vector using the protocol 3A described by Čermák et al., (2017). PDirect22c (35S::Csy4-P2A-AtCas9, 35S::gRNA-array) is composed by Csy4 system which allows the multiplexed editing. We assembled the vector expressing four sgRNAs (sgRNA_c1: GAACTCACTTGCCAACTCTG, sgRNA_c2: CAACTGCCAGAAATTCCAGC, sgRNA_c3: GAGAGCATTGATTTATGCTG and sgRNA_c4: TATGCTGAGGATCTTGTGGA) all targeting the *SEOc* gene for citrus. The same strategy was used to assemble a vector to edit it homologous in tobacco (*NtSEO1*). In this case, we inserted two sgRNAs (sgRNA_nt1: GCCTTTGATGGCATACTCGA and sgRNA_nt2: GGATACTTATTCGACAACAA). Through the heat shock method, the assembled vectors were introduced in competent *E. coli*. The cultures were plated into solid LB medium supplemented with 100 mg L⁻¹ kanamycin incubated at 37 °C overnight. The bacteria colonies were tested using the primers TC320: CTAGAAGTAGTCAAGGCGGC and TC089R:

GGAACCCTAATTCCCTTATCTGG. Plasmids from positive colonies were extracted using the PureYield™ Plasmid Miniprep System (Promega Corporation) and sequenced. SgRNAs were correctly cloned and the vectors were introduced into *Agrobacterium tumefaciens* strain GV3101 by electroporation.

3.3.3.2.Plant transformation

Citrus transformation was performed as previously described (Orbović and Grosser 2015), but some modifications were adopted. The co-incubated explants with recombinant *Agrobacterium tumefaciens* (*A. tumefaciens*) cells were cultivated in co-cultivation medium composed by MS medium plus 1 mg L⁻¹ of BAP and 8 g L⁻¹ of agar, pH 5.8. The same composition was used to prepare the regeneration medium and the appropriated antibiotics were added in the following concentrations: 100 mg L⁻¹ kanamycin and 300 mg L⁻¹ cefotaxime.

Tobacco transformation was carried out as reported by Gao et al., (2014). Briefly, leaf discs were infected by *A. tumefaciens* GV3101 harboring the CRISPR vector. Posteriorly, leaf discs were plated onto the same regeneration medium described for citrus.

3.3.3.3.Screening of transgenic and edited plants

Firstly, the emerging shoots had the genomic DNA extracted from a piece of leaf to test the presence of T-DNA. The transgenic shoots were tested by PCR analysis with a pair of primers TC320 and TC089R. To measure the frequencies of Cas9/gRNA-induced mutations, the sgRNA target sites were amplified by PCR from the extracted genomic DNA using a Phusion polymerase (New England BioLabs) and the primers SEOc_f: GGGAGGAGGAGATGCACTTG and SEOc_3r: GAAGGCCGAAATCCCCATATC for citrus. PCR products were purified with the Wizard® SV Gel and PCR Clean-Up System (Promega) following the manufacturer's instructions. The purified PCR products of wild type (WT) and transgenic lines of Hamlin were directly sequenced and tested in TIDE software (<http://shinyapps.datacurators.nl/tide/>) in order to track indels by decomposition from sanger sequencing data. None potential off-targets were detected for the four sites based on *Citrus sinensis* genome and CRISPR-P V 2.0 software (<http://crispr.hzau.edu.cn/CRISPR2/>).

All edited plants were directly grafted onto ‘Carrizo’ citrange rootstock plants [*C. sinensis* (L.) Osbeck × *Poncirus trifoliata* (L.) Raf.] for further analysis.

It was performed a PCR using Phusion polymerase (New England BioLabs) and the primers SEOt_f: GTCTGATGATCATGCCATGTCC and SEOt_r: ACTTGAGGGAAGCATGGTGTT to screen the edited plants from tobacco transformation experiment. The amplification patterns were evaluated on agarose gel 1%. and the PCR products were also purified with the Wizard® SV Gel and PCR Clean-Up System (Promega) following the manufacturer's instructions and sequenced (Sanger).

3.3.3.4. Immunoblotting

Total protein was isolated from transgenic and WT leaves in extraction buffer (10% SDS, 50% glycerol, 25% 2-mercaptoethanol, 0.02% bromophenol blue and 0.3125 M Tris HCl, pH approx. 6.8). Five discs of leaf with half centimeter diameter were macerated in the extraction buffer. The extracts were heated to 95 °C for 5 min and centrifuged at 10,000 × g for 10 min at 4 °C. The supernatant was applied in the Nitrocellulose PVDF membrane (Amersham) for the dot-blot analysis. The membrane was blocked overnight with blocking solution (PBS with 1 % BSA). After blocking, the membranes were incubated with Anti-CRISPR-Cas9 Rabbit Monoclonal Antibody (Boster Biological Technology, Pleasanton CA, USA, Catalog # M30929-1) for 1 hour using as dilution 1:3000. Washes with TBS-T were performed 3 times for 5 min each one before the next incubation steps. Anti-rabbit IgG secondary antibody conjugated with HRP (Jackson ImmunoResearch Laboratories, Inc. Code#111-0350144, 1:5000 dilution) was incubated for 15 min. The membrane was washed with TBS-T 3 times for 5 min each one before using alkaline phosphatase reagent (Sigma, Prod. No. B6404). For Blot visualization Bio-Rad® ChemiDoc MP® system was used and the image result was analyzed by Bio-Rad Image Lab™ software. One picture per minute was taken for 20 minutes. A representative picture was selected.

3.3.4. Results

3.3.4.1. Molecular cloning sgRNAs for *SEO* gene in pDirect22c and genetic transformation of citrus and tobacco

Before designing the sgRNAs for citrus, the target gene was sequenced. The sequence analysis revealed that *SEOc* has two copies in sweet orange Hamlin genome (Figure 1A and B). Three different alleles were found and as citrus is a diploid species, it leads us to believe that *SEOc* gene has more than one copy (Figure 1B). Moreover, *SEOc* has SNPs along of gene sequence, which compromised the sgRNA design.

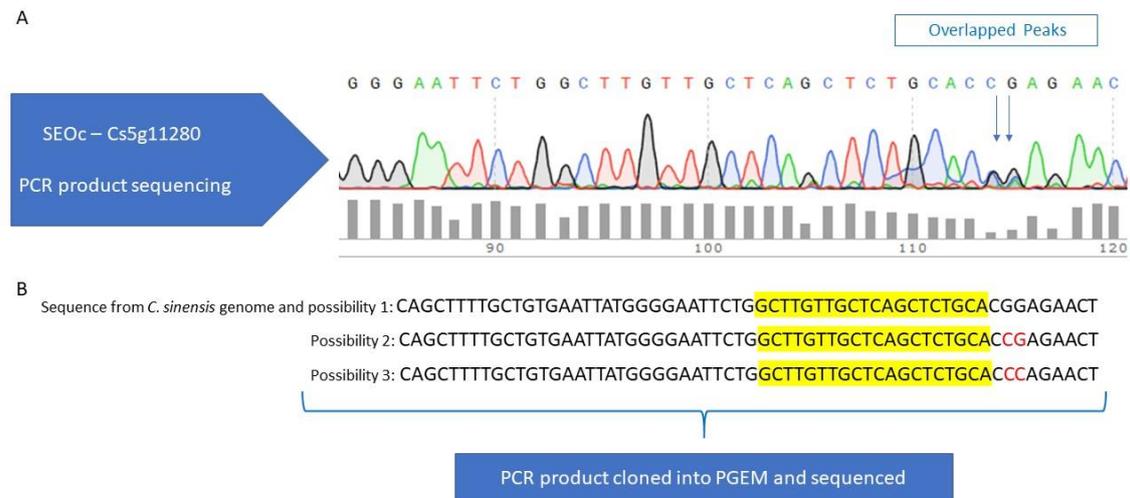


Figure 1. A: Scheme of PCR product sequence using the primers SEOc_f and SEOc_3r. Some parts of sweet orange Hamlin genome presented overlapped peaks demonstrating the presence of SNP. B: Confirmation of SNP identification by Cloned PCR product into PGEM T-easy (Promega) sequencing.

In order to edit all alleles and generate multiplex-edited plants, four different sgRNAs were designed in gene conserved regions (Figure 2A). Considering that the sgRNAs in Pdirect22c vector are not under the control of U6 promoter, the guides do not need to start with G nucleotide (Figure 2B and D). According to the methods previously described, the sgRNAs were cloned and transformed into *E. coli*. Colonies were analyzed by PCR and sequenced using primers TC320 and TC089 (Figure 2C).

TC320 and TC089R. Nine successfully transformed plants were identified, resulting in a transformation efficiency of 2.14%.

