

**UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE QUÍMICA
Programa Interunidades de Pós-Graduação em
Bioinformática**

DAVI TOSHIO INADA

**ANÁLISE DO METABOLOMA E TRANSCRIPTOMA DA CANA-DE-
AÇÚCAR AO LONGO DO CICLO DE MATURAÇÃO DA PLANTA**

Versão Corrigida da Dissertação

**SÃO PAULO
2016**

DAVI TOSHIO INADA

**ANÁLISE DO METABOLOMA E TRANSCRIPTOMA DA CANA-DE-
AÇÚCAR AO LONGO DO CICLO DE MATURAÇÃO DA PLANTA**

Dissertação de mestrado apresentada ao
Programa Interunidades de Pós-
Graduação em Bioinformática da
Universidade de São Paulo, como parte
dos requisitos para a obtenção do título
de Mestre em Bioinformática

Área de concentração: Bioinformática

Orientadora: Prof^a. Dr^a Glaucia Mendes
Souza

Co-orientador: Prof. Dr. André Fujita

SÃO PAULO

2016

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Assinatura: _____ Data ___/___/___

Nome: Davi Toshio Inada

Título: **Análise do metaboloma e transcrito da cana-de-açúcar ao longo do ciclo de maturação da planta.**

Dissertação apresentada ao Programa Interunidades de Pós-Graduação em Bioinformática da Universidade de São Paulo para obtenção do título de Mestre em Bioinformática.

Aprovado em: ___ / ___ / ___

Banca Examinadora

Prof^(a). Dr^(a). _____ Instituição: _____
Julgamento: _____ Assinatura: _____

Prof^(a). Dr^(a). _____ Instituição: _____
Julgamento: _____ Assinatura: _____

Prof^(a). Dr^(a). _____ Instituição: _____
Julgamento: _____ Assinatura: _____

DEDICATÓRIAS

Aos meus pais Koji Inada e Ivone Toshiko Ikeda Inada, dedico todas as conquistas realizadas ao longo da minha vida.

Aos meus amigos e colegas de trabalho, pelo companheirismo e pelas muitas risadas.

Aos meus amigos de Tomé-Açu e Belém, por todo o apoio que me deram ao longo de toda a jornada.

À família Yanagihara e a todos os membros da Igreja Aliança, por me acolherem calorosamente aqui em São Paulo, pelo amor, pelo carinho e por terem sido a minha família enquanto distante da minha terra natal.

AGRADECIMENTOS

À minha orientadora, Prof^a. Dr^a. Glaucia Mendes Souza, pela orientação no mestrado e por todo aprendizado adquirido ao longo do desenvolvimento do projeto. Agradeço ainda pelas inúmeras oportunidades concebidas para participar de importantes eventos científicos, que foram primordiais na minha formação pessoal e intelectual.

Ao meu co-orientador, Prof. Dr. André Fujita, por todo o apoio dado ao meu projeto e pelos inúmeros conselhos que me auxiliaram a direcionar minha pesquisa. Obrigado ainda pelos valiosos conselhos de vida, que carregarei ao longo de toda a minha carreira.

À Dr^a. Carolina Gimiliani Lembke Horta, mais conhecida como Carol, obrigado por toda ajuda ao longo do mestrado. Sua ajuda foi tão expressiva que seria impossível listá-las aqui. Que o mundo tenha mais pessoas como você, em especial o mundo acadêmico.

Ao Dr. Milton Yutaka Nishiyama Jr., por toda ajuda e suporte na bioinformática e pelas valiosas dicas que me ajudaram a direcionar meu estudo.

Ao Pablo, mais conhecido como Bruto ou Frango, obrigado pela grande amizade e pela grande colaboração no meu aprendizado. Graças a você pude aprender muito sobre programação em R e ainda melhorar o meu inglês.

Ao Ricardo Silva, muito obrigado pela grande ajuda que me deste no início do mestrado. Por ser a pessoa humilde que és sempre disposto a ajudar pessoas.

À Prof^a. Dr^a. Monalisa Carneiro, pela colaboração nos experimentos de campos, que propiciaram o material necessário para conduzir este estudo.

Ao Prof. Dr. Pio. Colepicolo, junto ao Leonardo e Eliezer, pela colaboração nas análises de metabólômica da cana-de-açúcar. Obrigado por toda a colaboração intelectual e suporte ao longo do desenvolvimento deste estudo.

Agradeço às entidades de fomento Comissão de Aperfeiçoamento de Pessoal do Nível Superior - CAPES (Cota Institucional, 06/2014 a 08/2014) e Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP (processo 2014/078287, 09/2014 a 04/2016 e temático FAPESP nº: 2008/52146-0) pelo suporte financeiro direcionado ao desenvolvimento do projeto de mestrado.

*"Quanto maiores somos em **humildade**, tanto mais próximos estamos da grandeza".*

(Rabindranath Tagore)

Inada D. T. Análise do metaboloma e transcriptoma da cana-de-açúcar ao longo do ciclo de maturação da planta. [dissertação]. São Paulo: Instituto de Química, Universidade de São Paulo; 2016.

RESUMO

A cana-de-açúcar contribui significativamente na economia brasileira, sendo utilizada como fonte alternativa de combustível e para a produção de açúcar, além de apresentar grande potencial para contribuir na substituição dos combustíveis fósseis causadores do efeito estufa. No Brasil, a genômica funcional está avançada em cana-de-açúcar. A identificação de um vasto repertório de genes associados ao acúmulo de sacarose e tolerância ao estresse hídrico levou à elaboração de hipóteses com potencial biotecnológico. O presente estudo visou analisar conjuntamente os dados de metabolômica e transcriptômica em tecido de folha +1 obtido ao longo do amadurecimento em campo, utilizando a plataforma LC-MS e microarranjos respectivamente. Para isso, ferramentas de bioinformática para identificação dos metabólitos e estimação de atividade das vias foram utilizadas, e foram iniciados estudo para a integração dos dados de transcriptoma e metaboloma. Um total de 115 compostos não ambíguos foi identificado (195 candidatos considerando-se a ambiguidade). No transcriptoma da cana-de-açúcar, foram detectados 1587 transcritos expressos, dos quais 841 foram selecionados por estarem descritos como pertencentes à pelo menos uma via metabólica. No metaboloma, observou-se grande diferença na atividade das vias metabólicas associadas ao metabolismo de carboidratos e metabolismo secundário ao longo do crescimento. A análise dos metabólitos pertencentes às vias com diferença significativa no nível de atividade revelou um perfil característico de resposta aos períodos de déficit hídrico. De forma similar, boa parte dos genes avaliados nas vias metabólicas com diferença significativa em nível de transcriptoma demonstrou estar

relacionados ao processo de resposta ao estresse abiótico. Os resultados obtidos no presente estudo forneceram informações importantes dos mecanismos moleculares que podem estar modulando o crescimento da cana-de-açúcar sob diferentes condições ambientais. Tais informações podem servir como base para um melhor planejamento em futuros plantios de cana-de-açúcar visando obter maior produtividade, bem como para a adoção da engenharia genética no desenvolvimento de cultivares melhoradas.

Inada D.T. Analysis of sugarcane metabolome and transcriptome along the maturation cycle [dissertation]. São Paulo: Instituto de Química, Universidade de São Paulo; 2016.

ABSTRACT

Sugarcane has a significant contribution in the Brazilian economy, it can be used as an alternative fuel source or for the sugar production and presents a great potential to overcome the challenges related to fossil fuels that cause the greenhouse effect. In Brazil, molecular studies have been conducted with the sugarcane. The identification of a vast repertoire of genes associated with the accumulation of sucrose and tolerance to water stress, led to the development of hypotheses on the potential application of genetic engineering; many of which require the analysis of the metabolome of sugarcane. This study we aimed to analyze together the metabolomics and transcriptomics from +1 leaf tissue data obtained during ripening by using LC-MS platform and microarray respectively. Bioinformatics tools for the identification of metabolites and pathways activity estimation were used, as well as for the integration of data from transcriptome and metabolome. A total of 115 unambiguous compounds were identified (195 candidates considering ambiguous candidates). In sugarcane transcriptome we identify a total of 1587 expressed transcripts of which 841 were selected because they are described as belonging to at least one metabolic pathway. The metabolome revealed a great difference in the activity of the metabolic pathways associated with the metabolism of carbohydrates and secondary metabolism during growth. The analysis of metabolites belonging to the routes with a significant difference in activity level revealed a characteristic response profile to drought. Similarly, some of the genes evaluated in the metabolic pathways with significant showed to be related to abiotic stress response. The results obtained in this study provided important information on the molecular mechanisms

that are probably modulating the growth of sugarcane under different environmental conditions. This information can serve as a basis for better planning for future sugarcane plantations to obtain higher productivity, as well as the adoption of genetic engineering in the development of improved cultivars.

LISTA DE ABREVIÇÕES

ANOVA (Análise de Variância)

EST (Etiqueta de Sequência Expressa)

CTW (Transformação Contínua Walevet)

DoE (Planejamento do Experimento)

ESI (Ionização por Eletro Spray)

FDR (*False Discovery Ratio*)

FMA (Fase Móvel A)

FMB (Fase Móvel B)

FT (Fatores de Transcrição)

GC-MS (Cromatografia Gasosa Acoplada à Espectrometria de Massas)

HCT (*Hydroxycinnamoyl-CoA Shikimate/quininate Hydroxycinnamoyl Transferase*)

HvBCAT (Aminotransferase de Aminoácidos de Cadeia Ramificada)

IPO (*Isotopologue Parameter Optimization*)

KEGG (*Kyoto Encyclopedia of Genes and Genomes*)

LC-MS (Cromatografia Líquida Acoplada à Espectrometria de Massas)

LSD (Menor Diferença Significante)

PAPi (*Pathway Activity Profiling*)

PK's (Proteínas Quinases)

PP's (Proteínas fosfatases)

qRT-PCR (PCR quantitativo em tempo real)

RNA-Seq (Sequenciamento de RNA)

SAS (Sugarcane Assembled Sequence)

SFF (Sacarose Fosfato Fosfatase)

SFS (Sacarose Fosfato Sintase)

SFS (Sacarose Fosfato Sintase)

SUCEST (*Dissecting the Sugarcane Expressed Sequence Tag*)

TCA (Ciclo do Ácido Cítrico)

LISTA DE ILUSTRAÇÕES

- Figura 1** - Fases de desenvolvimento da cana-de-açúcar até a sua maturação. 9
- Figura 2** - Desenho experimental dos plantios de cana-de-açúcar. O presente estudo é composto por dois plantios de mesmo modelo, com um total de 8 parcelas cada. Os plantios foram realizados em períodos subsequentes de 13 meses cada. O primeiro experimento se iniciou em um período de baixa pluviosidade, enquanto que o segundo plantio em um período de alta pluviosidade. Em cada experimento foram realizadas 4 coletas ao longo dos meses. Em cada coleta foram obtidas amostras de entrenós e folha +1. 17
- Figura 3** - Dados meteorológicos (temperatura e precipitação) mensais ao longo dos dois experimentos de campo. Em a), temos experimento com início em Abril de 2012. E em b), temos experimento com início em Outubro de 2012. Os meses destacados com a estrela em cor preta representam os períodos de plantio, enquanto que os meses destacados pelas setas em cor laranja, representam os meses da coleta de amostras (4 meses, 8 meses, 11 meses e 13 meses). 17
- Figura 4** - Estimação dos períodos que apresentaram déficit hídrico ao longo dos experimentos campo I e II. O primeiro plantio, destacado em cor verde, foi iniciado em Abril de 2012 (P) e finalizou em Maio de 2013. O segundo plantio, destacado na cor laranja, teve início em Outubro de 2012 (P) e finalizou em Novembro de 2013. Em ambos os experimentos de campo, amostras de diferentes tecidos e medidas morfológicas e fisiológicas foram coletadas quando as plantas apresentavam 4 meses, 8 meses, 11 meses e 13 meses de idade. No gráfico, os intervalos destacados em azul indicam um período favorável em termos de precipitação, enquanto que nos intervalos destacados em vermelho indicam período de déficit hídrico. 18
- Figura 5** - Extração de metabólitos. Metodologia adaptada de GIAVALISCO e colaboradores (GIAVALISCO et al., 2011). Foi utilizado MeOH:MTBE:H₂O (1:3:1) para realizar a separação dos metabólitos conforme a sua propriedade química. 26
- Figura 6** - Desenho experimental para comparação da atividade das vias metabólicas nos experimentos de campo I e II. As setas indicam as comparações realizadas entre cada ponto do experimento, sejam elas entre idades equivalentes ou ao longo do crescimento da planta. 41
- Figura 7** - Informações gerais do *pipeline* utilizado na análise dos dados do transcriptoma e metaboloma da cana-de-açúcar. Na análise do metaboloma da cana-de-açúcar as ferramentas XCMS, CAMERA e ProMetab foram utilizadas para a detecção de picos, anotação de íons e anotação de compostos respectivamente. Os dados do transcriptoma foram analisados pelas ferramentas disponíveis no banco de dados SUCEST- FUN. A análise estatística consistiu no cálculo do escore de atividade das vias a partir dos dados de metabolômica e transcriptômica. Também foi realizada uma análise de clusterização hierárquica sobre os escores obtidos. A comparação por teste t de *Student* foi aplicado sobre os escores no intuito de se evidenciar vias com diferença significativa ao longo do processo de maturação. Os dados foram integrados nas vias metabólicas para auxiliar na investigação dos possíveis eventos biológicos. 44
- Figura 8** - Medidas morfológicas da cana-de-açúcar coletadas ao longo do plantio de ambos os experimentos de campo (4 meses, 8 meses, 11 meses e 13 meses). O experimento I se refere ao plantio iniciado em Abril de 2012 e o experimento II se refere ao plantio realizado em Outubro de 2012. Para cada ponto de coleta de ambos os experimentos foram avaliados testes estatísticos de significância como o ANOVA e comparação múltipla de LSD (Fisher's

Least Significant Difference). Os pontos significativamente diferentes estão destacados com sinal de asterisco. 46

Figura 9 - Medidas fisiológicas da cana-de-açúcar coletadas ao longo do plantio de ambos os experimentos de campo (4 meses, 8 meses, 11 meses e 13 meses). O experimento I se refere ao plantio iniciado em Abril de 2012 e o experimento II se refere ao plantio realizado em Outubro de 2012. Para cada ponto de coleta de ambos os experimentos foram avaliados testes estatísticos de significância como o ANOVA e comparação múltipla de LSD (Fisher's Least Significant Difference). Os pontos significativamente diferentes estão destacados com sinal de asterisco. 47

Figura 10 - Distribuição do coeficiente de variação na abundância relativa entre as réplicas biológicas dos metabólitos detectados nos experimentos de campo I e II. Estimativa realizada sem a remoção de réplicas biológicas. 50

Figura 11 - Distribuição do coeficiente de variação na abundância relativa entre as réplicas biológicas dos metabólitos detectados nos experimentos de campo I e II; considerando-se a remoção de uma réplica biológica causadora de maior variação no CV. A linha em vermelha no eixo Y do gráfico, indica o ponto em que o coeficiente de variação é de 30%..... 51

Figura 12 - Quantidade de vias em que os genes de cana-de-açúcar expressos e metabólitos detectados participam, demonstrando assim um consenso de 60 vias metabólicas com pelo menos um gene e um metabólito identificado. 52

Figura 13 - Clusterização hierárquica dos escores obtidos para as vias mapeadas no experimento de campo. As colunas estão dispostas em ordem crescente para a idade da planta no momento da coleta, onde coleta 1, coleta 2, coleta 3 e coleta 4 referem-se respectivamente às idades de 4, 8, 11 e 13 meses da cana-de-açúcar. Nas linhas, as vias metabólicas foram agrupadas (A, B, C e D) conforme a similaridade entre os seus escores ao longo do experimento de campo I. 71

Figura 14 - Metabolismo associado às vias com diferença significativa na atividade nas comparações entre amostras de diferentes pontos de coleta do experimento de campo I. Cada coluna representa a quantidade de vias (N) com diferença significativa na atividade para a comparação entre amostras em idades diferentes (4 meses e 8 meses, 8 meses e 11 meses e 11 meses e 13 meses). Os valores relativos dentro da coluna representam a proporção do metabolismo em que as vias significativas estão associadas..... 72

Figura 15 - Clusterização hierárquica dos escores obtidos para as vias mapeadas no experimento de campo II. As colunas estão dispostas em ordem crescente para a idade da planta no momento da coleta, onde coleta 1, coleta 2, coleta 3 e coleta 4 referem-se respectivamente às idades de 4, 8, 11 e 13 meses da cana-de-açúcar. Nas linhas, as vias metabólicas foram agrupadas (A, B, C, D, E, F, G, H e I) conforme a similaridade entre os seus escores ao longo do experimento de campo II. 74

Figura 16 - Metabolismo associado às vias com diferença significativa na atividade nas comparações entre amostras de diferentes pontos de coleta do experimento de campo I. Cada coluna representa a quantidade de vias (N) com diferença significativa na atividade para a comparação entre amostras em idades diferentes (4 meses e 8 meses, 8 meses e 11 meses e 11 meses e 13 meses). Os valores relativos dentro da coluna representam a proporção do metabolismo em que as vias significativas estão associadas..... 76

LISTA DE QUADROS E TABELAS

- Quadro 1** - Amostras utilizadas no estudo do metaboloma da cana-de-açúcar. A abreviação Cn se refere às coletas realizadas em campo; C1 para coleta aos 4 meses, C2 para 8 meses, C3 para 11 meses e C4 para 13 meses. A abreviação F1 é designada para tecido de folha +1; Mn para informar que se tratam de amostras para análise do metaboloma.....20
- Quadro 2** - Amostras utilizadas para gerar os dados de transcriptômica pela tecnologia de microarranjo. Para cada ponto de coleta de ambos os experimentos de campo, foram selecionadas duas amostras. Dois *pools* de amostra foram gerados a partir de cada par, de forma que uma das amostras e um dos *pools* sejam hibridizados pelos corantes Cy3 e Cy5. As abreviações Cn se referem-aos pontos de coletas do campo I, ao passo que 2.Cn do campo II.....21
- Tabela 1** - Quantidade de vias, enzimas, reações e metabólitos descritos através de curadoria manual em espécies de plantas.....6
- Tabela 2** - Informações sobre gradiente da corrida cromatográfica para a coluna C18. As abreviações significam: FMB (Fase Móvel B) e FMA (Fase Móvel A). A corrida foi realizada a um fluxo de 0.4 ml/min durante um tempo total de 23 min, à temperatura de 40°C.....24
- Tabela 3** - Informações sobre gradiente da corrida cromatográfica para a coluna HILIC. FMB (Fase Móvel B) e FMA (Fase Móvel A). A corrida foi realizada a um fluxo de 0.5 ml/min durante um tempo total de 30 min, a temperatura de 40°C.....24
- Tabela 4** - Abundância relativa dos 115 compostos mais prováveis identificados em tecidos de folha +1 de cana-de-açúcar no experimento de campo I. Para cada metabólito identificado foi estimado a sua abundância relativa normalizada contra o padrão interno dos respectivos modos de aquisição (positivo/ negativo) e contra a concentração de proteínas. Os valores indicados à direita de cada intensidade normalizada representa o desvio padrão em relação à média entre as réplicas biológicas. A abreviação C1 indica "campo I"; M4, M8, M11 e M13 indicam as coletas aos 4 meses, 8 meses, 11 meses e 13 meses após o plantio..... 53
- Tabela 5** - Abundância relativa dos 115 compostos identificados em tecidos de folha +1 de cana-de-açúcar no experimento de campo II. Para cada metabólito identificado foi estimado a sua abundância relativa normalizada contra o padrão interno dos respectivos modos de aquisição (positivo/ negativo) e contra a concentração de proteínas. Os valores indicados à direita de cada intensidade normalizada representa o desvio padrão em relação à média entre as réplicas biológicas. A abreviação C2 indica "campo II"; M4, M8, M11 e M13 indicam as coletas aos 4 meses, 8 meses, 11 meses e 13 meses após o plantio.....59
- Tabela 6** - Metabólitos que apresentaram diferença significativa na abundância relativa ao longo do experimento de campo I. Os valores P e de correção (FDR) se referem ao teste de significância (P-valor < 0.05) para a análise de variância (ANOVA). Para cada metabólito com diferença significativa ao longo do experimento

também foram avaliadas as diferenças na abundância relativa nos intervalos descritos (4 meses, 8 meses, 11 meses e 13 meses de idade).66

Tabela 7 - Metabólitos que apresentaram diferença significativa na abundância relativa ao longo do experimento de campo II. Os valores P e de correção (FDR) se referem ao teste de significância (P-valor < 0.05) para a análise de variância (ANOVA). Para cada metabólito com diferença significativa ao longo do experimento também foram avaliadas as diferenças na abundância relativa nos intervalos descritos (4 meses, 8 meses, 11 meses e 13 meses de idade).67

SUMÁRIO

1 INTRODUÇÃO.....	1
1.1 A CANA-DE-AÇÚCAR E O SEU CONTEXTO ECONÔMICO	1
1.2 CIÊNCIAS ÔMICAS APLICADAS AO ESTUDO DE PLANTAS	3
1.2.1 A transcriptômica	3
1.2.2 A metabolômica	4
1.3 O CICLO DE MATURAÇÃO E O ACÚMULO DE SACAROSE NA CANA-DE-AÇÚCAR	8
1.4 MOTIVAÇÃO PARA O DESENVOLVIMENTO DE UM ESTUDO HOLÍSTICO DA CANA-DE-AÇÚCAR.....	12
2 OBJETIVOS	15
2.1 OBJETIVO GERAL.....	15
2.2 OBJETIVOS ESPECÍFICOS.....	15
3 MATERIAIS E MÉTODOS	16
3.1 EXPERIMENTO DE CAMPO.....	16
3.2 AMOSTRAS UTILIZADAS NO ESTUDO.....	19
3.3 TRANSCRIPTOMA DA CANA-DE-AÇÚCAR.....	21
3.3.1 Hibridização em microarranjo de oligonucleotídeos	21
3.3.2 Análise da expressão gênica.....	22
3.4. METABOLOMA DA CANA-DE-AÇÚCAR.....	23
3.4.1 Cromatografia líquida acoplada à espectrometria de massa	23
3.4.2 Extração de metabólitos	25
3.4.3 Quantificação de proteínas	27
3.5 ANÁLISE COMPUTACIONAL DOS DADOS	28
3.5.1 Otimização de parâmetros para detecção e agrupamento de picos.....	28
3.5.2 Identificação e agrupamento de picos	31
3.5.3 Anotação dos íons.....	34
3.5.4 Anotação dos compostos.....	37
3.5.5 Análise estatística dos compostos identificados	39
3.5.6 Análise da atividade das vias metabólicas a partir metaboloma	39
3.6 DEFINIÇÕES GERAIS DO PIPELINE DE ANÁLISE DO TRANSCRIPTOMA E METABOLOMA	41
4 RESULTADOS	45
4.1 EXPERIMENTO DE CAMPO.....	45

4.2 TRANSCRIPTÔMICA	48
4.2.1 Detecção dos genes significativamente expressos	48
4.3 METABOLÔMICA.....	49
4.3.1 Anotação dos picos de massa.....	49
4.3.2. Perfil da atividade das vias ao longo do experimento de campo I	70
4.3.3 Perfil da atividade das vias ao longo do experimento de campo II	73
5 DISCUSSÃO.....	76
5.1 RELAÇÃO ENTRE OS FATORES ABIÓTICOS E O DESENVOLVIMENTO DA CANA- DE-AÇÚCAR.....	76
5.1.1 NO EXPERIMENTO DE CAMPO I	76
5.1.2 NO EXPERIMENTO DE CAMPO II	83
6 CONCLUSÃO	87
7 PROPOSTAS PARA ESTUDOS FUTUROS.....	89

1 INTRODUÇÃO

1.1 A CANA-DE-AÇÚCAR E O SEU CONTEXTO ECONÔMICO

A cana-de-açúcar (*Saccharum* spp.) pertence à família Poaceae, assim como outros vegetais de importância alimentar como o milho, arroz, trigo e sorgo. O gênero *Saccharum* engloba diversas espécies: *S. officinarum*, *S. spontaneum*, *S. robustum*, *S. sinsensis* e *S. barberi*. Entretanto, no contexto econômico atual, apenas cultivares híbridas são utilizadas, devido à sua capacidade melhorada para o acúmulo de sacarose. Os híbridos são poliplóides e aneuplóides, apresentando cerca de 100 a 200 cromossomos; onde, dependendo da cultivar seu genoma pode chegar entre 769 a 926 Mpb (DE SIQUEIRA FERREIRA et al., 2013).

O Estado de São Paulo é o maior produtor de cana-de-açúcar, que é utilizada prioritariamente como fonte de sacarose para a produção de açúcar de consumo e bioetanol. Parte da extração também é destinada para a produção de bebidas alcoólicas como a cachaça. Além disso, o bagaço da cana-de-açúcar, um produto secundário do processamento, tem sido utilizado como biomassa para combustão, alimento para ruminantes e produção em escala industrial de etanol de segunda geração (EPE, 2015).

Em 2014, a produção de etanol no Brasil cresceu cerca de 4% e alcançou uma produção de 28,6 bilhões de litros, superando assim, o recorde de 2010 com 27,9 bilhões de litros (EPE, 2015). Apesar das recentes crises econômicas no setor, como a baixa competitividade no mercado do etanol em relação à gasolina, estima-se que a cana-de-açúcar sozinha, tenha uma contribuição direta e/ou indireta de cerca de 2,35% anuais no PIB nacional (DIEESE, 2015).

No mundo, mais de 100 países cultivam a cana-de-açúcar para a produção de açúcar, parte deles também utilizam para a produção de bioetanol pelo uso da sacarose para e bioeletricidade pelo uso do bagaço. Com a evolução crescente na tecnologia, a produção de bioetanol de 2ª geração tem demonstrado grande potencial de aplicação na economia de diversos países. Levando-nos à possibilidade de diminuir gradativamente o uso de combustíveis fósseis e direcionando a economia para um campo de maior sustentabilidade (Souza et. al, 2015). Diferente do combustível derivado do petróleo, o bioetanol é uma fonte renovável. Logo, os impactos ambientais são menores em comparação ao uso de fontes não renováveis como o petróleo (GOLDEMBERG; COELHO; GUARDABASSI, 2008; KIDO et al., 2012).

Dada a importância econômica da cana-de-açúcar no Brasil, foi implantado um dos mais notórios projetos de sequenciamento de ESTs (Etiquetas de sequências expressas). O projeto SUCEST (Sugarcane Expressed Sequence Tag), obteve como resultado cerca de 250.000 ESTs da cana-de-açúcar com a utilização de 37 bibliotecas diferentes de cDNA de tecidos distintos (meristema apical, flores, gemas, gemas vegetativas laterais, folhas maduras, raízes, casca do colmo, entrenós e sementes). Cerca de 43.000 clusters foram montados, dos quais 50% se apresentaram relacionados ao metabolismo de proteínas, comunicação celular, transdução de sinal, bioenergética e respostas ao estresse (SOUZA et al., 2001; VETTORE et al., 2001). Este projeto possibilitou ampliar o conhecimento molecular acerca da cana-de-açúcar contribuindo em estudos com potencial no melhoramento genético da planta para obtenção de maior produtividade (FERREIRA et al., 2016; LEMBKE et al., 2012; NISHIYAMA et al., 2014a; PAPINI-TERZI et al., 2009).

1.2 CIÊNCIAS ÔMICAS APLICADAS AO ESTUDO DE PLANTAS

1.2.1 A transcriptômica

O transcriptoma é descrito pelo conjunto de RNAs mensageiros (mRNA) e RNAs não codificantes (ncRNA); este segundo engloba o grupo de micro RNAs (miRNA), pequenos RNAs nucleolares (snoRNA) e pequenos RNAs nucleares (snRNA) (JALALI et al., 2013). A transcrição de genes é sujeita a uma intensa regulação, seja a nível transcricional ou pós-transcricional. A regulação a nível transcricional pode conferir diferentes níveis de expressão a um gene (LOVÉN et al., 2012), os quais podem ser mensurados e assim auxiliar na construção de novas hipóteses em um referido processo biológico (KOGENARU et al., 2012).

Atualmente, existe uma variedade de tecnologias capazes de quantificar a abundância de transcritos gênicos a partir de uma amostra. O *Northern blot* e o qRT-PCR (PCR quantitativo em tempo real) permitem a quantificação da abundância de um único gene por vez (MALONE; OLIVER, 2011). Apesar do qRT-PCR ser uma tecnologia robusta, é extremamente laborioso quando há a necessidade de se medir os níveis de expressão de uma grande quantidade de genes concomitantemente. Duas outras tecnologias são capazes de analisar o transcriptoma de um organismo em um único experimento: a tecnologia de microarranjo de DNA e o Sequenciamento de RNA (RNA-Seq) (MALONE; OLIVER, 2011). O microarranjo não possui a capacidade de identificar novos genes, alelos e isoformas; porém, apesar desta limitação, possui um menor custo e continua desempenhando um papel importante em estudos que visam observar a expressão de genes conhecidos em diferentes condições.

A tecnologia de microarranjo para a análise de expressão de genes consiste em fixar em uma lâmina sequências de oligonucleotídeos de DNA (sondas). Na

lâmina, milhões de sondas são depositadas. As sondas presentes na lâmina possuem a capacidade de hibridizar com as sequências complementares de mRNA (cDNA), as quais estão marcadas com moléculas fluorescentes. A abundância do transcrito é proporcional à intensidade de luz emitida, permitindo assim quantificar os níveis relativos da expressão de um gene (MALONE; OLIVER, 2011).

A aplicação da tecnologia de microarranjo é extensa, e engloba um vasto campo das ciências biológicas, servindo de ferramenta para o estudo do transcriptoma em humanos, plantas e microorganismos. Sua aplicação no estudo molecular de plantas tem fornecido um melhor entendimento dos mecanismos de resposta a fatores abióticos, como o estresse salino (ZHOU et al., 2016) e tolerância à seca (GONG et al., 2015); crescimento e acúmulo de moléculas com potencial de interesse, como a sacarose e lignina em cana-de-açúcar (FERREIRA et al., 2016; LEMBKE et al., 2012; PAPINI-TERZI et al., 2009).

1.2.2 A metabolômica

Metabólitos são produtos finais derivados do processo de controle da expressão gênica, tradução de proteínas e atividades enzimáticas (ALLWOOD; ELLIS; GOODACRE, 2008; ROCHFORT, 2005). Sendo de baixo peso molecular, com tamanho estimado < 900 Da (GOTTESFELD et al., 1997).

Os metabólitos podem ser divididos em duas classes. Os metabólitos primários correspondem à classe de compostos químicos essenciais para o crescimento e desenvolvimento do organismo, como os açúcares, aminoácidos, nucleotídeos, ácidos orgânicos e vitaminas. Os secundários correspondem três grandes grupos principais: alcalóides, terpenos e compostos fenólicos, que

assumem funções específicas desempenhando um papel crucial na estrutura e defesa para a sobrevivência do organismo das plantas (LAVINSKY et al., 2015).

Conforme as informações fornecidas pelo banco de dados biológicos PlantCyc v11.0 (<http://www.plantcyc.org/>), onde os dados são curados manualmente, é reportado a existência de 1213 vias metabólicas, 140581 enzimas, 3482 reações e 2802 metabólitos em plantas com as funções biológicas definidas. A quantidade de metabólitos descritos varia de acordo com a espécie de planta estudada. Em algumas gramíneas como o sorgo, milho e arroz, são descritos um total de 2276, 2525 e 2487 metabólitos respectivamente (tabela 1). Entretanto, sabemos que esta quantidade é na realidade muito maior, dado o conhecimento limitado da diversidade metabólica (AFENDI et al., 2012). O conhecimento científico acumulado indica que existem aproximadamente 270.000 espécies de plantas, sem contar a existência de espécies ainda não identificadas (SCOTLAND; WORTLEY, 2003). Dessa forma, estima-se que a diversidade de metabólitos existentes no reino das plantas gire em torno de 200.000 a 1.000.000 moléculas (AFENDI et al., 2012).

Tabela 1 - Quantidade de vias, enzimas, reações e metabólitos descritos através de curadoria manual em espécies de plantas.

Planta	Espécie	Vias	Enzimas	Reações	Metabólitos
Aegilops	<i>Aegilops tauschii</i>	540	6552	3059	2335
Arabidopsis	<i>Arabidopsis thaliana</i>	619	5472	3482	2802
Arroz	<i>Oryza sativa</i>	543	6198	3223	2487
Batata	<i>Solanum tuberosum</i>	558	5790	3122	2413
Brachypodium	<i>Brachypodium distachyon</i>	531	5319	3004	2285
Cassava	<i>Manihot esculenta</i>	567	6371	3183	2386
Cevada	<i>Hordeum vulgare</i>	540	4662	3033	2942
Chlamydomonas	<i>Chlamydomonas reinhardtii</i>	373	2768	2435	1750
Gramínea	<i>Panicum virgatum</i>	587	14006	3379	2581
Mamão	<i>Carica papaya</i>	538	4345	3111	2329
Milho	<i>Zea mays</i>	545	6778	3151	2525
Musgo	<i>Physcomitrella patens</i>	473	4952	2801	2035
Poplar	<i>Populus trichocarpa</i>	599	7996	3305	2509
Repolho	<i>Brassica rapa ssp.</i>	533	7825	3146	2331
Selaginella	<i>Selaginella moellendorffii</i>	481	4840	2811	2105
Setária	<i>Setaria italica</i>	549	6045	3028	2284
Soja	<i>Glycine max</i>	596	10607	3387	2636
Sorgo	<i>Sorghum bicolor</i>	540	5939	3017	2276
Spirodela	<i>Spirodela polyrhiza</i>	496	3638	2722	2045
Tomate	<i>Solanum lycopersicum</i>	588	6503	3288	2550
Trigo	<i>Triticum urartu</i>	551	6315	3093	2395
Uva	<i>Vitis vinifera</i>	557	5799	3182	2421

Fonte: PlantCyc v11.0 (2016)

O estudo destas pequenas moléculas pode ser conduzido por um conjunto de técnicas, sendo a espectrometria de massas uma das mais utilizadas. A espectrometria de massas pode ser dividida em duas metodologias conforme o

método de separação cromatográfica empregada: LC-MS (Cromatografia Líquida Acoplada à Espectrometria de Massas) e GC-MS (Cromatografia Gasosa Acoplada à Espectrometria de Massas). Os dados de espectrometria são adquiridos em espectros na forma da razão massa-carga (m/z). Cada espectro de massa é detectado em um tempo de retenção em particular e possui sua medida de intensidade, que está associada à abundância relativa do composto na amostra (ALONSO; MARSAL; JULIÀ, 2015).

