

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

**Perfis de expressão gênica temporal de cana-de-açúcar infectada por
ferrugem alaranjada**

Fernando Henrique Correr

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

**Piracicaba
2017**

Fernando Henrique Correr
Bacharel em Biotecnologia

**Perfis de expressão gênica temporal de cana-de-açúcar infectada por
ferrugem alaranjada**

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientador:

Prof. Dr. **GABRIEL RODRIGUES ALVES MARGARIDO**

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

Piracicaba
2017

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

Correr, Fernando Henrique

Perfis de expressão gênica temporal de cana-de-açúcar infectada por ferrugem alaranjada / Fernando Henrique Correr. -- versão revisada de acordo com a resolução CoPGr 6018 de 2011. -- Piracicaba, 2017 .

80 p.

Dissertação (Mestrado) -- USP / Escola Superior de Agricultura "Luiz de Queiroz".

1. RNA-Seq 2. Transcriptoma 3. *Saccharum* 4. *Puccinia kuehni* . I.
Título.

DEDICATÓRIA

*Aos meus pais, **Claudia e Valdecir.**
Por sempre me apoiarem e me fazerem refletir sobre minhas decisões.
Obrigado mãe, pelo seu exemplo de perseverança, força e fé.
Obrigado pai, por sempre ter a palavra certa, no momento certo.
Ao me ensinarem a importância do conhecimento,
você trilharam meus passos até aqui.*

Dedico

AGRADECIMENTOS

A Deus, Criador e Senhor de todas as coisas. Agradeço pela conclusão da dissertação e, muito mais, pelas pessoas incríveis aqui mencionadas, cuja importância estas páginas não são suficientes para descrever.

À minha família, de modo especial aos meus avós, Maria Cecília e Orlando, que contínua e ativamente me incentivam.

À minha namorada, Maria Clara Panciera, pelo seu amor, carinho e compreensão. Obrigado por me fazer sorrir em todos os desafios durante esses dois anos de mestrado. Obrigado pelas incríveis perguntas sobre a dissertação. Obrigado, por me ensinar a ter alegria em tudo que estou fazendo, seguindo cada passo de uma vez. Obrigado!

Ao Prof. Dr. Gabriel Rodrigues Alves Margarido por me acolher na ESALQ durante o estágio obrigatório e por ter me aceito como seu aluno de mestrado, foi uma honra. Sua amizade, bondade, confiança e paciência foram essenciais para a conclusão desta dissertação. Com você, pude aprender valores que nenhum livro, manuais de uso ou fóruns de dúvidas possuem.

Aos colegas do Laboratório de Bioinformática aplicada à Bioenergia: Amanda Oliveira, pela ajuda com modelos de relatórios, nas dúvidas de melhoramento genético e dores de cabeça no R; Leonardo Menegatto, pela companhia nos dois Congressos de Genética e conversas políticas durante a estrada; Lorena Batista, por sempre nos atualizar sobre a Pós-Graduação; Guilherme Hosaka, pelo grande auxílio na análise dos dados, pelas conversas sobre tecnologia e fisiologia vegetal; Francisca Araújo, por compartilhar seus conhecimentos dia a dia; Felipe Corrêa, pela parceria no aprendizado dos primeiros *heatmaps* com códigos não muito eficientes; Luan Clemente, por trazer novidades cotidianas e tomar conta do café; Tiago Corrêa, pela companhia durante as primeiras reuniões de Bioinformática; Guilherme Ambrosano, por também estar ajudando no processo de montagem do transcriptoma.

À Prof^a Dr^a Monalisa Sampaio Carneiro pela orientação no Laboratório de Biotecnologia de Plantas, por ter depositado tanta confiança para o desenvolvimento das bibliotecas de RNA e ter me incentivado a desenvolver este trabalho. Aos colegas Thiago, Fernanda, Catarina e Mariana, pela grande ajuda nas etapas de extração do RNA.

Ao Prof. Dr. Nelson Sidnei Massola Júnior e ao mestre Sergio Gregorio Pérez Gómez, pelo fornecimento das amostras e respostas aos questionamentos do experimento de infecção.

Ao Prof. Dr. Marco Aurelio Takita, por ter fornecido a estrutura de seu laboratório e colaborado com sua experiência nas etapas de quantificação de RNA.

Ao Prof. Dr. Luiz Lehmann Coutinho, grande auxiliador no planejamento das bibliotecas de sequenciamento.

À Prof. Dr^a Claudia Barros Monteiro Vitorello e ao Prof. Dr. Ricardo Antunes de Azevedo pelas valiosas contribuições nos resultados do enriquecimento funcional.

Ao Prof. Dr. Luis Eduardo Aranha Camargo e à Dr^a Mariana Cicarelli Cia que muito têm colaborado com seu conhecimento molecular para a realização das etapas finais do trabalho.

Ao Prof. Dr. Antonio Augusto Franco Garcia, por receptivamente me receber em seu laboratório, pelas valiosas lições sobre genética e por, juntamente ao Carlos, elaborar o template do \LaTeX .

Aos colegas do Laboratório de Genética Estatística: João, Carina, Guilherme, Adriana, Luis Felipe, Danilo, Letícia, Marianella, Rodrigo, Jhonathan, Matheus. Agradeço pela convivência, pelas aulas “particulares” em momentos tão cruciais e por sempre disponibilizarem de um pouco de tempo para me ajudarem com o R e \LaTeX . Também sou grato à Cristiane Taniguti, por todo o período de mestrado, onde compartilhamos disciplinas, trabalhos e preocupações.

Aos demais colegas do PPGGMP, professores e funcionários do LGN. Agradeço pela ajuda recebida dos funcionários da biblioteca e da pós-graduação.

À Universidade de São Paulo e à Escola Superior de Agricultura “Luiz de Queiroz” pela estrutura e serviços fornecidos. Principalmente pela oportuna formação através do Programa de Pós-Graduação.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela bolsa de estudo PROEX, e à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo apoio financeiro concedido através do processo nº 2015/17935-8.

Muito obrigado.

EPÍGRAFE

Nothing is impossible,
the word itself says
'I'm possible'!
Audrey Hepburn

O que vale na vida
não é o ponto de partida
e sim a caminhada.
Caminhando e semeando,
no fim terás o que colher.
Cora Coralina

SUMÁRIO

Resumo	8
Abstract	9
1 Introdução	11
1.1 A cana-de-açúcar no cenário atual	12
1.2 Importância da ferrugem alaranjada	12
1.3 Bases moleculares da infecção fúngica e resposta do hospedeiro	14
1.4 Melhoramento genético e recursos genômicos da cana-de-açúcar	15
1.5 O estudo do transcriptoma através do sequenciamento de RNA	17
1.6 Estudos de expressão diferencial e perfis de expressão gênica	18
1.7 Objetivos	19
2 Material e Métodos	21
2.1 Material vegetal	21
2.2 Seleção de amostras e planejamento experimental	21
2.3 Extração de RNA e sequenciamento	21
2.4 Pré-processamento das leituras	22
2.4.1 Diagnóstico das leituras pareadas	22
2.4.2 Controle da qualidade e remoção de rRNA	22
2.5 Alinhamento das leituras e construção da matriz de contagem	23
2.6 Análise de expressão diferencial	23
2.7 Análise de enriquecimento funcional	24
2.8 Perfis de expressão gênica	24
2.9 Mapas das vias moleculares	25
3 Resultados e Discussão	27
3.1 Controle de qualidade e pré-processamento das leituras	27
3.2 Expressão diferencial: Teste conjunto entre todos os pontos no tempo	28
3.3 Análise de expressão diferencial: o avançar da infecção pela comparação de momentos específicos	31
3.4 Perfis de expressão gênica temporal	34
3.5 Perfis de expressão gênica sob o ponto de vista da infecção de cana-de-açúcar por <i>Puccinia kuehni</i>	40
4 Conclusão	43
Referências	45
Apêndices	55

RESUMO

Perfis de expressão gênica temporal de cana-de-açúcar infectada por ferrugem alaranjada

A ferrugem alaranjada é uma doença foliar que afeta a cana-de-açúcar, tendo ganho destaque como uma das principais doenças da cultura. Causada pelo fitopatógeno biotrófico *Puccinia kuehni*, foi relatada em diversos países da América do Sul, incluindo o Brasil. Entender o comportamento desta doença é importante na compreensão de seus mecanismos de infecção. Para isso, o propósito deste trabalho é avaliar o transcriptoma, por sequenciamento de RNA, da resposta da cana-de-açúcar à infecção por *P. kuehni* durante o estabelecimento da doença. O material biológico utilizado correspondeu a folhas de cana-de-açúcar do cultivar suscetível SP89-1115, amostrados no momento da inoculação dos esporos (0 h), 12 horas após a inoculação (hai), 24 hai, 48 hai, 5 dias e 12 dias após a inoculação da ferrugem alaranjada. Após o pré-processamento das leituras, foi realizado o mapeamento de aproximadamente 800 milhões de leituras, utilizando como referência um transcriptoma de cana-de-açúcar obtido a partir do sequenciamento de seis genótipos distintos. Posteriormente, foi realizada a análise de expressão diferencial: *i*) em um teste do tipo Análise de Variância, *ii*) entre a comparação de cada tempo após a infecção contra 0 h e *iii*) pela comparação de tempos adjacentes. Por essa última estratégia, a maioria dos transcritos diferencialmente expressos ocorreu em 12 hai, 48 hai e 12 dai. Através do enriquecimento funcional foram evidenciados processos relacionados principalmente à fotossíntese e oxirredução, importantes na manutenção do metabolismo e sinalização celular sob perturbação. Os perfis de expressão gênica temporal dos transcritos anotados de acordo com termos da base *Gene Ontology*, por sua vez, mostraram ondas invertidas de repressão e estímulo entre as comparações 12 hai vs 0 h e 48 hai vs 24 hai. Através da segunda análise de expressão diferencial houve indícios de que os níveis de expressão dos transcritos eram similares entre 0 h e 48 hai. Portanto, *P. kuehni* pode ter agido no sentido de repressão das vias de resposta de defesa nas primeiras horas seguidas da inoculação, desfavorecendo mecanismos de acúmulo de fitormônios e deposição de lignina na parede celular. O cultivar de cana-de-açúcar pode ter sido prejudicado pela ausência de sistemas de reconhecimento e/ou no combate às proteínas efetoras do patógeno, de modo a ter sido modulado para o estabelecimento fúngico, havendo evidências do restabelecimento parcial da homeostase celular em 48 hai.

Palavras-chave: RNA-Seq, Transcriptoma, *Saccharum*, *Puccinia kuehni*

ABSTRACT

Temporal gene expression profiles of sugarcane infected by orange rust

Orange rust is a foliar disease that affects sugarcane, gaining importance among the main diseases of this crop. Caused by the biotrophic phytopathogen *Puccinia kuehnii*, it has been reported in several countries in South America, including Brazil. Understanding the behavior of this disease is very important in comprehending its mechanisms of infection. The purpose of this work is to evaluate the transcriptome, by RNA sequencing, of the sugarcane response to infection by *P. kuehnii* during the establishment of the disease. The biological material used corresponds to sugarcane leaves from susceptible cultivar SP89-1115, sampled at the time of spore inoculation (0 h), 12 hours post inoculation (hpi), 24 hpi, 48 hpi, 5 days and 12 days post inoculation of orange rust. After pre-processing the reads, we mapped approximately 800 million reads to a reference sugarcane transcriptome obtained from sequencing six different genotypes. Subsequently, we performed differential expression analysis: *i*) in an Analysis of Variance type test; *ii*) comparing each time after inoculation against 0 h and *iii*) by comparing adjacent times. By the latter strategy, most of the differentially expressed transcripts occurred at 12 hpi, 48 hpi and 12 dpi. Through functional enrichment analysis, processes related mainly to photosynthesis and oxidation were evidenced, which are important in metabolic maintenance and cellular signaling under perturbation. The temporal gene expression profiles of *Gene Ontology*-annotated transcripts, in turn, showed reverse repression and stimulus waves between the comparisons 12 hpi vs 0 h and 48 hpi vs 24 hpi. Through the second differential expression analysis we found evidence that the expression levels of the transcripts were similar between 0 h and 48 hpi. Therefore, *P. kuehnii* may have acted in the sense of repressing defense response pathways in the first hours following inoculation, disfavoring mechanisms of phytohormones signaling and lignin deposition in the cell wall. This sugarcane cultivar may have been harmed by the absence of recognition systems and/or in the control of pathogen effector proteins in order to have been modulated for fungal establishment, with evidence of partial reestablishment of cellular homeostasis at 48 hpi.

Keywords: RNA-Seq, Transcriptome, *Saccharum*, *Puccinia kuehnii*

1 INTRODUÇÃO

A cana-de-açúcar é uma das culturas agrícolas mais importantes no Brasil, movendo um setor de grande impacto agroindustrial, com destaque para o Estado de São Paulo, onde ocupa aproximadamente 6 milhões de hectares (UNICA, 2017). Historicamente, vem sendo utilizada principalmente para a produção de açúcar e, há quase cinco décadas, faz parte da cadeia produtiva do etanol, importante biocombustível no cenário nacional e internacional, devido à grande demanda por combustíveis renováveis (CHEAVEGATTI-GIANOTTO *et al.*, 2011; MACRELLI *et al.*, 2012). Cabe ressaltar que, além de utilizado como combustível na forma de álcool hidratado, o álcool na forma anídrica é empregado como aditivo para a gasolina (DANTAS *et al.*, 2013).

Historicamente, a fase de nobilitação do melhoramento genético da cana-de-açúcar foi direcionada principalmente ao acúmulo de sacarose nos colmos. Porém, notou-se a necessidade de serem incorporadas outras características, tal como capacidade de perfilhamento e melhor adaptação a diferentes ambientes. Posteriormente, métodos baseados na seleção recorrente, com o uso dos genitores considerados mais favoráveis, foram utilizados no melhoramento genético visando principalmente a produção e acúmulo de açúcar. Uma outra abordagem, porém, é a formação de várias populações pelo cruzamento de parentais contrastantes, permitindo selecionar os indivíduos superiores, em várias fases de avaliação e levando em conta os efeitos ambientais. Nos dois casos, além do conteúdo de sacarose e produção, a seleção permite a superação de limitações abióticas e resistência a doenças (CHEAVEGATTI-GIANOTTO *et al.*, 2011; JACKSON, 2005). Em relação ao último caráter, JOHNSON (1992) aponta que cada patógeno apresenta seus próprios obstáculos, assim como o hospedeiro é variável em sua resposta. Não obstante, em geral, é necessário verificar fontes de resistência, compreendendo a interação gene-a-gene entre os dois organismos. Essa relação é dinâmica, já que novos genes de virulência podem ser estabelecidos.

Entre as diversas formas de estresse biótico existentes para o cultivo da cana-de-açúcar, está a ferrugem alaranjada, causada por *Puccinia kuehni* (W. Krüger) E.J. Butler. É um patógeno morfológicamente e molecularmente distinto da ferrugem marrom (*P. melanocephala*), formando grupos arredondados de lesões foliares com distribuição irregular, a partir dos quais há dispersão de esporos de cor laranja, que dão o nome à doença. A ocorrência da primeira epidemia de *P. kuehni* foi no início do século XXI e afetou drasticamente os canaviais da Austrália onde se empregava a cultivar Q124, com perdas por volta de 45% (BRAITHWAITE *et al.*, 2009; MAGAREY *et al.*, 2001; VIRTUDAZO *et al.*, 2001). BARBASSO *et al.* (2010) relataram e confirmaram a primeira aparição da doença em canaviais brasileiros, a partir de observações durante o ano de 2009 sobre um cultivar resistente à *P. melanocephala*. Outras três cultivares responsáveis na época por cerca de 10% das áreas com plantio de cana-de-açúcar demonstraram-se altamente suscetíveis, entre elas a SP89-1115. Observando as bases bibliográficas, relatos confirmam a ferrugem alaranjada da cana-de-açúcar na Colômbia (CADAVID *et al.*, 2012), Louisiana (EUA) (GRISHAM *et al.*, 2013), Equador (GARCÉS *et al.*, 2014), República Dominicana (BRIGGS *et al.*, 2014) e, recentemente, na Argentina (FUNES *et al.*, 2016).

Os efeitos recíprocos entre as proteínas do patógeno e a planta permitem compreender as bases moleculares da suscetibilidade ou resistência à doença, corroborando os estudos epidemiológicos. Essas biomoléculas celulares podem ser compreendidas através de seus transcritos, elementos funcionais do genoma. Para isso, é feita a geração massiva de dados, por meio do sequenciamento das moléculas de RNA mensageiro. A partir daí, pode-se verificar as alterações nos níveis de expressão dos transcritos dada a ocorrência de uma perturbação. Essa estratégia tem importante aplicação no estudo da resposta molecular das plantas à infecção, pois a influência de estresses representa uma fonte de distúrbio que altera os níveis de expressão gênica. Visto que a célula é um sistema dinâmico, considera-se também o tempo como um fator preponderante na variação dos níveis de expressão gênica, verificando-se muitos transcritos de expressão não estacionária. Como os transcritos estão relacionados às proteínas que serão traduzidas

a partir de suas sequências, pode-se verificar através deles o comportamento de vias metabólicas. Com isso visa-se compreender os processos regulatórios existentes em um organismo ao longo de um período estabelecido de ação de um agente estressor (BAR-JOSEPH *et al.*, 2012; MARGUERAT & BÄHLER, 2010; VOGEL & MARCOTTE, 2012; WANG *et al.*, 2009).

Neste contexto, a investigação do RNA total coletado de folhas de cana-de-açúcar infectadas com *P. kuehni* é de grande importância. Em particular, a análise quantitativa dos transcritos ao longo do tempo, durante o período de infecção, pode fornecer evidências a respeito de mudanças nos genes expressos e, conseqüentemente, de sua influência na suscetibilidade do cultivar SP89-1115.

1.1 A cana-de-açúcar no cenário atual

No Brasil, de acordo com o quarto levantamento feito pela Companhia Nacional de Abastecimento (CONAB, 2016a), para a safra de 2015/16, o total de cana-de-açúcar colhida para processamento foi de 665,6 milhões de toneladas, dos quais cerca de 59,58% foram direcionados à produção de 30,5 bilhões de litros de álcool, sendo que os 40,42% restantes destinaram-se à produção de 33,5 milhões de toneladas de açúcar. De acordo com as atuais estimativas do mesmo órgão publicadas em Agosto de 2016, para a safra de 2016/17 espera-se uma elevação em torno de 19,3% na produção total de açúcar, com produção de 39,96 milhões de toneladas. Em contrapartida, espera-se uma redução de aproximadamente 9,7% na produção do etanol (CONAB, 2016b).

Tanto o açúcar quanto o álcool, os dois principais produtos comerciais, provêm do caldo da cana. Não obstante, o potencial da cultura se estende em função de mudanças no plantio e nas tecnologias de produção, criando-se produtos complementares como, por exemplo, a utilização do bagaço na co-geração de energia elétrica (bioeletricidade) e, também, na produção de etanol de segunda geração. Desenvolvem-se, também, pesquisas com bioplásticos e de aproveitamento dos demais resíduos, como palha e vinhaça. Trata-se de uma visão de sustentabilidade, na qual o setor busca seu desenvolvimento a partir do seu crescimento econômico com melhorias de práticas, inovação dos produtos e criação de novos programas socioambientais (DANTAS *et al.*, 2013; MACRELLI *et al.*, 2012; VIANA & PEREZ, 2013).

Dessa forma, para suprir toda a necessidade de matéria-prima para atender ao mercado dos produtos e subprodutos acima mencionados, é necessário aumentar a produtividade sem que se aumente a área de cultivo. Assim, buscam-se continuamente melhorias das técnicas de plantio e colheita, bem como do próprio potencial genético dos cultivares. Com isso, para a maior eficiência no uso da área cultivada, os programas de melhoramento genético de cana-de-açúcar devem atentar-se, além de elevada produção, à redução dos recursos naturais utilizados pela planta (SCORTECCI *et al.*, 2012).

1.2 Importância da ferrugem alaranjada

Em relação a formas bióticas de estresse responsáveis por problemas no cultivo, a cultura é afetada por pragas agrícolas e doenças de grande impacto, com destaque para três doenças fúngicas: o carvão, causado por *Sporisorium scitamineum*, a ferrugem marrom, ocasionada por *P. melanocephala*, e a ferrugem alaranjada, causada por *P. kuehni* (MOREIRA, 2013). A primeira doença atua principalmente sobre o meristema apical, enquanto as duas ferrugens se estabelecem nas folhas. Resultados indicam que a perda em produtividade causada pela *P. melanocephala*, agente causador da ferrugem marrom, é da ordem de 47% em cultivares suscetíveis. Para *S. scitamineum*, responsável pelo carvão, demonstrou-se uma perda de produtividade média de 0,78% para cada 300 touceiras infectadas por hectare (BURNQUIST & VEGA, 1995).

Morfológicamente, os urediniosporos pertencentes a *P. kuehni* geralmente são maiores do que os de *P. melanocephala* e podem ser diferenciados por possuírem um engrossamento na região apical,

apresentarem a superfície recoberta de espinhos de grande tamanho, ausência de paráfise e coloração direcionada ao laranja. Todavia, são similares pela possibilidade de os esporos serem transportados pelo vento, sendo que no caso da ferrugem marrom essas estruturas são dispersas a longas distâncias (VIRTUDAZO *et al.*, 2001; DIXON *et al.*, 2010; PURDY *et al.*, 1985; SOOD *et al.*, 2013).

Embora pertencentes ao mesmo gênero, *P. kuehni* e *P. melanocephala* mantiveram-se em clados distintos quando avaliadas as sequências de rDNA como, por exemplo, o espaçador transcrito interno 2 (ITS2 - *Internal transcribed spacer 2*) e a região 5' do DNA correspondente à subunidade maior do ribossomo. *P. melanocephala*, por exemplo, manteve-se mais próxima a *P. sorghi*, *P. graminis* e *P. striiformis*, agentes causadores da ferrugem do milho e das duas ferrugens do trigo, respectivamente (DIXON *et al.*, 2010). Além disso, verificando a diversidade de isolados de *P. kuehni*, BRAITHWAITE *et al.* (2009) obtiveram três clados, nos quais não foi possível explicar o posicionamento dos isolados de canas selvagens e de jardins. Porém, os isolados provenientes de canaviais comerciais permaneceram agrupados em um mesmo clado. Uma informação relevante foi que análises de microscopia revelaram características de *P. kuehni* para todos os isolados, dando suporte aos resultados moleculares.

Como uma forma de caracterizar visualmente o impacto sobre o hospedeiro, é notável que as ferrugens, como patógenos foliares, têm como principais sintomas a clorose e a necrose. SOOD *et al.* (2009) criaram uma escala para mensurar a infecção das ferrugens nas folhas de cana-de-açúcar, em que 0 significa ausência de sintomas, 1 a presença de manchas cloróticas, 2 quando há presença de lesões (na coloração da ferrugem), 3 quando há de uma a cinco pústulas com esporos e 4 quando o número de pústulas for superior a cinco. Esses autores também utilizaram essa escala de sintomas para caracterizar os cultivares quanto à resistência às ferrugens da cana-de-açúcar: um cultivar com valor de 0 a 1 é considerado resistente; na classe 2 é considerado moderadamente resistente; e nas classes 3 e 4 o cultivar é dito suscetível. No caso particular de *P. kuehni*, observou-se redução significativa e linear de medidas como a condutância estomática e a taxa de fotossíntese, nas porções infectadas de folhas classificadas nos níveis 2, 3 e 4 de infecção (ZHAO *et al.*, 2011).

Essa escala também foi empregada em um estudo realizado no programa de desenvolvimento de cultivares de cana-de-açúcar de Canal Point, na Flórida, no qual foi testada a metodologia de inoculação de SOOD *et al.* (2009), a fim de tornar mais eficiente a seleção de clones resistentes (SOOD *et al.*, 2013). CHAPOLA *et al.* (2016), por sua vez, consideraram que além dessa técnica, a inoculação pela pulverização manual é interessante nas etapas finais do melhoramento, pois as plantas são mantidas em ambiente controlado e, como vantagem, podem ser mensurados parâmetros como incubação, latência e número de pústulas por área foliar.

A ação do patógeno sobre os cultivares também depende de fatores ambientais, como a temperatura. Comparando-se o aumento de 18°C para 25°C, o cultivar moderadamente resistente RB855156 teve redução de 10 dias tanto no período de incubação quanto de latência, enquanto que no cultivar suscetível SP89-1115 as reduções nos dois períodos foram de 7 dias. Os autores notaram, também, que RB855156 poderia apresentar resistência horizontal, dado que a frequência da infecção e a área das pústulas foram maiores na temperatura mais elevada. Nesta temperatura, um fato importante foi que os sintomas aparentes e a esporulação em SP89-1115 ocorreram, respectivamente, dois e três dias antes do moderadamente resistente (GOMEZ, 2013).

Comumente em plantas resistentes, a morte de células do tecido vegetal nas proximidades da infecção leva ao acúmulo de fitoalexinas tóxicas ao patógeno e estimula a defesa nas células circundantes, o que representa um mecanismo de defesa localizado de uma interação planta-patógeno incompatível. Isto é resultado de uma intensa e específica relação de reconhecimento de genes de avirulência do patógeno por genes de resistência do hospedeiro, conhecida como resistência sistêmica adquirida. Esta, por sua vez, varia com o estágio desenvolvimental do hospedeiro, com constante reestabelecimento das relações entre raças do patógeno e os cultivares, em um escopo evolutivo (DANGL & JONES, 2001; KEEN, 1990;

WHALEN, 2005).

Para que a interação entre os produtos gênicos ocorra durante o processo de estabelecimento de uma doença sobre o hospedeiro, é necessário um período de estabelecimento do patógeno sobre o tecido. Diversas vias de sinalização são importantes, tanto para a ação dos sistemas de defesa do hospedeiro, quando este é resistente, quanto para promover a completa disseminação do patógeno, no caso da suscetibilidade (GREENBERG & YAO, 2004). Desse modo, o estudo do transcriptoma é fundamental para a compreensão desse estado fisiológico, pois permite a avaliação dos genes expressos pelos dois organismos, bem como de suas interações, a partir da avaliação dos elementos funcionais do genoma (WANG *et al.*, 2009). O grande impacto dos estresses bióticos sobre os cultivares suscetíveis (SCORTECCI *et al.*, 2012) justifica o entendimento de como essa relação se estabelece.

1.3 Bases moleculares da infecção fúngica e resposta do hospedeiro

KNOGGE (1996) aponta, de forma geral, a existência de três tipos de gravidade de infecção sobre o vegetal, correlacionados ao espectro de atuação dos fungos fitopatogênicos: *i*) Os de baixa virulência, quando há capacidade de se estabelecer em vários hospedeiros; *ii*) Os de alta virulência sobre um pequeno grupo de hospedeiros, mas que não são parasitas obrigatórios; *iii*) E os virulentos obrigatoriamente parasitas, cuja sobrevivência se dá com total dependência do organismo hospedeiro. Neste último pode se enquadrar *P. kuehni*, com comportamento biotrófico característico das ferrugens (BRAITHWAITE *et al.*, 2009; DUPLESSIS *et al.*, 2011; VIRTUDAZO *et al.*, 2001). KNOGGE (1996) considera ainda que os dois organismos envolvidos, patógeno e planta, realizam uma troca de sinais que reflete no desencadeamento dessa relação, ou seja, na resistência ou suscetibilidade vegetal, virulência ou avirulência do fungo.

Dado que o principal objetivo de um fungo fitopatogênico é conseguir nutrientes através da introdução de haustórios nas células hospedeiras, JONES & DANGL (2006) relatam dois principais sistemas imunes para que esse processo não seja bem-sucedido. O primeiro sistema imune ocorre pelo reconhecimento de padrões moleculares associados a patógenos (*pathogen-associated molecular patterns* - PAMPs) por receptores de reconhecimento transmembranares (PRR), que é conhecido por *plant-triggered immunity* (PTI). Porém, esse sistema não desencadeia uma indução da resposta hipersensitiva tão eficiente (Figura 1.1), podendo, dessa maneira, ser superado pela secreção de proteínas efetoras, causando a suscetibilidade desencadeada por efetores (*effector-triggered susceptibility* - ETS).

Caso a planta possua uma proteína de um gene de resistência, esta atuará sobre um efector produzido pelo fungo, desde que haja correspondência entre eles. Em vista disso, é determinado o segundo sistema de defesa, conhecido como imunidade desencadeada por efector (*effector-triggered immunity* - ETI), motivo pelo qual o gene do efector é conhecido como de avirulência. Porém, por se tratar de um sistema complexo, um novo efector pode induzir novamente a ETS e um novo gene hospedeiro pode agir para que se estabeleça a ETI (Figura 1.1). Portanto, trata-se de uma cascata de reconhecimento das moléculas secretadas pelo patógeno e genes de resistência da planta, um modelo *zig-zag*. Esta interação resultará em uma planta resistente quando a ação de ETS for inferior à PTI somada a uma eficaz ETI (JONES & DANGL, 2006).

O interesse em compreender as bases moleculares da infecção fúngica por meio do acúmulo de transcritos vem de trabalhos como o de LAWTON & LAMB (1987), no qual o mRNA produzido a partir de genes em condições normais era comparado ao uso de um tratamento com elicitor fúngico. De maneira similar, SCHMELZER *et al.* (1989) mantiveram a atenção sobre poucos genes específicos relacionados às vias de defesa, como a via da fitoalexina, na qual um dos compostos essenciais para sua formação é a fenilalanina amônio liase (*Phenylalanine ammonia-lyase* - PAL).