As we expected, when four sgRNAs were cloned, an amplicon of 653 bp was observed in transgenic plants (Figure 3). When more guides are added into the construction, we should consider more 116bp for each additional sgRNA (Čermák et al., 2017). Primers for *Cas9* were also tested in the transgenic plants in order to be sure that all T-DNA was inserted.

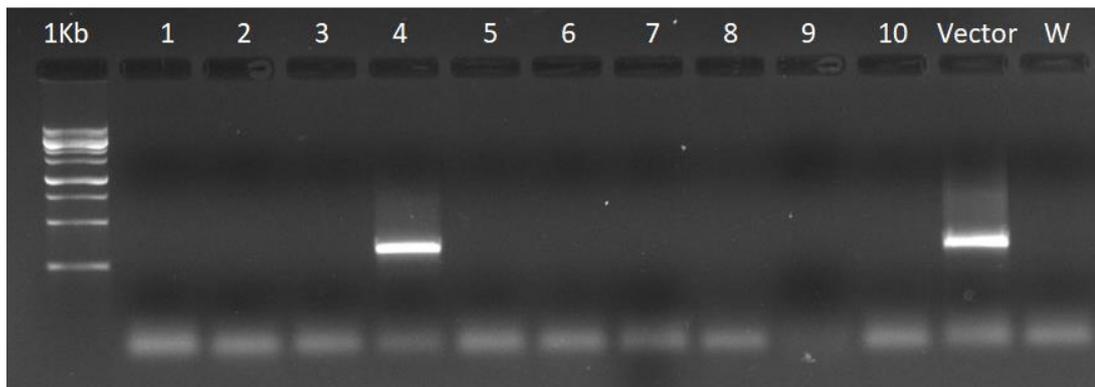


Figure 3. Representative image of screening of citrus transgenic plants. Gel agarose 1%. From left to right: Gene Ruler marker 1kb (Promega), 1-10: tested citrus plants (Amplicon size 653 bp), pDirect22c + 4 sgRNAs cloned as positive control and w: water as negative control.

For tobacco, the two target sites of *NtSEO1* were designed (sgRNA_nt1 and sgRNA_nt2). The target region was sequenced to certify that no SNP was present and could interfere with the recognition of sgRNA by Cas9 protein. The amplification pattern of transformed *E. coli* showed that two colonies had only one sgRNA cloned (Figure 4, lanes 1 and 3) and the others presented two sgRNAs successfully cloned (Figure 4, lanes 2 and 4). The two colonies which presented the expected size of the amplicon had the vector extracted and sequenced. Both presented the sgRNA_nt1 and sgRNA_nt2, but only one colony was randomly chosen to be used in tobacco transformation experiments. The expression cassette was transformed into tobacco using the leaf discs. Seventy-eight transgenic plants were obtained among 81 tested by PCR using the primers TC320 and TC089R (Figure 5). The amplicon size is dependent on the number of sgRNAs that were cloned into vector and since two sgRNA were cloned we expected an amplicon with 421 pb.

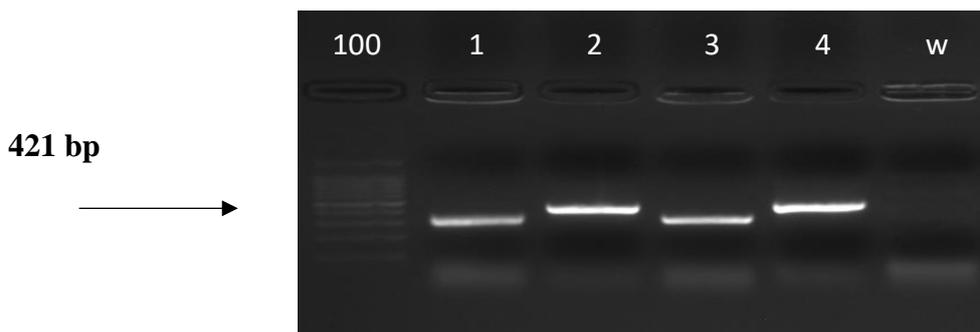


Figure 4. Confirmation of transformed *E. coli* colonies and vector pDirect22C customization with both sgRNAs for editing the *NtSEO1* gene. From left to right: Gene Ruler marker 100bp (Promega), 1- 4 tested colonies (Expected amplicon size 421 bp) and w: water as negative control.

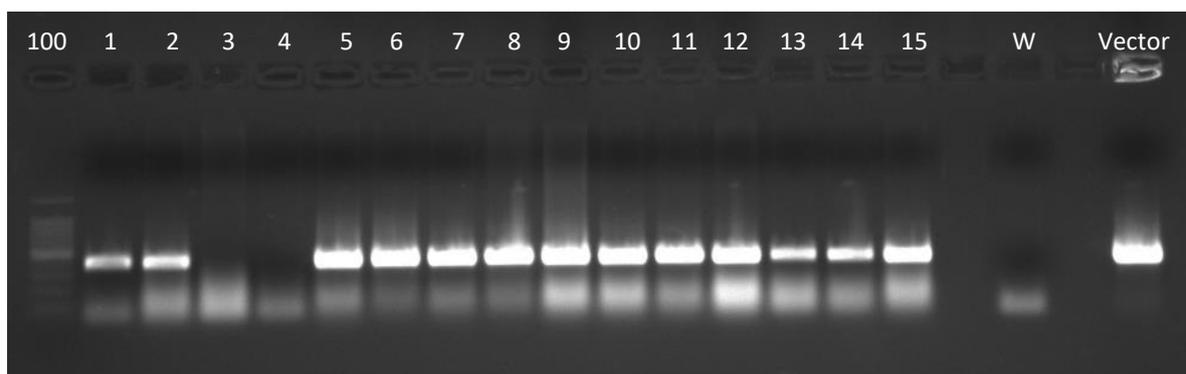


Figure 5. Representative image of PCR screening of tobacco transgenic plants. Gel agarose 1%. From left to right: Gene Ruler marker 100bp (Promega), 1-15: tested tobacco plants (Amplicon size 421 bp), pDirect22c + 2 sgRNAs cloned as positive control and w: water as negative control.

3.3.4.2. Immunoblotting of citrus transgenic leaves

The vector used in the experiment is part of a multipurpose toolkit to enable advanced genome engineering in plant proposed by Čermák et al. (2017). However, as the development of edited citrus plant using this tool was not previously described, we verified if the Cas9 protein was being translated into the transformed citrus plants. The protein from one representative plant was extracted from different leaves for immunoblotting using Cas-9 antibody. We analyzed different leaves to verify the possible occurrence of chimeric, since the development of chimeric in citrus transgenic plants is a factor that should be considered. In total, we analyzed six

transgenic leaves from the same plant and three different control plants (Wild type) (Figure 6A). The image was evaluated by the Bio-Rad Image Lab™ software (Figure 6B).

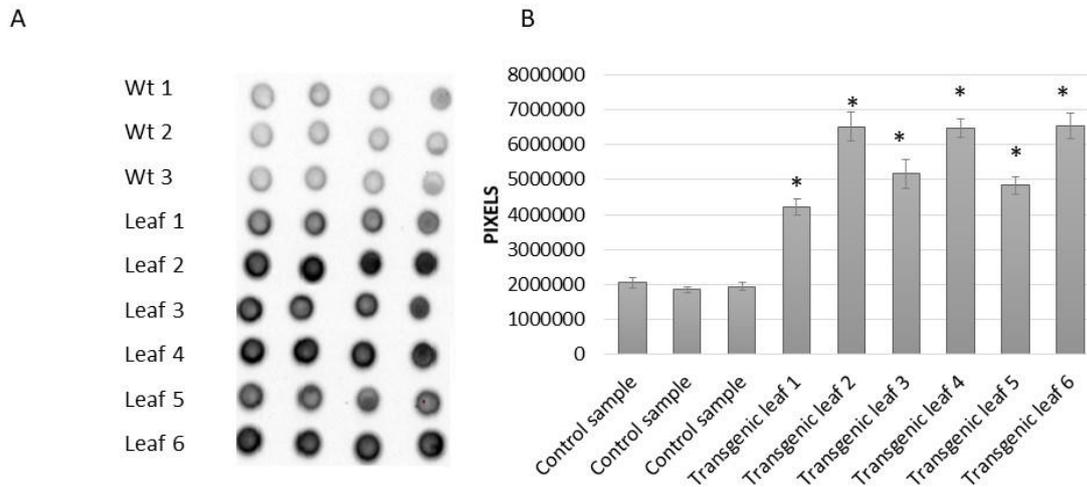


Figure 6. Immunoblotting of transgenic plants and pixels quantification. A. Image of transgenic plants immunoblotting. Each sample is represented by four technical repetition of six transgenic leaves from the same transgenic citrus plant. Three different WT plants were used as control. B. Pixels quantification. The blots were quantified by Image Lab software (Bio-Rad). *Samples that presented significant difference compared to wild type plants by *t*-Student analysis ($p > 0.05$).

