

**UNIVERSIDADE ESTADUAL PAULISTA
“JULIO DE MESQUITA FILHO”
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL**

**SELEÇÃO E VALIDAÇÃO DE GENES DE REFERÊNCIA
EM CONDIÇÕES DE FOTOPERÍODO ARTIFICIAL PARA
INDUÇÃO DO FLORESCIMENTO EM CANA-DE-AÇÚCAR.**

Paulo Henrique da Silva Santos

Biólogo

2020

**UNIVERSIDADE ESTADUAL PAULISTA “JULIO DE MESQUITA
FILHO”
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL**

**SELEÇÃO E VALIDAÇÃO DE GENES DE REFERÊNCIA
EM CONDIÇÕES DE FOTOPERÍODO ARTIFICIAL PARA
INDUÇÃO DO FLORESCIMENTO EM CANA-DE-AÇÚCAR.**

Paulo Henrique da Silva Santos
Orientadora: Dra. Luciana Rossini Pinto
Co-orientador: Prof. Dr. Michael dos Santos Brito

Dissertação apresentada à Faculdade de Ciências Agrárias e Veterinárias – UNESP, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Mestre em Agronomia (Genética e Melhoramento de Plantas).

S237s

Santos, Paulo Henrique da Silva

Seleção e validação de genes de referência em condições de fotoperíodo artificial para indução do florescimento em Cana-de-açúcar / Paulo Henrique da Silva Santos. -- Jaboticabal, 2020
82 p. : il., tabs.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal

Orientadora: Luciana Rossini Pinto

Coorientador: Michael dos Santos Brito

1. Genes de Referência. 2. Florescimento. 3. Cana-de-açúcar. I.
Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.



UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal



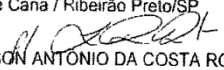
CERTIFICADO DE APROVAÇÃO

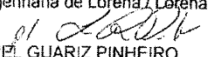
TÍTULO DA DISSERTAÇÃO: SELEÇÃO E VALIDAÇÃO DE GENES DE REFERÊNCIA EM CONDIÇÕES DE FOTOPERÍODO ARTIFICIAL PARA INDUÇÃO DO FLORESCIMENTO EM CANA-DE-AÇÚCAR

AUTOR: PAULO HENRIQUE DA SILVA SANTOS
ORIENTADORA: LUCIANA ROSSINI PINTO
COORIENTADOR: MICHAEL DOS SANTOS BRITO

Aprovado como parte das exigências para obtenção do Título de Mestre em AGRONOMIA (GENÉTICA E MELHORAMENTO DE PLANTAS), pela Comissão Examinadora:


Pesquisadora Dra. LUCIANA ROSSINI PINTO
IAC - Centro de Cana / Ribeirão Preto/SP


Prof. Dr. ELISSON ANTONIO DA COSTA ROMANEL (VIDEOCONFERÊNCIA)
Escola de Engenharia de Lorena / Lorena/SP


Prof. Dr. DANIEL GUARIZ PINHEIRO (VIDEOCONFERÊNCIA)
Departamento de Tecnologia / FCAV / UNESP - Jaboticabal

Jaboticabal, 03 de agosto de 2020

DADOS CURRICULARES DO AUTOR

PAULO HENRIQUE DA SILVA SANTOS – nascido em 13 de fevereiro de 1992 em Barrinha (SP), Brasil, iniciou o curso de Licenciatura e Bacharelado em Ciências Biológicas no ano de 2014, no Centro Universitário Barão de Mauá. Durante a graduação foi aluno de iniciação científica bolsista FAPESP no projeto “Caracterização dos alvos de regulação do fator de transcrição SHN por ChIPqPCR” no Laboratório de Biotecnologia do Instituto Agrônomo de Campinas (IAC-Centro de Cana-Ribeirão Preto-SP) sob a supervisão do pesquisador Dr. Michael dos Santos Brito no ano de 2016. Em 2017 recebeu o grau de Licenciatura Plena em Ciências Biológicas e também em 2017 o grau de Bacharel no mesmo curso pelo Centro Universitário Barão de Mauá. Em março de 2018 ingressou na Universidade Estadual Paulista “Júlio de Mesquita Filho” - Faculdade de Ciências Agrárias e Veterinárias - Campus de Jaboticabal no curso de Mestrado em Agronomia (Genética e Melhoramento de Plantas), sendo bolsista CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) durante o curso.

“Nós podemos explicar o azul-pálido desse pequeno mundo que conhecemos muito bem. Se um cientista alienígena, recém-chegado às imediações de nosso Sistema Solar, poderia fidedignamente inferir oceanos, nuvens e uma atmosfera espessa, já não é tão certo. Netuno, por exemplo, é azul, mas por razões inteiramente diferentes. Desse ponto distante de observação, a Terra talvez não apresentasse nenhum interesse especial. Para nós, no entanto, ela é diferente. Olhem de novo para o ponto. É ali. É a nossa casa. Somos nós. Nesse ponto, todos aqueles que amamos que conhecemos de quem já ouvimos falar, todos os seres humanos que já existiram, vivem ou viveram as suas vidas. Toda a nossa mistura de alegria e sofrimento, todas as inúmeras religiões, ideologias e doutrinas econômicas, todos os caçadores e saqueadores, heróis e covardes, criadores e destruidores de civilizações, reis e camponeses, jovens casais apaixonados, pais e mães, todas as crianças, todos os inventores e exploradores, professores de moral, políticos corruptos, "superastros", "líderes supremos", todos os santos e pecadores da história de nossa espécie, ali - num grão de poeira suspenso num raio de sol. A Terra é um palco muito pequeno em uma imensa arena cósmica. Pensem nos rios de sangue derramados por todos os generais e imperadores para que, na glória do triunfo, pudessem ser os senhores momentâneos de uma fração desse ponto. Pensem nas crueldades infinitas cometidas pelos habitantes de um canto desse pixel contra os habitantes mal distinguíveis de algum outro canto, em seus frequentes conflitos, em sua ânsia de recíproca destruição, em seus ódios ardentes. Nossas atitudes, nossa pretensa importância de que temos uma posição privilegiada no Universo, tudo isso é posto em dúvida por esse ponto de luz pálida. O nosso planeta é um pontinho solitário na grande escuridão cósmica circundante. Em nossa obscuridade, no meio de toda essa imensidão, não há nenhum indício de que, de algum outro mundo, virá socorro que nos salve de nós mesmos. (...)"

(Carl Sagan)

AGRADECIMENTOS

Ao Programa de Pós-graduação em Agronomia (Genética e Melhoramento de Plantas) da Faculdade de Ciências Agrárias e Veterinárias da Universidade Estadual Paulista “Júlio de Mesquita Filho” (FCAV/UNESP - Jaboticabal), pela oportunidade de realização do curso de Mestrado;

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

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), pelo auxílio financeiro (FAPESP 2016/02679-9; FAPESP 2018/08380-0) para execução do Projeto;

Aos meus orientadores Profa. Dra. Luciana Rossini Pinto e Prof. Dr. Michael dos Santos Brito pelas orientações, apoio e disponibilidade;

Aos colaboradores do projeto: Me. João Manechini, Prof. Dr. Renato Vicentini, Prof. Dr. Elisson Romanel e Prof. Dr. Stephen Jackson;

À Dra. Silvana Aparecida Creste Dias de Souza e ao Dr. Marcos Guimarães de Andrade Landell por ceder toda estrutura Centro de Cana do Instituto Agrônomo de Campinas, em Ribeirão Preto - SP para realização do projeto;

Aos colegas de pós-graduação Luiz e Carol, pelo companheirismo durante os eventos científicos e disciplinas;

Aos membros da banca do exame geral de qualificação e comissão examinadora: Profa. Dra. Luciana Rossini Pinto, Dra. Andressa Peres Bini, Dr. Marcel Fernando da Silva, Prof. Dr. Elisson Romanel e Prof. Dr. Daniel Guariz Pinheiro, muito obrigado pela disponibilidade e pelas correções;

A todos os colegas do Laboratório de Biotecnologia do Centro de Cana do Instituto Agrônomo de Campinas, em Ribeirão Preto, que de alguma maneira colaboraram para realização do trabalho;

E por último, porém não menos importante, toda a minha família pelo suporte durante os estudos.

Muito obrigado!

SUMÁRIO

Sumário

RESUMO	iii
ABSTRACT	iv
LISTA DE FIGURAS.....	v
LISTA DE TABELAS	vii
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA.....	4
2.1. A cultura cana-de-açúcar.....	4
2.2. Fotoperíodo.....	4
2.3. Aspectos Moleculares do Florescimento	6
2.4. Genes de referência em Cana-de-Açúcar	9
3. MATERIAIS E MÉTODOS.....	13
3.1. Material Biológico e Condições Experimentais:.....	13
3.2. Extração de RNA total e Síntese de cDNA	14
3.3. Seleção e Validação de Genes de Referência	14
3.4. Desenvolvimento de <i>primers</i> para genes diferencialmente expressos...16	
3.5. Estudo de expressão relativa do gene <i>ScFT3</i> ao longo do tempo.....18	
3.6. Condições da Reação de PCR quantitativa – PCR quantitativo em tempo real (RT-qPCR) :	19
4. RESULTADOS	19

4.1. Dados de PCR quantitativo em tempo real (RT-qPCR) dos candidatos a genes de referência.....	19
4.2. Análise de estabilidade dos candidatos a genes de referência por condição experimental	22
4.3. Validação de genes de referência.....	25
4.4. Validação de transcritos diferencialmente expressos.....	28
4.6. Estudo de expressão relativa ao longo do tempo (<i>timecourse</i>) gene ScFT3	36
5. DISCUSSÃO.....	38
5.1. Genes de Referência.....	38
5.2. Expressão relativa de genes alvos diferencialmente expressos.....	40
5.3. Expressão relativa de <i>ScFT3</i> ao longo do desenvolvimento	43
6. CONCLUSÕES	44
7. REFERÊNCIAS.....	45
Apêndice A.....	50
Tabela Suplementar 1A. Dados completos algoritmos <i>NormFinder</i> e <i>BestKeeper</i>	51
Tabela Suplementar 2A. Dados completos algoritmo <i>RefFinder</i>	57
Sequência dos transcritos utilizados para desenhar os <i>primers</i> dos genes diferencialmente expressos:	57

RESUMO

A cana-de-açúcar é uma cultura de alto impacto no agronegócio do Brasil, principalmente pela produção de açúcar, álcool e bioeletricidade. O florescimento é um processo fisiológico importante para os programas de melhoramento genético na realização de cruzamentos, e indesejável aos produtores por proporcionar, em algumas cultivares, perda de acúmulo de sacarose prejudicando diretamente o rendimento dos canaviais. Desta forma, o conhecimento da expressão dos genes envolvidos no processo de indução (transição do meristema vegetativo para reprodutivo) é de extrema importância para o desenvolvimento de novas cultivares de cana-de-açúcar. Entretanto, o estudo da expressão de genes via PCR quantitativo em tempo (RT-qPCR) requer a escolha correta de genes de referência para normalizar a expressão dos genes alvos e, consequentemente, proporcionar a sua correta interpretação. O objetivo deste trabalho foi identificar genes de referência estáveis para validação de genes diferencialmente expressos via RT-qPCR. A partir de um experimento de RNA-Seq que investigou genes envolvidos na indução floral fotoperiódica na cultivar comercial de cana-de-açúcar IACSP96-7569 foi selecionado os genes alvos para validação. Oito diferentes candidatos a genes de referência (*UBQ1*, *UBQ2*, *RPL*, *GAPDH*, *25SrRNA1*, *TUB*, *EF1- α* e *TIPS-41*) foram testados sob dois tratamentos fotoperiódicos (dia curto e dia longo), em dois momentos de desenvolvimento (correspondentes aos estágios iniciais da indução floral e formação de primórdios florais) em dois tecidos foliares (folha madura e folha imatura). A estabilidade dos candidatos a genes de referência em todas as condições foi avaliada pelos algoritmos *NormFinder*, *BestKeeper* e *RefFinder* e mostrou que *UBQ1* e *TUB* foram os genes de referência mais estáveis. O par de genes de referência mais estável (*UBQ1/TUB*) e o menos estável (*GAPDH/25SrRNA1*) foi utilizado para normalizar a expressão de dois transcritos identificados como expressos diferencialmente a partir dos dados de RNA-Seq, *PIL5* (“*PHYTOCHROME INTERACTING FACTOR 3-LIKE 5*”) e *LHP1* (“*LIKE HETEROCHROMATIN PROTEIN 1*”) evidenciando os seus efeitos na resposta da expressão destes alvos. A comparação entre os resultados de expressão diferencial para estes dois alvos via RNA-Seq (FPKM) e RT-qPCR mostrou certa divergência sinalizando a importância da utilização de genes de referência adequados para validar os resultados do RNA-Seq. Além da seleção e validação de genes de referência, 13 genes alvos incluindo dois transcritos de função desconhecida foram avaliados via RT-qPCR, bem como a expressão do gene ortólogo ao *FLOWERING LOCUS T (FT)* ao longo de seis tempos de coleta em folha madura. A seleção e validação de genes de referência foram essenciais para a correta interpretação dos dados de RT-qPCR e servirão como um guia nos estudos de florescimento em cana-de-açúcar, especialmente aqueles voltados para a indução fotoperiódica da floração.

Palavras-chave: Cana-de-Açúcar, Florescimento, PCR quantitativo em tempo real (RT-qPCR), Genes de Referência, Fotoperiodismo, RNA-Seq.

ABSTRACT

Sugarcane is an important economic crop worldwide, highlighting in Brazil's agribusiness mainly for sugar, ethanol, and bioelectricity production. The flowering induction process, marked by the apical meristem transition from vegetative to reproductive stage, negatively impacts sugarcane production and quality traits. Despite its importance, the knowledge of the genes involved in the flowering induction pathway as well as its expression in the beginning of the induction process is still scarce in this high complex polyploid crop. Reverse transcription quantitative real-time (RT-qPCR) is the most used technique for gene expression analysis and requires the selection of reliable reference genes to normalize the expression of target genes and for correct interpretation of the RT-qPCR data. The aim of this work was to identify stable reference genes for RT-qPCR validation of differentially expressed genes previously identified in a RNA-Seq experiment investigating genes involved with photoperiodic floral induction in the commercial sugarcane cultivar IACSP96-7569. Eight different candidate reference genes (*UBQ1*, *UBQ2*, *RPL*, *GAPDH*, *25SrRNA1*, *TUB*, *EF1- α* , and *TIPS-41*) were tested under two photoperiodic treatments (short day and long day), at two developmental time points (corresponding to the early stages of floral induction and floral primordia formation) in two leaf tissues (mature leaf and spindle leaf). The stability of the candidate reference genes in all conditions was evaluated with *NormFinder*, *BestKeeper*, and *RefFinder* algorithms and showed that *UBQ1* and *TUB* were the most stable reference genes. The most stable reference gene pair (*UBQ1/TUB*) and the least stable pair (*GAPDH/25SrRNA1*) were used to normalize the expression of two transcripts identified as being differentially expressed from our RNA-Seq data: *PIL5* (*PHYTOCHROME INTERACTING FACTOR 3-LIKE 5*) and *LHP1* (*LIKE HETEROCHROMATIN PROTEIN*) highlighting its effect on the expression of the targets. The comparison between the results of differential expression of these two target genes via RNA-Seq (FPKM) and RT-qPCR showed some divergence signaling the importance of using appropriate reference genes to validate the RNA-Seq results. Additionally, 13 target genes including two transcripts of unknown function were evaluated via RT-qPCR, as well as the developmental timecourse relative expression of *FLOWERING LOCUS T (FT)* gene over six time points in mature leaf. The selection and validation of reference genes were essential for the correct interpretation of RT-qPCR data and can be used as a guide in sugarcane flowering studies, especially those of photoperiodic flowering induction.

Keywords: Sugarcane, Flowering, Reverse transcription - quantitative polymerase chain reaction (RT-qPCR), Reference Genes, Photoperiodism, RNA-Seq.

LISTA DE FIGURAS

Figura 1. Esquema da via de indução do florescimento baseado em *Arabidopsis thaliana* e culturas cereais. Abaixo da figura do sol estão os genes relacionados com fotoperíodo, dentro do círculo estão os genes relacionados com relógio circadiano. As setas em verde indicam regulação positiva e em vermelha regulação negativa. Fonte: Glassop (2018).....7

Figura 2. Estádios de desenvolvimento do meristema apical da cultivar IACSP96-7569 submetido a tratamento fotoperiódico de indução. A. Seção longitudinal do meristema apical no estágio indutivo (coleta 7); B. Estágio de formação do primórdio da inflorescência (coleta 13).....14

Figura 3. Box Plot com os CTs médios na condição geral representando o perfil de variação de cada candidato. No gráfico é possível observar o *outlier* mínimo e máximo. O quadrante inferior da caixa representa o primeiro quartil (25%) e o superior o terceiro quartil (75%), linha horizontal na caixa representa a mediana. Gráfico desenvolvido pelo *template* cedido por vertex42.21

Figura 4. Curvas de dissociação dos candidatos a genes de referência em folha madura.21

Figura 5. Curvas de dissociação dos candidatos a genes de referência em folha imatura.....22

Figura 6. Dados de expressão relativa dos genes alvos pelo programa REST 2009. $P < 0.001$ (***), $P < 0.01$ (**), e $P < 0.05$ (*). Comparação entre a expressão gênica dos genes *PIL5* e *LHP1* (Dia curto em relação a dia longo) quanto normalizados com os pares de candidatos a normalizadores mais e menos estáveis. As barras de erro indicam erro padrão. U.A: Unidades arbitrárias.26

Figura 7. Dados em Log *fold change* de FPKM em folha madura na 7ª coleta advindos de RNA-Seq comparando a quantidade de transcritos encontrado dos genes *LHP1* e *PIL5* em tratamentos de dia curto (DC) e dia longo (DL).27

Figura 8. Dados referentes à expressão relativa de PCR quantitativo em tempo real (RT-qPCR) dos genes <i>PIL5</i> e <i>LHP1</i> em folha madura na 7 ^a coleta avaliado com REST 2009. $P < 0.001$ (***), $P < 0.01$ (**), e $P < 0.05$ (*). As barras de erro indicam erro padrão. U.A: Unidades arbitrárias.....	28
Figura 9. Curvas de dissociação dos genes alvos diferencialmente expressos em folha madura na 7 ^a coleta com os valores de eficiência (E) e coeficiente de correlação (R^2).	30
Figura 10. Curvas de dissociação dos genes alvos diferencialmente expressos em folha madura na 13 ^a coleta com os valores de eficiência (E) e coeficiente de correlação (R^2).	32
Figura 11. Curvas de dissociação dos genes alvos diferencialmente expressos em folha imatura na 7 ^a coleta com os valores de eficiência (E) e coeficiente de correlação (R^2).	34
Figura 12. Curvas de dissociação dos genes alvos diferencialmente expressos em folha imatura na 13 ^a coleta com os valores de eficiência (E) e coeficiente de correlação (R^2).	36
Figura 13. Expressão relativa do gene <i>ScFT3</i> ao longo do desenvolvimento do florescimento em folha madura avaliada com REST 2009 $P < 0.001$ (***), $P < 0.01$ (**), e $P < 0.05$ (*).	37

LISTA DE TABELAS

Tabela 1. Genes de referência com suas respectivas sequências, tamanho do amplicon (S) em pares de base (pb) e número de acesso.....	16
Tabela 2. Genes diferencialmente expressos com respectivos pares de <i>primers</i> e tamanho esperado do <i>amplicon</i> em pares de bases (pb).....	17
Tabela 3. Relação dos candidatos a genes de referência com suas respectivas melhores concentrações em μM , Eficiência (E) e Coeficiente de correlação (R^2) em pool de cDNA de folha madura e imatura.....	20
Tabela 4. Coeficiente de variação (CV) segundo algoritmo <i>BestKeeper</i> para os 8 candidatos a genes de referência, em folha madura e imatura, tempo de coleta, fotoperíodo de dia curto (DC) e dia longo (DL).....	23
Tabela 5. Dados resumidos dos algoritmos <i>NormFinder</i> e <i>BestKeeper</i> nas diferentes condições experimentais analisadas.	23
Tabela 6. Candidato a gene de referência mais e menos estável segundo dados do algoritmo <i>RefFinder</i> para as diferentes condições experimentais.	25
Tabela 7. Expressão relativa em genes alvos e transcritos de função desconhecida diferencialmente expressos em folha madura na 7 ^a coleta.	29
Tabela 8. Expressão relativa em genes alvos diferencialmente expressos em folha madura 13 ^a coleta.....	31
Tabela 9. Expressão relativa em genes alvos diferencialmente expressos em folha imatura 7 ^a coleta.....	33
Tabela 10. Expressão relativa em genes alvos diferencialmente expressos em folha imatura 13 ^a coleta.....	35

1. INTRODUÇÃO

A cana-de-açúcar é uma cultura de grande importância para o agronegócio brasileiro constituindo a matéria-prima para produção, principalmente de açúcar e etanol. Ao final da safra 2019/20 houve um aumento de 3,6% em relação à 2018/19, culminando em uma produção de 642,7 milhões de toneladas. A produção de açúcar foi de 29,8 milhões de toneladas e etanol total de cana-de-açúcar 34 bilhões de litros, um crescimento de 2,6% e 5,1%, respectivamente, quando comparado à safra anterior (Conab, 2020).

Em termos do cultivo da cana-de-açúcar, o florescimento é um processo fisiológico que prejudica diretamente o rendimento dos canaviais, uma vez que, a produção de estruturas florais redireciona os assimilados de carbono do colmo para inflorescência e resulta em perda de acúmulo de sacarose (Berding e Hurney, 2005) além de proporcionar, em algumas cultivares, a chamada isoporização que causa o secamento do interior do colmo (Capputto et al., 2007). Por outro lado, para os programas de melhoramento genético, o florescimento é de grande importância para permitir as hibridações e, conseqüentemente, o desenvolvimento de novas cultivares.

O processo do florescimento envolve várias etapas, marcadas por modificações do meristema apical, iniciando pela indução até a formação e emissão da inflorescência (Clements, 1975). A indução, mudança do estágio vegetativo para o reprodutivo é marcada por uma alteração na morfologia do meristema apical, isto é, alongamento do domo apical caracterizando-se como meristema indutivo propriamente dito (Clements, 1975; Rodrigues, 1995) e sendo esta modificação perceptível via observações em microscópio óptico de cortes histológicos do meristema apical.

O florescimento é um processo complexo dependente de vários fatores endógenos, tais como fitormônios, estímulos fotossensíveis e regulação gênica. E também fatores exógenos, tais como o fotoperíodo, temperatura, umidade, nutrição e latitude sendo que, dentre os fatores que influenciam a indução ao florescimento, o fotoperíodo é o mais importante. (Araldi et al., 2010).

Em relação aos aspectos moleculares de indução do florescimento, o gene *FLOWERING LOCUS T (FT)* descrito em *Arabidopsis thaliana* é um indutor do florescimento expresso nas células companheiras do floema foliar (Araki et al. 1998; Golembeski et al. 2015) enquanto o gene *TERMINAL FLOWER 1 (TFL1)* também descrito em *Arabidopsis*, atua de forma antagônica ao *FT* reprimindo o florescimento (Hanano & Goto, 2011). Possíveis ortólogos a estes dois genes em cana-de-açúcar (*ScFT* e *ScTFL1*) já foram identificados por Coelho et al. (2013; 2014). A expressão do gene *ScTFL1* no meristema apical, em folhas maduras e folhas imaturas (“spindle leaf”) de cana-de-açúcar foi relatado por Coelho et al. (2014). Estes autores observaram também variação na expressão do *ScFT* em condições de fotoperíodo indutivo e não indutivo em casa de vegetação, sem o uso de câmara de fotoperíodo. Em condição de fotoperíodo não indutivo foi observado a expressão do *ScFT* na folha imatura e região do meristema apical, sendo a expressão deste gene em folha madura observada também em condições de fotoperíodo indutivo.

Segundo Glassop et al. (2014;2018) a compreensão dos genes envolvidos na via de florescimento, em cana-de-açúcar, bem como de sua expressão poderá auxiliar na identificação de pontos-chaves desta via que poderão ser manipulados para assegurar o florescimento. Entretanto, a validação da expressão de genes-alvos via PCR quantitativo em tempo real (RT-qPCR) requer a escolha correta de genes de referência para normalizar a expressão dos mesmos e, consequentemente, proporcionar a sua correta interpretação de acordo com MIQE (“Minimum Information for Publication of Quantitative Real-Time PCR Experiments”) (Bustin et al., 2009). Em cana-de-açúcar, estudos da expressão de genes relacionados ao florescimento foram realizados em ensaios conduzidos em campo sob condições naturais de fotoperíodo, temperatura e umidade, em plantas amostradas no estágio de “vela”, i.e., primórdio da inflorescência, o qual é posterior ao estágio de início de indução propriamente dito (Papini-terzi et al. 2005; Medeiros et al. 2016).

O Programa de Melhoramento Genético de Cana-de-Açúcar do Instituto Agrônomo de Campinas (IAC), atualmente, é o único programa de melhoramento, no Brasil, que possui uma Câmara de Fotoperíodo, totalmente automatizada que permite induzir o florescimento em condições controladas de luz, temperatura e

umidade, simulando as condições fotoperiódicas naturais para o florescimento da cana-de-açúcar (Melloni et al., 2015). Recentemente, esta câmara de fotoperíodo foi utilizada para estudar o processo de indução fotoperiódica ao florescimento via RNA-Seq (“RNA-Sequencing”) da cultivar de cana-de-açúcar IACSP96-7569 (projeto FAPESP 2016/02679-9). Amostras foliares e os respectivos meristemas apicais desta cultivar foram coletadas ao longo de dois tratamentos fotoperiódicos indutivo e não indutivo. A dissecação do meristema apical seguida da avaliação histológica permitiu acompanhar as modificações ocorridas no meristema apical, tais como o início da indução (alongamento do domo apical) e o desenvolvimento do primórdio da inflorescência, associando as coletas foliares realizadas aos respectivos estágios de desenvolvimento. Os genes alvos diferencialmente expressos advindos do RNA-Seq, necessitam de validação de via RT-qPCR, para tanto é imprescindível uma normalização com bons genes de referência.

Neste sentido, o presente trabalho teve como objetivo identificar genes de referência que apresentam perfil de expressão estáveis em condição experimental de fotoperíodo indutivo (dia curto) e não indutivo (dia longo), em dois estádios do florescimento em cana-de-açúcar, respectivamente, no estágio inicial do processo de indução e, na formação do primórdio da inflorescência (observação da “vela”) tanto em folha madura como folha imatura (“spindle leaf”). Para tanto, foram utilizadas amostras foliares do ensaio mencionado anteriormente. Os genes de referência selecionados foram posteriormente utilizados para avaliar a expressão de dois genes alvos diferencialmente expressos via RNA-Seq envolvidos na regulação do florescimento *PIL5* (“*PHYTOCHROME INTERACTING FACTOR 3-LIKE 5*”) e *LHP1* (“*LIKE HETEROCHROMATIN PROTEIN1*”). Além disso, foram validados via RT-qPCR mais 13 genes alvos relacionados com florescimento identificados por RNA-Seq (Projeto FAPESP 2016/02679-9) bem como a avaliação da expressão relativa ao longo do tempo com um potencial gene florigeno em cana-de-açúcar *ScFT3*, expresso em folha madura.

2. REVISÃO DE LITERATURA

2.1. A cultura cana-de-açúcar

A cana-de-açúcar é uma gramínea da ordem *Poales*, família *Poaceae*, gênero *Saccharum*. É uma monocotiledônea com sistema fotossintético C4 que possui uma ampla variedade de finalidades, podendo ser utilizada desde alimentação animal, energia elétrica, produção de açúcar e biocombustíveis (Moraes et al., 2007). A cana-de-açúcar (*Saccharum* spp. hybrids) é uma cultura de grande importância econômica constituindo a matéria-prima para produção, principalmente de açúcar e etanol. De acordo com a FAO, espera-se que até 2027 um aumento em cerca de 83% na produção de açúcar por países em desenvolvimento. O Brasil continuará como principal produtor, provendo cerca de 50% da produção de açúcar mundial. No entanto haverá uma adiva competição, uma vez que a cultura também pode ser utilizada para produção de etanol. A cultura teve um grande papel no desenvolvimento do Brasil colônia e ainda o desempenha até hoje, como demonstrado pelos dados atuais da companhia nacional de abastecimento. Na década de 1970 em busca de uma alternativa a dependência de petróleo importado, foi criado o programa Proálcool dando ao Brasil o controle da tecnologia de produção de cana-de-açúcar, e desde então o país segue protagonista no cenário sucroalcooleiro (Moraes, 2007).

2.2. Fotoperíodo

O florescimento é um processo fisiológico que faz parte do ciclo vital da planta, onde tal mecanismo, naturalmente é necessário para que a espécie se perpetue por intermédio da reprodução. A adequada relação entre fatores endógenos e exógenos são necessários para que o florescimento ocorra (Araldi, 2010).