As proteínas codificadas pelos genes de resistência são tipificadas por dois domínios principais:

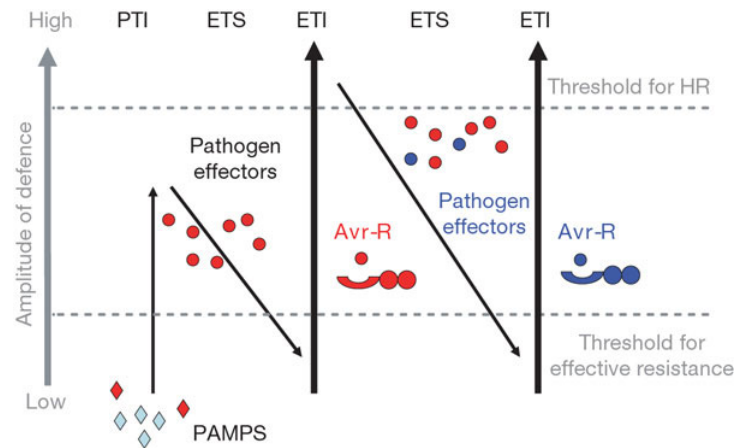


Figura 1.1. Interações moleculares no processo de interação planta-patógeno. O reconhecimento inicial dos PAMPs (*pathogen-associated molecular patterns*) conduz à PTI (*plant-triggered immunity*) e, caso efetores sejam produzidos, haverá ETI (*effector-triggered immunity*) se houver genes de resistência. A relação dos efetores fúngicos com os genes de resistência do hospedeiro é de um para um. Caso não haja uma correspondência por parte de um gene de resistência, o efector será chamado de virulento. Tais padrões constituem o chamado modelo *zig-zag*. Fonte: JONES & DANGL (2006).

nucleotide binding (NB) e *leucine-rich repeat* (LRR) (JONES & DANGL, 2006), além de um domínio variável na extremidade amino-terminal (TAKKEN *et al.*, 2006). A ação de um efector de um gene de avirulência (*Avr*) ocorre pela ação em proteínas substrato, como é o caso da clivagem da proteína PBS1 pelo *AvrPphB* em *Arabidopsis thaliana*. Isso permite a ativação indireta, por meio de sinalização celular, de uma proteína NBS-LRR. Além disso, as proteínas NBS-LRR são reguladas por outras proteínas, como é o caso de RIN4 com ação *trans* sobre RPM1 e RPS2 (ambas NBS-LRR), bem como podem ser estabilizadas por ação *trans* de HSP90 (BELKHADIR *et al.*, 2004; TAKKEN *et al.*, 2006). As interações formadas por uma proteína substrato podem ser: *i*) de regulação negativa quando essa proteína é ligada a uma proteína de resistência, a qual será liberada se um efector fizer a clivagem do substrato; *ii*) de ativação de um membro NBS-LRR, quando um *Avr* causar a fosforilação do substrato (TAKKEN *et al.*, 2006).

A sinalização é deveras relevante, com os fitormônios sendo os principais atuantes nesse processo. Eles são importantes não só no crescimento, mas o balanço de certas categorias de hormônios, bem como a interação entre elas, tem ação direta no comportamento do sistema de defesa das plantas (HUOT *et al.*, 2014; ROBERT-SEILANIANTZ *et al.*, 2011). Ademais, as vias de sinalização têm etapas conduzidas por transdutores de sinais dos estímulos de reconhecimento do patógeno, como as cascatas de cinase do tipo *mitogen-activated* (MAP). Os fatores de transcrição WRKY, por outro lado, atuam no sentido de estímulo ou repressão da transcrição dos genes de defesa (JONES & TAKEMOTO, 2004). Há um último mecanismo de sinalização, mais generalista entre diferentes formas de estresse, que são as chamadas espécies reativas de oxigênio (ROS), cujo acúmulo e propagação podem ser indicativos de perturbação nas células. Ademais, essa classe de moléculas possui diferentes alvos e relaciona-se com outras substâncias que promovem sua detoxificação e agem sobre outros alvos, aumentando a gama de processos relacionados com a defesa (LEHMANN *et al.*, 2015).

1.4 Melhoramento genético e recursos genômicos da cana-de-açúcar

O nome cana-de-açúcar representa espécies pertencentes a um grupo de gramíneas tropicais para o qual, no intuito de reunir caracteres de acúmulo de sacarose e robustez de espécies distintas, delinearam-se os programas de melhoramento genético para a obtenção de cultivares. São membros da

família Poaceae, tribo Andropogoneae, subtribo Saccharinae, gênero *Saccharum*. Cultivadas pelo acúmulo de sacarose, das quais são reconhecidas quatro espécies: *Saccharum officinarum*, *S. barberi*, *S. edule* e *S. sinense*; e outras duas tidas como selvagens: *S. spontaneum* e *S. robustum*, de ampla distribuição. Em complemento, notam-se outras espécies proximamente relacionadas, o que dá origem ao chamado complexo Saccharum, formado com os gêneros de gramíneas *Erianthus*, *Miscanthus*, *Sclerostachya* e *Narenga* (MOORE *et al.*, 2013; DANIELS & ROACH, 1987).

No tocante à organização genômica da cana-de-açúcar, a hibridação interespecífica entre as espécies *S. officinarum* e *S. spontaneum*, com menor contribuição das espécies *S. sinensi*, *S. robustum* e *S. barberi*, permitiu a obtenção dos cultivares modernos. Devido a recorrentes retrocruzamentos com *S. officinarum*, tais cultivares são caracterizadas geneticamente por altos níveis de ploidia, com frequente aneuploidia, encontrando-se usualmente de 10 a 13 conjuntos de cromossomos, com cerca de oito homólogos para cada cromossomo oriundo de *S. officinarum*, até 23% dos cromossomos provenientes de *S. spontaneum* e uma quantidade menor, até cerca de 13%, de recombinantes entre essas espécies (GRIVET & ARRUDA, 2002; PIPERIDIS *et al.*, 2010).

No que diz respeito à similaridade genética, KIM *et al.* (2014) fizeram uso de sequências paralogas e ortólogas para realizar comparações entre o gênero *Saccharum*, *Miscanthus*, sorgo (*Sorghum bicolor*) e milho. Os autores concluíram que a separação entre as subtribos Saccharinae e Sorghinae se deu há 5,4 milhões de anos. Sugeriram que, após a divergência em relação ao sorgo, *Saccharum* e *Miscanthus* compartilharam um evento de duplicação de todo o genoma e, além disso, em *Saccharum*, houve um evento de duplicação exclusivo no gênero. Os resultados estimam em 3,1 milhões de anos a divergência entre esses dois gêneros.

Embora não haja ainda um genoma de cana-de-açúcar completamente sequenciado e montado, DE SETTA *et al.* (2014) abordaram funcionalmente o genoma da espécie pelo uso de análises comparativas com *Expressed Sequence Tags* (ESTs) de cana-de-açúcar, arroz e sorgo. Estes autores salientaram também a presença de grande quantidade de elementos transponíveis, com destaque para retrotransposons de longa repetição terminal. De fato, estes elementos podem estar relacionados ao processo evolutivo da espécie, por atuar no aumento do tamanho do genoma. As regiões entre genes que não codificam elementos transponíveis variaram no comprimento e na presença desses elementos. Em complemento, houve picos de pequenos RNAs mapeados a essas regiões, indicando que os elementos transponíveis sofrem ação regulatória. Além disso, obteve-se uma riqueza de sequências codantes funcionais ligadas a diversos processos biológicos.

DE SETTA *et al.* (2014) corroboraram, também, a semelhança entre o genoma da cana-de-açúcar e o genoma do sorgo, uma gramínea diploide proximamente aparentada ao gênero *Saccharum*. A semelhança entre as duas espécies, frente a outras gramíneas, é ainda evidenciada por resultados que comprovam a sintenia entre elas (MING *et al.*, 1998; WANG *et al.*, 2010). NISHIYAMA *et al.* (2014) montaram *contigs* de transcritos de cana-de-açúcar provenientes do cultivar SP80-3280, de um acesso de *S. officinarum* e um de *S. spontaneum* e puderam alinhá-los com transcritos de várias espécies da família *Poaceae*, sendo que a porcentagem de alinhamento era maior conforme a proximidade da espécie com a cana-de-açúcar. Além disso, esses autores inferiram que diversas espécies do gênero *Saccharum* compartilham boa parte do genoma entre si.

Na tentativa de elucidar o genoma dessa importante gramínea, alguns trabalhos lançaram mão de estratégias no nível de DNA e RNA. Com técnicas de alto rendimento, a atenção foi centrada em usar plataformas de leituras longas (WANG *et al.*, 2010; DE SETTA *et al.*, 2014) ou de leituras mais curtas (BERKMAN *et al.*, 2013). GRATIVOL *et al.* (2014) restringiram a montagem ao genoma hipometilado, filtrando as regiões metiladas, de modo que a retirada de regiões de baixa complexidade (elementos transponíveis, rDNA e regiões centroméricas) favoreceu a montagem de regiões de baixo número de cópias, o que refletiu nas estimativas de qualidade da montagem. Por uma estratégia híbrida, combinando as

plataformas PacBio e Illumina, 178 BACs aleatórios resultaram em uma montagem de 19,2Mb contida em aproximadamente 2,45 mil *scaffolds* (OKURA *et al.*, 2016).

Em nível transcricional, o primeiro grande esforço foi o projeto SUCEST, que a partir de bibliotecas de diferentes tecidos em distintos estágios de desenvolvimento produziram mais de 200 mil ESTs (VETTORE *et al.*, 2003). Porém, nos últimos anos as tecnologias de sequenciamento e os algoritmos de montagem favoreceram a construção de transcriptomas *de novo*. Nesse tipo de abordagem se encontram o transcriptoma conjunto de seis genótipos construído por CARDOSO-SILVA *et al.* (2014), o transcriptoma dos cultivares IACSP04-065 e IACSP04-627, contrastantes quanto ao conteúdo de lignina de VICENTINI *et al.* (2015) e outro que considerou um cultivar híbrido e dois acessos ancestrais (NISHIYAMA *et al.*, 2014).

Novas pesquisas visando elucidar os mecanismos que levam ao aumento da produtividade agrícola permitem o acúmulo de informações estruturais e funcionais sobre o genoma. Desse modo, passam a ser melhor compreendidas as bases moleculares de algumas das características que são difíceis de incorporar ou que demandam muito tempo através do melhoramento clássico, como é o caso do aumento do conteúdo de sacarose das plantas de cana-de-açúcar e tolerância a estresses, bióticos ou abióticos (CHEAVEGATTI-GIANOTTO *et al.*, 2011).

1.5 O estudo do transcriptoma através do sequenciamento de RNA

O transcriptoma é definido como a coleção de todas as moléculas de RNA expressas em um dado momento, as quais incluem o RNA mensageiro e os demais RNA não codantes (DE KLERK *et al.*, 2014). Diversos estudos dos transcritos de um organismo concentraram-se principalmente no uso de microarranjos, em que moléculas de cDNA produzidas a partir da transcrição reversa de RNA hibridizam-se com curtas sequências pré-determinadas, estabelecidas de acordo com genes já identificados no genoma.

Porém, avanços recentes permitiram o desenvolvimento de novas plataformas utilizadas para o sequenciamento de DNA, as quais possibilitaram a obtenção de sequências em grande quantidade. Por ser possível sintetizar o cDNA a partir de mRNA através da transcrição reversa e, em seguida, proceder com o sequenciamento, originou-se a técnica que passou a ser conhecida por sequenciamento de RNA (RNA-Seq). Paralelamente, desenvolveram-se estratégias para o estudo dessas sequências por meio de ferramentas computacionais (OZSOLAK & MILOS, 2011). Previa-se que o progresso advindo dos estudos de larga-escala conduzisse a tecnologia do sequenciamento a patamares de maior aplicabilidade, tornando-se peça fundamental dos estudos genéticos modernos (LIU *et al.*, 2012).

Atualmente, o alto rendimento, custos reduzidos e a análise massiva paralela são os princípios que permitem a competitividade do sequenciamento de nova geração (LIU *et al.*, 2012). Como já ressaltado anteriormente, as principais plataformas de sequenciamento de DNA foram adaptadas aos métodos de sequenciamento de RNA. Há uma subdivisão em duas classificações principais: *i*) o método baseado em *tag* e *ii*) o método *shotgun*, os quais diferem quanto ao tamanho e especificidade do fragmento sequenciado. A partir da sobreposição destes fragmentos quando alinhados a um genoma de referência devidamente anotado, os transcritos expressos em uma determinada condição são identificados (DE KLERK *et al.*, 2014).

Genomas maiores, como é o caso da cana-de-açúcar, demandam sequenciamento com profundidade de cobertura mais elevada (WANG *et al.*, 2009). No sequenciamento de nova geração, preza-se pelo alto rendimento de sequências e, especificamente no caso do RNA, tal tendência revolucionou a forma de compreensão do transcriptoma. É possível quantificar níveis de expressão com maior acurácia, o que auxilia estudos de expressão diferencial, sendo também fundamental na descoberta de novos transcritos, a partir da identificação de isoformas (TRAPNELL *et al.*, 2009). Mesmo em eucariotos cujo genoma é de grande complexidade, pode-se amostrar o transcriptoma virtualmente em sua totalidade, o que resulta

em informações sobre a junção dos éxons, modificações pós-transcricionais e variações que podem ter ocorrido nos genes transcritos, a partir dos dados de outros indivíduos de uma mesma espécie (WANG *et al.*, 2009).

Devido à existência prévia de recursos genômicos e transcriptômicos, é possível realizar comparações de novos dados do genoma e/ou transcriptoma com aqueles já disponíveis em bancos de dados, sendo que as novas descobertas podem ser incorporadas a tais repositórios. Na ausência de um genoma de referência da cana-de-açúcar, é factível proceder o alinhamento das leituras utilizando um transcriptoma como referência. Visto que há três montagens de transcriptoma disponíveis (CARDOSO-SILVA *et al.*, 2014; NISHIYAMA *et al.*, 2014; VICENTINI *et al.*, 2015), pode-se explorar esses recursos para identificar os transcritos a partir das leituras alinhadas e medir as alterações nos níveis de expressão dos transcritos de referência. Em seguida, verificam-se suas anotações para entender se determinadas vias metabólicas são estimuladas ou reprimidas. Um cuidado necessário é selecionar a referência mais adequada aos dados, pois deve abranger os transcritos relacionados a uma determinada forma de perturbação que é objeto de estudo.

Todos esses processos são possíveis graças à disponibilidade de ferramentas de bioinformática para análise dos dados biológicos. Algumas delas, como o *Bowtie2* (LANGMEAD & SALZBERG, 2012), são utilizadas para o mapeamento de leituras à referência; outras lidam até mesmo com sítios de splicing, como o *TopHat2* (KIM *et al.*, 2013) e o *HISAT2* (KIM *et al.*, 2015), possibilitando o estudo de isoformas dos transcritos. Em relação à expressão diferencial dos genes, ferramentas como *Cufflinks* (TRAPNELL *et al.*, 2009), *DESeq2* (LOVE *et al.*, 2014) e *edgeR* (ROBINSON *et al.*, 2009) são interessantes em estudos de organismos submetidos a condições experimentais contrastantes, como estresses, verificando-se a mudança na expressão de genes relacionados às vias metabólicas de interesse.

Não obstante esses avanços, é ainda essencial planejar cuidadosamente o experimento. Visto que as plantas possuem um sistema dinâmico e são sensíveis às mudanças ambientais, é imprescindível analisar os dados adequadamente, de modo a evitar viés ou confundimento de efeitos. As pesquisas com plantas, quando relacionadas à expressão gênica, exigem cautela na condução do experimento e na coleta dos dados, tendo-se atenção para que outras fontes de variação sejam minimizadas, bem como para a definição de tecido(s) e/ou momento(s) para a coleta do material biológico (YANG & WEI, 2015).

Há na literatura trabalhos publicados referentes à investigação de transcritos de cana-de-açúcar. Um desses trabalhos demonstra que, embora clones de cDNA aleatórios não possuíssem especificidade ao tecido e/ou fase do desenvolvimento dos tecidos, havia variação na intensidade dos clones entre folhas e colmos. Isso permitiu atribuir aqueles que eram diferencialmente expressos para cada tipo de tecido, alguns deles com homologia a outros genes de plantas (CARSON *et al.*, 2002). Por sua vez, dentre 119 transcritos diferencialmente expressos, CASU *et al.* (2007) constataram quais eram exclusivos a cada estágio fenológico do colmo. Por conseguinte, foram identificados transcritos participantes do processo de formação da parede celular. Similarmente a esses resultados de microarranjos, VICENTINI *et al.* (2015) também encontraram alterações na biossíntese de lignina, mas com um conjunto de resultados superior a dois mil transcritos com expressão diferencial, obtidos pelo sequenciamento de entrenós de dois cultivares contrastantes para conteúdo desse composto.

Em relação aos estresses abióticos, NOGUEIRA *et al.* (2003) atentaram-se à influência do estresse térmico na composição de ESTs do cultivar SP80-3280. As maiores variações entre as temperaturas 4°C e 26°C, seja no sentido de estímulo ou repressão, foi identificado em 48h. Diferentes tempos de exposição a esse estresse viabilizaram a observação do comportamento desses transcritos sob a forma de perfis de expressão temporal. Já ROCHA *et al.* (2007) contemplaram o estresse hídrico, em que transcritos tardios responderam de forma exclusiva em 72 h e 120 h após o início da supressão do fornecimento de água. Todavia, quando avaliado o déficit nutricional de fósforo, a expressão diferencial dos transcritos foi mais breve. Além desses estresses, os autores expandiram os resultados sobre perturbações às categorias de

tratamento com hormônios, herbivoria e simbiose.

Quanto ao transcriptoma baseado em bibliotecas de diferentes cultivares, pôde-se já encontrar genes relacionados ao acúmulo de sacarose e resposta a doenças, além de diversos outros processos metabólicos fundamentais à fisiologia do vegetal (CARDOSO-SILVA *et al.*, 2014). NISHIYAMA *et al.* (2014) verificaram que, embora o número de transcritos genótipo-específicos fosse baixo entre a cultivar híbrida SP80-3280, *S. officinarum* e *S. spontaneum*, eles podem estar relacionados com mudanças importantes na região expressa que diferenciam os seus fenótipos.

Já a interação entre a cana-de-açúcar com um patógeno pode ser examinada pela composição de transcritos de uma biblioteca construída para essa finalidade. SCHAKER *et al.* (2016) propuseram um estudo de avaliação temporal através do uso de um transcriptoma de referência da cana-de-açúcar, combinado aos transcritos obtidos pela montagem das leituras não alinhadas a essa referência. Tais leituras foram geradas pelo sequenciamento de RNA de bibliotecas de cana infectada por carvão, em dois momentos de infecção, 120 horas e 200 dias após a inoculação com *S. scitamineum*. Os resultados indicaram alterações no metabolismo de amido e sacarose, fixação de carbono e outras vias do desenvolvimento vegetal. No tempo final, os genes dos processos celulares e metabólicos mostraram-se reprimidos, em detrimento ao surgimento de genes classificados para sistema imune e processos multiorganismos.

1.6 Estudos de expressão diferencial e perfis de expressão gênica

A técnica do RNA-Seq permite a quantificação da abundância dos transcritos, ou seja, fragmentos oriundos dos transcritos podem ser identificados e é possível determinar os níveis de expressão gênica (WANG *et al.*, 2009). Dessa forma, pode-se proceder à comparação da intensidade de expressão em condições diferentes entre si, isto é, conduzir análises de expressão gênica diferencial. Assim, os perfis do transcriptoma são comumente utilizados na busca de genes diferencialmente expressos, que são aqueles em que há diferença no nível de expressão entre diferentes condições (SONESON & DELORENZI, 2013).

Desta forma, quando é feita a imposição de estresse sobre um organismo, durante a condução de um experimento, espera-se que ocorra variação no nível de expressão dos transcritos em relação à condição não perturbada. Tais mudanças podem ser observadas ao longo do tempo e os resultados podem ser verificados por técnicas que mensurem o nível de expressão gênica, a exemplo da RT-PCR. FERNANDEZ *et al.* (2012), por esta metodologia, identificaram, em folhas infectadas, aproximadamente 7 mil *contigs* correspondentes ao patógeno causador da ferrugem do café. A partir desse conjunto, dentre aqueles preditos como genes, fez-se a seleção de 22 *contigs* e, com isso, foi averiguado quais eram seus níveis de expressão pela comparação de quatro tempos distintos. A indução foi mais pronunciada 18 horas após inoculação (hai), 7 dias após a inoculação (dai) e 21 dai.

Para o estudo adequado de dados oriundos de uma análise temporal, os pontos ao longo do tempo escolhidos para amostragem, bem como o intervalo entre eles, devem ser planejados a partir do impacto que o estresse causa no vegetal e o grau esperado de diferenciação entre os perfis de expressão nos diversos momentos. Paralelamente, o número de réplicas biológicas deve ser suficiente para que erros na condução do experimento e execução dos protocolos experimentais, bem como o efeito da variação ao acaso, sejam minimizados (YANG & WEI, 2015).

1.7 Objetivos

O objetivo principal deste trabalho é elucidar, ao longo do tempo, o impacto da infecção do agente patogênico *Puccinia kuehni* sobre o cultivar suscetível de cana-de-açúcar SP89-1115. Trata-se do estudo de um patossistema cuja interação nunca fora antes descrita através do enfoque da expressão diferencial. A suscetibilidade do hospedeiro é importante para averiguar quais os possíveis mecanismos de

resposta afetados pelo patógeno, bem como suas alterações com o desenvolver da infecção. Essa influência do tempo pode ser verificada mais detalhadamente pela obtenção de perfis, através dos níveis de expressão tanto de transcritos individuais, ou de grupos de transcritos anotados em uma mesma classe funcional. A metodologia foi embasada pelo uso de dados do transcriptoma de folhas infectadas do cultivar, os quais foram gerados a partir do sequenciamento de RNA. Para isso, foram cumpridos os seguintes objetivos específicos:

- Pré-processar as leituras brutas e realizar o controle de qualidade dos dados de sequenciamento;
- Quantificar as leituras de cada amostra alinhadas a cada transcrito;
- Detectar a presença de transcritos diferencialmente expressos;
- Realizar análise funcional para verificar quais vias biológicas são alteradas pelo desenvolvimento do patógeno sobre o tecido do hospedeiro;
- Obter perfis de expressão gênica temporal para transcritos anotados em categorias funcionais de interesse.

2 MATERIAL E MÉTODOS

2.1 Material vegetal

O material biológico utilizado no presente estudo foi composto por folhas de cana-de-açúcar infectadas por *P. kuehni*, obtidas a partir do trabalho de Mestrado de Sergio Gregorio Perez Gomez, sob orientação do Prof. Dr. Nelson Sidnei Massola Júnior, no Departamento de Fitopatologia e Nematologia da ESALQ – USP (GOMEZ, 2013). O principal objetivo deste estudo foi a avaliação de parâmetros epidemiológicos monocíclicos da ferrugem alaranjada sobre os cultivares de cana-de-açúcar SP89-1115 e RB85-5156, respectivamente caracterizados como suscetível e moderadamente suscetível, condicionados em duas temperaturas, 18°C e 25°C.

O experimento em que as folhas foram coletadas seguiu delineamento inteiramente casualizado com cinco repetições por tratamento. Coletaram-se esporos de *P. kuehni* a partir de folhas doentes com pústulas abertas e, posteriormente, as plantas foram inoculadas com solução de esporos após 40 dias de crescimento.

Todo o material foi conduzido em uma câmara de orvalho durante 24 horas e, decorrido este tempo, as plantas foram alocadas em câmaras de crescimento conforme as temperaturas pretendidas. Como controle, havia duas plantas sem inoculação.

A folha +1, pela definição do sistema de Kuijper, é aquela cujas barbelas visíveis (colarinho) estão na posição mais alta, ou seja, é a mais alta das folhas desdobradas em que se pode verificar o colarinho (GALLO, 1962). As folhas +1 de todas as plantas foram coletadas em 0 hai, 2 hai, 12 hai, 24 hai, 48 hai, 5 dai, 10 dai e 12 dai, colocadas em tubos Falcon e armazenadas em freezer -80°C. Em maio de 2014 o material foi transportado e permaneceu armazenado em um freezer -80°C no Laboratório de Biotecnologia de Plantas do Centro de Ciências Agrárias da Universidade Federal de São Carlos (UFSCar – CCA).

2.2 Seleção de amostras e planejamento experimental

Como o objetivo principal do presente trabalho é a avaliação dos perfis de expressão gênica temporal ao longo do processo de infecção, objetivou-se uma seleção dos pontos no tempo mais apropriados para o sequenciamento de RNA, a fim de se amostrar os momentos mais informativos para a avaliação do impacto da presença do patógeno sobre a expressão de genes do hospedeiro. Assim, decidiu-se utilizar as amostras provenientes do cultivar SP89-1115, condicionado a 25°C, nos tempos 0, 12, 24 e 48 horas, bem como 5 e 12 dias após a inoculação. Esse cultivar é considerado altamente suscetível à ferrugem alaranjada, desde os primeiros relatos no território brasileiro (BARBASSO *et al.*, 2010; CHAPOLA *et al.*, 2016).

Optou-se por amostrar mais pontos no início do processo, pois se espera que o estabelecimento da infecção ocorra majoritariamente nesses momentos. O tempo de duas horas não foi escolhido porque a formação das estruturas de infecção ainda está em seu início, já que o patógeno emite um tubo germinativo a partir da germinação do esporo e gradativamente origina as estruturas de infecção sobre o hospedeiro (STAPLES, 2000). Nota-se a necessidade de se amostrar o tempo inicial (0 horas), visto que este representa um momento em que a planta está em uma condição de ausência do estresse.

2.3 Extração de RNA e sequenciamento

Realizou-se a extração do RNA total para duas réplicas biológicas do tempo 0 hai e três de cada um dos demais tempos, obtendo-se um total de 17 amostras. Para tanto, utilizou-se o RNeasy Plant Mini Kit (Qiagen). Sabe-se que três réplicas biológicas são um número adequado para trabalhos

de sequenciamento de RNA, mas é notável que duas réplicas biológicas sejam suficientes para modelar e conferir poder estatístico aos testes (LIU *et al.*, 2014; YANG & WEI, 2015).

Para a quantificação do RNA, fez-se o uso da espectrofotometria em equipamento NanoDrop, requerendo-se um volume de 1 μL de cada amostra. Em relação à avaliação da qualidade do RNA, observaram-se as bandas 28S e 18S em gel de agarose 1% submetido a eletroforese. Por último, obteve-se o RNA Integrity Number (RIN) pelo equipamento Agilent 2100 Bioanalyzer, com a finalidade de avaliar a qualidade e concentração do RNA nas amostras, procedimento necessário para determinar se essas estavam adequadas às necessidades do sequenciamento.

O sequenciamento foi realizado na plataforma Illumina HiSeq 2500, no Laboratório de Biotecnologia Animal da ESALQ – USP, coordenado pelo Prof. Dr. Luiz Lehmann Coutinho. Cada uma das bibliotecas foi sequenciada no formato de leituras pareadas (2 x 125 pb). Essas bibliotecas foram subdivididas em duas *lanes*, para que eventuais problemas no sequenciamento das bibliotecas pudessem ser diagnosticados como oriundos de seu preparo ou do sequenciamento de *lanes* específicas. Não obstante, essa estratégia possibilita a manutenção da profundidade de cobertura desejada para cada biblioteca.

2.4 Pré-processamento das leituras

2.4.1 Diagnóstico das leituras pareadas

Os dados brutos de sequenciamento foram recebidos no formato FastQ, de modo a já serem adequados para a utilização em outras ferramentas. Para um controle de qualidade inicial, conjuntos de diagnósticos das sequências foram obtidos pelo *software* FastQC (ANDREWS, 2010). Entre os resultados fornecidos para as leituras de cada par, atenção foi voltada principalmente para os diagnósticos de qualidade das bases em ambas as extremidades, a distribuição da frequência de cada base nitrogenada ao longo das leituras, conteúdo de *k-mers* e conteúdo de GC (TURNER, 2014; DOHM *et al.*, 2008; HANSEN *et al.*, 2010).

O primeiro ponto investigado é interessante para ver a qualidade geral das leituras, pois é possível ver o *score* Phred de bases únicas ou intervalos de bases, conforme a posição dentro dos fragmentos. O segundo permite verificar a constância da porcentagem de cada base nos fragmentos sequenciados, visto que é esperado que a base de cada posição seja ocupada aleatoriamente, ou seja, por qualquer uma das quatro bases. Já o terceiro pode indicar a presença de sequências dos adaptadores, de cada par de leitura, necessários à formação de pontes utilizadas na amplificação por PCR que resultam nos *clusters* na superfície das *lanes*.

Além de uma visão geral do resultado da técnica de sequenciamento, é necessário considerar se as etapas de extração e construção das bibliotecas não introduziram contaminantes, como é o caso de moléculas de DNA, rRNA ou outros organismos que não fossem aquele em estudo. Dessa maneira, a verificação da distribuição do conteúdo GC pode apontar se as leituras atendem à distribuição esperada desse conteúdo, ou se há desvios nessa distribuição, havendo indícios de contaminação.

2.4.2 Controle da qualidade e remoção de rRNA

Em seguida, conforme a avaliação prévia, a remoção das sequências dos adaptadores Illumina TrueSeq, das bases terminais com *score* Phred de qualidade inferior a 20, das 12 primeiras bases em 5' e, também, leituras com bases ambíguas foi efetuada através do *cutadapt* (MARTIN, 2011). Para isso foram fornecidas simultaneamente as leituras de cada par, a fim de serem removidos os adaptadores *forward* e *reverse*.

Seguidamente, para a remoção de leituras provenientes de moléculas de rRNA, os pares de leituras foram alinhados às sequências de rRNA eucarióticos (18S e 28S) por meio do *sortMeRNA* (KOPYLOVA

et al., 2012). Aquelas que passaram pelo filtro, ou seja, não foram alinhadas aos rRNA, foram mantidas para as etapas posteriores.

2.5 Alinhamento das leituras e construção da matriz de contagem

Embora as folhas estivessem inoculadas com esporos do fungo, era esperado que a grande maioria dos dados fosse proveniente de sequências de cana-de-açúcar, espécie com ausência de um genoma de referência. Por isso foi utilizado o transcriptoma publicado por CARDOSO-SILVA *et al.* (2014). O conjunto de leituras que o originou adveio das cultivares IACSP96-3046, IACSP95-3018, SP81-3250, RB925345, SP80-3280 e RB835486. Como os indivíduos desses cultivares foram conduzidos em campo, possivelmente estavam sujeitos à ocorrência de estresses diversos. O alinhamento das leituras à referência foi realizado pelo *HISAT2* (KIM *et al.*, 2015), *software* que permite utilizar informações de sítios de *splicing*, caso estejam disponíveis para o organismo em questão. Não obstante, esse programa também é adequado para o alinhamento de leituras sem *gaps* intrônicos, como ocorre no alinhamento contra uma referência de transcriptoma.

