

**MÁRCIA HELENA SANTOS ESTEVES**

**MARCADORES MOLECULARES PARA DISTÚRBO FISIOLÓGICO EM  
EUCALIPTO**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Bioquímica Aplicada, para obtenção do título de *Magister Scientiae*.

Orientadora: Elizabeth Pacheco Batista Fontes

Coorientador: Pedro Augusto Braga dos Reis

**VIÇOSA – MINAS GERAIS  
2023**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade  
Federal de Viçosa - Campus Viçosa**

T

E79m  
2023 Esteves, Márcia Helena Santos, 1997-  
Marcadores moleculares para distúrbio fisiológico em  
eucalipto / Márcia Helena Santos Esteves. – Viçosa, MG, 2023.  
1 dissertação eletrônica (141 f.): il.

Inclui anexo.

Inclui apêndice.

Orientador: Elizabeth Pacheco Batista Fontes.

Dissertação (mestrado) - Universidade Federal de Viçosa,  
Departamento de Bioquímica e Biologia Molecular, 2023.

Referências bibliográficas: f. 43-47.

DOI: <https://doi.org/10.47328/ufvbbt.2023.375>

Modo de acesso: World Wide Web.

1. Marcadores genéticos. 2. Eucalipto - Fisiologia.  
3. Distúrbios fisiológicos. I. Fontes, Elizabeth Pacheco Batista,  
1957-. II. Universidade Federal de Viçosa. Departamento de  
Bioquímica e Biologia Molecular. Programa de Pós-Graduação  
em Bioquímica Aplicada. III. Título.

CDD 22. ed. 572.8


**MÁRCIA HELENA SANTOS ESTEVES**

**MARCADORES MOLECULARES PARA DISTÚRBO FISIOLÓGICO EM  
EUCALIPTO**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Bioquímica Aplicada, para obtenção do título de *Magister Scientiae*.

APROVADA: 28 de fevereiro de 2023.

Assentimento:

Documento assinado digitalmente  
 **MÁRCIA HELENA SANTOS ESTEVES**  
Data: 16/06/2023 15:16:57-0300  
Verifique em <https://validar.itl.gov.br>

---

Márcia Helena Santos Esteves  
Autora

Documento assinado digitalmente  
 **ELIZABETH PACHECO BATISTA FONTES**  
Data: 09/06/2023 21:02:34-0300  
Verifique em <https://validar.itl.gov.br>

---

Elizabeth Pacheco Batista Fontes  
Orientadora

## AGRADECIMENTOS

Primeiramente, agradeço a Deus por tudo, pela sabedoria e pelas conquistas que me proporcionou.

Agradeço os meus pais, minha irmã e meus avós que sempre estiveram ao meu lado em cada etapa da minha acadêmica e pessoal, sempre me apoiando e incentivando.

Ao meu namorado que sempre me apoiou e me deu forças durante todo o período do mestrado. Agradeço também as amigas que fiz no Bioagro, Raquel e Sâmara, que estiveram sempre ao meu lado me dando conselhos e deixando meu dia melhor.

À Universidade Federal de Viçosa e o programa de Bioquímica Aplicada, por proporcionar todo aprendizado durante o mestrado.

À minha orientadora do mestrado Elizabeth Fontes, que confiou no meu trabalho e sempre esteve disposta a me ajudar com seu vasto conhecimento.

Aos estudantes do Laboratório de Biologia Molecular de Plantas (LBMP), que contribuíram com ensinamentos em relação as técnicas executadas no laboratório. Ao Marco Aurélio que contribuiu com seus conhecimentos no projeto Eucalipto. Ao professor Pedro Augusto pela coorientação e sua disponibilidade em ajudar.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), concessão da bolsa de estudos.

À CAPES, FAPEMIG e FINEP pelos recursos que possibilitaram a execução do projeto.

À empresa BRACELL pela parceria no projeto de pesquisa realizado.

*“A persistência é o menor caminho para o êxito”.*  
(Charles Chaplin)

## RESUMO

ESTEVES, Márcia Helena Santos, M.Sc, Universidade Federal de Viçosa, fevereiro de 2023. **Marcadores moleculares para distúrbio fisiológico em eucalipto**. Orientadora: Elizabeth Pacheco Batista Fontes. Coorientador: Pedro Augusto Braga dos Reis.

O distúrbio fisiológico em *Eucalyptus grandis*, se apresenta como uma síndrome abiótica que afeta diretamente a qualidade da madeira, interferindo de forma significativa na economia brasileira, uma vez que o país é o principal produtor de madeira e celulose. Os principais sintomas relatados para esse distúrbio são a perda da dominância apical, seca do ponteiro, necrose do tronco principal e amarelecimento das folhas, sendo que alguns desses sintomas são também resultantes de outros estresses abióticos, como seca e estresse nutricional. Conseqüentemente, investigar a expressão diferencial de genes que respondem ao distúrbio fisiológico e desenvolver marcadores moleculares constituem ferramentas em potencial para melhor entendimento do distúrbio, gerando subsídios para o diagnóstico molecular, ainda na fase de muda. Para definir marcadores moleculares associados ao distúrbio fisiológico em eucalipto e analisar a presença de SNPs, o DNA dos caules de clones resistentes (1572, 1644, 1407 e 1404) e suscetíveis (520, 1296, 5204, 2361 e 5341) foi extraído e foram amplificados fragmentos sobrepostos de promotores dos genes selecionados (KASIII, KCSI, TIL1, MSH2, RPT2 e Galactose Oxidase), cuja expressão é indicativo do estado de distúrbio fisiológico. O sequenciamento de aproximadamente 2000 pb de regiões 5' flanqueadoras dos respectivos genes permitiu o alinhamento dos promotores sequenciados, possibilitando a identificação de polimorfismos (SNPs e Indels) nos diferentes genótipos, devido à variação natural. Entre os diversos SNPs e Indels identificadas, apenas dois SNPs nas posições -1366 e -1373,-1374 no promotor do gene RPT2 puderam ser associados ao caráter resistência e susceptibilidade. Experimentos estão em progresso para avaliar a capacidade de marcação genética desses SNPs com relação ao distúrbio fisiológico. Também foi identificado o transcriptoma induzido pelo distúrbio em folhas de eucalipto, sendo extraído RNA total de clones 520 sintomáticos, 520 assintomático e 1404 resistente. A qualidade do RNA extraído foi avaliada pelo Bioanalyser e submetido ao RNA-seq, pelo método Illumina. O contraste 520-SS vs 520-AS revelou uma predominância de genes regulados positivamente pelo distúrbio fisiológico relacionados com estresse abiótico. Consistente com os resultados previamente obtidos para o transcriptoma induzido pelo distúrbio fisiológico em caules de clones suscetíveis, a síndrome do distúrbio fisiológico em folhas promoveu a repressão de genes relacionados com a síntese de ácidos graxos de cadeia longa, que são essenciais para proteção e desenvolvimentos das plantas. Estes

resultados confirmam que os genes relacionados com o metabolismo de ácidos graxos podem ser utilizados como marcadores do distúrbio fisiológico. A análise dos genes diferencialmente regulados pelo contraste 1404-R vs 520-AS confirmou a repressão de genes do sistema imune da planta no clone resistente 1404, submetido às mesmas condições indutoras do distúrbio do clone suscetível 520. Esses últimos resultados corroboram o argumento de que os estímulos bióticos não são o gatilho do distúrbio fisiológico, sendo há muito tempo erroneamente considerado.

**Palavras-chave:** Distúrbio fisiológico. Eucalipto. Marcador molecular. RNA-seq

## ABSTRACT

ESTEVEES, Márcia Helena Santos, M.Sc., Universidade Federal de Viçosa, February, 2023. **Molecular markers for physiological disturbance in eucalyptus**. Adviser: Elizabeth Pacheco Batista Fontes. Co-adviser: Pedro Augusto Braga dos Reis.

The physiological disorder in *Eucalyptus grandis*, presents itself as an abiotic syndrome that directly affects the quality of the wood, interfering significantly in the Brazilian economy, since the country is the leading producer of wood and cellulose. The main symptoms reported for this disorder are the loss of apical dominance, point wilting, necrosis of the main trunk and yellowing of the leaves. Some of these symptoms also result from other abiotic stresses, such as drought and nutritional stress. Consequently, investigating differential expression of genes that respond to the physiological disorder and developing molecular markers are potential tools to better understanding the disorder, generating subsidies for molecular and more accurate diagnosis, still at the molting stage. To define molecular markers associated with the physiological disorder in eucalyptus and analyze the presence of SNPs, DNA from the stems of resistant clones (1572, 1644, 1407, and 1404) and susceptible clones (520, 1296, 5204, 2361, and 5341) was extracted. Overlapping fragments of the selected gene promoters (*KASIII*, *KCSI*, *TIL1*, *MSH2*, *RPT2*, and *Galactose Oxidase*), whose expression is indicative of the state of physiological disturbance, were amplified. The sequencing of approximately 2000-bp 5' flanking regions of the respective genes allowed the alignment of the sequenced promoters and the identification of polymorphisms (SNPs and Indels) in the different genotypes due to natural variation. Among the various SNPs and Indels identified, only two SNPs at positions -1366 and -1373, -1374 in the *RPT2* gene promoter could be associated with resistance and susceptibility traits. Experiments are in progress to assess the genetic marker capacity of these SNPs concerning a physiological disturbance. The disturbance-induced transcriptome in eucalyptus leaves was also identified. For this, total RNA was extracted from symptomatic 520, asymptomatic 520, and 1404 resistant leaves. The RNA quality was evaluated by the Bioanalyser and submitted to RNA-seq by the Illumina method. The contrast 520-SS vs 520-AS revealed a predominance of abiotic-related genes upregulated by the physiological disturbance. Consistent with the results of a previously identified transcriptome induced by the physiological disturbance in stems of the susceptible 520 clone, the state of physiological disturbance in leaves promoted the repression of genes related to the synthesis of long-chain fatty acids, which are essential for protection and development of plants. These results confirm the disorder marker nature of genes related to fatty acid metabolism. The analysis of the genes

differentially regulated by the contrast 1404-R vs 520-AS confirmed repression of genes of the plant's immune system in the resistant clone 1404, submitted to the same disorder-inducing conditions as the 520 susceptible clone. These latter results substantiate the argument that biotic stimuli are not the trigger of the physiological disorder, which has been long mistakenly considered.

**Keywords:** Physiological disorder. Eucalyptus. Molecular marker. RNA-seq

## SUMÁRIO

1. INTRODUÇÃO .....	10
2. REVISÃO BIBLIOGRAFICA.....	11
2.1 Eucalyptus grandis.....	11
2.2 Distúrbio fisiológico em eucalipto .....	13
2.3 Marcadores Moleculares .....	14
3. METODOLOGIA .....	16
3.1 Coleta do material vegetal.....	16
3.2 Extração de DNA e RNA das amostras de eucalipto .....	16
3.3 Verificação e preparo das amostras para sequenciamento do DNA dos caules de eucalipto.....	17
3.4 Sequenciamento e análise das amostras do DNA dos caules de eucalipto .....	20
3.5 Preparo das amostras de RNA de folhas de eucalipto para envio ao sequenciamento .....	20
3.6 Sequenciamento do transcriptoma - RNA (Seq).....	21
3.7 Bioinformática.....	22
3.8 Extração de RNA e síntese de cDNA para verificação de transcritos por PCR quantitativo (qRT-PCR) das folhas de eucalipto.....	22
4. RESULTADOS .....	24
4.2 Identificação do transcriptoma induzido pelo distúrbio fisiológico em folhas do clone suscetível 520 .....	30
4.3 Análise da expressão genica de genes candidatos diferencialmente expressos .....	37
5. DISCUSSÃO.....	39
5.1 Desenvolvimento de marcadores moleculares associados ao distúrbio fisiológico.....	39
5.2 Transcriptoma induzido pelo distúrbio fisiológico em folhas de eucalipto .....	40
6. REFERENCIAS .....	43
7. MATERIAL SUPLEMENTAR .....	48

## 1. INTRODUÇÃO

*Eucalyptus grandis* pertence à família das Myrtaceae, sendo uma das espécies mais cultivadas no Brasil, pois tem grande importância econômica para a produção de celulose e madeira. Alguns fatores abióticos como temperatura, condição hídrica, composição atmosférica e solo podem interferir no crescimento, desenvolvimento e qualidade dos eucaliptos cultivados, gerando prejuízos para empresas que fazem cultivo de grandes plantações de eucalipto (CÂMARA et al., 2018). Atualmente, a ocorrência do distúrbio fisiológico em eucaliptos em regiões litorâneas do estado do Espírito Santo e Bahia tem sido preocupante (SANTOS, 2021).

O distúrbio fisiológico em *Eucalyptus grandis* (DFE) é considerado uma doença que interfere na qualidade e produtividade da madeira. Esta anomalia tem sido descrita como um estado fisiológico de estresse que resulta em um conjunto de sintomas como a perda de turgor, perda da dominância apical, seca do ponteiro, encarquilhamento e amarelecimento das folhas, necrose do tronco principal, brotamento de galhos perpendicularmente ao tronco principal levando a perda da dominância apical e comprometimento do crescimento, podendo ocasionar a morte de genótipos mais susceptíveis (SANTOS, 2021).

Não se sabe ao certo o que desencadeia o aparecimento do distúrbio fisiológico, mas, é sugerido como hipótese para o aparecimento a interação de fatores ambientais adversos, como por exemplo, a influência do clima, condição hídrica, temperatura e salinidade dos solos de cultivo. Estes fatores abióticos podem desencadear o aparecimento de sintomas do distúrbio fisiológico em plantas suscetíveis, devido a modificações no metabolismo celular e expressão gênica (RODRIGUES, 2022). A realização do transcriptoma induzido pelo distúrbio fisiológico em troncos assintomáticos descartou a possibilidade de estresses bióticos como indutores do estado de distúrbio fisiológico (MARTINS, 2020).

Investigar genes diferencialmente expressos em clones de eucaliptos no estado de distúrbio fisiológico constitui uma alternativa para elucidação das bases moleculares que fundamentam o estabelecimento do estado de distúrbio fisiológico, pela análise das diferenças na expressão de genes em eucaliptos sintomáticos, assintomáticos e resistentes. Tal investigação é possível por meio da identificação do transcriptoma induzido pelo estado de distúrbio fisiológico que é baseado na tecnologia de sequenciamento de nova geração como RNA-seq. A análise bioinformática dos resultados do RNA-seq fornece informações precisas sobre a natureza dos transcritos expressos diferencialmente, possibilitando a informação sobre

o conjunto de genes regulados negativamente e positivamente (LUO; WANG; MA, 2021).

Além disso, comparações entre as sequências de DNA dos clones de eucalipto sucessível e resistentes podem contribuir para um melhor entendimento da doença. A presença de marcadores moleculares do tipo SNPs (*single nucleotide polymorphism*), podem justificar a suscetibilidade ou resistência de clones de eucalipto frente ao distúrbio fisiológico. Os SNPs são polimorfismos oriundos de alterações de uma única base no genoma e que são encontrados em regiões expressas e não expressas no genoma. Além disso, a identificação da presença de SNPs pode ser utilizada como uma forma de diagnóstico do distúrbio fisiológico, por meio do desenho de oligonucleotídeos (*primers*) que sejam capazes de amplificar a região de interesse (ZOLET, 2017).

Portanto é importante definir os marcadores moleculares associados ao estado da síndrome do distúrbio fisiológico em eucalipto, uma vez que já foi identificado previamente um conjunto de genes que são especificamente reprimidos no estabelecimento da síndrome. Além disso, é válido analisar a sequência dos promotores de genes marcadores do distúrbio fisiológico em eucalipto, na tentativa de mapear SNPs associado ao distúrbio fisiológico em clones suscetíveis. Adicionalmente, identificar o transcriptoma induzido pelo distúrbio em folhas a fim de compará-lo com aquele identificado em caules, com expectativa desses resultados permitir a identificação de genes diferencialmente expressos comumente associados ao estabelecimento do distúrbio fisiológico tanto em caules quanto em folhas.

## 2. REVISÃO BIBLIOGRÁFICA

### 2.1 *Eucalyptus grandis*

*Eucalyptus grandis*, originário da Austrália, Tasmânia e outras ilhas da Oceania, apresenta cerca de 700 espécies reconhecidas na botânica, sendo somente utilizadas 20 espécies com fins comerciais. É pertencente à família *Myrtaceae*, e ao sub-gênero *Symphyomyrtus*, se caracterizando por apresentar diversos híbridos naturais, com característica de árvores de grande porte e até mesmo arbustos. Sua descoberta é datada em 1788 por ingleses na Austrália (VALE et al., 2014).

Dentre as espécies utilizadas, o *Eucalyptus grandis*, *Eucalyptus saligna*, *Eucalyptus urophylla*, *Eucalyptus viminalis* e *Eucalyptus dunnii* são as mais utilizadas. Em relação a sua genômica, o eucalipto é diploide e possui 22 cromossomos e apresenta cerca de 600 a 700

milhões de pares bases, com um alto valor percentual de pares de GC com 57% da constituição genômica (YAN, 2019). Através do consórcio EUCAGEN foi possível concluir o sequenciamento completo do genoma do clone de *Eucalyptus grandis*, gerando 5,4 gigapares de bases obtidos por meio de bibliotecas de plasmídeos e fosmídeos (GOIS, 2021).

A introdução do plantio de Eucalipto no Brasil ocorreu em meados do ano de 1868 no estado do Rio Grande do Sul por Joaquim Francisco de Assis Brasil. A partir da introdução de seu cultivo, as plantações de eucalipto aumentaram consideravelmente no Brasil devido ao seu uso como fonte de biomassa para combustível e como matéria-prima para fabricação de celulose e papel (SANTAROSA; PENTEADO JÚNIOR; GOULART, 2014). O aumento do plantio de eucaliptos foi em decorrência de seu rápido crescimento, capacidade de adaptação às regiões de plantio e seu potencial econômico. Segundo estimativa, o Brasil no ano de 2019 apresentava cerca de 9,0 milhões de hectares de plantio florestal, desse total estimado, 7,83 milhões de hectares correspondem a plantação de eucalipto (IBA, 2020).

Em relação aos 7,83 milhões de hectares de cultivo de eucalipto no Brasil, esses estão distribuídas em diversos estados, mas é citado principalmente os estados de Minas Gerais com cerca de 24%, São Paulo com 17%, Mato Grosso do Sul com 16%. Além disso, o Brasil é conhecido como o maior exportador de celulose, exportando principalmente para China com cerca de 30% das exportações e Europa com 25% (IBA, 2020).

O eucalipto é capaz de fornecer diversos subprodutos através da utilização de suas folhas, madeira, cascas, flores, resinas e galhos. É possível obter do eucalipto óleos essenciais utilizados na indústria farmacêutica e em produtos de limpeza, além da sua importante utilização nos setores de papel e celulose, obtenção de energia por meio do carvão vegetal e produção de produtos sólidos da madeira, como a madeira maciça e industrializada (RODRIGUES, 2020). É frequentemente cultivado em regiões tropicais e subtropicais que apresentam latitude de cerca de 40° N e 45° S, além de uma altitude de aproximadamente 100 metros e com característica de solo com a salinidade de 2 a 3% (GOIS, 2021).

Como já mencionado, o cultivo de eucalipto proporciona diversos subprodutos; entretanto, esse gênero sofre condições ambientais adversas que são capazes de delimitar seu desenvolvimento, produtividade e crescimento. É citado como fatores abióticos que podem desencadear essas reações adversas, a incidência de luz e intensidade de luz ultravioleta, além da temperatura e condição hídrica, salinidade e exposição a metais pesados. Entretanto, não são todos os genótipos que sofrem com essas adversidades, devido principalmente a resistência e tolerância exibidas por alguns genótipos (RODRIGUES, 2022). Como exemplo é possível citar os clones 1572, 1644, 1407 e 1404 resistentes e os clones 2361, 5204, 520 e 5341 suscetíveis

ao distúrbio fisiológico em eucalipto cultivados no estado da Bahia pela empresa Bracell.

A condição hídrica pode limitar de forma significativa a produção de eucalipto. O estresse hídrico, ou seja, quando há baixa disponibilidade de água, pode influenciar diretamente no crescimento e desenvolvimento de eucaliptos uma vez que, afeta a fotossíntese, respiração, metabolismo de carboidratos e na absorção de íons (FERNANDES, 2021). Além disso, o aumento da temperatura pode gerar estresses fisiológicos devido a um aumento da fotorrespiração e do déficit de pressão de vapor, fatores que contribuem para um aumento evapotranspiração, fazendo que a planta perca água para atmosfera com taxas elevadas (FLORENCIO, 2021).

É verificado que a interação entre o ambiente e genoma pode desencadear a ocorrência de doença abiótica, como o distúrbio fisiológico, que pode interferir no metabolismo celular e consequentemente afetar a expressão gênica da planta, gerando modificações nas características fenotípicas. A ocorrência do distúrbio fisiológico gera grandes prejuízos em relação a produtividade e, portanto, tem impacto direto na economia. Por isso, é importante esclarecer melhor a interferência que o distúrbio fisiológico pode gerar em relação a morfologia, fisiologia e genética molecular do eucalipto (SANTOS, 2021).

## **2.2 Distúrbio fisiológico em eucalipto**

O distúrbio fisiológico é descrito como uma alteração de origem não patogênica, que pode ocorrer pela influência de fatores ambientais, nutricionais, temperatura e composição atmosférica, e promove alterações no metabolismo normal de um vegetal ou em seus tecidos. A primeira identificação de distúrbio fisiológico em eucalipto foi verificada no ano de 2008 em regiões do Espírito Santo e no extremo sul da Bahia, por técnicos da empresa brasileira Suzano SA, do setor de celulose (RODRIGUES, 2022).

Os sintomas observados no distúrbio fisiológico são a perda da dominância apical, redução do crescimento, perda de turgor, seca do ponteiro, necrose do tronco principal e formação de cascas necróticas no tronco, além de bifurcação da copa, aumento da brotação lateral, ercarquilhamento e amarelecimento das folhas. A ocorrência desse distúrbio já afetou diferentes estados no Brasil, embora tenha sido observado em maior prevalência nas regiões litorâneas no estado do Espírito Santo e Bahia (SILVA et al., 2010).

Segundo levantamento, os primeiros indícios do distúrbio fisiológico foram relatados no ano de 2000, no estado do Espírito Santo e Bahia. Os sintomas observados foram principalmente a presença de bifurcação do tronco principal, seca das folhas e ponteiro, além

da necrose no tronco principal, sintomas esses que podem levar a morte de plantas em genótipos mais suscetíveis (MAFIA, 2017).

Ainda não se sabe ao certo o que desencadeia o distúrbio fisiológico, mas estima-se que fatores abióticos, como região de plantio e clima, estão diretamente relacionados. Dessa forma, as plantações suscetíveis a essa desordem estão frequentemente ameaçadas pelas mudanças extremas no clima, seca prolongada ou excesso de chuvas, além de mudanças abruptas de temperatura. Os fatores citados, quando associados aos tipos de solo do plantio de eucalipto, podem contribuir fortemente para ocorrência do distúrbio fisiológico (CÂMARA et al., 2018).

Devido a sintomatologia apresentada, é provável que ocorra uma ativação de forma descontrolada das vias de sinalização da morte celular, acarretando em necrose e senescência desencadeada por fatores climáticos. A morte celular programada (PCD) pode ser definida como uma via de sinalização presente em plantas, sendo ativada por estímulos endógenos como os sinais de desenvolvimento e exógenos como o estresse. A PCD é muito importante em plantas pois é um mecanismo de defesa em relação aos estresses bióticos e abióticos (WANG, 2021).

Além disso, as plantas em geral, como o eucalipto, quando submetidas a situações adversas em relação a temperatura, estado hídrico e luminosidade por exemplo, promovem uma reprogramação da expressão gênica, gerando um conjunto de transcritos diferencialmente expressos em resposta aos estresses. O transcriptoma induzido pelos estresses fornecem informações importantes com relação ao conjunto de transcritos acumulados o que reflete a expressão diferencial dos genes regulados pelos estresses (SANTOS, 2021). Por meio da identificação do transcriptoma induzido pelo distúrbio fisiológico em eucaliptos é possível delimitar genes diferencialmente expressos que podem ser utilizados como marcadores do estado de distúrbio fisiológico.

Além disso, analisar a sequência do DNA e realizar comparações entre genótipos diferentes como por exemplo, entre clones de eucaliptos resistentes e suscetíveis, e associar as possíveis mutações pontuais (SNPs) aos genótipos suscetíveis ao distúrbio fisiológico constitui uma ferramenta para desenvolvimento de marcadores moleculares baseados em DNA. Estes marcadores moleculares podem ser utilizados como método de diagnóstico através das tecnologias de amplificação do DNA como a reação de cadeia polimerase (PCR) capaz de distinguir presença de SNPs em genótipos distintos.

### **2.3 Marcadores Moleculares**

Os marcadores moleculares podem ser definidos como sequencias DNA, RNA ou

proteínas, que podem ser identificados e utilizados na diferenciação de um ou mais indivíduos. Os marcadores moleculares podem ser identificados por técnicas diferentes, sendo que cada técnica se baseia em uma estratégia específica para detectar o polimorfismo, se diferenciando com relação a disponibilidade no genoma, níveis de polimorfismo encontrados, especificidade de locus, reprodutibilidade e custo (ZOLET, 2017).

Com relação a metodologia usada para identificação dos marcadores moleculares, é possível citar a metodologia por hibridização e os obtidos por reação de cadeia polimerase (PCR). A técnica utilizando PCR é uma das mais utilizadas pois tem a capacidade de amplificar grande quantidade de uma determinada sequência do DNA, sem requerer a clonagem dos fragmentos amplificados, partindo apenas de pequenas amostras da sequência alvo (OLIVEIRA, 2021).

É possível citar como metodologia de identificação de marcador molecular baseado em PCR o AFLP (*Amplified Fragment Length Polymorphism*), que consiste na amplificação de fragmentos de restrição produzidos por enzimas de restrição específica que são ligados a adaptadores oligonucleotídicos. Além disso, a utilização de sequenciamento de nova geração (*Next Generation Sequencing - NGS*), também tem sido muito utilizada, pois possibilita uma melhor identificação e análise de marcadores moleculares, devido a determinação rápida e de forma e eficiente da sequência de DNA e RNA de um organismo. Podemos citar como exemplo de plataforma de sequenciamento de nova geração a Illumina Hiseq, que realizam um sequenciamento com alta velocidade e geração de grande volume de dados (FILHO, 2019). Tal metodologia é capaz de realizar a identificação de polimorfismos de nucleotídeos únicos (SNPs) (ZIMMERMAN et al., 2020).

Os SNPs que são definidos como uma variação genética que ocorre em uma única posição nucleotídica em uma sequência de DNA (FADASON, 2021). Esse tipo de polimorfismo representa uma variação genética que apresenta maior ocorrência no genoma de eucariotos, estando presentes em regiões codificadoras como os éxons e regiões não codificadoras como os introns e regiões intergênicas. Os mesmos surgem a partir de mutações pontuais no DNA, como em transições quando ocorre trocas de bases purínicas (adenina e guanina) e pirimidínicas (citosina e timina) ou em transversões quando ocorre troca entre bases purínicas por pirimidínicas (WOLD et al., 2021).

É possível verificar uma maior presença de SNPs nas regiões não codificadoras, sendo esses denominados de SNPs não codificantes (ncSNPs). Em contrapartida, há uma menor presença de SNPs nas regiões codificadoras como nos éxons, sendo esses denominados SNPs codificadores (cSNPs). A presença de um SNP em um éxon de um gene pode acarretar

problemas em relação a função da proteína codificada, sendo, portanto, denominado SNP candidato, uma vez que pode estar associado a alguma característica fenotípica (NOGALES; DEDIEGO, 2019).

O uso dos SNPs como marcadores genéticos para realizar análises é muito vantajoso pois apresenta maior potencial para realizar varreduras no DNA realizando a detecção de regiões neutras ou adaptativas. Além disso, apresenta a vantagem do alto rendimento e custo moderado, podendo ser realizado em grande escala, gerando um custo-benefício maior quando comparado com demais marcadores moleculares (ZIMMERMAN et al., 2020).

### **3. METODOLOGIA**

#### **3.1 Coleta do material vegetal**

Para a realização dos testes, foram utilizados nove clones de eucaliptos assintomáticos, sendo quatro clones assintomáticos com genótipo resistente 1572, 1644, 1407 e 1404 e cinco clones assintomáticos suscetíveis para o distúrbio fisiológico 5341, 2361, 520, 1296 e 5204. Foi utilizado também, folhas dos clones de eucalipto 520 suscetíveis com sintomas, 520 suscetíveis sem sintomas e 1404 suscetível resistente. A coleta dos clones foi realizada por técnicos especializados na área de plantação de eucalipto da BRACELL, empresa localizada na Bahia-Alagoinhas. Durante o processo de obtenção das amostras, foram coletadas três repetições biológicas das quais apresentam um estágio de desenvolvimento entre 6 a 12 meses que estavam sob mesma condição climática e área de cultivo.

Após a realização da coleta das amostras, as mesmas foram congeladas imediatamente no campo com nitrogênio líquido e em seguida, congeladas com gelo seco e posteriormente, enviadas para o Laboratório de Biologia Molecular de Plantas/Bioagro/UFV localizado na cidade de Viçosa-MG.

#### **3.2 Extração de DNA e RNA das amostras de eucalipto**

A extração do DNA das amostras dos caules de eucalipto assintomáticos 1572, 1407, 1407, 1644, 520, 1296, 5204, 2361, 5341 foi realizada utilizando o protocolo CTAB 2% (CHANG, et al. 1993). Inicialmente, foi coletado com auxílio de lâmina estéril e adicionado em *epENDORF*, 100 mg de amostra dos caules que se encontravam sob refrigeração no -80 °C As amostras adicionadas no *epENDORF* foram maceradas em equipamento Tissuelyser e adicionado

1 ml de Tampão CTAB 2% pré-aquecido por 10 min a 65 °C.

Em seguida, as amostras foram homogeneizadas e incubadas a 65 °C por 25 minutos. Posteriormente, foram centrifugadas a 12000 x g por 30 s em temperatura ambiente, e o sobrenadante formado transferido para novo *epENDORF* e adicionado um volume de clorofórmio correspondente a 80% do sobrenadante obtido. As amostras foram centrifugadas por 15 min a 13000 x g, a 4 °C. Após centrifugação, o sobrenadante obtido foi transferido para novo *epENDORF* e adicionado um volume 1:1 de isopropanol gelado e deixado em gelo por 10 min para a precipitação do DNA.

As amostras foram centrifugadas a 15000 x g, por 15 min a 4 °C e o sobrenadante descartado. Em seguida, o pellet formado foi lavado com 1 ml de etanol 75% gelado e realizada centrifugação por 5 min a 4 °C, 12.000 x g. O líquido residual foi descartado e os pellets secados em temperatura ambiente. Após secagem completa, os pellets de DNA foram ressuspensas em 20 ul de água Mili-Q autoclavada e armazenados em freezer a -20 °C.

Para extração de RNA das folhas de eucalipto dos clones 520 assintomáticos, 520 sintomáticos e 1404 resistentes, foi utilizado o método de extração PureLink plant RNA, de acordo com as recomendações do fabricante (Invitrogen). Primeiramente, 50 mg de amostra congelada em nitrogênio foi macerada em Tissulyser com auxílios de esferas metálicas e adicionado 500 ul do reagente Plant RNA PureLink. As amostras foram homogeneizadas em vórtex e deixadas em posição horizontal por 5 min. Em seguida, foram centrifugadas a 12000 x g a 4 °C e o sobrenadante transferido para novo tubo *epENDORF*. Ao sobrenadante obtido, foi adicionado 100 ul de NaCl 5 M e 300 ul de clorofórmio, que foi homogeneizado em vórtex. Posteriormente, as amostras foram centrifugadas por 10 min a 12000 x a 4 °C e o sobrenadante obtido transferido para novo tubo *epENDORF* no qual foi adicionado um volume 1:1 de isopropanol gelado. Em seguida, as amostras foram incubadas em gelo por 10 min para precipitação do RNA e centrifugadas a 14000 x g a 4 C. O sobrenadante foi descartado e os pellets lavados com 1000 ul de etanol 75% gelado. Os *pellets* secaram em temperatura ambiente e foram ressuspensos em 20 ul de água DEPC e armazenados a -80 °c até envio ao sequenciamento.

### **3.3 Verificação e preparo das amostras para sequenciamento do DNA dos caules de eucalipto**

Após extração de DNA das amostras dos clones de eucalipto assintomáticos suscetíveis

e resistentes, a integridade e qualidade do DNA foi verificada por meio de eletroforese em gel de agarose 1% e, posteriormente, quantificado no espectrofotômetro, e selecionados pelos valores de relação OD 280/260 superiores a 1,8. Em seguida, foi adicionado 1 µL de RNase para degradação de RNA, e a eficiência do processo visualizado por eletroforese em gel de agarose 1%.

Após obtenção das amostras, fragmentos de promotores de genes selecionados foram amplificados por PCR, sendo estes genes envolvidos na síntese de ácidos graxos (*KASIII*, *KCSI*), genes envolvidos no reparo do DNA (*TILL1*, *MSH2*) e genes envolvidos em resposta a estímulos abióticos (*Galactose Oxidase* e *RPT2*). Foram utilizados os seguintes oligonucleotídeos conforme a Tabela 1.

A reação de PCR foi preparada para um volume final de 25 µl, utilizando o tampão 10X High Fidelity PCR Buffer, MgSO<sub>4</sub> 50 mM dNTP 10mM, primer forward e reverse 10mM, Platinum *Taq* DNA Polymerase High Fidelity (Invitrogen), água e DNA. As condições da reação foram estabelecidas com o seguinte ciclo: 94 °C por 30 segundos e 40 ciclos de desnaturação a 94 °C por 15 segundos. O anelamento foi realizado por 30 segundos com temperatura conforme TM (Melting Temperature) dos oligonucleotídeos e extensão por 1 minuto a 68 °C. Foi adicionada uma etapa a 4 °C por tempo infinito para maior conservação da amostra.

**Tabela 1.** Oligonucleotídeos utilizados para PCR

Primer	Eucalipto	Descrição	Sequência 5'—3'
Kas III_Fwd Kas III_688Rvs Kas III_501Fwd Kas III_1386Rvs Kas III_1192Fwd Kas III_1996Rvs KasIII_1110Fwd	Eucgr.E00458	3-Oxoacyl-[acyl-carrier-protein (ACP)] synthase III	CAA GAC TTAGTT CAT ATT AGT GCATG CAG ACA TGT CAA GCA ACC AC ATT CGA TGA CCC CAT TTA GCG CGG ATA TCT CCA ACC ATC GCG ATG AGT CTA AGT TTG ACC CAT CTC AAA TCA CTT GGA GCC TCC CTC CCCAGGTTTTCACTTCTC
KCSI_Fwd KCSI_820Rvs KCSI_640Fwd KCSI_1455Rvs KCSI_1208Fwd KCSI_2059Rvs	Eucgr.H00232	3-Oxoacyl-[acyl-carrier-protein (ACP)] synthase III C Term	GGAAGTGGGAATCGGACTAAT GAAGGTCTCGCTGGTACATATC AACCTCGTACCATCGTAATC GTCATTCTCACGCCTCTATC CTTACACCTCGGAGCTTTGT GAGATTCGACTGAGATAGGGAAAG
MSH2_Fwd MSH2_856Rvs MSH2_663Fwd MSH2_1480Rvs MSH2_1290Fwd MSH2_1950Rvs	Eucgr.B00700	MSH2; DNA mismatch repair protein MSH2	ACG AGG AGC ACA GAA AAT TG CGT CTG GGT TTG TGA TCA C ATG TTC GGG TTT AGG AGG TG CCT TCT GAA GCA CTC AAT ACT CG CTG TTA GCG TCT CGA GAG TC GAG AGA TTC TGC TCG TCA GTC
TILL1_Fwd TILL1_918Rvs TILL1_642Fwd TILL1_1572Rvs TILL1_1180Fwd TILL1_1991Rvs	Eucgr.B03882	OL2A,TIL1;DNA polymerase épsilon catalytic subunit	TAC GGT CCC TTC AGG TAC G GAT GTT GTG CCA TTG GTT ACT AGC CCA TGT GCT CGT AGA AAT GTG G GCA ACC ATA TGA TCT CTG TAG TCT G AGC TAA TAA GCA GTG ATC ACT GC CGA TCG GAG TAC GGT AGA GAC
RPT2_Fwd RPT2_954Rvs RPT2_453Fwd RPT2_1459Rvs RPT2_1136Fwd RPT2_2071Rvs RPT2_1110Fwd RPT2_ResFwd RPT2_SuscFwd RPT2_Res/SuscRvs	Eucgr.B03325	Root phototropism protein 2	CTT CGC CTG CTC TCT TCC T GAT TGG TTT CGC CAC ACA GC TCC CTT TCT AAA GTA GGT GGA TGG GGT AAA GAG AGA AGC AAC ATT GCC AGA GCG GGG ACT CCA AAA G GAG AGT CTG AGC TTT TCG GG TGAGTCTGACCGATGTACAATAA GAACTAAAGTTCATAGAT GAACTAAAGTTCATAGAC CCTTTAGTCACTTACATGCACATC
Galactose_Fwd Galactose_839Rvs Galactose_503Fwd Galactose_1417Rvs Galactose_1180Fwd Galactose_230Rvs	Eucgr.B03483	Galactose oxidase/kelch repeat superfamily protein	CTT ATG GTC GGT CAC AGT GTC T CAA CCC GGC AAG TTG AAG TAC CGT AGA GAG CTT GAA TAT CGC TG GAA GCT GCG TAG ATT ACA GAT GC CGG AAA ACC AGG TTC AAC CAC CGC TGA ATC TAA GTA AAT CTC GACC

Após o ensaio de PCR, os fragmentos amplificados foram verificados por meio de eletroforese em gel de agarose 1%, com objetivo de avaliar a efetividade da reação e integridade das amostras obtidas. Os oligonucleotídeos foram desenhados gerando amplicons correspondentes a 650 a 700 pb. Em seguida, as amostras foram purificadas utilizando o kit ExoSAP-IT PCR Clean-up Protocol, para remoção de oligonucleotídeos e nucleotídeos não usados.

De acordo com o protocolo ExoSAP-IT, 10 µl do produto de PCR foram homogeneizados com 1µl de reagente EXOSAP-IT diluído em uma proporção 1:4 com água Milli-Q autoclavada. Em seguida, as amostras foram incubadas no termociclador a 37 °C por 15 min para degradação dos oligonucleotídeos e nucleotídeos que restaram na reação. Posteriormente, foi realizada uma nova incubação a 80 °C por 15 min para inativação do reagente ExoSAP-IT. As amostras tratadas foram armazenadas no freezer a -20 °C, até envio para sequenciamento.

### **3.4 Sequenciamento e análise das amostras do DNA dos caules de eucalipto**

As amostras de PCR purificadas foram enviadas para o Laboratório de Patologia Florestal- Bioagro localizado na Universidade Federal de Viçosa. Para o sequenciamento, foram enviados 10µL de reação de PCR purificada, 10µL de cada oligonucleotídeo por amostra Forward e Reverse utilizado na amplificação do DNA, na concentração de 1mM. Além disso, foi encaminhado uma foto do gel de eletroforese da reação purificada e planilha de informações referentes as amostras enviadas.

Os dados do sequenciamento foram analisados utilizando o software Snap Gene 6.0.5, sendo que as análises de alinhamento das sequências *forward* e *reverse* com a sequência referência dos genes foram realizadas com o software Multalin Interface page-Florence Corpet. Com os resultados obtidos, as sequências consensus (*forward/reverse*) dos clones de eucalipto suscetíveis foram comparadas utilizando o Multalin Interface page-Florence Corpet a fim de detectar SNPs. Tendo SNPs, foram desenhados primers no Primer Express 3.0 (Life Technologies) capazes de distinguir as mutações identificadas.

### **3.5 Preparo das amostras de RNA de folhas de eucalipto para envio ao sequenciamento**

Primeiramente, a integridade da extração de RNA foi avaliada através de eletroforese em gel de agarose desnaturante 1% e quantificado em espectrofotômetro, verificando a pureza das extrações através do valor obtido da relação OD 260/280. Além disso, as amostras foram avaliadas no equipamento *Bioanalyser* que é responsável por realizar um controle de qualidade de amostras, como a integridade, pureza e tamanho das amostras.

Para utilização no equipamento *Bioanalyser*, as amostras foram diluídas para uma concentração de 5000 pg/µL e seguido o protocolo Agilent RNA 6000 pico Assay. Para realização do teste, foi utilizado um kit contendo uma matriz de separação que consiste em solução de polímero, solução corante fluorescente com o objetivo de intercalar a dupla fita de

DNA e o terceiro componente, marcadores de tamanho de RNA.

Antes de prosseguir para análise no equipamento *Bioanalyser*, o corante fluorescente foi acrescentado à matriz de separação e inserido nos canais de chip sob pressão com auxílio de uma seringa acoplada a um dispositivo específico. Em seguida, foi adicionado marcador no poço correspondente ao chip e as amostras acrescentadas nos poços. O chip foi agitado através de equipamento e introduzido no *Bioanalyser*.

Os resultados das separações foram avaliados por meio de software 2100 Expert que é capaz de corrigir os eletroferogramas através do alinhamento dos picos, cálculo de tamanho e de concentração dos fragmentos. Os resultados obtidos foram fornecidos em valores de RIN (*RNA integrity number*), que é capaz de estimar a integridade do RNA, sendo que esse valor varia entre 0 a 10, sendo selecionados RNAs com valores acima de 7 e quantidade suficiente para confecção da biblioteca, para sequenciamento na empresa MacroGen Inc. (Korea).

### **3.6 Sequenciamento do transcriptoma - RNA (Seq)**

O sequenciamento do RNA foi realizado através da tecnologia da plataforma Illumina HiSeq™ 2500, utilizando *paired-end*. Essa tecnologia possui um *flowcell*, que é uma célula de fluxo que apresenta diversas faixas, sendo que cada faixa possui a capacidade de produzir cerca de 250 milhões de leituras com comprimento de leitura de 35 a 250 bp. Além disso, todo processamento realizado pela Illumina HiSeq™ é realizado através de três etapas, sendo estas, preparo das bibliotecas de cDNA, geração de clusters e sequenciamento.

A construção da biblioteca de cDNA foi realizada somente com as amostras que passarem no controle de qualidade realizado pelo Bioanalyzer 2100 da Agilent Technologies ou seja, que apresentarem valor de RIN próximo ou superior a 7. Para a construção da biblioteca de cDNA, foi utilizado a metodologia *TruSeq mRNA library construction* conforme protocolo da Illumina. Com relação à etapa de geração de clusters, para que possa ser realizado o sequenciamento do RNA, os *templates* fita simples devem ser amplificados com objetivo de obter clusters clonais na *flowcell* de sequenciamento. Como a extremidade 5' dos fragmentos de cDNA foram hibridizados à *flowcell*, a extremidade 3' se apresenta livre para permitir o início do sequenciamento, processo que consiste em uma PCR que é realizada com a presença de um coquetel de oligonucleotídeos complementares às extremidades dos adaptadores que foram introduzidos. Em seguida, ocorreu um processo de desnaturação dos amplicons, promovendo a separação das fitas para promover o sequenciamento.