As observed in Figure 6, there was a significant difference in the detection of the Cas9 protein in the WT samples (Control samples) in relation to transgenic plant (Leaf 1, 2, 3, 4, 5 and 6), indicating that the transgenic plant was able to transcribe and translate Cas9 protein.

3.3.4.3. Identification of edited *SEOc* gene in citrus genotype and *NtSEO1* in tobacco

To further confirm the gene editing events and estimate editing efficiency, all transgenic plants and WT Hamlin had the target region amplified using the primers *SEOc_f*: GGGAGGAGGAGATGCACTTG and *SEOc_3r*: GAAGGCCGAAATTCCCATATC for citrus. All transgenic plants showed amplification patterns similar to WT plant, indicating that no significant deletions were produced in citrus genome (Figure 7).

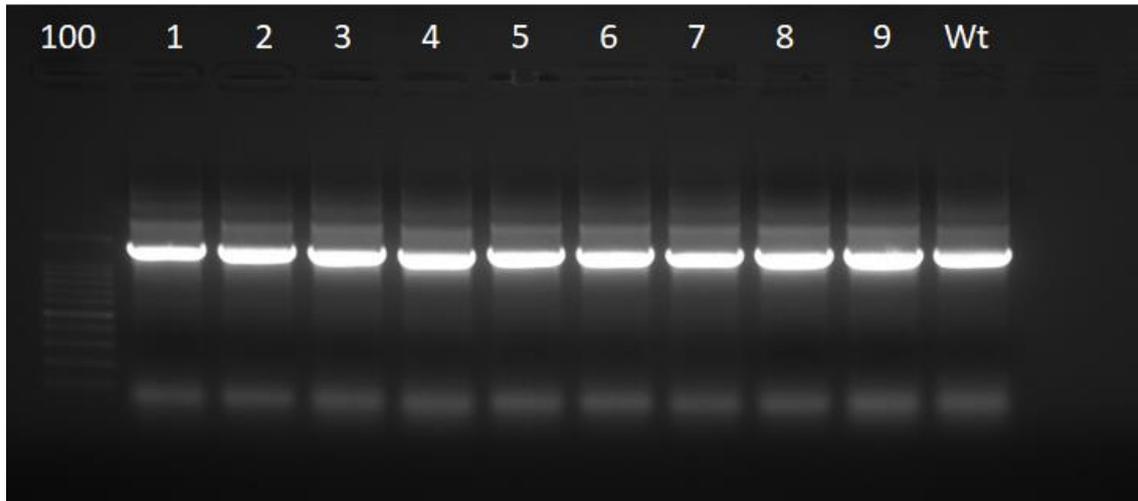


Figure 7. Sweet orange target region amplification. Gel agarose 1%. From left to right: Gene Ruler marker 100bp (Promega), 1-9: tested citrus transgenic plants and Wt = no transgenic plant (Amplicon size 1300 bp).

As four sgRNAs were cloned in the backbone vector, the sequence among the sgRNAs targets might have been deleted or expelled from citrus genome as well as exposed by other genome editing studies (Zsögön et al., 2018; Huang et al., 2020). The nine transgenic plants for Cas9 had the PCR product purified, directly sequenced and analyzed in TIDE software where is possible to track indels by decomposition using standard capillary sequencing reactions.

Based on sequencing results we identified only one putative edited plant. In this case, it was verified the presence of overlapped peaks in electropherogram from Sanger sequencing data (Figure 8A). The editing evidence was only verified at sgRNA_c2 target. This pattern was not detected in the other eight transgenic plants sequencing.

To calculate the efficiency of *SEOc* mutation caused by Cas9/sgRNA in Hamlin, we submitted the two standard capillary sequencing reactions into TIDE software. One was from the WT plant and the other was the sequence resulted from the transgenic plant that had the overlapped peaks. In this analysis, it was possible to determine the spectrum and frequency of small insertions and deletions (*indels*) generated in a pool of cells by genome editing tools. TIDE analysis revealed that we have achieved a 10.7% of mutations rate, so in this case there is a mixed of non-edited and edited cells (Figure 8B).

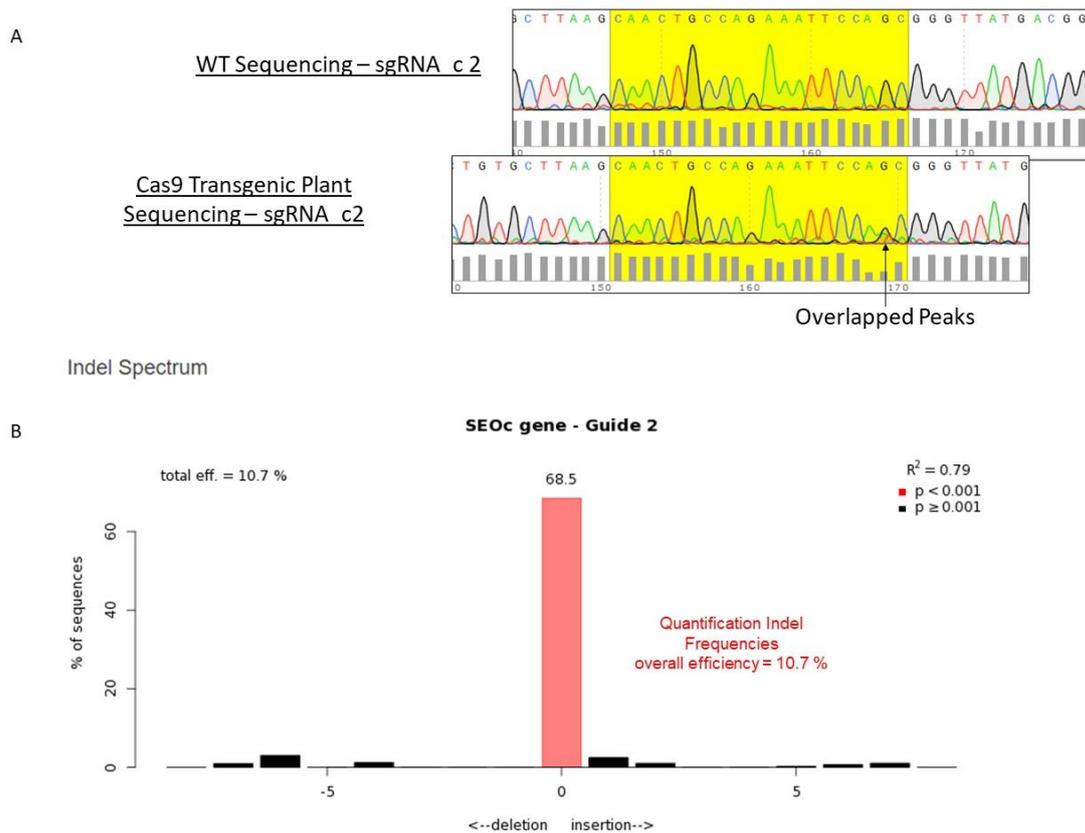


Figure 8. A: sgRNA_c2 target sequence in WT and one transgenic plant. The presence of overlapped peaks in the electropherogram from transgenic plant sequencing data evidencing the presence of indels at sgRNA_c2 target. B: Quantification of indel frequencies from TIDE analysis.

To determine the rate of mutagenesis onto tobacco we amplified the target region using the total genome extracted from all transgenic plants using the primers NtSEO1_f: CATGCTACATGGCACTACTGAT and NtSEO1_r: ACTTGAGGGAAGCATGGTGTT (Figure 9A). We observed that most of the plants presented the amplicon size similar to the wild tobacco (Figure 9B). However, two plants showed a different amplification pattern, presenting two bands (Figure 9B, lane 6). The shorter DNA fragment was sequenced, and we confirmed the chromosomal region between the two target sites was deleted (Figure 9C).