Por se tratar de um transcriptoma, o tipo de alinhamento utilizado foi o local, sendo inativados os módulos relacionados com a presença de íntrons do *HISAT2*. O próximo passo foi atribuir as leituras aos diversos transcritos e quantificar o nível de expressão de cada um, a fim de se obter uma matriz de contagem. Através do *HTSeq* (ANDERS & HUBER, 2010) foram feitas essas contagens para cada uma das amostras. Essa matriz de contagens conteve um transcrito por linha, representando-se a contagem de cada uma das amostras sequenciadas dos diversos tempos amostrais em uma coluna individual.

2.6 Análise de expressão diferencial

A fim de realizar os testes de expressão diferencial, foi utilizado o pacote *edgeR* (ROBINSON & SMYTH, 2008). Inicialmente, foram retirados os transcritos cujas contagens por milhão eram inferiores a um em pelo menos duas amostras (réplicas dos pontos temporais), de modo a eliminar transcritos não expressos, aqueles com baixo nível de expressão ou ainda aqueles expressos exclusivamente em uma única amostra. Nesses três casos os testes não teriam poder estatístico para detectar a expressão diferencial. Após isso, procedeu-se à normalização dos dados, com o intuito de eliminar a influência da cobertura de sequenciamento e também diferenças de composição entre as amostras. Visto que cada ponto temporal foi sequenciado em duas *lanes*, um gráfico do tipo *heatmap* foi utilizado para averiguar a similaridade entre todas as amostras sequenciadas.

Devido à investigação de seis pontos no tempo (0 hai, 12 hai, 24 hai, 48 hai, 5 dai e 12 dai), é possível testar hipóteses nulas envolvendo todos os tratamentos simultaneamente, bem como contrastes entre apenas dois momentos específicos. O *Fold Change* (FC) representa a razão da expressão média de um transcrito entre duas condições contrastantes. Usualmente se utiliza o logaritmo na base dois dessa razão, denotado por $\log_2\text{FC}$ (LFC), para mostrar quando há mudança na expressão do transcrito, seja no sentido de estímulo ou repressão (ANDERS & HUBER, 2010; LOVE *et al.*, 2014).

Sendo y_{gi} o número de leituras da i -ésima amostra mapeadas no gene g , o *edgeR* assume uma distribuição Binomial Negativa com média μ_{gi} e parâmetro de dispersão ϕ_g , isto é, $y_{gi} \sim NB(\mu_{gi}, \phi_g)$. Nesta distribuição tem-se que $\mu_{gi} = N_i \pi_{gi}$, em que π_{gi} representa a fração de fragmentos do gene g na i -ésima amostra e N_i o número total de leituras mapeadas na amostra i . Já a variância da distribuição é dada por $\text{Var}(Y_{gi}) = \mu_{gi} + \mu_{gi}^2 \phi_g$, de modo que há um componente linear associado à variabilidade experimental entre amostras, bem como um componente quadrático atribuído à variação biológica entre réplicas de uma mesma condição experimental.

O parâmetro ϕ_g é modelado de tal forma que sua raiz quadrada $\sqrt{\phi_g}$ representa o coeficiente de variação biológica (MCCARTHY *et al.*, 2012). Porém, como o número de amostras é pequeno, para se estimar ϕ_g empregou-se a verossimilhança ponderada, de modo que cada ϕ_g foi encolhido em direção a uma estimativa alisada da dispersão, sendo esta comum para transcritos de médias similares.

O teste F de quase-verossimilhança (LUN *et al.*, 2016) foi utilizado para testar a expressão diferencial, por ser preferido para quando o número de réplicas é pequeno, situação em que estimativa da dispersão torna-se mais incerta. Três tipos de abordagem para a expressão diferencial foram empregadas:

- (a) Um teste do tipo Análise de Variância (ANAVA), no qual foi reportado como diferencialmente expresso o transcrito que apresentasse diferença significativa entre pelo menos um par de pontos amostrais no tempo.
- (b) Genes diferencialmente expressos nos tempos 12 hai, 24 hai, 48 hai, 5 dai e 12 dai em relação ao tempo inicial (0 h), utilizado como ponto de referência.
- (c) Genes diferencialmente expressos quando comparados os tempos adjacentes (12 hai vs 0 h, 24 hai vs 12 hai, 48 hai vs 24 hai, 5 dai vs 48 hai e 12 dai vs 5 dai).

Como os testes de expressão diferencial foram realizados para um grande número de genes, um caso de múltiplos testes, é importante controlar o número falsos positivos, cuja probabilidade de ocorrência aumenta à medida que se eleva o número de testes. Optou-se pela correção de múltiplos testes pela *False Discovery Rate* (FDR), considerando independência entre os testes (BENJAMINI & HOCHBERG, 1995) e admitindo um nível de significância q^* , que representa a proporção aceita de falsos positivos entre os testes significativos. Em cada uma das três abordagens, os transcritos diferencialmente expressos foram aqueles com uma FDR inferior a 5%.

2.7 Análise de enriquecimento funcional

O transcriptoma de CARDOSO-SILVA *et al.* (2014) foi anotado através do BLAST2GO (CONESA *et al.*, 2005) seguindo as etapas de alinhamento contra a base de dados não-redundante do NCBI (NR), busca por domínios conservados com base nos resultados do InterProScan, mapeamento dos termos *Gene Ontology* (GO) e anotação destes termos aos transcritos. Na primeira etapa, para cada transcrito, foram considerados os alinhamentos com até 20 *hits* com *e-value* inferior a 1×10^{-3} . Foram utilizadas as anotações mais específicas possíveis para cada sequência, dentro das três classes: processos biológicos, funções moleculares e componentes celulares. Ao final foi elaborado um arquivo contendo o identificador dos transcritos, os termos GO associados a cada um deles, quando presentes, além de uma descrição da potencial função de cada transcrito.

Através do pacote *goseq* (YOUNG *et al.*, 2010) foram realizadas análises de enriquecimento funcional para os transcritos diferencialmente expressos, com correção para viés de tamanho. Assim, o comprimento de cada transcrito foi utilizado como fator de correção, visto que transcritos mais longos apresentam maior número de leituras mapeadas e, portanto, maior poder estatístico para detecção de expressão diferencial. Foram feitas análises de enriquecimento separadamente para os transcritos identificados a partir de dois tipos de testes realizados, a saber: o teste do tipo ANAVA e as comparações entre tempos adjacentes. Foram considerados enriquecidos os termos sobre-representados com valor-*p* ajustado para FDR igual ou inferior a 10%.

2.8 Perfis de expressão gênica

Os perfis de expressão gênica são importantes para entender o comportamento dos transcritos com o passar do tempo da infecção. Foram utilizadas duas estratégias com base em todos nos transcritos

diferencialmente expressos, tanto no teste do tipo ANAVA, quanto nos testes de tempos adjacentes:

a) Utilizou-se um gráfico de linhas dos transcritos diferencialmente expressos que foram anotados a cada termo GO. A pretensão foi verificar a existência de determinados padrões no tempo entre os transcritos de um termo, que colaborassem no entendimento do porquê foram enriquecidos somente em determinadas comparações. Além disso, foi possível perceber se transcritos pertencentes a processos diversos poderiam apresentar padrões similares de expressão. Essa forma de análise foi possível através da normalização das contagens dos transcritos entre as diversas amostras, de modo que a variância das contagens fosse unitária para todos os transcritos de interesse.

b) Fez-se o uso de *Heatmaps* para se obter um agrupamento hierárquico dos transcritos diferencialmente expressos conforme os LFC dos tempos 12 hai, 24 hai, 48 hai, 5 dai e 12 dai em relação a 0 h. Essa estratégia foi utilizada para termos GO e transcritos específicos de interesse. Como complemento, esse tipo de gráfico permitiu identificar os transcritos conforme a descrição da anotação. Os valores de estímulo e repressão foram truncados nos valores de 5 e -5, respectivamente.

2.9 Mapas das vias moleculares

A anotação do transcriptoma com o BLAST2GO incluiu informações sobre o *enzyme code* (EC) para diversos transcritos. Criando-se um arquivo de associação entre transcritos anotados com EC e fornecendo-se quais eram os transcritos diferencialmente expressos num dado contraste, foi possível separar os ECs diferencialmente expressos (em ambos os sentidos, isto é, com expressão estimulada ou reprimida) nos testes entre 12 hai vs 0 h e em 48 hai vs 24 hai. Com a intenção de se identificar mapas globais do metabolismo enriquecidos com genes diferencialmente expressos, a conversão dos ECs para a categoria de grupo ortólogo (KO) foi realizada através da obtenção da relação entre esses dois identificadores no KEGG API (<http://www.kegg.jp/kegg/docs/keggapi.html>). Posteriormente, os KOs de interesse foram obtidos pela associação dos transcritos diferencialmente expressos ao identificador K. No KEGG Mapper (http://www.genome.jp/kegg/tool/map_pathway2.html) utilizou-se a ferramenta de busca, colorindo em vermelho os genes estimulados e em azul os reprimidos.

3 RESULTADOS E DISCUSSÃO

3.1 Controle de qualidade e pré-processamento das leituras

O número de leituras sequenciadas em cada um dos pontos temporais pode ser encontrado na Figura 3.1 e na Tabela A.1. É possível notar que, à exceção do tempo 0 h, todas as demais amostras foram constituídas por uma quantidade de fragmentos entre 70 e 80 milhões. Essa circunstância pode ser explicada pela ausência de uma réplica no primeiro ponto amostral.

Cerca de 99,92% das leituras foram mantidas após o pré-processamento para a remoção das sequências de adaptadores Illumina TruSeq. A eliminação de bases das extremidades com score de qualidade Phred inferior a 20 resultou na remoção de 717.207.956 bases (0.63%). Adicionalmente, foram retiradas as 12 primeiras bases da extremidade 5', devido ao fato de que não havia uma proporção constante entre os quatro tipos de nucleotídeos nessas regiões. Tal situação é esperada para o sequenciamento de RNA, pois na construção das bibliotecas a fase de amplificação utiliza primers hexâmeros aleatórios, os quais introduzem viés de composição nucleotídica no início dessas leituras (HANSEN *et al.*, 2010).

Subsequentemente, aproximadamente 14,36% das leituras corresponderam a sequências de rRNA, pois alinharam-se às referências de rRNA 18S e 28S, as quais são presentes em organismos eucarióticos. Quando é observado o comportamento das amostras após a filtragem, verifica-se que o número de fragmentos sequenciados da primeira réplica de 0 h foi reduzido praticamente à metade. Nesse caso, a queda foi atribuída a uma baixa qualidade na primeira biblioteca de 0 h, possivelmente pela grande presença de rRNA na biblioteca (Tabela 3.1). O tamanho aproximado das bibliotecas sequenciadas de todas as amostras que compunham o transcriptoma de seis momentos distintos da infecção da cana-de-açúcar por ferrugem alaranjada, após a etapa de pré-processamento, foi de 86,6G pb.

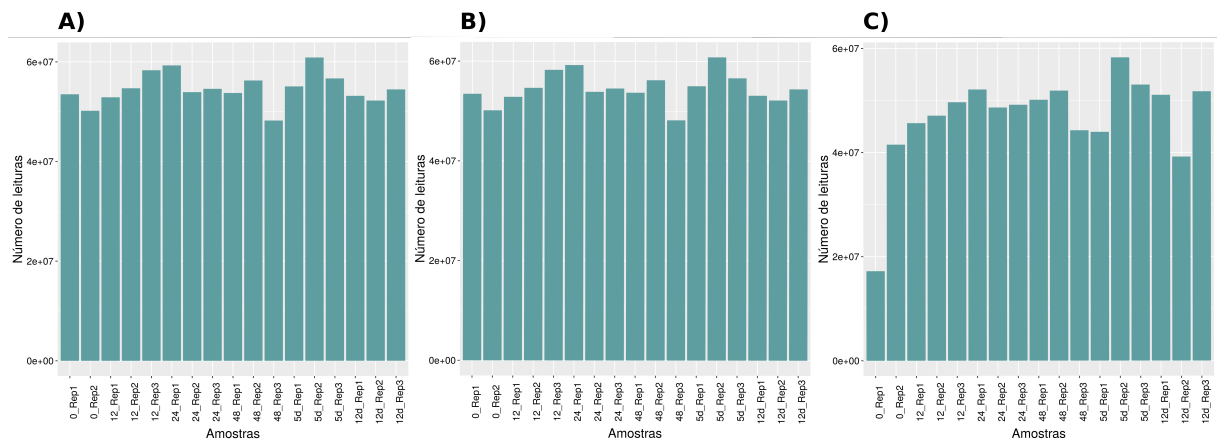


Figura 3.1. Número de leituras das amostras de cada um dos tempos. As figuras foram dispostas conforme as etapas do pré-processamento: (A) Número de leituras brutas oriundas do sequenciamento; (B) Número de leituras após a remoção de extremidades de baixa qualidade e de adaptadores; (C) Número de leituras após a remoção daquelas alinhadas a rRNAs.

A porcentagem média de alinhamento ao transcriptoma de referência foi de cerca de 65,70%, sendo notável que a primeira repetição do tempo 0 h distinguiu-se pela menor porcentagem de alinhamento. As demais amostras variaram de 51,56% a 71,67% (Tabela A.2). Portanto, manteve-se apenas uma das réplicas do tempo 0 h, já que a outra possuía grande porcentagem de leituras de rRNA e poderia viesar as análises e comprometer a interpretação dos resultados. Logo após, foi obtida a matriz de contagens com 72.269 linhas correspondentes aos transcritos e com as 32 amostras elencadas nas colunas.

Posteriormente, considerando-se apenas os transcritos que apresentaram pelo menos uma leitura alinhada, constatou-se que as duas *lanes* que compunham cada repetição eram muito similares (Figura

A.1 - A e Figura A.2) e, com isso, suas contagens foram somadas para cada um dos transcritos (Figura A.1 - B). Por conseguinte, para a expressão diferencial, foram utilizadas 16 amostras que eram representantes de uma das três réplicas dos pontos temporais, com exceção do ponto 0 h, no qual apenas uma réplica foi mantida. Embora essa não seja a situação ideal, os testes estatísticos de expressão diferencial podem lidar com esse problema, visto que estimativas de dispersão obtidas a partir de um conjunto de réplicas biológicas podem ser utilizadas para as demais amostras do experimento (CHEN *et al.*, 2016). É interessante notar que na Figura A.2 (e também em A.1 - B) seria esperado que as réplicas de um mesmo tempo se mantivessem mais próximas, mas houve uma mistura das réplicas de 48 h e 5 h, juntamente a 0 h.

3.2 Expressão diferencial: Teste conjunto entre todos os pontos no tempo

Para a análise de expressão diferencial, é necessário normalizar as contagens entre amostras, estimar a dispersão de acordo com as réplicas de cada ponto temporal e, por fim, decidir o teste estatístico mais adequado para verificar se os transcritos são diferencialmente expressos entre duas ou mais condições experimentais. A primeira estratégia aqui utilizada foi examinar a ocorrência de transcritos com diferenças significativas em pelo menos um ponto ao longo do tempo, através de um teste do tipo ANOVA. Assim, foi averiguado que cerca de 65% dos 28.641 transcritos foram detectados como diferencialmente expressos para uma FDR inferior a 5%. Dado esse conjunto, a etapa seguinte foi realizar a análise de enriquecimento funcional com a finalidade de observar quais rotas metabólicas eram influenciadas pelo estresse biótico. Para isso, seria necessário conhecer as funções moleculares, componentes celulares ou processos biológicos aos quais os transcritos da referência foram associados. Nesse transcriptoma, 26.821 transcritos foram anotados para 4.086 diferentes termos GO. Porém, como em diversos casos um transcrito se mostrou associado a mais de um termo, a anotação incluiu um total de 82.006 anotações de termos GO para os diversos transcritos.

Foram encontrados 47 termos enriquecidos (Figura 3.2 e Tabela A.3), nos quais é destacável principalmente a presença de processos relacionados com a ocorrência de oxirredução nas células (isto é, *oxidation-reduction process*, *oxidoreductase activity*, *protein disulfide oxidoreductase activity*) ou ao cloroplasto, suas membranas e estruturas, assim como a própria atividade fotossintética. Através da Figura 3.2 é possível perceber um grande número de transcritos diferencialmente expressos na categoria de processo de oxidação e redução, seguido dos processos metabólicos e regulação da transcrição. Porém, uma observação é que em algumas categorias com menor número de transcritos, como *chloroplast*, *chloroplast thylakoid membrane* e *chloroplast envelope*, a significância do teste foi mais elevada, porque a proporção de transcritos diferencialmente expressos entre aqueles que as compõem foi alta (Figura A.4).

Adicionalmente, outros termos podem indicar certa relação com o estresse oxidativo. Segundo as informações do AMIGO (<http://amigo.geneontology.org/amigo/landing>), o termo ligação ao íon de ferro é uma função molecular de interação seletiva a íons desse elemento, de tal forma que a oxidação do íon ferroso para o férrico (Fe^{+3}) pode ser inclusa nesse tipo de função. Complementarmente, o termo GO:0016705 (*oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen*) indica ação de transferência de elementos dentro da reação no sentido de reduzir o oxigênio, levando à produção das espécies reativas de oxigênio (ROS). *Iron-sulfur cluster assembly*, por sua vez, corresponde às estruturas [Fe-S] presentes em proteínas essenciais em inúmeras funções biológicas, principalmente na transferência de elétrons nas cadeias transportadoras e na constituição de sítios de ligação de proteínas (JOHNSON *et al.*, 2005).

Por fim, *cell redox homeostasis* pode estar relacionada com a ação de proteínas responsáveis por manter níveis normais de ROS, como é o caso da glutathiona-S-transferase (GST), superóxido-dismutase (SOD), catalases e peroxidases (INZÉ & MONTAGU, 1995). Entre os transcritos diferencialmente expres-

tos, foram observados 64 como GST, 40 peroxidases, sete SOD e cinco catalases. De fato, de forma geral, nota-se que o nível de expressão da SOD caiu de 0 a 12 horas, aumentando novamente às 48 h (Figura A.5). As catalases, por sua vez, exibiram um incremento em seus níveis de expressão já em 24 h (Figura A.6). São quatro classes de enzimas de grande importância porque as ROS são extremamente nocivas às outras moléculas que são componentes celulares básicos, como nucleotídeos, peptídeos e lipídios. Ferritinas e transferrinas, as proteínas estocadora e transportadora do ferro, respectivamente, são reguladas pós-transcricionalmente por esse elemento de modo que ele não colabore na oxirredução. Em contrapartida, a atividade de H_2O_2 sobre sinalizadores faz com que a proteína IRP1 (*iron-regulatory protein 1*) tenha conformação de repressora da ferritina (INZÉ & MONTAGU, 1995; RAY *et al.*, 2012).

Os termos da oxirredução e homeostase do sistema redox estão relacionados com a formação de ROS nas células. As ROS são moléculas que se originam a partir do oxigênio *triplet* (3O_2) pelo próprio mecanismo fotossintético das plantas, da respiração mitocondrial e também pela ação dos peroxissomos. São os íons superóxido ($O_2^{\bullet-}$), peróxido de hidrogênio (H_2O_2), hidroxil (HO^{\bullet}), hidroperoxila (HO_2^{\bullet}), oxigênio *singlet* (1O_2) e hidroxila (OH^-). Na ausência de distúrbios, o metabolismo celular as mantém reguladas, em baixos níveis. O desbalanço entre esses agentes oxidantes e as diversas moléculas antioxidantes pode ocorrer, além de diversas outras formas de estresse, pela infecção da planta por um patógeno. Como consequência as ROS acumulam-se nas células e tal explosão oxidativa se torna uma via de sinalização a fim de desencadear outros processos celulares controlados por genes de resistência (GADJEV *et al.*, 2006; SCHMITT *et al.*, 2014).

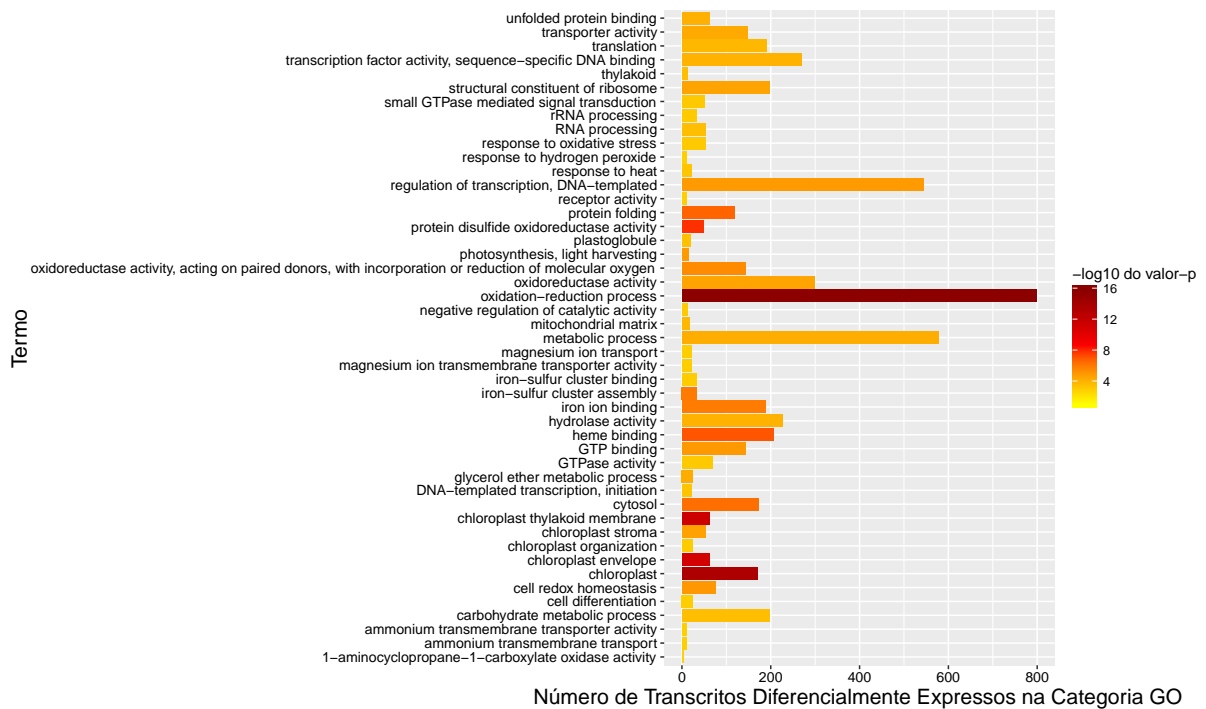


Figura 3.2. Termos funcionais enriquecidos e a quantidade dos transcritos diferencialmente expressos do teste do tipo ANOVA que os compõem. No eixo *y* estão os termos enriquecidos de genes diferencialmente expressos em resposta à infecção por *Puccinia kuenii*, em pelo menos um dos tempos, de acordo com a anotação do do transcriptoma de CARDOSO-SILVA *et al.* (2014). O número de transcritos diferencialmente expressos dentro cada categoria é indicado através do eixo *x*. O gradiente de cores elenca os valores-p na escala logarítmica.

LOARCE *et al.* (2016), em seus resultados de anotação de isogrupos, constataram transcritos referentes à geração das ROS, uma defesa precoce do hospedeiro. Adicionalmente, INZÉ & MONTAGU (1995) ressaltam a importância desse grupo na resposta de hipersensibilidade, na formação de lignina

na parede celular, sua relação com glicoproteínas ricas em hidroxiprolina (HRGP - *hydroxyproline-rich glycoproteins*) e genes da via das fitoalexinas.

Ao considerar as vias de defesa por ação da lignina, fitoalexinas e glicoproteínas ricas em hidroxiprolina, pela expressão dos genes HRGP, chalcona sintase (CHS - enzima essencial na síntese de fitoalexinas), PAL (enzima relacionada tanto às fitoalexinas quanto à via de lignina), LAWTON & LAMB (1987) verificaram um rápido estímulo de transcrição dos dois últimos já nos primeiros 5 minutos após a infecção, e uma resposta posterior de HRGP aproximadamente 7h após o tratamento. Em 4 hai, SCHMELZER *et al.* (1989) averiguaram a presença dos mRNAs de PAL, PR1 (*Pathogenesis-related protein*), e de 4CL (*4-coumarate:CoA ligase*). Embora sendo encontrados sítios de infecção da raça 1 do fungo *Phytophthora megasperma* f. sp. *glycinea* nas mudas de salsinha (*Petroselinum crispum* cv Hamburger Schnitt), esta planta demonstrava resistência por não ser o hospedeiro principal para essa doença, confirmada pela detecção da resposta de hipersensibilidade nesses locais. Nesse experimento, a quantidade de mRNA desses três genes era alta até 8 hai, enquanto que o transcrito BMT (*S-adenosyl-L-methionine:bergaptoil O-methyltransferase*), também da via das fitoalexinas, mostrava-se um pouco mais tardio (de 12 a 18 hai).

Ao que tudo indica, alguns dos mecanismos de resposta, como biossíntese de lignina e fitoalexina, são relativamente rápidos, agindo assim que o fungo inicia seu estabelecimento sobre as células internas da planta. Isso pode indicar que a presença de ROS pode ser evidenciada como um primeiro mecanismo de sinalização irrompendo essas vias.

As chamadas ondas de ROS são eventos de propagação de sinal dentro das células vegetais. A locomoção de algumas dessas moléculas as transformam em uma categoria de sinalizadores de alta mobilidade, com picos concomitantes a outros eventos de sinalização (MITTLER *et al.*, 2011). GADJEV *et al.* (2006), após a ocorrência de ROS, investigaram a ocorrência de transcritos com expressão diferencial. Entre os experimentos de microarranjos para o estresse oxidativo, um deles se caracterizava pela avaliação temporal da aplicação da toxina de *Alternaria alternata* em *A. thaliana*. Como estimulados foram encontrados os fatores de transcrição WRKY, glutationa-S-transferase, glicosil transferases, oxirredutases, entre outros. Os WRKY, por exemplo, são fatores de transcrição estimulados sob condições de estresse, já descritos na presença de patógenos em plantas (GADJEV *et al.*, 2006; JONES & TAKEMOTO, 2004; PANDEY & SOMSSICH, 2009; FANG *et al.*, 2016).

SCHMITT *et al.* (2014) apontam que o HO• é um mensageiro de longa distância, sendo mais reativo do que as demais ROS. Surge pela interação do peróxido de hidrogênio (também um mensageiro intracelular e intercelular) com o Fe⁺² em uma das etapas da reação de Fenton. Alguns autores também consideraram o H₂O₂ como um mensageiro secundário na comunicação de uma perturbação celular de causa biótica (INZÉ & MONTAGU, 1995).

Pode-se indagar a ocorrência do termo *photosynthesis, light harvesting*, que está dentro da categoria fotossíntese. Uma possível explicação é pelo fato de que o desenvolvimento do fungo se dá sobre o tecido foliar, perturbando a atividade fotossintética. SCHOLLES & ROLFE (1996) notaram queda na taxa fotossintética foliar, mas somente após os sintomas iniciais já estarem aparentes, sendo a queda mais proeminente em 8 e 11 dai. Todavia, embora a fotossíntese possa demorar a apresentar redução, a queda na transcrição dos genes pode ser mais rápida. Os demais termos obtidos na Tabela A.3 correspondem a processos mais gerais, ou seja, houve modificações no funcionamento das células do hospedeiro com o estabelecimento e desenvolvimento do fungo. Os cloroplastos e suas estruturas, por exemplo, contêm as proteínas dos fotossistemas I e II nas membranas do tilacoide. Mas essa organela é determinante também na síntese de componentes básicos das biomoléculas e, na infecção, atua na sinalização através da formação de ROS (pela transferência de elétrons oriundos da excitação pela luz), óxido nítrico e os hormônios ácido jasmônico e ácido salicílico (KANGASJÄRVI *et al.*, 2012).

Os resultados dessas análises forneceram uma ampla visão dos principais processos metabólicos

afetados ao longo do processo de infecção por *P. kuehni*. Em particular, a ferrugem alaranjada afetou os mecanismos ligados à fotossíntese, homeostase redox, transcrição e tradução. No entanto, uma limitação desta análise conjunta entre os pontos no tempo é a dificuldade de distinguir os genes com expressão maior ou menor em cada momento, pois aqui são comparados os níveis de expressão entre os seis tratamentos. Quando esse teste é realizado entre duas condições contrastantes, pode-se indicar se a diferença entre os níveis de expressão é significativa e se está no sentido de estímulo ou repressão. Essa última abordagem também poderia expandir o número e diversidade de termos encontrados entre dois tempos distintos, que refletem diferentes etapas do desenvolvimento de *P. kuehni* no tecido foliar.

3.3 Análise de expressão diferencial: o avançar da infecção pela comparação de momentos específicos

Uma outra maneira de averiguar alterações na expressão gênica é através de contrastes que comparam dois tratamentos, ou seja, são feitas múltiplas comparações para avaliar a expressão diferencial entre momentos distintos. A Tabela 3.1 mostra o número de transcritos diferencialmente expressos, quando feita a comparação de cada um dos pontos após a inoculação contra o tempo 0 h, o qual foi utilizado como referência, a uma FDR menor que 5%. De maneira geral, foi constatado que para ambos os casos de expressão diferencial (*up* e *down-regulated*), os números de transcritos afetados foram maiores em 12 hai, com decréscimo até 48 hai, quando voltaram a aumentar. Isto indica que houve dois momentos principais de resposta: uma resposta inicial entre 0 e 24 hai e outra mais tardia, a partir de 5 dai. Nota-se um período intermediário, às 48 hai, em que a expressão de muitos transcritos voltou aos níveis iniciais.