O sequenciamento foi realizado após a introdução da célula *flowcell* no equipamento

HiSeq 2500. No início do processo, ocorreu uma etapa de desnaturação e a introdução de nucleotídeos marcados com fluoróforos e se seguiu o processo de sequenciamento base a base. No momento em que ocorre a excitação dos nucleotídeos marcados utilizando lasers, uma luz é emitida e as imagens captadas. Sempre após a introdução do nucleotídeo no fragmento que está em processo de síntese, ocorre um processo de lavagem para retirada de reagentes e fluoróforos, até o fim da síntese da fita do sequenciamento *Read 1*. Em seguida, essa fita sofreu uma desnaturação e passou a apresentar a extremidade 3' livre, permitindo a síntese de uma fita molde, promovendo o sequenciamento *Read 2*.

### 3.7 Bioinformática

A qualidade das sequências dos arquivos no formato FASTQ proveniente da empresa Macrogen foi feita pela ferramenta FastQC. Em seguida, utilizando o programa Trimmomatic (BOLGER et al., 2014), foi realizada a remoção dos adaptadores Illumina e das sequências de baixa qualidade, isto é, sequências que possuíam bases com qualidade inferior a Q20. Na etapa posterior, para o mapeamento no genoma, utilizou-se o software STAR versão 2.7.10b (DOBIN et al. 2013) com o genoma de referência *Eucalyptus grandis* disponibilizado no site do Phytozome v13 (<https://phytozome-next.jgi.doe.gov>). Além disso, foi realizada a contagem das reads mapeadas por gene através da função featureCounts disponível no pacote Rsubread (LIAO et al., 2013) no sistema R/Bioconductor. Para a análise de expressão diferencial foi utilizado o pacote DESeq2 (LOVEL et al., 2014). A seleção dos genes diferencialmente expressos foi realizada através da utilização do critério p-valor ajustado pelo (“False discovery rate”- FDR) de  $< 0.1$  e  $\log_2\text{FoldChange} > |1.0|$ . Com relação ao enriquecimento gênico para as categorias do *Gene Ontology* (GO) foi utilizado os pacotes GSEABase e GOstats do R/Bioconductor (FALCON; GENTLEMAN, 2007). Para as vias do KEGG utilizou-se o pacote Pathview (LUO et al, 2013). A expressão diferencial dos genes será validada por meio de RT-PCR quantitativo.

### 3.8 Extração de RNA e síntese de cDNA para verificação de transcritos por PCR quantitativo (qRT-PCR) das folhas de eucalipto

Para extração de RNA das folhas de eucalipto e posterior síntese de cDNA, foi utilizado o protocolo de extração CTAB 2% [2%(v/v) CTAB(cethyltrimethylammonium bromide), 2%(v/v) PVP, 100 mM Tris-HCL, 25 mM EDTA, 20mM NaCl, e 2%  $\beta$ - mercaptoethanol], com algumas modificações (CHANG, PURYEAR e CAIRNEY, 1993).

Para extração foi utilizado 100 mg de material triturado em nitrogênio líquido com adição de uma pequena quantidade de PVP. Em seguida, as amostras foram transferidas para eppendorf e adicionado 1ml de tampão CTAB 2% pré-aquecido. As amostras foram homogeneizadas e incubadas por 25 minutos a 65 °C e centrifugadas 12.000 x g, por 1 minuto, a 4°C. Ao sobrenadante, foi adicionado um volume de 80% de clorofórmio e centrifugado 13.000 x g por 15 minutos a 4 °C. Um volume 1:1 de isopropanol foi adicionado ao sobrenadante para precipitação do RNA e as amostras foram incubadas em gelo por 10 minutos. Em seguida, as amostras foram centrifugadas a 14.000 x g, por 10 minutos a 4 °C, e o sobrenadante descartado e o pellet lavado com 1ml de etanol 75% gelado. Seguiu-se para centrifugação por 5 minutos a 12.000 x g a 4°C e o sobrenadante descartado e os pellets secados em temperatura ambiente e foram ressuspensos em 20 ul de água DEPC.

A integridade do RNA das folhas dos clones de eucaliptos foi avaliada por eletroforese em gel de agarose 1%, sendo quantificado por espectrofotometria. O RNA total isolado foi padronizado para 3µg e as amostras tratadas com DNase (Invitrogen) com modificações no protocolo. Para tratamento, foi utilizado 1ul de DNase e 1ul de Buffer por amostra, que foram incubadas a 37 °C por 10 minutos e posteriormente adicionado 1 ul de EDTA 25 mM e as amostras incubadas a 65 °C por 10 minutos. Em seguida, eficiência do tratamento foi avaliada por eletroforese em gel de agarose 1% e o cDNA sintetizado utilizando enzima de transcriptase reversa e oligo-dT. Para a avaliação dos transcritos por meio de RT-PCR em tempo real, foram utilizados os primers descritos na Tabela 2. Esses oligonucleotídeos foram selecionados conforme resultados obtidos no transcriptoma induzido realizado em folhas de eucalipto 520 assintomáticos e 520 assintomáticos para os genes KCS2, ANAC098, DRT102, bHLH, Leucine e TIR-NBS.

Para a reação de RT-PCR em tempo real, foi utilizado o equipamento 7500 Real Time PCR System (Applied Biosystems), através da metodologia de SYBR Green PCR Master Mix (Life Technologies). Foi utilizado como marcador endógeno da reação ACT (Actin-related protein 3). No preparo da reação, foi utilizado um volume de 10 µL, composto por template de cDNA, oligonucleotídeos (*primers* forward e reverse), listados na Tabela 2, o corante Fast SYBR Green PCR 2x Master mix (Applied Biosystems). Para programação do termociclador, foi utilizado 95 °C por 10 min para desnaturação inicial, 40 ciclos de 94 °C por 15 s para desnaturação e 60 °C por 1 min para a etapa de extensão e anelamento. Para verificação da especificidade dos resultados obtidos, foi avaliada as curvas de dissociação dos oligonucleotídeos geradas pelo termociclador. Em seguida, a expressão gênica obtida foi

quantificada utilizado método comparativo de Ct:  $2^{-\Delta Ct}$ .

**Tabela 2.** Lista de oligonucleotídeos que serão utilizados para qRT-PCR

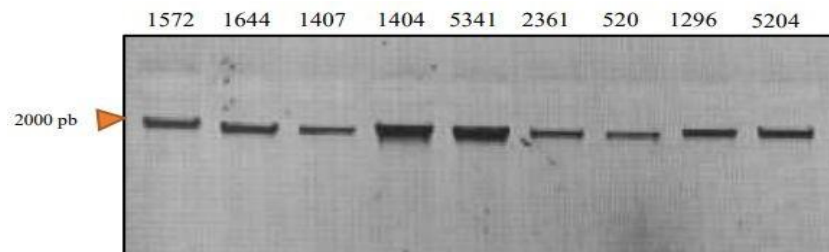
Primer	Eucalipto	Descrição	Sequência 5' ---3'
qRT-EG-KCS2-Fwd qRT-EG-KAS2-Rvs	Eucgr.D00056	3-ketoacyl-CoA synthase 2	GCCGAGCAACAAGAACAATAA CTCAGTTGGTCCGAGAGTTG
qRT-EG-ANAC098-Fwd qRT-EG-ANAC098-Rvs	Eucgr.F01170	NAC (No Apical Meristem) domain transcriptional regulator superfamily protein	TTCCAGAAGTCCAGCAACAG CTTAGGGATTGAGCTGTCGTAAG
qRT-EG-DRT102-Fwd qRT-EG-DRT102-Rvs	Eucgr.B01712	DNA-damage-repair/toleration protein (DRT102)	GCGCTACGCTCAAGGAC GAACTTGTTGGCGAGCATTG
qRT-EG-bHLH-Fwd qRT-EG-bHLH-Rvs	Eucgr.B01225	basic helix-loop-helix (bHLH) DNA-binding superfamily Protein	CAAGTGGTTGGCCCTTAGA GGAGATGGTGTAGGAGGTGATA
qRT-EG-Leucine-Fwd qRT-EG-Leucine-Rvs	Eucgr.K02588	Leucine-rich repeat receptor- like protein kinase family Protein	TGGAAAGAGGAAGCCTAAGAAC GCCGGCAATAATAGTCCAATTT
qRT-EG-TIR-NBS-Fwd qRT-EG-TIR-NBS-Rvs	Eucgr.H02116	Disease resistance protein (TIR- NBS-LRR class) family	CACAGACGTCCGCAATAACT TCTGCTAGCTCATCCAAACAC
qRT-EG-Actin-Fwd qRT-EG- Actin -Rvs	Eucgr.H02449	Actin-related protein 3	TAAGCATGACAAGGAACCAG TCAGGTCCAAGAAATCGT

## 4. RESULTADOS

### 4.1 Desenvolvimento de marcadores moleculares associados ao estado da síndrome do distúrbio fisiológico em eucalipto a partir de sequencias de promotores de genes selecionados

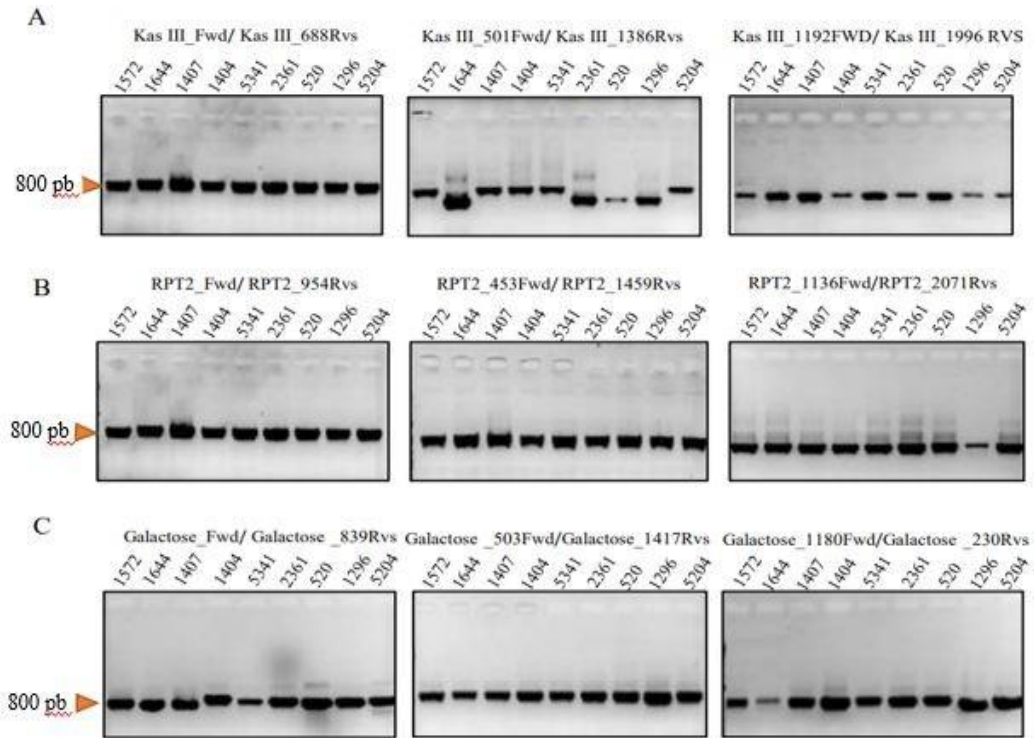
Embora o perfil de expressão de genes marcadores possa ser utilizado como genes repórteres do distúrbio fisiológico, os ensaios para determinação de expressão gênica são onerosos e de difícil realização. Assim, a proposta é analisar a sequência dos promotores dos genes marcadores selecionados a fim de identificar a presença de modificações nas sequencias que pudessem ser associadas ao distúrbio. Isto possibilitaria o desenvolvimento de marcadores moleculares utilizados em reações de PCR, menos onerosas e de fácil manuseio. Para análise da possível presença de marcadores moleculares nos promotores dos genes diferencialmente

expressos, inicialmente foi realizada a extração de DNA dos caules dos clones de eucalipto suscetíveis e resistentes utilizando o protocolo CTAB 2% com posterior quantificação e avaliação em gel de agarose 1% (Figura 1).



**Figura 1.** Integridade do DNA genômico de caules de eucalipto. DNA total foi extraído de caules dos clones 1572, 1644, 1407, 1404 2361, 520, 1296 e 5204 e visualizado em gel de agarose 1% corado com Brometo de Etídeo

Após extração das amostras, foi realizada Reação de Cadeia Polimerase (PCR) com os primers referentes aos promotores dos genes diferencialmente expressos *KasIII* e *KCSI* que são relacionados à produção de ácidos graxos, *MSH2* e *TILL1* envolvidos no reparo do DNA que foram identificados como regulados negativamente pelo distúrbio fisiológico em amostras de caules sintomáticos (Martins, 2020; patente BR BR102022025625-0, suplemento S1). Os promotores dos genes *RPT2* e *Galactose oxidase* envolvidos em respostas a estímulos abióticos foram selecionados como genes regulados positivamente pela condição de distúrbio fisiológico (Martins, 2020; patente BR BR102022025625-0, suplemento S1). Para análise dos promotores dos genes diferencialmente expressos, foram utilizados três pares de oligonucleotídeos com amplicon 800 pb, de forma a cobrir aproximadamente 2000 bp da região promotora, facilitando o sequenciamento completo de aproximadamente 2000 bp de sequencias na extremidade 5' flanqueando a região codificadora. O produto de PCR obtido foi visualizado em gel de agarose 1% para avaliar qualidade e integridade da reação correspondente à região promotora do gene *KASIII* (Figura 2A), *RPT2* (Figura 2B), *Galactose oxidase* (Figura 2C), *KCSI* (Figura S1A), *TILL1* (Figura S1B) e *MSH2* (Figura S1C).



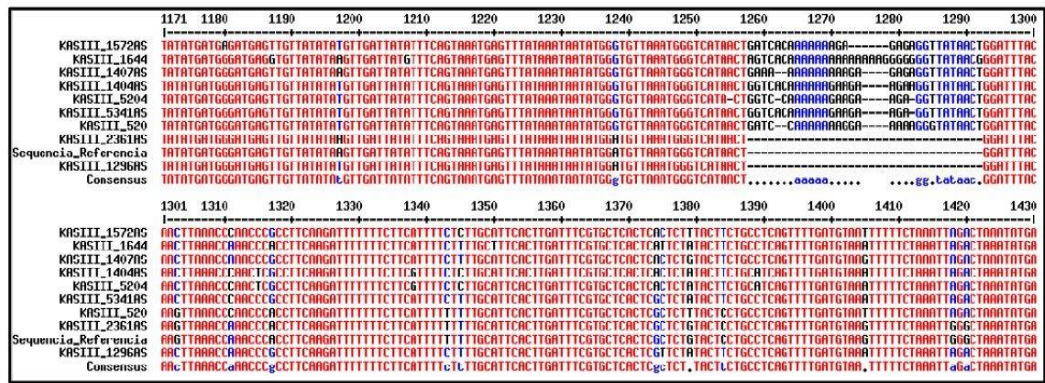
**Figura 2.** Amplificação de fragmentos sobrepostos de promotores dos genes indicados na figura. A) Eletroforese em gel de agarose 1% do produto de PCR com oligonucleotídeos específicos do promotor do gene *KASIII* visualizado por coloração de Brometo de Etídeo. B) Fragmentos sobrepostos do promotor do gene *RPT2*. C) Fragmentos sobrepostos do promotor do gene *Galactose Oxidase*. O número na nomenclatura do oligonucleotídeo corresponde aproximadamente à posição de anelamento no promotor do referido gene. Os clones 1572, 1644, 1407 e 1404 são resistentes e os clones 5341, 2361, 520, 1296 e 5204 são suscetíveis ao distúrbio fisiológico.

Os produtos amplificados foram purificados do gel e sequenciados (Figura 3). Os resultados do sequenciamento foram analisados pelo software Snap Gene 6.0.5 e o alinhamento das sequências senso e antisenso realizado por meio do software Multalin Interface page-Florence Corpet. Com o objetivo de identificar a presença de SNPs, procedeu-se ao alinhamento de sequências da região promotora dos genes *KASIII*, *KCSI*, *MSH2*, *TILL1*, *RPT2* e *Galactose oxidase* das amostras 1572, 1644, 1407, 1404, 5341, 2361, 520, 1296 e 5204 e do clone de referência *E. grandis*, obtidas do phytozome (Figuras S2-S7).

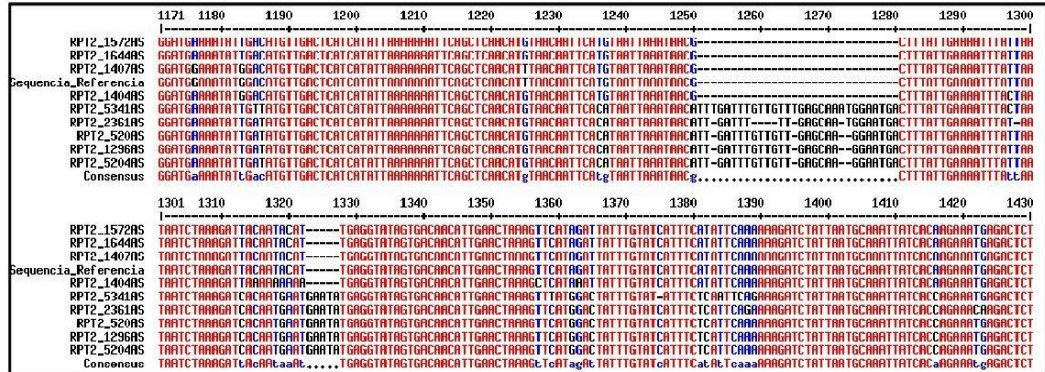
A análise comparativa de sequências revelou a presença de inserção na sequência promotora de algumas amostras suscetíveis e resistentes referentes ao gene *KASIII* (Figura 3A) e *RPT2* (Figura 3B). Nota-se, entretanto, discrepâncias entre o tamanho do fragmento amplificado e deleções/inserções detectadas pelo sequenciamento. Por exemplo, os fragmentos amplificados dos clones 1644 e 520 (painel central da Figura 2A) possuem aparentemente o número de pares de bases inferior, mas o sequenciamento dessa região dos promotores não

confirma possível deleção, o que contrasta com os resultados dos clones 2361 e 1293 (compare Figura 2A com o sequenciamento na Figura 3A). Já em contrapartida, na região promotora dos demais genes analisados como da Galactose oxidase não foram identificadas inserções, deleções ou SNPs que pudessem ser associadas com o caráter resistência e suscetibilidade (Figura 3C).

A



B



C

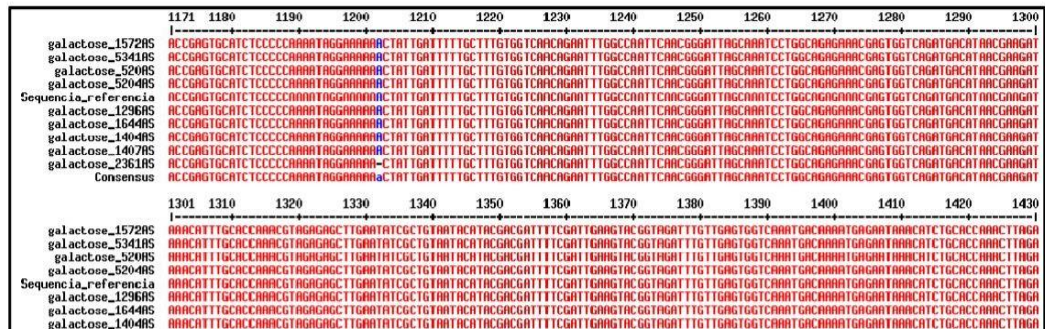
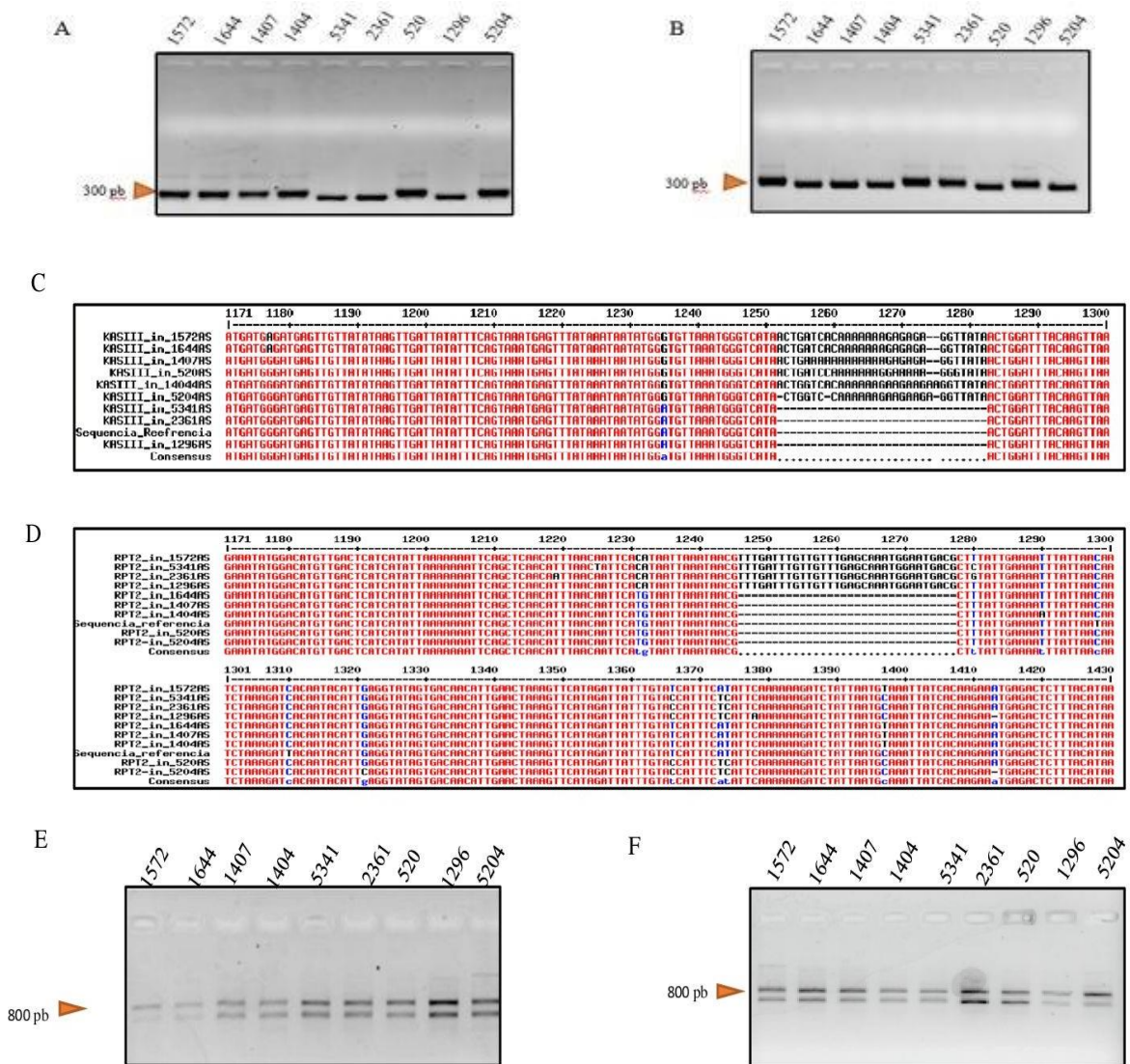


Figura 3: A) Alinhamento de sequência da região promotora do gene KASIII entre amostras resistentes 1572, 1644, 1407, 1404 e amostras suscetíveis 5341, 2361, 520, 1296 e 5204. Presença de inserção na região -1251 a 1292 pb. B) Alinhamento da região promotora do gene RPT2. Presença de inserção na região -1251 a -1280 pb e presença de SNP na posição -1366 pb. C) Alinhamento da região promotora do gene Galactose oxidase.

Além disso, na região promotora do gene *RPT2* além da inserção identificada, foi verificado a presença de um SNP devido à diferença de nucleotídeo na posição -1366 pb entre as amostras suscetíveis e resistentes, verificando a presença de uma timina nos clones de eucalipto resistentes e uma citosina nos clones suscetíveis. Também foram amplificados fragmentos dos promotores dos genes *KCSI*, *TILL1* e *MSH2* (Figura S1). Análise comparativa da sequência desses promotores não detectou polimorfismo associado ao caráter de resistência e suscetibilidade (Figuras S3, S5 e S7).

Com objetivo de confirmar a presença das inserções na região promotora do gene *KASIII* e *RPT2* das amostras dos clones de eucalipto, a referida região do promotor foi reamplificada utilizando oligonucleotídeos que geram amplicons de 300 pb flanqueando a região variante. Os fragmentos amplificados foram resequenciados para confirmação adicional de deleções/inserções, sendo detectadas diferenças de tamanho (número de pares de bases) entre amostras do promotor do gene *KASIII* (Figura 4A) e do promotor *RPT2* (Figura 4B). O sequenciamento dos fragmentos reamplificados confirmou a presença do SNP na posição -1366 pb, mantendo nesta posição Timina no promotor do gene *RPT2* dos clones resistentes e Citosina nos promotores de clones suscetíveis (Figura 4D).

Além disso, foi realizada PCR utilizando oligonucleotídeos específicos para região promotora do gene *RPT2* com a presença de SNP para possível genotipagem. Para isso, a sequência do oligonucleotídeo *forward* utilizado diferiu apenas na última base nitrogenada, sendo timina para amplificação de amostras resistentes (Figura 4E) e citosina para as amostras suscetíveis (Figura 4F). Conforme observado, ocorreu a amplificação de todas as amostras, tanto para os genótipos de amostras resistentes quanto suscetíveis, resultando em duas bandas amplificadas. Tal resultado é provavelmente devido a presença de outras regiões no genoma do *Eucalyptus grandis* capazes de amplificar com os oligonucleotídeos utilizados, conforme verificado no programa BLAST na função nucleotide Blast (Figure S8).



**Figura 4:** A) PCR com oligonucleotídeo de 300 pb referente a região do promotor KASIII. B) PCR com oligonucleotídeo de 300 pb referente ao promotor RPT2. C) Comparação do sequenciamento da região promotora do gene KASIII entre os clones de eucalipto. Confirmação da inserção na posição -1152pb. D) Comparação do sequenciamento da região promotora do gene RPT2 entre os clones de eucalipto com confirmação da presença de inserção na posição -1245 e SNP na posição -1366 pb. E) PCR de genotipagem para amostras resistentes ao distúrbio fisiológico. F) PCR de genotipagem para amostras suscetíveis ao distúrbio fisiológico.

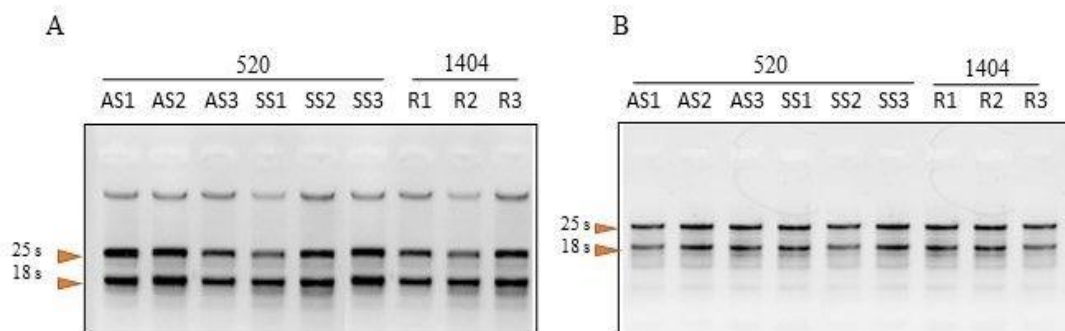
No promotor de *RPT2*, foram observados novos SNPs, que pudessem estar associados com o caráter resistência e suscetibilidade. nas posições 1373 e 1374, sendo a sequência AT em amostras resistentes e TC em amostras suscetíveis. Oligonucleotídeos específicos para estes SNPs foram desenhados para confirmar sua capacidade de marcação, mas ao analisar no programa BLAST na função nucleotide Blast, não foi verificada a especificidade do oligonucleotídeo (*primer*), para realizar genotipagem.

## 4.2 Identificação do transcriptoma induzido pelo distúrbio fisiológico em folhas do clone suscetível 520

A existência de genótipos suscetíveis e resistentes cultivados em condições climáticas e em estádios de desenvolvimento similares permite que se analise o transcriptoma induzido pelo distúrbio fisiológico no nível de campo. Em uma primeira etapa prévia desse projeto, foi identificado o transcriptoma induzido pelo “distúrbio fisiológico” em caules de genótipos contrastantes (resistentes x suscetíveis) em apresentando ou não sintomas aparentes.

Portanto, também foi objetivo dessa investigação a identificação do transcriptoma induzido pelo distúrbio em folhas a fim de compará-lo com aquele identificado em caules. A hipótese central seria a existência de genes diferencialmente expressos comumente associados ao estabelecimento do distúrbio fisiológico tanto em caules quanto em folhas.

Para o RNA-seq, RNA total foi extraído das folhas de eucalipto dos clones 520 assintomático, 520 sintomático e 1404 resistentes em triplicatas biológicas. A integridade do RNA total foi inicialmente avaliada por eletroforese em gel de agarose 1%, permitindo a identificação de duas bandas de rRNA ribossomal 25S e 18S e ausência de degradação visíveis do RNA durante processo de extração (Figura 5A). A quantificação em espectrofotômetro (Tabela 2) indicou uma razão A260/280 das amostras de RNA com variação entre 1,8 e 2,0, indicando boa qualidade das amostras.



**Figura 5.** Integridade do RNA total extraído de folhas de eucalipto, crescido no campo. A) Extração de RNA total de folhas de eucalipto dos clones 520 assintomático, 520 sintomático e 1404 resistente visualizado por coloração em Brometo de Etídeo em gel de agarose 1%. B) Visualização de extração de RNA total dos clones de eucalipto após tratamento com DNase.

**Tabela 2.** Quantificação total de RNA

<b>Amostra</b>	<b>Concentração (ng/ul)</b>	<b>OD260/280</b>
520 AS1	1694,8	1,84
520 AS2	1800	1,81
520 AS3	1344	1,83
520 SS1	1453	1,81
520 SS2	840	2,0
520 SS3	1200	1,98
1404 R1	1065	1,97
1404 R2	1337	1,8
1404 R3	1967	1,81

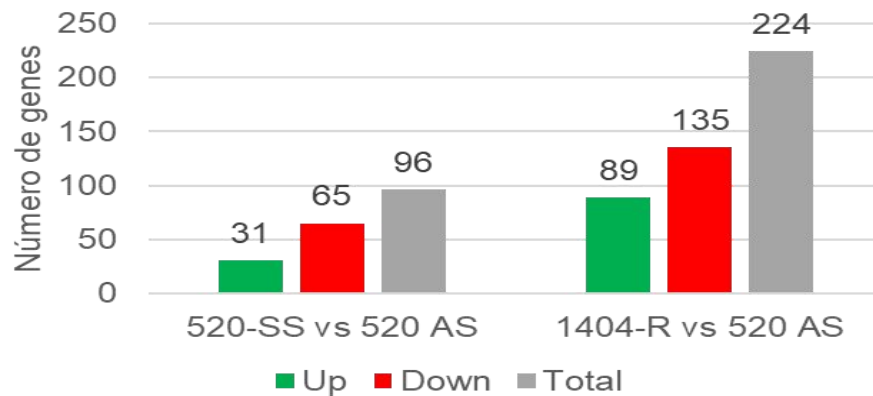
A qualidade do RNA extraído foi determinado por meio do *RNA integrity number* (RIN), com um cut off de 7 para RIN (Figura 5B) e Tabela 4. As amostras enviadas com valor de RIN superior à 7 prosseguiram para o sequenciamento.

**Tabela 3.** Controle de qualidade do RNA

<b>Clone</b>	<b>RIN</b>	<b>rRNA Ratio</b>
520 AS1	8,6	1,0
520 AS2	7,6	1,1
520 AS3	7,9	1,1
520 SS1	7,0	1,2
520 SS2	7,0	1,2
520 SS3	7,0	1,1
1404 R1	7,5	1,0
1404 R2	7,5	1,4
1404 R3	7,0	1,0

As amostras 520 AS correspondem aos clones suscetíveis assintomáticos e foram utilizadas como um parâmetro de referência para permitir a avaliação da variação global da expressão gênica em contraste com o genótipo 520 suscetível sintomático (520-SS) e 1404 resistente (1404-R). Os critérios para genes diferencialmente expressos foram p-valor de 0,05 e log<sub>2</sub>-fold-change >1,0 para de genes *up* regulados e log<sub>2</sub> fold-change < -1,0 para genes *down* regulados (Figura 6). O contraste 520 SS x 520 AS revelou 96 genes diferencialmente expressos, sendo 32,29% (31) regulados positivamente (Tabela S1) e 67,71% (65) regulados negativamente (Tabela S2). Entre os genes diferencialmente expressos, o número de genes

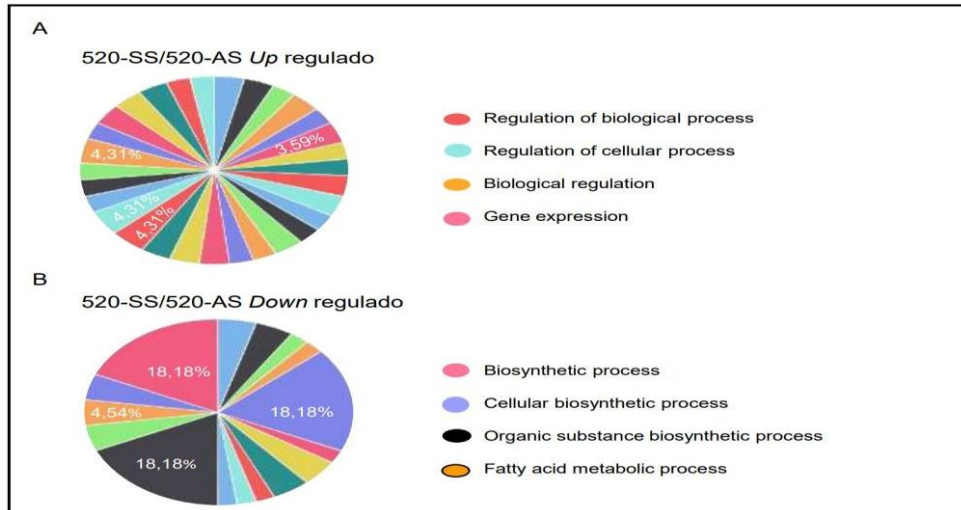
*down* regulados foi superior aos *up* regulados, consistentes com o fracionamento de genes diferencialmente expressos em resposta ao distúrbio fisiológico em caules do clone 520. Em se tratando do mesmo genótipo sujeito às mesmas condições de crescimento e em estágio de desenvolvimento similar, estes contrastes revelam o transcriptoma induzido pelo distúrbio fisiológico em folhas.



**Figura 6.** Análise da expressão gênica induzida pelo distúrbio fisiológico em folhas de clones de eucalipto. No eixo y é verificado os contrastes analisados, 520-SS vs 520-AS e 1404-R1. O número de genes diferencialmente expressos se encontra no eixo x.

A comparação entre o genótipo resistente e suscetível assintomático (1404-R vs 520-AS) revelou um total de 224 genes diferencialmente expressos (DE), sendo 39,73% (89) de genes regulados positivamente e 60,26% (135) regulados negativamente. O conjunto de genes diferencialmente expressos nesse cenário provavelmente resultam de diferenças naturais entre os genótipos, além de genes possivelmente responsáveis pelo caráter de resistência ao distúrbio fisiológico. Além disso, foi realizada análise através da função Gene Ontology (GO) utilizando o comando “Biological Process” utilizando  $\log_2$ -fold-change  $>1,0$  para de genes *up* regulados e  $\log_2$  fold-change  $< -1,0$  para genes *down* regulados, com p-valor de 0,05.

No contraste 520-SS x 520-AS (Figura 7A), dos 31 genes regulados positivamente, 5 genes (4.31%), Eucgr.B03208, Eucgr.C00780, Eucgr.F02691, Eucgr.H01200, Eucgr.H02116 e Eucgr.I01153, foram comumente enriquecidos em diversas categorias similares como regulação biológica, regulação de processos biológico, regulação de processo celular, metabolismo de RNA, expressão gênica, transcrição modelada por DNA, etc.



**Figura 7.** Categorização funcional de genes diferencialmente expressos baseado em processo biológico, a partir do contraste 520-SS vs 520-AS, conforme obtido na ontologia dos genes. A) Processos biológicos *down* regulado, cada cor representa um processo diferente. B) Processos biológicos *up* regulado no contraste analisado.

O gene Eucgr.B03208 está relacionado com os fatores de transcrição do tipo NAC, que em *Arabidopsis* já foi relacionado como mediador da sinalização de auxina para desenvolvimento de raiz lateral e presente em resposta a estresses abióticos (DE CAMARGO, 2007). Uma vez que o estresse abiótico é considerado uns dos possíveis desencadeadores do distúrbio fisiológico é possível que este gene regulador de transcrição seja funcional no distúrbio fisiológico. Além disso, a família de fatores de transcrição NAC está relacionado com desenvolvimento da planta, resposta ao estresse e senescência foliar, sendo a senescência característica principal observada em folhas de eucaliptos com distúrbio fisiológico, apresentando folhas amareladas e encarquilhadas (MA, 2020).

Com relação aos genes regulados positivamente Eucgr.C00780 e Eucgr.F02691, denominados *Ethylene Response Factor* (ERF) ou seja, responsável ao etileno, pertencem à superfamília AP2/ERF que possuem um domínio de ligação ao DNA do tipo AP2/ERF, envolvidos em regulação de respostas à estresse abiótico (YIN, 2022). Sabe-se que estresse biótico não desencadeia o distúrbio fisiológico em eucalipto (Martins, 2020), e os resultados do transcriptoma induzido pelo distúrbio fisiológico em eucalipto sugerem que o estresse abiótico seja fator determinante para o aparecimento desse distúrbio, o que seria consistente com a regulação positiva dos genes AP2/ERF-like em resposta ao estabelecimento do distúrbio fisiológico. A condição hídrica, irradiação solar, ambiente de cultivo e salinidade dos solos são exemplos de fatores que podem ser responsáveis por estresse abiótico (ROGRIGUES, 2022). Já o gene Eucgr.H01200 descrito como glutaredoxina monotiol está relacionado com resposta

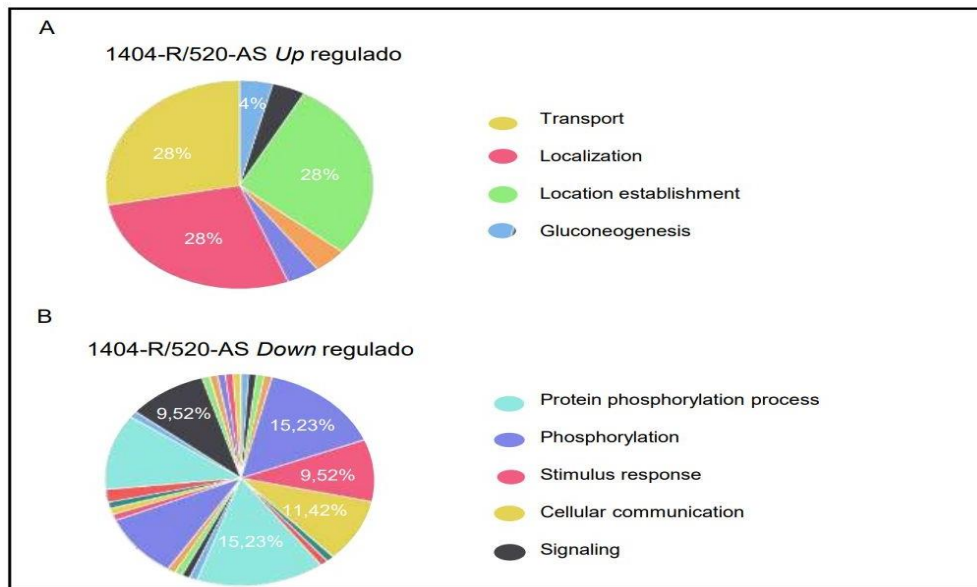
ao estresse oxidativo, atuando como mecanismo de defesa da planta. As glutaredoxinas são responsáveis por manter o ambiente redox celular, que pode sofrer alteração devido ao acúmulo de espécies reativas de oxigênio (ROS), desencadeado por estresses biótico ou abiótico (MALIK, 2020). Também consistente com esta observação seria a regulação positiva do gene Eucgr.I01153 descrito como *REDOX-RESPONSIVE TRANSCRIPTION FACTOR 1 (RRTF1)*. Em *Arabidopsis thaliana*, o possível gene ortólogo de Eucgr.I01153, possui papel crucial na regulação da rede redox por promover homeostase redox após a exposição a perturbações fotossintéticas, uma vez que a inativação de *RRTF1* torna as plantas sensíveis ao elevado estresse pela luz. É destacado também o papel do RRTF1 em resposta a estresse salino e fungos (SOLIMAN; MEYER, 2019). O gene Eucgr.H02116 é pertencente à família NBS-LRR com domínio TIR, e está relacionada com o sistema imune de plantas, devido a resistência a patógenos, mas nenhuma caracterização funcional desse gene em outras espécies de planta foi conduzida (ZHOU, 2022).

Na análise de processo biológico dos 65 genes regulados negativamente pelo contraste 520-SS x 520-AS (Figura 7B), a categoria processo biossintético celular foi enriquecida com 8 genes (18,18%), sendo que os genes envolvidos na biossíntese de substância orgânica predominaram com 4,54% envolvidos em processo metabólico de ácidos graxos. Os processos biossintéticos apresentaram genes em comum, sendo eles o Eucgr.A02390, Eucgr.C02487, Eucgr.D00056, Eucgr.D01144, Eucgr.F01170, Eucgr.F03588 e Eucgr.G03319. O gene Eucgr.A02390 é descrito como *Dehydration-Responsive Element-Binding Protein 2F (DREB2)*, que engloba fatores de transcrição envolvidos na tolerância à seca e ao estresse térmico, capazes de ativar a transcrição por meio do elemento responsivo à desidratação (DRE) (KIDOKORO, 2015). Diversos fatores de transcrição da subfamília DREB já foram identificados em diversas outras espécies de planta além de *Arabidopsis*, com atuação na tolerância ao estresse (MIZOI et al., 2013). O gene Eucgr.C02487, descrito como *Wrky Transcription Factor 32*, pertence à família dos genes WRKY, que é um grupo de fatores de transcrição específicos da planta, com papel fundamental em resposta a estímulos de estresse abiótico como temperatura, seca e irradiação UV (ZHANG, 2021). A regulação negativa desses fatores transcripcionais que respondem ao estresse abiótico pode sugerir que o distúrbio fisiológico não caracteriza um estado de estresse similar à temperatura, seca e radiação UV. Provavelmente, a regulação negativa possa ser utilizada como marcador específico do estado de distúrbio fisiológico.