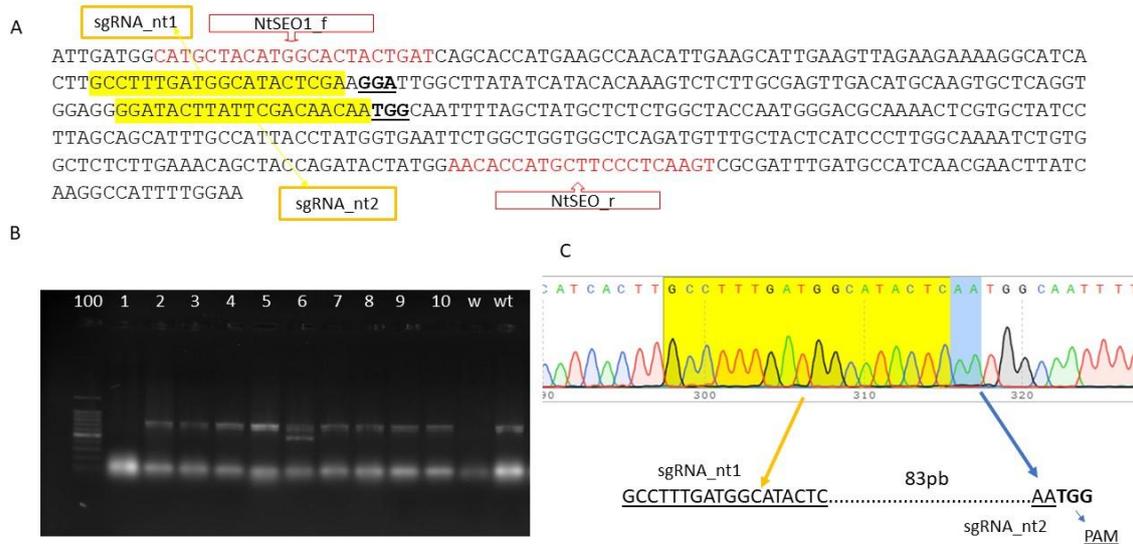


Figure 9. A: Representation of *NtSEO1* target sequencing, primers and sgRNAs positions: sgRNA_nt1, sgRNA_nt2 and the yellow sequences are the two targets at *NtSEO1*; the underlined nucleotides display the PAM sequences; NtSEO1_f and NtSEO1_r are the primers sequences used to amplify the target region; B: Target region amplification. Gel agarose 1%. From left to right: Gene Ruler marker 100bp (Promega), 1-10: tested tobacco transgenic plants, w= water as negative control and WT = no transgenic tobacco (Amplicon size from WT tobacco 750 bp approximately). C: Sequencing representation of the smaller band indicating a big deletion (83 nucleotides) between sgRNAs.

3.3.5. Discussion

CRISPR technology has been widely developed and tested to all living organisms or cells. Based on the assumptions of the technique, it seems that CRISPR is a technology easily applicable to any system and it promises to accelerate and improve many crops. It is true that CRISPR/Cas has significant advantages over other breeding technologies since it can precisely target genome sequences to be manipulated. Indeed, there is a solid base for some plants species, especially for model plants such as tobacco and *Arabidopsis*, achieving successful results (Wada et al., 2020).

To date, there are studies using CRISPR/Cas system in citrus, but most of them belong to a restricted group of researchers. In citrus, CRISPR/Cas9 system was firstly used to target the *CsPDS* gene in sweet orange via *Xcc*-facilitated agro infiltration (Jia and Wang 2014). Recently, Dutt et al., (2020) have successfully edited the *CsPDS* gene using citrus embryogenic cell cultures. CRISPR/Cas9 technology has also been applied to increase citrus canker

resistance through the modification of *CsLOB1* gene in Duncan grapefruit (Jia et al., 2017, 2016). *CsLOB1* gene was related to citrus canker susceptibility (Hu et al., 2014). The used strategy mutated only one allele, but it was enough to relieve the canker symptoms. Later, the edition of both alleles of *CsLOB1* promoters showed a high degree of resistance to citrus canker (Peng et al., 2017).

To crops such as citrus, some natural traits complicate and delay the development of all breeding technology, including CRISPR/Cas. For example, in this study it was observed a transformation efficiency of 2.14% and 96% for citrus and tobacco, respectively. Although numerous protocols have already been published to optimize the efficiency of transformation for different citrus varieties, many studies have continually been conducted (Sun et al., 2019). In addition to the low efficiency of the transformation process, the genetic transformation using citrus juvenile epicotyl or mature stem tissue usually result in chimeric plants. Indeed, it became necessary to obtain a large number of transgenic plants in order to find a non-chimeric plant. The edition rate can be very variable from 1% to 90% (Jia & Wang, 2014; Jia et al., 2016, 2017; Peng et al., 2017). This information associated with the fact that citrus genetic transformation process produce mostly chimeric shoots represent a real challenge using CRISPR in citrus (Domínguez et al., 2004; Dutt et al., 2020). In chimeric transgenic plants non-edited and edited cells will be mixed, composing a tissue.

In order to increase the editing, we adopted a system which can express multiple sgRNAs via *Csy4 ribonuclease*. But even using such strategy, it was observed the edition in only one target in one citrus plant and with a low frequency. The low frequency of edition can indicate the presence of a chimeric plant. In this case, most of the plant cells are not edited, so this specific plant will not achieve its full improvement potential. In addition, since most of the cells are not edited, the plant will likely transcribe and produce protein from *SEOc* gene. So, we cannot consider that the plant had *SEOc* significantly edited. On the other hand, the results showed that mutations at *NtSEO1* occurred at both target sites simultaneously, resulting in a significantly deletion (Figure 9B and C). These results are similar to the previously published studies using CRISPR/Cas technology for citrus and tobacco editing. In citrus, it was demonstrated that sgRNAs targeting two locations in the *csPDS* gene produced only insertion (1–2 bp), substitution (1 bp), or deletion (1–3 bp) mutations of few nucleotides (Dutt et al., 2020). Meanwhile, in tobacco, deletions and inversions of a 1.8-kb fragment between two target sites in the *NtPDS* locus were detected. Indel mutations were also detected at both sgRNA targets (Gao et al., 2015).

Divergences between sgRNA and genome sequence with other unknown factors either can affect the system efficiency or can edit one single allele (Jia et al., 2016). To access the information that precedes the guides design, it is necessary to have the genome under study sequenced and assembled. In addition, polymorphisms among varieties should be considered. Although the Citrus Clementine and sweet orange (*Citrus sinensis*) genome have already been sequenced and assembled, the sequence of the target gene may be different among oranges, since the citrus genome is highly polymorphic with several SNPs (Curtolo et al., 2020b).

The identification of mutations in the targeted region is a crucial step in CRISPR/Cas mutagenesis (Li et al., 2018). There are several technologies that allow the screening of CRISPR/Cas-induced mutations, among them: PCR, digestion with enzymes, multiplex ligation-dependent probe amplification-based method, High Resolution Melting (HRM), sequencing using Sanger or Illumina technology and Capillary electrophoresis (Samarut et al., 2016; Li et al., 2018). The occurrence of natural polymorphisms (SNP) combined with the obtained chimeric plants represent aspects which can compromise the edited plants identification by some techniques. For example: HRM is based on the fluorescence changes during the melting of the DNA duplex. In this case indels can be confused with SNP and present the same pattern.

To work with citrus protoplasts represents a real challenge, but it is extremely necessary when using CRISPR/Cas9 system. Protoplast transformation can be an option to enhance the genetic transformation efficiency and to avoid chimeric plants. Citrus protoplast regeneration is not a simple and easy process. Several aspects can affect the ability of protoplast-derived cells to express their totipotency and to develop into fertile plants, among the main parameters: source tissue, culture medium and environmental factors (Davey et al., 2005). In addition, regenerating protoplasts into a plant requires a long time, so it is really interesting to associate genes reporter, such as GFP, which allows the rapid assessment of CRISPR construction (Huang et al., 2020). Furthermore, it is possible to transfect pre-assembled ribonucleoproteins (RNPs) and to induce HR (homologous recombination) through the use of protoplasts as explant (Poddar et al., 2020). RNPs can produce genome modifications with the absence of insertion gene exogenous (Mao et al., 2019). In this case, the plants will not be considered as transgenics and they will not require a rigorous regulation process before being allowed for commercialization (Jia et al., 2017).

There are other challenges and limitations to the application of the CRISPR/Cas technology in citrus.

Those difficulties combined with a particularly complicated disease such as HLB represent a great challenge to citriculture.

The studied gene in this work can represent a target to the development of HLB tolerant citrus trees. Expression profiles from RNA-Seq analysis associated *SEOc* gene expression to HLB susceptibility, so in this work we aimed to knock it out from the sweet orange Hamlin sweet orange genome and its homologous in tobacco (*NtSEOI*) (Curtolo et al., 2018). However, we need to point out that due to the low frequency of edited cells in citrus *Cas9* transgenic plant, we cannot consider that the resulting plant had its gene significantly edited for application to control HLB. Despite that, we were able to successfully develop and apply the genome editing technology and show alternatives to improve CRISPR technology in further citrus studies.

3.3.6. Conclusion

We have shown that CRISPR/Cas9 system can be used to modify the citrus and tobacco genome. Considering different species, CRISPR/Cas9 can generate different mutation patterns. Optimizations are needed to increase the mutation rate in citrus, since the delivery method of the CRISPR/Cas9 system can directly affect the efficiency of mutation.