Fotoperíodo crítico é o momento onde o fotoperíodo deixa de ser não indutivo para indutivo. As plantas são subdivididas em relação ao fotoperíodo em três categorias: plantas de dias curtos, as quais florescem com regimes de fotoperíodo

abaixo do fotoperíodo crítico; plantas de dias longos, as quais florescem com regimes de fotoperíodo igual ou superior ao fotoperíodo crítico e; plantas neutras ou indiferentes, as quais florescem independentemente das condições de fotoperíodo (Jackson, 2009).

Um dos fatores mais importantes para indução ao florescimento é o fotoperíodo (Alexander, 1973), tanto que há estudos acerca desta característica iniciados há aproximadamente 100 anos atrás. Em 1920 Garner e Allard já discutiram a importante relação entre a duração do dia/noite para o florescimento. Em seu trabalho, Garner e Allard citaram Klebs que já em 1918 propôs que fatores exógenos como luz/escuro e temperatura afetavam na transição do estágio vegetativo para o reprodutivo. A classificação, hoje, tradicionalmente conhecida por plantas de dias curtos (PDC), plantas de dias longos (PDL) e plantas intermediárias, se iniciou a partir do trabalho de Garner e Allard, (1920).

O fotoperíodo tem relação direta com fitocromos, geralmente há duas formas. A forma mais instável FVE (fitocromo vermelho extremo) que é fisiologicamente ativa e FV (fitocromo vermelho) que é estável, porém fisiologicamente inativa. Em condições de luz FVE está alto e estimula o florescimento em PDL, entretanto, inibe o florescimento em PDC. Em condições de escuro, FVE está baixo e induz o florescimento em PDC e inibe em PDL. Em condições de luz o fitocromo FVE é predominante e no escuro FV. No entanto, FV não tem atividade fisiológica para induzir o florescimento, apenas FVE. (Rodrigues, 1995).

Hoje, se sabe que as melhores condições de fotoperíodo ocorrem nas regiões equatoriais em que há uma divisão de 12 horas de luz e 12 horas de escuro e não há bruscas variações na temperatura (Rodrigues, 1995). O tempo de florescimento de 16 genótipos brasileiros de cana-de-açúcar foi avaliado em três tratamentos fotoperiódicos (diminuição diária de 30s, 45s e 1 min partindo de um fotoperíodo de 12 h 55 min de luz). Nos tratamentos de 30s e 45s de decréscimo, onze genótipos emitiram a folha bandeira, no entanto, no tratamento de 1 min de decréscimo, apenas três genótipos emitiram. No tratamento com decréscimo de 45s/dia, o genótipo IACSP-7569 exibiu o menor tempo entre a emissão da folha bandeira e inflorescência, a folha bandeira foi emitida com 155 dias e a emissão da

inflorescência com 159 dias (Melloni et al., 2015). Assim sendo, este genótipo apresenta florescimento precoce e bom perfil para o florescimento.

A emissão da folha bandeira, também chamada de panícula ou flecha é uma condição particular de cana-de-açúcar, onde o surgimento desta inflorescência se dá a partir do último entrenó possuindo tamanho, cor e forma dependente de cada cultivar. Desta estrutura que possui milhares de florezinhas é que surgem as sementes de cana-de-açúcar, que só são viáveis se a planta se desenvolveu em local propício.

Em condições controladas em câmara de fotoperíodo foi verificado que o tratamento com adubação NPK influencia na emissão de panículas do florescimento em cana-de-açúcar. É utilizada comumente a técnica de submeter à planta ao déficit de NPK antes do início do tratamento fotoperiódico, uma vez que escassez de nitrogênio leva a melhora na iniciação floral. Entretanto, com novo trabalho realizado ficou evidente que é necessário readequar a técnica de cultivo do solo, uma vez que o regime de adubação contínua levou a um aumento de 21,2% na floração. Para este trabalho o fotoperíodo partiu de 12h 45 com decréscimo de 1 min ao dia até 11h 42 min (Hale et al., 2017). No entanto, o trabalho em questão ressaltou a necessidade experimentos futuro, uma vez que não há um consenso das necessidades nutricionais de cana-de-açúcar em florescimento.

2.3. Aspectos Moleculares do Florescimento

Em *Arabidopsis thaliana*, uma eudicotiledônea, foi realizada a caracterização molecular da via regulatória do florescimento (Araki et al. 1998; Golembeski et al. 2015). A regulação se inicia nas folhas com os fotorreceptores recebendo os estímulos fotossensíveis. Sob dias longos, o fator de transcrição *GIGANTEA* (*GI*), juntamente com o relógio circadiano, rege a expressão do gene *CONSTANS* (*CO*) (Song et al., 2015). A proteína *CO* estimula a expressão do gene *Flowering Locus T* (*FT*) (Puterill et. al., 1995; Song et al., 2015) que codifica a proteína *FT* nas folhas, que migra via vasos liberianos (floema) até o ápice do meristema apical (Corbesier et. al., 2007). No ápice do meristema apical, o *FT* irá interagir com a proteína *FD* que

é produzida pelo fator de transcrição da família bZIP, o *FLOWERING LOCUS D* (*FD*), a interação *FT/FD* leva a expressão dos genes *APETALA 1* (*AP1*), *LEAFY* (*LFY*) e *FRUITFULL* (*FUL*) que irá desencadear o florescimento propriamente dito com a produção dos órgãos florais (Abe et. al., 2005), como exemplificado na figura 1.

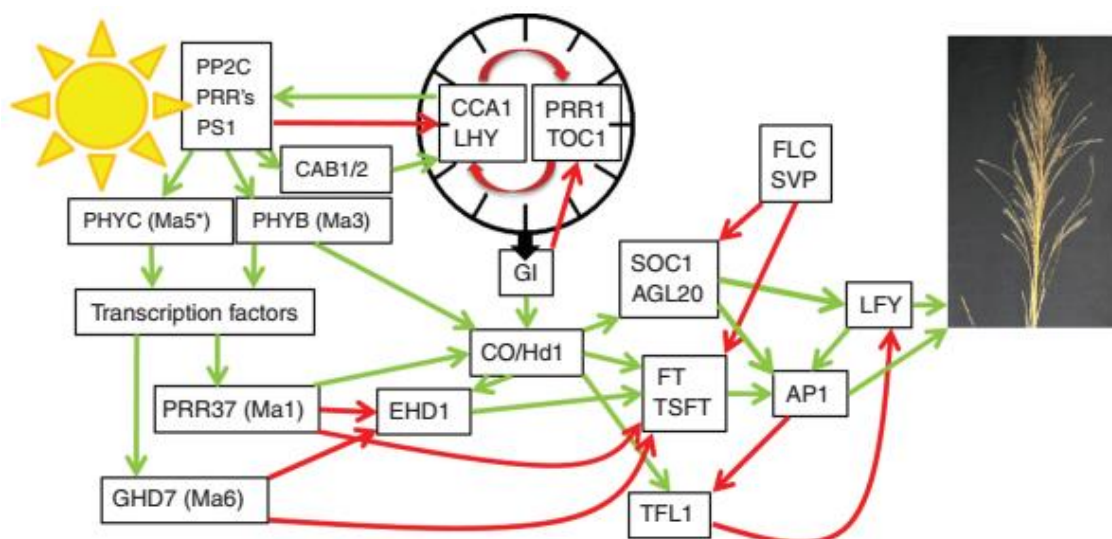


Figura 1. Esquema da via de indução do florescimento baseado em *Arabidopsis thaliana* e culturas cereais. Abaixo da figura do sol estão os genes relacionados com fotoperíodo, dentro do círculo estão os genes relacionados com relógio circadiano. As setas em verde indicam regulação positiva e em vermelha regulação negativa. Fonte: Glassop (2018).

A via de indução fotoperiódica *GI-CO-FT* é conservada tanto em dicotiledôneas como *Arabidopsis thaliana*, como em monocotiledôneas como gramíneas. No entanto, em gramíneas há uma via que não é encontrada em dicotiledôneas. Em arroz, uma planta de dia curto, os genes “Early heading date 1” (*Ehd1*) e “Grain number, plant height, e heading date 7” (*Ghd7*) atuam como promotor e repressor respectivamente no controle da expressão do florígeno *Hd3a*, ortólogo ao *FT*. *Ehd1* e *Ghd7* aparentam ser exclusivos de monocotiledôneas. (Wolabu e Tadege, 2016)

Em busca de um melhor entendimento desta rede gênica do florescimento em cana-de-açúcar, (Coelho et. al., 2013) fizeram algumas buscas e análises *in silico* no banco de dados SUCEST (“Sugarcane Expressed Sequence Tag”) e propuseram um modelo hipotético da rede gênica que controla a indução floral, onde o fator de

transcrição *GIGANTEA* regula as vias de fotoperíodo e relógio circadiano e assim, controla a passagem do estágio vegetativo para o reprodutivo por intermédio de dois mecanismos distintos, dependendo se a planta é considerada de dias curtos ou longos. Os autores também reportaram a presença de um ortólogo ao gene *FT*.

Os genes *FT/TFL1* fazem parte das subfamílias *FT-like* e *TFL1-like* que pertencem à família “Phosphatidyl Ethanolamine-Binding Proteins” (PEBPs). Em *Arabidopsis*, *FT* e *TFL1* possuem 98% de similaridade entre suas sequências de aminoácidos, porém, possuem funções antagônicas (Karlgrén et al., 2011). Os aminoácidos Tirosina e Histidina nas posições 85 e 88 respectivamente, auxiliam na diferenciação entre *FT* e *TFL1*, além disso, indicam também à determinação da proteína como indutora (*FT*) ou repressora (*TFL1*) do florescimento (Hanzawa; Money; Bradley, 2005). Neste sentido, a partir de um alinhamento de aminoácidos com a sequência do gene *ScFT3* juntamente com *FT* de *Arabidopsis thaliana* e outras gramíneas, foi observado que *ScFT3* seria um indutor, sendo assim um bom candidato de indução ao florescimento em cana-de-açúcar.

Em condições não indutivas em dias longos, a expressão do gene *ScFT1* ocorre em folha imatura e na região do ápice meristemático, porém, não em folhas maduras. Em condições indutivas de dias curtos, *ScFT1* é expresso em folhas maduras, de onde parte o sinal do florigeno para induzir o florescimento. O gene *ScTFL1* é expresso em ápice meristemático, folhas maduras, folhas imaturas e ao redor do meristema (“Spindle leaf”) (Coelho et al., 2014). Em *Arabidopsis*, com o gene *TFL1* foi feita a fusão com um ativador transcricional (VP16) e um domínio repressor (SRDX), a superexpressão TFL1-VP16 causou florescimento precoce, por outro lado, TFL1-SRDX causou atraso no florescimento. Este resultado apoia a hipótese de que *TFL1* interaja com gene *FD* no ápice do meristema apical reprimindo o florescimento (Hanano e Goto, 2011).

O gene *ScFT1* de cana-de-açúcar foi superexpresso em *Arabidopsis*, e ao contrário do esperado, causou atraso no florescimento e má formação no órgão floral, e foi inferido que *ScFT1* pode estar envolvido em atividades do meristema que regulam o tempo de florescimento e a produção de órgãos férteis, porém, estudos mais extensos são necessários para o melhor entendimento da ação de *ScFT1*.

(Coelho et al., 2014), ao que tudo indica *ScFT1* não é o ortólogo verdadeiro de *AtFT* em cana-de-açúcar.

Medeiros et al. (2016) identificaram genes diferencialmente expressos durante o florescimento, entre uma cultivar de florescimento precoce (SP81-3250) e de florescimento tardio (RB75126), em condições de campo. O ápice do meristema apical foi utilizado para avaliar os genes diferencialmente expressos. Dentre os genes avaliados, o gene *ScFT* mostrou expressão gênica similar em ambas cultivares, exceto em janeiro, em que a cultivar com florescimento precoce apresentou um aumento na abundância de transcritos de *ScFT*, enquanto na cultivar com florescimento tardio houve um decréscimo. Genes candidatos envolvidos nas vias de percepção de fotoperíodo e indução fotoperiódica, mostraram-se diferencialmente expressos em folhas madura e imatura ao longo de um ciclo de 24 horas. Também foi observado que a expressão relativa é maior em folha imatura tendendo a diminuir conforme o amadurecimento da folha indicando a variação tecido específica e temporal destes genes (Glassop et al., 2018). Ainda segundo os autores, os genes associados com relógio circadiano possuem um padrão no pico de expressão em um ciclo de 24 horas que segue a seguinte cascata: *LHY/CCA1*→*PRRs*→*TOC1*, este padrão no pico de expressão é observado em plantas de dias curtos, intermediário-curto e longo em diferentes culturas. No entanto, estes picos de expressão não podem ser sincronizados temporalmente em cana-de-açúcar, mas é interessante de se notar que ainda assim os picos de expressão ocorrem na mesma ordem.

2.4. Genes de referência em Cana-de-Açúcar

A utilização da análise quantitativa da transcrição reversa seguida da reação em cadeia da polimerase (RT-qPCR) tem se tornado o método mais comum e confiável em estudos de expressão gênica. Há basicamente dois métodos para quantificar a expressão gênica: absoluta e relativa. A quantificação absoluta calcula o número de cópias e usualmente relaciona o número de cópias a uma curva padrão. A quantificação relativa utiliza um controle endógeno para mensurar a

expressão relativa do gene de interesse, um destes métodos é o $\Delta\Delta CT$ (Schmittgen e Livak, 2001)

Entretanto, para utilizar a expressão relativa é necessária uma adequada normalização dos dados com um bom gene de referência, o qual deve apresentar uma expressão constante em diferentes tecidos, órgãos e estágios de desenvolvimento, independente das condições experimentais (Jain et al., 2006). No entanto, nenhum gene tem uma condição estável em todas as condições experimentais. Portanto, para a análise de expressão gênica é necessário validar genes de referência nas condições específicas na qual o experimento foi realizado antes de seu uso como gene de referência (ou também comumente chamado de normalizador), para permitir a estabilidade da expressão gênica (Bustin et al., 2000; Jain et al., 2006).

Até o presente momento, não foram relatados genes de referência avaliados em condições específicas de indução artificial de florescimento em cana-de-açúcar. No geral, poucos trabalhos foram conduzidos com a finalidade de definir genes de referência para esta cultura. Iskandar et al. (2004) foram um dos primeiros a propor algo neste sentido, ainda que sem a utilização de algoritmos estatísticos específicos, testaram para normalização os genes *GLICERALDEÍDO-3-FOSFATO DESIDROGENASE (GAPDH)*, *B-ACTINA (ACT)*, *B-TUBULINA (TUB)* e *RNA RIBOSSOMAL SUBUNIDADE 25 (25SrRNA)* em diferentes tecidos, cultivares e espécies de cana-de-açúcar. Dentre estes genes avaliados, o *GAPDH* mostrou-se como a melhor opção em estudos de expressão gênica relacionada com genes de acúmulo de sacarose. O *GAPDH* também foi apontado como o mais estável, com o auxílio da ferramenta *NormFinder*, para análise de expressão relativa de genes envolvidos na via metabólica da parede celular, em cultivares de cana-de-açúcar (Bottcher et al., 2013).

Uma extensa validação de genes de referência, envolvendo 13 genes candidatos avaliados em 5 cultivares de cana-de-açúcar, em diferentes tecidos e tratamentos foi realizada aplicando-se algoritmos estatísticos específicos: *NormFinder*, *BestKeeper*, *delta CT* e *RefFinder*. Os genes validados foram principalmente voltados para estresse abiótico. Constatou-se que para as condições de salinidade e tolerância a seca, os genes *GAPDH* e *FATOR DE ELONGAÇÃO*

EUCARIÓTICO 1A (eEF1-a) foram os melhores (Ling et al., 2014). Posteriormente, a combinação dos genes de referência *COMPLEXO ADAPTADOR DE CLATRINA* (CAC) e *CULINA* (CUL) foi validada como a melhor para estudos de estresse abiótico (Guo et al., 2014). Para estudos sobre estresse ao déficit hídrico, em raízes de cana-de-açúcar, ficou evidenciado pelos algoritmos *GeNorm*, *NormFinder* e *BestKeeper* que os genes mais estáveis foram *HISTONA* (H1), *TUB*, e *GAPDH* (Silva et al., 2014). Andrade et al., (2017) analisaram através das ferramentas *NormFinder*, *BestKeeper* e *RefFinder*, 10 candidatos a genes de referência em 2 tecidos de 2 cultivares de cana-de-açúcar submetidas a tratamentos de tolerância a seca, apontando os genes *GADPH* e *UBQ1* como ideais para os estudos de expressão envolvendo déficit hídrico em folhas e raízes. Os genes de referência *URIDILATO QUINASE* (UK) e *ENZIMAS DE CONJUGAÇÃO UBIQUITINA 18* (UBC18) foram definidos como os mais estáveis para normalizar dados de expressão relativa de transcritos diferencialmente expressos em condição de resposta da cana-de-açúcar ao SCMV (“*Sugarcane mosaic virus*”) pelos algoritmos *GeNorm*, *NormFinder*, *BestKeeper* e *RefFinder* (da Silva et al., 2019). Em um estudo de florescimento em cana-de-açúcar visando à identificação de genes diferencialmente expressos via biblioteca subtrativa, em cultivares contrastantes (florescimento precoce x tardio) em condições equatoriais foram identificados genes diferencialmente expressos validados por PCR quantitativo em tempo real (RT-qPCR). Dos dois genes normalizadores avaliados (*UBQ* e *EF1 α*), os valores de Ct (“Cycle Threshold”) mostraram-se consistentes e o gene *EF1 α* foi escolhido para normalizar os genes de interesse (Medeiros et al., 2016). Os genes de referência *GAPDH* e *ACT* foram utilizados como normalizadores na avaliação da expressão dos genes *ScFT* e *ScTFL1*, um indutor e um repressor do florescimento respectivamente, em cana-de-açúcar (Coelho et al., 2014).

De acordo com Vandesompele (2002), a escolha de um gene de referência inapropriado pode levar a erros nas interpretações de resultados e conclusões. Além disso, em alguns casos é necessária a utilização de mais de um gene de referência para uma maior confiabilidade dos resultados, sendo recomendada a utilização da média geométrica de dois genes mais estáveis (Vandesompele, 2002).

Atualmente, há vários programas disponíveis para a escolha de um bom normalizador. O *geNorm* (Vandesompele et al., 2002) trabalha com o “Valor M”, que indica estabilidade, onde elimina-se o gene de referência com maior valor de M e, recalcula o valor de M com os genes restantes, os pareando dois a dois até restar apenas dois genes. Os genes de referência com valores de M menor que 1,5 ($M < 1,5$) são considerados os mais estáveis.

O programa *NormFinder* (Andersen et al., 2004; Sandercock et al., 2017) fornece um valor de estabilidade para cada gene, valores de estabilidade mais baixos indicam menor estabilidade. Este programa compara a expressão relativa dos pares de genes para cada amostra em um experimento, e ranqueia o gene de referência por sua estabilidade de acordo com a repetibilidade da diferença de expressão entre todas as amostras.

O *BestKeeper* (Pfaffl et al., 2004) é um programa que também busca inferir a estabilidade da expressão dos genes de referência testados. Os genes com as melhores médias geométricas são combinados em um índice, e a correlação entre cada gene de referência candidato e o índice é calculado. A relação entre os genes de referência candidatos e o índice é avaliada com a correlação de Pearson (r), coeficiente de determinação (r^2) e o p -valor. O *RefFinder* (Xie et al., 2012) é uma plataforma *online* que tem inserida outros algoritmos (*geNorm*, *BestKeeper*, *NormFinder*) e que gera um ranking compreensivo que ranqueia, de maneira geral, os melhores genes de referência.

Outra abordagem é o ΔCT (Silver et al., 2006), um método similar ao adotado por Vandesompele (2002). A estabilidade dos genes de referência é ranqueada de acordo com a diferença da repetibilidade da expressão gênica dentre as amostras. O método ΔCT pode ser uma alternativa, diferentemente dos outros algoritmos que necessitam de múltiplos genes de referência, o que dificulta estudos onde a quantidade de mRNA é mínima.

3. MATERIAIS E MÉTODOS

3.1. Material Biológico e Condições Experimentais:

A cultivar de cana-de-açúcar (*Saccharum* spp.) IACSP96-7569 possui floração precoce e boa resposta ao fotoperiodismo em Câmara de Fotoperíodo (Melloni, 2015). Gemas da cultivar IACSP96-7569 foram plantadas em bandejas com substrato (Plantmax®) e mantidas em casa de vegetação. Após 60 dias foram transferidas para citropotes de 3,8 litros (diâmetro superior 15,3 cm, diâmetro inferior 5 cm e altura 35,3 cm) contendo areia, terra e substrato na mesma proporção e colocados nos vagões da câmara de fotoperíodo. Aos nove meses de idade, as plantas foram submetidas a dois tratamentos fotoperiódicos: 12 horas e 55 minutos com um decréscimo de 45 segundos por dia (dia curto- DC) e fotoperíodo constante de 13 horas (dia longo-DL) na câmara de fotoperíodo automatizada no “Centro de Cana” - IAC, localizado em Ribeirão Preto - SP, Brasil (21°12'34"S 47°52'29"W).

Semanalmente, foram coletados o meristema apical, a folha +1 (folha madura) e a folha imatura de três plantas (réplicas biológicas) de cada tratamento fotoperiódico. A folha +1 madura e a folha imatura foram imediatamente congeladas em nitrogênio líquido e armazenadas à -80°C até a extração de RNA Total. O meristema apical das respectivas plantas amostradas foi coletado e fixado em FAA 50% (formalina, ácido acético e álcool etílico) para a realização de cortes histológicos em micrótomo e confirmação do estágio de desenvolvimento (Figura 2). Após análise histológica dos meristemas apicais, foram selecionadas a sétima coleta (início da modificação do meristema apical) e décima terceira coleta (formação do primórdio floral) realizada a partir do início dos tratamentos fotoperiódicos.

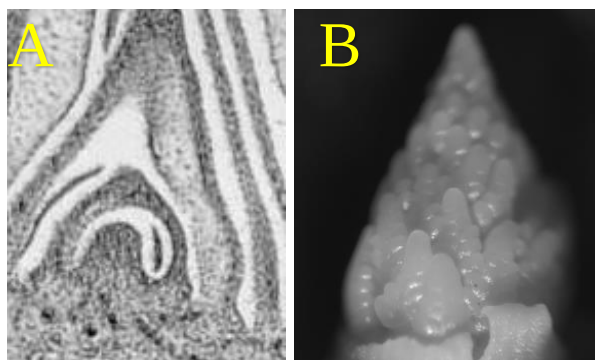


Figura 2. Estádios de desenvolvimento do meristema apical da cultivar IACSP96-7569 submetido a tratamento fotoperiódico de indução. A. Seção longitudinal do meristema apical no estágio indutivo (coleta 7); B. Estágio de formação do primórdio da inflorescência (coleta 13).

3.2. Extração de RNA total e Síntese de cDNA

A extração de RNA total foi realizada com o Kit PureLink RNA Mini Kit (ThermoFisher), quantificado com o NanoDrop 2000 (Thermo Fisher Scientific, Wilmington DE, EUA). A integridade do RNA foi verificada em gel de agarose 1,0% corado com brometo de etídio ($1 \mu\text{g mL}^{-1}$) em eletroforese (3 a 5mV/cm). A síntese de cDNA foi realizada com o kit QuantiNova Reverse Transcription (QIAGEN Strasse 1, 40724 Hilden, ALEMANHA). Seguindo as instruções do fabricante, todo o DNA genômico foi eliminado com *gDNA wipeout* (Incluso no kit) e em seguida, $1 \mu\text{g}$ de RNA foi definido como padrão para a síntese de cDNA utilizando oligo-dT e *random primers* (QuantiNova Reverse Transcription Mix) com ciclagem (42°C por 15 min e 95°C por 3 min). O cDNA então foi diluído 20 vezes e utilizado para os ensaios de RT-qPCR.

3.3. Seleção e Validação de Genes de Referência

Foram avaliados oito genes candidatos a genes de referência descritos em cana-de-açúcar: *Ubiquitina 1* (UBQ1), *Ubiquitina 2* (UBQ2), *Proteína ribossômica 60S L35-4* (RPL), *Gliceraldeído-3 fosfato desidrogenase* (GAPDH), *RNA ribossômico 25S* (25SrRNA1) e *Tubulina* (TUB) (Andrade et al., 2017); *Fator de alongamento 1-alpha* (EF1- α) (Medeiros et al., 2016) e *Proteína intrínseca tonoplástica* (TIPS-41)

(Ling et al., 2014) (Tabela 1). Todos os genes de referência foram escolhidos por já terem histórico na literatura de ter sido testado como normalizadores ou já foram usados em trabalhos relacionados com florescimento.

O cálculo da eficiência dos *primers* foi realizado em *pool* de tecidos (folha madura x folha imatura) e estimado por intermédio da fluorescência (ΔR_n) extraída do RT-qPCR pelo programa LinRegPCR 7.5 (Ruijter et al. 2009) Todos os genes de referência foram avaliados em três concentrações de *primers* (0,25; 0,4; 0,8 μ M) para a etapa de otimização dos *primers*, sendo a melhor concentração estabelecida pela análise da curva de dissociação aliada a eficiência dos *primers*. Posteriormente, todos os genes de referência foram testados em *pool* de cDNA (1:20) em sete condições experimentais. As condições foram por tecido (folha madura x folha imatura); estágio de desenvolvimento (início modificação do meristema apical correspondendo à sétima coleta e início da formação do primórdio floral correspondendo à décima terceira coleta); por tratamento (DC vs DL).

Os genes de referência foram testados via RT-qPCR e os dados de *Cycle Threshold* (Ct) e Eficiência (E) foram analisados por três algoritmos estatísticos visando a definição dos genes de referência mais estáveis: *NormFinder*, *BestKeeper*, além do algoritmo online *RefFinder* (<http://www.heartcure.com.au/reffinder/?type=reference>). Posteriormente, os genes de referência mais estáveis e menos estáveis foram validados ao avaliar a expressão relativa dos genes alvos *PIL5* e *LHP1*. Adicionalmente, para estes genes foi realizado um comparativo entre os dados de FPKM provenientes de RNA-Seq e dados da expressão relativa (RT-qPCR), somente na condição onde foi realizado o sequenciamento em larga escala (folha madura na 7ª coleta).

Tabela 1. Genes de referência com suas respectivas sequências, tamanho do *amplicon* em pares de base (pb) e número de acesso.

Gene (símbolo)	Nome do gene	Sequência do <i>primer</i> (5' - 3')	S (pb)	Referência	Número de acesso
<i>GAPDH</i>	<i>Gliceraldeído-3 fosfato desidrogenase</i>	F: TTGGTTTCCACTGACTTCGTT R: CTGTAGCCCCACTCGTTGT	122	Andrade et al. (2017)	CA254672
<i>TUB</i>	<i>Tubulina</i>	F: CTCCACATTCATCGGCAACTC R: TCCTCCTCTTCTTCCTCCTCG	237	Andrade et al. (2017)	CA222437
<i>UBQ1</i>	<i>Ubiquitina 1</i>	F: AGCCTCAGACCAGATTCCAA R: AATCGCTGTGGAACACTTGC	110	Andrade et al. (2017)	CA094944
<i>UBQ2</i>	<i>Ubiquitina 2</i>	F: CTTCTTCTGTCCCTCCGATG R: TCCAACCAAACCTGCTGCTC	158	Andrade et al. (2017)	CA093560
<i>RPL</i>	<i>Proteína ribossomal 60S L35-4</i>	F: CTGAAGACGGAGAGGGAAAA R: GGCGAAGAGAACTAACAC	264	Andrade et al. (2017)	CA127053
<i>25SrRNA1</i>	<i>Subunidade Ribossomal 25S</i>	F: ATAACCGCATCAGGTCTCCAAG R: CCTCAGAGCCAATCCTTTTCC	110	Andrade et al. (2017)	CA171131/ CO373883
<i>EF1α</i>	<i>Fator de elongação-1α</i>	F: AAGGCCCGTTATGATGAGTTGTG R: CAAAACCAGAGATTGGGACGAAAG	N/A	Medeiros et al. (2016)	AF331850
<i>TIPS-41</i>	<i>Proteína tonoplástica intrínseca</i>	F: CACCTGTTGAGGTTCTGCT R: CACAGCATCACTCCCACAGT	116	Ling et al. (2014)	CA228782.1

3.4. Desenvolvimento de *primers* para genes diferencialmente expressos

A partir do ensaio de RNA-Seq realizado (Artigo em preparação) em folha madura, na sétima coleta com meristema em transição de vegetativo para o reprodutivo foram obtidas sequências de genes diferencialmente expressos (Projeto FAPESP 2016/02679-9). O RNA-Seq utilizou o transcriptoma de cana-de-açúcar como referência (Vicentini et al., 2015) para mapear os transcritos. Com estas sequências (que se encontram disponíveis no final apêndice A) e o auxílio da ferramenta online “PrimerQuestTool”

(<https://www.idtdna.com/PrimerQuest/Home/Index>) foram desenhados os pares de *primers* para transcritos diferencialmente expressos: *Constans-Like 9* (COL9), “*Pseudo response regulador 1*” (PRR1) “*Late elongated hypocotyl*” (LHY), “*Phytochrome interacting factor 3-like 5*” (PIL5), *Agamous-Like 12* (AGL12), EID1-Like 1 (EID1), “*Cyclic dof factor 2*” (CDF2), “*Cyclic dof factor 3*” (CDF3), “*Pseudo response regulador 5*” (PRR5), *Agamous-Like 7* (AGL7) e “*Pseudo response regulador 7*” (PRR7) além dos transcritos desconhecidos 127290, 110622 e 177871 (Tabela 2).