O estudo no âmbito da metabolômica engloba dois tipos de abordagens: direcionado e não direcionado. O estudo direcionado é focado em um grupo específico de metabólitos, servindo como ferramenta de validação de hipóteses. O estudo não direcionado é uma análise. Este tipo de estudo envolve a produção de um grande volume de dados, aumentando assim a sua complexidade de processamento (NAZ, 2014). Logo, algoritmos computacionais específicos são necessários para o processamento dos dados.

A abordagem não direcionada vem sendo amplamente empregada em estudo de plantas. Por exemplo, Camañes et al. (2015) utilizaram uma abordagem não direcionada para o estudo da interação planta-patógeno, a fim de estabelecer um perfil metabólico do processo de infecção da bactéria *Pseudomonas syringae* em tomates e obter um melhor entendimento dos mecanismos de defesa da planta.

É possível ainda utilizar esta abordagem na elucidação dos mecanismos moleculares da planta em resposta a um estresse abiótico (CHEN; HOEHENWARTER, 2015; NAKABAYASHI; SAITO, 2015), ou ainda, evidenciar os processos regulatórios que ocorrem ao longo do crescimento e amadurecimento da planta (VERDIER et al., 2013).

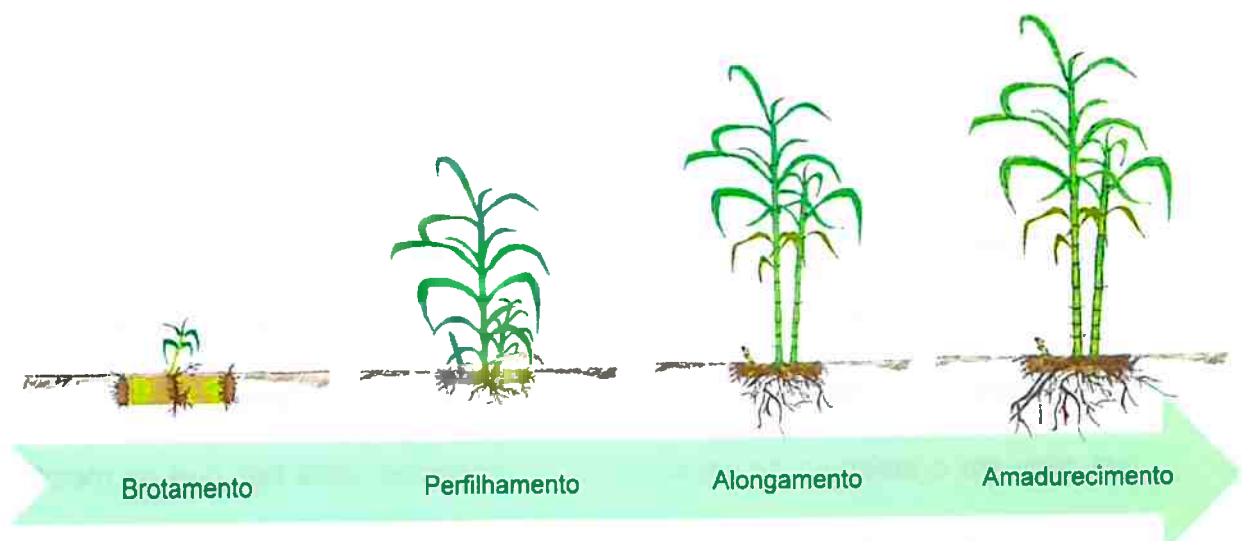
1.3 O CICLO DE MATURAÇÃO E O ACÚMULO DE SACAROSE NA CANA-DE-AÇÚCAR

O entendimento dos mecanismos envolvidos no acúmulo de sacarose na cana-de-açúcar é uma questão chave para o melhoramento de seu cultivo e produtividade. A sacarose é sintetizada pela ação de duas enzimas: a sacarose fosfato sintase (SFS) e a sacarose fosfato fosfatase (SFF). A SFS catalisa a reação reversível da UDP-glicose + frutose 6-fostato para sacarose 6-fosfato + UDP + H⁺. Logo após a reação mediada pela SFS a enzima SFF cataliza a reação de hidrólise da sacarose 6-fosfato para sacarose (WANG et al., 2013).

A sacarose produzida nas folhas fotossintetizantes (fonte) da cana-de-açúcar é a forma principal de carbono transportado até os tecidos dos entrenós jovens e maduros e demais tecidos dreno. Neste ponto ela pode seguir por dois caminhos diferentes: (I) ser armazenada nas células vacuolares ou então, (II) ser utilizada como substrato do processo respiratório, a fim de prover a energia necessária para o crescimento e manutenção da planta (WANG et al., 2013). Esta decisão para onde o fluxo de assimilados fotossintéticos será transportado irá depender da demanda dos tecidos dreno de acordo com o estágio de desenvolvimento da planta, ou ainda de fatores ambientais, como o déficit hídrico (WANG et al., 2013).

O crescimento da cana-de-açúcar é fortemente regulado pelas condições climáticas (temperatura e precipitação). Para um crescimento objetivando o acúmulo máximo de sacarose, a planta necessita de um período inicial com alta temperatura e precipitação média. Estas condições iniciais favorecem o crescimento vegetativo. Entretanto, para que a planta passe a acumular de forma efetiva a sacarose, as condições climáticas que se sucedem ao período de crescimento vegetativo devem ser de baixa temperatura e precipitação (CARDOZO; SENTELHAS, 2013).

O ciclo de crescimento e maturação da cana-de-açúcar incluem as seguintes etapas (figura 1): brotamento, perfilhamento, alongamento e amadurecimento (DOORENBOS & KASSAM, 1979). A primeira etapa consiste no brotamento das gemas presentes no tolete da cana-de-açúcar. Cada tolete apresenta em torno de 3-4 gemas nas regiões dos nós. Na etapa de perfilhamento ocorre um intenso crescimento das gemas formando estruturas denominadas rizomas, que são estruturas similares ao colmo aéreo da cana-de-açúcar adulta.



Fonte: Adaptado de Doorenbos & Kassam, 1979.

Figura 1 - Fases de desenvolvimento da cana-de-açúcar até a sua maturação.

Com os perfilhos desenvolvidos têm-se início à etapa de alongamento, onde se observa um elevado crescimento vertical da planta. A atividade fotossintética e a produção de açúcares são elevadas, porém, a relação de produção na fonte (folhas) e consumo nos drenos (tecidos em formação, colmo em processo de armazenamento de assimilados fotossintéticos), bem como a respiração permanecem em níveis similares, o que impede um acúmulo efetivo de açúcares pela planta nesta etapa do crescimento. Boa parte do carbono fixado pelo processo fotossintético é convertida em sacarose e amido, outra parte é designada para as

vias energéticas do processo respiratório, onde são produzidos ATP e os redutores bioquímicos pelas principais vias, como a via da glicólise, pentose fosfato, ciclo de Krebs e cadeia de transporte de elétrons.

O acúmulo efetivo da sacarose se inicia no período de maturação da cana-de-açúcar, quando o crescimento vertical da planta cessa. Nesta etapa, o envio de assimilados fotossintéticos para os tecidos dreno do colmo é maior que a sua demanda para o crescimento, permitindo assim que a cana-de-açúcar passe a acumular maior quantidade de sacarose no colmo (MCCORMICK; WATT; CRAMER, 2009). É postulado que quando o colmo apresenta saturação no armazenamento, a alta concentração de sacarose neste tecido age como sinalizador na inibição da fotossíntese (MCCORMICK; WATT; CRAMER, 2009). Um efeito similar de inibição da fotossíntese ocorre quando a cana-de-açúcar é submetida ao estresse hídrico ou baixa temperatura média (ROBERTO et al., 2015). Os mecanismos moleculares que regulam o acúmulo de sacarose são complexos, uma vez que os mesmos são dependentes da idade da planta, do genótipo e dos fatores abióticos (PAPINI-TERZI et al., 2009).

Estudos anteriores indicam uma sobreposição de possíveis vias associadas ao metabolismo de carboidratos com as vias de resposta à seca. Ao analisar a maturação do colmo em diferentes genótipos da cana, observou-se que muitos genes relacionados à sinalização de hormônios, resposta à seca, metabolismo da parece celular, transporte de açúcares, proteínas quinases (PKs), proteínas fosfatases (PPs) e fatores de transcrição (TF) podem estar associados à regulação do desenvolvimento da planta ao processo de acúmulo da sacarose e à resposta à seca (LEMBKE et al., 2012; PAPINI-TERZI et al., 2009). As PKs e PPs em particular

se destacam entre genes diferencialmente expressos durante este processo no colmo (PAPINI-TERZI et al., 2009)

No intuito de investigar parte dos genes responsivos ao estresse abiótico e seus possíveis mecanismos de facilitação do acúmulo de sacarose, Iskandar et al., (2011) selecionaram para estudo um conjunto de genes relacionados ao processo de estresse osmótico e hídrico, e observaram a abundância de seus transcritos nos tecidos de folha e colmo ao longo do crescimento. Ao comparar os tecidos, foram identificados genes diferencialmente expressos de proteínas associadas ao metabolismo de aminoácidos e poliols, transporte de açúcares, chaperonas e regulação transcricional.

Os genes codificantes TF's da família bZIP ou ERF se apresentaram mais expressos no colmo. Da mesma forma, os genes que codificam para transportadores de açúcares (PST5, PST7, PST2a e PST2b) também apresentaram maior expressão nos colmos. O gene PST5 apresentou um comportamento característico, diminuindo o nível de expressão em colmos mais jovens (ISKANDAR et al., 2011). Esta observação corroborou com observações anteriores, mostrando que a concentração de açúcares varia entre os tecidos de colmos jovens e maduros, onde a relação de concentração cresce proporcionalmente a idade do colmo (PAPINI-TERZI et al., 2009). Logo, o gene PST5 parece ser um modulador do processo de acúmulo de sacarose. Ainda, torna-se evidente a diferença na expressão de genes entre diferentes tecidos, bem como entre diferentes regiões de um mesmo tecido com idades de maturação diferentes (ISKANDAR et al., 2011; PAPINI-TERZI et al., 2009).

Recentemente, Mattiello et al., (2015) demonstraram resultado similar em folha da cana-de-açúcar. Em seu estudo, observaram que os níveis de expressão de

genes relacionados ao desenvolvimento variam significativamente de acordo com a localização no tecido da folha, sendo a expressão mais intensa na região do topo e decresce no sentido da base. Logo, torna-se evidente a grande complexidade dos mecanismos moleculares que regulam os processos biológicos. A cana-de-açúcar ainda apresenta muitos desafios no entendimento do processo regulatório do acúmulo de sacarose, trazendo assim novas perspectivas para estudos futuros.

1.4 MOTIVAÇÃO PARA O DESENVOLVIMENTO DE UM ESTUDO HOLÍSTICO DA CANA-DE-AÇÚCAR

O grupo vem realizando estudos de genômica funcional da cana-de-açúcar, gerando assim um grande contingente de dados que estão sendo armazenados no banco de dados SUCEST-FUN (<http://sucest-fun.org/wsapp/>). Este banco de dados visa agregar informações biológicas e fornecer meios para o desenvolvimento e aplicação de metodologias para integração e correlação entre as informações e análises.

Atualmente o banco de dados dispõe de ferramentas de análises de expressão gênica, como a identificação de genes expressos e diferencialmente expressos, integrados com análises de enriquecimento de termos baseado em vias, *Gene Ontology* e categorizações. A partir da expressão gênica também é possível estimar o nível de atividade de vias metabólicas e, através da análise de agrupamento hierárquico, identificar padrões de níveis de atividades das vias entre diferentes condições experimentais (NISHIYAMA et al., 2014b).

Como descrito acima, análise do transcriptoma da cana-de-açúcar revelou a expressão de genes associados ao teor de sacarose, fibra, produtividade e resposta à seca. Os dados apontaram para vias metabólicas e processos celulares envolvidos

na síntese de hormônios, lignina, transporte de água, fosforilação de açúcares, regulação da cromatina e da transcrição (FERREIRA et al., 2016; LEMBKE et al., 2012; PAPINI-TERZI et al., 2009). Entretanto, a validação de algumas das hipóteses formuladas requer a análise dos respectivos metabólitos envolvidos.

Ferramentas de análise de dados do metaboloma da cana-de-açúcar, com o fim de realizar a identificação dos possíveis metabólitos existentes em diferentes tecidos da planta, bem como estimar a atividade das vias metabólicas envolvidas partir da abundância dos metabólitos. A integração entre as informações fornecidas pelo transcriptoma e metaboloma da cana-de-açúcar poderá servir como uma importante ferramenta na investigação de futuras hipóteses, ressaltando-se a necessidade de implementá-las no banco de dados.

Um segundo desafio está relacionado à pequena quantidade de estudos no âmbito da metabolômica envolvendo a cana-de-açúcar. Em um primeiro estudo realizado em tecidos de entrenós jovens e maduros de duas variedades (N19 and US6656-15), foi identificado um total de 14 metabólitos pelo uso da tecnologia GC-MS. Ao verificar as diferenças na abundância dos metabólitos entre os tecidos de entrenós jovens e maduros (entrenós 3 e 7 respectivamente), constatou-se que a trealose se apresentava mais ativa em tecidos jovens, indicando assim um possível papel deste metabólito na sinalização de açúcares (BOSCH; ROHWER; BOTHA, 2003).

Posteriormente, um estudo similar foi realizado com a variedade Q117 (GLASSOP et al., 2007). Ao analisar o metaboloma de tecidos do colmo (região do meristema ao entrenó 2 e entrenós 4, 7 e 9) pelo uso da tecnologia GC-MS, um total de 55 metabólitos foi identificado, sendo boa parte deles aminoácidos e compostos

pertencentes ao ciclo do ácido cítrico (TCA). O metabólito trealose também apresentou maior associação ao acúmulo de sacarose (GLASSOP et al., 2007).

Todos os estudos descritos de metabolômica envolvendo a cana-de-açúcar foram realizados em tecidos do colmo, apesar da folha também assumir um importante papel na produção e regulação do transporte de açúcares, bem como no crescimento da planta (MCCORMICK; WATT; CRAMER, 2009). Mesmo nos estudos de transcriptômica da cana-de-açúcar, apenas recentemente foi realizado uma análise aprofundada da folha, uma vez que, a maioria dos estudos descritos na literatura envolvem tecido do colmo (MATTIELLO et al., 2015).

Tendo em vista a necessidade de implementação de ferramentas de bioinformática para análise de dados de metabolômica no bando de dados SUCEST-FUN, somado ao fato de existir uma quantidade limitada de estudos do metaboloma da cana-de-açúcar, iniciou-se o projeto que visa um estudo abrangente do desenvolvimento da cana-de-açúcar em campo (Temático FAPESP nº: 2008/52146-0).

Desta maneira, dois experimentos de campo foram realizados entre o período de Abril de 2012 e Novembro de 2013. Cada plantio foi realizado em épocas diferentes (informações detalhadas estão descritas na sessão de "Materiais e Métodos"), no intuito de se investigar os mecanismos moleculares atuantes no crescimento e acúmulo de sacarose na variedade comercial SP80-3280. O primeiro plantio foi realizado em um período de baixa precipitação e temperatura média (cana de ano-e-meio), ao passo que o segundo plantio teve início em um período de alta precipitação e temperatura média (cana de ano).

Em ambos os experimentos de campo foram coletadas amostras de tecido de raiz, colmo e folha da cana-de-açúcar, assim como medidas morfológicas e

fisiológicas ao longo do desenvolvimento da planta. Foi reunida uma quantidade de dados suficiente para o desenvolvimento de estudos no âmbito da biologia de sistemas, integrando-se transcriptômica, metabolômica e proteômica nesta variedade comercial. Assim, no presente estudo foi desenvolvido um *pipeline* de análise computacional de dados de metabolômica provenientes da plataforma LC-MS. Também foi iniciada a análise dos dados coletados em ambos os experimentos de campo do tecido de folha +1.

2 OBJETIVOS

2.1 OBJETIVO GERAL

- ✓ Utilização de ferramentas de bioinformática para o desenvolvimento de uma plataforma de análise de metabólitos de cana-de-açúcar e integração com dados de transcriptoma.

2.2 OBJETIVOS ESPECÍFICOS

- ✓ Analisar corridas de LC-MS geradas a partir de amostras de uma variedade comercial de cana-de-açúcar em alguns estágios do seu desenvolvimento;
- ✓ Aplicar ferramentas de bioinformática para identificação dos metabólitos;
- ✓ Aplicar e/ou implementar ferramentas de bioinformática para análise de vias alteradas em diferentes tecidos e pontos experimentais;
- ✓ Aplicar e/ou implementar ferramentas de bioinformática para a integração dos dados de transcriptoma e metaboloma.

3 MATERIAIS E MÉTODOS

3.1 EXPERIMENTOS DE CAMPO

Um total de dois experimentos de campo com duração de 13 meses cada foi realizado (figura 2), no Centro de Ciências Agrárias da Universidade Federal de São Carlos pelo grupo colaborador da Dr^a. Monalisa Carneiro com plantio em quatro parcelas. A área experimental apresentava solo latossolo vermelho eutroférico. Nas instalações, o solo foi adubado com 400kg/ha de NPK (5-25-25). Foi aplicada irrigação para formação do *stand*. A parcela experimental foi composta por quatro linhas (sulcos) com 10m com espaçamento de 1,35m entre os sulcos. Vale ressaltar que os experimentos de campo, bem como a coleta dos dados não fazem parte do presente projeto, portanto, não houve participação do autor

O primeiro experimento de campo realizado teve o seu plantio iniciado em Abril de 2012 e se estendeu até Maio de 2013 (figura 3, a). O período para o plantio foi cuidadosamente selecionado no intuito de se estabelecer uma safra característica para cana de ano e meio. Para tal, o plantio necessita ser realizado em um período inicial de baixa pluviosidade e temperatura média, seguindo por um período favorável de crescimento e novamente um período de déficit hídrico e baixa temperatura média, a fim de se induzir um maior acúmulo de sacarose no colmo.

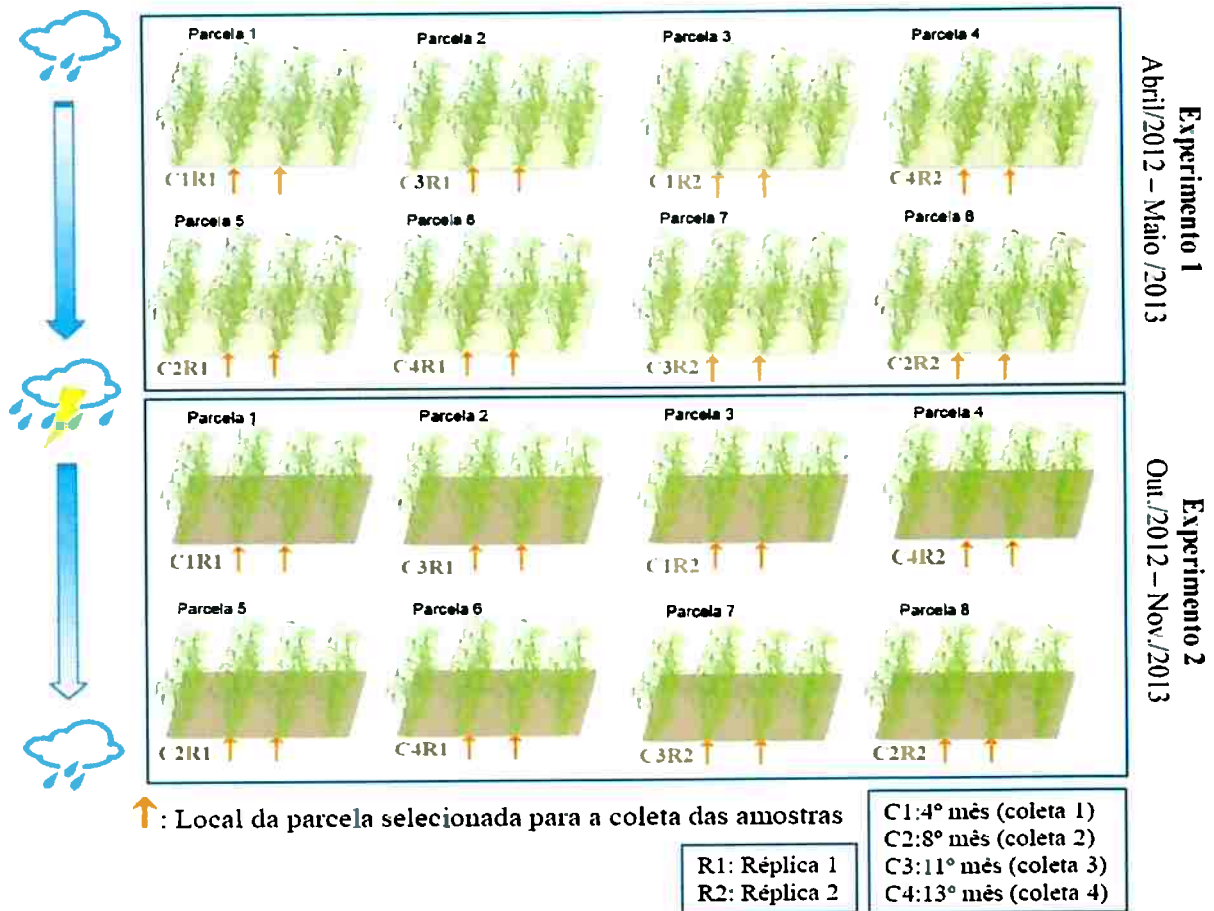


Figura 2 - Desenho experimental dos plantios de cana-de-açúcar. O presente estudo é composto por dois plantios de mesmo modelo, com um total de 8 parcelas cada. Os plantios foram realizados em períodos subsequentes de 13 meses cada. O primeiro experimento se iniciou em um período de baixa pluviosidade, enquanto que o segundo plantio em um período de alta pluviosidade. Em cada experimento foram realizadas 4 coletas ao longo dos meses. Em cada coleta foram obtidas amostras de entrenós e folha +1.

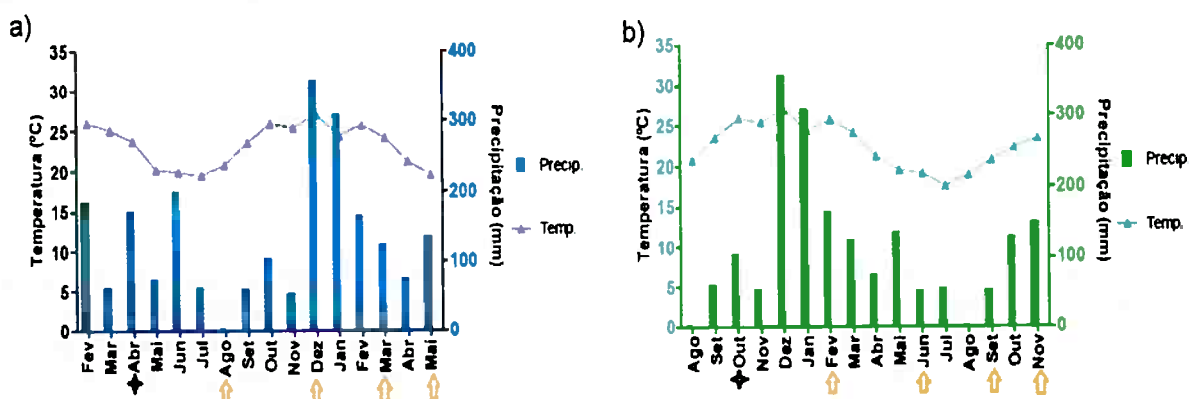
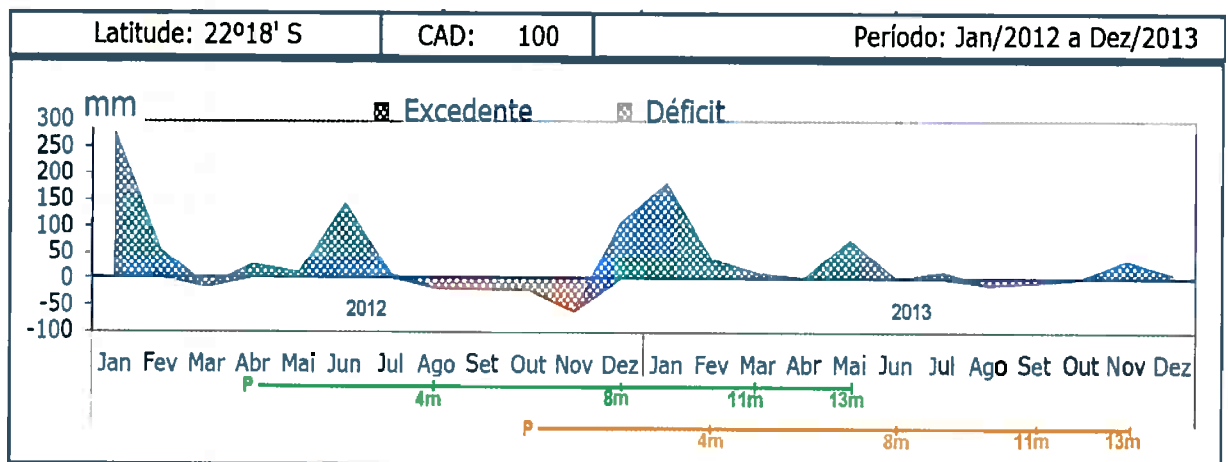


Figura 3 - Dados meteorológicos (temperatura e precipitação) mensais ao longo dos dois experimentos de campo. Em a), temos experimento com início em Abril de 2012. E em b), temos experimento com início em Outubro de 2012. Os meses destacados com a estrela em cor preta representam os períodos de plantio, enquanto que os meses destacados pelas setas em cor laranja, representam os meses da coleta de amostras (4 meses, 8 meses, 11 meses e 13 meses).

O segundo plantio, com as mesmas especificações técnicas, foi realizado no mês de Outubro de 2012, pouco após o início do primeiro experimento de campo, se estendendo até Novembro de 2013 (figura 3, b). O segundo experimento de campo foi planejado para uma safra característica de cana-ano, que possui um período médio de duração girando em torno de 12 meses. Diferente do primeiro experimento de campo, a safra de cana-ano necessita de um período inicial de crescimento sob alta precipitação e temperatura média, ocasionando assim um rápido crescimento. Após o rápido crescimento da planta, ela deve ser submetida a um período de baixa pluviosidade e temperatura média, para induzir um rápido acúmulo de sacarose. A partir das medidas de precipitação, também foram estimados os períodos que apresentaram déficit hídrico (Figura 4).



Fonte: UFSCar/CCA

Figura 4 - Estimação dos períodos que apresentaram déficit hídrico ao longo dos experimentos campo I e II. O primeiro plantio, destacado em cor verde, foi iniciado em Abril de 2012 (P) e finalizou em Maio de 2013. O segundo plantio, destacado na cor laranja, teve início em Outubro de 2012 (P) e finalizou em Novembro de 2013. Em ambos os experimentos de campo, amostras de diferentes tecidos e medidas morfológicas e fisiológicas foram coletadas quando as plantas apresentavam 4 meses, 8 meses, 11 meses e 13 meses de idade. No gráfico, os intervalos destacados em azul indicam um período favorável em termos de precipitação, enquanto que nos intervalos destacados em vermelho indicam período de déficit hídrico.

Visando realizar observações em diferentes épocas do ano e fases de crescimento da planta. Amostras de tecido (colmo, folha +1, entrenó imaturo e entrenós 5) foram coletadas dos dois experimentos de campo, quando as plantas apresentavam 4, 8, 11 e 13 meses de idade em cada um dos mesmos. Durante cada um dos experimentos também foram coletadas medidas de morfologia, fisiologia no intuito de se estabelecer futuras correlações com os dados moleculares. Assim, o presente estudo tem como objetivo elucidar os mecanismos moleculares que estariam modulando as alterações observadas em resposta aos fatores abióticos ao longo da maturação. Espera-se que um estudo integrado do transcriptoma e do metaboloma da cana-de-açúcar possa fornecer informações cruciais para a compreensão dos eventos observados.

3.2 AMOSTRAS UTILIZADAS NO ESTUDO

Para os experimentos de metaboloma foram utilizadas amostras de folhas +1 de oito indivíduos (4 por parcela) de cada um dos dois experimentos de campo, as quais foram misturadas dois a dois para formar 4 *pools* de amostras no intuito de se minimizar o *n* amostral e a variação no experimento (quadro 1). Para a análise do transcriptoma da cana-de-açúcar, foram utilizadas duas amostras biológicas para cada ponto de coleta do experimento de campo I e II. A partir de cada par de amostras de um ponto experimental, foram gerados 2 *pools*, de forma que uma das amostras e um dos *pools* foram hibridizados com um corante fluoróforo diferente, Cy3 e Cy5 (quadro 2).

Quadro 1 - Amostras utilizadas no estudo do metaboloma da cana-de-açúcar. A abreviação Cn se refere às coletas realizadas em campo; C1 para coleta aos 4 meses, C2 para 8 meses, C3 para 11 meses e C4 para 13 meses. A abreviação F1 é designada para tecido de folha +1; Mn para informar que se tratam de amostras para análise do metaboloma.

Amostras do campo I		<i>Pools</i>	Amostras do campo II	<i>Pools</i>
Parcela 1 4 meses	C1.F1.M8	I	2.C1.F1.M7	I
	C1.F1.M9	I	2.C1.F1.M9	I
	C1.F1.M10	II	2.C1.F1.M10	II
	C1.F1.M12	II	2.C1.F1.M11	II
Parcela 2 4 meses	C1.F1.M19	III	2.C1.F1.M19	III
	C1.F1.M21	III	2.C1.F1.M20	III
	C1.F1.M23	IV	2.C1.F1.M21	IV
	C1.F1.M24	IV	2.C1.F1.M22	IV
Parcela 1 8 meses	C2.F1.M7	I	2.C2.F1.M7	I
	C2.F1.M8	I	2.C2.F1.M8	I
	C2.F1.M9	II	2.C2.F1.M11	II
	C2.F1.M11	II	2.C2.F1.M12	II
Parcela 2 8 meses	C2.F1.M19	III	2.C2.F1.M20	III
	C2.F1.M21	III	2.C2.F1.M21	III
	C2.F1.M22	IV	2.C2.F1.M22	IV
	C2.F1.M23	IV	2.C2.F1.M23	IV
Parcela 1 11 meses	C3.F1.M7	I	2.C3.F1.M7	I
	C3.F1.M8	I	2.C3.F1.M9	I
	C3.F1.M10	II	2.C3.F1.M10	II
	C3.F1.M12	II	2.C3.F1.M11	II
Parcela 2 11 meses	C3.F1.M20	III	2.C3.F1.M19	III
	C3.F1.M22	III	2.C3.F1.M22	III
	C3.F1.M23	IV	2.C3.F1.M23	IV
	C3.F1.M24	IV	2.C3.F1.M24	IV
Parcela 1 13 meses	C4.F1.M7	I	2.C4.F1.M9	I
	C4.F1.M8	I	2.C4.F1.M10	I
	C4.F1.M9	II	2.C4.F1.M11	II
	C4.F1.M11	II	2.C4.F1.M12	II
Parcela 2 13 meses	C4.F1.M21	III	2.C4.F1.M19	III
	C4.F1.M22	III	2.C4.F1.M20	III
	C4.F1.M23	IV	2.C4.F1.M21	IV
	C4.F1.M24	IV	2.C4.F1.M22	IV

Quadro 2 - Amostras utilizadas para gerar os dados de transcriptômica pela tecnologia de microarranjo. Para cada ponto de coleta de ambos os experimentos de campo, foram selecionadas duas amostras. Dois *pools* de amostra foram gerados a partir de cada par, de forma que uma das amostras e um dos *pools* sejam hibridizados pelos corantes Cy3 e Cy5. As abreviações Cn se refere aos pontos de coletas do campo I, ao passo que 2.Cn do campo II.

Campo	Cy3	Cy5
Experimento de campo I	C1.F1.T1.T2 pool_F	pool_F C1.F1.T17.T18
	C2.F1.T1.T2 pool_F	pool_F C2.F1.T9.T10
	C3.F1.T1.T2 pool_F	pool_F C3.F1.T9.T10
	C4.F1.T1.T2 pool_F	pool_F C4.F1.T9.T10
Experimento de campo II	2.C1.F1.T1.T2 pool_F	pool_F 2.C1.F1.T17.T18
	2.C2.F1.T1.T2 pool_F	pool_F 2.C2.F1.T9.T10
	2.C3.F1.T1.T2 pool_F	pool_F 2.C3.F1.T9.T10
	2.C4.F1.T1.T2 pool_F	pool_F 2.C4.F1.T9.T10

3.3 TRANSCRIPTOMA DA CANA-DE-AÇÚCAR

3.3.1 Hibridização em microarranjo de oligonucleotídeos

A obtenção dos dados do transcriptoma das amostras de folha +1 foi realizada pela Dr.^a Carolina Lembke utilizando a tecnologia de microarranjo de oligonucleotídeos customizado da Agilent no formato 4 x 44K seguindo-se protocolo do fabricante. Este oligoarray customizado foi desenvolvido com oligos para detectar

a expressão de transcritos de cana-de-açúcar. A lâmina é composta por um total de 45.220 oligos, destes 21901 são representados em duplicata dentro do arranjo, onde 14522 hibridizam com transcritos senso e 7380 com transcritos anti senso.

Para a preparação das amostras, foi utilizado um total de 200 ng de RNA. A partir da amostra bruta, foram gerados cRNA e os mesmos foram marcados com os fluoróforos Cy3 e Cy5 para a hibridização nos *oligoarrays* customizados. O procedimento de marcação dos corantes foi realizado com o uso dos kits *Quick Amp Labeling Kit, Two-Color* (Agilent).

Para as demais etapas do procedimento de hibridização, também foram utilizados os kits: *Gene Expression Hybridization Kit, Gene Expression Wash Buffer Kit* and *RNeasy Mini Kit*. Após o processo de hibridização foi realizada a lavagem das lâminas e então a leitura no aparelho *GenePix 4000B scanner* (*Molecular Devices*). Os dados de fluorescência foram processados pelo programa *Feature Extraction 9.5.3* (Agilent).

3.3.2 Análise da expressão gênica

A análise computacional dos dados de transcriptômica foi realizada com o uso da plataforma *CaneTranscriptome* em SUCEST-FUN, desenvolvida pelo Dr. Milton Yutaka Nishiyama (disponível em <http://sucest-fun.org/>), qual permite o processamento dos dados de oligoarranjo de forma automatizada como descrito por Lembke et al., (2012).

Com os dados brutos, primeiramente foi realizada uma normalização para corrigir desvios intensidade-dependentes das intensidades detectadas para cada lâmina analisada, através da aplicação da função *Lowess*. Esta normalização foi feita com base nas intensidades das amostras controle (pool de todas as amostras

analisadas). Após a etapa de normalização, aplicamos o método HT-Self para a definição da lista dos genes diferencialmente expressos (LEMBKE et al., 2012; VÊNCIO; KOIDE, 2005). Para se determinar transcritos que estão expressos, identificamos aqueles que possuem sinal acima do ruído nas duas réplicas biológicas como descrito em Lembke et al 2012.