Tabela 3.1. Número de transcritos diferencialmente expressos em cada momento após a inoculação, em comparação ao tempo controle inicial, com subdivisão dentro dos tempos em transcritos estimulados (*Up*) e reprimidos (*Down*).

Regulação	12 hai		24 hai		48 hai		5 dai		12 dai	
	<i>Up</i>	<i>Down</i>	<i>Up</i>	<i>Down</i>	<i>Up</i>	<i>Down</i>	<i>Up</i>	<i>Down</i>	<i>Up</i>	<i>Down</i>
Transcritos Diferencialmente Expressos	3.537	4.643	2.947	3.578	951	1.032	1.529	1.767	2.864	3.048
Total	8.180		6.525		1.983		3.296		5.912	

Finalmente, outra possibilidade é fazer comparações de tempos adjacentes, isto é 12 hai vs 0 h, 24 hai vs 12 hai, 48 hai vs 24 hai, 5 dai vs 48 hai e 12 dai vs 5 dai, cujos resultados estão esquematizados na Tabela A.4 e exibidos na Figura 3.3. Estas comparações fornecem uma visão bastante detalhada sobre os processos que ocorreram a cada momento, se a infecção os fez potencialmente mais relevantes em algum ponto em detrimento de outros, ou se houve mudança no comportamento da expressão com o passar do tempo. Em três das comparações (12 hai vs 0 h, 48 hai vs 24 hai e 12 dai vs 5 dai) o número de transcritos diferencialmente expressos excederam 6.000, sendo que a quantidade foi maior tanto com transcritos com LFC positivo quanto aqueles com LFC negativo. Aqui, observa-se que as primeiras 48 horas após a infecção foram mais abundantes em transcritos diferencialmente expressos, que podem vir a ser os principais momentos de interação entre os tecidos foliares da cana-de-açúcar e *P. kuehni*.

O teste de enriquecimento funcional revelou os principais processos dos quais os transcritos diferencialmente expressos eram membros (Tabelas A.5 a A.9). Além do teste conjunto, considerando todos os transcritos diferencialmente expressos em determinada comparação, também foi analisada a subdivisão do sentido de expressão para serem encontrados termos sobre-representados somente entre os transcritos estimulados e, posteriormente, entre os reprimidos. Em quase todos os casos, foram encontrados termos com alguma relação à transcrição, regulação da transcrição e ribossomos. Conquanto sejam os processos mais basais, a partir dos quais as vias bioquímicas são iniciadas, não descrevem tão minuciosamente a

influência da perturbação sobre as células. As Tabelas de enriquecimento funcional do Apêndice III têm em destaque termos de potencial relevância ao entendimento da resposta da planta à infecção conforme o ciclo de desenvolvimento da ferrugem.

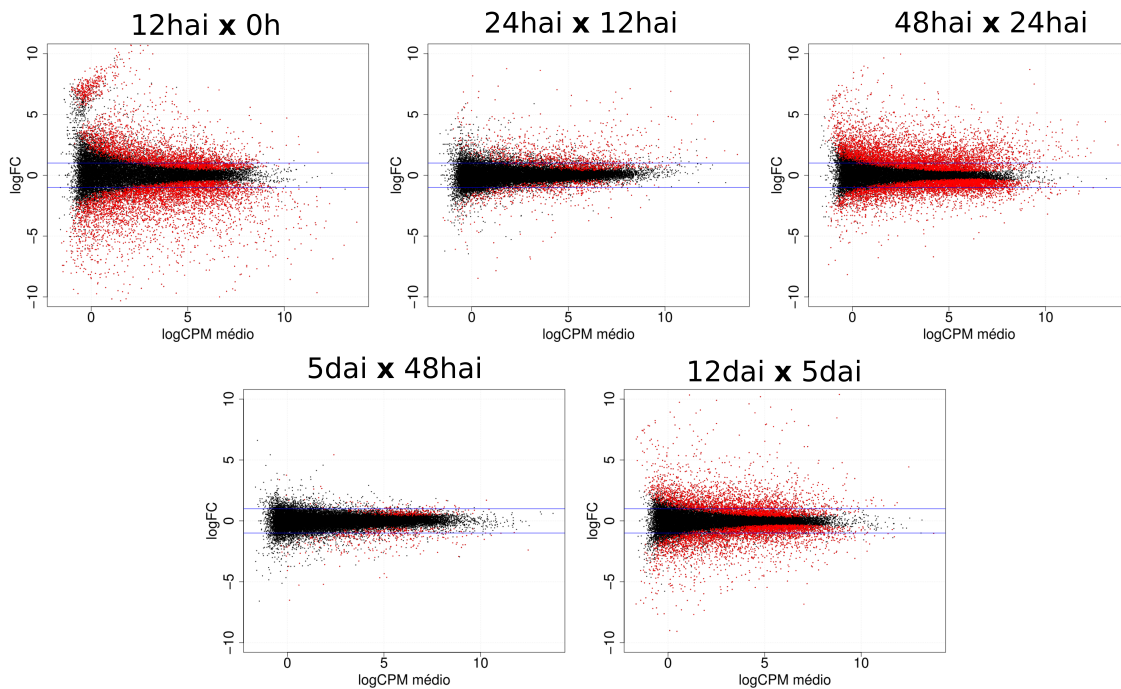


Figura 3.3. Gráficos dos valores de *Fold Change* em escala logarítmica (LFC), em função da média das contagens para cada comparação entre os tempos adjacentes. Os pontos representam cada um dos transcritos cujas contagens por milhão em pelo menos dois dos pontos temporais foi superior a um. No eixo x encontra-se a média das contagens por milhão entre dois tempos. No eixo y está o valor do LFC. As linhas horizontais azuis representam os LFC -1 e 1. Em vermelho estão destacados os transcritos diferencialmente expressos, estimulados e reprimidos.

A lignina é um dos polímeros constituintes da parede celular dos vegetais, resistente à degradação, cuja biossíntese passa por uma cadeia de transformação de compostos iniciais e intermediários por ação enzimática, claramente descrita por VANCE *et al.* (1980). Na comparação entre os tempos 0 h e 12 hai, nota-se a presença do termo GO:0009809, referente à síntese de lignina. Em particular, este termo funcional está enriquecido de genes reprimidos neste intervalo. Entre 24 hai e 48 hai, entretanto, esse termo é enriquecido de genes estimulados. Quando examinada a via dos fenilpropanoides pelos transcritos diferencialmente expressos dessa comparação, foram estimulados muitos dos elementos da via da lignina (Figura A.13 - destacados em verde), porém, ao fim, há uma peroxidase reprimida (EC1.11.1.7) impedindo a formação do polímero a partir dos cinamil álcoois (Figura A.13 - destacada em lilás). Curiosamente, apesar de não ser o único meio de defesa, a lignificação é uma resposta localizada, aumentando a quantidade desse composto na parede para formar uma barreira física. É induzida pela infecção, desde que estejam presentes os receptores para os efetores secretados por parte do patógeno. Além da falta de reconhecimento entre esses elementos, as plantas suscetíveis podem ter essa via suprimida pela ação do fungo nas matrizes de deposição da lignina (VANCE *et al.*, 1980; NICHOLSON & HAMMERSCHMIDT, 1992).

Na categoria de resposta a calor (GO:0009408), 12 dos 29 transcritos anotados com esse termo foram reprimidos 12 horas depois da infecção, sendo que 14 foram estimulados de 24 para 48 hai (Figura 3.6 - B). São as proteínas chaperonas (também conhecidas como *heat shock*) ou homólogos de genes bacterianos destas proteínas que são facilitadoras do dobramento proteico. Essas mesmas proteínas

podem ter função de ligação à *heat shock* (GO:0031072), sendo encontrados 10 transcritos estimulados correspondentes a esse tipo de atividade em 48 hai em relação à 24 hai. A Figura 3.4 mostra, para esses dois termos, uma comparação do LFC de cada tempo após a inoculação em relação a 0 h, trazendo o comportamento dos transcritos diferencialmente expressos que os compõem. Como nas 12 primeiras horas grande parte dessas proteínas possuíam níveis de expressão inalterados ou reprimidos, *P. kuehni* pode ter agido no sentido de suspender algumas dessas maquinarias de resposta. Por outro lado, a manutenção da síntese proteica, evidenciada pelos termos estimulados (Tabela A.5), pode ser essencial para o crescimento e reprodução desse fungo biotrófico, visto que depende do tecido vivo. Para beneficiar-se, é possível que ele tenha alterado o metabolismo hospedeiro (BERGER *et al.*, 2007).

Em concordância com os resultados obtidos, SWINDELL *et al.* (2007) comprovaram que diversos outros estresses não-térmicos induzem a expressão das proteínas *heat shock*. Em seus resultados dentre as famílias de proteínas *heat shock* do tecido foliar infectado de *A. thaliana*, destacaram-se proteínas das famílias Hsp70, Hsp90 e Hsp100. Com exceção da segunda, praticamente todas as demais famílias indicavam expressão diferencial por causa de estresse oxidativo artificialmente modulado. Essa última observação é interessante pois corrobora a hipótese de que os mecanismos de resposta às demais formas de perturbação ao organismo são possivelmente desencadeadas pelas ROS.

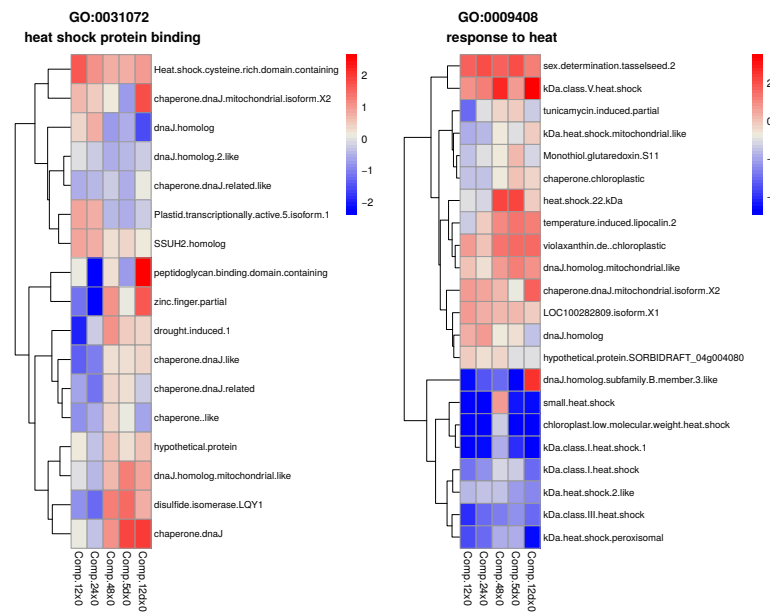


Figura 3.4. Heatmaps dos termos *response to heat* e *heat shock protein binding*. As cores representam os valores do LFC dos tempos após inoculação em relação a 0 h, os quais estão descritos no eixo *x*. Os transcritos diferencialmente expressos do termo estão alocados e descritos no eixo *y*. O dendograma agrupa esses transcritos de acordo com similaridade dos perfis de LFCs. Na barra de cores o vermelho representa valores de LFC positivos e o azul indica valores negativos.

Dado esse fato, é pertinente notar a presença de termos relativos às ROS, visto que os processos de oxirredução, ligação a íon de ferro e atividade de carreamento de elétrons foram sobre-representados nos transcritos reprimidos de 12hai em comparação a 0 h (Tabela A.5). Já em 24 vs 12 hai, eles estavam estimulados (Tabela A.6), mantendo essa característica até 48 hai (Tabela A.7).

Conforme ocorre a síntese das espécies reativas, posteriormente, são desencadeadas moléculas que atuam em sua detoxificação. Na resposta ao estresse oxidativo (Figura 3.5), grande parte dos transcritos representam sequências codificantes de peroxidases. Quatro dos transcritos associados à peroxidase foram fortemente estimulados em todos os pontos de comparação. Essas enzimas da classe oxirredutase permitem a formação de água a partir do peróxido de hidrogênio, juntamente à oxidação de algum com-

ponente. Entre todas as peroxidases, também havia algumas peroxidases de glutationa e peroxidases de ascorbato, ambas com atividade de redução de ROS ocorrente nos cloroplastos (KANGASJÄRVI *et al.*, 2012).

Na resposta ao peróxido de hidrogênio (Figura 3.5), os transcritos diferencialmente expressos foram os de catalase e de proteínas *heat shock*. As catalases também podem conduzir à formação de H_2O_2 , contudo, sua principal atividade é a catálise dessa molécula em água e oxigênio. Um dos transcritos anotado como essa classe enzimática era ativado logo a partir de 24 hai, enquanto que outro teve um pequeno estímulo em 12 dai. Em espécies resistentes do gênero *Lycopersicon*, um dos picos da atividade de peroxidase nas plantas ocorreu em 48 hai, sendo esse período concomitante ao desenvolvimento dos haustórios e, com isso, se daria o início da formação das estruturas germinativas de *Oidium neolyopersici*. Pelo que os resultados indicaram, apenas nas espécies resistentes eram observados altos níveis da peroxidase, bem como a presença da catalase (MLÍČKOVÁ *et al.*, 2004).

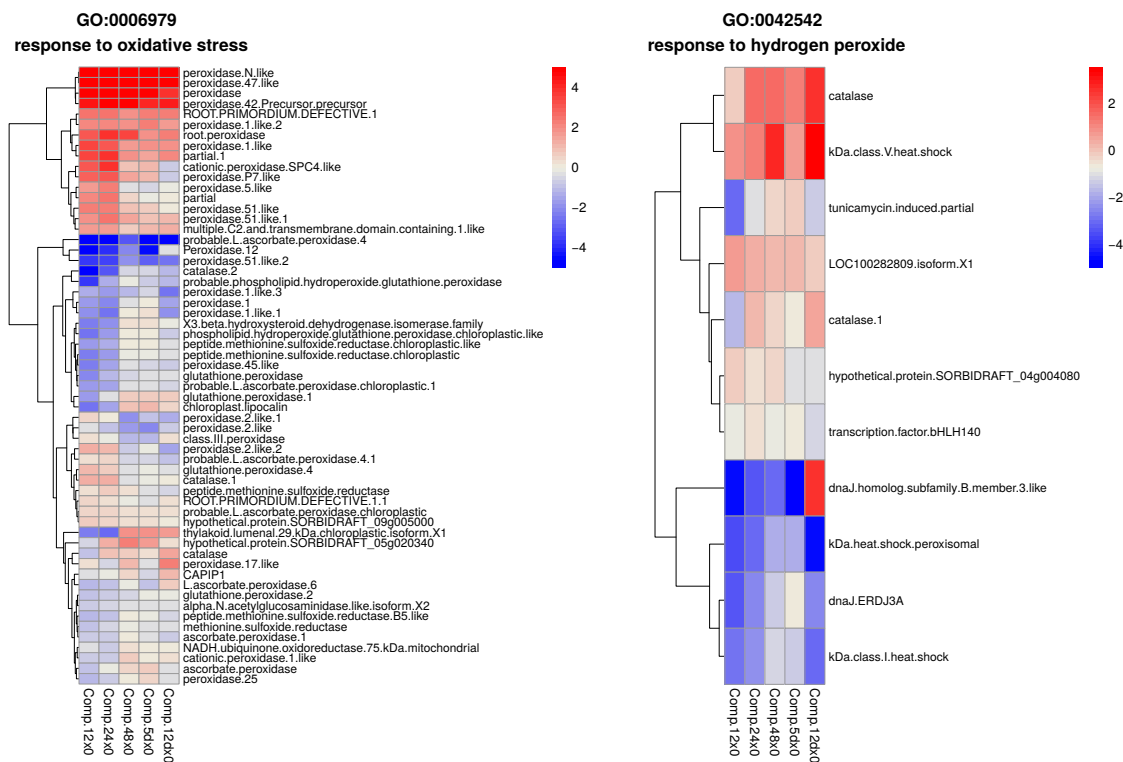


Figura 3.5. Heatmaps dos termos *response to oxidative stress* e *response to hydrogen peroxide*. As cores representam os valores do LFC dos tempos após inoculação em relação a 0 h, os quais estão descritos no eixo *x*. Os transcritos diferencialmente expressos do termo estão alocados e descritos no eixo *y*, o dendograma agrupa esses transcritos de acordo com similaridade dos LFCs dos tempos. Na barra de cores o vermelho representa valores de LFC positivos e o azul indica valores negativos.

3.4 Perfis de expressão gênica temporal

Embora cada uma dessas comparações tenha evidenciado aspectos particulares da resposta nos diversos momentos, é importante fazer uma ressalva a respeito dos termos enriquecidos detectados em cada teste. É possível perceber pelas tabelas de enriquecimento funcional do anexo III que alguns dos termos enriquecidos em transcritos com expressão reprimida de 12hai vs 0 h passaram a ser estimulados em 48 hai vs 24 hai (Tabelas A.5 e A.7). Essa inversão no sentido da expressão diferencial pode ser percebida principalmente nos perfis de expressão gênica assemelhando-se a uma onda de repressão seguida de expressão, ou vice-versa, principalmente em termos como a resposta ao estresse oxidativo, biossíntese

de lignina, fotorrespiração e fotossíntese (Figura 3.6 - A, C, E e F).

Gráficos da expressão de transcritos desejados, na forma de *heatmap*, foram elaborados no intuito de indicar variação dos demais tempos em relação ao tempo 0 h. O primeiro passo foi selecionar de acordo com proteínas que participassem do processo de infecção na formação da resposta de defesa. Foram levados em conta o LFC de todos os genes que passaram pelo filtro de contagens por milhão (cpm) e que foram considerados como transcritos diferencialmente expressos, resultantes da abordagem do tipo ANOVA e também da comparação entre os tempos adjacentes. Porém, ressalta-se que o LFC dos transcritos foi da comparação de um determinado tempo com 0 h.

O estudo das proteínas cinases mostrou-se particularmente relevante, pois eram sub-representadas (dados não incluídos no Apêndice III) entre os genes estimulados de 48 vs 24hai. Também não é notado um perfil muito claro do comportamento dos transcritos ao longo do tempo (Figuras A.7 e 3.6 - D). Isso pode ter ocorrido por haver muitos transcritos estimulados e reprimidos, o que evita enriquecimento em uma única direção. Uma segunda consideração é que grande parte dos transcritos associados a esse termo não era diferencialmente expressa: de um total de 691 transcritos, apenas 12,16% possuíam expressão diferencial. A explicação mais plausível é que se trata de um grande grupo, com diversas classes de proteínas reguladoras de muitos processos (STONE & WALKER, 1995), e que, embora cada uma regule genes como parte de um estímulo recebido, testá-las conjuntamente reduz a chance de essa classe ser significativa no enriquecimento funcional. O número de transcritos estimulados e reprimidos mostrou-se próximo para valores absolutos de LFC superiores a 0,1 nas comparações dos momentos contra 0 h.

Nos perfis de expressão da Figura 3.6 (A, B, C, E e F), é possível perceber o perfil antagônico entre os grupos de transcritos diferencialmente expressos em 12 hai vs 0 h e 48 hai vs 24 hai. Em particular, para as categorias de resposta a estresse oxidativo e cinases, em ambas essas comparações, nota-se a ocorrência simultânea de transcritos estimulados e reprimidos. Quando se atenta para a resposta ao estresse oxidativo, um número maior de transcritos é reprimido nas primeiras 12 horas, por estar presente entre os termos significativos da Tabela A.5. Todavia, passadas 48 horas da infecção, a quantidade de transcritos estimulados é praticamente equivalente à de reprimidos (Figura 3.6 - A) e, realmente, o número de transcritos estimulados (21) foi próximo ao de reprimidos (16). Em contrapartida, a resposta ao calor tem modificação mais aparente do sentido dos transcritos de 12 hai vs 0 h para 48 vs 24 hai (Figura 3.6 - B), dando indícios de que a expressão das chaperonas foi retardada por *P. kuehni*. Mesmo em relação à Figura 3.4 é visível que um dos dois grupos principais no agrupamento hierárquico de resposta a calor mostra estímulo dos transcritos, mas ainda com níveis inferiores em relação ao tempo 0 h. Portanto, o estresse levou à redução de apenas uma gama da resposta às ROS, em harmonia com a queda dos processos de oxirredução descritas nas seções anteriores. Considerando o quadro geral, pode-se supor que muitos dos processos resultem de respostas da planta, mas, como o genótipo mostrou-se suscetível, também é possível que o fungo esteja modulando essas respostas.

Entre os antioxidantes, o comportamento dos transcritos relacionados ao metabolismo da glutathione foi de divisão em quatro grupos (Figura A.8 - A). Nos grupos de transcritos estimulados, alguns mantiveram o nível de expressão durante todo o período investigado, com pequenas oscilações, enquanto que para outros, em certos momentos, o nível de expressão retornou ao basal (0 h). Houve um grupo de transcritos reprimidos nos dois primeiros tempos e um outro grupo com o estímulo mantido em quase todos os tempos. Quando a busca foi por ascorbato, a lista dos transcritos diferencialmente expressos indica dois grandes grupos, principalmente pelo comportamento das 24 horas que sucederam a infecção, sendo um deles com membros mais estimulados e outro com membros reprimidos. Em ambos os casos, houve transcritos cujos níveis de expressão retornaram aos patamares observados antes da inoculação. LOARCE *et al.* (2016) analisaram o um transcriptoma de *Avena sativa* após a inoculação com *Puccinia coronata*, no qual foram encontrados transcritos de enzimas pertencentes a processos de detoxificação oxidativa por glutathione e ascorbato. Por RT-qPCR foi possível detectar, ainda, que a glutathione S-transferase era

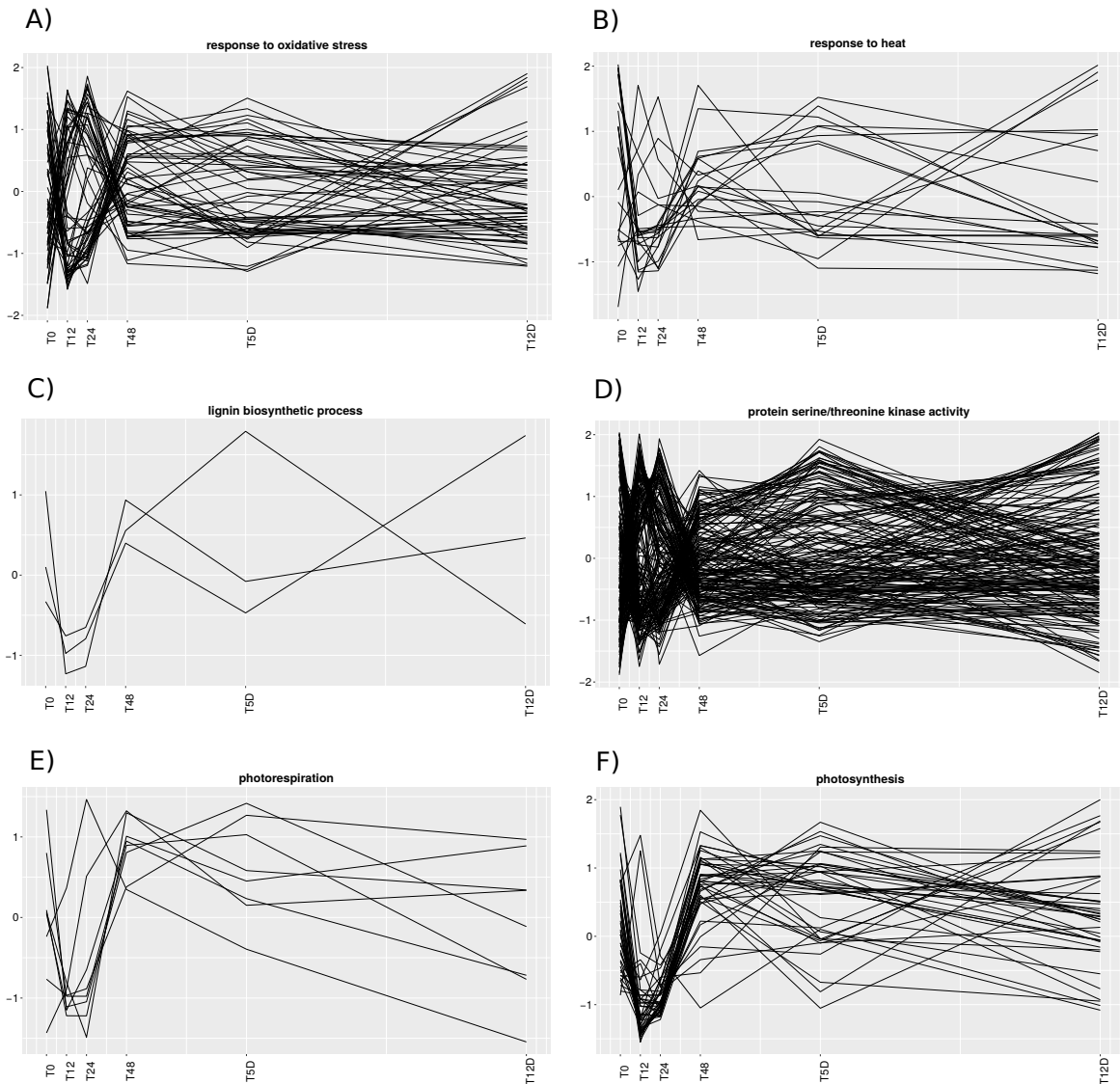


Figura 3.6. Perfis de expressão temporal dos transcritos diferencialmente expressos anotados com termos funcionais de interesse particular. Considerando apenas os genes diferencialmente expressos no teste ANOVA e entre comparações adjacentes, estão dispostos os perfis de expressão gênica dos termos *response to oxidative stress*, *response to heat*, *lignin biosynthetic process*, *protein serine/threonine kinase activity*, *photorespiration*, *photosynthesis*.

diferencialmente expressa na infecção de uma cultivar resistente em relação ao seu controle não inoculado, no tempo de 24 horas.

Como já dito, o cloroplasto é deveras importante em uma ampla gama de processos e por isso é um alvo do fungo. Conjuntamente ao processo de fotossíntese, há a ocorrência da fotorrespiração. Seu papel na defesa do hospedeiro é também fazer parte da regulação das ROS (LOARCE *et al.*, 2016). Nesse processo a enzima chave no ciclo de Calvin, a Ribulose 1,5-bifosfato carboxilase oxigenase (Rubisco), faz a acepção do oxigênio no lugar do gás carbônico (atividade de oxigenase) (KANGASJÄRVI *et al.*, 2012) e sua expressão é estimulada após 24 hai (Figura 3.7). A outra ribulose bifosfato de ação carboxilase mostrou-se reprimida até 24 hai, e novamente aos 12 dai. Essa última enzima, cujo transcrito era nomeado *comp189599_c1.seq1*, também estava inclusa no enriquecimento do processo de fotossíntese (Figura 3.8).

Explorando um pouco mais o processo de fotossíntese (Figura 3.8), nota-se que quase todos os transcritos sofreram repressão nas primeiras 24 horas, com aumento ou retorno à expressão em nível

basal de 48 hai em diante. Houve uma aglomeração dos transcritos referentes aos fotossistemas, com o comportamento notoriamente reprimido nos primeiros tempos. Como a cana-de-açúcar é uma planta de fixação de CO₂ pela via C₄, a presença do compostos enzimáticos relacionados à presença fosfoenolpiruvato era esperada, com um transcrito reprimido em praticamente todos os pontos e um outro com pequenas variações no sentido também de repressão.

Além da PAL, as cinamil álcool desidrogenases são fundamentais na via de biossíntese da lignina, nas etapas finais para formação de cinamil álcool. Pela ação do peróxido de hidrogênio, esse elemento forma radicais que serão incorporados à cadeia da lignina criando uma rede tridimensional (VANCE *et al.*, 1980; BOL *et al.*, 1990). Três delas estavam presentes no perfil de expressão do processo de biossíntese de lignina (*comp184920_c0_seq1*, *comp190954_c0_seq1*, *compt201163_c0_seq1*; Figura 3.6 - C), enquanto que na procura de outros transcritos diferencialmente expressos do transcriptoma com descrição relacionada a essas enzimas, foram encontradas outras cinco. Ao se pesquisar a anotação do transcriptoma, foi verificado que elas realmente não estavam anotadas para a biossíntese de lignina, mas a outros processos como ligação ao íon de zinco, atividade de oxirredutase e processo de oxirredução. O fato é que iniciaram a série temporal com repressão, com aumento dos níveis de expressão de 24 para 48 hai (Figura A.9).

Quitinases e β -1,3-glucanases são hidrolases que se enquadram na categoria de proteínas de resistência. Sua relevância se deve à composição de quitina das paredes celulares dos fungos, que agem como PAMP e, por isso, em plantas resistentes, é esperado o aumento da expressão dos genes de degradação desse composto nos primeiros momentos da infecção. Podem ser induzidas por ácido salicílico, etileno, entre outros compostos, e também pela própria infecção (BOL *et al.*, 1990; COLLINGE *et al.*, 1993; OLIVEIRA-GARCIA & VALENT, 2015). A maior parte das quitinases foram estimuladas, cinco delas de uma maneira mais proeminente em relação às demais (Figura A.10). Isso também ocorreu para três β -1,3-glucanases, embora vários dos transcritos indicassem repressão (Figura A.11).

Merece destaque também a presença de genes de resistência propriamente ditos, os fatores de transcrição que a eles se relacionem, como os WRKY, e a presença de domínios de repetições ricas em leucina, os quais podem caracterizar um gene de resistência. No agrupamento dos transcritos de resistência, o grupo que se encontra mais abaixo no gráfico caracteriza transcritos cuja expressão manteve-se abaixo dos níveis originais de 0 h em praticamente todos os tempos (Figura 3.9). Nele, são encontradas três NBS-LRR, uma RPS e algumas RPs, RGAs e RPPs, as quais são homólogas de proteínas de resistência. Em contrapartida, no grupo mais estimulado, havia apenas uma NBS-LRR entre diversos transcritos que originam proteínas das classes anteriormente citadas. É hipotetizado que elas possuem a capacidade de se ligar especificamente com os efetores patogênicos, com funções comprovadas na sinalização conduzindo à defesa (DEYOUNG & INNES, 2006).

