

Gabriel Garon Carvalho

CAFFEOYL SHIKIMATE ESTERASE (CSE): o elo perdido entre o metabolismo de ácidos clorogênicos e de lignina?

CAFFEOYL SHIKIMATE ESTERASE (CSE): the missing link between the metabolism of chlorogenic acids and lignin?

São Paulo

2023

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE BIOCÊNCIAS
DEPARTAMENTO DE BOTÂNICA

**CAFFEYOYL SHIKIMATE ESTERASE (CSE): o elo perdido entre o
metabolismo de ácidos clorogênicos e de lignina?**

**CAFFEYOYL SHIKIMATE ESTERASE (CSE): the missing link between
the metabolism of chlorogenic acids and lignin?**

Gabriel Garon Carvalho

Orientador: Prof. Dr. Igor Cesarino

Dissertação apresentada para a obtenção
de título de Mestre no Programa de Pós-
Graduação em Botânica, IB/USP

São Paulo
2023

FICHA CATALOGRÁFICA

Garon Carvalho, Gabriel

CAFFEOYL SHIKIMATE ESTERASE (CSE): o elo perdido entre o metabolismo de ácidos clorogênicos e de lignina? / Garon Carvalho Gabriel ; orientador Cesarino Igor -- São Paulo, 2023.

55 p.

Dissertação (Mestrado) -- Instituto de Biociências da Universidade de São Paulo. Ciências Biológicas (Botânica).

1. Biologia de sistemas. 2. Lignina. 3. ácido clorogênico. 4. Expressão heteróloga. 5. Cinética enzimática. I. Cesarino, Igor, orient. Título.

Bibliotecária responsável pela catalogação:
Elisabete da Cruz Neves - CRB - 8/6228

COMISSÃO JULGADORA

Prof(a). Dr(a).

Prof(a). Dr(a).

Prof(a). Dr(a).

Orientador(a)

AGRADECIMENTOS

Agradeço primeiramente aos meus pais, Vivian e Noracyl, por me apoiarem nesse trajeto, sempre me estimulando e me amparando.

Aos meus irmãos Alan e David, por me estimularem, e mesmo sem entender nada da pesquisa, sempre falar sobre ela com orgulho.

A minha companheira Kayne, por ter me suportado e ficado ao meu lado todo o período, me acalmando e me ajudando durante os momentos de estresse.

Ao Prof^o Dr. Igor Cesarino, por toda orientação e paciência. Pelas oportunidades proporcionadas, sempre me ajudando quando necessário, me guiando nessa jornada de ciência. Além de um ótimo profissional, uma ótima pessoa.

Ao Prof^o Dr. Luis Eduardo Soares Netto e ao Thiago Geronimo Pires Alegria, pelo acesso ao laboratório Laboratório Proteínas e Biologia Redox e todo auxílio tanto em conhecimento quanto na prática com experimentos importantes para o trabalho realizado.

Ao LigninLab, grupo de pesquisa coordenado pelo meu orientador, que mesmo com altos e baixos, somos um grupo unido, que nos ajudamos sempre que possível e também narguilamos para estreitar os laços.

A CAPES, pela bolsa de mestrado (processo 88887505942/2020-00).

Ao Departamento de Botânica e ao Instituto de Biociências, por toda a infraestrutura e recursos humanos que, direta- ou indiretamente, permitiram a realização deste trabalho.

RESUMO

Ácidos clorogênicos (CGAs) e lignina são produtos da via dos fenilpropanóides. Enquanto o primeiro está associado à respostas das plantas a estresses, o segundo é um biopolímero depositado nas paredes secundárias e responsável pela sustentação estrutural da planta e impermeabilização dos vasos do xilema, garantindo o transporte de água. Visto que a estrutura de CGAs é quimicamente semelhante àquela de alguns intermediários da via biossintética de lignina e visto que CGAs podem ser convertidos de volta à hidroxicinamoil-CoAs *in vitro*, estes compostos tem sido considerados autênticos intermediários para a produção de lignina. No entanto, ainda não está claro se CGAs são de fato remobilizados para a produção de monômeros de lignina ao longo do desenvolvimento ou em resposta a sinais ambientais. A recente descoberta que caffeoyl shikimate esterase (CSE) é uma enzima central para a biossíntese de lignina abre uma nova possibilidade de explorar a potencial interdependência entre os metabolismos de CGAs e lignina. Considerando que i) o substrato de CSE, cafeoil chiquimato (via de lignina), é estruturalmente muito similar ao cafeoil quinato (CGA); ii) que CSE recombinante de *Arabidopsis* é capaz de utilizar não somente cafeoil chiquimato mas também outros ésteres como substratos *in vitro*; e iii) que o silenciamento de *CSE* em álamo resultou no acúmulo de CGAs nas linhagens transgênicas, a hipótese deste trabalho é que CSE também é capaz de utilizar CGAs como substrato e, conseqüentemente, estaria envolvida no metabolismo de CGAs. Assim, o objetivo deste trabalho foi realizar um estudo funcional de CSE em batata (*Solanum tuberosum*), uma Solanaceae que acumula altos níveis de CGAs. Buscas por BLASTp seguidas de análises filogenéticas demonstraram que *StCSE* ocorre como gene único no genoma de batata, similar a outras espécies com genoma diploide. Análises transcriptômicas *in silico* e por RT-qPCR em diferentes tecidos de batata revelaram uma correlação positiva entre os padrões de expressão de *StCSE* e *StHCT*, um gene biossintético de lignina, enquanto uma correlação negativa foi observada para *StCSE* x *StHQT*, um gene biossintético de CGAs. *StCSE* recombinante foi produzida em *Escherichia coli* e purificada utilizando cromatografia por afinidade para a realização de ensaios enzimáticos *in vitro* com três substratos: cafeoil chiquimato, cafeoil quinato (um CGA monoéster) e 3,5-dicafeoil quinato (um CGA di-éster). Esses ensaios mostraram que *StCSE* não somente é capaz de utilizar CGAs como substrato *in vitro* mas que estes fenólicos são utilizados com maior eficiência do que o próprio cafeoil chiquimato. Esses resultados sugerem que CSE é uma enzima envolvida tanto no metabolismo de lignina quanto no metabolismo de CGAs, potencialmente atuando na mobilização de CGAs para a síntese de monômeros de lignina e/ou outros fenilpropanóides.

Palavras-chave: ácidos clorogênicos, caffeoyl shikimate esterase, expressão heteróloga, lignina, solanáceas.

ABSTRACT

Chlorogenic acids (CGAs) and lignins are products of the phenylpropanoid pathway. While the first is associated with plant responses to stress, the second is a biopolymer deposited in the secondary walls and responsible for the structural support of the plant and waterproofing the xylem vessels, ensuring water transport. Since the structure of CGAs is chemically similar to that of some intermediates of the lignin biosynthetic pathway and since CGAs can be converted back to hydroxycinnamoyl-CoAs *in vitro*, these compounds have been considered authentic intermediates for lignin production. However, it remains unclear whether CGAs are indeed remobilized for the production of lignin monomers throughout development or in response to environmental cues. The recent discovery that caffeoyl shikimate esterase (CSE) is a central enzyme for lignin biosynthesis opens a new possibility to explore the potential interdependence between the metabolisms of CGAs and lignin. Considering that i) the CSE substrate, caffeoyl shikimate (lignin pathway), is structurally very similar to caffeoyl quinate (CGA); ii) that recombinant CSE from *Arabidopsis* is capable of using not only caffeoyl shikimate but also other esters as substrates *in vitro*; and iii) that the silencing of CSE in poplar resulted in the accumulation of CGAs in transgenic lines, the hypothesis of this work is that CSE is also capable of using CGAs as a substrate and, consequently, would be involved in the metabolism of CGAs. Therefore, the aim of this work was to perform a functional study of CSE in potato (*Solanum tuberosum*), a Solanaceous species that accumulates high levels of CGAs. BLASTp searches followed by phylogenetic analyzes revealed that *StCSE* occurs as a single gene in the potato genome, similar to other species with a diploid genome. *In silico* and RT-qPCR transcriptomic analyzes using different potato tissues revealed a positive correlation between the expression patterns of *StCSE* and *StHCT*, a lignin biosynthetic gene, whereas a negative correlation was observed for *StCSE* x *StHQT*, a CGA biosynthetic gene. Recombinant *StCSE* was produced in *Escherichia coli* and purified using affinity chromatography to perform *in vitro* enzymatic assays with three substrates: caffeoyl shikimate, caffeoyl quinate (a CGA monoester) and 3,5-dicaffeoyl quinate (a CGA di-ester). These assays demonstrated that *StCSE* is not only capable of using CGAs as a substrate *in vitro*, but that these phenolics are used more efficiently than caffeoyl shikimate. These results suggest that CSE is an enzyme involved in both lignin metabolism and CGA metabolism, potentially in the mobilization of CGAs toward the synthesis of lignin monomers and/or other phenylpropanoids.

Key-words: chlorogenic acids, caffeoyl shikimate esterase, heterologous expression, lignin, solanaceous.

SUMÁRIO

RESUMO	5
ABSTRACT	6
1. INTRODUÇÃO	8
1.1 Ácidos Clorogênicos (CGAs)	9
1.2 Lignina.....	10
1.3 As vias biossintéticas de CGAs e lignina estão intimamente ligadas	12
1.4 CGAs são remobilizados para a produção de lignina?	15
1.5 CAFFEYOYL SHIKIMATE ESTERASE (CSE): o elo perdido entre o metabolismo de ácidos clorogênicos e de lignina?.....	16
2. HIPÓTESE	19
3. OBJETIVOS.....	20
4. MATERIAL E METODOS	21
4.1 Identificação de StCSE e análise filogenética	21
4.2 Análise de expressão gênica <i>in silico</i>	21
4.3 Análise de expressão gênica por RT-qPCR	22
4.4 Expressão heteróloga	22
4.5 Ensaio enzimático <i>in vitro</i>	24
5. RESULTADOS.....	26
5.1 O genoma de batata apresenta uma única cópia de <i>CSE</i>	26
5.2 <i>StCSE</i> é expresso em tecidos em lignificação e apresenta padrão de expressão semelhante à <i>StHCT</i> e distinto de <i>StHQT</i>	27
5.3 <i>StCSE</i> foi produzida em <i>E. coli</i> e purificada com sucesso	34
5.4 Ensaio enzimático <i>in vitro</i> demonstram o consumo de CGAs por <i>CSE</i>	35
6. DISCUSSÃO.....	40
6.1 Evolução dos genes <i>CSE</i> em um possível contexto de neofuncionalização	40
6.2 Lignina e ácidos clorogênicos: rotas biossintéticas interligadas e relações funcionais entre <i>CSE</i> , <i>HCT</i> e <i>HQT</i>	44
6.3 A eficiência catalítica de <i>CSE in vitro</i> suportam seu papel no metabolismo de CGAs.....	46
7. CONCLUSÃO	49
8. BIBLIOGRAFIA.....	50

1. INTRODUÇÃO

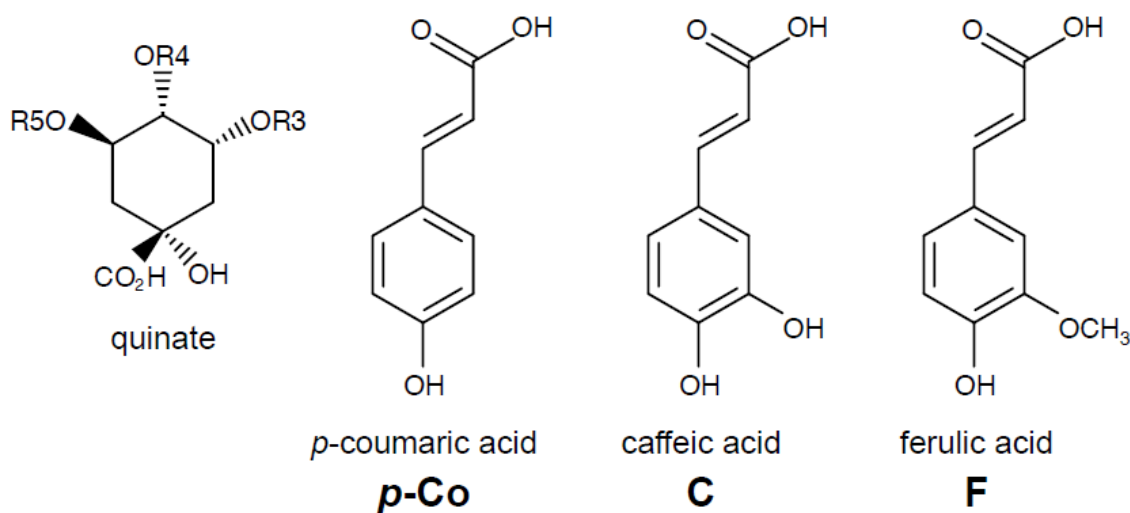
Há cerca de 450 milhões de anos atrás, durante o Período Ordoviciano Superior, as plantas pioneiras iniciaram a colonização do meio terrestre. Durante a conquista desse nicho, essas plantas foram confrontadas com uma série de novos desafios, incluindo a maior exposição à danosa radiação ultra-violeta (UV), falta de suporte estrutural (anteriormente providenciado pela flutuabilidade), disponibilidade errática de água e, eventualmente, coevolução de patógenos e herbívoros (Weng and Chapple, 2010). Dentre as diversas adaptações fisiológicas que evoluíram nestas plantas, o surgimento de vias metabólicas especializadas, classicamente conhecidas como metabolismo secundário, foi de crucial importância para o seu sucesso adaptativo. Particularmente, a via dos fenilpropanóides teve significativo destaque neste cenário evolutivo. O desenvolvimento da habilidade de desaminar o amino ácido fenilalanina e hidroxilar o anel aromático dos metabólitos resultantes resultou na capacidade de produzir e acumular fenilpropanóides simples com pico de absorbância máxima na região da luz UV (280-320 nm). Assim, esses novos atributos permitiram que as plantas pioneiras, especialmente seus vulneráveis esporos haploides, se tornassem resistentes à radiação UV e, conseqüentemente, fez a vida no meio terrestre possível (Weng and Chapple, 2010). Subseqüentemente, a emergência da capacidade de impregnar as paredes celulares com o polímero fenólico lignina permitiu a grande diversificação fenotípica e a real conquista dos mais diferentes ambientes terrestres por parte das plantas vasculares (Renault et al., 2019).

A via dos fenilpropanóides é responsável pela produção de um enorme conjunto de metabólitos secundários gerados a partir de alguns intermediários e produtos da via do ácido chiquímico. O principal precursor é o amino ácido aromático fenilalanina, que é inicialmente desaminado pela enzima fenilalanina amônia liase (PAL) gerando ácido trans-cinâmico. Subseqüentemente, a enzima cinamato 4-hidroxilase (C4H), localizada na membrana do retículo endoplasmático, realiza a hidroxilação da posição 4 do anel aromático do ácido cinâmico para gerar o ácido *p*-cumárico que, por sua vez, será utilizado como substrato para a produção de *p*-cumaroil-CoA pela atividade da 4-cumarato:CoA ligase (4CL). Estes três primeiros passos enzimáticos formam a chamada via geral dos fenilpropanóides (Tohge et al., 2013). Ademais, a diversificação da via ocorre a partir do metabólito *p*-cumaroil-CoA por meio da combinação de atividades de redutases, oxigenases e transferases, gerando todas as classes de fenilpropanóides, incluindo monolignóis (monômeros de lignina), flavonóides, cumarinas, estilbenos, auronas, entre outras (Vogt, 2010). Esses compostos participam não somente do desenvolvimento normal das plantas mas também contribuem para sua interação com o ambiente, sendo essenciais para as respostas das plantas a estresses bióticos e abióticos (Ferrer et al., 2008).

Ácidos clorogênicos (CGAs) e lignina são dois dos principais produtos da via dos fenilpropanóides. Enquanto CGAs são reportados como metabólitos importantes durante as respostas a estresses, a lignina é um biopolímero depositado nas paredes secundárias de tipos celulares especializados, garantindo força mecânica e hidrofobicidade aos tecidos de sustentação e condutores de seiva (Volpi e Silva et al., 2019). Apesar de largamente estudados, ainda existem muitas lacunas de conhecimento acerca do metabolismo destes compostos/polímeros.

1.1 Ácidos Clorogênicos (CGAs)

O termo “ácidos clorogênicos” é utilizado em referência a um grande grupo de compostos formados pela esterificação de um ácido trans-hidroxicinâmico e o ácido quínico. CGAs são caracterizados quimicamente de acordo com a identidade, número e posição de resíduos acil (**Figura 1.**; Clifford et al., 2017). O CGA individual mais comumente encontrado em plantas é o ácido 5-*O*-cafeoilquínico (5-CQA), mas outros mono-ésteres derivados de ácido caféico (ácidos cafeoil quínicos, CQAs), ácido *p*-cumárico (ácidos *p*-cumaroil quínicos, *p*CoQA) e ácido ferúlico (ácidos feruloil quínicos, FQA) também são relativamente bem distribuídos (Clifford, 2000). Ademais, também existem di-ésteres, tri-ésteres, tetra-ésteres e di-ésteres mistos entre ácido caféico e ferúlico ou mesmo envolvendo permutações de resíduos de ácido caféico com resíduos de ácidos alifáticos. Alguns destes compostos mais complexos e menos comuns são específicos de algumas famílias ou mesmo espécies de plantas. Finalmente, a lista de CGAs clássicos pode ser estendida quando considera-se também conjugados galoil do ácido quínico e conjugados cinamoil de derivados do ácido quínico como ácido chiquímico, butilquinato e ácido *muco*-quínico (Clifford, 2000; Clifford et al., 2017b). Dentre as funções biológicas dos CGAs em plantas destacam-se a proteção contra radiação UV (Clé et al., 2008) e seu papel na defesa química contra herbívoros e patógenos (Martínez et al., 2017; Kundu and Vadassery, 2019). Estes compostos são largamente encontrados em alimentos como tomate, berinjela, ameixas, mirtilos e alcachofras, e em bebidas como café e mate (Clifford et al., 2017b). Devido às suas propriedades antioxidantes, CGAs são considerados compostos nutracêuticos que trazem diversos benefícios à saúde humana, dentre eles ação protetora contra diabetes (Islam, 2006), hipertensão (Yamaguchi et al., 2008) e doenças degenerativas (Nakajima et al., 2007). Ademais, os efeitos benéficos dos CGAs são geralmente maiores em compostos com múltiplos anéis fenólicos, provavelmente devido ao maior número de hidroxilas que, por sua vez, determinam sua atividade antioxidante (Wang et al., 2003; Islam, 2006).



Name and Abbreviation	Number	R3	R4	R5
3- <i>O</i> -caffeoylquinic acid (3-CQA)	I	C	H	H
5- <i>O</i> -caffeoylquinic acid (5-CQA)	II	H	H	C
4- <i>O</i> -caffeoylquinic acid (4-CQA)	III	H	C	H
3- <i>O</i> - <i>p</i> -coumaroylquinic acid (3- <i>p</i> CoQA)	IV	<i>p</i> -CO	H	H
5- <i>O</i> - <i>p</i> -coumaroylquinic acid (5- <i>p</i> CoQA)	V	H	H	<i>p</i> -CO
4- <i>O</i> - <i>p</i> -coumaroylquinic acid (4- <i>p</i> CoQA)	VI	H	<i>p</i> -CO	H
3- <i>O</i> -feruloylquinic acid (3-FQA)	VII	F	H	H
5- <i>O</i> -feruloylquinic acid (5-FQA)	VIII	H	H	F
4- <i>O</i> -feruloylquinic acid (4-FQA)	IX	H	F	H
3,4-di- <i>O</i> -caffeoylquinic acid (3,4-diCQA)	X	C	C	H
3,5-di- <i>O</i> -caffeoylquinic acid (3,5-diCQA)	XI	C	H	C
4,5-di- <i>O</i> -caffeoylquinic acid (4,5-diCQA)	XII	H	C	C

Figura 1. Estrutura do ácido quínico e dos ácidos hidroxicinâmicos associados que juntos formam diferentes ácidos clorogênicos (numeração de acordo com a IUPAC).

1.2 Lignina

A lignina é um complexo polímero fenólico depositado majoritariamente na parede celular secundária de tipos celulares especializados como elementos traqueais e fibras do xilema, garantindo o suporte mecânico e o transporte de água e nutrientes ao longo da planta (Barros et al., 2015). A lignina também está envolvida na defesa da planta contra herbívoros e patógenos, podendo ter a produção ativada em sítios específicos em respostas a tais esses estresses (Cesarino, 2019). Paredes celulares lignificadas também podem ser encontradas nas células da esclereídes, no tecido endodérmico de raízes, e em células especializadas das anteras e de alguns tipos de frutos (Zhong and Ye, 2009). Este biopolímero é produzido pelo acoplamento oxidativo dos álcoois hidroxicinamil (de maneira geral chamados de “monolignóis”): álcoois *p*-cumaril, coniferil e sinapil, que diferem entre si pelo grau de metoxilação do anel aromático (Boerjan et

al., 2003). Ademais, metabólitos pertencentes a outras 10 classes de compostos fenólicos foram identificados como autênticos monômeros ocorrendo naturalmente no polímero, incluindo hidroxicinamatos, hidroxicinamaldeídos, estilbenos e tricina (Vanholme et al., 2019), embora ocorram em quantidades bem menores do que os monolignóis canônicos.