O gene Eucgr.D00056, regulado negativamente pelo distúrbio, é homólogo de *3-Ketoacyl-CoA Synthase 2 (KCS2)*, responsável pela biossíntese de ácidos graxos de cadeia

longa que são primordiais para a formação de ceras cuticulares que impedem a perda de água, lipídeos de membrana e esfingolipídeos de plantas (YANG, 2021). A identificação do gene *KCS2* *down* regulado pelo distúrbio em folhas é condizente com a observação prévia (MARTINS, 2020; patente BR BR102022025625-0, suplemento S1) de que o gene *KCS1* também é regulado negativamente pelo distúrbio em caules. Similarmente, o gene Eucgr.D01144 *Protein Eceriferum 1 (CER1)*, é *down* regulado pelo distúrbio em folhas, sendo relacionado com a biossíntese de lipídeos, mais precisamente na formação de cera cuticular, que promove a proteção contra perda hídrica das folhas (LIU, 2022). A regulação negativa dos genes relacionados a produção de cutina e cera cuticular, promove uma menor produção desses lipídeos, facilitando a perda de água pelas folhas. Além de explicar o fenótipo de murcha foliar observado pelo estabelecimento da síndrome do distúrbio fisiológico, a regulação negativa desses genes envolvidos em algum aspecto do metabolismo de ácidos graxos, lipídeos e cera cuticular é consistente com a *down*-regulação da síntese de ácidos graxos nos caules acometidos pelo distúrbio fisiológico. Provavelmente, o perfil de expressão diferencial desses genes nas folhas acometidas pelo distúrbio fisiológico possa ser usado para desenvolvimento de diagnósticos moleculares do distúrbio fisiológico em eucalipto. O gene Eucgr.F01170 é descrito como proteína da superfamília reguladora da transcrição ANAC098, que atuam regulando processos biológicos distintos, como o desenvolvimento de raízes e folhas e estão relacionados à resposta a estresses bióticos e abióticos (NOGUEIRA, 2005). O gene Eucgr.G03319 codifica a proteína ribossômica S6 (RPS6B), sendo ela uma proteína que sofre fosforilação desencadeada por estímulos fisiológicos e patológicos (RAJAMAKI, 2007).

No contraste 1404-R x 520-AS, dos 89 genes *up* regulados, apenas 8 genes constituíram categorias significativamente enriquecidas ( $p < 0.05$ ), sendo 7 classificados em categorias relacionadas com transporte e 1 único gene em gliconeogênese (Eucgr.E02131, *phosphoglucose isomerase 1*). A categoria transporte em geral incluem os genes Eucgr.F00506, Eucgr.H04444, Eucgr.H04447, Eucgr.H04448, Eucgr.H04527, Eucgr.L03170 e Eucgr.K00968, classificados na categoria biológica transporte, localização, e estabelecimento de localização (Figura 8A).



**Figura 8.** Categorização funcional de genes diferencialmente expressos baseado em processo biológico, a partir do contraste 1404-R vs 520-AS, conforme obtido na ontologia dos genes. A) Processos biológicos *down* regulado, cada cor representa um processo diferente. B) Processos biológicos *up* regulado no contraste analisado.

O gene Eucgr.F00506, é denominado como gene *Mate Efflux Family Protein* ou gene de Extrusão de Compostos Tóxicos e Multidrogas (*MATE*), pois são responsáveis por realizar a remoção de toxinas, garantindo condições ideais de sobrevivência da planta. Essa família de genes é responsável pelo efluxo de substâncias através de um gradiente eletroquímico de cátions e ânions na membrana celular, sendo responsável pelo transporte de metabólitos secundários, eliminação de xenobióticos, desintoxicação por metais pesados e alumínio, além da regulação da resistência a doenças (LI et al., 2019). Eucgr.K00968 também descrito em transporte é homólogo de Polyol Transporter, envolvido dessa forma, com o transporte de polióis. Os polióis possuem a capacidade de retirar espécies reativas de oxigênio (ROS) com intuito de regular a pressão osmótica (YOSHINO, 2019). Os demais genes categorizados em transporte, Eucgr.H04444, Eucgr.H04447, Eucgr.H04448, Eucgr.H04527 e Eucgr.L03170, correspondem a um grupo de proteínas que possuem domínio NTF2L comum, responsáveis pelo transporte de proteínas para o núcleo das células além de participar do processo de homeostase de oxigênio (ROS) em espécies nativas (LUO, 2022).

No contraste 1404-R x 520-AS (Figura 8B), 16 genes constituíram a categoria significativamente enriquecida de fosforilação de proteínas e 10 genes, a categoria transdução de sinais, resposta a estímulos ou comunicação celular. Com relação aos genes do processo de fosforilação de proteína e fosforilação foi identificados os genes Eucgr.A00693, Eucgr.C00989,

Eucgr.C02543, Eucgr.C03748, Eucgr.D01480, Eucgr.E03625, Eucgr.F01313, Eucgr.F01840, Eucgr.G00291, Eucgr.G01270, Eucgr.J00545, Eucgr.K00502, Eucgr.K02406, Eucgr.K03098, Eucgr.L01669, Eucgr.L03256, que codificam *receptor-like kinases* ou outro tipo de cinase.

Os genes citados são descritos como *Protein kinase domain* (Pkinase), codificando proteínas com domínio quinase, que realizam a catálise de proteínas quinases. Essas enzimas são responsáveis pela fosforilação de proteínas através da transferência de grupamento fosfato para resíduos de treonina, serina ou tirosina, sendo essenciais para a comunicação intracelular, regulação e transdução de sinais (CASTILLO, 2019).

Com relação ao grupo de genes incluídos no processo resposta à estímulos, os genes Eucgr.E02595, Eucgr.B01064, Eucgr.L03738, Eucgr.H01823, Eucgr.H02207, Eucgr.E02315, Eucgr.J01316 codificam proteínas de resistência (classe TIR-NBS-LRR), envolvidas no processo imune de plantas, conforme já descrito anteriormente (ZHOU, 2022). Também incluído nesta categoria, o gene Eucgr.E03204, um receptor transmembrana envolvido em processo de resistência a doenças (WANG, 2020). Estes resultados indicam que de fato o mecanismo de resistência ao distúrbio fisiológico não está associado ao sistema imune e, consequentemente, substanciam o argumento de que sinais abióticos são os determinantes predominantes que acionam a síndrome do distúrbio fisiológico. Também foram incluídos nesta categoria os genes Eucgr.H04573, exonuclease III, com função de fosfomonoesterase na posição 3' do DNA (LIU, 2020 e Eucgr.L01744, *Extra-Large Guanine Nucleotide-Binding Protein 2*, que está relacionado com transdução de sinais envolvidos no crescimento, desenvolvimento e respostas a estresses bióticos e abióticos (PETUTSCHNIG, 2022).

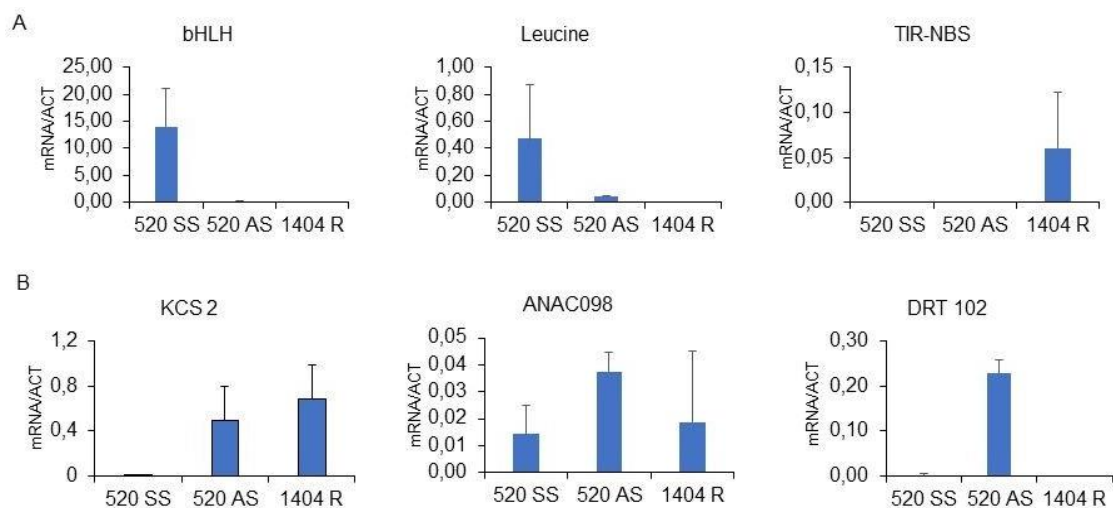
#### 4.3 Análise da expressão genica de genes candidatos diferencialmente expressos

Para validar os resultados obtidos no RNA-Seq, foi utilizado como metodologia de análise PCR quantitativo em tempo real (qRT-PCR) das amostras 520 SS, 520 AS e 1404 R. Para as análises foram selecionados três genes *up* e *down* regulados no contraste 520 SS x 520 AS. Os genes regulados positivamente escolhidos foram, bHLH (basic helix-loop-helix) relacionado com diferenciação celular e regulação do ciclo celular, Leucine pertencente a família de proteína quinase e TIR-NBS que está relacionado com resistência a patógenos. Com relação aos genes regulados negativamente, foram selecionados o gene KCS 2 envolvido na via da síntese de ácidos graxos, ANAC098 (No Apical Meristem) relacionado com fatores de transcrição envolvidos no desenvolvimento das plantas e DRT 102 (DNA-damage-repair/toleration protein) relacionado ao reparo de DNA. Como controle endógeno foi utilizado

o gene ACT (Actin-related protein 3).

De modo geral, todos os genes selecionados apresentaram acúmulos de transcritos consistentes com os resultados do RNA-Seq (Figura 9). Com relação aos genes positivamente regulados (9A) bHLH e Leucine, foi verificado que o acúmulo de transcritos foi superior na amostra 520 SS em contraste com a amostra 520 AS, confirmando o resultado obtido pelo RNA-seq. Com relação ao gene TIR-NBS não foi possível visualizar diferença entre o contraste 520 SS x 520 AS, mas foi possível confirmar o enriquecimento do gene em questão na amostra 1404 R (clone resistente) em contraste com o clone 520 AS (clone assintomático), conforme verificado no RNA-seq.

Com relação aos genes regulados negativamente (9B) KCS 2, ANAC098 e DRT 102, os três apresentaram um padrão de resultado esperado, com maior acúmulo de transcritos no clone 520 AS quando contrastado com o clone 520 SS. Além disso, os genes KCS 2 e ANAC098 se apresentaram regulado positivamente no contraste 1404R x 520 AS. No RNA-seq, foi verificado que o gene DRT102 é *down* regulado no contraste 520 AS x 1404 R, conforme confirmado pelo PCR quantitativo. Portanto, o resultado obtido pelo qRT-PCR confirma o resultado obtido pelo RNA-seq, validando dessa forma o ensaio de sequenciamento realizado.



**Figura 9:** Acúmulo de transcritos de genes marcadores em folhas de eucalipto de clone sintomático (520 SS), clone assintomático (520 AS) e clone resistente (1404 R). A expressão dos genes indicados na figura foi determinada por qRT-PCR e a expressão genica foi calculada pelo método do  $2^{-\Delta\Delta CT}$ . A) Expressão genica de genes regulados positivamente no contraste 520 SS x 520 AS. B) Expressão genica de genes regulados negativamente no contraste 520 SS x 520 AS. ACT (Actin-related protein 3) foi utilizado como controle endógeno. As barras indicam o desvio padrão da média, baseado em repetições técnicas e replicatas biológicas (n=3)

## 5. DISCUSSÃO

### 5.1 Desenvolvimento de marcadores moleculares associados ao distúrbio fisiológico

O distúrbio fisiológico é um estado de anomalia fisiológica que acomete clones suscetíveis no sul da Bahia, resultando no comprometimento da qualidade da madeira. Apesar de já ter sido identificado um componente genético no caráter resistência e suscetibilidade, as modificações moleculares associadas ao distúrbio não são conhecidas e marcadores moleculares que poderiam ser introduzidos em programas de melhoramento associados ao caráter suscetibilidade ou resistência não foram identificados.

A utilização de marcadores SNPs no cenário atual é crescente devido sua abundância no genoma, identificação mais barata, especificidade do *locus*, além da sua quantidade no genoma e a facilidade de identificação através de tecnologias de sequenciamento de próxima geração (NGS). A utilização dessa ferramenta de sequenciamento permite verificar a presença de SNPs, presença de inserções ou deleções e tamanho ou número de repetições em tandem ao longo de uma sequência de DNA em estudo. A análise desses aspectos fornece subsídios para identificação de marcadores moleculares, podendo ser utilizado para identificar uma variação genética em determinada região do genoma (KUMAR et al., 2021).

Nesta investigação, foram selecionados genes diferencialmente expressos pelo distúrbio fisiológico e analisados a presença de SNPs nos respectivos promotores que pudessem estar associados ao caráter resistência ou suscetibilidade. Foram utilizadas sequencias de promotores de quatro clones resistentes e cinco clones suscetíveis. Esta estratégia permitiu identificar SNPs, mas não *Indels*, comuns aos promotores de clones resistentes e de clones suscetíveis. Além de clones suscetíveis e resistentes apresentaram *Indels* nas mesmas posições de promotores avaliados, os tamanhos de fragmentos preditos não foram consistentes com a mobilidade eletroforética dos fragmentos amplificados. A não ser pela baixa qualidade de sequenciamento, nenhum outro argumento científico é passível de ser levantado para reconciliar esta aparente contradição. Um exemplo característico foi a presença de inserção na região promotora do gene *KASIII* de todos os clones de eucalipto resistentes ao distúrbio fisiológico 1572, 1644, 1407 e 1404. Entretanto, foi confirmada a presença de inserção em dois clones de eucaliptos suscetíveis assintomáticos 520 e 5204, não caracterizando a inserção exclusiva de clones de eucaliptos resistentes. O gene *KASIII* seria um ótimo candidato para marcador molecular pois, é regulado negativamente no estado de distúrbio fisiológico conforme verificado no transcriptoma de caule. Além disso, está envolvido na síntese de ácidos graxos de cadeia longa,

sendo importante para a condensação inicial da biossíntese de ácidos graxos (NOFIANI, 2019).

Por outro lado, alguns SNPs puderam ser associados ao caráter resistência e suscetibilidade, embora uma associação genética ao fenótipo não foi avaliada ainda. Por exemplo, foram observados SNPs nas posições -1366 (T/C) e -1373,-1374 (AT/TC) do promotor do gene *RPT2* conservados para clones resistentes e suscetíveis. *RPT2* é induzido por estímulos abióticos e regulado positivamente pelo distúrbio fisiológico em caules sintomáticos, sendo esse gene responsável pela transdução de sinal regulação do fototropismo em *Arabidopsis* (SAKAI et al., 2000). Os oligonucleotídeos desenhados para o SNP -1366 não reconheceram especificamente o respectivo promotor e amplificaram dois fragmentos, indicando não ser específico para esta região do genoma de eucalipto. Consequentemente, os experimentos de associação genética com o fenótipo, utilizando estes oligonucleotídeos marcadores não puderam ser realizados. Experimentos para seleção de oligonucleotídeos marcadores capazes de distinguir SNPs na posição -1366 e -1373,-1374 estão em progresso.

Existem outras metodologias com maior sensibilidade para realização da genotipagem, como por exemplo a *KASP (Kompetitive Allele-Specific PCR)* que permite realizar identificação de SNPs em regiões específicas, através da leitura de fluorescência terminal, utilizando oligonucleotídeos fluorescente-específicos que permitem a marcação bialélica de polimorfismos de nucleotídeo único em uma reação de qRT-PCR. Essa metodologia é altamente precisa e com custo acessível e, para que a metodologia de *KASP* seja eficiente, o desenho dos oligonucleotídeos devem levar em consideração as variações entre sequências não-alvo, garantindo a detecção dos alelos polimórficos alvo (STEELE, 2021).

## **5.2 Transcriptoma induzido pelo distúrbio fisiológico em folhas de eucalipto**

O distúrbio fisiológico em eucalipto atualmente interfere diretamente na qualidade da madeira, devido aos seus principais sintomas como a necrose do tronco principal, perda da dominância apical, clorose e encarquilhamento das folhas. Devido a grande produção de eucalipto no Brasil, principalmente na região sul da Bahia, é importante esclarecer as bases moleculares do distúrbio fisiológico com um objetivo de realizar diagnósticos precoces, evitando maiores perdas na produção de eucalipto. Nesta investigação, foi determinado o transcriptoma induzido pelo distúrbio fisiológico em folhas sintomáticas do clone 520. Além disso, foi determinado a variação global da expressão gênica em folhas do clone resistente 1404, crescido nas mesmas condições indutoras do distúrbio fisiológico como do clone 520, comparado com o clone 520 assintomático.

O transcriptoma induzido pelo distúrbio em folhas sintomáticas do clone 520 revelou um enriquecimento de genes regulados positivamente na categoria fatores de transcrição induzidos por estresses abióticos, como genes da família NAC e ERF, que estão de fato relacionados com respostas ao estresse abiótico em plantas. Além disso, foi verificado que o gene que codifica a proteína glutaredoxina monotiol e genes da família RRTF1, que promovem a manutenção de um ambiente redox quando desestabilizado por espécies reativas de oxigênio (ROS), induzido por estresses abióticos, foram também regulados positivamente. Estes resultados indicam que o ambiente redox é um componente importante no estabelecimento do estado de distúrbio fisiológico. Além disso, indicam que o estado do distúrbio fisiológico compartilha com outros estresses abióticos a regulação positiva de fatores de transcrição normalmente induzidos por estresses, não caracterizando uma resposta específica do distúrbio fisiológico.

Na análise de genes regulados negativamente pelo distúrbio fisiológico em folhas sintomáticas do clone 520 foi observado um enriquecimento da categoria “metabolismo de ácidos graxos”, sendo consistente com os resultados previamente obtidos para o transcriptoma induzido pelo distúrbio em caules sintomáticos do clone 520. A produção desses ácidos graxos de cadeia longa, cutina, suberina e ceras é importante para a manutenção da turgidez das folhas, uma vez que se trata de substâncias hidrofóbicas responsáveis pela impermeabilização de tecidos evitando a perda de água por meio da transpiração. Além disso, os ácidos graxos de cadeia longa são essenciais para o desenvolvimento das plantas por isso, um declínio na produção desses ácidos graxos afeta diretamente o desenvolvimento da planta (DE REZENDE, 2016).

Na análise dos genes diferencialmente expressos em caules de eucaliptos 520 suscetíveis realizado em estudo anterior (BR102022025625-0), foi verificado o enriquecimento de genes e fatores de transcrição relacionados ao estresse abiótico, sendo esse resultado compatível com os obtidos na análise dos genes DE em folhas dos clones 520. Além disso, foi destacado a repressão de genes relacionados a síntese de ácidos graxos de cadeia longa, sendo também verificado em caules suscetíveis. Estes resultados confirmam a hipótese da presença de genes diferencialmente expressos comumente associados ao estabelecimento do distúrbio fisiológico tanto em caules quanto em folhas. A observação de que a diminuição do acúmulo de ácidos graxos pode ocorrer em dois órgãos distintos, que apresentam sintomas fenótipos diferentes durante o estabelecimento do distúrbio, deve ser analisado em experimentos futuros como estímulo molecular que leva ao desenvolvimento dos sintomas do distúrbio fisiológico.

No contraste dos clones 1404 R vs 520 AS, crescidos sob condições indutoras do

distúrbio fisiológico, a categoria biológica antioxidante foi enriquecida para genes regulados positivamente, além de genes envolvidos em transporte e localização celular de proteínas (ROS) (LUO, 2022, PINTO, 2016). Possivelmente, a expressão aumentada desses genes contribui para caráter de resistência do clone 1404 em resposta aos estímulos que desencadeiam o distúrbio fisiológico. Em contraste, sob condições indutoras do distúrbio, houve uma repressão acentuada de genes envolvidos no sistema imune de plantas no clone resistente. Estes resultados descartam a possibilidade de que estímulos bióticos seja fator determinante para desencadear o distúrbio fisiológico, sendo absolutamente consistentes com os resultados do transcriptoma identificado nos caules do clone resistente comparado com o clone suscetível.

Em resumo, a análise da expressão diferencial de genes demonstrou que o distúrbio fisiológico induz mudanças moleculares tanto específicas como comuns a outros estresses abióticos, como seca, temperatura e luz UV. As mudanças moleculares conservadas induzidas no estado de distúrbio fisiológico incluem genes envolvidos no controle do estado redox, revelando em comum a outros estresses abióticos, o estado antioxidante com um dos determinantes do distúrbio. Como mudanças moleculares específicas do distúrbio, destaca-se a repressão de genes envolvidos no metabolismo de ácidos graxos, consistente com o transcriptoma induzido pelo distúrbio em caules sintomáticos. Além disso, foi possível comprovar uma sobreposição de mudanças moleculares na expressão gênica no transcriptoma induzido pelo distúrbio no caule e nas folhas. Uma vez que os sintomas do distúrbio divergem no caule e nas folhas, o transcriptoma sobreposto derivado de diferentes órgãos deve revelar estímulos moleculares comuns que podem levar ao estado do distúrbio fisiológico.

## 6. REFERENCIAS

- ALVARENGA, S. M. et al** (2009). Categorização funcional de sequências expressas envolvidas na defesa do cafeeiro a doenças. 2009.
- BERGE, K. V. et al.** (2019). RNA Sequencing Data: Hitchhiker's Guide to Expression Analysis. Annual Review of Biomedical Data Science. Vol. 2: p. 139-173.
- BOLGER A. M., Lohse M., Usadel B.** (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30,15, 2114-2120
- CASTILLO, J. E. F.** (2019). Análise estrutural de sítios de fosforilação em proteína. 2019. Tese de Doutorado. Universidade de São Paulo.
- CHANG, S.; PURYEAR, J.; CAIRNEY, J.** (1993). A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Report*, v. 11, p. 113-116.
- DE CAMARGO, S. R.** (2007). Identificação de Genes e Uso de Promotores Modulados por Etanol em Cana-de-açúcar. 2007. Tese de Doutorado. Instituto de Biologia.
- DE REZENDE, F. M. et al.** (2016). Vias de síntese de metabólitos secundários em plantas. *Laboratório de Ensino de Botânica*, v. 93, 2016.
- DOBIN A. et al.** (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013 Jan 1;29(1):15-21. doi: 10.1093/bioinformatics/bts635. Epub 2012 Oct 25. PMID: 23104886; PMCID: PMC3530905.
- FADASON, T. et al.** (2021). Atribuindo função aos SNPs: Considerações ao interpretar a variação genética. In: *Seminários em Biologia Celular e do Desenvolvimento*. Imprensa Acadêmica.
- FALCON, S.; GENTLEMAN, R.** (2007). Using GOstats to test gene lists for GO term association. *Bioinformatics*. 23(2), p. 257-258.
- FERNANDES, J. P. P.** (2021). Desempenho silvicultural e fisiológico de clones de eucalyptus sp. Sob duas condições hídricas em campo. Dissertação de Mestrado, FCAV, São Paulo.
- FILHO, Gil Monteiro** (2019). Sequenciamento de nova geração dos pontos de quebra do DNA para investigação dos mecanismos de formação em rearranjos genômicos. Tese de Doutorado. Universidade de São Paulo.
- FLORENCIO, G. W. L. et al.** (2021) Mudanças climáticas no eucalipto: modelos de desenvolvimento e zoneamento de aptidão climática. Dissertação de Mestrado em Meio Ambiente e Recursos Hídricos - Instituto de Recursos Naturais, Universidade Federal de Itajubá, Itajubá, p. 75.

- GOIS, A. F. D. de.** (2021). Caracterização do proteoma de folhas, caules e raízes de plantas de *Eucalyptus grandis* sob déficit hídrico. Tese de Doutorado. Universidade de São Paulo.
- HANDY, S. M. et al.** (2017). Evaluation of the Agilent Technologies bioanalyzer-based DNA fish identification solution. *Food Control*, v. 73, p. 627-633.
- HAMDI, K. et al.** (2020). Abscisic acid, stress, and ripening (TtASR1) gene as a functional marker for salt tolerance in durum wheat. *BioMed Research International*, v. 2020, 2020.
- IBA.** (2020). Indústria Brasileira de Árvores. Relatório IBA 2020. 202, 66 p.
- ILLUMINA, TRUSEQ, R. N. A.** (2011). TruSeq™ RNA and DNA Sample Preparation Kits v2. CHIP-SEQ RIBO-SEQ, p. 133.
- KIDOKORO, Satoshi e cols** (2016). Os fatores de transcrição do tipo DREB 1/CBF da soja funcionam no calor e na seca, bem como na expressão gênica responsiva ao estresse frio. *The Plant Journal*, v. 81, n. 3, pág. 505-518, 2015.
- KIM, B. M., et al.** (2013). Expression profile analysis of antioxidative stress and developmental pathway genes in the manganese-exposed intertidal copepod *Tigriopus japonicus* with 6K oligochip. *Chemosphere*, 92(9), 1214-1223.
- KVASTAD, L. et al.** (2021). The spatial RNA integrity number assay for in situ evaluation of transcriptome quality. *Communications biology*, v. 4, n. 1, p. 1-8, 2021.
- KUMAR, Sachin et al.** (2021). Advances in molecular markers and their use in genetic improvement of wheat In: Physiological, molecular, and genetic perspectives of wheat improvement. Springer, Cham, 2021. p. 139-174.
- LADO, S., et al.** (2020). Genome-wide diversity and global migration patterns in dromedaries follow ancient caravan routes. *Communications Biology* 3, 387. <https://doi.org/10.1038/s42003-020-1098-7>
- LIAO, Y., SMYTH, G. K.; SHI, W.** (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.*41(10), e108-e108.
- LI, Y., et al.**(2019). Genome-wide analysis of the MATE gene family in potato. *Molecular biology reports*, v. 46, n. 1, p. 403-414, 2019.
- LIU, Q. et al.** (2022). Two Arabidopsis MYB-SHAQKYF transcription repressors regulate leaf wax biosynthesis via transcriptional suppression on DEWAX. *New Phytologist*, 2022.
- LUO, L.; WANG, Q.; MA, F.** (2021). RNA-Seq transcriptome analysis of ileum in Taiping chicken supplemented with the dietary probiotic. *Tropical Animal Health and Production*, v. 53, n. 1, p. 1-12.
- LUO, Dong et al.** (2022). NUCLEAR TRANSPORT FACTOR 2-LIKE improves drought tolerance by modulating leaf water loss in alfalfa (*Medicago sativa* L.). *The Plant Journal*, v. 112, n. 2, p. 429-450, 2022.
- Luo, W.** (2013). “Pathview: an R/Bioconductor package for pathway-based data integration and visualization.” *Bioinformatics*, 29(14), 1830-1831. doi: [10.1093/bioinformatics/btt285](https://doi.org/10.1093/bioinformatics/btt285)

**MAFIA, R. G.** (2017). Etiologia e Manejo do Distúrbio Fisiológico do Eucalipto – DFE. In: II Seminário sobre sanidade florestal. p. 143, Belo Horizonte – MG.

**MA, L. et al** (2020). Involvement of NAC transcription factor NaNAC29 in *Alternaria alternata* resistance and leaf senescence in *Nicotiana attenuata*. *Plant diversity*, 2020.

**MALIK, W. A. et al** (2020). Genome-wide expression analysis suggests glutaredoxin genes response to various stresses in cotton. *International journal of biological macromolecules*, v. 153, p. 470-491, 2020.

**MIZOI, J. et al.** (2013). GmDREB2A; 2, a canonical DEHYDRATION-RESPONSIVE ELEMENT-BINDING PROTEIN2-type transcription factor in soybean, is posttranslationally regulated and mediates dehydration-responsive element-dependent gene expression. *Plant physiology*, v. 161, n. 1, p. 346-361, 2013.

**MORAZÁN, A. S. O.** (2019). Análises bioinformáticas de proteínas de resistência do tipo nucleotide binding site-leucine rich repeat (NBS-LRR) de café. 2019.

**NOGALES, A; DEDIEGO, M.** (2019) Host single nucleotide polymorphisms modulating influenza a virus disease in humans. *Pathogens* 8:168-189.

**NOGUEIRA, F. T. et al.** (2005). SsNAC23, a member of the NAC domain protein family, is associated with cold, herbivory and water stress in sugarcane. *Plant Science*, v. 169, n. 1, p. 93-106, 2005.

**NOFIANI, R. et al.** (2019). 3-Ketoacyl-ACP synthase (KAS) III homologues and their roles in natural product biosynthesis. *MedChemComm*, v. 10, n. 9, p. 1517-1530, 2019.

**OLIVEIRA, A. J. et al.** (2021). Principais marcadores moleculares. *Research, Society and Development*, v. 10, n. 15, p. e562101523633-e562101523633.

**PEREIRA, M. A.; IMADA, E. L.; GUEDES, R. L. M.** (2017). RNA-seq: Applications and Best Practices. In: MARCHI, F. A.; CIRILLO, P. D. R.; MATEO, E. C. *Applications of RNASeq and Omics Strategies – From Microorganisms to Human Health*.

**PETUTSCHNIG, E. et al.** (2022). EXTRA LARGE G-PROTEIN2 mediates cell death and hyperimmunity in the chitin elicitor receptor kinase 1-4 mutant. *Plant Physiology*, v. 189, n. 4, p. 2413-2431, 2022.

**PINTO, R. T.** (2016). Caracterização funcional de genes da família multdrug and toxic compound extrusion (MATE) envolvidos na adaptação de *Coffea arabica* à solos ácidos. 2016.

**RAJAMÄKI, M. L. et al** (2007). Differential requirement of the ribosomal protein S6 and ribosomal protein S6 kinase for plant-virus accumulation and interaction of S6 kinase with potyviral VPg. *Molecular Plant-Microbe Interactions*, v. 30, n. 5, p. 374-384, 2017.

**RODRIGUES, A. C. P.** (2021). Perfil de expressão gênica em híbridos de *Eucalyptus grandis* x *Eucalyptus urophylla* afetados pelo distúrbio fisiológico do eucalipto (DFE). Tese de Doutorado, Universidade Federal dos Vales do Jequitinhonha e Mucuri.

**RODRIGUES, B. P. et al.** (2022). Chemical and Kraft Pulping Properties of Young Eucalypt Trees Affected by Physiological Disorders. *Forests*, v. 13, n. 4, p. 504.

**SAKAI, T. et al.** (2000). RPT2. A signal transducer of the 491 phototropic response in

Arabidopsis. *Plant Cell* 12: 225-236.

**SANTAROSA, E.; PENTEADO JÚNIOR, J. F.; GOULART, ICG dos R.** (2014).

Transferência de tecnologia florestal: cultivo de eucalipto em propriedades rurais: diversificação da produção e renda. Embrapa Florestas-Fôlder/Folheto/Cartilha (INFOTECA-E), 2014.

**SANTOS, J. Naiara.** (2021). Expressão gênica em híbridos. afetados por distúrbio fisiológico. Programa de Pós-Graduação em Ciência Florestal, p. 13, 2021.

**SHAW, R.; TIAN, Xin; XU, Jian.** (2021). Single-cell transcriptome analysis in plants: advances and challenges. *Molecular Plant*, v. 14, n. 1, p. 115-126, 2021.

**SILVA, S.R. et al.** (2010) Palestra no 'VI Workshop de Melhoramento e XVI Reunião Técnica Científica do Programa Temático de Silvicultura e Manejo - PTSM' (Botucatu - SP): Distúrbios fisiológicos ocasionados pelo estresse hídrico em clones de eucalipto na Veracel: estudo de caso e hipóteses.

**SOLIMAN, E. R. S; MEYER, P.** (2019). Responsiveness and adaptation to salt stress of the redox-responsive transcription factor 1 (RRTF1) gene are controlled by its promoter. *Molecular biotechnology*, v. 61, n. 4, p. 254-260, 2019.

**SOUZA, B. de C.** (2021). Composição genética da raia viola *Pseudobatos percellens* (elasmobranchii: rhinobatidae) utilizando SNPs. Dissertação de Mestrado, Instituto de Biociências de Botucatu, Universidade Estadual Paulista “Júlio de [Type here] Mesquita Filho”, p. 49, São Paulo.

**SCHREIBER, K. J. et al** (2016). Multiple domain associations within the Arabidopsis immune receptor RPP1 regulate the activation of programmed cell death. *PLoS pathogens*, v. 12, n. 7, p. e1005769, 2016.

**STEELE, K.** (2021). Desenvolvimento de marcadores KASP para identificação de variedades de arroz basmati. *Métodos Analíticos de Alimentos*, v. 14, n. 4, pág. 663-673, 2021. See More

**VALE, A. B. et al.** (2014). Eucaliptocultura no Brasil: Silvicultura, manejo e ambiência. 551 p. Viçosa, MG:SIF.

**WANG, Y. et al.** (2021). A estrutura regulatória da morte celular programada no desenvolvimento em órgãos florais: uma revisão. *Plant Physiology and Biochemistry*, v. 158, p. 103-112, 2021.

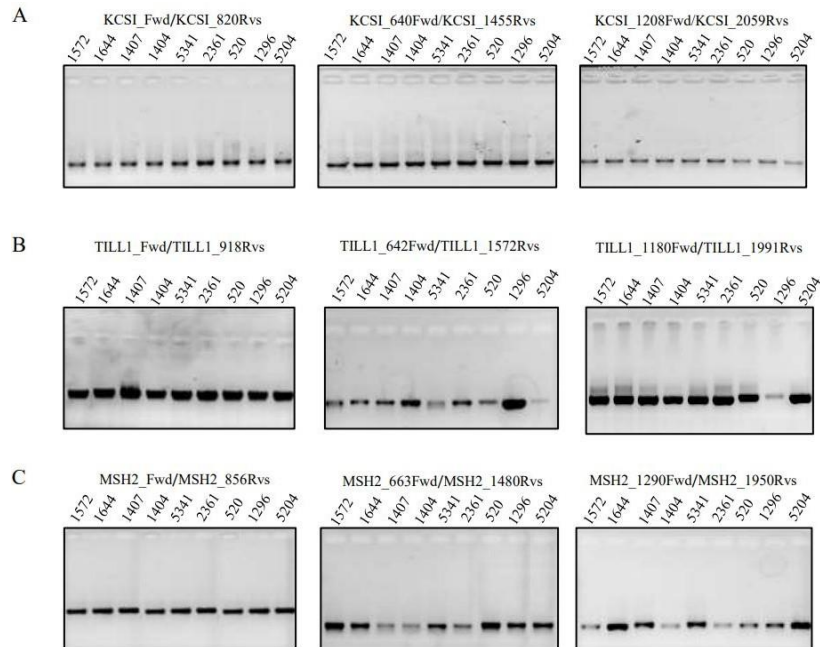
**WANG, Guodong et al.** (2010). The diverse roles of extracellular leucine-rich repeat-containing receptor-like proteins in plants. *Critical Reviews in Plant Science*, v. 29, n. 5, p. 285-299, 2010.

**WARSI, M. K. et al.** (2021). Regulon: Uma visão geral do sistema regulador da transcrição do estresse abiótico em plantas e papel nas plantas transgênicas. *Brazilian Journal of Biology*, v. 83, 2021.

**WOLD, J., et al.** (2021). Expanding the conservation genomics toolbox: incorporating structural variants to enhance functional studies for species of conservation concern. Preprints. Special Issue: Whole genome sequencing in molecular ecology, Vol 30, p 5949- 5965.

- YAN, H., W. et al.** (2019) Genome-Wide Characterization, Evolution, and Expression Profiling of VQ Gene Family in Response to Phytohormone Treatments and Abiotic Stress in *Eucalyptus grandis*. *International Journal of Molecular Sciences*, Vol. 20.
- YANG, T. et al.** (2021). The 3-ketoacyl-CoA synthase WFL is involved in lateral organ development and cuticular wax synthesis in *Medicago truncatula*. *Plant Molecular Biology*, v. 105, n. 1, p. 193-204, 2021.
- YIN, D. et al.** (2022). Os fatores de ligação do elemento responsivo ao etileno contribuem para a tolerância ao encharcamento, regulando os parâmetros fotossintéticos e fisiológicos na petúnia. *Ciência Rural*, v. 53, 2022.
- YOSHINO, K. et al.** (2019). The conservation of polyol transporter proteins and their involvement in lichenized Ascomycota. *Fungal biology*, v. 123, n. 4, p. 318-329, 2019.
- ZHANG, Lei et al.** (2021). Integrative analysis of metabolome and transcriptome reveals molecular regulatory mechanism of flavonoid biosynthesis in *Cyclocarya paliurus* under salt stress. *Industrial Crops and Products*, v. 170, p. 113823, 2021.
- ZHOU, Li et al.** (2022). A novel TIR-NBS-LRR gene regulates immune response to *Phytophthora* root rot in soybean. *The Crop Journal*, 2022.
- ZOLET, A. C. T. et al.** (2017) Marcadores moleculares na era genômica: metodologias e aplicações. Ribeirão Preto: Sociedade Brasileira de Genética, p. 181.
- ZIMMERMAN, S. J. et al** (2020). An empirical comparison of population genetic analyses using microsatellite and SNP data for a species of conservation concern. *BMC Genomics* 21, p. 1–16

## 7. MATERIAL SUPPLEMENTAR



**Figura S1:** A) Reação de cadeia polimerase com os oligonucleotídeos do promotor do gene KCSI. B) PCR com os oligonucleotídeos do promotor TILL 1 e C) PCR com os oligonucleotídeos do promotor MSH2. O número na nomenclatura do oligonucleotídeo corresponde aproximadamente à posição de anelamento no promotor do referido gene. Os clones 1572, 1644, 1407 e 1404 são resistentes e os clones 5341, 2361, 520, 1296 e 5204 são suscetíveis ao distúrbio fisiológico.

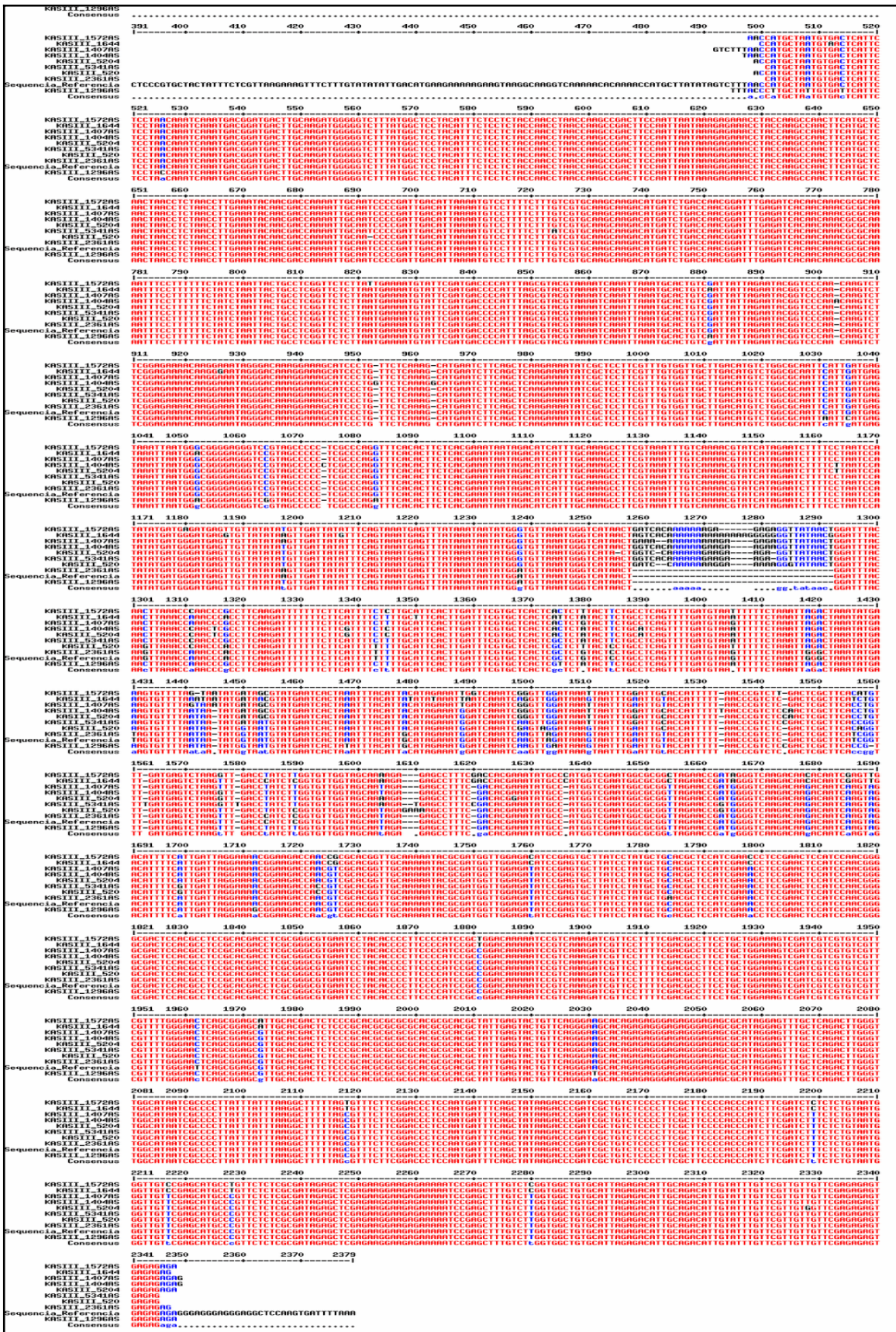


Figura S2: Sequenciamento do promotor do gene KASIII com 2000 pb, dos clones de eucalipto 1572. 1644, 1407, 1404, 5341, 2361, 520, 1296 e 5204.

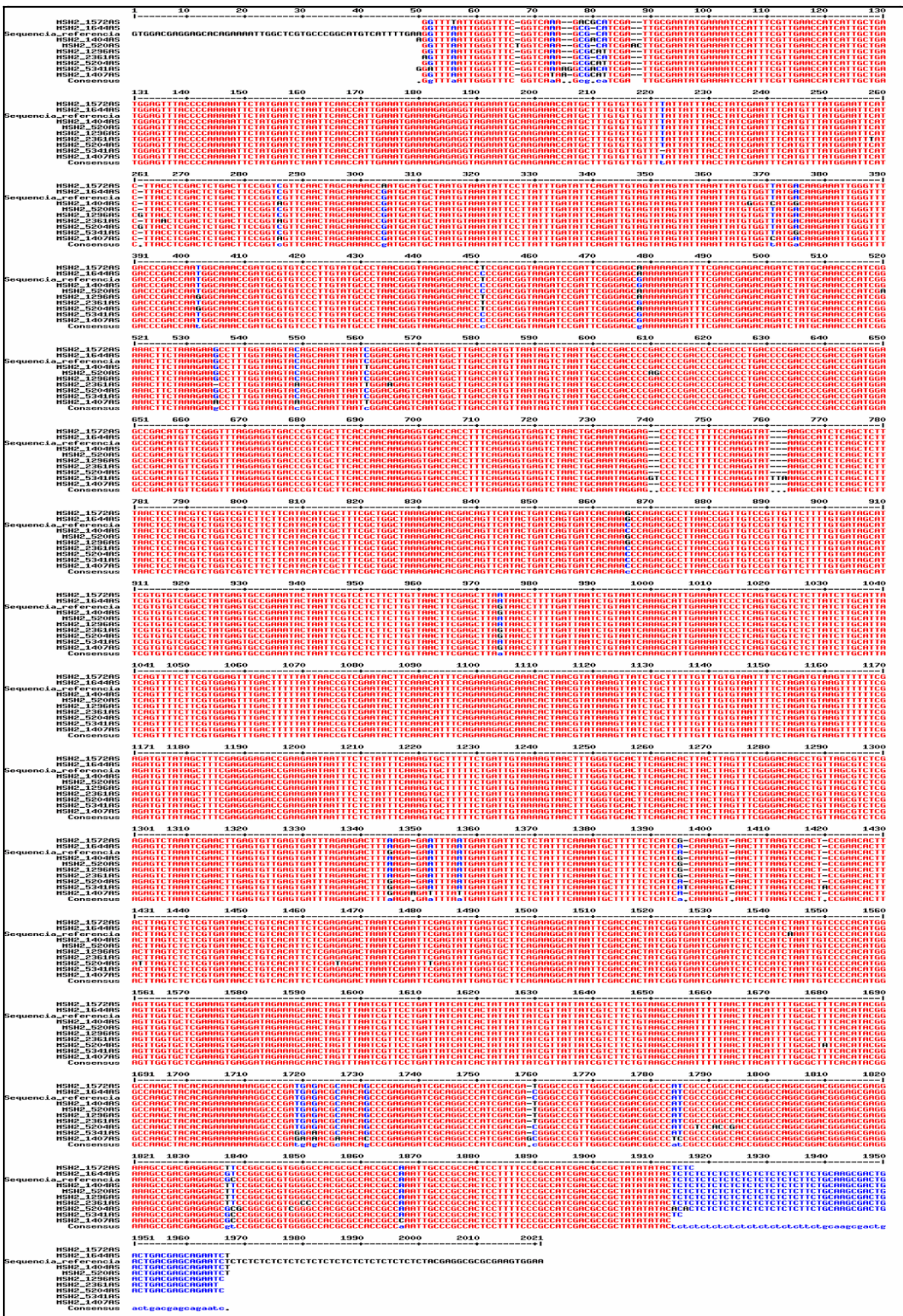


Figura S4: Sequenciamento do promotor do gene MSH2 com 2000 pb, dos clones de eucalipto 1572, 1644, 1407, 1404, 5341, 2361, 520, 1296 e 5204.