O desenvolvimento dos pares de *primers* evitaram domínios conservados e usou as regiões 5' e 3'UTR como referência. Foi adotado como parâmetros comprimento do *primer* variando de 17 a 22 pb, conteúdo de GC definido entre 35 a 65% e T_m (temperatura de *melting*) entre 59° e 65°C, com um *amplicon* variando entre 100 e 250 pb. A qualidade dos pares de *primers* foi verificada *in silico* pelo NetPrimer (www.premierbiosoft.com/netprimer), uma ferramenta gratuita em Web para analisar a qualidade dos *primers* quanto a formação de estruturas secundárias. A eficiência dos pares de *primers* também foi realizada pelo programa LinRegPCR 7.5 (RUIJTER, et al. 2009).

A diluição de cDNA a 1:20 e a concentração do par de *primer* à 0,25µM foram utilizadas para os ensaios de RT-qPCR realizados em triplicata técnica e biológica. Os dados de expressão relativa e análise estatística (teste de hipótese P (H₁)) foram realizados usando o software REST 2009 com 2000 iterações (Pfaffl, et al., 2002) e as diferenças foram consideradas significativas quando P < 0,001 (***), P < 0,01 (**), e P < 0,05 (*).

Tabela 2. Genes diferencialmente expressos com respectivos pares de *primers* e tamanho esperado do *amplicon* em pares de bases (pb).

Gene (símbolo)	Nome do gene	Sequência do <i>primer</i> (5' - 3')	Tamanho do <i>amplicon</i> (pb)
-------------------	--------------	--------------------------------------	--

COL9	<i>Constans-like 9</i>	F: CACTACTCTCACCAGTCAAACC R: CTGCAAATAGCAATGCGGTATC	141
PRR1	<i>Pseudo response regulator 1</i>	F: CTCAAGCACATACACCACCA R: ATGCCGATGACCACACATT	153
LHY	<i>Late elongated hypocotyl</i>	F: GTGTCTCTCCACACAGAGTTAAA R: TTGTCCGCATCTACATCACTAC	161
PIL5	<i>phytochrome interacting factor 3-like 5</i>	F: GGTCTGGCTCTGCATGAC R: AACCAGTTCGTCCCTGATTG	117
AGL12	<i>Agamous like-12</i>	F: GAGATGGGCTATTCCTTCTGAC R: CTCCTGAAGGGCTATGGTTTAT	149
EID1	<i>EID1-like 1</i>	F: TTCTGAGGACACAAAGGAAGAG R: CAAAGAGAAAGGCAGCTAGGA	166
CDF2	<i>Cyclic dof factor 2</i>	CTGTGATGGTGCCAGGTAAA GCACAAGTGGGTATGGAAATG	147
CDF3	<i>Cyclic dof factor 3</i>	F: TCAGGTTTCGACTGGAATGG R: AAGGAGATGAGAAGGCAGAAAG	151
PRR5	<i>Pseudo response regulator 5</i>	F: ACAGAAGCAGAACTGACTCG R: CCTTCAGTCTTACCAGTCCAAT	102
AGL7	<i>Agamous like-7</i>	F: GACGGTTCAGGCTCAGATT R: GCTTTAATGAACGCACACCTC	152
ScFT3*	<i>Flowerig Locus T 3</i>	-	93
PPR7	<i>Pseudo response regulator 7</i>	F: CAGTGGCAGTGGAAGTGAAA R: CATTGAGTCCGACACTGAAGTC	149
127290	<i>Transcrito desconhecido</i>	F: CGTTAGAGGGTGCTCAAGTT R: CTCAGCAGTTCTCTCTGTCTTT	175
110622	<i>Transcrito desconhecido</i>	F: GGAAGGCTAGAACAAGAGAATCA R: CAGCATGAGAAGAGCCAGTAG	169
177871	<i>Transcrito desconhecido</i>	F: GTCTCTCAATCTGCTGCCTATAC R: GAAGCTGGAGTTGTTGAGGT	196

(*) A sequência do *primer ScFT3* está sob sigilo, artigo em preparação.

3.5. Estudo de expressão relativa do gene *ScFT3* ao longo do tempo

A sequência do par de *primers* do *ScFT3* foi gentilmente cedida pelo Prof. Dr. Elisson Romanel da Escola de Engenharia de Lorena, Universidade de São Paulo (EEL/USP). A expressão relativa do gene *ScFT3* foi avaliada ao longo de 7 pontos

de coleta após o início dos tratamentos fotoperiódicos: 2º coleta (9 dias); 3º coleta (16 dias); 4º coleta (23 dias); 6º coleta (39 dias); 7º coleta (46 dias); 9º coleta (60 dias); e 13º coleta (88 dias). A extração do RNA total e síntese de cDNA foram realizados como descritos anteriormente (3.2). A expressão deste gene foi avaliada considerando três réplicas biológicas para cada tratamento (dia curto e dia longo). O RT-qPCR foi realizado em triplicata técnica e normalizado com genes de referência estáveis (*UBQ1/TUB*) previamente definidos. Os dados de expressão relativa foram calculados com programa REST 2009 e as diferenças foram consideradas significativas quando $P < 0,001$ (***), $P < 0,01$ (**), e $P < 0,05$ (*). Para a análise, o valor da expressão relativa de cada tempo de coleta foi plotado em um gráfico de linhas.

3.6. Condições da Reação de PCR quantitativa – PCR quantitativo em tempo real (RT-qPCR) :

Os ensaios de RT-qPCR foram realizados em um termociclador Bio-Rad IQ5 com reagente GoTaq® qPCR Master Mix Kit (Promega, EUA). A reação foi conduzida em um volume final de 10 μ L contendo 5 μ L de SYBR Green 2x, 3 μ L de cDNA (1:20) e os pares de *primers* com as concentrações à 0,25 μ M. Foram utilizados três controles negativos (branco) para cada conjunto de pares de *primers*. As condições de amplificação foram: 95°C por 3 min, seguido de 40 ciclos (95°C por 10s e 60°C por 30s), seguido pela curva de dissociação (de 55°C à 95°C). Para os genes de referência a etapa de eficiência e otimização foram realizadas em duplicata técnica, enquanto a avaliação das amostras individuais nas condições experimentais, e a validação foram realizadas em triplicata técnica.

4. RESULTADOS

4.1. Dados de PCR quantitativo em tempo real (RT-qPCR) dos candidatos a genes de referência

De maneira geral, dentre todas as condições, o gene mais expresso foi *25SrRNA1* com Ct mínimo de 13,56 e obteve também a menor média geométrica mínima de 19,42 ($\pm 2,185$). O gene menos expresso foi *EF1* com Ct máximo de 38,31, porém *UBQ1* foi quem obteve uma média geométrica máxima maior, de 34,06($\pm 1,768$), os perfis de variação estão representados na figura 3. Na etapa de otimização, os genes candidatos (GC) *UBQ1* e *TUB* tiveram sua curva de dissociação com menos ruídos de dímeros de *primers* na concentração de 0,25 μM ; *UBQ2*, *GAPDH* e *TIPS-41* na concentração de 0,4 μM e *RPL*, *25SrRNA1* e *EF1* na concentração de 0,8 μM (Tabela 3). A eficiência e coeficiente de correlação (R^2) dos pares de *primers* em folha madura variaram de 1,76 à 2,32 (E) e 0,97 à 0,99 (R^2); em folha imatura 1,94 à 2,26(E) e 0,99 (R^2) (Tabela 3).

Tabela 3. Relação dos candidatos a genes de referência com suas respectivas melhores concentrações em μM , Eficiência (E) e Coeficiente de correlação (R^2) em pool de cDNA de folha madura e imatura.

Gene	Folha madura			Folha imatura		
	μM	E*	R^2	μM	E*	R^2
<i>TUB</i>	0,25	2,136	0,979	0,25	1,990	0,999
<i>UBQ1</i>	0,25	1,934	0,995	0,25	2,000	0,999
<i>TIPS-41</i>	0,40	1,765	0,999	0,40	2,034	0,999
<i>UBQ2</i>	0,40	2,147	0,997	0,40	2,075	0,997
<i>GAPDH</i>	0,40	2,323	0,999	0,40	2,264	0,995
<i>25SrRNA1</i>	0,80	2,075	0,989	0,80	2,190	0,998
<i>RPL</i>	0,80	2,174	0,999	0,80	2,048	0,999
<i>EF1</i>	0,80	2,220	0,999	0,80	1,949	0,999

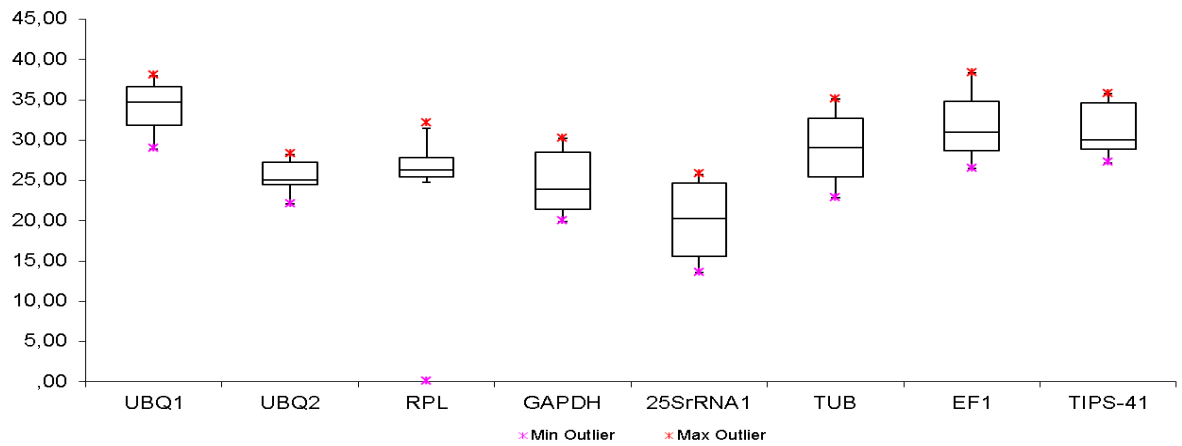


Figura 3. Box Plot com os CTs médios na condição geral representando o perfil de variação de cada candidato. No gráfico é possível observar o *outlier* mínimo e máximo. O quadrante inferior do caixa representa o primeiro quartil (25%) e o superior o terceiro quartil (75%), linha horizontal na caixa representa a mediana. Gráfico desenvolvido pelo *template* cedido por vertex42.

Todos os candidatos a gene de referência tanto em folha madura quanto em folha imatura apresentaram um pico único na curva de dissociação (Figuras 4 e 5), indicando a formação de um único produto durante o PCR quantitativo.

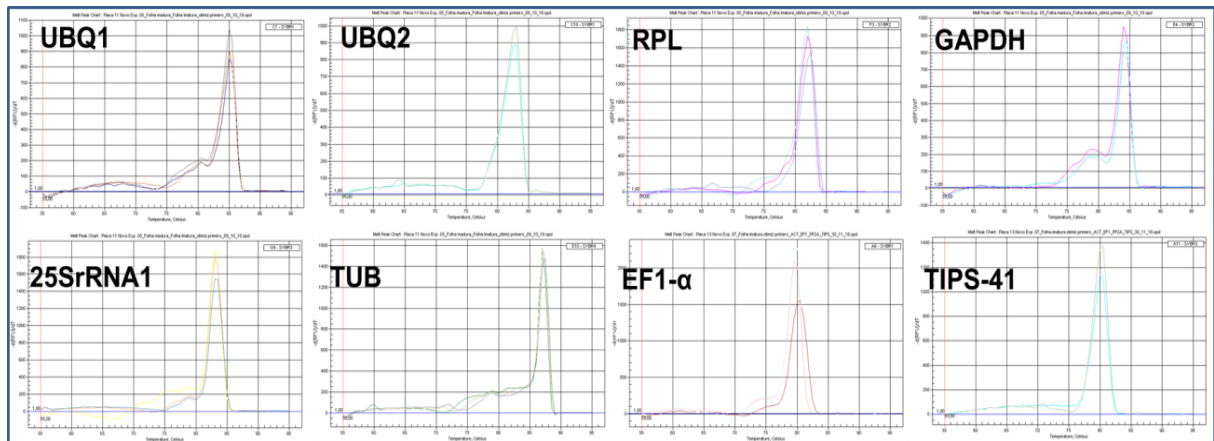


Figura 4. Curvas de dissociação dos candidatos a genes de referência em folha madura.

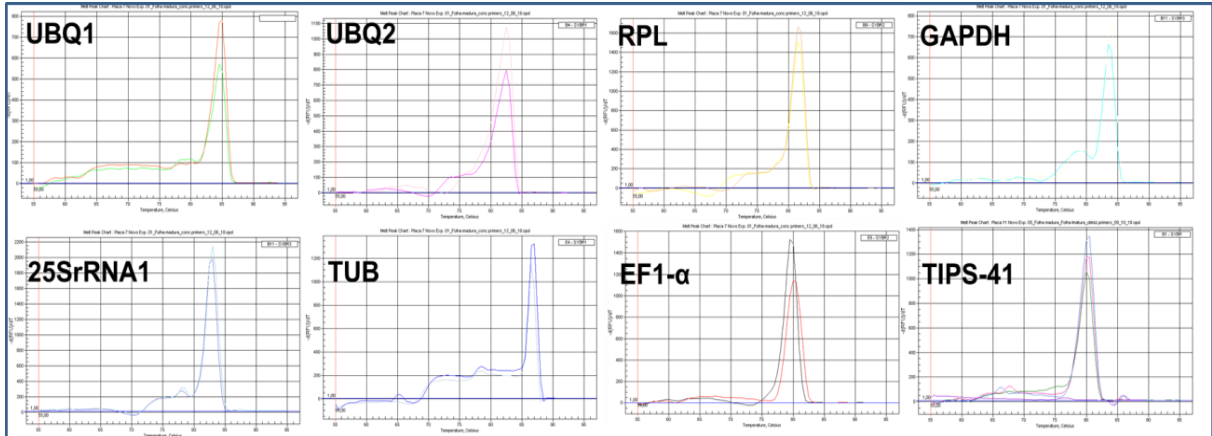


Figura 5. Curvas de dissociação dos candidatos a genes de referência em folha imatura.

4.2. Análise de estabilidade dos candidatos a genes de referência por condição experimental

O Coeficiente de Variação (CV) pode ser considerado como a primeira estimativa da estabilidade da expressão dos genes de referência. O CV determinado com auxílio do *BestKeeper* indicou o *UBQ2* como o gene mais estável para as condições de folha imatura (CV=6,18) e dia curto (CV=1,56). Por outro lado, os genes com a menor variação foram *RPL* na folha madura (CV=3,49), *EF1* na 7ª coleta (CV=3,00), *25SrRNA1* na 13ª coleta (CV=1,48) e *UBQ1* em dia longo (CV=5,18). Os genes com a maior variação foram *25SrRNA1* em folha madura (CV=22,60), folha imatura (CV=24,07), dia curto (CV=25,55) e dia longo (CV=21,13). Na 7ª e 13ª coletas o gene *GAPDH* foi o menos estável com valor de CV de 13,64 e 8,44, respectivamente (Tabela 4).

Tabela 4. Coeficiente de variação (CV) segundo algoritmo *BestKeeper* para os 8 candidatos a genes de referência, em folha madura e imatura, tempo de coleta, fotoperíodo de dia curto (DC) e dia longo (DL).

Rank	Folha madura		Folha imatura		7ª coleta		13ª coleta		DC		DL	
	Gene	CV	Gene	CV	Gene	CV	Gene	CV	Gene	CV	Gene	CV
1	<i>RPL</i>	3,49	<i>UBQ2</i>	6,18	<i>EF1</i>	3	<i>25SrRNA1</i>	1,48	<i>UBQ2</i>	1,56	<i>UBQ1</i>	5,18
2	<i>UBQ2</i>	4,7	<i>UBQ1</i>	8,14	<i>TIPS-41</i>	3,4	<i>UBQ1</i>	1,78	<i>RPL</i>	3,5	<i>RPL</i>	6,44
3	<i>TUB</i>	5,41	<i>RPL</i>	8,32	<i>RPL</i>	3,64	<i>EF1</i>	4,52	<i>TIPS-41</i>	7,3	<i>UBQ2</i>	7,6
4	<i>TIPS-41</i>	6,2	<i>EF1</i>	9,72	<i>UBQ2</i>	5,22	<i>UBQ2</i>	4,76	<i>TUB</i>	7,52	<i>TIPS-41</i>	8,41
5	<i>UBQ1</i>	6,37	<i>TIPS-41</i>	9,91	<i>UBQ1</i>	5,23	<i>TIPS-41</i>	5,95	<i>UBQ1</i>	9,34	<i>EF1</i>	10,62
6	<i>GAPDH</i>	11,34	<i>GAPDH</i>	13,27	<i>25SrRNA1</i>	5,71	<i>TUB</i>	7,19	<i>GAPDH</i>	10,05	<i>TUB</i>	13,55
7	<i>EF1</i>	11,76	<i>TUB</i>	13,59	<i>TUB</i>	10,12	<i>RPL</i>	7,54	<i>EF1</i>	11,17	<i>GAPDH</i>	14,8
8	<i>25SrRNA1</i>	22,6	<i>25SrRNA1</i>	24,07	<i>GAPDH</i>	13,64	<i>GAPDH</i>	8,44	<i>25SrRNA1</i>	25,55	<i>25SrRNA1</i>	21,13

O algoritmo do *BestKeeper* estima a estabilidade dos genes de referência de acordo com o desvio padrão dos dados brutos de CT. Segundo Pfaffl (2004), quando o desvio padrão é maior que 1, indica inconsistência na expressão do candidato a gene de referência. A correlação de Pearson é, no entanto, um melhor parâmetro que o desvio padrão para avaliar a estabilidade dos GC, uma vez que utiliza o índice do *BestKeeper* e realiza correlação de Pearson de cada GC com índice e assim tem mais robustez pra indicar qual melhor gene de referência do que o desvio padrão (Spiegelaere, et al., 2015). Por esta perspectiva, os genes mais estáveis foram *UBQ2* em folha imatura e dia longo (DL). O gene *RPL* foi o mais estável nas condições de folha madura, 7ª coleta, e dia curto (DC), porém, foi o menos estável em folha imatura. *EF1* foi o mais estável na 13ª coleta, enquanto *GAPDH* o menos estável em folha madura e DC (Tabela 5).

Tabela 5. Dados resumidos dos algoritmos *NormFinder* e *BestKeeper* nas diferentes condições experimentais analisadas.

Algoritmo	<i>NormFinder</i>		<i>BestKeeper</i>			
	Gene	Estabilidade	Gene	Desvio padrão	coef, de corr, [r]	p-value
Condição Experimental	Folha madura					
Melhor par	<i>RPL/EF1</i>	0,02	N,A	N,A	N,A	N,A
Gene mais estável	<i>TUB</i>	0,02	<i>RPL</i>	0,94	0,88	0,004
Gene menos estável	<i>25SrRNA1</i>	0,18	<i>GAPDH</i>	2,91	0,18	0,669
	Folha imatura					

Melhor Par	<i>UBQ1/GAPDH</i>	0,01	N,A	N,A	N,A	N,A
Gene mais estável	<i>TIPS-41</i>	0,04	<i>UBQ2</i>	1,53	0,92	0,001
Gene menos estável	<i>25SrRNA1</i>	0,15	<i>RPL</i>	2,25	0,82	0,013
Sétima coleta						
Melhor Par	<i>UBQ1/TUB</i>	0,02	N,A	N,A	N,A	N,A
Gene mais estável	<i>RPL</i>	0,01	<i>RPL</i>	0,96	0,99	0,001
Gene menos estável	<i>GAPDH</i>	0,09	<i>UBQ2</i>	1,27	0,34	0,418
Décima terceira coleta						
Melhor par	<i>RPL/EF1</i>	0,02	N,A	N,A	N,A	N,A
Gene mais estável	<i>UBQ2</i>	0,02	<i>EF1</i>	1,59	0,91	0,002
Gene menos estável	<i>25SrRNA1</i>	0,06	<i>UBQ1</i>	0,65	-0,76	0,03
Dia curto (DC)						
Melhor par	<i>RPL/EF1</i>	0,03	N,A	N,A	N,A	N,A
Gene mais estável	<i>UBQ1</i>	0,03	<i>RPL</i>	0,91	0,90	0,002
Gene menos estável	<i>25SrRNA1</i>	0,20	<i>GAPDH</i>	2,48	0,35	0,393
Dia longo (DL)						
Melhor par	<i>UBQ1/TUB</i>	0,02	N,A	N,A	N,A	N,A
Gene mais estável	<i>TIPS-41</i>	0,05	<i>UBQ2</i>	1,97	0,89	0,003
Gene menos estável	<i>25SrRNA1</i>	0,12	<i>25SrRNA1</i>	4,25	0,98	0,001

N.A: Não se aplica para este algoritmo. A tabela completa pode ser encontrada no material suplementar apêndice A, tabela Suplementar 1A.

Entretanto, segundo o algoritmo do *NormFinder*, *TIPS-41* foi o gene mais estável para duas condições experimentais (folha imatura e DL), enquanto *TUB*, *RPL*, *UBQ2* e *UBQ1* foram mais estáveis, respectivamente, em folha madura, 7º e 13º coleta e DC. O gene menos estável foi *25SrRNA1* para seis das sete condições experimentais, ao passo que, *GAPDH* foi para a 7º coleta (Tabela 5).

O algoritmo do *NormFinder* também provê uma combinação dos dois melhores genes de referência que podem ser usados como a melhor combinação de genes mais estáveis. *UBQ1/TUB* foi a combinação mais estável para a 7º coleta e DL, *RPL/EF1* em folha madura, 13º coleta e DC (Tabela 5).

De acordo com ranking compreensivo do *RefFinder*, *UBQ2* foi o gene de referência mais estável na 13º coleta, enquanto *TIPS-41* para folha imatura e DL. Já *EF1*, *TUB* e *UBQ1* para 7º coleta, folha madura, e DC, respectivamente (Tabela 6).

Tabela 6. Candidato a gene de referência mais e menos estável segundo dados do algoritmo *RefFinder* para as diferentes condições experimentais.

Algoritmo	<i>RefFinder</i>					
	<i>NormFinder</i>		<i>BestKeeper</i>		Ranking	
	Gene	Estabilidade	Gene	Estabilidade	Gene	Estabilidade
Folha Madura						
Gene mais estável	<i>TUB</i>	0,623	<i>RPL</i>	0,875	<i>TUB</i>	1,316
Gene menos estável	<i>25SrRNA1</i>	3,551	<i>25SrRNA1</i>	4,5	<i>25SrRNA1</i>	8
Folha imatura						
Gene mais estável	<i>TIPS-41</i>	0,364	<i>UBQ2</i>	1,625	<i>TIPS-41</i>	1,57
Gene menos estável	<i>UBQ1</i>	2,984	<i>25SrRNA1</i>	4,75	<i>TUB</i>	7,74
Sétima coleta						
Gene mais estável	<i>EF1</i>	0,231	<i>EF1</i>	0,875	<i>EF1</i>	1,316
Gene menos estável	<i>GAPDH</i>	3,359	<i>GAPDH</i>	3,125	<i>GAPDH</i>	8
Décima Terceira coleta						
Gene mais estável	<i>UBQ2</i>	0,704	<i>UBQ1</i>	0,656	<i>UBQ2</i>	1,732
Gene menos estável	<i>EF1</i>	2,653	<i>TUB</i>	2,375	<i>TUB</i>	7,238
DC						
Gene mais estável	<i>UBQ1</i>	1,614	<i>UBQ2</i>		<i>UBQ1</i>	1,565
Gene menos estável	<i>25SrRNA1</i>	3,908	<i>25SrRNA1</i>	5	<i>25SrRNA1</i>	8
DL						
Gene mais estável	<i>TIPS-41</i>	0,934	<i>UBQ1</i>	1,875	<i>TIPS-41</i>	1,414
Gene menos estável	<i>EF1</i>	2,489	<i>TUB</i>	4,25	<i>TUB</i>	7,238

A tabela completa pode ser encontrada no material suplementar apêndice A, tabela Suplementar 2A.

4.3. Validação de genes de referência

Os dados de expressão gênica de dois genes alvos (*PIL5* e *LHP1*) foram normalizados utilizando tanto o par mais estável (*UBQ1/TUB*), quanto o menos estável (*GAPDH/25SrRNA1*) de genes de referência (Figura 6). Em condições experimentais contrastantes como a 7^a e 13^a coletas, tratamentos fotoperiódicos (DC

e DL), e diferentes tecidos (folha madura e folha imatura) para avaliar o efeito na normalização dos dados perante pares de genes de referência estáveis e não estáveis.

Na folha madura, 7^a coleta (estágio inicial de indução), *PIL5* e *LHP1* foram induzidos (*up-regulated*) e significativamente mais expressos em DC do que DL quando os dados de expressão foram normalizados com o par mais estável (*UBQ1/TUB*). No entanto, a expressão relativa não foi significativa quando os dados foram normalizados com o par de genes de referência menos estável (*GAPDH/25SrRNA1*).

Na folha madura, 13^a coleta (estágio de primórdio floral), ambos os genes alvos *PIL5* e *LHP1* não foram significativamente expressos nem quando normalizados com o par mais estável ou menos estável (Figura 6).

Na folha imatura, 7^a coleta, não foi detectado diferença significativa na expressão relativa para *PIL5* e *LHP1* quando normalizados nem com o par mais estável (*UBQ1/TUB*) e nem com o menos estável (*GAPDH/25SrRNA1*). Entretanto, na 13^o coleta, *PIL5* e *LHP1* foram significativamente induzidos e mais expressos em dia curto (DC) do que dia Longo (DL) (Figura 6).

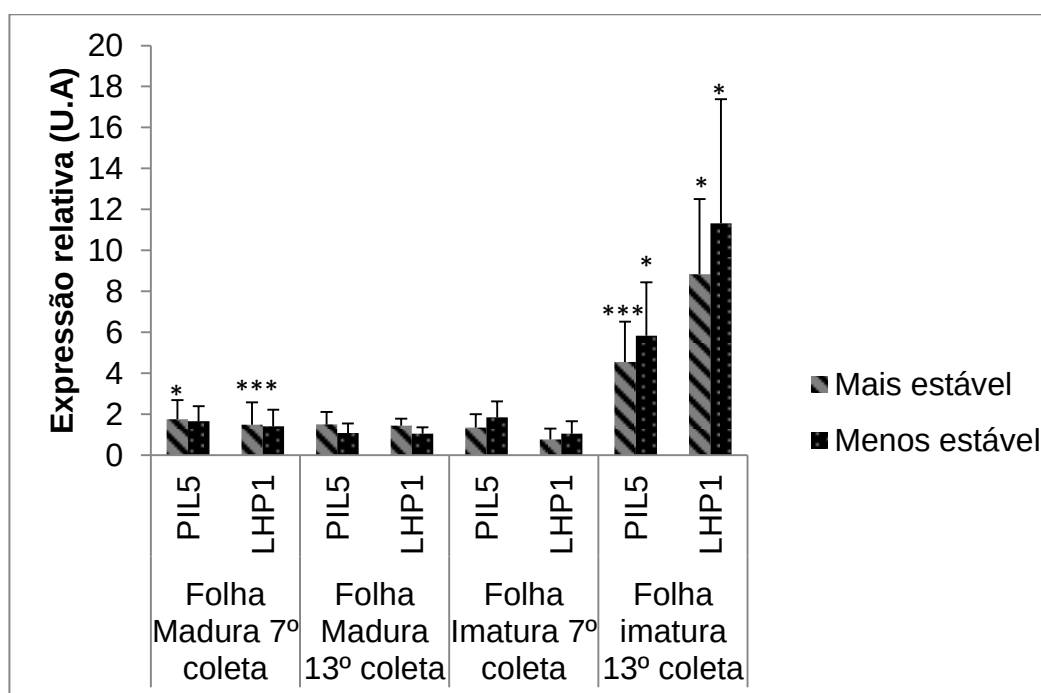


Figura 6. Dados de expressão relativa dos genes alvos pelo programa REST 2009. P <0.001 (***), P<0.01(**), e P< 0.05(*). Comparação entre a expressão gênica dos

genes *PIL5* e *LHP1* (Dia curto em relação a dia longo) quanto normalizados com os pares de candidatos a normalizadores mais e menos estáveis. As barras de erro indicam erro padrão. U.A: Unidades arbitrárias.