Os fatores de transcrição WRKY são uma classe ampla de proteínas regulatórias relatadas no desenvolvimento de sementes, dormência e germinação, no desenvolvimento e na senescência do vegetal, agindo também nos casos de estresse biótico e abiótico. Mostram-se pertinentes no processo de infecção, sendo reguladores do processo de defesa (FANG *et al.*, 2016; LI *et al.*, 2004; RUSHTON *et al.*, 2010). Esses resultados são baseados nas observações desses fatores em *A. thaliana* (PANDEY & SOMSSICH, 2009). É importante ressaltar que o aumento de expressão de alguns dos fatores de transcrição WRKY resultam em aumento da suscetibilidade da planta. Tal comportamento ocorre porque membros dessa família agem tanto na regulação positiva, quanto negativa da resposta de defesa (RUSHTON *et al.*, 2010). Além disso, outros desses fatores a regulam dependendo do tipo de nutrição do patógeno, isto é, necrotrófico ou biotrófico.

Pela Figura 3.10, verifica-se que houve um pequeno aumento do fator WRKY4 em três das comparações, o que estaria de acordo com os dados publicados para esse fator, o qual apresenta regulação negativa da resistência, isto é, poderia causar um aumento da suscetibilidade aos fungos biotróficos (PANDEY & SOMSSICH, 2009). No mesmo caso, os WRKY33, que também apresentam essa característica,

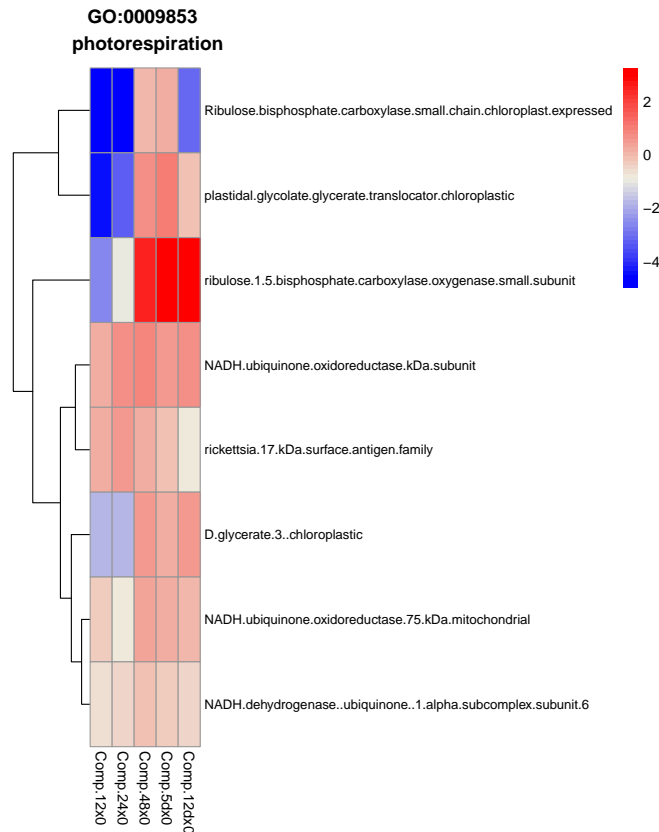


Figura 3.7. *Heatmap* dos transcritos diferencialmente expressos que compõem o termo fotorrespiração. Na barra de cores o vermelho representa valores de LFC positivos e o azul indica valores negativos.

possuíam dois membros muito estimulados, indicando que podem ter sido modulados para favorecer o patógeno. Em contrapartida, em situação infecção viral esse fator pode ser favorecido em casos de plantas resistentes, porque em mamão FANG *et al.* (2016) descobriram um gene *WRKY DNA-binding protein 33* relacionado ao ácido abscísico que estava estimulado no cultivar resistente.

Também de acordo com a bibliografia está o estímulo em todos os tempos após a infecção (Figura 3.10) de um regulador negativo WRKY40 (PANDEY & SOMSSICH, 2009). O WRKY40 é induzido por um brassinosteróide BZR1 e age para que a PTI não ocorra. Na transdução de sinal através de hormônios (Figura A.14), a biossíntese dos brassinosteróides tem um membro proteico estimulado, a fosfatase BSU1, cujo aumento permite a regulação negativa do repressor de BZR1, favorecendo a desfosforilação dessa proteína (HUOT *et al.*, 2014; GENDRON & WANG, 2007). De forma contrária, de 24 para 48 hai, BSU1 passou a ser reprimido, desfavorecendo a elongação e a divisão celular. Por isso, deve ser levado em consideração o cenário de que o fungo possa ter modulado membros dessa via no sentido de estimulá-la, de modo que cessão do estímulo em um tempo posterior resultaria na retomada dos níveis de transcrição de 0 h.

É possível que, devido à suscetibilidade, o crescimento vegetal tenha sido favorecido em detrimento àqueles processos necessários à resposta de defesa. De fato, a fim de que sejam produzidos os compostos de defesa, há demanda de recursos celulares, fazendo que estes tornem-se menos disponíveis para outros processos comuns às células. Isso causa um redirecionamento de toda a maquinaria celular para os processos imprescindíveis em um determinado momento. Essa compensação (*trade-off*) se faz necessária para priorizar a necessidade do vegetal conforme as condições ambientais a que está sujeito. Faz parte do mecanismo de defesa das plantas que algumas vias de formação dos hormônios de crescimento sejam reprimidas, como foram relatadas evidências da inibição da via da auxina, queda dos

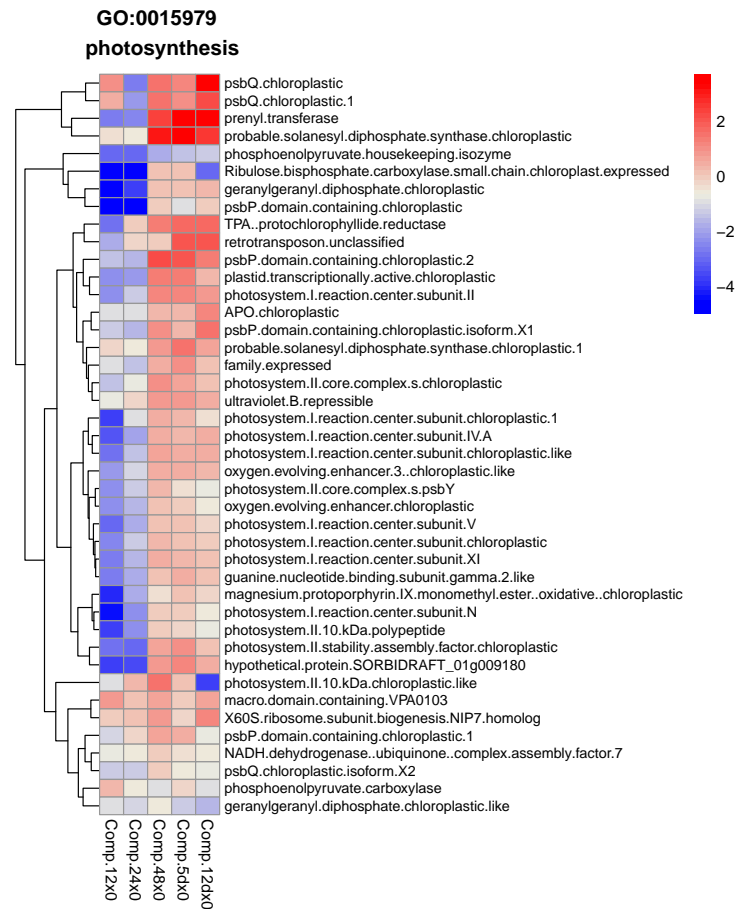


Figura 3.8. *Heatmap* dos transcritos diferencialmente expressos que compõem o termo fotossíntese. Na barra de cores o vermelho representa valores de LFC positivos e o azul indica valores negativos.

brassinosteróides e giberelinas (HUOT *et al.*, 2014).

Entre os transcritos anotados, um WRKY70 mostrou-se fortemente estimulado em todos os tempos em relação ao nível basal (Figura 3.10). O WRKY70 tem sua expressão aumentada para governar a oscilação dos fitormônios da resposta sistêmica adquirida, positivamente o ácido salicílico e negativamente o ácido jasmônico (LI *et al.*, 2004). Curiosamente, LOARCE *et al.* (2016) não encontraram domínios WRKY em sua biblioteca de hibridização subtrativa por supressão de uma cultivar de *Avena sativa* parcialmente resistente à ferrugem da folha.

Dado que há um balanço entre o crescimento e os mecanismos de defesa, os principais hormônios desses processos podem ser investigados. Quando se buscou por *auxin* no transcriptoma, a partir dos transcritos diferencialmente expressos é notável a presença de um grupo caracterizado pelo estímulo em praticamente todos os tempos (Figura A.15). Já ao se pesquisar *gibberellin*, a maioria dos transcritos anotados como seus receptores estava em níveis inferiores ao de 0 h (Figura A.16). Para o ácido salicílico, embora dois transcritos diferencialmente expressos estivessem ativados até as 24h (Figura A.17), não foi encontrada uma resposta a esse hormônio nas análises de enriquecimento funcional, o que pode ser justificado por poucos transcritos anotados com termos relativos a esse hormônio. Quando investigados os receptores NPR (*nonexpressor of pathogenesis related proteins*), que agem como reguladores da transcrição (HUOT *et al.*, 2014), notou-se que dois NPR3 estavam reprimidos (Figura A.18). Logo, a defesa mediada pelo ácido salicílico pode ter sido evitada pela supressão da expressão, por parte do patógeno, dos genes que originam o hormônio e aqueles que transmitem seu sinal (Figura A.14). De maneira similar, o estímulo da fosfatase PP2C (*protein phosphatase 2C*) (Figura A.14) pode ser um indício da repressão da

via do ácido abscísico em 12 hai. SCHAKER *et al.* (2016) relataram a expressão diferencial estimulada de auxina em 200 dai de uma cultivar de cana com resistência intermediária ao carvão, enquanto que havia repressão de resposta ao ácido salicílico, do ácido jasmônico e abscísico.

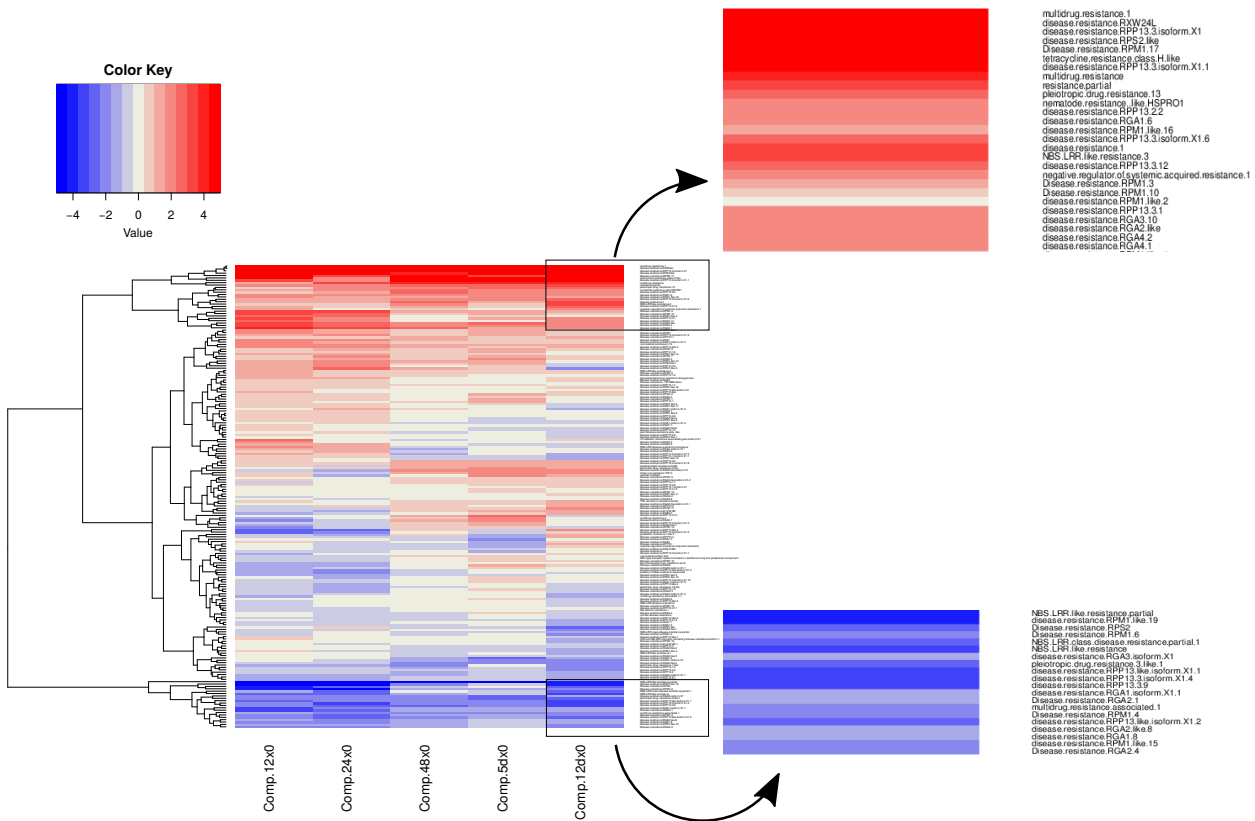


Figura 3.9. No *heatmap* dos transcritos diferencialmente expressos anotados como de resistência. Os dois quadros em destaque permitem verificar aqueles transcritos mais estimulados (acima) e os mais reprimidos (abaixo).

Relacionadas à PTI estão as cascatas de proteínas cinase do tipo MAP, cuja ativação se dá pelas ROS e que têm a capacidade de estimular essas espécies de oxigênio (PANDEY & SOMSSICH, 2009; MITTLER *et al.*, 2011; RAY *et al.*, 2012). Isso poderia ser explicado porque os cloroplastos são alvos de alguns tipos dessas enzimas (KANGASJÄRVI *et al.*, 2012). Algumas das proteínas WRKY são reguladas diretamente pela ligação das MAP cinases, como o WRKY33 governado pela ligação da MAP cinase MPK4 (PANDEY & SOMSSICH, 2009; RUSHTON *et al.*, 2010). Pela análise de vias KO todas as sete MAP cinases (MPK4, MPK3, MPK6, MPK1/2, MPK7/14, MKP2 e MPK8), elencadas em nove processos, foram reprimidas em 12 hai (Figura 3.11). Embora a repressão da MPK4 possa nortear indiretamente o aumento da resposta de defesa e morte celular, a queda de MPK3 e MPK6 pode estar diretamente envolvida na não ocorrência das respostas de defesa (precoce e tardia) e indiretamente no impedimento da morte celular e restrição à formação de etileno. Outra discrepância da redução da MPK4, em relação às outras duas é a possibilidade de atuar no aumento do acúmulo de ROS. As MPKs atuam em mais de uma via e, por vezes, em sentidos antagônicos. Um caso como este é a queda de MPK6, a qual na via do ácido jasmônico pode colaborar no estímulo à resposta a ferimentos, enquanto que na via do ácido abscísico demonstra atuar no sentido de repressão. Há indícios de resposta de defesa pelo aumento da quitinase (ChiB), o que seria esperado por algumas dessas enzimas serem diferencialmente expressas no sentido de estímulo (Figura A.10).

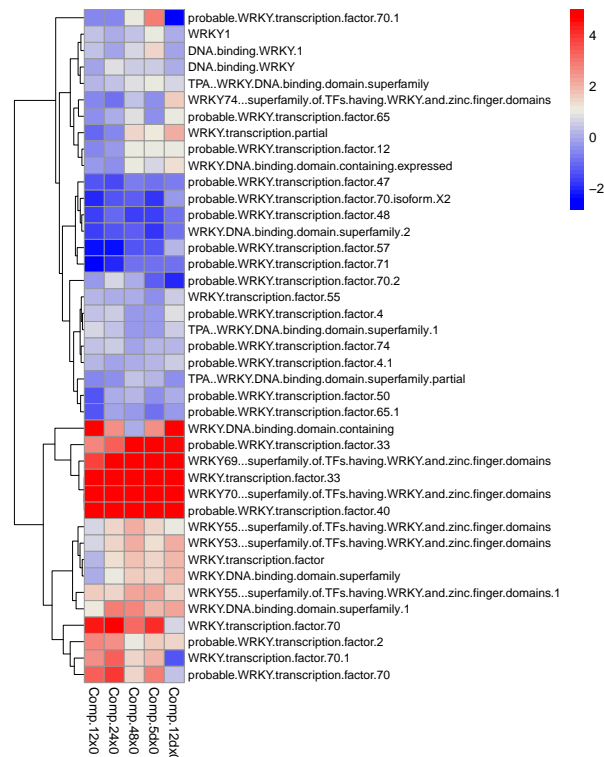


Figura 3.10. *Heatmap* dos transcritos anotados como fatores de transcrição WRKY ou contendo esse domínio. Na barra de cores o vermelho representa valores de LFC positivos e o azul indica valores negativos.

3.5 Perfis de expressão gênica sob o ponto de vista da infecção de cana-de-açúcar por *Puccinia kuehni*

OLIVEIRA-GARCIA & VALENT (2015) ressaltam o valor de ensaios de transcriptomas para corroborar, identificar e descobrir novos mecanismos de ação dos fungos fitopatogênicos para driblar os recursos de defesa da planta hospedeira. Os processos ligados à ocorrência de mecanismos de defesa na planta foram discutidos nos itens anteriores. Uma apreciação importante de LOARCE *et al.* (2016) é que os autores buscaram amostrar os tempos iniciais da infecção, considerando que de 12 a 24 hai ocorreria a pré-infecção com posterior formação do apressório. Em 48 hai, com a penetração efetiva, foi considerado que as vesículas e haustórios do fungo já tinham se desenvolvido. Ao se analisar 72 hai, já era esperado encontrar um certo acúmulo de compostos de defesa devido à sinalização para interromper a ação do patógeno.

GOMEZ (2013) visaram compreender os parâmetros epidemiológicos de *P. kuehni*, determinando os momentos de incubação, latência e esporulação. Nas plantas inoculadas do cultivar SP89-1115 a 25°C os sintomas foliares eram visíveis no 9º dia e, passados dois dias, a maior parte das pústulas tinham esporulado. No experimento de campo, a quantidade de pústulas sobre a superfície foliar foi menor (2,9 pústulas/cm²) do que com o inóculo controlado. Quanto ao comportamento do fungo, além do desenvolvimento de esporos, a formação de outras estruturas para a caracterização do patógeno também foi confirmada pelo uso da microscopia. Essas observações verificaram que os esporos germinaram e a infecção foi efetiva, porque em 24 hai os apressórios tinham se desenvolvido sobre as cavidades estomatais.

Neste trabalho, os transcritos mostraram que alguns processos relacionados com a defesa das plantas foram superados pelo patógeno. Houve um comportamento de intensa repressão nas primeiras 12 horas após a infecção. Em alguns dos casos, entre 12 e 24 horas, foi observado um estímulo não tão

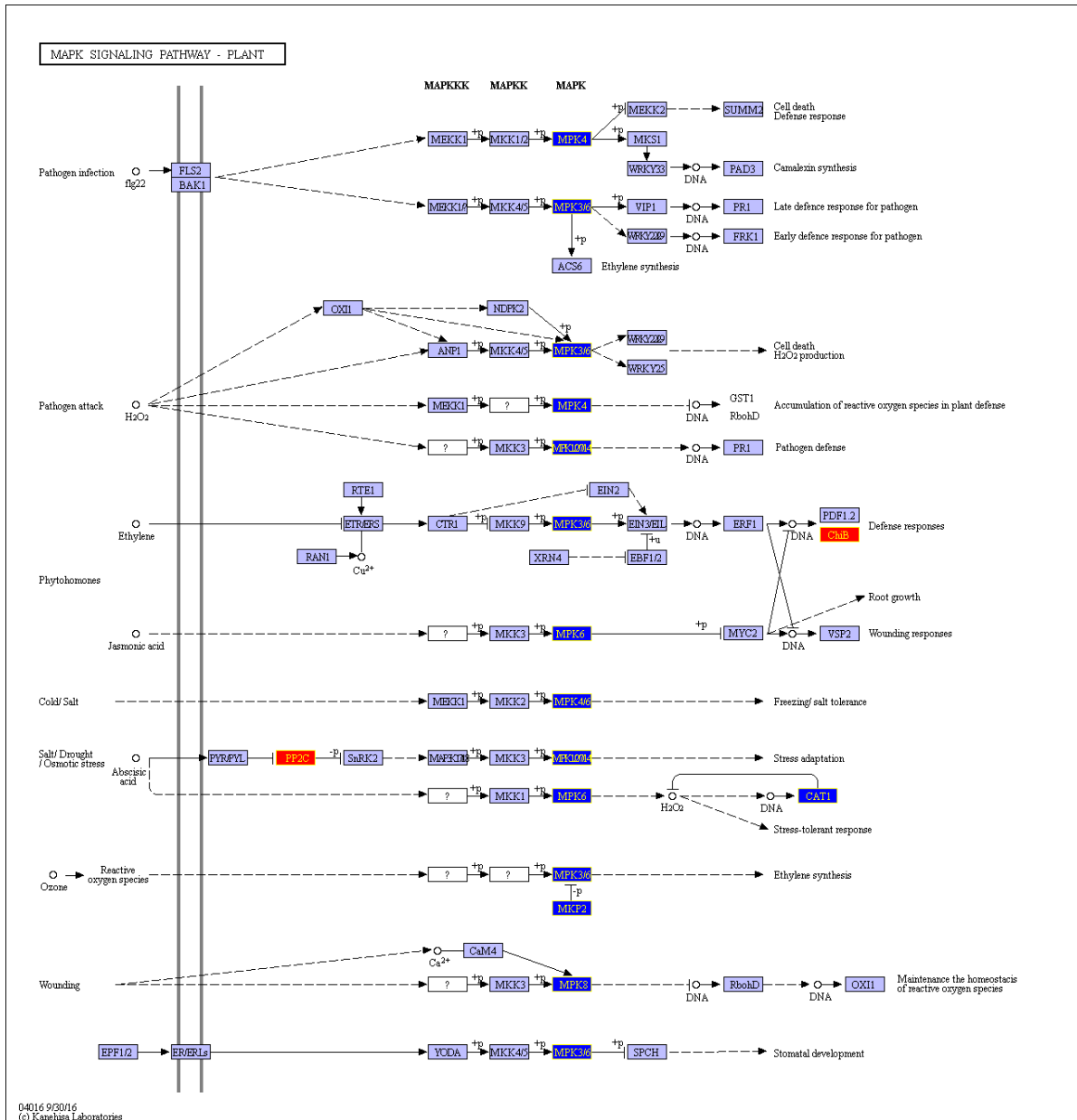


Figura 3.11. Via de sinalização através das cinases do tipo MAP em plantas na comparação 12 hai vs 0 h. Em azul estão as proteínas associadas aos transcritos reprimidos e em vermelho aqueles relativos aos estimulados.

proeminente da expressão (Figura 3.6 e Tabela A.6), que veio a ocorrer somente em 48 hai. Embora a comparação entre os dois últimos tempos tenha revelado transcritos diferencialmente expressos, os perfis notavelmente apontam os primeiros momentos após a inoculação como cruciais para a o estabelecimento da ferrugem.

Quando a expressão diferencial foi analisada entre cada ponto de amostragem e o tempo 0 h, observou-se que o menor número de transcritos diferencialmente expressos ocorreu às 48 hai (Tabela 3.1). Portanto, há maior similaridade entre esses dois momentos e isso se evidencia, também, pelos padrões invertidos dos perfis de expressão, isto é, a queda dos níveis de expressão em 12 hai e ativação de 24 para 48 hai. É concebível supor que *P. kuehni* tenha agido no sentido de reprimir muitas das vias de defesa e de sinalização durante o período de estabelecimento das estruturas de infecção, sendo um fungo efetivo tanto em inativar a PTI quanto em garantir a ETS (*effector-triggered susceptibility*) no cultivar SP89-1115.

Nesse primeiro ponto temporal após a infecção, alguns processos essenciais foram prejudicados, como a fotossíntese e aqueles a ela relacionados, como a ação do cloroplasto e suas estruturas. O restabelecimento das vias reprimidas, notada pelo estímulo dos transcritos em 48 hai e estabilização dos níveis até 5 dai, permite supor que: *i*) a planta respondeu tardiamente à infecção por ferrugem alaranjada; *ii*) o fungo já teve sucesso na infecção e não precisa mais influenciar tão fortemente o metabolismo da planta.

4 CONCLUSÃO

A análise de expressão diferencial a partir da avaliação dos diferentes momentos da infecção revelaram que a influência de *P. kuehni* sobre a cana-de-açúcar acarretou modificações substanciais nos níveis de expressão dos transcritos. A partir do teste conjunto, foi possível perceber quais os principais processos influenciados ao longo do desenvolvimento da doença. Além disso, a estratégia de comparação da expressão diferencial entre tempos adjacentes possibilitou encontrar as alterações dos componentes moleculares em vias metabólicas específicas. Algumas dessas modificações evidenciaram-se exclusivamente em dados momentos, enquanto que outras exibiram padrões antagônicos em tempos distintos.

Com base nos perfis de expressão temporal, pode-se observar que as mudanças metabólicas mais drásticas ocorreram principalmente nos tempos iniciais após a inoculação. O nível de expressão de diversos transcritos retornou aos níveis basais após o provável estabelecimento do fungo, o que potencialmente reflete menor influência do patógeno no metabolismo vegetal.

Por ser um cultivar suscetível, SP89-1115 pode ter sido desfavorecido em mecanismos como a sinalização através das espécies reativas de oxigênio e em barreiras estruturais como, por exemplo, a deposição de lignina na parede celular.

REFERÊNCIAS

- ANDERS, S. & W. HUBER, 2010 Differential expression analysis for sequence count data. *Genome Biology* **11**: R106.
- ANDREWS, S., 2010 FastQC: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc> .
- BAR-JOSEPH, Z., A. GITTER, & I. SIMON, 2012 Studying and modelling dynamic biological processes using time-series gene expression data. *Nature Reviews Genetics* **13**: 552–564.
- BARBASSO, D., H. JORDÃO, W. MACCHERONI, J. BOLDINI, J. BRESSIANI, & A. SANGUINO, 2010 First Report of *Puccinia kuehnii* , Causal Agent of Orange Rust of Sugarcane, in Brazil. *Plant Disease* **94**: 1170–1170.
- BELKHADIR, Y., R. SUBRAMANIAM, & J. L. DANGL, 2004 Plant disease resistance protein signaling: NBS-LRR proteins and their partners. *Current Opinion in Plant Biology* **7**: 391–399.
- BENJAMINI, Y. & Y. HOCHBERG, 1995 Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**: 289 – 300.
- BERGER, S., A. K. SINHA, & T. ROITSCH, 2007 Plant physiology meets phytopathology: plant primary metabolism and plant pathogen interactions. *Journal of Experimental Botany* **58**: 4019–4026.
- BERKMAN, P. J., R. E. CASU, J. STILLER, A. L. RAE, A. L. R. A. E. AITKEN, & K. S, 2013 Towards the sugarcane genome sequence and an understanding of polyploidy. *Proceedings of the International Society of Sugar Cane Technologists* **28**.
- BOL, J. F., H. J. M. LINTHORST, & B. J. C. CORNELISSEN, 1990 Plant Pathogenesis-Related Proteins Induced by Virus-Infection. *Annual review of phytopathology* **28**: 113–138.
- BRAITHWAITE, K. S., B. J. CROFT, R. C. MAGAREY, & T. SCHARASCHKIN, 2009 Phylogenetic placement of the sugarcane orange rust pathogen *Puccinia kuehnii* in a historical and regional context. *Australasian Plant Pathology* **38**: 380–388.
- BRIGGS, G. C., Z. NAKHID, A. T. ALLEYNE, J. AYATS, J. O. DESPRADEL, & W. ELIBOX, 2014 First Report of Orange Rust Disease of Sugarcane in the Dominican Republic. *Plant Disease* **98**: 1010–1010.
- BURNQUIST, W. L. & J. VEGA, 1995 Sugarcane diseases in southern Brazil: A brief report. Em *Sugarcane Germplasm Conservation and Exchange*, editado por B. J. Croft, C. M. Piggitt, E. S. Wallis, & D. M. Hogarth, pp. 59–61, Canberra: ACIAR Proceedings.
- CADAVID, M., J. C. ÁNGEL, & J. I. VICTORIA, 2012 First Report of Orange Rust of Sugarcane Caused by *Puccinia kuehnii* in Colombia. *Plant Disease* **96**: 143–143.
- CARDOSO-SILVA, C. B., E. A. COSTA, M. C. MANCINI, T. W. A. BALSALOBRE, L. E. COSTA CANESIN, L. R. PINTO, M. S. CARNEIRO, A. A. F. GARCIA, A. P. DE SOUZA, & R. VICENTINI, 2014 De novo assembly and transcriptome analysis of contrasting sugarcane varieties. *PLoS ONE* **9**: e88462.
- CARSON, D., B. HUCKETT, F. BOTHA, & J. VAN STADEN, 2002 Differential gene expression in sugarcane leaf and internodal tissues of varying maturity. *South African Journal of Botany* **68**: 434–442.