	781	790	800	810	820	830	840	850	860	870	880	890	900	910
Galac_15729S														
Galac_5341RS														
Galac_5204RS														
Sequencia_Referencia														
Galac_1296RS														
Galac_1404RS														
Galac_1407RS														
Galac_2361RS														
Consensus														
	911	920	930	940	950	960	970	980	990	1000	1010	1020	1030	1040
Galac_15729S														
Galac_5341RS														
Galac_5204RS														
Sequencia_Referencia														
Galac_1296RS														
Galac_1404RS														
Galac_1407RS														
Galac_2361RS														
Consensus														
	1041	1050	1060	1070	1080	1090	1100	1110	1120	1130	1140	1150	1160	1170
Galac_15729S														
Galac_5341RS														
Galac_5204RS														
Sequencia_Referencia														
Galac_1296RS														
Galac_1404RS														
Galac_1407RS														
Galac_2361RS														
Consensus														
	1171	1180	1190	1200	1210	1220	1230	1240	1250	1260	1270	1280	1290	1300
Galac_15729S														
Galac_5341RS														
Galac_5204RS														
Sequencia_Referencia														
Galac_1296RS														
Galac_1404RS														
Galac_1407RS														
Galac_2361RS														
Consensus														
	1301	1310	1320	1330	1340	1350	1360	1370	1380	1390	1400	1410	1420	1430
Galac_15729S														
Galac_5341RS														
Galac_5204RS														
Sequencia_Referencia														
Galac_1296RS														
Galac_1404RS														
Galac_1407RS														
Galac_2361RS														
Consensus														
	1431	1440	1450	1460	1470	1480	1490	1500	1510	1520	1530	1540	1550	1560
Galac_15729S														
Galac_5341RS														
Galac_5204RS														
Sequencia_Referencia														
Galac_1296RS														
Galac_1404RS														
Galac_1407RS														
Galac_2361RS														
Consensus														
	1561	1570	1580	1590	1600	1610	1620	1630	1640	1650	1660	1670	1680	1690
Galac_15729S														
Galac_5341RS														
Galac_5204RS														
Sequencia_Referencia														
Galac_1296RS														
Galac_1404RS														
Galac_1407RS														
Galac_2361RS														
Consensus														
	1691	1700	1710	1720	1730	1740	1750	1760	1770	1780	1790	1800	1810	1820
Galac_15729S														
Galac_5341RS														
Galac_5204RS														
Sequencia_Referencia														
Galac_1296RS														
Galac_1404RS														
Galac_1407RS														
Galac_2361RS														
Consensus														
	1821	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950
Galac_15729S														
Galac_5341RS														
Galac_5204RS														
Sequencia_Referencia														
Galac_1296RS														
Galac_1404RS														
Galac_1407RS														
Galac_2361RS														
Consensus														
	1951	1960	1970	1980	1990	2000	2010	2020	2030	2040	2050	2060	2070	2080
Galac_15729S														
Galac_5341RS														
Galac_5204RS														
Sequencia_Referencia														
Galac_1296RS														
Galac_1404RS														
Galac_1407RS														
Galac_2361RS														
Consensus														
	2081	2090	2100	2110	2120	2130	2140	2150	2160	2170	2180	2190	2200	2210
Galac_15729S														
Galac_5341RS														
Galac_5204RS														
Sequencia_Referencia														
Galac_1296RS														
Galac_1404RS														
Galac_1407RS														
Galac_2361RS														
Consensus														
	2211	2220	2230	2240	2250	2260	2270	2280	2290	2300	2310	2320	2330	2340
Galac_15729S														
Galac_5341RS														
Galac_5204RS														
Sequencia_Referencia														
Galac_1296RS														
Galac_1404RS														
Galac_1407RS														
Galac_2361RS														
Consensus														
	2341	2350	2360	2370	2380	2390	2400	2410	2420	2430	2440	2450	2460	2470
Galac_15729S														
Galac_5341RS														
Galac_5204RS														
Sequencia_Referencia														
Galac_1296RS														
Galac_1404RS														
Galac_1407RS														
Galac_2361RS														
Consensus														
	2471	2480	2490	2500	2510	2520	2530	2540	2550	2560	2570	2580	2590	2600
Galac_15729S														
Galac_5341RS														
Galac_5204RS														
Sequencia_Referencia														
Galac_1296RS														
Galac_1404RS														
Galac_1407RS														
Galac_2361RS														
Consensus														
	2601	2610	2620	2630	2640	2650	2660	2670	2680	2690	2700	2710	2720	2730
Galac_15729S														
Galac_5341RS														
Galac_5204RS														
Sequencia_Referencia														
Galac_1296RS														
Galac_1404RS														
Galac_1407RS														
Galac_2361RS														
Consensus														
	2731	2740	2750	2760	2770	2780	2790	2800	2810	2820	2830	2840	2850	2860
Galac_15729S														
Galac_5341RS														
Galac_5204RS														
Sequencia_Referencia														
Galac_1296RS														
Galac_1404RS														
Galac_1407RS														
Galac_2361RS														
Consensus														
	2861	2870	2880	2890	2900	2910	2920	2930	2940	2947				
Galac_15729S														
Galac_5341RS														
Galac_5204RS														
Sequencia_Referencia														
Galac_1296RS														
Galac_1404RS														
Galac_1407RS														
Galac_2361RS														
Consensus														

Figura S7: Sequenciamento do promotor do gene Galactose oxidase com 2000 pb, dos clones de eucalipto 1572, 1644, 1407, 1404, 5341, 2361, 520, 1296 e 5204.

A

Job Title **Nucleotide Sequence**

RID [XSTTJFVJ013](#) Search expires on 02-04 08:45 am [Download All](#) ▾

Program **BLASTN** [Citation](#) ▾

Database **nt** [See details](#) ▾

Query ID **lcl|Query\_25091**

Description **None**

Molecule type **nucleic acid**

Query Length **18**

Other reports [Distance tree of results](#) [MSA viewer](#) [?](#)

**Filter Results**

Organism *only top 20 will appear*  exclude

Type common name, binomial, taxid or group name

[+ Add organism](#)

Percent Identity  to  E value  to  Query Coverage  to

[Filter](#) [Reset](#)

**Descriptions** [Graphic Summary](#) [Alignments](#) [Taxonomy](#)

**Sequences producing significant alignments** [Download](#) ▾ [Select columns](#) ▾ [Show](#) 100 ▾ [?](#)

select all 100 sequences selected [GenBank](#) [Graphics](#) [Distance tree of results](#) [MSA Viewer](#)

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> uncharacterized LOC104449104 (LOC104449104), mRNA	<i>Eucalyptus gran...</i>	28.2	28.2	77%	1.3	100.00%	2613	XM_010083128.3
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> uncharacterized LOC120293220 (LOC120293220), transcript variant X...	<i>Eucalyptus gran...</i>	28.2	28.2	77%	1.3	100.00%	1232	XR_005650986.1
<input checked="" type="checkbox"/> <i>Eucalyptus grandis</i> isolate RLC_eoAngela_1 retrotransposon, partial sequence	<i>Eucalyptus gran...</i>	28.3	93.2	77%	5.2	100.00%	11280	KM198475.1
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> ubiquitin carboxyl-terminal hydrolase 5 (LOC104414259), mRNA	<i>Eucalyptus gran...</i>	28.3	28.3	72%	5.2	100.00%	3434	XM_039314585.1
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> TMV resistance protein N (LOC104457321), mRNA	<i>Eucalyptus gran...</i>	28.3	66.9	83%	5.2	100.00%	3933	XM_03820082.1
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> TMV resistance protein N-like (LOC108654789), transcript variant X3...	<i>Eucalyptus gran...</i>	28.3	28.3	72%	5.2	100.00%	9600	XM_018881287.2
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> TMV resistance protein N-like (LOC108654789), transcript variant X2...	<i>Eucalyptus gran...</i>	28.3	28.3	72%	5.2	100.00%	4657	XM_039300817.1
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> TMV resistance protein N-like (LOC108654789), transcript variant X1...	<i>Eucalyptus gran...</i>	28.3	28.3	72%	5.2	100.00%	4991	XM_039300816.1
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> disease resistance protein RUN1 (LOC104418700), mRNA	<i>Eucalyptus gran...</i>	28.3	28.3	72%	5.2	100.00%	2803	XM_038299203.1
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> uncharacterized LOC104445349 (LOC104445349), mRNA	<i>Eucalyptus gran...</i>	28.3	28.3	72%	5.2	100.00%	2650	XM_010090210.3
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> uncharacterized LOC104445198 (LOC104445198), transcript variant X...	<i>Eucalyptus gran...</i>	28.3	28.3	72%	5.2	100.00%	4282	XR_005651313.1
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> uncharacterized LOC104445198 (LOC104445198), transcript variant X...	<i>Eucalyptus gran...</i>	28.3	28.3	72%	5.2	100.00%	3885	XR_005651311.1
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> uncharacterized LOC104445198 (LOC104445198), transcript variant X...	<i>Eucalyptus gran...</i>	28.3	28.3	72%	5.2	100.00%	3933	XR_005651310.1
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> uncharacterized LOC104445198 (LOC104445198), transcript variant X...	<i>Eucalyptus gran...</i>	28.3	28.3	72%	5.2	100.00%	4321	XR_005651309.1

B

Job Title **Nucleotide Sequence**

RID [XSSXR7F7013](#) Search expires on 02-04 08:30 am [Download All](#) ▾

Program **BLASTN** [Citation](#) ▾

Database **nt** [See details](#) ▾

Query ID **lcl|Query\_53287**

Description **None**

Molecule type **nucleic acid**

Query Length **18**

Other reports [Distance tree of results](#) [MSA viewer](#) [?](#)

**Filter Results**

Organism *only top 20 will appear*  exclude

Type common name, binomial, taxid or group name

[+ Add organism](#)

Percent Identity  to  E value  to  Query Coverage  to

[Filter](#) [Reset](#)

**Descriptions** [Graphic Summary](#) [Alignments](#) [Taxonomy](#)

**Sequences producing significant alignments** [Download](#) ▾ [Select columns](#) ▾ [Show](#) 100 ▾ [?](#)

select all 100 sequences selected [GenBank](#) [Graphics](#) [Distance tree of results](#) [MSA Viewer](#)

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> uncharacterized LOC104449104 (LOC104449104), mRNA	<i>Eucalyptus gran...</i>	28.2	28.2	77%	1.3	100.00%	2613	XM_010083128.3
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> uncharacterized LOC120293220 (LOC120293220), transcript variant X...	<i>Eucalyptus gran...</i>	28.2	28.2	77%	1.3	100.00%	1232	XR_005650986.1
<input checked="" type="checkbox"/> <i>Eucalyptus grandis</i> isolate RLC_eoAngela_1 retrotransposon, partial sequence	<i>Eucalyptus gran...</i>	28.3	93.2	77%	5.2	100.00%	11280	KM198475.1
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> phosphoglycerate mutase-like protein 4 (LOC104418593), mRNA	<i>Eucalyptus gran...</i>	28.3	28.3	72%	5.2	100.00%	1221	XM_010029971.3
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> TMV resistance protein N (LOC104457321), mRNA	<i>Eucalyptus gran...</i>	28.3	66.9	83%	5.2	100.00%	3933	XM_03820082.1
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> TMV resistance protein N-like (LOC108654789), transcript variant X3...	<i>Eucalyptus gran...</i>	28.3	28.3	72%	5.2	100.00%	9600	XM_018881287.2
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> TMV resistance protein N-like (LOC108654789), transcript variant X2...	<i>Eucalyptus gran...</i>	28.3	28.3	72%	5.2	100.00%	4657	XM_039300817.1
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> TMV resistance protein N-like (LOC108654789), transcript variant X1...	<i>Eucalyptus gran...</i>	28.3	28.3	72%	5.2	100.00%	4991	XM_039300816.1
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> disease resistance protein RUN1 (LOC104418700), mRNA	<i>Eucalyptus gran...</i>	28.3	28.3	72%	5.2	100.00%	2803	XM_038299203.1
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> uncharacterized LOC104445349 (LOC104445349), mRNA	<i>Eucalyptus gran...</i>	28.3	28.3	72%	5.2	100.00%	2650	XM_010090210.3
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> ubiquitin carboxyl-terminal hydrolase 5 (LOC104414259), mRNA	<i>Eucalyptus gran...</i>	28.3	28.3	72%	5.2	100.00%	3434	XM_039314585.1
<input checked="" type="checkbox"/> <i>Eucalyptus grandis</i> chloroplast, complete genome	<i>Eucalyptus gran...</i>	24.5	105	72%	21	100.00%	180109	MG924399.1
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> lectin T (LOC104455177), mRNA	<i>Eucalyptus gran...</i>	24.3	24.3	88%	21	100.00%	1218	XM_010070009.1
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> anti-H <sub>2</sub> O lectin 1-like (LOC104455179), mRNA	<i>Eucalyptus gran...</i>	24.5	24.3	88%	21	100.00%	1155	XM_010070008.1
<input checked="" type="checkbox"/> PREDICTED: <i>Eucalyptus grandis</i> dual specificity protein kinase YAK1 homoloq (LOC104424263), partial...	<i>Eucalyptus gran...</i>	24.3	24.3	88%	21	93.75%	3333	XM_010038230.3

**S7:** A) Análise no BLAST com a sequência do oligonucleotídeo utilizado na genotipagem de amostras resistentes, em comparação com o genoma de *Eucalyptus grandis*. B) Análise no BLAST com a sequência do oligonucleotídeo utilizado na genotipagem de amostras suscetíveis, em comparação com o genoma de *Eucalyptus*

**Tabela S1.** Genes diferencialmente expressos (DE) *up* regulados no contraste 520 SS vs 520 AS

ID	Besemean	log2FC	Adj.t.p-value	Arabi Name	Arabi Symbol	Arabi Description
Eucgr.B01225	714,6140487	28,10140458		AT3G56970.1	BHLH038,ORG2	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
Eucgr.L00295	42,24863369	24,03724281		AT2G26695.1		Ran BP2/NZF zinc finger-like superfamily protein
Eucgr.F02404	26,52228522	23,33669919	1,2881E-06	AT4G31940.1	CYP82C4	cytochrome P450, family 82, subfamily C, polypeptide 4
Eucgr.D01509	24,62420988	23,24584691	1,37E-08	AT3G26040.1		HXXXD-type acyl-transferase family protein
Eucgr.J01495	18,09206877	22,78566468	5,47E-08	AT5G16080.1	AtCXE17,CXE17	carboxyesterase 17
Eucgr.F02697	16,43703594	22,61304358	7E-10			
Eucgr.K02588	16,33547025	22,56869321	1E-10	AT4G08850.1		Leucine-rich repeat receptor-like protein kinase family protein
Eucgr.B03208	13,85111511	22,43854576	4,8012E-06	AT1G69490.1	ANAC029,ATNAP,NA	NAC-like, activated by AP3/PI
Eucgr.L00724	13,50305477	22,33970137		AT2G29120.1	ATGLR2.7,GLR2.7	glutamate receptor 2.7
Eucgr.B02410	16,39714459	22,33061567		AT5G06900.1	CYP93D1	cytochrome P450, family 93, subfamily D, polypeptide 1
Eucgr.C00780	34,56958484	22,23933925		AT1G19210.1		Integrase-type DNA-binding superfamily protein
Eucgr.F02691	12,8238259	22,2388829	1,3E-09	AT5G61890.1		Integrase-type DNA-binding superfamily protein
Eucgr.H01200	12,7415195	22,16824344	6E-10	AT1G03020.1		Thioredoxin superfamily protein
Eucgr.I01153	14,6451455	22,05387478		AT4G34410.1	RRTF1	redox responsive transcription factor 1
Eucgr.E03604	22,33146422	19,63505062	2E-10	AT2G24960.2		
Eucgr.D00868	45,04363864	18,9556968	2,0544E-06	AT3G25820.1	ATTPS-CIN TPS-CIN	8-cineole terpene synthase-like sequence-1 terpene synthase-like sequence-1,8-cineole
Eucgr.C03485	28,35947729	18,77314298	2,2E-09	AT5G04700.1		Ankyrin repeat family protein
Eucgr.E02882	13,60412409	18,22936849	1,5044E-06	AT1G17020.1	ATSRG1,SRG1	senescence-related gene 1
Eucgr.F02832	11,98407458	17,94673643	5,21E-08	AT3G25940.1		TFIIB zinc-binding protein
Eucgr.L02612	17,81380403	17,92087847	3,07E-08	AT5G38280.1	PR5K	PR5-like receptor kinase
Eucgr.K00968	34,11343522	17,66948829	5,47E-08	AT2G20780.1		Major facilitator superfamily protein
Eucgr.G02364	21,86804742	17,2463464	1,89318E-05			
Eucgr.F00837	172,7989285	16,54860422	1,31904E-05	AT4G27190.1		NB-ARC domain-containing disease resistance protein
Eucgr.C03549	105,8953769	15,88877907	0,002480479	AT2G06050.2	DDE1,OPR3	oxophytodieneoate-reductase 3
Eucgr.H04444	10,65556061	15,69787717	0,010916026	AT1G27970.1	NTF2B	nuclear transport factor 2B
Eucgr.L03316	10,79690346	15,69667676	0,010916026			
Eucgr.H02116	20,20336037	15,65007653	0,011060286	AT5G36930.1		Disease resistance protein (TIR-NBS-LRR class) family
Eucgr.C03135	34,37607898	15,50757373	0,003052317	AT5G63050.1	EMB2759	embryo defective 2759
Eucgr.B02835	15,06854765	15,40070496	0,00208449	AT1G67000.1		Protein kinase superfamily protein

**Tabela S2.** Genes diferencialmente expressos (DE) *down* regulados no contraste 520 SS vs 520 AS.

EMS	Besemean	log2FC	Adj.t.p-value	Arabi Name	Arabi symbol	Arabi Description
Eucgr.A01319	84,17046346	-24,46370476	7,8E-09	AT1G04110.1	SDD1	Subtilase family protein
Eucgr.A02390	24,20809933	-22,41214044	3,52E-08	AT3G57600.1		Integrase-type DNA-binding superfamily protein
Eucgr.A02615	18,05757547	-21,97219226	0,00000002	AT3G51680.1		NAD(P)-binding Rossmann-fold superfamily protein
Eucgr.A02667	13,9873634	-22,02418613	1,6E-09	AT5G53210.1	SPCH	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
Eucgr.B01582	103,3715798	-16,46828623	0,002480479			
Eucgr.B01712	11,15769378	-14,45702347	0,030893136	AT3G04880.1	DRT102	DNA-damage-repair/toleration protein (DRT102)
Eucgr.B02645	96,86541035	-16,74676635	0,005038671	AT1G59960.1		NAD(P)-linked oxidoreductase superfamily protein
Eucgr.C00411	11,69964825	-17,70954614	1,59774E-05	AT1G69550.1		disease resistance protein (TIR-NBS-LRR class)
Eucgr.C00441	22939,54638	-18,53193463	5,7741E-06			
Eucgr.C02087	21,5719598	-16,6316391	0,000106035	AT1G27170.1		transmembrane receptors;ATP binding
Eucgr.C02487	10,97203911	-15,08548705	0,01911322	AT4G30935.1	ATWRKY32,WRKY32	WRKY DNA-binding protein 32
Eucgr.C03823	16,37633403	-18,22425975	3,07E-08	AT3G24550.1	ATPERK1,PERK1	proline extensin-like receptor kinase 1
Eucgr.D00056	71,35630363	-24,5435415	1,953E-07	AT1G04220.1	KCS2	3-ketoacyl-CoA synthase 2
Eucgr.D00475	65,49295799	-23,97304507	6,4E-09	AT1G29270.1		
Eucgr.D01144	13,02707775	-19,06430343	4E-10	AT1G02205.2	CER1	Fatty acid hydroxylase superfamily
Eucgr.D01195	582,889324	-12,12531022	0,013283852	AT3G26320.1	CYP71B36	cytochrome P450, family 71, subfamily B, polypeptide 36
Eucgr.D01340	308,9861501	-15,18521176	0,022120152	AT3G56190.1	ALPHA-SNAP2,ASNAP ALPHA-SNAP2,ASNAP	alpha-soluble NSF attachment protein 2
Eucgr.D01904	6104,493298	-10,6180182	0,005280797	AT4G11650.1	ATOSM34,OSM34	osmotin 34
Eucgr.E00183	15,19707678	-17,44526491	2,228E-07	AT4G00620.1		Amino acid dehydrogenase family protein
Eucgr.E00425	726,993942	-10,96111611	0,040131532	AT5G47500.1		Pectin lyase-like superfamily protein
Eucgr.E02867	1534,231817	-10,9311277	0,040131532	AT1G29670.1		GDSL-like Lipase/Acylhydrolase superfamily protein
Eucgr.E02942	13,62092209	-17,34174167	4,21556E-05	AT1G48175.1	emb2191	Cytidine/deoxycytidylate deaminase family protein
Eucgr.E03999	35,13171689	-22,85257543	0,000000004	AT3G17070.1		Peroxidase family protein
Eucgr.F01170	24,35847341	-22,33088774	2,93E-08	AT5G53950.1	ANAC098,ATCUC2,CU C2	NAC (No Apical Meristem) domain transcriptional regulator superfamily protein
Eucgr.F02843	18,95335899	-22,32292628	0	AT1G66370.1	AtMYB113,MYB113	myb domain protein 113
Eucgr.F03053	84,66185586	-24,37551478	6,83356E-20	AT3G50390.1		Transducin/WD40 repeat-like superfamily protein
Eucgr.F03306	30,37184136	-22,88059376	4E-10	AT5G53870.1	AtENODL1,ENODL1	early nodulin-like protein 1
Eucgr.F03588	38,37039826	-23,28844426	0	AT1G76420.1	ANAC031,CUC3,NAC36	NAC (No Apical Meristem) domain transcriptional regulator superfamily

					8	protein
Eucgr.F04098	101,4593581	-21,78092627	1E-10	AT4G15760.1	MO1	monooxygenase 1
Eucgr.G00367	23,56652876	-17,45741899	1,2172E-06	AT5G36930.2		Disease resistance protein (TIR-NBS-LRR class) family
Eucgr.G00379	34,3998387	-17,72137524	2,2E-09	AT1G27170.1		transmembrane receptors;ATP binding
Eucgr.G00438	16,11474277	-22,25501811	1,6E-09	AT3G19550.1		
Eucgr.G00766	23,86099117	-15,11332449	0,014161733			
Eucgr.G00791	11,8180463	-21,98370962	8,6668E-06	AT3G47210.1		Plant protein of unknown function (DUF247)
Eucgr.G00867	11,06599302	-17,64451185	0,000008564	AT5G64370.1	BETA-UP,PYD3	beta-ureidopropionase
Eucgr.G01294	695,2819198	-11,56413055	0,010916026	AT5G59845.1		Gibberellin-regulated family protein
Eucgr.G01327	219,2260605	-14,44224398	0,022591802	AT2G15220.1		Plant basic secretory protein (BSP) family protein
Eucgr.G02084	54,08619313	-23,35444429	5,20572E-17	AT2G01770.1	ATVIT1,VI T1	vacuolar iron transporter 1
Eucgr.G03319	14,26629873	-15,10796603	0,022120152	AT5G10360.1	EMB3010,R PS6B	Ribosomal protein S6e
Eucgr.H00209	17,32636107	-18,20903332	8,6668E-06	AT3G01680.1		
Eucgr.H01221	22,81862415	-22,26936578	2,27E-08	AT3G44540.1	FAR4	fatty acid reductase 4
Eucgr.H01222	28,2096902	-22,39969876	0	AT3G44540.1	FAR4	fatty acid reductase 4
Eucgr.H02961	22,92902575	-22,03209939	7,86E-08	AT5G55240.1	ATPXG2	ARABIDOPSIS THALIANA PEROXYGENASE 2
Eucgr.H03173	13,39309919	-15,20582596	0,022120152	AT4G13940.1	ATSAHH1, EMB1395	S-adenosyl-L-homocysteine hydrolase
Eucgr.H03907	26,08006087	-21,87446535	1,668E-07	AT3G09270.1	ATGSTU8, GSTU8	glutathione S-transferase TAU 8
Eucgr.H04527	1588,5329	-16,66616528	0,000106035	AT1G27970.1	NTF2B	nuclear transport factor 2B
Eucgr.H04839	19,90579383	-21,89896833	1E-10			
Eucgr.H05052	3198,575748	-12,56697736	0,000312402	AT5G63660.1	LCR74,PDF 2.5	Scorpion toxin-like knottin superfamily protein
Eucgr.H05101	14,91717186	-18,84551222	3,3106E-06	AT5G09220.1	AAP2	amino acid permease 2
Eucgr.I00763	23,08395578	-16,01336715	0,01911322	AT5G67270.1	ATEB1C,A TEB1H1,E B1 C	end binding protein 1C
Eucgr.I02479	14,94634473	-17,24329405	0,005145795	AT2G46600.1		Calcium-binding EF-hand family protein
Eucgr.I02543	66,32653954	-15,75032483	0,01185013	AT1G19715.3		Mannose-binding lectin superfamily protein
Eucgr.I02717	121,6361591	-24,53371853	2,2E-09	AT3G06470.1		GNS1/SUR4 membrane protein family
Eucgr.J00071	13,77214387	-19,75059373	2,8E-09	AT5G07010.1	ATST2A,S T2A	sulfotransferase 2 <sup>a</sup>

Eucgr.J00257	11,95066992	-21,98683674	8,6668E-06			
Eucgr.J01555	36,17267177	-23,05110161		0AT1G66370.1	AtMYB113, MYB113	myb domain protein 113
Eucgr.J01573	21,47063384	-21,40363671		0AT1G66370.1	AtMYB113, MYB113	myb domain protein 113
Eucgr.J02143	23,14273169	-19,74128272	6E-10	AT1G74790.1		Catalytics
Eucgr.K01068	14,79645727	-16,06075178	0,010916026	AT3G57380.1		Glycosyltransferase family 61 protein
Eucgr.K02946	24,81761467	-19,35329407	6E-10	AT5G01650.2		Tautomerase/MIF superfamily protein
Eucgr.L01034	32,8970132	-13,98944684	0,001584899	AT2G30660.1		ATP-dependent caseinolytic (Clp) protease/crotonase family protein
Eucgr.L01518	24,92930918	-15,19435535	0,010916026			

**Tabela S3.** Genes diferencialmente expressos (DE) up regulados no contraste 1404 R vs 520 AS.

ID	besemean	log2FC	Adj.t.p-value	Arabi Name	Arabi symbol	Arabi Description
Eucgr.A00308	209,1728603	7,856949297	0,027390805	AT5G19590.1		Protein of unknown function, DUF538
Eucgr.B00020	11,51864243	22,14688247	2,0083E-06	AT1G78170.1		
Eucgr.B01197	21,50718207	23,05467228	5,744E-07	AT2G45550.1	CYP76C4	cytochrome P450, family 76, subfamily C, polypeptide 4
Eucgr.B02410	16,39714459	16,70281415	3,2721E-06	AT5G06900.1	CYP93D1	cytochrome P450, family 93, subfamily D, polypeptide 1
Eucgr.B02835	15,06854765	22,519488	3E-10	AT1G67000.1		Protein kinase superfamily protein
Eucgr.C00007	85,09484412	24,72732831	4,74E-08	AT5G27680.1	RECQSIM	RECQ helicase SIM
Eucgr.C00008	23,35479466	22,04746032	2,3963E-06			
Eucgr.C00441	22939,54638	14,36517712	0,000278884			
Eucgr.C00780	34,56958484	19,15959201	1E-10	AT1G19210.1		Integrase-type DNA-binding superfamily protein
Eucgr.C01558	13,10662467	22,3325282	1,6209E-06	AT5G63950.1	CHR24	chromatin remodeling 24
Eucgr.C01650	19,17808448	22,88144856	7,263E-07	AT5G50140.1		Ankyrin repeat family protein
Eucgr.C01675	32,31420429	23,63414537	2,462E-07	AT1G18740.1		Protein of unknown function (DUF793)
Eucgr.C01737	51,02735779	24,42834594	7,41E-08	AT5G15500.2		Ankyrin repeat family protein
Eucgr.C01931	761,2465021	28,18267357	1E-10	AT1G28200.1	FIP1	FH interacting protein 1
Eucgr.C03112	15,08931549	22,54102703	1,1995E-06	AT2G41790.1		Insulinase (Peptidase family M16) family protein
Eucgr.C03135	34,37607898	23,82466649	1E-10	AT5G63050.1	EMB2759	embryo defective 2759
Eucgr.C03383	25,2370715	23,27415172	4,121E-07	AT5G51160.1		Ankyrin repeat family protein
Eucgr.C03485	28,35947729	23,38470045	3,37003E-16	AT5G04700.1		Ankyrin repeat family protein
Eucgr.C03549	105,8953769	25,7135359	0	AT2G06050.2	DDE1,OPR3	oxophytodienoate-reductase 3
Eucgr.D00360	101,6822926	25,28157684	1,94E-08	AT4G19090.1		Protein of unknown function (DUF594)
Eucgr.D00362	63,96234725	24,6685489	0,000000052	AT5G45540.1		Protein of unknown function (DUF594)
Eucgr.D00377	46,38529458	24,05741262	1,312E-07	AT4G19090.1		Protein of unknown function (DUF594)
Eucgr.D00386	84,08719131	24,63091073	5,44E-08	AT5G45470.1		Protein of unknown function (DUF594)
Eucgr.D00394	44,26741356	24,08211524	1,269E-07	AT5G45470.1		Protein of unknown function (DUF594)

Eucgr.D01319	36,67319994	23,87813259	1,687E-07	AT4G33070.1		Thiamine pyrophosphate dependent pyruvate decarboxylase family protein
Eucgr.D01509	24,62420988	15,35304986	0,003425628	AT3G26040.1		HXXXD-type acyl-transferase family protein
Eucgr.D01691	17,43715559	22,90406688	7,064E-07	AT5G41970.1		Metal-dependent protein hydrolase
Eucgr.E01920	11,01852336	21,83925059	3,2479E-06	AT4G21380.1	ARK3,RK3	receptor kinase 3
Eucgr.E02131	21,18370338	23,01701803	6,019E-07	AT4G24620.1 AT4G24620.2	PGI PGI1	phosphoglucose isomerase 1
Eucgr.E02336	13,39519985	22,39201675	1,4895E-06	AT1G79740.1		hAT transposon superfamily
Eucgr.E02882	13,60412409	22,25179036	0	AT1G17020.1	ATSRG1,SRG1	senescence-related gene 1
Eucgr.E03556	61,45183318	24,4175566	7,49E-08	AT1G69550.1		disease resistance protein (TIR-NBS-LRR class)
Eucgr.E03604	22,33146422	22,35889707	0	AT2G24960.2		
Eucgr.E03803	12,84794884	22,24882651	2,857E-07	AT1G69550.1 AT5G44510.1		disease resistance protein (TIR-NBS-LRR class) target of AVR-B operation 1
Eucgr.E04140	105,6710013	10,08205461	0,031840725	AT1G61480.1		S-locus lectin protein kinase family protein
Eucgr.F00506	12,10703765	22,21372712	1,26E-08	AT1G33080.1		MATE efflux family protein
Eucgr.F00823	129,5475467	8,761872186	0,031840725	AT4G26090.1	RPS2	NB-ARC domain-containing disease resistance protein
Eucgr.F00837	172,7989285	26,04874642	1,1679E-17	AT4G27190.1		NB-ARC domain-containing disease resistance protein
Eucgr.F01358	106,2494958	25,37142531	1E-10	AT2G24960.2		
Eucgr.F01492	324,5155411	7,345444721	0,040185592	AT4G31970.1	CYP82C2	cytochrome P450, family 82, subfamily C, polypeptide 2
Eucgr.F01876	101,1315215	24,80051609	6E-10			
Eucgr.F02691	12,8238259	17,51509151	6,4431E-06	AT5G61890.1		Integrase-type DNA-binding superfamily protein
Eucgr.F02697	16,43703594	16,93463938	2,64631E-05			
Eucgr.F02832	11,98407458	22,13040837	0	AT3G25940.1		TFIIB zinc-binding protein
Eucgr.F03146	31,5439176	22,92221454	6,3E-09	AT3G47570.1		Leucine-rich repeat protein kinase family protein
Eucgr.G00928	82,93733291	24,98957367	3,11E-08	AT4G15390.1		HXXXD-type acyl-transferase family protein
Eucgr.G01747	27,10201749	23,37382788	3,564E-07	AT4G08850.1		Leucine-rich repeat receptor-like protein kinase family protein
Eucgr.G02364	21,86804742	23,04903255	0			
Eucgr.G03328	46,46862461	24,05088909	1,317E-07			
Eucgr.H00193	572,6997912	12,54223294	0,017344492	AT1G04280.1		P-loop containing nucleoside triphosphate hydrolases superfamily protein
Eucgr.H01200	12,7415195	18,55224939	3,564E-07	AT1G03020.1		Thioredoxin superfamily protein
Eucgr.H01446	26,03099388	23,31059157	3,914E-07	AT5G17680.1		disease resistance protein (TIR-NBS-LRR class), putative
Eucgr.H01449	123,0794693	25,35316341	1,73E-08	AT5G17680.1		disease resistance protein (TIR-NBS-LRR class), putative

Eucgr.H02116	20,20336037	22,89851559	3,11E-08	AT5G36930.1		Disease resistance protein (TIR-NBS-LRR class) family
Eucgr.H04444	10,65556061	22,01444475	1,467E-07	AT1G27970.1	NTF2B	nuclear transport factor 2B
Eucgr.H04447	15,92475629	22,677688	9,842E-07	AT1G27970.1	NTF2B	nuclear transport factor 2B
Eucgr.H04448	165,9775695	26,24264598	3,8E-09	AT1G27970.1	NTF2B	nuclear transport factor 2B
Eucgr.H04527	1588,5329	12,38896781	0,005010921	AT1G27970.1	NTF2B	nuclear transport factor 2B
Eucgr.I00750	276,7453953	9,130995385	0,025139253	AT4G37370.1	CYP81D8	cytochrome P450, family 81, subfamily D, polypeptide 8
Eucgr.I00772	27,33115543	23,391561	3,529E-07	AT4G37360.1	CYP81D2	cytochrome P450, family 81, subfamily D, polypeptide 2
Eucgr.I01153	14,6451455	20,40279059	0	AT4G34410.1	RRTF1	redox responsive transcription factor 1
Eucgr.I02285	14,65995032	22,50572063	1,2591E-06	AT1G71120.1	GLIP6	GDSL-motif lipase/hydrolase 6
Eucgr.J01495	18,09206877	15,45511057	0,003898047	AT5G16080.1	AtCXE17,CXE17	carboxyesterase 17
Eucgr.J02578	51,13821942	24,43800394	7,35E-08	AT1G14130.1		2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
Eucgr.J02681	28,71978199	23,48104782	3,083E-07	AT4G02210.1		
Eucgr.K00692	32,24378792	23,63132748	9,39E-08	AT4G03500.1		Ankyrin repeat family protein
Eucgr.K00968	34,11343522	23,16106617	0	AT2G20780.1		Major facilitator superfamily protein
Eucgr.K00995	23,54288408	21,99540319	0,000002575	AT3G42170.1		BED zinc finger ;hAT family dimerisation domain
Eucgr.K01379	18,5896918	22,85606449	1,675E-07	AT5G48290.1		Heavy metal transport/detoxification superfamily protein
Eucgr.K01764	11,18909307	22,10327543	5,804E-07			
Eucgr.K02588	16,33547025	18,43955167	2,485E-07	AT4G08850.1		Leucine-rich repeat receptor-like protein kinase family protein
Eucgr.K02743	29,67238318	23,50326276	8,48E-08	AT5G63380.1		AMP-dependent synthetase and ligase family protein
Eucgr.K03095	59,06622072	23,63821866	5E-10	AT2G24960.2		
Eucgr.K03150	24,83707419	22,06766075	2,3487E-06	AT2G24960.2		
Eucgr.L00295	42,24863369	20,97741173	0	AT2G26695.1		Ran BP2/NZF zinc finger-like superfamily protein
Eucgr.L00700	57,45099518	24,46143021	2,58E-08	AT1G27170.2		transmembrane receptors;ATP binding
Eucgr.L00724	13,50305477	17,04235783	4,35E-08	AT2G29120.1	ATGLR2.7,GLR2.7	glutamate receptor 2.7
Eucgr.L01415	10,59490409	22,02729084	9,859E-07	AT5G37990.1		S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
Eucgr.L01649	318,6783144	26,99329284	4E-10			
Eucgr.L01936	497,0112285	26,86704206	3E-10	AT4G13440.1		Calcium-binding EF-hand family protein
Eucgr.L01937	287,169286	26,76038159	3E-10	AT4G13440.1		Calcium-binding EF-hand family protein
Eucgr.L02321	10,38309445	22,167596	8,189E-07	AT3G27420.1		
Eucgr.L02612	17,81380403	22,70299667	0	AT5G38280.1	PR5K	PR5-like receptor kinase
Eucgr.L02728	17,4541199	22,96574795	6,496E-07	AT1G14130.1		2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein

**Tabela S4.** Genes diferencialmente expressos (DE) *down* regulados no contraste 1404 R vs 520 AS

ID	besemean	log2FC	Adj.p-value	Arabi Name	Arabi symbol	Arabi Description
Eucgr.A00028	42,64588141	-22,83333915	1,09964E-18			
Eucgr.A00030	21,30840532	-21,35891097	6,85489E-16	AT3G59600.1	,NRPD8B,NR PE8B	RNA polymerase Rpb8
Eucgr.A00234	100,7929431	-24,71647691	5,94747E-21	AT3G07510.2		
Eucgr.A00235	21,2107183	-22,48739739	2,66628E-17	AT3G07510.2		
Eucgr.A00459	66,11787318	-23,33540869	3,60053E-20			
Eucgr.A00693	179,9084538	-25,39964698	3,4271E-22	AT5G38260.1		Protein kinase superfamily protein
Eucgr.A01016	64,96993718	-24,12474872	5,02605E-20			
Eucgr.B01064	186,3787758	-25,13754674	4,88676E-22	AT5G17680.1		disease resistance protein (TIR-NBS-LRR class), putative
Eucgr.B01132	247,9041396	-26,181602	1,38106E-23	AT1G03220.1		Eukaryotic aspartyl protease family protein
Eucgr.B01135	566,2434243	-7,713213224	0,025736027	AT1G03220.1		Eukaryotic aspartyl protease family protein
Eucgr.B01226	1220,169154	-12,68051895	0,031281436	AT3G56970.1	BHLH038,ORG2	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
Eucgr.B01230	51,19612371	-23,159874	1,62712E-18	AT3G14470.1		NB-ARC domain-containing disease resistance protein
Eucgr.B01384	19,49630547	-21,48875724	0	AT4G27650.1	PEL1	Eukaryotic release factor 1 (eRF1) family protein
Eucgr.B01744	5290,356906	-12,21451255	4,43739E-05	AT5G54160.1	ATOMT1 OMT1 ATOMT1,OMT1	O-methyltransferase 1
Eucgr.B01747	7376,096772	-30,54837898	4,10075E-32	AT5G54160.1	ATOMT1 OMT1	O-methyltransferase 1
Eucgr.B02219	287,9683262	-7,680488287	0,046606605	AT3G17210.1	ATHS1,HS1	heat stable protein 1
Eucgr.B02642	580,6209775	-24,47754059	3,838E-18	AT1G59960.1		NAD(P)-linked oxidoreductase superfamily protein
Eucgr.B02788	39,73914323	-23,35055525	5,18712E-20	AT3G50120.1		Plant protein of unknown function (DUF247)
Eucgr.B02913	15,57651618	-19,71650166	0	AT4G14060.1		Polyketide cyclase/dehydrase and lipid transport superfamily protein
Eucgr.B03326	382,6132892	-25,05947622	3,4412E-20	AT1G23760.1	JP630,PG3	BURP domain-containing protein
Eucgr.B03330	45,17548682	-22,01428476	1,12161E-16	AT1G23760.1	JP630,PG3	BURP domain-containing protein
Eucgr.B03654	42,53331986	-21,03481697	2,5E-09	AT2G44480.1	BGLU17	beta glucosidase 17
Eucgr.C00334	1868,966917	-28,25591595	3,14903E-27			
Eucgr.C00336	20,27002213	-21,94479026	6,68928E-16	AT5G11280.1		

Eucgr.C00702	141,0900679	-24,13144187	5,23458E-20	AT5G50140.1		Ankyrin repeat family protein
Eucgr.C00989	16,26599698	-20,86817402	0	AT5G25930.1		Protein kinase family protein with leucine-rich repeat domain
Eucgr.C01324	39,81790913	-22,51486663	5,87404E-17			
Eucgr.C01905	20,65312449	-21,93121137	0	AT3G16520.3	UGT88A1	UDP-glucosyl transferase 88A1
Eucgr.C02003	41,24346201	-18,55511251	2,361E-07	AT2G29120.1	ATGLR2.7, GLR2.7	glutamate receptor 2.7
Eucgr.C02073	21,19425853	-21,55412926	0			
Eucgr.C02260	20,02345795	-22,1696453	4,34837E-16	AT5G57280.1		S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
Eucgr.C02543	357,4987429	-9,548341481	0,003018921	AT1G69730.1		Wall-associated kinase family protein
Eucgr.C02704	31,08141325	-21,44435532	0	AT1G69550.1		disease resistance protein (TIR-NBS-LRR class)
Eucgr.C03353	25,44985515	-22,69371249	0	AT3G18670.1		Ankyrin repeat family protein
Eucgr.C03532	13,29515272	-21,83153674	0	AT3G18670.1		Ankyrin repeat family protein
Eucgr.C03642	23,76680153	-22,71479811	0	AT5G04730.1		Ankyrin-repeat containing protein
Eucgr.C03748	16,14994772	-20,58027948	0	AT3G47570.1		Leucine-rich repeat protein kinase family protein
Eucgr.C04045	18,53705581	-21,9451206	0			
Eucgr.D00056	71,35630363	-24,63361757	5,44E-08	AT1G04220.1		3-ketoacyl-CoA synthase 2
Eucgr.D01480	32,46752496	-23,19445144	4,634E-07	AT3G47110.1		Leucine-rich repeat protein kinase family protein
Eucgr.D01485	587,5475972	-26,07245391	5,94747E-21	AT3G20380.1		TRAF-like family protein
Eucgr.D01489	198,116912	-24,90877981	1,70888E-19	AT3G20360.1		TRAF-like family protein
Eucgr.D01532	170,4407898	-24,58478934	6,14998E-19	AT4G03500.1		Ankyrin repeat family protein
Eucgr.D02256	153,3699046	-25,23563506	8,51326E-20	AT3G61220.1		NAD(P)-binding Rossmann-fold superfamily protein
Eucgr.E00558	529,1935262	-25,66104458	3,42514E-20			
Eucgr.E00680	60,64275573	-22,89371577	1,44281E-16	AT1G53350.1		Disease resistance protein (CC-NBS-LRR class) family
Eucgr.E00715	40,30139853	-22,18253604	2,43466E-16	AT5G35450.1		Disease resistance protein (CC-NBS-LRR class) family
Eucgr.E01015	21,60704463	-21,5494333	0	AT1G50180.1		NB-ARC domain-containing disease resistance protein
Eucgr.E01320	34,21584364	-22,70282685	5,10683E-17	AT5G51980.1 AT5G51980.2		Transducin/WD40 repeat-like superfamily protein
Eucgr.E01669	21,43409524	-22,01203115	7,12029E-16	AT5G51160.1		Ankyrin repeat family protein
Eucgr.E01756	61,65714311	-23,26904676	8,32024E-16	AT1G64830.1		Eukaryotic aspartyl protease family protein
Eucgr.E01758	94,58255917	-24,83506375	1,39192E-17	AT1G64830.1		Eukaryotic aspartyl protease family protein
Eucgr.E02061	7687,26558	-8,369065472	0,009860822			
Eucgr.E02315	15,26826918	-20,77003731	0	AT5G17680.1		disease resistance protein (TIR-NBS-LRR class), putative
Eucgr.E02341	62,31165518	-23,76804657	1,76541E-17			
Eucgr.E02595	281,81697	-25,77485966	1,72778E-20	AT1G17615.1		Disease resistance protein (TIR-NBS class)