3.3.7. References

- Boava LP, Cristofani-Yaly M, Machado MA (2017) Physiologic, anatomic, and gene expression changes in *Citrus sunki*, *Poncirus trifoliata*, and their hybrids after “*Candidatus Liberibacter asiaticus*” infection. *Phytopathology* 107:590–599. <https://doi.org/10.1094/PHYTO-02-16-0077-R>
- Čermák T, Curtin SJ, Gil-Humanes J, et al (2017) A multipurpose toolkit to enable advanced genome engineering in plants. *Plant Cell* 29:1196–1217. <https://doi.org/10.1105/tpc.16.00922>
- Chen K, Wang Y, Zhang R, et al (2019) CRISPR/Cas Genome Editing and Precision Plant Breeding in Agriculture. *Annu Rev Plant Biol* 70:667–697. <https://doi.org/10.1146/annurev-arplant-050718-100049>
- Curtolo M, de Souza Pacheco I, Boava LP, et al (2020a) Wide-ranging transcriptomic analysis of *Poncirus trifoliata*, *Citrus sunki*, *Citrus sinensis* and contrasting hybrids reveals HLB tolerance mechanisms. *Sci Rep* 10:1–14. <https://doi.org/10.1038/s41598-020-77840-2>
- Curtolo M, Moreira Granato L, Aparecida T, et al (2020b) Expression Quantitative Trait Loci (eQTL) mapping for *callose synthases* in intergeneric hybrids of Citrus challenged with the

bacteria *Candidatus Liberibacter asiaticus*. <https://doi.org/10.1590/1678-4685-GMB-2019-0133>

Curtolo M, Soratto TAT, Gazaffi R, et al (2018) High-density linkage maps for *Citrus sunki* and *Poncirus trifoliata* using DArTseq markers. *Tree Genet Genomes* 14:. <https://doi.org/10.1007/s11295-017-1218-9>

Davey MR, Anthony P, Power JB, Lowe KC (2005) Plant protoplasts: Status and biotechnological perspectives. *Biotechnol. Adv.* 23:131–171

Domínguez A, Cervera M, Pérez RM, et al (2004) Characterisation of regenerants obtained under selective conditions after *Agrobacterium*-mediated transformation of citrus explants reveals production of silenced and chimeric plants at unexpected high frequencies. *Mol Breed* 14:171–183. <https://doi.org/10.1023/B:MOLB.0000038005.73265.61>

Dort EN, Tanguay P, Hamelin RC (2020) CRISPR/Cas9 Gene Editing: An Unexplored Frontier for Forest Pathology. *Front. Plant Sci.* 11:1126

Dutt M, Mou Z, Zhang X, et al (2020) Efficient CRISPR/Cas9 genome editing with Citrus embryogenic cell cultures. *BMC Biotechnol* 20:58. <https://doi.org/10.1186/s12896-020-00652-9>

Francischini FJB, Oliveira KDS, Astúa-Monge G, et al (2007) First Report on the Transmission of ‘*Candidatus Liberibacter americanus*’ from *Citrus* to *Nicotiana tabacum* cv. Xanthi . *Plant Dis* 91:631–631. <https://doi.org/10.1094/pdis-91-5-0631b>

Gao J, Wang G, Ma S, et al (2015) CRISPR/Cas9-mediated targeted mutagenesis in *Nicotiana tabacum*. *Plant Mol Biol* 87:99–110. <https://doi.org/10.1007/s11103-014-0263-0>

Granato LM, Galdeano DM, D’Alessandre NDR, et al (2019) *Callose synthase* family genes plays an important role in the Citrus defense response to *Candidatus Liberibacter asiaticus*. *Eur J Plant Pathol* 155:25–38. <https://doi.org/10.1007/s10658-019-01747-6>

Guha TK, Edgell DR (2017) Applications of alternative nucleases in the age of CRISPR/Cas9. *Int. J. Mol. Sci.* 18

Hu Y, Zhang J, Jia H, et al (2014) *Lateral organ boundaries 1* is a disease susceptibility gene for citrus bacterial canker disease. *Proc Natl Acad Sci U S A* 111:E521–E529. <https://doi.org/10.1073/pnas.1313271111>

- Huang TK, Armstrong B, Schindele P, Puchta H (2021) Efficient gene targeting in *Nicotiana tabacum* using CRISPR/SaCas9 and temperature tolerant LbCas12a. *Plant Biotechnol J*. <https://doi.org/10.1111/pbi.13546>
- Huang X, Wang Y, Xu J, Wang N (2020) Development of multiplex genome editing toolkits for citrus with high efficacy in biallelic and homozygous mutations. *Plant Mol Biol* 104:297–307. <https://doi.org/10.1007/s11103-020-01043-6>
- Jia H, Orbovic V, Jones JB, Wang N (2016) Modification of the PthA4 effector binding elements in Type I *CsLOB1* promoter using Cas9/sgRNA to produce transgenic Duncan grapefruit alleviating Xcc Δ pthA4: DCsLOB1.3 infection. *Plant Biotechnol J* 14:1291–1301. <https://doi.org/10.1111/pbi.12495>
- Jia H, Wang N (2014) Targeted Genome Editing of Sweet Orange Using Cas9/sgRNA. *PLoS One* 9:e93806. <https://doi.org/10.1371/journal.pone.0093806>
- Jia H, Zhang Y, Orbović V, et al (2017) Genome editing of the disease susceptibility gene *CsLOB1* in citrus confers resistance to citrus canker. *Plant Biotechnol J* 15:817–823. <https://doi.org/10.1111/pbi.12677>
- Li S, Liu S, Liu Y, et al (2018) HRM-facilitated rapid identification and genotyping of mutations induced by CRISPR/ Cas9 mutagenesis in rice. *Crop Breed Appl Biotechnol* 18:184–191. <https://doi.org/10.1590/1984-70332018v18n2a26>
- Ma X, Zhang X, Liu H, Li Z (2020) Highly efficient DNA-free plant genome editing using virally delivered CRISPR–Cas9. *Nat Plants* 6:773–779. <https://doi.org/10.1038/s41477-020-0704-5>
- Mao Y, Botella JR, Liu Y, Zhu JK (2019) Gene editing in plants: Progress and challenges. *Natl. Sci. Rev.* 6:421–437
- Miki D, Zhang W, Zeng W, et al (2018) CRISPR/Cas9-mediated gene targeting in *Arabidopsis* using sequential transformation. *Nat Commun* 9:1–9. <https://doi.org/10.1038/s41467-018-04416-0>
- Orbović V, Grosser JW (2015) Citrus transformation using juvenile tissue explants. *Methods Mol Biol* 1224:245–257. https://doi.org/10.1007/978-1-4939-1658-0_20
- Peng A, Chen S, Lei T, et al (2017) Engineering canker-resistant plants through CRISPR/Cas9-targeted editing of the susceptibility gene *CsLOB1* promoter in citrus. *Plant Biotechnol J*

15:1509–1519. <https://doi.org/10.1111/pbi.12733>

Pitino M, Allen V, Duan Y (2018) *Las45315* effector induces extreme starch accumulation and chlorosis as *Ca. Liberibacter asiaticus* infection in *Nicotiana benthamiana*. *Front Plant Sci* 9:. <https://doi.org/10.3389/fpls.2018.00113>

Poddar S, Tanaka J, Cate JHD, et al (2020) Efficient isolation of protoplasts from rice calli with pause points and its application in transient gene expression and genome editing assays. *Plant Methods* 16:151. <https://doi.org/10.1186/s13007-020-00692-4>

Samarut É, Lissouba A, Drapeau P (2016) A simplified method for identifying early CRISPR-induced indels in zebrafish embryos using High Resolution Melting analysis. *BMC Genomics* 17:547. <https://doi.org/10.1186/s12864-016-2881-1>

Sun L, Nasrullah, Ke F, et al (2019) Citrus genetic engineering for disease resistance: past, present and future. *Int. J. Mol. Sci.* 20

Tian Y, Liu X, Fan C, et al (2021) Enhancement of Tobacco (*Nicotiana tabacum* L.) Seed Lipid Content for Biodiesel Production by CRISPR-Cas9-Mediated Knockout of *NtAn1*. *Front Plant Sci* 11:2253. <https://doi.org/10.3389/fpls.2020.599474>

Wada N, Ueta R, Osakabe Y, Osakabe K (2020) Precision genome editing in plants: State-of-the-art in CRISPR/Cas9-based genome engineering. *BMC Plant Biol.* 20:234

Zsögön A, Čermák T, Naves ER, et al (2018) De novo domestication of wild tomato using genome editing. *Nat Biotechnol* 36:1211–1216. <https://doi.org/10.1038/nbt.4272>

4. Final consideration

The present work used RNA-seq analysis and eQTL mapping to expand the knowledge about citrus-CLas interaction.