Na folha madura 7^a coleta, onde foi realizado o RNA-Seq é possível ainda fazer um comparativo entre os dados de FPKM (Figura 7) e expressão relativa (Figura 8).

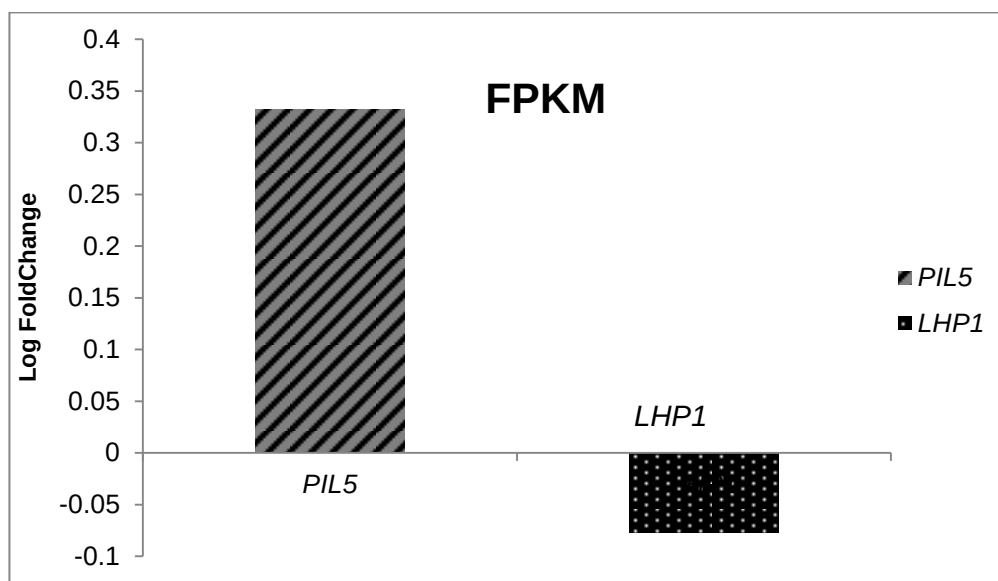


Figura 7. Dados em Log *fold change* de FPKM em folha madura na 7^a coleta advindos de RNA-Seq comparando a quantidade de transcritos encontrado dos genes *LHP1* e *PIL5* em tratamentos de dia curto (DC) e dia longo (DL).

O gene *PIL5* apresentou log *fold change* de 0,33 (Figura 7). Indicando que via FPKM o gene estaria induzido, mais expresso em dia curto em relação a dia longo. Comparando com a expressão relativa, o gene também se apresentou significativamente induzido (1,75), porém, apenas quando normalizado com genes de referência mais estáveis (Figura 8). O gene *LHP1* apresentou log *fold change* de -0,07 (Figura 7). Ou seja, por esta abordagem o gene mostrou-se reprimido. No entanto, quando este mesmo gene tem os dados de FPKM contrastados com os dados de expressão relativa (RT-qPCR), ele se apresentou significativamente como induzido (1,48) quando normalizado com genes de referência mais estáveis, e ainda, sem diferença significativa quando normalizados com genes de referência menos estáveis (Figura 8).

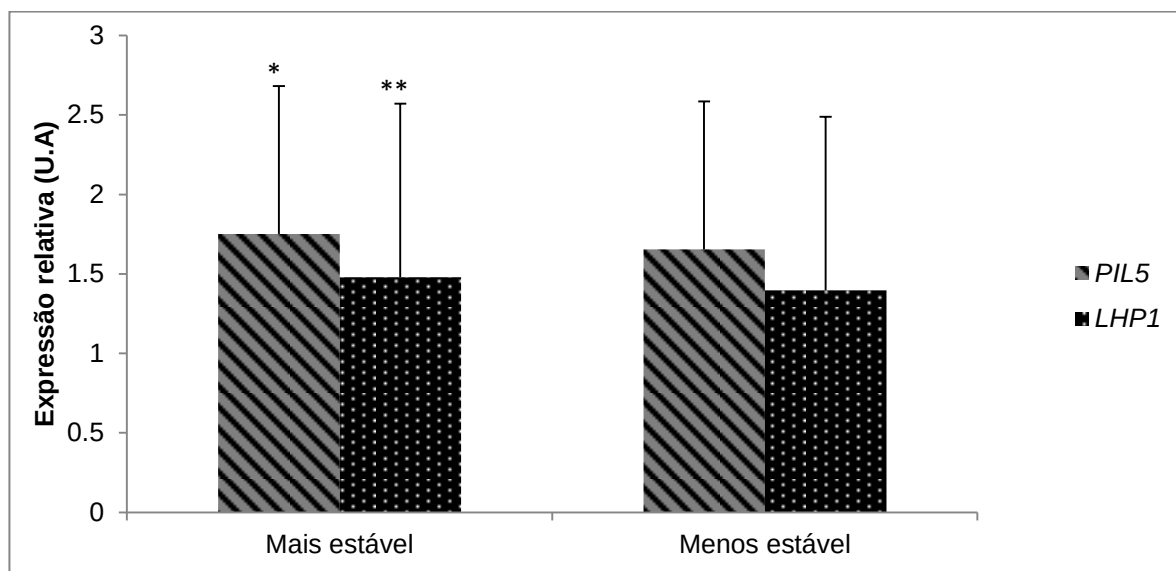


Figura 8. Dados referentes à expressão relativa de PCR quantitativo em tempo real (RT-qPCR) dos genes *PIL5* e *LHP1* em folha madeira na 7ª coleta avaliado com REST 2009. $P < 0.001$ (***) , $P < 0.01$ (**), e $P < 0.05$ (*). As barras de erro indicam erro padrão. U.A: Unidades arbitrárias.

4.4. Validação de transcritos diferencialmente expressos

Doze genes alvos e quatro transcritos de função desconhecida que se demonstraram diferencialmente expressos via RNA-Seq (FPKM) tiveram pares de *primers* desenvolvidos e sua expressão relativa (RT-qPCR) avaliada por intermédio do REST 2009 e normalizados com o par de genes de referência mais estável (*UBQ1/TUB*).

Em folha madura na 7ª coleta os genes diferencialmente expressos significativamente foram *PRR1* como reprimido (0,51) e como induzido *PIL5* (1,75); transcrito desconhecido 127290 (1,49); transcrito desconhecido 181875 (3,72); *CDF2* (2,44); *CDF3* (1,75) e *LHP1* (1,48) (Tabela 7).

Tabela 7. Expressão relativa em genes alvos e transcritos de função desconhecida diferencialmente expressos em folha madura na 7ª coleta.

Folha madura 7ª coleta							
Gene	Tipo	Eficiência da reação	Expressão relativa	Erro padrão	95% I.C	P(H1)	Resposta
<i>PRR1</i>	TRG	0,8	0,51	0,352 - 0,701	0,252 - 1,026	0	REPRIMIDO
<i>LHY</i>	TRG	0,8	1,47	0,812 - 2,722	0,489 - 3,803	0,081	
<i>AGL12</i>	TRG	0,8	1,14	0,666 - 2,093	0,358 - 2,437	0,481	
<i>CONSTANS</i>	TRG	0,7	0,57	0,230 - 1,596	0,112 - 2,645	0,09	
<i>PIL5</i>	TRG	0,6	1,75	0,930 - 3,279	0,513 - 5,696	0,021	INDUZIDO*
<i>177871</i>	TRG	0,7	0,81	0,408 - 1,561	0,220 - 2,401	0,38	
<i>127290</i>	TRG	0,7	1,49	0,890 - 2,178	0,693 - 2,617	0,013	INDUZIDO**
<i>110622</i>	TRG	0,8	0,92	0,504 - 1,659	0,290 - 2,459	0,703	
<i>181875</i>	TRG	0,6	3,72	1,541 - 7,898	1,064 - 17,378	0	INDUZIDO
<i>EID1</i>	TRG	0,9	1,21	0,799 - 1,786	0,543 - 2,360	0,162	
<i>CDF2</i>	TRG	0,9	2,44	1,149 - 5,278	0,678 - 8,818	0,001	INDUZIDO***
<i>CDF3</i>	TRG	1	1,75	1,264 - 2,454	0,983 - 2,840	0	INDUZIDO
<i>LHP1</i>	TRG	0,9	1,48	1,091 - 2,149	0,799 - 2,646	0,005	INDUZIDO**
<i>PRR5</i>	TRG	0,9	1,19	0,554 - 3,220	0,434 - 10,251	0,62	
<i>AGL7</i>	TRG	0,9	1,22	0,826 - 1,944	0,547 - 2,645	0,199	
<i>PRR7</i>	TRG	0,9	1,54	0,783 - 3,223	0,493 - 5,946	0,093	
<i>TUB</i>	REF	1,1	0,81	-	-	-	
<i>UBO1</i>	REF	0,9	1,23	-	-	-	

TRG: Gene alvo e REF: Gene de referência. P(H1) hipótese de nulidade: $P < 0.001$ (***), $P < 0.01$ (**), e $P < 0.05$ (*).

As curvas de dissociação em folha madura na 7ª coleta de maneira geral apresentaram um pico único, indicando amplificação de um único produto (Figura 9). No entanto, em alguns casos os genes apresentaram mais que um pico como, por exemplo: transcrito desconhecido 181875, *PIL5* e *COL9*, possivelmente dímeros de primers e formação de estruturas secundárias. No entanto, uma inespecificidade não pode ser descartada, impossibilitando a utilização dos dados sobre estes genes. A eficiência (E) variou de 1,66 a 2,04 e o R^2 se manteve sempre em 0,99.

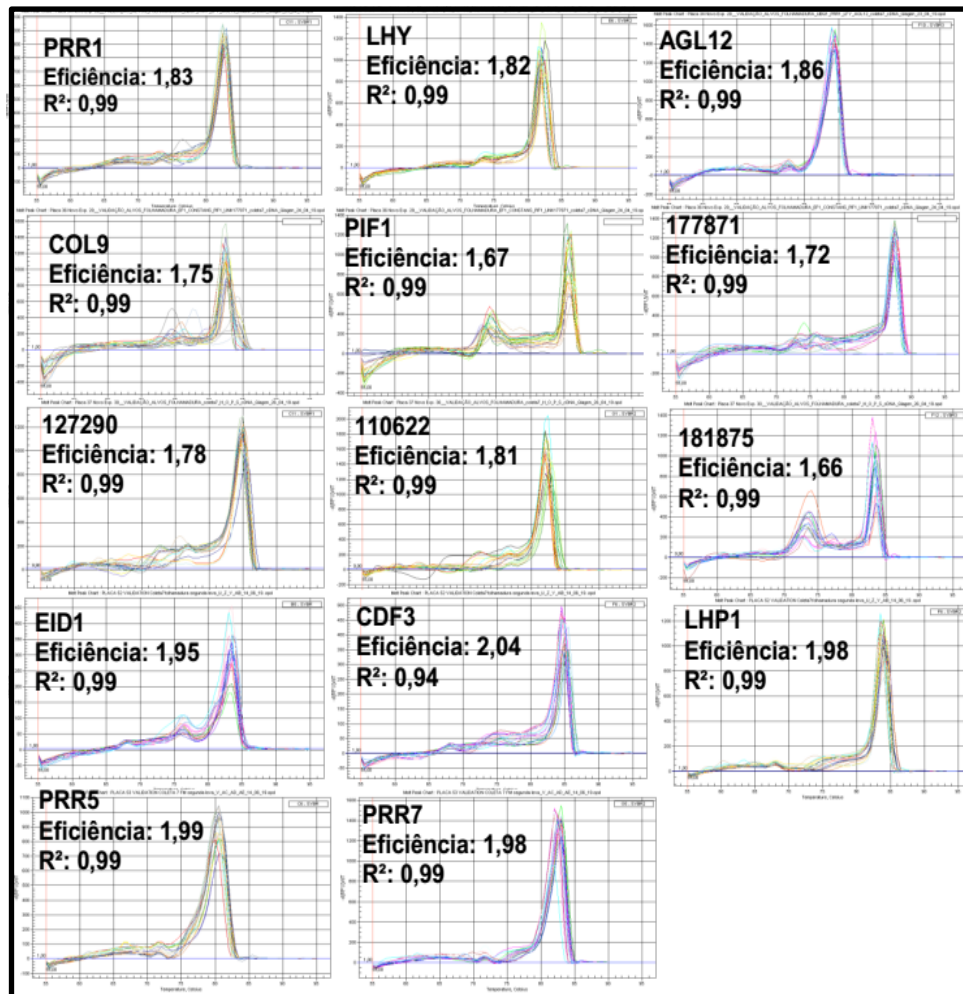


Figura 9. Curvas de dissociação dos genes alvos diferencialmente expressos em folha madura na 7ª coleta com os valores de eficiência (E) e coeficiente de correlação (R^2).

Em folha madura, na 13ª coleta, os genes que mostraram significância quanto a sua expressão diferencial foram: *COL9* (0,35) como reprimido, *AGL12* (1,93) como induzido; transcrito desconhecido 127290 (1,58) e transcrito desconhecido 110622 (2,01) como induzido; *PRR5* (1,56) e *PRR7* (1,90) como induzido (Tabela 8).

Tabela 8. Expressão relativa em genes alvos diferencialmente expressos em folha madura 13^a coleta.

Folha madura 13 ^a coleta							
Gene	Tipo	Eficiência da reação	Expressão relativa	Erro padrão	95% I.C	P(H1)	Resposta
<i>UBQ1</i>	REF	0,93	1,25	-	-	-	
<i>PRR1</i>	TRG	0,8	1,03	0,597 - 1,574	0,364 - 3,030	0,863	
<i>LHY</i>	TRG	0,7	0,80	0,390 - 1,709	0,129 - 2,623	0,501	
<i>AGL12</i>	TRG	0,8	1,93	1,018 - 4,529	0,698 - 6,379	0,008	INDUZIDO**
<i>CONSTANS</i>	TRG	0,6	0,35	0,146 - 0,709	0,107 - 1,637	0,004	REPRIMIDO**
<i>PIL5</i>	TRG	0,7	1,49	0,612 - 3,739	0,107 - 10,184	0,342	
<i>177871</i>	TRG	0,8	0,83	0,602 - 1,076	0,512 - 1,861	0,161	
<i>127290</i>	TRG	0,7	1,58	0,905 - 2,841	0,630 - 4,060	0,023	INDUZIDO*
<i>110622</i>	TRG	0,8	2,01	0,968 - 4,258	0,535 - 7,405	0,02	INDUZIDO*
<i>181875</i>	TRG	0,8	1,07	0,859 - 1,296	0,673 - 1,459	0,34	
<i>EID1</i>	TRG	0,7	0,80	0,264 - 1,981	0,150 - 4,714	0,524	
<i>CDF2</i>	TRG	0,9	1,14	0,548 - 2,267	0,439 - 3,362	0,535	
<i>CDF3</i>	TRG	1	0,78	0,155 - 4,073	0,095 - 21,515	0,652	
<i>LHP1</i>	TRG	0,9	1,44	0,343 - 3,607	0,149 - 23,498	0,439	
<i>PRR5</i>	TRG	0,9	1,56	0,853 - 2,494	0,566 - 4,366	0,036	INDUZIDO*
<i>AGL7</i>	TRG	0,9	1,29	0,903 - 1,796	0,580 - 2,158	0,057	
<i>PRR7</i>	TRG	1	1,90	0,970 - 3,761	0,545 - 8,371	0,026	INDUZIDO*
<i>TUB</i>	REF	1,1	0,79	-	-	-	

TRG: Gene alvo e REF: Gene de referência. P(H1) hipótese de nulidade: $P < 0.001$ (***), $P < 0.01$ (**), e $P < 0.05$ (*).

As curvas de dissociação ainda em folha madura na 13^a coleta (Figura 10) foram todas razoáveis, apresentando um pico único e até mesmo nos casos dos genes que apresentaram uma leve saliência na 7^a coleta (transcrito desconhecido 181875, *PIL5* e *COL9*), aqui se comportaram de maneira evidentemente específica. A Eficiência (E) foi de 1,68 à 2,2 e R^2 0,99.

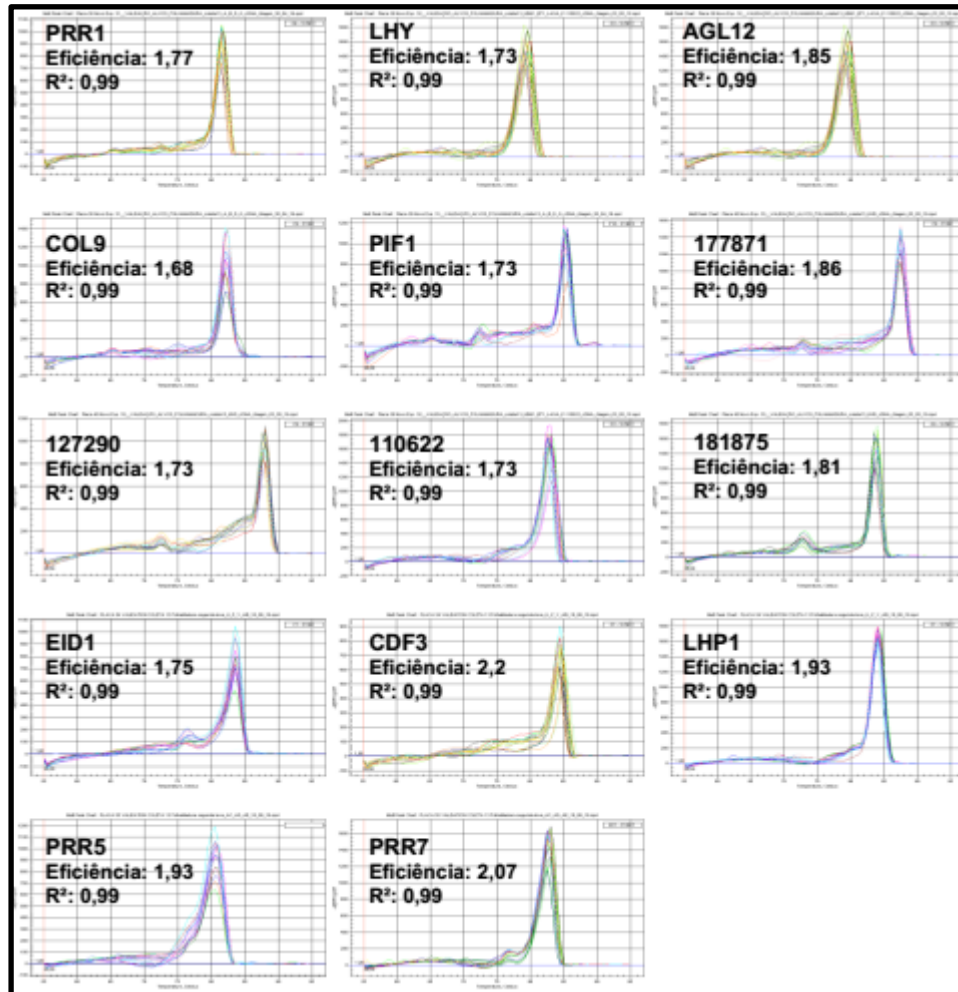


Figura 10. Curvas de dissociação dos genes alvos diferencialmente expressos em folha madura na 13ª coleta com os valores de eficiência (E) e coeficiente de correlação (R^2).

Em folha imatura na 7ª coleta os genes diferencialmente expressos significativamente foram *COL9* (0,41); *CDF3* (0,57) e *AGL7* (0,45) como reprimido e *AGL12* (1,84) e transcrito desconhecido 181875 (1,89) como induzido (Tabela 9).

Tabela 9. Expressão relativa em genes alvos diferencialmente expressos em folha imatura 7ª coleta

Folha imatura 7ª coleta						
Gene	Tipo	Eficiência da reação	Expressão relativa	Erro padrão	95% I.C	P(H1) Resposta
<i>UBQ1</i>	REF	1	1,025			
<i>PRR1</i>	TRG	0,8	0,70	0,425 - 1,128	0,241 - 1,741	0,066
<i>LHY</i>	TRG	0,6	1,20	0,736 - 2,039	0,504 - 3,159	0,31
<i>AGL12</i>	TRG	0,7	1,84	0,960 - 4,115	0,646 - 5,371	0,01 INDUZIDO*
<i>CO</i>	TRG	0,8	0,41	0,246 - 0,768	0,161 - 1,111	0,001 REPRIMIDO***
<i>PIL5</i>	TRG	0,7	1,33	0,663 - 2,692	0,402 - 4,824	0,246
<i>177871</i>	TRG	0,7	1,18	0,807 - 2,095	0,496 - 2,736	0,318
<i>127290</i>	TRG	0,7	1,22	0,812 - 2,053	0,492 - 3,045	0,249
<i>110622</i>	TRG	0,8	0,85	0,573 - 1,260	0,381 - 1,509	0,242
<i>181875</i>	TRG	0,7	1,89	0,861 - 4,000	0,358 - 5,495	0,035 INDUZIDO*
<i>EID1</i>	TRG	0,3	0,82	0,549 - 1,129	0,430 - 2,156	0,2
<i>CDF2</i>	TRG	0,7	1,48	0,830 - 2,241	0,533 - 6,059	0,071
<i>CDF3</i>	TRG	0,9	0,57	0,343 - 1,000	0,269 - 1,477	0,013 REPRIMIDO*
<i>LHP1</i>	TRG	0,8	0,76	0,541 - 1,076	0,404 - 1,902	0,068
<i>PRR5</i>	TRG	0,8	0,89	0,589 - 1,391	0,409 - 1,890	0,405
<i>AGL7</i>	TRG	0,9	0,45	0,208 - 0,966	0,089 - 1,644	0,013 REPRIMIDO*
<i>PRR7</i>	TRG	1	1,50	0,648 - 5,494	0,115 - 9,675	0,321
<i>TUB</i>	REF	0,99	0,97			

TRG: Gene alvo e REF: Gene de referência. P(H1) hipótese de nulidade: $P < 0.001$ (***), $P < 0.01$ (**), e $P < 0.05$ (*).

Na folha imatura na 7ª coleta, todas as curvas de dissociação foram específicas (Figura 11) exceto o transcrito desconhecido 181875. A eficiência (E) variou de 1,37 à 2,03 e R^2 : 0,99.

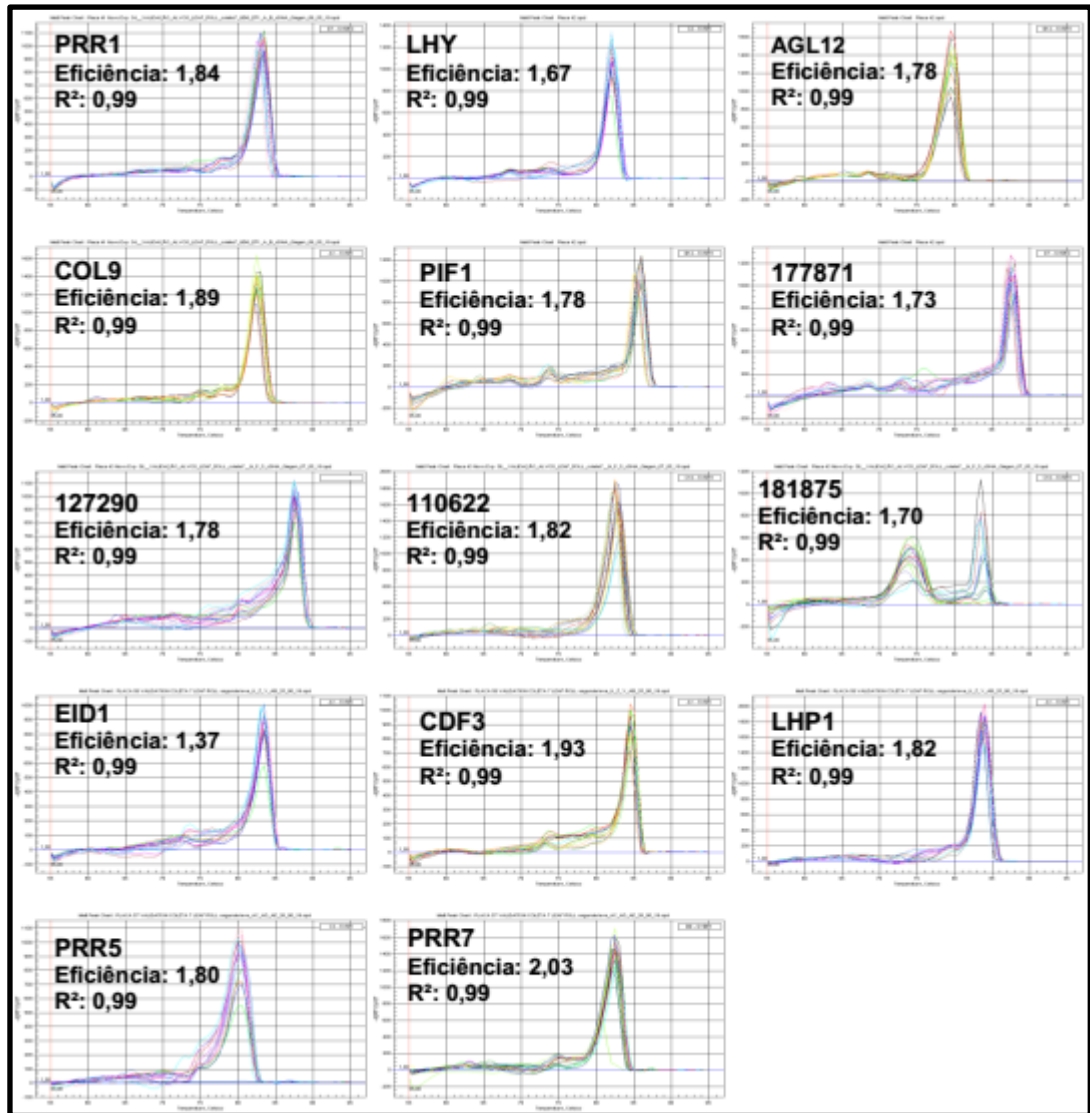


Figura 11. Curvas de dissociação dos genes alvos diferencialmente expressos em folha imatura na 7ª coleta com os valores de eficiência (E) e coeficiente de correlação (R^2).

Em folha imatura na 13ª coleta os genes diferencialmente expressos significativamente foram todos induzidos como *PRR1* (4,66); *LHY* (2,57); *AGL12* (1,98); *PIL5* (4,54); transcrito desconhecido 127290 (2,95); transcrito desconhecido 110622 (5,25); transcrito desconhecido 181875(3,82); *EID1* (2,28); *CDF2* (2,89); *CDF3* (2,80); *LHP1* (8,83); *PRR5* (1,83); *AGL7* (1,94) e *PRR7* (4,93) (Tabela 10).

Tabela 10. Expressão relativa em genes alvos diferencialmente expressos em folha imatura 13^a coleta

Folha imatura 13 ^a coleta							
Gene	Tipo	Eficiência da reação	Expressão Relativa	Erro padrão	95% I.C	P(H1)	Resposta
<i>UBQ1</i>	REF	1	1,70				
<i>PRR1</i>	TRG	0,7	4,66	2,146 - 0,121	1,138- 18,398	0	INDUZIDO
<i>LHY</i>	TRG	0,6	2,57	1,242 - 6,092	0,648- 12,785	0	INDUZIDO
<i>AGL12</i>	TRG	0,8	1,98	0,961 - 3,973	0,816- 12,058	0,009	INDUZIDO**
<i>CO</i>	TRG	0,8	0,71	0,283 - 1,542	0,193 - 3,252	0,254	
<i>PIL5</i>	TRG	0,7	4,54	1,971- 10,471	1,244- 24,818	0,001	INDUZIDO***
<i>177871</i>	TRG	0,7	1,07	0,410 - 2,601	0,302 - 4,648	0,778	
<i>127290</i>	TRG	0,7	2,95	1,487 - 6,342	0,925- 13,719	0	INDUZIDO
<i>110622</i>	TRG	0,8	5,25	1,906- 12,655	1,301- 38,008	0	INDUZIDO
<i>181875</i>	TRG	0,7	3,82	1,064- 14,045	0,463- 32,881	0,009	INDUZIDO**
<i>EID1</i>	TRG	0,8	2,28	0,926 - 5,207	0,456- 16,726	0,016	INDUZIDO*
<i>CDF2</i>	TRG	0,9	2,89	1,542 - 5,771	1,101- 12,111	0	INDUZIDO
<i>CDF3</i>	TRG	0,9	2,80	1,337 - 5,986	0,605- 12,704	0,002	INDUZIDO**
<i>LHP1</i>	TRG	0,9	8,83	3,670- 21,389	2,852- 56,233	0	INDUZIDO
<i>PRR5</i>	TRG	0,8	1,83	0,895 - 5,105	0,529 - 7,345	0,044	INDUZIDO*
<i>AGL7</i>	TRG	0,9	1,94	0,954 - 3,906	0,608 - 7,772	0,01	INDUZIDO*
<i>PRR7</i>	TRG	0,9	4,93	2,281- 13,766	1,273- 25,343	0	INDUZIDO
<i>TUB</i>	REF	0,99	0,58				

TRG: Gene alvo e REF: Gene de referência. P(H1) hipótese de nulidade: $P < 0.001$

(***), $P < 0.01$ (**), e $P < 0.05$ (*).