Eucgr.E02803	93,31583847	22,5880167	5,70773E-16	AT1G27170.1		transmembrane receptors;ATP binding
Eucgr.E03012	71,33086524	21,73353312	0	AT4G11460.1	CRK30	cysteine-rich RLK (RECEPTOR-like protein kinase) 30
Eucgr.E03204	274,9154331	9,165281805	0,02259009	AT1G27170.1		transmembrane receptors;ATP binding
Eucgr.E03205	217,7232198	25,79432449	1,34901E-20	AT1G27180.1		putative disease resistance protein (TIR-NBS-LRR class)
Eucgr.E03382	13,12848274	21,95805612	0	AT3G44710.1		Plant protein of unknown function (DUF247)
Eucgr.E03385	86,3573255	24,71633617	1,9389E-18	AT2G36430.1		Plant protein of unknown function (DUF247)
Eucgr.E03391	20,77765714	22,61974669	1,65624E-16	AT2G36430.1		Plant protein of unknown function (DUF247)
Eucgr.E03393	227,941533	22,82143921	1,65624E-16	AT2G36430.1		Plant protein of unknown function (DUF247)
Eucgr.E03394	38,21821098	23,04042786	0	AT5G22560.1		Plant protein of unknown function (DUF247)
Eucgr.E03625	36,00680597	21,81536914	6,68928E-16	AT4G21380.1 AT4G27290.1	ARK3 RK3	S-locus lectin protein kinase family protein receptor kinase 3
Eucgr.E03682	127,9283212	24,17898406	6,47421E-20	AT1G33720.1	CYP76C6	cytochrome P450, family 76, subfamily C, polypeptide 6
Eucgr.E03761	83,84505948	23,03415632	1,9389E-18	AT4G05200.1	CRK25	cysteine-rich RLK (RECEPTOR-like protein kinase) 25
Eucgr.E03952	208,8841533	13,1791636	0,03412191	AT1G62420.1		Protein of unknown function (DUF506)
Eucgr.F00011	150,9394175	23,06041381	9,21294E-17	AT2G26560.1	PLA IIA,PLA2A,PLP2	phospholipase A 2A
Eucgr.F00731	24,510542	21,30016352	0	AT5G25910.1	AtRLP52,RLP52	receptor like protein 52
Eucgr.F00839	41,33825687	22,74739194	3,838E-18	AT4G27190.1		NB-ARC domain-containing disease resistance protein
Eucgr.F01313	13,39306533	22,1674501	5,44E-08	AT3G14840.2		Leucine-rich repeat transmembrane protein kinase
Eucgr.F01840	181,3606374	22,4748384	4,6848E-18	AT1G21270.1	WAK2	wall-associated kinase 2
Eucgr.F03727	81,55697261	24,18632723	5,61612E-20	AT1G75980.1		Single hybrid motif superfamily protein
Eucgr.F03823	60,21963037	23,29513105	2,6696E-18			
Eucgr.F04312	548,0925584	8,052170926	0,033408298	AT1G27170.1		transmembrane receptors;ATP binding
Eucgr.G00274	114,477714	24,35925874	3,75788E-20	AT1G60095.1		Mannose-binding lectin superfamily protein
Eucgr.G00291	29,95409063	22,06879787	2,58338E-16	AT2G37710.1	RLK	receptor lectin kinase
Eucgr.G00791	11,8180463	22,05632886	2,3963E-06	AT3G47210.1		Plant protein of unknown function (DUF247)
Eucgr.G00913	12,09808766	18,97280119	7E-10	AT4G10490.1		2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
Eucgr.G01270	452,6515272	8,265173207	0,021750519	AT3G47570.1		Leucine-rich repeat protein kinase family protein
Eucgr.G01438	72,10329606	23,12483994	1,9204E-18			
Eucgr.G02885	1365,180104	13,1310564	0,000187597	AT1G09070.1	(AT)SRC2, SRC2	soybean gene regulated by cold-2
Eucgr.G03007	12,3594963	22,08355783	0	AT3G55700.1		UDP-Glycosyltransferase superfamily protein
Eucgr.H00427	121,423594	16,70302085	8,2985E-06	AT3G26040.1		HXXXD-type acyl-transferase family protein
Eucgr.H00747	263,7919226	14,86696977	0,005680808	AT3G25240.1		Protein of unknown function (DUF506)
Eucgr.H01823	16,28024338	22,33248347	0	AT1G72890.2		Disease resistance protein (TIR-NBS class)

Eucgr.H01977	13,2758027	-19,03554645	3,564E-07	AT3G50440.1	ATMES10 MES10	methyl esterase 10
Eucgr.H02207	16,8729879	-22,09728152	0	AT5G17680.1		disease resistance protein (TIR-NBS-LRR class), putative
Eucgr.H02212	155,4638332	-25,05291978	3,25689E-19	AT3G44630.1		Disease resistance protein (TIR-NBS-LRR class) family
Eucgr.H02216	426,3120373	-9,716389473	0,041272072			
Eucgr.H02393	38,44213154	-22,52921236	9,21294E-17			
Eucgr.H02394	45,23044876	-22,75104186	3,07098E-16			
Eucgr.H02812	261,0456338	-25,75290237	1,53073E-20	AT3G06240.1		F-box family protein
Eucgr.H02813	20,73282973	-21,65941678	0			
Eucgr.H02814	23,01486101	-21,31296646	0	AT3G57470.1		Insulinase (Peptidase family M16) family protein
Eucgr.H02869	26,63424511	-20,32648271	0	AT4G13580.1		Disease resistance-responsive (dirigent-like protein) family protein
Eucgr.H03963	21,37182803	-22,3446385	3,45684E-16			
Eucgr.H04053	21,02021007	-18,97929291	3,2269E-06	AT3G01680.1		
Eucgr.H04573	70,90317816	-23,13169244	9,21294E-17	AT3G05480.1	ATRAD9,RAD9	cell cycle checkpoint control protein family
Eucgr.I00047	786,0176542	-10,96720032	0,00683257	AT5G36950.1	DegP10	DegP protease 10
Eucgr.I00644	68,56649045	-23,2060922	9,21294E-17	AT1G19715.1		Mannose-binding lectin superfamily protein
Eucgr.I01196	55,9532261	-19,13162883	1E-10	AT2G23610.1	ATMES3,MES3	methyl esterase 3
Eucgr.I02586	35,71628913	-19,95026238	0	AT5G64410.1	ATOPT4,OPT4	oligopeptide transporter 4
Eucgr.J00257	11,95066992	-22,0782481	2,3453E-06			
Eucgr.J00366	16,37500884	-22,41379821	3E-10	AT5G41040.1		HXXXD-type acyl-transferase family protein
Eucgr.J00405	325,6732009	-26,42906947	1,37585E-20	AT5G22860.1		Serine carboxypeptidase S28 family protein
Eucgr.J00545	21,6594584	-21,47030217	0	AT4G21380.1	ARK3,RK3	receptor kinase 3
Eucgr.J01316	21,96650791	-21,32567102	0	AT1G27170.2		transmembrane receptors;ATP binding
Eucgr.J01326	17,61932242	-21,95004513	0	AT4G37170.1		Pentatricopeptide repeat (PPR) superfamily protein
Eucgr.J01871	26,93038517	-22,85181005	7,20454E-16	AT2G30140.2		UDP-Glycosyltransferase superfamily protein
Eucgr.J02285	54,52637018	-22,73375104	6,59382E-16	AT1G53350.1		Disease resistance protein (CC-NBS-LRR class) family
Eucgr.J02366	26,32176856	-23,06726445	4,87569E-16	AT3G50940.1		P-loop containing nucleoside triphosphate hydrolases superfamily protein
Eucgr.J02754	17,1407135	-20,06787308	0	AT5G45470.1		Protein of unknown function (DUF594)
Eucgr.J03021	11,12961987	-13,80591961	0,001501488	AT1G52820.1		2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
Eucgr.K00502	19,19276696	-21,01484158	0	AT5G39030.1		Protein kinase superfamily protein
Eucgr.K00636	352,3170699	-24,7835043	6,81961E-19	AT1G45616.1	AtRLP6,RLP6	receptor like protein 6
Eucgr.K02406	33,21936718	-15,11870756	0,000297956	AT1G35710.1		Protein kinase family protein with leucine-rich repeat domain
Eucgr.K03098	17,40876869	-20,84437169	0	AT2G37710.1	RLK	receptor lectin kinase
Eucgr.L00019	45,4946663	-22,7368767	2,69176E-18			

Eucgr.L00021	98,71544115	-23,52137082	7,34233E-20	AT2G30720.1		Thioesterase/thiol ester dehydrase-isomerase superfamily protein
Eucgr.L00579	18,1447381	-20,89742315	0	AT4G38225.3		
Eucgr.L00729	58,5188557	-23,02292608	2,29639E-16	AT5G11290.1		Plant protein of unknown function (DUF247)
Eucgr.L00863	22,78767681	-21,13330517	0	AT1G45616.1	AtRLP6,RLP6	receptor like protein 6
Eucgr.L01042	15,70971524	-20,82215538	0	AT2G41040.1		S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
Eucgr.L01052	640,4694303	-9,035020234	0,004472146	AT5G43910.2		pfkB-like carbohydrate kinase family protein
Eucgr.L01669	1374,420766	-12,44346725	0,000171856	AT5G38280.1	PR5K	PR5-like receptor kinase
Eucgr.L01744	25,03611174	-22,30727497	2,66171E-16	AT2G23460.1	ATXLG1,XLG1	extra-large G-protein 1
Eucgr.L02186	33,75755154	-22,62727684	7,99885E-17	AT3G04880.1	DRT102	DNA-damage-repair/toleration protein (DRT102)
Eucgr.L02775	24,97055553	-22,03874203	5,28602E-16	AT5G62740.1	ATHIR1,HIR1	SPFH/Band 7/PHB domain-containing membrane-associated protein family
Eucgr.L03256	134,8773967	-24,27535705	3,838E-18	AT1G11410.1		S-locus lectin protein kinase family protein
Eucgr.L03342	42,34585786	-23,07015361	4,87569E-16	AT5G52960.1		
Eucgr.L03738	89,45251438	-23,17604833	1,01341E-16	AT4G12010.1		Disease resistance protein (TIR-NBS-LRR class) family

**Suplemento S1- BR BR102022025625-0**



870220117662



15/12/2022

09:56

29409161955575915

**Pedido nacional de Invenção, Modelo de Utilidade, Certificado de Adição de Invenção e entrada na fase nacional do PCT**

Número do Processo: BR 10 2022 025625 0

Dados do Depositante (71)

---

Depositante 1 de 2

Nome ou Razão Social: UNIVERSIDADE FEDERAL DE VIÇOSA

Tipo de Pessoa: Pessoa Jurídica

CPF/CNPJ: 25944455000196

Nacionalidade: Brasileira

Qualificação Jurídica: Instituição de Ensino e Pesquisa

Endereço: Campus UFV, Pró-Reitoria de Pesquisa e Pós Graduação, sala 04.

Cidade: Vicosá

Estado: MG

CEP: 36570-900

País: Brasil

Telefone: (31) 3612 2334

Fax:

Email: [propriedadeintelectual@ufv.br](mailto:propriedadeintelectual@ufv.br)

---

**PETICIONAMENTO  
ELETRÔNICO**

Esta solicitação foi enviada pelo sistema Petição Eletrônica em 15/12/2022 às 09:56, Petição 870220117662

Depositante 2 de 2

Nome ou Razão Social: BRACELL BAHIA FLORESTAL LTDA

Tipo de Pessoa: Pessoa Jurídica

CPF/CNPJ: 15692999000154

Nacionalidade: Brasileira

Qualificação Jurídica: Pessoa Jurídica

Endereço: Rua Dr. José Tiago Correia s/n, bairro Alagoinhas Velha

Cidade: Alagoinhas

Estado: BA

CEP: 48030-300

País: BRASIL

Telefone:

Fax: Emal:

Dados do Pedido

---

Natureza Patente: 10 - Patente de Invenção (PI)

Título da Invenção ou Modelo de Oligonucleotídeos para o processo de diagnóstico molecular do Utilidade (54): distúrbio fisiológico em Eucalipto (*Eucalyptus* sp.), processo de diagnóstico e usos

Resumo: A presente invenção refere-se a oligonucleotídeos sintéticos para o diagnóstico molecular do distúrbio fisiológico em eucaliptos (*Eucalyptus* sp.), ao processo de diagnóstico, ao uso dos oligonucleotídeos para identificação de genes diferencialmente expressos pelo distúrbio fisiológico em caules, e produção de kits diagnósticos, e ao uso dos genes diferencialmente expressos para a identificação e diagnóstico do distúrbio fisiológico em eucaliptos. Os oligonucleotídeos utilizados neste processo biotecnológico permitem diagnosticar a síndrome do distúrbio fisiológico em clones de eucalipto a campo em diferentes estádios de desenvolvimento. A invenção pode ser aplicada no setor agroflorestral, nas áreas de genética vegetal, genética molecular e melhoramento genético Figura a publicar: 1

---

**PETICIONAMENTO  
ELETRÔNICO**

Esta solicitação foi enviada pelo sistema Peticionamento Eletrônico em 15/12/2022 às 09:56, Petição 870220117662

Dados do Inventor (72)

---

Inventor 1 de 6

Nome: ELIZABETH PACHECO BATISTA FONTES

CPF: 33289654672

Nacionalidade: Brasileira

Qualificação Física: Professor do ensino superior

Endereço: Rua A, No 36, Recanto da Serra

Cidade: Viçosa

Estado: MG

CEP: 36576-220

País: BRASIL

Telefone: (31) 988 642830

Fax:

Email: bbfontes@ufv.br

Inventor 2 de 6

Nome: KELLEN CRISTINA GATTI

CPF: 04207956778

Nacionalidade: Brasileira

Qualificação Física: Pesquisador

Endereço: Rua 8 de dezembro, 120 – Bairro Alagoinhas Velha

Cidade: Alagoinhas

Estado: BA

CEP: 48030-260

País: BRASIL

Telefone: (75) 998 168999

Fax:

---

**PETICIONAMENTO**  
**ELETRÔNICO** Petição Knowlengo@bracell.com

Esta solicitação foi enviada pelo sistema Peticionamento Eletrônico em 15/12/2022 às 09:56, Petição 870220117662

Inventor 3 de 6

Nome: PEDRO AUGUSTO BRAGA DOS REIS

CPF: 06860110660

Nacionalidade: Brasileira

Qualificação Física: Professor do ensino superior

Endereço: Campus UFV

Cidade: Viçosa

Estado: MG

CEP: 36570-900

País: BRASIL

Telefone: (31) 988 131715

Fax:

Email: pedroreis@ufv.br

Inventor 4 de 6

Nome: LAURA GONÇALVES COSTA MARTINS

CPF: 09938879616

Nacionalidade: Brasileira

Qualificação Física: Pesquisador

Endereço: Rua Independência, 268, apto 202, Bairro Santa Luzia

Cidade: Sete Lagoas

Estado: MG

CEP: 35700-350

País: BRASIL

Telefone: (31) 984 226666

Fax:

Email: lurcostlin@gmail.com

---

**PETICIONAMENTO  
ELETRÔNICO**

Esta solicitação foi enviada pelo sistema Petição Eletrônica em 15/12/2022 às  
09:56, Petição 870220117662

Inventor 5 de 6

Nome: MÁRCIA HELENA SANTOS ESTEVES

CPF: 13845076631

Nacionalidade: Brasileira

Qualificação Física: Pesquisador

Endereço: Rua Camilo Rosa de Oliveira, 430, Silvestre

Cidade: Viçosa

Estado: MG

CEP: 36576-342

País: BRASIL

Telefone: (31) 971 350749

Fax:

Email: marcia.esteves@ufv.br

Inventor 6 de 6

Nome: VIRGÍLIO ADRIANO PEREIRA LORIATO

CPF: 09733306789

Nacionalidade: Brasileira

Qualificação Física: Pesquisador

Endereço: BR 120 sítio cristais

Cidade: Viçosa

Estado: MG

CEP: 36578-899

País: BRASIL

Telefone: (31) 999 921369

Fax:

Email: virgilio.loriato@gmail.com

**PETICIONAMENTO  
ELETRÔNICO**

Esta solicitação foi enviada pelo sistema Peticionamento Eletrônico em 15/12/2022 às 09:56, Petição 870220117662

## Documentos anexados

---

Tipo Anexo	Nome
Comprovante de pagamento de GRU 200	Comprovante de pagamento.pdf
Procuração	Procuração Andrea INPI + Portaria.pdf
Procuração	Procurac,a~o Bracell.pdf
Relatório Descritivo	Relatório descritivo.pdf
Reivindicação	Reivindicac,o~es.pdf
Desenho	Figuras.pdf
Resumo	Resumo.pdf

## Sequências Biológicas

---

Declaro que a informação contida na 'Listagem de Sequências' apresentada em formato eletrônico está limitada ao conteúdo da matéria revelada pelas sequências de aminoácidos e/ou de nucleotídeos divulgadas no pedido de patente, conforme depositado

Tipos de Sequências Biológicas	Nome
Listagem de Sequências Biológicas em formato TXT	Listagem de sequências biológicas.txt

## Acesso ao Patrimônio Genético

---

Declaração Negativa de Acesso - Declaro que o objeto do presente pedido de patente de invenção não foi obtido em decorrência de acesso à amostra de componente do Patrimônio Genético Brasileiro, o acesso foi realizado antes de 30 de junho de 2000, ou não se aplica.

## Declaração de veracidade

---

Declaro, sob as penas da lei, que todas as informações acima prestadas são completas e verdadeiras.

[bb.com.br] - Boleto gerado pelo sistema MPAG. 22/09/2022 09:50:59

**INSTRUÇÕES:**

*A data de vencimento não prevalece sobre o prazo legal. O pagamento deve ser efetuado antes do protocolo. Órgãos públicos que utilizam o sistema SIAFI devem utilizar o número da GRU no campo Número de Referência na emissão do pagamento. Serviço: 200-Pedido nacional de*

*Invenção, Modelo de Utilidade, Certificado de Adição de Invenção e entrada na fase nacional do PCT*

*Clique aqui e pague este boleto através do Auto Atendimento Pessoa Física.*

*Clique aqui e pague este boleto através do Auto Atendimento Pessoa Jurídica.*

BANCO DO BRASIL		001-9	00190.00009 02940.916196 55575.915172 6 91450000007000		
Nome do Pagador/CPF/CNPJ/Endereço					
UNIVERSIDADE FEDERAL DE VICOSA CPF/CNPJ: 25944455000196					
CAMPUS UFV PRO REITORIA DE PESQUISA E POS GRADUACAO SALA 04 , VICOSA -MG CEP:36570900					
Sacador/Avalista					
Nosso-Número	Nr. Documento	Data de Vencimento	Valor do Documento	(-) Valor Pago	
29409161955575915	29409161955575915	21/10/2022	70,00		
Nome do Beneficiário/CPF/CNPJ/Endereço					
INSTITUTO NACIONAL DA PROPRIEDADE INDUST CPF/CNPJ: 42.521.088/0001-37					
RUA MAYRINK VEIGA 9 24 ANDAR ED WHITE MARTINS , RIO DE JANEIRO - RJ CEP: 20090910					
Agência/Código do Beneficiário				Autenticação Mecânica	
2234-9 / 333028-1					

BANCO DO BRASIL		001-9	00190.00009 02940.916196 55575.915172 6 91450000007000		
Local de Pagamento					
<b>PAGÁVEL EM QUALQUER BANCO ATÉ O VENCIMENTO</b>					
Nome do Beneficiário/CPF/CNPJ					
INSTITUTO NACIONAL DA PROPRIEDADE INDUST CPF/CNPJ: 42.521.088/0001-37					
Data do Documento					
22/09/2022	Nr. Documento	Espécie DOC	Aceite	Data do Processamento	Nosso-Número
	29409161955575915	DS	N	22/09/2022	29409161955575915
Uso do Banco					
29409161955575915	Carteira	Espécie	Quantidade	xValor	(-) Valor do Documento
	17	R\$			70,00
Informações de Responsabilidade do Beneficiário					
A data de vencimento não prevalece sobre o prazo legal.					
O pagamento deve ser efetuado antes do protocolo.					
Órgãos públicos que utilizam o sistema SIAFI devem utilizar o número da GRU n					
o campo Número de Referência na emissão do pagamento.					
Serviço: 200-Pedido nacional de Invenção, Modelo de Utilidade, Certificado de					
Adição de Invenção e entrada na fase nacional do PCT					
(-) Desconto/Abatimento					
(+/-) Juros/Multa					
(-) Valor Cobrado					
Nome do Pagador/CPF/CNPJ/Endereço					
UNIVERSIDADE FEDERAL DE VICOSA CPF/CNPJ: 25944455000196					
CAMPUS UFV PRO REITORIA DE PESQUISA E POS GRADUACAO SALA 04 ,					
VICOSA-MG CEP:36570900					
Sacador/Avalista					
Código de Baixa					
Autenticação Mecânica - Ficha de Compensação					



\_\_ SIAFI2022-DOCUMENTO-CONSULTA-CONGRU (CONSULTA GUIA DE RECOLHIMENTO DA  
 UNIAO 13/10/22 09:44 USUARIO : MARILUCE DATA EMISSAO : 13Out22  
 TIPO : 1 - PAGAMENTO NUMERO : 2022GR800229 UG/GESTAO EMITENTE : 154051 / 15268 -  
 UNIVERSIDADE FEDERAL DE VICOSA UG/GESTAO FAVORECIDA : 183038 / 18801 -  
 INSTITUTO NACIONAL DA PROPRIEDADE INDU RECOLHEDOR : 154051 GESTAO :  
 15268 CODIGO RECOLHIMENTO : 72200 - 6 COMPETENCIA: VENCIMENTO:  
 05Out22 DOC. ORIGEM: 154051 / 15268 / 2022NP002146 PROCESSO : 918118/2022-28  
 RECURSO : 1  
 (=)VALOR DOCUMENTO : 70,00 (-)DESCONTO/ABATIMENTO:  
 (-)OUTRAS DEDUCOES :  
 (+)MORA/MULTA :  
 (+)JUROS/ENCARGOS :  
 (+)OUTROS ACRESCIMOS :  
 (=)VALOR TOTAL : 70,00 NOSSO NUMERO/NUMERO REFERENCIA :  
 00029409161955575915 CODIGO DE BARRAS : 89610000000 0 70000001010 3  
 95523127220 9 00360640000 4 OBSERVACAO  
 PAGAMENTO DE TAXA (DEPOSITO DE PEDIDO DE PATENTE), PI 02, NOSSO NUMERO 2940916  
 1955575915, PROCESSO 918118/2022-28.  
 LANCADO POR : 92283535620 - MARILUCE UG : 154051 13Out2022 09:31 PF1=AJUDA PF3=SAI  
 PF2=DADOS ORC/FIN PF4=ESPELHO PF12=RETORNA



MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE FEDERAL DE VIÇOSA  
REITORIA

Campus Universitário - Viçosa, MG - 36570-900 - Tel (31) 3612-1001 - E-mail: reitoria@ufv.br

## PROCURAÇÃO

**OUTORGANTE: UNIVERSIDADE FEDERAL DE VIÇOSA**, pessoa jurídica de direito público, criada pelo Poder Público Federal, na forma de fundação pública de cunho educacional, com sede na Avenida P. H. Rolfs, s/nº, *Campus Universitário*, Viçosa, Minas Gerais, inscrita no CNPJ/MF sob o nº 25.944.455/0001-96, neste ato representada por seu Reitor, Professor Demetrius David da Silva, brasileiro, casado, portador da Carteira de identidade de nº 606795-ES, inscrito no CPF sob o nº 542.934.726-49;

**OUTORGADA: ANDREA DE OLIVEIRA BARROS RIBON**, brasileira, casada, professora universitária, portadora da carteira de identidade nº MG-4.324.867, inscrita no CPF sob o nº 830.354.866-20, com domicílio legal na Universidade Federal de Viçosa, localizada a Avenida P. H. Rolfs, s/nº, *Campus Universitário*, Viçosa, Minas Gerais, CEP 36570-900;

**PODERES: A OUTORGANTE** confere à **OUTORGADA** os poderes contidos na cláusula "AD NEGOTIA", objetivando representar a Universidade Federal de Viçosa junto ao **INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL - INPI**, com poderes amplos e gerais, podendo a **OUTORGADA**: assinar formulários e termos de autorização para cópia, promover, perante autoridade competente, prova ou contraprova pertinentes à titularidade de direitos de propriedade industrial e, ainda, tudo o mais que se fizer necessário para o bom e fiel cumprimento do presente mandato, podendo substabelecer, no todo ou em parte, a presente procuração, que terá a validade de 3 (três) anos.

Viçosa, 03 de maio de 2022.

  
\_\_\_\_\_  
Demetrius David da Silva  
Reitor



Ministério da Educação  
Universidade Federal de Viçosa  
Campus Viçosa Reitoria

**PORTARIA Nº 0300/2022, DE 03/05/2022**

O Reitor da Universidade Federal de Viçosa, no uso de suas atribuições, conferidas pelo Decreto de 23/05/2019, publicado no Diário Oficial da União de 24/05/2019, considerando o que consta do Processo SEI 23114.906696/2022-11, resolve

1. revogar a Portaria nº 0162/2021, de 24/03/2021;
2. delegar competência à Presidente da Comissão Permanente de Propriedade Intelectual, e do Núcleo de Inovação Tecnológica da Universidade Federal de Viçosa, ANDRÉA DE OLIVEIRA BARROS RIBON, matrículas 8331-3/UFV e 2381125-3/Siape, para representar a instituição junto aos seguintes órgãos: Ins tutos Nacional de Propriedade Industrial – INPI; Registro Nacional de Cultivares – RNC, e ao Serviço Nacional de Proteção de Cul vares - SNPC, podendo assinar formulários e termos de autorização para cópia; promover, perante autoridade competente, prova ou contraprova concernentes à tularidade de direitos de propriedade intelectual; realizar pe cionamentos, eletrônicos ou não; e realizar outras ações pertinentes ao desempenho da presente representação.

Publique-se e cumpra-se

*Demetrius David da Silva*

*Reitor*



Documento assinado eletronicamente por **DEMETRIUS DAVID DA SILVA, Reitor**, em 03/05/2022, às 14:41, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).

Portaria 0300/2022, de 03.05.2022 (0695136) SEI 23114.906696/2022-11 / pg. 1



A autenticidade deste documento pode ser conferida no site

[http://sei.dti.ufv.br/sei/controlador\\_externo.php?](http://sei.dti.ufv.br/sei/controlador_externo.php?)

[acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](http://sei.dti.ufv.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **0695136** e o código CRC **96678CA5**.

**Referência:** Processo nº 23114.906696/2022-11 SEI nº 0695136

*Campus Viçosa*

Av. Peter Henry Rolfs, s/nº, *Campus Universitário*  
36570-900 Viçosa/MG

*Campus Florestal*

Rodovia LMG-818, km 6  
35690-000 Florestal/MG

*Campus Rio Paranaíba*

Rodovia MG-230, Km 7, Zona Rural, Rodoviário  
38810-000 Rio Paranaíba/MG



## PROCURAÇÃO POR INSTRUMENTO PARTICULAR

**OUTORGANTE: BRACELL BAHIA FLORESTAL LTDA.**, com sede na Rua Dr. José Tiago Correia s/n, Bairro: Alagoinhas Velha, Município de Alagoinhas/BA, CEP: 48.030-300, inscrita no CNPJ sob nº 15.692.999/0001-54, com seus atos constitutivos arquivados na Junta Comercial do Estado da Bahia sob nº NIRE nº 292101350216, em 09/07/93, neste ato representada, na forma do parágrafo 5º da Cláusula 18 do seu Contrato Social, por seu por seus DIRETORES: **1º Diretor: Sr. WILLIAM TERAN TROELSEN**, brasileiro, solteiro, Administrador de Empresas, portador da cédula de identidade nº 08216805-95 SSP/BA, inscrito no CPF/MF sob o nº 007.534.925-63 e **2º Diretor: Sr. RENAN BODRA MACHADO**, brasileiro, casado, Administrador de Empresa, portador da cédula de identidade nº 23.931.221-1 SSP/SP, inscrito no CPF/MF 168.578.548-40, ambos com endereço profissional na sede da Outorgante.

**OUTORGADOS: UNIVERSIDADE FEDERAL DE VIÇOSA - UFV**, com sede na Avenida Peter Henry Rolfs, s/n, Campus Universitário, Viçosa – MG, inscrita no CNPJ sob o nº 25.944.455/0001-96, neste ato representada pela **Prof.ª Andrea de Oliveira Barros Ribon**, Presidente do Núcleo de Inovação Tecnologia (NIT).

**PODERES:** Através do presente instrumento de mandato, o **OUTORGANTE** nomeia e constitui como seus procuradores os **OUTORGADOS** acima qualificados para representá-la perante o **Instituto Nacional da Propriedade Industrial – INPI**, para o fim de requerer e processar direitos de propriedade intelectual face ao pedido de patente intitulado **"Oligonucleotídeos para o processo de diagnóstico molecular do distúrbio fisiológico em Eucalipto (*Eucalyptus sp.*), processo de diagnóstico e usos"**, para mantê-lo em vigor com amplos poderes para assinar petições e documentos, pagar taxas, anotar transferências, fazer prova de uso da invenção patenteada, apresentar oposições, recursos, réplicas, anotar, elaborar notificações extrajudiciais, e praticar para os fins mencionados, todos os atos necessários perante as autoridades administrativas competentes no Brasil e no exterior, em benefício da Outorgante, ratificando os atos já praticados, estando vedado o substabelecimento.

**VIGÊNCIA:** O presente mandato vigorará pelo período de 01 (um) ano, a contar da data de sua assinatura.

Alagoinhas, Bahia, 28 de novembro de 2022.

**BRACELL BAHIA FLORESTAL LTDA.**

  
**WILLIAM TERAN TROELSEN**

  
**RENAN BODRA MACHADO**  
Diretor



Bracell Bahia Florestal  
Rua Dr. José Tiago Correia, s/n  
CEP 48030-480 Alagoinhas Velha  
Alagoinhas - BA, Brasil  
Telefone: +55 75 3423-9900

[www.bracell.com](http://www.bracell.com)

## **Oligonucleotídeos para o processo de diagnóstico molecular do distúrbio fisiológico em Eucalipto (*Eucalyptus sp.*), processo de diagnóstico e usos.**

### **Campo da Invenção**

[001] A presente invenção refere-se a oligonucleotídeos sintéticos para o diagnóstico molecular do distúrbio fisiológico em eucaliptos (*Eucalyptus sp.*), ao processo de diagnóstico, ao uso dos oligonucleotídeos para identificação de genes diferencialmente expressos pelo distúrbio fisiológico em caules, e produção de kits diagnósticos, e ao uso dos genes diferencialmente expressos para a identificação e diagnóstico do distúrbio fisiológico em eucaliptos.

[002] Tais oligonucleotídeos utilizados neste processo biotecnológico permitem diagnosticar a síndrome do distúrbio fisiológico em clones de eucalipto a campo em diferentes estádios de desenvolvimento. A invenção pode ser aplicada no setor agroflorestal, nas áreas de genética vegetal, genética molecular e melhoramento genético.

### **Estado da técnica**

[003] O termo “distúrbio fisiológico” foi criado para descrever um estado de comprometimento fisiológico em eucalipto associado a um conjunto de sintomas e cuja incidência tem sido recorrente nas regiões sul e leste do Estado da Bahia no Brasil. Uma vez que o distúrbio fisiológico descreve um conjunto de sintomas ocorrendo nos caules e evoluindo para as folhas, este estado de anomalia fisiológica também é conhecido como síndrome do distúrbio fisiológico. Não se sabe exatamente a causa do aparecimento do fenômeno de “distúrbio fisiológico”, mas ela tem sido atribuída a fatores climáticos (precipitações com alternância de seca prolongada) e fatores genéticos.

[004] O "distúrbio fisiológico" está diretamente relacionado à perda de dominância apical levando à necrose do tronco principal que se torna quebradiço. Paralelamente, ocorre a emergência de um tronco secundário e os sintomas estão associados com a perda de turgidez foliar, amarelecimento, seca de ponteiros, além de multiplicação de ramos na copa. Estes sintomas evoluem para multiplicação de ramos ao longo do tronco e formação de cascas necróticas, enrugadas e secas, ao longo do tronco em árvores adultas e representam perdas significativas para o agronegócio.

[005] O conjunto de sintomas que delimitam o estado de distúrbio fisiológico são bem caracterizados na fase adulta. Porém, na fase de clones jovens, estes sintomas são facilmente confundidos com respostas típicas de estresses bióticos e estresses abióticos, o que dificulta a seleção no campo de clones jovens com a menor possibilidade de serem acometidos pelo distúrbio

fisiológico. Consequentemente, somente é possível certificar-se de que os clones juvenis de eucalipto no campo foram acometidos pelo distúrbio fisiológico quando se transformam em clones adultos, com a qualidade da madeira já comprometida. [006] Não existe tecnologia molecular descrita previamente para diagnóstico da síndrome do distúrbio fisiológico no eucalipto. Atualmente, o diagnóstico do distúrbio fisiológico em eucalipto é baseado no desenvolvimento de um conjunto de sintomas fenotípicos que também podem ser resultados de outros estresses abióticos e fisiológicos. Consequentemente, embora o distúrbio fisiológico reúna um conjunto característico de sintomas, o diagnóstico do distúrbio é dificultado porque vários sintomas do distúrbio são compartilhados com sintomas de outros estresses bióticos e abióticos.

[007] Entre os sintomas característicos do distúrbio fisiológico, a murcha e amarelecimento das folhas são sintomas também induzidos por seca e patógenos, assim como a seca de ponteiros e multiplicação de ramos na copa podem ser induzidos por estresse nutricional e deficiência de boro. Da mesma forma, a necrose do tronco principal na parte inferior levando à perda de dominância apical também é um sintoma que pode ser induzido por patógenos.

[008] Assim sendo, tem tornado absolutamente necessário o desenvolvimento de diagnósticos moleculares que precisamente possam confirmar o estado inicial de distúrbio fisiológico em clones de eucalipto mais jovens para que possam ser eliminados em estádios iniciais de desenvolvimento antes de atingirem a fase adulta, com a qualidade da madeira já comprometida.

[009] Estudos anteriores em eucaliptos definiram biomarcadores para seleção precoce de clones de eucalipto tolerantes à seca de ponteiros associada ao déficit hídrico, uma desordem fisiológica que pode ser confundida com o distúrbio fisiológico em termos de fenótipos [(1) Condé, S.A. Biomarcadores fenotípicos para seleção precoce de clones de eucalipto tolerantes à seca de ponteiros associada ao déficit hídrico. 2019. 41 f. Dissertação (Mestrado em Fitotecnia) - Universidade Federal de Viçosa, Viçosa. 2019; (2) Corrêa, T.R., Picolia, E.A.T., Souza, G.A., Condé, S.A., Silva, N.M., Lopes-Mattos, K.L.B., Resende, M.D.V., Zauza, E.A.V., Oda, S. 2017. Phenotypic markers in early selection for tolerance to dieback in Eucalyptus. *Industrial Crops and Products*, 107, 130-138].

[010] Diante da necessidade do desenvolvimento de meios para o diagnóstico do distúrbio fisiológico em clones jovens de eucalipto a presente invenção fornece oligonucleotídeos sintéticos para a identificação dos genes diferencialmente expressos pelo distúrbio e para o processo de diagnóstico molecular.

[011] Os oligonucleotídeos e o processo de diagnóstico, objetos da presente invenção, se diferenciam desses trabalhos prévios, descritos acima, em pontos básicos e relevantes. Inicialmente, a seca de ponteiros é um distúrbio identificado em eucalipto, porém, de bases fisiológicas e moleculares diferentes da síndrome do distúrbio fisiológico. Mesmo que nos estádios tardios do distúrbio fisiológico os sintomas nas folhas possam se assemelhar àqueles descritos para a seca de ponteiros, o conjunto de anomalias descritas no distúrbio fisiológico é muito mais complexo com bifurcações nos troncos não observados na seca de ponteiros. Esta semelhança fenotípica entre os sintomas das folhas ocorridos em estádios tardios do distúrbio fisiológico e aqueles sintomas desenvolvidos pela seca de ponteiros constitui uma justificativa adicional para o desenvolvimento de marcadores moleculares capazes de distinguir estas duas anomalias em eucalipto.

[012] Em segundo lugar, os biomarcadores desenvolvidos para seca de ponteiros descritos nos estudos anteriores foram de natureza fenotípica (desenvolvimento de sintomas), enquanto os oligonucleotídeos da presente invenção apresentam marcadores moleculares baseados na expressão de genes diferencialmente expressos pelo estabelecimento do distúrbio fisiológico, sendo, portanto, de caráter molecular e, portanto, mais eficaz e preciso.

[013] A cultura do eucalipto se encontra distribuída por quase todo o território brasileiro, sendo que grande parte da área de cultivo apresenta limitações ao desenvolvimento das plantas. Desta forma, uma maior produtividade no setor florestal tem demandado pesquisas à procura de materiais genéticos mais produtivos e adaptados às diferentes condições ambientais. As plantas se encontram frequentemente sob condições desfavoráveis ao seu desenvolvimento podendo sofrer estresses reversíveis ou irreversíveis causados por fatores bióticos ou abióticos. Fatores ambientais como radiação solar, solo, disponibilidade hídrica e temperatura podem ocasionar estresses em plantas, e desta forma, as plantas desencadeiam respostas moleculares para lidar com estas situações.

[014] A falta de definição de procedimentos para a identificação precoce do distúrbio fisiológico é agravada devido aos mecanismos moleculares envolvidos no desenvolvimento dos sintomas e na resposta de defesa da planta ao distúrbio fisiológico serem totalmente desconhecidos.

[015] A presente invenção se diferencia de todos os processos de diagnóstico do distúrbio fisiológicos já descritos previamente por que se baseia no conhecimento adquirido das variações globais de expressão de genes identificados como repórteres da condição do distúrbio fisiológico. Esta estratégia possibilita, como vantagem primordial, o diagnóstico de clones jovens suscetíveis

ao distúrbio fisiológico prevenindo o seu crescimento para fase adulta, não comprometendo assim a qualidade da madeira.

**[016]** O termo variação global de expressão gênica induzida pelo distúrbio fisiológico significa o conjunto de genes diferencialmente expressos em caules sintomáticos de clones suscetíveis comparados com caules normais assintomáticos crescidos sob as mesmas condições ambientais, preferencialmente no mesmo talhão. Um estudo prévio de variação global de expressão gênica utilizou o RNA extraído de folhas de eucalipto de clones híbridos com o objetivo de identificar vias bioquímicas e metabólicas afetadas pelo distúrbio fisiológico, porém nenhum gene marcador foi identificado nesse estudo anterior (Rodrigues, A.C.P. Perfil de expressão gênica em híbridos de *Eucalyptus grandis* x *Eucalyptus urophylla* afetados pelo distúrbio fisiológico do eucalipto [DFE] Tese Doutorado - Universidade Federal dos Vales de Jequitinhonha e Mucuri. 2020. 142 p.).

**[017]** Este estudo prévio difere da invenção aqui descrita para diagnóstico molecular do distúrbio fisiológico em diversos pontos relevantes: (i) para a definição do processo de diagnóstico da presente invenção foi utilizado RNA extraído de caules e não de folhas, conforme descrito anteriormente; (ii) os sintomas iniciais do distúrbio fisiológico são estabelecidos no caule, o que favorece a identificação de genes associados ao estado precoce do distúrbio fisiológico, caracterizando a principal vantagem do processo de diagnóstico molecular descrito, qual seja, a capacidade de diagnosticar clones jovens no início do aparecimento do distúrbio fisiológico; (iii) os genes diferencialmente expressos identificados pelo RNA-seq das folhas (estudo prévio) foram na sua maioria genes que respondem a estresses abióticos gerais não podendo ser utilizados como marcadores específicos do distúrbio fisiológico, o que fundamenta a presente invenção; (iv) as amostras sintomáticas e assintomáticas, utilizadas para identificação do transcriptoma induzido pelo distúrbio fisiológico em folhas, foram colhidas de diferentes regiões e, conseqüentemente, sujeitas a variações climáticas diferenciais.

**[018]** Nestas condições experimentais, não se pode concluir que o transcriptoma induzido resultante nas folhas seja específico do distúrbio fisiológico e não de estresses abióticos gerais, principalmente porque os sintomas do distúrbio fisiológico nas folhas se assemelham àqueles desenvolvidos por outros estresses abióticos.

**[019]** Em contraste, na presente invenção, as amostras de caules sintomáticos e assintomáticos, utilizadas para se determinar o transcriptoma induzido pelo distúrbio fisiológico, se desenvolveram e foram coletadas no mesmo talhão, e, portanto, sujeitas às mesmas condições climáticas. Esta

estratégia experimental elimina a ocorrência de estresses múltiplos, induzidos pelas condições climáticas diferentes de crescimento dos clones e, portanto, assegura que os sintomas observados são resultados diretos e específicos da síndrome do distúrbio fisiológico.

[020] Os genes que apresentam maiores níveis de transcritos no caule sintomático em comparação com o caule assintomático são ditos regulados positivamente (*up* regulados) ou induzidos pelo distúrbio fisiológico. Estes genes induzidos pelo distúrbio compõem o transcriptoma induzido pelo distúrbio. Os genes que possuem menores níveis de transcritos no caule sintomático em comparação com o caule assintomático são ditos regulados negativamente (*down* regulados) ou reprimidos pelo distúrbio fisiológico. Estes genes reprimidos pelo distúrbio fisiológico compõem o transcriptoma reprimido pelo distúrbio.

[021] O denominador comum de todos os sintomas, durante a evolução do distúrbio, parece ser a ativação descontrolada de vias de sinalização de morte celular, levando à necrose ou senescência induzida por fatores climáticos. A morte celular programada (PCD) é uma via de sinalização em plantas ativada por vários estímulos endógenos (como sinais de desenvolvimento) e exógenos (como estresses), sendo um processo altamente dinâmico e regulado, capaz de perceber estresse intracelular dentro de minutos e rapidamente montar uma resposta apropriada para lidar com o dano celular.