- CASU, R. E., J. M. JARMEY, G. D. BONNETT, & J. M. MANNERS, 2007 Identification of transcripts associated with cell wall metabolism and development in the stem of sugarcane by Affymetrix GeneChip Sugarcane Genome Array expression profiling. *Functional & Integrative Genomics* **7**: 153–167.
- CHAPOLA, R. G., H. P. HOFFMANN, & N. S. MASSOLA, 2016 Reaction of sugarcane varieties to orange rust (*Puccinia kuehni*) and methods for rapid identification of resistant genotypes. *Tropical Plant Pathology* **41**: 139–146.
- CHEAVEGATTI-GIANOTTO, A., H. M. C. DE ABREU, P. ARRUDA, J. C. BESPALHOK FILHO, W. L. BURNQUIST, S. CRESTE, L. DI CIERO, J. A. FERRO, A. V. DE OLIVEIRA FIGUEIRA, T. DE SOUSA FILGUEIRAS, M. D. F. GROSSI-DE SÁ, E. C. GUZZO, H. P. HOFFMANN, M. G. DE ANDRADE LANDELL, N. MACEDO, S. MATSUOKA, F. DE CASTRO REINACH, E. ROMANO, W. J. DA SILVA, M. DE CASTRO SILVA FILHO, & E. CÉSAR ULIAN, 2011 Sugarcane (*Saccharum X officinarum*): A Reference Study for the Regulation of Genetically Modified Cultivars in Brazil. *Tropical Plant Biology* **4**: 62–89.
- CHEN, Y., D. J. MCCARTHY, M. RITCHIE, M. D. ROBINSON, & G. K. SMYTH, 2016 edgeR: diferencial expression analysis of digital gene expression data.
- COLLINGE, D., K. KRAGH, J. MIKKELSEN, K. NIELSEN, U. RASMUSSEN, & K. VAD, 1993 Plant chitinases. *The Plant Journal* **3**: 31–40.
- CONAB, 2016a Acompanhamento da Safra Brasileira de Cana-de-Açúcar – Quarto Levantamento da safra 2015/16. *Companhia Nacional de Abastecimento* **2**: 76.
- CONAB, 2016b Acompanhamento da Safra Brasileira de Cana-de-Açúcar – Segundo Levantamento da safra 2016/17. *Companhia Nacional de Abastecimento* **3**: 72.
- CONESA, A., S. GÖTZ, J. M. GARCÍA-GÓMEZ, J. TEROL, M. TALÓN, & M. ROBLES, 2005 Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**: 3674–3676.
- DANGL, J. L. & J. D. G. JONES, 2001 Defence Responses To Infection. *Nature* **411**: 826–833.
- DANIELS, J. & B. T. ROACH, 1987 Chapter 2 - Taxonomy and Evolution. Em *Sugarcane Improvement through Breeding*, editado por D. J. Heinz, volume 11 de *Developments in Crop Science*, pp. 7–84, Elsevier.
- DANTAS, G. A., L. F. L. LEGEY, & A. MAZZONE, 2013 Energy from sugarcane bagasse in Brazil: An assessment of the productivity and cost of different technological routes. *Renewable and Sustainable Energy Reviews* **21**: 356–364.
- DE KLERK, E., J. T. DEN DUNNEN, & P. A. C. 'T HOEN, 2014 RNA sequencing: from tag-based profiling to resolving complete transcript structure. *Cellular and Molecular Life Sciences* **71**: 3537–3551.
- DE SETTA, N., C. MONTEIRO-VITORELLO, C. METCALFE, G. M. CRUZ, L. DEL BEM, R. VICENTINI, F. T. NOGUEIRA, R. CAMPOS, S. NUNES, P. C. TURRINI, A. VIEIRA, E. OCHOA CRUZ, T. C. CORRÊA, C. HOTTA, A. DE MELLO VARANI, S. VAUTRIN, A. DA TRINDADE, M. DE MENDONÇA VILELA, C. LEMBKE, P. SATO, R. DE ANDRADE, M. NISHIYAMA, C. CARDOSO-SILVA, K. SCORTECCI, A. A. GARCIA, M. CARNEIRO, C. KIM, A. H. PATERSON, H. BERGÈS, A. D'HONT, A. DE SOUZA, G. SOUZA, M. VINCENTZ, J. KITAJIMA, & M.-A. VAN SLUYS, 2014 Building the sugarcane genome for biotechnology and identifying evolutionary trends. *BMC Genomics* **15**: 540.

- DEYOUNG, B. J. & R. W. INNES, 2006 Plant NBS-LRR proteins in pathogen sensing and host defense. *Nature Immunology* **7**: 1243–1249.
- DIXON, L. J., L. A. CASTLEBURY, M. C. AIME, N. C. GLYNN, & J. C. COMSTOCK, 2010 Phylogenetic relationships of sugarcane rust fungi. *Mycological Progress* **9**: 459–468.
- DOHM, J. C., C. LOTTAZ, T. BORODINA, & H. HIMMELBAUER, 2008 Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Research* **36**.
- DUPLESSIS, S., C. A. CUOMO, Y.-C. LIN, A. AERTS, E. TISSERANT, C. VENEULT-FOURREY, D. L. JOLY, S. HACQUARD, J. AMSELEM, B. L. CANTAREL, R. CHIU, P. M. COUTINHO, N. FEAU, M. FIELD, P. FREY, E. GELHAYE, J. GOLDBERG, M. G. GRABHERR, C. D. KODIRA, A. KOHLER, U. KUES, E. A. LINDQUIST, S. M. LUCAS, R. MAGO, E. MAUCELI, E. MORIN, C. MURAT, J. L. PANGILINAN, R. PARK, M. PEARSON, H. QUESNEVILLE, N. ROUHIER, S. SAKTHIKUMAR, A. A. SALAMOV, J. SCHMUTZ, B. SELLES, H. SHAPIRO, P. TANGUAY, G. A. TUSKAN, B. HENRISSAT, Y. VAN DE PEER, P. ROUZE, J. G. ELLIS, P. N. DODDS, J. E. SCHEIN, S. ZHONG, R. C. HAMELIN, I. V. GRIGORIEV, L. J. SZABO, & F. MARTIN, 2011 Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *Proceedings of the National Academy of Sciences* **108**: 9166–9171.
- FANG, J., A. LIN, W. QIU, H. CAI, M. UMAR, R. CHEN, & R. MING, 2016 Transcriptome Profiling Revealed Stress-Induced and Disease Resistance Genes Up-Regulated in PRSV Resistant Transgenic Papaya. *Frontiers in Plant Science* **7**: 1–20.
- FERNANDEZ, D., E. TISSERANT, P. TALHINHAS, H. AZINHEIRA, A. VIEIRA, A. S. PETITOT, A. LOUREIRO, J. POULAIN, C. DA SILVA, M. DO CÉU SILVA, & S. DUPLESSIS, 2012 454-pyrosequencing of *Coffea arabica* leaves infected by the rust fungus *Hemileia vastatrix* reveals in planta-expressed pathogen-secreted proteins and plant functions in a late compatible plant-rust interaction. *Molecular Plant Pathology* **13**: 17–37.
- FUNES, C., S. G. PÉREZ GÓMEZ, D. D. HENRIQUEZ, V. DI PAULI, R. P. BERTANI, D. P. FONTANA, A. M. RAGO, C. M. JOYA, R. A. SOPENA, V. GONZÁLEZ, H. BABI, L. E. ERAZZU, M. I. CUENYA, & L. D. PLOPER, 2016 First Report of Orange Rust of Sugarcane Caused by *Puccinia kuehnii* in Argentina. *Plant Disease* **100**: 861.
- GADJEV, I., I. GADJEV, S. VANDERAUWERA, S. VANDERAUWERA, T. S. GECHEV, T. S. GECHEV, C. LALOI, C. LALOI, I. N. MINKOV, I. N. MINKOV, R. MITTLER, R. MITTLER, F. V. BREUSEGEM, F. V. BREUSEGEM, V. SHULAEV, V. SHULAEV, K. APEL, K. APEL, D. INZE, D. INZE, G. B. SCIENCES, G. B. SCIENCES, P. GENETICS, & P. GENETICS, 2006 Transcriptomic Footprints Disclose Specificity of Reactive Oxygen Species Signaling in *Arabidopsis*. *Society* **141**: 436–445.
- GALLO, J. R., 1962 Amostragem foliar em cana-de-açúcar para fins de análise foliar. *Bragantia* **21**: 899–921.
- GARCÉS, F. F., F. F. FIALLOS, E. SILVA, F. MARTINEZ, M. C. AIME, J. C. COMSTOCK, N. C. GLYNN, & L. A. CASTLEBURY, 2014 First Report of Orange Rust of Sugarcane Caused by *Puccinia kuehnii* in Ecuador. *Plant Disease* **98**: 842–842.
- GENDRON, J. M. & Z.-Y. WANG, 2007 Multiple mechanisms modulate brassinosteroid signaling. *Current Opinion in Plant Biology* **10**: 436–441.
- GOMEZ, S. G. P., 2013 *Quantificação de parâmetros monocíclicos da ferrugem alaranjada (Puccinia kuehnii) em cana-de-açúcar*. Dissertação de Mestrado, Universidade de São Paulo, Piracicaba.

- GRATIVOL, C., M. REGULSKI, M. BERTALAN, W. R. MCCOMBIE, F. R. DA SILVA, A. ZERLOTINI NETO, R. VICENTINI, L. FARINELLI, A. S. HEMERLY, R. A. MARTIENSSEN, & P. C. G. FERREIRA, 2014 Sugarcane genome sequencing by methylation filtration provides tools for genomic research in the genus *Saccharum*. *The Plant Journal* **79**: 162–172.
- GREENBERG, J. T. & N. YAO, 2004 The role of regulation of programmed cell death in plant-pathogen interactions. *Cellular Microbiology* **6**: 201–211.
- GRISHAM, M. P., J. W. HOY, J. S. HAUDENSHIELD, & G. L. HARTMAN, 2013 First Report of Orange Rust Caused by *Puccinia kuehnii* in Sugarcane in Louisiana. *Plant Disease* **97**: 426–426.
- GRIVET, L. & P. ARRUDA, 2002 Sugarcane genomics: Depicting the complex genome of an important tropical crop. *Current Opinion in Plant Biology* **5**: 122–127.
- HANSEN, K. D., S. E. BRENNER, & S. DUDOIT, 2010 Biases in Illumina transcriptome sequencing caused by random hexamer priming. *Nucleic Acids Research* **38**: e131–e131.
- HUOT, B., J. YAO, B. L. MONTGOMERY, & S. Y. HE, 2014 Growth-defense tradeoffs in plants: A balancing act to optimize fitness. *Molecular Plant* **7**: 1267–1287.
- INZÉ, D. & M. V. MONTAGU, 1995 Oxidative stress in plants. *Current Opinion in Biotechnology* **6**: 153–158.
- JACKSON, P. A., 2005 Breeding for improved sugar content in sugarcane. *Field Crops Research* **92**: 277–290.
- JOHNSON, D. C., D. R. DEAN, A. D. SMITH, & M. K. JOHNSON, 2005 Structure, Function, and Formation of Biological Iron-Sulfur Clusters. *Annu. Rev. Biochem* **74**: 247–281.
- JOHNSON, R., 1992 Past, present and future opportunities in breeding for disease resistance, with examples from wheat. *Euphytica* **63**: 3–22.
- JONES, D. A. & D. TAKEMOTO, 2004 Plant innate immunity - Direct and indirect recognition of general and specific pathogen-associated molecules. *Current Opinion in Immunology* **16**: 48–62.
- JONES, J. D. G. & J. L. DANGL, 2006 The plant immune system. *Nature* **444**: 323–329.
- KANGASJÄRVI, S., J. NEUKERMANS, S. LI, E. M. ARO, & G. NOCTOR, 2012 Photosynthesis, photorespiration, and light signalling in defence responses. *Journal of Experimental Botany* **63**: 1619–1636.
- KEEN, N. T., 1990 Gene-for-gene complementarity in plant-pathogen interactions. *Annual review of genetics* **24**: 447–463.
- KIM, C., X. WANG, T.-H. LEE, K. JAKOB, G.-J. LEE, & A. H. PATERSON, 2014 Comparative Analysis of *Miscanthus* and *Saccharum* Reveals a Shared Whole-Genome Duplication but Different Evolutionary Fates. *The Plant Cell* **26**: 2420–2429.
- KIM, D., B. LANGMEAD, & S. L. SALZBERG, 2015 HISAT: a fast spliced aligner with low memory requirements. *Nature methods* **12**: 357–360.
- KIM, D., G. PERTEA, C. TRAPNELL, H. PIMENTEL, R. KELLEY, & S. L. SALZBERG, 2013 TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* **14**: R36.
- KNOGGE, W., 1996 Fungal Infection of Plants. *The Plant cell* **8**: 1711–1722.

- KOPYLOVA, E., L. NOÉ, & H. TOUZET, 2012 SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**: 3211–3217.
- LANGMEAD, B. & S. L. SALZBERG, 2012 Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357–359.
- LAWTON, M. A. & C. J. LAMB, 1987 Transcriptional activation of plant defense genes by fungal elicitor, wounding, and infection. *Molecular and cellular biology* **7**: 335–341.
- LEHMANN, S., M. SERRANO, F. L'HARIDON, S. E. TJAMOS, & J.-P. METRAUX, 2015 Reactive oxygen species and plant resistance to fungal pathogens. *Phytochemistry* **112**: 54–62.
- LI, J., G. BRADER, & E. T. PALVA, 2004 The WRKY70 Transcription Factor: A Node of Convergence for Jasmonate-Mediated and Salicylate-Mediated Signals in Plant Defense. *The Plant Cell* **16**: 319–331.
- LIU, L., Y. LI, S. LI, N. HU, Y. HE, R. PONG, D. LIN, L. LU, & M. LAW, 2012 Comparison of Next-Generation Sequencing Systems. *Journal of Biomedicine and Biotechnology* **2012**: 1–11.
- LIU, Y., J. ZHOU, & K. P. WHITE, 2014 RNA-seq differential expression studies: More sequence or more replication? *Bioinformatics* **30**: 301–304.
- LOARCE, Y., E. NAVAS, C. PANIAGUA, A. FOMINAYA, J. L. MANJÓN, & E. FERRER, 2016 Identification of Genes in a Partially Resistant Genotype of *Avena sativa* Expressed in Response to *Puccinia coronata* Infection. *Frontiers in Plant Science* **7**: 1–15.
- LOVE, M. I., W. HUBER, & S. ANDERS, 2014 Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15**: 550.
- LUN, A. T. L., Y. CHEN, & G. K. SMYTH, 2016 It's DE-licious: A Recipe for Differential Expression Analyses of RNA-seq Experiments Using Quasi-Likelihood Methods in edgeR. *Em Methods in Molecular Biology*, volume 1418, pp. 391–416.
- MACRELLI, S., J. MOGENSEN, & G. ZACCHI, 2012 Techno-economic evaluation of 2nd generation bioethanol production from sugar cane bagasse and leaves integrated with the sugar-based ethanol process. *Biotechnology for Biofuels* **5**: 22.
- MAGAREY, R., T. WILLCOX, B. CROFT, & A. CORDINGLEY, 2001 Orange Rust, A Major Pathogen Affecting Crops of Q124 in Queensland in 2000. *Proc. Aust. Soc. Sugarcane Technol* **23**: 274–280.
- MARGUERAT, S. & J. BÄHLER, 2010 RNA-seq: from technology to biology. *Cellular and Molecular Life Sciences* **67**: 569–579.
- MARTIN, M., 2011 Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**: pp. 10–12.
- MCCARTHY, D. J., Y. CHEN, & G. K. SMYTH, 2012 Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research* **40**: 4288–4297.
- MING, R., S. C. LIU, Y. R. LIN, J. DA SILVA, W. WILSON, D. BRAGA, A. VAN DEYNZE, T. F. WENSLAFF, K. K. WU, P. H. MOORE, W. BURNQUIST, M. E. SORRELLS, J. E. IRVINE, & A. H. PATERSON, 1998 Detailed alignment of *Saccharum* and *Sorghum* chromosomes: Comparative organization of closely related diploid and polyploid genomes. *Genetics* **150**: 1663–1682.

- MITTLER, R., S. VANDERAUWERA, N. SUZUKI, G. MILLER, V. B. TOGNETTI, K. VANDEPOELE, M. GOLLERY, V. SHULAEV, & F. VAN BREUSEGEM, 2011 ROS signaling: The new wave? *Trends in Plant Science* **16**: 300–309.
- MLÍČKOVÁ, K., L. LUHOVÁ, A. LEBEDA, B. MIESLEROVÁ, & P. PEČ, 2004 Reactive oxygen species generation and peroxidase activity during *Oidium neolycopersici* infection on *Lycopersicon* species. *Plant Physiology and Biochemistry* **42**: 753–761.
- MOORE, P. H., A. H. PATERSON, & T. TEW, 2013 Sugarcane: The Crop, the Plant, and Domestication. Em *Sugarcane: Physiology, Biochemistry, and Functional Biology*, pp. 1–17, John Wiley & Sons Ltd, Chichester, UK.
- MOREIRA, A. S., 2013 *Ferrugem alaranjada da cana-de-açúcar no Brasil: estudo de populações do patógeno e comportamento varietal*. Tese de Doutorado, Universidade de São Paulo, Piracicaba.
- NICHOLSON, R. L. & R. HAMMERSCHMIDT, 1992 Phenolic Compounds and Their Role in Disease Resistance. *Annual Review of Phytopathology* **30**: 369–389.
- NISHIYAMA, M. Y., S. S. FERREIRA, P. Z. TANG, S. BECKER, A. PÖRTNER-TALIANA, & G. M. SOUZA, 2014 Full-length enriched cDNA libraries and ORFeome analysis of sugarcane hybrid and ancestor genotypes. *PLoS ONE* **9**: 1–14.
- NOGUEIRA, F. T. S., V. E. DE ROSA, M. MENOSSE, E. C. ULIAN, & P. ARRUDA, 2003 RNA expression profiles and data mining of sugarcane response to low temperature. *Plant physiology* **132**: 1811–1824.
- OKURA, V. K., R. S. C. DE SOUZA, S. F. DE SIQUEIRA TADA, & P. ARRUDA, 2016 BAC-Pool Sequencing and Assembly of 19 Mb of the Complex Sugarcane Genome. *Frontiers in Plant Science* **7**: 342.
- OLIVEIRA-GARCIA, E. & B. VALENT, 2015 How eukaryotic filamentous pathogens evade plant recognition. *Current Opinion in Microbiology* **26**: 92–101.
- OZSOLAK, F. & P. M. MILOS, 2011 RNA sequencing: advances, challenges and opportunities. *Nature Reviews Genetics* **12**: 87–98.
- PANDEY, S. & I. E. SOMSSICH, 2009 The Role of WRKY Transcription Factors in Plant Immunity. *Plant Physiology* **150**: 1648–1655.
- PIPERIDIS, G., N. PIPERIDIS, & A. D'HONT, 2010 Molecular cytogenetic investigation of chromosome composition and transmission in sugarcane. *Molecular Genetics and Genomics* **284**: 65–73.
- PURDY, L. H., S. KRUPA, & J. DEAN, 1985 Introduction of Sugarcane Rust into the Americas and Its Spread to Florida. *Plant Disease* **69**: 689.
- RAY, P. D., B. W. HUANG, & Y. TSUJI, 2012 Reactive oxygen species (ROS) homeostasis and redox regulation in cellular signaling. *Cellular Signalling* **24**: 981–990.
- ROBERT-SEILANIANTZ, A., M. GRANT, & J. D. JONES, 2011 Hormone Crosstalk in Plant Disease and Defense: More Than Just JASMONATE-SALICYLATE Antagonism. *Annual Review of Phytopathology* **49**: 317–343.
- ROBINSON, M. D., D. J. MCCARTHY, & G. K. SMYTH, 2009 edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–140.

- ROBINSON, M. D. & G. K. SMYTH, 2008 Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics* **9**: 321–332.
- ROCHA, F. R., F. S. PAPINI-TERZI, M. Y. NISHIYAMA, R. Z. VÊNIO, R. VICENTINI, R. D. DUARTE, V. E. DE ROSA, F. VINAGRE, C. BARSALOBRES, A. H. MEDEIROS, F. A. RODRIGUES, E. C. ULIAN, S. M. ZINGARETTI, J. A. GALBIATTI, R. S. ALMEIDA, A. V. FIGUEIRA, A. S. HEMERLY, M. C. SILVA-FILHO, M. MENOSSE, & G. M. SOUZA, 2007 Signal transduction-related responses to phytohormones and environmental challenges in sugarcane. *BMC Genomics* **8**: 71.
- RUSHTON, P. J., I. E. SOMSSICH, P. RINGLER, & Q. J. SHEN, 2010 WRKY transcription factors. *Trends in Plant Science* **15**: 247–258.
- SCHAKER, P. D. C., A. C. PALHARES, L. M. TANIGUTI, L. P. PETERS, S. CRESTE, K. S. AITKEN, M.-A. VAN SLUYS, J. P. KITAJIMA, M. L. C. VIEIRA, & C. B. MONTEIRO-VITORELLO, 2016 RNAseq Transcriptional Profiling following Whip Development in Sugarcane Smut Disease. *Plos One* **11**: e0162237.
- SCHMELZER, E., S. KRUGER-LEBUS, & K. HAHLBROCK, 1989 Temporal and spatial patterns of gene expression around sites of attempted fungal infection in parsley leaves. *The Plant cell* **1**: 993–1001.
- SCHMITT, F. J., G. RENGER, T. FRIEDRICH, V. D. KRESLAVSKI, S. K. ZHARMUKHAMEDOV, D. A. LOS, V. V. KUZNETSOV, & S. I. ALLAKHVERDIEV, 2014 Reactive oxygen species: Re-evaluation of generation, monitoring and role in stress-signaling in phototrophic organisms. *Biochimica et Biophysica Acta - Bioenergetics* **1837**: 835–848.
- SCHOLES, J. D. & S. A. ROLFE, 1996 Photosynthesis in localised regions of oat leaves infected with crown rust (*Puccinia coronata*): quantitative imaging of chlorophyll fluorescence. *Planta* **199**: 573–582.
- SCORTECCI, K. C., S. CRESTE, T. C. JR., M. A. XAVIER, M. G. A. LANDELL, A. FIGUEIRA, & V. A. BENEDITO, 2012 Challenges, Opportunities and Recent Advances in Sugarcane Breeding. *Plant Breeding* p. 352.
- SONESON, C. & M. DELORENZI, 2013 A comparison of methods for differential expression analysis of RNA-seq data. *BMC Bioinformatics* **14**: 91.
- SOOD, S., J. COMSTOCK, & R. RAID, 2013 Evaluation of sugarcane clones in the CP-cultivar program for resistance to *Puccinia kuehnii*, the pathogen of orange rust. Em *INTERNATIONAL SOCIETY OF SUGAR CANE TECHNOLOGISTS CONGRESS*, volume 28, capítulo Proc. Int., pp. 1–7, São Paulo.
- SOOD, S. G., J. C. COMSTOCK, & N. C. GLYNN, 2009 Leaf Whorl Inoculation Method for Screening Sugarcane Rust Resistance. *Plant Disease* **93**: 1335–1340.
- STAPLES, R. C., 2000 Research on the Rust Fungi During the Twentieth Century. *Annual Review of Phytopathology* **38**: 49–69.
- STONE, J. M. & J. C. WALKER, 1995 Plant Protein Kinase Families and Signal Transduction. *Plant Physiology* **108**: 451–457.
- SWINDELL, W. R., M. HUEBNER, & A. P. WEBER, 2007 Transcriptional profiling of Arabidopsis heat shock proteins and transcription factors reveals extensive overlap between heat and non-heat stress response pathways. *BMC Genomics* **8**: 125.
- TAKKEN, F. L., M. ALBRECHT, & W. I. TAMELING, 2006 Resistance proteins: molecular switches of plant defence. *Current Opinion in Plant Biology* **9**: 383–390.

- TRAPNELL, C., L. PACTHER, & S. L. SALZBERG, 2009 TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**: 1105–1111.
- TURNER, 2014 Assessment of insert sizes and adapter content in fastq data from NexteraXT libraries. *Frontiers in Genetics* **5**: 1–7.
- UNICA, 2017 UnicaData. Acessado em 3-02-2017.
- VANCE, C. P., T. K. KIRK, & R. T. SHERWOOD, 1980 Lignification as a Mechanism of Disease Resistance. *Annual Review of Phytopathology* **18**: 259–288.
- VETTORE, A. L., F. R. DA SILVA, E. L. KEMPER, G. M. SOUZA, A. M. DA SILVA, M. I. T. FERRO, F. HENRIQUE-SILVA, É. A. GIGLIOTI, M. V. F. LEMOS, L. L. COUTINHO, M. P. NOBREGA, H. CARRER, S. C. FRANÇA, M. BACCI, M. H. S. GOLDMAN, S. L. GOMES, L. R. NUNES, L. E. A. CAMARGO, W. J. SIQUEIRA, M. A. VAN SLUYS, O. H. THIEMANN, E. E. KURAMAE, R. V. SANTELLI, C. L. MARINO, M. L. P. N. TARGON, J. A. FERRO, H. C. S. SILVEIRA, D. C. MARINI, E. G. M. LEMOS, C. B. MONTEIRO-VITORELLO, J. H. M. TAMBOR, D. M. CARRARO, P. G. ROBERTO, V. G. MARTINS, G. H. GOLDMAN, R. C. DE OLIVEIRA, D. TRUFFI, C. A. COLOMBO, M. ROSSI, P. G. DE ARAUJO, S. A. SCULACCIO, A. ANGELLA, M. M. A. LIMA, V. E. DE ROSA, F. SIVIERO, V. E. COSCRATO, M. A. MACHADO, L. GRIVET, S. M. Z. DI MAURO, F. G. NOBREGA, C. F. M. MENCK, M. D. V. BRAGA, G. P. TELLES, F. A. A. CARA, G. PEDROSA, J. MEIDANIS, & P. ARRUDA, 2003 Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. *Genome Research* **13**: 2725–2735.
- VIANA, K. R. O. & R. PEREZ, 2013 Survey of sugarcane industry in Minas Gerais, Brazil: Focus on sustainability. *Biomass and Bioenergy* **58**: 149–157.
- VICENTINI, R., A. BOTTCHEER, M. DOS SANTOS BRITO, A. B. DOS SANTOS, S. CRESTE, M. G. DE ANDRADE LANDELL, I. CESARINO, & P. MAZZAFERA, 2015 Large-scale transcriptome analysis of two sugarcane genotypes contrasting for lignin content. *PLoS ONE* **10**: e0134909.
- VIRTUDAZO, E. V., H. NOJIMA, & M. KAKISHIMA, 2001 Taxonomy of Puccinia species causing rust diseases on sugarcane. *Mycoscience* **42**: 167–175.
- VOGEL, C. & E. M. MARCOTTE, 2012 Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews Genetics* **13**: 227–232.
- WANG, J., B. ROE, S. MACMIL, Q. YU, J. E. MURRAY, H. TANG, C. CHEN, F. NAJAR, G. WILEY, J. BOWERS, M.-A. VAN SLUYS, D. S. ROKHSAR, M. E. HUDSON, S. P. MOOSE, A. H. PATERSON, & R. MING, 2010 Microcollinearity between autopolyploid sugarcane and diploid sorghum genomes. *BMC genomics* **11**: 261.
- WANG, Z., M. GERSTEIN, & M. SNYDER, 2009 RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* **10**: 57–63.
- WHALEN, M. C., 2005 Host defence in a developmental context. *Molecular Plant Pathology* **6**: 347–360.
- YANG, C. & H. WEI, 2015 Designing microarray and RNA-Seq experiments for greater systems biology discovery in modern plant genomics. *Molecular Plant* **8**: 196–206.
- YOUNG, M. D., M. J. WAKEFIELD, G. K. SMYTH, & A. OSHLACK, 2010 Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome biology* **11**: R14.

ZHAO, D., N. C. GLYNN, B. GLAZ, J. C. COMSTOCK, & S. SOOD, 2011 Orange rust effects on leaf photosynthesis and related characters of sugarcane. *Plant Disease* **95**: 640–647.

APÊNDICES

Apêndice I

Informações sobre a porcentagem de alinhamento e o comportamento das amostras.

Tabela A.1. Estatísticas do sequenciamento e pré-processamento das bibliotecas. As colunas representam o número de leituras brutas após o sequenciamento, as leituras remanescentes após a remoção de extremidades de baixa qualidade e de adaptadores (primeira etapa do pré-processamento) e, por último, o número final após a remoção de leituras alinhadas a rRNAs (segunda etapa do pré-processamento)

Amostra	Réplica	Número Total de Leituras	Remoção de extremidades de baixa qualidade e adaptadores	Remoção de leituras alinhadas aos rRNA 18S e 28S
0 h	1	53.494.796	53.447.786	17.207.682
0 h	2	50.182.722	50.140.032	41.512.784
12 hai	1	52.890.694	52.845.208	45.643.024
12 hai	2	54.698.556	54.652.722	47.078.000
12 hai	3	58.316.470	58.266.628	49.650.526
24 hai	1	59.285.594	59.233.804	52.098.702
24 hai	2	53.917.270	53.871.410	48.634.720
24 hai	3	54.583.922	54.537.588	49.160.894
48 dai	1	53.744.654	53.698.142	50.112.458
48 dai	2	56.254.056	56.207.860	51.870.916
48 dai	3	48.214.660	48.175.122	44.251.474
5 dai	1	55.066.352	55.019.892	43.946.676
5 dai	2	60.867.374	60.815.688	58.255.994
5 dai	3	56.662.696	56.615.650	53.020.894
12 dai	1	53.167.696	53.122.696	51.050.042
12 dai	2	52.220.146	52.175.420	39.186.234
12 dai	3	54.457.816	54.411.456	51.728.196
Total	-	928.025.474	927.237.104	794.409.216

Tabela A.2. Porcentagem de alinhamento de leituras pareadas ao transcriptoma de referência, conforme a repetição e linha de cada uma das amostras

Amostra	Repetição	Linha	Alinhamento(%)
0h	1	1	46,06
0h	1	2	46,15
0h	2	1	65,26
0h	2	2	65,48
12hai	1	1	64,29
12hai	1	2	64,46
12hai	2	1	67,69
12hai	2	2	67,90
12hai	3	1	66,37
12hai	3	2	66,57
24hai	1	1	67,13
24hai	1	2	67,35
24hai	2	1	69,43
24hai	2	2	69,65
24hai	3	1	67,95
24hai	3	2	68,17
48hai	1	1	60,00
48hai	1	2	60,25
48hai	2	1	69,44
48hai	2	2	69,65
48hai	3	1	51,56
48hai	3	2	51,84
5dai	1	1	71,12
5dai	1	2	71,34
5dai	2	1	70,68
5dai	2	2	70,86
5dai	3	1	70,04
5dai	3	2	70,26
12dai	1	1	71,43
12dai	1	2	71,67
12dai	2	1	68,55
12dai	2	2	68,81
12dai	3	1	68,14
12dai	3	2	68,35

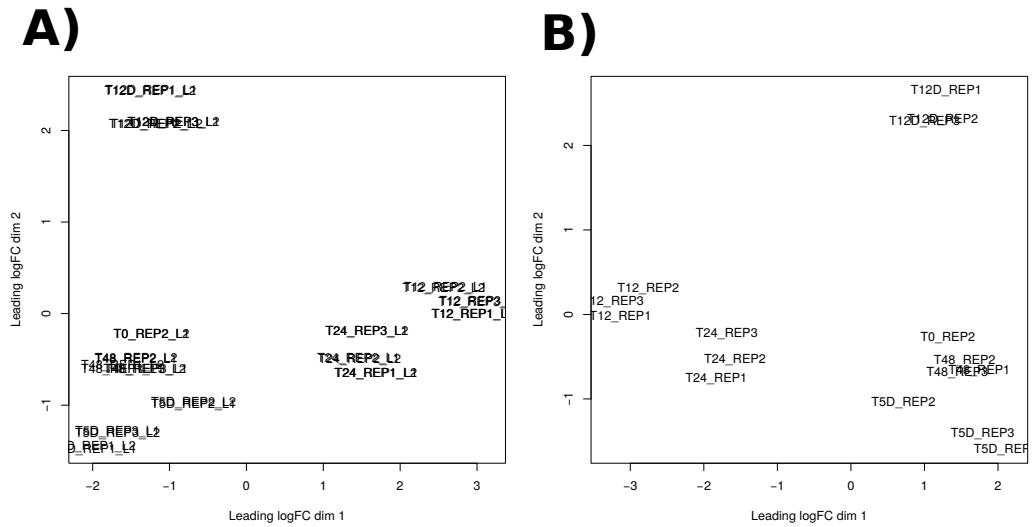


Figura A.1. Escalonamento multidimensional retratando a similaridade entre as amostras. (A) É possível observar a sobreposição entre as *lanes* de uma mesma biblioteca. Além disso, as réplicas dos tempos mantiveram-se próximas, sendo que 12 hai, 24 hai e 12 dai formaram grupos isolados de réplicas, distinguindo-se dos demais tempos. (B) Esse fato também foi visto após a união das *lanes* de uma mesma réplica biológica.