A wide-ranging transcriptomic analysis of *Poncirus trifoliata*, *Citrus sunki*, *Citrus sinensis* and contrasting hybrids revealed that HLB is indeed such a complex disease, since thousands of genes had their expression level affected during the CLas infection. The comparative study of transcriptomes allows to build a hypothetical model to understand the genetic mechanisms involved in different responses to CLas infection. In addition, some specific pathways and genes were linked with susceptibility, tolerance and resistance. The susceptibility response was mainly related to the down regulation of signaling receptors and the up regulation of genes related to gibberellin synthesis, callose and PP2 deposition. Meanwhile, the induction of signaling receptors, phenylpropanoids, cell wall-strengthened and GA related degradation genes seem to be linked with HLB tolerance response. We believe that an early and fast defense response may occur in resistance hybrids to CLas, since a low number of genes are modulated after 240 days post inoculation. Nevertheless, the induction of signaling receptors and upregulation of *Endochitinase B* were observed response in resistant hybrids.

Since physiological study and transcriptome analysis indicate the P protein and callose deposition on the phloem sieve plates seem to be the main alteration that determines the typical HLB symptoms, we performed the genetic mapping for *calloses synthases*. From the eQTL identified it was concluded that multiple regions can contribute to *Cscals* expression regulation and some eQTL have an epistatic effect for more than one *Cscals* gene, demonstrating again the complexity of the disease under study.

The regions identified from transcriptomics analysis or QTL mapping can be interesting targets for future studies of *Citrus* breeding programs to manipulate the genetic and molecular response during CLas infection.

However, HLB is an extremely complex and polygene disease. The *SEOc* gene, which was linked with HLB citrus susceptibility, was selected to be knocked out in the Hamlin genome and its homologous in tobacco (*NtSEOI*). The results demonstrated that CRISPR/Cas is an important tool for breeding programs and can assist in the development of HLB tolerant citrus commercial varieties. On the other hand, the effectiveness and efficiency are extremely variable when using CRISPR/Cas as genomic editing system for different species. The delivery method of endonuclease and sgRNA vector (vector or RNP complex) and explants source are aspects that should be considered before starting genomic editing assays.

5. References

- Albrecht U, Bowman KD (2012) Transcriptional response of susceptible and tolerant citrus to infection with *Candidatus Liberibacter asiaticus*. *Plant Sci* 185–186:118–130. <https://doi.org/10.1016/j.plantsci.2011.09.008>
- Balan B, Ibáñez AM, Dandekar AM, et al (2018) Identifying Host Molecular Features Strongly Linked With Responses to Huanglongbing Disease in Citrus Leaves. *Front Plant Sci* 9:277. <https://doi.org/10.3389/fpls.2018.00277>
- Barrangou R, Marraffini LA (2014) CRISPR-cas systems: Prokaryotes upgrade to adaptive immunity. *Mol. Cell* 54:234–244
- Boava LP, Cristofani-Yaly M, Machado MA (2017) Physiologic, anatomic, and gene expression changes in *Citrus sunki*, *Poncirus trifoliata*, and their hybrids after “*Candidatus liberibacter asiaticus*” infection. *Phytopathology* 107:590–599. <https://doi.org/10.1094/PHYTO-02-16-0077-R>
- Cai Q, Guy CL, Moore GA (1994) Extension of the linkage map in Citrus using random amplified polymorphic DNA (RAPD) markers and RFLP mapping of cold-acclimation-responsive loci. *Theor Appl Genet* 89:606–614. <https://doi.org/10.1007/BF00222455>
- Chen C, Bowman KD, Choi YA, et al (2008) EST-SSR genetic maps for *Citrus sinensis* and *Poncirus trifoliata*. *Tree Genet Genomes* 4:1–10. <https://doi.org/10.1007/s11295-007-0083-3>
- Chen K, Wang Y, Zhang R, et al (2019) CRISPR/Cas Genome Editing and Precision Plant Breeding in Agriculture. *Annu Rev Plant Biol* 70:667–697. <https://doi.org/10.1146/annurev-arplant-050718-100049>
- Clark K, Franco JY, Schwizer S, et al (2018) An effector from the Huanglongbing-associated pathogen targets citrus proteases. *Nat Commun* 9:1–11. <https://doi.org/10.1038/s41467-018-04140-9>
- Clark KJ, Pang Z, Trinh J, et al (2020) *Sec-delivered effector 1 (SDE1)* of “*Candidatus Liberibacter asiaticus*” promotes citrus huanglongbing. *Mol Plant-Microbe Interact* 33:1394–1404. <https://doi.org/10.1094/MPMI-05-20-0123-R>
- Cristofani M, Machado MA, Grattapaglia D (1999) Genetic linkage maps of *Citrus sunki* Hort. ex. Tan. and *Poncirus trifoliata* (L.) Raf. and mapping of citrus tristeza virus resistance gene.

Euphytica 109:25–32. <https://doi.org/10.1023/A:1003637116745>

Cuenca J, Aleza P, Vicent A, et al (2013) Genetically Based Location from Triploid Populations and Gene Ontology of a 3.3-Mb Genome Region Linked to Alternaria Brown Spot Resistance in Citrus Reveal Clusters of Resistance Genes. PLoS One 8:. <https://doi.org/10.1371/journal.pone.0076755>

Curtolo M, Cristofani-Yaly M, Gazaffi R, et al (2017) QTL mapping for fruit quality in Citrus using DArTseq markers. BMC Genomics 18:. <https://doi.org/10.1186/s12864-017-3629-2>

Curtolo M, Souza Pacheco I, Boava LP, et al (2020a) Wide-ranging transcriptomic analysis of *Poncirus trifoliata*, *Citrus sunki*, *Citrus sinensis* and contrasting hybrids reveals HLB tolerance mechanisms. Sci Rep 10:1–14. <https://doi.org/10.1038/s41598-020-77840-2>

Curtolo M, Moreira Granato L, Aparecida T, et al (2020b) Expression Quantitative Trait Loci (eQTL) mapping for *callose synthases* in intergeneric hybrids of Citrus challenged with the bacteria *Candidatus Liberibacter asiaticus*. <https://doi.org/10.1590/1678-4685-GMB-2019-0133>

Curtolo M, Soratto TAT, Gazaffi R, et al (2018) High-density linkage maps for *Citrus sunki* and *Poncirus trifoliata* using DArTseq markers. Tree Genet Genomes 14:. <https://doi.org/10.1007/s11295-017-1218-9>

Davis MJ, Mondal SN, Chen H, et al (2008) Co-cultivation of “*Candidatus liberibacter asiaticus*” with actinobacteria from citrus with Huanglongbing. Plant Dis 92:1547–1550. <https://doi.org/10.1094/PDIS-92-11-1547>

Deng H, Achor D, Exteberria E, et al (2019) Phloem Regeneration Is a Mechanism for Huanglongbing-Tolerance of “Bears” Lemon and “LB8-9” Sugar Belle® Mandarin. Front Plant Sci 10:277. <https://doi.org/10.3389/fpls.2019.00277>

Deng Z, Huang S, Xiao S, Gmitter FG (1997) Development and characterization of SCAR markers linked to the citrus tristeza virus resistance gene from *Poncirus trifoliata*. Genome 40:697–704. <https://doi.org/10.1139/g97-792>

Domínguez A, Cervera M, Pérez RM, et al (2004) Characterisation of regenerants obtained under selective conditions after *Agrobacterium*-mediated transformation of citrus explants reveals production of silenced and chimeric plants at unexpected high frequencies. Mol Breed 14:171–183. <https://doi.org/10.1023/B:MOLB.0000038005.73265.61>

Donkersley P, Silva FWS, Carvalho CM, et al (2018) Biological, environmental and socioeconomic threats to citrus lime production. *J. Plant Dis. Prot.* 125:339–356

Durham RE, Liou PC, Gmitter FG, Moore GA (1992) Linkage of restriction fragment length polymorphisms and isozymes in Citrus. *Theor Appl Genet* 84:39–48. <https://doi.org/10.1007/BF00223979>

Dutt M, Mou Z, Zhang X, et al (2020) Efficient CRISPR/Cas9 genome editing with Citrus embryogenic cell cultures. *BMC Biotechnol* 20:58. <https://doi.org/10.1186/s12896-020-00652-9>

Ettxeberria E, Gonzalez P, Achor D, Albrigo G (2009) Anatomical distribution of abnormally high levels of starch in HLB-affected Valencia orange trees. *Physiol Mol Plant Pathol* 74:76–83. <https://doi.org/10.1016/j.pmpp.2009.09.004>

Fang DQ, Federici CT, Roose ML (1997) Development of molecular markers linked to a gene controlling fruit acidity in citrus. *Genome* 40:841–849. <https://doi.org/10.1139/g97-809>

Ferreira A, da Silva MF, da Costa e Silva L, Cruz CD (2006) Estimating the effects of population size and type on the accuracy of genetic maps. *Genet Mol Biol* 29:187–192. <https://doi.org/10.1590/S1415-47572006000100033>