A biossíntese de monolignóis (**Figura 2**) ocorre no citoplasma em um processo que envolve mais de 10 passos catalíticos, realizados tanto por enzimas solúveis quanto por citocromos P450 associados à membrana do retículo endoplasmático (Ferreira et al., 2019). Utilizando a fenilalanina como precursor (assim como a maioria dos fenilpropanóides), as características estruturais dos monolignóis são estabelecidas por uma série de modificações do anel aromático, via reações de hidroxilação e metilação, e por transformações da cauda propano, via reações de esterificação e redução (Ralph et al., 2019). Após a sua síntese, os monolignóis são secretados para o apoplasto e previamente oxidados por peroxidases e/ou lacases antes de participarem de um processo de acoplamento oxidativo que dará origem a um polímero racêmico e oticamente inativo (Ralph et al., 2019). Uma vez incorporados ao polímero, estes monolignóis dão origem às subunidades de lignina *p*-hidroxifenil (H), guaiacil (G) e siringil (S), respectivamente, ligadas entre si, de maneira aleatória, tanto por ligações não-condensadas do tipo C-O-C éter como β -aryl éteres (β -O-4) quanto por ligações condensadas do tipo C-C como resinols (β - β) e fenilcumaranas (β -5). Vale ressaltar que após a geração dos radicais, a polimerização é puramente química, não controlada por enzimas ou proteínas (Ralph et al., 2019). Finalmente, o conteúdo e composição da lignina varia de acordo com o tecido, tipo celular, fase do desenvolvimento e espécie vegetal (Cesarino et al., 2016). Por exemplo, a lignina de gimnospermas é composta basicamente por unidades G com apenas traços de subunidades H, enquanto a lignina de eudicotiledôneas apresenta quantidades similares de G e S. Ademais, a lignina de gramíneas apresenta maiores níveis da subunidade H e incorporação de tricina (del Río et al., 2012).

O interesse pelo estudo do metabolismo da parede celular secundária tem crescido de maneira notória, visto que, como principal constituinte da biomassa vegetal, pode ter grande impacto na produção biocombustíveis e biomateriais nas biorrefinarias (Vanholme et al., 2013a). Nesse aspecto, a lignina é a principal responsável pela recalcitrância da biomassa vegetal (Chen and Dixon, 2007; Van Acker et al., 2013), sendo este um fenômeno que ocorre devido a estrutura física e as composições químicas da parede celular, que limitam o acesso de enzimas hidrolíticas aos seus substratos polissacarídicos, limitando a liberação de açúcares fermentáveis para produção de biocombustíveis e biomateriais. A lignina imobiliza as celulases, impedindo o acesso aos seus substratos, dificultando a sacarificação, processo no qual ocorre a hidrólise enzimática dos polissacarídeos de parede celular (Vanholme et al., 2012). Sendo

assim, o estudo acerca do metabolismo da parede celular secundária, em específico da lignina, tem ganhado uma enorme atenção. Contudo, o conhecimento acerca desta via metabólica ainda possui enormes lacunas.

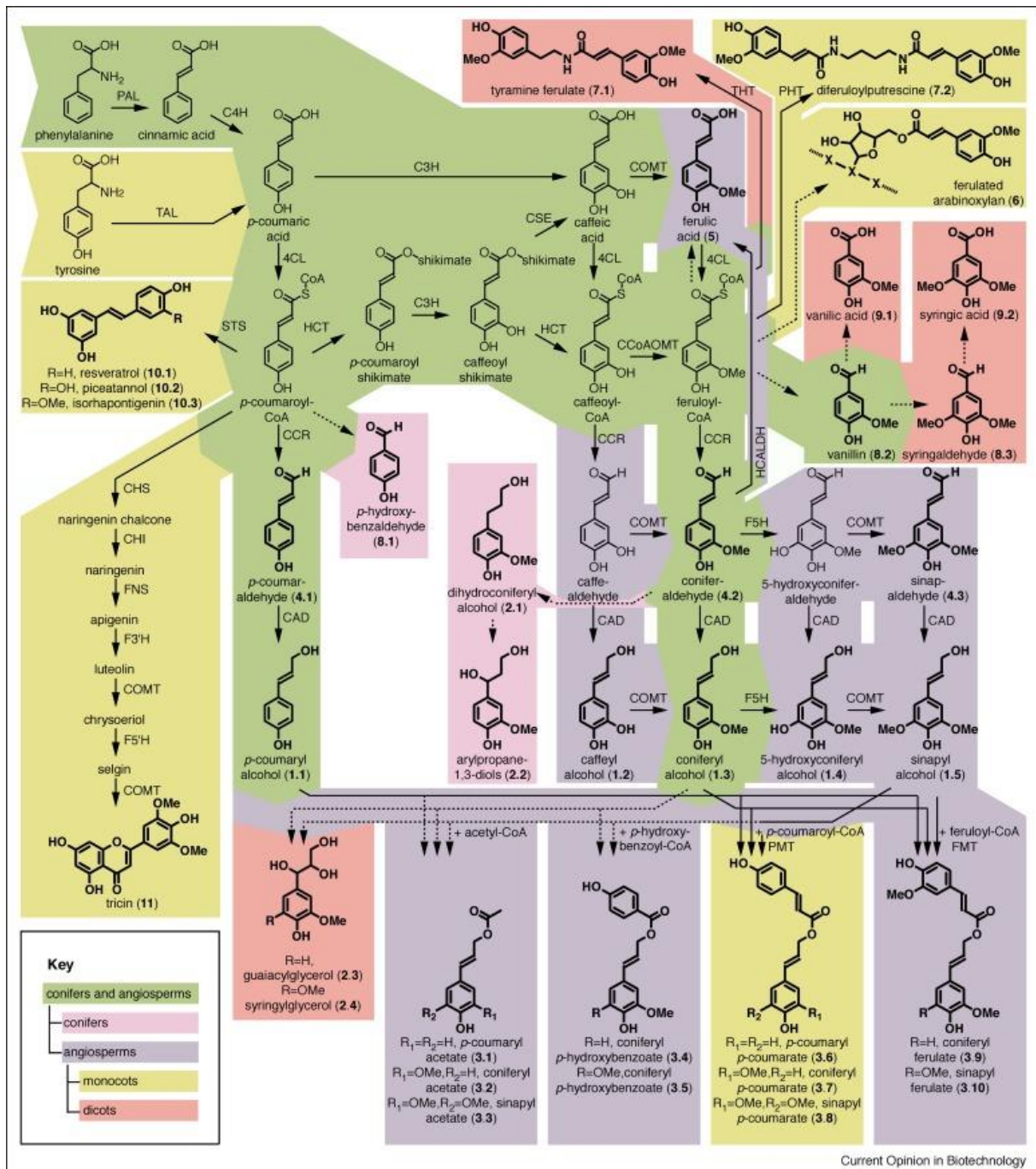


Figura 2: Representação esquemática da via de biossíntese de lignina. Fonte: Vanholme et al, 2019.

1.3 As vias biossintéticas de CGAs e lignina estão intimamente ligadas

As vias metabólicas que resultam na produção de CGAs e lignina compartilham uma série de enzimas e intermediários em comum (Figura 3). Ao final da via geral dos fenilpropanóides, o composto-chave *p*-cumaroyl-CoA pode ser utilizado para a formação da subunidade H da

lignina, ou seguir por ao menos dois caminhos paralelos, os chamados “*shikimate shunt*” e “*quininate shunt*” (**Figura 3**). O fluxo de carbono pelo “*shikimate shunt*” (Via em azul, **Figura 3**) inicia-se com a esterificação de *p*-cumaroil-CoA com chiquimato pela atividade da enzima hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase (HCT), produzindo *p*-cumaroil chiquimato que é posteriormente 3'-hidroxilado por *p*-cumaroil chiquimato 3'-hidroxilase (C3'H) resultando na produção de cafeoil chiquimato. Subsequentemente, a conversão de cafeoil chiquimato em cafeoil-CoA pode ser mediada ou pela reação reversa de HCT (Hoffmann et al., 2003; Hoffmann et al., 2004) ou pela combinação das atividades de cafeoil chiquimato esterase (CSE) e 4CL (**Figura 3**; Vanholme et al., 2013). A contribuição relativa de cada uma dessas vias em plantas que possuem *CSE* no seu genoma ainda é desconhecida. Cafeoil-CoA é então canalizado para a produção das unidades G e S de lignina por meio da atividade da enzima cafeoil-CoA 3-Ometiltransferase (CCoAOMT). Embora a maquinaria necessária para a produção de lignina tenha se diversificado intensamente ao longo da evolução das plantas, gerando diversos quimio-tipos do polímero, o “*shikimate shunt*” parece bem conservado e ativo em todas as traqueófitas (Renault et al., 2019).

A produção de CGAs pode ocorrer por 3 rotas metabólicas distintas. Na primeira, a enzima hydroxycinnamoyl-CoA:quininate hydroxycinnamoyl transferase (HQT) catalisa a formação de CGA a partir de ácido quínico e cafeoil-CoA, que por sua vez é produzido ao final do “*shikimate shunt*” (Via em azul, **Figura 3**). É claro, portanto, que uma das principais rotas para a produção de CGAs ocorre via “*shikimate shunt*”, que também é a rota principal para a biossíntese das subunidades G e S da lignina. Alternativamente, um cinamoil glicosídeo pode ser utilizado como intermediário ao invés de cafeoil-CoA pela ação da enzima hydroxycinnamoyl-D-glucose:quininate hydroxycinnamoyl transferase (HCGQT) (Via em verde, **Figura 3**; Barros et al., 2019; Xiong et al., 2019). Finalmente, na terceira rota proposta, HQT catalisa a formação de *p*-cumaroil quinato, que é subsequentemente hidroxilado por C3'H para produzir CGA (Via em vermelho, **Figura 3**). Esta terceira rota caracteriza o chamado “*quininate shunt*”. Visto que HQT é capaz de produzir CGAs tanto via cafeoil-CoA (“*shikimate shunt*”) quanto via *p*-cumaroil-CoA (“*quininate shunt*”), é difícil determinar qual rota biossintética de CGAs ocorre preferencialmente *in vivo*. Provavelmente, a rota majoritária depende da espécie vegetal e do tamanho dos pools relativos de cafeoil-CoA e *p*-cumaroil-CoA em um determinado tecido (Niggeweg et al., 2004). Dados de atividade *in vitro* de enzimas recombinantes sugerem que em tomate, alcachofra e chicória a rota majoritária para a produção de CGAs é via cafeoil-CoA (Niggeweg et al., 2004; Sonnante et al., 2010; Moglia et al., 2014; Legrand et al., 2016), enquanto em switchgrass (*Panicum virgatum*) a rota preferencial ocorre via *p*-cumaroil-CoA (Escamilla-Treviño et al., 2014). É importante notar que, embora HCT e HQT catalisem a

mesma reação (a formação de ésteres hidroxicinamoil com ácido chiquímico e quínico), estas enzimas apresentam especificidades distintas: enquanto HCT possui uma afinidade maior para o ácido chiquímico, HQT utiliza o ácido quínico de forma mais eficiente. Ademais, ambas são capazes de realizar a reação reversa, produzindo os ésteres de CoA a partir dos ésteres de chiquimato e/ou quinato (**Figura 3**; Hoffmann et al., 2003; Niggeweg et al., 2004; Escamilla-Treviño et al., 2014). Devido à sua maior afinidade por chiquimato e visto que seu silenciamento *in planta* resulta em menores teores de lignina sem nenhum efeito nos níveis de CGAs, HCT é considerada uma enzima envolvida especificamente no metabolismo de lignina (Hoffmann et al., 2003; Hoffmann et al., 2004; Shadle et al., 2007; Lallemand et al., 2012; Vanholme et al., 2013b; Escamilla-Treviño et al., 2014). Por outro lado, visto que HQT tem preferência por quinato e seu silenciamento afeta drasticamente os níveis de CGAs mas não resulta em alterações na deposição de lignina, sugere-se que esta enzima esteja especificamente envolvida no metabolismo de CGAs (Niggeweg et al., 2004; Sonnante et al., 2010; Lallemand et al., 2012; Escamilla-Treviño et al., 2014; Payyavula et al., 2015; Valiñas et al., 2015).

Urge ressaltar que nem todas as plantas que acumulam CGAs apresentam um ortólogo de *HQT/HCGQT* em seu genoma, sugerindo que vias ou enzimas alternativas podem ser empregadas. Em *Populus trichocarpa*, estudos baseados em QTLs sugerem uma função para PtHCT2 na biossíntese de CGAs (Zhang et al., 2018). Embora análises bioquímicas não tenham sido geradas para PtHCT2, estudos anteriores demonstraram que PtHCT1 e PtHCT6 são os membros da família envolvidos na biossíntese de lignina (Wang et al., 2018). Portanto, é possível que PtHCT2 apresente propriedades catalíticas distintas das PtHCTs envolvidas em lignificação e, na ausência de HQT em *Populus*, consiga utilizar quinato eficientemente e, consequentemente, seja a enzima responsável pela produção de CGAs. Um exemplo ainda mais surpreendente é em switchgrass, em que nem um ortólogo de HCT nem de HQT está aparentemente envolvido na síntese de CGAs. Nesta espécie, uma hidroxicinamoil transferase chamada de PvHCT-like 1 e filogeneticamente distante de HCTs e HQTs, apresenta preferência por ácido quínico como acceptor acil, funcionando como uma típica HQT (Escamilla-Treviño et al., 2014). No entanto, apesar desta evidência bioquímica, o envolvimento desta PvHCT-like 1 na biossíntese de CGAs ainda não foi confirmado por meio de genética reversa.

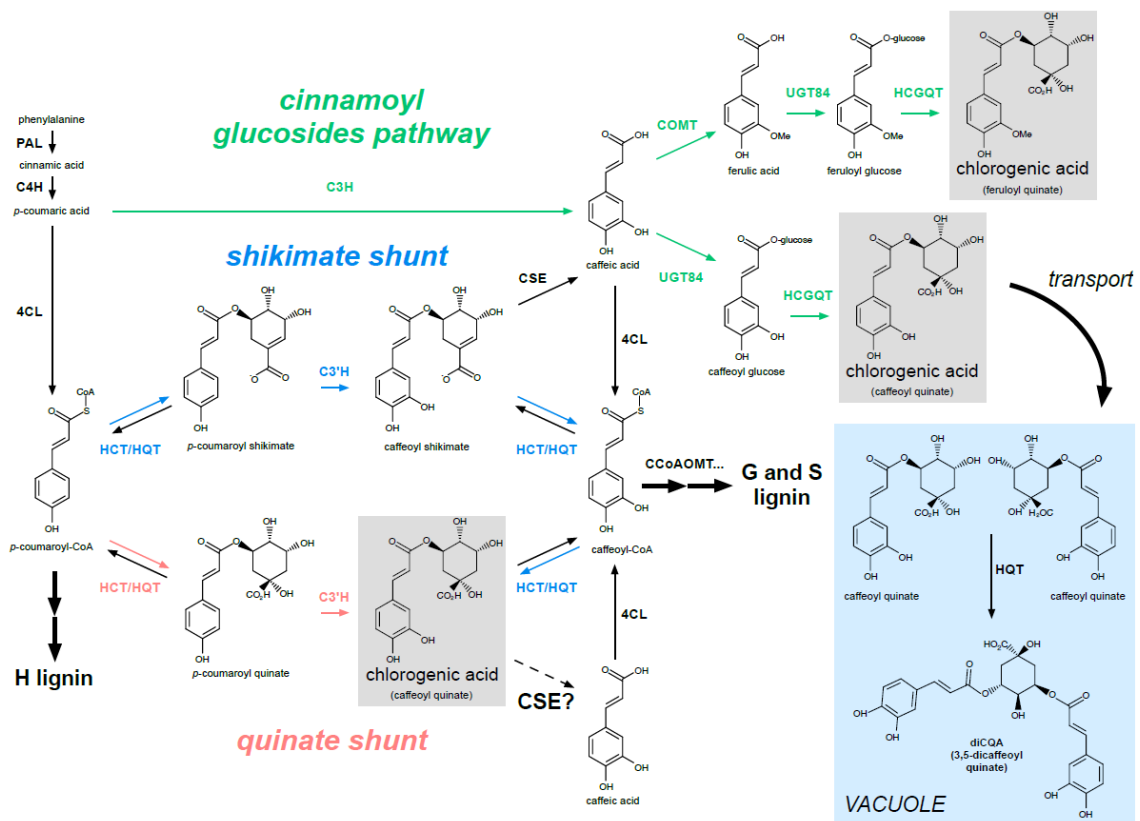


Figura 3: Vias metabólicas propostas para a biossíntese de CGAs e lignina. O “shikimate shunt” está marcado em azul e constitui uma rota comum entre CGAs e monolignols G e S. O “quinate shunt” está marcado em vermelho. A via alternativa que utiliza cinamoil glicosídeos como intermediários está marcada em verde. Em destaque com seta pontilhada, o reação proposta para CSE na conversão de CGA em ácido caféico. Fonte: Volpi e Silva et al, 2019.

1.4 CGAs são remobilizados para a produção de lignina?

CGAs são frequentemente sugeridos como autênticos intermediários da via biossintética de lignina (Aerts and Baumann, 1994; Joët et al., 2009; Zhang et al., 2018). No entanto, embora CGAs sejam estruturalmente muito semelhantes a alguns intermediários da via de lignina e apesar destes dois produtos do metabolismo fenólico compartilharem passos metabólicos em comum, evidências genéticas da remobilização de CGAs para a síntese de monolignóis são escassas. Isto ocorre porque as duas enzimas-chave tanto na produção de CGAs quanto de lignina, HCT e HQT, apresentam atividade reversível. Dessa forma, é difícil avaliar o efeito do silenciamento de HCT ou HQT na remobilização de CGAs (i.e. produção de CoA ésteres) uma vez que tal modificação genética necessariamente afeta a reação direta (*upstream*) que inicia tanto o “shikimate shunt” quanto o “quinate shunt” (i.e. produção de ésteres de chiquimato/quinate) (**Figura 3**).

Apesar de resultados discrepantes serem encontrados na literatura, diversos estudos sugerem uma potencial função de CGAs como esqueletos de carbono para a produção de

monolignóis. Obviamente, o fato de HQT (e, em menor escala, HCT) ser capaz de converter CGAs em hidroxicinamoil-CoAs *in vitro* sugere uma função semelhante *in vivo*. De fato, todas as HQTs recombinantes de diferentes espécies avaliadas até o momento apresentam parâmetros cinéticos eficientes para esta reação reversa (Niggeweg et al., 2004; Sonnante et al., 2010; Lallemand et al., 2012; Liu et al., 2019). Outra evidência é a dinâmica inversa de acúmulo de CGAs e lignina em diferentes tecidos de algumas espécies vegetais. No colmo de switchgrass, o teor de lignina aumenta de acordo com a maturidade dos entrenós, enquanto o oposto é observado para o teor de CGAs (Shen et al., 2009). O mesmo padrão foi observado durante o desenvolvimento de entrenós de sorgo (Ferreira et al., 2022). Durante o desenvolvimento inicial de plântulas de *Capsicum annuum*, a diminuição do conteúdo de CGAs em hipocótilos é acompanhada por um significativo aumento na deposição de lignina (Diaz et al., 1997). Um padrão semelhante foi observado durante o desenvolvimento de plântulas de *Coffea arabica* (Aerts and Baumann, 1994). Durante as primeiras 10 semanas de desenvolvimento, o conteúdo de CGAs em cotilédones em desenvolvimento é reduzido em mais de 70% sem nenhuma recuperação observada em outros órgãos da plântula. Concomitantemente, essa conspícua queda nos teores de CGAs coincidiu com um aumento na quantidade de polímeros fenólicos ligados à parede dos cotilédones, fazendo com que os autores sugerissem que CGAs são armazenados nas sementes de café para serem subsequentemente utilizados para a deposição de lignina durante o desenvolvimento cotiledonar (Aerts and Baumann, 1994). Finalmente, durante um estudo acerca do desenvolvimento dos grãos de café, Joët et al (2009) observaram que o programa transcricional relacionado à via dos fenilpropanóides muda da biossíntese de CGAs para a biossíntese de lignina ao longo do desenvolvimento do endosperma. A diminuição do conteúdo de CGAs foi acompanhada por picos de expressão de genes que codificam enzimas agindo *downstream* ao cafeoil quinato, dando suporte à ideia que CGAs funcionam como pools de carbono para a deposição de lignina ao final do desenvolvimento do endosperma (Joët et al., 2009). Apesar destas evidências, a remobilização de CGAs para a produção de monolignóis é contestada por experimentos que demonstraram que o silenciamento de *HCT* em plantas que acumulam CGAs ainda assim resulta em menores níveis de lignina e, conseqüentemente, em efeitos negativos no desenvolvimento da planta (Hoffmann et al., 2004; Nelson et al., 2017; Wang et al., 2018), sugerindo que o fluxo de carbono de CGAs para monolignóis é limitado.