Em folha imatura na 13^a coleta, o gene *PRR1* claramente apresentou uma curva de dissociação ruim (Figura 12), sendo inviável a utilização do mesmo. O transcrito desconhecido *181875* apresentou dois picos na curva de dissociação, no demais, todo o restante está com perfil único. A eficiência (E) variou de 1,68 à 1,97 e R^2 de 0,77 à 0,99.

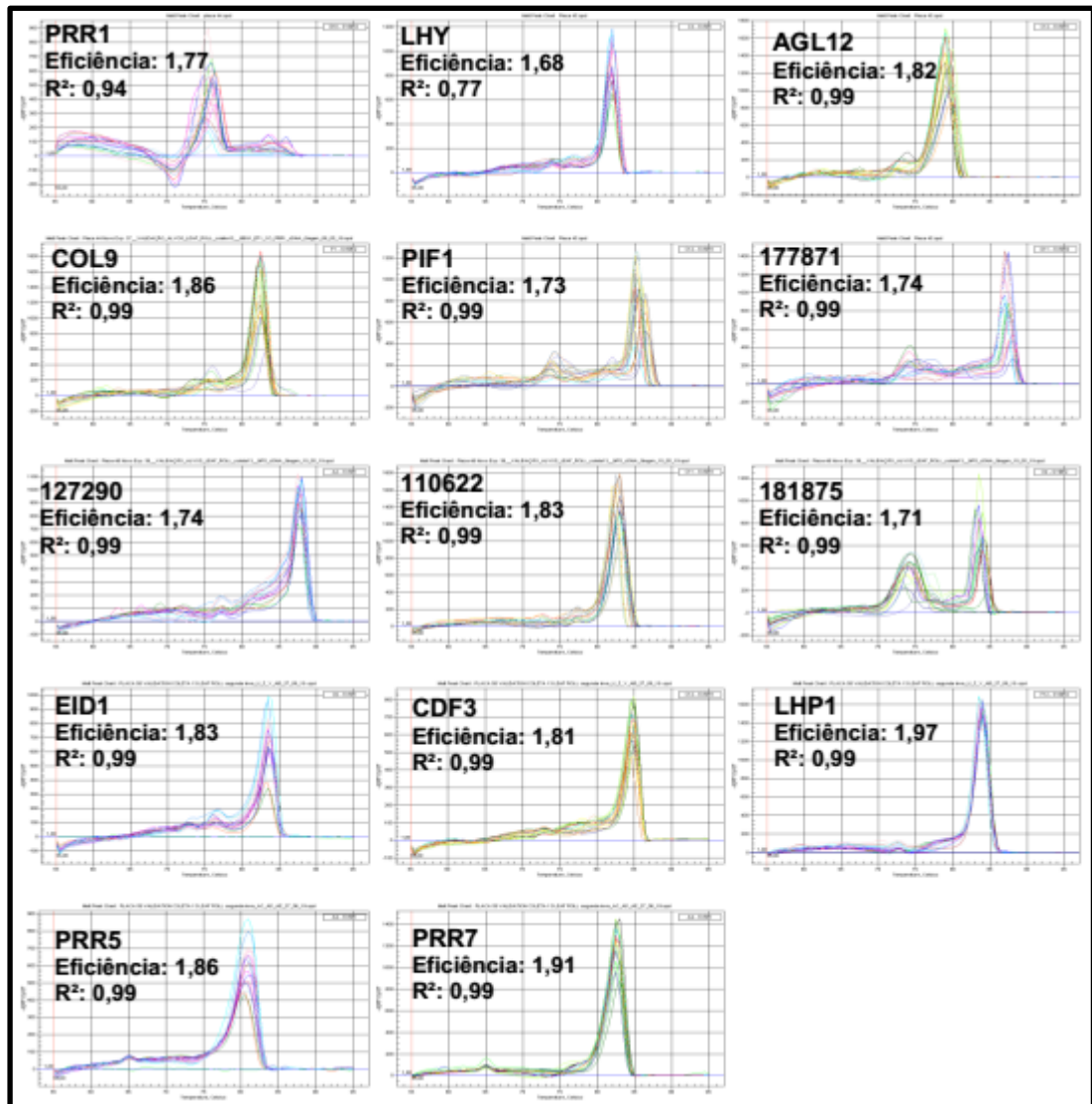


Figura 12. Curvas de dissociação dos genes alvos diferentemente expressos em folha imatura na 13ª coleta com os valores de eficiência (E) e coeficiente de correlação (R^2).

4.6. Estudo de expressão relativa ao longo do tempo (*timecourse*) gene ScFT3

O gene *ScFT3* possível ortólogo de *AtFT* de *Arabidopsis thaliana*, em cana-de-açúcar, teve sua expressão relativa avaliada ao longo de vários pontos de coleta tanto em regime fotoperiódico de dia curto como de dia longo.

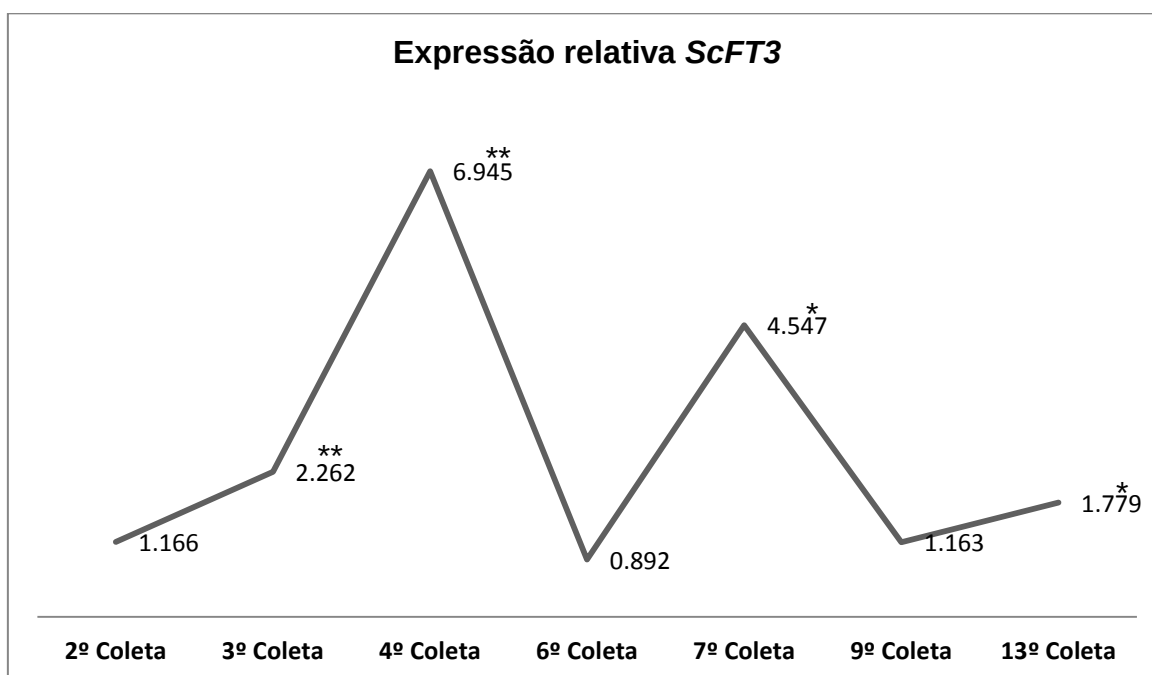


Figura 13. Expressão relativa do gene *ScFT3* ao longo do desenvolvimento do florescimento em folha madura avaliada com REST 2009 $P < 0.001$ (***), $P < 0.01$ (**), e $P < 0.05$ (*).

Analisando a expressão relativa do gene *ScFT3* ao longo do desenvolvimento do florescimento, em folha madura, nota-se que o *ScFT3* encontra-se induzido na 3ª coleta (2,262; $P < 0.01$); 4ª coleta (6,945; $P < 0.01$); 7ª coleta (4,547; $P < 0.05$) e 13ª coleta (1,779; $P < 0.05$), respectivamente, 16, 23, 46 e 88 dias após início do tratamento indutivo (dia curto).

Provavelmente, a 3ª, 4ª e 7ª coletas capturaram o acúmulo destes transcritos sendo que a 7ª coleta coincidiu com o alongamento do meristema apical observado em corte histológico, sinalizando o início da transição do meristema vegetativo em reprodutivo. A 13ª coleta representa o início do primórdio floral. Houve dois picos (Figura 16) indicando que a expressão relativa deste transcrito atingiu um nível alto de expressão em dois momentos distintos (4ª e 7ª coletas).

5. DISCUSSÃO

5.1. Genes de Referência

A seleção de genes de referência adequados é uma etapa essencial na normalização de dados de expressão gênica via PCR quantitativo em tempo real (RT-qPCR) pelo fato de impactar diretamente a análise e interpretação dos dados. Alguns genes de referência têm sido utilizados anteriormente em análise de expressão gênica envolvendo genes do florescimento em cana-de-açúcar, como por exemplo, *GAPDH* e *ACTINA*, os quais foram utilizados para normalizar a expressão de *ScFT1* e *ScTFL1* (Coelho et al., 2014). Medeiros et al. (2016) utilizou *EF1- α* para normalização de genes diferencialmente expressos em estádios iniciais e tardios do florescimento em cultivares de cana-de-açúcar, estes genes de referência não foram experimentalmente validados e não tiveram a estabilidade e consistência da expressão avaliados sob diferentes condições experimentais para o florescimento.

Subsequentemente, ao obter os dados de estabilidade de oito candidatos a genes de referência utilizando três algoritmos diferentes (*BestKeeper*, *NormFinder* e *RefFinder*) ficou evidenciado que *UBQ1* e *TUB* foram os mais estáveis e adequados para normalização de dados de expressão de genes alvos relacionados com a via da indução fotoperiódica em cana-de-açúcar.

O resultado da comparação entre *BestKeeper* e *NormFinder*, separadamente, com os dados integrados do *RefFinder*, nos aponta alguns candidatos a gene de referência comuns entre os algoritmos, mas não em todas as condições. Esta discrepância pode ser atribuída ao fato de que o *RefFinder* se utiliza apenas os dados brutos dos valores de CTs assumindo uma eficiência (E) de 100% (Spiegelaere, et al., 2015).

Finalmente, como o algoritmo *NormFinder* possibilita a estimação da variação da expressão gênica entre subgrupos (intra e inter grupos) de amostras (Andersen et al., 2004), se adequando a nossas condições experimentais (tecido, tempo de coleta e tratamentos fotoperiódicos) foi dado uma ênfase maior para a seleção de candidatos a gene de referência por este programa. Portanto, os candidatos a gene de referência *UBQ1* e *TUB* se enquadram como normalizadores na maioria das

condições experimentais, incluindo a qual o RNA-Seq foi realizado (folha madura, 7ª coleta e tratamento de dia curto). Sob esta perspectiva, a combinação *UBQ1/TUB* provou ser a melhor combinação de genes de referência.

Os genes de referência *UBQ1* e *TUB* além de boa estabilidade possuem uma média de CTs que fica próximo a dos genes alvos, algo que é requerido pra uma boa normalização (Andersen et al., 2004), isso evita sub ou superestimação na avaliação dos resultados. Vale ressaltar que, a seleção de genes de referência aqui realizada, sob uma ampla gama de condições experimentais, levando-se em consideração dois tecidos, em dois estádios de desenvolvimento e tratamentos fotoperiódicos, ainda não foi reportada na literatura.

Para demonstrar e validar a acurácia dos genes de referência selecionados (*UBQ1/TUB*), os mesmos foram utilizados para normalizar a expressão de dois genes alvos, *PIL5* e *LHP1*, ambos diferencialmente expressos em um ensaio de RNA-Seq, previamente realizado em nosso laboratório. A expressão destes dois genes alvos foi avaliada com três réplicas biológicas em triplicatas técnicas para três condições experimentais. Contrastando os dados de expressão dos genes alvos normalizados com genes de referência mais estáveis (*UBQ1/TUB*) assim como os menos estáveis (*GAPDH/25SrRNA*), ficou evidente que *UBQ1/TUB* representa a melhor combinação de genes de referência para normalização, em virtude dos mesmos serem aptos a distinguir, mesmo a variação mínima na expressão gênica (entre *PIL5* e *LHP1* em folha madura na 7ª coleta), a qual, não foi detectada quando utilizada a combinação *GAPDH/25SrRNA1*. Tal diferença na expressão gênica, mesmo mínima pode ser crucial para o entendimento de um processo fisiológico a nível molecular e reforça a necessidade de selecionar bons genes de referência e validar a estabilidade da expressão em condições apropriadas antes de ser utilizado para análise da expressão relativa de genes alvos.

Quanto á comparação entre FPKM e expressão relativa fica evidente a importância da validação de genes via (RT-qPCR), sendo a expressão relativa importante para verificar se de fato, o gene está induzido ou reprimido. No caso do *LHP1*, isso pode ser exemplificado, via FPKM, em que o gene se mostrou induzido. No entanto, via expressão relativa normalizando com genes de referências estáveis, ficou demonstrado que na verdade o gene estava reprimido.

5.2. Expressão relativa de genes alvos diferencialmente expressos

A maioria dos genes de indução floral são expressos em um ciclo de 24 horas e seu perfil de expressão que ativa a transição do meristema vegetativo para o reprodutivo. Genes associados às vias de percepção de fotoperíodo e indução floral foram avaliados sob dia longo (14 h de luz) (Glassop, 2018) servindo como parâmetro de comparação para a condição de dia longo do presente trabalho, entretanto, deve ser levado em consideração que os pares de *primers* desenvolvidos pelos autores tiveram intenção de captar expressão alélica cumulativa dos genes, diferente de nossa abordagem mais específica com pares de *primers* visando às regiões 5' e 3'UTRs ("Untranslated Regions" – Regiões Não Traduzidas). A cultivar (Q174) utilizada por Glassop et al., 2018 também pode ser pela diferença no número de alelos comparado com a cultivar brasileira IACSP96-7569, levando a diferença no perfil de expressão de alguns genes.

Na folha madura sétima coleta, o gene *PRR1* foi reprimido, mostrando-se significativamente mais expresso em DL do que DC. Em *Arabidopsis*, a família *PRR* consiste de cinco membros (*PRR9*, *PRR7*, *PRR5*, *PRR3* e *PRR1/TOC1*) que são envolvidos na função do relógio circadiano (Mizuno e Nakamichi, 2005). Os transcritos *PRRs* começam a se acumular após o amanhecer, um após o outro em uma ordem específica (*PRR9*–*PRR7*–*PRR5*–*PRR3*–*PRR1*) com intervalo de 2 a 3 horas, formando um perfil de expressão do tipo *loop* (Makino et al. 2001). Glassop et al., 2018 verificou que no ciclo de 24 horas, o gene *PRR1* em DL foi altamente expresso em folha imatura e folha madura (+1), atingindo seu pico máximo de expressão com cerca de 10 horas no regime de luz, seguido de um decréscimo após 2 horas no período escuro. No entanto, no presente trabalho a hora da coleta foi as 16h, período no qual, provavelmente, a expressão de *PRR1* já estava em declínio. Em folha madura na décima terceira coleta, os genes da família *PRR* que foram significativamente expressos foram *PRR5* e *PRR7* ambos induzidos, mais expressos em DC do que DL. Há de se notar que na sétima coleta (meristema indutivo em

transição do vegetativo para o reprodutivo) o “loop” de retroalimentação de PRRs está no fim, uma vez que tem uma ordem específica de expressão e *PRR1* é o último a ter o pico de expressão. Porém, na décima terceira coleta (primórdio floral) o “loop” está na metade, em vista que neste ponto de amostragem *PRR7* e *PRR5* estão sendo expressos. Em folha imatura, na décima terceira coleta, os genes *PRR7* e *PRR5* estão significativamente induzidos e seguem o mesmo padrão de expressão da folha madura na décima terceira coleta.

O gene *LHY* foi induzido de maneira significativa em folha imatura na décima terceira coleta, assim como no trabalho de Glassop et al., 2018 onde *LHY* se mostrou mais expresso em folhas imaturas. No entanto, segundo os autores o pico da expressão de *LHY* em um ciclo de 24h ocorre principalmente no período escuro, que é diferente do momento em que foi coletado a amostra do nosso experimento, por esta razão a expressão de *LHY* está no começo da ascensão no presente trabalho.

Em folha madura na sétima coleta, e em folha imatura na décima terceira coleta *PIL5* está induzida. É sabido que *PIL5* interage fisicamente (proteína-proteína) com *TOC1/PRR1* (Nakamura et al., 2007). O gene *PIL5* é um fator de transcrição bHLH (“basic helix-loop-helix”) que se liga nas regiões promotoras de *LHY* e *CCA1* em *Arabidopsis*, e quando suprimido, leva a uma maior expressão do gene *CONSTANS* (*CO*) e *FLOWERING LOCUS T* (*FT*) causando florescimento precoce (Oda et al., 2003). Na folha imatura na décima terceira coleta, onde houve expressão significativa de *PRR1* (induzida) pode ser um indicativo que esta interação ocorreu. Entretanto, em folha madura na sétima coleta, mesmo *PIL5* estando induzido, o *PRR1* está reprimido. *PIL5* também está relacionado com interação com fitocromos e atua de maneira fraca em *PhyA*. No entanto, atua de maneira forte sobre *PhyB* (Duek e Fankhauser, 2005). Na folha madura na sétima coleta (ponto de transição do meristema vegetativo para o reprodutivo), a indução de *PIL5* pode indicar interação com *PhyB* na resposta ao fotoperíodo indutivo.

O gene *LHP1* está induzido em folha madura na sétima coleta e folha imatura na décima terceira coleta. O fato de *LHP1* estar induzido na décima terceira coleta em folha madura faz sentido, do ponto de vista que o gene *ScFT3* não está com expressão significativa, possivelmente sendo reprimido pela indução de *LHP1*. Em

Arabidopsis análises moleculares de mutantes indicaram que *LHP1* controla o tempo de florescimento primeiramente reconhecendo e se ligando a H3K27me3 e interage diretamente com a repressão da cromatina do *FT* controlando sua expressão (Feng et al, 2017). Entretanto, em folha madura na sétima coleta, *LHP1* está induzido assim como o *ScFT3*, provavelmente, neste caso, o *LHP1* pode estar envolvido em outros papéis, uma vez que também tem função no desenvolvimento de raízes e outros processos (Feng et al, 2017).

Em folha madura na décima terceira coleta, e em folha imatura na décima terceira coleta, o gene *COL9* está significativamente reprimido. Em arroz, outra gramínea, já foi observado que *OsCOL9* atrasa o florescimento ao inibir a expressão de *EHD1*, *RFT* e *Hd3a* (LIU et al., 2016). O gene *ScFT3* que é o ortólogo de *Hd3a* em cana-de-açúcar está induzido nos pontos de coleta em que *COL9* está reprimido, o que indica uma possível conservação de função de *COL9* em gramíneas.

Em *Arabidopsis* *AGAMOUS-LIKE 12* (*XAL1*) é gene irmão dos genes relacionados *AGAMOUS* (*AG*) e é membro da família MADS-box que é um dos componentes chaves na rede de indução floral e desenvolvimento do florescimento. A expressão gênica sugere que *XAL1* é um regulador a montante (*upstream*) de *SOC*, *FT* e *LFY* (Tapia-lopez et al., 2008). Em cana-de-açúcar, *AGL12* se mostrou significativamente induzido em folha madura na décima terceira coleta e folha imatura na sétima e décima terceira coleta. O gene *ScFT3* está significativamente induzido em duas condições das quais *AGL12* também está (folha madura na décima terceira coleta e folha imatura na sétima coleta), Tapia-lopez et al., (2008) notaram via hibridização *in situ* do mRNA de *XAL1* que o gene se expressa dentre outros tecidos, no tecido vascular de *Arabidopsis*.

Em *Arabidopsis* há uma redundância entre os fatores de transcrição DOF (*CDF1*, *CDF2*, *CDF3* e *CDF5* como reguladores negativos de *CO* no tempo de florescimento (Fornara et al., 2009). O gene *CDF3* foi significativamente reprimido em folha imatura na sétima coleta. No entanto, esperava-se que o *CO* estaria induzido uma vez que não haveria a repressão de *CDF3*, o que não ocorreu. Na verdade, assim como *CDF3*, *CO* também está reprimido na folha madura na sétima coleta. Na folha madura na sétima coleta e folha imatura na décima terceira coleta, *CDF3* está significativamente induzido e nesta condição sim, o *CO* não apresenta

diferença de expressão, podendo ter sido regulado negativamente pela indução de *CDF3* assim como em *Arabidopsis*.

O gene *EID1* está relacionado com uma regulação negativa na via dependente de *PhyA* (Buche et al., 2000). Na ausência de *EID1*, *PhyA* responde como um receptor de luz vermelha sob irradiação contínua (Dieterle et al., 2001). O gene *EID1* está significativamente induzido em folha imatura na décima terceira coleta. Provavelmente, *EID1* está reprimindo *PhyA*, pois a folha imatura ainda não foi exposta ao meio externo para recepção de luz e, portanto, não há necessidade de fitocromos ativos.

A respeito dos transcritos desconhecidos diferencialmente expressos (127290, 117871 e 110622), são necessários para a caracterização destes genes, via transgenia ou estudo com mutantes em *Arabidopsis*.

5.3. Expressão relativa de *ScFT3* ao longo do desenvolvimento

As câmaras de fotoperíodo têm sido utilizadas com sucesso para promover o florescimento em cana-de-açúcar (Berding e Moore, 1991; Melloni et al. 2015, Hale et al. 2017) uma vez que aumenta a indução e o florescimento quando comparado com as condições naturais de indução (Hale et al. 2017).

O gene *ScFT3* mostrou-se induzido na 3^o (16 dias), 4^o (23 dias) e 7^a (46 dias) coleta indicando que nestas coletas ocorreu o estímulo da expressão do *FLOWERING LOCUS (FT)*. Em arroz a proteína Hd3a (ortólogo ao *FT* de *Arabidopsis*), migra através do floema para o meristema apical, onde interage com as proteínas 14-3-3 e OsFD1 (TSUJI et al., 2013). No meristema apical, são desencadeados os eventos que levam à formação do primórdio floral, representado pela 13^a coleta (88 dias) e está também induzido, uma vez que *ScFT3* possui um papel em estádios avançados do florescimento. Em *Arabidopsis* a interação de FT/FD no meristema apical leva a ativação dos genes *APETALA 1 (AP1)* e *FRUITFUL (FUL)* que são responsáveis pela identidade floral do meristema (Abe et. al., 2005). Um requerimento contínuo de estímulo à floração foi relatado por Julien (1971), que apontou que todas as etapas do desenvolvimento da floração (desde a

indução até a parte inicial do alongamento da inflorescência) são sensíveis ao fotoperíodo.

Morfologicamente, o meristema apical correspondente a 7ª coleta encontra-se em transição de vegetativo para o reprodutivo e, nesta coleta, na folha madura correspondente, observou-se um dos picos de expressão relativa de *ScFT3*, sugerindo provável ponto crítico para indução ao florescimento, pois é onde há maior indicativo correlacionando morfologia e expressão relativa.

6. CONCLUSÕES

Com base nos resultados aqui obtidos é possível concluir que os genes de referência *UBQ1/TUB* são os mais estáveis para normalizar dados de genes alvos relacionados com indução fotoperiódica em cana-de-açúcar via PCR quantitativo em tempo real (RT-qPCR). Uma vez que na etapa de validação ficou evidente sua sensibilidade para captar até mesmo pequenas diferenças na expressão relativa dos genes alvos. Na contramão, *GAPDH/25SrRNA1* são os menos estáveis para as mesmas condições propostas.

Dentre os genes alvos diferencialmente expressos, há de se notar que possa ocorrer uma relação entre os genes diferencialmente expressos de forma significativa com uma especificidade em estágio de desenvolvimento e tempo de fotoperíodo. Os genes alvos desconhecidos se mostram promissores, no entanto são necessários mais estudos para desvendar sua real função na indução ao florescimento em cana-de-açúcar.

A expressão relativa de *ScFT3* ao longo do desenvolvimento deixou claro que, este é um potencial florigeno e possui uma grande importância tanto nas etapas iniciais e indutivas do florescimento quanto em etapas mais avançadas do florescimento. De forma geral os dados aqui gerados são inéditos e compõe um novo capítulo nesta história visando o completo entendimento dos mecanismos relacionados ao processo de indução floral em cana de açúcar.

7. REFERÊNCIAS

- Abe M (2005) FD, a bZIP Protein Mediating Signals from the Floral Pathway Integrator FT at the Shoot Apex. **Science** 309:1052–1056.
- Alexander AG (1973) **Sugarcane physiology**. Amsterdam: Elsevier Scientific Publishing, 752p.
- Araki, Takashi, Yasushi Kobayashi, Hidetaka Kaya, and Masaki Iwabuchi. (1998) The flowering-time gene FT and regulation of flowering in Arabidopsis. **Journal of plant research** 111.2: 277-281.
- Berding N, Hurney AP (2005) Flowering and lodging , physiological-based traits affecting cane and sugar yield What do we know of their control mechanisms and how do we manage them ? **Field Crops Research** 92:261–275.
- Bottcher A, Cesarino I, et al. (2013) Lignification in Sugarcane: Biochemical Characterization, Gene Discovery, and Expression Analysis in Two Genotypes Contrasting for Lignin Content. **PLANT PHYSIOLOGY** 163:1539–1557.
- Büche C, Poppe C, Schäfer E, Kretsch T (2000) eid1 : A New Arabidopsis Mutant Hypersensitive in Phytochrome A–Dependent High-Irradiance Responses. **The Plant Cell** 12:547–558.
- Bustin S (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. **Journal of Molecular Endocrinology** 25:169–193.
- Caputo MM, Silva MA, Beauclair EGF, Gava GJC (2007) Acúmulo de sacarose, produtividade e florescimento de cana-de-açúcar sob reguladores vegetais. **InterCiência** 32:834-840.
- Carla P, Minow MAA, Chalfun-júnior A, Colasanti J (2014) Putative sugarcane FT / TFL1 genes delay flowering time and alter reproductive architecture in Arabidopsis. 5:1–12.
- Coelho CP, Netto APC, Colasanti J (2013) A proposed model for the flowering signaling pathway of sugarcane under photoperiodic control. 12:1347–1359.
- Conab- Companhia Nacional de Abastecimento. Cana-de-Açúcar, Safra 2020/2021. Primeiro levantamento, Maio de 2020. Disponível em: <<https://www.conab.gov.br/info-agro/safras/cana/boletim-da-safra-de-cana-de-acucar>>
- Corbesier L, Coupland G (2006) The quest for florigen: a review of recent progress. **Journal of Experimental Botany** 57:3395–3403.

Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, Searle I, Giakountis A, Farrona S, Gissot L, Turnbull C, et al. (2007) FT Protein Movement Contributes to Long-Distance Signaling in Floral Induction of Arabidopsis. **Science** 316:1030–1033.

Clements HF (1975) **Flowering of sugarcane: mechanics and control**. Hawaii: University Hawaii, 60 p. (Technical Bulletin.No. 92.).

de Andrade LM, dos Santos Brito M, Fávero Peixoto Junior R, Marchiori PER, Nóbile PM, Martins APB, Ribeiro RV, Creste S (2017) Reference genes for normalization of qPCR assays in sugarcane plants under water deficit. **Plant Methods** 13:28.

Dieterle M (2001) EID1, an F-box protein involved in phytochrome A-specific light signaling. **Genes & Development** 15:939–944.

Duek PD, Fankhauser C (2005) bHLH class transcription factors take centre stage in phytochrome signalling. **Trends in Plant Science** 10:51–54.

Fedoroff N, Schläppi M, Raina R (1995) Epigenetic regulation of the maize Spm transposon. **BioEssays** 17:291–297.

Feng J, Lu J (2017) LHP1 Could Act as an Activator and a Repressor of Transcription in Plants. **Frontiers in Plant Science** 8: 2041.

Fornara F, Panigrahi KCS, Gissot L, Sauerbrunn N, Rühl M, Jarillo JA, Coupland G (2009) Arabidopsis DOF Transcription Factors Act Redundantly to Reduce CONSTANS Expression and Are Essential for a Photoperiodic Flowering Response. **Developmental Cell** 17:75–86.

Garner WW, Allard HA (1920) Effect of the Relative Length of Day and Night and Other Factors of the Environment on Growth and Reproduction in Plants 1. **Monthly Weather Review** 48:415–415.

Glassop D, Rae AL, Bonnett GD (2014) Sugarcane Flowering Genes and Pathways in Relation to Vegetative Regression. **Sugar Tech** 16:235–240.