Nas análises do transcriptoma da variedade SP80-3280 da cana-de-açúcar, utilizamos a lista de genes expressos. Com a lista dos genes expressos e seus respectivos SAS, usamos o banco de dados SUCEST-FUN para agregar informações biológicas aos dados (possíveis enzimas associados aos SAS e as possíveis vias metabólicas em que os SAS podem estar inseridos). A partir desta lista, foi realizado o cálculo do escore de atividade de via e análise de significância sobre os escores.

Ainda, com a lista de enzimas indentificadas para cada SAS estimamos, os genes ortólogos de milho associados. O que nos possibilitou integrar os dados de transcriptômica e metabolômica nas vias metabólicas descritas pelo banco de dados biológico KEGG (<http://www.kegg.jp/>). O milho foi selecionado por ser uma planta gramínea, assim como a cana-de-açúcar, uma vez que o genoma completo da cana-de-açúcar ainda não existe.

3.4. METABOLOMA DA CANA-DE-AÇÚCAR

3.4.1 Cromatografia líquida acoplada à espectrometria de massa

As amostras foram analisadas no sistema de cromatografia UFLC (Shimadzu, Inc) acoplado ao espectrômetro de massas ESI-MicroTOF (BrukerDaltonics, Inc.). Para maximizar a cobertura do metaboloma, as colunas Ascentis Express HILIC-

Supelco (15 cm x 2.1 mm x 2.7 um) e C18- Phenomenex (10 cm x 2,1 mm x 2,7um) foram utilizadas na cromatografia e o espectrômetro de massas foi programado para leituras em modo positivo e negativo de ionização. A coluna HILIC foi utilizada para as aquisições em modo positivo e negativo, enquanto que a coluna C18 apenas para o modo de aquisição positivo. As Fases Móveis A (FMA) e B (FMB) utilizadas para a coluna C18 foram respectivamente: 0,1% de Ácido Fórmico + H₂O e 0,1% de FMA + ACN (Acetonitrila). Já as fases móveis A e B utilizadas para a coluna HILIC foram respectivamente: NH₄Ac a 50 mM (ph8) e FMA:ACN (5:95).

Informações do gradiente utilizado para a corrida cromatográfica de ambas as colunas podem ser encontradas na tabela 2 (coluna C18) e tabela 3 (coluna HILIC):

Tabela 2 - Informações sobre gradiente da corrida cromatográfica para a coluna C18. As abreviações significam: FMB (Fase Móvel B) e FMA (Fase Móvel A). A corrida foi realizada a um fluxo de 0.4 ml/min durante um tempo total de 23 min, à temperatura de 40°C.

Tempo (min)	FMB (%)	FMA (%)
0	1	99
1	1	99
13	35	65
14,5	70	30
15,5	99	1
19	99	1
19,5	1	99
23	1	99

Tabela 3 - Informações sobre gradiente da corrida cromatográfica para a coluna HILIC. FMB (Fase Móvel B) e FMA (Fase Móvel A). A corrida foi realizada a um fluxo de 0.5 ml/min durante um tempo total de 30 min, a temperatura de 40°C

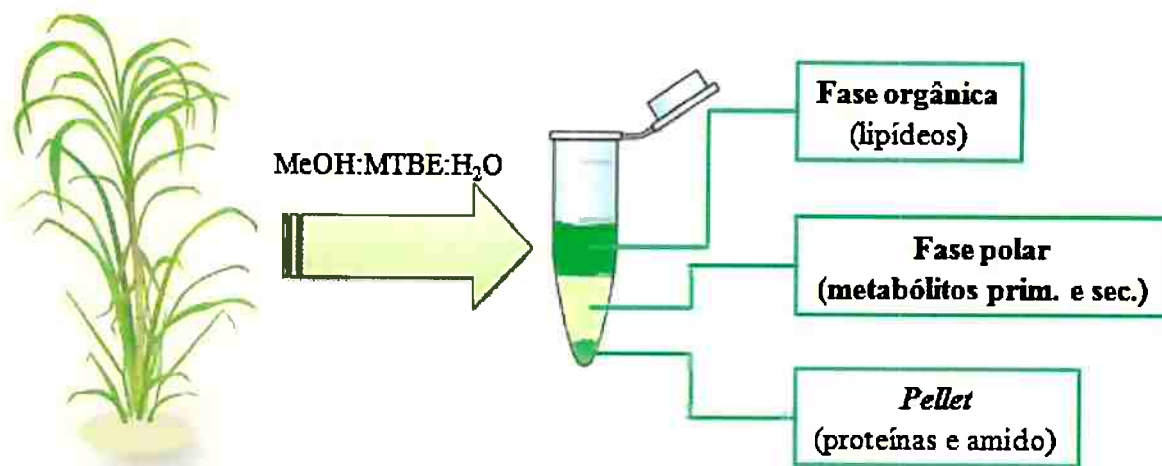
Tempo (min)	FMB (%)	FMA (%)
0	100	0
2	100	0
23	60	40
25	60	40
25,01	100	0
30	100	0

A fonte de ionização utilizada foi o ESI (Ionização por Eletro Spray). Para o modo positivo de aquisição o instrumento operou com a voltagem no capilar de 4500 V e diferença de potencial de -500 V. A temperatura usada na fonte de ionização foi de 220°C, e gás N₂ a uma pressão de 4 Bar, com fluxo de 10 L/min⁻¹. O espectro de massa para detecção foi configurada para um intervalo entre 50 m/z a 1000 m/z (m/z, razão entre a massa e carga da molécula). Para o modo negativo de aquisição, o instrumento operou com a voltagem no capilar de 3800 V e diferença de potencial de -500 V. A temperatura usada na fonte de ionização foi de 220°C, a uma pressão de 4 Bar e um fluxo de 12 L/min⁻¹. A faixa de massa para detecção foi configurado para um intervalo entre 50 m/z a 1100 m/z, utilizando-se formiato de sódio (Sigma) como calibrador.

3.4.2 Extração de metabólitos

Primeiramente, as amostras de folha +1 da cana-de-açúcar foram trituradas e homogeneizadas sob refrigeração constante no equipamento Tissue Lyser II (Qiagen), com o auxílio de esferas metálicas. Após a trituração, foram separados 50 mg de cada amostra em um tubo de 2 mL o qual foi mantido a -80°C.

Para a extração dos metabólitos, foi utilizado a metodologia adaptada de Giavalisco e colaboradores (GIAVALISCO et al., 2011), que consiste na utilização do Metanol juntamente ao éter metil *terc*-butílico e água (MeOH:MTBE:H₂O). Dessa forma é possível realizar uma extração em fases, conforme a figura 5.



Fonte: Adaptada de GIAVALISCO et al., 2011.

Figura 5 - Extração de metabólitos. Metodologia adaptada de GIAVALISCO e colaboradores (GIAVALISCO et al., 2011). Foi utilizado MeOH:MTBE:H₂O (1:3:1) para realizar a separação dos metabólitos conforme a sua propriedade química.

Na primeira etapa da extração foi adicionado 1 mL de MeOH:MTBE:H₂O (1:3:1) no tubo contendo a amostra juntamente com o padrão interno (15 µL de Glicose C13 a 150 nmol e 10 µL de 6-dimetilaminopurina a 2,5 mM). Então, realizou-se uma homogeneização por 1 minuto. Após homogeneizar, as amostras foram mantidas no gelo por 5 minutos. Esse processo foi repetido por três vezes consecutivas.

Na segunda etapa, adicionou-se 650 µL de MEOH:H₂O (1:3) seguindo-se de uma homogeneização por 30 segundos. Feito isso os tubos foram centrifugados a 4 °C por 5 minutos a uma velocidade de 10000 g. Após esta etapa, a amostra passa a apresentar duas fases distintas (fase polar e a fase não-polar) mais o *pellet* de proteínas. Foram retirados 1000 µL da fase polar de cada amostra, os quais foram secos a vácuo e estocados a -80°C. No momento da análise, as amostras foram ressuspensas em 200 µL de fase móvel A (0,1% de Ácido Fórmico + H₂O), e apenas 1 µL de cada amostra foi injetado para a análise. A fração protéica contida

no *pellet* das amostras foi estocada a -80°C para posterior quantificação de proteínas.

3.4.3 Quantificação de proteínas

A quantificação de proteínas foi realizada pelo método adaptado de Bradford (BRADFORD, 1976), que consiste na reação entre o corante de Bradford com as proteínas do meio. Conforme a concentração de proteínas no meio aumenta, a coloração pode variar entre o atenuado para o escuro. Com a leitura do espectro de luz, é possível inferir o valor aproximado de concentração protéica ao compará-lo a uma curva padrão de calibração.

A primeira etapa consiste em adicionar 1 mL de solução de NaOH (0.2 mol/L) no tubo contendo o *pellet* protéico. De forma equivalente são preparadas alíquotas com diferentes concentrações de proteína de soro bovino (1 mg/ml; 0,5 mg/ml; 0,25 mg/ml; 0,125 mg/ml e 0,6 mg/ml), que são utilizadas para a curva padrão. Estas alíquotas são submetidas às mesmas condições que as demais amostras biológicas.

Na segunda etapa, as amostras são aquecidas em banho seco por 5 minutos a 100°C . Depois, são refrigeradas no gelo por 1 minuto e centrifugadas por 1 minutos a uma velocidade de 5000 .rpm. Finalmente as amostras são lidas em leitor de placas (*Spectra Max*), e a concentração de proteínas solúveis é calculada com base na curva padrão.

3.5 ANÁLISE COMPUTACIONAL DOS DADOS

3.5.1 Otimização de parâmetros para detecção e agrupamento de picos

Devido à alta complexidade dos dados de metabolômica em estudos não direcionado utilizando diferentes plataformas de LC-MS (Cromatografia Líquida Acoplada ao Espectrômetro de Massas), uma boa parametrização dos algoritmos de detecção e agrupamento de picos de massa é dificultada quando o usuário não possui um conhecimento aprofundado a cerca das diferentes variáveis como: *mzdiff* (limite de tolerância para variação da *m/z*), *peakwidth* (largura do pico cromatográfico, estimado por uma janela no tempo de retenção), *prefilter* (filtro para eliminar picos de baixa qualidade, estimado pela sua intensidade), *fwhm* (largura à meia altura, usado para estimar o valor de sigma). Dessa forma, utilizar um algoritmo de otimização de parâmetros torna-se extremamente vantajoso (LIBISELLER et al., 2015).

No presente estudo, o programa XCMS (SMITH et al., 2006) foi utilizado para realizar as etapas iniciais de análise do metaboloma. A detecção de pico, o agrupamento e a correção do tempo de retenção dos mesmos exigem uma parametrização adequada, caso contrário poderemos estar adquirindo os dados de maneira ineficaz. Dessa forma, optamos por utilizar o otimizador de parâmetros IPO (*Isotopologue Parameter Optimization*), que realiza a otimização dos parâmetros de diferentes funções do programa XCMS (*xcmsSet()*, *group()* e *retCor()*).

A otimização de parâmetros foi realizada utilizando o planejamento de experimentos (DoE), que consiste em uma série de testes onde modificações específicas são realizadas nas variáveis de entrada com o intuito de se obter o melhor resultado. Todo este processo é realizado de forma semi-sequencial, através do planejamento experimental de Box-Behnken (BOX; BEHNKEN, 1960).

Para cada uma das etapas de processamento é calculado um escore que servirá como base de estimação dos melhores parâmetros. Primeiramente é calculado o escore para a detecção de picos, em seguida para a correção no tempo de retenção e finalmente para o agrupamento de picos. Na detecção de picos, são selecionados picos que pertencem a um isotopólogo devido ao seu maior grau de confiabilidade; estes são classificados como picos significativos (PS). Assim, o escore detecção de picos (EDP) é calculado pela fórmula:

$$EDP = \frac{PS^2}{\text{'número total de picos' - PBI}}$$

Onde EDP é a razão entre a quantidade de picos significativos (PS) elevado a 2 e o número total de picos menos o número de picos de baixa intensidade (PBI). A quantidade de picos significativos é ponderada empiricamente por um fator exponencial 2.

Para se obter a qualidade da correção do tempo de retenção para um determinado grupo, é calculado o escore da variação do tempo de retenção do grupo (VTRG), que é dado pela fórmula:

$$VTRG(x) = \sum_{n=1}^k \frac{|(média(x) - x_n)|}{k}$$

Onde x é o tempo de retenção de todos os picos dentro de um grupo, k é a quantidade de tempos de retenções e n o índice pontual no tempo de retenção de um pico individual no grupo. Assim, para cada grupo x é calculada a variação no tempo de retenção do grupo (VTRG). Com os valores individuais de cada VTRG é possível inferir a variação global no tempo de retenção (VGTR):

$$VGTR = \frac{1}{k} \times soma(VTRG)$$

Quanto menor o valor VGTR, melhor será a qualidade da correção do tempo de retenção resultado. Logo, para se criar um valor de escore que se adéque a um modelo de maximização, é estimado um escore de correção do tempo de retenção (ECR), que é dado pelo inverso do VGTR:

$$ECR = \frac{1}{VGTR}$$

O escore de agrupamento (SA) é baseado na classificação entre grupos significantes e não significantes. Os grupos significantes são aqueles que apresentam apenas um pico de massa de cada amostra analisada dentro do grupo, já os grupos não significantes são os que não obedecem a essa regra:

$$SA = \frac{'grupos significantes'^2}{'grupos não significantes'}$$

Finalmente é calculado de forma conjunta o valor da correção do tempo de retenção e agrupamento (VCRA):

$$VCRA = norma(ECR) + norma(SA)$$

O VCRA é dado pela soma dos valores normalizados ECR e AR (*norma* é uma função de normalização que converte os valores de ECR e SA para uma escala numérica entre 0 e 1). Esta normalização é feita para minimizar o impacto do balanceamento, é aplicado uma função de normalização sobre os valores de ambos os escores, convertendo-os em uma escala entre 0 e 1.

Os escores resultantes são avaliados em um modelo de superfície de resposta, onde a combinação de parâmetros que representa o melhor escore é

selecionada para o planejamento do experimento (DoE). Este processo é repetido enquanto a maximização do escore for possível.

3.5.2 Identificação e agrupamento de picos

Diversos softwares encontram-se disponíveis para a execução desta etapa, sejam eles de código aberto ou comerciais (KATAJAMAA; OREŠIČ, 2005; TAUTENHAHN; BÖTTCHER; NEUMANN, 2008). O pacote do programa XCMS (<http://www.bioconductor.org/.../xcms.html>) foi selecionado por ser gratuito e de código aberto, possuir uma vasta gama de referências na literatura e disponibilizar ferramentas que possibilitam a integração com diferentes programas de processamento de dados.

O *software* XCMS funciona basicamente processando os dados de metabolômica em três etapas subsequentes. Em primeira instância é realizada a detecção dos picos cromatográficos, onde os dados contínuos são discretizados e convertidos em formato centróide (cada íon é representado como um pico de massa). Dessa forma, é possível a redução da dimensão dos dados sem grande perda de informação, além disso, são eliminados os vieses causados por ruídos de dados contínuos (SMITH et al., 2006). Para a execução desta etapa, é possível escolher entre dois algoritmos distintos '*matchedFilter*' e '*CentWave*'. No presente estudo, o algoritmo '*Centwave*' foi selecionado para a detecção dos picos, por apresentar desempenho superior ao '*Matched filter*' (TAUTENHAHN; BÖTTCHER; NEUMANN, 2008).

O algoritmo *CentWave* de detecção de picos utiliza duas abordagens para a resolução do pico cromatográfico, a primeira abordagem é baseada na densidade de

picos por Regiões de Interesse (ROI) no domínio m/z , a segunda abordagem baseia-se na transformação contínua Walevet (CTW).

A detecção das ROI's consiste em seleccionar regiões de alta densidade de picos que apresentem um desvio padrão limite entre os valores de m/z , o valor de tolerância é determinado pela resolução do equipamento. A resolução do equipamento é medida em ppm (partes por milhão) e tem um valor μ , portanto, podemos dizer que a resolução do equipamento é de μppm . A distribuição dos picos de massa é dada ao longo de um tempo de retenção (em segundos), logo é possível estimar o comprimento de um pico cromatográfico. Devido ao fato da resolução e do comprimento de um pico cromatográfico serem conhecidos, é possível acessar facilmente regiões com pelo menos um número P_{min} de centróides com desvio padrão menor que a resolução μppm .

Tendo em mãos os dados em modo centróide, que possuem os *scans* numerados $s = 1, \dots, S$, a estimação da densidade de m/z é realizada da seguinte maneira:

1. Inicialização:

(a) Inicia-se uma lista de ROI's usando todos os valores m/z mz_i^s para o primeiro *scan*:

$$\forall i = 1, \dots, N, N = |mz^{s=1}|: ROI(i).valores(1) = mz_i^{s=1}$$

(b) Inicia-se o valor médio de m/z para cada região sendo processada:

$$ROI(i).médiamz = mz_i^{s=1}, i = 1, \dots, N, \quad N = |mz^{s=1}|$$

2. Para cada *scan* $s = 2, \dots, S$:

(a) Para cada valor $mz_i^s, i = 1, \dots, N, N = |mz^s|$ no *scan* atual s :

Existe $j, j = 1, \dots, J, J = |ROI|$ em que $|ROI(j).médiamz - mz_i^s| \leq \mu$

- Se sim: Adiciona mz_i^s para a ROI(j) e atualiza o valor da média de m/z

$$k = |ROI(j).valores| + 1, ROI(j).valores(k) = mz_i^s$$

$$ROI(j).médiamz = \frac{1}{K} \sum_{K=1}^K ROI(j).valores(K)$$

- Se não: Inicializa um novo ROI e adiciona para a lista

$$J = |ROI| + 1, ROI(J).valores(1) = mz_i^s, ROI(J).médiamz = mz_i^s$$

(b) Verifica e atualiza:

- Todas ROI's que não são estendidas na etapa 2(a) e que contenham pelo menos $Pmin$ centróides são removidas.

Com todas as regiões de interesse identificadas é necessário realizar a identificação dos picos cromatográficos. Para tal, é aplicado o algoritmo CWT (Transformação Contínua Walevet). Sua representação matemática é dada por:

$$(T^{WAV} f)(s, \tau) = \int_{-\infty}^{\infty} f(t) \psi_{s,\tau}(t) dt$$

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-\tau}{s}\right), s \in \mathbb{R}^+ - \{0\}, \tau \in \mathbb{R}$$

Onde, $f(t)$ é o sinal, ψ é a *walevet* mãe, s a escala e τ a translação. Como resultado da aplicação do algoritmo CWT sobre as intensidades das ROI's, temos uma matriz bidimensional de coeficientes T_{wav} . Para cada escala é estimado o máximo local dos coeficientes de CWT.

Após a detecção dos picos cromatográficos, é realizado o alinhamento dos mesmos. Esta etapa é crucial, pois permite a comparação entre amostras, uma vez que o tempo de retenção de um mesmo íon pode variar mesmo entre réplicas

biológicas de uma mesma condição. Esta variação é geralmente não uniforme durante todo o tempo de retenção.

O alinhamento é realizado em duas etapas. Primeiro, os picos identificados são agrupados baseados no seu valor de m/z e tempo de retenção. Após o agrupamento são selecionados os grupos de maior qualidade, dentre os quais possuem uma quantidade mínima de amostras sem picos atribuídos, para serem utilizados como bases para o alinhamento. Em seguida, o desvio do tempo de retenção no entorno destas bases a partir dos valores médios de cada grupo correspondente são correlacionados com o tempo de retenção. Após o primeiro alinhamento, algumas regiões no cromatograma podem apresentar regiões com picos não agrupados e que podem ser interpolados. Estes picos são agrupados novamente e alinhados para coincidir com os picos no tempo de retenção corrigido. Esta etapa é conduzida de forma iterativa até que o desvio no tempo de retenção seja suficientemente corrigido (SMITH et al., 2006).

Ao final desta etapa de processamento, é gerada uma matriz (objeto ou variável do R, a linguagem de programação utilizada) contendo todas as informações relacionadas ao pré-processamento (picos, grupos, tempo de retenção, intensidade e modo de aquisição) que serão utilizadas no decorrer das análises subsequentes.

3.5.3 Anotação dos íons

Para a identificação dos íons foi selecionado o pacote de análise do programa CAMERA (KUHLE et al., 2012). Este pacote foi desenvolvido especialmente para analisar dados de LC-MS de forma sequencial com o programa XCMS descrito anteriormente.

O algoritmo de anotação dos íons pelo programa CAMERA funciona basicamente em cinco etapas, tendo como dados de entrada os picos detectados pelo programa XCMS. A primeira etapa consiste em gerar pseudo-grupos de espectros dos compostos, onde o agrupamento é dado ao longo do tempo de retenção dos dados. Para tal, é selecionado o pico mais intenso da tabela de dados que ainda não está associado a um espectro e então é calculada uma janela específica para o tempo de retenção estimada pela FWHM (Largura à Meia Altura) do pico cromatográfico ao redor do centróide. Todos os compostos dentro desta janela são agrupados em um novo espectro. Esta etapa é conduzida de forma iterativa até que todos os picos estejam agrupados no espectro de algum composto.

A segunda etapa consiste em detectar todos os possíveis picos isotópicos, o que possibilita estimar a mudança de carga das moléculas. Esta detecção é realizada calculando-se uma matriz de distância par-a-par entre todas as m/z , a fim de se detectar todos os pares com 1,0033 z de diferença (valor de diferença conferida na carga entre o isótopo e sua molécula de maior abundância). Na terceira etapa, é utilizado um algoritmo de grafo para separar os compostos que podem estar co-eluídos em um tempo de retenção muito próximo. Para isso, é realizada uma correlação de *Pearson* entre todas as intensidades dos picos de massas de cada pico cromatográfico dos espectros. Finalmente, informações sobre picos isotópicos são atreladas a cada par calculado, sendo 1 para indicar a existência de um pico isotópico no par calculado e 0, caso contrário.

Nesta terceira etapa todas as informações são utilizadas para calcular um escore que represente a relação entre dois picos de massas x e y .

$$escore(x, y) = CAS_{XY} + ISO_{xy} + 1/N \sum_{i=1}^N CPS_{ixy}$$

O escore é uma combinação da correlação das intensidades entre as amostras (CAS) para estes dois picos, a informação binária 1 e 0 indicando presença ou ausência de isótopo (ISO) entre estes dois picos, e a correlação entre as intensidades dos espectros (CPS) calculadas para a amostra *i*. Dentro de um grafo, os picos de massas são representados como nós e os mesmos são conectados por arestas ponderadas pelo escore calculado. O grafo é então clusterizado de forma a evidenciar os compostos co-eluídos.

Em experimentos que utilizam o ESI como fonte ionizadora, frequentemente são gerados espécimes iônicos com ganho ou perda de carga, adutos e ainda perdas neutras devido à formação de fragmentos iônicos. Neste último caso, ainda que o processo de ESI seja considerado uma ionização fraca, é possível ocorrer a fragmentação de algumas moléculas. O processo mais comum observado é a perda de uma molécula de água ($[M+H-H_2O]^+$ ou $[M-H-H_2O]^-$). O programa CAMERA possui um banco de dados de íons observados experimentalmente, nos permitindo assim estimar determinados íons ao confrontar estas informações com os dados biológicos. A identificação de todos os possíveis íons é realizada na quarta etapa de processamento do algoritmo, mapeando-se todas as diferenças de *m/z* dos espectros contra o banco de dados de íons.

Na última etapa de processamento é realizado um mapeamento dos dados contra um banco contendo informações sobre o modo de ionização oposto ao modo a partir do qual os dados foram gerados originalmente. Este procedimento é realizado uma vez que alguns compostos podem ser ionizados em ambos os modos positivo e negativo. Dessa forma, é possível agregar uma informação complementar à anotação dos íons e fornecer uma evidência maior de que esta molécula se trata de um íon real.

3.5.4 Anotação dos compostos

Nesta última etapa de identificação dos compostos, utilizamos o programa ProbMetab (SILVA et al., 2014). Diferentemente da anotação convencional, que é realizada manualmente para cada massa de interesse, o ProbMetab é um pacote que possibilita a anotação de forma automatizada e integrada com diferentes pacotes de programas para processamento de dados de espectrometria de massas (por exemplo, XCMS, mzMatch, AStream). Além disso, o programa adota uma forma diferenciada de anotação, dispondo-se de um algoritmo de estimação probabilística, que possibilita agregar informações biológicas (reações bioquímicas) na predição do melhor candidato.

O modelo de anotação probabilística é uma expansão do modelo descrito por Rogers e colaboradores (ROGERS et al., 2009) e sua extensão (ROGERS et al., 2011). Em seu modelo, o mapeamento entre os picos de massa da análise com os compostos no banco de dados $m \rightarrow c$, onde $m \in \{1, \dots, M\}$, $c \in \{1, \dots, C\}$, sendo $M \leq C$; é representado por um vetor de coluna $\vec{z}_m \in \{1,0\}^C$, em que $|\vec{z}_m| = 1$. A distribuição conjunta de todos os mapeamentos dado a massa do espectro observado é: $f(\vec{z}_m, \dots, \vec{z}_M | x_1, \dots, x_M, I_1(t), \dots, I_M(t), \theta_{M+1}, \dots, \theta_N)$, sendo abreviada para $f(Z | \vec{x}, \vec{\theta})$. Nesta notação, Z representa todos os vetores de massa concatenados em colunas de uma matriz, x_m é o m -ésimo pico de massa observada e $I_m(t)$ a intensidade da massa (incluindo repetições experimentais). O valor $\vec{\theta}$ engloba todos os demais N parâmetros necessários. O amostrador de Gibbs é utilizado para acessar a distribuição conjunta de $f(\cdot)$ pelo uso das suas marginais: $f(\vec{z}_m | Z^{(-m)}, \vec{x}, \vec{\theta})$ onde $Z^{(-m)}$ representa Z sem a sua m -ésima coluna relacionado ao pico de massa em \vec{z}_m . De acordo com a estatística bayesiana, essa distribuição de

probabilidade $f(.)$ é proporcional ao produto: $L(x_m | Z, \vec{x}^{(-m)}, \vec{\theta}) \cdot p(\vec{z}_m | Z^{(-m)}, \vec{x}^{(-m)}, \vec{\theta})$, onde são $L(.)$ e $p(.)$, as funções *likelihood* e *priori* respectivamente.

O modelo descrito por Silva e colaboradores no programa Probmstab é uma expansão da função $L(.)$ em termos multiplicativos independentes:

$$L(x_m | Z, \vec{x}^{(-m)}, \vec{\theta}) \propto L_N(x_m | Z, \vec{\theta}_N) \cdot L_{rt}(x_m | Z, \vec{\theta}_{rt}) \cdot L_{iso}(x_m | Z, \vec{\theta}_{iso})$$

O modelo $L_N(.)$ realiza a estimação do ruído, enquanto que $L_{rt}(.)$ é o modelo de estimação do erro do tempo de retenção. Por fim, o modelo $L_{iso}(.)$ estima o erro no padrão isotópico. A expansão em termos independentes confere grande flexibilidade ao conjunto de $L(.)$. Assim, o modelo $L_{rt}(.)$ é opcionalmente agregado no cálculo da probabilidade, uma vez que o mesmo ainda é propenso a erro (DUNN et al., 2013). Dessa forma, podemos também optar por adicionar informações da abundância dos carbonos C^{12} e C^{13} para conferir uma razão de change (*odds ratio*) aprimorada para um dado composto em relação aos demais (WEBER et al., 2011).

O termo verossimilhança $L_N(.)$ é crucial, sendo este a base do modelo original $L(.)$ descrito por Rogers e colaboradores. Inicialmente o modelo de estimação do ruído foi apresentado como uma densidade de gaussiana com variância fixa $1/\gamma$ dado a precisão γ do espectrômetro de massas. No intuito de superar as limitações do modelo $L_N(.)$ original, Silva e colaboradores passaram a estimar o ruído de forma que:

$$L_N(x_m | Z, y, w) \propto 1_\omega * \left(1 - \Phi\left(\frac{1}{\sigma_\omega} \cdot \left(\left|\frac{x_m}{y_c} - 1\right| - \mu_\omega\right)\right)\right)$$

Onde 1_ω é a função indicadora para o conjunto ω formada por todos os j compostos candidatos no banco de dados em que $|x_m - y_j| \leq \omega$. Os valores μ_ω e σ_ω representam a média e o desvio padrão de todas as massas em y listadas em ω ; Φ

é a função da distribuição normal cumulativa; w é o parâmetro de limite na diferença de massas estabelecido pelo usuário.

O produto final do modelo probabilístico utilizado no programa ProbMetab é uma lista de candidatos ranqueados pelos seus valores de probabilidade. Diferente do modelo original onde $|\vec{z}_m| = 1$, o novo modelo possibilita que diferentes candidatos sejam atribuídos a uma mesma massa, desde que os mesmos estejam dentro das limitações $|x_m - y_j| \leq \omega$. Quando obtido apenas um candidato por massa a probabilidade agregada ao mesmo é 1. Por outro lado, quando mais de um candidato $c_1, \dots, c_{|\omega|}$ é atribuído a m , a probabilidade atribuída aos candidatos será uniforme e igual a $1/|\omega|$.

3.5.5 Análise estatística dos compostos identificados

No intuito de se investigar os metabólitos com diferenças significativas ao longo dos experimentos de campo I e II, realizamos o teste de ANOVA e teste LSD de Fisher (P-valor < 0,05) pelo programa Metaboanalyst 3.0 (<http://www.metaboanalyst.ca/>). Antes da aplicação dos testes estatísticos, os dados normalizados foram centralizados pela média e divididos pelo desvio padrão em cada um dos metabólitos anotados.

3.5.6 Análise da atividade das vias metabólicas a partir metaboloma

Dado o interesse em se elucidar as variações globais entre cada ponto de coleta de ambos os experimentos de campo, optamos por realizar uma análise de atividade de vias metabólicas. O programa PAPI (AGGIO; RUGGIERO; VILLAS-BÔAS, 2010) foi utilizado para este fim. Seu algoritmo utiliza as informações contidas no banco de dados KEGG (*Kyoto Encyclopedia of Genes and Genomes*)

para o cálculo do escore de cada via presente na amostra. Para tal, foi desenvolvido um banco de dados *in house*, de vias metabólicas presente apenas em plantas da família *Poaceae* (e.g., Arroz, Milho, Sorgo e Milho). O escore a seguir é calculado para cada via:

$$S(P) = \frac{(m_1 + m_2 + \dots + m_N)}{\frac{K}{M}}$$

Sendo S , o escore calculado para a via metabólica P ; m a intensidade normalizada dos metabólitos mapeados na via P ; K é a quantidade de compostos mapeados na via P e M é a quantidade total de metabólitos contidos na via P . Os valores obtidos para cada ponto experimental foram comparados entre si na busca por vias com diferença significativa na sua atividade, usando o teste ANOVA e teste LSD de Fisher (P -valor < 0,05). Assim, quanto maior o escore obtido maior será a atividade predita para a via metabólica em questão. No presente estudo, foram realizadas observações do comportamento das vias ao longo de cada experimento (figura 7).

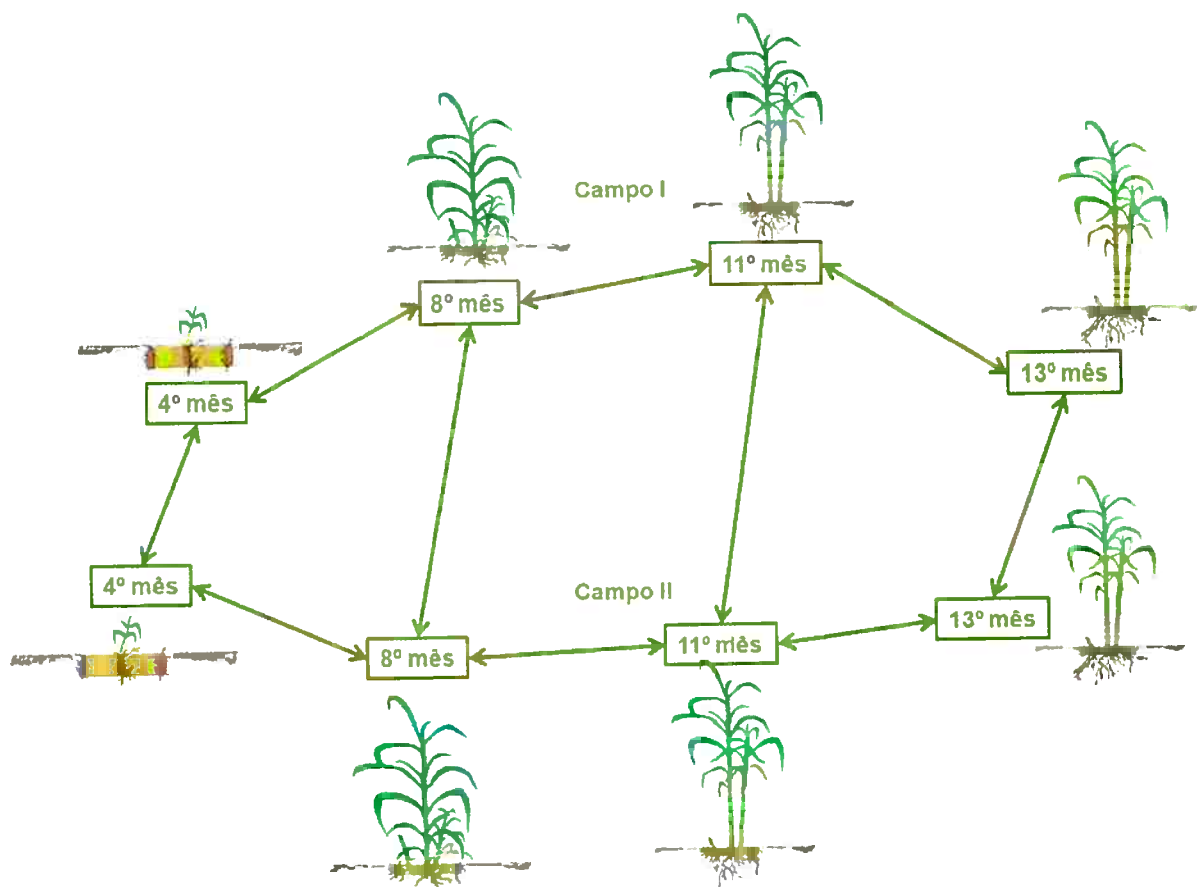


Figura 6 - Desenho experimental para comparação da atividade das vias metabólicas nos experimentos de campo I e II. As setas indicam as comparações realizadas entre cada ponto do experimento, sejam elas entre idades equivalentes ou ao longo do crescimento da planta.