1.5 CAFFEOYL SHIKIMATE ESTERASE (CSE): o elo perdido entre o metabolismo de ácidos clorogênicos e de lignina?

A recente descoberta de que CSE é uma enzima central na via biossintética de lignina mudou nossa visão acerca deste importante ramo do metabolismo dos fenilpropanóides. A chamada

“reação reversa” de HCT (i.e. conversão de cafeoil chiquimato em cafeoil-CoA) foi imediatamente aceita como um autêntico passo catalítico *in vivo* mesmo que tenha sido demonstrada somente *in vitro* (Hoffmann et al., 2003). No entanto, mesmo que mensurável *in vitro*, esta reação é extremamente ineficiente (Hoffmann et al., 2003; Escamilla-Treviño et al., 2014; Wang et al., 2014). Ademais, extratos protéicos preparados a partir de colmos de switchgrass não são capazes de formar cafeoil-CoA a partir de cafeoil chiquimato e CoA, mas produzem ácido caféico (produto de CSE), sugerindo que a reação reversa de HCT não ocorre *in vivo* e que este passo metabólico ocorre via CSE, ao menos nesta espécie (Escamilla-Treviño et al., 2014). Embora evidência genética do envolvimento de CSE na biossíntese de lignina tenha sido gerada para *Arabidopsis thaliana* (Vanholme et al., 2013b), *Medicago truncatula* (Ha et al., 2016b) e álamo (Saleme et al., 2017), algumas espécies, especialmente gramíneas, não apresentam ortólogos de CSE em seus genomas (Ha et al., 2016b). Aparentemente, plantas que apresentam um ortólogo *bona fide* de CSE em seus genomas também apresentam significativa atividade de CSE em extratos protéicos produzidos a partir de tecidos em lignificação (Ha et al., 2016b; Saleme et al., 2017), sugerindo que CSE é essencial para a deposição de lignina nestas espécies.

É muito interessante observar que o substrato de CSE, cafeoil chiquimato, é estruturalmente muito similar ao cafeoil quinato (CGA; **Figura 4**) e, portanto, é inevitável sugerir que CSE seria capaz de utilizar CGA como substrato endógeno. A principal diferença estrutural entre os ácidos chiquímico e quínico é a presença de uma ligação dupla entre C1 e C2 no ácido chiquímico e um grupo hidroxil em C1 no ácido quínico, resultando em diferentes geometrias no anel polirol que, conseqüentemente, afeta a especificidade de HCT/HQT (Lallemand et al., 2012). Até o momento, a função biológica de CSE foi avaliada somente no contexto da biossíntese de lignina. No entanto, algumas evidências sugerem uma potencial função de CSE na remobilização de CGAs. Primeiro, cafeoil chiquimato não foi o único substrato aceito por CSE recombinante de *Arabidopsis* em ensaios *in vitro*, sugerindo uma especificidade por substratos mais ampla para esta enzima. Compostos estruturalmente semelhantes como *p*-cumaroil chiquimato, cafeoil hexose e feruloil hexose também foram metabolizados (**Figura 5.**; Vanholme et al., 2013). Estes dados sugerem que, de modo similar a HCT e HQT, CSE é potencialmente capaz de utilizar ésteres de chiquimato e quinato como substratos, embora provavelmente com diferentes eficiências. Segundo, em um estudo acerca do metabolismo de CGAs em diferentes variedades de batata, Valiñas et al. (2015) observaram que a variedade com maiores níveis de lignina em seus tubérculos também apresentava os menores níveis de CGA e maior expressão de CSE. Terceiro, em álamo, uma espécie arbórea que produz CGAs independentemente de HQT (Zhang et al., 2018), o silenciamento de CSE

resultou em menor deposição de lignina e maiores níveis de unidades H no polímero, confirmando seu envolvimento no metabolismo de lignina (Saleme et al., 2017). Interessantemente, experimentos de determinação de perfis fenólicos mostrou um significativo acúmulo de dois CGAs, cafeoil quinato e feruloil quinato, nas linhagens transgênicas quando comparadas ao tipo selvagem. Embora perturbações na via dos fenilpropanóides frequentemente resulta em mudanças gerais no perfil fenólico, é inevitável sugerir que o acúmulo destes compostos observado nestas linhagens seria resultado de uma menor remobilização de CGAs para a produção de monolignóis devido à menor atividade de CSE. Posteriormente, o mesmo grupo demonstrou que a perda de função de CSE (via CRISPR-Cas9) em álamo resultou no acúmulo não somente de cafeoil chiquimato mas de *p*-coumaroil, 5-hidroxiferuloil e feruloil chiquimato. Ao realizarem um ensaio de cinética enzimática, foi demonstrada a capacidade de PtCSE em metabolizar todos estes ésteres de chiquimato.

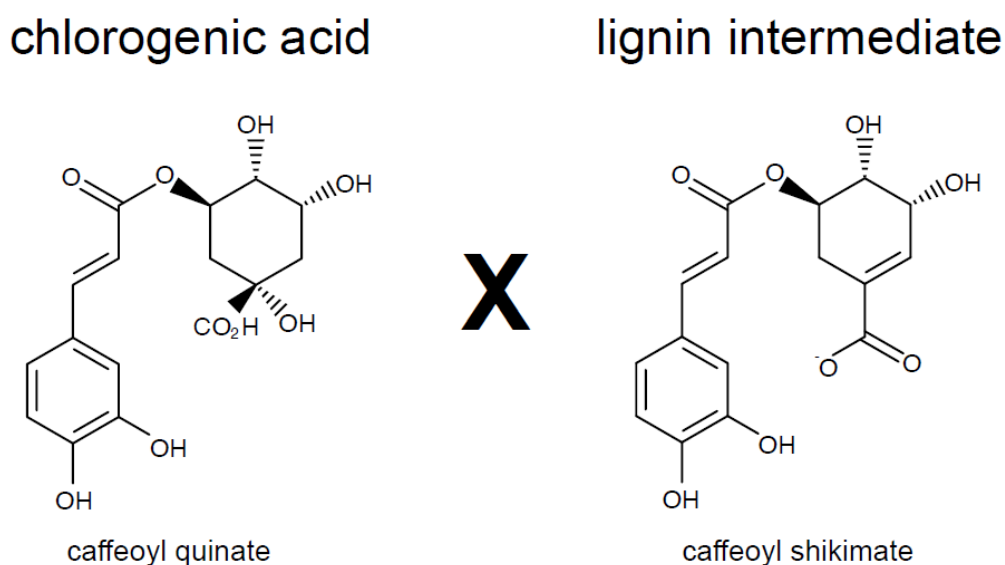


Figura 4: Estrutura dos metabólitos cafeoil quinato (CGA) e cafeoil chiquimato (considerado um autêntico intermediário da via de biossíntese de lignina).

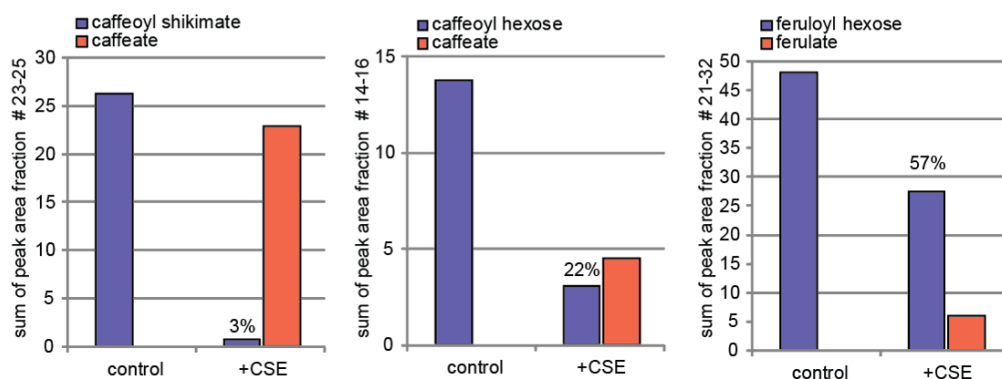


Figura 5: Ensaio enzimático demonstrando, que mesmo com menor eficiência, CSE é capaz de metabolizar não só ésteres de chiquimato, mas também cafeoil hexose e feruloil hexose.

2. HIPÓTESE

Considerando que i) o substrato de CSE, cafeoil chiquimato, é estruturalmente muito similar ao cafeoil quinato (CGA); ii) CSEs recombinantes de diferentes espécies foram capazes de utilizar não somente cafeoil chiquimato mas também outros ésteres como substratos *in vitro*; e iii) o silenciamento de *CSE* em álamo tenha gerado um acúmulo de CGAs nas linhagens transgênicas, a hipótese deste trabalho é que CSE também é capaz de utilizar CGAs como substrato e, conseqüentemente, estaria envolvida no metabolismo de CGAs (e potencialmente na sua remobilização à lignina) em espécies que acumulam tais compostos.

3. OBJETIVOS

O objetivo geral do projeto é realizar a caracterização funcional de *CSE* em batata (*Solanum tuberosum*) e avaliar seu possível envolvimento no metabolismo de lignina e de CGAs. Os objetivos específicos são:

- Caracterizar *StCSE* em batata, por meio de análise filogenética e
- Determinar o padrão de expressão gênica de *StCSE* *in silico* e *in vivo* por RT-qPCR
- Realizar a expressão heteróloga de *StCSE* em *Escherichia coli*, seguido de purificação da enzima recombinante;
- Realizar testes enzimáticos de *StCSE* recombinante com seu putativo substrato endógeno (cafeoil chiquimato) e diferentes ácidos clorogênicos
- Determinar os parâmetros cinéticos para todos os substratos testados

4. MATERIAL E METODOS

4.1 Identificação de StCSE e análise filogenética

Para identificar a sequência de CSE no genoma de batata (*Solanum tuberosum*), foi utilizada a sequência proteica de AtCSE (At1g52760) de *Arabidopsis thaliana* como iscas em buscas BLASTp na plataforma Phytozome (<http://phytozome.jgi.doe.gov/pz/portal.html>). As sequências de proteínas de CSEs de uma ampla gama de espécies de plantas também foram recuperadas de seu proteoma correspondente na plataforma Phytozome v12 usando buscas por BLASTp e a sequência de aminoácidos de AtCSE de *Arabidopsis* como isca. As sequências de proteínas foram alinhadas usando a versão online do servidor MAFFT v7 (<https://mafft.cbrc.jp/alignment/server/>) e o alinhamento resultante foi empregado para gerar uma árvore filogenética pelo método de máxima verossimilhança (*Maximum likelihood*) com a versão online do IQ-TREE (<http://iqtree.cibiv.univie.ac.at>). O melhor modelo de substituição foi determinado automaticamente pelo programa e a análise de suporte de ramos foi realizada a partir de 1.000 réplicas usando os testes SH-aLRT e bootstrap ultrarrápido. Uma árvore preliminar foi editada usando o programa FigTree (v.1.4.4) e depois formatada como figura pronta para publicação usando o Graphic para Mac. As árvores foram desenhadas em escala e os comprimentos dos ramos representam o número de substituições de aminoácidos por sítio.

4.2 Análise de expressão gênica *in silico*

O padrão de expressão dos genes *CSE* (*StCSE*, PGSC0003DMT400014303), *HQT* (*StHQT*, PGSC0003DMT400029079) e *HCT* (*StHCT*, PGSC0003DMT400036695) em diferentes tecidos de batata e em diferentes condições experimentais foi avaliado por meio da plataforma ePlant (http://bar.utoronto.ca/eplant_potato/). Após buscas individuais utilizando o número de acesso de cada gene, esta plataforma gera mapas de expressão nos diferentes tecidos da planta, que podem ser baixados como figuras. Além disso, os dados de expressão brutos foram diretamente baixados e utilizados para uma análise de correlação de Pearson, para verificar potenciais correlações no padrão de expressão entre os genes-alvo. O cálculo de correlação de Pearson foi realizado no Excel utilizando funções disponíveis no próprio programa. Correlações par-a-par (i.e. *StCSE* x *StHCT*, *StCSE* x *StHQT*, *StHCT* x *StHQT*) foram inicialmente determinadas com a função =PEARSON entre dois grupos de dados. Os valores dos coeficientes de correlação Pearson resultantes foram convertidos em valores t (estatística de teste) por meio da função =(r*SQRT(n-2))/(SQRT(1-r^2)), em que “r” representa o coeficiente de correlação Pearson e “n” representa o número de dados de expressão (i.e. número de tecidos em que o par de genes foi analisado). Finalmente, o valor *P* foi calculado a partir dos valores t usando o comando =2*T.DIST.RT(ABS(x),deg_freedom), em que “ABS” significa *absolute*

value of a number (valor absoluto), “deg_freedom” significa grau de liberdade (calculado por $n - 2$) e “x” equivale ao valor calculado pela equação anterior (valor t).

4.3 Análise de expressão gênica por RT-qPCR

Primeiramente, as plantas controle da variedade Desiree foram propagadas em garrafas (120×80 mm) contendo o meio Murashige and Skoog (1962, MS) suplementado com 20 gL⁻¹ de sacarose e 0,5 gL⁻¹ de ácido 2-(N-morfolino)etanosulfônico (MES), ajustado a um pH 5,7 antes da autoclavagem (120°C, 120min). As plantas foram mantidas em um fotoperíodo de 16h claro/8h escuro, com temperatura de 23°C/21°C. Após 2-3 semanas, quando elas demonstraram um bom desenvolvimento do sistema radicular, as plântulas foram transferidas para o solo e mantidas em uma estufa, com o fotoperíodo de 14h claro/10h escuro com a temperatura 23°C/21°C. O RNA foi extraído (Oñate-Sánchez and Vicente-Carbajosa, 2008) de folhas jovens, folhas maduras, casca do tubérculo, polpa do tubérculo e base do caule, e tratado com DNase (Roche). Após a extração, o cDNA foi sintetizado a partir de 2 µg de RNA total, utilizando a transcriptase reversa do vírus de mieloblastose aviária e primer oligo-(dT)15 (Promega) de acordo com as instruções do fabricante. A análise da expressão por RT-qPCR foi realizada em diferentes tecidos de batata para os *StCSE*, *StHQT* e *StHCT*. O experimento foi realizado usando LightCycler®480 System (Roche) nos parâmetros 0 s a 95°C, 45 ciclos de 95°C por 10 s, 55°C por 20 s min e 72°C por 30 s, utilizando a enzima LightCycler® 480 SYBR Green I Master (Roche). Além disso, foi utilizado como gene de referência para a padronização o gene constitutivo de Ubiquitina de batata (PGSC0003DMT400011938), e os cálculos de expressão relativa foram realizados de acordo com o método 2- $\Delta\Delta$ CT (Livak and Schmittgen, 2001).

4.4 Expressão heteróloga

Para realizar os ensaios enzimáticos, primeiramente o gene *StCSE* foi sintetizado já no plasmídeo para expressão em procariotos *pET-28a(+)* (**Figura 6**) junto à empresa GenOne (Rio de Janeiro, Brasil). Com o plasmídeo em mãos, foi realizada a transformação de duas linhagens de *Escherichia coli*, DH5 α (para estoque do plasmídeo em freezer -80°C) e BL21 (DE3) Δ SlyD, que contém o plasmídeo *pRARE2*, linhagem própria para a expressão de proteínas de eucariotos.

Para a expressão, utilizou-se meio de cultura Luria Bertani (triptona a 10 gL⁻¹, extrato de levedura 5 gL⁻¹ e cloreto de sódio 5 gL⁻¹) e kanamicina na concentração de 15 µg/mL. Primeiramente um pré-inóculo de 50 mL foi feito (mantido a 37°C, 200 rpm), e então um inóculo contendo 500 mL do meio de cultura Luria Bertani. A cultura foi mantida a 37 °C até atingir a OD 600 nm de 0,6, momento em que induziu-se a expressão da proteína pela adição

de isopropyl 1-thio- β -D-galactopyranoside (IPTG, 500 mM). A cultura foi mantida a 18°C, 200 rpm overnight.

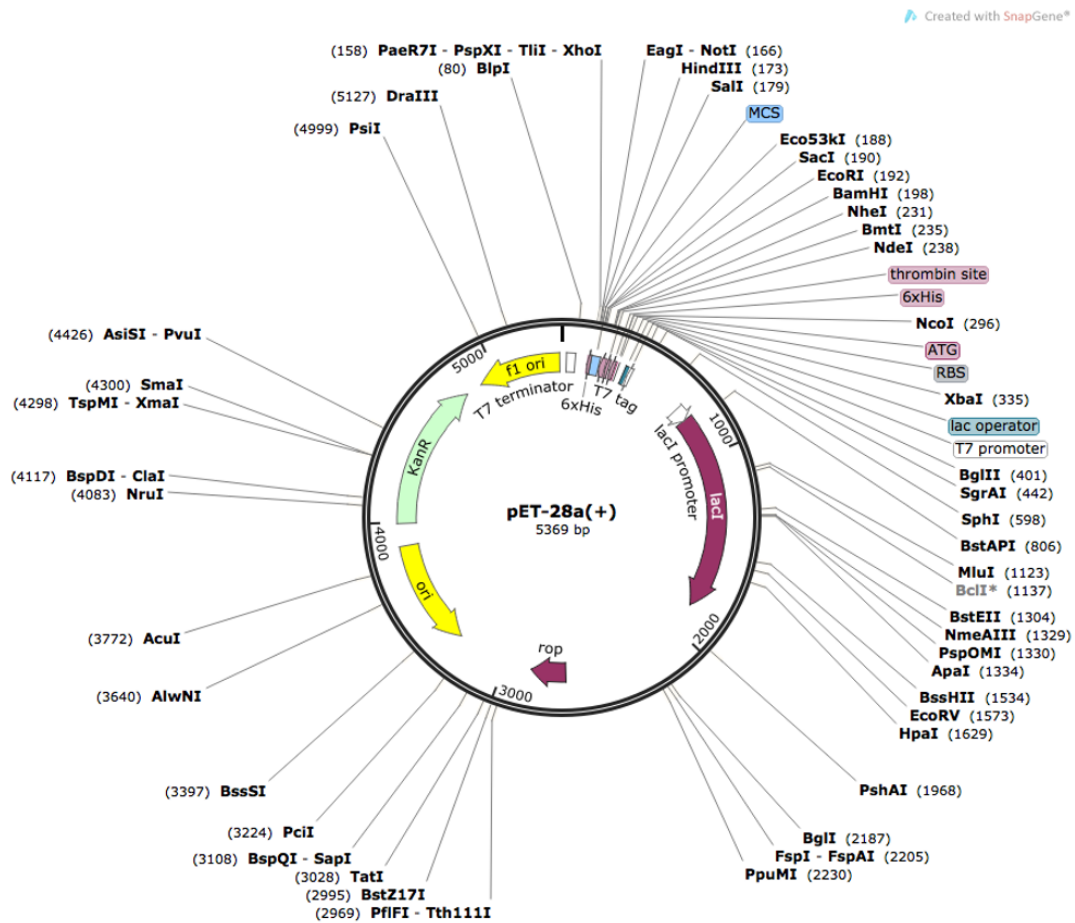


Figura 6. Mapa do vetor *pET-28a(+)*, utilizado para expressão heteróloga de StCSE. Os diferentes módulos funcionais do vetor estão destacados em cores distintas.

Após a expressão, realizou-se o preparo da proteína para a etapa de purificação. O inóculo então foi centrifugado a 5000 rpm por 10 minutos, o sobrenadante descartado e as bactérias ressuspensas em tampão fosfato NaPi 2M (fosfato de sódio monobásico 0,2 M com fosfato de sódio dibásico 0,2 M, pH 7,4), contendo um inibidor de protease SIGMAFAST™ Protease Inhibitor Tablets (Sigma Aldrich). Com isso, foi realizada a etapa da lise celular por sonicação (4 minutos de sonicação, 20 segundos a cada minuto), e adicionado streptomomicina (1% de concentração final) para precipitação dos ácidos nucleicos. Realizou-se então a filtração da proteína utilizando uma bomba a vácuo, seguida da etapa de purificação por IMAC (Immobilized Metal Ion Affinity Chromatography) em coluna de níquel. Uma bomba de fluxo líquido foi utilizada para adicionar fosfato de níquel (NiSO_4 0,1 M) na coluna contendo a resina Ni-NTA Superflow (1mL) e posteriormente a solução de proteína filtrada (a proteína recombinante apresenta cauda de histidina que, por sua vez, possui afinidade por níquel). Com

a coluna pronta, utilizou-se o sistema AKTA para a purificação. As amostras oriundas da purificação foram coletadas em eppendorfs de 2 mL, e preparadas para SDS-PAGE pela adição de 14 µL de cada amostra, 5 µL de BenchMark™ Protein Ladder (Invitrogen) e 1 µL de ditionitriol (DTT), seguindo de fervura por 3 minutos. A análise SDS-PAGE foi realizada em gel de poliacrilamida 14% (Laemmli, 1970). As frações que apresentaram bandas ao redor de 37 kDa no gel (e, portanto, que potencialmente apresentavam a proteína recombinante) foram misturadas, concentradas e submetidas à uma nova etapa de purificação utilizando uma coluna MonoQ, que tem como princípio a troca iônica. Posteriormente, uma análise por SDS-PAGE foi realizada com os produtos dessa nova purificação. As amostras que apresentaram a proteína de interesse tiveram a absorbância medida no espectrofotômetro e, com a equação fornecida pela Lei de Lambert-Beer ($A = \alpha lc$, sendo A a absorbância, α o coeficiente de extinção, l o comprimento de luz e c a concentração), a concentração foi determinada para assim serem estocadas sem glicerol em um freezer -20°C para o ensaio enzimático.

4.5 Ensaio enzimático *in vitro*

As reações enzimáticas foram realizadas com 3 substratos, caffeoil chiquimato (CS, fornecido pelo Dr. Ruben Vanholme, PSB/UGent, Bélgica), 3,5-dicafeoilquínico (3,5-DQA, SMB00131, Sigma Aldrich) e 5-cafeoilquínico (CQA, C3878, Sigma Aldrich). Todos foram ressuspensos em dimetilsulfóxido (DMSO, concentração estoque de 50 mM). As reações foram preparadas em 50 mM de tampão Tris-HCl pH 7,5 contendo 2 µg de enzima recombinante a uma temperatura de 25°C.