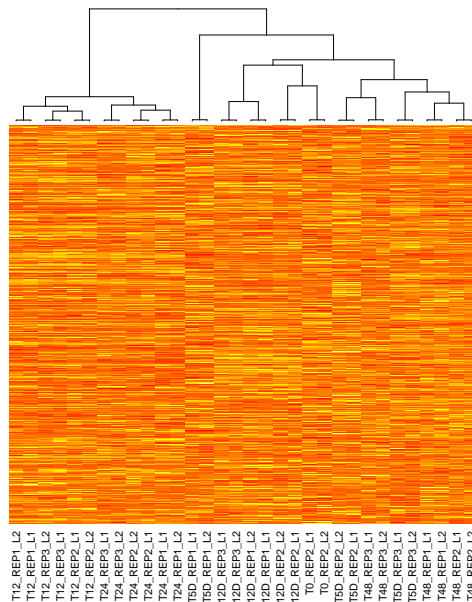


Figura A.2. *Heatmap* das amostras sequenciadas, obtido a partir das contagens normalizadas, em contagens por milhão (cpm), de 31.529 transcritos com leituras mapeadas. Nota-se grande semelhança entre ambas as corridas realizadas para cada amostra, em *lanes* distintas. Além disso, os tempos 48hai e 5dai mostraram-se agrupados.

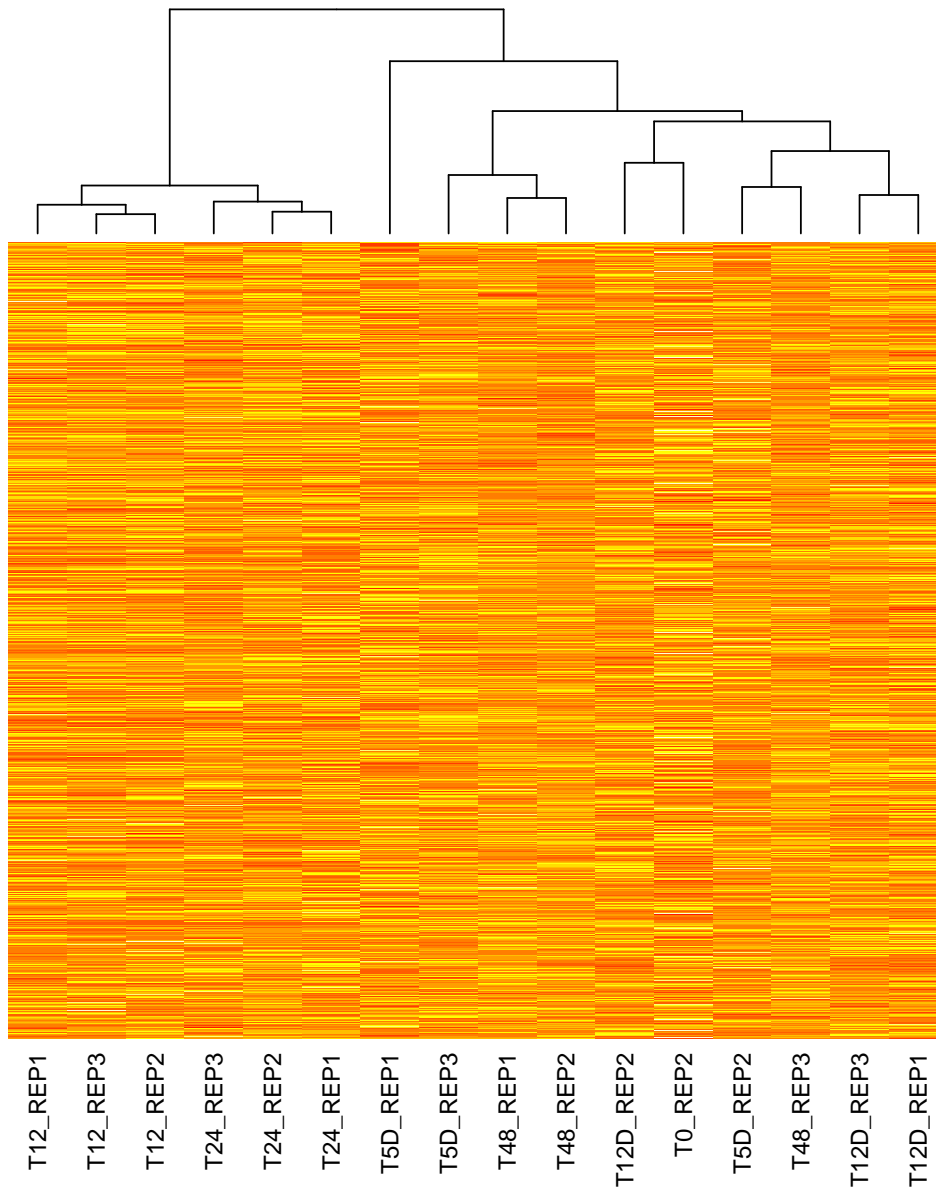


Figura A.3. *Heatmap* das amostras sequenciadas, obtido a partir das contagens normalizadas após a junção das *lanes*, em contagens por milhão (cpm), de 28.641 transcritos com leituras mapeadas. Verifica-se que as réplicas de 48 h formam grupos com 0 h e 5 dai, ao passo que as réplicas dos demais tempos permanecem próximas entre si.

Apêndice II

Informações complementares para a interpretação dos resultados provenientes da análise do enriquecimento funcional quando utilizado o teste do tipo ANAVA.

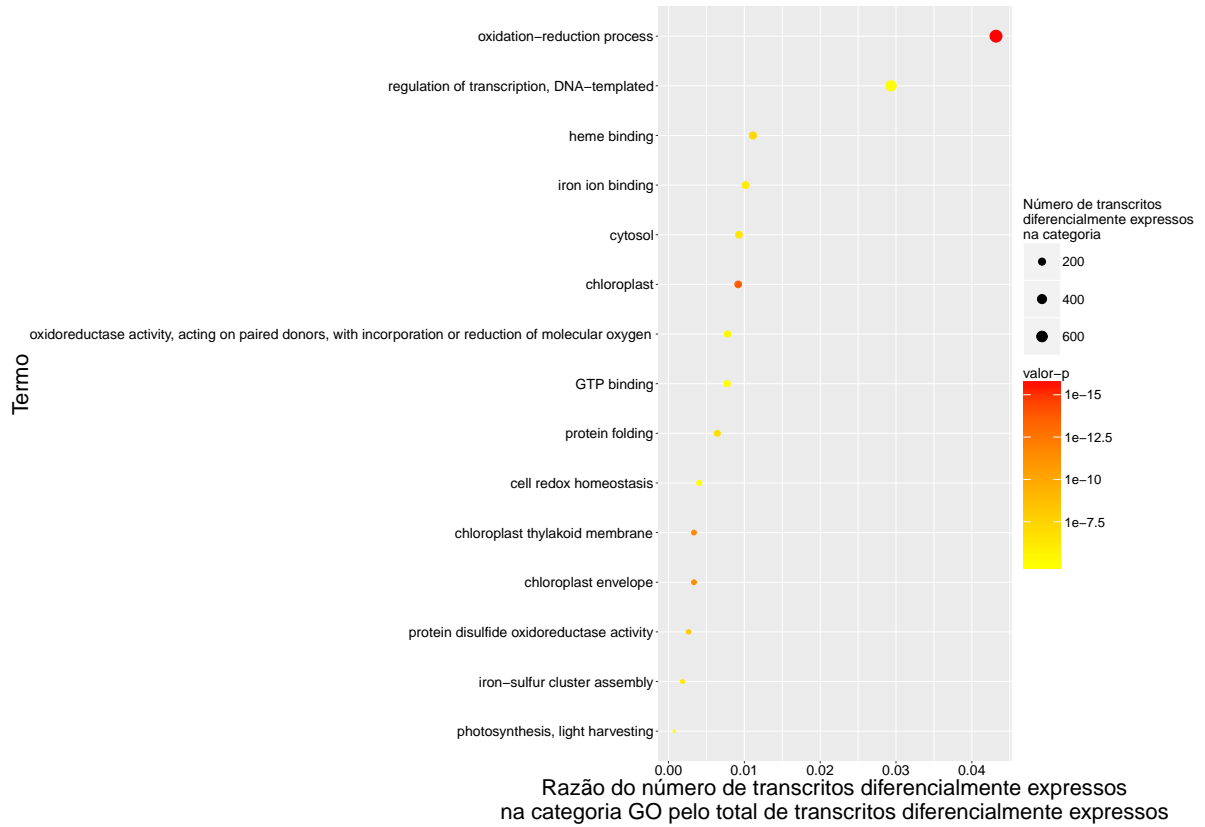


Figura A.4. Representação dos termos enriquecidos com transcritos diferencialmente expressos, incluindo informações sobre o número de transcritos diferencialmente expressos e valor- p . Foram considerados os 15 termos GO enriquecidos de menor valor- p , dispostos no eixo y . No eixo x os valores representam a razão entre o número de transcritos diferencialmente expressos em uma dada categoria pelo número total de transcritos diferencialmente expressos. O tamanho dos pontos representa o número de transcritos diferencialmente expressos e o gradiente vermelho-amarelo indica o valor- p das categorias.

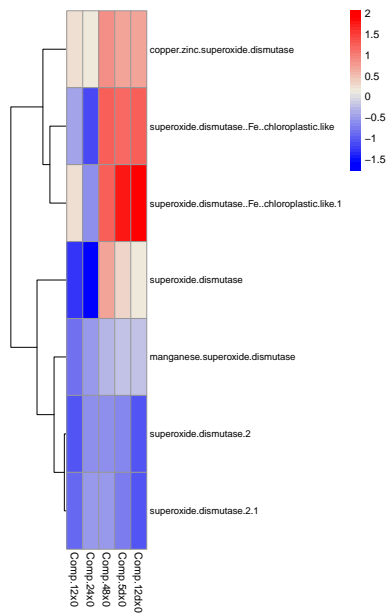


Figura A.5. Transcritos diferencialmente expressos anotados como superóxido-dismutase. Na barra de cores o vermelho representa valores de LFC positivos e o azul indica valores negativos.

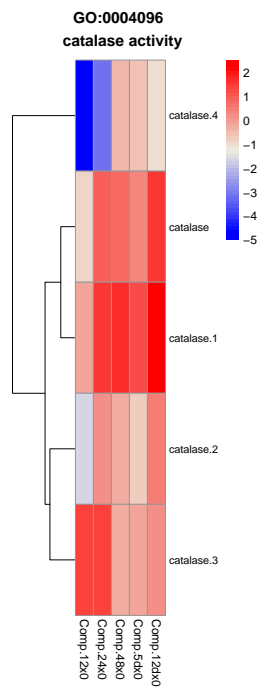


Figura A.6. Transcritos diferencialmente expressos anotados como relacionados à atividade de catalase. Na barra de cores o vermelho representa valores de LFC positivos e o azul indica valores negativos.

Tabela A.3. Categoria e descrição dos termos *Gene Ontology* sobre-representados da análise de enriquecimento funcional, fundamentada nos transcritos diferencialmente expressos do teste do tipo ANAVA.

Categoria	Descrição	valor-p (teste de sobre-representação)	número de transcritos na categoria	número de transcritos diferencialmente expressos
GO:0055114	oxidation-reduction process	1.902108e-16	1475	799
GO:0009507	chloroplast	2.430832e-14	234	170
GO:0009535	chloroplast thylakoid membrane	2.515823e-12	68	62
GO:0009941	chloroplast envelope	4.703717e-12	70	62
GO:0015035	protein disulfide oxidoreductase activity	1.018350e-08	64	49
GO:0020037	heme binding	6.098735e-08	400	206
GO:0006457	protein folding	1.537852e-07	172	119
GO:0005829	cytosol	3.456123e-07	286	172
GO:0016226	iron-sulfur cluster assembly	7.557684e-07	36	34
GO:0005506	iron ion binding	8.771207e-07	360	188
GO:0016705	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	3.005380e-06	271	144
GO:0045454	cell redox homeostasis	8.771332e-06	110	75
GO:0005525	GTP binding	9.049544e-06	220	143
GO:0006355	regulation of transcription, DNA-templated	9.946324e-06	1040	543
GO:0009765	photosynthesis, light harvesting	1.028048e-05	14	14
GO:0009570	chloroplast stroma	1.849390e-05	71	54
GO:0003735	structural constituent of ribosome	2.139142e-05	447	198
GO:0016491	oxidoreductase activity	2.435716e-05	559	298
GO:0005215	transporter activity	4.031418e-05	269	147
GO:0006662	glycerol ether metabolic process	4.844406e-05	33	25
GO:0008152	metabolic process	5.754971e-05	1278	577
GO:0051082	unfolded protein binding	6.909969e-05	90	62
GO:0016787	hydrolase activity	8.749778e-05	409	226
GO:0003700	transcription factor activity, sequence-specific DNA binding	8.764743e-05	509	269
GO:0006412	translation	1.441424e-04	431	190
GO:0005759	mitochondrial matrix	1.446374e-04	20	17
GO:0009579	thylakoid	2.359780e-04	13	13
GO:0006396	RNA processing	2.675891e-04	68	53
GO:0005975	carbohydrate metabolic process	2.799468e-04	362	197
GO:0010287	plastoglobule	2.828121e-04	21	20
GO:0006352	DNA-templated transcription, initiation	4.392841e-04	24	22
GO:0009408	response to heat	4.851669e-04	29	22
GO:0007264	small GTPase mediated signal transduction	6.260137e-04	80	50
GO:0006979	response to oxidative stress	6.352897e-04	97	54
GO:0003924	GTPase activity	6.367936e-04	99	68
GO:0043086	negative regulation of catalytic activity	7.258787e-04	18	12
GO:0006364	rRNA processing	7.655614e-04	40	33
GO:0009658	chloroplast organization	7.823319e-04	30	24
GO:0051536	iron-sulfur cluster binding	8.468335e-04	42	32
GO:0030154	cell differentiation	9.421451e-04	35	25
GO:0015095	magnesium ion transmembrane transporter activity	9.886053e-04	25	21
GO:0015693	magnesium ion transport	9.886053e-04	25	21
GO:0004872	receptor activity	1.074294e-03	12	10
GO:0042542	response to hydrogen peroxide	1.099653e-03	12	11
GO:0008519	ammonium transmembrane transporter activity	1.138341e-03	12	10
GO:0072488	ammonium transmembrane transport	1.138341e-03	12	10
GO:0009815	1-aminocyclopropane- 1-carboxylate oxidase activity	1.143742e-03	4	4

Apêndice III

Neste apêndice está inclusa a tabela contendo o número de transcritos diferencialmente expressos para as comparações de tempos adjacentes. Também são dispostas as tabelas de enriquecimento funcional considerando todos os transcritos diferencialmente expressos (*up + down*) ou somente aqueles estimulados (*up*) ou reprimidos (*down-regulated*).

Tabela A.4. Número de transcritos diferencialmente expressos, levando em conta a comparação entre tempos adjacentes, com indicação dos transcritos estimulados (*Up*) e reprimidos (*Down-regulated*).

Regulação	12 hai x 0h		24 hai x 12 hai		48 hai x 24 hai		5 dai x 48 hai		12 dai x 5 dai	
	<i>Up</i>	<i>Down</i>	<i>Up</i>	<i>Down</i>	<i>Up</i>	<i>Down</i>	<i>Up</i>	<i>Down</i>	<i>Up</i>	<i>Down</i>
Transcritos Diferencialmente Expressos	3.537	4.643	1.220	623	5.338	4.766	306	401	3.243	3.417
Total	8.180		1.843		10.104		707		6.660	

Tabela A.5. Categoria e descrição dos termos provenientes do enriquecimento funcional de 12 hai em relação a 0h. Essa avaliação foi feita separadamente para os transcritos estimulados, reprimidos e ambos os casos de maneira simultânea, indicados pela terceira coluna.

Categoria	Descrição	Sentido da expressão diferencial
GO:0003735	structural constituent of ribosome	<i>up</i>
GO:0006412	translation	<i>up</i>
GO:0005840	ribosome	<i>up</i>
GO:0032040	small-subunit processome	<i>up</i>
GO:0006364	rRNA processing	<i>up</i>
GO:0006396	RNA processing	<i>up</i>
GO:0004066	asparagine synthase (glutamine-hydrolyzing) activity	<i>up</i>
GO:0006529	asparagine biosynthetic process	<i>up</i>
GO:0022625	cytosolic large ribosomal subunit	<i>up</i>
GO:0055114	oxidation-reduction process	<i>down</i>
GO:0009507	chloroplast	<i>down</i>
GO:0016491	oxidoreductase activity	<i>down</i>
GO:0009535	chloroplast thylakoid membrane protein disulfide oxidoreductase activity	<i>down</i>
GO:0015035	activity	<i>down</i>
GO:0005506	iron ion binding	<i>down</i>
GO:0008152	metabolic process	<i>down</i>
GO:0020037	heme binding	<i>down</i>
GO:0009765	photosynthesis, light harvesting	<i>down</i>
GO:0015979	photosynthesis	<i>down</i>
GO:0016705	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen transferase activity, transferring hexosyl groups	<i>down</i>
GO:0016758	hexosyl groups	<i>down</i>
GO:0010287	plastoglobule	<i>down</i>
GO:0009055	electron carrier activity	<i>down</i>
GO:0009941	chloroplast envelope	<i>down</i>

Tabela A.5. Categoria e descrição dos termos provenientes do enriquecimento funcional de 12 hai em relação a 0h. Essa avaliação foi feita separadamente para os transcritos estimulados, reprimidos e ambos os casos de maneira simultânea, indicados pela terceira coluna.

Categoria	Descrição	Sentido da expressão diferencial
GO:0045454	cell redox homeostasis	<i>down</i>
GO:0016021	integral component of membrane	<i>down</i>
GO:0006662	glycerol ether metabolic process	<i>down</i>
GO:0005773	vacuole	<i>down</i>
GO:0009408	response to heat	<i>down</i>
GO:0006979	response to oxidative stress	<i>down</i>
GO:0015995	chlorophyll biosynthetic process	<i>down</i>
GO:0005747	mitochondrial respiratory chain complex I	<i>down</i>
GO:0009538	photosystem I reaction center	<i>down</i>
GO:0006629	lipid metabolic process	<i>down</i>
GO:0009638	phototropism	<i>down</i>
GO:0005829	cytosol	<i>down</i>
GO:0033180	proton-transporting V-type ATPase, V1 domain malate dehydrogenase (decarboxylating) (NAD ⁺) activity	<i>down</i>
GO:0004471	response to hydrogen peroxide	<i>down</i>
GO:0042542	carbohydrate metabolic process	<i>down</i>
GO:0005975	heat acclimation	<i>down</i>
GO:0010286	aleurone grain membrane	<i>down</i>
GO:0032578	lignin biosynthetic process cinnamyl-alcohol dehydrogenase activity	<i>down</i>
GO:0009809	sinapyl alcohol dehydrogenase activity	<i>down</i>
GO:0045551	chloroplast stroma	<i>down</i>
GO:0052747	potassium ion transmembrane transport	<i>down</i>
GO:0071805	oxidoreductase activity, acting on NAD(P)H, quinone or similar compound as acceptor	<i>down</i>
GO:0016655	NADH dehydrogenase activity	<i>down</i>
GO:0003954	response to high light intensity	<i>down</i>
GO:0009644	oxidation-reduction process	<i>up + down</i>
GO:0055114	structural constituent of ribosome	<i>up + down</i>
GO:0003735	chloroplast	<i>up + down</i>
GO:0009507	translation	<i>up + down</i>
GO:0006412	protein disulfide oxidoreductase activity	<i>up + down</i>
GO:0015035	oxidoreductase activity	<i>up + down</i>
GO:0016491	ribosome	<i>up + down</i>
GO:0005840	response to oxidative stress	<i>up + down</i>
GO:0006979	heme binding	<i>up + down</i>
GO:0020037	photosynthesis, light harvesting	<i>up + down</i>
GO:0009765	glycerol ether metabolic process	<i>up + down</i>
GO:0006662		<i>up + down</i>

Tabela A.5. Categoria e descrição dos termos provenientes do enriquecimento funcional de 12 hai em relação a 0h. Essa avaliação foi feita separadamente para os transcritos estimulados, reprimidos e ambos os casos de maneira simultânea, indicados pela terceira coluna.

Categoria	Descrição	Sentido da expressão diferencial
GO:0009535	chloroplast thylakoid membrane	<i>up + down</i>
GO:0045454	cell redox homeostasis	<i>up + down</i>
GO:0005506	iron ion binding	<i>up + down</i>
GO:0004066	asparagine synthase (glutamine-hydrolyzing) activity	<i>up + down</i>
GO:0006529	asparagine biosynthetic process	<i>up + down</i>
GO:0005975	carbohydrate metabolic process	<i>up + down</i>
GO:0008152	metabolic process	<i>up + down</i>
GO:0015979	photosynthesis	<i>up + down</i>
GO:0010287	plastoglobule	<i>up + down</i>
GO:0006629	lipid metabolic process	<i>up + down</i>
GO:0009055	electron carrier activity oxidoreductase activity, acting on	<i>up + down</i>
GO:0016705	paired donors, with incorporation or reduction of molecular oxygen	<i>up + down</i>
GO:0005773	vacuole	<i>up + down</i>
GO:0004602	glutathione peroxidase activity	<i>up + down</i>
GO:0005829	cytosol	<i>up + down</i>
GO:0016021	integral component of membrane	<i>up + down</i>
GO:0009941	chloroplast envelope	<i>up + down</i>
GO:0006873	cellular ion homeostasis	<i>up + down</i>
GO:0009408	response to heat	<i>up + down</i>
GO:0042542	response to hydrogen peroxide	<i>up + down</i>

Tabela A.6. Categoria e descrição dos termos provenientes do enriquecimento funcional de 24 hai em relação a 12 hai. Essa avaliação foi feita separadamente para os transcritos estimulados, reprimidos e ambos os casos de maneira simultânea, indicados pela terceira coluna.

Categoria	Descrição	Sentido da expressão diferencial
GO:0009765	photosynthesis, light harvesting	<i>up</i>
GO:0055114	oxidation-reduction process	<i>up</i>
GO:0009638	phototropism	<i>up</i>
GO:0003700	transcription factor activity, sequence-specific DNA binding	<i>up</i>
GO:0015035	protein disulfide oxidoreductase activity	<i>up</i>
GO:0006355	regulation of transcription, DNA-templated	<i>up</i>
GO:0004014	adenosylmethionine decarboxylase activity	<i>up</i>
GO:0006597	spermine biosynthetic process	<i>up</i>
GO:0005506	iron ion binding	<i>up</i>
GO:0020037	heme binding	<i>up</i>
GO:0008295	spermidine biosynthetic process	<i>up</i>
GO:0043565	sequence-specific DNA binding	<i>up</i>
GO:0022857	transmembrane transporter activity	<i>up</i>

Tabela A.6. Categoria e descrição dos termos provenientes do enriquecimento funcional de 24 hai em relação a 12 hai. Essa avaliação foi feita separadamente para os transcritos estimulados, reprimidos e ambos os casos de maneira simultânea, indicados pela terceira coluna.

Categoria	Descrição	Sentido da expressão diferencial
GO:0000989	transcription factor activity, transcription factor binding oxidoreductase activity, acting on	<i>up</i>
GO:0016705	paired donors, with incorporation or reduction of molecular oxygen protochlorophyllide reductase	<i>up</i>
GO:0016630	activity	<i>up</i>
GO:0006662	glycerol ether metabolic process	<i>up</i>
GO:0035066	positive regulation of histone acetylation	<i>up</i>
GO:0009055	electron carrier activity	<i>up</i>
GO:0004096	catalase activity	<i>up</i>
GO:0005773	vacuole	<i>up</i>
GO:0009535	chloroplast thylakoid membrane	<i>up</i>
GO:0009765	photosynthesis, light harvesting protein disulfide oxidoreductase	<i>up + down</i>
GO:0015035	activity	<i>up + down</i>
GO:0009638	phototropism	<i>up + down</i>
GO:0009535	chloroplast thylakoid membrane	<i>up + down</i>
GO:0009507	chloroplast	<i>up + down</i>
GO:0055114	oxidation-reduction process	<i>up + down</i>
GO:0005506	iron ion binding adenosylmethionine	<i>up + down</i>
GO:0004014	decarboxylase activity	<i>up + down</i>
GO:0006597	spermine biosynthetic process	<i>up + down</i>
GO:0006662	glycerol ether metabolic process	<i>up + down</i>
GO:0006355	regulation of transcription, DNA-templated	<i>up + down</i>
GO:0008295	spermidine biosynthetic process	<i>up + down</i>
GO:0006351	transcription, DNA-templated	<i>up + down</i>
GO:0003700	transcription factor activity, sequence-specific DNA binding	<i>up + down</i>
GO:0045454	cell redox homeostasis	<i>up + down</i>
GO:0009055	electron carrier activity	<i>up + down</i>

Tabela A.7. Categoria e descrição dos termos provenientes do enriquecimento funcional de 48 hai em relação a 24 hai. Essa avaliação foi feita separadamente para os transcritos estimulados, reprimidos e ambos os casos de maneira simultânea, indicados pela terceira coluna.

Categoria	Descrição	Sentido da expressão diferencial
GO:0009507	chloroplast	<i>up</i>
GO:0009535	chloroplast thylakoid membrane	<i>up</i>
GO:0055114	oxidation-reduction process	<i>up</i>
GO:0009941	chloroplast envelope	<i>up</i>
GO:0009570	chloroplast stroma	<i>up</i>
GO:0015979	photosynthesis	<i>up</i>
GO:0016491	oxidoreductase activity	<i>up</i>

Tabela A.7. Categoria e descrição dos termos provenientes do enriquecimento funcional de 48 hai em relação a 24 hai. Essa avaliação foi feita separadamente para os transcritos estimulados, reprimidos e ambos os casos de maneira simultânea, indicados pela terceira coluna.

Categoria	Descrição	Sentido da expressão diferencial
GO:0009765	photosynthesis, light harvesting	<i>up</i>
GO:0008152	metabolic process	<i>up</i>
GO:0009658	chloroplast organization	<i>up</i>
GO:0005506	iron ion binding	<i>up</i>
GO:0016758	transferase activity, transferring hexosyl groups	<i>up</i>
GO:0010207	photosystem II assembly	<i>up</i>
GO:0006457	protein folding	<i>up</i>
GO:0020037	heme binding	<i>up</i>
GO:0010287	plastoglobule	<i>up</i>
GO:0016705	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	<i>up</i>
GO:0009534	chloroplast thylakoid	<i>up</i>
GO:0005829	cytosol	<i>up</i>
GO:0009408	response to heat	<i>up</i>
GO:0016787	hydrolase activity isopentenyl diphosphate	<i>up</i>
GO:0019288	biosynthetic process, methylerythritol 4-phosphate pathway	<i>up</i>
GO:0016556	mRNA modification	<i>up</i>
GO:0015035	protein disulfide oxidoreductase activity	<i>up</i>
GO:0015995	chlorophyll biosynthetic process	<i>up</i>
GO:0009654	photosystem II oxygen evolving complex	<i>up</i>
GO:0004252	serine-type endopeptidase activity	<i>up</i>
GO:0031977	thylakoid lumen	<i>up</i>
GO:0051082	unfolded protein binding	<i>up</i>
GO:0005975	carbohydrate metabolic process	<i>up</i>
GO:0009536	plastid	<i>up</i>
GO:0009508	plastid chromosome	<i>up</i>
GO:0016226	iron-sulfur cluster assembly	<i>up</i>
GO:0035304	regulation of protein dephosphorylation	<i>up</i>
GO:0009538	photosystem I reaction center	<i>up</i>
GO:0000373	Group II intron splicing	<i>up</i>
GO:0009638	phototropism	<i>up</i>
GO:0000023	maltose metabolic process	<i>up</i>
GO:0009853	photorespiration	<i>up</i>
GO:0003690	double-stranded DNA binding	<i>up</i>
GO:0009706	chloroplast inner membrane	<i>up</i>
GO:0009522	photosystem I	<i>up</i>
GO:0006662	glycerol ether metabolic process	<i>up</i>
GO:0050662	coenzyme binding	<i>up</i>

Tabela A.7. Categoria e descrição dos termos provenientes do enriquecimento funcional de 48 hai em relação a 24 hai. Essa avaliação foi feita separadamente para os transcritos estimulados, reprimidos e ambos os casos de maneira simultânea, indicados pela terceira coluna.