3.6 DEFINIÇÕES GERAIS DO PIPELINE DE ANÁLISE DO TRANSCRIPTOMA E METABOLOMA

O *pipeline* de análise consistiu no uso de um vasto conjunto de ferramentas como indicado na figura 7. Os dados de transcriptoma da cana-de-açúcar adquiridos pelo uso da tecnologia de microarranjo foram processados pela plataforma *CaneTranscriptome* SUCEST-FUN (<http://sucest-fun.org/>) e então foram identificados os genes significativamente expressos para a estimação da atividade de vias metabólicas. A plataforma *CaneTranscriptome* do banco SUCEST-FUN está totalmente automatizada para analisar dados de transcriptômica, entretanto, ainda

carece de ferramentas para a análise de dados de metabolômica. Assim, foram agregadas diversas ferramentas desenvolvidas em linguagem de programação R (<https://www.r-project.org/>) para atender esta finalidade. As ferramentas de anotação de compostos foram selecionadas com a colaboração do pesquisador Dr.º Ricardo Roberto Silva, que colaborou no processo de construção do *pipeline* de análise dos dados de metabolômica. Os resultados da anotação são utilizados nas análises estatísticas, de forma similar ao conduzido sobre os dados de transcriptômica. Na etapa final dos processos, é possível a integrar ambas as ômicas em uma determinada via metabólica a fim de se conduzir uma análise conjunta dos dados.

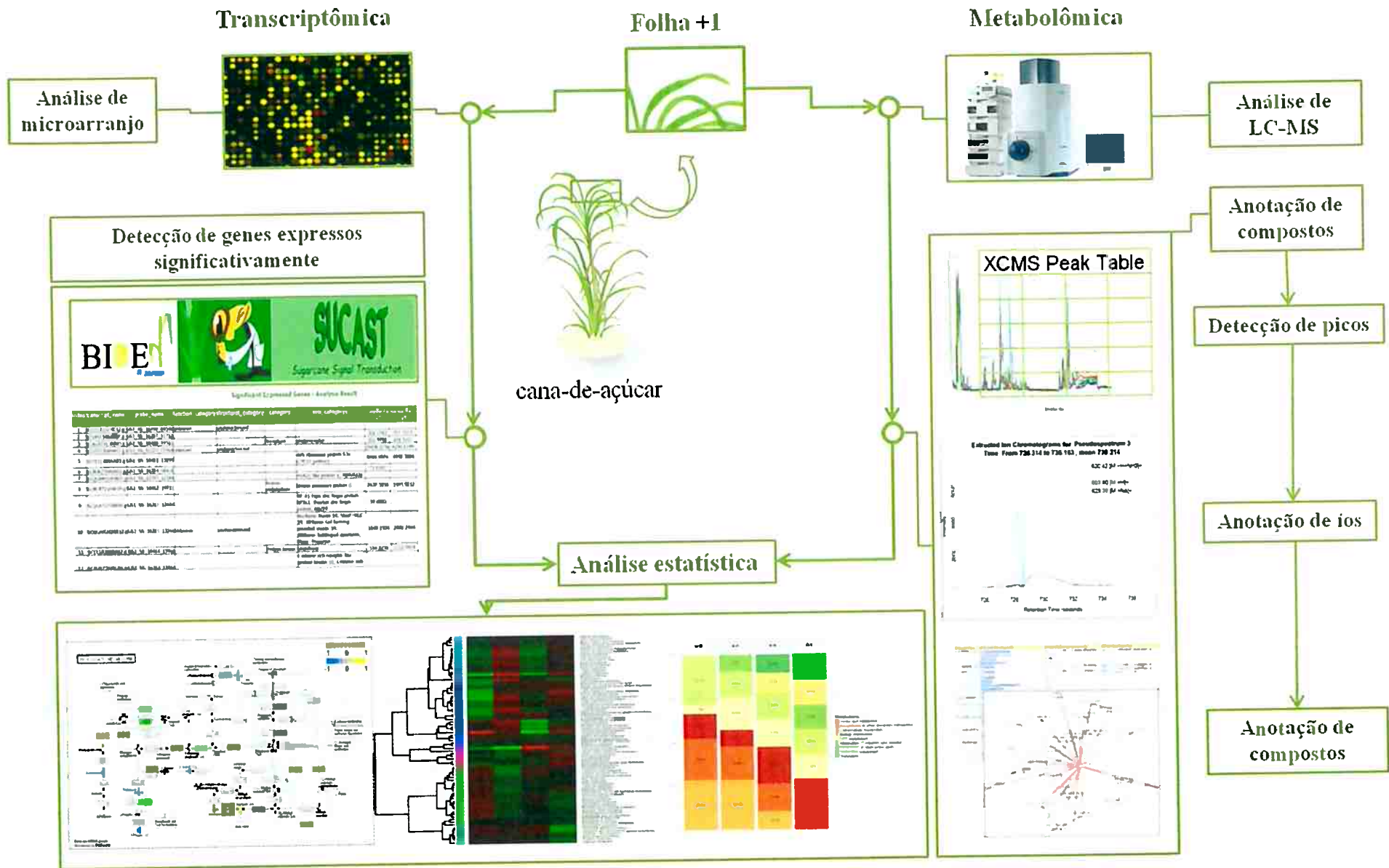


Figura 7 - Informações gerais do *pipeline* utilizado na análise dos dados do transcriptoma e metaboloma da cana-de-açúcar. Na análise do metaboloma da cana-de-açúcar as ferramentas XCMS, CAMERA e ProbMetab foram utilizadas para a detecção de picos, anotação de íons e anotação de compostos respectivamente. Os dados do transcriptoma foram analisados pelas ferramentas disponíveis no banco de dados SUCEST- FUN. A análise estatística consistiu no cálculo do escore de atividade das vias a partir dos dados de metabolômica e transcriptômica. Também foi realizada uma análise de clusterização hierárquica sobre os escores obtidos. A comparação por teste t de *Student* foi aplicada sobre os escores no intuito de se evidenciar vias com diferença significativa ao longo do processo de maturação. Os dados foram integrados nas vias metabólicas para auxiliar na investigação dos possíveis eventos biológicos.

4 RESULTADOS

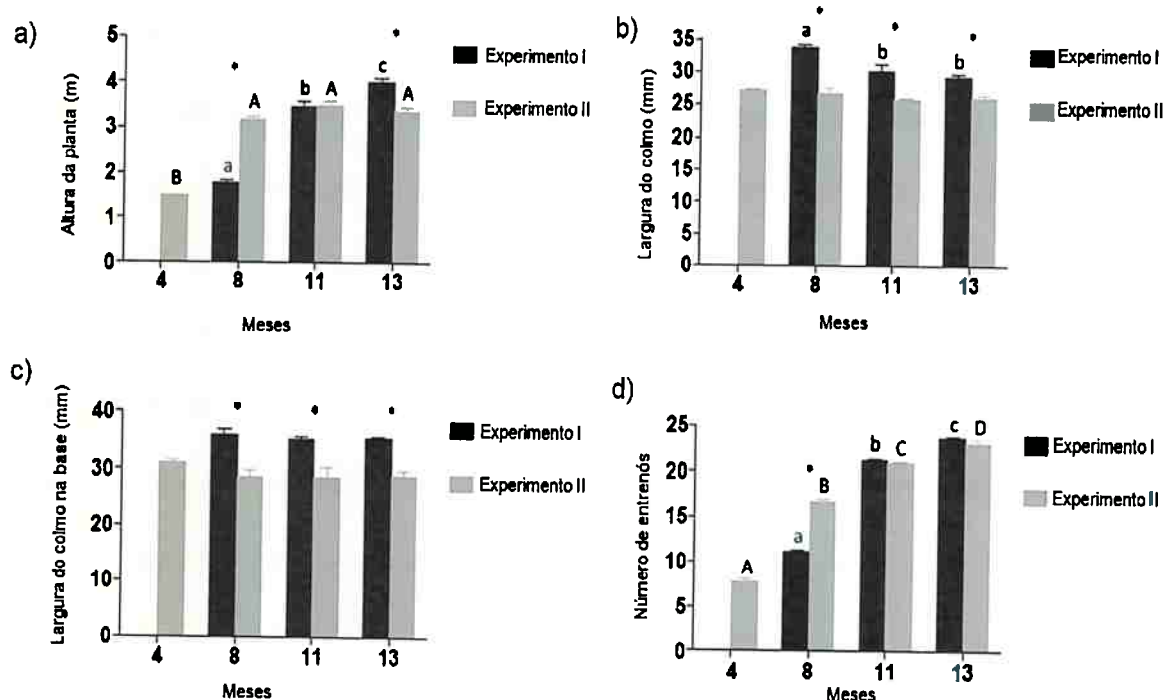
4.1 EXPERIMENTO DE CAMPO

O plantio do experimento de campo I foi realizado em Abril de 2012 e as coletas foram realizadas nos meses de Agosto de 2012, Dezembro de 2012, Março de 2013 e Maio de 2013 quando as plantas apresentavam respectivamente 4 meses, 8 meses, 11 meses e 13 meses de idade. Um segundo experimento de campo teve início em Outubro de 2012 com o mesmo período de duração que o campo I e coletas realizadas em idades equivalentes a 4, 8, 11 e 13 meses. Conforme apresentado nas figuras 3-a e figura 3-b, o experimento de campo I teve um período inicial de crescimento em condições desfavoráveis de temperatura e pluviosidade, enquanto o experimento de campo II teve um crescimento inicial sob condições favoráveis, com elevada temperatura média e precipitação.

Os dados obtidos indicam que no plantio iniciado em Abril de 2012, em um período sob influencia de um menor índice pluviométrico e baixa temperatura média, a cana-de-açúcar apresentou menor crescimento inicial (figura 8-a), com maior diâmetro no colmo (figura 8-b e figura 8-c), menor número de entrenós (figura 8d) e menor conteúdo de Brix (figura 9-a). Em contrapartida, no plantio iniciado em Outubro de 2012, que esteve sob influencia de alta temperatura média e maior regime pluviométrico (figura 3-a) para os primeiros 4 meses de crescimento, a cana-de-açúcar apresentou maior crescimento e produção de Brix.

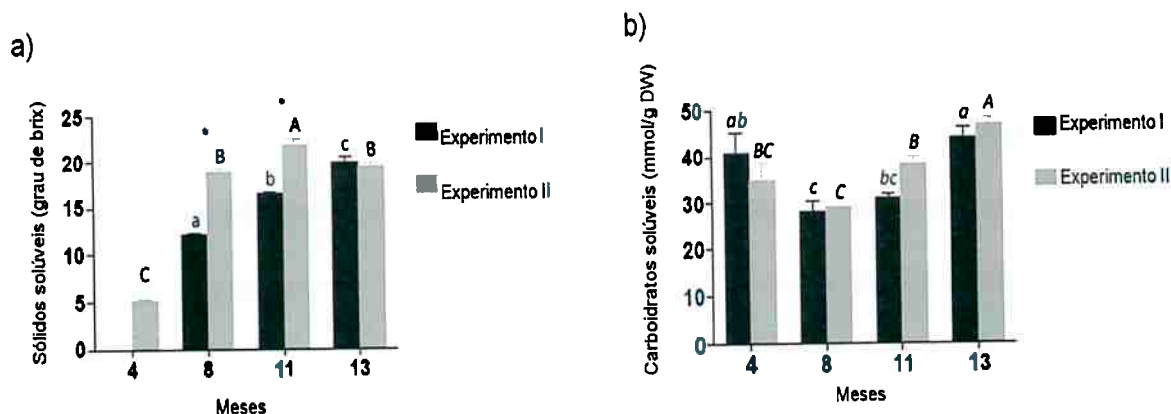
O grau de Brix mensurado ao longo de ambos os experimentos (figura 9-a) mostrou diferença significativa nas idades de 4 meses, 8 meses e 11 meses; aos 13 meses de idade o conteúdo de Brix não demonstrou diferença significativa. Em relação à altura da planta (figura 8-a), no experimento de campo I houve menor

crescimento nos primeiros 4 meses em comparação ao experimento de campo II. Esta diferença se manteve até os 8 meses, e aos 11 meses as plantas não apresentaram diferença significativa na altura. Porém, aos 13 meses, a cana-de-açúcar no experimento de campo I passou a apresentar maior altura.



Fonte: Dados obtidos pela Dr^a Monalisa Carneiro.

Figura 8 - Medidas morfológicas da cana-de-açúcar coletadas ao longo do plantio de ambos os experimentos de campo (4 meses, 8 meses, 11 meses e 13 meses). O experimento I se refere ao plantio iniciado em Abril de 2012 e o experimento II se refere ao plantio realizado em Outubro de 2012. Para cada ponto de coleta de ambos os experimentos foram avaliados testes estatísticos de significância como o ANOVA e comparação múltipla de LSD (Fisher's Least Significant Difference). Os pontos significativamente diferentes estão destacados com sinal de asterisco.



Fonte: Dados obtidos pela Dr^a Monalisa Carneiro.

Figura 9 - Medidas fisiológicas da cana-de-açúcar coletadas ao longo do plantio de ambos os experimentos de campo (4 meses, 8 meses, 11 meses e 13 meses). O experimento I se refere ao plantio iniciado em Abril de 2012 e o experimento II se refere ao plantio realizado em Outubro de 2012. Para cada ponto de coleta de ambos os experimentos foram avaliados testes estatísticos de significância como o ANOVA e comparação múltipla de LSD (*Fisher's Least Significant Difference*). Os pontos significativamente diferentes estão destacados com sinal de asterisco.

Conforme o observado em estudos anteriores, safras de cana de ano e meio, possuem uma taxa de crescimento inicial menor, entretanto, durante o período de maturação a taxa de crescimento é elevada e se mantém além dos 13 meses, se estabilizando próximos dos 16 meses de idade (MARCHIORI, 2004). Diferente do campo II, que é característico de uma safra de cana ano (safra de 12 meses), onde foi observada uma elevada taxa de crescimento inicial. Esta taxa de crescimento se estabiliza antes dos 10 meses, não havendo aumento posterior. Assim, apesar dos experimentos de campo I e 2 terem sido conduzidos ao longo de 13 meses, é esperado que no campo I ainda ocorra crescimento vertical da planta e um maior acúmulo de sólidos solúveis.

Outra importante diferença observada foi o diâmetro do colmo. No experimento de campo I, o colmo da cana-de-açúcar apresentou um diâmetro significativamente maior em relação à cana no experimento de campo II, tanto na base (figura 8-c) quanto à meia altura (figura 8-b). Tal medida se manteve constante

ao longo do experimento de campo II desde a etapa de perfilhamento. No experimento de campo I, não foi possível mensurar o diâmetro da cana aos 4 meses de idade, porém aos 8,11 e 13 meses, também se manteve constante. Esta diferença pode estar associado a um período de crescimento inicial sob baixa pluviosidade e temperatura média, o que ocasionou um retardamento na brotação e perfilhamento da cana, e maior particionamento para o engrossamento do colmo. Dada as observações realizadas, torna-se notório que o perfil de maturação da planta é fortemente definido nos quatro primeiros meses de crescimento e os fatores abióticos em que os mesmos estão submetidos.

4.2 TRANSCRIPTÔMICA

4.2.1 Detecção dos genes significativamente expressos

A análise de expressão gênica pela tecnologia de microarranjo resultou na detecção de um total de 1587 transcritos expressos nos tecidos de folha +1. Destes, apenas aqueles descritos como pertencentes à uma via metabólica foram selecionados para o presente estudo, uma vez que o transcriptoma foi adotado como uma metodologia complementar à análise do metaboloma da cana-de-açúcar, no intuito de se conduzir um estudo integrado entre as duas ômicas. Assim, de um total de 1587 transcritos, foram selecionados 720 transcritos associados a pelo menos uma via metabólica (tabela 1, material suplementar). Para integrar as informações de transcriptômica com os dados de metabolômica nas vias KEGG, utilizamos as informações de anotação gênica do milho como referência para identificar no KEGG a expressão dos genes de cana-de-açúcar. Este procedimento

foi necessário, uma vez que a cana-de-açúcar não possui o genoma completamente descrito nesta plataforma.

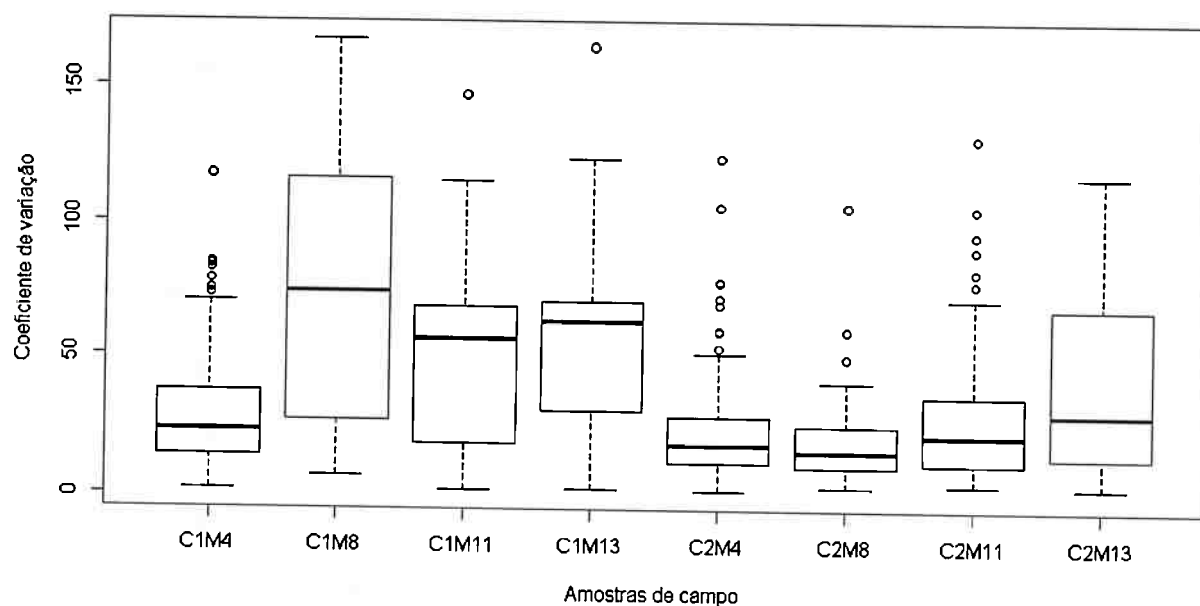
4.3 METABOLÔMICA

4.3.1 Anotação dos picos de massa

Com a aplicação do *pipeline* para identificação de compostos (figura 7), foram identificados um total de 115 metabólitos (tabela 4 do campo I e tabela 5 do campo II), a contar apenas com os candidatos de maior probabilidade. Ao considerarmos as ambiguidades presentes nas anotações, foram obtidos um total de 194 candidatos (tabela 1, material suplementar). Todos os metabólitos foram identificados em todos os pontos de coletas de ambos os experimentos de campo, não sendo observados candidatos pertencentes exclusivamente a um ponto experimental. A intensidade de cada um dos candidatos foi normalizada contra a intensidade do padrão interno específico do modo de ionização (6-Dimetilaminopurina para modo positivo e Glicose C13 para modo negativo de ionização). Em seguida foi realizada uma normalização com a concentração de proteínas solúveis, estimada pelo método adaptado de Bradford (BRADFORD, 1976).

A qualidade entre as réplicas biológicas dos 115 candidatos únicos anotados foi avaliada estimando-se o coeficiente de variação (CV; razão entre o valor do desvio padrão e a média multiplicado por 100). Observamos uma grande quantidade de metabólitos com um coeficiente de variação consideravelmente elevado (figura 10). O critério de avaliação estabelecido para o CV de tolerância máxima foi adotado como 30% conforme Silva e colaboradores (SILVA et al., 2012). Conforme a figura 10, ao conferirmos os metabólitos com um coeficiente de variação menor ou igual a 30% no campo I, verificamos um total de 73, 33, 47 e 28 com estas características

aos 4 meses, 8 meses, 11 meses e 13 meses respectivamente. No campo II, a quantidade de metabólitos observados foi de 90, 103, 81 e 59 para os pontos de 4 meses, 8 meses, 11 meses e 13 meses respectivamente.

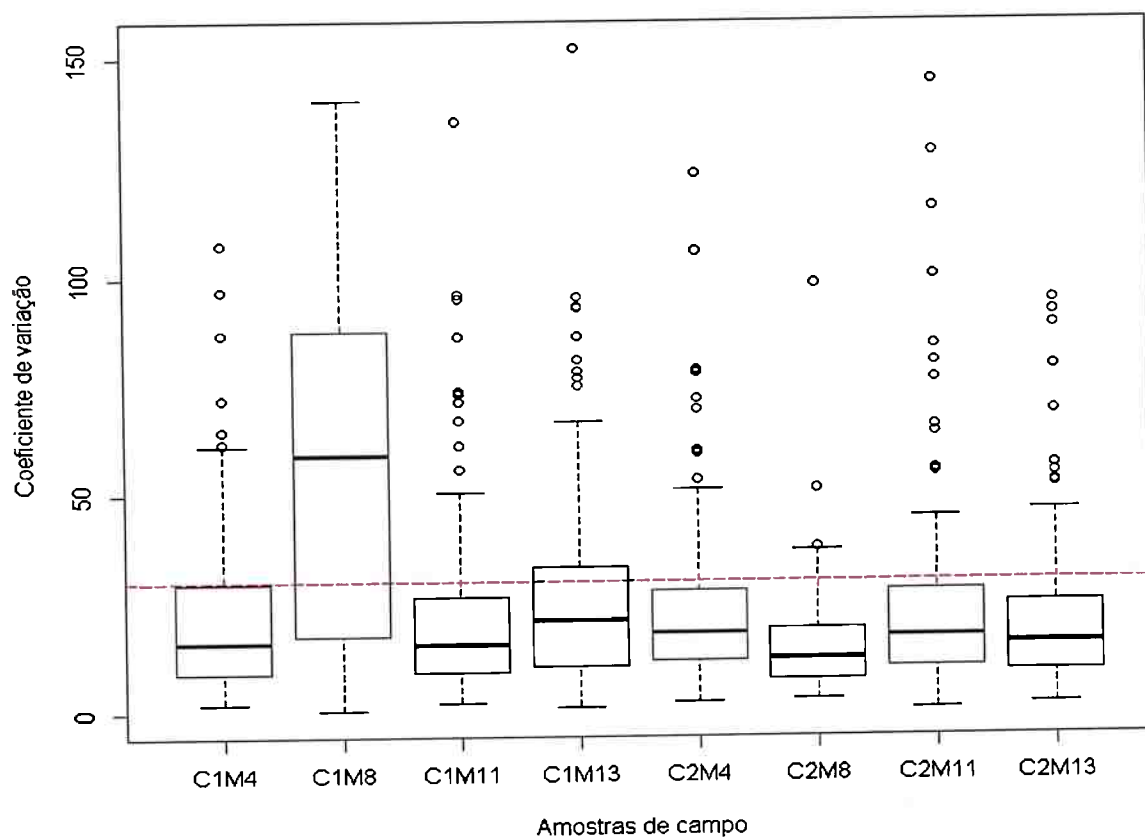


Coef. variação (%)	Campo I				Campo II			
	C1M4	C1M8	C1M11	C1M13	C2M4	C2M8	C2M11	C2M13
0 → 30	73	33	47	28	90	103	81	59
30 → 60	31	7	11	29	18	11	26	10
60 → 90	10	20	53	49	5	0	4	43
>90.00	1	55	4	9	2	1	4	3

Figura 10 - Distribuição do coeficiente de variação na abundância relativa entre as réplicas biológicas dos metabólitos detectados nos experimentos de campo I e II. Estimativa realizada incluindo-se todas as réplicas biológicas. No quadro anexo da figura, os valores em negrito destacam a quantidade de metabólitos com o coeficiente de variação < 30% em cada um dos pontos experimentais.

Uma segunda análise foi realizada excluindo-se a réplica causadora da maior variação em cada um dos pontos. Mantendo-se um total de três réplicas biológicas para cada ponto de coleta. Com isso, os dados passaram a apresentar menor variação, com exceção do ponto de coleta em 8 meses no experimento de campo I, em que grande parte dos dados ainda permaneceram com um coeficiente de

variação acima dos 30% (figura 11). No campo I, um total de 87, 49, 94 e 80 metabólitos passaram a apresentar um coeficiente de variação menor ou igual que 30% nos pontos de coleta em 4 meses, 8 meses, 11 meses e 13 meses respectivamente; para as amostras em períodos equivalentes no campo II passamos a observar 90, 107, 91 e 94 metabólitos respectivamente (figura 11).



Coef. variação (%)	Campo I				Campo II			
	C1M4	C1M8	C1M11	C1M13	C2M4	C2M8	C2M11	C2M13
0 → 30	87	49	94	80	90	107	91	94
30 → 60	20	10	10	23	18	7	15	16
60 → 90	6	36	7	8	5	0	5	3
> 90.00	2	20	4	4	2	1	4	2

Figura 11 - Distribuição do coeficiente de variação na abundância relativa entre as réplicas biológicas dos metabólitos detectados nos experimentos de campo I e II; considerando-se a remoção de uma réplica biológica causadora de maior variação no CV. A linha em vermelha no eixo Y do gráfico, indica o ponto em que o coeficiente de variação é de 30%. . No quadro anexo da figura, os valores em negrito destacam a quantidade de metabólitos com o coeficiente de variação < 30% em cada um dos pontos experimentais.

A lista de metabólitos de candidatos únicos (Tabelas 4 e 5) foi utilizada para o cálculo do escore da atividade das vias, enquanto que a lista contendo os candidatos ambíguos (tabela 2, material suplementar) foi utilizada posteriormente para integrar os metabólitos e genes detectados nas vias metabólicas, a fim de se estabelecer uma análise conjunta com o transcriptoma da cana-de-açúcar. A partir da lista dos 115 candidatos não ambíguos, também aplicamos a análise de variância (ANOVA) para verificar os compostos que apresentaram diferença significativa na abundância ao longo dos experimentos. No mais, ao mapearmos os genes expressos e os metabólitos identificados em suas possíveis vias biológicas, observamos uma consenso de 60 vias com pelo menos um metabólito e um transcrito participando na via (figura 12).

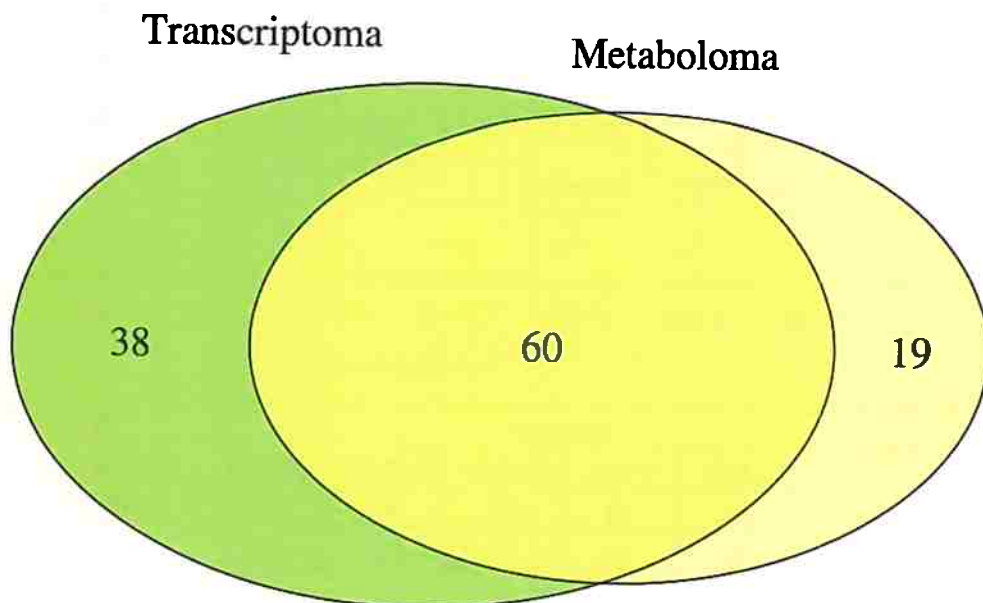


Figura 12 - Quantidade de vias em que os genes de cana-de-açúcar expressos e metabólitos detectados participam, demonstrando assim uma consenso de 60 vias metabólicas com pelo menos um gene e um metabólito identificado.

Tabela 4 - Abundância relativa dos 115 compostos mais prováveis identificados em tecidos de folha +1 de cana-de-açúcar no experimento de campo I. Para cada metabólito identificado foi estimado a sua abundância relativa normalizada contra o padrão interno dos respectivos modos de aquisição (positivo/ negativo) e contra a concentração de proteínas. Os valores indicados à direita de cada intensidade normalizada representa o desvio padrão em relação à média entre as réplicas biológicas. A abreviação C1 indica "campo I"; M4, M8, M11 e M13 indicam as coletas aos 4 meses, 8 meses, 11 meses e 13 meses após o plantio.

(Continua)

Nome do metabólito	Coluna/ modo ionização	C1M4	C1M8	C1M11	C1M13	Nº vias map
<i>2-Oxoglutarate</i>	Hilic, neg(-)	0,326 ± 0,026	0,176 ± 0,037	0,438 ± 0,06	0,22 ± 0,027	17
<i>L-Tyrosine</i>	Hilic, neg(-)	5,366 ± 0,098	4,748 ± 0,345	5,63 ± 0,159	4,133 ± 0,1	16
<i>Citrate</i>	C18, pos(+)	2,261 ± 0,677	3,636 ± 0,487	7,115 ± 0,988	6,755 ± 1,159	9
<i>(S)-2-Acetolactate</i>	Hilic, pos(+)	0,098 ± 0,009	0,041 ± 0,036	0,089 ± 0,007	0,08 ± 0,02	8
<i>(S)-Malate</i>	Hilic, pos(+)	0,846 ± 0,113	0,418 ± 0,367	0,684 ± 0,03	0,573 ± 0,22	8
<i>2-Oxoadipate</i>	Hilic, neg(-)	0,215 ± 0,031	0,167 ± 0,027	0,366 ± 0,075	0,346 ± 0,007	7
<i>cis-Aconitate</i>	Hilic, neg(-)	0,45 ± 0,225	0,261 ± 0,046	0,515 ± 0,491	0,491 ± 0,021	7
<i>L-Proline</i>	Hilic, pos(+)	0,083 ± 0,018	0,034 ± 0,03	0,067 ± 0,058	0,073 ± 0,057	7
<i>1-Deoxy-D-xylulose 5-phosphate</i>	Hilic, neg(-)	0,095 ± 0,013	0,05 ± 0,017	0,079 ± 0,022	0,054 ± 0,018	5
<i>3-Dehydroquinate</i>	Hilic, neg(-)	0,155 ± 0,007	0,119 ± 0,008	0,141 ± 0,027	0,108 ± 0,005	5
<i>3-Dehydroshikimate</i>	C18, pos(+)	3,639 ± 0,8	1,459 ± 0,182	1,757 ± 0,236	1,19 ± 0,326	5
<i>4-Methylthio-2-oxobutanoic acid</i>	C18, pos(+)	17,495 ± 1,006	14,471 ± 0,053	17,214 ± 0,321	12,63 ± 0,329	5
<i>Aminoimidazole ribotide</i>	C18, pos(+)	0,077 ± 0,047	0,262 ± 0,313	0,14 ± 0,102	0,255 ± 0,238	5
<i>2-C-Methyl-D-erythritol 4-phosphate</i>	Hilic, neg(-)	0,923 ± 0,207	0,389 ± 0,068	0,546 ± 0,095	0,496 ± 0,118	4
<i>2-Methylmaleate</i>	Hilic, neg(-)	0,368 ± 0,136	0,046 ± 0,018	0,411 ± 0,556	0,219 ± 0,334	4
<i>Acetyl phosphate</i>	Hilic, neg(-)	0,49 ± 0,072	0,315 ± 0,084	0,164 ± 0,157	0,059 ± 0,035	4
<i>Ascorbate</i>	C18, pos(+)	0,337 ± 0,063	0,342 ± 0,041	0,601 ± 0,171	0,471 ± 0,031	4
<i>Benzoate</i>	Hilic, neg(-)	0,141 ± 0,087	0,082 ± 0,013	0,132 ± 0,006	0,151 ± 0,073	4
<i>Chlorogenate</i>	C18, pos(+)	2,869 ± 2,781	4,857 ± 3,141	6,783 ± 1,086	4,589 ± 0,607	4
<i>D-Sorbitol</i>	Hilic, neg(-)	1,002 ± 0,253	1,136 ± 0,094	2,146 ± 0,457	1,761 ± 0,551	4
<i>Indolylmethyl-desulfoglucosinolate</i>	Hilic, neg(-)	5,603 ± 1,534	7,798 ± 2,603	12,889 ± 1,687	9,2 ± 1,76	4

Tabela 4 - Abundância relativa dos 115 compostos mais prováveis identificados em tecidos de folha +1 de cana-de-açúcar no experimento de campo I. Para cada metabólito identificado foi estimado a sua abundância relativa normalizada contra o padrão interno dos respectivos modos de aquisição (positivo/ negativo) e contra a concentração de proteínas. Os valores indicados à direita de cada intensidade normalizada representa o desvio padrão em relação à média entre as réplicas biológicas. A abreviação C1 indica "campo I"; M4, M8, M11 e M13 indicam as coletas aos 4 meses, 8 meses, 11 meses e 13 meses após o plantio.