Primeiramente, as reações foram analisadas por UHPLC-ESI-(TQD-WATERS) em parceria com o Prof. Dr. Paulo Mazzafera (Unicamp). O experimento foi realizado com o propósito de confirmar a utilização de ácidos clorogênicos por parte de StCSE recombinante, por meio da detecção da formação do produto ácido cafeico. Para isso, utilizou-se uma coluna C18 (50x 2.1mm partícula 1.7 micra) que tem como princípio a hidrofobicidade de hidrocarbonetos para realizar a separação da proteína. Como fase móvel A utilizou-se água ácida (0.1% ácido fórmico) e fase móvel B acetonitrila. Uma vazão de 0,2 mL/min foi utilizada. Os parâmetros da corrida foram 10% de fase móvel B no tempo inicial, aos 4 minutos 25% de fase móvel B, dos 8 aos 9 minutos 100% de fase móvel B, aos 9.1 minutos até 10 minutos 10% de fase móvel B. Foram injetados 7 µL de amostra, a uma temperatura de 35 °C. Para a espectrometria de massas, foi realizada uma ionização por ESI (eléctron spray ionization), modo negativo com capilar -3 KV, com cone -35 V. A temperatura da fonte foi 150 °C e temperatura de secagem foi 350 °C. Para o MS/MS de íons selecionados foi selecionada uma energia de colisão de 23 V.

Para os ensaios de cinética enzimática, cada substrato foi utilizado em 5 concentrações diferentes, cada concentração foi avaliada em 6 pontos de tempo diferentes, e para cada ponto de tempo foram realizadas 3 réplicas técnicas de 50 μL (totalizando 15 reações por substrato em cada concentração). As reações foram interrompidas pela adição de 5 μL de ácido acético glacial. Para o substrato CQA, as concentrações utilizadas foram 20, 50, 100, 500 e 1000 μM , e os pontos de tempo de reação foram 0, 5, 15, 30, 60 e 90 minutos. Para os substratos 3,5-DQA e CS, as concentrações utilizadas foram 20, 50, 100, 300 e 500 μM , e os pontos de tempo de reação foram 0, 5, 10, 15, 20 e 30 minutos. Após o término das reações com o uso de ácido acético, elas foram armazenadas em um freezer $-20\text{ }^{\circ}\text{C}$, e posteriormente avaliadas por cromatografia líquida de alta eficiência (HPLC). Para tal, foi utilizada uma coluna C18 ($150 \times 4,5\text{ mm}$, Luna® 3 μm), com temperatura mantida a $35\text{ }^{\circ}\text{C}$. Além disso, os tempos de corrida foram de 20 minutos. O comprimento de onda de leitura foi estabelecido entre 280 e 350 nm, pois o produto da reação (ácido cafeico) absorve em 320 nm. A fase móvel utilizada foi o ácido fórmico 10 mM (solvente A) e como solvente acetonitrila com ácido fórmico 10 mM (solvente B), uma vazão de 1 mL/min, tendo como parâmetro da corrida do tempo inicial - 12:50 min 25% solvente B, 12:50 - 15 min 100% solvente B, 15– 17:50 min 10% solvente B e 17:50 – 20 min 0% solvente B.

Após o término das reações enzimáticas e tendo o acesso às áreas do pico correspondente a cada substrato e produto nos cromatogramas resultantes, os valores foram convertidos em concentração (com base na equação da reta do produto), e para cada uma, um gráfico foi produzido para determinar a velocidade de formação de produto (utilizando o coeficiente angular da equação da reta gerada). Com a concentração dos produtos utilizados e a velocidade de reação correspondente, foi feito um gráfico de Michaelis-Menten, utilizando o GraphPad Prism 8, para cada uma das triplicatas de cada produto. Por fim, uma média foi tirada para os três valores obtidos para de K_m e K_{cat} para cada produto. Além disso, a eficiência da reação para cada produto foi calculada utilizando a média de K_{cat} dividida pela média de K_m .

5. RESULTADOS

5.1 O genoma de batata apresenta uma única cópia de *CSE*

Primeiramente foi realizada uma busca por gene(s) *CSE* no genoma de batata (*Solanum tuberosum*) utilizando a ferramenta blastP da plataforma digital Phytozome (<https://phytozome-next.jgi.doe.gov/>) e as sequências proteicas de *CSE* já caracterizados em *Arabidopsis thaliana*, *Medicago truncatula* e *Populus trichocarpa* como iscas. Como na maioria das espécies analisadas até o momento, foi encontrado somente uma cópia de *CSE* no genoma da batata (StCSE, PGSC0003DMP400009874). Com o objetivo de avaliar as relações filogenéticas na família e, em especial, do gene *StCSE* com seus putativos ortólogos em outras espécies, foi realizada uma análise filogenética utilizando o método de máxima verossimilhança e sequências proteicas de um set abrangente de espécies vegetais (**Figura 7**). Como demonstrado anteriormente, a maioria das monocotiledôneas não possui um ortólogo “bonafide” (verdadeiro) de *CSE* em seus genomas. Apenas algumas espécies deste grupo possuem gene(s) *CSE*, como arroz (*Oryza sativa*) e switchgrass (*Panicum virgatum*), que formam um grupo bem definido e separado das *CSEs* de eudicotiledôneas. No clado de *CSEs* de eudicotiledôneas, StCSE ficou mais próximo da SICSE de tomate (que ainda não foi caracterizado). As relações filogenéticas observadas para as proteínas *CSE* refletiram, em grande parte, as próprias relações filogenéticas entre as espécies utilizadas. Na árvore, as proteínas *CSE* correspondentes aos genes já caracterizados e com função em lignificação foram marcados em negrito.

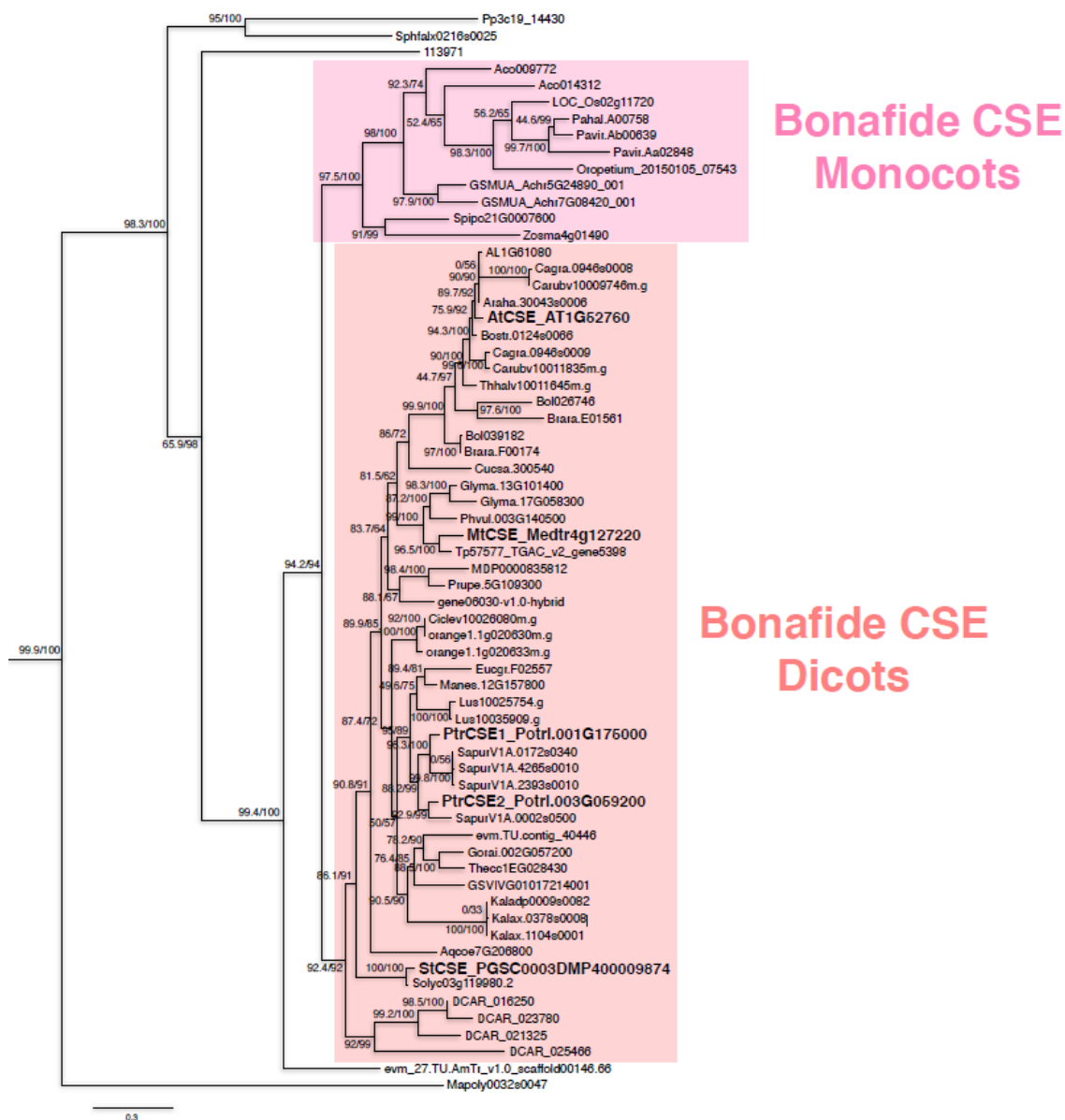
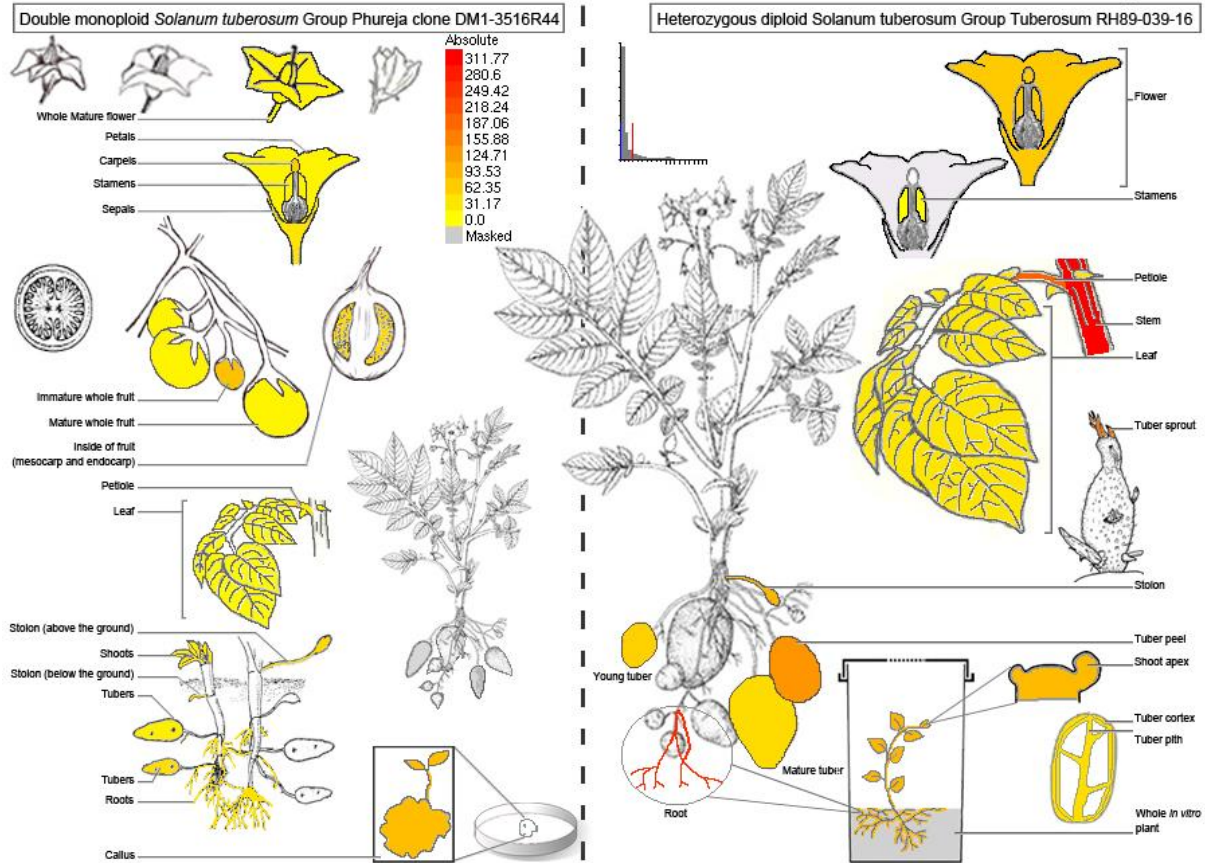


Figura 7: Análise filogenética comparativa da família de CSE de diversas espécies de Angiospermas. Membros pertencentes a espécies Eudicotiledôneas estão marcadas em vermelho, enquanto representantes de monocotiledôneas em rosa. As proteínas previamente caracterizadas de *Arabidopsis thaliana*, *Medicago truncatula* e álamo, assim como CSE de batata, estão marcadas em negrito. Números nos ramos representam os testes de Shimodaira–Hasegawa-like approximate likelihood ratio test (SH-aLRT) e ultrafast bootstrap approximation approach (UFBoot), respectivamente.

5.2 *StCSE* é expresso em tecidos em lignificação e apresenta padrão de expressão semelhante à *StHCT* e distinto de *StHQT*

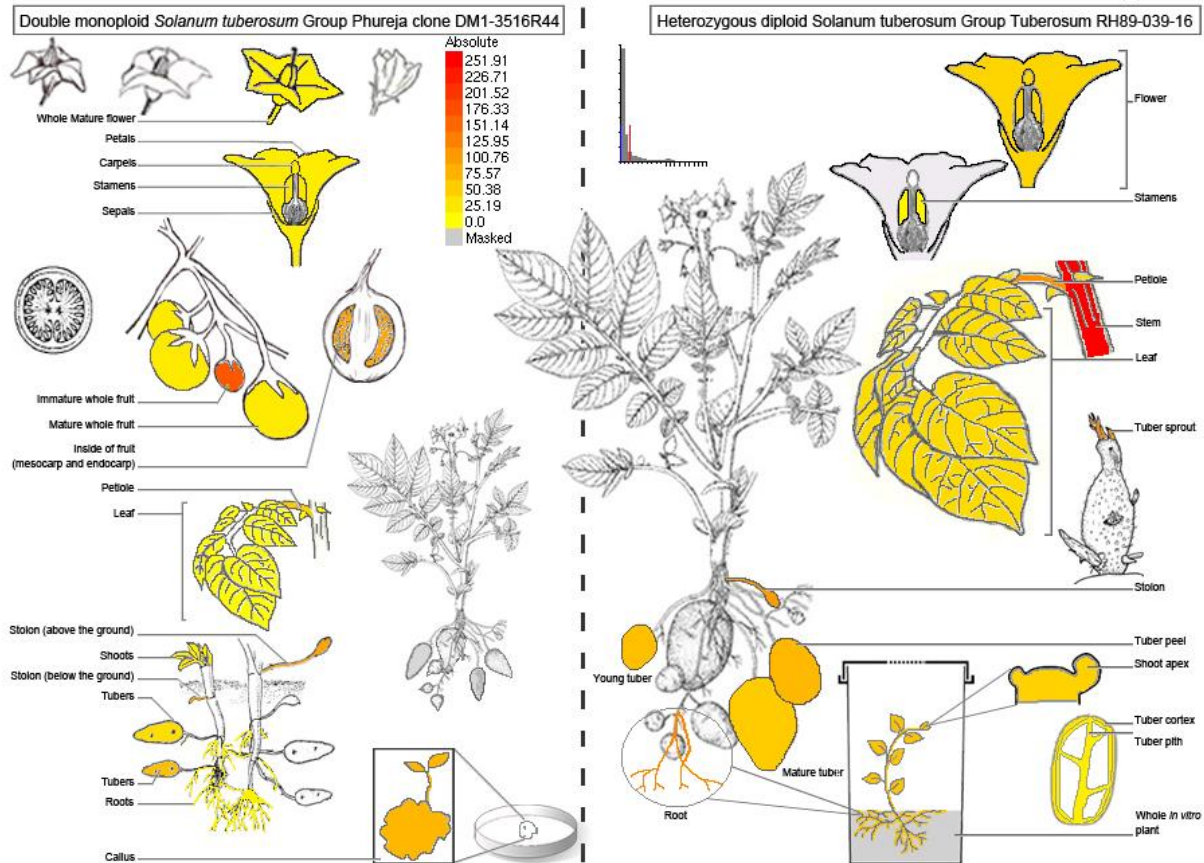
O próximo passo foi determinar o padrão de expressão de *StCSE* e outros genes relacionados aos chamados “*shikimate/quininate shunt*” (Volpi e Silva et al., 2019) em diferentes tecidos de batata, de modo a determinar se estes genes apresentam padrão de expressão semelhantes e, portanto, estariam potencialmente envolvidos nos mesmos processos biológicos. Além de *StCSE* (PGSC0003DMT400014303), foram analisados os genes *StHCT*

(PGSC0003DMT400036695), envolvido no *shikimate shunt* e na biossíntese de lignina, e *StHQT* (PGSC0003DMT400029079), envolvido no *quinate shunt* e no metabolismo de ácidos clorogênicos. Para a análise de expressão gênica, foram utilizadas duas abordagens distintas. Primeiramente, obteve-se o padrão de expressão *in silico* dos genes utilizando a plataforma ePlant (http://bar.utoronto.ca/eplant_potato/). Nesta plataforma, realiza-se uma busca por palavras-chave utilizando o número de acesso do gene de interesse, e um mapa de expressão visual é gerado a partir de dados de transcriptômica em larga escala, indicando os tecidos no qual o gene é expresso e os respectivos níveis de expressão. Com os mapas gerados, é possível notar que a expressão de *StCSE* (**Figura 8**) é maior no caule da batata, resultado também encontrado para *StHCT* (**Figura 9**) e esperado para genes relacionados com o metabolismo de lignina, uma vez que o caule é um tecido com alta deposição deste polímero fenólico. Observa-se também expressão relativamente alta de *StCSE* em raízes e casca dos tubérculos. Este dado é interessante, uma vez que a casca dos tubérculos é um tecido que produz altos teores de ácidos clorogênicos durante o seu desenvolvimento (Valiñas et al., 2015). O padrão de *StHCT* foi aparentemente semelhante ao de *StCSE*, porém com menos níveis de expressão, como pode ser observado na escala de intensidade de cores apresentadas no mapa (**Figura 9**).



Potato eFP Browser by A.Silva. Data from Transcriptome of the Reference Potato Genome *Solanum tuberosum* Group Phureja Clone DM1-3 516R44: Massa et al., 2011. PLoS One 6(10):e26801 and Genome sequence and analysis of the tuber crop potato: The Potato Genome Sequencing Consortium, 2011. Nature 475:10.1038/nature10158. RNA seq data are normalized by the FPKM method. The developmental tissues represent vegetative and reproductive organs from greenhouse-grown plants. Shoots and roots were from *in vitro*-grown plants. Callus (10-11 week old) was derived from leaves and stems. Images provided by Kelly Zarka, Michigan State University and adapted from: McIntosh, Thomas P. *The Potato: its History, Varieties, Culture and Diseases*. Great Britain: Oliver and Boyd, 1927. Vreugdenhil, Dick, et al. *Potato Biology and Biotechnology: Advances and Perspectives*. United Kingdom: Elsevier, 2007. Florida Center for Instructional Technology, 2009. *Setting up and running a school garden: A manual for teachers, parents and communities*. FAO Corporate Document Repository, 2005. *Solanum tuberosum L.*, Société Française d'Ethnopharmacologie, 2010.

Figura 8: Mapa de expressão do gene *StCSE* gerado na plataforma ePlant (http://bar.utoronto.ca/eplant_potato/). As cores indicam o tecido onde o gene foi expresso, sendo que é utilizado um gradiente que vai do amarelo ao vermelho (no qual o amarelo indica pouca expressão e o vermelho indica uma alta expressão). Nota-se que apenas no caule este gene é bastante expresso.