Fundecitrus. <https://www.fundecitrus.com.br/>. Accessed 31 Jan 2021b

García R, Asíns MJ, Forner J, Carbonell EA (1999) Genetic analysis of apomixis in Citrus and Poncirus by molecular markers. *Theor Appl Genet* 99:511–518. <https://doi.org/10.1007/s001220051264>

Gier RA, Budinich KA, Evitt NH, et al (2020) High-performance CRISPR-Cas12a genome editing for combinatorial genetic screening. *Nat Commun* 11:. <https://doi.org/10.1038/s41467-020-17209-1>

Gmitter FG, Xiao SY, Huang S, et al (1996) A localized linkage map of the citrus tristeza virus resistance gene region. *Theor Appl Genet* 92:688–695. <https://doi.org/10.1007/BF00226090>

Granato LM, Oliveira TS, Boscariol-Camargo RL, et al (2020) ‘*Candidatus Liberibacter asiaticus*’ putative effectors: in silico analysis and gene expression in citrus leaves displaying distinct huanglongbing symptoms. *Trop Plant Pathol* 45:646–657. <https://doi.org/10.1007/s40858-020-00382-5>

- Grattapaglia D, Sederoff R (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: Mapping strategy and RAPD markers. *Genetics* 137:1121–1137. <https://doi.org/10.1093/genetics/137.4.1121>
- Gulsen O, Uzun A, Canan I, et al (2010) A new citrus linkage map based on SRAP, SSR, ISSR, POGP, RGA and RAPD markers. *Euphytica* 173:265–277. <https://doi.org/10.1007/s10681-010-0146-7>
- Guo F, Yu H, Tang Z, et al (2015) Construction of a SNP-based high-density genetic map for pummelo using RAD sequencing. *Tree Genet Genomes* 11:1–11. <https://doi.org/10.1007/s11295-014-0831-0>
- He L, James MSJ, Radovcic M, et al (2020) Cas3 protein—a review of a multi-tasking machine. *Genes (Basel)*. 11
- Hilf ME, Lewis RS (2016) Transmission and propagation of '*Candidatus Liberibacter asiaticus*' by grafting with individual citrus leaves. *Phytopathology* 106:452–458. <https://doi.org/10.1094/PHYTO-09-15-0221-R>
- Hu Y, Zhang J, Jia H, et al (2014) *Lateral organ boundaries 1* is a disease susceptibility gene for citrus bacterial canker disease. *Proc Natl Acad Sci U S A* 111:E521–E529. <https://doi.org/10.1073/pnas.1313271111>
- Huang C-Y, Araujo K, Sánchez JN, et al (2021) A stable antimicrobial peptide with dual functions of treating and preventing citrus Huanglongbing. *Proc Natl Acad Sci* 118:e2019628118. <https://doi.org/10.1073/pnas.2019628118>
- Huang M, Roose ML, Yu Q, et al (2018) Construction of High-Density Genetic Maps and Detection of QTLs Associated With Huanglongbing Tolerance in Citrus. *Front Plant Sci* 9:1694. <https://doi.org/10.3389/fpls.2018.01694>
- Imai A, Nonaka K, Kuniga T, et al (2018) Genome-wide association mapping of fruit-quality traits using genotyping-by-sequencing approach in citrus landraces, modern cultivars, and breeding lines in Japan. *Tree Genet Genomes* 14:. <https://doi.org/10.1007/s11295-018-1238-0>
- Imai A, Yoshioka T, Hayashi T (2017) Quantitative trait locus (QTL) analysis of fruit-quality traits for mandarin breeding in Japan. *Tree Genet Genomes* 13:1–10. <https://doi.org/10.1007/s11295-017-1162-8>
- Ishino Y, Shinagawa H, Makino K, et al (1987) Nucleotide sequence of the iap gene,

responsible for alkaline phosphatase isoenzyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol* 169:5429–5433. <https://doi.org/10.1128/jb.169.12.5429-5433.1987>

Iwata H, Minamikawa MF, Kajiya-Kanegae H, et al (2016) Genomics-assisted breeding in fruit trees. *Breed. Sci.* 66:100–115

Jansen RC, Nap JP (2001) Genetical genomics: The added value from segregation. *Trends Genet.* 17:388–391

Jia H, Orbovic V, Jones JB, Wang N (2016) Modification of the *PthA4* effector binding elements in Type I *CsLOB1* promoter using Cas9/sgRNA to produce transgenic Duncan grapefruit alleviating Xcc Δ pthA4: DCsLOB1.3 infection. *Plant Biotechnol J* 14:1291–1301. <https://doi.org/10.1111/pbi.12495>

Jia H, Orbović V, Wang N (2019) <scp>CRISPR</scp> -LbCas12a-mediated modification of citrus. *Plant Biotechnol J* 17:1928–1937. <https://doi.org/10.1111/pbi.13109>

Jia H, Wang N (2014) Targeted Genome Editing of Sweet Orange Using Cas9/sgRNA. *PLoS One* 9:e93806. <https://doi.org/10.1371/journal.pone.0093806>

Jia H, Zhang Y, Orbović V, et al (2017) Genome editing of the disease susceptibility gene *CsLOB1* in citrus confers resistance to citrus canker. *Plant Biotechnol J* 15:817–823. <https://doi.org/10.1111/pbi.12677>

Johnson EG, Wu J, Bright DB, Graham JH (2014) Association of ‘*Candidatus Liberibacter asiaticus*’ root infection, but not phloem plugging with root loss on huanglongbing-affected trees prior to appearance of foliar symptoms. *Plant Pathol* 63:290–298. <https://doi.org/10.1111/ppa.12109>

Lima RPM, Curtolo M, Merfa MV, et al (2018) QTLs and eQTLs mapping related to citrandarins’ resistance to citrus gummosis disease. *BMC Genomics* 19:. <https://doi.org/10.1186/s12864-018-4888-2>

Ling P, Yu C, Deng Z, et al (1999) Citrus genome mapping with AFLP markers: In: *Plant and Animal Genome XIII Conference*. San Diego, CA, USA, p P189.

Ling P, Duncan LW, Deng Z, et al (2000) Inheritance of citrus nematode resistance and its linkage with molecular markers. *Theor Appl Genet* 100:1010–1017. <https://doi.org/10.1007/s001220051382>

- Liou PC, Gmitter FG, Moore GA (1996) Characterization of the Citrus genome through analysis of restriction fragment length polymorphisms. *Theor Appl Genet* 92:425–435. <https://doi.org/10.1007/BF00223689>
- Luna E, Pastor V, Robert J, et al (2011) Callose deposition: A multifaceted plant defense response. *Mol. Plant-Microbe Interact.* 24:183–193
- Luro F, Laigret F, Lorieux M, Ollitrault P (1996) Citrus genome mapping with molecular markers : two maps obtained by segregation analysis of progeny of one intergeneric cross. *Proc Int Soc Citric Vol 2*
- Machado MA, Cristofani-Yaly M, Bastianel M (2011) Breeding, genetic and genomic of citrus for disease resistance. *Rev Bras Frutic* 33:158–172. <https://doi.org/10.1590/s0100-29452011000500019>
- Mafra V, Martins PK, Francisco CS, et al (2014) *Candidatus Liberibacter americanus* induces significant reprogramming of the transcriptome of the susceptible citrus genotype
- Martinelli F, Uratsu SL, Albrecht U, et al (2012) Transcriptome profiling of citrus fruit response to huanglongbing disease. *PLoS One* 7:. <https://doi.org/10.1371/journal.pone.0038039>
- Michno JM, Wang X, Liu J, et al (2015) CRISPR/Cas mutagenesis of soybean and *Medicago truncatula* using a new web-tool and a modified Cas9 enzyme. *GM Crops Food* 6:243–252. <https://doi.org/10.1080/21645698.2015.1106063>
- Minamikawa MF, Nonaka K, Kaminuma E, et al (2017) Genome-wide association study and genomic prediction in citrus: Potential of genomics-assisted breeding for fruit quality traits. *Sci Rep* 7:1–13. <https://doi.org/10.1038/s41598-017-05100-x>
- Nica AC, Dermitzakis ET (2013) Expression quantitative trait loci: Present and future. *Philos. Trans. R. Soc. B Biol. Sci.* 368
- Ollitrault P, Terol J, Chen C, et al (2012) A reference genetic map of *C. clementina* hort. ex Tan.; citrus evolution inferences from comparative mapping. *BMC Genomics* 13:593. <https://doi.org/10.1186/1471-2164-13-593>
- Omura M, Shimada T (2016) Citrus breeding, genetics and genomics in Japan. *Breed. Sci.* 66:3–17
- Organização das Nações Unidas para Agricultura e Alimentação: FAO no Brasil | Food and