(Continuação)

Nome do metabólito	Coluna/ modo ionização	C1M4	C1M8	C1M11	C1M13	Nº vias map
<i>L-Pipecolate</i>	Hilic, pos(+)	1,057 ± 0,485	1,235 ± 1,073	2,064 ± 0,292	3,12 ± 0,673	4
<i>Nicotinate</i>	Hilic, pos(+)	0,314 ± 0,019	0,076 ± 0,066	0,107 ± 0,039	0,139 ± 0,024	4
<i>Orthophosphate</i>	Hilic, pos(+)	0,828 ± 0,248	0,285 ± 0,248	0,633 ± 0,099	0,597 ± 0,157	4
<i>Zeatin</i>	Hilic, neg(-)	0,124 ± 0,02	0,112 ± 0,013	0,209 ± 0,014	0,192 ± 0,025	4
<i>(6Z,9Z,12Z)-Octadecatrienoic acid</i>	Hilic, neg(-)	0,133 ± 0,042	0,059 ± 0,035	0,052 ± 0,01	0,049 ± 0,02	3
<i>(9Z,15Z)-(13S)-12,13-Epoxyoctadeca-9,11,15-trienoic acid</i>	Hilic, pos(+)	0,299 ± 0,06	0,192 ± 0,168	0,319 ± 0,022	0,239 ± 0,021	3
<i>1,7-Dimethylxanthine</i>	Hilic, neg(-)	0,531 ± 0,022	0,488 ± 0,089	1,327 ± 1,274	0,634 ± 0,124	3
<i>2-Oxo-6-methylthiohexanoic acid</i>	C18, pos(+)	1,326 ± 0,214	1,811 ± 0,639	2,701 ± 0,509	1,943 ± 0,632	3
<i>2-Oxo-8-methylthiooctanoic acid</i>	C18, pos(+)	2,75 ± 0,139	2,297 ± 0,074	2,702 ± 0,061	1,887 ± 0,052	3
<i>2-Oxo-9-methylthiononanoic acid</i>	Hilic, pos(+)	0,138 ± 0,038	0,029 ± 0,025	0,03 ± 0,008	0,023 ± 0,005	3
<i>2-Succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate</i>	Hilic, neg(-)	5,395 ± 1,971	6,665 ± 1,258	6,674 ± 0,896	7,072 ± 3,174	3
<i>5,10-Methenyltetrahydrofolate</i>	Hilic, neg(-)	0,081 ± 0,032	0,114 ± 0,083	0,201 ± 0,025	0,133 ± 0,01	3
<i>5-Amino-6-(1-D-ribitylamino)uracil</i>	Hilic, pos(+)	0,007 ± 0,002	0,009 ± 0,011	0,031 ± 0,006	0,13 ± 0,066	3
<i>5-Hydroxyconiferaldehyde</i>	Hilic, neg(-)	0,3 ± 0,084	0,286 ± 0,081	0,511 ± 0,032	0,402 ± 0,096	3
<i>8-[(1R,2R)-3-Oxo-2-((Z)-pent-2-enyl)cyclopentyl]octanoate</i>	Hilic, neg(-)	0,181 ± 0,012	0,079 ± 0,005	0,158 ± 0,019	0,106 ± 0,013	3
<i>Betaine</i>	Hilic, pos(+)	3,385 ± 0,673	1,29 ± 1,199	1,848 ± 0,216	2,686 ± 0,799	3
<i>Cinnamaldehyde</i>	Hilic, pos(+)	0,335 ± 0,149	0,126 ± 0,11	0,113 ± 0,014	0,094 ± 0,008	3
<i>cis-Zeatin</i>	Hilic, pos(+)	0,198 ± 0,047	0,11 ± 0,096	0,167 ± 0,044	0,196 ± 0,044	3
<i>Coniferyl aldehyde</i>	Hilic, pos(+)	0,286 ± 0,023	0,444 ± 0,384	0,587 ± 0,068	0,483 ± 0,046	3
<i>Desulfoglucotropeolin</i>	Hilic, neg(-)	0,984 ± 0,358	1,21 ± 0,215	1,265 ± 0,188	1,327 ± 0,544	3

Tabela 4 - Abundância relativa dos 115 compostos mais prováveis identificados em tecidos de folha +1 de cana-de-açúcar no experimento de campo I. Para cada metabólito identificado foi estimado a sua abundância relativa normalizada contra o padrão interno dos respectivos modos de aquisição (positivo/ negativo) e contra a concentração de proteínas. Os valores indicados à direita de cada intensidade normalizada representa o desvio padrão em relação à média entre as réplicas biológicas. A abreviação C1 indica "campo I"; M4, M8, M11 e M13 indicam as coletas aos 4 meses, 8 meses, 11 meses e 13 meses após o plantio.

(Continuação)

Nome do metabólito	Coluna/ modo ionização	C1M4	C1M8	C1M11	C1M13	Nº vias map
<i>Folate</i>	Hilic, neg(-)	0,48 ± 0,011	0,117 ± 0,043	0,273 ± 0,07	0,143 ± 0,049	3
<i>Methylmalonate</i>	Hilic, pos(+)	0,172 ± 0,03	0,082 ± 0,084	0,088 ± 0,011	0,068 ± 0,021	3
<i>Pelargonidin 3-O-glucoside</i>	Hilic, pos(+)	0,396 ± 0,122	0,214 ± 0,232	0,326 ± 0,051	0,207 ± 0,158	3
<i>Protochlorophyllide</i>	Hilic, neg(-)	0,143 ± 0,014	0,179 ± 0,016	0,25 ± 0,042	0,181 ± 0,018	3
<i>Riboflavin</i>	Hilic, pos(+)	0,169 ± 0,015	0,079 ± 0,072	0,125 ± 0,008	0,09 ± 0,018	3
<i>S-(4-Methylthiobutylthiohydroximoyl)-L-cysteine</i>	Hilic, neg(-)	0,473 ± 0,014	0,665 ± 0,098	0,947 ± 0,141	0,78 ± 0,045	3
<i>Sinapyl alcohol</i>	Hilic, pos(+)	1,474 ± 0,656	0,416 ± 0,424	0,49 ± 0,096	0,375 ± 0,281	3
<i>(-)-Epiarzelechin</i>	Hilic, pos(+)	3,845 ± 2,755	0,449 ± 0,63	0,16 ± 0,009	0,114 ± 0,001	2
<i>(S)-4-Hydroxymandelonitrile</i>	Hilic, pos(+)	0,118 ± 0,026	0,084 ± 0,073	0,049 ± 0,003	0,037 ± 0,005	2
<i>(S)-Allantoin</i>	Hilic, pos(+)	0,158 ± 0,026	0,025 ± 0,026	0,033 ± 0,009	0,017 ± 0,001	2
<i>1,2-Dihydroxy-5-(methylthio)pent-1-en-3-one</i>	C18, pos(+)	5,145 ± 4,475	8,158 ± 4,805	10,584 ± 1,384	7,308 ± 1,303	2
<i>13(1)-Oxo-magnesium-protoporphyrin IX 13-monomethyl ester</i>	Hilic, neg(-)	0,136 ± 0,021	0,171 ± 0,021	0,226 ± 0,038	0,154 ± 0,031	2
<i>2-Maleylacetate</i>	Hilic, pos(+)	0,165 ± 0,026	0,009 ± 0,009	0,031 ± 0,012	0,017 ± 0,006	2
<i>3,4-Dihydroxymandelaldehyde</i>	Hilic, pos(+)	0,082 ± 0,005	0,057 ± 0,049	0,074 ± 0,004	0,061 ± 0,009	2
<i>3-beta-D-Galactosyl-sn-glycerol</i>	Hilic, neg(-)	0,183 ± 0,019	0,056 ± 0,019	0,234 ± 0,142	0,175 ± 0,062	2
<i>4-Aminobenzoate</i>	Hilic, pos(+)	1,678 ± 0,15	0,988 ± 0,861	3,071 ± 0,329	3,616 ± 0,409	2
<i>4-Fumarylacetoacetate</i>	Hilic, pos(+)	0,353 ± 0,148	0,13 ± 0,114	0,121 ± 0,018	0,094 ± 0,007	2
<i>5-Hydroxyconiferyl alcohol</i>	C18, pos(+)	2,087 ± 0,428	1,613 ± 0,094	2,513 ± 0,407	0,97 ± 0,543	2
<i>7-Aminomethyl-7-carbaguanine</i>	Hilic, pos(+)	1,116 ± 0,099	0,502 ± 0,451	0,658 ± 0,034	0,378 ± 0,039	2

Tabela 4 - Abundância relativa dos 115 compostos mais prováveis identificados em tecidos de folha +1 de cana-de-açúcar no experimento de campo I. Para cada metabólito identificado foi estimado a sua abundância relativa normalizada contra o padrão interno dos respectivos modos de aquisição (positivo/ negativo) e contra a concentração de proteínas. Os valores indicados à direita de cada intensidade normalizada representa o desvio padrão em relação à média entre as réplicas biológicas. A abreviação C1 indica "campo I"; M4, M8, M11 e M13 indicam as coletas aos 4 meses, 8 meses, 11 meses e 13 meses após o plantio.

(Continuação)

Nome do metabólito	Coluna/ modo ionização	C1M4	C1M8	C1M11	C1M13	Nº vias map
<i>7-Methylthioheptanaldoxime</i>	Hilic, pos(+)	0,728 ± 0,06	0,324 ± 0,291	2,351 ± 0,121	1,397 ± 0,172	2
<i>Acetone cyanohydrin</i>	C18, pos(+)	1,907 ± 1,229	0,546 ± 0,141	1,117 ± 0,261	0,606 ± 0,062	2
<i>Adenosine</i>	Hilic, pos(+)	0,469 ± 0,261	2,195 ± 2,514	2,188 ± 1,336	1,649 ± 1,533	2
<i>Choline phosphate</i>	C18, pos(+)	0,644 ± 0,071	0,702 ± 0,004	1,111 ± 0,111	0,664 ± 0,189	2
<i>cis-beta-D-Glucosyl-2-hydroxycinnamate</i>	Hilic, pos(+)	0,285 ± 0,007	0,196 ± 0,171	0,332 ± 0,097	0,273 ± 0,086	2
<i>Cytidine</i>	Hilic, pos(+)	0,283 ± 0,029	0,128 ± 0,114	0,112 ± 0,008	0,079 ± 0,02	2
<i>D-Ribose</i>	Hilic, neg(-)	0,593 ± 0,072	0,294 ± 0,044	0,463 ± 0,024	0,361 ± 0,116	2
<i>Gibberellin A15 open lactone</i>	Hilic, pos(+)	0,185 ± 0,014	0,128 ± 0,113	0,147 ± 0,008	0,115 ± 0,007	2
<i>Guanosine</i>	Hilic, pos(+)	0,36 ± 0,072	0,164 ± 0,169	0,143 ± 0,015	0,08 ± 0,065	2
<i>Hypoxanthine</i>	Hilic, neg(-)	0,576 ± 0,084	0,111 ± 0,06	0,766 ± 0,209	0,399 ± 0,142	2
<i>Indolepyruvate</i>	Hilic, pos(+)	0,519 ± 0,558	0,063 ± 0,055	0,166 ± 0,092	0,161 ± 0,075	2
<i>Kaempferol 3-O-glucoside</i>	C18, pos(+)	1,829 ± 1,309	2,114 ± 1,777	1,811 ± 1,287	2,495 ± 0,587	2
<i>L-Cystine</i>	Hilic, pos(+)	1,572 ± 0,243	0,638 ± 0,554	0,574 ± 0,133	0,594 ± 0,074	2
<i>Leucodelphinidin</i>	Hilic, neg(-)	0,151 ± 0,086	0,054 ± 0,075	0,125 ± 0,092	0,166 ± 0,042	2
<i>Nicotinamide D-ribonucleotide</i>	Hilic, pos(+)	0,036 ± 0,007	0,025 ± 0,023	0,078 ± 0,021	0,102 ± 0,016	2
<i>Piceatannol</i>	Hilic, neg(-)	0,666 ± 0,019	0,599 ± 0,104	1,057 ± 0,083	0,906 ± 0,16	2
<i>Pyridoxine</i>	Hilic, pos(+)	0,085 ± 0,023	0,092 ± 0,081	0,055 ± 0,012	0,031 ± 0,004	2
<i>Pyropheophorbide a</i>	Hilic, pos(+)	2,189 ± 0,308	0,751 ± 0,713	1,624 ± 0,168	1,516 ± 0,16	2
<i>trans-Zeatin riboside monophosphate</i>	Hilic, neg(-)	0,668 ± 0,048	0,42 ± 0,061	0,631 ± 0,318	0,431 ± 0,092	2
<i>UDP-6-sulfoquinovose</i>	Hilic, pos(+)	0,183 ± 0,015	0,113 ± 0,099	0,07 ± 0,012	0,034 ± 0,029	2

Tabela 4 - Abundância relativa dos 115 compostos mais prováveis identificados em tecidos de folha +1 de cana-de-açúcar no experimento de campo I. Para cada metabólito identificado foi estimado a sua abundância relativa normalizada contra o padrão interno dos respectivos modos de aquisição (positivo/ negativo) e contra a concentração de proteínas. Os valores indicados à direita de cada intensidade normalizada representa o desvio padrão em relação à média entre as réplicas biológicas. A abreviação C1 indica "campo I"; M4, M8, M11 e M13 indicam as coletas aos 4 meses, 8 meses, 11 meses e 13 meses após o plantio.

(Continuação)

Nome do metabólito	Coluna/ modo ionização	C1M4	C1M8	C1M11	C1M13	Nº vias map
<i>(9Z)-Hexadecenoic acid</i>	Hilic, neg(-)	0,171 ± 0,01	0,101 ± 0,029	0,142 ± 0,023	0,119 ± 0,048	1
<i>1,3,7-Trimethyluric acid</i>	Hilic, neg(-)	0,655 ± 0,017	0,273 ± 0,055	0,524 ± 0,081	0,396 ± 0,116	1
<i>1-O-Sinapoyl-beta-D-glucose</i>	Hilic, neg(-)	0,027 ± 0,005	0,166 ± 0,12	0,06 ± 0,019	0,068 ± 0,002	1
<i>2'-Deoxyinosine 5'-phosphate</i>	C18, pos(+)	5,161 ± 0,77	0,731 ± 0,052	1,234 ± 0,823	0,914 ± 0,606	1
<i>2-Hydroxy-3-(4-hydroxyphenyl)propenoate</i>	C18, pos(+)	2,891 ± 0,239	2,356 ± 0,124	2,521 ± 0,204	2,152 ± 0,233	1
<i>3-Mercaptolactate</i>	Hilic, neg(-)	0,75 ± 0,121	0,568 ± 0,116	0,986 ± 0,066	0,853 ± 0,133	1
<i>3-Methoxyanthranilate</i>	Hilic, pos(+)	0,653 ± 0,129	0,476 ± 0,412	0,422 ± 0,032	0,248 ± 0,021	1
<i>4-Hydroxy-L-threonine</i>	Hilic, pos(+)	9,172 ± 1,309	4,187 ± 4,255	6,935 ± 0,663	5,114 ± 1,715	1
<i>6beta,7beta-Dihydroxykaurenoic acid</i>	Hilic, pos(+)	0,1 ± 0,017	0,027 ± 0,023	0,019 ± 0,004	0,016 ± 0,006	1
<i>Afzelin</i>	C18, pos(+)	0,812 ± 0,423	1,138 ± 0,746	1,007 ± 0,4	0,993 ± 0,038	1
<i>alpha-D-Galactosyl-(1->3)-1D-myo-inositol</i>	Hilic, neg(-)	19,784 ± 1,567	11,585 ± 2,986	22,93 ± 5,977	11,18 ± 3,261	1
<i>Benzene-1,2,4-triol</i>	C18, pos(+)	6,587 ± 1,103	3,29 ± 0,629	5,801 ± 0,495	3,471 ± 0,608	1
<i>Biliverdin</i>	Hilic, neg(-)	0,444 ± 0,065	0,445 ± 0,063	0,698 ± 0,041	0,463 ± 0,058	1
<i>D-Gal alpha 1->6D-Gal alpha 1->6D-Glucose</i>	C18, pos(+)	2,595 ± 1,51	1,932 ± 0,29	3,087 ± 0,204	1,641 ± 0,042	1
<i>Glutamate</i>	C18, pos(+)	0,854 ± 0,418	0,281 ± 0,059	0,459 ± 0,057	0,327 ± 0,07	1
<i>L-Selenocystathionine</i>	C18, pos(+)	0,414 ± 0,125	0,196 ± 0,026	0,262 ± 0,056	0,218 ± 0,054	1
<i>Molybdate</i>	Hilic, pos(+)	0,097 ± 0,019	0,056 ± 0,049	0,06 ± 0,007	0,043 ± 0,003	1
<i>Myrtenol</i>	Hilic, pos(+)	0,074 ± 0,009	0,023 ± 0,021	0,028 ± 0,002	0,025 ± 0,003	1
<i>N1,N5,N10-Tricoumaroyl spermidine</i>	Hilic, pos(+)	0,329 ± 0,019	0,418 ± 0,385	0,603 ± 0,026	0,413 ± 0,036	1
<i>Stachyose</i>	Hilic, pos(+)	0,018 ± 0,008	0,013 ± 0,011	0,021 ± 0,001	0,018 ± 0,005	1

Tabela 4 - Abundância relativa dos 115 compostos mais prováveis identificados em tecidos de folha +1 de cana-de-açúcar no experimento de campo I. Para cada metabólito identificado foi estimado a sua abundância relativa normalizada contra o padrão interno dos respectivos modos de aquisição (positivo/ negativo) e contra a concentração de proteínas. Os valores indicados à direita de cada intensidade normalizada representa o desvio padrão em relação à média entre as réplicas biológicas. A abreviação C1 indica "campo I"; M4, M8, M11 e M13 indicam as coletas aos 4 meses, 8 meses, 11 meses e 13 meses após o plantio.

(Conclusão)

Nome do metabólito	Coluna/ modo ionização	C1M4	C1M8	C1M11	C1M13	Nº vias map
<i>Syringin</i>	Hilic, pos(+)	0,406 ± 0,049	0,227 ± 0,198	0,263 ± 0,059	0,162 ± 0,004	1
<i>THF-L-glutamate</i>	Hilic, pos(+)	0,225 ± 0,025	0,115 ± 0,1	0,078 ± 0,013	0,057 ± 0,013	1
<i>Xanthurenic acid</i>	Hilic, pos(+)	0,071 ± 0,011	0,038 ± 0,034	0,035 ± 0,005	0,022 ± 0,006	1
<i>2-(S-Glutathionyl)acetyl glutathione</i>	Hilic, neg(-)	0,089 ± 0,009	0,05 ± 0,013	0,08 ± 0,032	0,087 ± 0,014	0
<i>2-Chloro-5-methyl-cis-dienelactone</i>	C18, pos(+)	0,707 ± 0,028	0,551 ± 0,014	0,683 ± 0,021	0,497 ± 0,019	0
<i>3-Keto-beta-D-galactose</i>	Hilic, neg(-)	0,27 ± 0,059	0,106 ± 0,016	0,138 ± 0,058	0,131 ± 0,023	0
<i>3-Ureidoisobutyrate</i>	Hilic, pos(+)	0,151 ± 0,012	0,083 ± 0,077	0,053 ± 0,016	0,027 ± 0,017	0
<i>5'-Deoxyadenosine</i>	Hilic, pos(+)	0,055 ± 0,007	0,043 ± 0,039	0,078 ± 0,011	0,199 ± 0,19	0
<i>5-Fluorodeoxyuridine</i>	C18, pos(+)	0,338 ± 0,185	1,218 ± 0,356	1,931 ± 0,373	1,425 ± 0,089	0
<i>6-Thiourate</i>	Hilic, neg(-)	2,86 ± 0,265	2,86 ± 0,631	3,899 ± 0,455	2,973 ± 0,588	0
<i>Benzo[a]pyrene-4,5-oxide</i>	Hilic, neg(-)	0,728 ± 0,252	0,74 ± 0,352	0,602 ± 0,072	0,389 ± 0,131	0
<i>cis-4-Carboxymethylenebut-2-en-4-olide</i>	Hilic, neg(-)	0,978 ± 0,092	1,01 ± 0,211	1,263 ± 0,149	0,981 ± 0,244	0
<i>cis-Acetylacrylate</i>	C18, pos(+)	0,986 ± 0,206	0,443 ± 0,088	0,745 ± 0,099	0,376 ± 0,162	0
<i>Glutathione episulfonium ion</i>	Hilic, neg(-)	0,213 ± 0,032	0,253 ± 0,063	0,33 ± 0,099	0,246 ± 0,064	0
<i>S-Adenosyl-4-methylthio-2-oxobutanoate</i>	Hilic, neg(-)	0,367 ± 0,051	0,225 ± 0,025	0,275 ± 0,078	0,248 ± 0,036	0

Tabela 5 - Abundância relativa dos 115 compostos identificados em tecidos de folha +1 de cana-de-açúcar no experimento de campo II. Para cada metabólito identificado foi estimado a sua abundância relativa normalizada contra o padrão interno dos respectivos modos de aquisição (positivo/ negativo) e contra a concentração de proteínas. Os valores indicados à direita de cada intensidade normalizada representa o desvio padrão em relação à média entre as réplicas biológicas. A abreviação C2 indica "campo II"; M4, M8, M11 e M13 indicam as coletas aos 4 meses, 8 meses, 11 meses e 13 meses após o plantio.

(Continua)

Nome do metabólito	Coluna/ modo ionização	C2M4	C2M8	C2M11	C2M13	Nº vias map
<i>2-Oxoglutarate</i>	Hilic, neg(-)	0,258 ± 0,04	0,082 ± 0,013	0,216 ± 0,065	0,126 ± 0,022	17
<i>L-Tyrosine</i>	Hilic, neg(-)	4,148 ± 0,108	3,651 ± 0,131	4,282 ± 0,199	3,817 ± 0,085	16
<i>Citrate</i>	C18, pos(+)	1,416 ± 0,721	1,484 ± 0,171	2,435 ± 0,261	2,395 ± 0,237	9
<i>(S)-2-Acetylactate</i>	Hilic, pos(+)	0,129 ± 0,03	0,072 ± 0,007	0,078 ± 0,01	0,078 ± 0,004	8
<i>(S)-Malate</i>	Hilic, pos(+)	0,689 ± 0,094	0,526 ± 0,032	0,509 ± 0,076	0,468 ± 0,038	8
<i>2-Oxoadipate</i>	Hilic, neg(-)	0,162 ± 0,032	0,252 ± 0,044	0,481 ± 0,165	0,25 ± 0,072	7
<i>cis-Aconitate</i>	Hilic, neg(-)	0,439 ± 0,053	0,228 ± 0,022	0,533 ± 0,164	0,203 ± 0,18	7
<i>L-Proline</i>	Hilic, pos(+)	0,059 ± 0,028	0,04 ± 0,005	0,099 ± 0,127	0,029 ± 0,016	7
<i>1-Deoxy-D-xylulose 5-phosphate</i>	Hilic, neg(-)	0,095 ± 0,014	0,045 ± 0,004	0,102 ± 0,044	0,046 ± 0,008	5
<i>3-Dehydroquininate</i>	Hilic, neg(-)	0,091 ± 0,023	0,086 ± 0,013	0,101 ± 0,006	0,109 ± 0,027	5
<i>3-Dehydroshikimate</i>	C18, pos(+)	1,174 ± 0,166	1,169 ± 0,209	1,698 ± 0,077	2,064 ± 0,26	5
<i>4-Methylthio-2-oxobutanoic acid</i>	C18, pos(+)	14,455 ± 2,194	12,517 ± 0,55	6,843 ± 0,545	5,824 ± 0,098	5
<i>Aminoimidazole ribotide</i>	C18, pos(+)	1,201 ± 0,274	0,197 ± 0,075	0,041 ± 0,005	0,289 ± 0,123	5
<i>2-C-Methyl-D-erythritol 4-phosphate</i>	Hilic, neg(-)	0,824 ± 0,05	0,471 ± 0,082	0,513 ± 0,054	0,282 ± 0,039	4
<i>2-Methylmaleate</i>	Hilic, neg(-)	0,782 ± 0,119	0,59 ± 0,131	0,045 ± 0,01	0,047 ± 0,032	4
<i>Acetyl phosphate</i>	Hilic, neg(-)	0,223 ± 0,081	0,02 ± 0,004	0,027 ± 0,003	0,021 ± 0,002	4
<i>Ascorbate</i>	C18, pos(+)	0,348 ± 0,065	0,33 ± 0,009	1,466 ± 0,241	0,603 ± 0,072	4
<i>Benzoate</i>	Hilic, neg(-)	0,084 ± 0,018	0,162 ± 0,021	0,061 ± 0,007	0,054 ± 0,008	4
<i>Chlorogenate</i>	C18, pos(+)	8,117 ± 3,202	6,568 ± 1,068	19,644 ± 2,36	12,869 ± 0,474	4

Tabela 5 - Abundância relativa dos 115 compostos identificados em tecidos de folha +1 de cana-de-açúcar no experimento de campo II. Para cada metabólito identificado foi estimado a sua abundância relativa normalizada contra o padrão interno dos respectivos modos de aquisição (positivo/ negativo) e contra a concentração de proteínas. Os valores indicados à direita de cada intensidade normalizada representa o desvio padrão em relação à média entre as réplicas biológicas. A abreviação C2 indica "campo II"; M4, M8, M11 e M13 indicam as coletas aos 4 meses, 8 meses, 11 meses e 13 meses após o plantio.

							(Continua)
Nome do metabólito	Coluna/ modo ionização	C2M4	C2M8	C2M11	C2M13	Nº vias map	
<i>D-Sorbitol</i>	Hilic, neg(-)	2,504 ± 0,4	1,521 ± 0,134	1,814 ± 0,348	1,26 ± 0,071	4	
<i>Indolylmethyl-desulfoglucosinolate</i>	Hilic, neg(-)	7,659 ± 1,05	4,728 ± 0,402	22,095 ± 5,116	16,008 ± 2,08	4	
<i>L-Pipecolate</i>	Hilic, pos(+)	1,045 ± 0,182	2,02 ± 0,563	0,525 ± 0,145	0,319 ± 0,072	4	
<i>Nicotinate</i>	Hilic, pos(+)	0,427 ± 0,332	0,046 ± 0,008	0,208 ± 0,241	0,213 ± 0,201	4	
<i>Orthophosphate</i>	Hilic, pos(+)	0,442 ± 0,113	0,402 ± 0,026	0,511 ± 0,014	0,408 ± 0,097	4	
<i>Zeatin</i>	Hilic, neg(-)	0,1 ± 0,028	0,107 ± 0,019	0,204 ± 0,048	0,143 ± 0,011	4	
<i>(6Z,9Z,12Z)-Octadecatrienoic acid</i>	Hilic, neg(-)	0,099 ± 0,037	0,076 ± 0,003	0,031 ± 0,007	0,064 ± 0,05	3	
<i>(9Z,15Z)-(13S)-12,13-Epoxyoctadeca-9,11,15-trienoic acid</i>	Hilic, pos(+)	0,141 ± 0,101	0,204 ± 0,014	0,146 ± 0,028	0,161 ± 0,027	3	
<i>1,7-Dimethylxanthine</i>	Hilic, neg(-)	1,105 ± 0,26	1,211 ± 0,097	1,989 ± 0,761	1,231 ± 0,464	3	
<i>2-Oxo-6-methylthiohexanoic acid</i>	C18, pos(+)	2,634 ± 0,814	1,577 ± 0,21	3,884 ± 0,79	3,319 ± 0,975	3	
<i>2-Oxo-8-methylthiooctanoic acid</i>	C18, pos(+)	2,201 ± 0,271	1,89 ± 0,058	1,071 ± 0,1	0,925 ± 0,009	3	
<i>2-Oxo-9-methylthiononanoic acid</i>	Hilic, pos(+)	0,031 ± 0,006	0,053 ± 0,01	0,054 ± 0,008	0,048 ± 0,025	3	
<i>2-Succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate</i>	Hilic, neg(-)	1,94 ± 0,719	7,981 ± 0,795	1,998 ± 0,716	4,789 ± 1,01	3	
<i>5,10-Methenyltetrahydrofolate</i>	Hilic, neg(-)	0,099 ± 0,007	0,124 ± 0,01	0,088 ± 0,067	0,121 ± 0,018	3	
<i>5-Amino-6-(1-D-ribitylamino)uracil</i>	Hilic, pos(+)	0,003 ± 0,001	0,015 ± 0,003	0,017 ± 0,007	0,017 ± 0,008	3	
<i>5-Hydroxyconiferaldehyde</i>	Hilic, neg(-)	0,275 ± 0,049	0,197 ± 0,02	0,779 ± 0,191	0,517 ± 0,073	3	
<i>8-[(1R,2R)-3-Oxo-2-[(Z)-pent-2-enyl]cyclopentyl]octanoate</i>	Hilic, neg(-)	0,182 ± 0,009	0,157 ± 0,009	0,068 ± 0,037	0,077 ± 0,023	3	
<i>Betaine</i>	Hilic, pos(+)	0,402 ± 0,097	1,575 ± 0,269	0,562 ± 0,163	0,762 ± 0,224	3	
<i>Cinnamaldehyde</i>	Hilic, pos(+)	0,072 ± 0,003	0,061 ± 0,011	0,031 ± 0,006	0,039 ± 0,009	3	

Tabela 5 - Abundância relativa dos 115 compostos identificados em tecidos de folha +1 de cana-de-açúcar no experimento de campo II. Para cada metabólito identificado foi estimado a sua abundância relativa normalizada contra o padrão interno dos respectivos modos de aquisição (positivo/ negativo) e contra a concentração de proteínas. Os valores indicados à direita de cada intensidade normalizada representa o desvio padrão em relação à média entre as réplicas biológicas. A abreviação C2 indica "campo II"; M4, M8, M11 e M13 indicam as coletas aos 4 meses, 8 meses, 11 meses e 13 meses após o plantio.

(Continua)

Nome do metabólito	Coluna/ modo ionização	C2M4	C2M8	C2M11	C2M13	Nº vias map
<i>cis</i> -Zeatin	Hilic, pos(+)	0,089 ± 0,018	0,087 ± 0,01	0,181 ± 0,019	0,112 ± 0,022	3
Coniferyl aldehyde	Hilic, pos(+)	0,376 ± 0,088	0,487 ± 0,031	0,661 ± 0,023	0,531 ± 0,048	3
Desulfoglucotropeolin	Hilic, neg(-)	0,406 ± 0,153	1,569 ± 0,179	0,469 ± 0,166	0,991 ± 0,21	3
Folate	Hilic, neg(-)	0,465 ± 0,125	0,134 ± 0,047	0,255 ± 0,045	0,137 ± 0,022	3
Methylmalonate	Hilic, pos(+)	0,1 ± 0,018	0,079 ± 0,004	0,071 ± 0,007	0,056 ± 0,009	3
Pelargonidin 3-O-glucoside	Hilic, pos(+)	0,016 ± 0,005	0,012 ± 0,002	0,666 ± 0,08	0,32 ± 0,025	3
Protochlorophyllide	Hilic, neg(-)	0,101 ± 0,012	0,137 ± 0,022	0,153 ± 0,004	0,13 ± 0,013	3
Riboflavin	Hilic, pos(+)	0,12 ± 0,002	0,077 ± 0,007	0,088 ± 0,011	0,072 ± 0,003	3
S-(4-Methylthiobutylthiohydroximoyl)-L-cysteine	Hilic, neg(-)	0,582 ± 0,064	0,502 ± 0,035	0,841 ± 0,141	0,657 ± 0,08	3
Sinapyl alcohol	Hilic, pos(+)	0,202 ± 0,062	0,126 ± 0,03	0,554 ± 0,036	0,442 ± 0,142	3
(-)-Epiarzelechin	Hilic, pos(+)	0,763 ± 0,597	0,092 ± 0,002	0,574 ± 0,366	0,252 ± 0,108	2
(S)-4-Hydroxymandelonitrile	Hilic, pos(+)	0,038 ± 0,006	0,019 ± 0,005	0,113 ± 0,022	0,073 ± 0,014	2
(S)-Allantoin	Hilic, pos(+)	0,023 ± 0,002	0,015 ± 0,002	0,027 ± 0,003	0,027 ± 0,005	2
1,2-Dihydroxy-5-(methylthio)pent-1-en-3-one	C18, pos(+)	11,272 ± 4,964	10,143 ± 1,447	29,378 ± 2,71	21,167 ± 0,297	2
13(1)-Oxo-magnesium-protoporphyrin IX 13-monomethyl ester	Hilic, neg(-)	0,101 ± 0,006	0,124 ± 0,026	0,073 ± 0,021	0,097 ± 0,009	2
2-Maleylacetate	Hilic, pos(+)	0,019 ± 0,002	0,014 ± 0,003	0,024 ± 0,003	0,022 ± 0,005	2
3,4-Dihydroxymandelaldehyde	Hilic, pos(+)	0,05 ± 0,009	0,044 ± 0,005	0,101 ± 0,005	0,077 ± 0,014	2
3-beta-D-Galactosyl-sn-glycerol	Hilic, neg(-)	0,329 ± 0,108	0,363 ± 0,014	0,285 ± 0,105	0,327 ± 0,144	2
4-Aminobenzoate	Hilic, pos(+)	0,912 ± 0,154	1,08 ± 0,071	1,457 ± 0,263	1,253 ± 0,095	2
4-Fumarylacetoacetate	Hilic, pos(+)	0,072 ± 0,004	0,062 ± 0,013	0,032 ± 0,008	0,039 ± 0,01	2

Tabela 5 - Abundância relativa dos 115 compostos identificados em tecidos de folha +1 de cana-de-açúcar no experimento de campo II. Para cada metabólito identificado foi estimado a sua abundância relativa normalizada contra o padrão interno dos respectivos modos de aquisição (positivo/ negativo) e contra a concentração de proteínas. Os valores indicados à direita de cada intensidade normalizada representa o desvio padrão em relação à média entre as réplicas biológicas. A abreviação C2 indica "campo II"; M4, M8, M11 e M13 indicam as coletas aos 4 meses, 8 meses, 11 meses e 13 meses após o plantio.

(Continuação)

Nome do metabólito	Coluna/ modo ionização	C2M4	C2M8	C2M11	C2M13	Nº vias map
<i>5-Hydroxyconiferyl alcohol</i>	C18, pos(+)	1,567 ± 0,399	0,997 ± 0,076	2,176 ± 0,444	2,068 ± 0,116	2
<i>7-Aminomethyl-7-carbaguanine</i>	Hilic, pos(+)	0,44 ± 0,042	0,25 ± 0,014	0,537 ± 0,025	0,35 ± 0,019	2
<i>7-Methylthioheptanaldoxime</i>	Hilic, pos(+)	1,572 ± 0,185	0,889 ± 0,058	1,886 ± 0,082	1,289 ± 0,1	2
<i>Acetone cyanohydrin</i>	C18, pos(+)	0,732 ± 0,237	0,32 ± 0,102	0,521 ± 0,523	0,854 ± 0,145	2
<i>Adenosine</i>	Hilic, pos(+)	0,064 ± 0,068	0,213 ± 0,209	0,23 ± 0,333	1,428 ± 1,311	2
<i>Choline phosphate</i>	C18, pos(+)	0,527 ± 0,122	0,493 ± 0,051	0,572 ± 0,092	0,467 ± 0,142	2
<i>cis-beta-D-Glucosyl-2-hydroxycinnamate</i>	Hilic, pos(+)	0,125 ± 0,022	0,135 ± 0,007	0,076 ± 0	0,14 ± 0,018	2
<i>Cytidine</i>	Hilic, pos(+)	0,089 ± 0,017	0,038 ± 0,013	0,067 ± 0,021	0,054 ± 0,009	2
<i>D-Ribose</i>	Hilic, neg(-)	0,66 ± 0,138	0,352 ± 0,025	0,603 ± 0,162	0,25 ± 0,013	2
<i>Gibberellin A15 open lactone</i>	Hilic, pos(+)	0,073 ± 0,039	0,072 ± 0,006	0,137 ± 0,012	0,112 ± 0,029	2
<i>Guanosine</i>	Hilic, pos(+)	0,147 ± 0,115	0,095 ± 0,02	0,077 ± 0,021	0,136 ± 0,01	2
<i>Hypoxanthine</i>	Hilic, neg(-)	0,907 ± 0,025	0,344 ± 0,02	0,38 ± 0,105	0,387 ± 0,017	2
<i>Indolepyruvate</i>	Hilic, pos(+)	0,148 ± 0,016	0,136 ± 0,034	0,266 ± 0,036	0,23 ± 0,039	2
<i>Kaempferol 3-O-glucoside</i>	C18, pos(+)	4,839 ± 1,435	1,158 ± 0,325	23,052 ± 5,475	4,678 ± 0,703	2
<i>L-Cystine</i>	Hilic, pos(+)	0,404 ± 0,016	0,356 ± 0,06	0,076 ± 0,032	0,13 ± 0,046	2
<i>Leucodelphinidin</i>	Hilic, neg(-)	0,108 ± 0,064	0,159 ± 0,042	0,111 ± 0,03	0,117 ± 0,019	2
<i>Nicotinamide D-ribonucleotide</i>	Hilic, pos(+)	0,082 ± 0,031	0,026 ± 0,004	0,029 ± 0,005	0,046 ± 0,011	2
<i>Piceatannol</i>	Hilic, neg(-)	1,056 ± 0,111	0,745 ± 0,102	0,441 ± 0,127	0,465 ± 0,05	2
<i>Pyridoxine</i>	Hilic, pos(+)	0,062 ± 0,019	0,023 ± 0,006	0,092 ± 0,077	0,084 ± 0,008	2
<i>Pyropheophorbide a</i>	Hilic, pos(+)	1,381 ± 0,125	0,823 ± 0,105	2,002 ± 0,199	1,024 ± 0,106	2

Tabela 5 - Abundância relativa dos 115 compostos identificados em tecidos de folha +1 de cana-de-açúcar no experimento de campo II. Para cada metabólito identificado foi estimado a sua abundância relativa normalizada contra o padrão interno dos respectivos modos de aquisição (positivo/ negativo) e contra a concentração de proteínas. Os valores indicados à direita de cada intensidade normalizada representa o desvio padrão em relação à média entre as réplicas biológicas. A abreviação C2 indica "campo II"; M4, M8, M11 e M13 indicam as coletas aos 4 meses, 8 meses, 11 meses e 13 meses após o plantio.