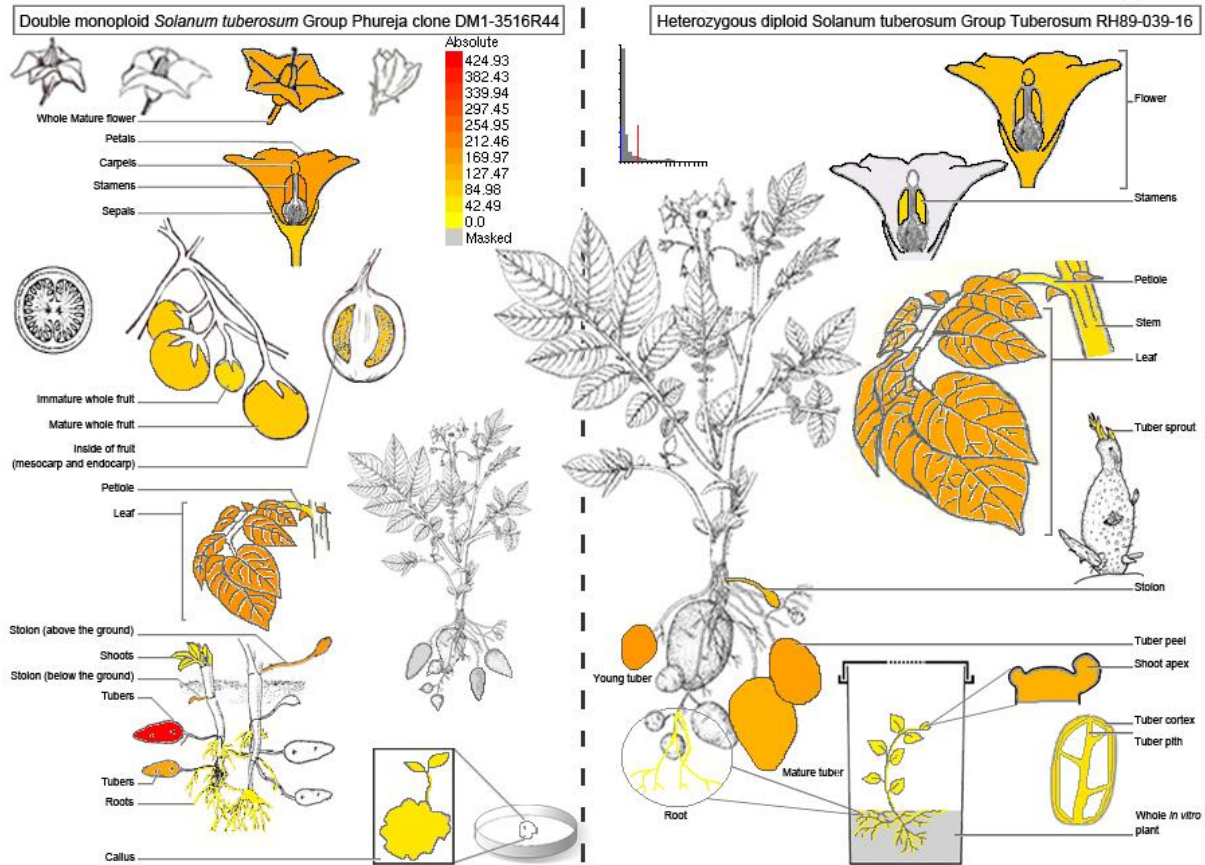


Potato eFP Browser by A.Silva. Data from Transcriptome of the Reference Potato Genome *Solanum tuberosum* Group Phureja Clone DM1-3 516R44: Massa et al., 2011. PLoS One 6(10):e26801 and Genome sequence and analysis of the tuber crop potato: The Potato Genome Sequencing Consortium, 2011. Nature 475:10.1038/nature10158. RNA seq data are normalized by the FPKM method. The developmental tissues represent vegetative and reproductive organs from greenhouse-grown plants. Shoots and roots were from *in vitro*-grown plants. Callus (10-11 week old) was derived from leaves and stems. Images provided by Kelly Zarka, Michigan State University and adapted from: McIntosh, Thomas P. *The Potato: Its history, Varieties, Culture and Diseases*. Great Britain: Oliver and Boyd, 1927. Vreugdenhil, Dick, et al. *Potato Biology and Biotechnology: Advances and Perspectives*. United Kingdom: Elsevier, 2007. Florida Center for Instructional Technology, 2009. *Setting up and running a school garden: A manual for teachers, parents and communities*. FAO Corporate Document Repository, 2005. *Solanum tuberosum L.*, Société Française d'Ethnopharmacologie, 2010.

Figura 9: Mapa de expressão do gene *StHCT* gerado na plataforma ePlant (http://bar.utoronto.ca/eplant_potato/). As cores indicam o tecido onde o gene foi expresso, sendo que é utilizado um gradiente que vai do amarelo ao vermelho (no qual o amarelo indica pouca expressão e o vermelho indica uma alta expressão). Nota-se que apenas no caule este gene é bastante expresso.

Já o padrão de expressão de *StHQT* foi significativamente distinto, apresentando maiores níveis de transcritos no tubérculo da batata, tanto na polpa quanto na pele, e em folhas jovens (**Figura 10**). Estes dados também corroboram o papel de HQT na síntese de ácidos clorogênicos, uma vez que estes compostos fenólicos são ativamente produzidos em tecidos jovens e nos tubérculos de batata (Payyavula et al., 2015). Para verificar se existe qualquer tipo de correlação (positiva ou negativa) entre o padrão de expressão dos 3 genes de interesse, os dados brutos de expressão utilizados pela plataforma ePlant foram baixados e usados no cálculo de correlação de Pearson feito de modo par a par (ou seja, *StCSE* x *StHCT*, *StCSE* x *StHQT* e *StHCT* x *StHQT*) (**Figura 11**). Vale ressaltar que dados baixados nesta plataforma são de 2 experimentos distintos com variedades distintas: o experimento 1 consiste em 13 tecidos/estágios de desenvolvimento e está nomeado como “Double Monoploid *Solanum tuberosum* Group Phureja clone DM1-3516R44”; o experimento 2 consiste em 18

tecidos/estágio de desenvolvimento e está nomeado como “Heterozygous diploid *Solanum tuberosum* Group Tuberosum RH89-039-16”. O cálculo de correlação de Pearson foi feito tanto para os experimentos separadamente (**Figura 11A e 11B**) quanto para todo o conjunto de dados (**Figura 11C**), que consiste em 31 pontos de expressão no total. Uma correlação positiva estatisticamente significativa foi observada para StCSE x StHCT tanto quando o cálculo foi realizado para os experimentos individualmente quanto para o cálculo completo (**Figura 11**). Este resultado sugere que *StCSE* e *StHCT* apresentam padrões de expressão semelhantes, corroborando a função de ambos os genes no metabolismo de lignina, uma vez que genes co-expressos tendem a participar dos mesmos processos biológicos (Delli-Ponti et al., 2021; Julca et al., 2023). Uma correlação negativa entre os padrões de expressão de StCSE e StHQT foi observada somente quando o cálculo de correlação de Pearson foi realizado para o conjunto completo de dados (**Figura 11C**). Tal correlação poderia ser explicada pelo fato do gene *StHQT* estar mais diretamente relacionado com a biossíntese de ácidos clorogênicos, processo que geralmente acontece em tecidos jovens, enquanto *StCSE* estaria mais relacionado com a biossíntese de lignina, cuja deposição ocorre com a maturação do tecido. Essa correlação negativa também corrobora a hipótese de que StCSE poderia estar envolvida com a mobilização de ácidos clorogênicos para a síntese de lignina, uma vez que o teor destes fenólicos solúveis tende a cair com a maturação (Volpi e Silva et al., 2019; Ferreira et al., 2022), justamente no momento do desenvolvimento em que há maiores níveis de expressão de genes biossintéticos de lignina (incluindo CSE).



Potato eFP Browser by A.Silva. Data from Transcriptome of the Reference Potato Genome *Solanum tuberosum* Group Phureja Clone DM1-3 516R44: Massa et al., 2011. PLoS One 6(10):e26801 and Genome sequence and analysis of the tuber crop potato: The Potato Genome Sequencing Consortium, 2011. Nature 475:10.1038/nature10158. RNA seq data are normalized by the FPKM method. The developmental tissues represent vegetative and reproductive organs from greenhouse-grown plants. Shoots and roots were from *in vitro*-grown plants. Callus (10-11 week old) was derived from leaves and stems. Images provided by Kelly Zarka, Michigan State University and adapted from: McIntosh, Thomas P. *The Potato: Its History, Varieties, Culture and Diseases*. Great Britain: Oliver and Boyd, 1927. Vreugdenhil, Dick, et al. *Potato Biology and Biotechnology: Advances and Perspectives*. United Kingdom: Elsevier, 2007. Florida Center for Instructional Technology, 2009. *Setting up and running a school garden: A manual for teachers, parents and communities*. FAO Corporate Document Repository, 2005. *Solanum tuberosum L.*, Société Française d'Ethnopharmacologie, 2010.

Figura 10: Mapa de expressão do gene *StHQT* gerado na plataforma ePlant (http://bar.utoronto.ca/eplant_potato/). As cores indicam o tecido onde o gene foi expresso, sendo que é utilizado um gradiente que vai do amarelo ao vermelho (no qual o amarelo indica pouca expressão e o vermelho indica uma alta expressão). Nota-se que apenas no tubérculo este gene é bastante expresso.

A)

Experimento 1			
	HCT x HQT	HCT x CSE	HQT x CSE
Coefficient (r)	-0,0284	0,598731	-0,362
N	14	14	14
T statistic	-0,09842	2,589505	-1,345243
Degree Freedom	12	12	12
p value	0,923223	0,023679	0,203418

B)

Experimento 2			
	HCT x HQT	HCT x CSE	HQT x CSE
Coefficient (r)	-0,271192	0,808093	-0,4042
N	18	18	18
T statistic	-1,127003	5,487387	-1,767628
Degree Freedom	16	16	16
p value	0,27636	4,97E-05	0,096186

C)

TOTAL			
	HCT x HQT	HCT x CSE	HQT x CSE
Coefficient (r)	-0,23785	0,822656	-0,35387
N	32	32	32
T statistic	-1,34125	7,925334	-2,07232
Degree Freedom	30	30	30
p value	0,189906	7,61E-09	0,046926

Figura 11: Cálculo da correlação de Pearson entre os genes utilizando os dados brutos obtidos da plataforma ePlant. Em destaque são os valores no qual o p -valor está menor que 0,05. A) Cálculo utilizando os dados do experimento 1, nota-se uma correlação positiva entre *StHCT* e *StCSE*. B) Cálculo utilizando os dados do experimento 2, nota-se uma correlação positiva entre *StHCT* e *StCSE*. C) Cálculo utilizando os dois conjuntos de dados, nota-se que com um maior número amostral, além de existir uma correlação positiva entre *StHCT* e *StCSE*, evidencia-se também uma correlação negativa entre *StHQT* e *StCSE*.

Após a análise *in silico*, foi realizada análise de expressão gênica dos genes *StHCT*, *StHQT* e *StCSE* por RT-qPCR utilizando diferentes tecidos de plantas adultas de batata da variedade Desiree. Neste experimento é possível observar que tanto *StHCT* quanto *StCSE* são mais expressos no caule que nos outros tecidos avaliados (**Figura 12**), o que é esperado para genes envolvidos no processo de lignificação e o que corrobora os dados observados na análise *in silico*. Diferentemente de *StHCT*, *StCSE* também apresentou níveis relativamente altos de transcritos na casca do tubérculo. Já o gene *StHQT* foi mais expresso na polpa e na casca do tubérculo (**Figura 12**), indicando seu envolvimento no metabolismo de ácidos clorogênicos. De forma geral, os dados de expressão por RT-qPCR corroboram largamente os dados de expressão *in silico* obtidos na plataforma ePlant. Ademais, a correlação de Pearson também foi calculada para os dados de RT-qPCR (**Figura 13**). Novamente, foi possível observar uma correlação positiva estatisticamente significativa entre os padrões de expressão de *StHCT* e *StCSE*. Urge ressaltar que o universo amostral dos dados de RT-qPCR ($n = 5$ tecidos/estágios de desenvolvimento) é muito menor do que aquele dos dados de expressão *in silico* ($n = 31$ tecidos/estágios de desenvolvimento), o que obviamente afeta a observação de alguma correlação entre a expressão dos genes-alvo. Conclui-se que, em ambos os experimentos de expressão gênica (tanto *in silico*, quanto por RT-qPCR), foi possível observar que *StHCT* e *StCSE* possuem um padrão de expressão semelhantes, sendo mais expressos no caule, muito provavelmente por se tratarem de genes relacionados ao metabolismo de lignina. Este padrão de expressão foi significativamente distinto daquele observado para o gene *StHQT*, que foi mais expresso em tubérculos, provavelmente por se tratar de um gene envolvido no metabolismo de ácidos clorogênicos.

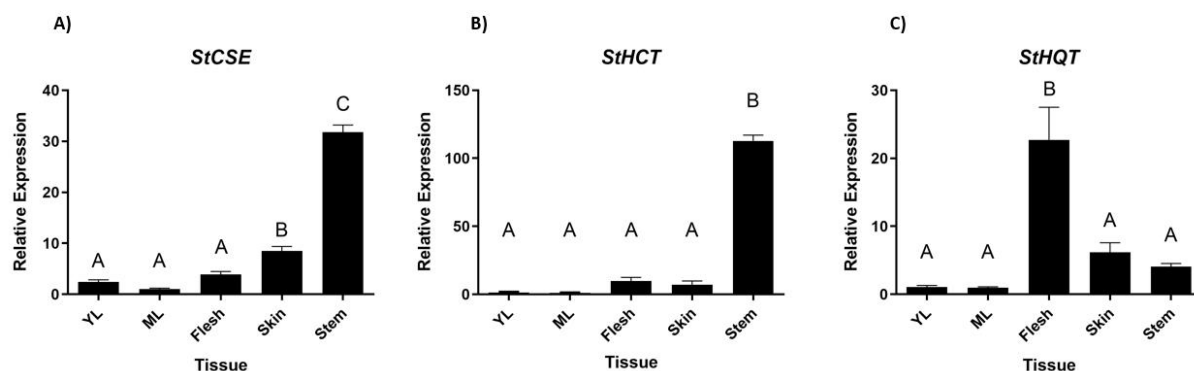


Figura 12: Análise de expressão gênica por RT-qPCR dos genes que codificam para: A) *StCSE*, B) *StHCT* e C) *StHQT*. Nota-se a semelhança do padrão de expressão entre os genes envolvidos no metabolismo de lignificação (*StCSE* e *StHCT*).

Correlação de Pearson RT-qPCR			
	HCT x HQT	HCT x CSE	HQT x CSE
Coefficient (r)	-0,1143337	0,98300793	-0,1319586
N	5	5	5
T statistic	-0,199339	9,27539219	-0,2305753
Degree Freedom	3	3	3
p value	0,85474357	0,00265212	0,83247403

Figura 13: Cálculo da correlação de Pearson entre os genes utilizando os dados brutos obtidos via RT-qPCR. Em destaque são os valores no qual o *p*-valor está menor que 0,05. Nota-se uma correlação positiva entre *StHCT* e *StCSE*.

5.3 *StCSE* foi produzida em *E. coli* e purificada com sucesso

A etapa de purificação da proteína foi realizada no equipamento AKTA utilizando cromatografia por afinidade por metais da enzima recombinante. O princípio desta purificação é utilizar a afinidade da cauda de histidina presente na proteína recombinante pelo níquel imobilizado na resina da coluna, permitindo a sua separação de outras proteínas da bactéria usada como sistema de expressão. O extrato proteico é passado no sistema AKTA, e as proteínas com afinidade pela resina ficam retidas na coluna, sendo gradualmente eluídas pela aplicação de um gradiente de concentração de um competidor pelo níquel imobilizado (no caso, imidazole). As diferentes frações eluídas do sistema AKTA são então avaliadas para a presença da proteína recombinante por meio de SDS-PAGE. Observou-se a proteína de interesse, que apresenta o peso molecular aproximado de 37 kDa, em quase todas as frações do AKTA, especialmente as frações 6 a 10 e as frações de 20 a 24 (**Figura 14**). Entretanto, como as frações contendo a proteína de interesse apresentaram muitas impurezas (muitas outras proteínas não-alvo), elas foram combinadas (com exceção das frações 12, 13, 14, 26 e 27, pois possuíam muitas impurezas) e submetidas a uma nova etapa de purificação utilizando uma coluna MonoQ, para separação de proteínas por

trocas iônicas (**Figura 15**). Com isso, as amostras de 4 a 20 foram selecionadas, quantificadas no espectrofotômetro e utilizadas para os ensaios enzimáticos.

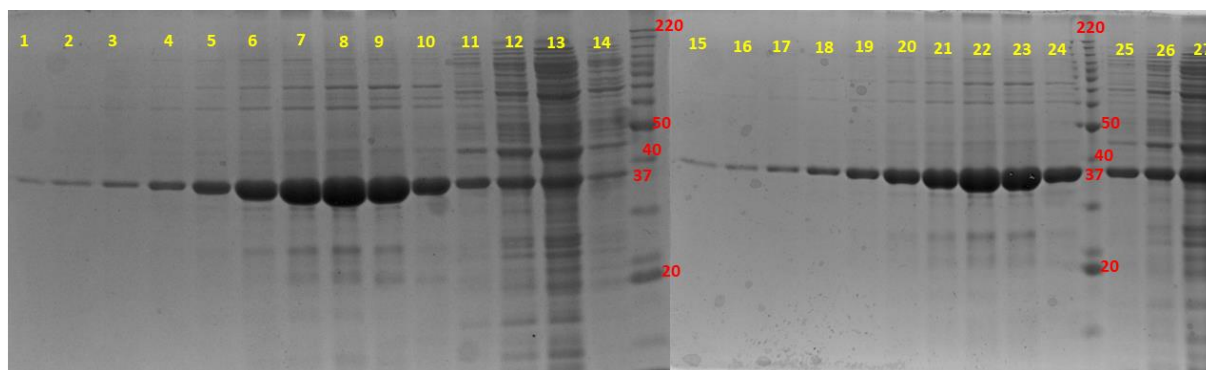


Figura 14: Gel de poliacrilamida 14% feito com as amostras da primeira etapa de purificação, utilizando cromatografia por afinidade por metais. A enzima possui aproximadamente 37 KDa. Em amarelo estão indicadas as frações proteicas e em vermelho o peso molecular em kDa.

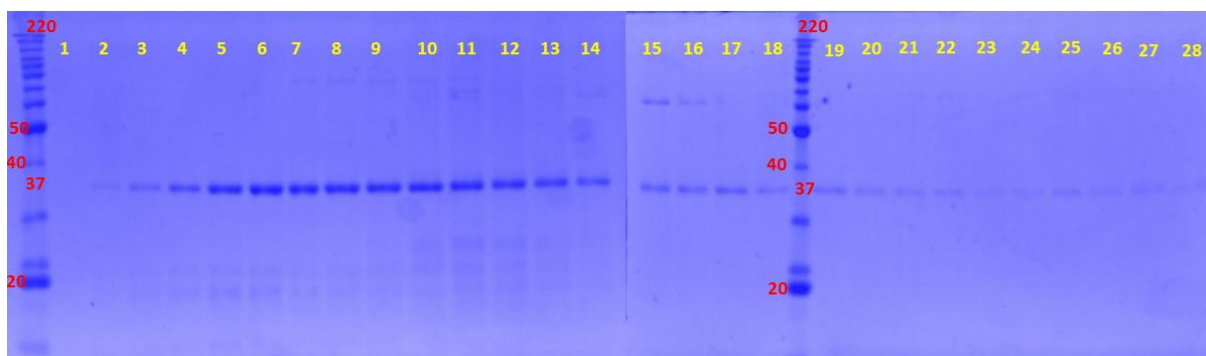


Figura 15: Gel de poliacrilamida 14% feito com as amostras da segunda etapa de purificação, utilizando troca iônica. Em amarelo estão indicadas as frações proteicas e em vermelho o peso molecular em kDa. Observa-se que as frações apresentam maior grau de pureza em relação à etapa anterior.

5.4 Ensaios enzimáticos *in vitro* demonstram o consumo de CGAs por CSE

Primeiramente, verificou-se a capacidade catalítica da enzima em metabolizar diferentes substratos, com foco principal em CGAs. Para tal, a enzima recombinante foi incubada por 2h a 30°C com os seguintes substratos: 1) caffeoyl shikimate (CS), usado como controle positivo, pois é o substrato endógeno da via biossintética de lignina previamente reportado para CSEs de outras espécies (Vanholme et al., 2013c; Ha et al., 2016a); 2) ácido 5-*O*-cafeoilquínico (CGA), visto que este é o principal CGA encontrado em plantas, incluindo espécies de solanáceas (Clifford et al., 2017a); e 3) ácido 3,5-dicafeoilquínico (3,5-DQA), um CGA duplamente acilado utilizado para verificar a capacidade da enzima em metabolizar diferentes CGAs, incluindo ésteres com mais de um motivo fenólico. Como controle negativo, as reações foram preparadas utilizando a proteína fervida (desnaturada, não-funcional). Ao término da reação, foi adicionado ácido acético nas amostras e posteriormente analisadas em UHPLC-MS para

verificação da formação do produto esperado, ácido caféico. Caso a enzima funcione com todos os substratos, o mesmo produto seria produzido, uma vez que todos os substratos apresentam o ácido caféico como motivo fenólico. Quando os padrões dos potenciais substratos e produto foram corridos na espectrometria de massas, o ácido caféico eluiu em um tempo de retenção por volta de 2.28 min, e o espectro de MS/MS demonstrou a presença dos principais fragmentos tendo m/z de 179 (íon parental) e 135 (**Figura 16**). Para os substratos, os fragmentos foram: CS, 335 (íon parental), 179 e 161; CGA, 353 (íon parental), 191 e 173; 3,5-DQA, 515 (íon parental), 353, 191 e 179 (**Figura 17**). Quando a enzima foi incubada com CS, observou-se no cromatograma a formação dos picos típicos referentes ao padrão de fragmentação de ácido caféico (**Figura 18A**), indicando que StCSE foi capaz de utilizar este substrato. Esses dados estando de acordo com a caracterização de CSE de outras espécies vegetais. Importante ressaltar que não há formação de ácido caféico nas reações controle, em que a enzima está desnaturada (**Figura 18B**), demonstrando que não há conversão espontânea de CS em ácido caféico. O mesmo resultado foi encontrado para CQA (**Figura 19**), evidenciando a formação de ácido caféico nas reações na qual a enzima não está desnaturada, e a não conversão do substrato nas reações controle. Para 3,5-DQA, o experimento demonstrou também que existe a metabolização do substrato (**Figura 20**), e a inexistência da conversão espontânea nas reações controle. Portanto, foi possível identificar a utilização de ácidos clorogênicos para formação de ácido caféico mediante a ação da enzima CSE.

