**Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”**

Identificação de genes de maracujá azedo diferencialmente expressos durante a interação com *Xanthomonas axonopodis*

Carla de Freitas Munhoz

Tese apresentada para obtenção do título de
Doutora em Ciências. Área de concentração:
Genética e Melhoramento de Plantas

**Piracicaba
2013**

Carla de Freitas Munhoz
Bióloga

**Identificação de genes de maracujá azedo diferencialmente expressos durante
a interação com *Xanthomonas axonopodis***

Orientadora:
Profa. Dra. **MARIA LUCIA CARNEIRO VIEIRA**

Tese apresentada para obtenção do título de
Doutora em Ciências. Área de concentração:
Genética e Melhoramento de Plantas

**Piracicaba
2013**

**Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP**

Munhoz, Carla de Freitas

Identificação de genes de maracujá azedo diferencialmente expressos durante a interação com *Xanthomonas axonopodis* / Carla de Freitas Munhoz.- - Piracicaba, 2013.

174 p: il.

Tese (Doutorado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2013.

1. Interação planta-patógeno 2. Expressão diferencial 3. Biblioteca subtrativa
4. Anotação funcional 5. Genes de defesa 6. qPCR 7. *Passiflora edulis* 8. Bacteriose
I. Título

CDD 634.425
M966i

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte -O autor"

DEDICATÓRIA

À Deus, pela graça da vida,
à minha mãe, Vanja,
ao meu noivo, Cícero,
aos meus familiares
e a todos os amigos que me apoiaram,
com amor dedico este trabalho.

AGRADECIMENTOS

À Escola Superior de Agricultura “Luiz de Queiroz” (ESALQ / USP) e ao Departamento de Genética, pela oportunidade de realizar o curso de Doutorado.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelas bolsas concedidas.

À Professora Dra. Maria Lucia Carneiro Vieira pela orientação, apoio, confiança e oportunidades concedidas durante toda a minha permanência no laboratório e durante o desenvolvimento deste trabalho.

À Professora Dra. Claudia de Barros Monteiro-Vitorello, da ESALQ/USP, pela amizade e colaboração durante as análises de bioinformática.

Ao Professor Dr. Luis Eduardo Aranha Camargo, da ESALQ/USP, pelas sugestões e colaboração durante o planejamento experimental.

Aos Professores do Programa de Pós-Graduação em Genética e Melhoramento de Plantas, pelos conhecimentos compartilhados.

À Dra. Danielle Gregorio Gomes Caldas, pela amizade, atenção e esclarecimentos durante a construção das bibliotecas.

Ao Dr. Rafael Augusto Arenhart, pela orientação nas análises de qPCR.

Ao Carlos Alberto de Oliveira, pela amizade e auxílio técnico durante a realização do trabalho.

Aos funcionários do Departamento de Genética pelo convívio amigável e pela colaboração durante a execução do trabalho.

Aos colegas do Laboratório de Biologia Molecular de Plantas (ESALQ/USP) Aline, Thais e Lígia e aos colegas do Laboratório de Biologia Celular e Molecular (CENA/USP) Gustavo e Milene, pela disponibilidade e atenção nos momentos de dúvidas.

Aos amigos do Laboratório de Genética Molecular de Plantas Cultivadas, Alessandra, Anselmo, Augusto, Bruna, Bruno, Carmelice, Cassia, Guilherme, Hanai, Helen, Luciane, Marcelo, Taislene e Zirlane, pelos auxílios, sugestões e companheirismo.

A todos que contribuíram para a realização deste trabalho.

SUMÁRIO

RESUMO-----	9
ABSTRACT-----	11
1 INTRODUÇÃO-----	13
2 REVISÃO BIBLIOGRÁFICA-----	17
2.1 A cultura do maracujá no Brasil e a ocorrência da bacteriose-----	17
2.2 A interação patógeno-hospedeiro-----	20
2.3 Moléculas envolvidas nas respostas de defesa vegetal-----	24
2.4 Identificação de genes de defesa contra patógenos-----	31
2.5 Ferramentas de Bioinformática para a anotação de sequências gênicas-----	34
3 MATERIAL E MÉTODOS-----	39
3.1 Material biológico-----	39
3.2 Obtenção e inoculação das mudas-----	39
3.3 Condução dos experimentos de inoculação-----	40
3.4 Construção das bibliotecas subtrativas-----	44
3.4.1 <i>Extração de RNA</i> -----	49
3.4.2 <i>Síntese de cDNA</i> -----	50
3.4.3 <i>Obtenção das mini-bibliotecas de cDNA</i> -----	52
3.5 Clonagem e sequenciamento dos fragmentos de cDNAs das bibliotecas-----	54
3.6 Processamento das sequências-----	56
3.7 Anotação funcional das bibliotecas subtrativas-----	58
3.8 Validação das bibliotecas subtrativas-----	60
3.8.1 <i>Síntese de cDNA</i> -----	60
3.8.2 <i>Seleção de sequências putativas codificadoras das bibliotecas e desenho de primers</i> -----	62
3.8.3 <i>Seleção e validação de genes de referência para a normalização da qPCR</i> -----	63
3.8.4 <i>Validação da expressão diferencial dos transcritos por qPCR</i> -----	64
3.9 Análise detalhada das sequências dos genes da lipoxigenase e da (+)-neomentol desidrogenase-----	66
4 RESULTADOS-----	67
4.1 Avaliação dos experimentos de inoculação-----	67
4.2 Construção das bibliotecas subtrativas-----	68
4.3 Clonagem dos fragmentos de cDNAs das bibliotecas subtrativas-----	71

4.4 Processamento das sequências -----	73
4.5 Anotação funcional das bibliotecas subtrativas -----	76
4.5.1 Anotação das bibliotecas com base na Plataforma PLAZA-----	76
4.5.2 Anotação das bibliotecas com base no programa computacional Blast2GO---	82
4.6 Validação das bibliotecas subtrativas -----	89
4.6.1 Seleção de sequências putativas codificadoras das bibliotecas e desenho de primers -----	89
4.6.2 Seleção e validação de genes de referência para a normalização da qPCR --	93
4.6.3 Validação da expressão diferencial dos transcritos por qPCR-----	96
4.7 Análise detalhada das sequências dos genes da lipoxigenase e da (+)-neomentol desidrogenase -----	104
5 DISCUSSÃO-----	115
5.1 Processamento e anotação das sequências das bibliotecas subtrativas -----	115
5.2 Proteínas envolvidas em respostas de defesa durante a interação maracujá azedo- <i>Xap</i> -----	120
5.3 Validação da expressão diferencial dos transcritos por qPCR -----	125
6 CONCLUSÕES -----	135
REFERÊNCIAS-----	137
ANEXOS-----	153

RESUMO

Identificação de genes de maracujá azedo diferencialmente expressos durante a interação com *Xanthomonas axonopodis*

O Brasil é o maior produtor mundial de maracujá azedo (*Passiflora edulis* f. *flavicarpa*) sendo esta a espécie de maior expressão comercial dentre as passifloras cultivadas. A bacteriose do maracujazeiro, causada por *Xanthomonas axonopodis* pv. *passiflorae* (*Xap*), é uma das doenças mais severas da cultura, acarretando grandes prejuízos aos produtores. Atualmente, é incipiente o conhecimento sobre a interação maracujá azedo-*Xap*. Diante disso, a identificação e a caracterização dos genes envolvidos no processo de defesa são passos importantes para dar suporte ao desenvolvimento de variedades resistentes. Assim, o objetivo deste trabalho foi identificar e caracterizar genes de maracujá azedo diferencialmente expressos durante a resposta de defesa à *Xap*, bem como mensurar a sua expressão. Para isso, foram construídas duas bibliotecas subtrativas de cDNA (*forward* e *reverse*) usando o método SSH a partir de transcritos de folhas, que foram inoculadas com o patógeno ou solução salina (controle). Após o sequenciamento dos clones, o processamento e a montagem das sequências, as unisequências foram anotadas através da Plataforma PLAZA e do programa computacional Blast2GO. Genes envolvidos em diversos processos biológicos foram selecionados para a validação das bibliotecas por PCR quantitativo. Usando a Plataforma PLAZA, 78 % (764) das unisequências mostraram similaridade com proteínas de *Arabidopsis thaliana*, enquanto 87 % (866) delas apresentaram similaridade com proteínas putativas de diversas espécies vegetais, quando se utilizou Blast2GO. Na biblioteca *forward*, foram identificadas 73 proteínas relacionadas à resposta de defesa, dentre as quais estão proteínas envolvidas na sinalização intracelular, na ativação da transcrição e regulação da expressão de genes de defesa, bem como proteínas de defesa, de resistência e relacionadas à patogênese (PRs). Dentre os 22 transcritos validados, 95 % foram diferencialmente expressos em pelo menos um dos três períodos avaliados; os genes mais expressos em resposta à infecção pelo patógeno são os que codificam as enzimas lipoxigenase, (+)-neomentol desidrogenase e quitinase, as quais participam diretamente nas respostas de defesa vegetal. Dos genes cuja expressão foi mais reprimida, dois codificam proteínas relacionadas à fotossíntese e dois codificam proteínas envolvidas na detoxificação da amônia e do H₂O₂. Nossos resultados sugerem que a planta utiliza um arsenal de transcritos para responder à infecção; entretanto, este arsenal não é eficiente para impedir a ação do patógeno e, consequentemente, o desenvolvimento da bacteriose nas condições estudadas. Nosso estudo é inédito e gerou informações sobre a reprogramação transcricional durante a interação maracujá azedo-*Xap*, o que constitui um importante passo para o melhor entendimento sobre este patossistema.

Palavras-chave: Interação planta-patógeno; Expressão diferencial; Biblioteca subtrativa; Anotação funcional; Genes de defesa; qPCR, *Passiflora edulis*; Bacteriose

ABSTRACT

Identification of differentially expressed genes during the yellow passion fruit–*Xanthomonas axonopodis* interaction

Brazil is the main producer of yellow passion fruit (*Passiflora edulis* f. *flavicarpa*) worldwide, which is the most widely commercialized crop among the cultivated passifloras. The bacterial leaf spot induced by *Xanthomonas axonopodis* pv. *passiflorae* (*Xap*) is one of the most severe diseases of the crop, causing great losses to producers. Currently, we understand very little about the yellow passion fruit-*Xap* interaction. Therefore, the identification and characterization of genes involved in the defense process are important steps to support the development of resistant varieties. Thus, the objective of this study was identify and characterize differentially expressed genes during the defense response to *Xap*, as well as to measure their expression. For that, we constructed two subtractive cDNA libraries (the forward and the reverse) by performing the SSH method from leaf transcripts, which were inoculated with the pathogen or saline solution (control). After sequencing the clones and sequence data processing, sequences were assembled into unique sequences, which were annotated using the PLAZA Platform and the computational program Blast2GO. Genes involved in several biological processes were selected to validate the libraries by quantitative PCR. When PLAZA was used for sequence similarity searches, 78 % (764) of the yellow passion fruit unique sequences showed similarity to proteins of *Arabidopsis thaliana*; when Blast2GO was used, 87 % (866) of the unique sequences showed similarities to putative proteins of several plant species. For the forward library, 73 proteins related to defense response were identified, such as those involved in intracellular signaling, transcription activation and regulation of defense gene expression, as well as defense and resistance proteins, and pathogenesis-related proteins (PRs). Of the 22 validated transcripts, 95 % were differentially expressed during at least one of the three periods evaluated; the genes up-regulated in response to the pathogen infection were those that code for the enzymes lipoxygenase, (+)-neomenthol dehydrogenase and chitinase, which participate directly in plant-defense responses. Out of down-regulated genes, two code for photosynthesis-related proteins, and two for ammonia and H₂O₂ detoxification. Our results suggest the plant uses an arsenal of transcripts to respond to infection; however, this arsenal is not effective to prevent pathogen action and consequently the occurrence of bacterial leaf spot under the evaluated conditions. The present study is the first to produce information on the transcriptional reprogramming during the passion fruit-*Xap* interaction, which represents an important step for a better understanding of this pathosystem.

Keywords: Plant-pathogen interaction; Differential expression; Subtractive library; Functional annotation; Defense genes; qPCR, *Passiflora edulis*; Bacterial leaf spot

1 INTRODUÇÃO

O controle das doenças de plantas tem exigido estudos na busca de métodos alternativos que proporcionem eficiência, mas, também, menor agressividade à saúde humana e ao equilíbrio dos agroecossistemas. Assim, para o futuro próximo, as estratégias mais adequadas de controle deverão ativar os mecanismos naturais de defesa das plantas, para que estas se protejam contra os patógenos, diminuindo ou mesmo dispensando o uso de defensivos tóxicos (ROMEIRO, 2008).

Dessa forma, o entendimento sobre os mecanismos bioquímicos, genéticos (herança) e moleculares que regulam a resposta das plantas à presença do patógeno é de grande importância para o desenvolvimento de tecnologias que ativem efetivamente os mecanismos naturais de defesa, aumentando a resistência a doenças.

Quando uma planta é hospedeira de um determinado patógeno, algumas variedades podem ser susceptíveis ou mais susceptíveis, enquanto outras podem ser resistentes ou mais resistentes ao mesmo patógeno. Isto depende do tipo e do número de genes (e alelos) de resistência presentes na planta, das condições ambientais prevalentes, além de outros fatores. Por isso, mesmo quando uma planta desenvolve uma doença provocada por um patógeno, ainda assim ela ativa diversas respostas de defesa. Na maioria dos casos, a planta consegue limitar a propagação do patógeno através de compostos químicos e estruturas de defesa pré-existentes (como fenóis, alcaloides, cutícula e estômatos) ou induzidos pela infecção (como, espécies reativas de oxigênio, fitoalexinas, lignificação e papilas) que bloqueiam a sua expansão (AGRIOS, 2005).

O Brasil destaca-se por ser o maior produtor e consumidor mundial de maracujá azedo (*Passiflora edulis* f. *flavicarpa*) e sua produção vem aumentando sensivelmente nos últimos anos. No cenário da fruticultura brasileira, a cultura do maracujá é uma excelente fonte de renda para a agricultura familiar, especialmente na região Nordeste, além de apresentar grande interesse dos setores comercial e agroindustrial (Veja PIRES; SÃO JOSÉ; CONCEIÇÃO, 2011). Dentre os problemas fitossanitários enfrentados pela monocultura do maracujazeiro, a bacteriose, causada por *Xanthomonas axonopodis* pv. *passiflorae* (Xap), é um dos mais

severos, dada à dificuldade de controle, principalmente, controle químico eficiente (JUNQUEIRA; JUNQUEIRA, 2007).

Embora a bacteriose do maracujazeiro cause grandes prejuízos à cultura, há poucos estudos que visam ao melhor entendimento sobre o patossistema maracujá azedo-*Xap*. Há indícios que a resistência à bacteriose é controlada por oligogenes (LOPES et al., 2006). Entretanto, não há variedades selecionadas para acumular e expressar alelos de resistência. Também, a literatura científica não dispõe de informações suficientes sobre o número e efeito dos genes envolvidos no processo de defesa a essa doença. Por isso, a identificação e a caracterização de genes envolvidos na defesa à *Xap* poderão ser úteis para o entendimento desse patossistema, bem como para acelerar os processos de seleção em programas de melhoramento.

Uma opção simples e adequada para o estudo de genes diferencialmente expressos durante a interação planta-patógeno é usar a metodologia de hibridização subtrativa por supressão (SSH), desenvolvida por Diathenko et al. (1996), a qual permite selecionar os genes expressos apenas em um dos tratamentos (plantas inoculadas ou não inoculadas com o patógeno) e não requer conhecimento prévio de sequências do genoma sob estudo.

Embora novas tecnologias de sequenciamento, como o *Illumina*[®] e o *454 Sequencing* (Roche), permitam o sequenciamento rápido e completo de todo o transcrito de tratamentos em análise, essas ainda são inviáveis para aplicação em espécies, como o maracujá azedo, que não apresentam dados genômicos suficientes e disponíveis, para servirem como referência durante a montagem das milhões de pequenas sequências geradas. Assim, mini bibliotecas subtrativas, sequenciadas pela metodologia de Sanger (SANGER et al., 1977), ainda são mais apropriadas nestes casos, permitindo a obtenção de sequências de cDNA de boa qualidade e de tamanhos adequados para a anotação.

Diversas ferramentas de bioinformática, como o BLAST (ALTSCHUL et al., 1990), programas de anotação automática, como o Blast2GO (CONESA et al., 2005), e uma infinidade de bancos de dados públicos de sequências nucleotídicas e de proteínas (FERNANDEZ-SUAREZ; GALPERIN, 2012), fornecem os subsídios necessários para atribuir informações funcionais às sequências gênicas identificadas nessas bibliotecas. Além disso, a RT-qPCR, a qual combina a transcrição reversa

com a reação em cadeia da polimerase, é um importante método para quantificar a expressão de genes, com elevada sensibilidade, especificidade e eficiente detecção de sinal (GUTIERREZ et al., 2008), possibilitando a confirmação da expressão diferencial desses genes.

Neste contexto, o presente trabalho teve os seguintes objetivos: (i) construir duas bibliotecas subtrativas de cDNA de maracujá azedo pelo método SSH, sendo uma a partir de transcritos induzidos em resposta à inoculação com *Xap* e outra controle; (ii) clonar e sequenciar os fragmentos de cDNA obtidos nas bibliotecas; (iii) anotar os genes candidatos diferencialmente expressos, utilizando ferramentas de bioinformática; (iv) validar o envolvimento de genes candidatos diferencialmente expressos no processo de defesa, por qPCR, em três períodos de interação, após a inoculação; (v) gerar um conjunto preliminar de informações sobre o interatoma maracujá-azedo-*Xap*.

Cabe comentar que, recentemente, nosso grupo de pesquisa, em parceria com o Centro Nacional de Recursos Genéticos Vegetais do INRA de Toulouse, França, construiu uma biblioteca genômica para o maracujá azedo inserida em BACs (<http://cnrgv.toulouse.inra.fr/Library/Passiflora>). Poucos milhares de sequências já foram obtidas e anotadas por Santos (2013), o que representa um enorme avanço, uma vez que há um escasso número de genes descritos para *Passiflora*. Somando-se a isso, os fragmentos de genes diferencialmente expressos aqui gerados poderão ser usados como sondas para localizá-los na biblioteca de BACs. Assim, o sequenciamento dos BACs contendo esses genes permitirá a identificação de suas sequências completas e a caracterização dos seus elementos regulatórios.

Também, a identificação dos genes diferencialmente expressos durante a interação maracujá azedo-*Xap* trará importantes contribuições à comunidade científica, já que estas sequências expressas poderão ser aproveitadas para: (a) o desenvolvimento de novos marcadores moleculares; (b) o mapeamento genético de locos quantitativos envolvidos na resposta de populações segregantes à infecção por *Xap*; (c) o desenvolvimento de sondas para ensaios de hibridização *in situ*, possibilitando a integração dos mapas genéticos (OLIVEIRA et al., 2008; PEREIRA et al., 2013), aos cromossomos de maracujá azedo ($2n = 18$, PENHA, 2012).

2 REVISÃO BIBLIOGRÁFICA

2.1 A cultura do maracujá no Brasil e a ocorrência da bacteriose

Maracujá é um termo indígena usado para se referir ao fruto (*cuia*) de plantas pertencentes ao gênero *Passiflora*. Este gênero pertence à família Passifloraceae, a qual apresenta cerca de 520 espécies descritas (MacDOUGAL; FEUILLET, 2004) e amplamente distribuídas no neotrópico. O Brasil destaca-se por ser um importante centro de diversidade da família, observando-se aproximadamente 148 espécies nativas (veja BERNACCI et al., 2013).

O Brasil é o principal produtor mundial de maracujás, bem como o maior consumidor da fruta fresca e processada. O seu cultivo ocorre em praticamente todos os Estados brasileiros, destacando-se as regiões Norte e Nordeste como as maiores produtoras, pois, por serem regiões de clima propício para a cultura, é possível se produzir a fruta durante quase o ano inteiro. No Brasil, a cultura é inerente a pequenas propriedades, pois apresenta uma boa opção como fonte de renda para a agricultura familiar. As exportações brasileiras, na forma de suco concentrado, são destinadas principalmente para os Estados Unidos e Países Baixos (veja PIRES et al., 2011).

Além do consumo dos frutos, que possuem apreciado sabor e aroma e são ricos em sais minerais e vitaminas A e C, o seu cultivo apresenta outras finalidades comerciais como a produção de flores, para o setor ornamental (revisado por ABREU et al., 2009), de folhas e ramos, para a indústria farmacêutica, por possuir substâncias com atividades calmantes e sedativas (DENG et al, 2010) e para a indústria de cosméticos (MICHALUN; MICHALUN, 2010).

As espécies cultivadas no Brasil são o maracujá azedo, *Passiflora edulis* f. *flavicarpa*, o maracujá doce, *P. alata*, e mais raramente o maracujá roxo, *Passiflora edulis* (revisado por VIEIRA; CARNEIRO, 2005).

O maracujá azedo é a espécie mais expressiva comercialmente dentre as passifloras cultivadas no Brasil, pois apresenta alta produtividade, vigor e qualidade dos frutos. É plantado em aproximadamente 95 % dos pomares comerciais e,

segundo dados do IBGE, sua produção foi de 718.789 toneladas em 2009, em uma área de cultivo de 50.795 ha, representando um aumento de 39,4 % em relação a produção de 2008 (Produção Agrícola Municipal, 2009). Nos últimos cinco anos, o consumo da fruta está aumentando no Brasil, atingindo 21 % do volume de frutos negociado pelo CEAGESP-SP (AGRIANUAL, 2012).

Com a demanda por frutos de maracujá e consequente aumento da área plantada em monocultura, naturalmente, houve maior incidência de doenças e pragas nos pomares. As doenças são causadas por fungos, vírus e bactérias. Tais problemas acarretam redução do tempo de exploração comercial dos pomares e, até mesmo, impossibilitam seu plantio em algumas regiões, o que exige esforços para que as perdas diminuam e se impeça a disseminação dos patógenos (veja SÃO JOSÉ; PIRES, 2011).

Dentre as doenças que atacam os maracujazeiros, a bacteriose causada por *Xanthomonas axonopodis* pv. *passiflorae* (Xap), também conhecida por crestamento bacteriano ou por mancha oleosa dos frutos, é uma das doenças mais severas, por não se dispor de controle químico eficiente, acarretando grandes perdas na produção de frutos (veja SANTOS FILHO et al., 2004; KIMATI et al., 2005). *Xanthomonas axonopodis* tem como hospedeiro várias espécies cultivadas, sendo que o patovar *passiflorae* específico do gênero *Passiflora* (LIBERATO, 2002); a bactéria pode ser encontrada nos pomares das espécies cultivadas *P. edulis* f. *flavicarpa*, *P. edulis*, *P. quadrangularis* e *P. alata* e nas silvestres, *P. cincinnata*, *P. nitida*, *P. amethystina* e *P. setacea* (JUNQUEIRA et al., 2005).

Temperaturas elevadas, acima dos 30°C, e alta umidade relativa do ar (veja VIANA et al., 2003) estabelecem as condições ambientais favoráveis ao desenvolvimento da doença. A bactéria é introduzida no pomar através de mudas contaminadas e se dissemina pela água da chuva, pelos instrumentos de poda e colheita, ou pelo próprio homem (SANTOS FILHO et al., 2004). Também, insetos que se alimentam de folhas infectadas, como o *Diabrotica speciosa* (vaquinha-brasileirinho) e o *Lagria villosa* (besouro-idiamim), podem ser vetores do patógeno (veja JUNQUEIRA; JUNQUEIRA, 2007).

A doença ataca a parte aérea da planta, apresentando formas de infecção localizada ou sistêmica, que podem ou não ocorrerem associadas. Na forma localizada, os sintomas foliares se iniciam no limbo com manchas angulares,

translúcidas, de coloração verde-escura e aspecto encharcado que evoluem para uma coloração pardacenta e seca, envoltas por um halo amarelo. Se a umidade permanecer superior a 80%, as lesões aumentam de tamanho, atingindo todo o limbo e ocasionando a seca e a queda das folhas. Já na forma sistêmica, a infecção pode avançar através das nervuras, evoluir para o pecíolo até atingir os vasos dos caules mais finos, o que provoca caneluras longitudinais e a seca dos órgãos, consequentemente ocorre intensa desfolha e a morte prematura da planta (veja SANTOS FILHO et al., 2004; KIMATI et al., 2005).

Nos frutos maduros, provoca manchas superficiais delimitadas, oleosas e de coloração verde-escura a marrom, podendo causar grandes lesões ao se unirem. Em condições favoráveis, a bactéria pode atingir a polpa, o que contamina as sementes, fermentando-as, e pode causar a queda dos frutos, inviabilizando o seu comércio (veja VIANA et al., 2003; JUNQUEIRA; JUNQUEIRA, 2007).

As medidas de prevenção são ainda as melhores estratégias de controle da bacteriose. Para isso, recomenda-se a eliminação das plantas infectadas e a esterilização da área de cultivo, o plantio de capim alto ao redor do pomar para servir de quebra-vento, bem como o uso de mudas e sementes saudáveis. Para eliminar a bactéria, tratam-se as sementes com água a 50°C durante 30 a 60 min. Também, são recomendadas medidas culturais como a troca do local de plantio para áreas distantes e não contaminadas, adubação e calagem apropriadas e controle de plantas invasoras e insetos (veja SANTOS FILHO et al., 2004; JUNQUEIRA; JUNQUEIRA, 2007).

Potenciais indutores de resistência para o controle da bacteriose do maracujazeiro foram indicados no trabalho de Boro, Beriam e Guzzo (2011). Os pesquisadores avaliaram a indução da resistência de maracujá azedo à *Xap* usando o indutor abiótico acibenzolar-S-metil (ASM) e os agentes bióticos proteína harpina, extraída de *Erwinia amylovora*, e glicoproteínas, extraídas de *Xap* e *X. campestris* pv. *campestris*. Nas plântulas as quais tiveram as folhas tratadas por aspersão com os indutores, 72 h antes da inoculação com *Xap*, o tratamento com ASM mostrou 70 % de proteção para as plantas, o tratamento com as glicoproteínas resultou em 72 % de proteção e o tratamento com a proteína harpina induziu a um aumento nos sintomas da doença. Já nas plântulas inoculadas que tiveram as sementes tratadas

com os indutores, os tratamentos com ASM e com proteína harpina conferiram proteção de até 90 % e 47 %, respectivamente, contra a doença.

Ao analisar uma coleção de 90 isolados de *Xap*, Munhoz et al. (2011) relataram haver alta variabilidade genética entre os isolados. Estes resultados, aliados a alta dificuldade de controle desta doença, reforçam a importância do uso de variedades resistentes. Dessa forma, são necessárias pesquisas que visem dar suporte à busca de fontes de resistência vegetal e à geração de variedades resistentes.

2.2 A interação patógeno-hospedeiro

Patógenos podem conviver na superfície aérea, caulinar ou radicular de uma planta hospedeira sem causar doença. Porém, a combinação de injúrias com ambiente favorável é determinante para a infecção da planta pelo patógeno. A doença, no entanto, é resultado de um processo em nível celular, fisiológico, metabólico, bioquímico, molecular e genético. Durante a interação patógeno-hospedeiro, o patógeno coloniza a planta a fim de se nutrir e se reproduzir; mas, o sucesso dessa colonização dependerá dos mecanismos de patogenicidade e virulência do patógeno e das respostas de defesa da planta (WULFF, 2008).

Os principais gêneros de bactérias fitopatogênicas são *Acidovorax*, *Agrobacterium*, *Burkholderia*, *Clavibacter*, *Curtobacterium*, *Erwinia*, *Leifsonia*, *Pseudomonas*, *Ralstonia*, *Rhodococcus*, *Streptomyces*, *Xanthomonas* e *Xylella*. Grande parte dessas bactérias possui fatores de virulência estruturais em sua superfície externa ou são secretados pela célula para o ambiente (WULFF, 2008). Tais fatores de virulência, quando reconhecidos pelas plantas, são denominados ‘padrões moleculares associados a patógenos’ ou PAMP (*Pathogen-Associated Molecular Patterns*) (NÜRNBERGER; BRUNNER, 2002): são enzimas bacterianas degradadoras de parede celular, toxinas, moléculas sinalizadoras, hormônios, sideróforos, flagelos, *pili*, lipopolissacarídeos, etc (FINLAY; FALKOW, 1997).

Microrganismos desenvolveram estratégias especializadas para suprimir as respostas de defesa da planta e induzir a susceptibilidade em hospedeiros outrora resistentes. Para isso, secretam proteínas, como as toxinas, as proteínas efetoras e

outros fatores, capazes de entrar na célula hospedeira e modificar a fisiologia do hospedeiro, promovendo a sua colonização. Dentre os processos de defesa vegetal que são alvos dessas moléculas estão a morte celular programada, a defesa baseada na parede celular, a sinalização hormonal, a expressão de genes de defesa e outras respostas basais. Muitos efetores almejam mais de um componente no hospedeiro, bem como, há evidências de que algumas dessas proteínas são capazes de interferir em múltiplos processos vegetais. Bactérias Gram-negativas usam sistemas de secreção especializados para facilitar este processo, como o sistema de secreção tipo III (SST3), enquanto mollicutes e bactérias gram-positivas podem se utilizar apenas das vias de secreção gerais (ABRAMOVITCH; MARTIN, 2004; TSENG; TYLER; SETUBAL, 2009; DESLANDES; RIVAS, 2012).

O gênero *Xanthomonas* é composto quase que exclusivamente por espécies fitopatogênicas, causando importantes doenças em todas as áreas cultivadas do mundo. As espécies que constituem este gênero são Gram-negativas e penetram na planta hospedeira apenas através de ferimentos ou por aberturas naturais, como estômatos, hidatódios, lenticelas ou nectários (SWINGS; CIVEROLO, 1993). *Xanthomonas spp.* contém genes codificadores de proteínas secretadas por todos os sistemas conhecidos em bactérias Gram-negativas, porém, a maioria desses sistemas de secreção em *Xanthomonas* e sua importância na virulência não são conhecidos. No entanto, sabe-se que o SST3 é essencial para a sua patogenicidade. A maioria das proteínas secretadas pelo SST3, denominadas efetores tipo III, também conhecidas por proteínas de avirulência (Avr), são translocadas diretamente para o interior da célula vegetal, mas a ação molecular da maioria delas ainda não foi elucidada (KAY; BONAS, 2009).

O conhecimento sobre a interação planta-*Xanthomonas spp.* tem tido considerável progresso, no entanto, ainda existem muitas questões a serem respondidas. De modo geral, a espécies patogênicas e seus patovares apresentam um alto grau de especificidade ao hospedeiro e muitas mostram especificidade ao tecido que invadem, podendo este ser do sistema vascular ou do mesófilo. Análises genômicas comparativas de *Xanthomonas spp.* tem mostrado genes conservados, *clusters* gênicos associados à virulência e diferenças no conteúdo genético, sendo que algumas dessas diferenças podem estar relacionadas à especificidade ao hospedeiro e/ou tecido. Também, essas análises revelaram que todos os genomas

de *Xanthomonas spp.* tem numerosos genes que codificam enzimas putativas para a degradação da parede celular vegetal e que a maioria dos genomas sequenciados contém um conjunto de nove genes que codificam efetores tipo III. Além disso, estes genomas se mostram altamente plásticos, de rápida evolução, como, por exemplo, *X. albilineans*, que sofreu considerável redução no tamanho do seu genoma, o que pode estar associado com o seu estilo de vida limitado ao xilema (Revisado por RYAN et al., 2011).

Para elucidar, as proteínas AvrBs3 ou TAL (*Transcription Activator-Like*) pertencem a uma grande família de efetores identificada em *Xanthomonas spp.*, secretadas por SST3. Estes efetores atuam como ativadores transcricionais no núcleo da célula hospedeira, modulando a expressão dos genes da planta. Estas proteínas contêm um sinal de localização nuclear, para guiá-la ao núcleo da célula vegetal, um domínio de ativação C-terminal, necessário para ativar a expressão dos genes da planta, e um domínio de repetição central, que direciona a ligação ao DNA. Nas plantas suscetíveis, os TAL ativam a expressão de genes de suscetibilidade para ajudar na virulência da bactéria. Nas plantas resistentes, estes efetores são atraídos a induzir a expressão de genes de resistência, porém, o reconhecimento pelas proteínas R induzem reações específicas de defesa (BOCH; BONAS, 2010).

Quando o patógeno penetra na parede celular vegetal, encontra receptores de membrana que reconhecem os PAMP. Ao perceber um microorganismo na superfície celular, a planta ativa a imunidade desencadeada por PAMP, denominada PTI (*PAMP-Triggered Immunity*), a qual, normalmente, inibe a infecção antes de o patógeno ganhar a célula. Este tipo de resposta, que também pode ser desencadeada por um reconhecimento fraco de proteínas efetoras, é conhecida como defesa basal e resulta na ativação da expressão de genes de defesa ou na indução tardia da morte celular (ABRAMOVITCH; MARTIN, 2004; CHISHOLM et al., 2006).

Uma vez que os patógenos adquiriram a capacidade de suprimir as defesas primárias, através da interferência no sistema de reconhecimento da membrana plasmática vegetal ou pela secreção de proteínas efetoras no citoplasma da célula hospedeira, as plantas desenvolveram um mecanismo mais especializado para detectá-los, o qual leva à indução de um segundo nível de resposta de defesa. Este mecanismo, denominado imunidade desencadeada por efector ou ETI (*Effector-*

Triggered Immunity), envolve o reconhecimento direto ou indireto das proteínas efetoras através de proteínas vegetais de resistência (proteínas R). Estas respostas são qualitativamente similares às respostas basais; no entanto, há um forte reconhecimento das proteínas efetoras e uma rápida resposta que suprime a reprodução do patógeno. Contudo, se a planta hospedeira não possuir o gene R capaz de reconhecer a proteína efetora do invasor, o patógeno é capaz de causar a doença (ABRAMOVITCH; MARTIN, 2004; CHISHOLM et al., 2006; TAMELING; TAKKEN, 2008).

Após o reconhecimento do patógeno pelas proteínas R, uma cascata de sinalização, associada com proteínas MAPK (*Mitogen Activated Protein Kinases*), induz a transcrição de genes de defesa, promovendo uma série de eventos conhecidos como resposta de hipersensibilidade (HR), os quais caracterizam o mecanismo de defesa local que auxilia no confinamento do patógeno e na prevenção da sua proliferação. Segundo Dmitriev (2003), Veronese et al. (2003) e Nürnberger et al. (2004), dentre tais eventos estão:

- (a) a deposição de calose para o reforço da parede celular;
- (b) a explosão oxidativa pela geração de espécies reativas de oxigênio (ERO) e óxido nítrico (NO), que provocam a morte das células infectadas;
- (c) a produção de fitoalexinas que apresentam atividade antimicrobiana;
- (d) a produção de moléculas reguladoras das respostas de defesa, como ácido salicílico, ácido jasmônico e etileno;
- (e) a síntese de proteínas relacionadas à patogênese (proteínas PR), possivelmente associadas à resistência não-específica do hospedeiro ao patógeno.

Assim, sem o reconhecimento específico do patógeno ou sem a ativação eficiente de tais respostas de defesa, a colonização pelo patógeno pode evoluir para níveis que dificilmente poderão ser vencidos pela planta. Todavia, se a planta ativa seus mecanismos de defesa locais, estes sinais de respostas são amplificados para as células vizinhas, levando a uma resposta imune sistêmica. Esta resposta sistêmica protege a planta contra subseqüentes ataques do mesmo patógeno, pois as proteínas PRs são produzidas e mantidas no tecido vegetal, atuando como uma “memória imunológica” (CURSINO-SANTOS; DEFINA; MARTINEZ-ROSSI, 2003; DMITRIEV, 2003).

As plantas são dotadas de diversos tipos de sistemas de defesa que lhes conferem resistência a patógenos. Vários desses sistemas não funcionam separadamente, mas, interagem, intensificando e suplementando um ao outro. Dentre os vários tipos de classificações usados para diferenciar as respostas de defesa, um deles classifica a resistência de plantas a patógenos em duas categorias: resistência qualitativa e resistência quantitativa (DMITRIEV, 2003, KOU; WANG, 2010). A resistência qualitativa é mediada por um ou poucos genes de resistência (R), que permite o reconhecimento de raças distintas de patógenos e desencadeiam reações de defesa no sítio da infecção (EULGEM, 2005). A resistência quantitativa ou poligênica permite à planta utilizar-se da combinação de mecanismos de defesa pré-existentes ou induzidos para defender-se parcial ou quase completamente dos seus patógenos. Por isso, quando as condições para a infecção e o desenvolvimento de uma doença são favoráveis, o genótipo da planta e o do patógeno serão determinantes para a ocorrência (em estágio moderado ou grave) ou não da doença (AGRIOS, 2005).

2.3 Moléculas envolvidas nas respostas de defesa vegetal

A ativação do sistema de defesa de plantas se dá quando receptores específicos reconhecem a penetração do patógeno e transmitem esta informação ao interior da célula. Esta transdução de sinais intracelular compreende uma complexa e bem coordenada rede de sinalização, que controla as respostas de defesa da planta ao ataque de patógenos e envolve diversos tipos de moléculas (BOLLER; HE, 2009; SPOEL; DONG, 2012).

Quinases pertencem a uma superfamília de proteínas com centenas a milhares de cópias e são componentes de, basicamente, todas as funções celulares. Estas proteínas usam o γ -fosfato do ATP para fosforilar resíduos de serina e treonina (serina/treonina quinases - STK) ou tirosina (receptor tirosina quinase - RTK) de outras proteínas (LEHTI-SHIU; SHIU, 2012). Nas respostas de defesa, as proteínas quinases são ativadas nas células vegetais ao redor do sítio de invasão do patógeno e regulam a fosforilação e a atividade de outras proteínas (DMITRIEV, 2003).

Os receptores quinases de plantas (RLKs ou RPKs) são os principais receptores de sinais extracelulares e atuam em variadas vias de sinalização, incluindo as respostas a sinais microbianos durante simbioses e defesas (ANTOLÍN-LLOVERA et al., 2012). Os RLKs contêm um domínio serina/treonina quinase C-terminal intracelular ou um versátil domínio N-terminal extracelular, apresentando uma significativa diversidade estrutural, sendo que a maioria dos RLKs possuem repetições ricas em leucina (LRR-RLKs) (AFZAL; WOOD; LIGHTFOOT, 2008). Também, existem os receptores quinases citoplasmáticos (RLCKs), os quais não possuem o domínio extracelular e são específicos de plantas (JURCA; BOTTKA; FEHÉR, 2007).

Na defesa de plantas, os RLKs atuam como receptores de reconhecimento de padrões (PRRs), os quais são localizados na membrana plasmática e percebem diversos MAMP (padrões moleculares associados a microrganismos) e PAMP, bem como podem reconhecer peptídeos vegetais secundários induzidos por perigo e sinais de parede celular (TENA; BOUDSOCQ; SHEEN, 2011; ANTOLÍN-LLOVERA et al., 2012).

Os RLKs desencadeiam uma rápida e convergente rede de sinalização controlada pelas proteínas quinases ativadas por cálcio e pelas cascatas de sinalização das MAPK (*Mitogen activated protein kinases*). Esta rede de sinalização das proteínas quinases tem funções específicas e sobrepostas, controlando a atividade e a síntese de uma infinidade de fatores de transcrição, enzimas, hormônios e peptídeos e químicos antimicrobianos, contribuindo assim para a resistência contra bactérias, fungos e oomicetos (TENA; BOUDSOCQ; SHEEN, 2011).

A cascata de sinalização das MAPK (*Mitogen activated protein kinases*) é uma via de sinalização conservada nos eucariotos e atua em diversos processos como crescimento e diferenciação celular, regulação do ciclo celular, respostas a estresses e apoptose. Neste contexto, estudos já mostraram que a cascata de MAPKs em plantas induz a expressão de genes de defesa e a morte celular, em resposta ao reconhecimento de patógenos. Esta cascata é o principal componente a jusante de receptores de membrana, transduzindo os estímulos extracelulares em respostas intracelulares. É composta por três módulos de quinases, quais sejam MAPKKK-MAPKK-MAPK, e para cada um desses módulos existem vários membros

de quinases, que contribuem para a especificidade do sinal transmitido (WRZACZEK; HIRT, 2001; ZHANG; KLESSIG, 2001).

Um componente importante que atua como segundo mensageiro universal nas cascatas de transdução de sinais é o cálcio. Nas plantas, seu nível intracelular é modulado em resposta a vários estímulos, como hormônios, luz, danos mecânicos, estresses abióticos e elicitores de patógenos, sendo considerado um mediador primário essencial e conservado durante as respostas de defesa (CHENG, 2002; BOUDSOCQ et al., 2010).

Diversas classes de proteínas sensores se ligam ao Ca^{2+} , dentre elas as proteínas quinases dependentes de cálcio (CDPKs) e as calmodulinas. As CDPKs são proteínas quinases Ca^{2+} sensores essenciais na sinalização para a reprogramação transcricional durante a imunidade inata de plantas (CHENG, 2002; BOUDSOCQ et al., 2010). Boudsocq et al. (2010) propõem que as CDPKs sejam o ponto de convergência da sinalização desencadeada por vários MAMP, uma vez que observaram que CDPKs e cascatas de MAPK atuam diretamente em quatro ciclos de regulação mediado por MAMP para controlar genes envolvidos na síntese de peptídeos e metabólitos de defesa, modificação da parede celular e sinalização de redox.

Por sua vez, as calmodulinas (CaM e CMLs) constituem uma grande família de proteínas sensores ligadas a Ca^{2+} , as quais transduzem o sinal do aumento do Ca^{2+} citosólico à proteínas alvos em diversas cascatas de transdução de sinais, incluindo a sinalização durante interação planta-patógeno e a resposta imune inata. Neste caso, o sinal do fluxo de Ca^{2+} no citosol, pelo reconhecimento inicial do patógeno, ativa a CaM e/ou a CML, as quais induzem a síntese de ácido nítrico (NO) como um passo intermediário na cascata de sinalização durante a percepção do patógeno, levando a respostas de defesa, incluindo a resposta de hipersensibilidade (HR) (MA et al., 2008).

Já com relação às proteínas com atividade ATPase, as enzimas H^+ -ATPases tem a função de estabelecer o gradiente eletroquímico da membrana plasmática através da hidrólise de ATP e do transporte de prótons para fora da célula. Dessa forma, controlam os principais processos de transporte na planta, como a absorção de nutrientes pela raiz e o carregamento de xilema ou floema, bem como a abertura estomatal, o alongamento das células e a regulação do pH celular. Porém, o

envolvimento dessas enzimas em diferentes processos fisiológicos sugere a existência de múltiplas funções reguladoras que integram algumas vezes sinais contraditórios a partir da célula ou do meio ambiente. Também, a atividade das H⁺-ATPases pode ser positiva ou negativamente influenciada pela sua fosforilação e localização (DUBY; BOUTRY, 2009).

Durante a resposta de defesa vegetal, a atividade das H⁺-ATPases de membrana plasmática são reguladas de forma dinâmica, provavelmente contribuindo com a despolarização da membrana, induzida por PAMP, e com a alcalinização do apoplasto em conjunto com o efluxo de ânions e o influxo de prótons. Em contrapartida, pode ser necessária a inibição dessas enzimas durante alguns tipos de respostas de defesa. Assim, durante o reconhecimento inicial de patógenos há uma modulação espacial e temporal complexa da sua atividade. Também, a atuação da H⁺-ATPase está relacionada com a proteína de sinalização RIN4 para regular a abertura estomatal durante a penetração bacteriana no tecido foliar. Por outro lado, patógenos possuem mecanismos para manipular a atividade dessas enzimas durante sua infecção (ELMORE; COAKER, 2011).

Ainda nesse contexto, as proteínas da superfamília ABC transportadoras (*ATP-binding cassette*), as quais são encontradas em todos os seres vivos, também atuam como bombas ativas primárias intrínsecas de membranas. Estas proteínas são encontradas na membrana plasmática e tonoplasto, em maior proporção, como também nas membranas mitocondriais, plastidiais e peroxomais, e transportam diversos tipos de substratos como lipídios, fitormônios, carboxilatos, metais pesados, metabólitos secundários, toxinas, dentre outros. Além de atuarem na detoxificação, as ABC transportadoras participam de diversos processos fisiológicos, o que permite à planta se adaptar a mudanças ambientais e lidar com estresses bióticos e abióticos (VERRIER et al., 2008; CROUZET et al., 2013).

Quanto à regulação da síntese protéica nos vegetais, esta frequentemente é feita por diferentes fatores de transcrição. Por exemplo, membros das famílias de fatores de transcrição de *Arabidopsis*, TGA-bZIP, ERF, Myb, Whirly e WRKY, estão envolvidos na regulação da expressão de genes de defesa ao se ligarem aos promotores desses genes. Esta regulação pode se dar por ativação ou da repressão da transcrição gênica. Os elementos alvos desses fatores, e seus derivados, intermediam a expressão de genes ao responder a vários estímulos relacionados à

defesa, como à infecção de patógeno, a elicitores e a moléculas mensageiras (SA, ET e AJ) (EULGEM, 2005).

Na família de fatores de transcrição WRKY, seus membros possuem ao menos uma região conservada de ligação ao DNA, o domínio WRKY, que contém a sequência peptídica conservada 'WRKYGQK' e um motivo dedo de zinco. Este domínio geralmente se liga ao elemento W box (C/TTGACC/T), o qual é altamente representado nos promotores de genes de resposta de defesa de *Arabidopsis*. Também, os elementos W Box frequentemente estão presentes nos promotores dos próprios genes *WRKY*, o que mostra que estes genes estão sobre controle direto, positivo ou negativo, dos próprios fatores WRKY, via um mecanismo específico de *feedback*. Assim, esses fatores de transcrição constituem uma rede de regulação complexa e altamente interconectada durante as respostas de defesa vegetal (LIPPOK et al., 2007; PANDEY; SOMSSICH, 2009).

Os fatores de transcrição ERF pertencem a uma subfamília da família APETALA2 (AP2)/EREBP (*Ethylene-Responsive-Element-Binding Protein*) exclusiva de plantas. Estas proteínas apresentam o domínio conservado ERF que se ligam aos *cis*-elementos GCC Box, encontrados em muitos promotores de genes das proteínas PR (SINGH; FOLEY; OÑATE-SÁNCHEZ, 2002). O fator de transcrição ERF1 de *Arabidopsis* é responsável pelo aumento da resistência da planta a vários tipos de patógenos. Sua transcrição é ativada rapidamente pelas vias de sinalização do etileno ou jasmonato, ou pode ser ativado sinergicamente por ambos os hormônios, regulando, assim, a expressão de genes de resposta ao patógeno, prevenindo a propagação da doença (LORENZO et al., 2002).

Também, diversas outras proteínas estão envolvidas na regulação de genes de defesa vegetal. A enzima glutamato-cisteína ligase (GCL), por exemplo, catalisa o primeiro passo da via de biossíntese da glutathione e atua na regulação do ambiente redox intracelular. Ela catalisa a formação da γ -glutamilcisteína dependente de ATP a partir de glutamato e cisteína (JEZ; CAHOON; CHEN, 2004). O metabolismo da glutathione pode determinar o nível de expressão de genes de defesa controlados por diversas vias de sinalização, antes e durante o estresse (BALL et al., 2004).

Em contrapartida, as proteínas MLO pertencem a uma classe de proteínas transmembrana e funcionam como um regulador negativo de resistência a doenças e morte celular. Diversos estudos em cevada tem mostrado a participação do gene

Mlo na modulação da resistência a fungo, sendo esta baseada em múltiplos componentes e mecanismos de defesa. Genótipos de cevada (homozigotos recessivos) para esse gene são altamente resistentes ao patógeno, apresentando resposta de defesa e morte celular espontâneas e deposição de parede celular; genótipos com o alelo dominante remetem à suscetibilidade ao fungo, pois se tornam ineficientes para desencadear a cascata de eventos que resultam na morte celular (DEVOTO et al., 2003; ZIEROLD; SCHOLZ; SCHWEIZER, 2005).

Além disso, durante o estresse biótico, especialmente nas interações incompatíveis, a rápida produção e o acúmulo de espécies reativas de oxigênio (ERO) são um dos primeiros sinais após o reconhecimento do patógeno, desempenhando diversas e importantes funções. As ERO, principalmente o H_2O_2 , possuem ação antimicrobiana direta, catalisam o reforço mecânico da parede celular, participam como segundo mensageiro na via de sinalização da superóxido sintase e no desencadeamento da resposta de hipersensibilidade (HR), bem como, são necessárias para a ativação de genes de resistência e para a síntese de componentes antimicrobianos (DMITRIEV, 2003).

Em plantas, as ERO, como o radical superóxido ($O_2^{\bullet-}$), o peróxido de hidrogênio (H_2O_2) e o radical hidroxila ($OH^{\bullet}$) são continuamente produzidas como subprodutos de várias vias metabólicas aeróbicas, localizadas principalmente nos cloroplastos, mitocôndrias e peroxissomos. O rápido aumento das ERO, conhecido como explosão oxidativa, pode ocorrer por diversos mecanismos. Porém, essas moléculas são altamente tóxicas e seu acúmulo por um longo período pode causar danos em lipídeos, proteínas e no DNA (RESENDE; SALGADO; CHAVES, 2003; MITTLER et al., 2004; SOARES; MACHADO, 2007).

No entanto, a maquinaria celular vegetal apresenta um alto grau de controle da toxicidade dessas moléculas, utilizando-as, então, como importantes moléculas sinalizadoras. Para a proteção contra esses agentes oxidantes, as plantas exibem várias respostas bioquímicas, que incluem a ação de mecanismos enzimáticos, como aumento na atividade da superóxido dismutase (SOD), da catalase (CAT), da glutathione peroxidase (GPX), das peroxidases (POX), juntamente com a atuação de moléculas antioxidantes como o ácido ascórbico e a glutathione (RESENDE; SALGADO; CHAVES, 2003; MITTLER et al., 2004; SOARES; MACHADO, 2007).

Estas enzimas são oxidoredutases e algumas delas, como as peroxidases, são consideradas proteínas PR.

O desencadeamento da HR, durante interações incompatíveis, é ativado por ERO com importante participação do óxido nítrico (NO), molécula a qual tem demonstrado ser um componente essencial nas respostas imunes vegetais (BELLIN et al., 2013). A HR ocorre em quase todos os tipos de respostas de defesa e é acompanhada por um tipo complexo de morte celular programada, ao redor do sítio de infecção, e relacionado com a indução de respostas de defesa local e sistêmica. A morte celular associada a HR restringe o crescimento do patógeno, e outros fatores, como a síntese proteica e o ácido salicílico, se mostram necessários para a sua indução (HEATH, 2000; LAM; KATO; LAWTON, 2001).

Quanto às proteínas vegetais atuantes nas respostas a patógenos, as proteínas de defesa são características do sistema imune de diversos grupos de organismos. As plantas tem a capacidade de acumular essas proteínas e peptídeos como um dos recursos para revidar a invasão de um organismo estranho. Assim, genes que codificam proteínas e peptídeos antimicrobianos têm importante função nas interações planta-patógeno e a indução da transcrição desses genes está relacionada com o fenômeno de indução de resistência. Sabe-se que muitas dessas proteínas ou peptídeos antimicrobianos são tóxicos para alguns microrganismos e inofensivos para outros, porém, ainda não se sabe ao certo a função dessas moléculas especificamente em cada interação planta-patógeno (VERONESE et al., 2003).