Glassop D, Rae AL (2018) Expression of sugarcane genes associated with perception of photoperiod and floral induction reveals cycling over a 24-hour period. **Functional Plant Biology** 46:314.

Hanano S, Goto K (2011) Arabidopsis TERMINAL FLOWER1 Is Involved in the Regulation of Flowering Time and Inflorescence Development through Transcriptional Repression. **The Plant Cell** 23:3172–3184.

Hanzawa Y, Money T, Bradley D (2005) A single amino acid converts a repressor to an activator of flowering. **Proceedings of the National Academy of Sciences** 102:7748–7753.

Iskandar HM, Simpson RS, Casu RE, Bonnett GD, Maclean DJ, Manners JM (2004) Comparison of reference genes for quantitative real-time polymerase chain reaction analysis of gene expression in sugarcane. **Plant Molecular Biology Reporter** 22:325–337.

- Jackson SD (2009) Plant responses to photoperiod. **New Phytologist** 181:517–531.
- Jain M, Nijhawan A, Tyagi AK, Khurana JP (2006) Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. **Biochemical and biophysical research communications** 345:646–651.
- Karlgrén A, Gyllenstrand N, Källman T, Sundström JF, Moore D, Lascoux M, Lagercrantz U (2011) Evolution of the PEBP Gene Family in Plants: Functional Diversification in Seed Plant Evolution. **Plant Physiology** 156:1967–1977.
- Liu H, Gu F, Dong S, Liu W, Wang H, Chen Z, Wang J (2016) CONSTANS-like 9 (COL9) delays the flowering time in *Oryza sativa* by repressing the Ehd1 pathway. **Biochemical and Biophysical Research Communications** 479:173–178.
- Livak KJ, Schmittgen TD (2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2- $\Delta\Delta$ CT Method. **Methods** 25:402–408.
- Makino S, Matsushika A, Kojima M, Oda Y, Mizuno T (2001) Light Response of the Circadian Waves of the APRR1/TOC1 Quintet: When Does the Quintet Start Singing Rhythmically in Arabidopsis? **Plant and Cell Physiology** 42:334–339.
- Medeiros AL, Furtado CM, Leite FS, Souto VS, de Setta N, Van Sluys M-A, Kitajima JP, Costa APP, Benedito VA, Scortecci KC (2016) Molecular Genetic Dissection of Sugarcane Flowering under Equatorial Field Conditions. **Tropical Plant Biology** 9:252–266.
- Melloni MLG, Melloni MNG, Scarpari MS, Garcia JC, Landell MGA, Pinto LR (2015) Flowering of Sugarcane Genotypes under Different Artificial Photoperiod Conditions. **American Journal of Plant Sciences** 06:456–463.
- Mitchell AL, Attwood TK, Babbitt PC, Blum M, Bork P, Bridge A, Brown SD, Chang H-Y, El-Gebali S, Fraser MI, et al. (2019) InterPro in 2019: improving coverage, classification and access to protein sequence annotations. **Nucleic Acids Research** 47:D351–D360.
- Mizuno T, Nakamichi N (2005) Pseudo-Response Regulators (PRRs) or True Oscillator Components (TOCs). **Plant and Cell Physiology** 46:677–685.
- Moraes MAFD de (2007) Introdução - As profundas mudanças institucionais ao longo da história da agroindústria canavieira e os desafios atuais. **Economia Aplicada** 11.
- Nakamura Y, Kato T, Yamashino T, Murakami M, Mizuno T (2007) Characterization of a Set of Phytochrome-Interacting Factor-Like bHLH Proteins in *Oryza sativa*. **Bioscience, Biotechnology, and Biochemistry** 71:1183–1191.
- OCDE-FAO. (Org.). Chapter 5. Sugar. In: **OECD/FAO (2019): OECD-FAO Agricultural Outlook 2019-2028**. Paris: OCDE Publishing, 2019. Cap. 5. p. 139-139. Disponível em: <<https://doi.org/10.1787/19991142>>. Acesso em: 01 nov. 2019.

Oda A, Fujiwara S, Kamada H, Coupland G, Mizoguchi T (2004) Antisense suppression of the Arabidopsis PIF3 gene does not affect circadian rhythms but causes early flowering and increases FT expression. **FEBS Letters** 557:259–264.

Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. **Nucleic Acids Research** 29:45e – 45.

Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper – Excel-based tool using pair-wise correlations. **Biotechnology Letters** 26:509–515.

Premier biosoft (California). Bioinformatics Company. **Netprimer**. 1994. Disponível em: <<http://www.premierbiosoft.com/index.html>>. Acesso em: 13 set. 2018

Putterill J, Robson F, Lee K, Simon R, Coupland G (1995) The CONSTANS gene of arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. **Cell** 80:847–857.

Rodrigues JD (1995) **Fisiologia da cana-de-açúcar**. Botucatu: UNESP, 1995. 100p. (Apostila).

Ruijter JM, Ramakers C, Hoogaars WMH, Karlen Y, Bakker O, van den Hoff MJB, Moorman AFM (2009) Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. **Nucleic Acids Research** 37:e45–e45.

Sandercock DA, Coe JE, Di Giminiani P, Edwards SA (2017) Determination of stable reference genes for RT-qPCR expression data in mechanistic pain studies on pig dorsal root ganglia and spinal cord. **Research in Veterinary Science** 114:493–501.

Silver, N., Best, S., Jiang, J. *et al.* (2006) Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. **BMC Molecular Biol** 7:33

Song YH, Shim JS, Kinmonth-Schultz HA, Imaizumi T (2015) Photoperiodic Flowering: Time Measurement Mechanisms in Leaves. **Annual Review of Plant Biology** 66:441–464.

Tapia-López R, García-Ponce B, Dubrovsky JG, Garay-Arroyo A, Pérez-Ruíz R V., Kim S-H, Acevedo F, Pelaz S, Alvarez-Buylla ER (2008) An AGAMOUS -Related MADS-Box Gene, XAL1 (AGL12), Regulates Root Meristem Cell Proliferation and Flowering Transition in Arabidopsis. **Plant Physiology** 146:1182–1192.

Tsuji, H.; Taoka, K.; Shimamoto, K. (2013) Florigen in rice: complex gene network for florigen transcription, florigen activation complex, and multiple functions. **Current Opinion in Plant Biology** 16:228–235.

Vandesompele J, Preter K De, Poppe B, Roy N Van, Paepe A De (2002) Accurate normalization of real-time quantitative RT -PCR data by geometric averaging of multiple internal control genes. **Genome Biology** 3:1–19.

Vicentini R, Bottcher A, Brito M dos S, dos Santos AB, Creste S, Landell MG de A, Cesarino I, Mazzafera P (2015) Large-Scale Transcriptome Analysis of Two Sugarcane Genotypes Contrasting for Lignin Content (S Amancio, Ed). **PLOS ONE** 10:e0134909.

Xie F, Xiao P, Chen D, Xu L, Zhang B (2012) miRDeepFinder: a miRNA analysis tool for deep sequencing of plant small RNAs. **Plant Molecular Biology** 80:75–84.

Wolabu TW, Tadege M (2016) Photoperiod response and floral transition in sorghum. **Plant Signaling & Behavior** 11:e1261232.

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL (2012) Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. **BMC Bioinformatics** 13:134.

Apêndice A

Tabela Suplementar 1A. Dados completos algoritmos *NormFinder* e *Bestkeeper*.

Condição Experimental		<i>NormFinder</i>		<i>BestKeeper</i>		
	Gene	Estabilidade	Gene	Desvio padrão	Coef. de corr.	[r]
Folha madura	<i>TUB</i>	0,02	<i>RPL</i>	0,94	0,55	
	<i>UBQ1</i>	0,04	<i>UBQ2</i>	1,23	0,81	
	<i>TIPS-41</i>	0,04	<i>TUB</i>	1,66	0,87**	
	<i>EF1</i>	0,06	<i>TIPS-41</i>	1,9	0,97***	
	<i>UBQ2</i>	0,06	<i>UBQ1</i>	2,18		
	<i>RPL</i>	0,08	<i>GAPDH</i>	2,91	0,18	
	<i>GAPDH</i>	0,11	<i>EF1</i>	3,86	0,98***	
	<i>25SrRNA1</i>	0,18	<i>25SrRNA1</i>	4,59	0,35	
Melhor par	<i>RPL/EF1</i>	0,02	<i>N.A</i>	<i>N.A</i>	<i>N.A</i>	
Gene mais estável	<i>TUB</i>	0,02	<i>RPL</i>	0,94	0,55	
Spindle Leaf	<i>TIPS-41</i>	0,04	<i>UBQ2</i>	1,53	0,81**	
	<i>EF1</i>	0,04	<i>RPL</i>	2,25	0,82**	
	<i>GAPDH</i>	0,06	<i>UBQ1</i>	2,78	0,84**	
	<i>UBQ1</i>	0,06	<i>EF1</i>	2,98	0,89**	
	<i>RPL</i>	0,06	<i>TIPS-41</i>	3,16	0,96***	
	<i>UBQ2</i>	0,07	<i>GAPDH</i>	3,14	0,95***	
	<i>TUB</i>	0,08	<i>TUB</i>	3,73	0,92***	
	<i>25SrRNA1</i>	0,15	<i>25SrRNA1</i>	4,74	0,96***	
Melhor par	<i>UBQ1/GAPDH</i>	0,01	<i>N.A</i>	<i>N.A</i>	<i>N.A</i>	
Gene mais estável	<i>TIPS-41</i>	0,04	<i>UBQ2</i>	1,53	0,81**	
Sétima coleta	<i>RPL</i>	0,01	<i>UBQ2</i>	1,27	0,34	
	<i>EF1</i>	0,03	<i>TIPS-41</i>	0,99	0,56	
	<i>TIPS-41</i>	0,05	<i>EF1</i>	0,85	0,85**	
	<i>UBQ1</i>	0,05	<i>UBQ1</i>	1,66	0,55	
	<i>25SrRNA1</i>	0,05	<i>25SrRNA1</i>	0,88	0,65	
	<i>TUB</i>	0,07	<i>RPL</i>	0,96	0,99***	
	<i>UBQ2</i>	0,08	<i>TUB</i>	2,67	0,84**	
	<i>GAPDH</i>	0,09	<i>GAPDH</i>	3,16	0,85**	
Melhor par	<i>UBQ1/TUB</i>	0,02	<i>N.A</i>	<i>N.A</i>	<i>N.A</i>	
Gene mais estável	<i>RPL</i>	0,01	<i>UBQ2</i>	1,27	0,34	
Décima terceira coleta	<i>UBQ2</i>	0,02	<i>UBQ1</i>	0,65	-0,76	
	<i>RPL</i>	0,05	<i>25SrRNA1</i>	0,37	-0,66	
	<i>TIPS-41</i>	0,06	<i>EF1</i>	1,59	0,067	
	<i>TUB</i>	0,06	<i>UBQ2</i>	1,26	0,94***	
	<i>EF1</i>	0,06	<i>TIPS-41</i>	1,99	0,64	
	<i>GAPDH</i>	0,06	<i>TUB</i>	2,28	0,83**	
	<i>UBQ1</i>	0,06	<i>GAPDH</i>	2,21	0,96***	
	<i>25SrRNA1</i>	0,06	<i>RPL</i>	2,09	0,89**	

Melhor par	<i>RPL/EF1</i>	0,02	<i>N.A</i>	<i>N.A</i>	<i>N.A</i>
Gene mais estável	<i>UBQ2</i>	0,02	<i>UBQ1</i>	0,65	-0,76
DC	<i>UBQ1</i>	0,03	<i>RPL</i>	0,91	0,27
	<i>EF1</i>	0,05	<i>UBQ2</i>	0,39	0,67
	<i>TUB</i>	0,05	<i>GAPDH</i>	2,48	0,35
	<i>TIPS-41</i>	0,06	<i>TUB</i>	2,12	0,79*
	<i>UBQ2</i>	0,07	<i>TIPS-41</i>	2,21	0,75*
	<i>RPL</i>	0,08	<i>EF1</i>	3,5	0,94***
	<i>GAPDH</i>	0,11	<i>UBQ1</i>	3,19	0,98***
	<i>25SrRNA1</i>	0,2	<i>25SrRNA1</i>	5,08	0,91**
Melhor par	<i>RPL/EF1</i>	0,03	<i>N.A</i>	<i>N.A</i>	<i>N.A</i>
Gene mais estável	<i>UBQ1</i>	0,03	<i>RPL</i>	0,91	0,27
DL	<i>TIPS-41</i>	0,05	<i>UBQ1</i>	1,77	0,81**
	<i>GAPDH</i>	0,05	<i>UBQ2</i>	1,97	0,81**
	<i>EF1</i>	0,06	<i>RPL</i>	1,81	0,83**
	<i>RPL</i>	0,06	<i>TIPS-41</i>	2,71	0,94***
	<i>UBQ1</i>	0,06	<i>EF1</i>	3,4	0,85**
	<i>UBQ2</i>	0,06	<i>GAPDH</i>	3,66	0,99***
	<i>TUB</i>	0,07	<i>TUB</i>	4,06	0,94***
	<i>25SrRNA1</i>	0,12	<i>25SrRNA1</i>	4,25	0,98***
Melhor par	<i>UBQ1/TUB</i>	0,02	<i>N.A</i>	<i>N.A</i>	<i>N.A</i>
Gene mais estável	<i>TIPS-41</i>	0,05	<i>UBQ1</i>	1,77	0,81**

Tabela Suplementar 2A. Dados completos algoritmo *RefFinder*.

Condição Experimental	RefFinder									
	NormFinder		BestKeeper		GeNorm		Delta CT		Ranking	
	Gene	Estab.	Gene	Estab.	Gene	Estab.	Gene	Estab.	Gene	Estab.
Folha madura	<i>TUB</i>	0,62	<i>RPL</i>	0,875	<i>TUB</i>	1,246	<i>TUB</i>	2,157	<i>TUB</i>	1,316
	<i>TIPS-41</i>	0,99	<i>UBQ2</i>	1,281	<i>TIPS-41</i>	1,246	<i>TIPS-41</i>	2,446	<i>TIPS-41</i>	2
	<i>UBQ1</i>	1,49	<i>TUB</i>	1,75	<i>UBQ2</i>	1,647	<i>UBQ1</i>	2,698	<i>UBQ2</i>	3,13
	<i>UBQ2</i>	1,91	<i>TIPS-41</i>	1,938	<i>RPL</i>	1,728	<i>UBQ2</i>	2,8	<i>RPL</i>	3,162
	<i>RPL</i>	2,13	<i>UBQ1</i>	2,125	<i>UBQ1</i>	2,013	<i>RPL</i>	2,864	<i>UBQ1</i>	3,873
	<i>EF1</i>	2,37	<i>GAPDH</i>	2,875	<i>EF1</i>	2,352	<i>EF1</i>	3,089	<i>EF1</i>	6,236
	<i>GAPDH</i>	3,42	<i>EF1</i>	3,75	<i>GAPDH</i>	2,671	<i>GAPDH</i>	3,816	<i>GAPDH</i>	6,735
	<i>25SrRNA1</i>	3,55	<i>25SrRNA1</i>	4,5	<i>25SrRNA1</i>	2,964	<i>25SrRNA1</i>	3,846	<i>25SrRNA1</i>	8
	N,A	N,A	N,A	N,A	N,A	N,A	N,A	N,A	N,A	N,A
Melhor par	<i>TUB</i>	0,623	<i>RPL</i>	0,875	<i>TIPS-41</i>	1,246	<i>TUB</i>	2,157	<i>TUB</i>	1,316
Folha imatura	<i>TIPS-41</i>	0,364	<i>UBQ2</i>	1,625	<i>EF1</i>	1,188	<i>TIPS-41</i>	2,03	<i>TIPS-41</i>	1,57
	<i>GAPDH</i>	1,512	<i>RPL</i>	2,375	<i>TIPS-41</i>	1,188	<i>GAPDH</i>	2,43	<i>EF1</i>	2,45
	<i>EF1</i>	1,727	<i>UBQ1</i>	3	<i>UBQ1</i>	1,21	<i>EF1</i>	2,45	<i>GAPDH</i>	3,16
	<i>UBQ2</i>	1,912	<i>EF1</i>	3	<i>25SrRNA1</i>	1,59	<i>UBQ1</i>	2,64	<i>UBQ2</i>	3,31
	<i>RPL</i>	2,187	<i>GAPDH</i>	3,125	<i>GAPDH</i>	2,11	<i>UBQ2</i>	2,66	<i>UBQ1</i>	3,66
	<i>25SrRNA1</i>	2,375	<i>TIPS-41</i>	3,125	<i>UBQ2</i>	2,314	<i>RPL</i>	2,82	<i>RPL</i>	4,74
	<i>TUB</i>	2,606	<i>TUB</i>	3,75	<i>RPL</i>	2,476	<i>25SrRNA1</i>	2,91	<i>25SrRNA1</i>	6,29
	<i>UBQ1</i>	2,984	<i>25SrRNA1</i>	4,75	<i>TUB</i>	2,626	<i>TUB</i>	3,08	<i>TUB</i>	7,74
	N,A	N,A	N,A	N,A	N,A	N,A	N,A	N,A	N,A	N,A
Melhor par	<i>TIPS-41</i>	0,364	<i>UBQ2</i>	1,625	<i>EF1</i>	1,188	<i>TIPS-41</i>	2,03	<i>TIPS-41</i>	1,57
Sétima coleta	<i>RPL</i>	0,231	<i>EF1</i>	0,875	<i>25SrRNA1</i>	0,518	<i>EF1</i>	1,557	<i>EF1</i>	1,316

	<i>EF1</i>	0,231	<i>25SrRNA1</i>	0,938	<i>TIPS-41</i>	0,518	<i>RPL</i>	1,582	<i>25SrRNA1</i>	2,213
	<i>TIPS-41</i>	1,149	<i>RPL</i>	1	<i>EF1</i>	0,622	<i>25SrRNA1</i>	1,782	<i>RPL</i>	2,632
	<i>25SrRNA1</i>	1,207	<i>TIPS-41</i>	1	<i>RPL</i>	0,743	<i>TIPS-41</i>	1,792	<i>TIPS-41</i>	2,632
	<i>UBQ1</i>	1,816	<i>UBQ2</i>	1,156	<i>UBQ1</i>	0,949	<i>UBQ1</i>	2,223	<i>UBQ1</i>	5,233
	<i>UBQ2</i>	1,847	<i>UBQ1</i>	1,75	<i>UBQ2</i>	1,367	<i>UBQ2</i>	2,427	<i>UBQ2</i>	5,733
	<i>TUB</i>	2,019	<i>TUB</i>	2,625	<i>TUB</i>	1,73	<i>TUB</i>	2,598	<i>TUB</i>	7
	<i>GAPDH</i>	3,359	<i>GAPDH</i>	3,125	<i>GAPDH</i>	2,193	<i>GAPDH</i>	3,581	<i>GAPDH</i>	8
Melhor par	N,A	N,A	N,A	N,A	N,A	N,A	N,A	N,A	N,A	N,A
Melhor gene	<i>EF1</i>	0,231	<i>EF1</i>	0,875	<i>TIPS-41</i>	0,518	<i>EF1</i>	1,557	<i>EF1</i>	1,316
	<i>UBQ2</i>	0,704	<i>25SrRNA1</i>	0,219	<i>RPL</i>	1,302	<i>UBQ2</i>	1,872	<i>UBQ2</i>	1,732
	<i>GAPDH</i>	1,346	<i>UBQ1</i>	0,656	<i>GAPDH</i>	1,302	<i>GAPDH</i>	2,224	<i>GAPDH</i>	2,3
	<i>25SrRNA1</i>	1,413	<i>UBQ2</i>	1,125	<i>UBQ2</i>	1,546	<i>25SrRNA1</i>	2,264	<i>25SrRNA1</i>	2,59
Décima terceira	<i>RPL</i>	1,825	<i>EF1</i>	1,625	<i>TIPS-41</i>	1,786	<i>RPL</i>	2,467	<i>RPL</i>	3,13
coleta	<i>TIPS-41</i>	2,05	<i>TIPS-41</i>	2	<i>25SrRNA1</i>	2,038	<i>UBQ1</i>	2,625	<i>UBQ1</i>	4,356
	<i>UBQ1</i>	2,058	<i>RPL</i>	2,156	<i>UBQ1</i>	2,162	<i>TIPS-41</i>	2,675	<i>TIPS-41</i>	4,949
	<i>TUB</i>	2,362	<i>GAPDH</i>	2,25	<i>TUB</i>	2,315	<i>TUB</i>	2,85	<i>EF1</i>	6,727
	<i>EF1</i>	2,653	<i>TUB</i>	2,375	<i>EF1</i>	2,508	<i>EF1</i>	3,085	<i>TUB</i>	7,238
Melhor par	N,A	N,A	N,A	N,A	N,A	N,A	N,A	N,A	N,A	N,A
Melhor gene	<i>UBQ2</i>	0,704	<i>UBQ1</i>	0,656	<i>RPL</i>	1,302	<i>UBQ2</i>	1,872	<i>UBQ2</i>	1,732
	<i>UBQ1</i>	1,614	<i>UBQ2</i>	0,625	<i>UBQ1</i>	1,035	<i>UBQ1</i>	2,868	<i>UBQ1</i>	1,565
	<i>UBQ2</i>	1,693	<i>RPL</i>	0,875	<i>EF1</i>	1,035	<i>UBQ2</i>	2,934	<i>UBQ2</i>	2
	<i>TUB</i>	1,769	<i>TIPS-41</i>	2,094	<i>TIPS-41</i>	2,181	<i>EF1</i>	2,995	<i>EF1</i>	3,201
DC	<i>TIPS-41</i>	1,93	<i>TUB</i>	2,25	<i>UBQ2</i>	2,551	<i>TUB</i>	3,062	<i>TIPS-41</i>	3,834
	<i>EF1</i>	1,984	<i>GAPDH</i>	2,375	<i>TUB</i>	2,695	<i>RPL</i>	3,144	<i>TUB</i>	3,936
	<i>RPL</i>	2,306	<i>UBQ1</i>	3,25	<i>RPL</i>	2,75	<i>TIPS-41</i>	3,15	<i>RPL</i>	4,356
	<i>GAPDH</i>	3,604	<i>EF1</i>	3,5	<i>GAPDH</i>	3,001	<i>GAPDH</i>	4,047	<i>GAPDH</i>	6,435
	<i>25SrRNA1</i>	3,908	<i>25SrRNA1</i>	5	<i>25SrRNA1</i>	3,299	<i>25SrRNA1</i>	4,196	<i>25SrRNA1</i>	8
Melhor par	N,A	N,A	N,A	N,A	N,A	N,A	N,A	N,A	N,A	N,A
Melhor gene	<i>UBQ1</i>	1,614	<i>UBQ2</i>		<i>UBQ1</i>	1,035	<i>UBQ1</i>	2,868	<i>UBQ1</i>	1,565

Sequência dos transcritos utilizados para desenhar os pares de *primers* dos genes diferencialmente expressos:

>transcript_137897_CONSTANS-like 9:

CCAAAACAACCAGAAGTTTCAACCAAAAAAAAAAGAATAGGACAAAATTTTACCCTCCAATA
 TTGCGGTTTGGTACAAGAGGACGTATTTGAGGTTTGAGCACATGCTCACTTGCTCAGTAGTCGGT
 ACACAATACAGATAATAAATAACATCGAACCAGTCTGGAGGTTGGTTTGATCAGAGTGAGCACAG
 AAAGTTGGTGTCCACTGTCCAGTTACATGGGTTCCATCACCAACTTAAAAAGATCAAAGTACCAC
 CAGTTTGACTCCGACCTGGTAATTACACTCATAATAGCTGAAACACTCCAGATCCCAGCAAATTTT
 ACATCATCATAATAATAGAAGGAACTTCACATGTCACATCTAAAGCCATCTTTGTCTAATCTGGAC
 GAACACCTCCAAAAACACCGCAGCCGCCAGCCCATGTACATGGACCGATGCGAGTCAGCAGAAT
 GGAGATCAAATGCAACAAATCTGTTAGCTGGGTTGAGGCTAATTCACAAAAGGTGAAGAATGA
 CTCAAATCGCTCCAGCAAAAAGAAACGGAAGATGCGGTTTCCATCCACTGTACAGGAGAAAAAG
 GAGGATCCTCATACAAAACCTTCTGAAACATCCAAATATTTTTGTATTTTTTATCATGCTGACGGT
 TTTCTCCCTTCTTTTTTGTCTTTTTCGAAGGCGCAATCTGGCCCAATAGCTTGGATTGCCAGTGCT
 TTTGCTCTAGCTGCAACTAGCCTTGGTTTAAGTTTGCTACTAATCTAAAACGCTTCAGTAACTCCTT
 GTTTGGCAGAGAGGATCGTAGTCATATGCTTCACCAGCCTTAACAAATCGTCCCTTGACCCTCTTT
 CGCACATCTGCTCTAGCCTTGCGAGAAGCATACCTGATCTTCTGTGCAACTTTCTTCTCTTCTC
 TTCTCCTTGTATCGTGTGATAGCGCTGTCTCTACTACCTCCAGCAATTGATCCCTCAGGGCCAGG
 AGGCTGCCAAGGAGGCTCACCCATAAGGAGCAGTGATGATACCACACAATCTTGGTGGTCTCCA
 AACTACTCTCACCAAGTCAAACCGGAAAAGGAAAGAGACAGACTAGACCTAGCCTGCCTCAAAGT
 GATACAAACACTGGAATCACCTTTCACTCCTGGGTTGACATCCCAGAGTCAGCAGATACCGCAT
 TGCTATTTGCAGTTGCTTCAACTTGGATTGCTCGGTGGAATCCCAGCAGGTGCTTCCTTCACTT
 CAAAGTAACTGTCTATTCCAGCATCATCAAACAGTTGCTCTGTTTGAATATGAGAGGTACCAAAGA
 GTTCTTCGTAATTTTCGAAAGCAAGGTCCACATCATCCACACAAAATCTTCATATATGCTGTCCTT
 ACTGAACATATCTTTATCTGGTGTATAGGGCACCTTGTCTGGTGTGGAATCCACAGATCCAGCTG
 TCTGATCAGTAGCAAGTGAAGCAGGTTCACTCCAGCTCCTGAAGAGGAACCTACCAAAGGTTTG
 TCACAAGTACCGAGATCACTTATAAGTGTGCACTAGCTATATCTAACATGTCCCCTGCAGCATT
 TCTTGATTGCTGACACCGCTGTCGCTGATGCTCATCATGCTTATACCTTGCTCACAGTTCGGCTCA
 GGTGCTACATTAGGAATATCTGAAACGAACGACCATATTCTAGAAAGTTCTGCAGATGATGGACAT
 CCTGAGTAGCAATTTATGGTCTGCCTTTTGTGTCCGGCAGCTGAGGATCCTGCAATGTGGCCATT
 CCAATCACAAATTTTGGCAAAGCGAGGCATTCTCTGCTAGACAGCGGACCATAGCAGGCTGTGAA
 GCACACCGGTCACATAAAAGGGTCCTTGTATGGCGCCGAGAAAGTGCAATTAGCTGAATGCACATT
 ACGGTCGCATGATAAACACAAGGATGCTGCATCCGATCGGCAGTAAACCATGGACCTCTGCTCC
 CCACAGAAATCGCAGAGAGCACCCATGACCTTTCTATCAGCAAACCTTCAAACCAGCTTCAGGAG
 ATGTATGATGCCAACAACTCTTCTTCACAGTGTTTCCCGATCAATAGTTAAGGATCCTGTGCAGT
 ATCGATAAGCAGCTATTAGAAAGAATTACTACTGCATCAGTAATCGTTCCGAATCTTTTACAACAG