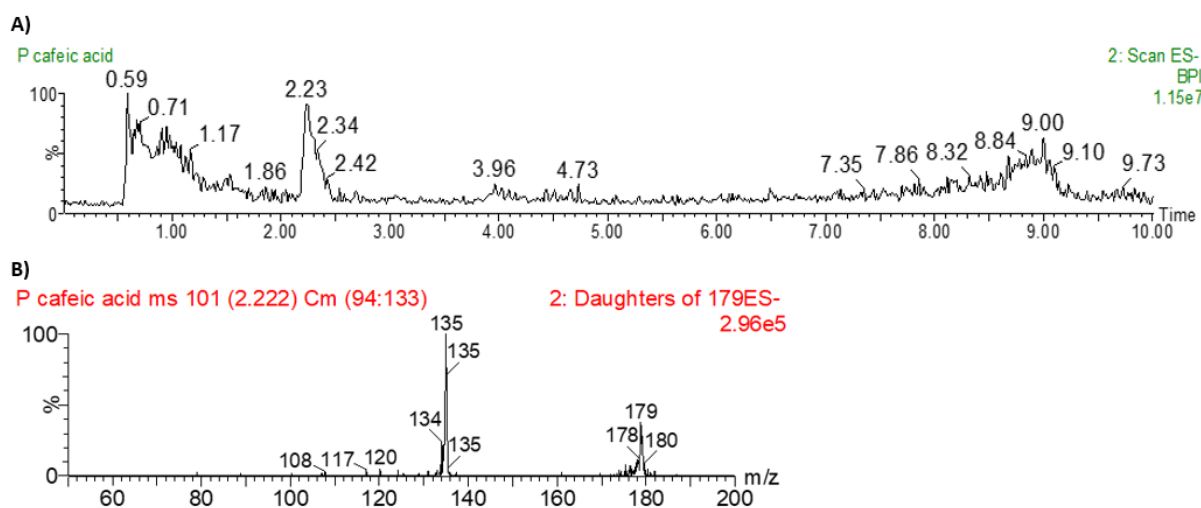


Figura 16: Análise realizada em UHPLC-MS para o padrão comercial de ácido caféico. A) Gráfico demonstrando que para a eluição do ácido caféico o tempo de retenção é por volta de 2.23 minutos. B) Espectro de MS/MS demonstrou a presença dos principais fragmentos tendo m/z de 179 (íon parental) e 135.

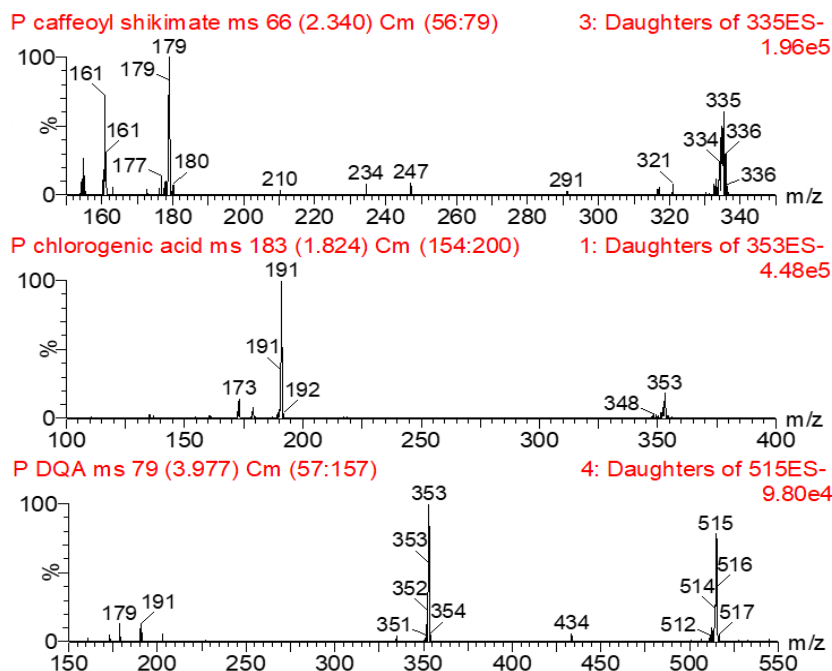


Figura 17: Espectro MS/MS dos padrões comerciais (ou sintetizados) dos substratos utilizados nas reações com StCSE. Os fragmentos encontrados foram: CS, 335 (ion parental), 179 e 161; CGA, 353 (ion parental), 191 e 173; 3,5-DQA, 515 (ion parental), 353, 191 e 179.

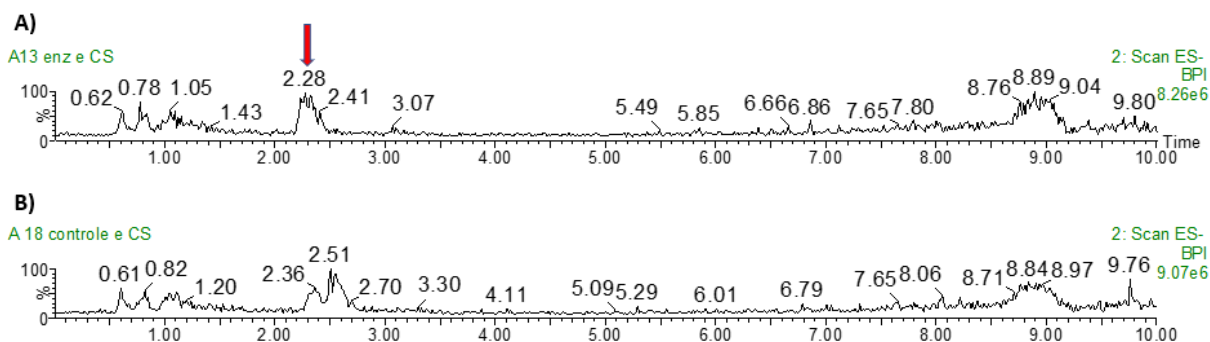


Figura 18: Análise em UHPLC-MS realizada para demonstra se a enzima metaboliza CS. A) Corrida referente à reação com a enzima funcional, em que é possível ver o aparecimento de ácido caféico no tempo de retenção 2.28 min. B) Reação realizada com a enzima desnaturada utilizada como controle negativo, em que o pico do produto é inexistente.

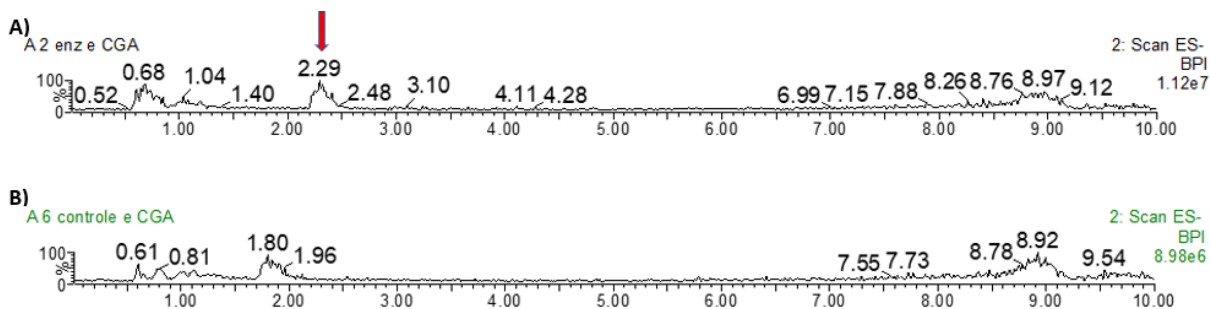


Figura 19: Análise em UHPLC-MS realizada para demonstra se a enzima metaboliza CQA. A) Corrida referente a reação com a enzima funcional, em que é possível ver o aparecimento do ácido cafeico no tempo de retenção

2.28min. B) Amostra com a enzima desnaturada utilizada como controle negativo, em que o pico do produto é inexistente.

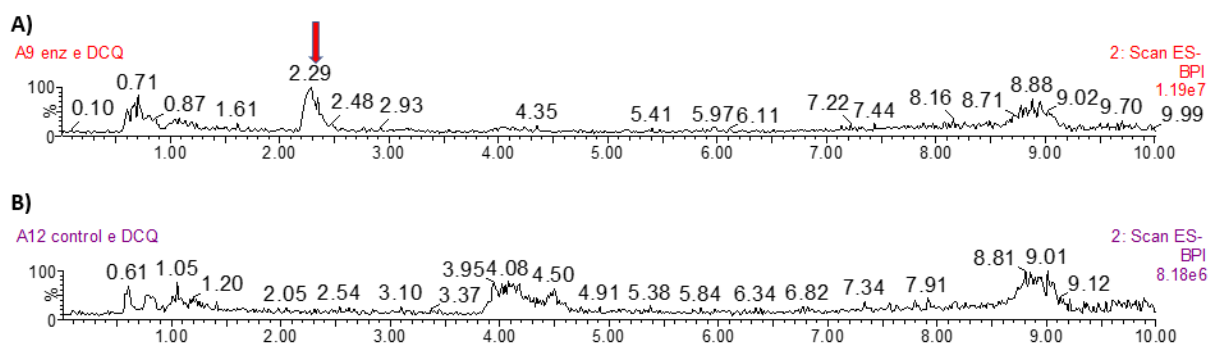


Figura 20: Análise em UHPLC-MS realizada para demonstra se a enzima metaboliza 3,5-DQA. A) Amostra com a enzima funcional, possível ver o tempo de retenção do ácido cafeico em 2.29 min, indicando que houve o consumo. B) Amostra com a enzima desnaturada, o pico do produto é inexistente.

Uma vez confirmada a capacidade de StCSE em metabolizar os três substratos testados, a próxima etapa foi verificar com que eficiência cada um dos substratos é usado pela enzima. A premissa é que, para considerarmos que CGAs podem ser potenciais substratos endógenos de CSE e que, conseqüentemente, CSE estaria envolvida no metabolismo de CGAs *in planta*, espera-se que a enzima consiga metabolizar estes substratos com eficiência semelhante (ou melhor) à observada para CS. Para tal, experimentos de cinética enzimática foram conduzidos, acompanhando a formação do produto em HPLC. Primeiramente determinamos os valores de K_m , este parâmetro expressa a afinidade da enzima por um determinado substrato, sendo que quanto menor o seu valor, maior a afinidade. Neste caso, encontramos os valores de $K_m = 6,12 \times 10^{-9}$ M para CQA, para 3,5-DQA um $K_m = 2,86 \times 10^{-8}$ M e para CS um $K_m = 7,93 \times 10^{-8}$ M (**Figura 21**). Avaliando esses valores, temos que CQA foi o substrato pelo qual a enzima teve a maior afinidade e 3,5-DQA foi o substrato pela qual a enzima teve a menor afinidade. Após o cálculo de K_m , obtivemos a velocidade máxima (V_{max}), que seria a velocidade na qual a enzima está saturada, ou seja, todas as enzimas presentes na reação estão com seus sítios ativos ocupados. Os valores encontrados foram $V_{max} = 1,60 \times 10^{-12}$ $M \times s^{-1}$ para CQA, $V_{max} = 1,47 \times 10^{-11}$ $M \times s^{-1}$ para 3,5-DQA e $V_{max} = 3,28 \times 10^{-11}$ $M \times s^{-1}$ para CS (**Figura 21**). Com os dados de V_{max} , foi possível calcular o K_{cat} , parâmetro que expressa a quantidade de produto produzida quando se atinge a velocidade máxima da reação. Nesta etapa foram encontrados $K_{cat} = 99,13$ $1 \times s^{-1}$ para CQA, $K_{cat} = 130,56$ $1 \times s^{-1}$ para 3,5-DQA e $K_{cat} = 167,23$ $1 \times s^{-1}$ para CS (**Figura 21**). É possível observar que embora a enzima apresente uma maior afinidade por CQA, ao atingir sua velocidade máxima, CS e 3,5-DQA tiveram mais produtos produzidos, o que poderia ser explicado por maiores valores de V_{max} para tais substratos. Por fim, calculou-se a eficiência de cada reação, utilizando K_{cat}/K_m . Esse parâmetro é a constante de especificidade, que no nosso

caso demonstra o número de vezes que os substratos são catalisados por uma mesma enzima. Assim, quanto maior este valor, maior a eficiência para um determinado substrato. Os valores encontrados foram $K_{cat}/K_m = 1,62 \times 10^{10} \text{ M}^{-1}\text{s}^{-1}$ para CQA, $K_{cat}/K_m = 4,56 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$ para 3,5-DQA, e $K_{cat}/K_m = 2,11 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$ para CS (**Figura 21**). Como conclusão, observou-se que, além da enzima ser capaz de metabolizar os 3 substratos, ela tem uma afinidade e uma eficiência maior para CQA em relação a CS (o substrato endógeno pertencente à via biossintética de lignina), e valores muito próximos para 3,5-DQA. Todos os dados foram obtidos utilizando GraphPad Prism 8, programa que além de gerar o gráfico de Michaelis-Menten, também calcula o K_m , K_{cat} e a V_{max} . Vale ressaltar, que os valores finais são as médias dos gráficos gerados de cada triplicata experimental. Ao observar os gráficos (**Figura 22**), é possível notar que os 3 substratos se comportaram de acordo com o esperado na teoria, onde existe um platô indicando que todas as enzimas em algum momento estão com seus sítios ativo ocupados pelo substrato.

Valores obtidos na cinética				
Substrato	K_m	V_{max}	K_{cat}	K_{cat}/K_m
CQA	6,12E-09	1,6E-12	99,13	1,62E+10
3,5-DQA	2,86E-08	1,47E-11	130,5667	4,56E+09
CS	7,93E-08	3,28E-11	167,2333	2,11E+09

Figura 21: Valores de K_m , V_{max} e K_{cat} encontrados para os substratos CGA, 3,5-DQA e CS. Nota-se uma eficiência e uma afinidade maior da enzima em relação ao substrato CQA quando comparado com o substrato endógeno CS.

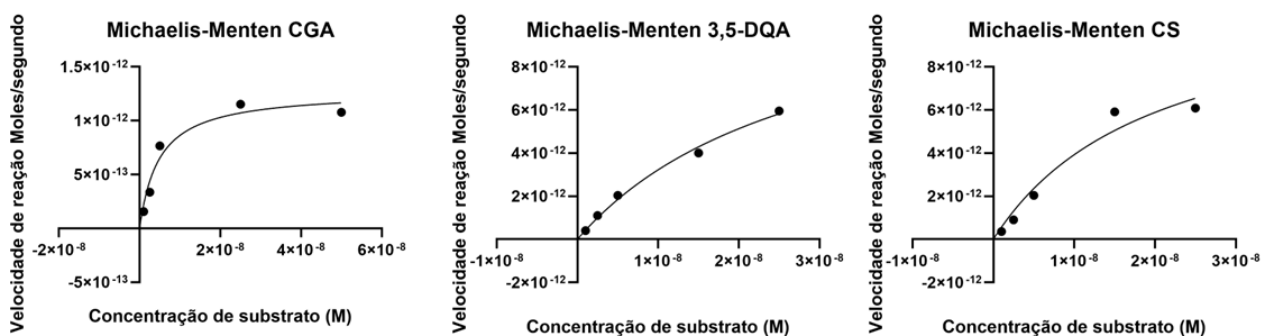


Figura 22: Gráficos de Michaelis-Menten para cada substrato. Ressalta-se que estes são gráficos representativos de uma réplica das três utilizadas para calcular os valores finais, que foram as médias de K_m , V_{max} e K_{cat} .

6. DISCUSSÃO

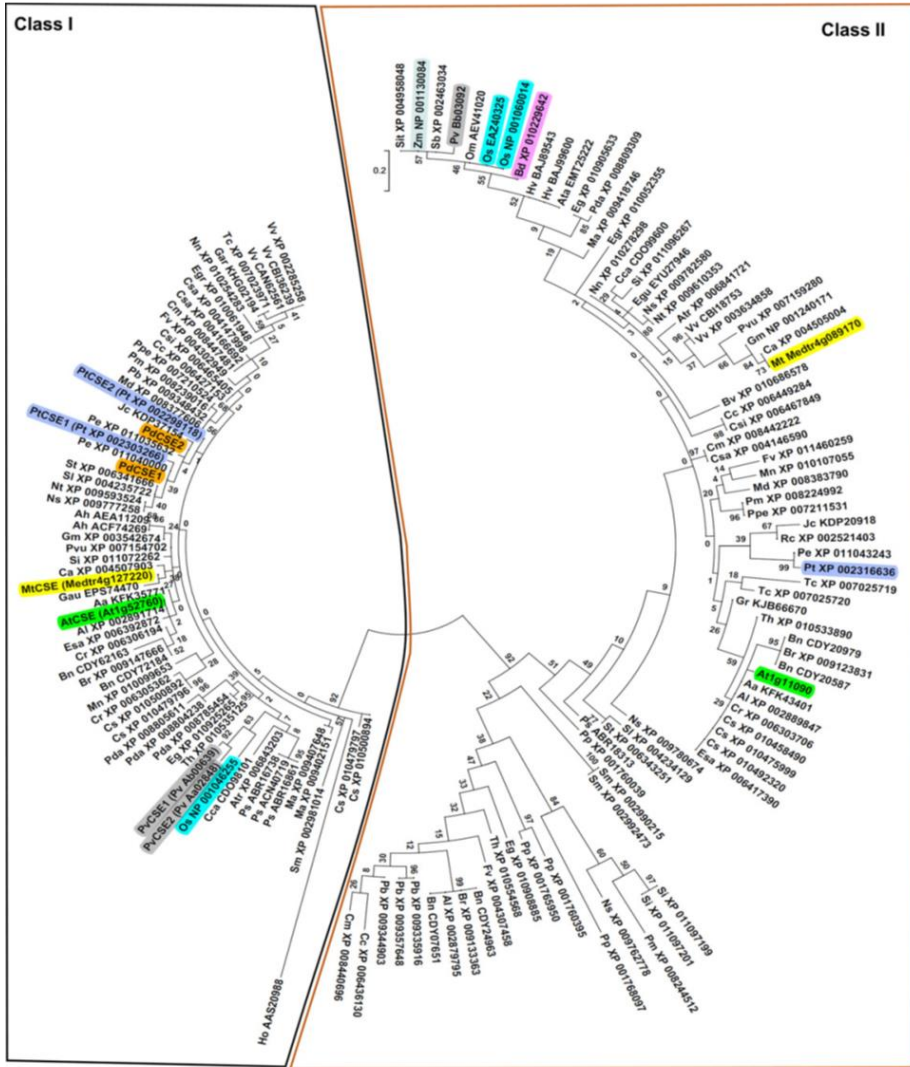
6.1 Evolução dos genes *CSE* em um possível contexto de neofuncionalização

Após a descoberta do gene *CSE* em *Arabidopsis* (Vanholme et al., 2013c), vários pesquisadores começaram buscas pelo gene em diversas espécies vegetais com o intuito de verificar se a função desta enzima no processo de lignificação é conservada. Genes considerados *bona fide* *CSE* foram encontrados largamente no genoma de espécies eudicotiledôneas, mas aparentemente a maioria das monocotiledôneas carece de um gene *CSE* autêntico em seus genomas. Com poucas exceções, *CSE* normalmente ocorre como cópia única no genoma das plantas, como observado em *Arabidopsis* e *Medicago*, enquanto poucas espécies (provavelmente aquelas com genoma poliploide) apresentam mais de uma cópia, como o álamo que possui 2 genes *CSE* largamente redundantes (de Vries et al., 2021). Em batata, também foi encontrado apenas uma cópia de *CSE*, assim como é observado na maioria dos genomas diploides.

Ao analisar o histórico de trabalhos com foco na caracterização de *CSE* em diferentes espécies vegetais, constata-se que a enzima foi primeiramente identificada como uma lisofosfolipase (lysoPL2) em *Arabidopsis thaliana*, em um trabalho sobre tolerância ao estresse oxidativo causado por cádmio (Gao et al., 2010). Lisofosfolípidos são intermediários do metabolismo dos fosfolípidos produzidos em condições de estresse, e as lisofosfolipases hidrolisam as ligações éster destes lisofosfolípidos para produzir ácidos graxos e derivados de glicerolfosfato, atuando no processo de desintoxicação da planta sob estresse (Wang and Dennis, 1999; Gao et al., 2010). LysoPL2 foi identificada em ensaios de duplo-híbrido em levedura e co-imunoprecipitação como uma proteína que interage fisicamente com acyl-CoA-binding protein 2 (ACBP2), uma proteína conhecida por seu papel no metabolismo de fosfolípidos (Gao et al., 2009). A superexpressão de lysoPL2 em *Arabidopsis* resultou em plantas mais tolerantes ao H₂O₂ e ao estresse induzido por cádmio quando comparadas com plantas tipo selvagem, sugerindo envolvimento de lysoPL2 no reparo fosfolipídico após peroxidação lipídica decorrente do estresse induzido por metal pesado (Gao et al., 2010). Um estudo posterior também demonstrou a natureza bifuncional de lysoPL2 no metabolismo de fosfolípidos, apresentando tanto atividade de acil-transferase quanto de acil-hidrolase para substratos monoacilglicéris (Vijayaraj et al., 2012). Após a descoberta e caracterização de *CSE* com função relacionada a lignificação (Vanholme et al., 2013), estudos estritamente focados em lignina praticamente ignoraram a função previamente reportada para a enzima no metabolismo de fosfolípidos e caracterizaram filogeneticamente a família gênica *CSE* em membros *CSE bona fide* (classe I) e *CSE-like* (classe II) (Fig. 23A) (Ha et al., 2016b). No entanto, um estudo filogenético e funcional em *Arabidopsis* demonstrou que *CSE/lysoPL2* é

um membro da família de monoacilglicerol lipases (MAGL) (Fig. 23B) (Kim et al., 2016), que catalisam a última etapa da degradação do triacilglicerol, que é a hidrólise do monoacilglicerol (MAG) em ácido graxo e glicerol. Interessantemente, neste mesmo trabalho, os pesquisadores não conseguiram observar atividade de MAGL/lysoPL para a CSE recombinante (no trabalho chamada MAGL3), o que nos leva a propor que CSE é fruto de um processo de neofuncionalização dentro da família MAGL, em que a enzima perdeu sua capacidade de metabolizar fosfolipídios e adquiriu uma nova função na biossíntese de lignina. Alternativamente e em vista da observação da atividade de CSE com substratos fosfolipídicos por 2 trabalhos independentes (Gao et al., 2010; Vijayaraj et al., 2012), CSE pode ser um caso de bifuncionalização, podendo participar de ambas as vias metabólicas.