Já as proteínas de resistência, conhecidas por proteínas R, atuam em outro nível de defesa, geralmente relacionadas com a resistência qualitativa. Neste caso, o patógeno consegue subverter o reconhecimento das PAMPs pela planta, através de seus receptores de membrana, e implantar moléculas efetoras (fatores de avirulência Avr) no interior da célula hospedeira (JONES; DANGL, 2006; KOU; WANG, 2010).

As proteínas R são capazes de reconhecer raças distintas de patógenos biotróficos e desencadear reações de defesa no sítio da infecção. Assim, proteínas R são capazes de fazer um forte reconhecimento dos fatores de avirulência Avr do patógeno, desencadeando uma rápida resposta de hipersensibilidade, com morte celular programada, limitando o estabelecimento da infecção do patógeno. Neste

tipo de resistência, quando existe uma proteína R que reconhece a proteína de avirulência do patógeno invasor, a interação é incompatível, sendo a planta resistente e o patógeno avirulento. Quando a proteína R não é capaz de reconhecer a proteína de avirulência, a interação é compatível, pois o patógeno é virulento e a planta suscetível (ABRAMOVITCH; MARTIN, 2004; EULGEM, 2005).

Por fim, as proteínas relacionadas à patogênese, conhecidas como proteínas PR, formam um diverso grupo de proteínas presentes em pequenas quantidades, porém, são altamente sintetizadas em resposta ao ataque de patógenos ou a estresses, especialmente durante a HR. São classificadas de acordo com suas funções e propriedades, se apresentam extremamente ácidas ou extremamente básicas e, por isso, são altamente solúveis e reativas. Dentre as suas diversas funções, algumas proteínas PR induzem a síntese de etileno, outras induzem a produção de enzimas oxidativas e proteínas envolvidas na modificação e reforço da parede celular; também, podem sintetizar compostos antimicrobianos, tais como as fitoalexinas, ou destruir os componentes da parede celular do patógeno, ou serem inibidoras de proteases, que neutralizam enzimas específicas do patógeno (AGRIOS, 2005).

Dentre as 14 famílias de proteínas PR, as mais conhecidas são as PR1 (antioomicetos e antifúngicas), as PR2 (β -1,3-glucanases), as PR3 (quitinases), as PR4 (antifúngicas) e PR6 (inibidoras de proteases), dentre outras proteínas PR como defensinas, tioninas, lisozimas, lipoxigenases, proteinases, quitosanases e peroxidases. Por exemplo, as β -1,3-glucanases e as quitinases quebram a estrutura de quitina da parede celular de muitos fungos fitopatogênicos; já lipoxigenases e peroxidases de lipídios produzem metabólitos antimicrobianos e moléculas sinalizadoras secundárias como o ácido jasmônico (AGRIOS, 2005).

2.4 Identificação de genes de defesa contra patógenos

O funcionamento de um sistema biológico depende, quase sempre, da atuação de uma ou mais proteínas e de suas interações com outros tipos de moléculas ou estruturas. Estas interações, por sua vez, dependem do reconhecimento de estruturas específicas contidas nas moléculas proteicas. Sendo as proteínas o

produto da expressão dos genes, a análise de transcritomas é uma abordagem importante para se obter informações sobre as possíveis proteínas produzidas por um organismo, na condição estudada, permitindo, também, a caracterização de centenas de genes na mesma análise (BINNECK, 2004).

O entendimento sobre a base molecular da resistência de plantas a doenças tem sido o foco de inúmeras pesquisas. A identificação das defesas de plantas suscetíveis ao ataque de patógenos tem proporcionado novos insights sobre processos essenciais para a patogênese bacteriana e para a imunidade da planta, mostrando um grande potencial para a melhoria da resistência a doenças (ABRAMOVITCH; MARTIN, 2004).

Para isso, é necessário o isolamento e a identificação dos genes envolvidos na resposta de plantas à infecção por patógenos e no consequente desenvolvimento de doença. Assim, no intuito de identificar genes relacionados à defesa em diversas interações planta-patógeno, tem sido construídas bibliotecas de cDNA por seleção diferencial, a partir de mRNAs isolados de plantas inoculadas, tratadas com elicitores ou indutores de resistência ou a partir de cultura de células.

Uma metodologia alternativa para o isolamento de genes diferencialmente expressos, denominada hibridização subtrativa por supressão (SSH), desenvolvida por Diathenko et al. (1996), se mostra adequada e eficiente para a detecção de genes de defesa, pois permite selecionar apenas os genes do hospedeiro expressos no momento do reconhecimento do patógeno, e não requer conhecimento prévio das sequências do genoma em estudo. Esta metodologia é baseada na amplificação preferencial, por PCR, de sequências diferencialmente representadas em duas populações de cDNA, enquanto a supressão impede a amplificação de sequências comuns. Os cDNAs obtidos são clonados, gerando uma mini-biblioteca enriquecida de genes diferencialmente expressos, fornecendo um amplo material para a comparação de populações de mRNAs (DIATCHENKO et al., 1996).

Além disso, o método SSH tem a vantagem de apresentar baixa taxa de falsos positivos e usar procedimentos menos complexos; também, é eficaz para normalizar a representação da população de cDNAs diferencialmente expressos, pois ao controlar o tempo da hibridização, as fitas simples de cDNAs altamente abundantes podem ser reduzidas ao mesmo nível daquelas menos abundantes. Assim, se mostra vantajoso em relação a outras técnicas desenvolvidas para a

identificação de genes diferencialmente expressos, como, por exemplo, *Representational Difference Analysis* (RDA), *Differential Display* (DD) e *Serial Analysis of Gene Expression* (SAGE), que identificam uma pequena quantidade de genes, excluindo da análise os genes pouco representados (STEIN; THIES; HOFMANN, 1997; MAHALINGAM et al., 2003).

Para validar a expressão diferencial dos genes obtidos em bibliotecas subtrativas, a transcrição reversa (RT) aliada à PCR em Tempo Real (qPCR) tem sido bastante empregada. A RT-qPCR é uma metodologia que permite a quantificação dos níveis de transcrição gênica, baseando-se na amplificação de moléculas de DNA/cDNA e na obtenção dos sinais de fluorescência produzidos durante cada ciclo de amplificação, ao usar corantes fluorescentes intercalantes de DNA de dupla fita, como o SYBR Green. Esta técnica se mostra eficiente devido a sua rapidez de execução, sendo extremamente sensível, pois é capaz de detectar pequenas quantidades de produto gênico; também, é precisa e reprodutível, além de apresentar, relativamente, baixo custo (UDVARDI; CZECHOWSKI; SCHEIBLE, 2008; BUSTIN et al., 2009).

Exemplificando, durante a interação compatível *Oryza meyeriana* Baill-*Xanthomonas oryzae* pv. *oryzae*, foi construída uma biblioteca subtrativa pelo método SSH para identificar genes responsáveis pela resistência do arroz à uma raça altamente virulenta do patógeno. Neste trabalho, foram aleatoriamente escolhidas e analisadas 288 sequências e, após a busca por sequências homólogas no GeneBank (<http://www.ncbi.nlm.nih.gov/>), as supostas funções dos genes foram agrupadas em cinco categorias: estresse e defesa, transdução de sinais, fotossíntese e fotorrespiração, transporte de materiais e metabolismo. Os resultados indicaram a ocorrência de uma série de processos moleculares após a infecção da planta e os dados gerados devem ajudar novas pesquisas a estabelecer as relações entre os genes diferencialmente expressos e suas funções biológicas (CHENG et al., 2009).

Para a identificação de genes relacionados à resistência da videira ao fungo *Elsinoe ampelina*, causador da antracnose, foi construída uma biblioteca subtrativa pelo método SSH usando um genótipo resistente ao patógeno. As 461 unisequências obtidas foram analisadas no banco de dados do Gene Ontology (<http://www.geneontology.org/>), sendo 310 delas atribuídas a 22 termos de função

molecular e 317 atribuídas a 43 termos de processos biológicos. Vinte genes atribuídos a diferentes categorias funcionais foram selecionados para a validação da expressão diferencial por RT-qPCR e também analisados em um genótipo de videira suscetível à antracnose. Os perfis de expressão desses genes, avaliados em ambos os genótipos às 0, 3, 6, 12, 24, 48, 72 e 120 h.a.i., mostraram que a maioria dos genes teve expressão rápida e fortemente induzida em resposta ao patógeno no genótipo resistente, enquanto, no genótipo suscetível, a indução da expressão desses genes foi mais lenta e limitada. Os resultados deste estudo produziram informações valiosas para se analisar a importância dos vários mecanismos de resistência à antracnose, regulados pelos produtos destes genes, tanto em genótipos resistentes como em susceptíveis (GAO et al., 2012).

Finalizando, o estudo de transcritomas, durante a interação planta-patógeno, tem trazido importantes avanços na busca de genes de resistência a doenças em culturas vegetais, pois, apesar dos progressos na caracterização molecular dos interatomas, ainda se conhece muito pouco sobre a co-evolução patógeno-hospedeiro. Estes estudos possibilitam a identificação (e o uso) de genes de resistência, especialmente ao combinar vários receptores capazes de reconhecer efetores altamente virulentos, levando a uma resistência mais durável. Também, conhecendo uma grande quantidade de proteínas efetoras, podem ser buscadas novas fontes de resistência em plantas silvestres (DODDS; RATHJEN, 2010).

2.5 Ferramentas de Bioinformática para a anotação de sequências gênicas

A atribuição de funções a sequências putativas de genes se tornou possível a partir da busca de suas similaridades com sequências previamente depositadas em banco de dados, utilizando-se, para tal, ferramentas de bioinformática. Os bancos de dados públicos, que armazenam grande quantidade de dados de sequências de nucleotídeos e de proteínas, são o GenBank, do *National Center for Biotechnology Information* (NCBI) (<http://www.ncbi.nlm.nih.gov/>), o DDBJ (*DNA Data Bank of Japan*) (<http://www.ddbj.nig.ac.jp>) e o EMBL-EBI (*European Molecular Biology Laboratory - European Bioinformatics Institute*) (<http://www.ebi.ac.uk>), os quais, juntos, formam o INSDC, *International Nucleotide Sequence Database collaboration*.

Estes bancos trabalham em cooperação para estabelecer formatos de dados, metadados e protocolos que facilitem a submissão segura dos dados aos bancos e proporcionar o intercâmbio contínuo de dados em todo o mundo, fazendo do INSDC uma das alianças globais de arquivos de dados biológicos mais duradouras (NAKAMURA et al., 2012).

Também, atualmente, existem centenas de bancos de dados de biologia molecular *on line*, os quais disponibilizam informações específicas para determinadas espécies ou grupos de organismos, de sequências nucleotídicas ou proteicas, e de grupos de moléculas ou de elementos gênicos ou genômicos (FERNANDEZ-SUAREZ; GALPERIN, 2012). Por exemplo, o UNIPROT (*Universal Protein Resource*) (<http://www.uniprot.org>) é um banco de dados que integra, interpreta e padroniza dados de inúmeras fontes para formar o maior e mais completo catálogo de sequências proteicas e de anotação funcional. Assim, o UNIPROT disponibiliza uma base de conhecimento de sequências proteicas com livre acesso, estável e abrangente, sendo essas sequências totalmente classificadas e ricamente anotadas e acuradas (THE UNIPROT CONSORTIUM, 2012).

Já a Plataforma PLAZA 2.5 (<http://bioinformatics.psb.ugent.be/plaza/>) é um banco de dados *on line* com diversas ferramentas e visualizações interativas, de fácil uso, que permite explorar informações genômicas de 25 espécies vegetais, as quais cobrem uma ampla variedade taxonômica. Os dados disponíveis consistem em anotações estruturais e funcionais de genes, famílias de genes homólogos, múltiplos alinhamentos de sequência, árvores filogenéticas e regiões colineares dentro e entre espécies (VAN BEL et al., 2011).

A ferramenta mais conhecida para a busca de sequências similares nos bancos de dados é a BLAST (*Basic Local Alignment Search tool*) (ALTSCHUL et al. 1990), disponível no GeneBank. Os resultados fornecidos por essa ferramenta, a partir de uma sequência alvo, apresentam pontuações (*scores*) e a significância estatística dessas pontuações, em valores conhecidos como *e-values*. Isto que permite separar os alinhamentos biologicamente significativos dos resultados espúrios. Assim, um alinhamento é mais significativo quanto menor for seu *e-value* (Veja SETUBAL, 2004).

Diferentes tipos de busca de similaridades são oferecidos pela ferramenta BLAST, como, por exemplo, o BlastN, que é usado para pesquisar sequências de

nucleotídeos similares a uma sequência alvo, o BlastX, que traduz uma sequência nucleotídica alvo em produtos proteicos, em todos os possíveis quadros de leitura, e as compara com as proteínas do banco de dados e o BlastP, que pesquisa sequências de proteínas similares a uma sequência proteica alvo.

A anotação de dados de sequências de proteínas, contendo a localização e o tamanho de domínios conservados, tornou-se uma ferramenta indispensável na análise comparativa de genes e genomas. Assim, é possível verificar se a sequência analisada possui domínios conservados através do *Conserved Domain Database* (CDD), que é uma combinação de sequências múltiplas alinhadas, representando domínios proteicos conservados durante a evolução molecular. Esta ferramenta, que faz parte do sistema de recuperação de dados (Entrez) do NCBI, utiliza informações geradas pelo próprio NCBI e de fontes externas, como Pfam, SMART, COG, PRK e TIGRFAM (<http://www.ncbi.nlm.nih.gov/cdd>; MARCHLER-BAUER et al., 2002; MARCHLER-BAUER et al., 2003).

Outro banco de dados importante para a anotação de sequências gênicas é o *Gene Ontology Consortium* (GO), que visa uniformizar a representação de genes. O GO fornece um vocabulário de termos padronizados para descrição e anotação de produtos gênicos e oferece ferramentas para acesso e processamento desses dados. Durante a anotação de uma sequência alvo, é possível obter informações sobre a atividade e a localização do produto do gene putativo com os termos GO, e obter referências e indicações dos tipos de evidências disponíveis que suportam tais anotações. As ontologias estão organizadas em função molecular, processo biológico e componente celular (<http://www.geneontology.org>).

O Blast2GO é um programa computacional disponível gratuitamente via Java Web Start (<http://www.blast2go.de>), que permite a anotação automática e de alto rendimento de sequências, especialmente para espécies não modelo, e a atribuição de funcionalidade aos dados anotados. Para isso, o programa usa a ferramenta BLAST para encontrar sequências homólogas às sequências de entrada, em bancos de dados públicos ou em bancos personalizados; aos resultados do BLAST são mapeados os GO termos, através dos arquivos de anotação fornecidos pelo *GO Consortium*; as sequências são então anotadas através da seleção dos GO termos que atingem uma pontuação mínima, a qual leva em consideração vários critérios, como a similaridade entre as sequências do alinhamento, o código de evidência da

anotação e a relação entre os GO termos. O programa também possibilita uma análise estatística das informações sobre a função do gene (CONESA et al., 2005).

Contextualizando, não existem variedades comerciais de maracujá azedo resistentes à *Xap* e poucos são os estudos sobre este patossistema. Também, a literatura científica não dispõe de informações suficientes sobre os genes envolvidos no processo de defesa à bacteriose. Dessa forma, o presente trabalho se propôs: a identificar genes de maracujá azedo diferencialmente expressos durante a interação com *Xap*, construindo, para isso, bibliotecas subtrativas usando a metodologia SSH; caracterizar esses genes através de comparações em bancos de dados públicos com auxílio de ferramentas de bioinformática; quantificar a expressão de determinados genes usando a metodologia RT-qPCR.

3 MATERIAL E MÉTODOS

3.1 Material biológico

Neste trabalho, foram utilizados o acesso de maracujá azedo 'IAPAR123', moderadamente resistente à *Xanthomonas axonopodis* pv. *passiflorae* (Xap), e o isolado bacteriano AP302, da coleção de isolados bacterianos estudados por Munhoz et al. (2011), que apresentou o maior grau de agressividade contra o genótipo 'IAPAR123' dentre alguns isolados testados (dados não publicados). Cabe mencionar que, a partir desse mesmo acesso de maracujá azedo, nosso grupo construiu uma biblioteca genômica inserida em BACs, que está depositada no INRA/Toulouse (<http://cnrgv.toulouse.inra.fr/Library/Passiflora>).

3.2 Obtenção e inoculação das mudas

As mudas de maracujá azedo foram obtidas por propagação vegetativa, através de estaquia. Foram produzidas estacas de 15 cm contendo uma gema, as quais foram plantadas em tubetes contendo substrato organomineral à base de vermiculita. Após 30 dias, foi realizada uma adubação com NPK (14-14-14) e após 40 dias as mudas foram transplantadas para vasos plásticos (1 L) contendo o mesmo substrato. Foram realizadas adubações mensais com 1 g sulfato de amônio (21 % de N). As mudas foram mantidas em casa de vegetação, telada.

Para o preparo do inóculo, o isolado AP302 (preservado em glicerol 60% a -80°C) foi repicado em meio nutriente líquido (NL) e incubado durante 24 h, sob agitação, a 28°C. Após o crescimento, 100 µL da suspensão bacteriana foram plaqueados em meio nutriente ágar (NA), fazendo-se o espalhamento com alça de Drigalsky; as placas foram incubadas durante 24 h a 28°C. Após o crescimento da cultura, foram adicionados 20 mL de solução salina (NaCl 0,85 %) autoclavada para suspender as colônias na solução. Depois de homogeneizada, a suspensão bacteriana foi diluída e padronizada em espectrofotômetro para uma densidade ótica de 0,3 e comprimento de onda de 600 nm ($OD_{600}=0,3$), o que corresponde a uma concentração aproximada de 100 milhões de unidades formadoras de colônia por mililitro (10^8 UFC/mL).

Para a inoculação de *X. axonopodis* pv. *passiflorae* nas folhas de maracujá azedo, foi recortado um disco de cada folha, usando-se um perfurador de papel, e pelo qual foi passada uma haste com ponta de algodão embebida na suspensão bacteriana, permitindo o contato da bactéria com o tecido foliar (Figura 1). Nas plantas controles fez-se o mesmo procedimento, porém, usando apenas a solução salina (NaCl 0,85 %).

Figura 1 – Procedimento de inoculação das folhas de maracujá azedo com *Xap* ou solução salina (controle)

3.3 Condução dos experimentos de inoculação

Foram realizados dois experimentos de inoculação, sendo um para a obtenção de transcritos para a construção das bibliotecas SSH e outro para a obtenção de transcritos para a validação das bibliotecas por qPCR. O primeiro experimento foi conduzido no mês de Dezembro de 2010, enquanto o segundo foi realizado em Outubro de 2012 (Figura 2).

Figura 2 – Condições das mudas de maracujá utilizadas no primeiro (A) e segundo (B) experimentos de inoculação com *Xap* ou solução salina (controle)

No primeiro experimento foram inoculadas 18 mudas, sendo 9 delas inoculadas com *Xap* (tratamento com *Xap*) e 9 com solução salina (tratamento controle). Cada muda teve três folhas inoculadas, isto para garantir uma quantidade suficiente de tecido para a extração de RNA, uma vez que a expressão diferencial de genes durante a resposta de defesa se dá próxima ao local de inoculação.

As coletas das folhas inoculadas foram feitas às 6, 12 e 24 h após a inoculação (h.a.i.), sendo que em cada período foram coletadas as folhas de três mudas inoculadas com *Xap* e de três com solução salina.

De cada folha, foi retirado um pequeno fragmento de tecido (2 x 2 cm). Os fragmentos das três folhas inoculadas de cada muda foram misturados, garantindo uma quantidade aproximada de 200 mg de tecido ao redor dos discos de inoculação (Figuras 3 e 4). Os fragmentos foliares foram imediatamente congelados em nitrogênio líquido, macerados, aliquotados em quatro microtubos e armazenados em ultra freezer a -80 °C até a extração do RNA.

No segundo experimento procedeu-se da mesma forma, porém, com as seguintes variações: foram inoculadas duas folhas de cada muda e as coletas das folhas foram realizadas às 6 horas e aos 5 e 9 dias após a inoculação (d.a.i.), conforme esquematizado na figura 5.

Figura 3 – Processo de coleta das folhas e retirada do fragmento ao redor do disco onde realizou-se a inoculação

Figura 4 – Esquema do primeiro experimento de inoculação com *Xap* ou solução salina (controle)

Figura 5 – Esquema do segundo experimento de inoculação com *Xap* ou solução salina (controle)

Durante o período de condução dos experimentos, a temperatura média e a umidade relativa média na casa de vegetação foram iguais a 29°C $\pm$ 6 e 60 % $\pm$ 15 durante o primeiro experimento e a 30°C $\pm$ 10 e 66 % $\pm$ 26 durante o segundo experimento. Os dados foram coletados uma vez ao dia.

Quatro mudas extras foram inoculadas com *Xap* e quatro com solução salina, concomitantemente, visando avaliar a presença da bactéria no tecido vegetal, durante os períodos de avaliação (6, 12 e 24 h.a.i. e 5 e 9 d.a.i.), bem como após 14 dias, período em que os sintomas da lesão já estão visíveis. Para isso, foram feitas PCRs com *primers* específicos para a detecção de *Xap* no tecido foliar, conforme metodologia descrita por Munhoz et al. (2011), além da avaliação visual dos sintomas. Este conjunto de *primers*, denominado Xapas, foi desenhado na região ITS 16S-23S e amplifica um fragmento de 520 pb, que é específico para *Xap* e um fragmento de 641 pb, que detecta bactérias do gênero *Xanthomonas* e serve como controle da reação de amplificação. O DNA foi extraído das folhas coletadas às 6, 12 e 24 h.a.i. e 5, 9 e 14 d.a.i., de acordo com a metodologia convencional (DOYLE; DOYLE, 1987), usando CTAB. As PCRs foram conduzidas em um volume final de 20 μ L, contendo 40 ng de DNA, 1 x de tampão de reação (Promega), 0,1 μ M dos *primers reverse* (interno e externo), 0,2 μ M do *primer forward*, 0,2 mM de cada dNTP, 2,0 mM de MgCl₂ e 1 U de Go *Taq* Flex (Promega). As condições de amplificação foram as seguintes: 94°C por 5 min., 35 ciclos de 94°C por 1 min., 60°C por 1 min. e 72°C por 1 min. Ao final da ciclagem, a reação foi submetida a uma extensão final a 72°C por 5 min. Utilizou-se o termociclador *Gen Amp PCR System* 9700 (ABI). Os amplicons foram separados por eletroforese a 3 V/cm em gel de agarose 1,2 % e visualizados e fotodocumentados sob luz UV, após coloração com SYBR Safe (0,5x) (Invitrogen).

3.4 Construção das bibliotecas subtrativas

Para a construção das bibliotecas, foram utilizados os transcritos de folhas de maracujá azedo obtidos no primeiro experimento de inoculação, ou seja, de tecidos coletados às 6, 12 e 24 h.a.i dos tratamentos de inoculação com *Xap* e com solução salina (Figura 4).

O cDNA utilizado para a construção das bibliotecas foi sintetizado utilizando-se o kit 'SMARTer™ PCR cDNA Synthesis' (Clontech), conforme os procedimentos descritos no manual do fabricante. Este kit possibilita a síntese de cDNAs de alta qualidade, necessários para a construção das mini bibliotecas de cDNA.

O kit 'SMARTer™ PCR cDNA Synthesis' utiliza um *primer* oligo(dT) modificado, denominado *primer* CDS, e o oligonucleotídeo SMARTer II para a síntese da primeira fita de cDNA. O *primer* CDS inicia a síntese do cDNA ao se ligar no início da cauda poli A do RNA mensageiro. Quando a enzima transcriptase reversa SMARTScribe™ atinge o terminal 5' do RNA mensageiro, sua atividade transferase terminal adiciona alguns nucleotídeos 'C' na extremidade 3' do cDNA. O oligonucleotídeo SMARTer II, que possui uma sequência oligo-G (3') e uma sequência adicional (5'), se liga aos nucleotídeos adicionais, o que faz com que a transcriptase reversa continue a transcrição do mRNA até a extremidade 5' do oligonucleotídeo. A síntese de cDNA de dupla fita ocorre por PCR a longa distância (LD PCR), utilizando-se um *primer* que possui uma sequência de nucleotídeos igual a uma região comum ao *primer* CDS e ao SMARTer II (Figura 6). Assim, ocorre a amplificação exponencial apenas dos cDNAs de fita simples que apresentam as sequências complementares completas dos mRNAs presentes na amostra de transcritos.

Figura 6 – Representação diagramática da síntese de cDNA segundo consta no manual do kit 'SMARTer™ PCR cDNA Synthesis'. O oligonucleotídeo SMARTer II A, o *primer* CDS e o *primer* PCR contém trechos de sequências idênticas. TR = transcriptase reversa

A construção das bibliotecas subtrativas foi feita com base na metodologia de hibridização subtrativa por supressão (SSH), descrita por Diatchenko et al. (1996), utilizando para isso o kit 'PCR-Select™ cDNA Subtraction' (Clontech), de acordo com os procedimentos descritos no manual do fabricante. A hibridização subtrativa permite a comparação entre duas populações de mRNAs, denominadas *tester* (população de mRNAs da qual se deseja obter os transcritos diferencialmente expressos) e *driver* (população de mRNA de referência), e a obtenção de transcritos

diferencialmente expressos na população *tester*, após a remoção das sequências comuns por hibridizações entre moléculas.

A metodologia de hibridização subtrativa por supressão é composta por seis passos, conforme apresentado na figura 7 (I-VI, a-d): (I) cDNAs de dupla fita são sintetizados a partir das populações de mRNAs *tester* e *driver* e digeridos, separadamente, com a enzima *Rsa*I. (II) Adaptadores Ad1 e Ad2R são ligados separadamente a duas alíquotas de cDNAs *tester*. As extremidades 5' dos adaptadores são sintetizadas sem os grupos fosfatos, permitindo que cada adaptador seja adicionado apenas às extremidades 5' das fitas de cDNA. (III) As amostras ligadas a adaptadores são misturadas a um excesso de 30 x de cDNAs *driver*. Estas misturas resultam no primeiro processo de hibridização para normalizar e enriquecer as sequências diferencialmente expressas entre moléculas *tester* de fita simples, ou seja, as concentrações de cDNA de alta e baixa abundância tornam-se aproximadamente iguais. Esta normalização ocorre porque o processo de hibridização o qual gera homohíbridos (b) e heterohíbridos (c) é mais rápido em moléculas mais abundantes. (IV) As reações da primeira hibridização são misturadas e a elas adicionam-se mais moléculas *driver* de fita simples, iniciando o segundo processo de hibridização cinética. Este processo permite a hibridização dos cDNAs menos abundantes (ligados a adaptadores) que permaneceram em fita simples na primeira etapa de hibridização (e), bem como produz outros tipos de híbridos. (V) As terminações das respectivas moléculas híbridas são preenchidas a fim de que sejam gerados fragmentos de cDNA dos genes diferencialmente expressos, que possam ser preferencialmente amplificados. (VI) Duas etapas de PCR são realizadas para amplificar preferencialmente os genes diferencialmente expressos. O efeito de supressão por PCR acontece devido às longas repetições terminais invertidas que estão presentes nos adaptadores. Quando o mesmo adaptador está ligado a ambas as extremidades de um fragmento de cDNA, as repetições invertidas formam uma estrutura estável em forma de *loop* que não pode ser amplificada exponencialmente. Isto porque a ligação intramolecular da longa repetição terminal invertida é altamente favorecida e mais estável que a ligação intermolecular do *primer* PCR curto. Assim, os fragmentos de cDNA indesejados são inibidos durante a amplificação exponencial das sequências alvo.

Figura 7 - Esquema da hibridização subtrativa por supressão, conforme apresentado no manual do usuário do kit 'PCR-Select™ cDNA Subtraction'. As linhas sólidas correspondem aos cDNAs *testers* ou *drivers*, digeridos pela enzima *RsaI*. Os retângulos pretos correspondem à porção externa dos adaptadores Ad1 e Ad2R, os quais tem sequência idêntica a do *primer* P1. Os retângulos brancos correspondem à porção interna do adaptador Ad1, o qual tem sequência idêntica à do *primer* NP1 para a *nested* PCR. Os retângulos cinzas correspondem à porção interna do adaptador Ad2R, o qual tem sequência idêntica à do *primer* NP2R para a *nested* PCR

3.4.1 Extração de RNA

Foi extraído o RNA total dos tecidos foliares, separadamente, por período de coleta (6, 12 e 24 h). Após a extração, o RNA total foi quantificado e os três produtos gerados para cada tratamento foram misturados em quantidades iguais, formando dois *bulks* de RNAs, ou seja, um *bulk* de transcritos do tratamento com *Xap* e outro *bulk* com transcritos do tratamento controle, tal como ilustrado na figura 4.

Para a extração do RNA total foi utilizado o método baseado no reagente TRIZOL (Invitrogen) conforme descrito a seguir: aos microtubos contendo aproximadamente 100 mg de tecido foliar macerado, foi adicionado 1 mL de reagente TRIZOL. As misturas foram homogeneizadas e incubadas por 5 min., em gelo, para permitir a completa dissociação das nucleoproteínas. Em seguida, foram adicionados 0,2 mL de clorofórmio e os tubos invertidos vigorosamente. As amostras foram incubadas por 5 min. em temperatura ambiente e, depois, centrifugadas a 11.000 rpm por 15 min., a 4°C. A centrífuga utilizada foi da marca Eppendorf®, rotor F-45-30-11. Após a centrifugação, foi coletada a fase aquosa superior, contendo o RNA, a qual foi transferida para novos tubos. Esta solução foi precipitada com 0,5 mL de isopropanol e centrifugada a 11.000 rpm por 10 min. a 4°C; após a centrifugação, o sobrenadante foi descartado, restando nos tubos um *pellet* contendo o RNA precipitado. O *pellet* foi lavado com 1 mL de etanol 75 % e foi feita uma nova centrifugação a 9.500 rpm por 5 min. a 4°C. Após esta última centrifugação, o sobrenadante foi descartado e o *pellet* foi parcialmente seco e dissolvido em água tratada com DEPC 0,01%.

Ao final da extração, todas as amostras foram tratadas com DNase, adicionando-se em cada tubo 3 U da enzima RQ1 DNase (Promega) e 3 µL de tampão de reação 10 x (Promega). As amostras foram incubadas a 37°C por 30 min.

A concentração e a qualidade das amostras de RNA total foram estimadas por espectrofotometria, no equipamento NanoDrop 2000 (Thermo), nos comprimentos de onda 260 nm e 280 nm. Foram aceitas apenas amostras de RNA com razão A_{260}/A_{280} entre 1,8 e 2,1. A integridade dos RNAs foi confirmada por eletroforese (3 V/cm) em gel de agarose 1,5 % contendo formaldeído 2 % em tampão TAE 1 x. As amostras foram visualizadas e fotodocumentadas sob luz UV, após coloração com SYBR Safe (0,5 x) (Invitrogen).

3.4.2 Síntese de cDNA

Para a síntese da primeira fita de cDNA foram preparadas reações utilizando 1 µg de RNA total de cada *bulk* (figura 4) e do controle do kit. Primeiramente, foram adicionados ao RNA 2,7 µM do primer '3' SMART CDS Primer II A' e água para completar o volume final (4,5 µL). Esta mistura foi incubada a 72°C por 3 min. e a 42°C por 2 min. A seguir, as reações foram completadas com 1 x de 'First-Strand Buffer', 2,5 mM de DTT, 1 mM de cada dNTP, 1,2 µM de 'SMARTer II A Oligonucleotide', 1 U de 'RNase Inhibitor' e 1 U de 'SMARTScribe™ Reverse Transcriptase', totalizando um volume final de 10 µL / reação. As reações foram incubadas a 42°C por 90 min. e a 70°C por 10 min. e diluídas em 40 µL de tampão TE.

Para as reações de síntese da segunda fita do cDNA por LD PCR foram utilizados 30 µL do cDNA de fita simples, 1 x de 'Advantage 2 PCR Buffer', 0,2 mM de cada dNTP, 0,24 µM do '5' PCR Primer II A', 1 x de 'Advantage 2 Polymerase Mix' e água para completar o volume final (300 µL). Antes da amplificação, esta mistura foi aliqüotada em 3 tubos com 100 µL com a finalidade de determinar o número ideal de ciclos. Os três tubos de cada *bulk* foram submetidos à ciclagem a 95°C por 1 min. em diferentes números de ciclos a 95°C por 15 s, 65°C por 30 s e 68°C por 6 min.

Conforme apresentado na figura 8, ao atingir 15 ciclos, os tubos foram retirados do termociclador e estocados a 4°C. De um dos tubos foram retirados 30 µL da reação que retornou ao termociclador para completar os 27 ciclos. Porém, desta reação, foram sendo retirados 5 µL a medida em que se atingiu 18, 21, 24 e 27 ciclos. Estas alíquotas (5 µL) foram submetidas à eletroforese em gel de agarose 1,2 %, com tampão TAE 1 x, usando-se um marcador de peso molecular 1 kb. Posteriormente, foram coradas com SYBR Safe (0,5 x) e fotodocumentadas sob luz UV visando à identificação do número de ciclos em que se obteve o melhor perfil de amplificação. Após esta etapa, os tubos estocados a 4°C deveriam retornar ao termociclador para completar o número de ciclos, porém, não foi necessário, pois as reações apresentaram melhor perfil de amplificação já com 15 ciclos. Para interromper as reações, foram adicionados 2 µL de EDTA 0,5 M.

As três alíquotas do produto da LD PCR foram misturadas e purificadas em coluna cromatográfica CHROMA SPIN-1000. Antes da purificação, 7 µL do produto

de reação foram reservados e, juntamente com uma alíquota da amostra purificada, foram submetidos à eletroforese em gel de agarose 1,2 %. As amostras foram coradas com SYBR Safe (0,5 x) para a visualização da concentração dos cDNAs de fita dupla antes e depois dessa purificação.

Figura 8 - Representação esquemática do protocolo visando à otimização do número de ciclos de amplificação para a síntese da segunda fita do cDNA, conforme consta no manual do kit 'SMARTer™ PCR cDNA Synthesis'

3.4.3 Obtenção das mini-bibliotecas de cDNA

A construção das mini-bibliotecas enriquecidas com genes de maracujá azedo diferencialmente expressos em resposta à *Xap* constou de duas subtrações. Para a subtração *forward* foi utilizado como *tester* o cDNA oriundo do *bulk* de transcritos do tratamento com *Xap* e, como *driver*, o cDNA oriundo do *bulk* de transcritos do tratamento controle, visando identificar os genes expressos apenas nas mudas inoculadas com o patógeno. Para a subtração *reverse*, foi utilizado como *tester* o cDNA do *bulk* do tratamento controle e, como *driver*, o cDNA do *bulk* do tratamento com *Xap*, visando identificar os genes expressos apenas nas mudas sadias.

Primeiramente, os cDNAs de fita dupla (ds-cDNA) purificados foram digeridos com a enzima de restrição *Rsa*I. Esta etapa é necessária para gerar fragmentos curtos com extremidades abruptas, favorecendo a ligação dos adaptadores e a subtração.

Para a hidrólise enzimática utilizou-se todo o produto de cDNA purificado (exceto uma alíquota de 10 µL que foi utilizada para se fazer a eletroforese) adicionando-se ao produto 1 x de '*Rsa*I restriction buffer' e 15 U da enzima *Rsa*I. A reação foi incubada a 37°C por 3 h e, para finalizá-la, adicionaram-se 8 µL de EDTA 0,5 M. Para cada *bulk*, os 10 µL do cDNA íntegro e 10 µL do cDNA digerido foram submetidos à eletroforese em gel de agarose 1,2 % e as amostras coradas com SYBR Safe (0,5 x) para a confirmação da digestão.

O produto da hidrólise foi purificado com fenol:clorofórmio:álcool isoamílico (25:24:1) e clorofórmio:álcool isoamílico (24:1), precipitado com NH₄OAc 4 M e etanol 95 %, lavado com etanol 80 % e dissolvido em 5,5 µL de água ultrapura.

As próximas etapas descrevem os procedimentos para a realização da subtração *forward*. Porém, os procedimentos para a subtração *reverse* são os mesmos, invertendo-se apenas as amostras usadas como *tester* e *driver*.

Para a ligação de adaptadores na amostra *tester* (amostra oriunda das mudas inoculadas com *Xap*), diluiu-se 1 µL do cDNA digerido e purificado em 5 µL de água ultrapura. Duas reações foram preparadas, uma para cada tipo de adaptador, utilizando 2 µL do cDNA *tester* diluído, 1 x de tampão de ligação, 400 U de enzima DNA ligase, 2 µM de adaptador (Adaptor 1 ou Adaptor 2R) e água ultrapura para completar um volume final (10 µL). As reações foram incubadas a 16°C *overnight*;

em seguida, adicionou-se 1 μ L de EDTA/Glycogen e as reações foram aquecidas a 72°C por 5 min. para a inativação da DNA ligase.

Para a primeira hibridização, também duas reações foram preparadas, uma para cada amostra de *tester* ligada a um tipo de adaptador. Em cada reação, foram adicionados 1,5 μ L do cDNA digerido e purificado da amostra *driver* (amostra oriunda das mudas inoculadas com solução salina), 1,5 μ L da amostra *tester* ligada ao adaptador e 1 x de 'hybridization buffer'. As reações foram incubadas a 98°C por 1,5 min. e a 68°C por 8 h.

Para a segunda hibridização procedeu-se da seguinte forma: quando a primeira hibridização estava terminando, concomitantemente desnaturou-se 1 μ L do cDNA digerido e purificado da amostra *driver*, misturado a 1 x de 'hybridization buffer' e água ultrapura para completar um volume final (4 μ L); esta reação foi incubada a 98°C por 1,5 min. Esta nova alíquota da amostra *driver* desnaturada e as outras duas reações da primeira hibridização foram misturadas em um único tubo, que foi incubado a 68°C *overnight*. Ao final da segunda hibridização, foram adicionados 200 μ L de 'dilution buffer' e a reação foi novamente aquecida a 68°C por 7 min.

Na primeira PCR, foi usado 1 μ L do produto da segunda hibridização, 1 x de 'Advantage[®] 2 PCR Buffer' (Clontech), 1 mM de cada dNTP, 0,4 μ M do 'PCR Primer 1', 1 x do 'Advantage[®] 2 Polymerase Mix' (Clontech) e água ultrapura para completar um volume final (25 μ L). A reação foi incubada a 75°C por 5 min. para a extensão dos adaptadores e em seguida se deu início ao programa de amplificação: 94°C por 25 s e 27 ciclos de 94°C por 10 s, 66°C por 30 s e 72°C por 1,5 min.

Na *Nested* PCR, diluíram-se 3 μ L do produto da primeira PCR em 27 μ L de água ultrapura e utilizou-se 1 μ L desta diluição como *template*. À reação foram adicionados também 1 x de 'Advantage[®] 2 PCR Buffer', 1 mM de cada dNTP, 0,4 μ M do 'Nested PCR Primer 1', 0,4 μ M de 'Nested PCR Primer 2R', 1 x do 'Advantage[®] 2 Polymerase Mix' e água ultrapura para completar um volume final (25 μ L).

Alíquotas da primeira PCR e da *nested* PCR foram submetidas à eletroforese em gel de agarose 2 % e coradas com SYBR Safe (0,5 x) para se analisar a qualidade das reações. O produto da *nested* PCR é resultado da amplificação

preferencial de transcritos diferencialmente expressos, ou seja, é uma biblioteca subtrativa.

3.5 Clonagem e sequenciamento dos fragmentos de cDNAs das bibliotecas

Antes da clonagem dos fragmentos de cDNA das bibliotecas subtrativas, todo o conteúdo da *nested* PCR foi aplicado em um gel de agarose, em um único poço. Após a eletroforese, o rastro de fragmentos foi retirado do gel em várias porções, a fim de separá-los por tamanho, visando evitar a ligação preferencial dos fragmentos menores nos vetores de clonagem. Estas porções de cDNA imersas na agarose foram purificadas com o kit 'Illustra GFX PCR DNA and Gel Band Purification' (GE), de acordo com os procedimentos descritos no manual do fabricante, obtendo-se conjuntos de cDNAs de diferentes tamanhos.

Os conjuntos de cDNAs foram ligados no plasmídeo pGEM[®]-T Easy (Promega), sendo usados, em cada reação, 150 ng de inserto, 50 ng de vetor, 1 x de 'Rapid Ligation Buffer' e 3 U de T4 DNA ligase (Promega). A reação foi incubada a 4°C *overnight*.

Do produto da ligação inserto-vetor, foram usados 2 µL para a transformação em 70 µL de cultura de células eletrocompetentes de *E. coli*, com o auxílio do eletroporador Gene Pulser II (BioRad). As células eletroporadas foram misturadas a 1 mL de meio SOC (20 g de triptona, 5 g de extrato de levedura, 0,5 g de NaCl, 20 mL de deglicose 1M e água q.s.p. 1 L) e crescidas sob agitação a 37°C por 1 h.

Para o cultivo das células transformadas, foram plaqueados 120 µL da suspensão bacteriana, 100 µL de X-Gal (2%) e 100 µL de IPTG (20%) em placas de Petri com 70 mL de meio LB sólido contendo ampicilina (50 µg/mL). As placas, que foram feitas em duplicata para cada produto transformado, foram incubadas sem vedação a 37°C por 14 h.

As colônias brancas foram repicadas e cultivadas individualmente em 200 µL de meio *Circle Grow* contendo 12 % de glicerol e ampicilina (100 mg/mL), em placas acrílicas de 96 poços tipo 'ELISA'. Estas placas foram incubadas a 37°C, *overnight*, sob agitação (125 rpm) e, em seguida, foram preservadas em ultra freezer a -80°C.

Para a extração dos plasmídeos foram repicados 10 µL das culturas preservadas em 1 mL de meio *Circle Grow* contendo ampicilina (50 µg/mL), em placas plásticas de 96 poços tipo '*deep well*'. As colônias foram crescidas a 37°C, por 22 h, sob agitação (250 rpm). Após o crescimento das colônias, procedeu-se a extração dos plasmídeos, por *miniprep*, em placas de 96 poços pelo método de lise alcalina (SAMBROOK; RUSSELL, 2001).

De cada placa de 96 poços, contendo os produtos da *miniprep*, foram selecionadas 24 amostras para as suas quantificações em NanoDrop 2000 e visualização em gel de agarose 1 %. Para a confirmação da presença e do tamanho dos insertos, essas mesmas amostras foram usadas para a amplificação por PCR, utilizando-se os *primers* M13 *forward* (5'GTTTTCCCAGTCACGAC3') e *reverse* (5'GAGCGGATAACAATTTTCACACAGG3'), os quais amplificam um fragmento de 272 pb do vetor, além do inserto. Nessas reações foram usados 2 µL dos produtos da *miniprep*, 1 x de tampão de reação, 1,5 mM de MgCl₂, 0,2 mM de cada dNTP, 0,5 µM de cada *primer*, 1 U de Go *Taq* Flex (Promega) e água ultrapura para completar um volume final (25 µL). As condições de amplificação foram as seguintes: 94 °C por 2 min., 35 ciclos de 94°C por 30 s, 60°C por 1 min. e 72°C por 1,5 min. e extensão final a 72°C por 10 min. em termociclador *Gen Amp PCR System* 9700 (ABI). Os amplicons foram separados por eletroforese a 3 V/cm em gel de agarose 1,2 % e visualizados e fotodocumentados sob luz UV após coloração com SYBR Safe (0,5 x) (Invitrogen).

O sequenciamento dos insertos foi realizado com base na metodologia de Sanger (SANGER et al., 1977). Para o sequenciamento dos insertos oriundos da biblioteca *forward* foram realizadas PCRs usando-se 1 µM do *primer* M13 *forward* (5'GTTTTCCCAGTCACGAC3'), 4 µL do produto da *miniprep* (aproximadamente 400 ng), 2 µL do mix de reação do kit DYEnamic ET Cycle Sequencing (GE) e 2 µL de tampão (Tris-HCl 200mM, pH 9.0, e MgCl₂ 5mM). As reações foram submetidas ao termociclador *Gen Amp PCR System* 9700 (ABI) por 30 ciclos de 95°C por 20 s e 60°C por 1 min. Os produtos da PCR foram purificados com acetato de sódio 1,5 M + EDTA 0,25 M, precipitados com etanol 95 % e lavados com etanol 70 %. As reações precipitadas foram enviadas ao Laboratório de Genômica e Biologia Molecular de Plantas da ESALQ/USP para serem submetidas à eletroforese capilar no sequenciador ABI Prism 3100 (ABI).

Para o sequenciamento dos insertos oriundos da biblioteca *reverse*, 5 µL do produto da *miniprep* (aproximadamente 500 ng) juntamente com 2 µL do *primer* M13 *forward* (5'GTAAAACGACGGCCAGT3') a 5 µM foram enviados ao Centro de Estudos do Genoma Humano da Universidade de São Paulo (<http://genoma.ib.usp.br/>). Neste Centro foram realizadas as PCRs, com o kit 'BigDye[®] Terminator v3.1 Cycle Sequencing' (ABI), e a eletroforese capilar, no sequenciador ABI 3730 DNA Analyser (ABI).

3.6 Processamento das sequências

Primeiramente, as sequências de cada biblioteca foram analisadas, separadamente, usando-se o programa computacional *CodonCode Aligner* v. 3.5.7 (CodonCode Corporation). Esta suíte de aplicativos faz a leitura dos arquivos de saída do sequenciador e permite a visualização dos eletroferogramas. Em seguida, analisa a qualidade das bases lidas (função '*Call bases with phred*'), elimina as bases iniciais e finais que possuem baixa qualidade (função '*Clip Ends*'), e, finalmente, exclui as sequências do vetor (função '*Trim Vector*') juntamente com as sequências dos *primers* dos adaptadores. Antes de realizar o '*base calling*', o arquivo de parâmetros do *phred* (EWING et al., 1998) foi alterado de acordo com o sequenciador que originou as sequências ('ABI_3100' ou 'ABI_3730'). O critério utilizado no '*end clipping*' foi o padrão do programa, que adota índice *Phred* de qualidade '20' como limite para considerar a base como sendo verdadeira, ou seja, a probabilidade da leitura da base estar incorreta é de 1 em 100. Ao final do processamento, foi feita a montagem das sequências (função '*Assemble*') para a detecção de sequências redundantes, obtendo-se, assim, os *contigs* (Figura 9).

Posteriormente, as sequências originais (arquivos de saída do sequenciador) obtidas nas duas bibliotecas foram reunidas e submetidas a um *pipeline* de processamento de pré-anotação, para uma nova análise. Este novo processamento visou também a avaliação da qualidade das sequências, mascaramento dos vetores e dos *primers* dos adaptadores, aparamento das pontas de baixa qualidade e filtragem por tamanho. Este *pipeline* foi desenvolvido utilizando o programa

computacional *Egene* (DURHAM et al., 2005), que é um sistema de geração modular de *pipelines*.

Sintetizando, a primeira etapa desse processamento foi o procedimento de ‘*base calling*’ e a avaliação da qualidade das bases, ou seja, foi realizada a interpretação dos eletroferogramas, tendo sido atribuída a cada um dos picos a identidade de uma base nitrogenada, baseada na forma e intensidade do pico, e um valor de qualidade a cada uma das bases (EWING; GREEN, 1998; RICHTERICH, 1998).

O ‘*base calling*’ foi feito usando-se o programa *Phred* (EWING et al., 1998), com base nos parâmetros padrão do sistema. Os arquivos de saída do *Phred* foram convertidos em arquivos FASTA formatados, por meio do PHD2FASTA.

Os arquivos ‘.fasta’, contendo as sequências, foram encaminhados à etapa seguinte, a qual constou da identificação e do mascaramento de trechos de sequências que apresentaram alta identidade com os *primers* e vetores usados na etapa de amplificação ou de clonagem dos fragmentos. O mascaramento desses trechos foi realizado pela ferramenta *CrossMatch* (<http://www.phrap.org/phredphrapconsed.html>), usando os parâmetros de penalidades e alinhamento mínimos, bem como o banco de referência de *primers* e vetores.

A etapa seguinte consistiu na eliminação de bases que apresentaram valores baixos de qualidade. Neste caso, os valores de *Phred* foram computados para todas as bases e comparados com um limite mínimo pré-estabelecido de 20, aceitando-se, então, apenas sequências que possuíam uma porcentagem de bases válidas (*Phred* acima de 20) superior a 85%.

Após o mascaramento das sequências, foi realizada a excisão dos trechos mascarados ou os que apresentaram índices de qualidade inferiores aos parâmetros previamente estipulados. Daí, as sequências que continham um número de bases insuficiente (< 100 pb) foram filtradas, a fim de eliminar aquelas que poderiam gerar viés nas análises.

Por fim, as sequências foram submetidas ao programa CAP3 (HUANG; MADAN, 1999) para a formação de *contigs*, eliminando-se assim as redundantes (Figura 9).

3.7 Anotação funcional das bibliotecas subtrativas

As sequências das duas bibliotecas, processadas com o auxílio do programa *CodonCode Aligner*, foram anotadas utilizando a Plataforma PLAZA v. 2.0 (PROOST et al., 2009). Nesta anotação, as sequências de cDNA foram comparadas com o banco de dados de sequências protéicas da espécie *Arabidopsis thaliana*, com base na ferramenta BlastX. O valor 10^{-5} (1e-05) foi adotado como máximo para aceitar o resultado do alinhamento.

A partir do resultado da anotação da biblioteca *forward*, uma análise manual foi realizada visando obter informações adicionais sobre as proteínas que apresentaram similaridade com as sequências anotadas. Para tanto, o *Universal Protein Resource* (UniProt) (www.uniprot.org/), o *Gene Ontology* (www.geneontology.org/) e a literatura científica da área foram explorados.

Concomitantemente, foram criadas categorias funcionais para alocar as proteínas a partir das informações obtidas, com base nas categorias apresentadas pelo COG (*Clusters of Orthologous Groups of proteins*) e pelo MIPS (*Munich Information Center for Protein Sequences*). Cada proteína foi atribuída a uma categoria funcional (Figura 9).

Em outra análise manual, as informações obtidas sobre as proteínas serviram de base para selecionar todas aquelas que estão envolvidas direta ou indiretamente no processo de defesa de plantas, segundo a literatura da área. As proteínas selecionadas foram analisadas no banco de dados STRING 9.0 (SZKLARCZYK et al., 2011), com a finalidade de identificar possíveis interações. Nestas análises, foram usados os códigos do banco TAIR (SWARBRECK et al., 2007) atribuídos às sequências durante a anotação realizada na plataforma PLAZA. Assim, as interações identificadas tiveram por base aquelas informações disponíveis para *A. thaliana*.

Uma nova anotação das sequências das duas bibliotecas foi realizada, porém, desta vez, usando os resultados do processamento realizado pelo *pipeline* de processamento de pré-anotação e o programa Blast2GO® v. 2.6.2 (CONESA et al., 2005). Nesta análise, foi usado o banco de dados GeneBank e a ferramenta BlastX, aceitando-se um *expect value* máximo de 10^{-6} e um número máximo de 20

hits por entrada. Após o mapeamento dos termos do '*Gene Ontology* (GO)' e a anotação das sequências, foi utilizada a função '*GOSlim*' para reduzir o vocabulário da anotação, usando o subconjunto de termos '*Plant*' (Figura 9).