Categoria	Descrição	Sentido da expressão diferencial
	transferase activity, transferring	
GO:0016765	alkyl or aryl (other than methyl) groups	<i>up</i>
GO:0016311	dephosphorylation	<i>up</i>
GO:0019252	starch biosynthetic process	<i>up</i>
GO:0009809	lignin biosynthetic process	<i>up</i>
	cinnamyl-alcohol	
GO:0045551	dehydrogenase activity	<i>up</i>
	sinapyl alcohol dehydrogenase activity	
GO:0052747		<i>up</i>
GO:0010027	thylakoid membrane organization	<i>up</i>
GO:0006546	glycine catabolic process	<i>up</i>
GO:0031248	protein acetyltransferase complex	<i>up</i>
GO:0009707	chloroplast outer membrane	<i>up</i>
GO:0005978	glycogen biosynthetic process	<i>up</i>
GO:0019438	aromatic compound biosynthetic process	<i>up</i>
GO:0001927	exocyst assembly	<i>up</i>
GO:0016655	oxidoreductase activity, acting on NAD(P)H, quinone or similar compound as acceptor	<i>up</i>
GO:0015238	drug transmembrane transporter activity	<i>up</i>
GO:0006855	drug transmembrane transport	<i>up</i>
GO:0033014	tetrapyrrole biosynthetic process	<i>up</i>
GO:0061077	chaperone-mediated protein folding	<i>up</i>
GO:0045454	cell redox homeostasis	<i>up</i>
GO:0031072	heat shock protein binding	<i>up</i>
GO:0006535	cysteine biosynthetic process	<i>up</i>
GO:0003735	from serine structural constituent of ribosome	<i>down</i>
GO:0006412	translation	<i>down</i>
GO:0000166	nucleotide binding	<i>down</i>
GO:0005840	ribosome	<i>down</i>
GO:0006338	chromatin remodeling	<i>down</i>
GO:0009507	chloroplast	<i>up + down</i>
GO:0009535	chloroplast thylakoid membrane	<i>up + down</i>
GO:0009941	chloroplast envelope	<i>up + down</i>
GO:0055114	oxidation-reduction process	<i>up + down</i>
GO:0009765	photosynthesis, light harvesting	<i>up + down</i>
GO:0009570	chloroplast stroma	<i>up + down</i>
GO:0015095	magnesium ion transmembrane transporter activity	<i>up + down</i>
GO:0015693	magnesium ion transport	<i>up + down</i>
GO:0009658	chloroplast organization	<i>up + down</i>
GO:0015979	photosynthesis	<i>up + down</i>
GO:0005829	cytosol	<i>up + down</i>

Tabela A.7. Categoria e descrição dos termos provenientes do enriquecimento funcional de 48 hai em relação a 24 hai. Essa avaliação foi feita separadamente para os transcritos estimulados, reprimidos e ambos os casos de maneira simultânea, indicados pela terceira coluna.

Categoria	Descrição	Sentido da expressão diferencial
GO:0010207	photosystem II assembly	<i>up + down</i>
GO:0008152	metabolic process	<i>up + down</i>
GO:0006457	protein folding	<i>up + down</i>
GO:0006412	translation	<i>up + down</i>
GO:0003735	structural constituent of ribosome	<i>up + down</i>
GO:0016491	oxidoreductase activity	<i>up + down</i>
GO:0005525	GTP binding	<i>up + down</i>
GO:0020037	heme binding	<i>up + down</i>
GO:0015035	protein disulfide oxidoreductase activity	<i>up + down</i>
GO:0010287	plastoglobule	<i>up + down</i>
GO:0004427	inorganic diphosphatase activity	<i>up + down</i>
GO:0005840	ribosome	<i>up + down</i>
GO:0006979	response to oxidative stress	<i>up + down</i>
GO:0031977	thylakoid lumen	<i>up + down</i>
GO:0006662	glycerol ether metabolic process	<i>up + down</i>
GO:0051082	unfolded protein binding	<i>up + down</i>
GO:0009116	nucleoside metabolic process	<i>up + down</i>

Tabela A.8. Categoria e descrição dos termos provenientes do enriquecimento funcional de 5 dai em relação a 48 hai. Essa avaliação foi feita separadamente para os transcritos estimulados, reprimidos e ambos os casos de maneira simultânea, indicados pela terceira coluna.

Categoria	Descrição	Sentido da expressão diferencial
GO:0000786	nucleosome	<i>down</i>
GO:0055114	oxidation-reduction process	<i>down</i>
GO:0000786	nucleosome	<i>up + down</i>
GO:0009408	response to heat	<i>up + down</i>

Tabela A.9. Categoria e descrição dos termos provenientes do enriquecimento funcional de 12 dai em relação a 5 dai. Essa avaliação foi feita separadamente para os transcritos estimulados, reprimidos e ambos os casos de maneira simultânea, indicados pela terceira coluna.

Categoria	Descrição	Sentido da expressão diferencial
GO:0006364	rRNA processing	<i>up</i>
GO:0001522	pseudouridine synthesis	<i>up</i>
GO:0032040	small-subunit processome	<i>up</i>
GO:0009982	pseudouridine synthase activity	<i>up</i>
GO:0030515	snoRNA binding	<i>up</i>
GO:0008519	ammonium transmembrane transporter activity	<i>up</i>

Tabela A.9. Categoria e descrição dos termos provenientes do enriquecimento funcional de 12 dai em relação a 5 dai. Essa avaliação foi feita separadamente para os transcritos estimulados, reprimidos e ambos os casos de maneira simultânea, indicados pela terceira coluna.

Categoria	Descrição	Sentido da expressão diferencial
GO:0072488	ammonium transmembrane transport	<i>up</i>
GO:0005730	nucleolus	<i>up</i>
GO:0003723	RNA binding	<i>up</i>
GO:0003924	GTPase activity	<i>up</i>
GO:0006396	RNA processing	<i>up</i>
GO:0005215	transporter activity	<i>up</i>
GO:0006412	translation	<i>up</i>
GO:0003735	structural constituent of ribosome	<i>up</i>
GO:0030870	Mre11 complex	<i>up</i>
GO:0006810	transport	<i>up</i>
GO:0008168	methyltransferase activity	<i>up</i>
GO:0006626	protein targeting to mitochondrion	<i>up</i>
GO:0031428	box C/D snoRNP complex	<i>up</i>
GO:0003697	single-stranded DNA binding	<i>up</i>
GO:0020037	heme binding	<i>up</i>
GO:0009640	photomorphogenesis	<i>up</i>
GO:0008308	voltage-gated anion channel activity	<i>up</i>
GO:0008152	metabolic process	<i>down</i>
GO:0009611	response to wounding	<i>down</i>
GO:0016758	transferase activity, transferring hexosyl groups	<i>down</i>
GO:0016021	integral component of membrane protein disulfide oxidoreductase	<i>down</i>
GO:0015035	activity	<i>down</i>
GO:0051537	2 iron, 2 sulfur cluster binding	<i>down</i>
GO:0055114	oxidation-reduction process	<i>down</i>
GO:0006810	transport	<i>up + down</i>
GO:0008519	ammonium transmembrane transporter activity	<i>up + down</i>
GO:0072488	ammonium transmembrane transport	<i>up + down</i>
GO:0005215	transporter activity	<i>up + down</i>
GO:0003924	GTPase activity	<i>up + down</i>
GO:0055114	oxidation-reduction process	<i>up + down</i>
GO:0042542	response to hydrogen peroxide	<i>up + down</i>
GO:0015035	protein disulfide oxidoreductase activity	<i>up + down</i>
GO:0020037	heme binding	<i>up + down</i>
GO:0032040	small-subunit processome	<i>up + down</i>
GO:0001522	pseudouridine synthesis	<i>up + down</i>
GO:0016705	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	<i>up + down</i>
GO:0030515	snoRNA binding	<i>up + down</i>
GO:0030488	tRNA methylation	<i>up + down</i>

Tabela A.9. Categoria e descrição dos termos provenientes do enriquecimento funcional de 12 dai em relação a 5 dai. Essa avaliação foi feita separadamente para os transcritos estimulados estimulados, reprimidos e ambos os casos de maneira simultânea, indicados pela terceira coluna.

Categoria	Descrição	Sentido da expressão diferencial
GO:0006364	rRNA processing	<i>up + down</i>

Apêndice IV

Aqui são encontrados os *heatmaps* de transcritos de potencial relevância no processo de resposta à infecção no estudo do perfil temporal de expressão gênica.

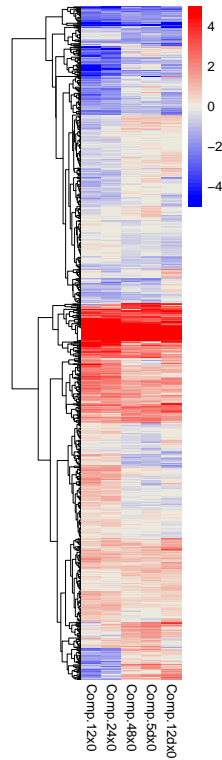


Figura A.7. Comportamento temporal dos transcritos anotados como cinases após a inoculação da cana-de-açúcar com ferrugem alaranjada. *Heatmap* do LFC dos transcritos diferencialmente expressos da comparação de tempos adjacentes e, também, do teste do tipo ANAVA. Na barra de cores o vermelho representa valores de LFC positivos e o azul indica valores negativos.

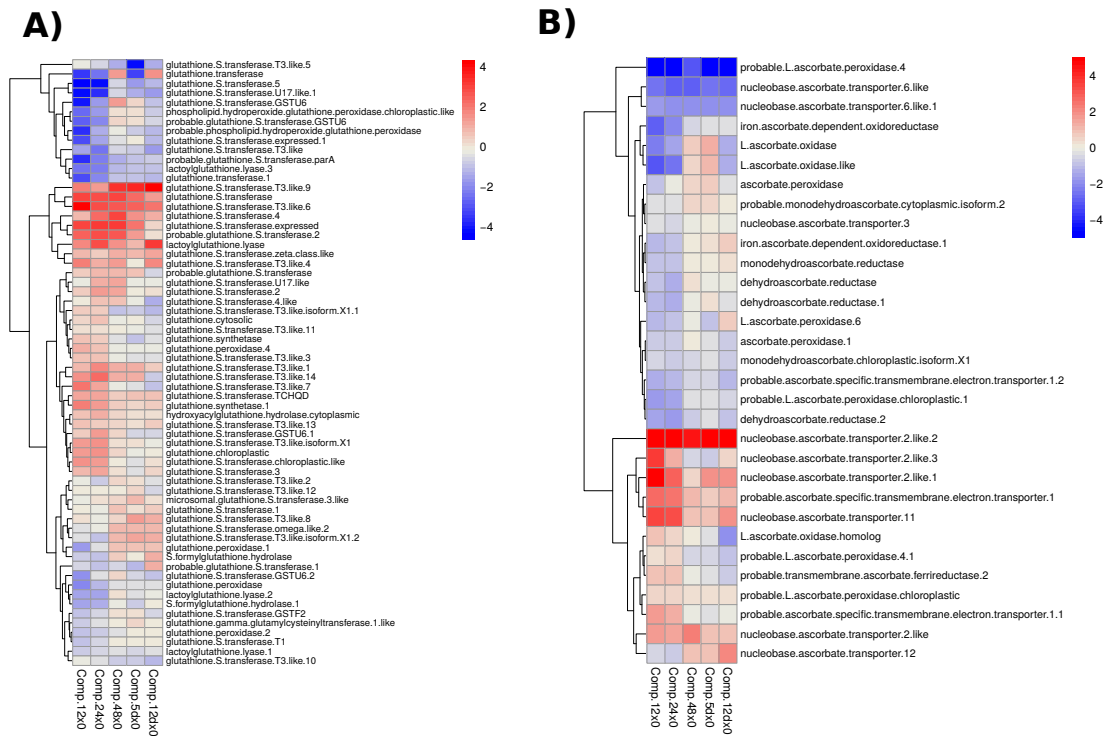


Figura A.8. Heatmaps dos transcritos diferencialmente expressos relacionados às vias de síntese de (A) Glutaciona e (B) Ascorbato. Na barra de cores o vermelho representa valores de LFC positivos e o azul indica valores negativos.

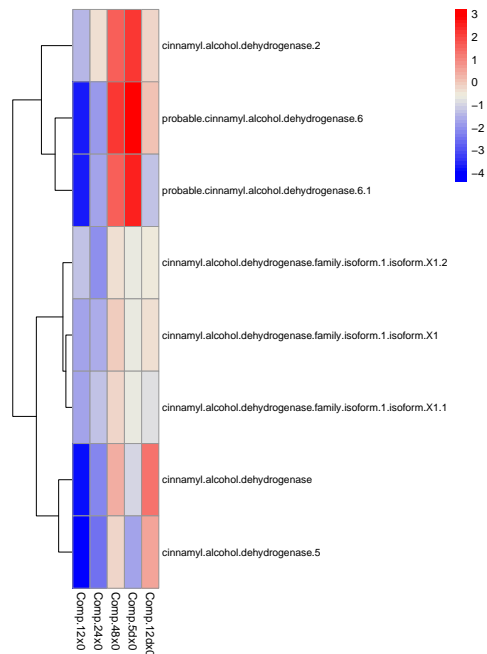


Figura A.9. Heatmap das cinamil álcool desidrogenases encontradas no transcriptoma. Nas primeiras 24 horas a expressão de todas elas eram inferiores aos níveis encontrados em 0h. Após 48 horas todas elas demonstraram estímulo, mas nem todas mantiveram os níveis de expressão elevados até 5 dai. Na barra de cores o vermelho representa valores de LFC positivos e o azul indica valores negativos.

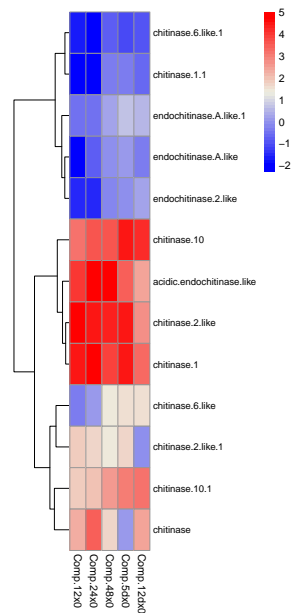


Figura A.10. *Heatmaps* dos transcritos diferencialmente expressos identificados como quitinase. Na barra de cores o vermelho representa valores de LFC positivos e o azul indica valores negativos.

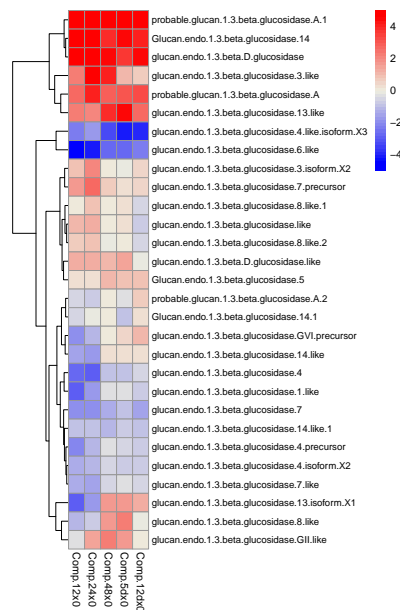


Figura A.11. *Heatmaps* dos transcritos diferencialmente expressos identificados como β -1,3-glucanases. Na barra de cores o vermelho representa valores de LFC positivos e o azul indica valores negativos.

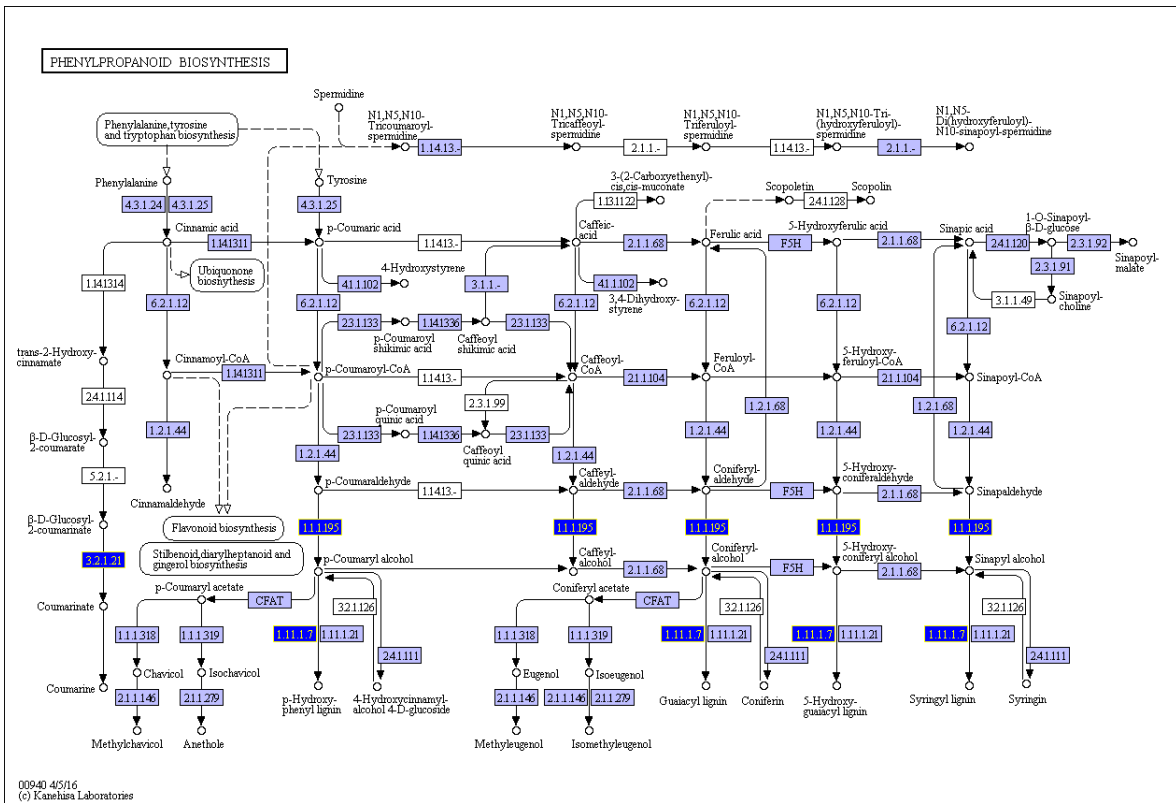


Figura A.12. Via dos fenilpropanoides, na comparação 12 hai vs 0 h. São destacadas as proteínas reprimidas, em azul, conforme os transcritos diferencialmente expressos associados aos identificadores do KEGG Orthology (KO).

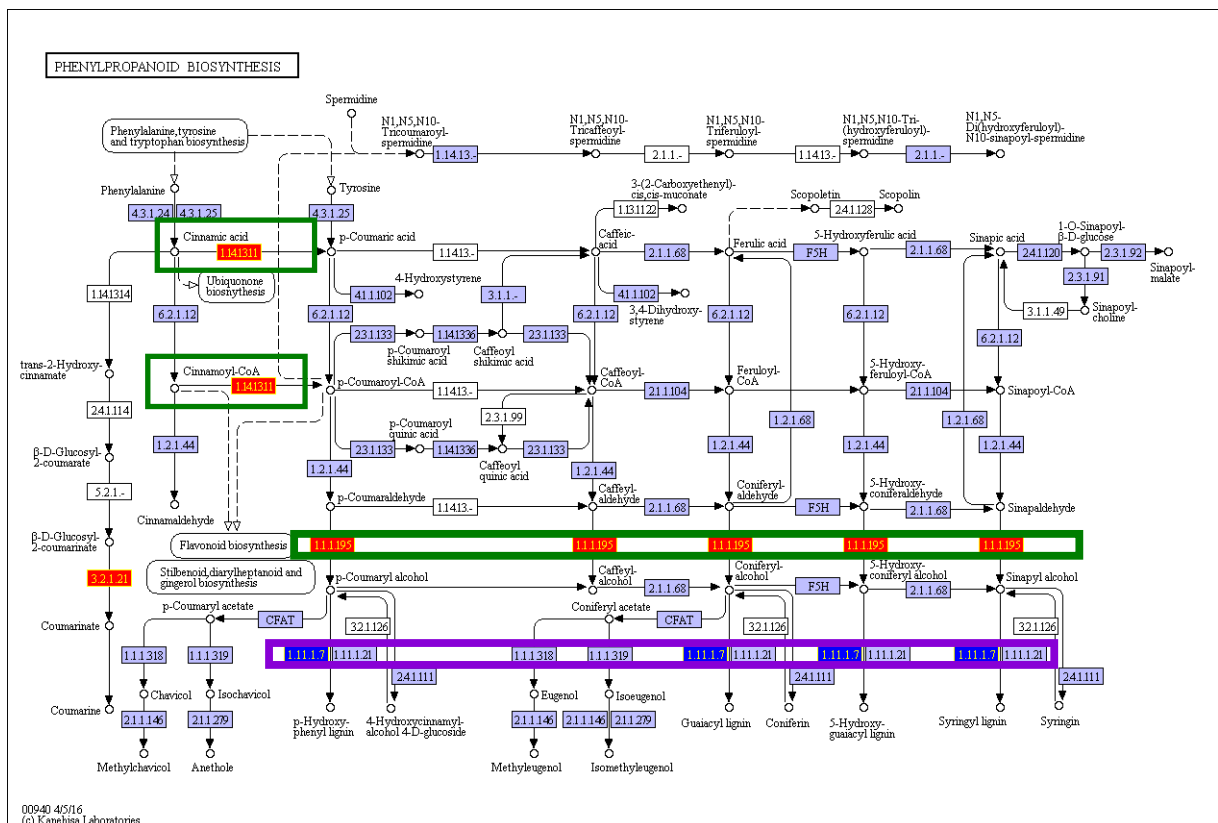


Figura A.13. Via dos fenilpropanoides, na comparação 48 hai vs 24 hai. São destacadas as proteínas estimuladas, em vermelho, e reprimidas, em azul, conforme os transcritos diferencialmente expressos associados aos identificadores do KEGG Orthology (KO). Os quadros verdes e o quadro lilás destacam as proteínas da via de biossíntese da lignina.

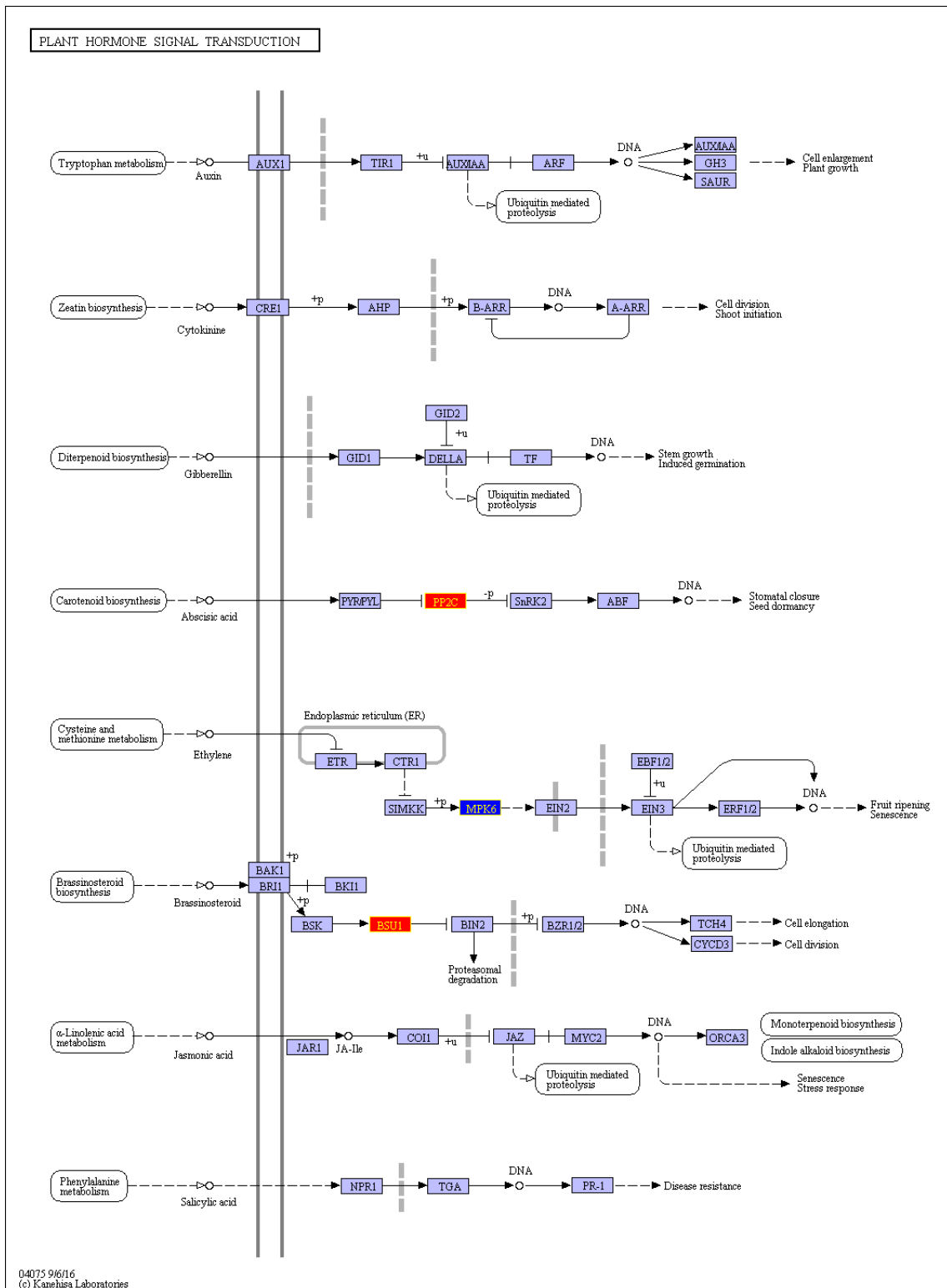


Figura A.14. Via da transdução de sinal em plantas na comparação de 12 hai vs 0h. Destacam-se as proteínas estimuladas, em vermelho, e reprimidas, em azul, conforme os transcritos diferencialmente expressos associados aos identificadores do KEGG Orthology (KO).

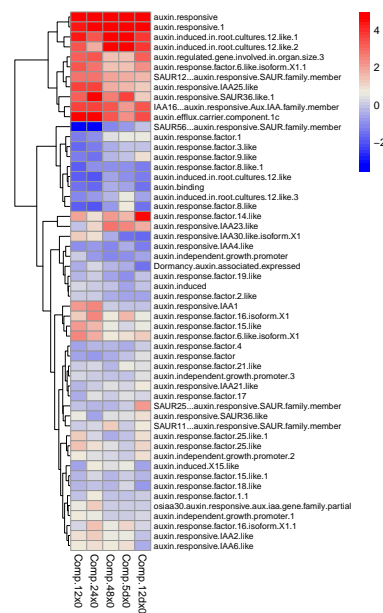


Figura A.15. Heatmaps dos transcritos diferencialmente expressos com descrição associada às auxinas. Na barra de cores o vermelho representa valores de LFC positivos e o azul indica valores negativos.

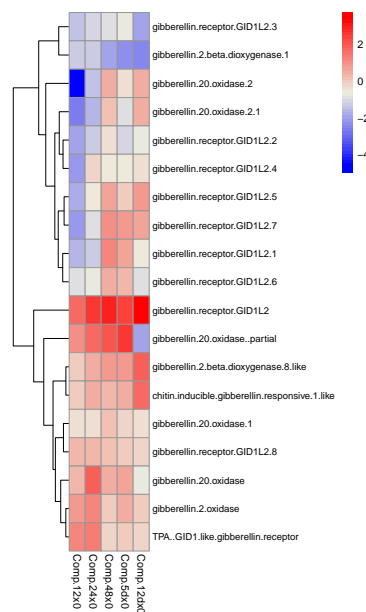


Figura A.16. Heatmaps dos transcritos diferencialmente expressos com descrição associada às giberelinas. Na barra de cores o vermelho representa valores de LFC positivos e o azul indica valores negativos.

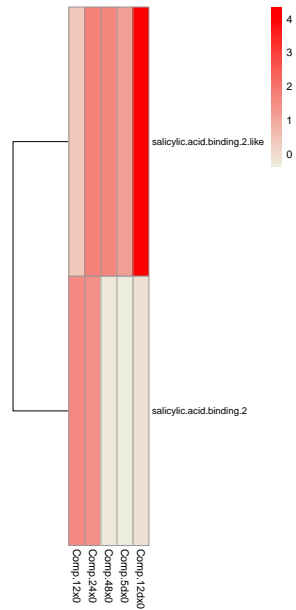


Figura A.17. Heatmaps dos transcritos diferencialmente expressos com descrição associada ao ácido salicílico. Na barra de cores o vermelho representa valores de LFC positivos e o azul indica valores negativos.

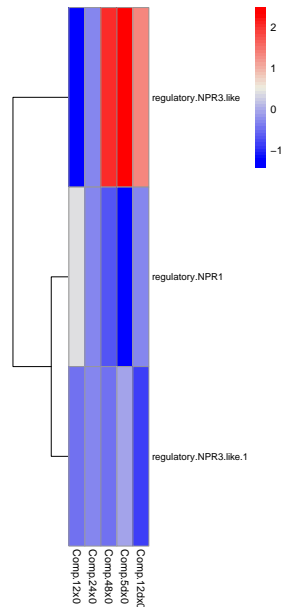


Figura A.18. Heatmaps dos transcritos diferencialmente expressos identificados como NPR (*nonexpressor of pathogenesis related proteins*). Na barra de cores o vermelho representa valores de LFC positivos e o azul indica valores negativos.