A)



B)

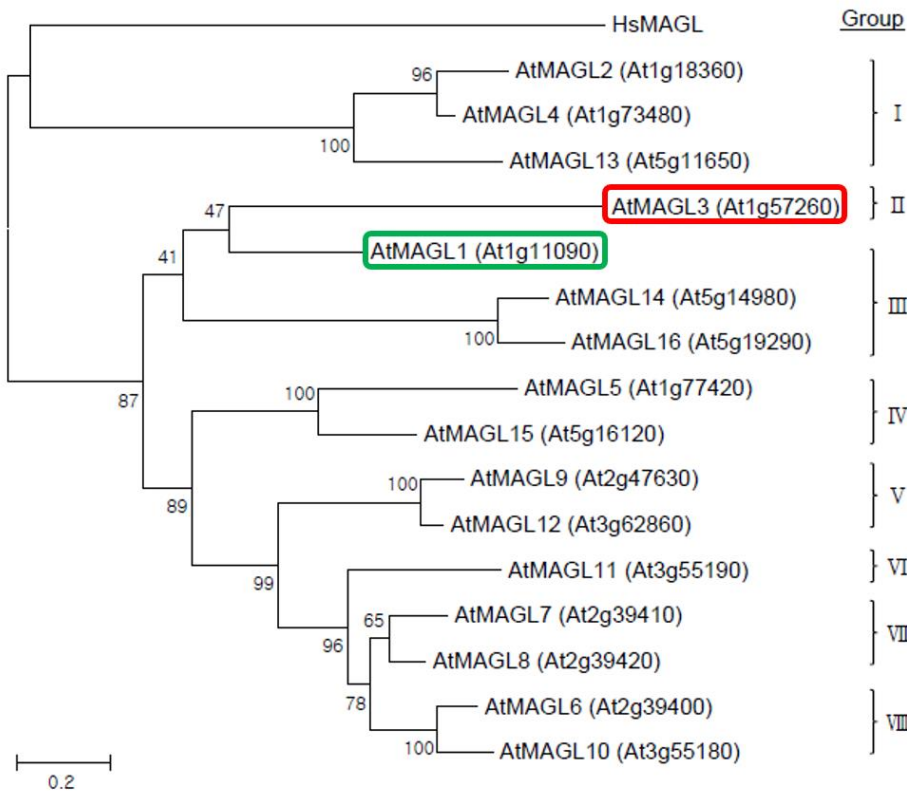


Figura 23: Classificação filogenética de CSE em diferentes contextos científicos. A) Estudos abordando lignificação classificaram CSE como uma nova enzima central à biossíntese de lignina, observando basicamente dois clados: CSE *bona fide* (ou classe I; aquelas que estariam de fato envolvidas no metabolismo de lignina) e CSE-like (ou classe II; proteínas filogeneticamente próximas mas sem função demonstrada para o processo de lignificação). Fonte: Ha et al. (2016). B) Classificação filogenética de CSE quando considerada membro da família das monoacilglicerol lipases (MAGLs). Kim et al (2016) demonstrou que CSE (aqui chamada de AtMAGL3 e marcada em vermelho; observar que, nesta figura, há um erro no número de acesso do gene, que deveria ser At1g52760) é membro do grupo II de MAGLs, juntamente com AtMAGL1 (At1g11090; marcada em verde) que aparece como uma CSE-like de classe II na filogenia de Ha et al. (2016).

A bifuncionalização de CSE foi recentemente sugerida (Dixon and Barros, 2019), levando em consideração a sua localização subcelular na membrana plasmática e a possível interação com a proteína ACBP2. Neste cenário (Fig. 24) o metabolismo de lignina seria deslocado para uma região mais próxima da membrana, facilitando o transporte dos monolignóis para o apoplasto, onde a polimerização da lignina ocorre. Interessante observar que existem outros casos de bifuncionalização em que uma enzima relacionada à lignificação também está relacionada ao metabolismo de resposta ao estresse oxidativo e desintoxicação de H₂O₂ da célula. A seguir, descreve-se o envolvimento de outras enzimas biossintéticas de lignina no metabolismo de H₂O₂, como descrito por Dixon e Barros (2019), sendo cada caso identificado por um número. H₂O₂ é gerado em múltiplos compartimentos intracelulares (incluindo cloroplasto, mitocôndria e peroxissomo), sendo a principal fonte a NADPH oxidase (NOX) associada à membrana plasmática com seu sítio catalítico voltado para o apoplasto. (1) NOX é ativada pela ligação da enzima biossintética de lignina cinamoil CoA redutase (CCR) à RAC1 GTPase associada à membrana plasmática, e CCR juntamente com outra enzima biossintética de lignina chamada cinamil álcool desidrogenase (CAD) formam um complexo protéico que pode induzir o transporte de monolignóis para o apoplasto via difusão passiva 'dirigida' (PD) ou por transportadores de membrana. O H₂O₂ apoplástico derivado do superóxido via superóxido dismutase (SOD) é usado para oxidação de monolignóis por peroxidases de classe III, permitindo a polimerização da lignina. (2) H₂O₂ e outras ROS causam dano à membrana plasmática pela oxidação de lipídios, e esse dano pode ser reparado pela lysoPL2, que também é tem função na biossíntese de monolignóis atuando como CSE. CSE/lysoPL2 está ligada à membrana plasmática pela interação com ACBP2, e essa localização pode ajudar ainda mais a mover a biossíntese de monolignóis para o lado interno da membrana plasmática para difusão dos monômeros de lignina direcionada ao apoplasto. (3) O H₂O₂ gerado no citosol ou transportado do apoplasto via aquaporinas (AqP) pode ser desintoxicado pela ascorbato peroxidase citosólica (APX), com conversão do ascorbato em monodesidroascorbato. Essa enzima também é bifuncional e funciona para converter o ácido *p*-cumárico em ácido cafeico

(C3H/APX) (Barros et al., 2019). Um gradiente de concentração através da membrana plasmática parece ser necessário para permitir que o H₂O₂ gerado no apoplasto atue como um sinal para coordenar a biossíntese de monolignóis no citosol e sua polimerização no próprio apoplasto.

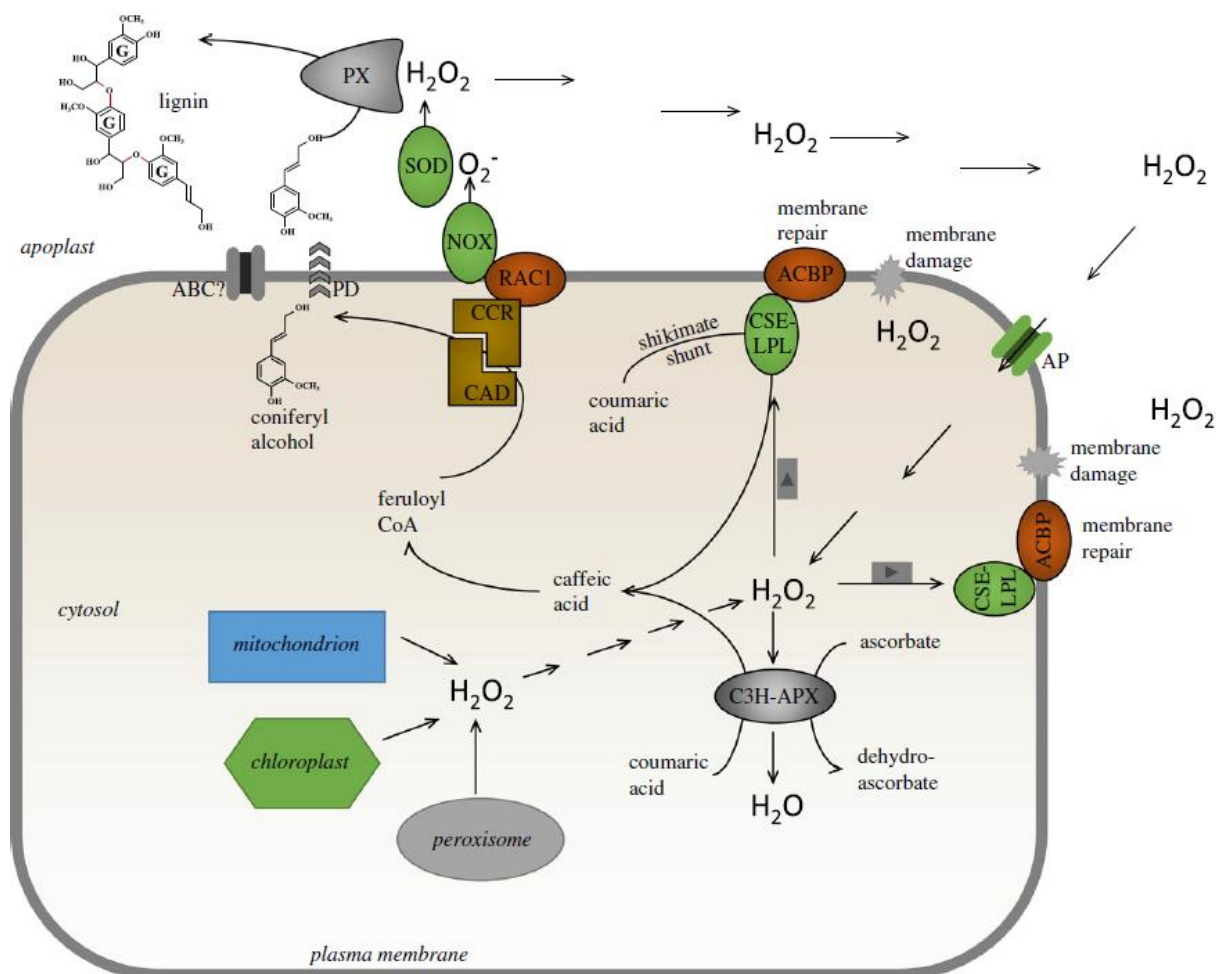


Figura 24: Modelo que liga a biossíntese de monolignóis (aqui exemplificado pelo álcool coníferílico apenas para simplificação) à geração e desintoxicação de peróxido de hidrogênio. Este modelo inclui os 3 casos de bifuncionalização em que enzimas da biossíntese de lignina também atuam no metabolismo de resposta ao estresse e desintoxicação de H₂O₂. Fonte: Dixon & Barros (2019).

6.2 Lignina e ácidos clorogênicos: rotas biossintéticas interligadas e relações funcionais entre CSE, HCT e HQT

As vias biossintéticas que resultam na produção de CGAs e lignina compartilham intermediários e enzimas em comum, além de envolverem rotas paralelas equivalentes denominadas *shikimate shunt* (para lignina) e *quinate shunt* (para CGAs). Nesse contexto, 3 enzimas aparentemente desempenham funções-chave que vão determinar para qual rota, e em que momento do desenvolvimento, os precursores da via geral dos fenilpropanóides serão direcionados: HCT, HQT e CSE. Enquanto HCT canaliza os ésteres de CoA (e.g., *p*-cumaroil-

CoA e cafeoil-CoA) para o *shikimate shunt* e, conseqüentemente, para a biossíntese de monômeros de lignina, HQT canaliza os mesmos precursores para o *quinato shunt* para a produção e acúmulo de CGAs. CSE seria considerada a porta de saída dos ésteres de chiquimato para a biossíntese de lignina, mas neste trabalho também propomos que esta enzima poderia participar da remobilização de CGAs (ésteres de quinato) para gerar precursores para a produção de lignina. Desse modo, estas três enzimas estariam altamente correlacionadas por utilizarem os mesmos substratos para a biossíntese de compostos distintos (no caso de HCT x HQT) e a mesma “porta de saída” para a biossíntese de monômeros de lignina (sendo esta a enzima CSE). Neste trabalho utilizando batata como um organismo modelo, exploramos os perfis de expressão desses três genes para identificar possíveis correlações, utilizando diferentes tecidos vegetais em diferentes fases do desenvolvimento e utilizando uma abordagem *in silico* e outra *in vivo*. Nossas premissas foram: (i) genes envolvidos no metabolismo de lignina apresentariam perfis semelhantes, resultando em uma correlação positiva significativa; (ii) em vista do perfil inverso de acúmulo de CGAs e lignina em tecidos vegetais (Aerts and Baumann, 1994; Diaz et al., 1997; Shen et al., 2009; Ferreira et al., 2022), genes envolvidos na biossíntese e mobilização de CGAs apresentariam correlação negativa, refletindo padrões de expressão opostos (ao menos em alguns tecidos). Ao realizar o estudo com os dados de expressão disponíveis na plataforma ePlant (http://bar.utoronto.ca/eplant_potato/), oriundos de 2 experimentos de transcriptômica em larga escala distintos e totalizando 32 pontos amostrais, observou-se uma correlação positiva nos dados de expressão de *StHCT* e *StCSE*, indicando que estes genes apresentam padrão de expressão similar. Uma correlação positiva significativa também foi observada para estes genes quando calculada para os dados de RT-qPCR, que incluíram somente 5 pontos amostrais. Tais observações corroboram nossa primeira premissa, além de corroborar fortemente o papel de ambos os genes no metabolismo de lignina, uma vez que a expressão de genes relacionados ao processo de lignificação é altamente coordenada (Zhao et al., 2010; Ruprecht et al., 2011; Vanholme et al., 2013c; Ferreira et al., 2022). Ademais, uma correlação negativa foi observada nos dados de expressão de *StCSE* e *StHQT*, embora ela tenha sido significativa somente quando calculada usando todos os 32 pontos amostrais dos dados *in silico*. Embora essa observação não corrobore totalmente nossa segunda premissa, uma vez que não houve correlação significativa entre os padrões de expressão de *StHCT* e *StHQT*, a correlação negativa entre os padrões de expressão de *StCSE* e *StHQT*, juntamente com o fato de *StCSE* ser capaz de utilizar CGAs como substrato *in vitro*, nos permite especular que *StHQT* participa da biossíntese de CGAs em alguns tecidos jovens ou de reserva, enquanto a expressão de *StCSE* é ativada em fases mais finais do desenvolvimento vegetal, de modo a induzir a mobilização de esqueletos de carbono na forma de hidroxicinamatos para a biossíntese de

lignina ou outros biopolímeros (ver próxima seção). Alinhados com os nossos resultados, outros estudos também demonstraram uma correlação negativa entre genes relacionados ao metabolismo de CGAs e lignina, especialmente durante o desenvolvimento do fruto de café (Koshiro et al., 2007; Lepelley et al., 2007; Joët et al., 2009). Por outro lado, um estudo recente que avaliou o metabolismo de CGAs em casca e polpa de tubérculos de diferentes variedades de batata reportou dados distintos (Valiñas et al., 2015). Neste estudo, os pesquisadores observaram uma correlação positiva entre os dados de expressão gênica de *StHQT* e *StCSE*, enquanto ambos os genes apresentaram correlação negativa com *StHCT*. Embora a correlação negativa entre *HCT* e *HQT* fosse esperada, é interessante observar uma correlação negativa entre dois genes da biossíntese de lignina (*HCT* e *CSE*). Provavelmente esta discrepância com nossos dados possa ser explicada pela diferença no desenho experimental entre os trabalhos: enquanto nosso trabalho avaliou a expressão gênica em diferentes tecidos da planta e em diferentes estágios do desenvolvimento, o trabalho de Valiñas et al (2015) utilizou somente dados de expressão de casca e polpa de tubérculos de diferentes variedades de batata. Apesar desta discrepância, urge notar que os dados de expressão obtidos neste trabalho corroboram o papel de *HCT* e *CSE* no processo de lignificação, sendo ambos preferencialmente expressos em caule (um tecido reconhecido como um sítio de lignificação), e também de *HQT*, que é preferencialmente expresso em tecidos de tubérculos, que produzem grandes quantidades de CGAs (Payyavula et al., 2015; Valiñas et al., 2015).

6.3 A eficiência catalítica de *CSE in vitro* suportam seu papel no metabolismo de CGAs

Trabalhos anteriores já haviam demonstrado que *CSEs* recombinantes apresentavam grande promiscuidade *in vitro* em termos de substratos. No trabalho original de caracterização de *CSE* em *Arabidopsis*, Vanholme et al (2013) utilizaram um experimento em que o extrato fenólico de mutantes perda-de-função *cse2* foram incubados com a enzima recombinante, seguido de análise de LC-MS para identificar quais compostos eram consumidos pela enzima. Neste experimento, além de cafeoil chiquimato, os compostos cafeoil glicose e feruloil glicose também tiveram o seu conteúdo diminuído pela atividade da enzima, sugerindo a capacidade desta em metabolizar outros tipos de ésteres de hidroxicinamatos além de cafeoil chiquimato. Em um trabalho posterior do mesmo grupo, enzimas recombinantes para os dois parálogos de *CSE* em álamo (*Populus tremula* × *alba*) foram capazes de hidrolisar todos os quatro hidroxicinamoil-chiquimatos (i.e., *p*-cumaroil, cafeoil, feruloil e sinapoil-chiquimato), embora com diferentes afinidades (de Vries et al., 2021). Neste presente trabalho, expandimos a diversidade metabólica dos substratos de enzimas *CSE* ao testarmos a atividade enzimática *in vitro* com diferentes CGAs. *CSE* recombinante de batata foi capaz de hidrolisar eficientemente

os CGAs cafeoil quinato (5QA) e 3,5-dicafeoil quinato (3,5-DQA). Considerando todos esses resultados, podemos concluir sobre a diversidade estrutural dos substratos de CSE: (i) em termos de motivos fenólicos, essas enzimas aceitam substratos com diferentes níveis de hidroxilação/metilação/metoxilação em diferentes posições do anel aromático; (ii) essas enzimas aceitam ácidos fenólicos conjugados com diferentes compostos, podendo ser chiquimato, quinato e glicose; (iii) essas enzimas aceitam ésteres com mais de um motivo fenólico.

O fato de StCSE ter sido capaz de utilizar 3,5-DQA traz algumas considerações importantes. Até recentemente, pouco se sabia sobre a síntese desses di-ésteres, ao contrário do sólido conhecimento gerado acerca da produção dos monoésteres pela atividade de HQT/HCT. Em um trabalho pioneiro em tomateiro, Moglia et al (2014) observou a produção de diCQAs quando CGA (ácido 5-*O*-cafeoilquínico) foi incubado com extratos proteicos de frutos de tomate. O processo de purificação levou à identificaram da própria HQT como a enzima responsável por essa atividade, que foi caracterizada como “clorogenato:clorogenato transferase”. Ensaios de atividade *in vitro* com a enzima recombinante mostrou que essa atividade clorogenato:clorogenato transferase era ótima em pH baixo, o que é tipicamente encontrado no interior de vacúolos. Posteriormente, observou-se que a enzima está localizada tanto no citoplasma (onde aparentemente ela vai funcionar na produção dos mono-ésteres) quanto no vacúolo (onde ela está envolvida na produção dos di-ésteres). Essas observações sobre o papel de HQT na síntese de ácidos dicafeoil-quínicos e sua localização nos vacúolos têm implicações na atividade de CSE com 3,5-DQA demonstrada pelo atual trabalho. O artigo original que caracterizou CSE como lysoPL2 reportou a localização subcelular da enzima como sendo majoritariamente presente na membrana plasmática, com algum sinal aparecendo também no retículo endoplasmático quando as células foram plasmolisadas (Gao et al., 2010). Uma vez que (i) a localização no vacúolo é difícil de ser monitorada (geralmente a fluorescência é baixa e deve ser minuciosamente observada) e, (ii) o fato dos autores terem usado fusão com GFP (que é menos estável em pH baixos do que outras proteínas repórteres como RFP) (Moglia et al 2014), não é possível concluir se CSE e 3,5-DQA encontram-se necessariamente em compartimentos subcelulares distintos e, portanto, se a atividade enzimática observado no presente trabalho tem relevância fisiológica. No entanto, o fato de mono-ésteres também serem encontrados no vacúolo (Moglia et al 2014) e do processo de maturação de diferentes tecidos ser acompanhado por uma conspícua diminuição no conteúdo de CGAs em diferentes espécies vegetais (como descrito anteriormente) sugerem um processo de remobilização de CGAs do vacúolo para o citoplasma em momentos-chave do desenvolvimento, permitindo sua metabolização por CSE e consequente canalização dos hidroxicinamatos para a produção de

lignina ou outros compostos fenólicos. A observação de que a eficiência de StCSE utilizando CGAs é tão alta ou até maior do que utilizando cafeoil chiquimato como substrato reforça a ideia de que CGAs são seus substratos *in vivo* e que CSE pode ser considerada uma enzima atuando no metabolismo de CGAs *in planta*.