Figura 9 – Etapas envolvidas no processamento das sequências pelo programa computacional CodonCode Aligner v. 3.5.7 e pelo pipeline de processamento de pré-anotação e etapas envolvidas na anotação pela plataforma PLAZA e pelo programa computacional Blast2GO

3.8 Validação das bibliotecas subtrativas

3.8.1 Síntese de cDNA

Foram sintetizados novos cDNAs, a partir dos RNAs totais dos fragmentos foliares coletados no primeiro e no segundo experimentos de inoculação. Estes RNAs foram extraídos usando o mesmo protocolo descrito no item 3.4.1. Estes novos cDNAs foram usados nas etapas envolvidas na validação das bibliotecas, que serão descritas nos itens 3.8.2, 3.8.3 e 3.8.4.

Para a síntese de cDNA de fita simples (ss-cDNA) foi utilizado 1 µg de RNA total, 1 x de tampão da M-MLV RT (Promega), 200 U de M-MLV *Reverse Transcriptase* (Promega), 25 U de inibidor de RNase RNasin (Promega), 0,25 µM de *primer* CDS (5'AAG CAG TGG TAT CAA CGC AGA GTA CTT TTT TTT TTT TTT TTT TTT TTT TTT TTN V3'), 0,6 mM de cada dNTP e água ultrapura tratada com DEPC (0,01 %) para completar o volume final (25 µL). As reações foram incubadas a 42°C por 1 h.

Com a finalidade de testar a qualidade do cDNA sintetizado e a ausência de DNA genômico de maracujá azedo, foram feitas PCRs com um par de *primers* que amplifica um fragmento do gene *NDID* (*NADP-dependent isocitrate dehydrogenase*) (F:5'CAAGGTCGCTAATCCCATC3' e R: 5'TTTCTGGTAAGCCGTGTTTC3') e tendo como *template* o cDNA ou DNA genômico. Este produto de amplificação, quando se usa como *template* cDNA de maracujá azedo, é de 598 pb (PEREIRA et al., 2013). Quando se usa DNA genômico de maracujá azedo, a PCR produz dois fragmentos maiores de, aproximadamente, 1.000 e 2.000 pb, devido à presença de íntrons. Assim, quando o produto da PCR, tendo como *template* o cDNA, é um amplicon de 598 pb, isto significa que a síntese do cDNA foi satisfatória e que na amostra de RNA não houve contaminação por DNA genômico.

Para as PCRs foram utilizados 5 µL de ss-cDNA diluído 20 x (ou 20 ng de DNA genômico), 1 x de tampão de reação, 1 mM de MgCl₂, 0,2 mM de cada dNTP, 0,3 µM de cada *primer*, 1,2 U de Go *Taq* Flex (Promega) e água ultrapura para completar um volume final (20 µL). As condições de amplificação foram: 95°C por 5 min., 35 ciclos de 95°C por 40 s, 55°C por 40 s e 72°C por 1 min. e extensão final a 72°C por 8 min. em um termociclador *Gen Amp PCR System 9700* (ABI). Os

amplicons foram separados por eletroforese a 3 V/cm em gel de agarose 1 % e visualizados e fotodocumentados sob luz UV após coloração com SYBR Safe (0,5 x) (Invitrogen).

A eficiência da transcrição reversa (RT) também foi analisada. Para isto, selecionou-se a sequência gênica F-contig33, oriunda da biblioteca SSH, por ser uma das maiores sequências da biblioteca e por apresentar a região CDS completa. A partir desta sequência, foram desenhados dois pares de *primers* para a amplificação por qPCR, sendo que um par de *primers* amplifica um fragmento de 196 pb da região 3' e o outro par amplifica um fragmento de 175 pb da região 5' da sequência (Figura 10). A região entre os dois fragmentos é de 665 pb.

>F-contig33

```
5' ATGGGGACCCACCTCGCCTCATTCTTGGTTTCAGAGAGAGTAGCCGAATACAAAACCGCTACTCCGGTCTCTA
GAATACTGTTTTGCACCGAACACCGCCGCTTCTTCGGTCTCGTTTTCTTCCGAGGAACCATGGGAAAAGGCCCC
GGGCTTTACACCGATATCGGCAAGAGAGCTAGAGGATCTTCTCTACAAGGATTACCAGAGCGACCACAAGTTCACA
CTCACTACTTATACCTCAACTGGAGTCGCAATCACTTCTAATGGGCTGAAAAAAGGCGATCAGTTTTTGGCTGAT
GTGAGCACTCAGCTGAAGAACAATAACATCACTTCTGATGTCAAGGTTGACACGAACCTCAATCTTTAGCAACC
ATCACGGTGGATGAACCTCGCTCCTGGACTCAAGACAATATTTAGCTTCAGAGTTCCTGATCAGAGGTCTGGAAG
GTGGAACCTGCAATACCAGCATGAATATGCAGGAATTAGTGCTGCCCTTGGGCTGACTGCAAATCCCATTTGTTAAC
TTCTCAGGCGTTGTTGGCAGTGAGGCTATTTCTCTGGGCACTGATGTCTCATTTCGACACTGCTTCTGGAACTTT
ACCCAAATACATGCAGGGTTGAGTTATACCAAAGATGACCTTATTGCTTCTCTGGCAGTGAACGATAAGGGTGAC
ACTGTTATCGCTTCTTATTATCACATAGTGAGCCCATTAACCAATACAGCTGTTGGTGCGGAACCTGGCTCATAAC
TTTTCAAGCAACGAAAACACGCTGACTATTGGAACACAGCATGCACTCGATCCCTTGACGACATTGAAGGCTCGT
GTGAACAGCTTTGGCAGAGCCAGTGCTCTCATTTCAGCATGAGTGGCGCCCAAATCTCTCTTCACAATCTCAGGA
GAAGTGGACACAAGGGCAATTGAAAAGAGTGCTAAGGTTGGACTAGCCGTGGCTCTTAAGCCATAGGTCGTGATC
TGGAAGAGTGGGCTTGCCTTTCTTCAGTAGCAGAGAGCAATAATGTGATGGTGAATGATACGAGTAACCATTTT
CCAAATGGGACGTTATATCCGAGTTATTGTGTCATTGATTTTTGCATATTTTGGTGTTATGGATTGTAGAAAA
TAATTGTAGCTGGTGCT3'
```

Figura 10 – Sequência de cDNA obtida na biblioteca *forward*, a qual foi usada para desenhar dois pares de *primers*, destacados em negrito. A região sublinhada representa a sequência codante completa, que codifica uma porina

As reações foram feitas, separadamente, para cada amostra de cDNA, contendo 3,0 µL de ss-cDNA diluído 20x, 5,1 µL de GoTaq® qPCR Master Mix (Promega), 0,3 µM de cada *primer* e água ultrapura para completar um volume final de 10 µL, e submetidas ao equipamento StepOne Real-Time PCR System (ABI), nas seguintes condições: 95°C por 5 min.; 40 ciclos de 95°C por 5 s e 60°C por 1 min. com captação do sinal de fluorescência; por último, para a obtenção da curva de *Melt*, fez-se um passo a 95°C por 15 s, 60°C por 1 min. e captação de fluorescência a cada 0,5°C até a temperatura atingir 95°C.

A curva de *Melt* é uma curva que registra o decaimento da fluorescência à medida que as duplas fitas são dissociadas com o aumento da temperatura, e é

utilizada para avaliar a qualidade da PCR quando se utilizam corantes intercalantes de DNA. Assim, quando a curva apresenta apenas um pico, isto indica que a PCR produziu apenas um único tipo de fragmento. Dois ou mais picos na curva de *Melt* indicam produtos de amplificação inespecíficos, como dímero de *primers*, inviabilizando a reação para uma análise quantitativa.

Após a determinação dos valores dos ciclos de quantificação (C_q) com auxílio do programa LinRegPCR (RAMAKERS et al., 2003), calculou-se o ΔC_q entre as duas reações (cada uma com um par de *primers*) de cada amostra de cDNA sintetizada, e considerou-se a RT eficiente e o cDNA apropriado para o uso quando $C_{q3'}-C_{q5'} < 2$ (KAKAR et al., 2008).

3.8.2 Seleção de sequências putativas codificadoras das bibliotecas e desenho de *primers*

Para a validação das duas bibliotecas por PCR em Tempo Real, tendo como base as anotações e as informações sobre as respectivas proteínas possivelmente codificadas pelos transcritos das bibliotecas, foram selecionadas de cada biblioteca 50 sequências putativas que codificam proteínas relacionadas ao processo de defesa de plantas, bem como proteínas representantes de diversas outras categorias funcionais.

Foram desenhados *primers* com o auxílio do programa GeneRunner (<http://www.generunner.net/>), obedecendo-se aos seguintes parâmetros: tamanho do amplicon de 100 a 200 pb; tamanho do *primer* de 18 a 22 nucleotídeos; temperatura de *Melting* de 58°C a 62°C; porcentagem de 'GC' do *primer* entre 40 e 60 %; ausência de estruturas secundárias como dímeros e *hairpins*.

Os *primers* desenhados foram testados quanto à amplificação de fragmentos únicos, de tamanhos esperados, preparando-se para isso PCRs convencionais e tendo como *template* ss-cDNA, conforme as condições de reação descritas no item 3.8.1.

3.8.3 Seleção e validação de genes de referência para a normalização da qPCR

Na busca de genes de *P. edulis* f. *flavicarpa* para servirem de normalizadores nas reações de quantificação por PCR em Tempo Real, foram testados alguns dos genes estudados por Pereira et al. (2013). Também, se utilizou a sequência do gene de *P. edulis* *EF1a1* (*Translation elongation factor 1a-1*), disponível no *GeneBank* (GB nº DQ447160.1), para o desenho de um par de *primers* para qPCR.

A partir dos *primers* disponibilizados por Pereira et al. (2013) foram feitas reações nas mesmas condições descritas no item 3.8.1. Os fragmentos amplificados foram retirados do gel de agarose e purificados com o kit 'Illustra GFX PCR DNA and Gel Band Purification' (GE). Os produtos da purificação foram quantificados e usados como *template* para a realização das PCRs de sequenciamento. Essas reações foram feitas conforme descrito no item 3.5, porém, usando 50 ng dos produtos de PCR purificados. Para cada fragmento gênico, as reações de sequenciamento foram feitas nos dois sentidos, *forward* e *reverse*.

As novas sequências obtidas, agora a partir de transcritos, foram usadas para o desenho de *primers* para qPCR, conforme descrito no item 3.8.2. Estes novos *primers* foram testados quanto à amplificação de fragmentos únicos, de tamanhos esperados, conforme descrito no item 3.8.1.

Para calcular a estabilidade dos candidatos a genes de referência, foram feitas reações de amplificação usando os *primers* que produziram os amplicons esperados e os cDNAs sintetizados a partir dos transcritos dos dois experimentos de inoculação, separadamente.

Neste caso, foram usadas as nove amostras de cDNA do tratamento com *Xap* e as nove amostras de cDNA do tratamento controle. Ou seja, para o primeiro experimento de inoculação (Figura 4), os cDNAs foram sintetizados a partir do RNA dos fragmentos foliares de cada muda, sem fazer *bulks*; já para o segundo experimento, o cDNA foi sintetizado de acordo com o esquema da figura 5.

As reações foram submetidas ao equipamento StepOne Real-Time PCR System (ABI), contendo 3,0 µL de ss-cDNA diluído 20 x, 5,1 µL de GoTaq® qPCR Master Mix (Promega), 0,3 µM de cada *primer* e água ultrapura para completar um volume final (10 µL). As condições de ciclagem foram as seguintes: 95°C por 5 min; 40 ciclos de 95°C por 5 s e 60°C por 1 min. com captação do sinal de fluorescência;

por último, para a obtenção da curva de *Melt*, fez-se um passo a 95°C por 15 s, 60°C por 1 min. e captação de fluorescência a cada 0,5°C até a temperatura atingir 95°C. Foram utilizadas três repetições biológicas e três replicatas técnicas.

Os valores dos ciclos de quantificação (C_q) e das eficiências de amplificação foram determinados usando o programa LinRegPCR (RAMAKERS et al., 2003), a partir dos dados brutos de amplificação. As eficiências de cada gene e os C_q médios das triplicadas técnicas foram utilizados para o cálculo do ' Q ', o qual transforma os valores dos C_q em quantidades relativas à amostra de maior expressão, usando a fórmula $Q = E^{(C_q \text{ menor} - C_q \text{ da amostra})}$, onde ' E ' é a eficiência da reação e ' $C_q \text{ menor}$ ' se refere a amostra de maior expressão, por apresentar o menor C_q .

Os valores de ' Q ' foram utilizados para a análise da estabilidade da expressão de cada gene candidato e para a análise da melhor combinação de genes normalizadores para a condição experimental estudada. Para isso, foram utilizados dois programas computacionais que utilizam o método do C_q , quais sejam, o geNorm v. 3.5 (VANDESOMPELE et al., 2002) e o NormFinder (ANDERSEN; JENSEN; ØRNTTOFT, 2004). O programa geNorm utiliza o método *pairwise*, ou seja, o nível da expressão dos genes é determinado em pares e comparados com o nível de expressão dos demais genes. Já o programa NormFinder utiliza método do *model-based*, o qual seleciona os melhores genes normalizadores baseando-se na variação intergrupo e intragrupo.

3.8.4 Validação da expressão diferencial dos transcritos por qPCR

A validação das bibliotecas foi conduzida em duas etapas. Na primeira, foram feitas PCRs quantitativas utilizando-se todos os *primers* que apresentaram boa qualidade no teste de amplificação, descrito no item 3.8.2. Nestas reações foram usados os dois *bulks* de cDNAs do primeiro experimento de inoculação com *Xap* ou solução salina, conforme esquematizado na figura 4. Foram feitas quatro repetições técnicas para cada *bulk* e a determinação da razão da expressão de cada gene foi feita para o *bulk* de cDNAs do tratamento com *Xap* em relação o *bulk* de cDNAs do tratamento controle.

Na segunda etapa, as PCRs quantitativas foram feitas com alguns dos genes analisados na primeira etapa e com os cDNAs obtidos das mudas do segundo

experimento de inoculação, conforme esquematizado na figura 5. A determinação da razão da expressão de cada gene, para cada período de avaliação, foi feita para os cDNAs do tratamento com *Xap* em relação aos cDNAs do tratamento controle. Utilizaram-se três repetições biológicas, com três repetições técnicas cada, por período de avaliação.

As reações foram preparadas com 3,0 µL de ss-cDNA diluído 20 x, 5,1 µL de GoTaq® qPCR Master Mix (Promega), 0,3 µM de cada *primer* e água ultrapura para completar um volume final (10 µL), e submetidas ao equipamento StepOne Real-Time PCR System (ABI) nas seguintes condições: 95°C por 5 min; 40 ciclos de 95°C por 5 s e 60°C por 1 min. com captação do sinal de fluorescência; por último, para a obtenção da curva de *Melt*, fez-se um passo a 95°C por 15 s, 60°C por 1 min. e captação de fluorescência a cada 0,5°C até a temperatura atingir 95°C.

Os dados brutos (sem a correção da linha de base) dos níveis de fluorescências foram exportados do equipamento e analisados no programa computacional LinRegPCR (RAMAKERS et al., 2003), para a obtenção da eficiência média (E) das curvas de amplificação e dos valores dos ciclos de quantificação (C_q). Este programa corrige a linha de base e faz uma análise de regressão linear para cada curva de amplificação, definindo os pontos mais bem correlacionados em cada curva (janela de linearidade). Isto permite o cálculo dos valores do *threshold* e do C_q , sendo que para cada gene foi padronizado um único *threshold*. Também, o programa calcula a eficiência da amplificação de cada curva baseando-se na inclinação (*slope*) da reta ($E = 10^{\text{slope}}$), considerando como ideais os valores de eficiência entre $1,8 \leq E \leq 2$ e uma correlação (R) entre os pontos das curvas $\geq 0,995$.

Para os cálculos dos níveis de expressão relativos de cada transcrito (utilizando-se como normalizadores os genes descritos no 3.8.3), bem como nas análises estatísticas, usou-se a metodologia implementada no programa computacional REST (*Relative Expression Software Tool*) v. 2.0.13 (PFAFFL; HORGAN; DEMPFLÉ, 2002). Este programa calcula a razão da expressão dos genes, utilizando um modelo matemático baseado nas eficiências das PCRs, para os genes alvos e normalizadores, e nos desvios dos C_q entre as amostras dos grupos 'tratado' e 'controle'. Para isso, aplica um teste de realocação fixa aos pares (*Pair Wise Fixed Reallocation Randomisation Test*), com *bootstrap* de 2.000 permutações, para a determinação dos níveis de significância.

3.9 Análise detalhada das sequências dos genes da lipoxigenase e da (+)-neomentol desidrogenase

A partir dos resultados da segunda etapa de validação das bibliotecas, dois transcritos que apresentaram maior diferença na expressão entre os tratamentos de inoculação com *Xap* e com solução salina, aos 5 e 9 d.a.i., e que codificam, respectivamente, as enzimas lipoxigenase e (+)-neomentol desidrogenase, foram selecionados para a comparação com sequências similares de outras espécies vegetais.

As sequências dos transcritos foram traduzidas com o auxílio da ferramenta *Translate* do *ExPASy* (<http://web.expasy.org/translate/>) e as respectivas sequências de aminoácidos foram selecionadas pelo *frame* que mais bem caracterizava um quadro aberto de leitura (ORF). As sequências de aminoácidos de maracujá azedo foram então comparadas com sequências protéicas disponíveis no *GeneBank*, através da ferramenta BlastP.

As sequências protéicas de diversas espécies vegetais que apresentaram similaridade com as sequências de maracujá azedo foram selecionadas para se realizar um alinhamento múltiplo, pelo programa MEGA v. 5.10 (TAMURA et al., 2011), usando o algoritmo ClustalW. A partir do alinhamento foi computada uma estimativa de distância evolutiva por distância aos pares, com 1.000 reamostragens por *bootstrap*, usando o método *p-distance* como modelo de substituição. Em seguida, foi construída uma árvore filogenética utilizando-se o método *Neighbor-Joining*.

4 RESULTADOS

4.1 Avaliação dos experimentos de inoculação

Conforme descrito em Materiais e Métodos, foram realizados dois experimentos nos quais mudas de maracujá azedo, do genótipo 'IAPAR123', foram inoculadas com o isolado AP302 de *X. axonopodis* pv. *passiflorae* ou com solução salina (NaCl 0,85 %), para controle, visando à construção e validação de bibliotecas subtrativas.

Concomitantemente, mudas extras foram usadas para avaliar a eficiência da inoculação e, para isto, foram coletadas amostras das folhas inoculadas para o diagnóstico molecular da doença. Assim, as PCRs feitas com os DNAs das folhas inoculadas com *Xap* e com solução salina, em todos os períodos de avaliação (6, 12 e 24 h.a.i. e 5 e 9 d.a.i.) e aos 14 d.a.i., e utilizando o conjunto de *primers Xapas*, apresentaram os amplicons específicos, de 520 pb, que confirmam a presença do patógeno, apenas nas plantas inoculadas com *Xap*. Neste caso, foi possível observar bandas mais intensas nos períodos de 5, 9 e 14 d.a.i., indicando a proliferação do patógeno. Nas folhas inoculadas com solução salina não houve a amplificação dos fragmentos de DNA esperados, confirmando a ausência da bactéria (Figura 11).

Figura 11 – Eletroforese em gel de agarose mostrando os produtos da PCR usando o conjunto de *primers Xapas* e DNA de maracujá azedo como *template*. Canaletas 1 a 8, amostras do primeiro experimento, coletadas às 6, 12 e 24 h.a.i e aos 14 d.a.i., e do segundo experimento, coletadas às 6 h.a.i. e aos 5, 9 e 14 d.a.i., respectivamente, e inoculadas com *Xap*; Canaletas 9 a 16, amostras controle, inoculadas com solução salina; Canaleta 17, DNA do isolado AP302; Canaleta 18, reação de amplificação sem *template*. A seta indica o amplicon de 520 pb, específico de *Xap*. M, Marcador de peso molecular 100 pb (Invitrogen)

Aos 14 dias, as folhas inoculadas com *Xap* apresentaram o sintoma característico da bacteriose, ou seja, uma região necrosada, com aspecto encharcado, circundada por um halo clorótico. Por outro lado, as folhas inoculadas com solução salina não apresentaram nenhum sinal, além da injúria causada pelo disco perfurado (figura 12).

Figura 12 – Folhas de maracujá azedo 14 dias após a inoculação *Xanthomonas axonopodis* pv. *passiflorae* (A e C) e com solução salina (B e D), do primeiro (A e B) e segundo (C e D) experimentos de inoculação

4.2 Construção das bibliotecas subtrativas

As seis amostras de RNA total extraídas para a construção das bibliotecas apresentaram boa qualidade, ou seja, foi possível observar em agarose a presença de bandas definidas, sem sinal de degradação, além das duas bandas mais intensas, características de RNA ribossômico, 28S e 18S (Figura 13). As leituras no NanoDrop mostraram concentrações entre 1.200 e 2.300 ng/ μ L e razões A_{260}/A_{280} entre 1,80 e 1,92.

Figura 13 – Eletroforese em gel de agarose do RNA total extraído de tecidos foliares inoculados com *Xap* às 6, 12 e 24 h.a.i (canaletas 1, 2 e 3) e de tecidos foliares inoculados com solução salina às 6, 12 e 24 h.a.i (canaletas 4, 5 e 6)

A síntese de cDNA de dupla fita por LD PCR apresentou melhor resultado com 15 ciclos de amplificação (Figura 14A), sendo que após esta fase a reação atingiu seu ponto de saturação, apresentando, na eletroforese em gel de agarose, rastros mais fracos em relação aos rastros obtidos com 15 ciclos.

A purificação dos cDNAs em coluna cromatográfica apresentou o resultado recomendado no manual do kit, ou seja, após a eletroforese em gel de agarose, o rastro do produto purificado teve intensidade aproximadamente 50 % menor em relação à intensidade do rastro do produto que não sofreu purificação, já que uma alíquota foi reservada e submetida à eletroforese antes da purificação.

A digestão dos cDNAs com a enzima *Rsa*I também foi satisfatória, pois as amostras antes da digestão (Figura 14B-1 e 14B-3) apresentam um rastro mais intenso, de fragmentos com amplitude de 0,5 a 10 Kb, e as amostras após a digestão (Figura 14B-2 e 14B-4) apresentam um rastro mais fraco, de fragmentos com amplitude de 0,1 a 2 kb, conforme previsto no manual do kit.

Figura 14 – **A**, Eletroforese em gel de agarose mostrando, em duplicata, o produto da síntese de cDNA fita dupla por LD PCR com 15 ciclos de amplificação, do *bulk* das mudas inoculadas com *Xap* (1 e 2) e do *bulk* das mudas inoculadas com solução salina (3 e 4). **B**, Eletroforese em gel de agarose mostrando os ds-cDNAs dos *bulks* das mudas inoculadas com *Xap* e com solução salina, antes (1 e 3) e após a digestão com a enzima *RsaI* (2 e 4). M1 e M2, Marcadores de peso molecular 1 Kb e 100 pb, respectivamente (Invitrogen)

A eficiência da etapa de ligação dos adaptadores, da primeira e segunda hibridização só foi confirmada a partir dos resultados da primeira PCR e da *Nested* PCR, para cada biblioteca, separadamente. Isto porque, conforme mostra a figura 15, a primeira PCR gerou os produtos de amplificação esperados, de acordo com o manual do kit, ou seja, gerou um rastro de fragmentos amplificados com amplitude de tamanho entre 0,2 e 2,2 kb, aproximadamente. Esta amplitude de tamanho dos fragmentos corrobora o resultado da digestão dos cDNAs com a enzima *RsaI*, que gerou fragmentos entre 0,1 a 2 kb. Os fragmentos foram um pouco maiores, após a amplificação, devido à presença dos adaptadores. A *nested* PCR também mostrou o resultado esperado, gerando o mesmo rastro de fragmentos, porém com maior intensidade (figura 15), confirmando a amplificação dos produtos gerados na primeira PCR.

Figura 15 – Eletroforese em gel de agarose mostrando os produtos da primeira PCR e da *nested* PCR durante a construção das bibliotecas SSH *forward* (A) e *reverse* (B). Resultados da primeira PCR, para os produtos gerados nas bibliotecas (canaletas 1A e 1B) e para o controle do kit (canaletas 2A e 3B), e da *nested* PCR, para os produtos gerados nas bibliotecas (canaletas 3A e 2B) e para o controle do kit (canaletas 4A e 4B). M1 e M2, marcadores de peso molecular 1 Kb e 100 pb, respectivamente (Invitrogen)

4.3 Clonagem dos fragmentos de cDNAs das bibliotecas subtrativas

Dos produtos da *nested* PCR das bibliotecas *forward* e *reverse*, submetidos à eletroforese em gel de agarose, obtiveram-se conjuntos de cDNAs de diferentes tamanhos (figura 16).

Figura 16 – Eletroforese em gel de agarose mostrando os produtos da *nested* PCR (NP) e os conjuntos de fragmentos de cDNAs purificados das bibliotecas *forward* (A) e *reverse* (B). F1-F5, conjuntos de fragmentos da biblioteca *forward*; R1-R3, conjuntos de fragmentos da biblioteca *reverse*. M, Marcador de peso molecular 100 pb (Invitrogen)

As placas com a cultura dos clones obtidos a partir dos conjuntos de fragmentos menores apresentaram maior rendimento, ou seja, apresentaram grande número de colônias por placa, sendo a maioria das colônias brancas. Já para os

conjuntos de fragmentos maiores, o rendimento da clonagem foi baixo, apresentando poucas colônias por placa; isto porque fragmentos maiores reduzem a eficiência da transformação e da multiplicação dos plasmídeos pela bactéria.

O tamanho dos fragmentos de cDNA de cada conjunto e o número de clones obtidos a partir da transformação destes fragmentos estão apresentados na tabela 1.

Tabela 1 - Conjuntos de fragmentos de cDNA purificados do gel de agarose a partir do produto da *nested* PCR para as bibliotecas *forward* (F) e *reverse* (R), tamanho dos fragmentos e número de clones por conjunto

Conjunto de fragmentos de cDNA	Tamanho (em pb)	Número de clones
F1	400 a 650	240
F2	650 a 900	240
F3	900 a 1300	240
F4	1.300 a 1.700	192
F5	1.700 a 2.100	48
R1	370 a 650	180
R2	650 a 1.200	192
R3	1.200 a 1.600	12

A quantificação de uma amostra de 25 % dos produtos da *mini-prep*, feita em NanoDrop 2000, apresentou uma concentração média de 210 ng/μL. As PCRs feitas com essa mesma amostra, para confirmar a presença do vetor + inserto, apresentaram o fragmento único esperado em 95 % das reações. As demais apresentaram mais de um fragmento ou nenhum produto de amplificação (Figura 17).

Figura 17 – Eletroforese em gel de agarose mostrando 5 μL do produto da *miniprep* (A) e os fragmentos gerados pela PCR usando os produtos da *miniprep* e o *primer* M13 (B), para 24 clones da biblioteca *forward*. M, Marcador de peso molecular 100 pb (Invitrogen)

De acordo com os resultados apresentados, a clonagem dos fragmentos de cDNAs das duas bibliotecas apresentou um alto rendimento e os produtos da extração dos plasmídeos apresentaram qualidade e concentrações adequadas para o sequenciamento dos clones. As PCRs e a eletroforese capilar para o sequenciamento dos clones também apresentaram alta qualidade e rendimento, conforme será apresentado no item de processamento das sequências, a seguir.

4.4 Processamento das sequências

O primeiro processamento das sequências das bibliotecas, usando o programa computacional *CodonCode Aligner*, mostrou um alto aproveitamento das sequências, e estas apresentaram boa qualidade, já que os eletroferogramas mostraram picos altos e definidos, com valores de qualidade *Phred* acima de 20 (Figura 18).

Figura 18 – Interface gráfica do programa computacional *CodonCode Aligner*, mostrando os eletroferogramas de nove sequências da biblioteca *reverse*

Para a biblioteca *forward*, dos 960 clones sequenciados, aproveitaram-se 92 % das sequências. Da montagem destas 881 sequências, obteve-se a formação de 100 *contigs*, sendo que 70 deles foram formados por apenas duas sequências

cada, 23 deles por 3 a 11 sequências e apenas três *contigs* foram formados por mais de 20 sequências.

Para a biblioteca *reverse*, foram sequenciados 384 clones e obtiveram-se 97 % das sequências aproveitadas. Após a montagem das 373 sequências, formaram-se 41 *contigs*, dos quais 31 foram formados por apenas duas sequências cada e os outros 10 *contigs* foram formados por até 7 sequências (Figura 19).

Na tabela 2 estão sumarizados os resultados obtidos neste primeiro processamento das sequências das bibliotecas *forward* e *reverse*.

Tabela 2 – Resultado do processamento das sequências das bibliotecas *forward* e *reverse*, usando o programa computacional *CodonCode Aligner*

	biblioteca <i>forward</i>	biblioteca <i>reverse</i>
Nº de sequências aproveitadas (x)	881	373
Nº de sequências únicas	572	269
Nº de sequências que formaram <i>contigs</i>	309	104
Nº de <i>contigs</i>	100	41
Nº de unisequências para anotação (y)	672	310
Redundância da biblioteca $100 \times [1 - (y/x)]$	23,7 %	16,9 %
Tamanho médio das unisequências em pares de bases e respectiva amplitude	562 (114 a 1418)	607 (128 a 1070)

Figura 19 – Interface gráfica do programa computacional *CodonCode Aligner*, mostrando um *contig* formado por quatro sequências da biblioteca *reverse*

No segundo processamento das sequências, usando, agora, o *pipeline* de processamento de pré-anotação, e analisando o conjunto das 1.344 sequências das duas bibliotecas, obtiveram-se 1.273 (94,7 %) sequências de boa qualidade, de acordo com os parâmetros estabelecidos, ou seja, com tamanho > 100 pb e *Phred* > 20. O tamanho das sequências antes da montagem dos *contigs* variou de 100 a 690 pb, com média de 492 pb (Figura 20).

A proporção de GC observado nestas sequências foi de 45,56 %. Santos (2013), ao sequenciar *BAC-ends* de uma biblioteca genômica em BACs, desenvolvida também para o genótipo 'IAPAR123' de *P. edulis* f. *flavicarpa*, observou uma proporção de GC de 41,92 % nas sequências totais (gênicas e genômicas). Esta diferença observada na porcentagem de GC é esperada, já que sequências expressas apresentam um maior teor de GC em contraste com o seu conteúdo total no genoma (MADIGAN et al. 2003).

Figura 20 – Ilustração gráfica mostrando a saída de sete sequências após a análise pelo *pipeline* de pré-anotação. Cada barra em verde representa uma base, o eixo 'y' indica o valor de qualidade de *Phred* e o eixo 'x' indica o número de bases da sequência

A montagem das 1.273 sequências resultou em 857 sequências únicas e em 141 *contigs*, formados por 416 sequências, sendo que cada *contig* foi formado por 2 até 19 sequências. Portanto, a análise em conjunto das bibliotecas apresentou 998 (857+141) unisequências passíveis de anotação, com tamanho médio de 508,1 pb, e uma redundância de 21,6 %.

Destas 998 unisequências, 683 são provenientes da biblioteca *forward*, 274 da biblioteca *reverse* e 41 unisequências são oriundas de 41 *contigs* formados por sequências provenientes de ambas as bibliotecas.

4.5 Anotação funcional das bibliotecas subtrativas

4.5.1 Anotação das bibliotecas com base na Plataforma PLAZA

Biblioteca *forward*

A anotação das 672 unisequências da biblioteca *forward*, feita com base na Plataforma PLAZA, mostrou que estas apresentaram similaridade com 525 sequências proteicas de *A. thaliana*. As 525 proteínas resultantes foram atribuídas a

15 categorias funcionais, tendo por base as informações obtidas nesta anotação e as informações depositadas em bancos de dados específicos como, por exemplo, o UniProt, e na literatura científica (Figura 21).

Figura 21 – Número de proteínas atribuídas a cada categoria funcional, resultante da anotação das sequências da biblioteca *forward* na Plataforma PLAZA

A categorização das proteínas apresentada na figura 21 se deu mais por uma questão de organização das informações, uma vez que muitos produtos gênicos podem exercer diferentes funções, dependendo da circunstância. Por outro lado, a observação de diversas categorias funcionais corrobora o fato de que o processo de defesa contra patógenos é complexo e envolve a atuação conjunta de diversos mecanismos. Assim, observaram-se proteínas que possivelmente atuam durante a resposta de defesa vegetal, como as proteínas envolvidas nas categorias ‘mecanismos de defesa’, ‘percepção e transdução de sinais’, ‘transporte’ e ‘detoxificação’, como também proteínas envolvidas na manutenção celular e no desenvolvimento vegetal.

A partir desta anotação, foram selecionadas 73 proteínas que tem algum envolvimento na resposta de defesa de plantas (AGRIOS, 2005) (Tabela 3). Pode-se identificar proteínas envolvidas no desencadeamento de mecanismos de defesa não específicos, quais sejam, duas proteínas da cascata de sinalização da MAP quinase e diversas proteínas envolvidas em sinalização intracelular, como proteínas quinases (22), ATPases (18), fosforilases (1) e fosfolipases (2). Foram também identificadas sete proteínas relacionadas com a ativação e a regulação da transcrição de genes de defesa e três proteínas de defesa/resistência, bem como cinco enzimas oxidoredutases, duas proteínas envolvidas em morte celular e 11 proteínas relacionadas à patogênese (PR) (Figura 22).

Estas 73 proteínas também foram analisadas usando o banco de dados STRING, o que permitiu identificar 27 interações proteicas, representadas por setas pretas na figura 22, com base nas interações de proteínas de *A. thaliana*.

Tabela 3 - Proteínas relacionadas à resposta de defesa vegetal, selecionadas a partir da anotação da biblioteca *forward* com base na Plataforma PLAZA

			(continua)
Unisequência	Nº de acesso	Proteína	E-value
<i>Transdução de sinal intracelular</i>			
F-P2E05	AT4G01370	Mitogen-activated protein kinase 4 (MPK4)	2,00E-61
F-P6G10	AT4G29810	Mitogen-activated protein kinase kinase 2 (MK1)	2,00E-63
F-P3C04	AT4G33430	Brassinosteroid Insensitive 1 associated receptor kinase 1	2,00E-15
F-P9F07	AT5G03730	Serine/threonine-protein kinase CTR1	2,00E-06
F-P10A04	AT1G53440	Probable LRR receptor-like serine/threonine-protein kinase	2,00E-14
F-P6D01	AT3G23750	Leucine-rich repeat receptor-like protein kinase	2,00E-96
F-P7A04	AT5G11020	Ser/thr specific protein kinase-like protein	3,00E-53
F-P9B08	AT2G02800	Protein kinase 2B, chloroplastic	7,00E-91
F-P9C06	AT1G25390	Protein kinase family protein	3,00E-18
F-P9H12	AT2G25760	Protein kinase family protein	9,00E-93
F-P10A10	AT3G03940	Protein kinase family protein	1,00E-09
F-P10C11	AT1G54610	Protein kinase family protein	1,00E-40
F-P9B10	AT1G03920	Protein kinase family protein	1,00E-41
F-P2C12	AT4G19110	Protein kinase family protein	5,00E-27
F-P8H01	AT3G62220	Serine/threonine protein kinase	9,00E-22
F-P5G10	AT5G49660	Leucine-rich repeat receptor-like protein kinase	5,00E-59
F-Contig91	AT1G16670	Protein kinase family protein	6,00E-32
F-P5E01	AT2G44830	Protein kinase family protein	3,00E-47
F-P8E05	AT4G34100	Serine/threonine-protein kinase BRI1-like 2	2,00E-62
F-P6C09	AT3G58760	Protein kinase family protein	8,00E-71
F-P5G12	AT5G63710	Leucine-rich repeat receptor-like protein kinase	1,00E-50
F-P9G05	AT1G50700	Calcium-dependent protein kinase 33	1,00E-96
F-P8E06	AT3G10660	Calcium-dependent protein kinase 2	3,00E-19
F-P9G11	AT1G78290	Serine/threonine-protein kinase SRK2C	1,00E-18
F-P8H08	AT1G10130	Calcium-transporting ATPase 3, endoplasmic reticulum-type	1,00E-71
F-P7B10	AT3G05530	RPT5a ATPase/ calmodulin binding	2,00E-94
F-P3F06	AT2G24520	ATPase 5, plasma membrane-type	4,00E-36
F-P8D03	AT4G30190	ATPase 2, plasma membrane-type	1,00E-33
F-Contig9	ATCG00130	ATP synthase subunit b, chloroplastic	9,00E-71
F-P3D12	AT4G28070	AFG1-like ATPase family protein	4,00E-85
F-Contig71	AT4G38510	V-ATPase subunit B2	5,00E-102
F-P8D09	AT2G36910	ABC transporter B family member 1	5,00E-86
F-P1H10	AT4G25450	ABC transporter B family member 28	2,00E-56
F-P3B12	AT5G39040	ABC transporter B family member 27	8,00E-61
F-P4A12	AT3G21090	ABC transporter G family member 15	1,00E-72
F-P3C01	AT3G28345	ABC transporter B family member 15	6,00E-35
F-P6A03	AT1G27940	ABC transporter B family member 13	2,00E-57
F-P5A12	AT1G15210	ABC transporter G family member 35	3,00E-97
F-P9G01	AT4G33460	ABC transporter I family member 10, chloroplastic	3,00E-32
F-P9A05	AT5G64840	ABC transporter F family member 5	1,00E-86

Tabela 3 - Proteínas relacionadas à resposta de defesa vegetal, selecionadas a partir da anotação da biblioteca *forward* com base na Plataforma PLAZA

Unisequência	Nº de acesso	Proteína	(conclusão)
			<i>E-value</i>
F-P8E09	AT3G10670	ABC transporter I family member 6, chloroplastic	7,00E-41
F-P6H04	AT1G02890	AAA-type ATPase family protein	4,00E-28
F-P9B11	AT5G09650	Pyrophosphatase 1, chloroplastic	1,00E-51
F-P3G09	AT3G15730	Phospholipase D alpha 1	2,00E-88
F-P5A05	AT4G38690	1-Phosphatidylinositol phosphodiesterase-related	3,00E-73
<i>Ativação de gene de defesa / regulação da transcrição</i>			
F-P7H10	AT2G38470	WRKY transcription factor 33	7,00E-30
F-Contig80	AT4G16150	Calmodulin-binding transcription activator 5	5,00E-25
F-P10A05	AT1G53910	Ethylene-responsive transcription factor RAP2-12	2,00E-18
F-P6B07	AT1G50640	Ethylene-responsive transcription factor 3	9,00E-06
F-P4G06	AT4G23100	Glutamate-cysteine ligase, chloroplastic	5,00E-98
F-P10F04	AT1G61560	MLO-like protein 6	1,00E-30
F-P2F11	AT2G25110	Stromal cell-derived factor 2-like protein (SDF2)	2,00E-50
<i>Proteína de defesa / resistência</i>			
F-P10H08	AT3G61220	(+)-neomenthol dehydrogenase	2,00E-34
F-P4C05	AT4G27220	Probable disease resistance protein At4g27220	5,00E-13
F-P4E05	AT1G27170	Disease resistance protein	3,00E-20
<i>Oxidoreductase</i>			
F-P9H04	AT4G35090	Catalase-2	2,00E-82
F-P7C03	AT4G15760	Monooxygenase	1,00E-31
F-P6G12	AT2G35660	Monooxygenase	5,00E-49
F-Contig46	AT2G45570	Cytochrome P450	3,00E-34
F-P2E09	AT1G13710	Cytochrome P450	8,00E-33
<i>Morte celular</i>			
F-P1F04	AT1G29690	Constitutively activated cell death 1 (CAD1)	3,00E-41
F-Contig33	AT3G01280	Voltage-dependent anion channel 1 (VDAC)	1,00E-111
<i>Proteína relacionada à patogênese</i>			
F-P4D09	AT3G16920	Chitinase	2,00E-63
F-Contig2	AT3G12500	Class I chitinase	9,00E-107
F-P4C03	AT2G40880	Cysteine proteinase inhibitor 3	1,00E-13
F-P7G06	AT3G45140	Lipoxygenase 2, chloroplastic	4,00E-64
F-P6H11	AT2G25080	Glutathione peroxidase 1, chloroplastic	2,00E-25
F-Contig61	AT4G21960	Peroxidase 42	3,00E-60
F-P9E12	AT1G71695	Peroxidase 12	5,00E-53
F-P9D06	AT5G05340	Peroxidase 52	6,00E-53
F-Contig27	AT3G04720	Hevein-like protein (PR4)	4,00E-49
F-P5E02	AT1G70850	MLP-like protein 34 (PR10)	6,00E-19
F-P7F09	AT4G35000	L-ascorbate peroxidase 3, peroxisomal	3,00E-47

Figura 22 - Representação esquemática da célula vegetal, mostrando as proteínas possivelmente relacionadas à resposta de defesa à bactéria (*Xap*). As cores dos nomes das proteínas se referem às respectivas categorias destacadas em retângulos, com exceção daquelas em preto, que mostraram interação com as demais proteínas. As setas pretas se referem às interações proteicas sugeridas pelo banco de dados STRING. Os círculos amarelos representam os efetores bacterianos

As 147 sequências que não apresentaram similaridade à *A. thaliana* foram comparadas com aquelas do banco de dados do NCBI, usando a ferramenta BlastX. Desta análise, 42 sequências apresentaram similaridade a proteínas com funções conhecidas ou proteínas preditas de outras espécies vegetais, sendo que dos *top hits*, 12 sequências foram similares a proteínas de *Ricinus communis* e 10 a proteínas de *Populus trichocarpa*.

Porém, para 105 sequências não se obteve qualquer resultado ou poucos resultados significativos foram obtidos com a ferramenta BlastN. Quatro sequências apresentaram similaridades a RNAs mensageiros, duas a DNA mitocondrial, duas a DNA cloroplastidial, uma a DNA plastidial e uma a RNA transportador.

Essas 105 sequências também foram traduzidas com o auxílio da ferramenta *Translate* do *ExPASy* (<http://web.expasy.org/translate/>) para identificar possíveis quadros abertos de leitura (ORF). Dos seis possíveis *frames* de leitura, 98 sequências apresentaram todos os *frames* interrompidos por códons de terminação e sete sequências apresentaram um único *frame* sem interrupção, caracterizando uma região codante. Estes resultados sugerem que estas 98 sequências sejam regiões gênicas regulatórias, que não são traduzidas, 5' ou 3' UTR.

Biblioteca *reverse*

A anotação das 310 unisequências da biblioteca *reverse*, feita com base na Plataforma PLAZA, mostrou que estas apresentaram similaridade com 239 sequências proteicas de *A. thaliana*. As proteínas identificadas nesta anotação estão relacionadas a grande parte das categorias da figura 21, dentre elas, 'mecanismos de defesa', 'percepção e transdução de sinais' e 'transporte'. Ou seja, também foram identificados transcritos que codificam diferentes tipos de proteínas, incluindo os relacionados à resposta de defesa, como, por exemplo, transcritos que codificam proteína quitinase, proteína com domínio CC-NBS-LRR, proteína de resposta à auxina, dentre outras. Porém, decidiu-se não atribuir os transcritos anotados desta biblioteca a categorias funcionais.

Assim, sugere-se que os transcritos diferencialmente expressos obtidos na biblioteca *reverse* sejam normalmente expressos na planta sadia; já na presença do

patógeno, esses transcritos tiveram a expressão suprimida, podendo, alguns deles, serem alvos do patógeno.

4.5.2 Anotação das bibliotecas com base no programa computacional Blast2GO

Em uma nova análise, as 998 unisequências, oriundas do processamento em conjunto das sequências das duas bibliotecas, foram anotadas usando o programa computacional Blast2GO. Neste programa usou-se a ferramenta BlastX para encontrar sequências similares às unisequências das bibliotecas. Em seguida, fez-se o mapeamento, que é a recuperação dos termos do GO associados aos *hits* obtidos após a pesquisa feita pelo BlastX, e a anotação, que é a seleção dos termos do GO a partir do conjunto dos termos obtidos no passo de mapeamento.

Nesta análise, obtiveram-se os seguintes resultados: 760 unisequências foram mapeadas e anotadas; 68 foram mapeadas, mas não anotadas; 38 tiveram identificação pelo BlastX, mas não foram mapeadas e anotadas; 132 não tiveram identificação pelo BlastX. Os resultados do BlastX estão apresentados no Anexo A. Também foi possível identificar as espécies vegetais que apareceram com maior frequência entre os *top-hits* (ocorrência mais relevante por alinhamento efetuado), bem como a distribuição dos níveis de similaridade entre os alinhamentos (*hits*) (Figura 23).

Figura 23 – Número de *top-hits* por espécie vegetal (A) e a similaridade (em %) obtida em todos os alinhamentos (*hits*) (B) durante a busca por sequências similares realizada com a ferramenta Blastx

Das 760 sequências mapeadas e anotadas, todas foram associadas a pelo menos um termo do *GO* (ocorrências), havendo sequência associada a até 30

termos do *GO*, sendo estas ocorrências distribuídas em diversos níveis de ontologia. Todas as ocorrências foram, primeiramente, distribuídas nos termos ontológicos do nível hierárquico '1', que é o mais abrangente e contém os termos basais 'processos biológicos', 'funções moleculares' e 'componentes celulares' (Figura 24). Na anotação com base nos termos do *Gene Ontology*, uma única proteína pode estar associada a mais de um termo ontológico do mesmo nível, uma vez que uma proteína pode atuar em vários processos biológicos, com diferentes funções e localizações.

Figura 24 – Distribuição das ocorrências nos termos ontológicos de nível '1' das 760 unisequências mapeadas e anotadas segundo o programa computacional Blast2GO

Separando por biblioteca, das 683 unisequências da biblioteca *forward*, 498 (73%) foram mapeadas e anotadas, segundo a análise feita com base no Blast2GO. Das 274 unisequências da biblioteca *reverse*, 231 (84%) foram mapeadas e anotadas. Já das 41 unisequências oriundas de transcritos presentes nas duas bibliotecas, 31 foram mapeadas e anotadas. A distribuição das ocorrências nos termos ontológicos de nível '2' está apresentada nos gráficos da figura 25, levando em conta as três categorias (termos de nível '1') 'processos biológicos', 'funções moleculares' e 'componentes celulares'. A distribuição das ocorrências nos termos dos demais níveis está apresentada na tabela 4.

Figura 25 – Distribuição das ocorrências nos termos ontológicos de nível '2', levando em conta as três categorias 'processos biológicos', 'funções moleculares' e 'componentes celulares', entre as unisequências mapeadas e anotadas da biblioteca *forward* (A), da biblioteca *reverse* (B) e das unisequências oriundas das duas bibliotecas (C), segundo o programa computacional Blast2GO

Tabela 4 - Distribuição das ocorrências nos termos ontológicos dos níveis de '3' a '8', levando em conta as três categorias 'processos biológicos', 'funções moleculares' e 'componentes celulares', entre as unisequências mapeadas e anotadas da biblioteca forward (A), da biblioteca *reverse* (B) e das unisequências oriundas das duas bibliotecas (C), segundo o programa computacional Blast2GO

(continua)

Nível hierárquico	Termo ontológico	A	B	C
<i>Processos Biológicos</i>				
3	regulation of biological quality	16	5	0
3	cellular developmental process	5	1	1
3	catabolic process	100	37	5
3	establishment of localization	0	0	7
3	regulation of biological process	41	18	0
3	primary metabolic process	242	99	13
3	cell growth	5	1	0
3	cell cycle	3	3	0
3	anatomical structure development	21	9	2
3	cell death	2	2	0
3	response to endogenous stimulus	29	9	0
3	response to abiotic stimulus	37	18	3
3	biosynthetic process	129	51	9
3	secondary metabolic process	19	7	10
3	cellular homeostasis	12	4	0
3	macromolecule metabolic process	123	55	5
3	tropism	2	1	0
3	nitrogen compound metabolic process	72	29	4
3	establishment of localization	62	27	0
3	multicellular organismal development	49	17	4
3	cell communication	6	3	0
3	reproductive process	13	6	1
3	cellular component organization	36	17	1
3	cellular metabolic process	177	73	0
3	cellular response to stimulus	39	16	0
3	response to stress	56	30	2
3	response to biotic stimulus	21	12	0
3	response to external stimulus	13	8	0
4	transport	62	27	7
4	lipid metabolic process	41	14	2
4	cellular biosynthetic process	37	14	2
4	cellular macromolecule metabolic process	97	36	4
4	regulation of metabolic process	3	3	0
4	generation of precursor metabolites and energy	34	17	5
4	regulation of cellular component size	5	1	0
4	carbohydrate metabolic process	52	26	3
4	homeostatic process	12	4	0
4	cellular nitrogen compound metabolic process	72	29	4
4	regulation of anatomical structure size	5	1	0

Tabela 4 - Distribuição das ocorrências nos termos ontológicos dos níveis de '3' a '8', levando em conta as três categorias 'processos biológicos', 'funções moleculares' e 'componentes celulares', entre as unisequências mapeadas e anotadas da biblioteca forward (A), da biblioteca *reverse* (B) e das unisequências oriundas das duas bibliotecas (C), segundo o programa computacional Blast2GO

(continuação)				
Nível hierárquico	Termo ontológico	A	B	C
4	protein metabolic process	115	50	5
4	macromolecule biosynthetic process	37	14	2
4	response to extracellular stimulus	5	2	0
4	cell differentiation	5	1	1
4	post-embryonic development	31	10	3
4	regulation of cellular process	39	16	0
4	signal transduction	39	16	4
4	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	72	29	5
4	photosynthesis	16	6	1
4	anatomical structure morphogenesis	12	5	1
4	developmental process involved in reproduction	12	6	2
4	gene expression	39	17	2
4	macromolecule modification	55	19	0
4	embryo development	14	3	4
5	cellular protein metabolic process	90	33	2
5	cellular macromolecule biosynthetic process	37	14	0
5	nucleic acid metabolic process	8	3	0
5	regulation of macromolecule metabolic process	3	3	0
5	regulation of cell size	5	1	1
5	reproductive structure development	12	6	2
6	translation	37	14	0
6	DNA metabolic process	8	3	0
6	regulation of gene expression	3	3	1
6	flower development	12	6	2
6	protein modification process	55	19	1
7	regulation of gene expression, epigenetic	3	3	0
<i>Função molecular</i>				
3	sequence-specific DNA binding transcription factor activity	17	13	2
3	signal transducer activity	16	7	5
3	lipid binding	3	4	0
3	nucleic acid binding	79	29	0
3	chromatin binding	1	1	3
3	carbohydrate binding	9	4	4
3	protein binding	80	39	3
3	hydrolase activity	112	39	0
3	transferase activity	93	47	0
3	oxygen binding	1	0	0
4	hydrolase activity, acting on acid anhydrides	2	2	1
4	hydrolase activity, acting on ester bonds	7	16	2

Tabela 4 - Distribuição das ocorrências nos termos ontológicos dos níveis de '3' a '8', levando em conta as três categorias 'processos biológicos', 'funções moleculares' e 'componentes celulares', entre as unisequências mapeadas e anotadas da biblioteca forward (A), da biblioteca *reverse* (B) e das unisequências oriundas das duas bibliotecas (C), segundo o programa computacional Blast2GO

(continuação)

Nível hierárquico	Termo ontológico	A	B	C
4	DNA binding	29	14	0
4	RNA binding	35	12	1
4	transferase activity, transferring phosphorus-containing groups	33	1	0
4	receptor binding	1	2	6
4	nucleotide binding	114	56	1
5	hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides	2	1	1
5	translation factor activity, nucleic acid binding	10	4	0
5	kinase activity	33	16	0
5	nuclease activity	7	2	0
6	pyrophosphatase activity	2	1	0
7	nucleoside-triphosphatase activity	2	1	0
8	motor activity	2	1	0
<i>Componente celular</i>				
3	membrane-bounded organelle	267	119	19
3	non-membrane-bounded organelle	47	27	3
3	organelle part	33	16	2
3	cell part	373	166	26
4	intracellular	334	145	23
4	plasma membrane	45	22	0
4	external encapsulating structure	17	7	1
4	organelle lumen	32	16	2
4	intracellular part	315	142	23
4	intracellular organelle part	33	16	2
4	cell periphery	60	27	1
5	cell wall	16	7	1
5	cytoplasm	271	124	18
5	thylakoid	37	9	6
5	ribonucleoprotein complex	35	17	2
5	intracellular organelle	281	130	22
5	cytoplasmic part	233	108	17
6	nucleoplasm	23	13	2
6	nucleolus	9	4	0
6	mitochondrion	54	23	2
6	vacuole	34	12	2
6	endoplasmic reticulum	13	4	0
6	Golgi apparatus	8	2	0
6	cytosol	20	11	0
6	ribosome	35	17	2
6	plastid	140	68	11
6	microbody	16	6	1

Tabela 4 - Distribuição das ocorrências nos termos ontológicos dos níveis de '3' a '8', levando em conta as três categorias 'processos biológicos', 'funções moleculares' e 'componentes celulares', entre as unisequências mapeadas e anotadas da biblioteca *forward* (A), da biblioteca *reverse* (B) e das unisequências oriundas das duas bibliotecas (C), segundo o programa computacional Blast2GO

(conclusão)				
Nível hierárquico	Termo ontológico	A	B	C
6	intracellular membrane-bounded organelle	267	119	19
6	intracellular non-membrane-bounded organelle	47	27	3
6	nuclear part	33	16	2
6	intracellular organelle lumen	32	16	2
7	nucleus	71	33	4
7	peroxisome	16	6	1
7	cytoskeleton	7	6	1
7	nuclear lumen	32	16	2

Para a categoria 'processos biológicos', os termos ontológicos de nível '2' mais representados nas três análises foram 'cellular process' e 'metabolic process'. Estes são termos mais amplos, pois, de acordo com as definições disponibilizadas pelo *Gene Ontology Consortium*, as proteínas relacionadas participam de algum processo biológico em nível celular e de algum processo metabólico, seja ele catabólico ou anabólico, esperando-se assim, serem esses termos os mais representados.