Agriculture Organization of the United Nations. <http://www.fao.org/brasil/pt/>. Accessed 31 Jan 2021

Pang XM, Hu CG, Deng XX (2007) Phylogenetic relationships within Citrus and its related genera as inferred from AFLP markers. *Genet Resour Crop Evol* 54:429–436. <https://doi.org/10.1007/s10722-006-0005-5>

Peng A, Chen S, Lei T, et al (2017) Engineering canker-resistant plants through CRISPR/Cas9-targeted editing of the susceptibility gene *CsLOB1* promoter in citrus. *Plant Biotechnol J* 15:1509–1519. <https://doi.org/10.1111/pbi.12733>

Permyakova N V., Sidorchuk Y V., Marenkova T V., et al (2019) CRISPR/Cas9-mediated *gfp* gene inactivation in Arabidopsis suspension cells. *Mol Biol Rep* 46:5735–5743. <https://doi.org/10.1007/s11033-019-05007-y>

Pickar-Oliver A, Gersbach CA (2019) The next generation of CRISPR–Cas technologies and applications. *Nat. Rev. Mol. Cell Biol.* 20:490–507

Pitino M, Allen V, Duan Y (2018) *LasA5315* effector induces extreme starch accumulation and chlorosis as *Ca. Liberibacter asiaticus* infection in *Nicotiana benthamiana*. *Front Plant Sci* 9:. <https://doi.org/10.3389/fpls.2018.00113>

Pitino M, Armstrong CM, Cano LM, Duan Y (2016) Transient expression of *Candidatus liberibacter asiaticus* effector induces cell death in *Nicotiana benthamiana*. *Front Plant Sci* 7:982. <https://doi.org/10.3389/fpls.2016.00982>

Raga V, Bernet GP, Carbonell EA, Asins MJ (2012) Segregation and linkage analyses in two complex populations derived from the citrus rootstock Cleopatra mandarin. Inheritance of seed reproductive traits. *Tree Genet Genomes* 8:1061–1071. <https://doi.org/10.1007/s11295-012-0486-7>

Ratner HK, Sampson TR, Weiss DS (2016) Overview of CRISPR-Cas9 biology. *Cold Spring Harb Protoc* 2016:1023–1038. <https://doi.org/10.1101/pdb.top088849>

Rawat N, Kumar B, Albrecht U, et al (2017) Genome resequencing and transcriptome profiling reveal structural diversity and expression patterns of constitutive disease resistance genes in Huanglongbing-tolerant *Poncirus trifoliata* and its hybrids. *Hortic Res* 4:1–8. <https://doi.org/10.1038/hortres.2017.64>

Sampaio Passos O, da Silva Souza J, Costa Bastos D, et al (2019) Citrus Industry in Brazil with

Emphasis on Tropical Areas. In: Citrus - Health Benefits and Production Technology. IntechOpen.

Schadt EE, Monks SA, Drake TA, et al (2003) Genetics of gene expression surveyed in maize, mouse and man. *Nature* 422:297–302. <https://doi.org/10.1038/nature01434>

Song G, Jia M, Chen K, et al (2016) CRISPR/Cas9: A powerful tool for crop genome editing. *Crop J.* 4:75–82

Soratto TAT, Curtolo M, Marengo S, et al (2020) QTL and eQTL mapping associated with host response to *Candidatus Liberibacter asiaticus* in citrandarins. *Trop Plant Pathol.* <https://doi.org/10.1007/s40858-020-00372-7>

Sugiyama A, Omura M, Shimada T, et al (2014) Expression quantitative trait loci analysis of carotenoid metabolism-related genes in citrus. *J Japanese Soc Hortic Sci* 83:32–43. <https://doi.org/10.2503/jjshs1.CH-054>

Swartz DC, Jinek M (2018) Cas9 versus Cas12a/Cpf1: Structure–function comparisons and implications for genome editing. *Wiley Interdiscip. Rev. RNA* 9

Velasquez Guzman JC, Basu S, Rabara R, et al (2018) Liposome Delivery System of Antimicrobial Peptides against Huanglongbing (HLB) Citrus Disease. *Biophys J* 114:266a. <https://doi.org/10.1016/j.bpj.2017.11.1540>

Wang N, Pierson EA, Setubal JC, et al (2017) The *Candidatus Liberibacter*–Host Interface: Insights into Pathogenesis Mechanisms and Disease Control. *Annu Rev Phytopathol* 55:451–482. <https://doi.org/10.1146/annurev-phyto-080516-035513>

Wang N, Trivedi P (2013) Citrus huanglongbing: A newly relevant disease presents unprecedented challenges. *Phytopathology* 103:652–665

Wang Y, Zhou L, Yu X, et al (2016) Transcriptome Profiling of Huanglongbing (HLB) Tolerant and Susceptible Citrus Plants Reveals the Role of Basal Resistance in HLB Tolerance. *Front Plant Sci* 7:933. <https://doi.org/10.3389/fpls.2016.00933>

Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: A revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10:57–63

Weber CA, Moore GA, Deng Z, Gmitter FG (2003) Mapping freeze tolerance quantitative trait loci in a *Citrus grandis* x *Poncirus trifoliata* F1 pseudo-testcross using molecular markers. *J*

Am Soc Hortic Sci 128:508–514. <https://doi.org/10.21273/jashs.128.4.0508>

Weiss KM, Clark AG (2002) Linkage disequilibrium and the mapping of complex human traits. *Trends Genet.* 18:19–24

Xu Q, Chen LL, Ruan X, et al (2013) The draft genome of sweet orange (*Citrus sinensis*). *Nat Genet* 45:59–66. <https://doi.org/10.1038/ng.2472>

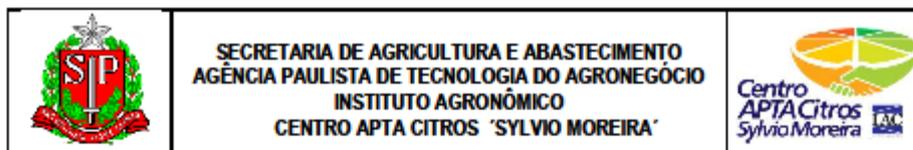
Yu Y, Bai J, Chen C, et al (2017) Identification of QTLs controlling aroma volatiles using a “Fortune” x “Murcott” (*Citrus reticulata*) population. *BMC Genomics* 18:646. <https://doi.org/10.1186/s12864-017-4043-5>

Yu Y, Chen C, Gmitter FG (2016) QTL mapping of mandarin (*Citrus reticulata*) fruit characters using high-throughput SNP markers. *Tree Genet Genomes* 12:. <https://doi.org/10.1007/s11295-016-1034-7>

Zhou G, Jian J, Wang P, et al (2018) Construction of an ultra-high density consensus genetic map, and enhancement of the physical map from genome sequencing in *Lupinus angustifolius*. *Theor Appl Genet* 131:209–223. <https://doi.org/10.1007/s00122-017-2997-y>

6. Anexos

6.1. Declaração de Bioética



DECLARAÇÃO DE BIOÉTICA

Eu, Raquel Luciana Boscarior Camargo, presidente da CIBio do Centro de Citricultura Sylvio Moreira - IAC, o qual possui CQB nº 417/16, declaro que o projeto de tese "Estratégias de identificação de genes alvos e edição de genoma de laranja doce (*Citrus sinensis* L. Osb.) para tolerância ao *Huanglongbing*", desenvolvido pela aluna Maiara Curtolo, no Programa de Pós Graduação em Genética e Biologia Molecular, área de concentração de Genética Vegetal e Melhoramento, da UNICAMP (Universidade Estadual de Campinas), foi realizado dentro das normas de bioética e biossegurança determinadas por essa comissão.

Cordeirópolis, 01 de março de 2021.

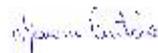
Dra. Raquel Luciana Boscarior Camargo
Pesquisadora (PqC-VI)
Presidente da CIBio do Centro de Citricultura - IAC

6.2. Declaração sobre direitos autorais

Declaração

As cópias de artigos e capítulos de livros de minha autoria ou de minha co-autoria, já publicados ou submetidos para publicação em revista científicas ou anais de congressos sujeitos a arbitragem, que constam na minha Tese de Doutorado, intitulada “Estratégias de identificação de genes alvos e edição de genoma de laranja doce (*Citrus sinensis* L. Osb.) para tolerância ao Huanglongbing”, não infringem os dispositivos da Lei n. 9.610/98, nem o direito autoral de qualquer editora.

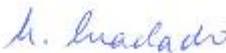
Assinatura:



Nome do(a) aluno(a): Maiara Curtolo

RG: 47110490-5

Assinatura:



Nome do(a) orientador(a): Marcos Antonio Machado

RG: MG 215 101

Campinas, 8 de março de 2021