(Continuação)

Nome do metabólito	Coluna/ modo ionização	C2M4	C2M8	C2M11	C2M13	Nº vias map
<i>trans-Zeatin riboside monophosphate</i>	Hilic, neg(-)	0,345 ± 0,045	0,246 ± 0,011	0,511 ± 0,078	0,29 ± 0,045	2
<i>UDP-6-sulfoquinovose</i>	Hilic, pos(+)	0,006 ± 0,004	0,014 ± 0,003	0,016 ± 0,004	0,024 ± 0,006	2
<i>(9Z)-Hexadecenoic acid</i>	Hilic, neg(-)	0,163 ± 0,051	0,181 ± 0,02	0,036 ± 0,003	0,11 ± 0,06	1
<i>1,3,7-Trimethyluric acid</i>	Hilic, neg(-)	0,924 ± 0,087	0,288 ± 0,019	1,085 ± 0,869	0,325 ± 0,034	1
<i>1-O-Sinapoyl-beta-D-glucose</i>	Hilic, neg(-)	0,026 ± 0,009	0,025 ± 0,004	0,727 ± 0,101	0,425 ± 0,046	1
<i>2'-Deoxyinosine 5'-phosphate</i>	C18, pos(+)	3,098 ± 0,239	1,172 ± 0,435	0,07 ± 0,016	0,181 ± 0,066	1
<i>2-Hydroxy-3-(4-hydroxyphenyl)propenoate</i>	C18, pos(+)	2,383 ± 0,645	2,155 ± 0,475	1,854 ± 0,194	2,289 ± 0,151	1
<i>3-Mercaptolactate</i>	Hilic, neg(-)	0,649 ± 0,138	0,594 ± 0,065	0,836 ± 0,145	0,547 ± 0,112	1
<i>3-Methoxyanthranilate</i>	Hilic, pos(+)	0,228 ± 0,032	0,153 ± 0,012	0,613 ± 0,131	0,394 ± 0,056	1
<i>4-Hydroxy-L-threonine</i>	Hilic, pos(+)	7,526 ± 0,915	5,607 ± 0,154	5,526 ± 0,427	4,04 ± 0,538	1
<i>6beta,7beta-Dihydroxykaurenoic acid</i>	Hilic, pos(+)	0,022 ± 0,005	0,028 ± 0,005	0,018 ± 0,01	0,019 ± 0,007	1
<i>Afzelin</i>	C18, pos(+)	1,623 ± 0,269	1,084 ± 0,277	3,96 ± 0,618	1,805 ± 0,122	1
<i>alpha-D-Galactosyl-(1->3)-1D-myo-inositol</i>	Hilic, neg(-)	12,729 ± 3,013	5,958 ± 0,684	21,791 ± 5,423	11,732 ± 0,92	1
<i>Benzene-1,2,4-triol</i>	C18, pos(+)	5,839 ± 0,68	3,17 ± 0,408	6,866 ± 0,535	4,877 ± 0,466	1
<i>Biliverdin</i>	Hilic, neg(-)	0,307 ± 0,071	0,426 ± 0,109	0,212 ± 0,094	0,356 ± 0,039	1
<i>D-Gal alpha 1->6D-Gal alpha 1->6D-Glucose</i>	C18, pos(+)	2,007 ± 0,407	1,82 ± 0,139	2,198 ± 0,558	1,536 ± 0,108	1
<i>Glutamate</i>	C18, pos(+)	0,54 ± 0,026	0,322 ± 0,09	0,433 ± 0,009	0,289 ± 0,03	1
<i>L-Selenocystathionine</i>	C18, pos(+)	0,195 ± 0,014	0,197 ± 0,005	0,303 ± 0,068	0,134 ± 0,026	1
<i>Molybdate</i>	Hilic, pos(+)	0,046 ± 0,002	0,041 ± 0,004	0,045 ± 0,009	0,042 ± 0,005	1
<i>Myrtenol</i>	Hilic, pos(+)	0,042 ± 0,005	0,029 ± 0,004	0,03 ± 0,002	0,022 ± 0,005	1

Tabela 5 - Abundância relativa dos 115 compostos identificados em tecidos de folha +1 de cana-de-açúcar no experimento de campo II. Para cada metabólito identificado foi estimado a sua abundância relativa normalizada contra o padrão interno dos respectivos modos de aquisição (positivo/ negativo) e contra a concentração de proteínas. Os valores indicados à direita de cada intensidade normalizada representa o desvio padrão em relação à média entre as réplicas biológicas. A abreviação C2 indica "campo II"; M4, M8, M11 e M13 indicam as coletas aos 4 meses, 8 meses, 11 meses e 13 meses após o plantio.

Nome do metabólito	Coluna/ modo ionização					(Conclusão)
		C2M4	C2M8	C2M11	C2M13	Nº vias map
<i>N1,N5,N10-Tricoumaroyl spermidine</i>	Hilic, pos(+)	0,323 ± 0,055	0,412 ± 0,044	0,331 ± 0,016	0,345 ± 0,025	1
<i>Stachyose</i>	Hilic, pos(+)	0,02 ± 0,002	0,017 ± 0,001	0,019 ± 0,006	0,013 ± 0,002	1
<i>Syringin</i>	Hilic, pos(+)	0,189 ± 0,025	0,161 ± 0,007	0,231 ± 0,013	0,187 ± 0,022	1
<i>THF-L-glutamate</i>	Hilic, pos(+)	0,051 ± 0,006	0,052 ± 0,002	0,045 ± 0,005	0,047 ± 0,004	1
<i>Xanthurenic acid</i>	Hilic, pos(+)	0,016 ± 0,002	0,008 ± 0,003	0,02 ± 0,003	0,008 ± 0,004	1
<i>2-(S-Glutathionyl)acetyl glutathione</i>	Hilic, neg(-)	0,049 ± 0,014	0,049 ± 0,014	0,13 ± 0,038	0,082 ± 0,033	0
<i>2-Chloro-5-methyl-cis-dienelactone</i>	C18, pos(+)	0,535 ± 0,079	0,516 ± 0,022	0,648 ± 0,035	0,589 ± 0,028	0
<i>3-Keto-beta-D-galactose</i>	Hilic, neg(-)	0,266 ± 0,013	0,097 ± 0,018	0,194 ± 0,018	0,076 ± 0,007	0
<i>3-Ureidoisobutyrate</i>	Hilic, pos(+)	0,044 ± 0,055	0,083 ± 0,042	0,014 ± 0,003	0,012 ± 0,003	0
<i>5'-Deoxyadenosine</i>	Hilic, pos(+)	0,055 ± 0,015	0,103 ± 0,024	0,092 ± 0,06	0,04 ± 0,007	0
<i>5-Fluorodeoxyuridine</i>	C18, pos(+)	0,774 ± 0,188	0,854 ± 0,028	1,252 ± 0,176	0,9 ± 0,063	0
<i>6-Thiourate</i>	Hilic, neg(-)	2,423 ± 0,395	2,193 ± 0,255	3,064 ± 0,556	2,071 ± 0,38	0
<i>Benzo[a]pyrene-4,5-oxide</i>	Hilic, neg(-)	0,615 ± 0,364	0,392 ± 0,076	0,674 ± 0,079	0,577 ± 0,179	0
<i>cis-4-Carboxymethylenebut-2-en-4-olide</i>	Hilic, neg(-)	0,754 ± 0,069	0,717 ± 0,083	0,878 ± 0,114	0,662 ± 0,097	0
<i>cis-Acetylacrylate</i>	C18, pos(+)	0,742 ± 0,099	0,393 ± 0,05	0,95 ± 0,11	0,637 ± 0,048	0
<i>Glutathione episulfonium ion</i>	Hilic, neg(-)	0,413 ± 0,068	0,218 ± 0,019	0,309 ± 0,027	0,343 ± 0,085	0
<i>S-Adenosyl-4-methylthio-2-oxobutanoate</i>	Hilic, neg(-)	0,252 ± 0,033	0,223 ± 0,024	0,316 ± 0,015	0,225 ± 0,03	0

Para analisar os metabólitos que variaram significativamente ao longo dos experimentos, realizamos o teste de ANOVA e como complemento o teste t (significância, P-valor < 0,05), com o objetivo de avaliar pontualmente as diferenças na abundância relativa dos metabólitos ao longo do tempo. Informações detalhadas dos resultados no teste t podem ser acessados nos materiais suplementares. A análise de variância dos metabólitos ao longo dos experimentos I e II, revelou um total de 48 metabólitos (tabela 6) com diferença significativa entre as médias, no campo I e 77 metabólitos (tabela 7) com diferença significativa entre as médias no campo II.

O cálculo de atividade das vias metabólicas foi realizado usando-se todos os 115 candidatos únicos identificados e informado o valor da quantidade de metabólitos mapeados em cada via metabólica durante o cálculo. Para a discussão dos dados e elaborações de hipóteses a cerca das possíveis interações biológicas dos metabólitos ao longo dos experimentos, consideramos apenas os metabólitos com diferença significativa nos valores de média entre os pontos de coleta, que apresentaram um coeficiente de variação abaixo de 30%. Dessa forma, foi possível trabalhar com uma margem maior de confiança sobre os dados, visto que alguns metabólitos (figura 11) apresentaram grande variação nos experimentos de campo.

Tabela 6 - Metabólitos que apresentaram diferença significativa na abundância relativa ao longo do experimento de campo I. Os valores P e de correção (FDR) se referem ao teste de significância (P-valor < 0.05) para a análise de variância (ANOVA). Para cada metabólito com diferença significativa ao longo do experimento também foram avaliadas as diferenças na abundância relativa nos intervalos descritos (4 meses, 8 meses, 11 meses e 13 meses de idade).

Metabólitos	P-valor	FDR	4m-8m	8m-11m	11m-13m
-Epiarfaelechin	0,03192	0,06463			up
1-Deoxy-D-xylulose 5-phosphate	0,04963	0,08911	up		
1,3,7-Trimethyluric acid	0,0015	0,00745	up	down	
2-C-Methyl-D-erythritol 4-phosphate	0,00567	0,01918	up		
2-Chloro-5-methyl-cis-dienelactone	< 0,005	0,00019	up	down	up
2-Hydroxy-3-4-hydroxyphenylpropenoate	0,01279	0,03501	up		
2-Maleylacetate	< 0,005	0,00019	up		
2-Oxo-8-methylthiooctanoic acid	< 0,005	0,00019	up	up	
2-Oxo-9-methylthiononanoic acid	0,00085	0,00514	down		
2-Oxadipate	0,00105	0,00602		down	
2-Oxoglutarate	0,00018	0,00191	up	down	up
2'-Deoxyinosine 5'-phosphate	< 0,005	0,0009	up		
3-Dehydroquininate	0,01896	0,04739	up		
3-Dehydroshikimate	0,00068	0,00492	up		
3-Keto-beta-D-galactose	0,00701	0,02239	up		
3-Mercaptolactate	0,01055	0,03111		down	
4-Aminobenzoate	0,00078	0,00497		down	
4-Methylthio-2-oxobutanoic acid	< 0,005	0,00029	up	down	up
5-Fluorodeoxyuridine	0,00075	0,00497	down		
5-Hydroxyconiferaldehyde	0,02378	0,0547		down	
5-Hydroxyconiferyl alcohol	0,00844	0,02622		down	up
6beta,7beta-Dihydroxykaurenoic acid	0,00033	0,00294	up		
7-Aminomethyl-7-carbaguanine	0,02099	0,05028			up
7-Methylthioheptanaloxime	< 0,005	0,00019			up
8-1R,2R-3-Oxo-2-Z-pent-2-enylcyclopentyl-octanoate	< 0,005	0,0008	up	down	
alpha-D-Galactosyl-1-3-1D-myo-inositol	0,01144	0,03207	up	down	up
Benzene-1,2,4-triol	0,00126	0,00657	up	down	up
Biliverdin	0,00155	0,00745		down	up
Choline phosphate	0,00313	0,01185		down	up
cis-Acetylacrylate	0,00319	0,01185	up	down	up
Citrate	0,0003	0,00288	down	down	
D-Ribose	0,00499	0,01778	up	down	
D-Sorbitol	0,01986	0,04858		down	

Tabela 6 - Metabólitos que apresentaram diferença significativa na abundância relativa ao longo do experimento de campo I. Os valores P e de correção (FDR) se referem ao teste de significância (P-valor < 0.05) para a análise de variância (ANOVA). Para cada metabólito com diferença significativa ao longo do experimento também foram avaliadas as diferenças na abundância relativa nos intervalos descritos (4 meses, 8 meses, 11 meses e 13 meses de idade).

(Conclusão)

Metabólitos	P-valor	FDR	4m-8m	8m-11m	11m-13m
Folate	< 0,005	0,0008	up	down	
Glutamate	0,04223	0,07708		down	
Hypoxanthine	0,00219	0,00944	up	down	
Indolylmethyl-desulfoglucosinolate	0,01054	0,03111		down	
L-Selenocystathionine	0,02833	0,05924	up		
L-Tyrosine	< 0,005	0,0009	up	down	up
Myrtenol	0,00184	0,00845	down		
Nicotinamide D-ribonucleotide	0,0023	0,00944		down	
Nicotinate	0,00043	0,00327	up		
Piceatannol	0,00224	0,00944		down	
Protochlorophyllide	0,0051	0,01778	down		
Pyropheophorbide a	0,0162	0,04234	up		
S-4-Methylthiobutylthiohydroximoyl-L-cysteine	0,00121	0,00657	up	down	
S-Adenosyl-4-methylthio-2-oxobutanoate	0,04027	0,07593	up		
S-Allantoin	< 0,005	0,0008	up		

Tabela 7 - Metabólitos que apresentaram diferença significativa na abundância relativa ao longo do experimento de campo II. Os valores P e de correção (FDR) se referem ao teste de significância (P-valor < 0.05) para a análise de variância (ANOVA). Para cada metabólito com diferença significativa ao longo do experimento também foram avaliadas as diferenças na abundância relativa nos intervalos descritos (4 meses, 8 meses, 11 meses e 13 meses de idade).

(Continua)

Metabólitos	p.value	FDR	m4m8	m8m11	m11m13
1-Deoxy-D-xylulose 5-phosphate	0,02871	0,04288	up		
1-O-Sinapoyl-beta-D-glucose	< 0,005	2,9E-05		down	up
1,2-Dihydroxy-5-methylthiopent-1-en-3-one	0,00012	0,00067		down	up
2-C-Methyl-D-erythritol 4-phosphate	< 0,005	0,00026	up		up
2-Chloro-5-methyl-cis-dienelactone	0,03198	0,04655		down	
2-Maleylacetate	0,03913	0,05488		down	
2-Methylmaleate	< 0,005	0,00018		up	
2-Oxo-6-methylthiohexanoic acid	0,02805	0,04244		down	
2-Oxo-8-methylthiooctanoic acid	< 0,005	0,00018		up	
2-Oxadipate	0,01675	0,02752	down		
2-Oxoglutarate	0,00244	0,00562	up		
2-S-Glutathionylacetyl glutathione	0,01859	0,02969		down	

Tabela 7 - Metabólitos que apresentaram diferença significativa na abundância relativa ao longo do experimento de campo II. Os valores P e de correção (FDR) se referem ao teste de significância (P-valor < 0.05) para a análise de variância (ANOVA). Para cada metabólito com diferença significativa ao longo do experimento também foram avaliadas as diferenças na abundância relativa nos intervalos descritos (4 meses, 8 meses, 11 meses e 13 meses de idade).

Metabólitos	(Continuação)				
	p.value	FDR	m4m8	m8m11	m11m13
<i>2'-Deoxyinosine 5'-phosphate</i>	< 0,005	< 0,005	up	up	down
<i>3-Dehydroshikimate</i>	0,00101	0,00276		up	
<i>3-Keto-beta-D-galactose</i>	< 0,005	< 0,005	up	down	up
<i>3-Methoxyanthranilate</i>	0,00026	0,00101	up	down	
<i>3,4-Dihydroxymandelaldehyde</i>	0,00022	0,00098		down	
<i>4-Fumarylacetoacetate</i>	0,00203	0,00489		up	
<i>4-Hydroxy-L-threonine</i>	0,0006	0,00182	up		up
<i>4-Methylthio-2-oxobutanoic acid</i>	< 0,005	0,00031			down
<i>5-Amino-6-1-D-ribitylamouracil</i>	0,03835	0,05445	down		
<i>5-Fluorodeoxyuridine</i>	0,01024	0,01784		down	up
<i>5-Hydroxyconiferaldehyde</i>	0,00058	0,00182		down	
<i>5-Hydroxyconiferyl alcohol</i>	0,00548	0,01126		down	
<i>7-Aminomethyl-7-carbaguanine</i>	< 0,005	0,00014	up	down	up
<i>7-Methylthioheptanaloxime</i>	< 0,005	0,00031	up	down	up
<i>8-1R,2R-3-Oxo-2-Z-pent-2-enylcyclopentylactanoate</i>	0,00054	0,00182	up	up	
<i>9Z-Hexadecenoic acid</i>	0,0099	0,01771		up	
<i>Acetyl phosphate</i>	0,00059	0,00182	up		up
<i>Afzelin</i>	< 0,005	0,0004		down	up
<i>alpha-D-Galactosyl-1-3-1D-myo-inositol</i>	0,00196	0,00489		down	up
<i>Aminoimidazole ribotide</i>	< 0,005	0,00042		up	down
<i>Ascorbate</i>	< 0,005	0,00019		down	up
<i>Benzene-1,2,4-triol</i>	0,00017	0,00083	up	down	up
<i>Benzoate</i>	< 0,005	0,00043	down	up	
<i>Betaine</i>	0,00041	0,00151	down	up	
<i>Chlorogenate</i>	0,00023	0,00098		down	up
<i>Cinnamaldehyde</i>	0,00071	0,0021		up	
<i>cis-Acetylacrylate</i>	0,00024	0,00098		down	up
<i>cis-Aconitate</i>	0,03071	0,04528		down	
<i>cis-beta-D-Glucosyl-2-hydroxycinnamate</i>	0,00234	0,00548		up	down
<i>cis-Zeatin</i>	0,00057	0,00182		down	up
<i>Citrate</i>	0,02345	0,03595		down	
<i>Coniferyl aldehyde</i>	0,00142	0,00378		down	up

Tabela 7 - Metabólitos que apresentaram diferença significativa na abundância relativa ao longo do experimento de campo II. Os valores P e de correção (FDR) se referem ao teste de significância (P-valor < 0.05) para a análise de variância (ANOVA). Para cada metabólito com diferença significativa ao longo do experimento também foram avaliadas as diferenças na abundância relativa nos intervalos descritos (4 meses, 8 meses, 11 meses e 13 meses de idade).

Metabólitos	(Conclusão)				
	p.value	FDR	m4m8	m8m11	m11m13
<i>D-Ribose</i>	0,00431	0,00901	up		up
<i>D-Sorbitol</i>	0,00297	0,0066	up		
<i>Desulfoglucotropeolin</i>	0,00014	0,00074	down	up	down
<i>Folate</i>	0,00145	0,00378	up	down	up
<i>Gibberellin A15 open lactone</i>	0,03338	0,04798		down	
<i>Glutamate</i>	0,00096	0,00276	up		up
<i>Glutathione episulfonium ion</i>	0,01946	0,03065	up	down	
<i>Hypoxanthine</i>	< 0,005	< 0,005	up		
<i>Indolepyruvate</i>	0,00298	0,0066		down	
<i>Indolylmethyl-desulfoglucosinolate</i>	0,00025	0,00098	up	down	
<i>Kaempferol 3-O-glucoside</i>	< 0,005	0,00039	up	down	up
<i>L-Cystine</i>	< 0,005	0,00026		up	
<i>L-Pipecolate</i>	0,0006	0,00182	down	up	
<i>L-Selenocystathionine</i>	0,00357	0,00775			up
<i>L-Tyrosine</i>	0,00172	0,00438		down	up
<i>Myrtenol</i>	0,00394	0,00839	up		
<i>Nicotinamide D-ribonucleotide</i>	0,01071	0,01838	up		
<i>Pelargonidin 3-O-glucoside</i>	< 0,005	< 0,005		down	up
<i>Piceatannol</i>	0,00024	0,00098		up	
<i>Protochlorophyllide</i>	0,0118	0,01996			up
<i>Pyropheophorbide a</i>	< 0,005	0,00031	up	down	up
<i>Riboflavin</i>	< 0,005	0,0005	up		
<i>S-2-Acetolactate</i>	0,0077	0,01528	up		
<i>S-4-Hydroxymandelonitrile</i>	0,00015	0,00077		down	
<i>S-4-Methylthiobutylthiohydroximoyl-L-cysteine</i>	0,00885	0,01669		down	
<i>S-Adenosyl-4-methylthio-2-oxobutanoate</i>	0,00804	0,01546		down	up
<i>S-Allantoin</i>	0,00602	0,01214	up	down	
<i>Sinapyl alcohol</i>	0,0006	0,00182		down	
<i>Syringin</i>	0,00929	0,01724		down	up
<i>trans-Zeatin riboside monophosphate</i>	0,00101	0,00276		down	up
<i>Xanthurenic acid</i>	0,00204	0,00489		up	up
<i>Zeatin</i>	0,00986	0,01771		down	

4.3.2. Perfil da atividade das vias ao longo do experimento de campo I

Após os escores serem calculados para todas as vias mapeadas (tabela 3 do material suplementar), foi aplicada uma análise de clusterização hierárquica sobre os escores fixando-se a idade da planta em uma série temporal de 4 meses, 8 meses, 11 meses e 13 meses respectivamente, a fim de se observar o padrão de comportamento das vias ao longo do processo de maturação da cana (figura 13).

Grande parte das vias demonstrou alternância na atividade ao longo do tempo, onde parte delas aumentou ou diminuiu entre 4 e 8 meses, alternando novamente entre 8 e 11 meses e entre 11 e 13 meses (grupos A e D). Por outro lado, algumas vias apresentaram atividade elevada em 4 e 8 meses, e baixa atividade em 11 e 13 meses (grupo B). De forma similar, porém contrário ao B, o grupo C apresentou baixa atividade em 4 e 8 meses e alta atividade em 11 e 13 meses. Por fim, as vias identificadas em E apresentaram atividade elevada em 4, 8 e 11 meses e baixa atividade em 13 meses.

A clusterização hierárquica dos escores permitiu avaliar o comportamento geral das vias metabólicas ao longo do crescimento da planta, entretanto, não nos fornece quais vias podem estar significativamente alteradas no decorrer deste período. Para analisar tais diferenças ao longo do desenvolvimento da cana-de-açúcar, foi realizado um total de três comparações (teste t, P-valor < 0,05): 4 meses com 8 meses, 8 meses com 11 meses e 11 meses com 13 meses.

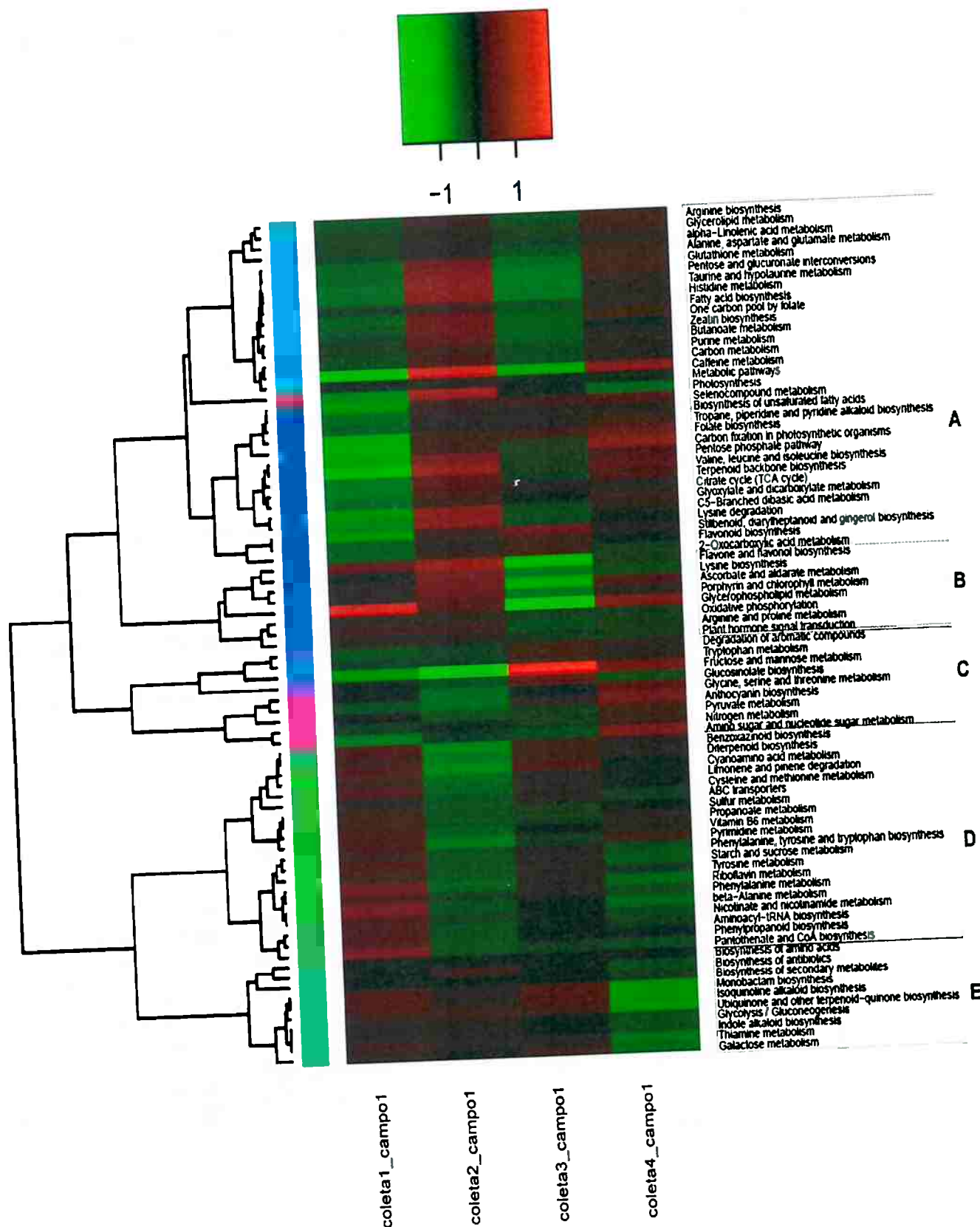


Figura 13 - Clusterização hierárquica dos escores obtidos para as vias mapeadas no experimento de campo. As colunas estão dispostas em ordem crescente para a idade da planta no momento da coleta, onde coleta 1, coleta 2, coleta 3 coleta 4 referem-se respectivamente às idades de 4,8, 11 e 13 meses da cana-de-açúcar. Nas linhas, as vias metabólicas foram agrupadas (A, B, C e D) conforme a similaridade entre os seus escores ao longo do experimento de campo I.

A partir da lista das vias metabólicas com diferença significativa no escore de atividade de via em cada uma das comparações (tabelas de comparação por teste t do material suplementar), as vias metabólicas foram agrupadas conforme o processo metabólico ao qual as mesmas estão relacionadas, possibilitando assim observar o comportamento global das vias que apresentaram alteração significativa na atividade ao longo do tempo no campo I. Conforme a figura 14, podemos observar que entre 4 meses e 8 meses grande parte das vias alteradas estão relacionadas ao metabolismo de carboidrato, aminoácido e de biossíntese de metabólitos secundários. As vias relacionadas ao metabolismo de carboidrato e de biossíntese de metabólitos secundários também apresentaram alteração considerável entre 8 meses e 11 meses. Bem como entre 11 meses e 13 meses, mostrando assim a importância das vias associadas no processo de desenvolvimento na cana-de-açúcar no experimento de campo I.

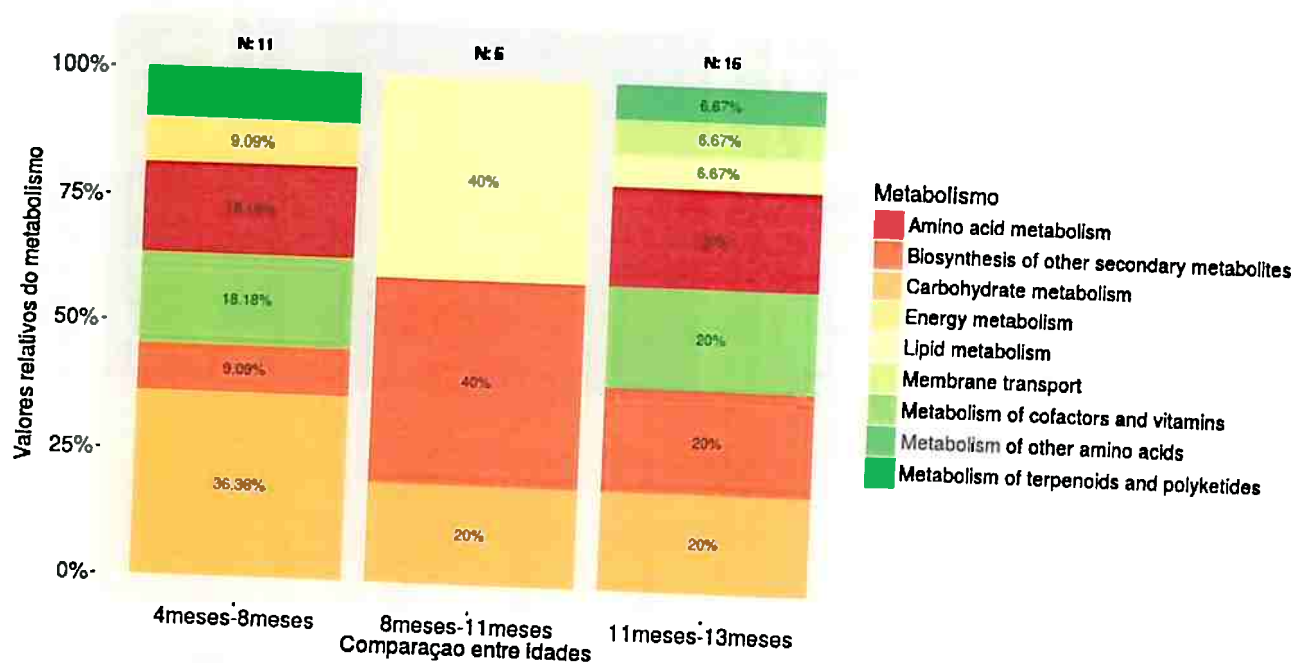


Figura 14 - Metabolismo associado às vias com diferença significativa na atividade nas comparações entre amostras de diferentes pontos de coleta do experimento de campo I. Cada coluna representa a quantidade de vias (N) com diferença significativa na atividade para a comparação entre amostras em idades diferentes (4 meses e 8 meses, 8 meses e 11 meses e

11 meses e 13 meses). Os valores relativos dentro da coluna representam a proporção do metabolismo em que as vias significativas estão associadas.

Com base nas observações realizadas e nas evidências biológicas encontradas no experimento de campo I, selecionamos um conjunto de vias metabólicas capazes de explicar parte de alguns dos importantes fenômenos observados que podem estar associados ao processo de resposta ao estresse hídrico e ao crescimento da planta: a via de biossíntese de fenilpropanóide (map00940), metabolismo de amido e sacarose (map00500) e via do ácido cítrico (map00020). Os detalhes de cada uma das vias citadas estão disponíveis na sessão 9.2 dos materiais anexos (figuras anexas 1, 2 e 3). Alguns dos metabólitos e genes identificados também foram analisados separadamente do seu contexto de vias metabólicas.

4.3.3 Perfil da atividade das vias ao longo do experimento de campo II

A análise de clusterização hierárquica também foi aplicada sobre os scores do experimento de campo II, fixando-se a idade da planta em uma série temporal de 4 meses, 8 meses, 11 meses e 13 meses respectivamente, a fim de se observar o padrão de comportamento das vias ao longo do processo de maturação da cana (figura 11). A clusterização das vias evidenciou 8 grupos com padrões diferentes ao longo do experimento. Os grupos descritos em A e E apresentaram aumento de atividade entre os intervalos de 4 a 8 meses. Neste mesmo período, o grupo das vias em C e H apresentou diminuição de atividade, enquanto que nos demais grupos não houve mudança no padrão de atividade, se mantendo baixo (F, G e I) ou elevado (B e D).