O terceiro termo mais representado é o 'response to stimulus', o qual engloba proteínas que detectam estímulos que resultam na mudança de estado ou de atividade de uma célula ou organismo. Juntando-se os termos 'response to stimulus' e 'signaling', estão as proteínas envolvidas na detecção da invasão de um patógeno e na transdução de sinais intracelulares que desencadeiam as respostas de defesa vegetal.

Porém, o nível '2' ainda é abrangente e termos mais específicos relacionados à resposta de defesa vegetal, para a categoria 'processos biológicos', só foram encontrados a partir do nível '3'. Neste nível, duas proteínas de cada biblioteca foram associadas ao termo 'cell death' e 21 da biblioteca *forward* e 12 da *reverse* foram associadas ao termo 'response to biotic stimulus'. Do nível '4', foram associadas 39 e 16 proteínas da biblioteca *forward* e *reverse*, respectivamente, ao termo 'signal transduction' (Tabela 4).

Para a categoria ‘funções moleculares’ os termos ontológicos de nível ‘2’ mais representados nas três análises foram ‘binding’ e ‘catalytic activity’. Estes termos também são amplos; por isso, é esperado serem os termos mais representados neste nível hierárquico. Proteínas associadas ao termo ‘binding’ facilitam interações de moléculas, de forma não covalente e em regiões específicas. Já proteínas associadas ao termo ‘catalytic activity’, catalisam diversos tipos de reações bioquímicas a temperaturas fisiológicas.

Já para a categoria ‘componentes celulares’, os termos de nível ‘2’ ‘cell’, ‘membrane’ e ‘organelle’ foram os mais representados nas três análises. Estes termos representam as localizações menos específicas nas células e as proteínas relacionadas a esses termos atuam no ambiente celular (‘cell’), na membrana plasmática, especificamente (‘membrane’) ou no núcleo, mitocôndrias, plastos, vacúolos, vesículas, ribossomos e citoesqueleto (‘organelle’).

Ao comparar os gráficos de termos de nível ‘2’ da biblioteca *forward* e *reverse* (Figura 25A e 25B), observa-se que a maioria dos termos associados às proteínas da biblioteca *forward* está presente na *reverse*. Também, estes termos são relacionados à maioria das atividades celulares, confirmando que qualquer atividade realizada pela célula, como, por exemplo, a resposta de defesa, envolve toda a sua complexa e coordenada maquinaria. Os gráficos da figura 25C apresentam um número menor de termos, uma vez que esta análise envolveu poucas unisequências, porém, todos os termos dos gráficos da figura 25C também estão presentes nos gráficos das figuras 25A e 25B.

4.6 Validação das bibliotecas subtrativas

4.6.1 Seleção de sequências putativas codificadoras das bibliotecas e desenho de primers

A partir da anotação e da categorização funcional das duas bibliotecas, foram selecionadas 100 sequências atribuídas a diversas categorias funcionais, conforme apresentado nas figuras 21 e 25, sendo que 30 % delas foram selecionadas a partir das categorias relacionadas à resposta de defesa em plantas. Do total das sequências selecionadas (100), foi possível desenhar 74 pares de *primers*, segundo

os parâmetros pré-estabelecidos para a qPCR. Desses, 42 foram desenhados para amplificar sequências oriundas da biblioteca *forward* e 32 da *reverse*.

O teste de amplificação com esses *primers*, por PCR convencional, utilizando como *template* cDNA e DNA genômico de maracujá azedo ('IAPAR123'), mostrou que 69 pares de *primers* produziram um amplicon único, de tamanho esperado, quando se usou cDNA como *template*. Porém, em alguns casos, quando se usou DNA genômico como *template*, os fragmentos foram maiores que o esperado, ou apresentaram mais de um fragmento, isto devido à presença de íntrons (Figura 26). Assim, nesta etapa, obteve-se um aproveitamento de 93 % dos pares de *primers*, evidenciando-se a alta qualidade dos *primers* desenhados.

Na tabela 5 estão apresentadas as sequências dos 69 pares de *primers* que produziram o amplicon predito, tendo cDNA como *template*, e o nome da provável proteína codificada pelo transcrito cuja sequência foi usada para desenhar o respectivo par de *primers*.

Figura 26 – Eletroforese em gel de agarose mostrando os produtos das reações de amplificação com cinco pares de *primers* (tabela 5), visando à confirmação dos amplicons preditos. Canaletas 1, *Template* de cDNA de maracujá azedo; Canaletas 2, *Template* de DNA genômico de maracujá azedo; Canaletas 3, Controle, sem *template*. M, Marcador de peso molecular 100 pb (Invitrogen)

Tabela 5 – Sequências nucleotídicas dos *primers* desenhados para a validação das bibliotecas SSH e respectivas proteínas codificadas pelos transcritos, cujas sequências foram usadas para desenhar os pares de *primers*. As letras F e R no início do *primer* indicam que ele foi desenhado a partir de uma sequência da biblioteca *forward* ou *reverse*, respectivamente

(continua)

Cód.	Primer	Sequências F (1ª linha) e R (2ª linha) (5'→3')	Tamanho do amplicon	Proteína
1	F-P3C4	CGAGGAGAGGAGAACTACC GCATCGCCTTCTTCATTACC	186 pb	Brassinosteroid insensitive 1-associated receptor kinase 1
2	F-Contig20	TGACGAAATCCAGAGTAAGAC ACAGTGAATGAGGTGGGCTC	153 pb	Oxygen-evolving enhancer protein 1-1, chloroplastic
3	F-P5G4	CCGACTGACTGTGATGGCT CTCACCTGCTCTTATCCTG	188 pb	Phosphoribulokinase, chloroplastic
4	F-P7H10	GACCTCTCAACTTTCTTCTTAG CCGCAATACCACCCTGTTAG	155 pb	WRKY transcription factor 33

Tabela 5 – Sequências nucleotídicas dos *primers* desenhados para a validação das bibliotecas SSH e respectivas proteínas codificadas pelos transcritos, cujas sequências foram usadas para desenhar os pares de *primers*. As letras F e R no início do *primer* indicam que ele foi desenhado a partir de uma sequência da biblioteca *forward* ou *reverse*, respectivamente

(continuação)

Cód.	Primer	Sequências F (1ª linha) e R (2ª linha) (5'→3')	Tamanho do amplicon	Proteína
5	F-P2F11	GTGGTCAGCAGTCCGTTAC CACTCTCAATCGCAACTC	108 pb	Stromal cell-derived factor 2-like protein
6	F-P9A6	TCTACGCCTGGGGTTACTG TGCCTTTTGGTTTGTATGGAC	161 pb	Lipoxygenase 2, chloroplastic
7	F-Contig33	GAAGCGATAACAGTGTACACC TTCTCAGGCGTTGTTGGCAG	164 pb	Mitochondrial outer membrane protein porin 1
8	F-P9F9	CGGAGATGAGGATGAAGATG ACCTGCTGCCTATGAGTGAG	165 pb	Cellulose synthase A catalytic subunit 1 [UDP-forming]
9	F-P4E5	CAAATGACTCTGGCTTCTTC CAATCCGTGTCTCAAAAACG	119 pb	Disease resistance protein / Tir-nbs-Irr
10	F-Contig2	CGTCTCTAGTATTCTCGCTG AGCACACGCACGACAGTTG	173 pb	Class I chitinase
11	F-P2E5	ACTGATAGAGAAAAACCGAC ATCATACGACCTCCCCAAAG	136 pb	Mitogen-activated protein kinase 4
12	F-P6G10	CATAGAAAGACTGGTAGCAC GAAGGGAAGTAGCGGAATTG	178 pb	Mitogen-activated protein kinase kinase 2
13	F-P9H4	CCTGTCTATTGTCCGTTTCTC TTCCCACAAGATCAAAGTTAC	121 pb	Catalase-2
14	F-P7E12	CTTGCTCTTGTGCTGATG TCTTCCACAGTATCCTAACG	109 pb	Protein Transport inhibitor response 1
15	F-P4E11	CATTGCTCTTCTCTCCTG GCTTTGATAGGTCTGGTG	133 pb	Universal stress protein (USP) family protein
16	F-Contig56	AGTGTAGTCTCAGCCAGTAG GCCAGCATTGACTCCTTTTC	161 pb	Glutamine synthetase, chloroplastic/ mitochondrial
17	F-P1A5	GGAGATATGGAAGTGCTAAAG ACAAGGAATGCCCGTTATGC	157 pb	Sulfate adenylyltransferase
18	F-P2C10	GACCCAAACCCAGGAAAATC GGCTCTCTCTTCTCTCTC	151 pb	SUMO-conjugating enzyme UBC9
19	F-P4D5	AGTTTCCAGCGATTTACAGAC GCCTCCAATCCAATAGTTCC	191 pb	F-box/LRR-repeat protein 4
20	F-Contig23	CTGTGAAGTGAACGCAACC TGATGCTATTGTGTCCCTC	99 pb	Peroxisomal fatty acid beta-oxidation protein
21	F-Contig19	GGAGTCTAACACCGTATGAG AAAATGTGTCCGAGCAAACC	140 pb	ATP citrate lyase
22	F-P6H3	CAGCAACAGATAAAGGGAAC GACTGAAAGACGAGGATGAC	119 pb	Myb family transcription factor
23	F-Contig21	GCTGCCACAAGTTAATGATG GCACGGTTCCTTTGAGTCC	150 pb	30S ribosomal protein 3-1, chloroplastic
24	F-Contig9	CCATTGATTTCTGCCGTTTC GCTAATGCTACAACCACTCC	212 pb	ATP synthase subunit b, chloroplastic
25	F-Contig11	CGTTGATGTTAATGCTGCTAC ATCGTATAAGCCTGCCCTC	157 pb	Hydroxypyruvate reductase
26	F-P9G5	GGACACGACAACAGTGG GTCAACATCAGCAGCCTCC	118 pb	Calcium-dependent protein kinase 33
27	F-P3H3	ATGGATGTGGGCTTCAAACG CTTGGTATTGTCAAGTATC	174 pb	Cation/calcium exchanger 3
28	F-P4A12	CCTCAATATGCTAACACACC ACAACACTCTTCTTCCAG	126 pb	ABC transporter G family member 15
29	F-P9H3	AGCGGTCAAGTGGATAACAAC GCTCTTCGGAATCTTCTTTG	170 pb	Alpha-L-arabinofuranosidase 1
30	F-P9E12	GTCTTGGTGGGGTCTTGG CCTGCTTGAACCTTTCTTG	142 pb	Peroxidase 12
31	F-Contig17	CGTATTATGGATCTTGGTTTCG GATCTACTACTACCTCGTCC	185 pb	Cytochrome c oxidase subunit 2
32	F-P1H11	CCTTCTGCTCTTCCCTTCG AATACAACCCAACCTTCTTCC	98 pb	Nucleoredoxin 1
33	F-Contig1	CTTGCCCTCCTTCTCTCTG GTGTCAATTTTACTGCCAACG	154 pb	Inibidores de tripsina e proteases – miraculin
34	F-Contig3	TTCTCTTCGCCTCCATTTTG TCCACTATCGTAAGGTTCTG	206 pb	Inibidores de tripsina e proteases – miraculin

Tabela 5 – Sequências nucleotídicas dos *primers* desenhados para a validação das bibliotecas SSH e respectivas proteínas codificadas pelos transcritos, cujas sequências foram usadas para desenhar os pares de *primers*. As letras F e R no início do *primer* indicam que ele foi desenhado a partir de uma sequência da biblioteca *forward* ou *reverse*, respectivamente

(continuação)

<i>Cód.</i>	<i>Primer</i>	<i>Sequências F (1ª linha) e R (2ª linha) (5'→3')</i>	<i>Tamanho do amplicon</i>	<i>Proteína</i>
35	F-P4G6	CATCAAACCTTGTGCGGAGG GTCTTTAAGTCCCCATTTAGG	114 pb	Glutamate-cysteine ligase, chloroplastic
36	F-P1F4	CTGGTTTATATTGATGAAGACC GATGCTTCTCTTGCCCTTAG	108 pb	CAD1 (constitutively activated cell death 1)
37	F-P10A5	TTATTTCTGACCAAGGGAGC CGAAGAGCAATTCATCAACC	126 pb	Ethylene-responsive transcription factor
38	F-P10H8	ATTGGTCTGGAGACTGTGC AACATTTCGTGAAGCCCATGC	123 pb	(+)-neomenthol dehydrogenase
39	R-P3B4	GTCATCTTGATATTGGGAACAC CGTCACATTTTGAAGGTTTCG	182 pb	CC-NBS-LRR resistance protein
40	R-P4G8	ACTGCTCACTTTGCCTACC CCCAGTTTCATCTTCCGTG	174 pb	Phosphatidylglycerolphosphate synthase 1
41	R-P2A12	GGGCTGGTTCTTGCTCTTC CCTCTTTGTGTGTCTCAATG	159 pb	Ferredoxin 1
42	R-P1B6	GCAAAATCCCAATATGACCACC CTTATGAATGCCCCCTCTAC	164 pb	GTP-binding protein
43	R-P1A5	CTCCCTACAACCTCCGTTTC CAATAATCCCAGTGTTCTCG	178 pb	Harpin-induced protein
44	R-P2G11	ATGGAGGCTGCTAATAACAAC GTATTTAGTTGGCGAACGAAG	206 pb	Heat shock transcription factor hsf1
45	R-P4F3	CTGTTTCGTCTTCCACCTTG GATGAGTATGTTTACCAACCTG	195 pb	Homeobox-leucine zipper protein athb-6
46	R-P3B6	CTTTTGGATCAGGCTGCTC CTTGGAAGACGCTCACTTC	177 pb	Kinesin-like protein
47	R-P2B8	CATACAACAATTTGGGAGGAAG CACATTTGATGGCGAACTGAG	189 pb	Leucine-rich repeat transmembrane protein kinase
48	R-P4D5	GTCGCTTATGACTTGCTCTG GTCAGTATCCACTCCTTG	180 pb	Lipoxygenase 2
49	R-P2F12	TGACTGCTGCTTCCACGAC GTAATAACTTGCCCTGCTGC	159 pb	Mitotic checkpoint protein
50	R-P3H11	CATCAGGCACAACCGAGTC TTACACCGATCATTTCTTGG	113 pb	Acyl-CoA oxidase 1
51	R-P3F2	GTAGTAGAGAGTGTTAGAGAG ACTAGCGGATAAGTGTCTGC	160 pb	20S proteasome beta subunit F1
52	R-Contig23	GGATGATTTGAGGTGTCTGC TACCCTGTTCTCCTGGTGC	142 pb	Leucine-rich repeat protein kinase family protein
53	R-P1D2	GTTTCGTGATGATGCCGCTG CCTCCCCAATTATACACACC	171 pb	Beta-amylase 1
54	R-P3C2	GGCAATTTATCGTCAGAGTTC ATCTCCCAATTTCTCGCATC	101 pb	Receptor protein kinase clavata 1
55	R-P4A10	GCCAAACGGACAATCATACAG GAGGTTAGAAGCAGAAGAGTC	153 pb	Ribonuclease t2
56	R-P2H10	CAGTAAAGGTGGTGGTAGG TTCTGGTAGAGGGCAAAAAC	166 pb	RNA helicase
57	R-Contig16	CAGCGGAGTAAGAGCGTTG ACGACAGGTTCCCCAGTAAG	193 pb	Root cap 1 (RCP1)
58	R-P3G12	GGAGAACCCAAACAGACC ATCCATCAGACGCATTAGC	174 pb	Shaggy-related protein kinase
59	R-P2F4	GGTTCATCTACACTTCATTCC CAATTATGCCACATACTTGAGC	114 pb	Stearoyl-acp desaturase
60	R-P3C1	TGATGGAGAGATGGAGAAGG CATTCCGATCTGTGGCACC	111 pb	Strubbelig-receptor family 7 protein
61	R-P4B7	CGTTCTTAATAGCAGTGAGG CAGCAGGATTGAAGTGTGG	187 pb	Superoxide dismutase
62	R-P4D8	GATTTTCCAACTCCACCTTC AGCGATTATGTTGTAGAGACC	102 pb	Ubiquitin-conjugating enzyme
63	R-P3F11	CGGGACGGAGATTCTGATTC CTCTTCTTCTCAATGAGGTAG	166 pb	Inorganic H pyrophosphatase family protein
64	R-P4B2	CTCGCTCCATTATCTGAAGG GAAACCAGTGTAGGGAATCTG	182 pb	Pheophorbide a oxygenase

Tabela 5 – Sequências nucleotídicas dos *primers* desenhados para a validação das bibliotecas SSH e respectivas proteínas codificadas pelos transcritos, cujas sequências foram usadas para desenhar os pares de *primers*. As letras F e R no início do *primer* indicam que ele foi desenhado a partir de uma sequência da biblioteca *forward* ou *reverse*, respectivamente (conclusão)

Cód.	Primer	Sequências F (1ª linha) e R (2ª linha) (5'→3')	Tamanho do amplicon	Proteína
65	R-P3H3	TTCTTGGTGGTGGCTATATC CACTCCCATTGTTTTCTTCG	106 pb	Receptor protein Crinkly 4
66	R-P4B8	AGCCTTACACGGTCCAAAC CTCTTGCTCCATTGTTCTTG	183 pb	Xyloglucan endotransglycosylase hydrolase
67	R-P4H8	ACAACCACTCACTCATACAC GCGATACCGTTCAGTTTGC	167 pb	Glyceraldehyde-3-phosphate dehydrogenase
68	R-P1G8	TGGGTGCTAATCTTTCTTCAG AACGCTCTGCTCAGTCTTG	191 pb	F-box family protein
69	R-Contig14	GCGACCAATATAGGAAAAATGC CGACTCATACAACCTGAAATGC	155 pb	Cysteine proteinase rd21a

4.6.2 Seleção e validação de genes de referência para a normalização da qPCR

A partir das sequências expressas de maracujá azedo, para os genes ERS (*Ethylene response sensor*), SUS (*Sucrose synthase*), ACCS (*1-aminocyclopropane-1-carboxylate synthase*), NDID (*NADP-dependent isocitrate dehydrogenase*) e ncpGS (*Glutamine synthetase GS2*), estudados por Pereira et al. (2013), e para o gene EF1a1 (*Translation elongation factor 1a-1* - GB nº DQ447160.1), foram desenhados *primers*, segundo os parâmetros pré-estabelecidos para a qPCR, com o intuito de servirem como normalizadores nos experimentos de análise de expressão relativa de genes.

As PCRs convencionais, usando os *primers* desenhados, produziram um fragmento único, do tamanho esperado, para os genes ERS, ACCS, NDID, ncpGS e EF1a1; porém, para o par de *primers* complementar ao gene SUS não foi observado nenhum produto (Figura 27).

Figura 27 – Eletroforese em gel de agarose mostrando os produtos das reações com os *primers* desenhados a partir das sequências expressas dos genes ERS, ACCS, ncpGS, NDID e EF1a1, e usando como template DNA genômico (canaletas b) ou cDNA (canaletas c) de maracujá azedo ou reação controle, sem *template* (canaletas a). M, Marcador de peso molecular 100 pb (Invitrogen)

As qPCRs, feitas com os *primers* desenhados a partir das sequências dos genes ERS, ACCS, NDID, ncpGS e EF1a1 e usando os cDNAs obtidos a partir dos transcritos do primeiro experimento de inoculação, resultaram em curvas de amplificação e eficiências adequadas, porém, nas análises das curvas de *Melt*, o produto da reação com o *primer* do gene ACCS apresentou um pico extra no decaimento da fluorescência, o que significa que o par de *primers* gerou produtos secundários, inviabilizando a utilização desse gene como normalizador (Figura 28).

Figura 28 – Curvas de amplificação (A) e curvas de *Melt* (B), geradas durante a qPCR usando o par de *primers* desenhado para amplificar um fragmento do gene ERS e cDNA de maracujá azedo

As reações com os *primers* dos genes ERS, NDID, ncpGS e EF1a1 apresentaram eficiências iguais a 2, ou seja, de 100 %, e os C_q médios das amostras para cada gene foram de 22,3, 21,0, 26,5 e 19,1, respectivamente. A partir das eficiências de cada gene e dos C_q médios das triplicadas técnicas de cada amostra, foram calculados os valores de Q , obtendo-se, assim, as quantidades relativas de expressão.

Os valores de Q foram usados para a análise da estabilidade desses genes, através do programa geNorm, obtendo-se os seguintes valores de M : 0,87 para o gene ncpGS, 0,49 para o gene ERS e 0,23 para os genes NDID e EF1a-1. O valor de ' M ' é definido como a variação aos pares de um gene, comparada com todos os outros genes da análise; assim, genes com valores de ' M ' menores apresentam expressões mais estáveis. Neste caso, a análise mostrou que os genes mais estáveis foram o ERS, NDID e o EF1a1.

Também, o programa calculou o número ótimo de genes normalizadores, baseando-se nos valores de expressão dos quatro genes analisados, obtendo-se

que o número ideal foi igual a 3, uma vez que a variação média apresentada por três genes em relação a quatro foi de 0,145, sendo que o valor máximo recomendado é de 0,15 (VANDESOMPELE et al., 2002).

A análise dos mesmos dados pelo programa NormFinder também mostrou que os genes com expressões mais estáveis foram o ERS, NDID e EF1a1 e que a melhor combinação entre dois genes foi entre o ERS e o EF1a-1. O gene ncpGS apresentou alta variação intragrupo e intergrupos.

Diante dos resultados apresentados pelos dois programas computacionais, foram escolhidos os genes ERS, NDID e EF1a1 para servirem de normalizadores nas qPCRs de validação das bibliotecas SSH para o primeiro experimento (primeira etapa de validação). As sequências dos *primers* para os três genes estão apresentadas na tabela 6.

Tabela 6 – Sequências nucleotídicas dos *primers* desenhados para a amplificação de fragmentos de genes normalizadores nas qPCRs

<i>Gene</i>	<i>Sequência dos primers 5'→3'</i>
ERS	F: GTATCTTGTGCTGTTGTGTC R: CCATCTCCCTGTCAAGTTC
NDID	F: GTCGTCACTCTCTCTTTACG R: TCATTTTCATCACCGTCCATC
EF1a1	F: GTTAAGGATTTGAAGCGTGG R: ATGTGTGATGTGTGGCAGT

As qPCRs, usando os cDNAs obtidos a partir dos transcritos do segundo experimento, foram feitas apenas com os *primers* que amplificam fragmentos dos genes ERS, NDID e EF1a1. Estas reações também resultaram em curvas de amplificação adequadas e eficiências médias por gene iguais a 100 %. A partir das eficiências de cada gene e dos C_q médios das triplicadas técnicas de cada amostra, foram calculados os Q , obtendo-se, assim, as quantidades relativas de expressão.

Os valores de Q foram usados para a análise da estabilidade desses genes, através do programa geNorm, obtendo-se os seguintes valores de M : 0,678, 0,731 e 0,982 para os genes EF1a1, ERS e NDID, respectivamente. O cálculo do número ótimo de genes normalizadores, baseando-se nos valores de expressão dos três

genes, resultou em uma variação média para dois genes em relação a três de 0,32, excedendo o valor máximo recomendado de 0,15.

O resultado da análise da estabilidade de expressão desses genes no programa NormFinder concordou com aquele obtido anteriormente (geNorm). Assim, optou-se por usar apenas o gene EF1a1 como normalizador na quantificação da expressão gênica, durante a segunda etapa de validação das bibliotecas. Isto devido ao fato de o gene EF1a1 ter apresentado maior estabilidade de expressão e a pouca estabilidade da combinação entre os genes testados.

Genes de referência são usados como controles internos para normalizar a expressão dos genes alvos, evitando-se assim erros nas análises de qPCR. Porém, havia um consenso em se usar genes envolvidos em processos celulares básicos, como rRNA 18S, actina, tubulina e GAPDH, sem analisar previamente sua estabilidade. Porém, esses genes, além de serem diferencialmente regulados em cada espécie vegetal, também são suscetíveis a alterações em um mesmo organismo. Por isso, um gene de referência não deve ser considerado universal (DIE et al., 2010).

O surgimento de algoritmos estatísticos para a análise da estabilidade de genes, implementados nos programas aqui utilizados (geNorm e NormFinder), trouxe mais confiabilidade para o uso de um gene de referência; da mesma forma, se tem sugerido o uso de mais de um normalizador.

Concluindo, a diferença na estabilidade dos três genes de maracujá azedo (EF1a1, ERS e NDID) na segunda condição experimental inviabilizou o uso de dois deles (ERS e NDID) como normalizadores.

4.6.3 Validação da expressão diferencial dos transcritos por qPCR

Primeira etapa da validação das bibliotecas:

Os RNAs extraídos das alíquotas preservadas dos fragmentos foliares do primeiro experimento apresentaram boa qualidade, ou seja, foi possível observar a presença de bandas definidas, sem sinal de degradação, e as duas bandas mais intensas, características de RNA ribossômico, 28S e 18S (Figura 29). As leituras das 18

amostras de RNA total, no NanoDrop, mostraram concentrações entre 160 e 440 ng/ μ L e razões A_{260}/A_{280} entre 1,90 e 2,04.

Figura 29 – Eletroforese em gel de agarose do RNA total extraído das folhas oriundas do primeiro experimento de inoculação. Canaletas de 1 a 9, triplicatas biológicas de cada tempo de coleta (6, 12 e 24 h) do tratamento com *Xap*. Canaletas de 10 a 18, triplicatas biológicas de cada tempo de coleta (6, 12 e 24 h) do tratamento controle. M, Marcador de peso molecular 100 pb (Invitrogen)

A síntese do cDNA em *bulk* das nove amostras de RNA total do tratamento com *Xap* e das nove amostras do tratamento controle, apresentou a qualidade esperada no teste de eficiência com os conjuntos de *primers* F-Contig33_3' e F-Contig33_5'. Os valores de ΔC_q ($C_q3'-C_q5'$) foram iguais a 0,80 e 0,83 relativamente aos tratamentos com *Xap* e controle. Também, no teste de amplificação com o par de *primers* que amplifica um fragmento do gene *NDID*, os produtos das PCRs com os cDNAs apresentaram um fragmento único de 598 pb, confirmando a ausência de contaminação por DNA genômico nos RNAs (Figura 30).

Figura 30 – Eletroforese em gel de agarose do produto da PCR com o par de *primers* NDID. Canaletas 1 e 2, *template* de cDNA do *bulk* do tratamento com *Xap* diluídos 8 x e 16 x, respectivamente; Canaletas 3 e 4, *template* de cDNA do *bulk* do tratamento controle diluídos 8 x e 16 x, respectivamente; Canaletas 5 e 6, *template* de DNA genômico de maracujá azedo; Canaleta 7, reação controle, sem *template*. M, Marcador de peso molecular 100 pb (Invitrogen)

Os resultados dos produtos das qPCRs para os 69 genes alvos e para os três genes endógenos foram registrados pelo programa *StepOne* v2.1 (ABI). Isto permitiu analisar a qualidade das curvas de amplificação, que se mostraram consistentes nas repetições técnicas, e as de *Melt*, que apresentaram um pico único e uniforme de decaimento da fluorescência, para cada gene analisado, mostrando que os produtos amplificados estavam adequados para a posterior análise de quantificação relativa da expressão (Figura 31).

Figura 31 – Curvas de amplificação (A) e Curvas de *Melt* (B), do produto das qPCRs com os *bulks* de cDNAs e com os pares de *primers* F-P3C4, F-P9A6, F-P9H4, EF1a1, ERS e NDID

Os dados brutos foram exportados e corrigidos usando o programa LinRegPCR, resultando em valores de eficiência média, por gene, acima de 1,9 e em valores de C_q para cada repetição técnica dos dois *bulks*. Esses valores foram inseridos no programa REST, para todos os genes alvos e endógenos, sendo que os C_q do *bulk* do tratamento com *Xap* foram inseridos na categoria '*Samples*' (tratado) e os C_q do *bulk* do tratamento controle foram inseridos na categoria '*Controls*' (controle). Assim, obteve-se uma razão de expressão dos genes alvos para o *bulk* de cDNAs do tratamento com *Xap* em relação ao tratamento controle, que está apresentada na figura 32.

Figura 32 – Razão da expressão (E) dos 69 genes alvos (1 a 69, tabela 5) para o *bulk* de cDNAs do tratamento com *Xap* em relação ao *bulk* de cDNAs do tratamento controle, considerando um nível de significância de 5 %. Em azul, indicam-se os genes mais expressos no tratamento com *Xap*, em vermelho, os genes mais expressos no tratamento controle e, em preto, são indicados os genes em que não houve diferença de expressão entre os tratamentos

Os genes 1 a 38 foram identificados na biblioteca *forward*; por isso, esperava-se que a maioria deles fosse mais expressa no tratamento com *Xap*. Os genes 39 a 69 são oriundos da biblioteca *reverse*, esperando-se, então, o inverso, ou seja, que houvesse maior expressão no tratamento controle.

Dos 38 genes da biblioteca *forward*, sete apresentaram maior expressão no tratamento com *Xap* (valores em azul, figura 32) e um gene foi mais expresso no tratamento controle (valor em vermelho, figura 32). Já dos 31 genes da biblioteca *reverse*, apenas três foram mais expressos no tratamento controle (valores em

vermelho, figura 32) e seis foram mais expressos no tratamento com *Xap* (valores em azul, figura 32).

Os genes da biblioteca *forward* com expressão significativamente maior no tratamento com *Xap* (valores em azul, figura 32) codificam as proteínas *Brassinosteroid insensitive 1-associated receptor kinase 1*, *Glutamine synthetase*, *30S ribosomal protein 3-1*, *Hydroxypyruvate reductase*, *Peroxidase 12* e *Cytochrome c oxidase subunit 2*, as quais estão envolvidas, respectivamente, na resposta de defesa à bactéria, síntese de aminoácido, biossíntese proteica, produção/conversão de energia e óxido-redução (as duas últimas). O gene com menor expressão no tratamento com *Xap* ($E = 0,83$) codifica a proteína *3-Phosphoribulokinase chloroplastic*, que está envolvida resposta de defesa à bactéria.

Neste caso, conforme o esperado, foi observado um gene de resposta de defesa mais expresso no tratamento com *Xap*. No entanto, outros diversos genes relacionados à resposta de defesa não apresentaram diferença na sua expressão e um, ao contrário do esperado, foi menos expresso.

Já os genes da biblioteca *reverse* com maior expressão no tratamento controle (valores em vermelho, figura 32) codificam as proteínas *Receptor protein kinase clavata 1*, *RNA helicase* e *F-box family protein*, que estão envolvidas na percepção e transdução de sinais, processamento de RNA ribossômico e ligação a carboidrato, respectivamente. Já os genes com menor expressão (valores em azul, figura 32) codificam as proteínas *CC-NBS-LRR resistance protein*, *GTP-binding protein*, *Leucine-rich repeat transmembrane protein kinase*, *Shaggy-related protein kinase*, *Heat shock transcription factor hsf1* e *Beta-amylase 1*, as quais estão envolvidas com resposta de defesa, transdução de sinais, resposta a estresse e metabolismo de carboidrato.

No caso da biblioteca *reverse*, os três genes mais expressos no tratamento controle sugerem que na planta inoculada com *Xap* estes genes tiveram sua expressão inibida pelo patógeno. E ao contrário do esperado, seis genes foram mais expressos no tratamento com *Xap*, sendo a maioria deles relacionados à resposta de defesa, o que era esperado para os genes da biblioteca *forward*.

Mesmo com pequenas diferenças, foi possível observar genes diferencialmente expressos entre os dois tratamentos, principalmente genes

relacionados à resposta de defesa, cuja expressão foi induzida durante a interação com *Xap*.

Segunda etapa da validação das bibliotecas:

Os RNAs extraídos dos fragmentos foliares do segundo experimento também apresentaram boa qualidade, ou seja, foi possível observar a presença de bandas definidas e as duas bandas mais intensas, características de RNA ribossômico, 28S e 18S (Figura 33). As leituras das 18 amostras de RNA total, no NanoDrop, mostraram concentrações entre 330 e 990 ng/ μ L e razões A_{260}/A_{280} entre 1,80 e 2,04.

Figura 33 - Eletroforese em gel de agarose do RNA total extraído das folhas oriundas do segundo experimento de inoculação. Canaletas de 1 a 9, triplicatas biológicas de cada tempo de coleta (6 h, 5 e 9 d) do tratamento com *Xap*. Canaletas de 10 a 18, triplicatas biológicas de cada tempo de coleta (6 h, 5 e 9 d) do tratamento controle. M, Marcador de peso molecular 100 pb (Invitrogen)

A síntese dos cDNAs das nove amostras de RNA total do tratamento com *Xap* e das nove amostras do tratamento controle, apresentou a qualidade esperada no teste de eficiência com os conjuntos de *primers* F-Contig33_3' e F-Contig33_5'. Os valores de ΔC_q ($C_q3' - C_q5'$) das 18 amostras foram entre 1,4 a 2,0, sendo os cDNAs considerados apropriados para o uso. Também, no teste de amplificação com o par de *primers* que amplifica um fragmento do gene *NDID*, os produtos das PCRs com os cDNAs apresentaram um fragmento único de 598 pb, confirmando a ausência de contaminação por DNA genômico nos RNAs (Figura 34).

Figura 34 - Eletroforese em gel de agarose do produto da PCR com o par de *primers* NDID. Canaletas de 1 a 9, *template* de cDNA das triplicatas biológicas de cada tempo de coleta (6 h, 5 e 9 d) do tratamento com *Xap*; Canaletas de 10 a 18, *template* de cDNA das triplicatas biológicas de cada tempo de coleta (6 h, 5 e 9 d) do tratamento controle; Canaletas 19 e 20, *template* de DNA genômico de maracujá azedo; Canaleta 21, reação controle, sem *template*. M, Marcador de peso molecular 100 pb (Invitrogen)

Os resultados dos produtos das qPCRs para 22 genes alvos e para o gene endógeno foram registrados pelo programa *StepOne* v2.1 (ABI). Isto permitiu analisar a qualidade das curvas de amplificação, que se mostraram consistentes nas repetições técnicas, e as de *Melt*, que apresentaram um pico único e uniforme de decaimento da fluorescência, para cada gene analisado, mostrando que os produtos amplificados estavam adequados para a posterior análise de quantificação relativa da expressão (Figura 35).

Figura 35 – Curvas de amplificação (A) e Curvas de *Melt* (B), do produto da qPCR com os cDNAs do segundo experimento de inoculação e com o par de *primers* F-P9A6, que amplifica um fragmento do gene da lipoxigenase 2

Os dados brutos foram exportados e corrigidos usando o programa LinRegPCR, resultando em valores de eficiência média, por gene, acima de 1,9 e em valores de C_q para cada repetição técnica de cada amostra de cDNA. Esses valores foram inseridos no programa REST, para todos os genes alvos e endógenos, sendo que os C_q das amostras do tratamento com *Xap* foram inseridos na categoria

‘*Samples*’ (tratado) e os C_q das amostras do tratamento controle foram inseridos na categoria ‘*Controls*’ (controle). Assim, obteve-se uma razão de expressão dos genes alvos para as amostras de cDNAs de cada período de coleta do tratamento com *Xap* em relação ao tratamento controle, que está apresentado na figura 36.

Figura 36 – Razão da expressão (E) de 22 genes alvos (1 a 64, tabela 5) para os cDNAs do tratamento com *Xap* em relação aos cDNAs do tratamento controle, considerando um nível de significância de 5 %. Em azul, indicam-se os genes mais expressos no tratamento com *Xap*, em vermelho, os genes mais expressos no tratamento controle e, em preto, são indicados os genes em que não houve diferença de expressão entre os tratamentos

Dos 22 genes avaliados após 6 horas de interação prata-patógeno (6 h.a.i.) um gene foi mais expresso (em azul, figura 36) e 12 genes foram menos expressos (em vermelho, figura 36) em relação ao tratamento com *Xap*, ou seja, 59 % dos genes apresentaram diferença na expressão entre os tratamentos. O gene mais expresso é um receptor quinase brassinosteróide, envolvido no reconhecimento de PAMP, o que sugere a sua atuação no reconhecimento do patógeno, no início da

infecção. Dos genes menos expressos, cinco tem funções relacionadas à resposta de defesa, os quais devem ter sido alvos do patógeno, que conseguiu inibi-los já no período inicial da infecção.

Aos 5 d.a.i., dos 22 genes, oito foram mais expressos (em azul, figura 36) e dois menos expressos (em vermelho, figura 36) no tratamento com *Xap*, ou seja, para 45 % dos genes observou-se expressão diferencial. Os oito genes mais expressos são conhecidos por estarem envolvidos na resposta de defesa, mostrando que após 5 dias da inoculação a planta já tinha reconhecido e estava respondendo à presença do patógeno.

Já aos 9 d.a.i., dos 22 genes, seis foram mais expressos (em azul, figura 36) e dez menos expressos (em vermelho, figura 36) no tratamento com *Xap*. Isto significa que 72 % dos genes responderam diferencialmente à presença do patógeno. Dos genes mais expressos, cinco foram os mesmos detectados aos 5 d.a.i., sendo que dois tiveram um decréscimo na expressão e três um acréscimo na expressão. Isto mostra que após 9 dias de interação, a planta continua respondendo à presença do patógeno, porém, o desenvolvimento da doença começa a entrar em estágio mais avançado. Neste período, iniciam-se os sintomas de clorose.

4.7 Análise detalhada das sequências dos genes da lipoxigenase e da (+)-neomentol desidrogenase

Conforme comentado no item anterior, Os genes que codificam as enzimas lipoxigenase e (+)-neomentol desidrogenase foram os que apresentaram maior nível de expressão durante a interação com *Xap*. Por isso, as sequências dos transcritos desses genes foram escolhidas para a realização de uma análise de similaridade com sequências de outras espécies vegetais depositadas em bancos públicos.

O transcrito de maracujá azedo que codifica a enzima lipoxigenase (clones R-P4D5 e F-P9A6) possui 1.337 nucleotídeos, sendo 853 nucleotídeos da região codante e 484 da região 3' UTR (*untranslated region*). A região codante foi traduzida *in silico* resultando em 284 aminoácidos, os quais foram comparados com sequências proteicas disponíveis no *GeneBank*, através da ferramenta BlastP. Os alinhamentos resultantes mostraram alta similaridade com sequências de

lipoxigenases identificadas em diversas espécies vegetais. A partir daí, foram selecionadas sequências de nove espécies vegetais com alta similaridade àquela de maracujá azedo (Tabela 7) para a estimativa de distância evolutiva.

Tabela 7 – Sequências da enzima lipoxigenase com similaridade à respectiva sequência de maracujá azedo, selecionadas para a estimativa de distância evolutiva. Código de acesso das sequências, identidade e pontuação do alinhamento, segundo a ferramenta BlastP

nº de acesso	espécie	Similaridade (%)	e-value
gi 224106626	<i>Populus trichocarpa</i>	75,36	7,00E-151
gi 225450227	<i>Vitis vinifera</i>	74,29	2,00E-150
gi 255544332	<i>Ricinus communis</i>	71,43	7,00E-140
gi 308943877	<i>Camellia sinensis</i>	70,00	8,00E-141
gi 6911852	<i>Arabidopsis thaliana</i>	65,49	4,00E-128
gi 356500370	<i>Glycine max</i>	64,89	8,00E-131
gi 124014020	<i>Phaseolus vulga</i>	63,57	3,00E-122
gi 218190241	<i>Oryza sativa</i> Indica Group	62,32	2,00E-125
gi 238014768	<i>Zea mays</i>	58,74	3,00E-122

A enzima lipoxigenase 2, de *A. thaliana* (AtLOX2) (gi|6911852|), é uma 13-lipoxigenase e, dentre as sequências alinhadas, é a que está melhor anotada no banco de dados do UniProt (<http://www.uniprot.org/uniprot/P38418>). A AtLOX2 é formada por 896 resíduos de aminoácidos, apresenta um peptídeo de trânsito para o cloroplasto entre as posições 1 a 56 da sua cadeia linear e sítios de ligação a ferro nas posições 554, 559, 746, 750 e 896. Também, apresenta um domínio PLAT entre as posições 79 a 199 e um domínio lipoxigenase entre as posições 202 a 896.

O domínio PLAT (*Polycystin-1, Lipoxxygenase, Alpha-Toxin*) é encontrado em proteínas de membrana ou proteínas associadas a lipídios, e forma um beta-sanduiche composto por duas folhas com quatro cadeias cada. Já o domínio lipoxigenase pertence à classe de dioxigenases contendo ferro, as quais catalisam a hidroperoxidação de lipídios, e tem uma estrutura cis,cis-1,4-pentadieno (<http://www.ebi.ac.uk/interpro/entry/IPR013819>).

A sequência da AtLOX2 foi comparada com a estrutura tridimensional do modelo 2iuj (<http://swissmodel.expasy.org/>), apresentando 44 % de similaridade. Esta estrutura é formada por um monômero e um ferro ligante (Figura 37).

Figura 37 – Representação da estrutura tridimensional da enzima modelo 2iuj (<http://swissmodel.expasy.org/>) (A) com 44 % de similaridade à enzima lipoxigenase 2 de *A. thaliana* (gi|6911852|), contendo um ferro ligante; destacam-se o domínio lipoxigenase em vermelho e o domínio PLAT em amarelo (B)

O alinhamento múltiplo de 286 resíduos para as 10 sequências de lipoxigenase, feito pelo algoritmo ClustalW e implementado no programa MEGA, está apresentado na figura 38. Já a estimativa de distância evolutiva, gerada a partir do alinhamento múltiplo das sequências de lipoxigenase, permitiu a construção da árvore filogenética apresentada na figura 39.

<i>P. edulis</i>	1	W	A	R	N	Y	S	I	E	L	S	S	V	A	Y	D	L	L	W	R	F	D	T	E	A	L	P	A	D	L	I	A	R	G	M	A	V	E	D	K	S	A	K	H	G	L	K	L	T	I	D	D	Y	P	Y	A	S	D	G
<i>P. trichocarpa</i>	501	S	P	G	K	Y	C	M	E	L	S	S	V	A	Y	D	K	L	W	R	F	D	T	E	A	L	P	A	D	L	I	R	R	G	M	A	V	E	D	P	T	A	R	H	G	L	K	L	T	I	E	D	Y	P	F	A	N	D	G
<i>V. vinifera</i>	620	S	P	G	K	Y	A	I	E	L	S	S	A	A	Y	D	Q	L	W	R	F	D	M	E	A	L	P	A	D	L	I	R	R	G	M	A	V	E	D	P	T	A	E	H	G	L	K	L	T	I	E	D	Y	P	F	A	N	D	G
<i>R. communis</i>	619	N	P	G	K	Y	S	E	L	S	S	V	A	Y	D	K	L	W	R	F	D	T	E	A	L	P	G	D	L	I	K	R	G	M	A	V	E	D	P	T	A	K	H	G	L	K	L	T	I	E	D	Y	P	F	A	N	D	G	
<i>C. sinensis</i>	618	S	P	G	K	Y	S	E	L	S	S	V	A	Y	D	Q	L	W	R	F	D	L	Q	A	L	P	A	D	L	I	S	R	G	M	A	V	E	D	Q	T	A	P	H	G	L	R	L	T	I	E	D	Y	P	F	A	N	D	G	
<i>A. thaliana</i>	613	W	P	G	K	Y	A	E	L	S	S	A	V	Y	G	K	L	W	R	F	D	Q	E	G	L	P	A	D	L	I	K	R	G	L	A	E	E	D	K	T	A	E	H	G	V	R	L	T	I	P	D	Y	P	F	A	N	D	G	
<i>G. max</i>	635	S	P	R	K	Y	S	M	E	I	S	S	V	A	Y	D	Q	L	W	Q	F	D	L	Q	A	L	P	N	D	L	I	F	R	G	M	A	V	A	D	P	N	A	P	H	G	L	K	L	T	I	E	D	Y	P	F	A	N	D	G
<i>P. vulgaris</i>	619	V	P	G	K	H	S	V	L	L	S	S	I	V	Y	D	K	H	W	R	F	D	H	Q	A	L	P	K	D	L	I	H	R	G	M	A	V	E	D	P	T	A	P	H	G	L	K	L	T	I	E	D	Y	P	Y	A	N	D	G
<i>O. sativa</i>	612	W	P	G	R	Y	A	M	E	L	S	S	V	A	A	T	W	R	F	D	A	E	A	L	P	E	D	L	V	R	R	G	L	A	V	R	Q	E	D	G	E	--	L	E	L	T	I	K	D	Y	P	Y	A	N	D	G			
<i>Z. mays</i>	23	W	P	G	R	Y	A	V	E	L	S	S	A	A	Y	G	A	T	W	Q	F	D	T	E	A	L	P	N	D	L	V	K	R	G	L	A	V	R	R	D	D	G	--	L	E	L	T	I	K	D	Y	P	Y	A	H	D	G		
<i>P. edulis</i>	59	L	I	L	W	D	A	I	K	E	W	I	S	D	Y	V	K	F	Y	L	D	D	S	K	V	G	Y	D	Q	E	L	Q	A	W	T	E	V	R	T	K	G	H	A	D	K	K	D	E	P	W	W	P	V	L	K	T	R		
<i>P. trichocarpa</i>	559	L	V	L	W	D	A	I	K	E	W	V	G	D	Y	V	K	H	Y	P	E	A	S	M	V	E	S	D	K	E	L	Q	A	W	T	E	V	R	A	K	G	H	E	D	K	K	D	E	P	W	W	P	V	L	K	T	Q		
<i>V. vinifera</i>	678	L	V	L	W	D	A	I	K	Q	W	V	R	D	Y	V	N	H	Y	P	D	P	S	L	V	E	S	D	K	E	L	Q	G	W	T	E	V	R	T	K	G	H	A	D	K	K	D	E	P	W	W	P	V	M	K	T	P		
<i>R. communis</i>	677	L	E	L	W	D	A	I	K	Q	W	V	T	D	Y	V	N	H	Y	P	E	A	S	Q	V	K	S	D	N	E	L	Q	A	F	W	E	E	V	R	T	K	G	H	G	D	K	K	D	E	P	W	W	P	V	L	N	T	Q	
<i>C. sinensis</i>	676	L	L	V	W	D	A	I	K	Q	W	V	T	D	Y	V	K	H	Y	Q	D	A	S	F	I	Q	S	D	K	E	L	Q	A	W	T	E	I	R	T	V	G	H	G	D	K	K	D	E	P	W	W	P	V	L	K	T	P		
<i>A. thaliana</i>	671	L	I	L	W	D	A	I	K	E	W	V	T	D	Y	V	K	H	Y	P	D	E	E	L	I	T	S	D	E	E	L	Q	G	W	S	E	V	R	N	I	G	H	G	D	K	K	D	E	P	W	W	P	V	L	K	T	Q		
<i>G. max</i>	693	L	L	I	W	D	A	I	K	E	W	V	S	E	Y	V	N	H	Y	P	S	S	S	T	I	E	F	D	Q	E	L	Q	A	W	T	E	I	R	T	V	G	H	G	D	K	S	E	E	P	W	W	P	N	L	K	T	P		
<i>P. vulgaris</i>	677	L	V	L	W	D	A	L	K	T	W	F	T	E	Y	V	D	H	X	Y	A	X	S	A	A	V	Q	S	D	T	E	L	Q	A	W	E	E	V	I	T	L	G	H	A	D	K	K	D	E	P	W	W	P	E	L	K	T	K	
<i>O. sativa</i>	668	L	L	V	W	S	I	K	Q	W	A	S	D	Y	I	D	F	Y	Y	K	S	D	E	E	V	A	C	D	E	E	V	R	A	W	W	E	E	V	R	T	K	G	H	A	D	K	K	D	E	P	W	W	P	A	V	D	T	R	
<i>Z. mays</i>	79	L	M	V	W	N	T	I	R	Q	W	A	A	D	Y	V	N	V	Y	K	S	D	E	A	V	A	A	D	P	E	L	K	A	F	W	D	E	V	R	N	V	G	H	G	D	K	K	D	E	P	W	W	P	V	L	D	T	R	
<i>P. edulis</i>	117	D	D	L	I	H	V	L	T	T	I	T	W	V	A	S	A	H	H	A	A	V	N	F	G	Q	Y	E	Y	G	G	Y	F	P	N	R	P	S	I	A	R	I	N	M	P	T	E	D	F	S	--	E	E	E	F	K	E	F	
<i>P. trichocarpa</i>	617	E	N	L	V	H	V	L	T	T	I	I	W	V	T	S	G	H	H	A	A	V	N	F	G	Q	Y	M	Y	G	G	Y	F	P	N	R	P	T	I	A	R	T	N	M	P	T	E	S	P	S	--	D	E	E	W	K	L	F	
<i>V. vinifera</i>	736	E	D	L	I	H	V	L	T	T	I	I	W	V	T	A	G	H	H	A	A	V	N	F	G	Q	Y	V	Y	A	G	Y	F	P	N	R	P	T	I	A	R	T	N	M	P	T	E	D	P	S	--	D	E	E	F	K	N	F	
<i>R. communis</i>	735	V	D	L	I	Q	V	L	T	T	I	I	W	V	T	S	G	H	H	A	A	V	N	F	G	Q	Y	V	Y	A	G	Y	F	P	N	R	P	T	M	A	R	T	N	M	P	T	E	E	P	S	--	E	Q	E	M	E	L	F	
<i>C. sinensis</i>	734	Q	D	L	I	G	I	L	T	T	M	I	W	V	T	S	G	H	S	A	V	N	F	G	Q	Y	M	Y	A	G	Y	F	P	N	R	P	T	I	A	R	T	K	M	P	T	E	D	P	T	--	D	E	E	W	K	C	F		
<i>A. thaliana</i>	729	D	D	L	I	G	V	T	T	I	A	W	V	T	S	G	H	H	A	A	V	N	F	G	Q	Y	G	Y	G	Y	F	P	N	R	P	T	T	T	R	I	R	M	P	T	E	D	P	T	--	D	E	A	L	K	E	F			
<i>G. max</i>	751	K	D	L	I	E	I	T	T	I	A	W	V	S	S	A	H	H	A	A	V	N	F	A	Q	Y	T	Y	G	G	Y	F	P	N	R	P	T	I	V	R	N	N	I	P	T	E	D	P	S	--	K	E	E	L	E	K	L		
<i>P. vulgaris</i>	735	E	D	L	V	E	I	V	T	T	I	A	W	T	A	S	G	H	H	A	A	V	N	F	G	Q	T	F	A	G	Y	F	P	N	R	P	T	I	V	R	N	N	M	P	I	E	D	P	T	--	D	P	E	W	E	L	F		
<i>O. sativa</i>	726	D	G	L	I	G	V	L	T	T	I	M	W	V	T	S	G	H	H	A	A	V	N	F	G	Q	Y	H	Y	G	Y	F	P	N	R	P	T	V	M	R	K	N	M	P	V	E	E	N	K	--	E	E	V	M	K	K	F		
<i>Z. mays</i>	137	D	S	L	A	E	T	L	T	T	I	M	W	V	T	S	G	H	S	A	V	N	F	G	Q	Y	H	F	G	Y	F	P	N	R	P	T	T	I	R	K	S	M	P	V	E	E	E	A	G	R	D	E	E	M	K	K	F		
<i>P. edulis</i>	173	L	R	K	P	E	D	T	L	L	K	C	F	P	S	Q	I	Q	A	T	K	V	M	A	V	L	N	V	L	S	H	S	P	D	E	E	Y	I	A	D	K	L	E	A	Y	W	E	D	E	A	V	I	K	A	A	F	E	R	
<i>P. trichocarpa</i>	673	L	K	K	P	E	L	S	L	L	K	C	F	P	T	Q	L	Q	A	T	K	V	M	A	V	L	N	V	L	S	H	S	P	D	E	E	Y	I	G	E	K	T	E	P	S	W	E	E	N	P	V	I	K	A	A	F	E	K	
<i>V. vinifera</i>	792	L	H	K	P	E	I	A	L	L	K	C	F	P	S	Q	I	Q	A	T	K	I	M	A	V	L	D	V	L	S	H	S	P	D	E	E	Y	L	G	D	Q	M	E	P	S	W	T	E	N	P	I	I	K	A	A	F	E	R	
<i>R. communis</i>	791	L	K	K	P	E	Y	T	L	L	K	C	F	P	S	Q	I	Q	A	T	K	V	M	A	V	L	D	V	L	S	G	H	S	P	E	E	E	Y	I	G	D	T	L	E	P	S	W	E	A	D	P	V	I	K	T	A	Y	E	R
<i>C. sinensis</i>	790	I	N	K	P	E	V	A	L	L	M	C	F	P	S	Q	I	Q	A	T	K	V	M	A	V	L	D	V	L	S	N	H	S	P	D	E	E	Y	L	G	K	D	M	E	A	S	W	I	E	N	P	I	I	K	A	A	F	E	R
<i>A. thaliana</i>	785	Y	E	S	P	E	K	V	L	L	K	T	P	S	Q	K	Q	A	T	L	V	M	V	T	L	D	L	S	T	H	S	P	D	E	E	Y	I	G	E	Q	Q	E	A	S	W	A	N	E	P	V	I	N	A	A	F	E	R		
<i>G. max</i>	807	I	N	N	P	E	K	T	F	L	E	S	L	P	S	Q	I	Q	A	T	L	V	M	V	L	N	L	S	N	H	S	P	D	E	E	Y	I	G	Q	Y	E	Q	S	W	V	E	N	Q	T	I	K	A	A	F	E	R			
<i>P. vulgaris</i>	791	L	Q	K	P	E	K	T	M	L	K	S	F	P	S	Q	I	Q	A	A	T	V	M	T	V	L	D	I	L	S	N	H	S	P	D	E	E	Y	L	G	M	S	V	E	P	A	W	E	E	E	P	L	V	N	A	A	F	E	K
<i>O. sativa</i>	782	M	E	M	P	E	H	V	L	D	T	M	P	S	K	M	Q	A	I	T	I	M	A	T	L	D	I	L	S	H	S	P	D	E	E	Y	M	G	E	H	A	E	P	A	W	L	A	E	P	R	V	K	A	A	F	E	R

Figura 39 – Árvore de distância computada usando o método *p-distance* (distância representada em número de substituições de aminoácidos por loco). A análise envolveu 10 sequências de aminoácidos, de diferentes espécies, de uma porção da enzima lipoxigenase e foi conduzida no programa computacional MEGA

O alinhamento mostrou uma alta identidade entre as sequências, sendo que as posições dos sítios de ligação a ferro, caracterizados para a AtLOX2 nas posições 746, 750 e 896, são conservadas em todas as demais sequências (Figura 38). Assim, diante da alta similaridade detectada entre a sequência de maracujá azedo e a da AtLOX2 (65 %), sugere-se que a primeira também seja uma lipoxigenase 2, embora este alinhamento represente apenas 1/3 da sequência completa da enzima.