Neste trabalho, sugerimos que CSE participe da remobilização de CGAs para a biossíntese de lignina. No entanto, é importante destacar que a hidrólise de CGAs e consequente liberação de hidroxicinamatos podem estar envolvidas na produção de outros compostos fenólicos. Embora CGAs sejam considerados compostos de defesa, não podemos destacar que estresses bióticos e abióticos possam desencadear a sua remobilização para a síntese de fenilpropanóides que sejam mais efetivos em resposta a estresses específicos (Barber et al., 2000; König et al., 2014). Outra possibilidade, especialmente no caso de tubérculos de batata, é que a hidrólise de CGAs seja ativada para a produção de suberina, um heteropolímero constituído por uma fração polialifática associada a compostos fenólicos (basicamente hidroxicinamatos) (Figueiredo et al., 2016). Como o ácido ferúlico é o principal composto fenólico presente na suberina, é possível que a hidrólise de CGAs em alguns tecidos sirva para gerar este hidroxicinamato, principalmente via 3-O-metilação do ácido cafeico que foi liberado pela quebra de CGAs. De fato, esta ideia já foi sugerida pelo trabalho de Valiñas et al (2015), que observou uma forte correlação entre a biossíntese de CGAs e a formação de suberina no tubérculo de diferentes variedades de batata. Aqui, propomos que o papel de CSE na remobilização de CGAs para diferentes ramos do metabolismo fenólico vai depender (i) do tecido/órgão da planta; (ii) da espécie vegetal; (iii) do estágio de desenvolvimento; e (iv) de possíveis interações da planta com agentes causadores de estresses bióticos e abióticos.

7. CONCLUSÕES E PERSPECTIVAS

Neste trabalho, iniciamos experimentos para avaliar um potencial papel da enzima biossintética de lignina CSE no metabolismo de CGAs. Para tal, utilizamos uma abordagem integrativa incluindo análises filogenéticas, transcriptômicas e de atividade enzimática *in vitro*. As principais conclusões desta dissertação são:

- 1) O genoma de batata apresenta somente uma cópia de CSE genuína, similar ao observado na maioria das espécies eudicotiledôneas diploides;
- 2) A expressão de *StCSE* se correlaciona positivamente com a expressão do gene biossintético de lignina *StHCT* e negativamente com a expressão do gene biossintético de CGAs *StHQT* em diferentes tecidos de batata;
- 3) *StCSE* recombinante é capaz de utilizar não somente o intermediário da via de lignina cafeoil chiquimato mas também diferentes CGAs como substrato *in vitro*, incluindo compostos com mais de um motivo fenólico;
- 4) *StCSE* é tão (ou mais) eficiente em utilizar CGAs quanto cafeoil chiquimato, o que reforça uma potencial função dessa enzima no metabolismo de CGAs *in vivo*;
- 5) A atividade de *StCSE* utilizando CGAs como substrato reforça o cenário evolutivo de surgimento dessas enzimas em um processo de neofuncionalização seguido de diversificação funcional, sendo que até o momento essas enzimas foram vinculadas a três processos biológicos: defesa contra estresse oxidativo, metabolismo de lignina e metabolismo de ácidos clorogênicos.

8. BIBLIOGRAFIA

- Aerts RJ, Baumann TW** (1994) Distribution and utilization of chlorogenic acid in *Coffea* seedlings. *J Exp Bot* **45**: 497–503
- Barber MS, McConnell VS, DeCaux BS** (2000) Antimicrobial intermediates of the general phenylpropanoid and lignin specific pathways.
- Barros J, Escamilla-Trevino L, Song L, Rao X, Serrani-Yarce JC, Palacios MD, Engle N, Choudhury FK, Tschaplinski TJ, Venables BJ, et al** (2019) 4-Coumarate 3-hydroxylase in the lignin biosynthesis pathway is a cytosolic ascorbate peroxidase. *Nat Commun* **10**: 1994
- Barros J, Serk H, Granlund I, Pesquet E** (2015) The cell biology of lignification in higher plants. *Ann Bot* **115**: 1053–1074
- Boerjan W, Ralph J, Baucher M** (2003) LIGNIN BIOSYNTHESIS. *Annu Rev Plant Biol* **54**: 519–546
- Cesarino I** (2019) Structural features and regulation of lignin deposited upon biotic and abiotic stresses. *Current Opinion in Biotechnology* **56**: 209–214
- Cesarino I, Simões MS, dos Santos Brito M, Fanelli A, da Franca Silva T, Romanel E** (2016) Building the wall: recent advances in understanding lignin metabolism in grasses. *Acta Physiol Plant* **38**: 269
- Chen F, Dixon RA** (2007) Lignin modification improves fermentable sugar yields for biofuel production. *Nat Biotechnol* **25**: 759–761
- Clé C, Hill LM, Niggeweg R, Martin CR, Guisez Y, Prinsen E, Jansen MAK** (2008) Modulation of chlorogenic acid biosynthesis in *Solanum lycopersicum*; consequences for phenolic accumulation and UV-tolerance. *Phytochemistry* **69**: 2149–2156
- Clifford MN** (2000) Chlorogenic acids and other cinnamates - nature, occurrence, dietary burden, absorption and metabolism. *J Sci Food Agric* **11**
- Clifford MN, Jaganath IB, Ludwig IA, Crozier A** (2017a) Chlorogenic acids and the acyl-quinic acids: discovery, biosynthesis, bioavailability and bioactivity. *Nat Prod Rep* **34**: 1391–1421
- Clifford MN, Jaganath IB, Ludwig IA, Crozier A** (2017b) Chlorogenic acids and the acyl-quinic acids: discovery, biosynthesis, bioavailability and bioactivity. *Nat Prod Rep* **34**: 1391–1421
- Delli-Ponti R, Shivhare D, Mutwil M** (2021) Using Gene Expression to Study Specialized Metabolism—A Practical Guide. *Front Plant Sci* **11**: 625035
- Diaz J, Barcelo AR, De FM** (1997) Changes in shikimate dehydrogenase and the end products of the shikimate pathway, chlorogenic acid and lignins, during the early development of seedlings of *Capsicum annuum*. **7**
- Dixon RA, Barros J** (2019) Lignin biosynthesis: old roads revisited and new roads explored. *Open Biol* **9**: 190215

- Escamilla-Treviño LL, Shen H, Hernandez T, Yin Y, Xu Y, Dixon RA** (2014) Early lignin pathway enzymes and routes to chlorogenic acid in switchgrass (*Panicum virgatum* L.). *Plant Mol Biol* **84**: 565–576
- Ferreira SS, Goeminne G, Simões MS, Pina AV de A, Lima LGA de, Pezard J, Gutiérrez A, Rencoret J, Mortimer JC, del Río JC, et al** (2022) Transcriptional and metabolic changes associated with internode development and reduced cinnamyl alcohol dehydrogenase activity in sorghum. *Journal of Experimental Botany* **73**: 6307–6333
- Ferreira SS, Simões MS, Carvalho GG, de Lima LGA, Svartman RM de A, Cesarino I** (2019) The lignin toolbox of the model grass *Setaria viridis*. *Plant Mol Biol*. doi: 10.1007/s11103-019-00897-9
- Ferrer J-L, Austin MB, Stewart C, Noel JP** (2008) Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. *Plant Physiology and Biochemistry* **46**: 356–370
- Figueiredo R, Cesarino I, Mazzafera P** (2016) Suberin as an Extra Barrier to Grass Digestibility: a Closer Look to Sugarcane Forage. *Tropical Plant Biol* **9**: 96–108
- Gao W, Li H-Y, Xiao S, Chye M-L** (2010) Acyl-CoA-binding protein 2 binds lysophospholipase 2 and lysoPC to promote tolerance to cadmium-induced oxidative stress in transgenic *Arabidopsis*: Lysophospholipase in cadmium and oxidative stress. *The Plant Journal* no-no
- Gao W, Xiao S, Li H, Tsao S, Chye M** (2009) *Arabidopsis thaliana* acyl-CoA-binding protein ACBP2 interacts with heavy-metal-binding farnesylated protein AtFP6. *New Phytologist* **181**: 89–102
- Ha CM, Escamilla-Trevino L, Serrani Yarce JC, Kim H, Ralph J, Chen F, Dixon RA** (2016a) An essential role of caffeoyl shikimate esterase in monolignol biosynthesis in *Medicago truncatula*. *Plant J* n/a-n/a
- Ha CM, Escamilla-Trevino L, Yarce JCS, Kim H, Ralph J, Chen F, Dixon RA** (2016b) An essential role of caffeoyl shikimate esterase in monolignol biosynthesis in *Medicago truncatula*. *Plant J* **86**: 363–375
- Hoffmann L, Besseau S, Geoffroy P, Ritzenthaler C, Meyer D, Lapierre C, Pollet B, Legrand M** (2004) Silencing of Hydroxycinnamoyl-Coenzyme A Shikimate/Quinate Hydroxycinnamoyltransferase Affects Phenylpropanoid Biosynthesis. *Plant Cell* **16**: 1446–1465
- Hoffmann L, Maury S, Martz F, Geoffroy P, Legrand M** (2003) Purification, Cloning, and Properties of an Acyltransferase Controlling Shikimate and Quinate Ester Intermediates in Phenylpropanoid Metabolism. *J Biol Chem* **278**: 95–103
- Islam S** (2006) Sweetpotato (*Ipomoea batatas* L.) Leaf: Its Potential Effect on Human Health and Nutrition. *Journal of Food Science* **71**: R13–R121
- Joët T, Laffargue A, Salmona J, Doubeau S, Descroix F, Bertrand B, de Kochko A, Dussert S** (2009) Metabolic pathways in tropical dicotyledonous albuminous seeds: *Coffea arabica* as a case study. *New Phytologist* **182**: 146–162

- Julca I, Tan QW, Mutwil M** (2023) Toward kingdom-wide analyses of gene expression. *Trends in Plant Science* **28**: 235–249
- König S, Feussner K, Kaefer A, Landesfeind M, Thurow C, Karlovsky P, Gatz C, Polle A, Feussner I** (2014) Soluble phenylpropanoids are involved in the defense response of *A. rabidopsis* against *V. verticillium longisporum*. *New Phytologist* **202**: 823–837
- Koshiro Y, Jackson MC, Katahira R, Wang M-L, Nagai C, Ashihara H** (2007) Biosynthesis of Chlorogenic Acids in Growing and Ripening Fruits of *Coffea arabica* and *Coffea canephora* Plants. *Zeitschrift für Naturforschung C* **62**: 731–742
- Kundu A, Vadassery J** (2019) Chlorogenic acid-mediated chemical defence of plants against insect herbivores. *Plant Biol J* **21**: 185–189
- Laemmli UK** (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **227**: 680–685
- Lallemand LA, Zubieta C, Lee SG, Wang Y, Acajjaoui S, Timmins J, McSweeney S, Jez JM, McCarthy JG, McCarthy AA** (2012) A Structural Basis for the Biosynthesis of the Major Chlorogenic Acids Found in Coffee. *Plant Physiol* **160**: 249–260
- Legrand G, Delporte M, Khelifi C, Harant A, Vuylsteker C, Mörchen M, Hance P, Hilbert J-L, Gagneul D** (2016) Identification and Characterization of Five BAHD Acyltransferases Involved in Hydroxycinnamoyl Ester Metabolism in Chicory. *Front Plant Sci*. doi: 10.3389/fpls.2016.00741
- Lepelley M, Cheminade G, Tremillon N, Simkin A, Caillet V, McCarthy J** (2007) Chlorogenic acid synthesis in coffee: An analysis of CGA content and real-time RT-PCR expression of HCT, HQT, C3H1, and CCoAOMT1 genes during grain development in *C. canephora*. *Plant Science* **172**: 978–996
- Liu Q, Yao L, Xu Y, Cheng H, Wang W, Liu Z, Liu J, Cui X, Zhou Y, Ning W** (2019) In vitro evaluation of hydroxycinnamoyl CoA:quinic acid hydroxycinnamoyl transferase expression and regulation in *Taraxacum officinale* in relation to 5-caffeoylquinic acid production. *Phytochemistry* **162**: 148–156
- Livak KJ, Schmittgen TD** (2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* **25**: 402–408
- Martínez G, Regente M, Jacobi S, Del Rio M, Pinedo M, de la Canal L** (2017) Chlorogenic acid is a fungicide active against phytopathogenic fungi. *Pesticide Biochemistry and Physiology* **140**: 30–35
- Moglia A, Lanteri S, Comino C, Hill L, Knevitt D, Cagliero C, Rubiolo P, Bornemann S, Martin C** (2014) Dual Catalytic Activity of Hydroxycinnamoyl-Coenzyme A Quinic Transferase from Tomato Allows It to Moonlight in the Synthesis of Both Mono- and Dicafeoylquinic Acids. *Plant Physiol* **166**: 1777–1787
- Nakajima Y, Shimazawa M, Mishima S, Hara H** (2007) Water extract of propolis and its main constituents, caffeoylquinic acid derivatives, exert neuroprotective effects via antioxidant actions. *Life Sciences* **80**: 370–377
- Nelson RS, Stewart CN, Gou J, Holladay S, Gallego-Giraldo L, Flanagan A, Mann DGJ, Hisano H, Wuddineh WA, Poovaiah CR, et al** (2017) Development and use of a

- switchgrass (*Panicum virgatum* L.) transformation pipeline by the BioEnergy Science Center to evaluate plants for reduced cell wall recalcitrance. *Biotechnol Biofuels* **10**: 309
- Niggeweg R, Michael AJ, Martin C** (2004) Engineering plants with increased levels of the antioxidant chlorogenic acid. *Nat Biotechnol* **22**: 746–754
- Oñate-Sánchez L, Vicente-Carbajosa J** (2008) DNA-free RNA isolation protocols for *Arabidopsis thaliana*, including seeds and siliques. *BMC Research Notes* **1**: 93
- Payyavula RS, Shakya R, Sengoda VG, Munyaneza JE, Swamy P, Navarre DA** (2015) Synthesis and regulation of chlorogenic acid in potato: Rerouting phenylpropanoid flux in *HQT*-silenced lines. *Plant Biotechnol J* **13**: 551–564
- Ralph J, Lapierre C, Boerjan W** (2019) Lignin structure and its engineering. *Current Opinion in Biotechnology* **56**: 240–249
- Renault H, Werck-Reichhart D, Weng J-K** (2019) Harnessing lignin evolution for biotechnological applications. *Current Opinion in Biotechnology* **56**: 105–111
- del Río JC, Rencoret J, Prinsen P, Martínez ÁT, Ralph J, Gutiérrez A** (2012) Structural Characterization of Wheat Straw Lignin as Revealed by Analytical Pyrolysis, 2D-NMR, and Reductive Cleavage Methods. *J Agric Food Chem* **60**: 5922–5935
- Ruprecht C, Mutwil M, Saxe F, Eder M, Nikoloski Z, Persson S** (2011) Large-Scale Co-Expression Approach to Dissect Secondary Cell Wall Formation Across Plant Species. *Front Plant Sci*. doi: 10.3389/fpls.2011.00023
- Saleme M de LS, Cesarino I, Vargas L, Kim H, Vanholme R, Goeminne G, Van Acker R, Fonseca FC de A, Pallidis A, Voorend W, et al** (2017) Silencing *CAFFEOYL SHIKIMATE ESTERASE* Affects Lignification and Improves Saccharification in Poplar. *Plant Physiol* **175**: 1040–1057
- Shadle G, Chen F, Srinivasa Reddy MS, Jackson L, Nakashima J, Dixon RA** (2007) Down-regulation of hydroxycinnamoyl CoA: Shikimate hydroxycinnamoyl transferase in transgenic alfalfa affects lignification, development and forage quality. *Phytochemistry* **68**: 1521–1529
- Shen H, Fu C, Xiao X, Ray T, Tang Y, Wang Z, Chen F** (2009) Developmental Control of Lignification in Stems of Lowland Switchgrass Variety Alamo and the Effects on Saccharification Efficiency. *Bioenerg Res* **2**: 233–245
- Sonnante G, D'Amore R, Blanco E, Pierri CL, De Palma M, Luo J, Tucci M, Martin C** (2010) Novel Hydroxycinnamoyl-Coenzyme A Quinate Transferase Genes from Artichoke Are Involved in the Synthesis of Chlorogenic Acid. *Plant Physiol* **153**: 1224–1238
- Tohge T, Watanabe M, Hoefgen R, Fernie AR** (2013) The evolution of phenylpropanoid metabolism in the green lineage. *Critical Reviews in Biochemistry and Molecular Biology* **48**: 123–152
- Valiñas MA, Lanteri ML, ten Have A, Andreu AB** (2015) Chlorogenic Acid Biosynthesis Appears Linked with Suberin Production in Potato Tuber (*Solanum tuberosum*). *J Agric Food Chem* **63**: 4902–4913

- Van Acker R, Vanholme R, Storme V, Mortimer JC, Dupree P, Boerjan W** (2013) Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in *Arabidopsis thaliana*. *Biotechnol Biofuels* **6**: 46
- Vanholme B, Desmet T, Ronsse F, Rabaey K, Breusegem FV, Mey MD, Soetaert W, Boerjan W** (2013a) Towards a carbon-negative sustainable bio-based economy. *Front Plant Sci*. doi: 10.3389/fpls.2013.00174
- Vanholme R, Cesarino I, Rataj K, Xiao Y, Sundin L, Goeminne G, Kim H, Cross J, Morreel K, Araujo P, et al** (2013b) Caffeoyl Shikimate Esterase (CSE) Is an Enzyme in the Lignin Biosynthetic Pathway in *Arabidopsis*. *Science* **341**: 1103–1106
- Vanholme R, Cesarino I, Rataj K, Xiao Y, Sundin L, Goeminne G, Kim H, Cross J, Morreel K, Araujo P, et al** (2013c) Caffeoyl Shikimate Esterase (CSE) is an enzyme in the lignin biosynthetic pathway in *Arabidopsis*. *Science* **341**: 1103–1106
- Vanholme R, De Meester B, Ralph J, Boerjan W** (2019) Lignin biosynthesis and its integration into metabolism. *Current Opinion in Biotechnology* **56**: 230–239
- Vanholme R, Morreel K, Darrah C, Oyarce P, Grabber JH, Ralph J, Boerjan W** (2012) Metabolic engineering of novel lignin in biomass crops. *New Phytol* **196**: 978–1000
- Vogt T** (2010) Phenylpropanoid Biosynthesis. *Molecular Plant* **3**: 2–20
- Volpi e Silva N, Mazzafera P, Cesarino I** (2019) Should I stay or should I go: are chlorogenic acids mobilized towards lignin biosynthesis? *Phytochemistry* **166**: 112063
- de Vries L, Brouckaert M, Chanoca A, Kim H, Regner MR, Timokhin VI, Sun Y, De Meester B, Van Doorselaere J, Goeminne G, et al** (2021) CRISPR-Cas9 editing of CAFFEYOYL SHIKIMATE ESTERASE 1 and 2 shows their importance and partial redundancy in lignification in *Populus tremula* × *P. alba*. *Plant Biotechnol J* **19**: 2221–2234
- Wang A, Dennis EA** (1999) Mammalian lysophospholipases. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **1439**: 1–16
- Wang JP, Matthews ML, Williams CM, Shi R, Yang C, Tunlaya-Anukit S, Chen H-C, Li Q, Liu J, Lin C-Y, et al** (2018) Improving wood properties for wood utilization through multi-omics integration in lignin biosynthesis. *Nat Commun* **9**: 1579
- Wang JP, Naik PP, Chen H-C, Shi R, Lin C-Y, Liu J, Shuford CM, Li Q, Sun Y-H, Tunlaya-Anukit S, et al** (2014) Complete Proteomic-Based Enzyme Reaction and Inhibition Kinetics Reveal How Monolignol Biosynthetic Enzyme Families Affect Metabolic Flux and Lignin in *Populus trichocarpa*. *Plant Cell* **26**: 894–914
- Wang M, Simon JE, Aviles IF, He K, Zheng Q-Y, Tadmor Y** (2003) Analysis of Antioxidative Phenolic Compounds in Artichoke (*Cynara scolymus* L.). *J Agric Food Chem* **51**: 601–608
- Weng J-K, Chapple C** (2010) The origin and evolution of lignin biosynthesis: Tansley review. *New Phytologist* **187**: 273–285
- Xiong W, Wu Z, Liu Y, Li Y, Su K, Bai Z, Guo S, Hu Z, Zhang Z, Bao Y, et al** (2019) Mutation of 4-coumarate: coenzyme A ligase 1 gene affects lignin biosynthesis and

increases the cell wall digestibility in maize brown midrib5 mutants. *Biotechnol Biofuels* **12**: 82

Yamaguchi T, Chikama A, Mori K, Watanabe T, Shioya Y, Katsuragi Y, Tokimitsu I (2008) Hydroxyhydroquinone-free coffee: A double-blind, randomized controlled dose–response study of blood pressure. *Nutrition, Metabolism and Cardiovascular Diseases* **18**: 408–414

Zhang J, Yang Y, Zheng K, Xie M, Feng K, Jawdy SS, Gunter LE, Ranjan P, Singan VR, Engle N, et al (2018) Genome-wide association studies and expression-based quantitative trait loci analyses reveal roles of HCT2 in caffeoylquinic acid biosynthesis and its regulation by defense-responsive transcription factors in *Populus*. *New Phytol* **220**: 502–516

Zhao Q, Gallego-Giraldo L, Wang H, Zeng Y, Ding S-Y, Chen F, Dixon RA (2010) An NAC transcription factor orchestrates multiple features of cell wall development in *Medicago truncatula*: *Medicago NST* gene. *The Plant Journal* no-no

Zhong R, Ye Z (2009) Secondary Cell Walls. eLS. doi: 10.1002/9780470015902.a0021256