[022] Isto somente é possível porque uma gama de vias de sinalização celulares não relacionadas convergem para ativação do processo de morte celular programada sinalizando para uma diversidade de condições. De uma forma geral, a PCD ou autofagia é requerida para permitir que organismos sésseis como as plantas se protejam contra estresses bióticos e abióticos. Acredita-se que várias vias de sinalização celular de adaptação das plantas ao meio ambiente são, de alguma forma, integradas com a sinalização de morte celular programada (Batoko, H., Dagdas, Y., Baluska, F., Sirko, A. 2017. Understanding and exploiting autophagy signaling in plants. *Essays in Biochemistry* 61, 675-685).

[023] Este conjunto de informações, associados aos sintomas típicos do “distúrbio fisiológico”, é um indicativo forte de que vias de sinalização de morte celular integradas com estresses que permitem às plantas a se adaptarem com o meio ambiente estão sendo inibidas ou ativadas irregularmente nos genótipos suscetíveis em resposta a fatores climáticos. Neste caso, os fenótipos observados são provavelmente resultantes de um somatório dos efeitos genéticos, ambientais e da interação genótipo x ambiente.

[024] Embora tenham sido identificados genótipos resistentes e suscetíveis, as causas da ocorrência do distúrbio fisiológico e os genes associados às respostas moleculares alteradas são totalmente desconhecidas. A elucidação desse fenômeno requer um conhecimento pleno das vias de sinalização de morte celular ativadas, sua integração com respostas a estresses abióticos, além dos genes e proteínas envolvidas na regulação dessas vias.

[025] Devido à abrangência prevista da resposta molecular, este conhecimento somente é possível por meio de estudos moleculares de amplitude genômica. Determinar o transcriptoma induzido pelo distúrbio fisiológico em genótipos suscetíveis e contrastá-lo com genótipos resistentes pode fornecer informações valiosas sobre os mecanismos moleculares que levam ao aparecimento da anomalia.

[026] Assim sendo, a presente invenção propõe um processo para o diagnóstico molecular do distúrbio fisiológico por meio de oligonucleotídeos sintéticos, no qual inicialmente foi determinada a variação global de expressão gênica em genótipos contrastantes para resistência e suscetibilidade ao distúrbio fisiológico.

[027] A invenção utiliza o transcriptoma induzido pelo distúrbio fisiológico, identificado pela análise de sequenciamento do RNA de alta performance de amostras de troncos de clones com sintomas do distúrbio fisiológico em comparação com as amostras de troncos assintomáticos do mesmo genótipo, crescidas sob as mesmas condições do meio ambiente e colhidas nas plantações de eucaliptos de 12 a 16 meses de idade.

[028] Esta variação global de expressão gênica representa um quadro de reprogramação gênica específica do estabelecimento do distúrbio fisiológico em troncos de clones de eucalipto suscetíveis e, portanto, genes específicos que variam a expressão em resposta ao distúrbio fisiológico foram selecionados como genes marcadores do distúrbio baseado na sua expressão em comparação com a expressão em amostras controles do mesmo genótipo, porém derivada de troncos não sintomáticos.

[029] A presente invenção refere-se a oligonucleotídeos sintéticos utilizados no processo de diagnóstico molecular, que se diferenciam de todos os demais descritos previamente porque se baseia no conhecimento adquirido das variações globais de expressão de genes identificados como repórteres da condição do distúrbio fisiológico em troncos de eucalipto.

[030] A invenção possibilita, como vantagem primordial, o diagnóstico de clones jovens suscetíveis ao distúrbio fisiológico. Uma vez certificado nestes clones jovens o estabelecimento do

estado de distúrbio fisiológico, estes poderão ser eliminados, prevenindo, portanto, o seu crescimento para fase adulta não comprometendo assim a qualidade do produto madeirável.

### Descrição das figuras

[031] As figuras 1 a 15 demonstram resultados de experimentos inéditos que fundamentaram o desenvolvimento da invenção. As figuras 16 a 20 demonstram a aplicação da invenção, objeto do atual pedido de patente. **Figura 1.** Representação esquemática das amostras de eucalipto. A planta 1 representa o genótipo suscetível com sintomas, a planta 2 o mesmo genótipo, porém (ainda) sem sintomas, ao passo que a planta 3, representa genótipo resistente.

**Figura 2.** Caule sintomático (SS) e emergente (ES) crescidos no campo do clone 520.

**Figura 3.** Gráficos de qualidade por base obtido através do FastQC. A) Pontuação Phred de qualidade, evidenciando boa qualidade das sequências geradas. B) Percentagens dos adaptadores em relação à posição nos fragmentos em bp. C) conteúdo da sequência por bases.

**Figura 4.** D) Gráfico de qualidade por base obtido através do FastQC, após remoção de adaptadores, (E) trimagem e limpeza das sequências de fragmentos.

**Figura 5.** Variação global na expressão gênica induzida pelo distúrbio fisiológico em eucalipto O eixo “x” indica o número de genes diferencialmente expressos. O eixo “y” indica contrastes entre os clones 520-S1 versus 520-S3, 520-S2 versus 520-S3 e 1404-R1 versus 520-S3. A análise foi realizada em troncos sintomáticos e assintomáticos do clone suscetível 520 e do clone resistente 1404.

**Figura 6.** Categorização funcional de genes diferencialmente expressos baseado em processo biológico, a partir do contraste 520-S1 versus 520S3. A figura ilustra a distribuição de genes diferencialmente expressos entre as categorias funcionais conforme definido em Ontologia de Genes processo biológico. Cada processo biológico é representado por uma cor e um percentual no gráfico. São demonstrados dois gráficos, um para genes diferencialmente expresso regulados positivamente (*up-regulated*) e outro para genes regulados negativamente (*down-regulated*). Apenas as categorias mais bem representadas são apresentadas ao lado da figura. **Figura 7. F-G)**

Repressão da síntese de ácidos graxos de cadeia longa em caules 520 sintomáticos no contraste 520-S1 versus 520-S3. Genes em vermelho são reprimidos e em verde induzidos, utilizando-se os critérios de  $\log_2 \text{fold change} = 1$  e p-valor de 0.05.

**Figura 8.** Genes envolvidos no reparo de DNA diferencialmente expressos em 520-S1 comparado com 520-S3. Genes em vermelho são reprimidos e em verde induzidos, utilizando-se os critérios de  $\log_2 \text{fold change} = 1$  e pvalor de 0.05.

**Figura 9.** Quantificação de transcritos extraídos de amostras de caules de eucalipto. H) Extração de RNA total dos clones 520, 1404, 2361, 1407, e das mudas correspondentes. Separação eletroforética de RNA extraído de caules e folhas foi realizada em gel de agarose 1% e visualizado por coloração em Brometo de Etídeo. As amostras são definidas na Figura 1. I) Análise de RT-PCR com base no gene endógeno *eEF2*. J) Eficiência dos oligonucleotídeos utilizados como controle endógeno para reação de RTqPCR.

**Figura 10.** Acúmulo de transcritos de genes envolvidos na síntese de ácidos graxos e reprimidos pelo distúrbio fisiológico em caules sintomáticos. O RNA total foi isolado de amostras SS (S1) sintomáticas do clone suscetível 520, amostras AS (S3, controle) assintomáticas do clone 520 e amostras ASR (R1), assintomáticas do clone resistente 1404, e a expressão dos genes indicados na figura foi determinada por RT-qPCR.

**Figura 11.** Acúmulo de transcritos de genes envolvidos no sistema de reparo do DNA e reprimidos pelo distúrbio fisiológico em caules sintomáticos. O RNA total foi isolado de amostras SS (S1) sintomáticas do clone suscetível 520, amostras AS (S3, controle) assintomáticas do clone 520 e amostras ASR (R1), assintomáticas do clone resistente 1404, e a expressão dos genes indicados na figura foi determinada por RT-qPCR. **Figura 12.** Acúmulo de transcritos de genes induzidos por estresses abióticos e pelo distúrbio fisiológico em caules sintomáticos. O RNA total foi isolado de amostras SS (S1) sintomáticas do clone suscetível 520, amostras AS (S3, controle) assintomáticas do clone 520 e amostras ASR (R1), assintomáticas do clone resistente 1404, e a expressão dos genes indicados na figura foi determinada por RT-qPCR.

**Figura 13.** Acúmulo de transcritos de genes envolvidos na síntese de fenilpropanóides e induzidos pelo distúrbio fisiológico em caules sintomáticos. O RNA total foi isolado de amostras SS (S1)

sintomáticas do clone suscetível 520, amostras AS (S3, controle) assintomáticas do clone 520 e amostras ASR (R1), assintomáticas do clone resistente 1404, e a expressão dos genes indicados na figura foi determinada por RT-qPCR.

**Figura 14.** Acúmulo de transcritos de genes diferencialmente expressos pelo distúrbio fisiológico em caules sintomáticos. O RNA total foi isolado de amostras SS (S1) sintomáticas do clone suscetível 520, amostras AS (S3, controle) assintomáticas do clone 520 e amostras ASR (R1), assintomáticas do clone resistente 1404, e a expressão dos genes indicados na figura foi determinada por RT-qPCR.

**Figura 15.** Curvas de dissociação dos oligonucleotídeos utilizados para a quantificação da expressão dos genes marcadores. Foi realizado o teste de eficiência dos oligonucleotídeos utilizando quantidades cDNA variando de 0,08ng a 50ng. A curva de dissociação foi calculada, sendo observado um único pico, que comprova a especificidade dos oligonucleotídeos (*primers*).

**Figura 16.** Extração de RNA total de amostras de troncos assintomáticos e sintomáticos dos clones indicados na figura. Separação eletroforética de RNA extraído de caules e folhas foi realizada em gel de agarose 1% e visualizado por coloração em Brometo de Etídeo.

**Figura 17.** Acúmulo de transcritos de genes Galactose oxidase (Eucgr.B03483) e RPT2 (Eucgr.B03325), induzidos pelo distúrbio fisiológico em caules sintomáticos dos clones indicados na figura. O RNA total foi isolado de amostras SS (sintomáticas de clones suscetíveis), amostras AS (assintomáticas, controle) e amostras assintomáticas, ASR, dos clones resistentes 1572, 1404 e 1407. A expressão dos genes indicados na figura foi determinada por RT-qPCR.

**Figura 18.** Acúmulo de transcritos de genes diferencialmente expressos pelo distúrbio fisiológico em caules sintomáticos dos clones indicados na figura. O RNA total foi isolado de amostras SS (sintomáticas de clones suscetíveis), amostras AS (assintomáticas, controle) e amostras assintomáticas, ASR, dos clones resistentes 1572, 1404 e 1407.

**Figura 19.** Acúmulo de transcritos de genes marcadores TIL1 e MSH2, reprimidos pelo distúrbio fisiológico em caules sintomáticos dos clones indicados na figura. O RNA total foi isolado de amostras SS (sintomáticas de clones suscetíveis), amostras AS (assintomáticas, controle) e amostras assintomáticas, ASR, dos clones resistentes 1572, 1404 e 1407.

**Figura 20.** Acúmulo de transcritos de genes marcadores KAS\_III e KCS1, reprimidos pelo distúrbio fisiológico em caules sintomáticos dos clones indicados na figura. O RNA total foi isolado de amostras SS (sintomáticas de clones suscetíveis), amostras AS (assintomáticas, controle) e amostras assintomáticas, ASR, dos clones resistentes 1572, 1404 e 1407.

### **Descrição resumida da invenção**

[032] A presente invenção refere-se aos oligonucleotídeos sintéticos utilizados para o diagnóstico molecular do distúrbio fisiológico em troncos de eucalipto crescidos no campo, baseado no nível de expressão de genes marcadores, selecionados por meio do transcriptoma induzido pelo distúrbio fisiológico em troncos sintomáticos de eucalipto, suscetível ao distúrbio fisiológico. Um aspecto importante da invenção é a determinação do transcriptoma induzido pelo distúrbio fisiológico em amostras de troncos sintomáticos. Outro aspecto importante da invenção é o uso tanto de genes marcadores regulados negativamente quanto genes marcadores regulados positivamente pelo distúrbio fisiológico. A presente invenção refere-se também ao processo de diagnóstico e ao uso dos oligonucleotídeos e dos genes diferencialmente expressos para a produção de kits de diagnóstico.

[033] O uso desses oligonucleotídeos no processo de diagnóstico molecular em amostra de caules de eucalipto permite distinguir o estado precoce de distúrbio fisiológico dos sintomas compartilhados pelos estresses abióticos e bióticos em amostras de clones jovens sintomáticos eucalipto e confirmar o diagnóstico fenotípico de distúrbio fisiológico em amostras de troncos sintomáticos de eucaliptos na transição do estágio juvenil para adulto e na fase adulta desenvolvidos nas áreas de plantio.

### **Descrição detalhada da Invenção**

[034] A presente invenção refere-se a oligonucleotídeos sintéticos para o diagnóstico do distúrbio fisiológico em eucalipto. Tais oligonucleotídeos são representados pelas sequências SEQ ID N° 1

a SEQ ID N° 32. A presente invenção refere-se ainda ao processo de diagnóstico e ao uso dos oligonucleotídeos e dos genes diferencialmente expressos para a produção de kits de diagnóstico do distúrbio fisiológico em eucaliptos. O processo de diagnóstico compreende as etapas a seguir:

- Utilização dos 32 oligonucleotídeos (SEQ ID N° 1 a SEQ ID N° 32) específicos dos dezesseis genes marcadores;
- Colheita de caules sintomáticos e assintomáticos de clones de eucalipto, três réplicas biológicas no mínimo;
  - Congelamento imediato em nitrogênio líquido;
  - Extração de RNA total dos caules sintomáticos e assintomáticos;
  - Síntese do DNA complementar com a enzima reverse transcriptase;
- Quantificação do nível de transcrito dos seis genes marcadores nos caules sintomáticos e assintomáticos por meio de RT-qPCR (RT-PCR em tempo real), utilizando três réplicas técnicas;
- Determinação do nível de transcrito dos seis genes marcadores nos clones sintomáticos e assintomáticos;
  - Análise do perfil comparativo de expressão dos genes marcadores.

[035] A presente invenção pode ser mais bem compreendida através da descrição a seguir, que apresenta formas de concretização não limitantes da mesma.

### **Experimento de demonstração**

[036] As amostras definidas para sequenciamento do RNA e, portanto, para identificação do transcriptoma induzido pelo distúrbio fisiológico foram definidas como amostras S1 (ou SS), S2 (ou ES) e S3 (ou AS) que indicam o caule primário necrosado, o caule secundário eminente de plantas sintomáticas e o caule primário, na mesma altura dos nós, de plantas assintomáticas, respetivamente (Figura 1). A amostra R1 (ASR) corresponde a caules de plantas resistentes. Estas amostras foram coletadas na mesma área de plantio.

[037] O transcriptoma induzido pelo distúrbio fisiológico em caules de eucalipto foi determinado utilizando amostras do clone suscetível 520, correspondentes a caules sintomáticos (S1 ou SS), e caules assintomáticos (S3 ou AS), obtidos de plantas crescidas no mesmo talhão, sob as mesmas condições ambientais. Nestas condições, uma comparação direta entre o transcriptoma de caules sintomáticos (amostra 520-S1) e do caule suscetível assintomático (amostras 520-S3) resultou na

identificação do transcriptoma induzido pelo distúrbio fisiológico, ou seja, na identificação do conjunto de genes que são diferencialmente expressos pelo estabelecimento da síndrome do distúrbio fisiológico.

### **Sequenciamento do RNA (RNA-seq) de nova geração**

[038] Foram utilizados dois clones de *Eucalyptus grandis*, 520, e 1404 como material contrastante para sequenciamento de RNA. O clone 520 é classificado como suscetível - identificado no campo com sintomas típicos de distúrbio fisiológico, enquanto o clone 1404 é resistente ao distúrbio fisiológico.

[039] As amostras coletadas em condições de campo foram definidas por regiões de amostragem no tronco, assim denominadas: SS (caule sintomático, S1) - tronco necrosado do clone 520, ES (caule emergente, S2) - tronco emergente, AS (caule assintomático, S3)-controle, oriundo de plantas suscetíveis sem sintomas. Quanto ao clone resistente 1404, foram coletadas as amostras ASR (caule do clone resistente, R1) similar às amostras AS, porém referente ao genótipo resistente 1404 (Figura 1). Os sintomas do distúrbio fisiológico em caules do clone 520 é mostrado na Figura 2.

[040] Foram coletadas três repetições biológicas de cada amostra, plantadas na mesma área de cultivo e sujeitas às mesmas condições climáticas, sob o estágio de desenvolvimento de 4 a 7 meses. Imediatamente após a coleta, ainda em campo, as amostras foram congeladas em nitrogênio líquido, e posteriormente acondicionadas em gelo seco.

[041] A extração de RNA total dos clones 520 e 1404 foi realizada conforme descrito. Foram utilizados 0,2 g de cada amostra armazenada em -80°C e então macerada em nitrogênio líquido com auxílio de almofariz e pistilo. Ao material macerado foi adicionado 1 mL do tampão de extração, contendo CTAB (brometo de cetiltrimetilamônio) 2% (p/v), PVP 2% (p/v), Tris-HCL 100 mM, EDTA 25 mM, NaCl 20 mM, e β-mercaptoetanol 2% (v/v). As amostras foram incubadas por 20 minutos a 65 °C. A esse material adicionou-se 1 mL de clorofórmio, seguida de centrifugação a 15.000 x g, por 10 minutos, a 4°C. Após uma segunda extração do sobrenadante com igual volume de clorofórmio, RNA total foi precipitado do sobrenadante com igual volume de isopropanol (aproximadamente 400 µL), por 10 minutos a temperatura ambiente e, posteriormente, por centrifugação por 10 minutos a 4 °C (12.000 x g). Após a centrifugação, o sobrenadante foi descartado, o sedimento foi lavado com 1 mL de etanol a 80% preparado com água DEPC (Dietilpirocarbonato), e os tubos foram conduzidos para outra centrifugação por 1

minuto a 4 ° C (12.000 x g). O líquido residual foi removido do tubo com uma pipeta e seco sobre papel absorvente, antes de ser ressuspense em 20 µL de água tratada com DEPC.

[042] A quantificação do RNA foi monitorada, pelo espectrofotômetro, pela relação de absorvância 260/280 e a integridade destes RNAs foram verificadas por eletroforese em gel de agarose desnaturante 1,5%, e também pelo equipamento *2100 Bioanalyser* (Agilent Technologies), observando-se a pureza e o tamanho das amostras. O RNA total de cada uma das amostras foi diluído para 5000 pg/µL para a visualização da sua integridade no instrumento *2100 Bioanalyser* (Agilent).

[043] Para isso, foi seguido o protocolo Agilent RNA 6000 pico Assay, de acordo com os procedimentos recomendados com relação à utilização do equipamento e à corrida das amostras, que foi efetuada usando um chip específico de RNA, juntamente com os kits de análise de RNA. Cada kit é composto por uma matriz de separação que é uma solução de polímero, uma solução de corante fluorescente de intercalação à fita dupla de DNA e marcadores de tamanho de RNA.

[044] De acordo com o protocolo de análise deste kit, adiciona-se o corante fluorescente à matriz de separação a qual é introduzida nos canais do chip sob pressão aplicada por uma seringa acoplada a um dispositivo próprio. Adiciona-se o marcador no poço adequado do chip. As amostras são adicionadas logo após, a cada um dos poços e o chip é então agitado em sistema próprio semelhante a um vortex, antes de ser inserido no instrumento “*Bioanalyzer*”.

[045] O resultado das separações foi registrado para avaliação através do software “2100 Expert” que contém um pacote de ferramentas para correções dos eletroferogramas com alinhamento dos picos, cálculo de tamanho e de concentração dos fragmentos.

[046] Como resultados, são fornecidos os valores de RIN (*RNA integrity number*) de cada uma das amostras, que estima a integridade do RNA, juntamente com o eletroferograma dos rRNAs 28S, 18S e 5S, cuja integridade é essencial para a construção da biblioteca de RNAseq. A escala do RIN varia de 0 a 10, considerando o RNA com um número menor que 6 de baixa qualidade.

[047] Amostras do clone 520 (suscetível) de eucalipto, de plantas em estado de acometimento, designadas caule necrosado (S1), e caule emergente (S2); de plantas assintomáticas (caules S3), além de amostras resistentes do clone 1404 designadas caule de clones resistentes (R1), foram utilizadas para extração do RNA total (Figura 1). A qualidade do RNA total foi avaliada, utilizando *Bioanalyzer 2100* da Agilent Technologies, antes de proceder ao envio das amostras para sequenciamento. A qualidade das amostras foi determinada mediante o método RIN (*RNA integrity*

*number*) que estima a integridade do RNA.

[048] As amostras que obtiveram um RIN próximo ou maior do que 7,0 seguiram para sequenciamento. Desta forma, as amostras de caule S1 (SS1), S2 (ES2), S3 (AS3) do clone 520 e R1 (ASR1) do clone 1404 prosseguiram, e foram enviados para o sequenciamento (Tabela 1). As amostras SS1.1, SS1.2, SS1.3; ES2.1, ES2.2, ES2.3; AS3.1, AS3.2, AS3.3 e ASR1.1, ASR1.2, ASR1.3. com três replicatas biológicas, pertencem aos clones 520 (suscetível) e 1404 (resistente) que foram enviadas para seu sequenciamento na empresa Macrogen Inc. (Korea).

[049] Para o sequenciamento do RNA, utilizou-se a tecnologia da plataforma *Illumina HiSeq™ 2500*, no modo *paired-end*. Cada *flowcell* contém várias faixas com cada faixa produzindo 250 milhões de leituras (ou seja, até 500 GB / *flowcell*) com o comprimento de cada sequência lida de 35 bp a 250 (3402Bioinformatics). O processo padrão de sequenciamento no HiSeq consiste em etapas: preparo das bibliotecas, geração de clusters e sequenciamento

<b>Tabela 1- Controle de qualidade do RNA</b>			
Clones	Amostras		Rin
Clone 520 SS1-1	Caule com sintomas		7.0
Clone 520 SS1-2	Caule com sintomas		8.1
Clone 520 SS1-3	Caule com sintomas		8.0
Clone 520 SE2-1	Caule emergente		8.2
Clone 520 SE2-2	Caule emergente		7.6
Clone 520 SE3-3	Caule emergente		7.3
Clone 520 AS3-1	Caule sem sintomas		8.6
Clone 520 AS3-2	Caule sem sintomas		7.6
Clone 520 AS3-3	Caule sem sintomas		7.5
Clone 1404 ASR1-1	Caule resistente		7.8
Clone 1404 ASR1-2	Caule resistente		7.0
Clone 1404 ASR1-3	Caule resistente		7.8

## **Avaliação da qualidade de leituras geradas pelo sequenciamento, com FastQC, TRIMAGEM e limpeza das sequências de Reads**

[050] O software de conversão “*Illumina bcl2fastq2*” foi utilizado para conversão dos dados gerados no sequenciamento para o formato de sequência *fastq*. Posteriormente, foi realizado o controle de qualidade das *reads*. As *reads* representam sequências curtas provenientes do sequenciamento (em número de milhões), para isso utilizou-se a ferramenta “*FastQC*”. Este software analisa parâmetros como a qualidade da sequência por nucleotídeo (Q20 e Q30), por conteúdo de GC, por conteúdo de nucleotídeos não determinados (N) e outros. *FastQ* usa o phred quality score para determinar os Q20 e Q30 score, a partir da probabilidade de que uma base na leitura tenha sido atribuída de modo errôneo durante o sequenciamento. E então, foi feita a remoção dos adaptadores Illumina, utilizados para ancoragem dos oligonucleotídeos (*primers*) de sequenciamento.

[051] A remoção dos adaptadores é de grande importância, devido ao fato que podem ser incorporadas na montagem dos genes ou no mapeamento do genoma de referência, contribuindo desta forma para uma montagem errônea das sequências. As sequências dos adaptadores foram eliminadas com o software “*Trimmomatic*”. Na trimagem dos dados, para remoção das leituras e sequências de baixa qualidade, foi utilizado o “*FASTX-Toolkit*”. A análise de qualidade das sequências foi realizada antes e após a limpeza das reads, pelo programa “*FastQC*”.

[052] De acordo com os resultados do sequenciamento, em relação à qualidade das leituras geradas, foram obtidos gráficos através do programa *FastQC*, por ferramentas de bioinformática, indicando os valores de qualidade de cada base no formato FastQ. Três parâmetros são mostrados, Q20, Q30 e conteúdo de GC% (Figura 3). Q20 indica a porcentagem das leituras com 99% ou mais de acurácia, ou seja, a probabilidade de que uma vez em 100 vezes uma base seja atribuída de forma errada nos picos do cromatograma durante o sequenciamento.

[053] O resultado de Q30 representa 99.9% ou mais de acurácia. Quando um sequenciamento alcança um *phred quality* Q30, praticamente todas as leituras serão perfeitas, com zero erros e ambiguidades. As cores de fundo representam a qualidade da amostra, sendo verde muito boa, laranja mediana e vermelha pobre em qualidade. Assim, o eixo Y representa o índice de qualidade, enquanto o eixo X indica a posição em read (bp) (Figure 3A).

[054] Os gráficos correspondentes às amostras SS, ES, AS (clone 520) e ASR (clone 1404), demonstram alta pontuação Phred de qualidade (índice de qualidade por base) (Figura 3A). Mesmo

assim, foi realizado a limpeza dos dados do sequenciamento, devido a presença de adaptadores nas sequências, de trechos que foram mal sequenciados e que apresentam baixa qualidade, ou seja, distorção na proporção de conteúdos de bases no início das sequências (Figura 3B e 3C). Para isso, foi efetuada a remoção destes adaptadores, trimagem e limpeza das sequências de reads e novos gráficos foram gerados de qualidade por base, tornando evidente a boa qualidade das amostras conforme Figura 4, o que indica um alto grau de confiança nas sequências geradas

### **Transcriptoma induzido pelo distúrbio fisiológico**

[055] O alinhamento das *reads* foi feito baseado no genoma de referência de *E. grandis* (versão 2.0). O software é específico para alinhamento de sequências curtas de grandes genomas, com maior eficiência e acurácia, de uma maneira rápida e consumindo pouca memória devido à forma como este indexa os dados. Sabendo-se que o número de *reads* mapeados para cada tipo de RNA é linearmente correlacionado com a sua abundância dentro da célula (número de *reads* é proporcional ao nível de expressão do gene), a próxima etapa a partir do arquivo resultante do alinhamento, foi reunir as *reads* numa unidade com algum significado biológico. A contagem das *reads* mapeadas por gene foi feita pela função *featureCounts* implementada no pacote “*Rsubread*” disponível no programa R/Bioconductor.

[056] Essas contagens precisam ser normalizadas para a remoção de viés técnicos decorrentes dos passos necessários para análise do RNA-seq, portanto, é preciso equacionar os dados para uma medida mais correta da taxa de expressão dos RNAs. Para a normalização dos *counts* foram utilizados pacotes disponíveis no Bioconductor.

[057] Para análise de expressão gênica diferencial (DGE) utilizou-se o pacote R/Bioconductor DESeq2, onde o valor “p” foi corrigido pela taxa de descoberta falsa. O valor de p significativo foi ajustado para 0,05. A análise de expressão diferencial utiliza o número de fragmentos (*reads*) reportados para cada gene em cada amostra. Diversos modelos estatísticos baseados no teste exato da binomial negativa e nos modelos lineares generalizados foram testados com base nos resultados de contagem dos fragmentos (*reads*), para então obter os genes diferencialmente expressos.

[058] Posteriormente, novas regiões do genoma onde ocorrem expressão foram identificadas por meio do programa Genome Browser e, portanto, novos genes foram estudados. Análises de expressão diferencial de exons foram estudadas usando o pacote DEXSeq (um pacote do Bioconductor) para a detecção de novas formas alternativas de ‘splicing’. Esses dados foram

anotados usando o banco de dados *Phytozome 12.1 (Eucalyptus grandis v2.0)* e armazenados em um banco de dados relacional, PostgreSQL 10, e uma interface de acesso foi implementada em PHP.

**[059]** Para o refinamento das análises, primeiramente, foi feita uma análise de enriquecimento gênico usando o teste exato de Fisher para categorias da ontologia gênica (GO), pelo pacote GOstats, do R / Bioconductor e das vias metabólicas presentes na base dados “*Plant Metabolic Network*”. Por meio dos ortólogos identificados, redes de interação proteína-proteína e de interações genéticas foram construídas. Usando a biblioteca de construção de grafos Boost Graph Library (BGL) em C++ foi construída *in silico* uma rede dinâmica onde diversas operações tais como centralidades, caminhos mais próximos e caminhos de máxima significância poderão ser usados para inferir novas vias de sinalização regulação.

**[060]** A amostra 520-S3, oriunda de plantas suscetíveis (clone 520) sem sintomas (caule suscetível assintomático – amostra S3 ou AS3), foi utilizada como controle de referência, e desta forma, permitiu examinar a variação global na expressão gênica dos genótipos contrastantes (caule necrosado, sintomático – amostra S1 ou SS, caule emergente, sintomático – amostra S2 ou ES e caule resistente – amostra R1 ou ASR) associada ao estabelecimento do distúrbio fisiológico (Figura 5). Usando um p-valor de 0,05 e  $\log_2\text{-fold-change} > 1,0$  para genes regulados positivamente (*up*) e valores de  $\log_2\text{ fold-change} < -1,0$  para genes regulados negativamente (*down*), foi observado que, o estado de distúrbio fisiológico (520-S1 vs 520S3) exibiu 4286 genes diferencialmente expressos (DE), dos quais 62,2 % (2666) foram regulados negativamente (*down*) nas amostras S1 em relação à S3, e 37,8% (1620) regulados positivamente (*up*), diferentemente do contraste 520-S2 versus 520-S3, que apresentou apenas 1131 genes expressos diferencialmente, sendo 58% (757) regulados positivamente (*up*) e 42% (554) regulados negativamente (*down*).

**[061]** Esta alta divergência entre os genes diferencialmente expressos no clone 520-S1 (tronco necrosado, sintomático) quando comparado ao clone 520-S2 (tronco emergente sintomático) evidencia que a amostra S2 representa um estágio de transição entre o estado de distúrbio fisiológico (responsável por uma expressão gênica massiva no caule necrosado) e o estado de recuperação (controle –S3) com a emergência de novo tronco.

**[062]** Uma alta percentagem de genes regulados negativamente presentes nas amostras 520-S1 quando contrastado às amostras controle, 520-S3 (plantas suscetíveis sem sintomas), indica que o distúrbio fisiológico induziu uma rede complexa de respostas adaptativas no caule do clone

suscetível, tal fato faz apologia ao fenótipo encontrado no campo. Uma vez que estas amostras são provenientes de clones de mesmo genótipo, em estágio de desenvolvimento similar, sob as mesmas condições ambientais de crescimento, os genes diferencialmente expressos nas amostras sintomáticas S1(SS) em relação aos da amostra assintomáticas S3 (AS) provavelmente representam o transcriptoma induzido pelo distúrbio fisiológico. Consequentemente, os genes diferencialmente expressos pelo contraste S1 *versus* S3 estão relacionados com os sintomas responsáveis pelo distúrbio.

**[063]** Por outro lado, a variação global da expressão gênica observada pelo contraste 1404-R1 *versus* 520-S3 deve refletir as mudanças no transcriptoma resultante do background genético dos genótipos, da resposta diferenciada dos referidos genótipos a condições ambientais de crescimento e da condição de resistência ao distúrbio fisiológico. Uma comparação das mudanças no transcriptoma dos clones 1404-R1 *versus* 520S3 identificou 51,36% (1979) de genes regulados positivamente (*up*) e 48,63 % (1874) de genes com regulação negativa (*down*), totalizando 3853 genes diferencialmente expressos no referido contraste.

**[064]** Estes genes diferencialmente expressos (DE) possivelmente compõem o transcriptoma induzido pelo estado de resistência ao distúrbio fisiológico e o transcriptoma diferenciado entre os dois genótipos (1404-genótipo resistente, 520-genótipo suscetível), ou seja, os genes DE devem representar um somatório entre efeito da resistência mais o efeito do genótipo.

**[065]** A princípio, o distúrbio fisiológico tem sido associado à morte celular induzida por fatores climáticos, uma vez que os caules se apresentam com sinais visíveis de necrose, quebradiços e as folhas, com clorose, amarelecimento e sinais de murcha. A ativação descontrolada de vias de sinalização de morte celular desencadeada por estresses múltiplos provoca uma série de alterações fisiológicas, metabólicas, homeostáticas e de desenvolvimento.

**[066]** Desta forma, para facilitar a interpretação dos resultados e análise dos genes associados às respostas moleculares alteradas, foram categorizados funcionalmente os genes diferencialmente expressos via Gene Ontology (GO), utilizando os critérios  $\log_2\text{-fold-change} > 1,0$  para genes regulados positivamente (*up*),  $\log_2\text{-fold-change} < -1,0$  para genes regulados negativamente (*down*) e p-valor de 0,05. Os resultados para a classe “processos biológicos”, representados de forma gráfica na Figura 6. **[067]** O transcriptoma induzido pelo distúrbio fisiológico (520-S1 *versus* 520-S3) deve representar a sintomatologia do fenômeno. Entre os genes com expressão diferencial positiva (*up*), as categorias biológicas mais enriquecidas de genes DE foram processo de oxidação-

redução (16.7%), transportadores-transmembrana (8%) e regulação de transcrição (7.4%) (Figura 6).

**[068]** Os genes envolvidos em processos de oxidação-redução fazem parte do sistema antioxidativo de defesa da planta, normalmente regulados positivamente em situações de estresses que produzem espécies reativas de oxigênio (ROS).

**[069]** Exemplos nesta categoria, regulados positivamente pelo estado de distúrbio fisiológico, incluem os genes da super família de peroxidases (Eucgr.A01166, Eucgr.A01166, Eucgr.I01176, Eucgr.L02460); superóxido dismutase (Eucgr.E00505, Eucgr.F01775, Eucgr.F01776, Eucgr.H04426, Eucgr.I01176, Eucgr.I02673); peroxidase (Eucgr.A01166, Eucgr.I01176, Eucgr.I02673, Eucgr.L02460); catalase (Eucgr.F01775, Eucgr.F01776), e Eucgr.H02433, Eucgr.H03936, Eucgr.I00009, Eucgr.I01803, Eucgr.J00610, Eucgr.J02207, Eucgr.K00212, Eucgr.K02121, Eucgr.K02125, Eucgr.K02858, Eucgr.L01277, Eucgr.L01765).

**[070]** Destacadamente, nesta categoria, foi a super representação de genes Cytochrome P450 monooxigenase entre os genes induzidos (*upregulated*) (Eucgr.A00523, Eucgr.A01644, Eucgr.A02188, Eucgr.B00199, Eucgr.F00151, Eucgr.F00153, Eucgr.F00173, Eucgr.F00225, Eucgr.F00492, Eucgr.F01386, Eucgr.F01387, Eucgr.F01492, Eucgr.F01493, Eucgr.F01548, Eucgr.F02072, Eucgr.F02396, Eucgr.F02404, Eucgr.H01953, Eucgr.H01954, Eucgr.H01958, Eucgr.H01963, Eucgr.H04841, Eucgr.H04878, Eucgr.H04879, Eucgr.I00335, Eucgr.I00744, Eucgr.I00749, Eucgr.I00783, Eucgr.I01215, Eucgr.I01217, Eucgr.J00084, Eucgr.J00095, Eucgr.J01402, Eucgr.J02384, Eucgr.K02083, Eucgr.K02211, Eucgr.K02212, Eucgr.K02924, Eucgr.L01403, Eucgr.L02217, Eucgr.L02324, Eucgr.L03497), estes genes CYP representam aproximadamente 12,23% da superfamília de CYP em eucalipto e têm sido descritos como componentes relevantes do sistema de detoxicação de todos os organismos, sendo que diversos membros são regulados positivamente em diversas condições de estresses fisiológicos.

**[071]** Estes resultados são consistentes com um estado de estresse associado ao distúrbio fisiológico. CYP450 também está envolvido em diversos processos metabólicos, relacionados à síntese de esteróides e detoxificação de compostos tóxicos, consistente com a regulação positiva das vias de biossíntese de brassinosteróides, diterpenóides e flavonóides, que se observa especificamente em caules suscetíveis sintomáticos

(Tabelas 2, 3 e 4).

<b>Tabela 2. Genes da via de biossíntese de flavonóides (regulados positivamente)</b>						
<b>TAIR10</b>	<b>ID</b>	<b>Normalização</b>	<b>log2foldchange</b>	<b>p-valor</b>	<b>EC</b>	<b>Descrição</b>
AT2G40890	Eucgr.A02188	1018.30	1.73	0.01	-	CYP98A3
AT5G07990	Eucgr.I00335	266.76	2.66	0.00	1.14.13.21	TT7
AT5G07990	Eucgr.K02211	162.37	3.36	0.00	-	TT7
AT5G13930	Eucgr.D01635	927.34	4.65	0.00	2.3.1.74	TT4

<b>Tabela 3. Genes da via de biossíntese de diterpenóides</b>						
<b>TAIR10</b>	<b>ID</b>	<b>Normalização</b>	<b>log2foldchange</b>	<b>pvalor</b>	<b>EC</b>	<b>Descrição</b>
AT1G02400	Eucgr.K03561	71.86	2.37	0.00	1.14.11.13	GA2OX6
AT2G32440	Eucgr.K02924	110.01	2.30	0.00	-	CYP88A3
AT2G32440	Eucgr.D02168	174.45	3.01	0.00	-	CYP88A3
AT2G32440	Eucgr.D02171	52.91	3.17	0.00	-	CYP88A3
AT5G25900	Eucgr.F02396	47.78	1.34	0.03	1.14.13.78	GA3

[072] Embora não categorizado em função biológica, um conjunto de genes de morte celular foi regulado positivamente. Neste caso, a via de morte celular induzida por múltiplos estresses e mediada pelas proteínas NRPs foi especificamente induzida nos troncos sintomáticos e necróticos com um aumento considerável na expressão do componente a jusante, a enzima VPE (“*vacuolar processing enzyme*”), que age como um efetor de um tipo de morte celular programada específica de plantas mediada pelo colapso do vacúolo.

[073] Também relevante, foi a indução robusta do homólogo do fator de transcrição ANAC029 (AtNAP), mediador de morte celular programada induzida por senescência foliar natural e em resposta a múltiplos estresses, podendo estar associado ao fenótipo de morte celular dos troncos

sintomático. Adicionalmente, como gene alvo controlado por ANAC029, destaca-se SAG113, uma phosphatase 2C, que regula o movimento de estômatos e a perda de água nas folhas, um fenótipo observado em folhas de plantas suscetíveis apresentando os sintomas de distúrbio fisiológico.

[074] Estes resultados indicam que o estabelecimento do estado de distúrbio fisiológico induz o fenômeno de morte celular, culminando em necrose no caule apical e, provavelmente, levando à murcha das folhas, em plantas acometidas pelo distúrbio fisiológico.

<b>Tabela 4. Genes da via de biossíntese de brassinosteróide (regulados positivamente)</b>						
<b>TAIR10</b>	<b>ID</b>	<b>Normalização</b>	<b>log2foldchange</b>	<b>pvalor</b>	<b>EC</b>	<b>Descrição</b>
AT2G26710	Eucgr.C04102	53.46	2.08	0.00	1.14.13.173	BAS1
AT2G26710	Eucgr.C04098	25.06	1.92	0.00	-	BAS1
AT2G26710	Eucgr.L02324	102.88	2.12	0.04	1.14.13.173	BAS1
AT2G26710	Eucgr.C04098	25.06	1.92	0.00	1.14.13.173	BAS1
AT3G30180	Eucgr.J01402	6.07	2.68	0.03	-	BR6OX2
AT4G36380	Eucgr.J02384	283.25	1.40	0.00	1.14.13.112	CYP90D1

[075] Considerando os genes reprimidos (*down-regulated*) nos caules sintomáticos (520-S1), predominaram as seguintes categorias biológicas, metabolismo de carboidratos (15.5%), biogêneses e organização de componentes celulares (12.5%), componentes de movimento de microtúbulos (9.2%), organização de organelas (7.1%), metabolismo de polissacarídeos (4.0%), Replicação do DNA e reparo (4.7%) (Figura 6). [076] Entre os genes regulados negativamente (*down*), na categoria biogênese, destaca-se uma repressão massiva de genes envolvidos na síntese de ácidos graxos, incluindo ácidos graxos de cadeia longa e ceras (Figura 7, Tabela 5). Estes genes amplamente reprimidos incluem *FATB* que codifica ACIL-ACP THIOESTERASES B de Ácidos graxos (*FATTY ACYL-ACP THIOESTERASES B*), *ACC1*, que codifica ACETIL-COA CARBOXILASE (*ACETYL-COA CARBOXYLASE*), *LACS2*, DESENVOLVIMENTO DE RAÍZES LATERAIS 2 (*LATERAL ROOT DEVELOPMENT 2*), também designado ACIL-COA

SINTETASE 2 DE CADEIA LONGA (*LONG-CHAIN ACYL-COA SYNTHETASE 2*), *FAB1*, BIOSÍNTESE 1 DE ÁCIDOS GRAXOS (*FATTY ACID BIOSYNTHESIS 1*), *MOD1*, MORTE MOSAICO 1 (*MOSAIC DEATH 1*), também designado ENOIL-ACP REDUTASE (*ENOYL-ACP REDUCTASE 1*), *EMB3147*, EMBRIO DEFECTIVO 3147 (*EMBRYO DEFECTIVE 3147*), também designado, MALONIL COA-ACP MOLONILTRANSFERASE (*MALONYL COA-ACP MALONYLTRANSFERASE*), *FATA1* que codifica ACIL-ACP TIOESTERASE 1 (*FATA ACYL-ACP THIOESTERASE 1*), e *FTM1*, TRANSIÇÃO FLORAL NO MERISTEMA 1 (*FLORAL TRANSITION AT THE MERISTEM 1*).

[077] Estes genes, essenciais para síntese de ácidos graxos de cadeia longa, estão envolvidos em diversas etapas de desenvolvimento e crescimento de plantas, como síntese de ceras e cutina nos caules e folhas, formação da raiz lateral, controle de divisão celular na região de meristemas.

[078] Reduções ou bloqueios na síntese de ácidos graxos saturados têm sido associados a declínios na taxa de crescimento de plantas, na síntese de ceras e cutina, um dos principais componentes da cutícula da planta e indução de morte celular. Em plantas silenciadas para o gene *FATB*, o carregamento total de ceras na folha é reduzido 20% e, nos caules, a redução chega a 50%, o que implica *FATB* no suprimento de ácidos graxos saturados para biossíntese de ceras.

**Tabela 5.** Repressão de genes envolvidos na síntese de ácidos graxos no contraste 520-S1 versus 520-S3

ID	Normaliza ção	log2foldcha nge	p-valor	Descrição	TAIR10
Eucgr.C010 82	377.56	-2.30	0.00	FATB	AT1G08510
Eucgr.C041 39	2356.73	-1.20	0.00	FATB	AT1G08510
Eucgr.H046 66	130.54	-3.04	0.00	AT1G24360	AT1G24360
Eucgr.F004 84	3036.56	-1.25	0.02	ACC1	AT1G36160
Eucgr.G004 90	665.95	-3.14	0.00	LACS2	AT1G49430
Eucgr.C007 78	194.52	-1.39	0.02	FAB1	AT1G74960
Eucgr.A024 48	453.80	-1.04	0.01	MOD1	AT2G05990
Eucgr.J018 59	363.50	-1.03	0.02	EMB3147	AT2G30200
Eucgr.H002 96	5.89	-5.89	0.02	FTM1	AT2G43710
Eucgr.H007 74	952.19	-2.68	0.00	FaTA	AT3G25110

[079] Mutantes no gene *ACC1* exibem divisões celulares descoordenadas, uma organização aberrante da região apical do embrião e desenvolvimento anormal de caules e raízes, sendo deficientes à tolerância ao congelamento após aclimação ao frio.