Bannenberg et al. (2008) demonstraram que as lipoxigenases AtLOX-2, AtLOX-3, AtLOX-4 e AtLOX-6 são 13S-lipoxigenases e que os tecidos vegetais com expressão predominante de alguma das 13S-lipoxigenases produzem principalmente o ácido 13S-hidroperoxioctadecatrienóico, usado na síntese do ácido jasmônico.

Sanier et al. (2012) estudaram um gene de lipoxigenase 2 de algodão, o *GhLOX2*, e confirmaram que este gene codifica uma 13-lipoxigenase e apresenta um peptídeo de trânsito para o cloroplasto. Os autores observaram que o gene *GhLOX2* está altamente relacionado com a resposta de hipersensibilidade, pois foi altamente expresso nas primeiras horas após inoculação com *X. campestris* pv.

malvacearum, e que o tratamento de cotilédones de algodão com metil jasmonato e ácido salicílico também induziram a expressão deste gene.

A estimativa de distância evolutiva mostrou que a sequência de maracujá azedo apresentou maior similaridade com as sequências de lipoxigenase das espécies *P. trichocarpa* (75 %), *V. vinífera* (73 %) e *R. communis* (71 %). Estes resultados estão de acordo com o esperado, pois, dentre as espécies envolvidas na análise, *P. trichocarpa* e *R. communis* são as espécies filogeneticamente mais próximas ao maracujá, uma vez que elas pertencem a mesma Ordem, a *Malpighiales*, embora pertençam a famílias diferentes. Também, cabe ressaltar que dentre as espécies filogeneticamente próximas ao maracujá, estas são as que apresentam mais informações sobre a sua genômica, ou seja, grande parte de suas sequências estão disponíveis nos bancos de dados.

Já a sequência do transcrito de maracujá azedo que codifica a enzima (+)-neomentol desidrogenase (clone F-P10H8) possui 529 nucleotídeos, sendo 492 nucleotídeos da região codante e 37 da região 5' UTR (*untranslated region*). A região codante foi traduzida *in silico* resultando em 164 aminoácidos, os quais foram comparados com sequências proteicas disponíveis no *GeneBank*, através da ferramenta BlastP. Assim, foram selecionadas sequências de nove espécies vegetais com alta similaridade àquela de maracujá azedo (Tabela 8) para a estimativa de distância evolutiva.

Tabela 8 – Sequências da enzima (+)-neomentol desidrogenase com similaridade à respectiva sequência de maracujá azedo, selecionadas para a estimativa de distância evolutiva. Código de acesso das sequências, identidade e pontuação do alinhamento, segundo a ferramenta BlastP

nº de acesso	espécie	Similaridade (%)	e-value
gi 224104921	<i>Populus trichocarpa</i>	82,99	2,00E-80
gi 462401256	<i>Prunus persica</i>	75,32	2,00E-76
gi 255561230	<i>Ricinus communis</i>	75,00	1,00E-82
gi 449452688	<i>Cucumis sativus</i>	71,17	1,00E-75
gi 147777305	<i>Vitis vinífera</i>	71,15	7,00E-72
gi 357462063	<i>Medicago truncatula</i>	67,68	7,00E-72
gi 218200284	<i>Oryza sativa</i> Indica Group	61,29	1,00E-59
gi 356504773	<i>Glycine max</i>	53,90	8,00E-45
gi 15233062	<i>Arabidopsis thaliana</i>	50,64	4,00E-42

De acordo com as informações obtidas nos bancos de dados para a sequência gi|15233062| de *Arabidopsis thaliana*, a enzima (+)-neomentol desidrogenase, que tem como nomes alternativos mentona:neomentol redutase ou short-chain dehydrogenase/reductase 1 (AtSDR1), é composta por 296 resíduos de aminoácidos e apresenta uma região de ligação a coenzima NADP entre os resíduos 11 a 34 (<http://www.uniprot.org/uniprot/Q9M2E2>) e uma região 'short-chain dehydrogenase/reductase (SDR)' (<http://www.ebi.ac.uk/interpro/entry/IPR002198>) entre os resíduos 9 a 171. Sua sequência foi comparada à estrutura tridimensional do modelo 3o26A (<http://swissmodel.expasy.org/>), que possui um ligante NADP, apresentando 56 % de similaridade (Figura 40).

Figura 40 – Representação da estrutura tridimensional da enzima modelo 3o26A (<http://swissmodel.expasy.org/>) (A) com 56 % de similaridade à enzima (+)-neomentol desidrogenase de *A. thaliana* (gi|15233062|); destacam-se em vermelho (B) a região 'Short-chain dehydrogenase/reductase (SDR)' com uma região de ligação a NADP

O alinhamento múltiplo de 170 resíduos para as 10 sequências de (+)-neomentol desidrogenase, feito pelo algoritmo ClustalW e implementado no programa computacional MEGA está apresentado na Figura 41. Já a estimativa de distância evolutiva, gerada a partir do alinhamento múltiplo das sequências de (+)-neomentol desidrogenase, permitiu a construção da árvore filogenética apresentada na figura 42.

		Motivo 1	
<i>P. edulis</i>	1	-MEANGKPLVGRYAVVT GANKGIG LETVRQLASQG-VTVVLTARSEKRGMDAISRIH	
<i>P. trichocarpa</i>	1	MATNGDHSRVEKYAVVT GANKGIG LETVRQIASKG-VTVVLTARDEKRGTDATCMLH	
<i>R. communis</i>	1	-MKAHGNQSAEKYAVVT GANKGIG FETVRQLASRG-VTVVLTARNEKRGVDATSMHL	
<i>P. persica</i>	1	-MGSNEGGAERYAVVS GANKGIG YEVVRQLAACQGVVVLTARDEKRGDRDTSQLH	
<i>C. sativus</i>	1	-METNGNHVTERYGAVT GANKGIG FETAKQLASEG-ITVILTARNEQRGLEAVSKLH	
<i>V. vinifera</i>	1	MEVNMKHAAVEKCAVVT GANKGIG LETVRQLAAQG-VRVLTARDEERGIQATSSLH	
<i>M. truncatula</i>	1	-MESNGRYVAQRYAVVT GANKGIG LEIVKQLAFLG-VTVVLTARNDRTRGRDAITKLH	
<i>O. sativa</i>	67	AVNHDHETIPERLAVVT GGNRGIG LEVCRQLALQG-VTVILTARDEKRGKDAVESLC	
<i>G. max</i>	1	-----MGEATERYAVVT GANKGIG LEIVRQLASAG-IKVLLTARNEKKGLQALETLK	
<i>A. thaliana</i>	1	-----MAEETPRYAVVT GANKRGIGFEICRQLASEG-IRVVLT SRDENRGLEAVETLK	
		Motivo 2	
<i>P. edulis</i>	56	--SMG-FTNVLFHQLDVLDQVSAESLANFIKEK FGRLDILVNNAA ASGVVVDEERLR	
<i>P. trichocarpa</i>	57	--NMG-LTNVVFHQLDVLDPVSIESLAKFIKDRFGRLDIL VNNAG ASGVVVDEERLR	
<i>R. communis</i>	56	--QMG-LTNVVFHQLDVLDPVSIHSLANFIQNTFGRLDIL VNNAG ASGVVVDEHLR	
<i>P. persica</i>	57	--HLG-FSNVLFHQLDVLDPFISIQLANFIDNRFGRDLIL VNNAG ASGAVVDEEALK	
<i>C. sativus</i>	56	--EIG-LTNVVFHQLDVLDPDSIQSLAKFIADKFGRLDIL VNNAG ASGVVVDEGLR	
<i>V. vinifera</i>	57	--KLG-FSNVIFHQLDVVDPASIRSLADFIRHQFGKLDIL VNNAG ASGVIVDEQGLK	
<i>M. truncatula</i>	56	--QTG-LSNVMFHQLDVLDALSIESLAKFIQHKFGRLDIL VNNAG ASCVEVDKEGLK	
<i>O. sativa</i>	123	--HESNLSNII FHQLD ILDGNSRASLARYINSRFGKLDIL VNNAG VGGVAVDQDGLR	
<i>G. max</i>	52	DSGLS--HLVLFHQVDVADATNVASLADFVKSFGKLDIL VNNAG IGGVVDDTDLI	
<i>A. thaliana</i>	52	KELEISDQSL L FHQLDVADPASITSLAEFVK TQFGKLDILVNNAG IGGIITDAEALR	
<i>P. edulis</i>	110	AMNIDPQTWLSGKATN-MVQVAIKT TYETAE CLNTN FYGV KRVTEALL PLVQL SS	
<i>P. trichocarpa</i>	111	AMNIDPETWLSGKATN-MVQV IKTSYETAVE CLNTN FYGV QRLTEALL PLLQL SP	
<i>R. communis</i>	110	ALNIDPETWLSGKAVN-LLQEV MKT TYEKAE CLNTN FYGV RR LTEALL PLLQL ST	
<i>P. persica</i>	111	SLNIDPTTWLSGKAVN-KIQCAIK FTY EKAE CLNTN Y YGA KRLTEALL PLLKL SP	
<i>C. sativus</i>	110	AMNIDFSSWLSGKATN-LVQSV IKTNCE KA EGLNTN Y YGL KNVTEALL PLLQK SL	
<i>V. vinifera</i>	111	ALNIDPASWLSGKATN-LVQAV IKQTYE KA CLNTN Y YGC KRVTEALL PLLKL ST	
<i>M. truncatula</i>	110	ALNVDPATWLAGKVSNTLLQGVLTQTY KKA E CLNTN Y YGV KRV TM ALL PLLQL SP	
<i>O. sativa</i>	178	ALNIDPRVWLSGKAVN-LIQSV IVQTY DEAV KCLNTN Y YGL KWITEALL PLLQK SP	
<i>G. max</i>	107	TTAIMNRGAIP-----EDNGTKG ITHTYELAE CLQIN YGA KKT TESLM PLL QL SD	
<i>A. thaliana</i>	109	AGAGKEG-----FKWDEI ITET YEL TEECIK IN YGP KRMCEAFI PLLKL SD	

Figura 41 – Alinhamento de 10 sequências de aminoácidos, de uma região da enzima (+)-neomentol desidrogenase, das espécies *Passiflora edulis* f. *flavicarpa*, *Populus trichocarpa* (acesso 224104921), *Prunus persica* (acesso 462401256), *Ricinus communis* (acesso 255561230), *Cucumis sativus* (acesso 449452688), *Vitis vinifera* (acesso 147777305), *Medicago truncatula* (acesso 357462063), *Oryza sativa* Indica Group (acesso 218200284), *Glycine max* (acesso 356504773) e *Arabidopsis thaliana* (acesso 15233062). Em cor alaranjada estão destacadas as 41 posições conservadas com um único tipo de resíduo, em azul as 33 posições conservadas que apresentam um grupo de resíduos ‘forte’ e em marrom as 8 posições conservadas que apresentam um grupo de resíduos ‘fraco’ (‘forte’: score > 0,5 e fraco: score < 0,5, de acordo com a matriz Gonnet Pam250). Na sequência de *A. thaliana*, em negrito entre os resíduos 11 a 34, há uma região de ligação de NADP. Motivo 1, domínio de ligação à coenzima; Motivo 2, domínio estrutural de função indefinida. Os números indicam a posição do primeiro aminoácido da linha

Figura 42 – Árvore de distância computada usando o método *p-distance* (distância representada em número de substituições de aminoácidos por loco). A análise envolveu 10 sequências de aminoácidos, de diferentes espécies, de uma porção da enzima (+)-neomentol desidrogenase e foi conduzida no programa computacional MEGA

Na estimativa de distância evolutiva, a sequência de maracujá azedo apresentou 80 % de similaridade à sequência de *P. trichocarpa*, espécie filogeneticamente mais próxima ao maracujá, e também alta similaridade às sequências de *R. communis* (75 %), *P. persica* (70 %), *C. sativus* (70 %), *V. vinífera* (69 %) e *M. truncatula* (68 %). Os valores de similaridade foram menores comparativamente quando se usou *O. sativa* (58 %), *G. max* (55 %) e *A. thaliana* (51 %).

A (+)-neomentol desidrogenase pertence a uma grande família de enzimas, as 'Short-chain dehydrogenases/reductases' (SDRs), que apresentam uma alta diversidade funcional. A identidade entre estas enzimas é de apenas 15 a 30 %, porém, apresentam regiões específicas que indicam um padrão comum no seu dobramento. As SDRs possuem por volta de 250 resíduos e catalisam reações de oxidação/redução dependentes de NAD(H) ou NADP(H) (KALLBERG et al., 2002).

O alinhamento apresentado na figura 41 representa uma porção de aproximadamente 50 % da sequência completa da enzima (+)-neomentol desidrogenase, mostrando regiões conservadas entre as diversas sequências analisadas, como a região de ligação a NADP, entre os resíduos 11 a 34 para a sequência de *A. thaliana*.

Choi et al. (2008) estudaram a enzima de pimenta mentona:(1)-(3S)-neomentol redutase, CaMNR1, que é ortóloga à AtSDR1 de *A. thaliana* (mesma sequência usada no alinhamento da figura 41). Esta enzima catalisa a redução da mentona, tendo como cofator o NADP, para produzir o neomentol com atividade antimicrobiana, e também possui atividade catalítica para a oxidação do neomentol. Os autores observaram similaridade de 55 a 57 % entre as sequências proteicas CaMNR1 e AtSDR1 e identificaram na CaMNR1 domínios conservados 'short-chain dehydrogenase', sendo eles: um domínio de ligação à coenzima, um domínio estrutural de função indefinida e um elemento de sítio ativo.

Dos três motivos descritos acima, o domínio de ligação à coenzima (motivo 1, figura 41) e o domínio estrutural de função indefinida (motivo 2, figura 41) puderam ser identificados no alinhamento e se mostraram conservados entre as sequências. Já o elemento de sítio ativo se localiza em uma região não apresentada no alinhamento.

Nesta análise, embora a sequência AtSDR1 de *A. thaliana* tenha apresentado a menor similaridade com àquela de maracujá azedo, os 51 % de similaridade entre as duas sequências e a identificação dos dois domínios conservados são uma forte evidência de que a sequência aqui identificada é uma enzima da família SDR, ou seja, uma (+)-neomentol desidrogenase.

5 DISCUSSÃO

5.1 Processamento e anotação das sequências das bibliotecas subtrativas

O resultado dos processamentos das sequências oriundas das bibliotecas mostraram um alto percentual de aproveitamento, aproximadamente 94 % (1.254 e 1.273/1.344), devido à qualidade do sequenciamento, e uma grande quantidade (982/1.254 e 998/1.273 ou 78 %) de unisequências passíveis de anotação. Comparativamente, Da et al. (2012) obtiveram um menor aproveitamento dos clones sequenciados (302/443 ou 68 %) de uma biblioteca SSH de citros para a identificação de genes envolvidos na forma e estrutura de folhas, porém, uma grande quantidade de unisequências (73%, 221/302) foi obtida. Já Feng et al. (2012b), analisando uma biblioteca SSH de ervilha durante interação com *Peronospora viciae* f. sp. *pisi*, obtiveram um aproveitamento de 96,6 % (399/413) dos clones sequenciados e uma percentagem de apenas 51,8 % (207/399) de unisequências. Estas diferenças possivelmente se devem à qualidade do sequenciamento e ao número de clones sequenciados em relação ao número total de transcritos presentes em cada biblioteca.

As redundâncias obtidas nas bibliotecas desenvolvidas no presente trabalho se mostraram baixas, em média 21,6 %, significando que muito mais clones poderiam ter sido sequenciados, permitindo a obtenção de mais sequências expressas, ainda não identificadas. Por outro lado, Feng et al. (2012b) obtiveram cerca de duas vezes mais (48,1 %) de redundância dos clones sequenciados da biblioteca de ervilha, tendo uma maior representação de genes, em relação a todos os possíveis genes contidos na biblioteca.

O processamento das sequências aqui apresentadas, conforme descrito no item 3.6 da seção Materiais e Métodos, foi realizado de duas formas, ou seja, com o auxílio do programa computacional *CodonCode Aligner* e pelo *pipeline* de processamento de pré-anotação. Foi observada uma pequena diferença nos resultados das duas formas de processamento, devido aos parâmetros utilizados em cada análise e, principalmente, porque na primeira análise as sequências das duas

bibliotecas foram processadas separadamente enquanto na segunda análise as sequências foram processadas juntas.

Após o processamento, a busca por sequências similares aos transcritos de maracujá azedo resultou em uma grande porcentagem ($> 70\%$) de sequências com alta similaridade a proteínas disponíveis nos bancos de dados públicos. Na análise feita com base na Plataforma PLAZA, das 982 unisequências, sendo 672 da biblioteca *forward* e 310 da *reverse*, 78 % delas apresentaram similaridade com proteínas de *Arabidopsis thaliana*. Já na análise feita com base no programa computacional Blast2GO, do total das 998 unisequências das duas bibliotecas (683 da *forward*, 274 da *reverse* e 41 *contigs* com sequências de ambas as bibliotecas), 87 % apresentaram similaridade com proteínas putativas de diversas espécies vegetais disponíveis nos bancos de dados. O fato de os transcritos terem sido comparados com o banco de dados proteico de *A. thaliana*, quando se usou a Plataforma PLAZA e com o banco de dados proteico, não redundante, com sequências de diversas espécies vegetais, quando se usou o programa Blast2GO, explica a diferença nos resultados das duas análises de anotação. *A. thaliana* foi escolhida para análise, mesmo não sendo filogeneticamente próxima ao maracujá azedo, por ser uma espécie modelo e pelo fato de o banco de dados apresentar bons resultados de anotação para sequências proteicas.

Na análise feita com o programa Blast2GO (Figura 23), 41 % dos *top-hits* apresentaram similaridade entre as sequências de maracujá azedo e as de *Populus trichocarpa* e 27 % entre as sequências de maracujá azedo e as de *Ricinus communis*. Isto decorre do fato dessas duas espécies serem as mais estudadas do ponto de vista genômico, dentre as mais próximas filogeneticamente do maracujá azedo, segundo dados do 'Angiosperm Phylogeny Website v. 12' (<http://www.mobot.org/MOBOT/research/APweb/>). Santos (2013) também observou resultados similares ao anotar 767 sequências codificadoras de maracujá azedo, oriundas do sequenciamento de BAC-ends de uma biblioteca construída para o genótipo 'IAPAR123' (<http://cnrgv.toulouse.inra.fr/Library/Passiflora>). Neste caso, 28,8 % dos *top-hits* apresentaram similaridade com *P. trichocarpa* e 16,6 % com *R. communis*. Por se tratar de uma espécie pouco estudada, com limitadas informações de sequências expressas disponíveis, pode-se afirmar que nossos resultados superaram as expectativas de sucesso nas comparações de sequências.

Dos resultados da anotação e categorização funcional das sequências da biblioteca *forward*, ou seja, da atribuição em 15 categorias, apresentadas na figura 21, nota-se que as proteínas estão relacionadas a quase todos os processos celulares. Nesta classificação, a categoria menos representada (1,5 %, 8/525) foi a de 'metabolismo do DNA' e a mais representada (8,8 %, 46/525) foi a de 'transporte'; entretanto, 19,8 % (104/525) das sequências foram atribuídas à categoria das proteínas 'não classificadas'. A atribuição de transcritos diferencialmente expressos durante a interação planta-patógeno a diversas categorias funcionais também pode ser observada em outros trabalhos relacionados. Por exemplo, Shi et al. (2010) categorizaram transcritos de feijão diferencialmente expressos em resposta à *X. axonopodis* pv. *phaseoli* em 'função metabólica' (32 %), 'fotossíntese ou transporte celular' (8 %), 'resposta de defesa' (8 %), 'regulação enzimática ou da transcrição' (10 %) e 'resposta à estímulos' (10 %), sendo o restante não classificado. Transcritos de arroz selvagem diferencialmente expressos em resposta à *X. oryzae* pv. *oryzae*, foram atribuídos a categorias relacionadas a 'stress e defesa' (9 %), 'transdução de sinal' (2,9 %), 'fotossíntese e fotorrespiração' (8,7 %) e 'metabolismo e transporte de materiais' (16,7 %), porém, 52 % dos transcritos com outras diversas funções foram alocados em um único grupo e 9,7 % não puderam ser classificados (CHENG et al., 2009). Interessante destacar que nos trabalhos acima citados e na presente pesquisa, as categorias relacionadas à 'defesa' apresentaram baixa porcentagem de transcritos, ou seja, entre 7 e 9 %.

Para a interação couve-flor-*X. campestris* pv. *campestris*, estudada pela mesma metodologia aqui desenvolvida, foram sequenciados 280 clones, obtendo-se 202 unisequências. Na busca por similaridade nos bancos de dados, usando BlastX, 86,7 % das unisequências apresentaram similaridade a genes de plantas com funções conhecidas. Baseando-se nas categorias do MIPS (*Munich Information Center for Protein Sequences*), para *Arabidopsis*, Jiang et al. (2010) atribuíram essas sequências a 11 categorias funcionais, dentre elas metabolismo, defesa celular, energia, transdução de sinais, transporte, etc. Comparativamente, os autores obtiveram resultados similares aos observados no presente trabalho, ou seja, 87 % das unisequências de maracujá azedo apresentaram similaridade a proteínas vegetais, as quais foram atribuídas a 15 categorias correspondentes às usadas por Jiang et al. (2010).

Da mesma forma, visando comparar o perfil transcricional de dois genótipos de *Glycine tomentella* em resposta ao fungo patogênico *Phakopsora pachyrhizi*, sendo um genótipo resistente e outro suscetível ao patógeno, foram construídas quatro bibliotecas SSH, das quais foram sequenciados 1.536 clones. A montagem das sequências resultou em 1.060 unisequências que, na busca por similaridade nos bancos de dados, por BlastN, apresentaram similaridade a 646 (61 %) genes de plantas com funções conhecidas e 414 (39 %) unisequências apresentaram similaridade a proteínas hipotéticas ou a genes com funções desconhecidas. Os genes com funções identificadas codificam proteínas envolvidas no metabolismo, defesa celular, energia, síntese proteica e transporte celular. Soria-Guerra et al. (2010) identificaram genes constitutivamente expressos em altos níveis no genótipo resistente e em baixos níveis no genótipo suscetível. Vale ressaltar que foi identificado um grande número de lipoxigenases na biblioteca construída para o genótipo suscetível em relação às demais bibliotecas, o que está de acordo com o observado na biblioteca construída na presente pesquisa, em que foi observado um transcrito da lipoxigenase altamente expresso no genótipo de maracujá azedo, suscetível ao patógeno.

Outros trabalhos relacionados à interação planta-patógeno, que usaram o programa Blast2GO nas análises de anotação de sequências, também obtiveram uma maior associação de proteínas aos termos mais gerais de nível '2', como se observa nos gráficos da figura 25, que ilustra os resultados da presente pesquisa.

Visando estudar a interação canola-*Plasmodiophora brassicae*, se obtiveram 407 unisequências de canola e se observou uma redundância da ordem de 45 % da biblioteca SSH. Ao anotar estas sequências usando Blast2GO, 369 unisequências foram identificadas com a ferramenta BlastX, sendo que destas, 336 foram mapeadas e anotadas e cada uma foi associada a até 23 termos do GO. Da distribuição das ocorrências em termos do nível 2, para a categoria 'processos biológicos', os mais representados foram 'cellular process' (232), 'metabolic process' (222) e 'response to stimulus' (117), destacando-se nove unisequências associadas ao termo 'immune system process'. Para a categoria 'função molecular' os termos mais representados foram 'binding' (212) e 'catalytic activity' (184) e para a categoria 'componente celular' foram 'cell' (278) e 'organelle' (220). Dos sete genes validados por qPCR, cinco foram mais expressos (*cysteine proteinase*, 2,36 X; *cysteine*

proteinase, 2.61 X; *peroxidase* 27, 1,38 X; *Sulphotransferase*, 1,48 X; *Meprin / TRAF homology protein*, 1,15 X) e dois menos expressos (*CBL-interacting protein kinase* 9, 1,45 X; *aspartyl protease*, 1,47 X) na planta inoculada, sugerindo a importância desses genes na resposta de defesa (FENG; HWANG; STRELKOV, 2012c).

Com o intuito de identificar transcritos de banana responsivos ao fungo *Mycosphaerella musicola*, foram construídas duas bibliotecas usando a metodologia 454 GS-FLX Titanium, sendo uma para o genótipo resistente e outra para o suscetível ao patógeno. Os autores obtiveram 36.384 unisequências para a biblioteca do genótipo resistente e 35.269 para a do genótipo suscetível. Dessas, cerca de 10.000 unisequências, tanto do genótipo resistente como do suscetível, apresentaram similaridade a proteínas com funções conhecidas, quando comparadas no banco de dados usando BlastX. A atribuição dessas sequências a termos ontológicos de nível '2', usando Blast2GO, mostrou que para a categoria 'processos biológicos', os termos que continham maior número de unisequências, para ambas as bibliotecas foram: 'cellular process' (5.530 e 5.649 do genótipo resistente e suscetível, respectivamente), 'metabolic process' (5.276 e 5.418), 'biological regulation' (1.397 e 1.377), 'response to stimulus' (1.343 e 1.383) e 'localization' (1.179 e 1.227). Para a categoria 'função molecular' os termos 'binding' (4.732 e 4.863), 'catalytic activity' (4.659 e 4.771) e 'transporter activity' (678 e 708) foram os mais evidentes; já para a categoria 'componente celular' foram 'cell' (7.580 e 7.725) e 'organelle' (5.589 e 5.735). Dessas análises, foram identificados em ambos os genótipos, muitos genes possivelmente envolvidos na imunidade desencadeada por efetores (ETI) e na imunidade desencadeada por PAMP (PTI) (PASSOS et al., 2013).

Ao comparar os resultados das anotações automáticas, realizadas pelo programa Blast2GO, dos trabalhos acima descritos (interação canola-*P. brassicae* e banana-*M. musicola*) com os do presente trabalho, observa-se que os termos mais representados de nível '2' foram os mesmos. Ou seja, para a interação maracujá azedo-*Xap*, na categoria 'processos biológicos', os termos mais representados foram igualmente 'cellular process', 'metabolic process', 'response to stimulus' e 'localization'; para a categoria 'funções moleculares' foram 'binding', 'catalytic activity' e 'transporter activity'; já para 'componentes celulares', os termos mais representados foram 'cell', 'membrane' e 'organelle'. O estudo sobre a interação

canola-*P. brassicae* apresentou um termo a menos para cada uma das três categorias comparativamente ao presente trabalho enquanto no estudo banana-*M. musicola* houve, adicionalmente, o termo 'biological regulation' como um dos mais representados.

Tal como observado no presente trabalho e nos demais acima descritos, que avaliaram o perfil transcricional de espécies vegetais em resposta a patógenos, as proteínas putativas, codificadas pelos transcritos, foram atribuídas a diversas categorias funcionais relacionadas a maioria das atividades celulares. Isto mostra que o processo de defesa contra patógenos é complexo e envolve a atuação conjunta de diversos mecanismos, bem como, concomitantemente à resposta de defesa, a célula vegetal continua realizando seus processos basais.

5.2 Proteínas envolvidas em respostas de defesa durante a interação maracujá azedo-*Xap*

A anotação das sequências obtidas na biblioteca *forward* permitiu a identificação de 73 proteínas relacionadas a algum tipo de resposta de defesa, as quais foram atribuídas a categorias, de acordo com a sua provável função, (tabela 3 e figura 22) e possíveis interações. Porém, sabendo-se que a interação maracujá azedo-*Xap* é compatível, pode-se afirmar que a resposta de defesa desencadeada pela planta é basal. Na defesa basal, a planta reconhece padrões moleculares conservados associados a patógenos (PAMPs), como flagelina e quitina, ou reconhece fracamente fatores de avirulência (Avr), através de receptores de membrana específicos PRR (receptores de reconhecimento de padrões), desencadeando respostas aparentemente genéricas. Este sistema de defesa limita a extensão da doença causada por raças virulentas de patógenos, durante as interações compatíveis, porém não evita o desenvolvimento do patógeno e, conseqüentemente, da doença (ABRAMOVITCH; MARTIN, 2004; EULGEM, 2005; SPOEL; DONG, 2012).

Dentre as respostas desencadeadas durante a defesa basal estão: o fluxo de íons, a combustão oxidativa, a ativação da cascata de sinalização de MAP quinases, mudanças transcricionais com ativação de genes de defesa, a produção de

componentes antimicrobianos, como proteínas relacionadas à patogênese (PR) e fitoalexinas, e a morte celular tardia (ABRAMOVITCH; MARTIN, 2004; SPOEL; DONG, 2012).

Assim, dessas 73 proteínas, a maior parte (62 %) foi atribuída a categoria ‘transdução de sinais intracelular’. Dentre elas, foram identificadas duas proteínas da cascata de sinalização das MAP quinases, sendo uma MPKK2 e uma MPK4. Em *Arabidopsis*, a MPKK2 ativa a MPK4 durante a cascata de sinalização em resposta a estresse biótico enquanto a MPK4 atua como um regulador nas respostas de defesa, pois participa na repressão da resistência dependente de ácido salicílico (SA) e na ativação de genes de defesa dependente de ácido jasmônico (JA). A MPK4, ativada em resposta a um patógeno, fosforila o substrato MKS1 (*MAP kinase 4 substrate 1*) que, por sua vez, interage com os fatores de transcrição WRKY25 e WRKY33. Estes fatores de transcrição atuam na regulação da expressão de genes que codificam proteínas de defesa (ANDREASSON et al., 2005; QIU et al., 2008a; QIU et al., 2008b). Em concordância a isso, aqui foram identificadas interações da MPKK2 com a MPK4 e o WRKY33 (Figura 22).

As 22 proteínas quinases atribuídas à mesma categoria englobam diversos tipos de receptores quinases, como o BAK1, serina/treonina quinases, proteínas quinases com repetições ricas em leucina e proteínas quinases dependentes de cálcio. Também, foi aqui observada uma rede de interações desses receptores, tendo como base as informações de interações proteicas de *A. thaliana* (Figura 22). O BAK1, receptor quinase 1 com repetições ricas em leucina associado ao brassinosteróide insensível 1 (BRI1), de *A. thaliana*, atua conjuntamente com o receptor do hormônio de desenvolvimento brassinosteróide (BR), BRI1, e com o receptor de flagelina, FLS2, e regula a contenção da morte celular programada induzida pela infecção microbiana, atuando independente do hormônio BR (KEMMERLING et al., 2007; OH et al., 2010).

O gene *Pm21*, que codifica uma proteína serina/treonina quinase, confere resistência durável e de amplo espectro ao oídio do trigo, foi transferido por translocação cromossômica a partir de *Haynaldia villosa*, e tem se tornado um recurso fundamental nos programas de melhoramento de resistência ao oídio em todo o mundo (CAO et al., 2011). O gene de pimenta *CaPIK1* codifica um receptor proteína quinase citoplasmático e sua transcrição é ativada pela infecção por *X. campestris*

pv. *vesicatoria*. O *CaPIK1* modula a sinalização necessária para a defesa dependente do ácido salicílico em resposta à infecção pelo patógeno, induzindo a geração de ERO e, conseqüentemente, a morte celular (KIM; HWANG, 2011). Em contrapartida, o efector tipo III AvrAc, de *X. campestris* pv. *campestris* provoca o aumento da virulência e a inibição da imunidade da planta ao inativar os sítios ativos dos receptores quinases citoplasmáticos de *Arabidopsis* BIK1 e RIPK, reduzindo a atividade quinase e inibindo a via de sinalização desses receptores (FENG et al., 2012a).

Também, outras 18 proteínas aqui atribuídas na categoria 'transdução de sinais intracelular' apresentam atividade de ATPases, sendo 10 delas pertencentes à superfamília das proteínas ABC transportadoras (*ATP-binding cassette*). Proteínas ATPases de membrana participam da sinalização em resposta a diversos tipos de estímulos, contribuindo para a sinalização durante as respostas de imunidade, tanto na PTI como na ETI, e para a regulação da abertura dos estômatos, controlando a entrada de patógenos. Porém, estas enzimas são alvos de patógenos, o que pode causar o aumento da suscetibilidade das plantas (ELMORE; COAKER, 2011). Em um estudo sobre o gene *NtPDR1* de *Nicotiana tabacum*, que codifica a proteína PDR (*Pleiotropic Drug Resistance*), pertencente a subfamília ABCG transportadora, foi observada a indução da sua expressão pelo metil jasmonato e por elicitores, e o transporte pelas PDR de diterpenos, que são moléculas antimicrobianas, sugerindo a participação dessa proteína na defesa da planta contra tratamentos bióticos (CROUZET et al., 2013).

Na categoria 'ativação de genes de defesa/regulação da transcrição' foram alocados alguns fatores de transcrição, como o WRKY33 e o ERF1, bem como a enzima glutamato-cisteína ligase e a proteína MLO. O fator de transcrição WRKY33 de *Arabidopsis* reage a sinais de defesa vegetal e a vários tipos de fitopatógenos, bem como é dependente do PAD4, que é um regulador à montante do ácido salicílico, promovendo rápida ativação da transcrição desencadeada por patógeno ou por PAMP (LIPPOK et al., 2007). Já o fator de transcrição responsivo ao etileno 1 (ERF1), sugere-se que seja um importante integrador dos sinais de etileno e jasmonato durante a regulação das respostas de defesa desencadeadas por esses hormônios (LORENZO et al., 2002).

Por sua vez, a enzima glutamato-cisteína ligase (GCL) catalisa o primeiro passo da via de biossíntese da glutathione e atua na regulação do ambiente redox intracelular (JEZ; CAHOON; CHEN, 2004). Ao identificar o gene *PAD2*, que codifica a γ -glutamylcisteína sintetase (GSH1), também envolvida na biossíntese da glutathione, foi verificado que níveis adequados desta enzima são importantes em *Arabidopsis* para limitar a dispersão de *Pseudomonas syringae* e para estabelecer a resistência a *P. brassicae*. Assim, níveis adequados de GSH são importantes para o acúmulo de componentes relacionados à resistência e para o estabelecimento da resistência a vários patógenos (PARISY et al., 2006). Já a proteína MLO pertence a uma classe do tipo transmembrana que funciona como um regulador negativo de resistência a doenças e à morte celular (DEVOTO et al., 2003; ZIEROLD; SCHOLZ; SCHWEIZER, 2005).

Na categoria 'proteínas de defesa/resistência' está a enzima (+)-neomentol desidrogenase, a qual participa da síntese do neomentol, uma proteína de defesa de ação antimicrobiana que tem sua síntese induzida em resposta ao ataque de patógenos (CHOI et al., 2008). Também, duas proteínas de resistência aqui identificadas, sugere-se que possuam outras funções na interação compatível maracujá azedo-*Xap*, ou que não desempenharam eficientemente suas funções de proteínas R.

Já na categoria 'oxidoreductases' estão quatro enzimas monoxigenases, sendo duas delas citocromo P450, além da enzima de detoxificação do H_2O_2 catalase. O gene de arroz *SL* codifica uma proteína da família citocromo P450 monoxigenase, a qual catalisa a 5-hidroxilação do triptofano para sintetizar serotonina. Demonstrou-se que a expressão desse gene é induzida por quitina e infecção fúngica (*Magnaporthe grisea*) enquanto a serotonina induziria respostas imunes, incluindo a expressão de genes de defesa e a morte celular (FUJIWARA et al., 2010). Também, foi evidenciado que a CYP71A13, uma citocromo P450 monoxigenase, participa da biossíntese da camalexina, a partir de triptofano. A camalexina é uma fitoalexina encontrada em *A. thaliana*, cuja alta produção é induzida em resposta ao ataque de patógenos como *Pseudomonas syringae* e fungos necrotróficos; além disso, esta proteína está envolvida na resistência ao fungo *Alternaria brassicicola* (NAFISI et al., 2007).

Foram aqui atribuídas a categoria ‘morte celular’ duas proteínas, uma CAD1 (*Constitutively Activated Cell Death 1*) e uma VDAC (*Voltage-dependent anion channel 1*). O gene *CAD1* foi isolado e caracterizado em um estudo sobre os mecanismos que ativam a imunidade vegetal. Plantas de *Arabidopsis* com um único mutante *cad1* recessivo apresentaram imunidade mediada por ácido salicílico constitutivamente ativada. Os autores propuseram que quando a planta não está sob o ataque de patógenos, o gene *CAD1* atua como um regulador negativo da imunidade, reprimindo a via da morte celular mediada pelo ácido salicílico; entretanto, quando a planta é infectada, a função de regulador negativo do *CAD1* é inibida, permitindo a ativação da morte celular e restringindo o crescimento e a dispersão do patógeno. Completando, a CAD1 parece atuar como um mediador que reconhece os sinais da infecção (MORITA-YAMAMURO et al., 2005).

Por fim, 11 proteínas da interação maracujá azedo-*Xap* estão na categoria ‘proteínas relacionadas à patogênese’, dentre elas duas quitinases, uma lipoxigenase e cinco peroxidases. As quitinases quebram a estrutura de quitina da parede celular de muitos fungos fitopatogênicos, sendo que em plantas, elas exercem atividade antimicrobiana direta, pois, além de inibir o desenvolvimento dos fungos, os oligossacarídeos resultantes da degradação da quitina atuam como elicitores do sistema de defesa da planta (KEZUKA et al., 2010; GROVER, 2012; YEOH et al., 2013). As lipoxigenases são importantes nas respostas de defesa, pois, participam de diversas vias metabólicas que produzem moléculas sinalizadoras secundárias, como o ácido jasmônico, e moléculas com atividade antimicrobiana, além de atuarem no desencadeamento da resposta de hipersensibilidade (HR) (FEUSSNER; WASTERACK, 2002; PORTA; ROCHA-SOSA, 2002; BANNENBERG et al., 2008). As peroxidases também participam de vários processos durante as respostas de defesa, como no metabolismo de ERO, de ERN (espécies reativas de nitrogênio), da auxina e da lignina, bem como da reticulação de componentes da parede celular e da síntese de fitoalexina (ALMAGRO et al., 2009).

Conclui-se que as proteínas acima citadas são atuantes em respostas de defesa vegetais e, possivelmente, a maioria atuou durante a resposta de defesa basal do maracujá azedo contra *Xap*.

5.3 Validação da expressão diferencial dos transcritos por qPCR

Na primeira etapa de validação das bibliotecas, a partir dos *bulks* de transcritos de folhas coletadas após 6, 12 e 24 h da inoculação com *Xap* ou solução salina, foram observados 17 genes diferencialmente expressos (24,6 % dos genes analisados), sendo que esta diferença de expressão foi de 1,15 a 1,67 X entre os dois tratamentos. As diferenças obtidas foram pequenas, porém significativas, observando-se, principalmente, genes relacionados à resposta de defesa mais expressos no tratamento com *Xap*. Isto sugere que, nos períodos avaliados, a planta estava iniciando a sua reprogramação transcricional em resposta ao patógeno, sendo este tempo (até 24 h), portanto, curto, por se tratar de um genótipo suscetível à *Xap*.

No mesmo contexto, para identificar genes de couve-flor diferencialmente expressos em resposta a *X. campestris* pv. *campestris*, foi construída uma biblioteca SSH usando uma linhagem resistente ao patógeno como *tester* e uma suscetível como *driver*. Das 715 unisequências relacionadas a alguma função putativa, atribuídas a 11 categorias funcionais, os pesquisadores selecionaram 12 delas, de diferentes categorias, para a validação da biblioteca por qPCR. Estes transcritos tiveram seus níveis de expressão avaliados em ambos os genótipos, resistente e suscetível, inoculados com o patógeno às 0, 2, 6, 12, 24, 48, 72 e 96 h.a.i. Como resultado, a expressão desses genes na linhagem suscetível foi mais tardia, enquanto na resistente a reação foi mais rápida. Também, os níveis de expressão desses genes foram relativamente maiores na planta resistente em relação à suscetível. Desses, o gene LTP, que codifica uma proteína PR envolvida no transporte de lipídios, se destacou por apresentar alto nível de expressão em ambos os genótipos (até 60 X), a partir das 24 h.a.i. (JIANG et al., 2010). Estes resultados concordam com os aqui obtidos, uma vez que se observou baixa indução da expressão dos genes de maracujá azedo suscetível à *Xap* nos períodos iniciais da interação (*bulk* dos transcritos de 6, 12 e 24 h.a.i.).

Além disso, diversos outros trabalhos sobre interação planta-patógeno produziram resultados precisos e informativos ao validar genes obtidos em bibliotecas SSH por qPCR. Por exemplo, durante a interação ervilha-*Peronospora viciae* f. sp. *pisi* foram identificados 11 unisequências similares a genes de defesa.

Ao validar cinco desses genes por qPCR, todos tiveram um aumento de expressão no tratamento inoculado, com razões de expressão que variaram entre 3,6 X a 127,1 X, sendo este último valor referente ao transcrito que codifica a proteína de resistência a míldio RPP27 (FENG et al. 2012b). Também, entre os oito genes de algodão relacionados à resposta de defesa contra o fungo *Verticillium dahliae*, sete foram mais expressos no tratamento inoculado do genótipo resistente, comparativamente ao controle, na maioria dos tempos avaliados (6, 12, 24, 36, 48 e 96 h.a.i.), enquanto apenas três genes foram mais expressos no tratamento inoculado do genótipo suscetível, em um ou dois dos tempos avaliados (ZHANG et al. 2012). Os resultados deste último trabalho evidenciam que nem todos os genes de uma biblioteca subtrativa por SSH apresentam diferença de expressão entre os tratamentos, e que esses genes são fracamente induzidos nas primeiras horas de uma interação compatível, concordando com os resultados obtidos na presente pesquisa. Concluindo, na interação compatível maracujá azedo-*Xap* nem todos os genes validados por qPCR foram diferencialmente expressos, sendo que as diferenças de expressão foram baixas nos períodos iniciais da interação (*bulk* dos transcritos de 6, 12 e 24 h.a.i.).

Relembrando, a metodologia da hibridização subtrativa por supressão (SSH) foi escolhida neste trabalho por permitir a seleção (através de dois passos de hibridização), e a amplificação (por PCR) de fragmentos de cDNAs oriundos de transcritos presentes apenas no tratamento alvo, em relação a um tratamento controle. Dentre as possibilidades de uso deste método, ele permite identificar genes envolvidos nas interações planta-patógeno. Além disso, uma das principais vantagens da SSH, em relação a outros métodos subtrativos, é presença de uma etapa de normalização dos transcritos, na qual reduz a frequência de cDNAs altamente abundantes a níveis similares a cDNAs de genes pouco expressos. Isto facilita a identificação de transcritos pouco abundantes e aumenta a possibilidade de se selecionar os transcritos diferencialmente expressos. Esta etapa de normalização é particularmente importante, pois alguns genes ativados por stress são altamente transcritos pela indução de uma variedade de estímulos, mascarando importantes transcritos, como os relacionados à patogênese, que são expressos em baixos níveis e podem ter a sua expressão limitada a um tecido ou a um tipo de célula.

No entanto, relativamente ao método SSH, Rebrikov et al. (2000) propõem que podem ocorrer amplificações de transcritos não diferencialmente expressos, já que algumas moléculas de cDNA redundantes escapam, por acaso, durante a etapa de hibridização com o *driver* e, por isso, são amplificadas nas etapas de PCR. Embora os autores considerem esta possibilidade muito improvável, estimando que na maioria dos casos apenas uma única molécula de cada tipo de cDNA redundante esteja presente entre as milhares de outras moléculas de cDNA usadas na PCR, após a hibridização subtrativa, um grande excesso de sequências redundantes em relação ao cDNA alvo pode, cumulativamente, resultar em um alto número de cada molécula não diferencialmente expressa.

Dessa forma, explica-se a alta frequência de transcritos não diferencialmente expressos relatada em diversos trabalhos que desenvolveram bibliotecas SSH (CHENG et al., 2009; GUZZO; HAKAKAVA; TSAI, 2009; QI et al., 2010; DA et al., 2012). Porém, como nas bibliotecas maracujá azedo-*Xap*, não foi feita uma seleção diferencial para identificar os clones contendo transcritos diferencialmente expressos, a observação de transcritos com o mesmo nível de expressão, nos dois tratamentos, durante a validação das bibliotecas por qPCR, pode ser considerado um resultado esperado. Também, a observação de transcritos idênticos nas duas bibliotecas pode ser explicada pela permanência dessas moléculas redundantes.

Por outro lado, a reduzida proporção de transcritos diferencialmente expressos, identificada na primeira etapa de validação, pode ser explicada pela pouca quantidade de cDNA alvo e pela grande quantidade de cDNA redundante durante a construção das bibliotecas, levando a falhas durante as etapas de hibridização e produzindo um grande número de moléculas espúrias. Isto pressupondo que os tempos escolhidos (até 24 h após a inoculação) para a identificação de transcritos diferencialmente expressos, não tenham sido suficientes para uma resposta efetiva da planta, já que o genótipo de maracujá azedo se mostrou suscetível. Dessa forma, a planta pode ter induzido a transcrição de um baixo número de genes ou produzido uma baixa quantidade de cada transcrito ao responder a um fraco reconhecimento da presença do patógeno. Assim, sugere-se que os resultados obtidos na primeira etapa de validação das bibliotecas se devem a motivos biológicos e não a problemas na metodologia empregada.