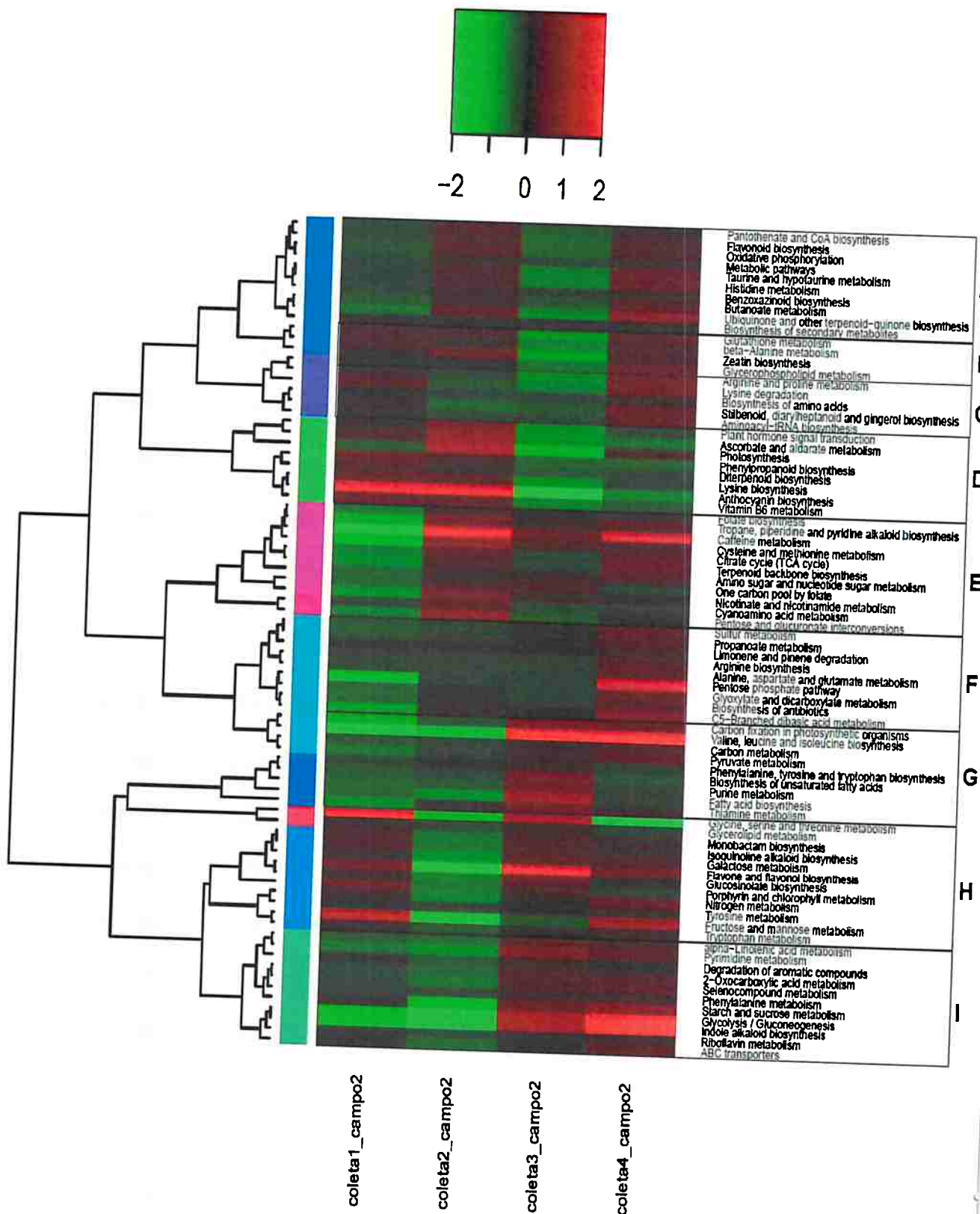


Figura 15 - Clusterização hierárquica dos escores obtidos para as vias mapeadas no experimento de campo II. As colunas estão dispostas em ordem crescente para a idade da planta no momento da coleta, onde coleta 1, coleta 2, coleta 3 e coleta 4 referem-se respectivamente às idades de 4,8, 11 e 13 meses da cana-de-açúcar. Nas linhas, as vias metabólicas foram agrupadas (A, B, C, D, E, F, G, H e I) conforme a similaridade entre os seus escores ao longo do experimento de campo II.

Entre os intervalos de 8 a 11 meses, houve diminuição na atividade nos grupos descritos em A, B e D, enquanto que em G, H e I houve aumento da atividade no decorrer deste período. Em C e F, a atividade das vias permaneceu baixa, e em E alta. Por fim, no intervalo entre 11 e 13 meses, os grupos E, G, H e I mantiveram um nível elevado de atividade. Os grupos A, B, C e F apresentaram aumento na atividade, ao passo que em D as vias mantiveram-se em baixa atividade.

De forma similar à análise realizada com os dados do experimento de campo I, as vias metabólicas também foram agrupadas conforme o metabolismo em que a mesmas estão relacionadas, possibilitando assim observar o comportamento global das vias que apresentaram alteração significativa na atividade ao longo do tempo (figura 16). O experimento de campo II apresentou maior quantidade de vias com diferença significativa na atividade ao longo do crescimento. Foi possível observar que as vias metabólicas relacionadas ao metabolismo de carboidratos e de biossíntese de metabólitos secundários foram as que apresentaram maior diferença ao longo do crescimento da cana-de-açúcar.

No experimento de campo II selecionamos para a discussão as vias de metabolismo de amido e sacarose (map00500) e biossíntese de fenilpropanóides (map00940). Parte das vias selecionadas, assim como no experimento de campo I, também formam capazes de explicar alguns dos fenômenos biológicos observados, e estão relacionados ao processo de resposta ao déficit hídrico e crescimento da cana-de-açúcar. Detalhes das vias podem ser acessados na sessão 9.3 dos materiais anexos (figuras anexas 4, 5 e 6).

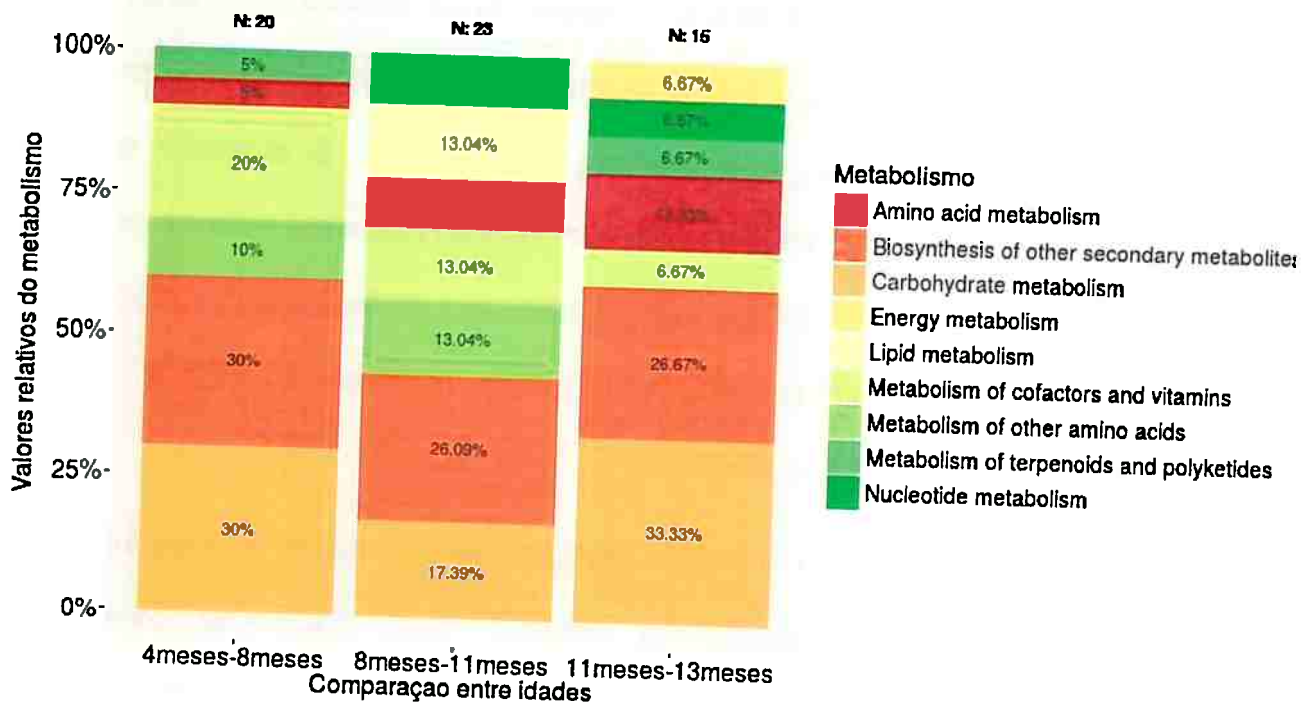


Figura 16 - Metabolismo associado às vias com diferença significativa na atividade nas comparações entre amostras de diferentes pontos de coleta do experimento de campo I. Cada coluna representa a quantidade de vias (N) com diferença significativa na atividade para a comparação entre amostras em idades diferentes (4 meses e 8 meses, 8 meses e 11 meses e 11 meses e 13 meses). Os valores relativos dentro da coluna representam a proporção do metabolismo em que as vias significativas estão associadas.

5 DISCUSSÃO

5.1 RELAÇÃO ENTRE OS FATORES ABIÓTICOS E O DESENVOLVIMENTO DA CANA-DE-AÇÚCAR

5.1.1 NO EXPERIMENTO DE CAMPO I

O experimento de campo I possui a característica de plantio de cana ano e meio, ou seja, o plantio foi realizado em um período de baixa temperatura e precipitação (com indicativo para déficit hídrico, conforme a figura 4), o que conferiu uma taxa de crescimento inicial baixa. Entre 4 meses e 8 meses, houve aumento na temperatura média e precipitação, alcançando o máximo valor em 8 meses. Apenas a partir deste ponto a cana-de-açúcar deu início ao processo de alongamento tardio (figura 8). Tais observações deixam evidente

que o estresse hídrico afetou significativamente o desenvolvimento da cana-de-açúcar.

O estresse abiótico como o déficit hídrico, salinidade e baixa temperatura é conhecido por causar alterações no metabolismo de carboidratos e aminoácidos em plantas. O déficit hídrico em particular é conhecido por causar aumento da concentração de açúcares solúveis e aminoácidos livres na folha (RAMOS; PARSONS; SPRENT, 2005; ROSA et al., 2009). Os açúcares solúveis são importantes mensageiros primários que controlam o crescimento e metabolismo em plantas. Logo, condições ambientais desfavoráveis podem causar desequilíbrio na relação fonte-dreno da planta, tendo como consequência a superexpressão de diferentes enzimas relacionadas à biossíntese de carboidratos (ROSA et al., 2009).

Em estudos anteriores, alguns açúcares solúveis como a rafinose e trealose foram identificados em maior concentração em folha de plantas sob déficit hídrico (KRASENSKY; JONAK, 2012). A sacarose também é conhecida por aumentar a sua concentração em folhas, propiciando maior tolerância à seca (FU; HUANG; FRY, 2010). Em nosso estudo, o metabólito trealose e a sacarose (candidatos ambíguos que apresentam o mesmo valor de massa m/z) apresentaram uma abundância relativa 2 vezes maior aos 4 meses (tabela 4).

Identificamos também o monossacarídeo D-xilose seu ambíguo L-arabinose com uma abundância relativa aproximadamente 2 vezes maior aos 4 meses (tabela 4). Estes monossacarídeos não importantes precursores na síntese da parede celular. Estudos indicam que os monossacarídeos precursores da parede celular possuem diferentes funções na manutenção da homeostasia da planta em condições de estresse abiótico, não apenas servido

como componentes estruturais (BALSAMO et al., 2015; YAMADA et al., 2010). O acúmulo destes monossacarídeos na folha é conhecido por conferir maior tolerância à seca, assumindo um papel osmoprotetor (BALSAMO et al., 2015). Na via de metabolismo de amido e sacarose (figura anexa 2) também observamos elevada expressão das enzimas trealose 6-fosfato sintase (ec: 2.4.1.15, sas: SCCCL3002E04.b) e amido fosforilase (ec: 2.4.1.1, sas: SCUTLR2030E06.g) , cerca de 2 vezes e 3 vezes mais espessas aos 4 meses respectivamente (tabela 1 do material suplementar).

A trealose 6-fosfato sintase é uma enzima responsável por catalisar a síntese de trealose 6-fosfato, utilizando os precursores UDP-glicose e glicose 6-fosfato. O composto trealose 6-fosfato possui importante papel na regulação de vias do metabolismo de carboidratos, uma vez que o mesmo atua como um provável sinalizador em resposta às variações ambientais. Em *Arabidopsis t.* a superexpressão do gene que codifica para a enzima TPS é observada quando a planta é submetida ao estresse hídrico. Uma resposta similar foi observada em folhas de batata da variedade *Solanum tuberosum* L. (KONDRÁK et al., 2012).

Já a enzima amido fosforilase é responsável por catalisar a reação de quebra do amido. A expressão do gene codificante para esta enzima está associada ao processo de resposta à seca. Em linhagens de *Lolium multiflorum* tolerantes à seca, foi observada alta expressão do gene que codifica para a enzima amido fosforilase quando submetidas ao déficit hídrico, ao passo que em linhagens sensíveis à seca não foi observado o mesmo comportamento (PAN et al., 2016).

Além dos açúcares, alguns aminoácidos também são conhecidos como responsivos ao estresse hídrico (RAMOS; PARSONS; SPRENT, 2005). Em nosso experimento, identificamos a L-valina 3 vezes mais abundante aos 4 meses (tabela 4). Embora este metabólito tenha apresentado um alto coeficiente de variação aos 8 meses (baixa qualidade de detecção), podemos inferir que o aumento da abundância de L-valina na folha pode estar associado ao período de restrição hídrica e baixa temperatura média aos 4 meses. Em folhas de trigo submetido a regime de estresse hídrico foi observado um aumento significativo na concentração de aminoácidos, dentre eles a valina (BOWNE et al., 2012). Logo, uma resposta similar poderia estar ocorrendo na cana-de-açúcar entre este período.

Elevados níveis de aminoácidos na folha podem levar ao dano tecidual, dessa forma, a planta utiliza de mecanismos específicos para preveni-lo. Um destes mecanismos envolve a superexpressão do gene codificante para a enzima aminotransferase de aminoácidos de cadeia ramificada (HvBCAT, ec: 2.6.1.42, SAS: SCCCLR1001C07.g). Esta enzima tem a função de degradar o excesso de aminoácidos livres na folha, evitando assim o dano tecidual (MALATRASI et al., 2006). Em nosso experimento, foi possível identificar a expressão deste gene 2 vezes maior aos 4 meses (tabela 4), evidenciando assim uma possível resposta na folha da cana-de-açúcar ao acúmulo de aminoácidos.

Outro metabólito, a betaína, foi identificado 3 vezes mais abundante aos 4 meses (entretanto, com baixa qualidade de detecção aos 8 meses). Ela é conhecida por atuar como um agente osmoprotetor em condições diversas de estresse abiótico. Sua alta concentração em folhas de trigo é conhecida por

conferir tolerância às baixas temperaturas (ALLARD et al., 1998). Em sorgo, a betaina confere resistência ao estresse salino (GRIEVE; MAAS, 1984), ao passo que em tabaco foi observado que o acúmulo deste aminoácido propiciou tanto a resistência às baixas temperaturas como também ao estresse salino (HOLMSTRÖM et al., 2000).

Além das alterações no metabolismo de carboidratos e aminoácidos, também realizamos importantes observações no metabolismo secundário. A via de biossíntese de fenilpropanóide (figura anexa 1) não apresentou diferença significativa no escore de atividade estimado para o período inicial do plantio (intervalo entre 4 meses e 8 meses), entretanto observamos que alguns dos metabólitos participantes desta via possam estar desempenhando um importante papel no processo de desenvolvimento da cana-de-açúcar.

A coniferina, um importante monolignol precursor da lignina, apresentou uma abundância relativa aproximadamente 2 vezes maior aos 4 meses no campo I (tabela 4). Este metabólito exerce uma gama de funções primordiais na sobrevivência da planta, mantendo a integridade estrutural, além de permitir o transporte de água e solutos através dos tubos do sistema vascular (LIU; MIAO; ZHANG, 2011). Assim como outras classes de monolignóis, a coniferina é produzida no citosol da célula, sendo necessário transportá-la até a parede celular para que a mesma seja polimerizada na síntese de lignina. Estudos indicam que o transporte da coniferina é conduzido de forma ativa por canais transportadores específicos, onde os mesmos são fortemente regulados pela disponibilidade de ATP no meio (MIAO; LIU, 2010). Pode-se postular que a alta abundância de coniferina na folha aos 4 meses possa estar associada à baixa produção de lignina.

Esta hipótese é reforçada pela maior abundância do metabólito 2-oxoglutarato aos 4 meses em relação a 8 meses, aproximadamente 2 vezes maior (tabela 5). Este metabólito participa de um vasto conjunto de vias, tais como no ciclo do ácido cítrico (ver figura anexa 6), biossíntese de fenilpropanóides, biossíntese de hormônios, metabolismo de aminoácidos e consequentemente na assimilação do nitrogênio, pela ação enzimática da glutamato sintase (ARAÚJO et al., 2014).

Ao focarmos especificamente para a via do ácido cítrico (figura anexa 3), é sugestivo que a sua maior abundância aos 4 meses em relação a 8 meses possa estar associada à uma baixa atividade enzimática da enzima 2-oxoglutarato desidrogenase (ec: 1.2.4.2), apesar do gene que codifica para esta proteína não ter sido detectado no estudo (necessária posterior validação). Esta enzima é responsável por catalisar a reação reversível do 2-oxoglutarato juntamente da tiamina difosfato para formar succinato semialdeído-tiamina difosfato e uma molécula de gás carbônico, sendo crucial na continuidade do processo respiratório da planta para a formação de derivados energéticos como o ATP (ARAÚJO et al., 2008).

No experimento conduzido por Araújo e colaboradores (ARAÚJO et al., 2008) em batata, a inibição da enzima 2-oxoglutarato desidrogenase acarretou em diversas alterações como: expressiva diminuição da respiração da planta, bem como do fluxo para a parede celular; e aumento da concentração de 2-oxoglutarato e alanina na folha. Dessa forma, com uma possível diminuição de ATP disponível no meio, o transporte de monolignóis poderá ser comprometido, levando ao seu acúmulo na folha.

Tais observações parecem corroborar com as observações realizadas aos 4 meses, com a ocorrência de déficit hídrico e consequente inibição do crescimento. As condições desfavoráveis de crescimento neste período ocasionaram o acúmulo de 2-oxoglutarato. Dessa forma, com a produção energética prejudicada, o transporte de monolignóis para a síntese de lignina na parede celular também passaram a ser prejudicada.

Entre os intervalos de 8 meses até 11 meses, foi observado diminuição nos valores de temperatura e precipitação no campo I. Neste período a concentração dos compostos D-sorbitol e sacarose se apresentou aproximadamente 2 vezes maior em comparação aos 8 meses. Lembrando que para a sacarose foram identificados 3 candidatos ambíguos, a maltose, alfa-trealose e celobiose.

O sorbitol e a sacarose são os principais produtos fotossintéticos produzidos pelas plantas para suprir a demanda de esqueleto carbônico e energia no processo de fonte-dreno. O acúmulo de sorbitol na folha é conhecido por conferir maior tolerância à estresses abióticos, além de manter o turgor das células em plantas (DEGUCHI et al., 2006). Entretanto, em experimentos conduzidos com cana-de-açúcar, o acúmulo excessivo deste metabólito nas folhas promove inibição do crescimento (CHONG et al., 2007) e o acúmulo em folhas de tabaco causou dano tecidual (SHEVELEVA et al., 1998).

Aos 11 meses, a cana-de-açúcar ainda se encontrava em crescimento. Dessa forma, a grande abundância do metabólito D-sorbitol neste período contraria os dados de estudos anteriores. Em relação à sacarose, a maior abundância na folha pode ser não em decorrência de uma resposta ambiental,

mas sim pelo intenso crescimento que a planta ainda apresentava neste período.

Entre os intervalos de 11 meses e 13 meses a cana-de-açúcar ainda apresentou crescimento vertical (figura 8) e aumento do grau Brix de forma significativa (figura 9). Este período foi marcado por uma gradativa diminuição da temperatura média, chegando próximo aos 18°C. Também se observou um baixo índice de precipitação (figura 3). A cana-de-açúcar se encontrava em pleno estágio de maturação com aumento do grau Brix e crescimento esperado para além dos 13 meses, uma vez que o plantio do campo I possui a característica de cana ano-e-meio.

Não foi possível identificar um padrão metabólico distinto para este período, capaz de explicar as observações morfológicas e fisiológicas da cana-de-açúcar.

5.1.2 NO EXPERIMENTO DE CAMPO II

O experimento de campo II foi planejado para recriar um plantio de cana ano. Portanto, o plantio foi realizado em um período no qual o brotamento e desenvolvimento inicial da plantas ocorreria em um período de elevada precipitação e temperatura (figura 3). Entre o plantio e a coleta aos 4 meses foram observados os valores máximos de temperatura e precipitação (figura 3). Logo, diferente do campo I, aos 4 meses no campo II a cana-de-açúcar já se encontrava em fase de alongamento com intenso crescimento vertical, entretanto, com um diâmetro do colmo menor em comparação ao campo I (figura 8).

Entre os intervalos da primeira coleta com a segunda coleta (4 meses a 8 meses) houve diminuição na temperatura média e precipitação, chegando

próximo ao valor estimado para déficit hídrico (figura 4). Neste período a cana-de-açúcar alcança seu crescimento vertical máximo, e o acúmulo de sacarose no colmo estimado pelo grau Brix também é alto (figura 9). Tais observações sugerem que a cana encontra-se em início de processo de maturação, diferente do experimento de campo I, onde o processo de maturação iniciou apenas aos 11 meses de idade.

Observamos, uma abundância relativa de sacarose 2 vezes maior aos 4 meses (figura anexa 5). Entretanto, é sugestivo que a sua grande abundância aos 4 meses no experimento de campo II possa ser em decorrência da alta atividade enzimática da sacarose fosfato sintase (EC: 2.4.1.14), diferente do campo I aos 4 meses, onde o acúmulo na folha foi em decorrência de uma resposta ao estresse hídrico. O gene codificante para esta enzima não foi identificado no experimento de transcriptômica, logo será necessária futura validação para a análise desta hipótese.

A atividade da SFS possui forte correlação com o acúmulo de sacarose na folha durante o crescimento da planta (GALTIER et al., 1993; KOMOR, 2000). Além disso, sua atividade é fortemente modulada pela disponibilidade de luz e pela concentração de CO₂ no meio (GALTIER et al., 1993). A primeira coleta do experimento de campo II (4 meses de idade), que se situa no intervalo entre os meses de Dezembro e Março, corresponde ao verão. Neste período há maior incidência da luz solar, uma vez que os dias passam a ser mais longos, além do aumento nos valores de temperatura média e precipitação (figura 3). Conforme o esperado, a cana-de-açúcar apresentou intenso crescimento, alcançando altura máxima aos 8 meses. Tal observação corrobora com a hipótese de que a sacarose apresenta maior concentração

aos 4 meses em decorrência de uma elevada produção para suprir a demanda energética do crescimento da planta.

Aos 8 meses observamos um perfil similar ao período de 4 meses do campo I, que teve aumento na abundância dos compostos L-valina e betaína em resposta ao déficit hídrico (3 vezes mais abundante em 8 meses). Neste período, os valores de precipitação e temperatura média diminuíram até os valores próximos de déficit hídrico (figura 4). O que sugere uma resposta similar ao discutido anteriormente nos dados do experimento de campo I.

Entre 8 e 11 meses, foi observado o menor valor médio para a precipitação e temperatura no experimento de campo II; com a contínua diminuição da precipitação, o plantio passou a estar sob regime de déficit hídrico aos 11 meses (figura 4). Neste período a cana-de-açúcar alcançou seu máximo teor de Brix (figura 9) e a sua altura permaneceu constante (figura 8). A cana-de-açúcar alcançou sua maturação plena, uma vez que foi observado o início do florescimento aos 13 meses (dados não disponíveis).

Aos 11 meses o metabólito afzelina, um glicosídeo flavonóide, demonstrou uma abundância 3 vezes maior em comparação aos 8 meses (tabela 5). Em *Arabidopsis thaliana*, o silenciamento do gene HCT (do inglês, hydroxycinnamoyl-CoA shikimate/quininate hydroxycinnamoyl transferase), responsável pela biossíntese de unidades guaiacila e siringila da lignina, ocasiona elevado acúmulo de flavonóides e antocianina nas folhas. O acúmulo destes glicosídeos possui a capacidade de inibir o transporte de auxinas e tem como consequência secundária o retardo no crescimento da planta. A inibição do gene HCT também causa uma profunda alteração no metabolismo de fenilpropanóides, por inibir a biossíntese de lignina (BESSEAU et al., 2007).

O gene codificante para a enzima HCT não foi detectada na análise do transcriptoma, logo será necessária futura validação. Entretanto, é possível que o acúmulo da afzelina nas folhas aos 11 meses possa estar associado à inibição do crescimento da planta já que nesse período a planta já alcançara o máximo de crescimento vertical (não houve diferença significativa com a altura aos 13 meses, conforme a figura 8-a).

Na via de biossíntese de fenilpropanóides (figura anexa 4) identificamos uma grande variação em seus metabólitos detectados: o metabólito álcool sinapílico (3,75 vezes mais abundante em 11 meses), *Caffeyl alcohol*, cis-beta-D-Glucosyl-2-hydroxycinnamate e cinamaldeído (2 vezes mais abundantes em 8 meses), 5-Hidroxi coniferaldeído e seu ambíguo ferulato e coniferina (3,8 vezes mais abundantes em 11 meses), 1-O-Sinapoyl-beta-D-glucose (34 vezes mais abundante em 11 meses).

A coniferina e o álcool sinapílico são compostos precursores da lignina, sendo o primeiro um composto glicosilado, ao passo que o segundo apresenta-se em sua forma não glicosilada. A coniferina e o álcool sinapílico são produzidas no citosol da célula e transportadas até a parede celular para serem polimerizadas na síntese de lignina (MIAO; LIU, 2010). Além dos precursores de lignina também foi observada maior abundância de 2-oxoglutarato aos 11 meses, cerca de 2 vezes maior em relação aos 8 meses. Curiosamente, uma resposta similar de inibição do crescimento também foi observada no experimento de campo I aos 4 meses (conforme discutido anteriormente). Em ambos os casos o plantio encontra-se sob regime de déficit hídrico (figura 4).

Entre 11 e 13 meses no experimento de campo II, é observado pequeno aumento nos valores médios de temperatura e precipitação, se mantendo

próximo aos valores observados entre 4 meses e 8 meses. A cana-de-açúcar manteve a sua altura sem diferença significativa desde os 8 meses, entretanto, houve diminuição no grau Brix. Conforme discutido anteriormente, a cana-de-açúcar estava iniciando o período de florescimento, justificando assim a observação na diminuição do grau Brix (CARDOZO; SENTELHAS, 2013).

6 CONCLUSÃO

O presente estudo teve como objetivo avaliar o metaboloma da cana-de-açúcar ao longo de dois experimentos de campo através de uma abordagem conjunta do crescimento da planta e do seu transcriptoma transcriptômica; a fim de elucidar possíveis mecanismos moleculares responsáveis por regular o crescimento da planta em diferentes condições ambientais e ao longo da maturação. No primeiro experimento, o plantio foi realizado em período de baixa precipitação e temperatura média, ao passo que no segundo experimento, o plantio da cana-de-açúcar foi realizado em período de alta temperatura média e precipitação.

A análise do tecido de folha +1 demonstrou um padrão distinto no comportamento das vias metabólicas ao longo do experimento. Observou-se grande diferença na atividade das vias metabólicas relacionadas ao metabolismo de carboidratos e de biossíntese de metabólitos secundários.

A análise dos metabólitos pertencentes às vias com diferença significativa no nível de atividade entre as comparações intra experimento revelou um perfil característico de resposta aos períodos de déficit hídrico. De forma similar, os genes avaliados nas vias metabólicas com diferença

significativa demonstraram estar relacionados ao processo de resposta ao estresse abiótico, apresentando elevado nível de expressão em ambos os experimentos de campo.

O mecanismo de resposta ao déficit hídrico demonstrou ser similar em diferentes estágios de crescimento da cana-de-açúcar. Durante o período de déficit hídrico foi observado demasiado acúmulo de aminoácidos e açúcares na folha da cana-de-açúcar, bem como o acúmulo de monolignóis e monossacarídeos precursores na síntese de lignina e parede celular respectivamente. Além disso, o acúmulo do composto 2-oxoglutarato e de glicosídeos flavonóis sugerem estar associados à inibição do crescimento como consequência secundária ao déficit hídrico.

Os presente trabalho foi importante para o desenvolvimento de uma plataforma de análise do metaboloma de cana-de-açúcar e os resultados obtidos forneceram informações importantes dos mecanismos moleculares que podem estar modulando o crescimento da cana-de-açúcar sob diferentes condições ambientais. Tais informações podem servir como base para aumento da produtividade e desenvolvimento de estratégias de melhoramento com engenharia genética para a obtenção de cultivares melhoradas.

As hipóteses formuladas neste estudo necessitam ser validadas por metodologias específicas. No caso do metaboloma, conduzir um estudo direcionado aos metabólitos de interesse visando a elucidação estrutural bem como a quantificação da abundância em cada um dos pontos experimentais, em especial os candidatos ambíguos. Para o transcriptoma, será necessário a quantificação dos níveis de expressão dos genes de interesse, pelo uso de uma metodologia de maior sensibilidade, como o qPCR.

7 PROPOSTAS PARA ESTUDOS FUTUROS

- Realizar estudo direcionado dos metabólitos que apresentaram ambiguidade a fim identificar a estrutura de cada um dos candidatos ambíguos. Para isso, será utilizado uma tecnologia de espectrometria de massas com capacidade de realizar a fragmentação das moléculas alvos (MS/MS);

- Validar a expressão dos genes identificados com diferença significativa nas comparações realizadas, que possuem importante papel na resposta ao período de déficit hídrico. Bem como, realizar um estudo direcionado dos genes que não foram identificados no estudo, mas que possam estar modulando alguns dos processos descritos na sessão de "Discussão" dos dados;

- Também será necessário a adoção de um *pipeline* computacional para a análise de dados de LC-MS/MS.

8 REFERÊNCIA BIBLIOGRÁFICA

AFENDI, F. M. et al. KNApSACK family databases: Integrated metabolite-plant species databases for multifaceted plant research. **Plant and Cell Physiology**, v. 53, n. 2, 2012.

AGGIO, R. B. M.; RUGGIERO, K.; VILLAS-BÔAS, S. G. Pathway Activity Profiling (PAPi): from the metabolite profile to the metabolic pathway activity. **Bioinformatics (Oxford, England)**, v. 26, n. 23, p. 2969–76, 2010.

ALLWOOD, J. W.; ELLIS, D. I.; GOODACRE, R. **Metabolomic technologies and their application to the study of plants and plant-host interactions** *Physiologia Plantarum*, 2008.

ALONSO, A.; MARSAL, S.; JULIÀ, A. Analytical methods in untargeted metabolomics: state of the art in 2015. **Frontiers in bioengineering and biotechnology**, v. 3, n. March, p. 23, 2015.

ARAÚJO, W. L. et al. 2-Oxoglutarate: linking TCA cycle function with amino acid, glucosinolate, flavonoid, alkaloid, and gibberellin biosynthesis. **Frontiers in plant science**, v. 5, n. October, p. 552, 2014.

BALSAMO, R. et al. Leaf biomechanical properties in *Arabidopsis thaliana* polysaccharide mutants affect drought survival. **Journal of Biomechanics**, v. 48, n. 15, p. 4124–4129, 2015.

BOSCH, S.; ROHWER, J. M.; BOTHA, F. C. The sugarcane metabolome. **Proceedings of the Annual Congress - South African Sugar Technologists' Association**, n. No.77, p. 129–133, 2003.

BOX, G. E. P.; BEHNKEN, D. W. Some New Three Level Designs for the Study of Quantitative Variables. **Technometrics**, v. 2, n. 4, p. 455–475, 1960.

BRADFORD, M. Rapid and Sensitive Method for Quantification of Microgram Quantities of Protein utilizing principle of Protein-Dye-Binding. **Analytical Biochemistry**, v. 72, n. 1-2, p. 248–254, 1976.

CAMAÑES, G. et al. An untargeted global metabolomic analysis reveals the biochemical changes underlying basal resistance and priming in *Solanum lycopersicum*, and identifies 1-methyltryptophan as a metabolite involved in plant responses to *Botrytis cinerea* and *Pseudomonas sy.* **Plant Journal**, v. 84, n. 1, p. 125–139, 2015.

CARDOZO, N. P.; SENTELHAS, P. C. Climatic effects on sugarcane ripening under the influence of cultivars and crop age. **Sci. agric. (Piracicaba, Braz.)**, v. 70, n. 6, p. 449–456, 2013.

CHEN, Y.; HOEHENWARTER, W. Changes in the Phosphoproteome and Metabolome Link Early Signaling Events to Rearrangement of Photosynthesis and Central Metabolism in Salinity and Oxidative Stress Response in *Arabidopsis*. **Plant physiology**, v. 169, n. 4, p. 3021–33, 2015.

CHONG, B. F. et al. Growth and metabolism in sugarcane are altered by the creation of a new hexose-phosphate sink. **Plant Biotechnology Journal**, v. 5, n. 2, p. 240–253, 2007.

DE SIQUEIRA FERREIRA, S. et al. Biofuel and energy crops: high-yield Saccharinae take center stage in the post-genomics era. **Genome biology**, v. 14, n. 6, p. 210, 2013.

DEGUCHI, M. et al. An engineered sorbitol cycle alters sugar composition, not growth, in transformed tobacco. **Plant, Cell and Environment**, v. 29, n. 10, p. 1980–1988, 2006.

DUNN, W. B. et al. **Mass appeal: Metabolite identification in mass spectrometry-focused untargeted metabolomics** *Metabolomics*, 2013.

FERREIRA, S. S. et al. Co-expression network analysis reveals transcription factors associated to cell wall biosynthesis in sugarcane. **Plant Molecular Biology**, v. 91, n. 1-2, p. 15–35, 2016.

GIAVALISCO, P. et al. Elemental formula annotation of polar and lipophilic metabolites using ¹³C, ¹⁵N and ³⁴S isotope labelling, in combination with high-resolution mass spectrometry. **Plant Journal**, v. 68, n. 2, p. 364–376, 2011.

GLASSOP, D. et al. Changes in the sugarcane metabolome with stem development. Are they related to sucrose accumulation? **Plant & cell physiology**, v. 48, n. 4, p. 573–584, 2007.

GOLDEMBERG, J.; COELHO, S. T.; GUARDABASSI, P. **The sustainability of ethanol production from sugarcane** *Energy Policy*, 2008.

GONG, L. et al. Transcriptome profiling of the potato (*Solanum tuberosum* L.) plant under drought stress and water-stimulus conditions. **PLoS ONE**, v. 10, n. 5, 2015.

GOTTESFELD, J. M. et al. **Regulation of gene expression by small molecules.** *Nature*, 1997.