[080] O gene *LACS2* catalisa a síntese de omega-hidroxi graxo acil-CoA, intermediário na via de síntese de cutina, já o *MOD1* codifica enoil-ACP redutase, um componente do complexo sintase de ácidos graxos.

[081] Redução na expressão de *MOD1* causa morte celular prematura, um fenótipo característico de caules 520-S1. *MOD1* possuem nível de lipídeo reduzido e defeitos morfológicos pleiotrópicos incluindo folhas cloróticas e com formato anormal. Embora todas as mutações nos genes da biossíntese de ácidos graxos tenham sido geradas e analisadas na espécie vegetal *Arabidopsis*

*thaliana*, o fenótipo de plantas de eucalipto exibindo distúrbio fisiológico pode ser amplamente explicado pela redução na expressão de genes essenciais na biossíntese de ácidos graxos de cadeia longa.

**[082]** Na categoria Replicação do DNA e Reparo, predomina a regulação negativa de genes relacionados ao sistema de reparo do DNA, indicando que falhas no sistema podem ter sido o estímulo para ativação de vias de morte celular nos caules sintomáticos (520-S1) (Figura 8, Tabela 6). As plantas são particularmente vulneráveis a fatores que danificam o DNA. Desta forma, elas evoluíram uma rede complexa de mecanismos para detecção de danos no DNA e reparos, dedicados a assegurar a estabilidade genômica por meio de remoção das lesões no DNA e reconstituição da informação genética original.

**[083]** Se o sistema de reparo de DNA é comprometido, então a instabilidade genômica gera um sinal para ativar vias de morte celular. O aparato de reparo celular, conta com a presença de proteínas que após reconhecimento do sítio danificado, desencadeiam a interrupção do ciclo celular, enquanto enzimas relacionadas diretamente com a correção do dano são recrutadas. Este cenário é consistente com as variações de expressão gênica observadas nos caules sintomáticos pelo estabelecimento do estado de distúrbio fisiológico.

**[084]** Coletivamente, estes resultados sugerem que o distúrbio fisiológico no clone 520, genótipo suscetível, está associado à alteração negativa no sistema de reparo do DNA, o que provavelmente ativa eventos de morte celular programada no clone 520, desencadeando processos que envolvem principalmente resposta ao estresse abiótico, síntese de alcalóide, flavonoides e em contrapartida, diminuição de processos relacionados com desenvolvimento, replicação do DNA e metabolismo.

**[085]** Condições severas de estresses abióticos rapidamente induzem um conjunto de genes associados com morte celular programada (PCD) em plantas, como uma última tentativa de sobrevivência das células.

**Tabela 6. Genes envolvidos no reparo de DNA down-regulados em 520-S1 comparado com S3**

ID	Normalização	log2foldchange	p-valor	Descrição	TAIR10
<b>03430 Reparo de pareamento incorreto</b>					
Eucgr.B03882	81.33	-2.9	0	TIL1	AT1G08260
Eucgr.B03112	178.49	-6.07	0	AT1G13635	AT1G13635
Eucgr.J01377	14.4	-2.06	0.04	ARP	AT2G41460
Eucgr.K01848	49.67	-1.07	0.02	AT3G12040	AT3G12040
Eucgr.J00843	176.61	-1.42	0.05	UNG	AT3G18630
Eucgr.I00173	97.14	-1.01	0.01	POLD4	AT5G63960
<b>03420 Reparo por excisão de nucleotídeos</b>					
Eucgr.B03882	81.33	-2.9	0	TIL1	AT1G08260
Eucgr.C02095	34.23	-2.25	0	AT1G03810	AT2G24490
Eucgr.H01039	44.9	-1.03	0.01	CYCH;1	AT5G27620
Eucgr.I00173	97.14	-1.01	0.01	POLD4	AT5G63960
<b>03410 Reparo por excisão de bases</b>					
Eucgr.B03882	81.33	-2.9	0	TIL1	AT1G08260
Eucgr.B03112	178.49	-6.07	0	AT1G13635	AT1G13635
Eucgr.J01377	14.4	-2.06	0.04	ARP	AT2G41460
Eucgr.K01848	49.67	-1.07	0.02	AT3G12040	AT3G12040
Eucgr.J00843	176.61	-1.42	0.05	UNG	AT3G18630
Eucgr.I00173	97.14	-1.01	0.01	POLD4	AT5G63960

[086] Em condições de estresse prolongado, o estado de PCD pode evoluir para necrose, caracterizado pelos sintomas observados em caules acometidos pelo distúrbio fisiológico. Consistente com estas observações, genes envolvidos na via de sinalização de morte celular foram altamente induzidos nos troncos sintomáticos e necróticos (S1) em relação ao controle assintomático (S3). As principais proteínas expressas em S1, envolvidas com a morte da célula incluem as proteínas NRP (*N-rich proteins*), responsável por uma regulação positiva de NAC81, um efetor da resposta de morte celular, na cultura da soja, genes NAC, relacionados a processos de desenvolvimento, sinalização auxina, defesa, respostas ao estresse abiótico e senescência foliar, que são consideradas efetores centrais de apoptose.

[087] Além desses, também se destaca em S1 a enzima de processamento vacuolar VPE que exibem atividade de caspase-1-like. Deste modo, os resultados indicam que o denominador comum de todos os sintomas do distúrbio, induzido por elementos climáticos, parece ser a ativação descontrolada de vias de sinalização de morte celular, levando à necrose ou senescência.

### Validação dos resultados do RNA-seq

[088] Os resultados de RNA-Seq foram validados por RT-PCR quantitativo em tempo real (RT-qPCR) das amostras 520-S1, 520-S3, 1404-R1 confirmando a expressão diferencial de um subconjunto de genes responsivos ao estresse da síndrome do distúrbio fisiológico. Foram selecionados três genes da via da síntese de ácidos graxos, três genes envolvidos no sistema de reparo do DNA, três genes da síntese de fenilpropanóide, além de alguns genes diferencialmente expressos em S1 ou R1, escolhidos aleatoriamente (Tabela 7). Como controle endógeno foram testados os genes *eEF2* (*Eukaryotic translation elongation factor 2*- fator de alongação 2 de tradução de eucarioto), *ACT* (*Actin-related protein 3*- proteína 3 relacionada com actina) e *30S* (*30S ribosomal protein S5, chloroplastic* – proteína ribossomal S5 cloroplastídica).

Tabela 7. Genes diferencialmente expressos avaliados por qRT-PCR						
Locus	Gene	Descrição	$Log_2$ <i>ldCh nge</i>	p-valor <i>a</i>	Contrast <i>e</i>	Processo biológico
Eucgr.E00 458	KAS_III	3-Oxoacyl- [acylcarrier-protein (ACP)] synthase III	0.78 -0.61	0.014302789 0.037650496	R1-S1 S1- S3	Biossíntese de ácidos graxos
Eucgr.F00 484	ACC1	acetyl-CoA carboxylase 1, synthesis of malonyl-CoA	-1.25	0.01830475	S1-S3	Biossíntese de ácidos graxos
Eucgr.H00 232	KCS1	3-Oxoacyl- [acylcarrier-protein (ACP)] synthase III C Term	4.23 -3.68	1.52E-05 3.80E-05	R1-S1 S1- S3	Biossíntese de ácidos graxos

Eucgr.K00 927	ADH	GroES-like zincbinding dehydrogenase family protein	1.88 -1.48	0.000750402 0.003124045	R1-S1 S1- S3	Degradação de ácidos graxos
Eucgr.K00 013	PGL1	NagB/RpiA/CoA transferase-like superfamily protein	2.12 1.89	6.95E-05 0.000688734	R1-S3 R1- S1	Via da pentose fosfato
Eucgr.D01 028	RSW10	NagB/RpiA/CoA transferase-like superfamily protein	-1.1 0.98	0.006591108 0.006342078	R1-S1 S1- S3	
Eucgr.B03 112	Glycosyla se	DNA glycosylase superfamily protein	7.56 -6.07	1.56E-09 1.87E-07	R1-S1 S1- S3	Reparo por excisão de bases
Eucgr.B03 882	TIL1	POL2A,TIL1; DNA polymerase epsilon catalytic subunit	3.42 -2.9 -1.6	2.25E-13 5.27E-12 0.000655981	R1-S1 S1-S3 S2-S3	Reparo por excisão de nucleotídeo ,
Eucgr.B00 700	MSH2	MSH2; DNA mismatch repair protein MSH2	0.69 2.32 -1.63 -1.28	1.25E-18 1.62E-13 6.38E-12 9.06E-07	R1-S3 R1-S1 S1-S3 S2-S3	Reparo de pareamento incorreto
Eucgr.D00 242	BARD1	ATBRCA1,BRCA1; breast cancer susceptibility1	3.56 -3.63 -1.64	3.29E-07 4.45E-09 0.022923577	R1-S1 S1-S3 S2-S3	Recombinaç ão homóloga
Eucgr.B03 483	Galactose oxidase	Galactose oxidase/kelch repeat superfamily protein	-5.61 5.98 3.46	0.036537475 6.3766993e- 09 0.006836687	R1-S1 S1-S3 S2-S3	
Eucgr.B03 325	RPT2	Root phototropism protein 2	-10.91 9.57 4.98	1.634768e-05 2.836317e-45 4.374303e-11	R1-S1 S1-S3 S2-S3	Resposta a estímulo abiótico
Eucgr.J022 88	GAPDH	glyceraldehyde 3phosphate dehydrogenase	2.63 -2.260	0.000123073 0.00027011	R1-S1 S1-S3	Fixação de carbono
Eucgr.K00 756	Ubiquitin	small ubiquitin-like modifier 2	-1.00 -0.93	1.368476e-07 2.114482e-06	R1-S1 R1-S3	Proteolises mediada por ubiquitina

[089] Os oligonucleotídeos (*primers*) utilizados para detectar os níveis de transcrições com RT-PCR em tempo real foram projetados usando o Primer Express 3.0 (Life Technologies) com base nos genes selecionados. A seleção dos 17 genes de interesse foi realizada a partir dos dados de contrastes entre clone 1404 R1 versus clone 520 S3, clone 520 S1 versus clone 520 S3, com base nos valores de *Log2 Fold Change* e *p* valor.

[090] Para quantificação do nível de transcrito, por RT-PCR quantitativo, do gene Eucgr.E00458, *KAS\_III (3-Oxoacyl-[acyl-carrier-protein (ACP)] synthase III)* foram utilizados os oligonucleotídeos iniciadores (SEQ ID No 1) e (SEQ ID No 2), do gene Eucgr.F00484, *ACCI (acetyl-CoA carboxylase 1, synthesis of malonyl-CoA)*, os oligonucleotídeos (SEQ ID No 3) e (SEQ ID No 4), do gene Eucgr.H00232, *KCSI (3-Oxoacyl-[acyl-carrier-protein (ACP)] synthase III C Term)*, os oligonucleotídeos (SEQ ID No 5) e (SEQ ID No 6), do gene Eucgr.K00927, *ADH (GroES-like zinc-binding dehydrogenase family protein)*, foram usados os oligonucleotídeos (SEQ ID No 7) e (SEQ ID No 8), do gene Eucgr.K00013, *PGL1 (NagB/RpiA/CoA transferase-like superfamily protein)*, os oligonucleotídeos (SEQ ID No 9) e (SEQ ID No 10), do gene Eucgr.D01028, *RSW10 (NagB/RpiA/CoA transferase-like superfamily protein)*, foram usados os oligonucleotídeos (SEQ ID No 11) e (SEQ ID No 12), do gene Eucgr.B03112, *Glycosylase (DNA glycosylase superfamily protein)*, os oligonucleotídeos (SEQ ID No 13) e (SEQ ID No 14), do gene Eucgr.B03882, *TIL1 (DNA polymerase epsilon catalytic subunit)*, os oligonucleotídeos (SEQ ID No 15) e (SEQ ID No 16), do gene Eucgr.B00700, *MSH2 (DNA mismatch repair protein MSH2)*, os oligonucleotídeos (SEQ ID No 17) e (SEQ ID No 18), do gene Eucgr.D00242, *BARD1 (breast cancer susceptibility)*, os oligonucleotídeos (SEQ ID No 19) e (SEQ ID No 20), do gene Eucgr.B03483, *Galactose oxidase (Galactose oxidase/kelch repeat superfamily protein)*, os oligonucleotídeos (SEQ ID No 21) e (SEQ ID No 22), do gene Eucgr.B03325, *RPT2 (Root phototropism protein 2)*, os oligonucleotídeos (SEQ ID No 23) e (SEQ ID No 24), do gene Eucgr.J02288, *GAPDH (glyceraldehyde 3-phosphate dehydrogenase)*, os oligonucleotídeos (SEQ ID No 25) e (SEQ ID No 26), do gene Eucgr.C03857, *(NAD(P)(+) Aldehyde dehydrogenase* , foram usados os oligonucleotídeos (SEQ ID No 27) e (SEQ ID No 28), do gene Eucgr.I01176, *Peroxidase 10*, os oligonucleotídeos (SEQ ID No 29) e (SEQ ID No 30), do gene Eucgr.D00471, *Alcohol dehydrogenase related*, os oligonucleotídeos (SEQ ID No 31) e (SEQ ID No 32).

[091] Os genes endógenos usados como normalizados e os respectivos oligonucleotídeos para o RT-PCR quantitativo foram: Eucgr.K00756,

*Ubiquitin (small ubiquitin-like modifier 2)* com (SEQ ID No 33) e (SEQ ID No 34), Eucgr.B02912, *eEF2 (Eukaryotic translation elongation factor 2)* com (SEQ ID No 35) e (SEQ ID No 36), Eucgr.H02449, *Actin (Actin-related protein 3)* com (SEQ ID No 37) e (SEQ ID No 38) e Eucgr.D00609, *30S (30S ribosomal protein)* com (SEQ ID No 39) e (SEQ ID No 40).

[092] Anterior às análises de RT-qPCR foram realizados testes para verificar a eficiência de amplificação dos oligonucleotídeos (*primers*) desenhados, sendo utilizadas as condições dos oligonucleotídeos com eficiência maior ou igual a 95 %. As reações de PCR em tempo real foram realizadas em um equipamento 7500 Real Time PCR System (Applied Biosystems) utilizando a metodologia *SYBR Green PCR Master Mix* (Life Technologies). Como referência endógena, foram testados os genes *eEF2*, *ACT* e *30S*.

[093] As reações de RT-qPCR foram preparadas em um volume final de 10 µL, contendo molde de cDNA, oligonucleotídeos específicos (*forward* e *reverse*) e corante Fast *SYBR Green PCR 2x Master mix* (Applied Biosystems). As condições de termociclagem foram: 95 °C por 10 minutos e 40 ciclos de desnaturação a 94°C por 15 segundos, seguido de anelamento e extensão a 60 °C por 1 minuto. A especificidade dos produtos amplificados foi analisada pelas curvas de dissociação geradas pelo equipamento. A expressão gênica foi quantificada utilizando o método comparativo de Ct:  $2^{-\Delta Ct}$ .

[094] A especificidade e eficiência dos oligonucleotídeos foram determinadas por eletroforese em gel de agarose (RT-PCR) (Figura 9) e pela análise da curva de fusão (*melting*). Cada uma das curvas de *fusão* desses genes evidenciaram um único pico em RT-qPCR e as curvas de amplificação das nove repetições (três repetições biológicas x três repetições técnicas) para cada gene apresentou excelente repetibilidade. Esses resultados indicam alta especificidade dos oligonucleotídeos para seus respectivos genes-alvo.

[095] O cálculo da eficiência de amplificação da PCR foi feito para os três genes endógenos, e variaram de 0,9 a 1,18, atendendo aos requisitos para análises com PCR quantitativo (Figura 9J). O RNA total das amostras avaliadas foi analisado por gel de agarose. Além do clone suscetível 520 e o resistente 1404, foram incluídas amostras dos clones 2536 e 1407, respectivamente suscetível e resistente ao distúrbio fisiológico (Figura 9H). A integridade do cDNA, preparado a partir do RNA total, foi avaliada pela amplificação do fragmento de 1000 bp do cDNA do gene *eEF2*, utilizado como controle endógeno (Figura 9I).

[096] De uma maneira geral, todos os genes selecionados apresentaram acúmulos de transcritos

consistentes com os resultados do RNA-Seq. O nível de transcritos dos genes envolvidos na síntese de ácido graxo, *KASIII*, *ACCI* e *KCSI*, foi muito inferior em SS (S1) quando comparados com AS (S3), confirmando a repressão global da síntese de ácidos graxos em caules sintomáticos do clone 520 (Figura 10). Não se observaram variações significativas desses transcritos em 1404-ASR (R1), confirmando os resultados do RNA-Seq. A expressão gênica foi calculada pelo método do  $2^{-\Delta CT}$ . *eEF2* (fator de alongação 2 de eucarioto) foi utilizado como controle endógeno. As barras indicam o desvio padrão da média, baseado em repetições técnicas e replicatas biológicas (n=3).

[097] Quanto aos genes envolvidos no reparo de DNA, *DNA Glycosylase*, *TILL* e *MSH2*, o perfil de expressão por RT-qPCR foi similar ao identificado por RNA-Seq (Figura 11). Comparados com 520-AS, os níveis de transcritos desses genes foram inferiores em 520-SS e superiores em 1404-ASR (R1). A expressão gênica foi calculada pelo método do  $2^{-\Delta CT}$ . *eEF2* (fator de alongação 2 de eucarioto) foi utilizado como controle endógeno. As barras indicam o desvio padrão da média, baseado em repetições técnicas e replicatas biológicas (n=3).

[098] Também consistentes com os resultados do RNA-Seq foram os resultados de expressão dos genes envolvidos em resposta a estímulos abióticos, *Galactose oxidase* e *RPT2*, cujos níveis de transcritos nas amostras SS do clone 520, mensurados por RT-qPCR, foram superiores em relação a AS (Figura 12). Além disso, os três genes envolvidos na síntese de fenilpropanóides foram induzidos pelo distúrbio fisiológico, confirmando os dados do RNA-seq (Figura 13). A expressão gênica foi calculada pelo método do  $2^{-\Delta CT}$ . *eEF2* (fator de alongação 2 de eucarioto) foi utilizado como controle endógeno. As barras indicam o desvio padrão da média, baseado em repetições técnicas e replicatas biológicas (n=3).

[099] Exceto por *PGL1*, cujo nível de transcrito por RT-qPCR foi inferior em 520-SS quando comparado com 520-AS, os demais genes selecionados, *BARI*, *ADH*, *RSW10*, *GAPDH*, *UBQ*, acumularam um nível de transcrito inferior nas amostras sintomáticas, determinado por RT-qPCR, em concordância com os resultados do RNA-Seq (Figura 14). A expressão gênica foi calculada pelo método do  $2^{-\Delta CT}$ . *eEF2* (fator de alongação 2 de eucarioto) foi utilizado como controle endógeno. As barras indicam o desvio padrão da média, baseado em repetições técnicas e replicatas biológicas (n=3).

## **Desenvolvimento do processo para diagnóstico do distúrbio fisiológico em troncos de eucalipto baseado na expressão diferencial de genes marcadores**

[100] No perfil do transcriptoma reprimido pelo distúrbio fisiológico, genes envolvidos no sistema de reparo de DNA foram extensivamente reprimidos. Estes resultados foram determinados por RNA-seq e validados por RTqPCR. As plantas são particularmente vulneráveis a fatores que danificam o DNA. Desta forma, elas evoluíram uma rede complexa de mecanismos para detecção de danos no DNA e reparos, dedicados a assegurar a estabilidade genômica por meio de remoção das lesões no DNA e reconstituição da informação genética original (Eggleston, A.K. 2007. DNA replication and repair. *Nature*, 447, 923).

[101] Se o sistema de reparo de DNA é comprometido, então a instabilidade genômica gera um sinal para ativar vias de morte celular. Consequentemente, a instabilidade do genoma causada por um sistema ineficiente de reparo do DNA, em condições de estresse abiótico, pode ser o sinal de ativação das vias de morte celular induzida pelo distúrbio fisiológico. Sendo assim, uma característica molecular da síndrome do distúrbio fisiológico é a repressão de genes específicos no processo de reparo do DNA.

[102] O presente pedido de patente reivindica que a utilização de oligonucleotídeos sintéticos representados pelas SEQ ID N° 1 a SEQ ID N° 32, para mensurar a expressão dos genes descritos na Tabela 6 pode ser usada para o processo de diagnóstico do distúrbio fisiológico e são baseados nos denominados genes marcadores do distúrbio fisiológico. Os níveis dos transcritos correspondentes aos genes marcadores em caules sintomáticos inferiores aqueles de caules assintomáticos do mesmo genótipo revela o estabelecimento do distúrbio fisiológico.

[103] Um método de quantificar os níveis de transcritos é o real time PCR ou RT-PCR quantitativo ou RT-qPCR. Esta técnica está amplamente difundida na literatura e largamente utilizada para quantificar expressão gênica. O RNA total extraído das amostras é convertido em cDNA usando reverse transcriptase comerciais. Kits comerciais contendo todos os componentes da reação da polimerase em cadeia para quantificação dos níveis de cDNA são disponíveis.

[104] Entretanto, há necessidade de se desenharem e sintetizarem oligonucleotídeos iniciadores (*primers*) que sejam complementares à sequência do cDNA no sentido direto (*forward*) e no sentido inverso (*reverse*) dos genes de interesse do genoma de todas as cultivares e clones de eucalipto cultiváveis. Além disso, têm de ser capazes de discriminar o gene de interesse dos outros membros de famílias multigênicas.

[105] Esta invenção reivindica as sequências dos oligonucleotídeos iniciadores (*primers forward e reverse*) desenhados para amplificação do cDNA dos transcritos dos genes *KAS\_III* (Eucgr.E00458), SEQ ID No 1 e SEQ ID No 2; *ACCI* (Eucgr.F00484), SEQ ID No 3 e SEQ ID No 4; *KCSI* (Eucgr.H00232), SEQ ID No 5 e SEQ ID No 6; *ADH* (Eucgr.K00927), SEQ ID No 7 e SEQ ID No 8; *PGLI* (Eucgr.K00013), SEQ ID No 9 e SEQ ID No 10; *RSW10* (Eucgr.D01028), SEQ ID No 11 e SEQ ID No 12; *Glycosylase* (Eucgr.B03112), SEQ ID No 13 e SEQ ID No 14; *TILI* (Eucgr.B03882), SEQ ID No 15 e SEQ ID No 16; *MSH2* (Eucgr.B00700), SEQ ID No 17 e SEQ ID No 18; *BARD1* (Eucgr.D00242), SEQ ID No 19 e SEQ ID No 20; *Galactose oxidase* (Eucgr.B03483), SEQ ID No 21 e SEQ ID No 22; *RPT2* (Eucgr.B03325), SEQ ID No 23 e SEQ ID No 24; *GAPDH* (Eucgr.J02288), SEQ ID No 25 e SEQ ID No 26; *NAD(P)(+) Aldehyde dehydrogenase* (Eucgr.C03857), SEQ ID No 27 e SEQ ID No 28; *Peroxidase 10* (Eucgr.I01176), SEQ ID No 29 e SEQ ID No 30; *Alcohol dehydrogenase related* (Eucgr.D00471), SEQ ID No 31 e SEQ ID No 32, identificados na sequência da região codificadora dos respectivos genes.

[106] Estes conjuntos de oligonucleotídeos iniciadores (*primers*) de cada gene apresentaram curvas de dissociação consistentes para alinhamento específico para gene único (Figura 15). Portanto, são capazes de discriminar o gene alvo de outros membros da mesma família no genoma de eucalipto.

[107] Também relevante foi a repressão expressiva dos genes envolvidos na síntese de ácidos graxos, incluindo ácidos graxos de cadeia longa e ceras, nas amostras 520-SS. Estes genes, essenciais para síntese de ácidos graxos de cadeia longa, estão envolvidos em diversas etapas de desenvolvimento e crescimento de plantas, como síntese de ceras e cutina nos caules e folhas, formação da raiz lateral, controle de divisão celular na região de meristemas (Bonaventure, G., Salas, J.J., Pollard, M.R., Ohlrogge, J.B. 2003. Disruption of the FATB gene in Arabidopsis demonstrates an essential role of saturated fatty acids in plant growth. *The Plant Cell*,15, 1020-1033). Reduções ou bloqueios na síntese de ácidos graxos saturados têm sido associados a declínios na taxa de crescimento de plantas, na síntese de ceras e cutina, um dos principais componentes da cutícula da planta e indução de morte celular [(1) Bonaventure, G., Salas, J.J., Pollard, M.R., Ohlrogge, J.B. 2003. Disruption of the FATB gene in Arabidopsis demonstrates an essential role of saturated fatty acids in plant growth. *The Plant Cell*,15, 1020-1033; (2) Baud, S., Guyon, V., Kronenberger, J., Wuillème, S., Miquel, M., Caboche, M., Rochat, C. 2003. Multifunctional acetyl-CoA carboxylase 1 is essential for very long chain fatty acid elongation and embryo development

in Arabidopsis. The Plant Journal, 33, 75-86; (3) Schnurr, J., Shockey, J. 2004. The acyl-CoA synthetase encoded by LACS2 is essential for normal cuticle development in Arabidopsis. The Plant Cell, 16, 629-642).

[108] O transcriptoma induzido pelo distúrbio fisiológico divergiu consideravelmente das amostras suscetíveis assintomáticas do clone 520. Nas modificações positivas de expressão gênica induzidas pelo distúrbio fisiológico, observou-se um enriquecimento de genes diferencialmente expressos nas categorias biológicas do sistema antioxidativo (catálase, peroxidase, desidrogenases, superóxido dismutase), morte celular (caspases, VPE, SAG, genes associados à senescência) e síntese de flavonóide. Genes regulatórios, normalmente envolvidos em respostas típicas a estresses abióticos, também foram enriquecidos nas modificações positivas de expressão gênica, como genes da família NAC, APE, WRKY de transfatores transcricionais.

[109] Este conjunto de genes positivamente regulados categorizam as respostas majoritárias de plantas a estresses abióticos. De fato, estresses abióticos primários, resultante de mudança de temperatura, salinidade, seca e alagamento, levam a estresse osmótico secundário como efeito comum, que aciona o sistema oxidativo de plantas pelo acúmulo e produção de espécies reativas de oxigênio (ROS) (Gill, S.S., Tuteja, N. 2010. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. Plant physiology and biochemistry, 48, 909-930).

[110] Para o desenvolvimento do processo de diagnose molecular também foram selecionados genes regulados positivamente (*up*) pelo distúrbio fisiológico e classificados como genes induzidos por estresses abióticos (seca, osmótico, alta temperatura). Estes genes não devem ser marcadores específicos do estado de distúrbio fisiológico, mas em combinação com a expressão dos genes selecionados que são negativamente regulados pelo distúrbio fisiológico resultam em um diagnóstico molecular do estabelecimento do distúrbio fisiológico. Este processo para diagnose molecular de distúrbio fisiológico em eucalipto reivindica que a utilização dos oligonucleotídeos específicos na quantificação da expressão dos genes *Galactose Oxidase* e *RPT2* são componentes da invenção.

[111] Este processo de diagnose molecular reivindica que a utilização dos oligonucleotídeos específicos na quantificação da expressão dos genes *Galactose Oxidase* e *RPT2* em combinação com a expressão dos genes *TILI*, *MSHII*, *KASIII* e *KCSI* possa ser usado para diagnóstico do distúrbio fisiológico e serão denominados genes marcadores. Os níveis dos transcritos

correspondentes aos genes marcadores em caules sintomáticos em relação àqueles de caules assintomáticos do mesmo genótipo revela o estabelecimento do distúrbio fisiológico.

[112] Resultado positivo do processo de diagnose molecular do distúrbio fisiológico implica que os níveis de transcritos ou a expressão dos genes *TIL1*, *MSHII*, *KASIII* e *KCSI* no caule dito sintomático (SS) é menor do que aqueles do caule assintomático (AS) do mesmo clone, ou genótipo. A expressão desses genes regulados negativamente pelo distúrbio fisiológico está associada ao estabelecimento do estado de distúrbio fisiológico. Além disso, os níveis de transcritos ou a expressão dos genes *Galactose Oxidase* e *RPT2* devem ser superiores nas amostras de caules dito sintomáticos do que nas amostras de caules assintomáticos. A expressão induzida desses genes de estresse abiótico em geral indica que o caule dito sintomático não evoluiu para um estado de necrose geral irreversível, característico de necroses em resposta a estresses abióticos prolongados.

[113] Esta invenção também reivindica que a utilização dos oligonucleotídeos específicos na quantificação da expressão dos genes *Galactose Oxidase* e *RPT2* em combinação com a expressão de pelo menos um gene discriminado na Tabela 5 e pelos menos um gene discriminado na Tabela 6 possam ser usados para diagnóstico do distúrbio fisiológico, e serão denominados genes marcadores. Os níveis dos transcritos correspondentes aos genes marcadores em caules sintomáticos em relação àqueles de caules assintomáticos do mesmo genótipo revela o estabelecimento do distúrbio fisiológico.

[114] O resultado positivo da diagnose implica em expressão aumentada de *Galactose Oxidase* e *RPT2* nas amostras de caules dito sintomáticos em relação às amostras dos caules assintomáticos e expressão diminuída dos genes da Tabela 5 e Tabela 6 em amostras de caules dito sintomáticos em relação às amostras de caules assintomáticos do mesmo clone ou genótipo.

[115] Esta invenção utiliza o método real time RT-PCR ou RT-PCR quantitativo ou RT-qPCR para quantificar os níveis de transcritos de genes marcadores. Trata-se de uma técnica amplamente difundida na literatura e largamente utilizada para quantificar expressão gênica. O RNA total extraído das amostras é convertido em cDNA usando reverse transcriptase comerciais. Kits comerciais contendo todos os componentes da reação da polimerase em cadeia para quantificação dos níveis de cDNA são disponíveis.

[116] Entretanto, há necessidade de se desenharem e sintetizarem oligonucleotídeos iniciadores (*primers*) que sejam complementares à sequência do cDNA no sentido direto (*forward*) e no sentido inverso (*reverse*) dos genes marcadores do genoma de todas os cultivares e clones de eucalipto

cultiváveis. Além disso, têm de ser capazes de discriminar o gene de interesse dos outros membros de famílias multigênicas.

[117] Os clones (genótipos) de eucalipto analisados no experimento estão especificados na Tabela 8 e as amostras coletadas para validar o processo desenvolvido para diagnóstico do distúrbio fisiológico em eucalipto são especificadas na Tabela 10. Foram coletadas amostras, com idades de 16 meses ou 8 meses (especificadas na designação das amostras). Destacando que os clones utilizados para comprovação molecular do estado de distúrbio fisiológico se encontravam em idade suficientemente avançada para ser realizado o registro do conjunto de sintomatologia específica do distúrbio fisiológico.

<b>Tabela 8. Clones e amostras utilizadas para validar o método de diagnóstico</b>		
Clones Assintomáticos (AS)		Clones sintomáticos (SS)
Suscetíveis	Resistentes	Suscetíveis
520	1572	520
1296	1404	1296
5204	1644	5204
2361	1407	2361
5341		5341

#### **Extração do RNA total das amostras sintomáticas e assintomáticas dos troncos de eucalipto**

[118] O RNA total foi preparado de todas as amostras. Foram utilizados 0,2 g de cada amostra armazenada em -80°C e então macerada em nitrogênio líquido com auxílio de almofariz e pistilo. Ao material macerado foi adicionado 1 mL do tampão de extração, contendo CTAB(*cethyltrimethylammonium bromide*) 2% (p/v), PVP 2% (p/v), Tris-HCL 100 mM, EDTA 25 mM, NaCl 20 mM, e  $\beta$ -mercaptoethanol 2% (v/v). As amostras foram incubadas por 20 minutos a 65 °C.

[119] A esse material adicionou-se 1 mL de clorofórmio, seguida de centrifugação a 15.000 x g, por 10 minutos, a 4°C. Após uma segunda extração do sobrenadante com igual volume de clorofórmio, RNA total foi precipitado do sobrenadante com igual volume de isopropanol (aproximadamente 400  $\mu$ L), por 10 minutos a temperatura ambiente e, posteriormente, por

centrifugação por 10 minutos a 4 ° C (12.000 x g).

[120] Após a centrifugação, o sobrenadante foi descartado, o sedimento foi lavado com 1 mL de etanol a 80% preparado com água DEPC (Dietilpirocarbonato), e os tubos foram conduzidos para outra centrifugação por 1 minuto a 4 ° C (12.000 x g). O líquido residual foi removido do tubo com uma pipeta e seco sobre papel absorvente, antes de ser ressuspensão em 20 µL de água tratada com DEPC.

[121] A Figura 16 mostra um exemplo de eletroforese do RNA em géis de agarose, demonstrando a qualidade do RNA a ser utilizado para análise de expressão gênica dos genes marcadores e confirmação do estado de distúrbio fisiológico.

[122] Somente os clones geneticamente suscetíveis, clones 1296, clone 520, clone 5204, clone 2361, clone 5341, poderão estar acometidos pelo distúrbio fisiológico, o que poderá ser registrado em caules sintomáticos, como controles positivos.

[123] Embora os clones 1572, 1404 e 1407 não sejam acometidos pelo distúrbio e, conseqüentemente, não apresentem sintomas de distúrbio fisiológico, estes clones resistentes ao distúrbio fisiológico também foram incluídos na demonstração experimental do processo de diagnóstico molecular do distúrbio fisiológico, como controles negativos.

[124] Como o processo de diagnóstico molecular se baseia na expressão diferencial de genes marcadores do distúrbio fisiológico, os níveis do transcrito dos genes marcadores nos caules sintomáticos deverão obrigatoriamente ser comparados com seus níveis de transcritos de caules assintomáticos do mesmo material genético.

[125] A quantificação do RNA foi monitorada, pelo espectrofotômetro, pela relação de absorvância 260/280 e a integridade destes RNAs foi verificada por eletroforese em gel de agarose desnaturante 1,5 %, e também pelo equipamento 2100 *Bioanalyser* (Agilent Technologies), observando-se a pureza e o tamanho das amostras. O RNA total de cada uma das amostras foi diluído para 5000 pg/µL para a visualização da sua integridade no instrumento 2100 *Bioanalyser* (Agilent). Para isso, foi seguido o protocolo Agilent RNA 6000 pico Assay, de acordo com os procedimentos recomendados com relação à utilização do equipamento e à corrida das amostras, que foi efetuada usando um *chip* específico de RNA, juntamente com os kits de análise de RNA. Cada kit é composto por uma matriz de separação que é uma solução de polímero, uma solução de corante fluorescente de intercalação à fita dupla de DNA e marcadores de tamanho de RNA.

### **Análise da expressão dos genes marcadores por RT-PCR quantitativo (RT-qPCR)**

[126] A presente invenção para o processo de diagnóstico molecular do distúrbio fisiológico baseia-se na utilização de oligonucleotídeos específicos na análise de variação da expressão de genes marcadores pelo estabelecimento de distúrbio fisiológico. Os genes marcadores selecionados pela análise do transcriptoma induzido pelo distúrbio foram: *Galactose Oxidase* e *RPT2*, com expressão ou níveis de transcritos aumentados pelo distúrbio, *TIL1*, *MSHIII*, *KASIII* e *KCSI* com níveis de transcritos diminuídos pelo distúrbio.

[127] Com o objetivo de comprovar a efetividade do processo de diagnóstico molecular do distúrbio fisiológico, procedeu-se a extração do RNA total de amostras sintomáticas e assintomáticas (Figura 16) e, em seguida, a quantificação por RT-PCR quantitativo (RT-qPCR ou real time RT-PCR) da expressão dos níveis de mRNA dos genes marcadores em comparação com os níveis de expressão de troncos assintomático do mesmo genótipo. Amostras de clones resistentes assintomáticos foram incluídas nos ensaios como controles negativos.

[128] A extração de RNA dos clones 1296, 520, 5204, 2361, 5341, 1572, 1404, 1644 e 1407 foi realizada pelo método CTAB, seguido por precipitação com isopropanol, conforme descrito. A integridade do RNA foi avaliada por eletroforese em gel de agarose 1,5%. O RNA foi quantificado no espectrofotômetro e a concentração foi expressa em  $\mu\text{g} \cdot \mu\text{L}^{-1}$ . O RNA total isolado (4 $\mu\text{g}$ ) foi tratado com DNase livre de RNase (Invitrogen) e utilizados para obtenção do cDNA, usando a enzima transcriptase reversa M-MLV (Invitrogen) e oligo-dT, de acordo com determinação do fabricante.

[129] As reações de RT-qPCR foram preparadas em um volume final de 10  $\mu\text{L}$ , contendo molde de cDNA, oligonucleotídeos (*primers*) específicos (*forward* e *reverse*) e corante Fast *SYBR Green PCR 2x Master mix* (Applied Biosystems). As condições de termociclagem foram: 95 °C por 10 minutos e 40 ciclos de desnaturação a 94°C por 15 segundos, seguido de anelamento e extensão a 60 °C por 1 minuto. A especificidade dos produtos amplificados foi analisada pelas curvas de dissociação geradas pelo equipamento. A expressão gênica foi quantificada utilizando o método comparativo de Ct:  $2^{-\Delta\text{Ct}}$ .

[130] Para avaliar os níveis de expressão dos seis genes marcadores selecionados foram utilizados os oligonucleotídeos (*primers*) específicos de cada gene. Como controle endógeno para normalização dos dados de RTqPCR foi utilizado o gene de eucalipto *eEF2* (*Eukaryotic translation*

*elongation factor 2*; Eucgr.B02912) que codifica o fator de alongação de tradução 2. As sequências dos oligonucleotídeos específicos para *eEF2* são as SEQ ID No 35 e SED ID No 36.

[131] Conforme os resultados obtidos pelo RNA-seq e validados por RTqPCR, os níveis dos transcritos dos genes *Galactose Oxidase* e *RPT2* foram superiores em troncos sintomáticos quando comparados com troncos assintomáticos do mesmo clone (genótipo). A expressão gênica foi calculada pelo método do  $2^{-\Delta CT}$ . *eEF2* foi utilizado como controle endógeno. As barras indicam o desvio padrão da média, baseado em repetições técnicas e replicatas biológicas (n=3). Estes genes são induzidos por estresses abióticos de uma maneira geral e, portanto, foram incluídos no diagnóstico molecular como um exemplo de genes induzidos pelo distúrbio fisiológico como comprovante de que a maquinaria de transcrição está funcionando nas amostras sintomáticas, não representando um estágio de necrose irreversível (Figura 17).

[132] Esta interpretação foi observada com os três genes selecionados da síntese de fenilpropanóides também induzidos pelo distúrbio fisiológico nos troncos sintomáticos dos clones 520 e 2361, mas não nos troncos sintomáticos dos clones 1296, 5204 e 5341 (Figura 18). A expressão gênica foi calculada pelo método do  $2^{-\Delta CT}$ . *eEF2* foi utilizado como controle endógeno. As barras indicam o desvio padrão da média, baseado em repetições técnicas e replicatas biológicas (n=3).

[133] Conforme os resultados obtidos pelo RNA-seq e validados por RTqPCR, os níveis dos transcritos dos genes *TILL1* e *MSHII*, envolvidos no sistema de reparo do DNA, *KASIII* e *KCSI*, envolvidos na síntese de ácidos graxos de cadeia longa, foram inferiores em troncos sintomáticos quando comparados com troncos assintomáticos do mesmo clone (genótipo) (Figuras 19 e 20). A expressão gênica foi calculada pelo método do  $2^{-\Delta CT}$ . *eEF2* foi utilizado como controle endógeno. As barras indicam o desvio padrão da média, baseado em repetições técnicas e replicatas biológicas (n=3).

[134] Tomando em conjunto, a repressão dos genes do sistema de reparo do DNA e dos genes da síntese de ácidos graxos, além da indução dos genes *Galactose Oxidase* e *RPT2* caracterizam o estabelecimento do distúrbio fisiológico em todos os troncos sintomáticos dos clones analisados. Portanto, o conjunto do perfil de expressão dos genes marcadores nos troncos sintomáticos em comparação com o nível de expressão nos troncos assintomáticos de todos os clones suscetíveis analisados valida o processo de diagnóstico molecular descrito.

## **Conclusão**

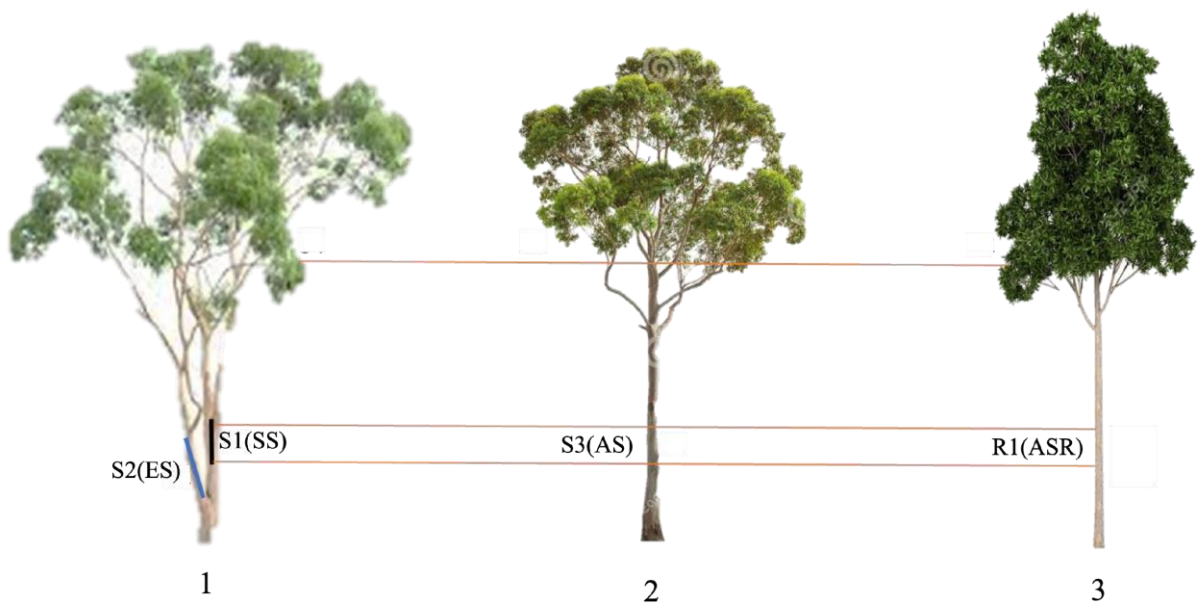
[135] Os oligonucleotídeos sintéticos e o processo de diagnóstico molecular descritos e requeridos nesta invenção se diferenciam de todos os demais processos anteriores descritos porque se baseiam no conhecimento adquirido das variações globais de expressão de genes identificados como repórteres da condição do distúrbio fisiológico em troncos de eucalipto. Estes oligonucleotídeos utilizados neste processo de diagnóstico molecular possibilitam como vantagem primordial a identificação de clones jovens suscetíveis ao distúrbio fisiológico. Uma vez certificado nestes clones jovens o estabelecimento do estado de distúrbio fisiológico, eles poderão ser eliminados do cultivo, prevenindo, portanto, o seu crescimento até a fase adulta não comprometendo assim a qualidade da madeira.