Na segunda etapa de validação, a partir dos transcritos das folhas coletadas às 6 h e aos 5 e 9 dias após a inoculação, 21 dos 22 genes analisados apresentaram diferença na sua expressão entre os tratamentos (*Xap* versus controle), em pelo menos um dos períodos avaliados. Nesta etapa, as maiores diferenças de expressão ocorreram aos 5 e 9 d.a.i., tendo sido observados genes altamente expressos ou altamente reprimidos em resposta à presença do patógeno.

O gene que se destacou por apresentar um aumento drástico na sua expressão foi o que codifica a enzima lipoxigenase (nº 6, figura 36). Aos 5 d.a.i., houve um incremento na expressão da ordem de 477 X, em resposta à presença do patógeno, relativamente ao controle. Aos 9 d.a.i., houve uma redução na sua expressão, mas ainda assim, 265 X maior no tratamento com *Xap*. Interessantemente, às 6 h.a.i. não se observou diferença na sua expressão, evidenciando-se, portanto, uma resposta tardia da célula vegetal.

As lipoxigenases (LOX) são dioxigenases não-heme contendo ferro, que catalisam a incorporação de oxigênio molecular a ácidos graxos polinsaturados contendo um sistema (Z,Z)-1,4-pentadieno para produzir hidroperóxidos. O oxigênio pode ser adicionado a qualquer uma das extremidades do sistema pentadieno, sendo que, em plantas, os mais comuns são os ácidos linoleico e linolênico, produzindo os 9- e 13-hidroperóxidos. Estes produtos servem como precursores da síntese de oxilipinas, as quais apresentam diversas funções nas células animais e vegetais (PORTA; ROCHA-SOSA, 2002; BANNENBERG et al., 2008).

Por isso, em plantas, as lipoxigenases são classificadas como 9- e 13-lipoxigenases. Os produtos de suas vias desempenham diversas funções, como a defesa contra patógenos, os quais podem atuar na síntese de diferentes compostos com função sinalizadora, como o ácido jasmônico, no desencadeamento da resposta de hipersensibilidade (HR), bem como apresentar atividade antimicrobiana (FEUSSNER; WASTERACK, 2002; PORTA; ROCHA-SOSA, 2002; BANNENBERG et al., 2008).

Muitos trabalhos relatam o aumento da produção das lipoxigenases durante a patogênese. Em um estudo sobre expressão gênica, usando a técnica de microarranjos, em dois genótipos de arroz, um resistente e outro suscetível à *X. oryzae* pv. *oryzae*, Grewal, Gupta e Das (2012) identificaram uma lipoxigenase 10 X mais expressa no genótipo resistente em relação ao suscetível, 1 h.a.i. do patógeno.

Os autores concluíram que o ácido jasmônico, que é um dos produtos da via das lipoxigenases, tem importante função na sinalização durante a interação incompatível arroz-*Xoo*.

Em pimenta, foi observado um aumento da expressão do gene *9-LOX*, após o seu tratamento com dois isolados de *X. campestris* pv. *vesicatoria*, um virulento e outro avirulento. Na interação compatível, ou seja, com o isolado virulento, o pico de expressão foi mais tardio, após 20 h.a.i.; entretanto, na incompatível, o gene apresentou um pico de expressão a partir de 5 h.a.i., mantendo-se altamente expresso nos demais períodos avaliados (10, 15, 20 e 25 h.a.i.). Os autores concluíram que o gene *9-LOX* é altamente expresso durante a resposta de hipersensibilidade (HR) na interação incompatível pimenta-*Xanthomonas*, sugerindo que a LOX é essencial no estabelecimento dessa resposta e que a morte celular em decorrência da HR se deve, em parte, à peroxidação lipídica iniciada pela LOX (HWANG; HWANG, 2009).

Similarmente, em algodão, ao estudarem o gene *GhLOX2*, durante as interações com isolados de *X. campestris* pv. *malvacearum*, Sanier et al. (2012) observaram que durante a HR, na interação incompatível, o gene apresentou picos significativos de expressão. Aos 30 e 60 minutos após a inoculação observaram-se aumentos de 9,5 X e 5,9 X, e às 8 h.a.i. de 5,8 X, comparativamente ao controle. Na interação compatível, o gene se expressou 2,2, 0,8 e 1,6 X mais, respectivamente, aos mesmos tempos de interação. Os autores concluíram que a função da lipoxigenase na defesa do algodão contra *Xcm* está relacionada à HR, já que o aumento na expressão de *GhLOX2* ocorreu durante os eventos desta resposta.

No presente trabalho, o gene que codifica a enzima (+)-neomentol desidrogenase (nº 38, figura 36) também se destacou por apresentar aumento na sua expressão da ordem de 6 X e 29 X aos 5 e 9 d.a.i., respectivamente, em resposta ao patógeno. Esta enzima atua na via metabólica do neomentol, um monoterpreno de ação antimicrobiana do grupo dos terpenóides. Os terpenóides formam um diverso grupo de metabólitos secundários, os quais são conhecidos por terem sua síntese induzida em resposta ao ataque de patógenos ou estresse (RINGER et al., 2003; CHOI et al., 2008).

Em pimenta, observou-se que a expressão de *CaMNR1* ((1)-(3S)-neomenthol reductase) foi altamente induzida durante a interação incompatível com *X.*

campestris pv. *vesicatoria*, a partir de 5 h.a.i. Já na presença de um isolado virulento, *CaMNR1* foi fracamente induzido durante as primeiras horas de interação. As plantas de pimenta, que tiveram este gene silenciado, se tornaram suscetíveis à infecção pelos isolados avirulento e virulento, apresentando um baixo nível de expressão de genes relacionados à defesa, sugerindo que o produto do gene *CaMNR1* age como um fator de sinalização em ambas as respostas de imunidade, desencadeadas por PAMP ou por efetores (PTI ou ETI) (CHOI et al. 2008). Neste mesmo trabalho, o nível de expressão de *CaMNR1* foi correlacionado com a expressão de genes PR responsivos à ácido salicílico e ácido jasmônico, em pimenta e *Arabidopsis*. Sugere-se que a *CaMNR1* ((1)-(3S)-neomenthol reductase) está envolvida em uma ampla variedade de respostas de defesa, através da regulação da expressão de genes de defesa.

Já o gene da enzima quitinase (nº 10, figura 36) apresentou expressão reprimida às 6 h.a.i., sendo 1,67 X menos expresso devido à presença do patógeno. Já aos 5 d.a.i., a expressão foi altamente induzida (17,35 X) em resposta à *Xap*; porém, não foram observadas diferenças de expressão aos 10 d.a.i. Quitinases catalisam a degradação da quitina, uma proteína presente em diversos tipos de organismos, principal componente da parede celular de fungos. Nas plantas, a indução da expressão do gene da quitinase acontece quando elas são atacadas por patógenos, pois exerce atividade antimicrobiana direta ao degradar a quitina da parede celular de fungos patogênicos, inibindo o seu desenvolvimento, e os oligossacarídeos resultantes da degradação da quitina atuam como elicitores do sistema de defesa da planta. Apesar das quitinases serem classificadas como proteínas relacionadas à patogênese (PR-3), nem todas exercem função de defesa (KEZUKA et al., 2010; GROVER, 2012; YEOH et al., 2013).

Durante a interação incompatível mandioca-*X. axonopodis* pv. *Manihotis*, duas quitinases foram identificadas entre os 199 genes diferencialmente expressos em resposta a *Xam*. Oito genes foram validados por qPCR, que foram igualmente avaliados na interação compatível mandioca-*Xam*. Na interação compatível, as duas quitinases apresentaram pouca diferença de expressão nas primeiras horas de interação (12, 24 e 48 h.a.i.). Após 15 dias, a hevamine quitinase e a quitinase classe III foram altamente induzidas (18 X e 50 X, respectivamente) nas plantas inoculadas. Na interação incompatível, a hevamine quitinase foi altamente induzida

em todos os períodos avaliados (12, 24 e 48 h.a.i. e 7 e 15 d.a.i.) e a quitinase classe III apresentou um pico de expressão de aproximadamente 180 X aos 7 d.a.i. e de 30 X aos 15 d.a.i. (LOPEZ et al. 2005).

Embora na presente pesquisa e em outras relacionadas (SANA et al., 2010; DAURELIO et al., 2011), incluindo o trabalho sobre a interação mandioca-*Xam* (LOPEZ et al. 2005), se observou a indução de quitinases em resposta a bactérias patogênicas, a função dessas enzimas durante a interação planta-bactéria ainda não está elucidada.

O gene da proteína MAP-quinase 2 (nº 12, figura 36) teve seu nível de expressão aumentado em 1,55 X e 1,93 X aos 5 e 9 dias de interação com *Xap*, respectivamente, enquanto o gene da proteína MAP-quinase 4 (nº 11, figura 36) não apresentou diferença de expressão às 6 h.a.i. e aos 5 d.a.i. e foi reprimido pelo patógeno aos 9 d.a.i. Estas proteínas estão associadas à resposta a estresses bióticos e abióticos e fazem parte da cascata de sinalização das MAP-quinases, desencadeada pelo reconhecimento de patógenos (QIU et al., 2008b).

Em uma biblioteca de expressão construída para caracterizar a interação incompatível *Fortunella margarita*-*X. citri* subsp. *citri*, foi identificado um gene de uma MAP-quinase 3, o qual, durante a validação por qPCR, se mostrou 5,66 X e 14,1 X mais expresso na planta inoculada em relação à controle, respectivamente às 6 e 24 h.a.i. Já aos 3 e 5 d.a.i., a razão de expressão da MAP-quinase 3 foi reduzida a 1,22 X e 0,58 X, respectivamente (KHALAF et al. 2011). Em outra biblioteca de expressão, foi identificado um grande número de genes diferencialmente expressos durante a interação incompatível tomate-*X. campestris* pv. *vesicatoria*. No grupo de genes que apresentou aumento de expressão gradual até às 12 h.a.i. e leve redução da expressão nos tempos subsequentes, estavam inclusos genes de proteínas MAP-quinases (GIBLY et al. 2004).

Concluindo, a ação das MAP-quinases se dá nas horas iniciais após a infecção, quando a planta é resistente ao patógeno, contrariamente aos dados obtidos no presente trabalho, no qual o reconhecimento do patógeno se mostrou tardio. Como discutido anteriormente, a MAP-quinase 2 ativa a MAP-quinase 4, durante a cascata de sinalização em resposta a estresse biótico. Possivelmente, no nosso estudo, o produto do gene da MAP-quinase 4 tenha sido alvo do patógeno, por isso detectou-se uma inibição da sua expressão.

Os genes da peroxidase 12 e da glutamato-cisteína ligase (nº 30 e 35, respectivamente, figura 36) também tiveram um aumento de expressão em resposta à *Xap*; as razões de expressão desses genes foram 1,78 X e 2,80 X aos 5 d.a.i. e 3,10 X e 2,27 X aos 9 d.a.i., respectivamente; porém, o gene da glutamato-cisteína ligase foi reprimido às 6 h.a.i., se expressando 2,17 X menos no tratamento com *Xap*. Peroxidades vegetais da classe III estão entre as proteínas induzidas durante as respostas de defesa e desempenham importante papel em várias reações metabólicas, catalisando a oxidação de um único elétron de diversos substratos usando para isso o H_2O_2 ; também são consideradas PR e estão envolvidas no metabolismo da auxina, lignina e formação da suberina, reticulação de componentes da parede celular, síntese de fitoalexina e metabolismo de ERO e ERN (espécies reativas de nitrogênio) (ALMAGRO et al., 2009). Já a enzima glutamato-cisteína ligase atua na via metabólica da glutatona, sendo que o metabolismo da glutatona pode regular a expressão de genes de defesa (BALL et al., 2004).

Dentre os diversos genes que tiveram expressão induzida em resposta à *X. citri* subsp. *citri*, identificados em uma biblioteca de expressão, uma peroxidase 12 foi 11,37 X mais expressa em *Fortunella crassifolia*, espécie resistente ao cancro cítrico, comparativamente ao controle, 5 d.a.i.. Porém, em *Citrus sinensis*, que é suscetível à doença, não houve indução da expressão desse gene em resposta ao patógeno; também, outros genes relacionados ao processo de ligação a cátions apresentaram o mesmo perfil de expressão, levando à sugestão de que esse processo pode ser parcialmente responsável pela diferença na tolerância ao cancro dessas duas espécies (FU et al. 2012). Isto contraria o observado na interação maracujá azedo-*Xap*, já que o gene da peroxidase apresentou aumento de expressão aos 5 e 9 d.a.i.

No trabalho sobre a interação tomate-*X. campestris* pv. *vesicatoria*, mencionado anteriormente, foi observado um grupo de 36 genes de tomate, incluindo os da peroxidase e catalase, com expressão inibida por *Xcv* já no início da infecção. O nível de expressão médio deste grupo de genes começou a diminuir 8 h.a.i., atingindo expressão mínima às 12 h.a.i, e ligeiro aumento às 18 h.a.i. (GIBLY et al. 2004). Contrariamente, o gene da catalase identificado na presente biblioteca (nº 13, figura 36) apresentou inibição da sua expressão apenas no 9º dia de interação com *Xap*, em uma razão 7,7 X menor relativamente ao controle. Isto mostra que este gene não teve atuação na resposta de defesa à *Xap* e passou a ser

inibido pelo patógeno nos estágios mais avançados da interação, já quando ocorrem indícios da doença.

Da mesma forma que o gene da catalase, outros genes também tiveram sua expressão reprimida durante a interação com *Xap*. O gene da proteína potencializadora envolvendo oxigênio 1-1 (*Oxygen-evolving enhancer protein 1-1, chloroplastic*) (nº 2, figura 36), apresentou inibição de expressão já às 6 h.a.i (1,67 X), sendo que aos 9 d.a.i. apresentou uma alta inibição (10 X) sob tratamento com *Xap*. Este gene foi o que apresentou maior inibição de sua expressão dentre os analisados. Esta proteína faz parte do complexo cloroplastidial fotossistema II, estabilizando o grupo Mn, que é o principal local de separação da água durante a fotossíntese, sendo necessária para aumentar o nível de oxigênio fotossintético (MARQUÉS; DURAN-VILA; DARÒS, 2011).

Outro gene relacionado ao processo fotossintético cuja expressão se mostrou reprimida foi o da fosforribuloquinase (nº 3, figura 36). Às 6 h.a.i. e aos 5 d.a.i. de interação maracujá azedo-*Xap*, a redução da expressão foi de 1,78 X e 1,92, respectivamente; já aos 9 d.a.i., a inibição foi maior, com uma redução de 9,1 X. A enzima cloroplastidial fosforribuloquinase faz parte do ciclo de Calvin, atuando na via de redução das pentoses fosfato, responsável pela fixação do CO₂ (MICHELS; WEDEL; KROTH, 2005).

Em uma análise do proteoma de folhas de *Arabidopsis* infectadas por *Pseudomonas syringae*, foi identificada uma proteína do complexo fotossistema II, que diminuiu significativamente 1,5 h.a.i. e uma fosforribuloquinase, que diminuiu durante o reconhecimento de PAMP, às 6 h.a.i.; os autores consideraram este comportamento relacionado à modulação transcricional da maquinaria fotossintética relatado durante diversas respostas de defesa (JONES et al. 2006). Também, o gene da fosforribuloquinase, bem como outros genes de proteínas envolvidas no Ciclo de Calvin, que foram observados em uma biblioteca de expressão construída usando cDNA-AFLP, tiveram expressão inibida durante a interação compatível trigo-*Puccinia striiformis* f. sp. *tritici*, em todos os períodos avaliados (12, 18, 24, 48, 72 e 120 h.a.i.), após validação por qPCR (WANG et al. 2009). Esses resultados estão de acordo com os obtidos na presente pesquisa, sendo que a inibição dos genes da proteína potencializadora envolvendo oxigênio 1-1 e o da fosforribuloquinase sugere

que os processos envolvidos na fotossíntese foram altamente afetados durante a interação com *Xap*.

Além disso, o gene da glutamina sintetase (nº 16, figura 36) foi inibido em todos os períodos de avaliação, com pico máximo de inibição aos 9 d.a.i., se expressando 7,14 X menos na planta tratada com *Xap*. Esta enzima detoxifica a amônia gerada pela fotorrespiração, que é um metabólito altamente reativo, catalisando a conversão, dependente de glutamato e ATP, da amônia tóxica para glutamina não tóxica (TAIRA et al., 2004).

A enzima glutamina sintetase (GS1) foi usada como um dos marcadores durante um estudo sobre a manutenção do N em tabaco sob diferentes tipos de estresses (PAGEAU et al. 2006). O perfil de expressão da enzima GS1, avaliado por *Northern blot*, em folhas inoculadas com três isolados de *P. syringae*, sendo dois avirulentos e um virulento, mostrou indução da expressão nas plantas inoculadas com os isolados avirulentos, em todos os períodos avaliados (2, 5, 9 e 21 h.a.i. e 5 d.a.i.); já nas plantas inoculadas com o isolado virulento, a expressão não foi induzida nas primeiras horas, mas apresentou uma leve inibição aos 5 d.a.i. Isto indica o envolvimento desse gene na resposta de defesa durante as interações incompatíveis, sendo alvo do patógeno durante as interações compatíveis, tal como no presente caso.

Os resultados dessa etapa de validação mostram que nas primeiras horas de interação (6 h.a.i.), período crítico para o reconhecimento do patógeno, a planta ainda não estabelece uma resposta eficiente. Aos 5 dias de interação sugere-se que o sistema de defesa da planta esteja mais ativo, pois vários genes tiveram sua expressão induzida, enquanto apenas dois foram reprimidos pelo patógeno. Já aos 9 dias de interação, há uma maior alteração transcricional nas folhas de maracujá azedo, notando-se genes com expressão altamente reprimida apenas neste período. Neste caso, possivelmente a maquinaria transcricional começou a entrar em colapso, uma vez que deste período em diante a doença entra em um estágio mais avançado, levando à clorose e necrose dos tecidos.

6 CONCLUSÕES

Do ponto de vista metodológico pode-se concluir que:

- (a) o método de hibridização subtrativa por supressão (SSH) se presta para identificar genes de maracujá azedo diferencialmente expressos durante a interação com *Xanthomonas axonopodis* pv. *passiflorae* (*Xap*);
- (b) A plataforma PLAZA e o programa computacional Blast2GO são adequados para a anotação automática de sequências expressas de maracujá azedo;
- (c) A técnica qPCR é sensível e se presta para a quantificação relativa da expressão de genes de maracujá azedo diferencialmente expressos em resposta à *Xap*;

Do ponto de vista da interação maracujá azedo-*Xap* pode-se concluir que:

- (a) A reprogramação transcricional nas folhas de maracujá azedo em resposta à infecção pelo patógeno inclui genes envolvidos em diversos processos biológicos e que desempenham diferentes funções nas células, sendo que os picos de reprogramação transcricional são tardios, ou seja, ocorrem aos 5 e 9 dias após a inoculação;
- (b) Os genes de maracujá azedo mais expressos em resposta à *Xap* codificam as enzimas lipoxigenase (477 x), (+)-neomentol desidrogenase (29 X) e quitinase (17 X), às quais são sabidamente relacionadas às respostas de defesa vegetal;
- (c) Os genes de maracujá azedo cuja expressão é mais reprimida em resposta à *Xap* codificam uma proteína potencializadora envolvendo oxigênio e uma fosforribuloquinase, ambas relacionadas à fotossíntese, como também uma glutamina sintetase e uma catalase, envolvidas na detoxificação da amônia e do H₂O₂, respectivamente.

REFERÊNCIAS

- ABRAMOVITCH, R. B.; MARTIN, G. B. Strategies used by bacterial pathogens to suppress plant defenses. **Current Opinion in Plant Biology**, London, v. 7, n. 4, p. 356–364, 2004.
- ABREU, P.P.; SOUZA, M.M.; SANTOS, E.A.; PIRES, M.V.; PIRES, M.M.; ALMEIDA, A.A.F. Passion flower hybrids and their use in the ornamental plant market: perspectives for sustainable development with emphasis on Brazil. **Euphytica**, Wageningen, v. 166, p. 307-315, 2009.
- AFZAL, A. J.; WOOD, A. J.; LIGHTFOOT, D. A. Plant Receptor-Like Serine Threonine Kinases: Roles in Signaling and Plant Defense. **Molecular Plant-Microbe Interactions**, Saint Paul, v. 21, n. 5, p. 507–517, 2008.
- AGRIANUAL 2012. **Anuário da agricultura brasileira**. São Paulo: FNP - Consultoria e Comércio, 2011. p. 482.
- AGRIOS, G. N. **Plant pathology**. 5 ed. Boston: Elsevier Academic Press, 2005. 922 p.
- ALMAGRO, L.; GOMEZ ROS, L. V.; BELCHI-NAVARRO, S.; BRU, R.; ROS BARCELO, A.; PEDRENO, M. A. Class III peroxidases in plant defence reactions. **Journal of Experimental Botany**, Oxford, v. 60, n. 2, p. 377–390, 2009.
- ALTSCHUL, S. F.; GISH, W.; MILLER, W.; MYERS, E. W.; LIPMAN, D. J. Basic local alignment search tool. **Journal of molecular biology**, London, v. 215, n. 3, p. 403–410, 1990.
- ANDERSEN, C. L.; JENSEN, J. L.; ØRNTTOFT, T. F. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. **Cancer research**, Baltimore, v. 64, n. 15, p. 5245–5250, 2004.
- ANDREASSON, E.; JENKINS, T.; BRODERSEN, P.; THORGRIMSEN, S.; PETERSEN, N. H. T.; ZHU, S.; QIU, J.-L.; MICHEELSEN, P.; ROCHER, A.; PETERSEN, M.; NEWMAN, M.-A.; BJØRN NIELSEN, H.; HIRT, H.; SOMSSICH, I.; MATTSSON, O.; MUNDY, J. The MAP kinase substrate MKS1 is a regulator of plant defense responses. **The EMBO Journal**, Oxford, v. 24, n. 14, p. 2579–2589, 2005.
- ANTOLÍN-LLOVERA, M.; RIED, M. K.; BINDER, A.; PARNISKE, M. Receptor Kinase Signaling Pathways in Plant-Microbe Interactions. **Annual Review of Phytopathology**, Palo Alto, v. 50, n. 1, p. 451–473, 2012.
- BALL, L.; ACCOTTO, G.-P.; BECHTOLD, U.; CREISSEN, G.; FUNCK, D.; JIMENEZ, A.; KULAR, B.; LEYLAND, N.; MEJIA-CARRANZA, J.; REYNOLDS, H.; KARPINSKI, S.; MULLINEAUX, P. M. Evidence for a Direct Link between Glutathione Biosynthesis and Stress Defense Gene Expression in *Arabidopsis*. **The Plant Cell Online**, Rockville, v. 16, n. 9, p. 2448–2462, 2004.

BANNENBERG, G.; MARTÍNEZ, M.; HAMBERG, M.; CASTRESANA, C. Diversity of the Enzymatic Activity in the Lipxygenase Gene Family of *Arabidopsis thaliana*. **Lipids**, Champaign, v. 44, n. 2, p. 85–95, 2008.

BELLIN, D.; ASAI, S.; DELLEDONNE, M.; YOSHIOKA, H. Nitric Oxide as a Mediator for Defense Responses. **Molecular Plant-Microbe Interactions**, Saint Paul, v. 26, n. 3, p. 271–277, 2013.

BERNACCI, L.C.; CERVI, A.C.; MILWARD-DE-AZEVEDO, M.A.; NUNES, T.S.; IMIG, D.C.; MEZZONATO, A.C. **Passifloraceae**. Disponível em: <<http://floradobrasil.jbrj.gov.br/jabot/floradobrasil/FB182>>. Acesso em: 30 jun. 2013.

BINNECK, E. As ômicas: Integrando a bioinformação. **Biotecnologia Ciência & Desenvolvimento**, Brasília, v. 32, p. 28–37, 2004.

BOCH, J.; BONAS, U. *Xanthomonas* AvrBs3 Family-Type III Effectors: Discovery and Function. **Annual Review of Phytopathology**, Palo Alto, v. 48, n. 1, p. 419–436, 2010.

BOLLER, T.; HE, S. Y. Innate Immunity in Plants: An Arms Race Between Pattern Recognition Receptors in Plants and Effectors in Microbial Pathogens. **Science**, Washington, v. 324, n. 5928, p. 742–744, 2009.

BORO, M. C.; BERIAM, L. O. S.; GUZZO, S. D. Induced resistance against *Xanthomonas axonopodis* pv. *passiflorae* in passion fruit plants. **Tropical Plant Pathology**, Brasília, v. 36, p. 74–80, 2011.

BOUDSOCQ, M.; WILLMANN, M. R.; MCCORMACK, M.; LEE, H.; SHAN, L.; HE, P.; BUSH, J.; CHENG, S.-H.; SHEEN, J. Differential innate immune signalling via Ca^{2+} sensor protein kinases. **Nature**, London, v. 464, n. 7287, p. 418–422, 2010.

BUSTIN, S. A.; BENES, V.; GARSON, J. A.; HELLEMANS, J.; HUGGETT, J.; KUBISTA, M.; MUELLER, R.; NOLAN, T.; PFAFFL, M. W.; SHIPLEY, G. L.; VANDESOMPELE, J.; WITTEWER, C. T. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. **Clinical Chemistry**, Baltimore, v. 55, n. 4, p. 611–622, 2009.

CAO, A.; XING, L.; WANG, X.; YANG, X.; WANG, W.; SUN, Y.; QIAN, C.; NI, J.; CHEN, Y.; LIU, D.; WANG, X.; CHEN, P. From the Cover: Serine/threonine kinase gene Stpk-V, a key member of powdery mildew resistance gene Pm21, confers powdery mildew resistance in wheat. **Proceedings of the National Academy of Sciences**, Washington, v. 108, n. 19, p. 7727–7732, 2011.

CHENG, S.-H. Calcium Signaling through Protein Kinases. The *Arabidopsis* Calcium-Dependent Protein Kinase Gene Family. **Plant Physiology**, Minneapolis, v. 129, n. 2, p. 469–485, 2002.

CHENG, Z.; TANG, K.; YAN, H.; FU, J.; YING, F.; HUANG, X. Analysis of differentially expressed genes in *Oryza meyeriana* in response to infection by

Xanthomonas oryzae pv. *oryzae*. **Plant Breeding**, Farnham Royal, v. 129, n. 4, p. 393–399, 2009.

CHISHOLM, S. T.; COAKER, G.; DAY, B.; STASKAWICZ, B. J. Host-Microbe Interactions: Shaping the Evolution of the Plant Immune Response. **Cell**, Cambridge, v. 124, n. 4, p. 803–814, 2006.

CHOI, H. W.; LEE, B. G.; KIM, N. H.; PARK, Y.; LIM, C. W.; SONG, H. K.; HWANG, B. K. A Role for a Menthone Reductase in Resistance against Microbial Pathogens in Plants. **Plant Physiology**, Minneapolis, v. 148, n. 1, p. 383–401, 2008.

CONESA, A.; GOTZ, S.; GARCIA-GOMEZ, J. M.; TEROL, J.; TALON, M.; ROBLES, M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. **Bioinformatics**, Oxford, v. 21, n. 18, p. 3674–3676, 2005.

CROUZET, J.; ROLAND, J.; PEETERS, E.; TROMBIK, T.; DUCOS, E.; NADER, J.; BOUTRY, M. NtPDR1, a plasma membrane ABC transporter from *Nicotiana tabacum*, is involved in diterpene transport. **Plant Molecular Biology**, Dordrecht, v. 82, n. 1-2, p. 181–192, 2013.

CURSINO-SANTOS, J.R.; DEFINA, T.P.A; MARTINEZ-ROSSI, N.M. **Os segredos das plantas e de seus patógenos na era molecular**. Ribeirão Preto: Sociedade Brasileira de Genética, 2003. 45 p.

DA, X.; YU, K.; SHEN, S.; ZHANG, Y.; WU, J.; YI, H. Identification of differentially expressed genes in a spontaneous altered leaf shape mutant of the navel orange [*Citrus sinensis* (L.) Osbeck]. **Plant Physiology and Biochemistry**, Paris, v. 56, p. 97–103, jul. 2012.

DAURELIO, L. D.; PETROCELLI, S.; BLANCO, F.; HOLUIGUE, L.; OTTADO, J.; ORELLANO, E. G. Transcriptome analysis reveals novel genes involved in nonhost response to bacterial infection in tobacco. **Journal of Plant Physiology**, Stuttgart, v. 168, n. 4, p. 382–391, 2011.

DENG, J.; ZHOU, Y.; BAI, M.; LI, H.; LI, L. Anxiolytic and sedative activities of *Passiflora edulis* f. *flavicarpa*. **Journal of Ethnopharmacology**, Lausanne, v. 128, n. 1, p. 148–153, 2010.

DESLANDES, L.; RIVAS, S. Catch me if you can: bacterial effectors and plant targets. **Trends in Plant Science**, Oxford, v. 17, n. 11, p. 644–655, 2012.

DEVOTO, A.; HARTMANN, H. A.; PIFFANELLI, P.; ELLIOTT, C.; SIMMONS, C.; TARMINO, G.; GOH, C.-S.; COHEN, F. E.; EMERSON, B. C.; SCHULZE-LEFERT, P.; PANSTRUGA, R. Molecular Phylogeny and Evolution of the Plant-Specific Seven-Transmembrane MLO Family. **Journal of Molecular Evolution**, New York, v. 56, n. 1, p. 77–88, 2003.

DIATCHENKO, L.; LAU, Y. F.; CAMPBELL, A. P.; CHENCHIK, A.; MOQADAM, F.; HUANG, B.; LUKYANOV, S.; LUKYANOV, K.; GURSKAYA, N.; SVERDLOV, E. D.; SIEBERT, P. D. Suppression subtractive hybridization: a method for generating

differentially regulated or tissue-specific cDNA probes and libraries. **Proceedings of the National Academy of Sciences**, Washington, v. 93, n. 12, p. 6025–6030, 1996.

DIE, J. V.; ROMÁN, B.; NADAL, S.; GONZÁLEZ-VERDEJO, C. I. Evaluation of candidate reference genes for expression studies in *Pisum sativum* under different experimental conditions. **Planta**, Berlin, v. 232, n. 1, p. 145–153, 2010.

DMITRIEV, A. P. Signal Molecules for Plant Defense Responses to Biotic Stress. **Russian Journal of Plant Physiology**, New York, v. 50, p. 417–425, 2003.

DODDS, P. N.; RATHJEN, J. P. Plant immunity: towards an integrated view of plant–pathogen interactions. **Nature Reviews Genetics**, London, v. 11, n. 8, p. 539–548, 2010.

DOYLE, J. J.; DOYLE, J. L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. **Phytochemical Bulletin**, Irvine, v. 19, p. 11–15, 1987.

DUBY, G.; BOUTRY, M. The plant plasma membrane proton pump ATPase: a highly regulated P-type ATPase with multiple physiological roles. **Pflügers Archiv - European Journal of Physiology**, Berlin, v. 457, n. 3, p. 645–655, 2009.

DURHAM, A. M.; KASHIWABARA, A. Y.; MATSUNAGA, F. T. G.; AHAGON, P. H.; RAINONE, F.; VARUZZA, L.; GRUBER, A. EGene: a configurable pipeline generation system for automated sequence analysis. **Bioinformatics**, Oxford, v. 21, n. 12, p. 2812–2813, 2005.

ELMORE, J. M.; COAKER, G. The Role of the Plasma Membrane H⁺-ATPase in Plant-Microbe Interactions. **Molecular Plant**, Saint Paul, v. 4, n. 3, p. 416–427, 2011.

EULGEM, T. Regulation of the *Arabidopsis* defense transcriptome. **Trends in Plant Science**, Oxford, v. 10, n. 2, p. 71–78, 2005.

EWING, B.; GREEN, P. Base-calling of automated sequencer traces using phred. II. Error probabilities. **Genome research**, Cold Spring Harbor, v. 8, n. 3, p. 186–194, 1998.

EWING, B.; HILLIER, L.; WENDL, M. C.; GREEN, P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. **Genome research**, Cold Spring Harbor, v. 8, n. 3, p. 175–185, 1998.

FENG, F.; YANG, F.; RONG, W.; WU, X.; ZHANG, J.; CHEN, S.; HE, C.; ZHOU, J.-M. A *Xanthomonas* uridine 5'-monophosphate transferase inhibits plant immune kinases. **Nature**, London, v. 485, n. 7396, p. 114–118, 2012a.

FENG, J.; CHANG, K. F.; HWANG, S. F.; STRELKOV, S. E.; CONNER, R. L.; GOSSEN, B. D.; MCLAREN, D. L.; CHEN, Y. Y. Analysis of expressed sequence tags derived from pea leaves infected by *Peronospora viciae* f. sp. *pisi*. **Annals of Applied Biology**, Warwick, v. 161, n. 3, p. 214–222, 2012b.

FENG, J.; HWANG, S.-F.; STRELKOV, S. E. Analysis of expressed sequence tags derived from a compatible *Plasmodiophora brassicae*–canola interaction. **Canadian Journal of Plant Pathology**, Ontario, v. 34, n. 4, p. 562–574, 2012c.

FERNANDEZ-SUAREZ, X. M.; GALPERIN, M. Y. The 2013 Nucleic Acids Research Database Issue and the online Molecular Biology Database Collection. **Nucleic Acids Research**, Oxford, v. 41, n. D1, p. D1–D7, 2012.

FEUSSNER, I.; WASTERACK, C. The Lipoxygenase Pathway. **Annual Review of Plant Biology**, Palo Alto, v. 53, n. 1, p. 275–297, 2002.

FINLAY, B. B.; FALKOW, S. Common themes in microbial pathogenicity revisited. **Microbiology and molecular biology reviews**, New York, v. 61, n. 2, p. 136–169, 1997.

FU, X.Z.; GONG, X.Q.; ZHANG, Y.X.; WANG, Y.; LIU, J.H. Different Transcriptional Response to *Xanthomonas citri* subsp. *citri* between Kumquat and Sweet Orange with Contrasting Canker Tolerance. **Plos One**, San Francisco, v. 7, n. 7, p. e41790, 2012.

FUJIWARA, T.; MAISONNEUVE, S.; ISSHIKI, M.; MIZUTANI, M.; CHEN, L.; WONG, H. L.; KAWASAKI, T.; SHIMAMOTO, K. Sekiguchi Lesion Gene Encodes a Cytochrome P450 Monooxygenase that Catalyzes Conversion of Tryptamine to Serotonin in Rice. **Journal of Biological Chemistry**, Bethesda, v. 285, n. 15, p. 11308–11313, 2010.

GAO, M.; WANG, Q.; WAN, R.; FEI, Z.; WANG, X. Identification of genes differentially expressed in grapevine associated with resistance to *Elsinoe ampelina* through suppressive subtraction hybridization. **Plant Physiology and Biochemistry**, Paris, v. 58, p. 253–268, 2012.

GIBLY, A.; BONSHIEN, A.; BALAJI, V.; DEBBIE, P.; MARTIN, G. B.; SESSA, G. Identification and Expression Profiling of Tomato Genes Differentially Regulated During a Resistance Response to *Xanthomonas campestris* pv. *vesicatoria*. **Molecular Plant-Microbe Interactions**, Saint Paul, v. 17, n. 11, p. 1212–1222, 2004.

GREWAL, R. K.; GUPTA, S.; DAS, S. *Xanthomonas oryzae* pv *oryzae* triggers immediate transcriptomic modulations in rice. **BMC Genomics**, London, v. 13, n. 1, p. 49, 2012.

GROVER, A. Plant Chitinases: Genetic Diversity and Physiological Roles. **Critical Reviews in Plant Sciences**, London, v. 31, n. 1, p. 57–73, 2012.

GUTIERREZ, L.; MAURIAT, M.; PELLOUX, J.; BELLINI, C.; WUYTSWINKEL, O. VAN. Towards a Systematic Validation of References in Real-Time RT-PCR. **The Plant Cell Online**, Rockville, v. 20, n. 7, p. 1734–1735, 2008.

GUZZO, S. D.; HAKAKA, R.; TSAI, S. M. Identification of Coffee Genes Expressed During Systemic Acquired Resistance and Incompatible Interaction with

Hemileia vastatrix. **Journal of Phytopathology**, Berlin, v. 157, n. 10, p. 625–638, 2009.

HEATH, M. C. Hypersensitive response-related death. **Plant Molecular Biology**, Dordrecht, v. 44, p. 321–334, 2000.

HUANG, X.; MADAN, A. CAP3: A DNA Sequence Assembly Program. **Genome Research**, Cold Spring Harbor, v. 9, n. 9, p. 868–877, 1999.

HWANG, I. S.; HWANG, B. K. The Pepper 9-Lipoxygenase Gene CaLOX1 Functions in Defense and Cell Death Responses to Microbial Pathogens. **Plant Physiology**, Minneapolis, v. 152, n. 2, p. 948–967, 2009.

JEZ, J. M.; CAHOON, R. E.; CHEN, S. *Arabidopsis thaliana* Glutamate-Cysteine Ligase. **Journal of Biological Chemistry**, Bethesda, v. 279, n. 32, p. 33463–33470, 2004.

JIANG, H.; SONG, W.; LI, A.; YANG, X.; SUN, D. Identification of genes differentially expressed in cauliflower associated with resistance to *Xanthomonas campestris* pv. *campestris*. **Molecular Biology Reports**, Dordrecht, v. 38, n. 1, p. 621–629, 2010.

JONES, A. M. E.; THOMAS, V.; BENNETT, M. H.; MANSFIELD, J.; GRANT, M. Modifications to the *Arabidopsis* Defense Proteome Occur Prior to Significant Transcriptional Change in Response to Inoculation with *Pseudomonas syringae*. **Plant Physiology**, Minneapolis, v. 142, n. 4, p. 1603–1620, 2006.

JONES, J. D. G.; DANGL, J. L. The plant immune system. **Nature**, London, v. 444, n. 7117, p. 323–329, 2006.

JUNQUEIRA, N.T.V.; ANJOS, J.R.N.; JUNQUEIRA, L.P.; SHARMA, R.D. Principais doenças e pragas. In: MANICA, I.; BRANCHER, A.; SANZONOWICZ, C.; ICUMA, I.M.; AGUIAR, J.L.P.; AZEVEDO, J.A. VASCONCELLOS, M.A.S.; JUNQUEIRA, N.T.V. (Ed). **Maracujá-doce**. Tecnologia de produção, pós-colheita e mercado. Porto Alegre: Cinco Continentes, 2005. p. 113-144.

JUNQUEIRA, N. T. V.; JUNQUEIRA, K. P. Manejo das principais doenças do maracujazeiro. In: SUSSEL, A. A. B.; MEDEIROS, F. H. V.; RIBEIRO JÚNIOR, P. M.; UCHOA, C. N.; AMARAL, D. R.; MEDEIROS, F. C. L.; PEREIRA, R. B.; SANTOS, J.; LIMA, L. M.; ROSWALKA, L. C. (Ed) **Manejo integrado de doenças de fruteiras**. Brasília: Sociedade Brasileira de Fitopatologia, 2007. p. 87-105

JURCA, M. E.; BOTTKA, S.; FEHÉR, A. Characterization of a family of *Arabidopsis* receptor-like cytoplasmic kinases (RLCK class VI). **Plant Cell Reports**, Berlin, v. 27, n. 4, p. 739–748, 2007.

KAKAR, K.; WANDREY, M.; CZECHOWSKI, T.; GAERTNER, T.; SCHEIBLE, W.-R.; STITT, M.; TORRES-JEREZ, I.; XIAO, Y.; REDMAN, J. C.; WU, H. C.; CHEUNG, F.; TOWN, C. D.; UDVARDI, M. K. A community resource for high-throughput quantitative RT-PCR analysis of transcription factor gene expression in *Medicago truncatula*. **Plant Methods**, London, v. 4, n. 1, p. 18, 2008.

KALLBERG, Y.; OPPERMAN, U.; JÖRNVALL, H.; PERSSON, B. Short-chain dehydrogenases/reductases (SDRs). **European Journal of Biochemistry**, Berlin, v. 269, n. 18, p. 4409–4417, 2002.

KAY, S.; BONAS, U. How *Xanthomonas* type III effectors manipulate the host plant. **Current Opinion in Microbiology**, Oxford, v. 12, n. 1, p. 37–43, 2009.

KEMMERLING, B.; SCHWEDT, A.; RODRIGUEZ, P.; MAZZOTTA, S.; FRANK, M.; QAMAR, S. A.; MENGISTE, T.; BETSUYAKU, S.; PARKER, J. E.; MÜSSIG, C.; THOMMA, B. P. H. J.; ALBRECHT, C.; VRIES, S. C. DE; HIRT, H.; NÜRNBERGER, T. The BRI1-Associated Kinase 1, BAK1, Has a Brassinolide-Independent Role in Plant Cell-Death Control. **Current Biology**, London, v. 17, n. 13, p. 1116–1122, 2007.

KEZUKA, Y.; KOJIMA, M.; MIZUNO, R.; SUZUKI, K.; WATANABE, T.; NONAKA, T. Structure of full-length class I chitinase from rice revealed by X-ray crystallography and small-angle X-ray scattering. **Proteins: Structure, Function, and Bioinformatics**, New York, v. 78, n. 10, p. 2295–2305, 2010.

KHALAF, A. A.; GMITTER, F. G.; CONESA, A.; DOPAZO, J.; MOORE, G. A. *Fortunella margarita* Transcriptional Reprogramming Triggered by *Xanthomonas citri* subsp. *citri*. **BMC Plant Biology**, London, v. 11, n. 1, p. 159, 2011.

KIM, D. S.; HWANG, B. K. The pepper receptor-like cytoplasmic protein kinase CaPIK1 is involved in plant signaling of defense and cell-death responses. **The Plant Journal**, Oxford, v. 66, n. 4, p. 642–655, 2011.

KIMATI, H.; AMORIM, L.; REZENDE, J.A.M.; BERGAMIM FILHO, A. CARMARGO, L.E.A. (Ed). **Manual de fitopatologia**. Doenças das plantas cultivadas. 4.ed. v. 2. São Paulo: Agronômica Ceres, 2005. 663 p.

KOU, Y.; WANG, S. Broad-spectrum and durability: understanding of quantitative disease resistance. **Current Opinion in Plant Biology**, London, v. 13, n. 2, p. 181–185, 2010.

LAM, E.; KATO, N.; LAWTON, M. Programmed cell death, mitochondria and the plant hypersensitive response. **Nature**, London, v. 411, n. 6839, p. 848–853, 2001.

LEHTI-SHIU, M. D.; SHIU, S.-H. Diversity, classification and function of the plant protein kinase superfamily. **Philosophical Transactions of the Royal Society B: Biological Sciences**, London, v. 367, n. 1602, p. 2619–2639, 2012.

LIBERATO, J.R. Controle das doenças causadas por fungos, bactérias e nematóides em maracujazeiro. In: ZAMBOLIM, L.; VALE, F.X.R.; MONTEIRO, A.J.A.; COSTA, H. (Ed). **Controle de doenças de plantas fruteiras**. Viçosa: Suprema, 2002. v. 2. p. 699–825.

LIPPOK, B.; BIRKENBIHL, R. P.; RIVORY, G.; BRÜMMER, J.; SCHMELZER, E.; LOGEMANN, E.; SOMSSICH, I. E. Expression of *AtWRKY33* Encoding a Pathogen- or PAMP-Responsive WRKY Transcription Factor Is Regulated by a Composite DNA

Motif Containing W Box Elements. **Molecular Plant-Microbe Interactions**, Saint Paul, v. 20, n. 4, p. 420–429, 2007.

LOPES, R.; LOPES, M. T. G.; CARNEIRO, M. S.; PINA MATTA, F. DE; CAMARGO, L. E. A.; VIEIRA, M. L. C. Linkage and mapping of resistance genes to *Xanthomonas axonopodis* pv. *passiflorae* in yellow passion fruit. **Genome**, Ottawa, v. 49, n. 1, p. 17–29, 2006.

LOPEZ, C.; SOTO, M.; RESTREPO, S.; PI-GU, B.; COOKE, R.; DELSENY, M.; TOHME, J.; VERDIER, V. Gene expression profile in response to *Xanthomonas axonopodis* pv. *manihotis* infection in cassava using a cDNA microarray. **Plant Molecular Biology**, Dordrecht, v. 57, n. 3, p. 393–410, 2005.

LORENZO, O.; PIQUERAS, R.; SÁNCHEZ-SERRANO, J. J.; SOLANO, R. Ethylene Response Factor1 Integrates Signals from Ethylene and Jasmonate Pathways in Plant Defense. **The Plant Cell**, Rockville, v. 15, n. 1, p. 165–178, 2002.

MA, W.; SMIGEL, A.; TSAI, Y.-C.; BRAAM, J.; BERKOWITZ, G. A. Innate Immunity Signaling: Cytosolic Ca²⁺ Elevation Is Linked to Downstream Nitric Oxide Generation through the Action of Calmodulin or a Calmodulin-Like Protein. **Plant Physiology**, Minneapolis, v. 148, n. 2, p. 818–828, 2008.

MacDOUGAL, J. M.; FEUILLET, C. Systematics. In: ULMER, T.; MacDOUGAL, J. M. (Ed). **Passiflora**: passionflowers of the world. Portland: Timber Press, 2004. p 27-31.

MADIGAN, M.T.; MARTINKO, J.M. **Brock biology of microorganisms**. 10 ed. Harlow: Pearson-Prentice Hall, 2003. 1385p.

MAHALINGAM, R.; GOMEZ-BUITRAGO, A.M.; ECKARDT, N.; SHAH, N.; GUEVARA-GARCIA, A.; DAY, P.; RAINA, R.; FEDOROFF, N.V. Characterizing the stress/defense transcriptome of *Arabidopsis*. **Genome Biology**, London, v. 4, p. R20.1-R20.14, 2003.

MARCHLER-BAUER, A.; ANDERSON, J.B.; DEWEESE-SCOTT, C.; FEDOROVA, N.D.; GEER, L.Y.; HE, S.; HURWITZ, D.I.; JACKSON, J.D.; JACOBS, A.R.; LANCZYCKI, C.J.; LIEBERT, C.A.; LIU, C.; MADEJ, T.; MARCHLER, G.H.; MAZUMDER, R.; NIKOLSKAYA, A.N.; PANCHENKO, A.R.; RAO, B.S.; SHOEMAKER, B.A.; SIMONYAN, V.; SONG, J.S.; THIESSEN, P.A.; VASUDEVAN, S.; WANG, Y.; YAMASHITA, R.A.; YIN, J.J.; BRYANT, S.H. CDD: a curated Entrez database of conserved domain alignments. **Nucleic Acids Research**, Oxford, v. 31, n. 1, p. 383–387, 2003.

MARCHLER-BAUER, A.; PANCHENKO, A.R.; SHOEMAKER, B.A.; THIESSEN, P.A.; GEER, L.Y.; BRYANT, S.H. CDD: a database of conserved domain alignments with links to domain three-dimensional structure. **Nucleic Acids Research**, Oxford, v. 30, n. 1, p. 281–283, 2002.

MARQUÉS, J.; DURAN-VILA, N.; DARÒS, J.-A. The Mn-binding proteins of the photosystem II oxygen-evolving complex are decreased in date palms affected by

brittle leaf disease. **Plant Physiology and Biochemistry**, Paris, v. 49, n. 4, p. 388–394, 2011.

MICHALUN, V.; MICHALUN, N. **Dicionário de Ingredientes para a Cosmética e Cuidados da Pele**. Tradução de SILVA, M. São Paulo: Editora Senac São Paulo, 2010. 358 p.

MICHELS, A. K.; WEDEL, N.; KROTH, P. G. Diatom Plastids Possess a Phosphoribulokinase with an Altered Regulation and No Oxidative Pentose Phosphate Pathway. **Plant Physiology**, Minneapolis, v. 137, n. 3, p. 911–920, 2005.

MITTLER, R.; VANDERAUWERA, S.; GOLLERY, M.; BREUSEGEM, F. VAN. Reactive oxygen gene network of plants. **Trends in Plant Science**, Oxford, v. 9, n. 10, p. 490–498, 2004.

MORITA-YAMAMURO, C.; TSUTSUI, T.; SATO, M.; YOSHIOKA, H.; TAMAOKI, M.; OGAWA, D.; MATSUURA, H.; YOSHIHARA, T.; IKEDA, A.; UYEDA, I.; YAMAGUCHI, J. The *Arabidopsis* Gene CAD1 Controls Programmed Cell Death in the Plant Immune System and Encodes a Protein Containing a MACPF Domain. **Plant and Cell Physiology**, Kyoto, v. 46, n. 6, p. 902–912, 2005.

MUNHOZ, C. F.; WEISS, B.; HANAI, L. R.; ZUCCHI, M. I.; FUNGARO, M. H. P.; OLIVEIRA, A. L. M.; MONTEIRO-VITORELLO, C. B.; VIEIRA, M. L. C. Genetic Diversity and a PCR-Based Method for *Xanthomonas axonopodis* Detection in Passion Fruit. **Phytopathology**, Saint Paul, v. 101, n. 4, p. 416–424, 2011.

NAFISI, M.; GOREGAOKER, S.; BOTANGA, C. J.; GLAWISCHNIG, E.; OLSEN, C. E.; HALKIER, B. A.; GLAZEBROOK, J. *Arabidopsis* Cytochrome P450 Monooxygenase 71A13 Catalyzes the Conversion of Indole-3-Acetaldoxime in Camalexin Synthesis. **The Plant Cell Online**, Rockville, v. 19, n. 6, p. 2039–2052, 2007.

NAKAMURA, Y.; COCHRANE, G.; KARSCH-MIZRACHI, I.; ON BEHALF OF THE INTERNATIONAL NUCLEOTIDE SEQUENCE DATABASE COLLABORATION. The International Nucleotide Sequence Database Collaboration. **Nucleic Acids Research**, Oxford, v. 41, n. D1, p. D21–D24, 2012.

NÜRNBERGER, T.; BRUNNER, F. Innate immunity in plants and animals: emerging parallels between the recognition of general elicitors and pathogen-associated molecular patterns. **Current Opinion in Plant Biology**, London, v. 5, n. 4, p. 318–324, 2002.

NÜRNBERGER, T.; BRUNNER, F.; KEMMERLING, B.; PIATER, L. Innate immunity in plants and animals: striking similarities and obvious differences. **Immunological reviews**, Copenhagen, v. 198, p. 249–266, 2004.

OH, M.-H.; WANG, X.; WU, X.; ZHAO, Y.; CLOUSE, S. D.; HUBER, S. C. Autophosphorylation of Tyr-610 in the receptor kinase BAK1 plays a role in brassinosteroid signaling and basal defense gene expression. **Proceedings of the National Academy of Sciences**, Washington, v. 107, n. 41, p. 17827–17832, 2010.