GAATAAGAAAATTGTTATAAGTACTACTATTTTATGAACTAGTACCCATACATGGTGTGAAAAAAT
 GTACATGATAGTATGACACTGAGTCACTGACGTGATGACAGAAATTTGTTACTGGCAGATACCCT
 GACCATGTCACCCACCCAAACAGCAGGTTGCTTTTTTTTTTCAACTTCATCCAAGAGAAACACTCAA
 ACATGTTCCACACATCAACATCTTACATCAGCTACCAGCAATATAAAATCATGGATAATATAGATTT
 ACTTCCTAATTAGGACATATAACCACCATCTGCATTGAGACAATTCTAGGTATCCATTGACTCATTA
 ATCCGTGGACTCCGCCTCAATTTGCCGACCGGGCATAAAGAACTGCGAAACCCACTACAAAAGT
 GCAGGCAACTAGTCTTGGAATCTTCTGTTCTCCGTGCATATTATAATTCAAAGCCCCCACTCACTG
 CTTGCAAGAACACAAGTATATATTCCACGAGGTCACTAGGTACTACGAATCCTAGAAGAGATAATG
 TGTAATAAAAAATAGAAGCCAACCATCACAGACACAAATCATCGTCATCCCTGACACTGGTTTCAT
 TAGAAAATCCTATTCTATAATATCCTACTTTTGAAGCAACCAATGACAGCACGCACAATACTATG
 CTAAGGAGGAAACATATGAAACATGAATGATTGGACCAGTGAAGTGAGCCAACAAACACAGGTCTG
 GTGATAATTTTATTTAAGTTATACTATTCCGAATAAGCAGAGAGGGTTACTTTTCCACCAGCAAAAT
 AATCTTAGAGTTGTCTAGAAGATTTGAGCTAGAAAAAAAATCTGATGTGTGAAAAGCAAGAAGTA
 AAGTTGACTTTATTGGAAAAGTATTTGGGAAGGAAAAATGCAAATATTTTCAGGAAGGGTGCTC
 ATACGGTCATACGAATGCATAACAAAAAGATTACCCTTATATTAAGAAAGATTATAAAGTATGAAC
 ATATCTTAAAGCAGTAGATTACATCTTTGCACCTAAGACTGGATACAGTTACAAGACAAACCAAA
 TTTTGACGATTCATATGAGAAACAAAAACATAATGTTTCTTAGAAGAAAAAATAAAATCTTGTAT
 GAAAAAATAAGATATGACTTTGAGCACTGCAGAAGCAGTAGAGATTAGAGCGAGAGGAGGCAATA
 CGATGATCGTTTCAGATCAAGGGGGAAAAAACCAGCACAAAGCAGTTTTCTTACCGTATTCATGTCT
 CCCGTGTAAGTGAACGACTGAAACGAGTACAAAAAGACAACAAAGCAAAAGAAAAACATA
 ACCCAAGTGACGAAGAAAAATACAAACAGACATTACTAACACCCTTAATCGGCATCAAACCAACA
 AAAATCCAAGGCAGCAGTATCTGGATCACCGAAAAGAGAAGGAAACGGGGAACGCGATAAACCG
 GTCGGTAGAAAAAAAAGATACGTATTTTGCTGAATCAGAGCTGGAACCAGGACGTTGGGA

>transcript_165796_PRR1:

AGCGAAGCACCGCAGCAGAGAAGAGCGGACGCCACCGCCGCCCTCAATAGATTTGGCGCGGA
 TAATGGTGCGGGCCGGACCGCGACGGGCGGAGCAGGGCGCCGCGGCCCCGGCGGCCGCGAG
 CTGATAGTAGTATCCCTGCTAACCAGCCATCTGCAGGCACGCGCGGTGGTAGCCCGCGAGGGC
 CCCGGGTTGATTGGCGGTACCGGCGGCCGCGGGGATGGAGGAGGAGGCGTCTGCAGATCCAC
 CGACCGCTTCGTCGTCTCTGCTCTGTGTCTCTGGTTCCGGCCGCTCCTGAATCCGACGCTTCGG
 TTCGTGTCCCCTTGCCGGAGGCCGGAGGCGTGGTTGGTGTGTCTGGGGGTTTGGATATGTTG
 GCGGCGGCGAGGGGGATCGTGTGGGCAGCGCCGGCGGAGGCGGTTTGGGGTTGGGGGAGG
 CCAGCAGTTTGTGGACCGGAGCAAGGTGAGGATTCTTCTCTGCGACGGCGACGCCACCAGCTCC
 CGGGAGGTGCTGCGGCTCCTCTGCAACTGCTCATACCATGTAACCTGCGCCAAGTCTCCGCGGC
 AGGTGATCAATATTCTCAACTACGAGGGCGGCGAGATCGACATCATCCTGGCTGAGGTTGATCTG
 CCGGTCTCCAAGTGCTTCAAGATGCTCAAGTACATTGCCAGGAACAAGGACCTGCGCCACATCC
 CCATCATCATGATGTCCAACAGAGATGAGGTGTCTGTTGTTGTCAAGTGCTTGC GGCTTGGCGCG
 GCTGAGTACCTGGTGAAGCCGTTGCGCACGAATGAGCTGCTGAACCTTTGGACCCATGTGTGGC
 GGCGGAGACGGATGCTTGGTTTGCCCGAGAAAAACTTCTTCAATGACAACTTCGAGTTGGTGCTC

TCAGAACCTAGTGATGCCAATACCAATAGCACCACCCTCCTCTCAGACGAGACAGACGATAGGCC
TAAAGAAAATATGAATCAAGAAACAGGCACCTCAAATCAACTTGAGTACGAGTCTAATCCTTCTGT
TGCGGAGCCCCGACCAGAGAGACAAGATGGAGGGTCTACCAGGCTCTGTCCTAGATGCCAGTCAA
GCATCCTCTCCGGGAAGAATGTTTTACGCCCAATAAAAACTAACCTGAGGATTGCTGAGTCGTC
TGCATTTCTAGCATATGTTAAGACAAGCACCCCAACCACCAGCTCATTTGATAGCGAATTACAAAG
AGGTGGCAGTCGGTTAGATTCTTTGGATAACCAGGGTAATTGCTCTAGTGCAACCGATAGAAGTG
ACACCGGTACTGATGTAAATATTCGGAATAAAGAAGCTTTTGAGATGCCTGTGCAGTACCCTATG
GTATGCTTTTCTTCCTCTAACATACATATGGAGCGAAGCAATGAAGGCCATAATGATACTTCAGGA
ACTCCACCTGTATACCATTTCCCATTTTATTATCCAGGGATGGTAGAGCACAACATGGCTGTTTCA
TCAGTGCAAAATTTTCAAGCAAACATAAACAATGCTCAAGCACATACACCACCAACGATGCTTCCT
CAGTACAATGTTTATCCCCAATGCCATGGTTTACCTATGATATCATCATATCAGTTTAATCCTGCTG
GCATGAATATGCACTCAAGTCATTTACCAACGCAAAATGTGTGGTCATCGGCATCAAGCACACCA
ATGCCTGAGGAAACATGCAAGCCCTCTGAAAGGAGGGCTGCAGCACTTGCCAAATTCAGGCAGA
AAAGGAAAGAGCGCTGCTTTGACAAGAAGGTGAGGTATGTGAATCGGAAGAACTTGCTGAAACA
AGGCCTAGAGTGCGAGGTGAGTTTGTAGACAGGCAAGCAATACGGATATCATCAGCACTGGAG
ATGATATCTCTGAAGACGAAGACGACGATCCATCGTCCAGGGAGGTAGATATGGTTTCCTCTCCA
GAGTAGTGAGAGGTCTCTATGTGAAAAAATAGACTGCAGATGATTACTGCTCACTTTGTTACA
ACTCAGTAGACCTAGGCAATACATTCTTCTACTGGCAACAAGGCAATACACCCTTTTTGCTCACAC
ACATATAACCTTTTTGTCTCCAAACGTGTGGCTGTTACTCTGGTTACTTTAATGGAGAAATCATCT
TTTGATTAAGTGTAGTTGTAGTTTTGCGAGTTATTGATGCACAGACCAGTCTAATCAATATGGT

>transcript_137926_LHY:

TGAGGCCCGCCGTGCGTGGGAGGTGGGCGGACGGGGGAGCGAGGCGGCGGGACCGAGGTTG
CGTTGGGTTTCGATTTTCGATTCGGCTGCGTCTCCTGAGAGGACGGCGGCTGGGATTTGTTGGTGG
CGGTGTGTGCTTGTTAACTTCCACGGGCGTGAGTGTTTCCTTGGGAGAACTACTTGATCTAGCC
ATTTAGCAATTCCGTTGAGGTTGTTCTAGAAGGGATATCTTGGGTGCTATTTGATTGTTTTGGAG
CTGTGAGATTTTGTACTGGTACTGCCTGTTACCGTGCTGAACAGTCCACTCTTCAGTCAGTACTCA
TCTAAGCGAATTTACAGAAGTGTTTGGCCTTAGAAGAGTTGAAGATAAGGAGCGGAGCTAGGTTT
AACTCTTCTGACTTGTTGTTGTGTGCAACGGATCAACTACGCTGAGTTTTTCTCTACAGTAACA
GCAATTGGATCAATTCAAGTCGGTGAGTGGAATTGATCCTTTTGAAGATTCTGATCCCAGTTCAAT
TCATGACTAGAGTCGTGAATTCTATAGGTTTTCTCGTGTAATTTAACTAATTTGGCTGCTCTTT
TGGGTCTGGGCGCACAAATTTGAAATATTGAAATATTGGGAGTGCTGATGTGATATAGAGGTCTA
GGAGGAGATATGAGGTACTACTTGCCTCAGATCCAGCGTTTGGTTGGAGATATGGAAGGAGGAG
TTCTCTTGTTTATTGATGACCTCAACTCCAAGTGACAAAGCGAACAGCTGTTGCACATTTCCCTTC
TGTTTCTGTTTATTTTTGCCGGATTTGGAGATGGAGGTGAATTCCTCTGGTGAGGAAACGGTGGT
CAAGGTAAGAAAGCCATACACAATAACAAAGCAGCGGGAGCGGTGGACAGAGGCTGAGCACAAA
CGGTTCTTGAAGCCTTGAACTTTATGGCAGAGCATGGCAGCGTATAGAAGAGCATGTTGGGAC
AAAGACGGCCGTGCAGATCAGAAGTCACGCTCAAAAGTTCTTCACCAAGTTGGAAAAGGAAGCAA

TGAACAATGGTACTTCTCCGGGGCAAGCCCATGACATTGACATACCTCCACCACGGCCTAAAAGA
AAGCCTAACAGCCCATATCCTCGAAAAAGTGGTCTCAGCTCCGAGACAACAACCAAAGAACTTCC
AAATGACAAGTCAACAAAACCAAATATGCCCTTGAGCAATGGAAATGTACAAATGGCAGGCGATG
CATCTCTTCAGAAATTTCAAAGGAAGGAGTTGTCTGAAAAAGGAAGTTGCTCGGAAGTTCTTAATC
TCTTCCGTGATGCCCCATCTGCATCATTTTCTTCAGTTAACAAAAGCTCTTCGAATCATGGTGCAC
CCAGGGGGGATCGAGGCAAGTAAAACAGAAATCCGAGATATGACCATCATGGAAAATAATTCTCTT
AACCCCAACACGCAAGAGGATGTAAAGGAGATCAGTGATCAGGAAATGGAAAGGCTTAATGGTAT
CCAAATCAGCTCTAAATGTGAACATTCTCATGAAGGATATTTGGATCTCTCAATGCAACAAATGAA
GCTAAAGCCAAAGTCTGTGGAGACAACATATGTGGACAAACAACTGCAAGAACTTCACACTCCC
TAGCGGAGAGAAATGGGACAGCTAGCATTCTGGTCACTGCAACTGAAGGAACTCATTCTGATCAA
ACAAGTGATCAAGTGGGAATCAATGGAAGCATGAACCCATGCATCCATCCAATGCTTTCTGCAGA
TCCAAAATTTGATAGCAGTGCCACACCACAACCTTTTCTCATAATTATGCAGCCTTTGCTCCAAT
GATGCAGTGCAACTGCAACCAAGATACCTACAGGTCATTTCGTCAACATGTCATCCACCTTCTCCA
GCATGCTTGTTTCCACATTGTTGTCAAACCCTGCCATCCATGCAGCTGCCAGGCTCGCAGCATCA
TACTGGCCAACAGCTGAAGGTAACACACCTATTGATCCGAATCAAGAAAATCCTGCAGATGGTGT
TCAAGGAAGGAACATAGGCTCTCCTCCAAGCATGGCTTCTATTGTAGCAGCTACAGTTGCTGCAG
CATCTGCATGGTGGGCAACACAAGGTCTTCTCCCTTTCTTTGCTCCACCCATGGCTTTTCCATTTG
TACCAGCTCCTAGCGCTGCCTTTCCACAGTTGATGTTCCACTACCTTCAGAGAAAGACAGAGAT
TGCCCAGTTGAAAAATGCACAGAAGGAATGCCAAGAAGCTCTAAAACAGGGACAGTTTGAAGGTTT
AAGAGTTGCCGCTTCTTCAATGTCTGATGGGAGTGGAAAAGGCGAGGTGTCTCTCCACACAGAG
TTAAAGATATCTCCTGTCCAGAATGCTGATGCCACACCTACCACAGGAGCTGATACAAGTGATGC
ATTCAGGAATAAGAAAAAGCAGGATCGCTCTTCATGCGGTTCTAACACACCTTCAAGTAGTGATGT
AGATGCGGACAATGTTCTGAGAAGGAAGACAATGCTAATGAGAAGGTGAAGCAAGCCTCCTGC
AGCAACAACCTCTTCAGCTGGTGACACTAACCACCGCAGATTTAGAAGCAGTGGAAGCACAAAGTGA
TTCATGGAAGGAAGTTTCCGAAGAGG

>transcript_112776_PIL5:

CCGAGAGGTTGTGCACTTCGGCGGCGCGGCTCCTCCGCGCCGTCGTCATCTTCTGCGG
CGGCTTGCGGGCCAACATGGCGGCGGACTCCGACTCCGCATCCTCGCTGGGGCTCTCCGAGTC
CTCGGTGTCGAGCCGCTTGCGCTTGCTGCTCCGGTTGCTGGTGTGTCGGCTCGGTGCTGGTG
GTGGTGGTGGCGGCGGCGGTGCCAAAGCAGTAGTTGGACCGCCCCGACGAGGACGCCTCGTT
GCCCCTGCCGCCCATGGCGTTCGCGTTCGGGCGCCCCGGCGCGCCACCGCGCGCTGCACCA
GCACCTGGTTGCTCCCGCAGATGCTCGACCCGATCACGCTCAGCATCGACGACGCGCCGGCCT
CCGCCGCCGCCGCGCAGCCGACTTCCCCGCGCTCCCCGCCGGACGACGAGGTCCGAGAGG
TCGCCGGCATTGTCGCCGCCGCCGTTGGCCAGGCTCATCGTCTGCTGCCTGGAGCAGGACGCG
GGCGTCGACTTGGGCGGCGGCATCAGGTCGCACGGGTCTCGGACACCGGGCACGCCCCGCG
CCTGTGCGCCTCCGCTGGAACGTCGTCCCTGCACGACTTGCCGCCGGCGGCTTCTGTGCTGGT
GCCACCAGAGGCCACCAGCGCCGCCGCCGCCACCCCTGCCGGTGCCGGTGCCAGTGCCTCGT

AGAAGAACTCCGAGAAGATGTCCTTGTCTGAGCGAGTCGGCCAGCGAGAAGGGGAACACAGGC
 CGCCCTCGTCTTCCTGCGGTACCGACGCGGGGGCTGGAGGCGCCACCACCGCAGCCTTCTCGG
 GCCTCGGCGGCACCTTCCGGTGGGTCTGGCTCTGCATGACGACATGGCCGTTGCACCAGAGCA
 GCTCAATGAGGTCATCGTCTTCGCCGAGCGGTCTGGAGGTGTCTCCCATGTTGCTCCAATCAGG
 GACGAACTGGTTCATCGCTCAGATGAGCACCACAACCATCGATGAGCAGCCAGTCACTTTGTGTT
 CGATTCAATCAGAGTGACGCTTTCTCCTTGTTCTGCTGAGAGCTTGAATCCGGCTTACCGA
 GGGGGGGCAGAGCATGTCCACAGCACAGACAGCTGCTGCAGGCCTCTGTGACTGACTGCGACT
 CTCGCTTCTTCCTATCTACCAGTATGTAGAGCCTAGAGCTGGAGCTGTGATGAGAAGCTACGGGT
 GAGGGAGAAGGAGAAGTAGATAGCGCGACGGACGGGGAAGGGAGGAAGGGGAAGGAAGAGTCG
 ATTTTATGCAGTCCACCTCACGCAGACGCAGAAGCAGCATTGCGGTGTGGCTTGAATCGGAGTG
 GATTCTCCCTTCCCCGTCCGTCTCGCTATCTACTTC

transcript_22706_AGAMOUS:

TCCTCTATCAGTGGGACTGTACTAAAATGCCATTCTGTTGATCCGCTACCACCATGCCTACGTCA
 AGCAAACACTCTGTTCAACTATCTTTTCCTGGAGAACTTCGTTAGCAGCTTTCAACATGCCTTCC
 TTGCTTTTCAGTGCTTGTATCTCTTGAATCATTATCTGCATCTTTGCAGAGCGGATGTTGTACATCC
 ATATCTCCAAGTACCTCTCTAGGGCATTCAAGTTCTTCAACATTCATGTGTTCAATTTGCCCTGTTCC
 CGTAGATGTACCTGAGTCCTTTCTGGAGCAGATTGATTTCTGTTGCAGCACCATGGTTTCTTGTT
 TTGGGTCCATTCTGTTGTGCGCGGAGCTGTGCGCAGCTTCTCCGCTGGCAGTCTTGACCTCTC
 GATCAGCTCTTCCATGGTCCCAGTGGTGGCGAGGTCGTAGAGCTTGCCGTGTGCGGAGAAGATG
 ATGATGCCGATGTGGGCGTCGCAGAGGACGGAGAGCTCCCGGGCCTTCTTGAGCAGCCCGGCG
 CGGCGCTTGCAGAAGGTCACCTGCCGTGCACCGGGTCTCGATGCGCCGCAGCTGCACCTTG
 CCCC GCGCCATGGGTGCCGATGGATTGCCTGGTCCCTTTTGCCTTGCCTTGGCTAGCCCGTCGG
 TCGTCACCTAGCAGCAGCAGCAAAAGGTTGTCGTCTTCCTCGATCCCTGCACACAACCTCCCCT
 AGTTTGGACTAGCTTCGAGGTTGGCGAGGCTCTCTCCCTTGCTATCTCCTCTCCCTTGTTGCT
 CCTCT

>transcript_48940_EID1:

GGTGATTGGCGCCGCGTCTTATTCGGCCCTTTCGCGCCGGGGACTGGCGCCCACCAGGAGCAG
 GATTGAAGGGGCTATATAGAGTTGGGCACAGTTGACCTGATCGCCTGAGACTGAAATGTCATTGG
 TGGTGAAGCGGTACCGATGTACGCACTCTGCGAGCTGTGCTTGCTCAAGGGCCATGTCAGGGA
 GGATGCGATGTATCTCGTGTTCAAACACATGAATTGGAACCCGAAGATGATTGCTATCTTATCCAG
 CGTATGTAAGTGGTTTGATGAATTTGCCAAGAGGGTGCTCTGGAAGGAATTCTGCAATGCCAGGG
 CACCAAAGATGATGAAGGATCTGCACTCTGATGGAAGCCATATTGTTGATGGCAACTGGAAGGCC
 CTAGGGAAGCTTCTTATCTATTGCGCAGGGTGCCCTAGTGGAGACCTGTTTAGTAATATCCAGGA
 TCCAATCCCAGGGCACTTTGTCTACAGGACCCGTTTTCTAGGACATTGGGGAAGAGCTTATTGG
 CGCCACAGTGCAAGTCTGATGCGTTGTATGTCTCAGAGTCTTGTGAACATCTTGACCAGGGTGAG
 GAAGGTGACTTGGGCCTCTTCCGGGGTATCTTCAAGTCATTTGCTGCATCGAAGATGAAGAAGAT
 GTTAATTGAGAGGCAGGCAAAGTTCCATCCGAATGAGATGTGTCCTTATTGTAAACAAAGCTATG

GAACTTGATGCAGCCAAATATGATACCAAGTAGTGCTGCCGTGAGATTGGGTGCAGAAGATGGTT
 CTGCTGAGTACTATGTATGCCTTAATGGGCACATCCTTGGTTCATGCACTCTGATACCCATCTCCG
 ATTCTGAGGACACAAAGGAAGAGTGATCAAGACTTGAATTAGGTTGCTGTGTCAAATCCTGTGC
 TTTTGACTAGGGCGTGAGCAGCGATAAGGACGAGCCCTGTCGTGCGCGTTGTACTTAACTGTAAC
 TGAGTACCAGTGTTCTCCTAGCTGCCTTTCTCTTTGGGGTGGGGGTTGACAGATTCTCGTAGT
 TGCTTGTAGCAGTAGAAAAAGAAGAGTTGCTGTATCGTGTGTTGCTGTATTTATGAATGCCTCATT
 CAGGTATATTTGCAGTGCACCTGCATACTGCATTTGTTACTGTGTAATGCATGTACCTTAAATGT
 ACAGACTTTTCTTGTATCCAGTACAGCCTGTTGTGCTCTGGACCTGTGGTTGTGCTC

>transcript_129242_PRR7:

GGAGACGGTGGTTTCTGACTTCCTTTCTCTACTCATACCAAATTCTTTGTTTCGTAAGTGCCTCAA
 TCTCTATTGTGCGTTTTGTGTACGGGGAATTTTCGTATATAAATCCTTATTATATGAGCTGTTTAATG
 TTTACAGGAGTCGGAAGCAACTGATTTTGAGCGTCGCAGAGCACATGGCTTAAGAATAATTTCT
 TCCAGCCCGGCTACATATTCGTAGTTGGATTCAGCATATGCATTTTTTATGCTGACATAGGATGTG
 CTCCACCATCTGTGAACTCCCTCTAGTTGCACAAACCAACCACTGCGGTGCTCTCTCTCTCTCG
 AATCAGGCTAATGGGTAGCGCTTGCCAAGCTGGCACGGACGGGTCTTCCCGCAAGGATGTGTTG
 GGGATAGGGAATGTCACCTTAGATAATGGCCACCATGAGGCTGAAGCTGATGCAGATGAATGGA
 GGGAAAAGGAAGATGACTTAGCCAATGGGCGCAGTGCGCCACCGGGCATGCAGCAGGTGGATG
 AACAGAAGGAGCAACAAGGACAAAGCATTCACTGGGAGAGGTTTCTACCTGTGAAGACACTGAG
 AGTCTTGCTGGTGGAGAATGATGACTCTACTCGTCAGGTGGTCAGTGCCCTGCTCCGTAAGTGCT
 GCTATGAAGTTATCCCTGCTGAAAATGGTTCACATGCATGGCGATATCTTGAAAATCTGCAGAACA
 ACATTGACCTTGTATTGACTGAGGTTTTCATGCCTTGTCTATCTGGCATCGGTCTGCTTAGCAAAA
 TCACTAGTCACAAAATTTGCAAGGACATTCCTGTGATTATGATGTCTTCAAATGACTCTATGAGTAT
 GGTGTTTTAAGTGTTTGTGCAAGGGAGCAGTTGACTTCTTGGTAAAGCCACTACGTAAGAATGAGC
 TTAAGAACCTTTGGCAGCACGTTTGGAGGCGATGCCACAGTTCCAGTGGCAGTGGAAGTGAAAG
 TGGCATCCAGACACAGAAGTGTGCCAAACCAATACTGGTGACGAGTATGAGAACGACAGTGAC
 AGCAATCATGATGATGAAGAAAATGATGACGACGACGATGACGACTTCAGTGTGCGACTCAATGC
 TAGGGATGGAAGTGATAATGGCAGTGGTACTCAAAGCTCATGGACAAAGCGTGCTGTGGAGATT
 GACAGTCCACAACCTATGTCTCCTGATCAACTAGCAGATCCACCTGATAGTACATGTGCACAAGT
 AATTCACCCCAAATCAGAGATATGCAGTAACAAGTGGCTACCGACAGCAAACAAAAGGAATGGCA
 AGAAACATAAGGAGAATAAAGATGAATCTATGGGAAGATACCTAGAAATAGGTGCTCCTAGGAAC
 TCAAGTGCAGAATATCAATCATCTCTCAATGACGTATCTGTAAATCCAACAGAAAAACGACATGAG
 ACTCACATGCCCAATGCAAATCCAAAAAGAAAATGATGGCAGAAGATGATTGTACAGACATACC
 TAGTGAAACAAATACTGAAACTGCTGATTTAATTAGCTCAATAGCCAGAAACACAGAAGGCCAACA
 AGCAGTACGAGCCGTTGATGCACCTGATGGCCCTTCCAAGATGCCCGATGGAAATGATAAGAAT
 CATGATTCTCATATCGAGGTGACACCCCATGAGTTGGGTTTGAAGAGATTGAGAACAGATGGAGC
 TACAGCCGAAATCCATGATGAGCGAAATATTCTGAAAAGATCAGATCAGTCAGCCTTACCAGGT

GCAAAACATAATATCAGTGTCTACTAGGTTATTAACATAACAGTAAGTTGCATGCTAACTCGTTA
TGATTTTTGCAAGGTACCATACATCTGTGGCTTCCAATCAAGGTGGAGCAAGATGTGGGGAAAGC
TCTTACCACAAGATAACAGTTCTGAGGCTGTGAAAACGGACTCTACATGCAAGATGAAGTCAAA
TTCAGATGCTGCTCCAATAAAGCAGGGATCCAATGGCAGTAGCAACAACGATGTGGGCTCCAGTA
CAAAGAATGTTGTTGCAAAGCCTTCGGCTAACAGGGAGAGAGTAACGTCACCATCAGCCATCAAA
TCTACCCAGCATGCCTCAGCATTTCTACTATACATAATCAAACATCACCAGCTAATCTGGTTGGG
AAAGACAAAGCTGATGAAGGAATTTCCAATGCAGTGAAAATGAGCCACCCAACAGAGGTTCCACA
AAGCTGCGTCCAGCATCATCATCAGTGCATTATTACCTCCATGTTATGACACAGAAACAGCTATC
CATTGACCGTGGATCATCAGATGTTCAAGTGTGGTTCATCAAATGTGTTTGATCCTCCTGTTGAAGG
ACATGCTGCTAACTACAGTGTGAATGGGGGTGTCTCAGTTGGTCATAATGGGTGCAATGGGCAG
AATGGAACGAGCGCTGTCCCAATATTGCAACACCAAACATAGAGAGTGTTAATGGTACCATGAG
CCAAAATATCGCTGGAGGTGGCATTGTAAGTGGGAGTGGGAGTGGCAATGATGTTTATCAGAATC
GGTTCCCCCAACGAGAAGCTGCATTGAACAAATTCAGACTGAAGCGGAAAGATCGGAACTTCGG
TAAAAGGTTTCGCTACCAAAGCAGGAAGAGGCTTGCTGAGCAGCGGCCACGGGTCCGTGGACA
GTTTGTGCGACAATCTGGGCAAGAAGATCAAGCAGCGCAAGGTTTCAAGAAAGATGACACGGACCT
TCTAATACAACATGGCATCTTCGTCGACAGATCTGAGAATTAGGCGGTTCTTCCCATTGCGCCTCT
TCATCGACAGATCTGAGAATTAAGTGGTTCCTTCCCATTGCTGATAGTAACCCATTGATTCCCAACG
TTAGAGTACTCAATAAATAAGACTCTAGTCGTCAATTTCTGTTGTGCGACTCCAGATGATCACCTA
GTGCTATGGAGTCAAACATGTAATATTTATATATGTTGAGCCTTGGCGGTTTGCAATCTTCATTTT
CCTTGCAAGTATATTCTTGGCCAAGTGTCTTAAGGAGGAGCAAACCAGCAGCGTCATTGCTATTA
TATTACTATACATGCTAATATCAGGATTGAGCTGGAATTCTGGAGTAGTTCATCCATTCATTGCAA
GACCGTGTTCCAAATCAACACAAGTAGGGGTTTGTCTGTTGCTTATAATCTGTACTTTT
TAGCTTGTTTTTT

>transcript_146760_LHP1

TTGCTTAACGACCATGCTATCAAATTTGCTGATCAAAAACTCCAATCACCAGCACAGCAGGTACA
ATGAAGCAACTAACTTCTTAGCATATAATACACACAATAGCAAGCAGTTCGTTTGGCAAACCATGC
TATCAATATTTTACTTGCACTCAGAAAGTCAGAATGACCACTAAACGAGAAGAGTGCAGACTAATC
ATCTACTTACTCAGATAGAGCGACAGATATATCTAGCACAAAAAATACTACATTTAAATACATCTC
CATCGGTTCCAAACATAGGTGCAGCACAAGCAACTGAAACACAGCGGCGTCAATGCATCCTTCTC
ACGAGGTGGGGTTATAGCGAAGGTGTTGCTCATAGTAATTAATAAGCACCAGGGGCTCCTTAGCT
TTCAACTGCTTGTCTATCCACGAATACCTCCTCTCCGTCAGACCTGAGTGCCTTGAATGTTATTGCA
ACTTGCTGCACACCATCCATAACTGTGGCAAAATAGCGAACTGGCTTGATGATCTTAATGATCTTA
GGCATATGAATTTGGTTACCAGAGTCCTCCCCCTTGTTCTTATCACCAGTTTCTCCTTCTGTGGAA
TCAACAGTCTCATTGCCTGGCTTGTCACTTGTGCGATCACGGATATCTCCTTGCTCCTGTGTTACT
TCATCCTGCTTGAACCTTCTGACACATCCAGACTTGCCTTCTTGGCACCTGTCACCTGTCCACC
CTGAGGTGGGGTTGCTCGTATTGAATTGTCCACCTTCAATGAACCATTAGCAGGGTGCTGCACAT
CTTGTTGGTCAGTCAAGCTAACAGATAAGGGAGCCCCTTGAGACAGAAACCCAACCTGAGACCACA
CCAGAACTACCCTCCTGTACTACTCTTTGCCAAGCAAGTTCAGTATGATCAAGGCCCCCAGG
TGACGTCTTGTTACTATTACAATTTGCTCTCCTACTACTTGTCTGCGAGGGCAGTTTTTTCACTTCT