ISKANDAR, H. M. et al. Identification of drought-response genes and a study of their expression during sucrose accumulation and water deficit in sugarcane culms. **BMC Plant Biology**, v. 11, n. 1, p. 12, 2011.

JALALI, S. et al. Systematic Transcriptome Wide Analysis of lncRNA-miRNA Interactions. **PLoS ONE**, v. 8, n. 2, 2013.

KATAJAMAA, M.; OREŠIČ, M. Processing methods for differential analysis of LC/MS profile data. **BMC bioinformatics**, v. 6, p. 179, 2005.

KIDO, É. A. et al. **New Insights in the Sugarcane Transcriptome Responding to Drought Stress as Revealed by Supersage** *The Scientific World Journal*, 2012.

KOGENARU, S. et al. RNA-seq and microarray complement each other in transcriptome profiling. **BMC genomics**, v. 13, n. 1, p. 629, 2012.

KUHL, C. et al. CAMERA: An integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. **Analytical Chemistry**, v. 84, n. 1, p. 283–289, 2012.

LAVINSKY, A. O. et al. Partitioning between primary and secondary metabolism of carbon allocated to roots in four maize genotypes under water deficit and its effects on productivity. **The Crop Journal**, v. 3, n. 5, p. 379–386, 2015.

LEMBKE, C. G. et al. **Identification of sense and antisense transcripts regulated by drought in sugarcane** *Plant Molecular Biology*, 2012.

LIBISELLER, G. et al. IPO: a tool for automated optimization of XCMS parameters. **BMC Bioinformatics**, v. 16, n. 1, p. 118, 2015.

- LOVÉN, J. et al. Revisiting global gene expression analysis. *Cell*, v. 151, n. 3, p. 476–482, 2012.
- MALONE, J. H.; OLIVER, B. Microarrays, deep sequencing and the true measure of the transcriptome. *BMC biology*, v. 9, n. 2, p. 34, 2011.
- MATTIELLO, L. et al. Physiological and transcriptional analyses of developmental stages along sugarcane leaf. *BMC Plant Biology*, v. 15, n. 1, p. 300, 29 dez. 2015.
- MCCORMICK, A. J.; WATT, D. A.; CRAMER, M. D. Supply and demand: Sink regulation of sugar accumulation in sugarcane. *Journal of Experimental Botany*, v. 60, n. 2, p. 357–364, 2009.
- NAKABAYASHI, R.; SAITO, K. **Integrated metabolomics for abiotic stress responses in plants** *Current Opinion in Plant Biology* Elsevier Ltd, , 2015.
- NAZ, S. Analytical protocols based on LC – MS , GC – MS and CE – MS for nontargeted metabolomics of biological tissues. *Bioanalysis*, v. 6, n. 12, p. 1657–1677, 2014.
- NISHIYAMA, M. Y. et al. Full-length enriched cDNA libraries and ORFeome analysis of sugarcane hybrid and ancestor genotypes. *PLoS ONE*, v. 9, n. 9, 2014.
- PAPINI-TERZI, F. S. et al. Sugarcane genes associated with sucrose content. *BMC genomics*, v. 10, p. 120, 2009.
- ROBERTO, G. G. et al. Variação da fotossíntese e dos teores de carboidratos induzidos por etefom e déficit hídrico na fase de maturação da cana-de-açúcar. *Bragantia*, p. 379–386, 2015.
- ROCHFORT, S. **Metabolomics reviewed: A new “omics” platform technology for systems biology and implications for natural products research** *Journal of Natural Products*, 2005.
- ROGERS, S. et al. Probabilistic assignment of formulas to mass peaks in metabolomics experiments. *Bioinformatics*, v. 25, n. 4, p. 512–518, 2009.
- ROGERS, S. et al. Bayesian Approaches for Mass Spectrometry-Based Metabolomics. In: **Handbook of Statistical Systems Biology**. Chichester, UK: John Wiley & Sons, Ltd, 2011. v. 15p. 467–476.
- SCOTLAND, R. W.; WORTLEY, A. H. How many species of seed plants are there? *Taxon*, v. 52, n. 1, p. 101–104, 2003.
- SHEVELEVA, E. et al. Sorbitol-6-phosphate dehydrogenase expression in transgenic tobacco. High amounts of sorbitol lead to necrotic lesions. *Plant physiology*, v. 117, n. 3, p. 831–9, 1998.

SILVA, A. R. DA et al. Avaliação do coeficiente de variação experimental para caracteres de frutos de pimenteiras. **Revista Ceres**, v. 58, n. 2, p. 168–171, 2012.

SILVA, R. R. et al. ProbMetab: An R package for Bayesian probabilistic annotation of LC-MS-based metabolomics. **Bioinformatics**, v. 30, n. 9, p. 1336–1337, 2014.

SOUZA, G. M., R. L. VERDADE, L. M. JOLY, C. A. NETTO, P. E. A. CRUZ, C. H. B. CANTARELLA, H. CHUM, H. L. CORTEZ, L. A. B. DIAZ-CHAVEZ, R. FERNANDES, E. FINCHER, G. B. FOUST, T. GOLDEMBERG, J. NOGUEIRA, L. A. H. HUNTLEY, B. J. JOHNSON, F. X. KAFFKA, S. KARP, A. LEAL, M. R. L. V. LONG, S. P. LYND, L. R. MACEDO, I. C. MACIEL FILHO, R. NASSAR, A. M. , *et al.* ; Bioenergy & Sustainability. Policy Brief.. **SCOPE**, v. 1, p. 6, 2015.

SMITH, C. A. et al. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. **Analytical chemistry**, v. 78, n. 3, p. 779–787, 2006.

TAUTENHAHN, R.; BÖTTCHER, C.; NEUMANN, S. Highly sensitive feature detection for high resolution LC/MS. **BMC bioinformatics**, v. 9, p. 504, 2008.

VERDIER, J. et al. A regulatory network-based approach dissects late maturation processes related to the acquisition of desiccation tolerance and longevity of *Medicago truncatula* seeds. **Plant physiology**, v. 163, n. 2, p. 757–74, 2013.

VETTORE, A. L. et al. **The libraries that made SUCESTGenetics and Molecular Biology**, 2001.

WANG, J. et al. Carbon partitioning in sugarcane (*Saccharum* species). **Frontiers in plant science**, v. 4, p. 201, 2013.

WEBER, R. J. M. et al. Characterization of isotopic abundance measurements in high resolution FT-ICR and Orbitrap mass spectra for improved confidence of metabolite identification. **Analytical Chemistry**, v. 83, n. 10, p. 3737–3743, 2011.

YAMADA, K. et al. Functional analysis of an *Arabidopsis thaliana* abiotic stress-inducible facilitated diffusion transporter for monosaccharides. **Journal of Biological Chemistry**, v. 285, n. 2, p. 1138–1146, 2010.

ZHOU, Y. et al. Transcriptome Analysis of Salt Stress Responsiveness in the Seedlings of Dongxiang Wild Rice (*Oryza rufipogon* Griff.). **PLOS ONE**, v. 11, n. 1, p. e0146242, 11 jan. 2016.

9 ANEXOS

9.1 SÚMULA CURRICULAR

DADOS PESSOAIS

Nome: Davi Toshio Inada

Local e data de nascimento: 07/04/1992, Tomé-Açu (PA)

EDUCAÇÃO

Universidade Federal do Pará (UFPA), Brasil, 2014.
Graduação em Biomedicina (Bacharelado).

FORMAÇÃO COMPLEMENTAR

Curso de gestão para lab.- NBR/ISO/IEC 17.025:2005 (Carga horária: 20h).
Instituto Evandro Chagas, IEC, Brasil (2013).

Curso de Verão em Bioinformática (Carga horária: 40h).
Universidade de São Paulo, USP, Brasil (2013).

Química Medicinal: Oportunidades e Desafios (Carga horária: 8h).
Universidade Federal do Pará, UFPA, Brasil (2012).

Linux Básico e Redes Linux (Carga horária: 40h).
Universidade Federal do Pará, UFPA, Brasil (2011).

I Workshop Internacional Em Bioinformática (Carga horária: 35h).
Instituto Evandro Chagas, IEC, Brasil (2011).

Curso de língua japonesa grau 1 ao grau 10.
Escola de língua japonesa de Tomé-Açu, Brasil (1997-2006).

OCUPAÇÃO

Bolsista de Mestrado, FAPESP (processo: 2014/078287), Setembro de 2014 a
Maio de 2016.

PUBLICAÇÕES

APRESENTAÇÃO DE TRABALHOS

INADA, D. T.; LEMBKE, C. G.; NISHIYAMA JUNIOR, M. Y.; VILLELA, L. Z.;
SILVA, R. R.; SOUZA, G. M.
Analysis of sugarcane transcriptoma and metabolite changes along the plant
maturation cycle. 2015. (Outra, Apresentação de Trabalho)
Áreas do conhecimento : Bioinformática

Referências adicionais : Brasil/Inglês. Meio de divulgação: Vários; Local: Faculdade de Ciências Farmacêuticas; Cidade: Ribeirão Preto - SP; Evento: First Brazilian Workshop on Bioinformatics/Chemometrics for Metabolomics; Inst.promotora/financiadora: Núcleo de Pesquisas em Produtos Naturais e Sintéticos

INADA, D. T.; VILLELA, L. Z.; LEMBKE, C. G.; NISHIYAMA-JR, M. Y.; FUJITA, A.; SOUZA, G. M.

Metabolomics of sugarcane leaves along two experimental fields for maturation cycle study, 2015. (Congresso,Apresentação de Trabalho)

Palavras-chave: Bioinformatica

Referências adicionais : Brasil/Português. Meio de divulgação: Impresso; Local: Universidade de São Paulo; Cidade: São Paulo; Evento: X-Meeting 2015 - 11th International Conference of the AB3C + Brazilian Symposium of Bioinformatics; Inst.promotora/financiadora: AB3C

SILVA, F. R.; CARDOSO, J. F.; INADA, D. T.; SILVA, S. P.; SOUSA JUNIOR, E. C.; VIANEZ JUNIOR, J. L. S. G.; NUNES, M.R.T.

COMPARATIVE MODELING AND MOLECULAR DYNAMICS OF THE ENVELOPE PROTEIN CEPA BE AN423429 VIRUS ENCEPHALITIS OF SAINT LOUIS, 2013. (Congresso,Apresentação de Trabalho)

Referências adicionais : Brasil/Inglês. Meio de divulgação: Impresso; Local: UFPE; Cidade: Recife, PE - Brazil; Evento: X-meeting/BSB 2013 Conference; Inst.promotora/financiadora: AB3C

SOUSA JUNIOR, E. C.; SANTOS, M. C.; FERREIRA, D. L.; INADA, D. T.; FILIZZOLA, E. M. A.; MEDEIROS, R. C.; MELLO, W. A.

EPIDEMIOLOGIA MOLECULAR DOS VÍRUS INFLUENZA E METAPNEUMOVÍRUS, 2012. (Congresso,Apresentação de Trabalho)

Palavras-chave: Bioinformatica, Epidemiologia moleciular

Áreas do conhecimento : Genética

Referências adicionais : Brasil/Português. Meio de divulgação: Impresso; Local: Hangar Centro de Convenções e Feiras da Amazônia; Cidade: Belém- PA; Evento: XVI CONGRESSO MÉDICO AMAZÔNICO; Inst.promotora/financiadora: Sociedade Médico-Cirúrgica do Pará

SILVA, D. E. A.; Inada, Davi Toshio; VIANEZ JUNIOR, J. L. S. G.; CARDOSO, J. F.; LIMA, C. P. S.; SOUSA JUNIOR, E. C.; NUNES, K. N. B.; Nunes, Márcio Roberto T.; VASCONCELOS, P. F. C.

FULL-LENGTH GENOME DETERMINATION OF BIMITI VIRUS USING THE PIROSSEQUENCING METHOD, 2012. (Congresso,Apresentação de Trabalho)

Referências adicionais : Brasil/Inglês. Meio de divulgação: Impresso; Local: Rafain Palace Hotel & Convention Center; Cidade: Foz do Iguaçu- PR; Evento: XXIII Congresso Brasileiro de Virologia, VII Encontro de Virologia do Mercosul; Inst.promotora/financiadora: Sociedade Brasileira de Virologia

CARDOSO, J. F.; NUNES, M.R.T.; SOUSA JUNIOR, E. C.; Inada, Davi Toshio; OLIVEIRA, R. L. S.; NASCIMENTO, S. B.; PINHEIRO, S. S.; SILVA, J. L.; VIANEZ JUNIOR, J. L. S. G.

MOLECULAR MODELING AND DOCKING STUDY OF DENDUE VIRUS 4 NS2B/ NS3 PROTEIN USING PEPTIDIC INHIBITORS, 2012. (Congresso,Apresentação de Trabalho)

Referências adicionais : Brasil/Português. Meio de divulgação: Impresso; Local: Unicamp Conversion Center; Cidade: Campinas- São Paulo; Evento: 8th International Conference of the Brazilian Association for Bioinformatics and Computational Biology - X-Meeting; Inst.promotora/financiadora: AB3C

SOUSA JUNIOR, E. C.; CARDOSO, J. F.; VIANEZ JUNIOR, J. L. S. G.; OLIVEIRA, R. L. S.; INADA, D. T.; LIMA, C. P. S.; SILVA, F. R.; NUNES, K. N. B.; Nunes, Márcio Roberto T.; VASCONCELOS, P. F. C.

SEQUENCING AND MODELING OF THE ENVELOPE PROTEIN OF A YELLOW FEVER VIRUS STRAIN (BEAR378600) ISOLATED IN THE BRAZILIAN AMAZON, 2012. (Congresso,Apresentação de Trabalho)

Referências adicionais : Brasil/Inglês. Meio de divulgação: Impresso; Local: Rafain Palace Hotel & Convention Center; Cidade: Foz do Iguaçu- PR; Evento: XXIII Brazilian Congress of Virology & VII Mercosul Meeting of Virology (XXIII CBV); Inst.promotora/financiadora: Sociedade Brasileira de Virologia

INADA, D. T.; SOUSA JUNIOR, E. C.; CARDOSO, J. F.; SILVA, F. R.; OLIVEIRA, R. L. S.; NUNES, M.R.T.

GENOMICS ANNOTATION OF VIRAL STRAINS DENV4 (DENGUE VIRUS TYPE 4) COMING FROM PIROSEQUENCING, 2011. (Outra,Apresentação de Trabalho)

Áreas do conhecimento : Genética Molecular e de Microorganismos

Referências adicionais : Brasil/Inglês. Meio de divulgação: Impresso

Este trabalho foi premiado com a Menção honrosa no "I Workshop Internacional em Bioinformática do Instituto Evandro Chagas"; Local: Seção de Arbovirologia e Febres Hemorrágicas; Cidade: Belém; Evento: I Workshop Internacional em Bioinformática do Instituto Evandro Chagas; Inst.promotora/financiadora: Centro de Inovações Tecnológicas do Instituto Evandro Chagas

SOUSA JUNIOR, E. C.; CARDOSO, J. F.; OLIVEIRA, R. L. S.; SILVA, F. R.; INADA, D. T.; NUNES, M.R.T.

MOLECULAR MODELING AND DOCKING OF YELLOW FEVER VIRUS NS2B-NS3 PROTEIN, 2011. (Outra,Apresentação de Trabalho)

Palavras-chave: Bioinformática, modelagem molecular, dengue, iec

Áreas do conhecimento : Genética Molecular e de Microorganismos

Referências adicionais : Brasil/Inglês. Meio de divulgação: Impresso

Este trabalho recebeu a premiação de menção honrosa no I Workshop Internacional em Bioinformática do Instituto Evandro Chagas,; Local: Seção de Arbovirologia e Febres Hemorrágicas; Cidade: Ananindeua; Evento: I Workshop Internacional em Bioinformática do Instituto Evandro Chagas; Inst.promotora/financiadora: Centro de Inovações Tecnológicas do Instituto Evandro Chagas

CARDOSO, J. F.; SOUSA JUNIOR, E. C.; INADA, D. T.; SILVA, F. R.; OLIVEIRA, R. L. S.; NUNES, M.R.T.; SILVA, A. L. C.

MOLECULAR MODELING OF DENGUE VÍRUS 4 NS3 PROTEIN, 2011.
(Outra,Apresentação de Trabalho)

Palavras-chave: Bioinformática, modelagem molecular, dengue, iec

Áreas do conhecimento : Genética Molecular e de Microorganismos

Referências adicionais : Brasil/Inglês. Meio de divulgação: Impresso

Este trabalho foi premiado como 1º lugar no I Workshop Internacional em Bioinformática do Instituto Evandro Chagas.; Local: Seção de Arbovirologia e Febres Hemorrágicas; Cidade: Ananindeua; Evento: I Workshop Internacional em Bioinformática do Instituto Evandro Chagas; Inst.promotora/financiadora: Centro de Inovações Tecnológicas do Instituto Evandro Chagas

OLIVEIRA, R. L. S.; SOUSA JUNIOR, E. C.; SILVA, F. R.; CARDOSO, J. F.; INADA, D. T.; NUNES, M.R.T.; SILVA, J. L.

MOLECULAR MODELING OF THE BUNYAMWERA ORTHOBUNYAVIRUS GC GLYCOPROTEIN, 2011. (Outra,Apresentação de Trabalho)

Palavras-chave: Bioinformática, modelagem molecular, dengue, iec

Áreas do conhecimento : Genética Molecular e de Microorganismos

Referências adicionais : Brasil/Inglês. Meio de divulgação: Impresso; Local: Seção de Arbovirologia e Febres Hemorrágicas; Cidade: Ananindeua; Evento: I Workshop Internacional em Bioinformática do Instituto Evandro Chagas; Inst.promotora/financiadora: Centro de Inovações Tecnológicas do Instituto Evandro Chagas

SILVA, F. R.; SOUSA JUNIOR, E. C.; CARDOSO, J. F.; OLIVEIRA, R. L. S.; INADA, D. T.; NUNES, M.R.T.; SILVA, J. L.

MOLECULAR MODELING OF THE VIRUS OROPOUCHE GC GLYCOPROTEIN, 2011. (Outra,Apresentação de Trabalho)

Palavras-chave: Bioinformática, modelagem molecular, dengue, iec

Áreas do conhecimento : Genética Molecular e de Microorganismos

Referências adicionais : Brasil/Inglês. Meio de divulgação: Impresso; Local: Seção de Arbovirologia e Febres Hemorrágicas; Cidade: Ananindeua; Evento: I Workshop Internacional em Bioinformática do Instituto Evandro Chagas; Inst.promotora/financiadora: Centro de Inovações Tecnológicas do Instituto Evandro Chagas

ARTIGOS

LIMA, A. R. J. ; SIQUEIRA, A. S. ; DOS SANTOS, B. G. S. ; DA SILVA, F. D. F. ; **INADA, D. T.** ; LIMA, C. P. ; CARDOSO, J. F. ; VIANEZ-JUNIOR, J. L. S. G. ; NUNES, M. R. T. ; GONCALVES, E. C. . Draft Genome Sequence of Rhodobacter sp. Strain CACIA 14H1, a Heterotrophic Bacterium Obtained from a Nonaxenic Culture of a Cyanobium Species. Genome Announcements, v. 2, p. e01116-13-e01116-13, 2014.

Garza, Daniel Rios ; Thompson, Cristiane C. ; Loureiro, Edvaldo Carlos Brito ; Dutilh, Bas E. ; **Inada, Davi Toshio** ; Junior, Edivaldo Costa Sousa ; Cardoso, Jedson Ferreira ; Nunes, Márcio Roberto T. ; de Lima, Clayton Pereira Silva ; Silvestre, Rodrigo Vellasco Duarte ; Nunes, Keley Nascimento Barbosa ; Santos, Elisabeth C. O. ; Edwards, Robert A. ; Vicente, Ana Carolina P. ; de Sá

Morais, Lena Lillian Canto ; Bereswill, Stefan . Genome-Wide Study of the Defective Sucrose Fermenter Strain of *Vibrio cholerae* from the Latin American Cholera Epidemic. *Plos One*, v. 7, p. e37283, 2012.

Figura anexa 1: Via da biossíntese de fenilpropanóides. Na via estão mapeados todos os possíveis compostos identificados no metaboloma da cana-de-açúcar e todos os possíveis genes expressos detectados no experimento de microarranjo. Cada proteína/ metabólito apresenta os seus respectivos valores de expressão e abundância relativa aos 4 meses, 8 meses, 11 meses e 13 meses; representados por um gradiente de cor crescente entre os valores -1 e 1. As barras indicam as proteínas (com o nível de expressão do gene que a codifica) e os círculos os metabólitos.

Figura anexa 2: Via do metabolismo de amido e sacarose. Na via estão mapeados todos os possíveis compostos identificados no metaboloma da cana-de-açúcar e todos os possíveis genes expressos detectados no experimento de microarranjo. Cada proteína/ metabólito apresenta os seus respectivos valores de expressão e abundância relativa aos 4 meses, 8 meses, 11 meses e 13 meses; representados por um gradiente de cor crescente entre os valores -1 e 1. As barras indicam as proteínas (com o nível de expressão do gene que a codifica) e os círculos os metabólitos.

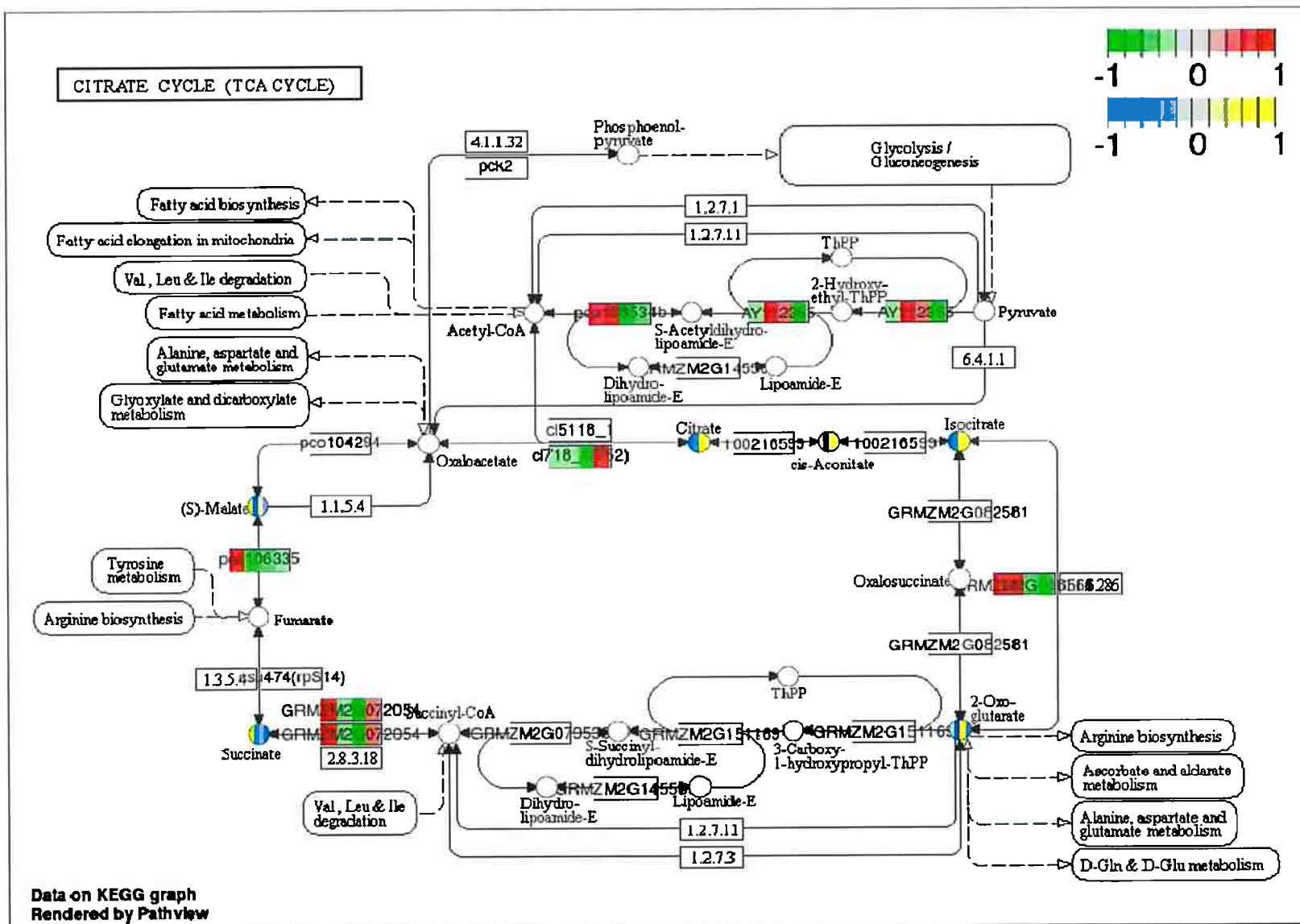
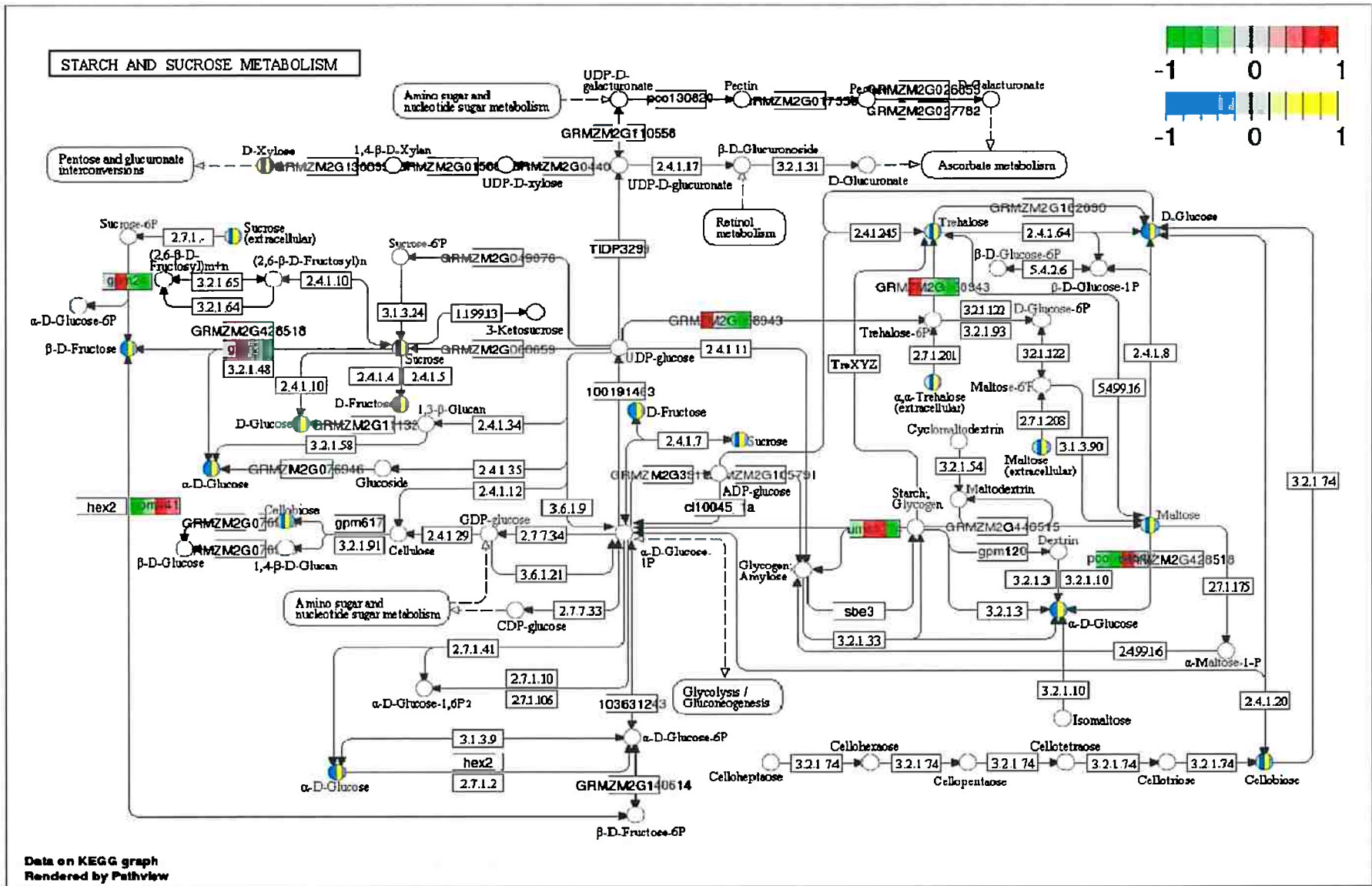


Figura anexa 3: Via do metabolismo de amido e sacarose. Na via estão mapeados todos os possíveis compostos identificados no metaboloma da cana-de-açúcar e todos os possíveis genes expressos detectados no experimento de microarranjo. Cada proteína/ metabólito apresenta os seus respectivos valores de expressão e abundância relativa aos 4 meses, 8 meses, 11 meses e 13 meses; representados por um gradiente de cor crescente entre os valores -1 e 1. As barras indicam as proteínas (com o nível de expressão do gene que a codifica) e os círculos os metabólitos.

Figura anexa 4: Via de biossíntese de fenilpropanóide. Na via estão mapeados todos os possíveis compostos identificados no metaboloma da cana-de-açúcar e todos os possíveis genes expressos detectados no experimento de microarranjo. Cada proteína/ metabólito apresenta os seus respectivos valores de expressão e abundância relativa aos 4 meses, 8 meses, 11 meses e 13 meses; representados por um gradiente de cor crescente entre os valores -1 e 1. As barras indicam as proteínas (com o nível de expressão do gene que a codifica) e os círculos os metabólitos.



Data on KEGG graph
Rendered by Pathview

Figura anexa 5: Via do metabolismo de amido e sacarose. Na via estão mapeados todos os possíveis compostos identificados no metaboloma da cana-de-açúcar e todos os possíveis genes expressos detectados no experimento de microarranjo. Cada proteína/ metabólito apresenta os seus respectivos valores de expressão e abundância relativa aos 4 meses, 8 meses, 11 meses e 13 meses; representados por um gradiente de cor crescente entre os valores -1 e 1. As barras indicam as proteínas (com o nível de expressão do gene que a codifica) e os círculos os metabólitos.

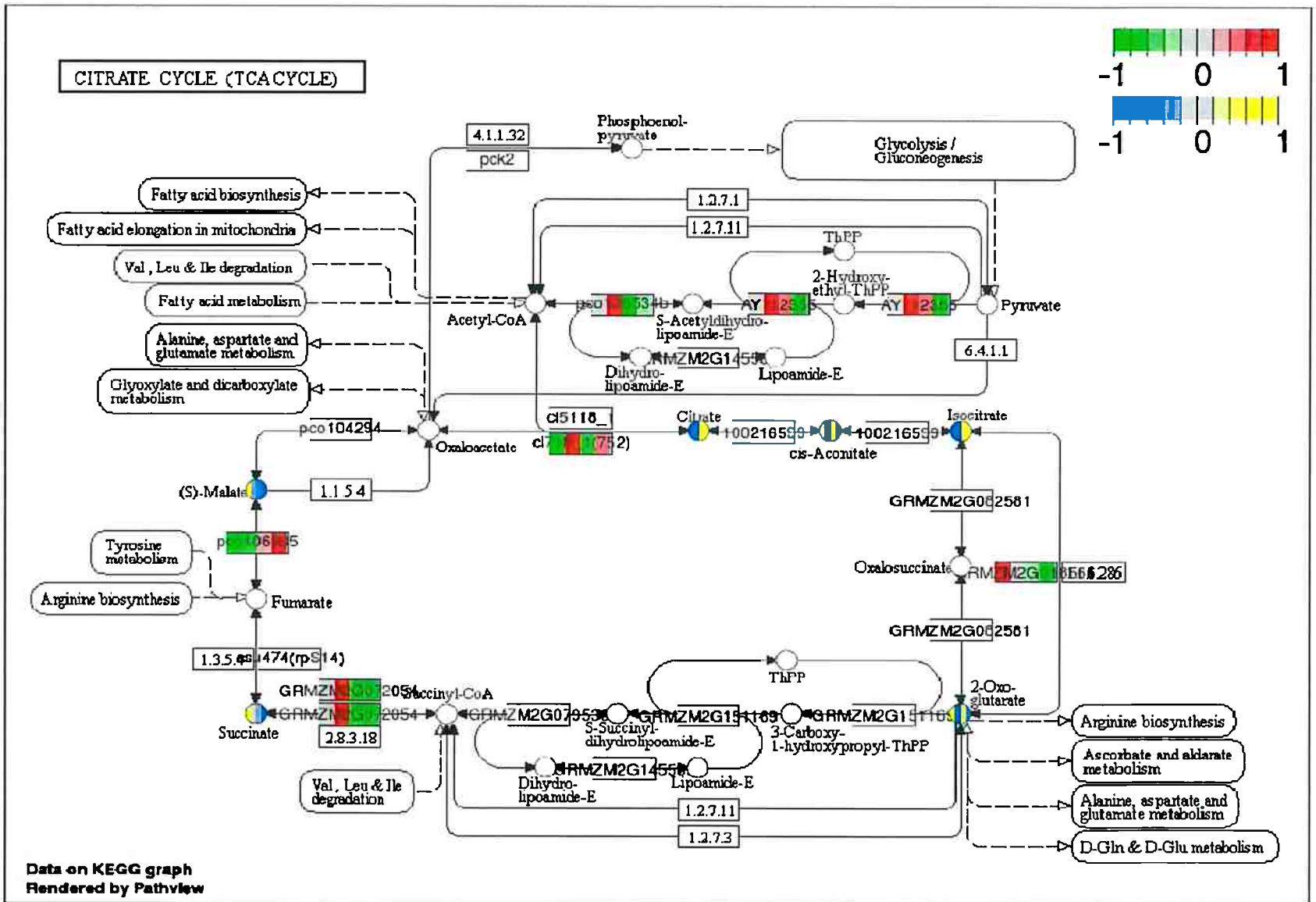


Figura anexa 6: Via do ciclo do ácido cítrico. Na via estão mapeados todos os possíveis compostos identificados no metaboloma da cana-de-açúcar e todos os possíveis genes expressos detectados no experimento de microarranjo. Cada proteína/ metabólito apresenta os seus respectivos valores de expressão e abundância relativa aos 4 meses, 8 meses, 11 meses e 13 meses; representados por um gradiente de cor crescente entre os valores -1 e 1. As barras indicam as proteínas (com o nível de expressão do gene que a codifica) e os círculos os metabólitos.

MATERIAIS SUPLEMENTARES

Os materiais suplementares deste estudo podem ser obtidos no CD anexo ao documento, ou então no link:

<https://www.dropbox.com/sh/q0wqg9xj6twpzxr/AAANkzfRykyQoYMstCXp7dLMa?dl=0>