OLIVEIRA, E.J.; VIEIRA, M.L.C.; GARCIA, A.A.F.; MUNHOZ, C.F.; MARGARIDO, G.R.A.; CONSOLI, L.; MATTA, F.P.; MORAES, M.C. An integrated molecular map of yellow passion fruit based on simultaneous maximum-likelihood estimation of linkage and linkage phases. **Journal of the American Society for Horticultural Science**, Alexandria, v. 133, p. 35–41, 2008.

PAGEAU, K.; REISDORF-CREN, M. LE; MOROT-GAUDRY, J.-F.; MASCLAUX-DAUBRESSE, C. LINE. The two senescence-related markers, GS1 (cytosolic glutamine synthetase) and GDH (glutamate dehydrogenase), involved in nitrogen mobilization, are differentially regulated during pathogen attack and by stress hormones and reactive oxygen species in *Nicotiana tabacum* L. leaves. **Journal of Experimental Botany**, Oxford, v. 57, n. 3, p. 547–557, 2006.

PANDEY, S. P.; SOMSSICH, I. E. The Role of WRKY Transcription Factors in Plant Immunity. **Plant Physiology**, Minneapolis, v. 150, n. 4, p. 1648–1655, 2009.

PARISY, V.; POINSSOT, B.; OWSIANOWSKI, L.; BUCHALA, A.; GLAZEBROOK, J.; MAUCH, F. Identification of PAD2 as a γ -glutamylcysteine synthetase highlights the importance of glutathione in disease resistance of *Arabidopsis*. **The Plant Journal**, Oxford, v. 49, n. 1, p. 159–172, 2006.

PASSOS, M. A. N.; CRUZ, V. O. DE; EMEDIATO, F. L.; TEIXEIRA, C. C. DE; AZEVEDO, V. C. R.; BRASILEIRO, A. C. M.; AMORIM, E. P.; FERREIRA, C. F.; MARTINS, N. F.; TOGAWA, R. C.; PAPPAS, G. J.; SILVA, O. B. DA; MILLER, R. N. Analysis of the leaf transcriptome of *Musa acuminata* during interaction with *Mycosphaerella musicola*: gene assembly, annotation and marker development. **BMC Genomics**, London, v. 14, n. 1, p. 78, 2013.

PENHA, H.A. **Construção de uma biblioteca genômica de *Passiflora edulis* f. *flavicarpa* inserida em BACs (*Bacterial Artificial Chromosome*) e mapeamento cromossômico usando hibridação in situ fluorescente**. 2012. 129p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2012.

PEREIRA, G. S.; NUNES, E. S.; LAPERUTA, L. D. C.; BRAGA, M. F.; PENHA, H. A.; DINIZ, A. L.; MUNHOZ, C. F.; GAZAFFI, R.; GARCIA, A. A. F.; VIEIRA, M. L. C. Molecular polymorphism and linkage analysis in sweet passion fruit, an outcrossing species. **Annals of Applied Biology**, Warwick, v. 162, n. 3, p. 347–361, 2013.

PFAFFL, M. W.; HORGAN, G.H.; DEMPFLER, L. Relative expression software tool (REST[®]) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. **Nucleic Acids Research**, Oxford, v. 30, n. 9, p. 36e–36, 2002.

PIRES, M.M.; GOMES, A.S.; MIDDLEJ, M.M.B.C.; SÃO JOSÉ, A.R.; ROSADO, P.L.; PASSOS, H.D.B. Caracterização do mercado de maracujá. In: PIRES, M.M.; SÃO JOSÉ, A.R.; CONCEIÇÃO, A. O. (Org). **Maracujá: avanços tecnológicos e sustentabilidade**. Ilhéus: Editus, 2011. p. 21-67.

PIRES, M.M; SÃO JOSÉ, A.R.; CONCEIÇÃO, A. O. (Org). **Maracujá: avanços tecnológicos e sustentabilidade**. Ilhéus: Editus, 2011. 238 p.

PORTA, H.; ROCHA-SOSA, M. Plant Lipxygenases. Physiological and Molecular Features. **Plant Physiology**, Minneapolis, v. 130, n. 1, p. 15–21, 2002.

PROOST, S.; BEL, M. VAN; STERCK, L.; BILLIAU, K.; PARYS, T. VAN; PEER, Y. VAN DE; VANDEPOELE, K. PLAZA: A Comparative Genomics Resource to Study Gene and Genome Evolution in Plants. **The Plant Cell**, Rockville, v. 21, n. 12, p. 3718–3731, 2009.

QI, X.; XU, Q.; SHEN, L.; ALFANDI, M.; LUO, J.; CHEN, X. Identification of differentially expressed genes between powdery mildew resistant near-isogenic line and susceptible line of cucumber by suppression subtractive hybridization. **Scientia Horticulturae**, Amsterdam, v. 126, n. 1, p. 27–32, 2010.

QIU, J.-L.; FIIL, B. K.; PETERSEN, K.; NIELSEN, H. B.; BOTANGA, C. J.; THORGRIMSEN, S.; PALMA, K.; SUAREZ-RODRIGUEZ, M. C.; SANDBECH-CLAUSEN, S.; LICHOTA, J.; BRODERSEN, P.; GRASSER, K. D.; MATSSON, O.; GLAZEBROOK, J.; MUNDY, J.; PETERSEN, M. *Arabidopsis* MAP kinase 4 regulates gene expression through transcription factor release in the nucleus. **The EMBO Journal**, Oxford, v. 27, n. 16, p. 2214–2221, 2008a.

QIU, J.-L.; ZHOU, L.; YUN, B.-W.; NIELSEN, H. B.; FIIL, B. K.; PETERSEN, K.; MACKINLAY, J.; LOAKE, G. J.; MUNDY, J.; MORRIS, P. C. *Arabidopsis* Mitogen-Activated Protein Kinase Kinases MKK1 and MKK2 Have Overlapping Functions in Defense Signaling Mediated by MEKK1, MPK4, and MKS1. **Plant Physiology**, Minneapolis, v. 148, n. 1, p. 212–222, 2008b.

RAMAKERS, C.; RUIJTER, J. M.; DEPREZ, R. H. L.; MOORMAN, A. F. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. **Neuroscience Letters**, Limerick, v. 339, n. 1, p. 62–66, 2003.

REBRIKOV, D. V.; BRITANOVA, O. V.; GURSKAYA, N. G.; LUKYANOV, K. A.; TARABYKIN, V. S.; LUKYANOV, S. A. Mirror orientation selection (MOS): a method for eliminating false positive clones from libraries generated by suppression subtractive hybridization. **Nucleic Acids Research**, Oxford, v. 28, n. 20, p. 90e–90, 2000.

RESENDE, M. L. V.; SALGADO, S. M. L.; CHAVES, Z. M. Espécies Ativas de Oxigênio na Resposta de Defesa de Plantas a Patógenos. **Fitopatologia Brasileira**, Brasília, v. 28, p. 123–130, 2003.

RICHTERICH, P. Estimation of errors in “raw” DNA sequences: a validation study. **Genome research**, Cold Spring Harbor, v. 8, n. 3, p. 251–259, 1998.

RINGER, K. L.; MCCONKEY, M. E.; DAVIS, E. M.; RUSHING, G. W.; CROTEAU, R. Monoterpene double-bond reductases of the (–)-menthol biosynthetic pathway: isolation and characterization of cDNAs encoding (–)-isopiperitenone reductase and

(+)-pulegone reductase of peppermint. **Archives of Biochemistry and Biophysics**, New York, v. 418, n. 1, p. 80–92, 2003.

ROMEIRO, R. Indução de resistência em plantas a patógenos. In: PASCHOLATI, S.F.; LEITE, B.; STANGARLIN, J.R.; CIA, P. (Ed). **Interações planta-patógeno: fisiologia, bioquímica e biologia molecular**. Piracicaba: Fealq, 2008. p. 411-431.

RYAN, R. P.; VORHÖLTER, F.-J.; POTNIS, N.; JONES, J. B.; SLUYS, M.-A. VAN; BOGDANOVE, A. J.; DOW, J. M. Pathogenomics of *Xanthomonas*: understanding bacterium–plant interactions. **Nature Reviews Microbiology**, London, v. 9, n. 5, p. 344–355, 2011.

SAMBROOK, J.; RUSSELL, D. W. **Molecular cloning: a laboratory manual**. 3. ed. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 2001. 2344 p.

SANA, T. R.; FISCHER, S.; WOHLGEMUTH, G.; KATREKAR, A.; JUNG, K.; RONALD, P. C.; FIEHN, O. Metabolomic and transcriptomic analysis of the rice response to the bacterial blight pathogen *Xanthomonas oryzae* pv. *oryzae*. **Metabolomics**, New Yourk, v. 6, n. 3, p. 451–465, 2010.

SANGER, F.; AIR, G. M.; BARRELL, B. G.; BROWN, N. L.; COULSON, A. R.; FIDDES, J. C.; HUTCHISON, C. A.; SLOCOMBE, P. M.; SMITH, M. Nucleotide sequence of bacteriophage ϕ X174 DNA. **Nature**, London, v. 265, n. 5596, p. 687–695, 1977.

SANIER, C.; SAYEGH-ALHAMDI, M.; JALLOUL, A.; CLERIVET, A.; NICOLE, M.; MARMEY, P. A 13-lipoxygenase is Expressed Early in the Hypersensitive Reaction of Cotton Plants to *Xanthomonas campestris* pv. *malvacearum*. **Journal of Phytopathology**, Berlin, v. 160, n. 6, p. 286–293, 2012.

SANTOS, A.A. **Exploração de uma biblioteca genômica de *Passiflora edulis* f. *flavicarpa* por sequenciamento de BAC-ends**. Piracicaba, 2013. 141 p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2013.

SANTOS FILHO, H.P.; LARANJEIRA, F.F.; SANTOS, C.C.F.; BARBOSA, C.J. Doenças do maracujazeiro. In: LIMA, A.A.; CUNHA, M.A.P. (Ed). **Maracujá: produção e qualidade na passicultura**. Cruz das Almas: Embrapa Mandioca e Fruticultura, 2004. p. 239-280.

SÃO JOSÉ, A.R.; PIRES, M.M. Aspectos gerais da cultura do maracujá no Brasil. In: PIRES, M.M; SÃO JOSÉ, A.R.; CONCEIÇÃO, A. O. (Org). **Maracujá: avanços tecnológicos e sustentabilidade**. Ilhéus: Editus, 2011. p. 13-19.

SETUBAL, J.C. Bioinformática. In: MIR, L. (Ed.). **Genômica**. São Paulo: Editora Atheneu, 2004. p. 105-117.

SHI, C.; CHAUDHARY, S.; YU, K.; PARK, S. J.; NAVABI, A.; MCCLEAN, P. E. Identification of candidate genes associated with CBB resistance in common bean

HR45 (*Phaseolus vulgaris* L.) using cDNA-AFLP. **Molecular Biology Reports**, Dordrecht, v. 38, n. 1, p. 75–81, 2010.

SINGH, K. B.; FOLEY, R. C.; OÑATE-SÁNCHEZ, L. Transcription factors in plant defense and stress responses. **Current Opinion in Plant Biology**, London, v. 5, n. 5, p. 430–436, 2002.

SOARES, A. M. DOS S.; MACHADO, O. L. T. Defesa de plantas: Sinalização química e espécies reativas de oxigênio. **Revista Trópica – Ciências Agrárias e Biológicas**, Chapadinha, v. 1, p. 9–19, 2007.

SORIA-GUERRA, R. E.; ROSALES-MENDOZA, S.; CHANG, S.; HAUDENSHIELD, J. S.; ZHENG, D.; RAO, S. S.; HARTMAN, G. L.; GHABRIAL, S. A.; KORBAN, S. S. Identifying differentially expressed genes in leaves of *Glycine tomentella* in the presence of the fungal pathogen *Phakopsora pachyrhizi*. **Planta**, Berlin, v. 232, n. 5, p. 1181–1189, 2010.

SPOEL, S. H.; DONG, X. How do plants achieve immunity? Defence without specialized immune cells. **Nature Reviews Immunology**, London, v. 12, n. 2, p. 89–100, 2012.

STEIN, O.D. VON; THIES, W-G.; HOFMANN, M. A high throughput screening for rarely transcribed differentially expressed genes. **Nucleic Acids Research**, Oxford, v. 25, n. 13, p. 2598-2602, 1997.

SWARBRECK, D.; WILKS, C.; LAMESCH, P.; BERARDINI, T. Z.; GARCIA-HERNANDEZ, M.; FOERSTER, H.; LI, D.; MEYER, T.; MULLER, R.; PLOETZ, L.; RADENBAUGH, A.; SINGH, S.; SWING, V.; TISSIER, C.; ZHANG, P.; HUALA, E. The *Arabidopsis* Information Resource (TAIR): gene structure and function annotation. **Nucleic Acids Research**, Oxford, v. 36, p. D1009–D1014, 2007.

SWINGS, J.G.; CIVEROLO, E.L. **Xanthomonas**. London: Chapman & Hall, 1993. 399 p.

SZKLARCZYK, D.; FRANCESCHINI, A.; KUHN, M.; SIMONOVIC, M.; ROTH, A.; MINGUEZ, P.; DOERKS, T.; STARK, M.; MULLER, J.; BORK, P.; JENSEN, L. J.; MERING, C. V. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. **Nucleic Acids Research**, Oxford, v. 39, p. D561–D568, 2011.

TAIRA, M.; VALTERSSON, U.; BURKHARDT, B.; LUDWIG, R. A. *Arabidopsis thaliana* GLN2-Encoded Glutamine Synthetase Is Dual Targeted to Leaf Mitochondria and Chloroplasts. **The Plant Cell Online**, Rockville, v. 16, n. 8, p. 2048–2058, 2004.

TAMELING, W. I. L.; TAKKEN, F. L. W. Resistance proteins: scouts of the plant innate immune system. **European Journal of Plant Pathology**, Dordrecht, v. 121, n. 3, p. 243–255, 2008.

TAMURA, K.; PETERSON, D.; PETERSON, N.; STECHER, G.; NEI, M.; KUMAR, S. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood,

Evolutionary Distance, and Maximum Parsimony Methods. **Molecular Biology and Evolution**, Chicago, v. 28, n. 10, p. 2731–2739, 2011.

TENA, G.; BOUDSOCQ, M.; SHEEN, J. Protein kinase signaling networks in plant innate immunity. **Current Opinion in Plant Biology**, London, v. 14, n. 5, p. 519–529, 2011.

THE UNIPROT CONSORTIUM. Update on activities at the Universal Protein Resource (UniProt) in 2013. **Nucleic Acids Research**, Oxford, v. 41, n. D1, p. D43–D47, 2012.

TSENG, T.-T.; TYLER, B. M.; SETUBAL, J. C. Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. **BMC Microbiology**, London, v. 9, p. S2, 2009.

UDVARDI, M. K.; CZECHOWSKI, T.; SCHEIBLE, W.-R. Eleven Golden Rules of Quantitative RT-PCR. **The Plant Cell Online**, Rockville, v. 20, n. 7, p. 1736–1737, 2008.

VAN BEL, M.; PROOST, S.; WISCHNITZKI, E.; MOVAHEDI, S.; SCHEERLINCK, C.; VAN DE PEER, Y.; VANDEPOELE, K. Dissecting Plant Genomes with the PLAZA Comparative Genomics Platform. **Plant Physiology**, Minneapolis, v. 158, n. 2, p. 590–600, 2011.

VANDESOMPELE, J.; PRETER, K. DE; PATTYN, F.; POPPE, B.; ROY, N. VAN; PAEPE, A. DE; SPELEMAN, F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. **Genome biology**, London, v. 3, n. 7, p. 34-34.11, 2002.

VERONESE, P.; RUIZ, M. T.; COCA, M. A.; HERNANDEZ-LOPEZ, A.; LEE, H.; IBEAS, J. I.; DAMSZ, B.; PARDO, J. M.; HASEGAWA, P. M.; BRESSAN, R. A.; NARASIMHAN, M. L. In Defense against Pathogens. Both Plant Sentinels and Foot Soldiers Need to Know the Enemy. **Plant Physiology**, Minneapolis, v. 131, n. 4, p. 1580–1590, 2003.

VERRIER, P. J.; BIRD, D.; BURLA, B.; DASSA, E.; FORESTIER, C.; GEISLER, M.; KLEIN, M.; KOLUKISA OGLU, Ü.; LEE, Y.; MARTINOIA, E.; MURPHY, A.; REA, P. A.; SAMUELS, L.; SCHULZ, B.; SPALDING, E. P.; YAZAKI, K.; THEODOULOU, F. L. Plant ABC proteins – a unified nomenclature and updated inventory. **Trends in Plant Science**, Oxford, v. 13, n. 4, p. 151–159, 2008.

VIANA, F. M. P.; FREIRE, F. C. O.; CARDOSO, J. E.; VIDAL, J. C. **Principais doenças do maracujazeiro na Região Nordeste e seu controle**. Fortaleza: Embrapa Agroindústria Tropical, 2003. 12 p.

VIEIRA, M.L.C.; CARNEIRO, M.S. *Passiflora* spp. passionfruit. In: LITZ, R.E. (Ed). **Biotechnology of fruit and nut crops**. Oxford: CABI Publishing, 2005. p. 436-453.

WANG, X.; TANG, C.; ZHANG, G.; LI, Y.; WANG, C.; LIU, B.; QU, Z.; ZHAO, J.; HAN, Q.; HUANG, L.; CHEN, X.; KANG, Z. cDNA-AFLP analysis reveals differential

gene expression in compatible interaction of wheat challenged with *Puccinia striiformis* f. sp. *tritici*. **BMC Genomics**, London, v. 10, n. 1, p. 289, 2009.

WRZACZEK, M.; HIRT, H. Plant MAP kinase pathways: how many and what for? **Biology of the cell**, Paris, v. 93, n. 1-2, p. 81–87, 2001.

WULFF, N.A. Bactérias causadoras de doenças em vegetais. In: PASCHOLATI, S.F.; LEITE, B.; STANGARLIN, J.R.; CIA, P. (Ed). **Interações planta-patógeno: fisiologia, bioquímica e biologia molecular**. Piracicaba: Fealq, 2008. p. 51-114.

YEOH, K.-A.; OTHMAN, A.; MEON, S.; ABDULLAH, F.; HO, C.L. Sequence analysis and gene expression of putative oil palm chitinase and chitinase-like proteins in response to colonization of *Ganoderma boninense* and *Trichoderma harzianum*. **Molecular Biology Reports**, Dordrecht, v. 40, n. 1, p. 147–158, 2013.

ZHANG, S.; KLESSIG, D. F. MAPK cascades in plant defense signaling. **Trends in plant science**, Oxford, v. 6, n. 11, p. 520–527, 2001.

ZHANG, W.-W.; JIAN, G.-L.; JIANG, T.-F.; WANG, S.-Z.; QI, F.-J.; XU, S.-C. Cotton gene expression profiles in resistant *Gossypium hirsutum* cv. *Zhongzhimian* KV1 responding to *Verticillium dahliae* strain V991 infection. **Molecular Biology Reports**, Dordrecht, v. 39, n. 10, p. 9765–9774, 2012.

ZIEROLD, U.; SCHOLZ, U.; SCHWEIZER, P. Transcriptome analysis of *mlo* - mediated resistance in the epidermis of barley. **Molecular Plant Pathology**, London, v. 6, n. 2, p. 139–151, 2005.

ANEXOS

ANEXO A

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

(continua)

Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	Similaridade (%)	e-value
F-P6B11	XP_002311536	rna-dependent rna polymerase	484	56	93	7,99E-50
F-P8C08	CAA09896	rna-dependent rna polymerase 1	518	90	88	1,78E-75
F-P9C07	CBI40245	neutral invertase	458	100	84	8,48E-78
F-P8E02	XP_002312110	ribulose biphosphate carboxylase oxygenase activase chloroplast	602	50	82	3,74E-35
F-P5B08	BAA22071	sucrose-phosphate synthase	548	82	89	6,67E-76
F-P9B06	XP_002326480	lil3 protein	496	54	67	2,52E-09
F-P1A01	XP_002522136	serine carboxypeptidase-like 18-like	436	99	79	2,04E-54
F-P4A01	XP_002305956	protein decapping 5	537	99	73	5,17E-40
F-P2A01	XP_002330883	protein	532	86	88	2,62E-74
F-P7A01	XP_002299674	probable plastid-lipid-associated protein chloroplastic	604	91	83	8,66E-85
F-P4A02	XP_002516879	chromatin-remodeling protein 11	508	38	90	3,52E-17
F-P9A02	XP_002325841	predicted protein [P. trichocarpa]	467	49	78	2,93E-11
F-P7A02	XP_003637430	zinc finger protein constans-like 16-like	367	38	95	2,80E-14
F-P8A02	XP_002305200	fructokinase 3	529	100	95	2,24E-95
F-P4A03	XP_002300248	indole-3-acetic acid-amido synthetase	523	72	97	1,72E-71
F-P6A03	XP_002510564	abc transporter b family member 13-like	506	91	82	2,97E-61
F-P9A03	XP_002520135	porin voltage-dependent anion-selective channel protein	538	66	94	8,17E-62
F-P2A03	XP_002525743	class i chitinase	491	32	76	1,41E-08
F-P3A03	CAN80513	atp-dependent clp protease atp-binding subunit 1	343	99	92	8,52E-50
F-P7A03	XP_002522015	g-type lectin s-receptor-like serine threonine-protein kinase	505	73	77	1,69E-38
F-P8A03	XP_002522020	gds1 esterase lipase at5g33370-like	551	78	90	1,03E-75
F-P5A04	XP_002530486	pollen-specific c2 domain containing protein	552	100	86	3,05E-92
F-P10A04	XP_002522279	protein	300	46	91	1,42E-13
F-P2A04	XP_002326066	protein	374	67	89	4,04E-21
F-P7A04	XP_002515456	probable receptor-like protein kinase at1g80640-like	493	92	83	4,56E-67
F-P8A04	XP_002529889	subtilase family protein	596	91	94	3,00E-109
F-P1A05	AEK64517	sulfate adenyltransferase	417	99	96	2,73E-82
F-P4A05	XP_002334016	linoleate 13s-lipoxygenase 2-chloroplastic-like	561	100	86	9,52E-77
F-P5A05	XP_002528284	1-phosphatidylinositol phosphodiesterase-related protein	587	81	94	7,43E-96
F-P9A05	XP_002310031	abc transporter family protein	552	99	97	5,06E-96
F-P10A05	AAX84670	ap2 erf domain-containing transcription factor	537	70	88	2,22E-51
F-P2A05	XP_002301880	peroxisomal membrane protein	484	61	83	6,58E-36
F-P3A05	BAH03319	beta-galactosidase	470	50	89	6,17E-28
F-P5A06	AFU96803	ribosomal protein s11	526	63	97	2,31E-60
F-P6A06	XP_002527764	protein argonaute 4	488	100	96	1,04E-91
F-P9A06	ACS28586	lipoxygenase	543	41	97	3,45E-36
F-P10A06	XP_002321884	molybdopterin cofactor	608	84	77	4,83E-68
F-P2A06	AAF24826	methyltransferase nsun6-like	354	31	97	7,15E-09
F-P7A06	XP_002338954	methionine synthase	504	61	96	4,54E-57
F-P9A07	XP_003547043	ccr4-not transcription complex	227	78	96	2,25E-28
F-P10A07	AAM94350	gag-pol polyprotein	365	72	75	2,36E-14

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

(continuação)						
Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	Similaridade (%)	e-value
F-P3A07	BAA95409	acetylornithine deacetylase	376	100	88	4,54E-54
F-P1A08	XP_002523967	50s ribosomal protein l3	453	75	68	1,04E-13
F-P5A08	XP_002309484	abscisic insensitive 1b	522	87	94	6,72E-91
F-P9A08	XP_002524280	tpr repeat-containing thioredoxin	593	95	82	7,42E-70
F-P5A09	XP_002327389	tpv38 tmem64 family membrane	550	97	87	2,22E-57
F-P6A09	CAC12883	60s ribosomal protein l11	573	93	97	1,97E-100
F-P9A09	XP_002309490	cytokinin riboside 5 - monophosphate	522	84	95	1,09E-83
F-P7A09	XP_002323585	omega-3 fatty acid desaturase	479	99	71	2,24E-55
F-P8A09	XP_002299190	eri1 exoribonuclease	508	93	82	1,91E-69
F-P4A10	ACU18420	magnesium-protoporphyrin ix monomethyl ester	548	53	94	1,46E-50
F-P9A10	XP_003552463	2-phosphoglycolate phosphatase 1	524	53	97	2,10E-38
F-P10A10	XP_002325416	casein kinase i isoform alpha	445	41	91	2,59E-09
F-P8A10	XP_002522810	pinus taeda anonymous locus	455	100	98	3,59E-83
F-P1A11	XP_002528192	coatomer subunit beta -1	355	47	92	4,72E-19
F-P4A11	BAD94507	glutamine synthetase	604	58	96	6,80E-58
F-P5A11	XP_002301732	50s ribosomal protein l20	468	81	97	1,59E-75
F-P10A11	XP_002528100	ca2+ antiporter cation exchanger	515	100	91	1,24E-69
F-P2A11	XP_002280297	oxidoreductase glyr1-like	486	94	84	1,46E-65
F-P3A11	XP_002534443	protein	530	29	82	2,99E-08
F-P4A12	XP_002323321	white-brown-complex abc transporter family	530	100	89	5,69E-89
F-P5A12	CBI36070	abc transporter family pleiotropic drug resistance protein	630	98	96	2,23E-117
F-P6A12	XP_002316801	protein	529	99	74	2,29E-43
F-P9A12	XP_002317107	transparent testa 12 protein	563	87	92	6,10E-85
F-P10A12	CBI24749	ankyrin repeat-containing protein	504	85	85	9,30E-50
F-P2A12	XP_002316526	protein	454	29	86	1,02E-07
F-P3A12	CBI40528	tetratricopeptide repeat-containing	475	35	98	5,46E-21
F-P7A12	ACA24497	nucleic acid binding protein	546	38	78	2,35E-06
F-P8A12	XP_002524275	dna binding	526	81	57	5,56E-21
F-P1B01	XP_002320646	glutaredoxin	409	46	95	5,76E-28
F-P5B01	XP_002308853	phosphatidylinositol 3- and 4-kinase-like protein	521	99	58	2,80E-22
F-P6B01	XP_002298903	sodium calcium exchanger family	562	30	91	1,62E-08
F-P10B01	XP_002271666	photosystem ii reaction center psb28 protein	540	99	84	4,79E-77
F-P2B01	CAN65842	nuclease harbi1-like	507	96	77	1,15E-59
F-P8B01	ABK96172	g-box binding factor 1	367	59	84	1,85E-09
F-P5B02	XP_002317500	protein	566	100	96	1,19E-117
F-P2B02	XP_002298012	2-c-methyl-d-erythritol - cyclodiphosphate synthase	465	99	96	1,52E-95
F-P7B02	CAN64310	uncharacterized membrane	513	70	92	1,52E-32
F-P6B03	XP_002327405	wd repeat-containing	536	99	98	4,76E-90
F-P9B03	AEQ94140	glutamine partial	375	46	87	1,53E-21
F-P10B03	XP_002311648	predicted protein [P. trichocarpa]	406	46	72	1,43E-06
F-P3B03	XP_002519546	nucleotide binding	435	34	87	5,70E-11
F-P7B03	XP_002521574	dna-3-methyladenine glycosylase ii	528	95	80	3,78E-55
F-P8B03	XP_002324756	far upstream element-binding	531	88	89	1,76E-60
F-P4B04	NP_001234138	profilin	540	69	95	1,19E-69
F-P5B04	ABH02907	r2r3-myb transcription factor	557	88	90	1,41E-90
F-P6B04	CBI31917	dead-box atp-dependent rna helicase	535	44	75	6,22E-17

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

(continuação)						
Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	Similaridade (%)	e-value
F-P3B04	XP_002533575	upf0308 protein chloroplastic	487	31	82	1,02E-20
F-P7B04	AAF35185	lipid transfer protein	554	64	82	1,32E-41
F-P1B05	XP_002527180	zinc finger (ubiquitin-hydrolase)	529	86	87	4,75E-50
F-P4B05	Q41596	class i chitinase	459	71	61	9,29E-16
F-P5B05	XP_003529492	PREDICTED: uncharacterized protein [Glycine max]	579	72	59	1,59E-23
F-P9B05	XP_002301851	histone-lysine n-methyltransferase suvr4	541	100	83	4,70E-82
F-P10B05	CAN67413	glycolate oxidase	531	94	96	1,35E-100
F-P4B06	XP_002317910	protein	590	89	75	2,04E-51
F-P5B06	XP_002311964	udp-glycosyltransferase 1	483	94	71	3,32E-38
F-P6B06	AAL06516	translocase of chloroplast chloroplastic-like	554	100	90	3,79E-92
F-P10B06	XP_002312786	porin voltage-dependent anion-selective channel protein	548	75	94	6,25E-75
F-P2B06	CBI30656	magnesium-chelatase subunit chloroplastic-like	426	74	97	1,21E-48
F-P7B06	XP_002285897	v-type proton atpase subunit b2	597	99	98	1,23E-129
F-P6B07	XP_003519443	ethylene-responsive transcription factor 3-like	590	43	69	9,46E-03
F-P3B07	ADR71306	50s ribosomal protein l20	493	64	75	2,06E-23
F-P7B07	ABK93854	geminivirus rep-interacting motor	541	71	97	9,73E-77
F-P6B08	XP_002514883	protein	536	66	67	9,88E-24
F-P9B08	XP_002315799	protein kinase	517	99	99	1,73E-112
F-P2B08	XP_002329164	protein	374	99	92	1,22E-61
F-P1B09	XP_002518283	nodulin 21 -like transporter protein	496	33	79	1,79E-11
F-P4B09	XP_002303534	predicted protein [P. trichocarpa]	587	89	73	9,05E-51
F-P3B09	XP_002299991	asparagine synthetase	468	91	94	1,32E-77
F-P7B09	XP_002300442	solanesyl diphosphate synthase	556	79	80	1,08E-51
F-P8B09	XP_002312571	zinc knuckle (cchc-type) family	543	97	61	1,44E-31
F-P1B10	XP_002319430	protein	411	99	91	6,19E-66
F-P4B10	AAW48295	pore-forming toxin-like protein hfr-2	520	65	60	1,61E-10
F-P5B10	XP_002309559	coiled-coil domain-containing protein 25-like	546	53	95	1,02E-20
F-P9B10	XP_002306862	protein kinase	522	59	83	1,02E-32
F-P10B10	XP_002513951	phosphate synthetase	564	95	98	4,90E-115
F-P2B10	XP_002306038	udp-glycosyltransferase 83a1	392	100	84	2,79E-61
F-P7B10	ACJ85402	26s protease regulatory subunit 6a	545	99	100	2,74E-119
F-P1B11	XP_002520110	sucrose transport	452	53	67	6,87E-05
F-P5B11	XP_002519042	ring finger protein	319	90	96	2,81E-43
F-P9B11	XP_002313207	inorganic pyrophosphatase	485	72	76	4,18E-37
F-P2B11	XP_002521824	anthocyanidin -o-glucosyltransferase-like	498	89	86	2,56E-66
F-P3B11	XP_002329673	pumilio homolog 5	470	61	73	2,49E-20
F-P7B11	XP_002518020	set domain-containing protein	594	98	73	4,55E-66
F-P10B12	XP_002518948	b3 domain-containing protein	546	46	65	1,60E-08
F-P3B12	XP_002324121	abc transporter family protein	457	99	94	1,91E-77
F-P7B12	XP_002323194	14 kda zinc-binding protein	474	70	96	5,55E-65
F-P1C01	XP_002520024	conserved hypothetical protein	410	41	83	4,63E-15
F-P4C01	AAZ79656	beta-glucanase	577	99	89	1,20E-94
F-P6C01	XP_002308119	subtilisin-like serine protease	572	97	86	6,99E-92
F-P9C01	XP_002312695	myb-like transcription factor-like	541	60	80	2,49E-32
F-P10C01	XP_002304106	protein phosphatase 2c 16	525	100	82	2,48E-56
F-P3C01	XP_002514211	multidrug resistance protein	457	100	88	3,54E-64

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

(continuação)						
Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	Similaridade (%)	e-value
F-P7C01	XP_002302076	single-strand binding protein	484	77	79	4,06E-44
F-P4C02	XP_002529195	dna binding	484	33	79	3,49E-04
F-P6C02	XP_002321580	at2g26730-like partial	532	47	85	1,60E-28
F-P10C02	ACZ74051	translation initiation factor 6	536	58	96	8,90E-59
F-P7C02	XP_002530714	nudix hydrolase chloroplastic-like	501	90	68	6,97E-32
F-P1C03	XP_002514060	stress responsive alpha-beta barrel domain protein	361	100	71	4,68E-14
F-P4C03	BAA95416	cysteine proteinase inhibitor	604	68	65	1,23E-19
F-P6C03	XP_002303641	protein	593	87	78	1,62E-64
F-P3C03	XP_002310077	atp-dependent clp protease proteolytic subunit	399	35	89	4,88E-07
F-P7C03	XP_002515156	monooxygenase 1	427	93	75	2,52E-47
F-P4C04	XP_002306799	beach domain-containing protein lvs-a-like	515	80	79	6,97E-37
F-P5C04	XP_002310919	phosphoprotein phosphatase	606	99	99	3,75E-137
F-P6C04	XP_002309411	letm1-like protein	520	99	79	3,81E-71
F-P3C04	XP_002520361	somatic embryogenesis receptor-like kinase 1	469	46	84	7,55E-15
F-P7C04	XP_002532648	conserved hypothetical protein	284	72	70	2,06E-05
F-P8C04	XP_002315956	phytochrome and flowering time regulatory protein 1	553	100	86	6,85E-69
F-P4C05	XP_002307025	cc-nbs-lrr resistance protein	522	100	70	2,39E-41
F-P5C05	XP_002523168	maternal effect embryo arrest 22	533	80	68	4,06E-28
F-P6C05	XP_002325430	exosome component 10-like	466	98	65	2,55E-33
F-P2C05	ABF48719	protein	504	99	92	3,59E-91
F-P4C06	XP_002509949	glycoside hydrolase family 28 protein	545	100	95	1,10E-101
F-P5C06	XP_002313719	protein	444	99	92	1,10E-57
F-P2C06	AFK37576	brain protein 44	370	41	92	1,18E-16
F-P3C06	XP_002314140	transcription initiation factor	376	100	60	7,08E-27
F-P7C06	XP_002307690	purple acid phosphatase	389	75	88	7,39E-43
F-P4C07	XP_002314500	actin-related protein c2b	523	27	97	1,03E-23
F-P6C07	CBI34773	wd-40 repeat family protein	565	22	88	2,87E-15
F-P10C07	AFK40307	at1g18740 f6a14_15	449	76	78	2,61E-30
F-P2C07	XP_002313579	rna recognition motif-containing	465	100	59	1,92E-25
F-P7C07	XP_002318082	soul-like protein	520	96	86	2,61E-84
F-P8C07	XP_002306249	photosystem i reaction center subunit n	483	32	100	3,13E-37
F-P1C08	ABK95838	oxygen-evolving enhancer protein 1	470	63	95	6,52E-52
F-P10C08	XP_002521988	conserved hypothetical protein	113	98	97	9,20E-11
F-P3C08	XP_002310868	duf246 domain-containing protein	507	69	81	6,07E-43
F-P7C08	XP_002298022	predicted protein [P. trichocarpa]	528	60	96	3,48E-54
F-P6C09	XP_002520399	integrin-linked protein kinase family	481	94	86	1,72E-72
F-P2C09	AAL25839	non-specific lipid-transfer protein	449	71	66	2,06E-19
F-P1C10	XP_002511263	callose synthase 10	535	24	88	6,25E-06
F-P10C10	XP_002306908	gaga-binding transcriptional activator	516	99	83	2,00E-64
F-P2C10	XP_003526688	ubiquitin-conjugating enzyme e2	543	70	100	1,62E-82
F-P3C10	XP_002319182	exosome component 10-like	402	99	90	1,65E-54
F-P7C10	XP_002310568	pinus taeda anonymous locus	181	93	87	1,27E-21
F-P1C11	XP_002301360	signal peptide peptidase-like 2b-like	443	57	95	3,46E-41
F-P9C11	XP_002520074	zinc-binding protein	558	70	90	2,67E-54
F-P10C11	XP_002516867	probable serine threonine-protein kinase at1g54610-like	516	81	92	4,13E-60

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

(continuação)						
Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	Similaridade (%)	e-value
F-P3C11	XP_002883406	isoleucyl-trna synthetase	390	55	69	2,66E-09
F-P8C11	CBI21935	acyl- oxidase acx3	546	99	93	2,34E-104
F-P5C12	XP_002320779	digalactosyldiacylglycerol synthase	566	100	95	1,34E-103
F-P2C12	XP_002304373	serine threonine protein kinase	486	43	82	3,45E-31
F-P3C12	XP_002299770	predicted protein [P. trichocarpa]	434	47	76	9,44E-06
F-P7C12	XP_002298324	peptide transporter ptr1-like	523	69	91	1,55E-61
F-P4D01	XP_002336224	ubiquitin carboxyl-terminal hydrolase 5-like	494	100	79	2,21E-50
F-P5D01	XP_002519595	u3-55k-like protein 2	540	93	89	1,15E-85
F-P6D01	XP_002324686	receptor-like kinase	476	100	98	2,78E-100
F-P10D01	XP_002525758	light-harvesting complex ii protein lhcb3	410	100	100	5,25E-87
F-P2D01	XP_002317377	protein	433	100	92	1,88E-80
F-P7D01	XP_002307988	pyrroline-5-carboxylate reductase	520	94	93	5,70E-92
F-P8D01	XP_002301626	glutamate receptor	559	99	96	1,19E-90
F-P1D02	AAC16011	class i chitinase	454	36	75	1,83E-11
F-P9D02	XP_002310753	f-box family protein	566	86	71	2,05E-43
F-P10D02	XP_002527835	auxin response factor 1	487	100	89	3,50E-75
F-P3D02	CAA67155	ferredoxin-dependent glutamate synthase 1	539	100	95	1,62E-114
F-P7D02	XP_002325221	ein3-binding f-box protein	339	97	88	8,89E-37
F-P8D02	XP_002309725	protein	487	76	95	2,74E-67
F-P4D03	XP_002509676	histone deacetylase complex subunit sap18	516	72	90	7,15E-59
F-P3D03	AAF79822	elongation factor 1-	406	78	62	1,93E-16
F-P8D03	XP_002324397	plasma membrane h+ atpase	269	99	96	1,66E-31
F-P1D04	XP_003519171	protein	327	88	95	2,34E-50
F-P4D04	XP_002309124	glycerol-3-phosphate acyltransferase 6	504	99	97	2,94E-100
F-P5D04	XP_002510918	40s ribosomal protein	466	41	92	5,07E-26
F-P6D04	XP_002311194	guanylyl cyclase	481	56	52	8,78E-08
F-P10D04	XP_002513274	gtp-binding protein	580	100	97	1,54E-113
F-P3D04	XP_002526899	yippee-like protein	323	33	94	9,07E-07
F-P7D04	XP_002509581	elongation factor chloroplastic-like	539	81	75	1,65E-26
F-P8D04	XP_002312401	protein	497	77	79	3,62E-30
F-P1D05	XP_002519374	gtp-binding protein	510	97	74	1,79E-52
F-P4D05	XP_002516815	transport inhibitor response 1	507	99	94	1,56E-89
F-P5D05	CBI40244	alkaline neutral invertase	488	92	99	1,21E-94
F-P10D05	XP_003539275	hexose transporter	593	100	95	2,38E-113
F-P3D05	CBI14927	protein	444	99	95	5,61E-82
F-P4D06	XP_002513189	swap domain-containing partial	549	91	92	9,63E-94
F-P5D06	ABF51620	myo-inositol-1-phosphate synthase	538	99	100	1,13E-118
F-P6D06	XP_002304601	molybdenum cofactor sulfurase family protein	523	69	90	5,60E-60
F-P9D06	XP_002328991	peroxidase 4	515	76	93	7,18E-68
F-P3D06	NP_001237940	amp dependent	382	96	72	4,72E-39
F-P8D06	XP_002319454	aspartic proteinase	557	100	88	8,56E-96
F-P4D07	CBI18357	high chlorophyll fluorescence phenotype 173 protein	502	100	97	1,13E-102
F-P5D07	XP_002299650	tyrosine-specific transport	532	78	94	9,97E-77
F-P9D07	XP_002318216	tripeptidyl peptidase ii	493	100	96	4,74E-92
F-P10D07	XP_002319667	lung seven transmembrane	541	100	85	6,16E-89
F-P2D07	XP_002515313	ctp synthase	502	38	92	1,54E-28

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

(continuação)						
Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	Similaridade (%)	e-value
F-P5D08	XP_002266350	thioredoxin y	511	51	86	2,03E-33
F-P6D08	XP_002330375	aldehyde dehydrogenase	529	48	84	1,88E-30
F-P9D08	AEA50860	partial	524	33	98	5,10E-13
F-P10D08	XP_002300622	chromatin remodeling complex	105	100	94	3,67E-05
F-P2D08	XP_002330491	pollen-specific protein sf21-like	418	29	92	1,40E-07
F-P1D09	XP_002298779	adp-ribosylation factor-like protein	537	94	96	3,38E-108
F-P4D09	AEZ67304	chitinase-like protein	530	71	94	1,98E-70
F-P5D09	CAN62820	phox domain-containing protein	313	58	78	1,01E-13
F-P3D09	XP_002320487	1-aminocyclopropane-1-carboxylate oxidase	395	81	93	8,16E-55
F-P8D09	XP_003625677	abc transporter b family member 1-like	518	100	96	2,44E-100
F-P5D10	BAF98287	isopentenyl-diphosphate delta-isomerase	528	71	95	3,17E-72
F-P6D10	XP_002298808	bromodomain-containing protein	505	63	69	5,65E-15
F-P9D10	AEK05572	abscisic insensitive 1b	588	100	67	9,73E-37
F-P10D10	XP_002518737	o-acyltransferase wsd1-like	460	44	80	1,75E-18
F-P3D10	XP_002262829	calcium-transporting atpase plasma membrane-type-like	438	30	81	5,76E-04
F-P1D11	XP_002322436	arginine serine-rich-splicing factor rsp40	419	42	98	2,34E-25
F-P4D11	XP_002513881	acetyl- carboxylase	529	27	89	3,76E-08
F-P5D11	XP_002319194	predicted protein [P. trichocarpa]	453	52	69	2,68E-07
F-P9D11	XP_002269478	metal tolerance protein c1	495	47	84	3,20E-18
F-P7D11	XP_002519195	amp dependent	518	100	93	7,09E-91
F-P8D11	XP_002304129	subtilase family protein	583	94	95	2,68E-102
F-P5D12	XP_002525098	nedd8-conjugating enzyme ubc12	561	84	97	2,42E-102
F-P9D12	XP_002510912	fk506-binding protein	563	87	79	1,34E-52
F-P10D12	ADR82372	pectinacetylesterase family protein	492	98	90	2,60E-87
F-P2D12	XP_002525726	syntaxin 1b 2 3	474	87	98	2,02E-66
F-P3D12	XP_002533431	afg1-like protein	571	99	96	1,90E-116
F-P7D12	XP_002524350	tryptophan aminotransferase-related protein 4-like	510	100	86	1,30E-76
F-P8D12	XP_002318112	protein kinase family protein	486	65	93	3,02E-48
F-P1E01	XP_002285135	pentatricopeptide repeat-containing	434	62	98	7,44E-54
F-P4E01	XP_002525348	equilibrative nucleoside transporter 4-like	542	63	93	1,15E-56
F-P5E01	XP_002302494	protein kinase g11a-like	523	100	84	1,23E-72
F-P9E01	XP_002269218	pinus taeda anonymous locus	472	29	95	3,43E-12
F-P2E01	CBI32664	udp-glucuronic acid decarboxylase	503	67	97	1,29E-63
F-P3E01	XP_002320284	heme binding protein	479	67	87	2,39E-46
F-P7E01	XP_002325998	dihydrolipoyllysine-residue acetyltransferase component	562	60	86	3,28E-43
F-P8E01	XP_002523319	kinesin light	525	50	57	2,02E-06
F-P5E02	AAK14060	mlp-like protein 28	348	57	81	1,87E-18
F-P6E02	XP_002531721	hydroxypyruvate reductase	570	82	98	1,02E-86
F-P10E02	XP_002319276	predicted protein [P. trichocarpa]	592	100	88	2,22E-90
F-P7E02	XP_002514115	zinc finger cch domain-containing protein 64	554	100	84	7,26E-86
F-P6E03	XP_002278872	homeobox-leucine zipper protein hat5-like	612	38	84	3,24E-08
F-P10E03	CAN65806	aspartyl protease family protein	497	100	60	5,06E-32
F-P3E03	XP_002525907	autophagy-related protein 8f	467	68	96	1,90E-57
F-P8E03	XP_002310187	cold acclimation protein	370	88	93	3,77E-52
F-P1E04	XP_002305660	predicted protein [P. trichocarpa]	533	73	83	9,48E-27

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

(continuação)						
Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	Similaridade (%)	e-value
F-P4E04	XP_003543688	wd repeat-containing	524	99	94	4,52E-88
F-P5E04	XP_002320506	f-box family protein	530	100	88	8,06E-93
F-P6E04	XP_002520708	ndx1 homeobox protein	464	95	73	1,74E-45
F-P10E04	EEE68481	60s ribosomal protein l7a	532	100	97	2,49E-112
F-P7E04	XP_002329276	predicted protein [P. trichocarpa]	425	90	57	7,78E-13
F-P8E04	AAU93594	40s ribosomal protein s9	409	84	99	3,67E-68
F-P4E05	XP_002337974	transmembrane receptors atp binding protein	462	81	76	2,50E-32
F-P5E05	XP_002530734	transducin wd40 domain-containing	526	27	97	4,19E-16
F-P10E05	AAL68843	aminoalcoholphosphotransferase	422	53	95	3,59E-34
F-P2E05	NP_001242213	mitogen-activated protein kinase 4	377	99	95	2,06E-69
F-P3E05	CBI16673	nuclear antigen homolog	575	100	78	2,57E-45
F-P8E05	XP_003518705	ring u-box domain-containing	472	44	98	2,25E-38
F-P4E06	XP_002330999	multiple c2 and transmembrane domain-containing protein 2-like protein	484	73	95	1,74E-53
F-P5E06	XP_002524643	calcium dependent protein kinase 1	512	87	77	2,70E-56
F-P8E06	XP_002267735	uncharacterized protein	567	95	53	1,67E-21
F-P4E07	NP_001119115	hexose transporter	578	21	92	5,11E-07
F-P5E07	XP_002305649	probable carboxylesterase 2-like	559	88	96	1,64E-70
F-P9E07	XP_002510250	alpha-l-fucosidase 1	601	90	82	7,37E-69
F-P2E07	AAC98456	aluminum-activated malate transporter 9-like protein iq-domain 1	419	49	94	1,03E-31
F-P3E07	XP_002328660	protein iq-domain 1	527	44	83	4,01E-10
F-P7E07	XP_002525175	rwp-rk domain-containing protein	175	75	95	1,61E-13
F-P4E08	XP_002301919	early flowering 4	474	100	86	1,73E-72
F-P5E08	XP_002299809	aspartic proteinase nepenthesin-1	614	23	93	1,06E-15
F-P6E08	XP_002532530	gdp-l-galactose phosphorylase	562	61	91	2,01E-49
F-P9E08	XP_002530359	PREDICTED: uncharacterized protein [Vitis vinifera]	551	100	92	3,20E-88
F-P10E08	XP_002269529	26s proteasome regulatory subunit	555	68	70	6,69E-31
F-P8E08	XP_002300175	cysteine synthase	501	64	95	1,44E-49
F-P6E09	XP_003554689	nucleic acid binding	508	78	96	6,59E-65
F-P9E09	XP_002278933	cytochrome p450 78a3-like	588	82	59	2,46E-21
F-P2E09	XP_002514632	luminidependens	467	57	92	1,77E-39
F-P3E09	XP_002520657	iron-sulfur cluster assembly protein	558	57	92	1,88E-46
F-P8E09	XP_002267203	nucleotidyltransferase family partial	463	86	66	4,48E-30
F-P1E10	XP_002329093	peptide chain release factor erf subunit 1	511	33	87	1,60E-13
F-P5E10	NP_199599	formamidase-like protein	513	100	99	1,70E-112
F-P6E10	XP_002525256	protein	489	85	89	1,44E-71
F-P3E10	XP_002297912	long chain acyl- synthetase 2-like	512	47	95	2,45E-29
F-P7E10	XP_002313075	universal stress protein family	546	40	90	1,11E-25
F-P4E11	XP_002534082	cytochrome p450	515	41	92	1,19E-29
F-P5E11	XP_002531093	rna recognition motif-containing	525	67	85	1,31E-43
F-P9E11	XP_002519921	40s ribosomal protein	585	58	69	2,86E-26
F-P3E11	NP_001239976	cytosolic fe-s cluster assembly factor nbp35	521	74	93	2,84E-47
F-P7E11	ACJ84230	af375428_1at4g29590 t16l4_100	504	76	98	3,25E-74
F-P1E12	XP_002514141	peroxidase 12	520	91	87	2,04E-82
F-P9E12	XP_002302162	phosphatidylserine decarboxylase	502	91	91	7,10E-81
F-P2E12	NP_001234166	transport inhibitor response 1	536	53	86	6,69E-34
F-P7E12	CBI32265	leucine aminopeptidase 1	471	66	96	1,71E-53
F-P8E12	XP_002325134		508	95	85	1,97E-58