GAGGCATGTATGCCTTGCATGCTCCGGGCCTCCGACCGAGGTGGTCGGCCACGTTTGCCACGA
 GAAGGGTTAGGATCGGATGTTGGAGTTGTTGCAGTCTTACGCTTTCGCTTTCGAGAGGACCTGG
 GCGACCTTGACCGCTTATTGAAAGCATCAACAATGTGGAACAGGCCTTCAAGTTTTCAAGGGGT
 TCCCATGTGTTGCACTCTCCGGCCACCCACGCCACTTGACGAGGTACTGAAGCTGGCCCTTGC
 GGAGGCGGCGGGCGGCGGATGGCCTCAATCTCGTAGAAGCCCTCGGCCAGCTTCGGCGGCTCCA
 CTGCTTCCTGTGCCGGCGTCTCCGCCTCAGCCTCGGCCTGATCCTCCCCACCAGACCCCTCCTC
 CGTCTGTCCCTCCCCCTCTTCCTCGTCGCCCTCCTCAGTCTGTTCCCTCCGCCTCTTCTTGCTCCT
 CCTCTAGAGCCGCCGTGCGATAGCCGGACTCCTTCCTGTTCCGCCCCATGGCGAATCTTGTTTC
 ACTCGACGGCCCGGCCTCACGGTCACTCGTGCTACCCAAATCAAAGGCGGATCTTGACAGATT
 TTGCTGATGAGCGCTCTGGTTGTTGGAGCCGCCTTTCCTTGGAGTGGTTCGGGAGGAACAAGGC
 TCG

>transcript_151675_PRR5:

CCGACGCCGCCGCCGGTGGAGGCGCCCCGCCGGAGGCGGAGGAGGCGCAGAGGCAGG
 CGCCGGCGCGTCGTCGTCCCCCGCGTCCGCGGCGGGGGTGAACGGGCGGGCGCTGGTGCGG
 TGGGACCAGATCCTCCCGCGCCGCTCCCTGCGGGTCCTCCTCGTCGAGCACGACGACTCCACG
 CGCCAGGTCGTACCGCGCTCCTCCGCAAGTGCGGCTACCGCGTGCGCGCGGTGCGGGACGG
 GATGAAGGCGTGGGAGGTCATGCGGGCGCGCGCCTACGCCTTCGACCTCGTCCTCACCGAGGT
 CGCCATGCCCTCGCTCTCCGGCATCCAGTGCTCTCCAGGATCGTCGCCGCCGACGAGTGCAA
 GAACATCCCCGTCATCATGATGTGTCGCCAAGATTCTATTGGCACGGTGCTCAAGTGCATGCAGA
 AGGGTGCAGTGGACTTCCTTGTAACCATTGAGAAAGAACGAGCTGCGTAACTTATGGCAACAT
 GTGTGGAGGCGGCATGTGATGAATTGCCAGACAAATGGATCAGAAAATAATGCAGCCAGCAATC
 ATATAAGTGCCAATGTTGCTAATGGGTCAAAGACAGGTGAAAACAGCGATGAGGAAAGTGATGCC
 CAAAGCTTTGGTAGCAAGAGGGGAGACTGAAATTCAGAGTGTTGAGAAGTTACCAGACATTCGTAG
 AGATGAAGATGAAGTGGCTGGCTCATCCAAAAAACTGAATCACAAAATAAGTCTTATGATGTCAG
 AGTAAACACAAAAGTGGATGCGTCAAAAGATAGCAATGGTGCTCCAAGTGGAAGTGAGAAGAATG
 TACATTCCAAATGCCTTAATGGTATTACTTCTGCAAAGTGGCTGAGCAAATTATGGATAATGCCT
 TGAGGATCACGGATGCAAGTTCACGTATCCAAGCAATCTTGGCAAAGATTTGGCTATGACTGAG
 CCAGCAGCTGACAGAAAATGCCAATCTTCAGTCATGGAAAATAATGCTGTACAGAAAATAACCTT
 GGTGAAAAGTCAAAGGGTGCTGCAATAGGTCATGCTGATTCTTGCCCTCCCAGTTTCTGGAGAC
 CAACTTGGGAAAGCAACACCATCTTAATGGTTACAAAAACCAAGAATTCAGGGAAAAAGATATCTT
 CAATCACTCGAATTCCTCTGCCTTTTCAAGATATGGCAATAAAAGGGTAGAACCATCAGGTGAGAT
 GCAGTTTTTTCTTCTCTCTGCATCACTGTGCAAGAGCATGTTTCATGGCAAGGACCCAGTTTTCCA
 ACCCAATGGGGTGTTGCTGCCTCCCAATGAACATAATACAGGTGAGAGTACAAGGCAAGCTCGA
 ATTACTTTGGACAGTAGCACGGAGGGTGCTGCCATCATGTGTTCCAGTAGTGCTAGGGAAGATG
 CTGGTGCAAGTAGTTCTTCCCACAGAAAGGACAGCATGAGCCATCCCTCGTATGGGTTTATACCT
 GTTTC AATTCCTGTTGGTCCTGGGATGCCCTACCATTACGGTGCAATTCTGCAGCCAGTATACTA
 CCCACAGGGTACTCTTATGCACTGTGGTTCAGCTGGAATCAACAAGGCAGCAATCCAGCACGCC
 TCTGGTCAACCTAATTACCATGAAGCTCCTGGTAAACCATCTCAGGTTGATGAGCACAAACAGTC
 AGAGGAAAACCAACAGTTGCATCATTCAAGACAAATTCTTCGGGAATCAGGAGAACCAACTGACA

TGGCAAGAGCTCACATGGATCATGCTAACCAGAGTGCAAGCTGCAGCCAGGATATCTGTAAAGG
 AAGTGGATGCACGGGTAGTGGTGAAGCTGATATCAATACACATACAATGGTTGCCCTAGAGAGTG
 GTAACGAAAGTGGCATCCAAAATAGTGATAGGTCTTGCCGTGAAGCCGCGTTGATGAAGTTTCGC
 ATGAAAAGGAAAGATAGATGCTTCGAGAAAAAGGTTAGATATCATAGCAGGAAGAAGCTTGCAGA
 GCAAAGGCCAAGAGTTAAGGGACAATTTGTAAGCCAAAAGCTGAAATCAGCTACAGCAACAGAAG
 CAGAAACTGACTCGTGACATGCGTTAATTTGTAAGTGAATTTTGATTCTTCCCTTCCCACCAGGT
 GTGGGCAATTGGACTGGTAAGACTGAAGGCTGGCAAAAGTTTTCCCTTCCATGGTTAGCTTTTC
 TGTACAAGGATGAACTCTTTAGTCTTTATGTCTAGTGTTTGTGGCATTCAAGCCTTATTTTGTAAG
 GTACATCAAATTGACAA

>Transcript_159838_CDF3:

TTGAGGATGTAAGTCTTCCAATGGTACAGTAATGGTGCCTGCAGCCTTGGACAGCGGCAGCACCTA
 CGTTGCCAATCATTGTACAGAATATAGAACTTTTTTTCTCTTAATATAGCATCGTTTTTGAATAATA
 TCACACAGGTTTCAGTAATACTCGATCATGTTGTCTCCTGAAAAGATCGCGAACGTGACAGAGCTG
 CTGGATTTGCTTGCAGGGCTCGAGCGGCGCCGGATATCTGCTCCTGCCTTTCAGGTTTCGACTG
 GAATGGTCTGAACATGCCTCGATCACCAGGTTCAATCCCAAGGGTGGTCCAGATTGAGCTCTTTG
 CAGCTTCGACAGGGTCATCGATCCGAAGCGTTTTGGGGATCCATAAGCATCTTCTGCCTTCTCA
 TCTCCTTGAGGTTTGGAGTCCCTGGAGTGCTTTCCTAGGACAGGAGAGCCGCTGTGCGAACATG
 TGCTGCTGCTTGTGGAGACGATGCTGACACGGTAGAGCCAGGTCCGAGCCATGGTGCCTTCCA
 TGCTCCGTTAGGCCAAGGAGTAATGAATGGCCAAACCGAAGGCGACATTACCGGAACTGGAATT
 GGACCGCAGAATCCCGGTGGCAGTACAGGCACAATTGGTGGCACATTCCATTGAACATTATTGCT
 TGTGGTGCCATTTTTTGAAGAATTTGCTGCTTCGGCTGGGGCTGGGCACACTGCTGCTGCCATG
 GTGGGAATGCCATTCCATGCTGGAGTCCATGGGTACACAAAAGGCGGACCAGGGAAGAATGGTA
 TAGGATGCACAGAAGTGACTCCATTGCTTTGCCCAACAATTCCATTTTTATGTGCACTCGCTGCTC
 CTTTTGTGATTCATTCCCTGGACCATCTGATGATGTAGTAGAAGGTGGGCACGTCTGGTTTTCTC
 CATTTTTGAGCTGTGCTGCTAAGGCAGGGTTGGAATCATAATCTGCTCTCCTCCAATCTTCAGCA
 CAGAGGCCATGGAGTTGCAGAGAGGGGAATCAGGCCCAAAGCTGACTGCCGCTTGATTTCTTT
 TATGGACAACGGAAGAGGGAAGCCTCTCCCCAGGAGTAGCTATACTGCTGCCTGGAATCAAT
 ATGCTGCAGGAATTCGCACTAGAGCTCTTGCTCTTGCGCCTTCCAGCACCGACAGGGATGTTTCT
 CATGCTTCCACCTGCAGTCCAGTACCTCTGACAACTCTTGCAAAAATGCCTTGGTTGCTTAATGTT
 GTAATTGTTATAGTAACAGAAGTTTGTATCCATGCTATTGCAACGAGGGCATGGCAGAATTTTGTG
 TGGCTTCTCAGGACCTTCTCTCACTCGGTACATCACCTCATTCTTGGCTATCTCCAAATTGGA
 CTCTGATCCCGATATTCGTCATTGCTGGGTGTCTGATTCTCACACTCACTGGAATTGTTAAGGCT
 AGAACTATTCGTCTGATCAATGTCTAATTCCTGATGTGTGCATGGCTTTTCCCTAGTGACCCTCAT
 GTCCCTCTATCTTCACATAAATCCTTTTCCGGAGACTGCGGAGGTGGGTGAGGCGCCGCCATG
 TGGGATTCCGGTGGTGGGCGAGGGGACGCGGCTCCGGCGATCTCCATACCATACAGGCAAGAT

CTGACAAGTGGAGTTGGCGACGCCAGAAAGGGTAAAAAGATTTGCCCAATTTGTGTCACGCGTTC
GTGCAACCACGCCCACACAATCTTTCAATGGCT

>Transcript_142865_CDF2:

CTCCAGCTTTGTTCTCTCTCGCCGCTTCGAGAGCGCGTGAGCTGCAGCGGCTTTGCGG
GCAGGAGCAGGACGAGGATGTCGGATCACAAGGATCCTGGGATCAAGCTCTTCGGCCGGGCGA
TCCCGTTGGGGCCGGAACCCGCGCCGGGAACACGGAAGCGGAGGACCCGCCACCCCGCCAC
GAACCGCCGCGCGATGAGCTGCAGCCACGGGCGCCGGAGGTTGCCGCGGCGGCGGCGGATGA
GGATCAACACAATGAGAAAGAAGAGAAACCTGCTAGTGAAATGGTTGACATGCCACAAGAGAAG
GGTAAAGAAATCAAGTTGACACGCCACAAGAGGAGAAGGATAATGAAATGAAGGTCGATGCGC
CACAAAAGGAACATGATGATGAAATGAAAATTGATGCACAGCAAGAGAAAAAAGATGAACAGATG
GAAGTCAATGCATCACCAATGCATGAAAATATAGAACCAGCCAATTTACCTCCCTCAGAGCATAAG
AAAGAAGACGAGGGCCTGATGAACGGTACTGAAGATAAAGCAGCATCAGATCCAAAGGGGGAGA
ATGAGAAGACATCAAATGAGGAATCAGGCCAGGACAAGGCACTTAAGAAGCCAGATAAGATTCTA
CCTTGTCTCGGTGCAACAGCATGGATACAAAGTTCTGCTATTACAACAATTACAATGTTAATCAA
CCAAGGCACTTCTGTAAGAATTGCCAGCGGTACTGGACTGCAGGGGGCACTATGAGGAATATAC
CTGTTGGTGCTGGGCGGCGCAAAAGTAAGAATGCATCATTGCACTACCGTCAATTACTGATGGCC
CCTGATTGTGTGCTGGGGTCTAGAGTGGACATATCAAAGTCAGTGCTCCCCGAAGCTCTTGTATC
TCCACCTGCCCCGATACAGCCAACCAGTCGAAATGAAACAGTCCTAAAATTTGGGCCTGAGGTGC
CACTTTGTGAATCAATGGTGTGTCAGCATTGAACATTGATGAGCAGAATGTAAACAACCCTGGATCA
GCACCAAGAGGTGAAAAATAGGGAAGATAACCCTGGATCTGGCACATCATACAATGGTGTGCCTG
AAAACATGGTTCCCGTCGATAAGAATGGAGCACCAGTTCATTGTAACGGAGTTGCCCCAGTGCCT
CAGTATTACCTTGAACCCCTTTTCATGTACCCTTGGAATGTAGGATGGAACAATGTCCCTGTGAT
GGTGCCAGGTAAAAGTATGCCCCGAATCTGCTTCTGCTTCTGATTGAGAGAGCTGCAGTACTAGTT
CAGCTCTATGGATGAACTCTCCCATGATGCCAGCCTCAAGACTTCCTAGTCCAGCATTTCATAC
CCAATTGTGCCACCTGTGCTTTGGGGCTGTTTGTCTGGATGGCCAGCCACGACTTGGAACATAC
CATGGATCAGAACTAACGGTTGCGTGTCCCCATCGTCATCAAGCAACAGCAGCTGCTCGGGCAA
TGGGTCTCCTACTCTGGGTAAACATTCCAGGGACTCTAATCCACTGAAAGAGGAGAAGAGAGAGA
AATCACTGTGGGTTCCCAAGACACTCCGGATCGATGATCCTGATGAAGCTGCAAAGAGTTCAATA
TGGGCCACCCCTTGGCATCAAACCCGGAGACCCTGGCACTTTCAAGCCTTTCCAGTCCAAGGTGCG
AGAGCAAGGGCCAGAGATCAGATGCTGCTCAGGTTTTGCAGGCGAATCCAGCAGCATTATCACG
CTCGCAATCGTTCCAGGAGAGCTCTTGACTTCATACCTACAGTCTATCTTATAGTTTTGGTACTGT
AATCCTTACATGCTAGTAGAGTTTTGCGGTATGTGAAAATATCAGCAGAGAGCTTCTTGTATAGCC
ACCTAGTGGCTCTTGCAAAGAGCTTCCATGAGAACTTTGAGATATGTATGTATTGTAAATACTTG
ACATACT

>transcript_127290_Desconhecido:

TTTTTCCAGTGCTCCATGATTAGCTGTGGACTCTCCGAAGACGCAAGAGGGCGGCGCGGACAGC
 CGGACACGCACGCCATGGCCTCGCCGGTCGTCGCAGCTATTTAACCTCGCCGCCACCTCCAGC
 CAATGGCCAGATACGCACGCGCACTCCTGCGGCGTCGAGAGGTGGAAGAAGAAGAAGGCCAAAG
 GCATGGCGAAGCTGCAGGCGGCCGCGGCTGCTGGTGATCCTGCTCGCTCTCTTTCTCCTCGT
 CGTCGCGCCCCGCTCAAGCCGCTCGTTTTACAAGGGGCTTCAGAGCGACGAGGATCGGTGAGGC
 TCCGGCGTTAGAGGGTGCTCAAGTTGCAGCAGAGGATAACTGGAGAAGCAACGCGGTGGTTGAA
 GAGATGTTTGGAAGGATGGCCCTGCAGATAACTGACTACCCAGGCTCAGGCCCGAACGATCGTC
 AACTCCCAAAGCACCGGGGCCCATGACGACAAAGACAGAGAGAACTGCTGAGCCCTGAAGATTG
 TGTACAAAGATTTCAAGCTAATCTGGGGTCTTTTGGTCTGGAACAAATTCAAGCTAACCATGGATC
 GGCTAGCGCTTTGAAAATCGTAGTAGGGACCTGCAAGTGTAACATACATGTAGTGACCTGTGC
 AGCCTGAATTCCTACGCTTTCTGTTCTGTAACACGTTATAGCTGATTAGCTGTAGACTAGTAAGAA
 TTGTATGAACAAAACTGCTGAAAATTCCTTGAAAAGACAGGTATCCAGCCGTCCAGGAGATGTG
 GGCATGTGCAGTTAGCGGGCGCTCTGAATATTCAGACCTTAAACCCTTCGATCAGTTCTTC

>transcript_177871_Desconhecido:

GGTAAAGAGCACAATGGATGACCCAGAGTTGAATGCTCCTTCCAAGAATGATCCAGAGG
 AGCCAAAGCAGGCAACAAATAGGATACTGAAGCGTACCTCATTCGCTCAAGGCACTATTGCC
 ACTGCGACCGGCTTCCTCACCGCAGCTTTCTCGGTGCGCAAAGATGCTCTGTTGCACCGGCACA
 TCCTGGTGGCCGGCGGTTGCTTCTGGTCATCGCGTACCTTTCTGCGTCTCTGCTGGTCTACCT
 GAAATTGTTCTTGTGTCAGGACACAGGCAGCTGCATAAGGGGCACATCCGGTTCATCCAGGTCCTCT
 GCATCATCGGCGGGCGCGGCTGGTCGCGACGAACTCCCTGCTGCTCCTCCTCATAAGTGAAGA
 CAGCATCTTGTTGTCTCTCAATCTGCTGCCTATACAGGGCCTCATCGGCATTCTCGCGTACCACG
 CGACGCCGACGGAGGACAGCATGCGCGACGAGGCGTTTCGAGGCGGGGATCAAGAGCGGGCCG
 AAGGTTGCCCTGTTGCGCACCGCGACGGCCTTCGCTGTGCAGACCACGCTCCTGTTGCGCTACC
 TCAACAACTCCAGCTTCCGGGCGCTGGGCCACCGGTTGACCTCTCTGTGTCCTTCCTGGCGGC
 GGCCTTCAGCGTGTTCTGCTGCTGGCCACCTGCATGCCGCTTGGGTACAGGACTAGTGCCGC
 GAGGGACAAGGTGCTGTCCCTCGTGAGGTACCTGAAGGACGCCGTCATCGCCTTGCTCGCGGT
 CACCGCGTTGACTATTGCCAAGGAGTTTCTCGGCGGCGACACCGTGCTCGCACTCTTCCCGGAG
 ATCACCGTGGCTGCCATGTATTATGCCGTGAACCTGTTCACTGACGACACTGCAGAGCAGGGCC
 AGCGTGAGATCGCTGAGAACAAAGATGGAGATGTTGCCGACTGCCGTCGTGGCCACGTTGCGCTT
 CGGCATGCTGGGTGCGGCCTACGCCGCGCTGCTTGGAACGCCGGAGTACGACATGTACACGAA
 GGCGCTGGTGTTACCTTGCTCTCCGCCGTGCTGTCCAGCCTGGGACGGGTGGCGGGGCCCT
 CTGCAGCCCCCGGCACGACAAGAACGCGGCGGCGTGCGTGGTGTTCTGAGCAACATCCTCCC
 CATCGTGGAGATGCTGGTGGCCGTCCCGCTCGCCGCAAAGGTCGCCAGTAATGCCCTCCCCAC
 GTCTTAGGAACAAAACTATCTGTATTCTGTTTAACTACGAGGCACCGTTGATTTACCTTACAGTA
 AATGTTTGTATGGCAGTCATATTATTTTGTATGTCAGTTAAACAATGTTGTTACTACTAGCAAACA
 ACCACAAACAACTAGCATAACAAAATTCTTCCAGAAAGAATGACTGCCTGAGTAAACCAGATAGAA
 GGTTC

>transcript_110622_Desconhecido:

TGCACAGCTAGCATAGCAGTGCCATTCTGGCAAGGGCGGGAGAGAGGATGAGGACGAGCAGCG
 GCGGCTTTGCCGGCCTGGTGGTGTGCGTCGCCACCGTGTGCCTGCTCCTGCCGACGGCGAGCC
 GCGCCCAGCTGAAGGTGGGGTTCTACCACACCACCTGCCCCAACGCCGAGGCGCTCGTGCGTC
 AGGTGCTCACGGCTGCTTTCGCCAACAACTCCGGCGTCGCACCCGGGCTCATCCGCCTCCATTT
 CCATGACTGCTTTGTCAGGGTACGTAAAAGACATATTATCCGCAATAAACACGCATGCAAGTGGT
 GTCGACCAATGAGTGATGGAAATATGCACTACTGGAACTCGAATACTTCCGTGGGTGTGCAATT
 TTTCTGTGCGTTATTTTAGGCACGCACGGAAAAATTCCGATTGTTCTGTGGGTTTCCAAATATACC
 CACAGAAAAATTCGAAAACCTACAGAAAAATTATAACATTTTTCTGTGCGTGAGAGTAACCACAG
 GAAAAATATAAACACACAGAAAAATTAACTTTTTCTGTGCGTTTTCTTAAGCCTCACAGAAAAAT
 TTATACACGCACGGAAATATATGTATTATTCTGTGGGTCATTTTATACACGCACAGAAAAAACATA
 CTAACACACAGAAAAATTCTAACGCGCCATCCGATTTGGCGTTTCGCACCCTCGAATTTGGCGTTG
 GCGCTTTCCGTTCTGGCGTTGGCGCGTTCTCACGAAAAGAAAGGATATTTAATGAATAAAAAA
 GGAAAAGTCCTAAACCCAACCAGCACACGCACTGCACCCACGCACGGACTCCATCCCAACACAC
 ACACGCACCGCACCCAAACACGCGCACACGCACCGCTGGGCTTCCTCACGCAGGAGCGAGCAG
 CACGCCGCGGCCGGCTTCCTCCCCAGCAGCACACCCTCCGCCGCCGCCACAGCAGCTCCTCTC
 CAATGGCGTTGGCAGCCCCTAGATCTGCGCCTCCCTAAGCATCCCACAGGCTGCCGCCAAGACT
 GCGAGCCCTCCCCTTCCCTAGGCCGCCGCGAGCTGCGCCCCCTTCGGTGCGCCGGCCTAGGTCC
 ACCGGCTGCCGGGCGCTGGGCACTGGGTGCCGGCCAGCAGCCGGTGCTCTGGGACGCATCCC
 CAAGGACCTCCCCCTCGACTGCTCGTCTCGACGCAACGACCCTCCTCTTCTGTGATCTCCTG
 CGCCCTGCTCTTCCATCCTCCAGATCTGGCCACCAGTGGGTCAAGGTAACTCCTCTCTTCACTG
 ATTCAATTTTTGTATGAGATGTGATGCCGGTTTTATCTGAGTATCCAGTTGTATGAATTTCCAGTT
 GAAGTCTGCTTCCCTCAAGATTTGGTCGCCGGTGTGTATGCTTCAAATCTGAATCGAGTTTTAG
 GTTATCTGTTGTCCATTTTTTTCATGGCAGTGACTATGAAAAGTGTAGACTAGCCAGAACCCC
 GCGGCAATTGCCGCGGGCTAGGCAGAGTTAGTTGACATAAATTTTTCTCTAAAGCATATATGC
 ATCAATTATTTGTTAGTTGACATCTCTTAAATTTGATTGAAGCAATTTCTCATTTGCTAAATTTATCT
 TGACGAACCCATAAATGCCTGAGATTGCAGCACACAGTCTTGGTCATATTGATGTACAATTATGTC
 ATGTTAAATATCTGTGCTCAATATGAACAAAGTGAATGGCAATATCCAAGATAACTCACCCTGA
 CCTATGAGGACCTGGGGACAAAACGAATTTTTGTAGCTCCAAATGGCAGATCTATGCTGATTGG
 AGCTTGATACATTTCTGATATATACAATCCCCGGTACTGAAGCAGCATAAAGTTAATAGTTAGATA
 CATTGCTAGTCATTTTGTGTTTGGTCCATAGCTGAAGACACTGCAATGGGATTCCATAGCTGAGGT
 TCGCTAGGATTAGAGTGTAGAGAAGGATAAATAAAGCATGTTTTACTTATCTATTCTGAATATATAT
 GATTCTTATTGATATTTTCATGACATGATGTATCAAGTCCGTGTGGATGTTGTCAAGGAGTACTAC
 AGCAAAGTAAACACATACAAATAAAAAATGCAGTAGCGTGCCATATTGAGCTTGAGTATGAGCAA
 TACAAGCAAGGACAGCTTGAGGGTTTCAGTGATGAGGTGTGGCTCCTCCTTTGTGATTATTGGTG
 CTCAGATGAGTACAAGGAAAAACGTAAGAGGGGCCAAATATCTCGGTCAAGTAATGAAGATATTG

CTCAAACCGAGGAGGATCTAGACCATTACAGAGACACAACAAGTACTGGAAGCTAAGTTTGGA
CCATAGAAAGCAACCCCTGTTAATACATTTGCTGCGATGAAATCTGGCTTGAAAAGCCTGGATAG
CAATGGTAGATGTGGTCCAATCAGTAGCCGCAAAGCACAAAGACGCATGGATGGGTACCTTACT
GCTGTTAAGAGGATAGAATGTACAAACGAAGAGGTTGAGATGATGCAGACATGGAAGTGGATG
GGCATGAGGTTGAGCAACAGCTATATTTGAATGGGAAGGCTCTATATGATGTGTCTAAGGGTAGC
CCGAGAAAGCATGGACGTCTTGCCATAGCAAATGGTGCGGTTAAGTCTTCAGATGTCAGGGCAG
CCGGAAGGGAAAGAAGTATGCGTCCATCAAATTCAGTGACTATGCAAAACATGTCTCGTGAAATG
GAAGAGTTACGTCGTGCAAATGGAAGGCTAGAACAAGAGAATCATGATAAGGAGAATGCACTGCA
GCAAAGCAAAGTTCTTG TAGGCTTGGTATTGAACCTTTATCAGGAGACAGGAAGGGTTGTGCCAG
AAGAAGCGTTACGACAGTTGTCCGCTACACAAGCAATTGCTACTGGCTCTTCTCATGCTGAATCT
GCCAACAATACTGATGATGTTTGCCACAACAATGGGGATCCTCATGCTGCAAATGGTGGCAATCA
AGGTTCCGACAACAATGGAGATCATGCACCGATGTGATGGTTGTGTTTAGGTGGAAAACATGGAT
GTGCTGGTGGTTTCTTCTTTTGTAGAGCTGTGCTTTTGCACATTAAGTAGGTGGCTCTAATCT
GTTAGCCTTAGTTCTATATGACACAATGATGCATGCTGATACTTCATTCATCTTGGCTGAACTTG
AACCGATGAGATGAGGTAGCAATGCTTTTGAATGTATTGATGTCGGAATATTTGGACATCATCTAT
ATGTATTAAACCTAGGTGCCGGAATATTTGGACTTCATCTTTACATGTATGCTATTTATTCTGTGAT
GATGAAATTTCTGTATGGACAGTCAGTATAATATGATGTTGCTACTGTTTTTAAATAATTATTATATC
CAATTATGGCATGGGAAAATGTATATATTCTGTCGGTGGTTTGAATTTTTCTGTGCGTGGTTTTA
AACCGACAGAATAATGGTAATTTTTCTGTGAGTACACCTAGTTTTTCTGACAGATCTACACCGACA
GAAAAACTAATTTTTCTGTGCGGTGTTGATTTTTTCTGTGCGTAACAACACACGGAAAATTGTTTT
TGATCCGGCGGAAACCAATTTTTCTGTGAGTGCTTGCCCCACGGAAATATGTTTTTCTGACGGTG
AACCCACGGAAAAATTCAAATTTTTCTGTCATCCTATTTCCGTGGGTATTTTTCTGTGCGTCACCGT
CAGAAAAACATTTTTCCGTGGGTATACGCATTTTTCCGTGGGTTTTTGACCCACGGAAAAATCCG
AGTTTCCAGTAGTGCAAGCACTGCAATAAAAATCTTCATGTCAAATAACAATAACAAGAAATTAT
AAATGAGGCTCGAGTAGCCTGAATCATAGTAACCCGGAGAGATGAGGCTTTTACAGAACTATTT
GGGATGACGCAATGAAGTTAGTATCAGTAATTTGAGATCCTTGTGATATTCTTGTGAGTGCACTAA
CCTGGTAGGAGACATATTATCAATCATATGTGTTACAAGCCAAATAAAAAAAGACTAAGAAATTTA
GAAGCTATGCTTTATGGATATTTGAAAAAAAAAACATATG