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

(continuação)						
Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	Similaridade (%)	e-value
F-P1F01	XP_002520193	upf0392 protein rcom_0530710-like	517	98	72	3,94E-54
F-P4F01	XP_002528196	transcription factor, putative	517	100	72	3,43E-25
F-P5F01	XP_002314045	af370611_1 protein	549	65	76	6,11E-40
F-P9F01	XP_002318767	auxin response factor 2	531	76	75	1,02E-42
F-P3F01	XP_002300381	ribosomal rna small subunit methyltransferase	585	100	90	2,91E-103
F-P7F01	XP_002529460	ribonuclease 3-like protein 2-like	469	99	83	1,68E-56
F-P8F01	XP_002324966	syntaxin 32	496	62	82	1,44E-29
F-P1F02	XP_002306118	dihydrofolate reductase-thymidylate synthase	550	100	98	2,55E-121
F-P4F02	CBI20877	esterase lipase domain-containing	504	99	91	3,83E-93
F-P6F02	XP_002528465	2-oxoglutarate e1 component	485	69	85	1,01E-45
F-P2F02	NP_001234234	threonine dehydratase chloroplastic-like	469	42	78	1,59E-17
F-P8F02	XP_002530384	allene oxide cyclase	560	23	92	1,21E-14
F-P4F03	XP_002311277	yth domain family protein	503	92	92	1,96E-85
F-P9F03	XP_003529483	telomerase-binding protein est1a	521	44	75	2,38E-14
F-P10F03	XP_002298482	2og-fe oxygenase family protein	429	34	83	3,49E-07
F-P2F03	XP_002317586	dna polymerase	455	98	90	6,49E-79
F-P3F03	XP_002306073	2-hydroxy-6-oxo-6-phenylhexa- -dienoate hydrolase	383	78	74	7,12E-27
F-P1F04	XP_002305709	protein	443	92	90	1,13E-58
F-P4F04	XP_002324538	probable gmp synthase	484	76	92	3,71E-60
F-P5F04	XP_002516674	ubiquitin ligase sinat3	406	100	97	1,56E-86
F-P6F04	AAC49847	potassium transporter	466	26	95	4,51E-10
F-P10F04	XP_002314276	mlo-like protein 6	360	100	88	4,71E-51
F-P7F04	XP_002300771	ribophorin ii family protein	506	96	83	1,14E-66
F-P8F04	AAW48295	pore-forming toxin-like protein hfr-2	506	44	68	7,77E-10
F-P1F05	XP_002526551	histone-lysine n-methyltransferase	557	85	66	4,86E-25
F-P10F05	XP_002309233	peptidase m28 family protein	536	57	93	1,36E-41
F-P3F05	AFK38756	s-adenosyl methionine synthetase 2	526	58	97	8,31E-56
F-P7F05	XP_002517206	hormone-sensitive lipase	600	100	80	1,89E-87
F-P8F05	AAG01888	sucrose phosphate synthase	559	45	94	1,11E-42
F-P4F06	XP_002313362	nuclear cap-binding protein	557	31	86	1,56E-17
F-P2F06	CBI37638	protein	204	51	91	6,90E-09
F-P3F06	CAG28306	proton-exporting atpase	569	56	95	5,13E-41
F-P7F06	XP_002527865	auxin-regulated protein	483	91	82	1,97E-56
F-P8F06	ABR25841	40s ribosomal protein s8	551	69	87	7,63E-51
F-P4F07	XP_002520440	conserved hypothetical protein	557	91	75	4,14E-61
F-P5F07	XP_002322899	ctr1-like protein kinase	559	55	67	5,38E-16
F-P3F07	AFR40706	partial	469	99	90	3,85E-89
F-P5F08	XP_003521531	transport inhibitor response 1	501	99	91	3,62E-87
F-P6F08	XP_002317760	cyclic nucleotide-gated ion channel	541	69	80	6,69E-42
F-P2F08	XP_002313432	telomere repeat-binding protein 5	446	40	90	7,73E-18
F-P3F08	AFK33626	s-adenosylmethionine synthetase	515	83	98	1,20E-94
F-P7F08	XP_002509843	homeobox protein bel1 homolog	482	56	77	4,11E-17
F-P1F09	XP_003605958	b-cell receptor-associated 31-like	452	100	84	2,27E-52
F-P5F09	XP_002318349	predicted protein [P. trichocarpa]	543	38	70	1,17E-03
F-P9F09	AAO25536	cellulose synthase	532	99	92	3,47E-83
F-P3F09	ABK93956	acyl carrier protein	539	76	88	4,92E-62
F-P7F09	AAD30294	ascorbate peroxidase	496	83	91	6,69E-62
F-P1F10	XP_002312978	extensin-like protein	524	59	86	4,25E-44

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

(continuação)						
Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	Similaridade (%)	e-value
F-P5F10	XP_003631566	amino acid	536	100	92	1,77E-96
F-P9F10	AFU95166	ribosomal protein l36	518	21	100	3,82E-09
F-P10F10	BAJ61942	ubiquitin-like protein	513	67	100	3,45E-69
F-P2F10	XP_002532229	transparent testa 12	442	88	80	5,02E-48
F-P3F10	NP_001237530	protein	493	95	85	5,63E-78
F-P1F11	ABK92551	homeobox protein knotted-1-like	444	78	89	6,29E-41
F-P4F11	XP_002530392	protein acclimation of photosynthesis to environment	407	71	86	4,65E-22
F-P6F11	XP_002516130	secologanin synthase-like	469	100	82	4,37E-62
F-P9F11	ABU85518	ribosomal protein s4	505	41	95	2,51E-29
F-P2F11	XP_003603304	stromal cell-derived factor 2-like	356	100	90	1,09E-63
F-P3F11	XP_002520507	tropinone reductase	539	78	80	7,45E-52
F-P4F12	XP_002316371	gata transcription factor 9-like	551	77	66	3,10E-24
F-P5F12	XP_002526854	peroxisomal membrane protein	610	54	96	4,12E-58
F-P6F12	XP_002320236	thylakoid membrane phosphoprotein	554	66	88	1,53E-51
F-P9F12	XP_002519286	dihydrolipoamide dehydrogenase	577	67	95	1,02E-67
F-P10F12	XP_002524867	oas-tl4 cysteine synthase	494	51	92	3,09E-40
F-P8F12	XP_002509816	ww domain-containing oxidoreductase-like isoform 2	506	75	92	1,61E-64
F-P1G01	XP_002514774	cohesin subunit	372	85	61	2,05E-12
F-P4G01	XP_002526972	protein	498	99	89	7,19E-86
F-P6G01	ACJ11726	serine partial	441	99	98	1,10E-93
F-P9G01	XP_002531325	metal abc	532	46	79	1,01E-20
F-P10G01	XP_002313701	gcip-interacting family	583	46	78	1,59E-25
F-P3G01	XP_002528983	u-box domain-containing protein	528	100	93	1,00E-80
F-P4G02	XP_002512253	cysteine synthase	558	82	96	5,25E-91
F-P6G02	XP_002524110	gibberellin receptor	562	93	87	4,07E-93
F-P10G02	XP_002529166	glutaredoxin family protein	514	72	57	2,92E-11
F-P3G02	NP_001236219	photosystem ii 10 kda polypeptide	576	68	91	6,60E-55
F-P7G02	AEK69741	triosephosphate isomerase	560	67	92	2,59E-46
F-P1G03	NP_001032145	acyl-coenzyme a oxidase 2	431	100	97	1,23E-83
F-P4G03	XP_002334885	protein	446	48	97	2,27E-22
F-P7G03	CBI30590	dof zinc finger protein	394	52	80	9,75E-18
F-P4G04	XP_002525341	acyl- oxidase	503	50	94	2,04E-34
F-P5G04	ABK32930	phosphoribulokinase	496	53	94	5,23E-44
F-P6G04	XP_002314265	sit4 phosphatase-associated family protein	571	99	91	7,73E-86
F-P9G04	XP_002306554	f-box family protein	539	100	88	2,64E-89
F-P9G05	XP_002532559	calcium-dependent protein kinase 33	591	99	96	1,98E-120
F-P10G05	XP_002299409	embryonic flower 2	577	51	94	5,13E-49
F-P3G05	AFR42124	alpha- partial	504	67	99	1,14E-70
F-P7G05	XP_002315298	light-harvesting complex i protein lhca1	591	30	96	1,68E-25
F-P4G06	AAG13459	gamma-glutamylcysteine synthetase	515	100	94	3,50E-104
F-P6G06	XP_002882497	enoyl- hydratase 3-hydroxyacyl-dehydrogenase	570	37	85	6,49E-16
F-P2G06	XP_002307813	plastid movement impaired1	314	99	86	6,91E-38
F-P7G06	XP_002297796	protein	499	99	87	5,85E-75
F-P4G07	NP_001238183	secretory peroxidase	492	76	96	7,92E-63
F-P5G07	XP_002512281	translation initiation factor eif-2b epsilon subunit	579	62	92	1,60E-63

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

(continuação)						
Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	Similaridade (%)	e-value
F-P6G07	XP_002324038	predicted protein [P. trichocarpa]	578	79	54	1,83E-13
F-P10G07	XP_002525258	casein kinase ii alpha subunit	490	46	92	1,95E-33
F-P7G07	XP_002297666	esterase lipase domain-containing protein	533	92	69	3,20E-45
F-P4G08	XP_002332050	protein phosphatase 2c 60	528	49	93	3,60E-38
F-P6G08	XP_002525391	atp binding	491	46	77	1,80E-19
F-P9G08	XP_002330220	rna binding	350	45	96	1,34E-06
F-P3G08	XP_002305670	lysine ketoglutarate reductase trans-splicing related 1	501	80	84	5,44E-59
F-P7G08	XP_003605914	transparent testa 12 protein	500	51	91	3,31E-30
F-P4G09	XP_002525647	peptidyl-prolyl cis-trans isomerase	537	99	94	4,52E-97
F-P5G09	XP_002532314	s-acyltransferase tip1	575	100	86	6,49E-98
F-P6G09	XP_002318164	serine carboxypeptidase-like 45	521	100	95	2,14E-87
F-P9G09	XP_002325134	leucine aminopeptidase	588	71	89	4,88E-69
F-P3G09	XP_002327529	phospholipase d alpha	461	99	100	6,85E-89
F-P1G10	BAD83547	transport membrane protein	531	39	85	2,55E-20
F-P4G10	XP_002301836	glutathione-dependent formaldehyde dehydrogenase	450	51	77	3,78E-28
F-P5G10	XP_002301126	receptor-like protein kinase hsl1-like	487	100	87	9,41E-66
F-P6G10	XP_002324775	mitogen-activated protein kinase kinase 2	479	100	93	3,36E-88
F-P2G10	XP_002514149	transcripton factor	459	101	67	3,71E-43
F-P3G10	XP_002319031	mitochondrial inner membrane translocase subunit tim44-related protein	491	40	66	2,02E-07
F-P5G11	XP_002314140	proline-rich protein	494	98	90	5,11E-83
F-P10G11	XP_002517918	beta-galactosidase 3	563	87	78	3,78E-34
F-P1G12	XP_002310279	probable lrr receptor-like serine threonine-protein kinase at5g63710-like	539	100	97	1,52E-107
F-P5G12	XP_002300217	zeaxanthin epoxidase	533	88	94	7,26E-73
F-P6G12	XP_002531782	aminophospholipid atpase	545	47	89	2,48E-36
F-P9G12	XP_002514445	nad-dependent epimerase dehydratase	492	91	93	7,91E-74
F-P10G12	XP_002530987	brca1 brca2-containing complex subunit 3	550	80	96	1,40E-85
F-P3G12	XP_002299607	cdk5rap3-like protein	502	100	76	6,87E-63
F-P7G12	XP_002331003	zinc finger	553	105	88	8,16E-81
F-P5H01	XP_002268140	b-box 32 protein	528	33	91	2,33E-21
F-P10H01	XP_002316440	pinus taeda anonymous locus	499	67	68	2,08E-21
F-P2H01	XP_002513740	upf0420 protein c16orf58-like	441	54	90	9,73E-32
F-P3H01	XP_002309136	50s ribosomal protein l13	472	100	81	1,67E-58
F-P5H02	XP_002532120	protein	518	85	97	1,41E-78
F-P10H02	XP_002309815	acetylornithine deacetylase	581	24	97	1,84E-14
F-P2H02	BAA95409	arginine serine-rich splicing	500	79	52	4,64E-17
F-P7H02	XP_002308490	glyceraldehyde-3-phosphate dehydrogenase	404	26	100	7,35E-07
F-P4H03	AFE48357	beta galactosidase 1	545	99	99	5,01E-103
F-P5H03	ABK96254	protein	479	100	94	3,41E-93
F-P6H03	XP_002329734	alpha-l-arabinofuranosidase	477	52	82	5,26E-28
F-P9H03	XP_002523672	ca2+ antiporter cation exchanger	469	100	92	1,31E-77
F-P3H03	XP_002534466	protein decapping 5	544	90	81	1,19E-63
F-P7H03	XP_002311428	mitochondrial translational initiation	460	100	73	9,88E-33
F-P8H03	XP_002300479	upf0420 protein c16orf58 homolog	521	67	59	4,21E-05
F-P1H04	XP_002525869		478	99	98	1,26E-84

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

(continuação)

Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	Similaridade (%)	e-value
F-P4H04	XP_003553926	dcd (development and cell death) domain protein	485	84	98	3,77E-88
F-P5H04	XP_002328915	selenium-binding protein	505	91	90	4,17E-83
F-P6H04	ABK92596	atp binding	577	30	94	2,44E-25
F-P9H04	AAX88799	catalase	316	99	100	2,75E-61
F-P2H04	XP_002310840	transmembrane and coiled-coil domain-containing protein	500	56	84	4,63E-32
F-P8H04	XP_002329736	dihydropyrimidine dehydrogenase	582	67	85	8,99E-55
F-P6H05	XP_002323987	purple acid phosphatase	496	100	93	1,03E-89
F-P3H05	ABN46984	peroxidase	547	95	92	1,58E-92
F-P8H05	XP_002533345	late embryogenesis abundant protein lea14-5	536	90	96	7,88E-94
F-P4H06	XP_002510514	5'-3' exoribonuclease 3-like	533	100	94	1,97E-98
F-P5H06	XP_002519539	plasma membrane intrinsic protein	493	77	92	7,10E-58
F-P6H06	XP_002323470	chloroplast rna-binding protein 29	537	55	85	7,64E-35
F-P10H06	XP_002311837	nuclear transcription factor y subunit c-9	526	71	88	1,54E-38
F-P2H06	ABK96632	thioredoxin-like domain-containing	450	91	94	8,76E-80
F-P4H07	XP_002528404	pentatricopeptide repeat-containing	572	100	88	1,82E-94
F-P6H07	XP_002280268	beta-glucosidase-like sfr2	563	100	81	9,03E-73
F-P9H07	XP_002325360	gds1 esterase lipase at5g33370-like	514	53	87	2,94E-33
F-P10H07	XP_002515578	nudix hydrolase 24	594	98	70	6,96E-55
F-P9H08	XP_003523184	staphylococcal nuclease domain-containing protein 1-like	529	90	89	3,50E-80
F-P3H08	XP_002517092	homocysteine s-methyltransferase	528	75	88	2,10E-63
F-P7H08	XP_002298535	monodehydroascorbate reductase	574	100	95	6,73E-111
F-P8H08	AAT77893	calcium-transporting atpase endoplasmic reticulum-type-like conserved hypothetical	487	97	92	2,21E-65
F-P4H09	XP_002510707	conserved hypothetical	482	90	74	4,38E-26
F-P9H09	XP_002532081	ribose phosphate pyrophosphokinase	506	79	86	4,69E-62
F-P2H09	XP_003526921	nucleotide binding	507	64	75	1,39E-19
F-P3H09	CAN74770	o-linked c transferase like protein	493	83	97	1,04E-76
F-P7H09	XP_002338082	polyubiquitin 2	442	53	100	2,74E-42
F-P8H09	XP_002512251	protein	401	93	73	3,21E-30
F-P1H10	AEA86345	abc transporter b family member 28	460	100	96	2,87E-71
F-P4H10	XP_002318264	alpha-l-fucosidase 1-like	501	100	94	1,49E-96
F-P5H10	XP_002328967	ac007230_9ests gb	155	99	86	2,27E-15
F-P9H10	XP_002307071	gtp-binding protein	488	50	96	2,30E-40
F-P2H10	CAN78103	protein suppressor of gene silencing-like protein	445	62	65	3,66E-05
F-P3H10	XP_002512346	beta-n-acetylglucosaminyltransferase	489	61	97	5,63E-53
F-P7H10	AEV66272	probable wrky transcription factor 33	508	100	85	2,02E-73
F-P8H10	CAL07992	transketolase 1	548	99	98	1,97E-119
F-P1H11	XP_002525369	probable nucleoredoxin 1-like	472	83	85	4,85E-58
F-P4H11	XP_002310349	beta-amyrin synthase	448	61	92	1,09E-38
F-P9H11	ACV60361	type 2 ribosome-inactivating protein	562	83	64	1,99E-32
F-P10H11	XP_002526335	senescence-associated protein	529	78	98	3,75E-59
F-P2H11	XP_002300909	zinc finger	460	97	89	3,04E-72
F-P3H11	AFU93068	1-deoxy-d-xylulose 5-phosphate synthase	423	79	75	2,96E-39
F-P8H11	XP_002509843	homeobox protein bel1 homolog	513	100	82	9,89E-43
F-P6H12	CBI26938	atp synthase beta subunit	496	99	99	2,12E-105
F-P9H12	XP_002516524	protein	612	100	99	7,81E-124

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

(continuação)						
Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	Similaridade (%)	e-value
F-P10H12	XP_002306856	protein	545	55	81	2,10E-33
F-P2H12	XP_002319980	root phototropism protein 2	502	64	85	4,28E-40
F-P8H12	XP_002318635	40s ribosomal protein	554	47	96	5,95E-45
F-P1C12	XP_003588355	cell wall-associated hydrolase	384	88	61	3,17E-23
F-P2D04	AFU96047	ribosomal protein l2	487	78	92	1,06E-64
F-P4D08	YP_006666139	cytochrome c oxidase subunit 2	551	85	84	2,10E-57
F-P7F12	BAD94507	glutamine synthetase	524	68	97	3,06E-60
F-P1G11	BAA10929	cytochrome p450 like_tbp	485	33	88	8,60E-29
F-P7A11	ADJ67447	photosystem ii cp43 chlorophyll apoprotein	478	54	97	9,17E-39
F-P6B09	AFU96504	atp synthase cf0 subunit i	562	89	97	1,50E-105
F-P8C03	NP_084841	ribosomal protein s7	563	79	97	2,45E-94
F-P1C05	ACN66393	atp synthase beta subunit	346	100	99	1,31E-66
F-P1E11	AEP95750	aminomethyltransferase	544	81	91	1,80E-79
F-P5F02	ABU85499	cytochrome f	527	100	97	6,97E-110
F-P7H06	XP_002515161	conserved hypothetical protein	616	24	90	1,90E-11
F-P6H11	XP_002515414	glutathione peroxidase	535	60	80	2,99E-26
F-Contig2	XP_002298508	p8mtcp1	514	51	95	3,18E-42
F-Contig5	Q41596	class i chitinase	566	74	61	3,16E-26
F-Contig6	XP_002519229	atp citrate lyase a-subunit	835	43	96	4,96E-73
F-Contig8	ADC94855	dna binding	690	63	88	2,53E-51
F-Contig9	XP_002521339	rossmann-fold nad -binding domain-containing protein	437	53	87	2,31E-26
F-Contig10	XP_003635583	lil3 protein	520	77	91	8,45E-74
F-Contig11	YP_005090236	hypothetical chloroplast rf1	557	89	66	1,56E-31
F-Contig13	XP_002298160	replication factor c subunit 1-like	568	60	77	1,64E-09
F-Contig14	XP_002527344	cbl-interacting serine threonine-protein kinase 8	467	100	90	7,47E-86
F-Contig16	XP_002321160	iq domain-containing protein	598	97	75	3,57E-67
F-Contig18	ABK63195	wound-induced protein win2	512	107	68	2,45E-63
F-Contig19	XP_002325337	rna-binding protein 42-like	505	37	98	4,56E-27
F-Contig20	XP_002534347	rubisco subunit binding-protein alpha subunit	595	100	99	4,22E-108
F-Contig22	XP_002313619	short-chain dehydrogenase reductase	479	88	90	2,61E-71
F-Contig24	XP_002336529	aim1 protein	553	99	94	4,11E-92
F-Contig26	ADR71297	at4g32020 f10n7_170	565	98	64	1,59E-26
F-Contig27	XP_002308990	galactokinase	545	95	82	9,61E-52
F-Contig33	NP_001241257	protein	583	72	56	4,82E-14
F-Contig34	XP_002301348	(+)-neomenthol dehydrogenase	521	85	88	3,58E-68
F-Contig35	XP_002313160	seed maturation protein pm23	475	99	75	1,03E-45
F-Contig37	ABY19575	actin	524	70	99	1,40E-79
F-Contig38	XP_002262872	serine hydroxymethyltransferase	611	99	94	4,45E-120
F-Contig40	XP_002321645	cytochrome p450	536	60	91	6,15E-52
F-Contig43	XP_002326049	protein	520	21	91	2,22E-07
F-Contig44	CAL07991	glycogen synthase kinase-3	593	53	94	4,52E-56
F-Contig47	XP_003626526	thylakoid membrane phosphoprotein 14 chloroplast	804	18	100	2,79E-16
F-Contig48	XP_002329316	protein	579	95	81	8,02E-60
F-Contig49	AAX84672	aldo keto	598	81	96	8,80E-93
F-Contig50	XP_002307284	adp-ribosylation factor 1	544	74	98	2,69E-82
F-Contig51	XP_002334345	rieske domain containing expressed	521	100	95	9,29E-103
F-Contig52	XP_002332075	protein	508	89	82	3,83E-55

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

(continuação)						
Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	Similaridade (%)	e-value
F-Contig53	XP_002326118	germin-like protein	504	56	88	1,35E-39
F-Contig59	XP_002316361	atp binding	446	84	86	3,66E-54
F-Contig61	XP_003534985	c2 and gram domain-containing	772	61	94	2,80E-85
F-Contig62	XP_002309918	ras-related protein rab-18-b	596	89	97	5,58E-114
F-Contig66	XP_002329556	protein	572	97	68	3,40E-53
F-Contig67	XP_002330044	dead-box atp-dependent rna helicase 3	989	100	96	0.0
F-Contig68	CBI30828	o-acetyltransferase-like protein	669	17	92	2,95E-07
F-Contig69	AAU94363	glycine dehydrogenase	661	63	92	1,38E-75
F-Contig70	XP_002519198	calmodulin-binding transcription activator	529	100	69	8,42E-40
F-Contig71	XP_002315948	30s ribosomal protein 3-1	715	76	81	5,35E-71
F-Contig72	XP_002311056	nadp-dependent malate dehydrogenase	582	74	81	1,28E-46
F-Contig74	YP_006666139	cytochrome oxidase subunit 2	544	38	69	1,25E-09
F-Contig75	XP_002531721	hydroxypyruvate reductase	678	70	97	1,84E-87
F-Contig77	XP_002519280	e3 ubiquitin-protein ligase upl6	614	41	91	2,52E-33
F-Contig81	XP_002528466	40s ribosomal protein s15a	854	31	100	1,01E-47
F-Contig82	ABU96744	acyl acp-thioesterase	537	32	81	3,56E-07
F-Contig83	XP_002327389	tpv38 tmem64 family membrane protein slr0305	853	51	81	4,69E-27
F-Contig84	XP_003556440	ubiquitin carboxyl-terminal hydrolase family 1 protein	571	72	88	3,28E-53
F-Contig85	XP_002515139	cathepsin b-like cysteine protease	1197	87	85	0.0
F-Contig86	P46280	elongation factor chloroplastic-like	708	100	96	1,64E-130
F-Contig87	XP_002529138	serine threonine protein phosphatase 2a regulatory subunit	880	87	96	3,13E-168
F-Contig88	XP_002278777	protein	541	100	71	9,79E-66
F-Contig89	XP_003565852	pore-forming toxin-like protein hfr-2	480	53	60	5,39E-08
F-Contig90	XP_002511195	sap domain-containing protein	422	94	80	2,56E-27
F-Contig91	XP_002283426	mitochondrial-processing peptidase	550	71	85	1,89E-42
F-Contig138	AFS35502	rna polymerase beta subunit	392	98	98	1,88E-77
F-Contig139	AFU95067	photosystem i p700 apoprotein a2	393	57	100	2,32E-51
F-Contig140	ABU85539	nadh-plastoquinone oxidoreductase	767	62	98	3,43E-103
R-P1A11	XP_002268815	gtp-binding protein	500	86	96	5,77E-90
R-P1A12	AEQ94144	iron-inhibited abc transporter	471	99	97	4,78E-92
R-P1A01	AFV60741	transcription factor lim1	518	82	96	1,10E-89
R-P1A02	ABK94828	ribosome biogenesis gtp-binding protein	592	100	74	5,51E-63
R-P1A03	XP_002513166	conserved hypothetical protein	491	26	86	3,06E-08
R-P1A05	XP_002322794	vamp protein sec22	510	42	87	9,84E-25
R-P1A06	CAN78264	gcn5-related n-acetyltransferase family-like protein	571	88	86	1,84E-75
R-P1A07	XP_002308166	40s ribosomal protein	417	84	97	8,77E-63
R-P1A08	XP_002523587	zeaxanthin epoxidase	480	77	76	9,46E-20
R-P1B10	AEK64520	beta tubulin	365	99	100	2,82E-78
R-P1B12	ACA04028	galactinol synthase	342	50	80	1,76E-08
R-P1B02	XP_002515011	at1g67700 f12a21_30	352	84	77	1,19E-19
R-P1B03	XP_002510649	class i chitinase	328	39	93	6,55E-11
R-P1B04	XP_002314295	predicted protein [P. trichocarpa]	446	100	77	3,91E-25
R-P1B05	XP_002328924	shikimate kinase like 1	637	65	83	5,62E-54
R-P1B06	XP_002526048	gtp-binding protein gb2	550	99	97	3,18E-119
R-P1B07	XP_002530900	ring u-box domain-containing	494	83	80	4,76E-53
R-P1B08	CBI22256	unnamed protein [Vitis vinifera]	356	29	85	4,78E-03

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

(continuação)						
Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	Similaridade (%)	e-value
R-P1C10	XP_002872554	translation initiation factor if-2	563	100	90	1,01E-84
R-P1C02	XP_002308235	40s ribosomal protein	399	97	100	1,98E-64
R-P1C06	XP_002274043	double-stranded dna-binding-like	423	87	95	5,87E-41
R-P1C07	XP_002317252	rni-like superfamily protein	492	98	51	3,87E-11
R-P1C09	XP_002331331	cytochrome p450	407	93	75	1,45E-37
R-P1D10	XP_002309474	mechanosensitive ion channel domain-containing protein	446	100	94	3,08E-62
R-P1D02	XP_002311706	protein	347	97	96	1,29E-61
R-P1D07	XP_002519198	calmodulin-binding transcription activator	572	61	76	7,74E-30
R-P1D09	XP_002328949	tom1-like protein 2	577	42	86	6,88E-15
R-P1E10	XP_002513228	linoleate 13s-lipoxygenase 2-chloroplastic-like	487	97	85	5,92E-68
R-P1E11	XP_002297641	f-box family protein	446	75	78	1,62E-31
R-P1E12	XP_002312502	protein	340	74	89	1,05E-31
R-P1E01	XP_002270877	protein	434	67	87	1,19E-40
R-P1E02	ABY19575	actin	548	76	99	1,10E-92
R-P1E05	XP_002301123	serine threonine-protein kinase ctr1-	504	99	98	5,29E-108
R-P1E06	XP_003601798	60s ribosomal protein l17	483	85	94	1,83E-81
R-P1E08	XP_002310148	polcalcine jun	366	89	91	1,30E-48
R-P1E09	XP_002534464	nadph--cytochrome p450 reductase-like	387	98	85	2,41E-60
R-P1F11	ADL36693	gata transcription factor 5	260	48	85	5,21E-08
R-P1F01	XP_002532886	constans-2	523	77	92	4,87E-50
R-P1F02	NP_001239866	dcd (development and cell death)	505	24	78	2,28E-05
R-P1F04	XP_002532831	50s ribosomal protein l29	600	73	80	2,53E-55
R-P1F07	XP_002517672	haloalkane dehalogenase 2-like	382	79	68	2,34E-18
R-P1G10	XP_002324330	chorismate mutase	536	92	81	2,45E-78
R-P1G02	XP_002302794	glycogen synthase kinase-3 -3	246	79	96	4,89E-30
R-P1G03	ADR71286	40s ribosomal protein s5	469	88	89	1,53E-73
R-P1G04	XP_003536739	dihydroorotate dehydrogenase	559	100	91	4,08E-86
R-P1G08	XP_002304470	f-box family protein	635	56	75	2,77E-35
R-P1G09	ABK94888	protein lurp-one-related 8	363	47	91	2,83E-17
R-P1H10	XP_002330302	predicted protein [P. trichocarpa]	517	38	69	9,37E-08
R-P1H11	XP_002312089	heat shock	620	100	100	3,88E-135
R-P1H01	XP_002312333	lojap-related protein	431	94	87	3,23E-66
R-P1H02	XP_002306693	fructose- - cytosolic	398	87	95	5,43E-64
R-P1H03	XP_002517058	atp sulfurylase	428	100	93	1,06E-79
R-P1H04	XP_002518917	protein	536	100	85	2,53E-88
R-P1H05	XP_002532336	rhodanese-like domain-containing	410	100	93	7,60E-73
R-P1H06	XP_002314284	exosome complex component mtr3	518	94	90	2,86E-79
R-P1H07	XP_002302180	gt-2 factor	449	75	88	1,70E-44
R-P1H09	XP_002336676	predicted protein [P. trichocarpa]	334	75	59	2,31E-06
R-P2A10	XP_003634351	low quality protein: probable nad h-dependent oxidoreductase 1-like	408	55	82	2,47E-22
R-P2A12	AEI83424	ferredoxin- chloroplast	488	40	87	1,09E-06
R-P2A01	XP_002325025	vacuolar protein sorting-associated	469	65	87	5,66E-41
R-P2A03	XP_002511689	dna repair protein rad51-like protein	421	81	89	9,50E-33
R-P2A04	XP_002511018	embryonic flower 2	457	42	93	2,67E-15
R-P2A07	CAN79954	nucleotidyltransferase domain containing expressed	419	114	61	3,52E-31
R-P2A09	XP_002532451	udp-glycosyltransferase 1	400	100	84	5,14E-61
R-P2B01	XP_002314566	small heat shock protein	453	62	74	1,27E-22

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

(continuação)						
Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	Similaridade (%)	e-value
R-P2B02	XP_002302388	predicted protein [P. trichocarpa]	504	85	62	3,80E-26
R-P2B03	XP_002522846	trafficking protein particle complex	536	100	98	3,15E-116
R-P2B04	XP_002328502	hydroxycinnamoyl-coenzyme a shikimate quinate	539	63	85	1,61E-49
R-P2B07	ABK96170	protein	509	100	92	4,13E-96
R-P2B08	XP_002310677	probable lrr receptor-like serine threonine-protein kinase	441	100	89	3,99E-63
R-P2B09	XP_002514086	uncharacterized membrane	426	54	86	9,12E-24
R-P2C12	XP_002515306	chaperone protein	400	39	90	7,22E-03
R-P2C01	XP_002323124	heavy-metal-associated domain--containing protein	449	100	84	4,00E-53
R-P2C03	ACT87980	ascorbate peroxidase	340	56	96	4,32E-29
R-P2C04	XP_002319998	oligopeptidase a	436	75	95	1,04E-47
R-P2C06	XP_002264858	seed specific protein of balanced nutritional quality	580	63	54	5,74E-14
R-P2C07	XP_002338413	protein	278	78	93	1,91E-25
R-P2C08	XP_002330702	protein	406	100	81	1,87E-48
R-P2C09	XP_002285378	acetyl- carboxylase biotin-containing subunit	512	77	61	1,01E-15
R-P2D10	CAA44478	orf 873	366	47	75	1,67E-07
R-P2D11	ABK94763	mangrin	486	77	80	3,50E-45
R-P2D01	XP_002320429	sialyltransferase-like protein	577	98	93	1,52E-113
R-P2D04	XP_002530384	mangrin	404	98	89	5,09E-64
R-P2D05	AAL25839	lipid-transfer protein	419	62	74	1,34E-26
R-P2D06	XP_002524642	carbonic anhydrase	399	100	78	8,40E-27
R-P2D07	ABK94571	glycosyltransferase family 14	496	100	92	4,82E-91
R-P2D09	XP_002265080	malate dehydrogenase	344	55	96	3,66E-29
R-P2E11	XP_002515557	protein	365	77	91	2,72E-42
R-P2E12	XP_002512737	chaperone protein dnaJ	461	51	80	2,59E-18
R-P2E01	XP_002520925	photosystem ii stability assembly factor hcf136	491	100	75	2,76E-58
R-P2E02	XP_002305703	p-loop containing nucleoside triphosphate hydrolase-like protein	429	29	100	2,96E-12
R-P2E03	XP_002533495	60s ribosomal protein	544	73	97	1,96E-71
R-P2E04	XP_002524053	bel1-like homeodomain protein 1	428	88	84	2,67E-31
R-P2E05	XP_002315454	predicted protein [P. trichocarpa]	494	92	56	3,46E-19
R-P2E06	XP_002526925	probable carboxylesterase 11-like	489	99	93	1,12E-95
R-P2E07	XP_002519480	s-adenosylmethionine partial	406	95	99	5,98E-65
R-P2E09	CAL07988	elongation factor 1- partial	489	99	99	1,54E-105
R-P2F10	XP_002331201	exocyst complex component 7	298	100	98	2,24E-38
R-P2F11	XP_002301399	gh3 family protein	391	100	87	3,40E-64
R-P2F12	XP_002527504	mitotic checkpoint protein	429	100	96	5,99E-92
R-P2F01	XP_002319251	ac009360_1est gb	377	98	93	1,84E-63
R-P2F02	XP_002301052	predicted protein [P. trichocarpa]	492	76	62	2,62E-08
R-P2F03	XP_002318804	30s ribosomal protein s9	350	56	93	1,12E-27
R-P2F04	BAF02241	acyl-	501	99	98	1,45E-110
R-P2F06	XP_002326766	methyltransferase pmt20	484	100	98	6,86E-101
R-P2F07	Q6YW46	elongation factor 1-gamma	505	99	83	2,10E-50
R-P2F08	XP_002327350	protein	417	101	70	5,87E-40
R-P2G10	XP_002528465	oxoglutarate dehydrogenase	577	41	82	1,89E-21
R-P2G11	XP_002281028	heat shock factor	594	35	94	6,06E-39
R-P2G12	XP_002516469	Protein	463	97	89	6,51E-71
R-P2G01	BAM14583	glutathione s-transferase	434	76	79	1,74E-39

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

(continuação)						
Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	Similaridade (%)	e-value
R-P2G04	XP_002316869	Protein	494	86	80	4,75E-63
R-P2G05	XP_002321300	transducin wd40 domain-containing protein	529	44	91	2,04E-40
R-P2G06	XP_003518718	kelch repeat-containing f-box family	433	68	89	9,85E-45
R-P2G07	XP_002325907	thioredoxin f	412	55	60	7,00E-07
R-P2G08	XP_002327818	pentatricopeptide repeat-containing	427	83	78	1,16E-40
R-P2H10	XP_002304361	dead-box atp-dependent rna helicase 5	522	99	97	5,26E-106
R-P2H11	ABM91067	xyloglucan endotransglucosylase hydrolase	541	44	83	5,95E-21
R-P2H01	XP_002520216	iaa-amino acid hydrolase ilr1	461	27	95	9,21E-11
R-P2H02	XP_002308503	predicted protein [P. trichocarpa]	487	86	57	6,05E-15
R-P2H03	XP_002264345	cysteine proteinase rd19a	589	53	98	1,55E-60
R-P2H04	AAG38521	cysteine proteinase inhibitor 1-like	167	102	75	7,25E-11
R-P2H06	XP_002512793	otu domain-containing protein	446	59	89	3,63E-38
R-P2H09	XP_002866476	cxe carboxylesterase	360	99	81	1,12E-47
R-P3A10	XP_002326240	enolase	622	93	95	3,21E-120
R-P3A01	XP_002299268	dcn1-like protein 4	628	76	95	1,06E-86
R-P3A02	XP_002314695	lactoylglutathione lyase	643	65	88	4,28E-70
R-P3A05	XP_002519977	ribosomal protein l17-like protein	625	82	84	1,28E-80
R-P3A06	XP_002315762	cdc73 domain protein	594	99	70	6,02E-46
R-P3A07	XP_002517985	translocon at the inner envelope membrane of chloroplasts 20-iv	646	72	59	1,95E-23
R-P3A08	ABK96199	endo- β -glucanase	650	100	81	2,86E-93
R-P3B12	XP_002299371	protein toc75	616	86	91	6,87E-92
R-P3B01	ABK96772	universal stress protein a-like	648	76	84	3,70E-73
R-P3B02	XP_003523861	PREDICTED: uncharacterized protein [Glycine max]	578	28	72	4,73E-03
R-P3B03	XP_002534327	7-methylxanthosine synthase 1-like	594	64	77	2,68E-45
R-P3B04	XP_002297736	cc-nbs-lrr resistance protein	648	87	66	2,04E-37
R-P3B05	XP_002314975	galactinol synthase	635	78	88	1,36E-82
R-P3B06	XP_002282749	kinesin-like protein	585	87	97	5,82E-101
R-P3B07	XP_002512541	rna binding protein	552	100	95	4,79E-102
R-P3B09	ABK95374	enzyme of the cupin superfamily	614	77	78	6,37E-60
R-P3C10	BAD94507	glutamine synthetase	564	57	97	8,61E-51
R-P3C11	ABK94370	ankyrin repeat domain	623	52	85	3,04E-68
R-P3C12	XP_002299329	saur family protein	645	35	74	6,89E-12
R-P3C01	XP_002326624	strubbelig-receptor family 7 protein	576	53	76	2,41E-28
R-P3C02	XP_002308597	protein	585	74	91	1,25E-71
R-P3C05	AFK38756	s-adenosylmethionine synthetase	606	84	97	1,00E-97
R-P3C04	CBI22801	zinc finger and scan domain-containing protein	538	69	86	2,04E-59
R-P3C07	XP_003553752	benzyl alcohol o-benzoyltransferase-like	588	54	67	1,52E-22
R-P3C09	XP_002275105	heparan-alpha-glucosaminide n-acetyltransferase-like	605	100	86	8,47E-107
R-P3D10	XP_002517036	beta-ketoacyl-coa synthase family protein	557	39	90	3,04E-24
R-P3D11	CAA38252	glycine dehydrogenase	595	67	90	1,92E-75
R-P3D12	XP_002520117	protein kinase	610	75	89	2,55E-74
R-P3D01	CBI24120	abc transporter d family member chloroplastic-like	618	90	88	3,26E-93
R-P3D02	ACJ72158	udp-glycosyltransferase 1	601	99	66	2,08E-50
R-P3D03	XP_002314638	actin-related protein 9-like	592	99	85	7,71E-103
R-P3D04	XP_002306698	tubby-like f-box protein 8-like	647	83	93	4,35E-105

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

(continuação)						
Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	Similaridade (%)	e-value
R-P3D05	XP_002528785	uncharacterized mitochondrial protein ymf11-like	631	79	81	2,61E-62
R-P3D06	XP_002326656	peroxisomal acyl- oxidase 1 ^a	561	94	93	7,72E-101
R-P3D07	XP_002525198	thioredoxin-like protein	590	94	74	7,99E-72
R-P3D08	CCA66009	non-ltr retroelement reverse transcriptase	598	48	62	3,31E-14
R-P3E11	XP_002302817	npl4-like protein 2	592	100	92	1,58E-98
R-P3E01	XP_002509684	homeobox protein knotted-1-like protein	605	89	98	4,14E-117
R-P3E02	XP_002317691	wd40 repeat-containing protein smu1	677	67	97	1,30E-94
R-P3E04	XP_003529454	transcription factor pur-alpha	595	52	100	5,57E-60
R-P3E05	XP_003523479	probable peptide nitrate transporter	614	100	70	2,65E-52
R-P3E07	ACC59195	cellulose synthase catalytic subunit	230	100	84	3,41E-29
R-P3F10	XP_002304961	protein	592	75	76	1,60E-46
R-P3F11	XP_002325187	vacuolar h+-translocating inorganic pyrophosphatase	643	85	91	6,53E-96
R-P3F12	AEZ56251	lysophosphatidic acid acyltransferase	551	29	94	3,55E-06
R-P3F01	XP_002512841	ring finger protein 10-like	596	84	72	5,74E-35
R-P3F02	NP_001238324	proteasome subunit beta type-1	677	94	96	1,07E-127
R-P3F03	XP_002326971	protein	540	100	95	1,88E-116
R-P3F04	XP_003527586	upf0301 protein plut_0637-like	312	70	89	3,87E-29
R-P3F08	XP_002326230	threonine dehydratase biosynthetic	498	99	80	1,53E-66
R-P3G11	AFK34175	metal tolerance protein	610	46	95	1,25E-48
R-P3G12	XP_002317404	shaggy-related protein kinase eta	572	48	98	6,26E-49
R-P3G02	XP_002326588	acyl- -binding protein	530	50	96	4,82E-46
R-P3G05	XP_002310227	predicted protein [P. trichocarpa]	519	27	67	2,24E-03
R-P3G06	XP_002510751	conserved hypothetical protein [Ricinus communis]	582	94	76	2,23E-55
R-P3G07	XP_002302309	structural constituent of cell	622	55	76	5,73E-07
R-P3G08	XP_002526308	fructose-bisphosphate aldolase	537	91	89	3,94E-81
R-P3G09	XP_002324407	alcohol dehydrogenase	533	78	84	3,94E-47
R-P3H11	XP_002303363	transcription factor myb sant peroxisomal acyl- oxidase 1a	638	99	93	2,55E-117
R-P3H12	ACU19625	probable fructose-bisphosphate aldolase chloroplastic-like	369	36	84	9,82E-07
R-P3H03	XP_002533445	serine threonine-protein kinase-like protein acr4	607	100	83	1,96E-76
R-P3H07	ABK96635	galactokinase like protein	518	100	91	2,01E-93
R-P3H08	XP_002314721	bzip transcription factor bzip109	615	77	87	5,05E-76
R-P4A10	XP_002316136	ribonuclease 1	594	49	86	1,34E-40
R-P4A02	XP_002309250	receptor-like protein kinase	565	20	94	2,07E-07
R-P4B11	XP_002315914	uncharacterized amino-acid permease -like	634	44	89	8,66E-30
R-P4B12	XP_003607185	ubiquitin-conjugating enzyme e2 10	581	32	100	2,94E-29
R-P4B02	XP_002331482	pheophorbide a oxygenase	634	38	93	3,94E-53
R-P4B03	ACS88359	cellulose synthase	681	100	99	2,44E-149
R-P4B05	AEX09184	glycolate oxidase	666	48	97	2,00E-55
R-P4B07	XP_002512288	superoxide dismutase	601	76	91	4,06E-82
R-P4B08	ADO24300	xyloglucan endotransglycosylase hydrolase	631	100	91	3,11E-128
R-P4C11	ADL36680	f-box family protein	591	79	65	3,58E-26
R-P4C01	XP_002511077	chromophore lyase 3-like	614	59	95	3,92E-70
R-P4C04	XP_002515859	dna primase small subunit	650	83	92	3,08E-101
R-P4C07	XP_002513378	polyubiquitin 2	656	70	96	1,15E-81

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

(continuação)						
Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	Similaridade (%)	e-value
R-P4D10	XP_002509740	e3 ubiquitin-protein ligase	578	80	90	5,85E-60
R-P4D11	XP_002300801	photosystem ii reaction center w chloroplast	586	69	80	9,26E-44
R-P4D12	ADC53691	heat shock protein 90-2	581	100	99	9,25E-87
R-P4D01	XP_002263076	protein hira	608	93	73	1,34E-63
R-P4D02	XP_003550159	electron carrier electron transporter iron ion binding protein	592	56	81	1,07E-42
R-P4D05	XP_002314229	linoleate 13s-lipoxygenase 2-chloroplastic-like	612	95	87	6,71E-95
R-P4D06	XP_002266924	protein valine-tolerant 1	536	58	95	3,94E-49
R-P4D08	XP_002321579	ubiquitin-conjugating enzyme	616	65	97	6,66E-78
R-P4E11	XP_002521429	transmembrane protein 184c	673	86	97	1,00E-121
R-P4E01	BAJ61942	ubiquitin-like protein	564	68	99	1,36E-78
R-P4E02	ACJ85676	microtubule-associated protein 65-7	565	40	80	2,81E-22
R-P4E04	XP_002334565	s-ribonuclease binding protein 1	602	99	87	1,82E-81
R-P4E05	XP_002522772	ribosomal rna small subunit methyltransferase	631	99	85	1,33E-98
R-P4E06	ZP_01467486	class i	530	33	71	9,80E-06
R-P4E07	AEB77871	glutathione s-transferase	592	73	75	2,82E-51
R-P4E09	XP_002517085	40s ribosomal protein	503	100	95	7,26E-103
R-P4F12	XP_002311252	early response to dehydration 15	582	74	74	7,52E-40
R-P4F01	XP_002303660	pinus taeda anonymous locus	592	100	76	5,71E-66
R-P4F03	ABK95823	homeobox-leucine zipper protein hat5-like	607	90	91	1,86E-87
R-P4F06	XP_002327247	Protein	552	69	69	2,69E-31
R-P4F07	XP_002509995	glutamyl-trna reductase	610	78	75	5,54E-42
R-P4G10	XP_002308562	glycine dehydrogenase	516	99	91	7,32E-70
R-P4G01	XP_002306956	eukaryotic translation initiation factor 3 subunit c-like	572	61	86	4,75E-38
R-P4G06	ABX57881	chaperone protein	591	75	99	6,07E-63
R-P4G07	XP_002321410	dipeptidyl peptidase	633	97	92	6,75E-121
R-P4G08	XP_002316331	cdp-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase	644	88	91	3,23E-103
R-P4G09	XP_002312650	photosystem ii core complex proteins chloroplastic-like	657	80	90	4,56E-60
R-P4H10	CBI27775	nucleolysin tiar-like	560	99	92	1,10E-100
R-P4H02	NP_001236143	photosystem ii subunit q-2	649	65	91	5,51E-51
R-P4H07	XP_002517749	nodulin-like protein	645	100	70	5,64E-34
R-P4H08	ABW89104	glyceraldehyde-3-phosphate dehydrogenase	659	87	98	1,12E-89
R-Contig28	ABK96359	ribulose biphosphate carboxylase oxygenase activase chloroplast	589	100	91	1,58E-98
R-Contig94	XP_002511463	betaine-aldehyde dehydrogenase	252	90	89	4,79E-30
R-Contig95	AAF35185	lipid transfer protein	477	75	83	1,63E-42
R-Contig96	XP_002318863	dead-box atp-dependent rna helicase 7-like	560	100	94	1,01E-105
R-Contig97	XP_002514419	dnaj protein homolog	419	98	98	1,40E-39
R-Contig99	XP_002321926	predicted protein [P. trichocarpa]	532	48	78	9,91E-14
R-Contig100	XP_002521598	Protein	554	68	96	4,00E-70
R-Contig102	AAF35184	lipid transfer protein	604	60	80	2,23E-40
R-Contig103	ABY19575	Actin	540	83	99	4,78E-101
R-Contig105	XP_002517430	Protein	480	43	83	6,94E-16
R-Contig108	XP_002285246	cytokinin riboside 5 - monophosphate phosphoribohydrolase log1-like	665	71	96	6,08E-81
R-Contig109	XP_002306429	transcription factor	597	56	92	8,91E-40

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

(continuação)						
Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	Similaridade (%)	e-value
R-Contig111	XP_002515309	protein disulfide	689	34	93	3,40E-32
R-Contig113	XP_002511031	orange protein	618	98	95	5,58E-101
R-Contig115	XP_002517918	structural constituent of cell	624	56	84	9,70E-42
R-Contig117	AEV59651	chloroplast photosystem ii light harvesting complex protein	786	35	98	5,04E-50
R-Contig119	XP_002325437	cationic amino acid transporter	639	99	91	1,62E-92
R-Contig123	XP_002319805	fertilization-independent endosperm	653	88	97	6,00E-120
R-Contig125	XP_002326950	cysteine proteinase rd21a	812	89	84	1,52E-99
R-Contig126	ABK92581	actin	753	16	100	4,42E-13
R-Contig127	XP_002533878	pdi-like protein	589	72	89	9,31E-69
R-Contig128	XP_002529380	leucine aminopeptidase 1	746	80	72	2,40E-46
R-Contig129	XP_002513613	pyruvate kinase	553	88	78	4,41E-53
R-Contig130	XP_002283566	chlorophyll a b binding	751	84	99	6,45E-143
R-Contig131	ABD93208	probable fructose-bisphosphate aldolase chloroplastic-like	630	40	87	6,56E-34
R-Contig134	XP_002524029	root cap 1	745	96	72	6,78E-52
R-Contig135	BAH10641	mpbq msbq methyltransferase 2	647	93	92	2,29E-93
F+R-Contig1	AFZ78580	cellulose synthase-like protein h1	611	55	70	1,18E-24
F+R-Contig25	ADZ75466	oxygen-evolving enhancer protein chloroplast	745	93	92	1,76E-136
F+R-Contig29	CBI32274	c-type u1 snrnp	795	29	97	6,83E-29
F+R-Contig30	XP_002510217	sterol carrier protein 2-like	603	61	87	6,76E-51
F+R-Contig31	XP_002307516	protein gigantea	732	100	87	2,63E-117
F+R-Contig32	DAA33861	mip pip subfamily	736	91	94	1,29E-141
F+R-Contig36	CAA56286	ferredoxin-dependent glutamate synthase	634	99	95	1,45E-113
F+R-Contig39	XP_002512622	cbs domain-containing protein	795	48	94	7,46E-63
F+R-Contig42a	XP_002306220	class i chitinase	1079	80	76	2,87E-122
F+R-Contig42b	ABK96336	carbonic anhydrase	548	97	91	1,00E-103
F+R-Contig45	XP_002518477	atp synthase gamma chain	793	84	92	6,74E-113
F+R-Contig54	XP_002519996	60s ribosomal protein	623	88	92	1,06E-97
F+R-Contig60	XP_002299414	f chain improved model of plant photosystem i	456	68	73	1,71E-15
F+R-Contig63	XP_002514702	dna binding	586	91	59	5,26E-27
F+R-Contig73	Q41596	class i chitinase	859	47	61	9,65E-24
F+R-Contig80	CAN71829	hfr-2-like partial	595	71	62	2,45E-28
F+R-Contig92	DAA51874	actin	495	74	100	1,67E-80
F+R-Contig93	NP_200614	vamp-like protein ykt61	496	66	97	2,61E-65
F+R-Contig98	XP_002512790	glyceraldehyde-3-phosphate dehydrogenase	510	68	87	1,38E-44
F+R-Contig101	XP_002518072	elongation factor 1- partial	597	98	98	9,53E-130
F+R-Contig106	XP_002305603	thiamine biosynthetic enzyme	541	76	95	2,34E-58
F+R-Contig107	XP_002304601	molybdopterin cofactor	845	38	95	2,63E-51
F+R-Contig110	XP_002318639	nucleobase-ascorbate transporter 6	612	84	98	9,13E-95
F+R-Contig112	XP_002521232	ribulose- -bisphosphate carboxylase oxygenase small subunit	828	57	91	1,06E-93
F+R-Contig114	XP_002318093	rna-binding protein cp31	719	84	63	2,94E-29
F+R-Contig116	XP_002315054	predicted protein	567	68	64	1,04E-16
F+R-Contig118	CAN81811	class i chitinase	1049	33	67	4,38E-24
F+R-Contig120	BAA82843	miraculin homologue	789	48	53	1,69E-02
F+R-Contig121	AAQ96377	miraculin-like protein	1089	48	51	2,07E-04
F+R-Contig122	XP_002330074	at1g56580 f25p12_18	552	67	94	6,46E-60
F+R-Contig132	XP_002327977	glycoprotease m22 family protein	674	66	99	3,72E-93

Tabela A1 – Resultados da busca por similaridade no banco de dados proteico não redundante 'NR' e seus respectivos *top-hits*, feita através do programa computacional Blast2GO

Clone	Nº de acesso	Proteína	tamanho (pb)	Cobertura (%)	(conclusão)	
					Simila- ridade (%)	<i>e-value</i>
F+R-Contig133	XP_002513654	beta-carotene hydroxylase	592	69	76	9,02E-42
F+R-Contig136	ABR26094	conserved protein	1090	31	89	6,32E-54
F+R-Contig137	XP_003637074	cell wall-associated hydrolase	657	93	77	3,07E-88
F+R-Contig141	BAH11228	atp synthase cf0 subunit iii	1021	26	97	4,01E-26