[136] Nenhum método molecular de diagnóstico de distúrbio fisiológico foi previamente descrito, principalmente devido ao desconhecimento sobre as mudanças moleculares relacionadas à variação global de expressão gênica induzida especificamente pelo distúrbio fisiológico em eucalipto. Para o desenvolvimento da presente invenção, foi inicialmente determinado o transcriptoma induzido pelo distúrbio fisiológico e, em seguida, selecionado um conjunto de genes cuja variação da expressão gênica de uma forma conjunta representa um perfil de expressão específico do distúrbio fisiológico.

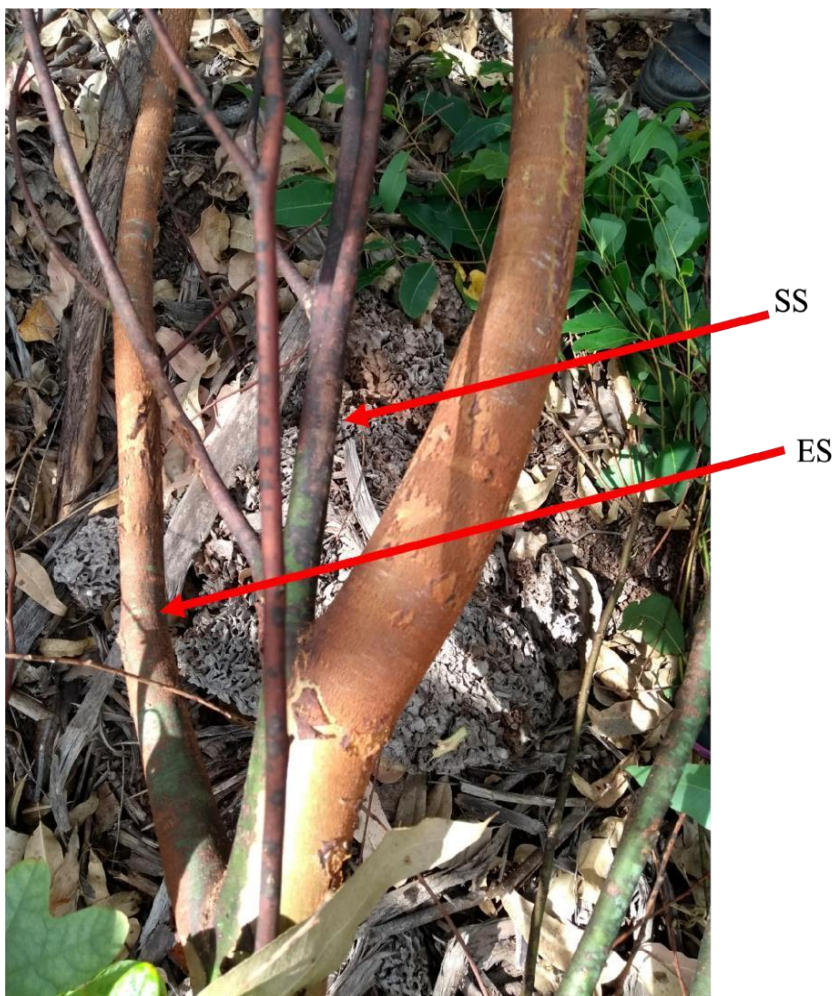
[137] Estes genes, listados nas tabelas 4 e 5, foram designados genes marcadores por indicarem o estado do distúrbio fisiológico por meio da variação da expressão gênica. A invenção também reivindica o uso de sequência dos oligonucleotídeos específicos para determinar a expressão específica dos referidos genes por RT-PCR quantitativo para o diagnóstico molecular do distúrbio fisiológico.

## Reivindicações

- 1. Oligonucleotídeos** caracterizados por serem as SEQ ID N° 1 a SEQ ID N° 32.
  
- 2. Processo de diagnóstico molecular do distúrbio fisiológico em eucalipto por meio dos oligonucleotídeos definidos na reivindicação 1**, caracterizado por compreender as seguintes etapas:
  - Utilização dos 32 oligonucleotídeos (SEQ ID N° 1 a SEQ ID N° 32) específicos dos dezesseis genes marcadores;
  - Colheita de caules sintomáticos e assintomáticos de clones de eucalipto, três réplicas biológicas no mínimo;
    - Congelamento imediato em nitrogênio líquido;
    - Extração de RNA total dos caules sintomáticos e assintomáticos;
  - Síntese do DNA complementar com a enzima reverse transcriptase;
  - Quantificação do nível de transcrito dos seis genes marcadores nos caules sintomáticos e assintomáticos por meio de RT-qPCR (RT-PCR em tempo real), utilizando três réplicas técnicas;
  - Determinação do nível de transcrito dos seis genes marcadores nos clones sintomáticos e assintomáticos;
    - Análise do perfil comparativo de expressão dos genes marcadores.
  
- 3. Uso dos oligonucleotídeos** definidos na reivindicação 1 caracterizado por ser para identificação dos genes diferencialmente expressos pelo distúrbio fisiológico em eucaliptos, diagnóstico do distúrbio fisiológico e produção de kits diagnóstico.
  
- 4. Uso dos genes** *KAS\_III* (Eucgr.E00458); *ACC1* (Eucgr.F00484); *KCS1* (Eucgr.H00232); *ADH* (Eucgr.K00927); *PGL1* (Eucgr.K00013); *RSW10* (Eucgr.D01028); *Glycosylase* (Eucgr.B03112); *TIL1* (Eucgr.B03882); *MSH2* (Eucgr.B00700); *BARD1* (Eucgr.D00242); *Galactose oxidase* (Eucgr.B03483); *RPT2* (Eucgr.B03325); *GAPDH* (Eucgr.J02288); *NAD(P)(+) Aldehyde dehydrogenase* (Eucgr.C03857); *Peroxidase 10* (Eucgr.I01176); *Alcohol dehydrogenase related* (Eucgr.D00471) diferencialmente expressos.



**Figura 1**



**Figura 2**

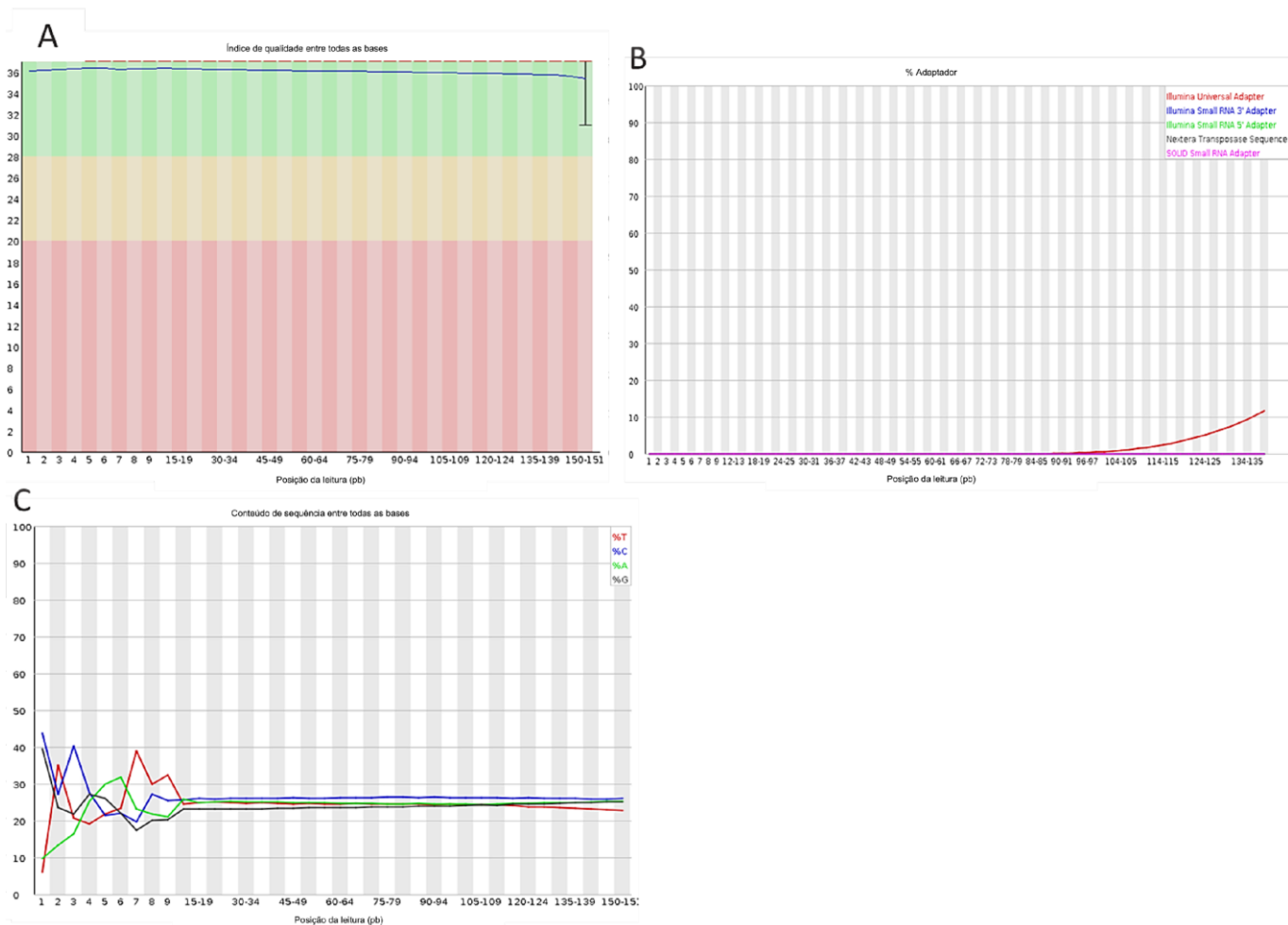
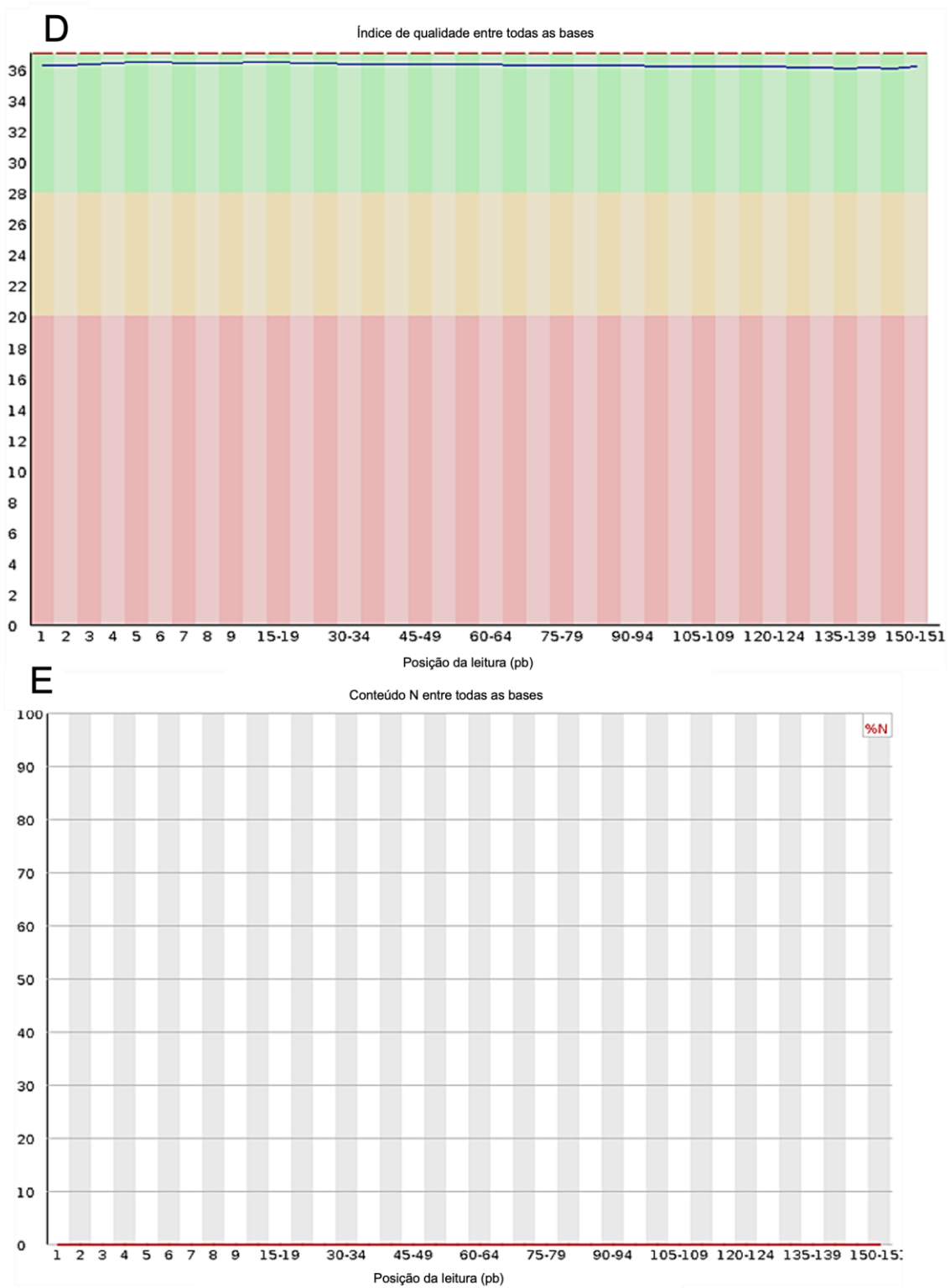
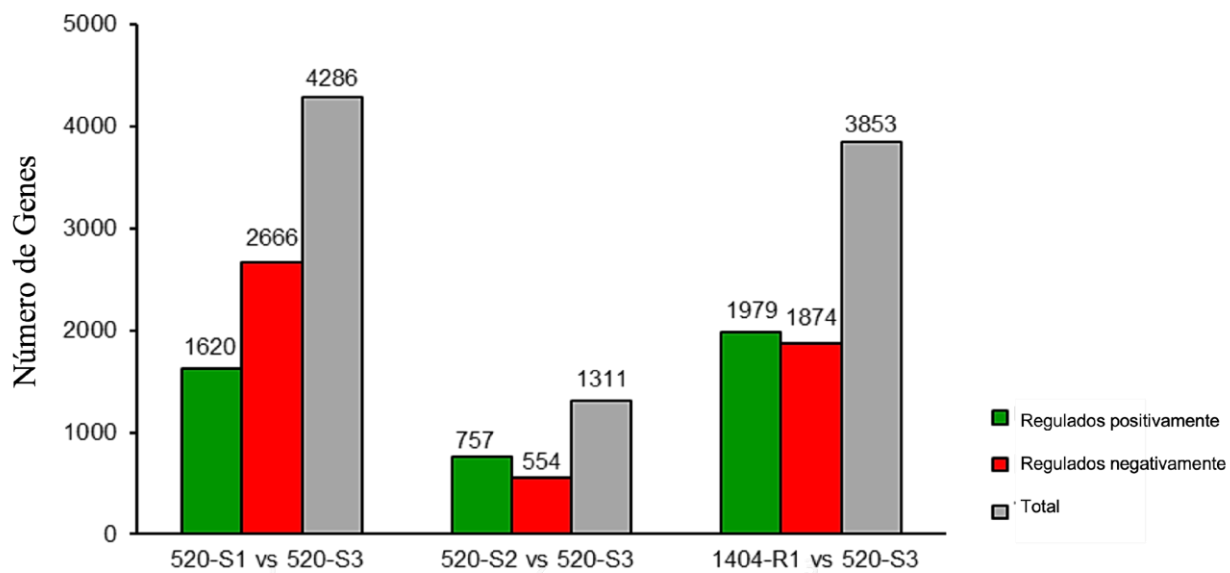
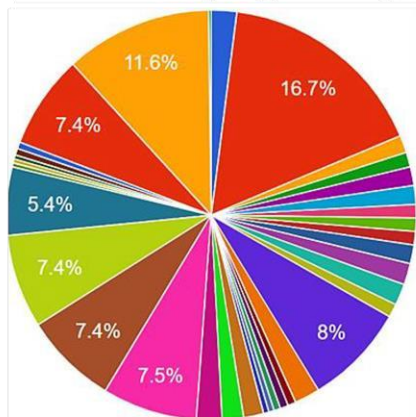


Figura 3

**Figura 4**

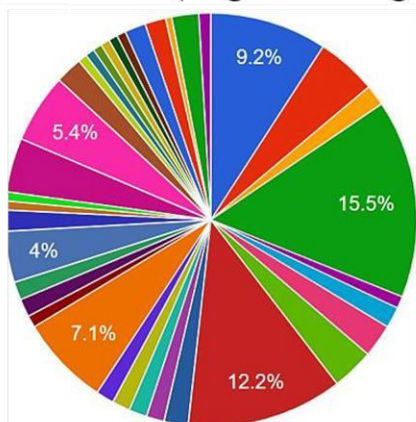
**Figura 5**

### S1-S3 (Regulados positivamente)



- Processo de oxidação-redução
- Transporte transmembrana
- Regulação de processo biossintético
- Regulação da biossíntese do RNA
- Regulação de transcrição em DNA molde
- Metabolismo de carboidratos
- Regulação do Metabolism de compostos contendo nucleobases
- Estabelecimento de localização

### S1-S3 (Regulados negativamente)



- Movimento baseado em microtubulos
- Metabolismo de carboidratos
- Organização ou biogênesis de componente celular
- Organização de organelas
- Processo metabólico de glucano celular
- Metabolismo de carboidrato

**Figura 6**



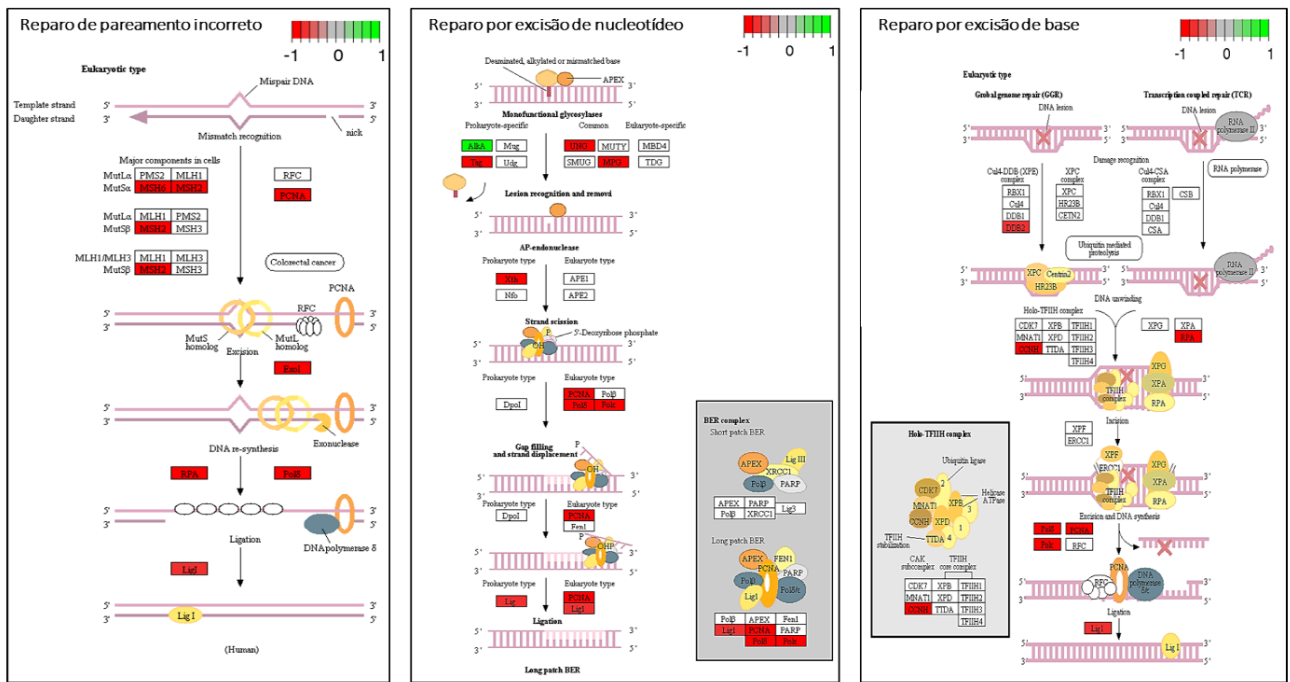
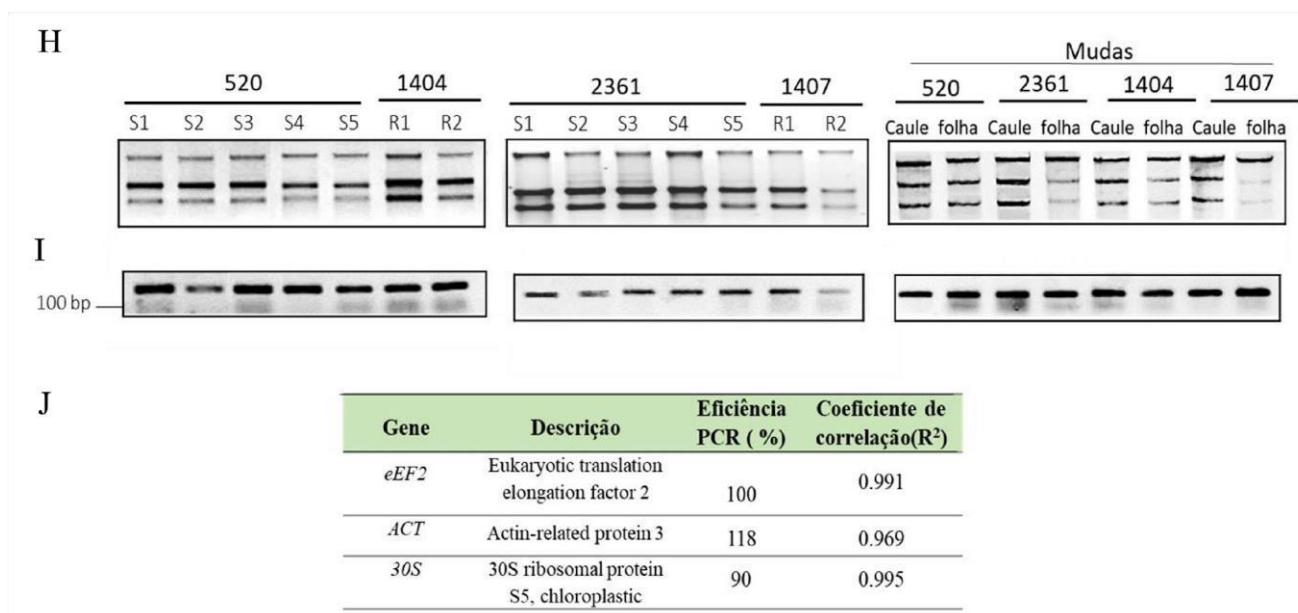


Figura 8



**Figura 9**

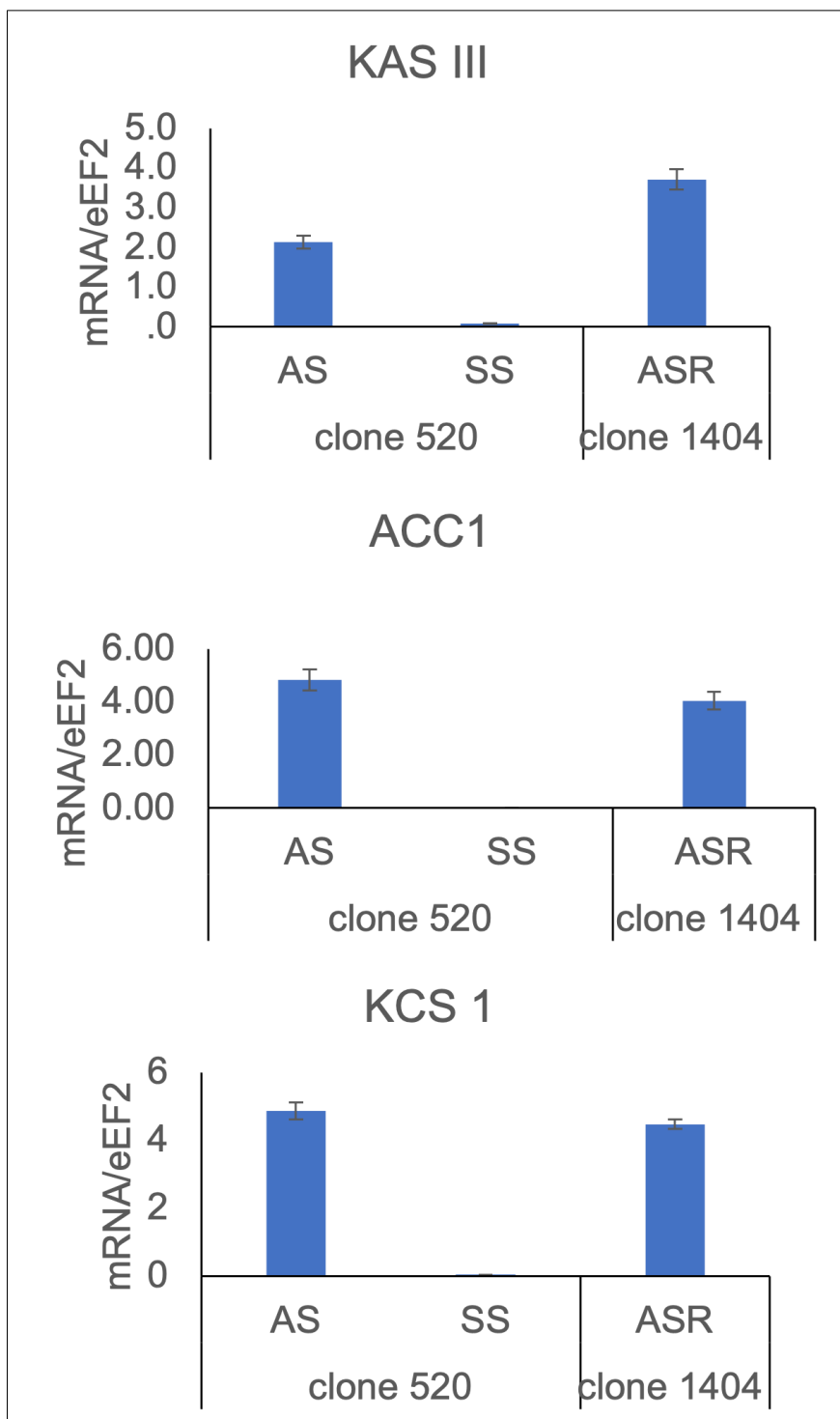
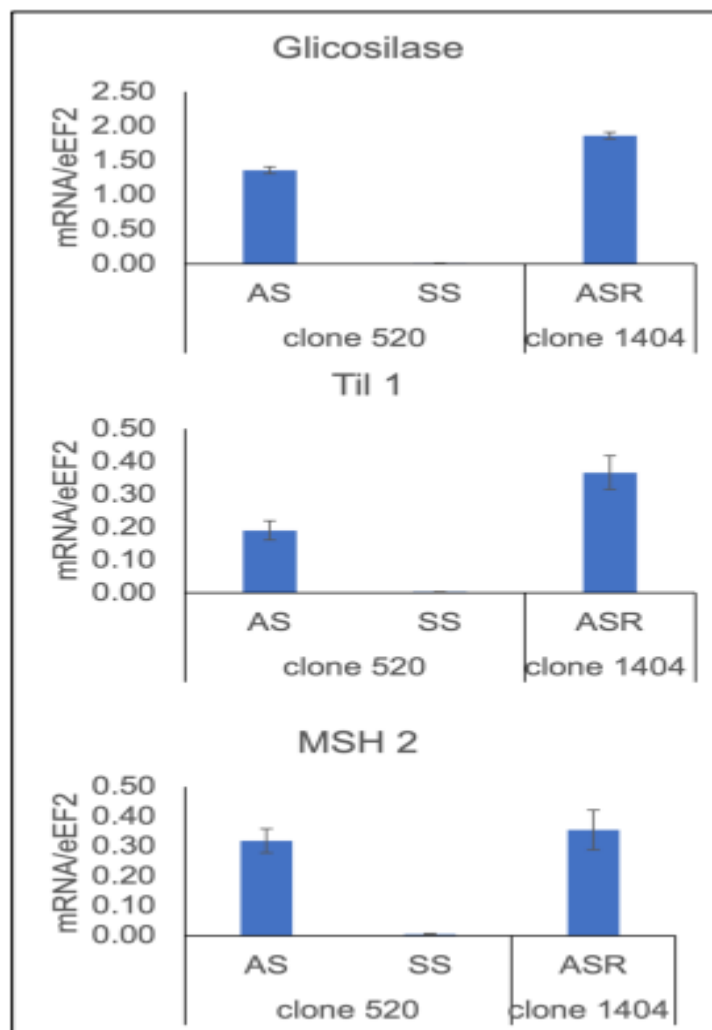


Figura 10

**Figura 11**

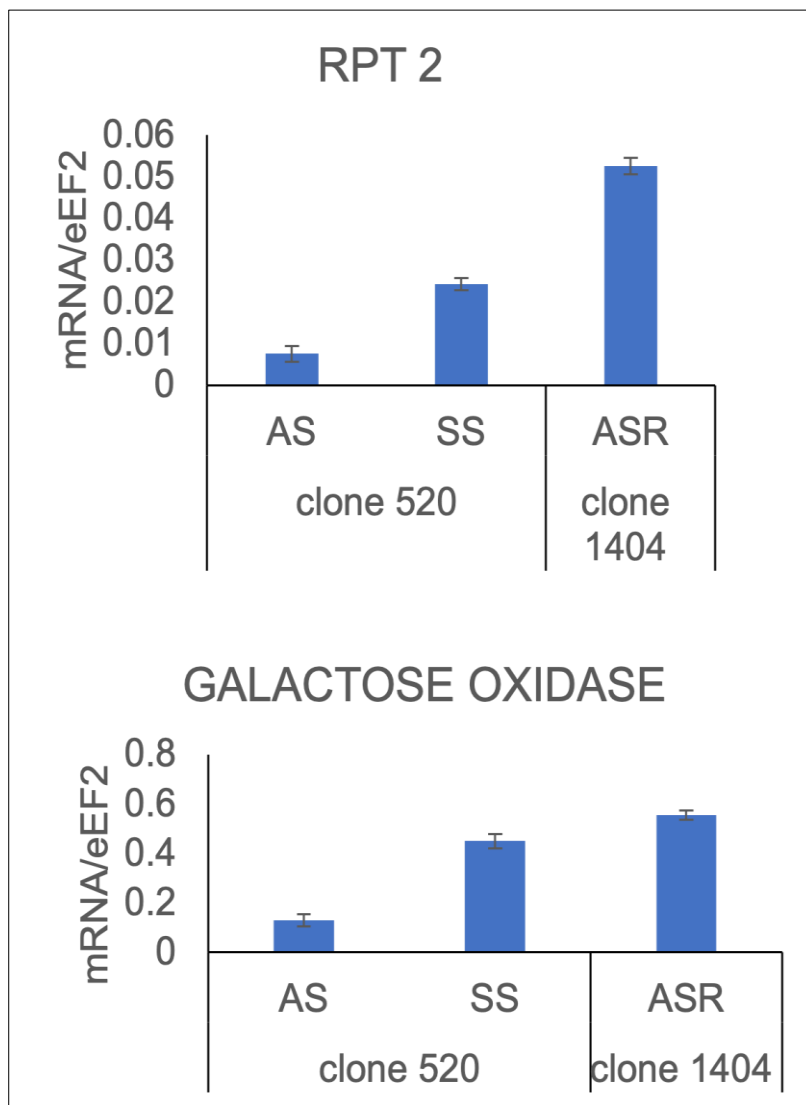


Figura 12

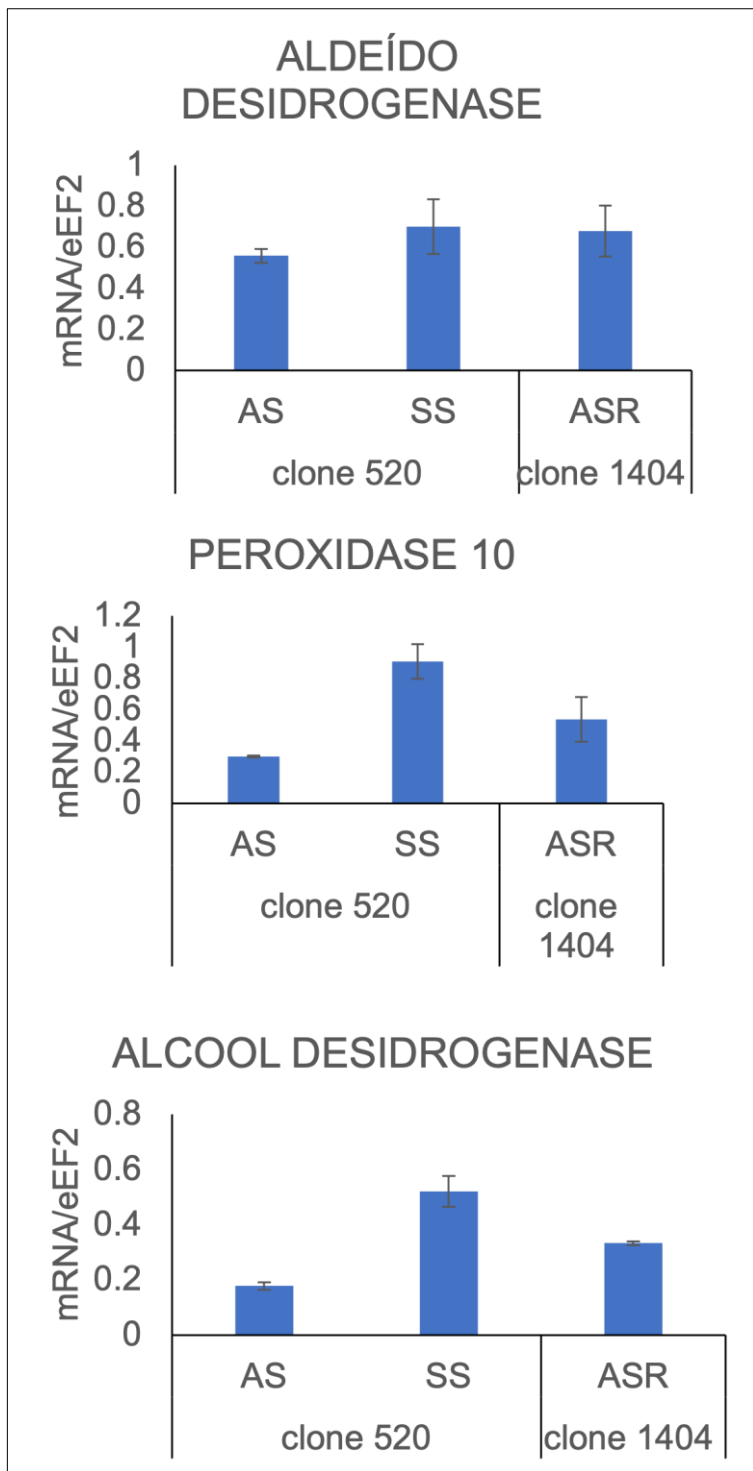
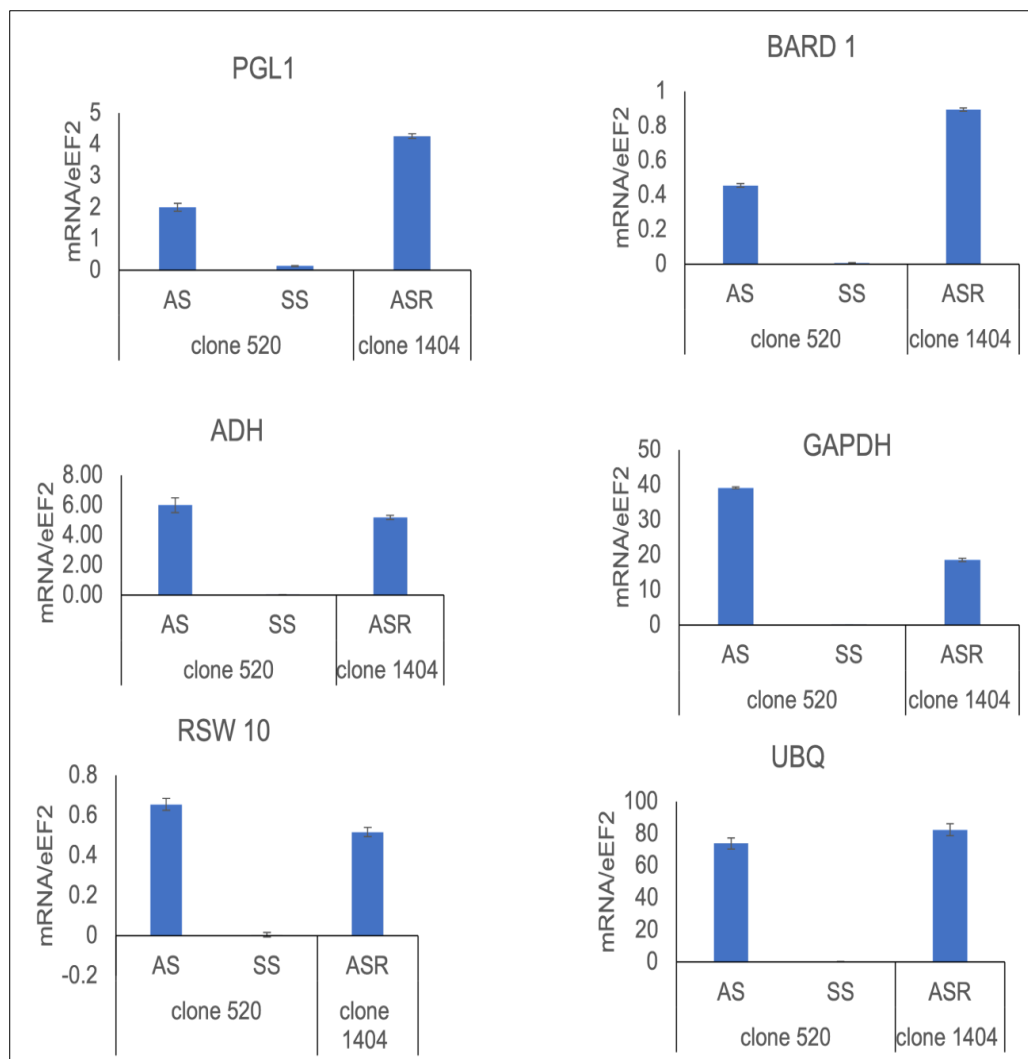


Figura 13

**Figura 14**

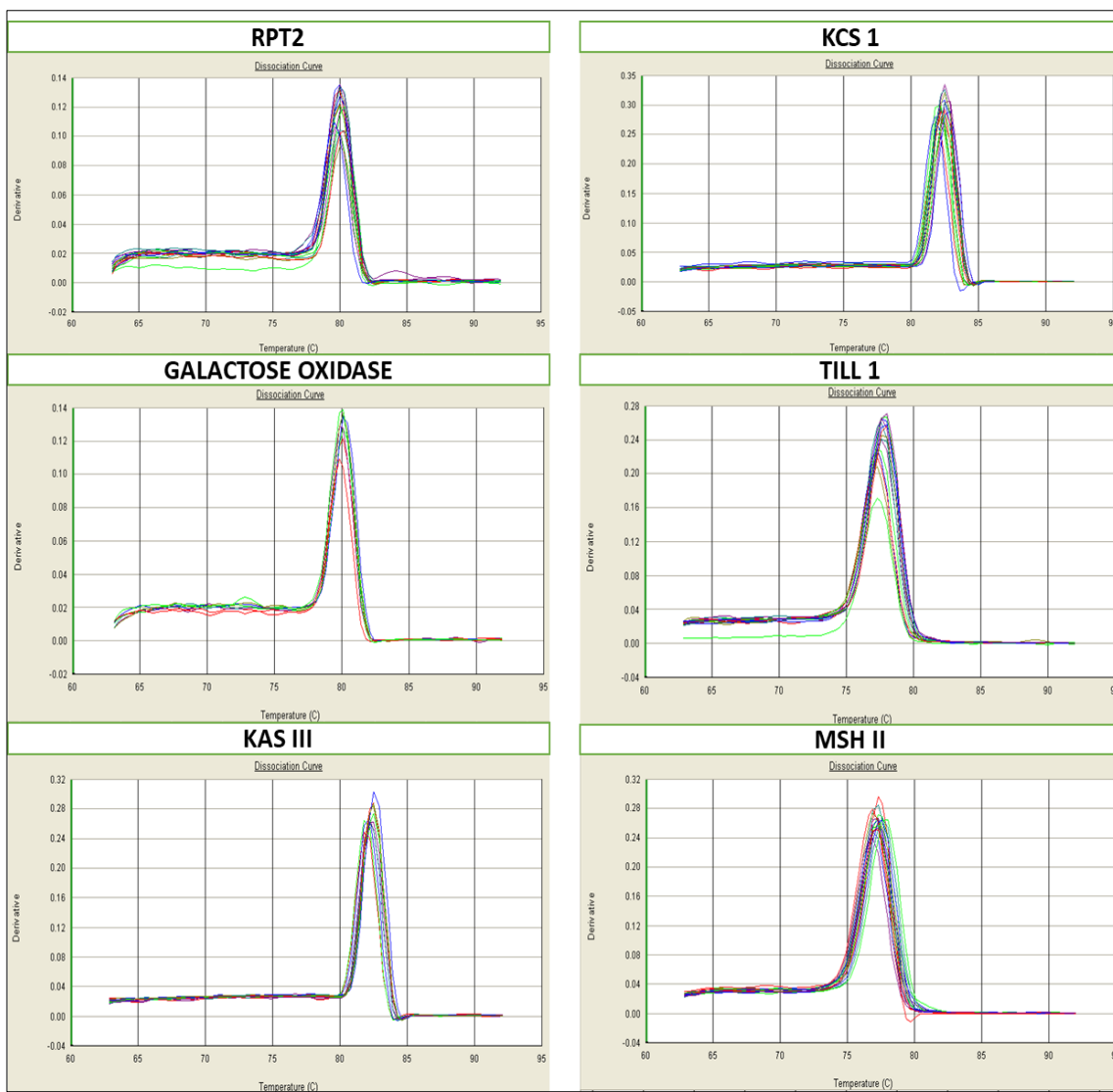
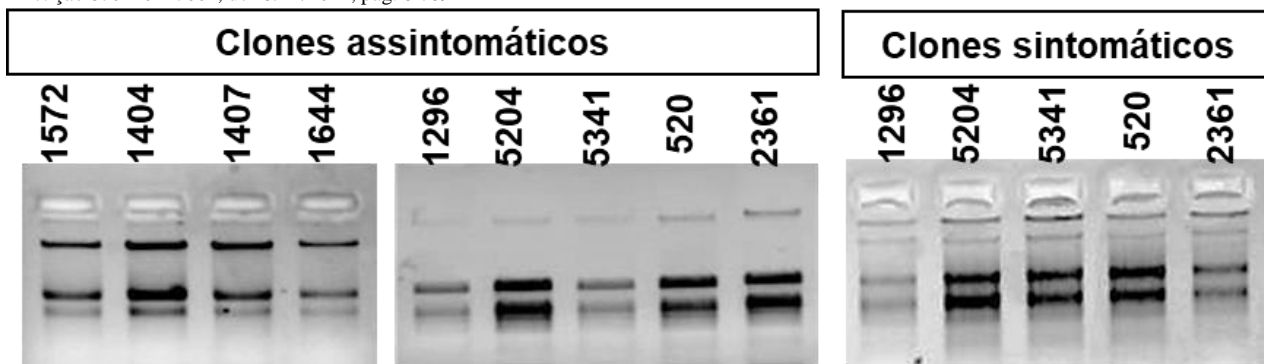
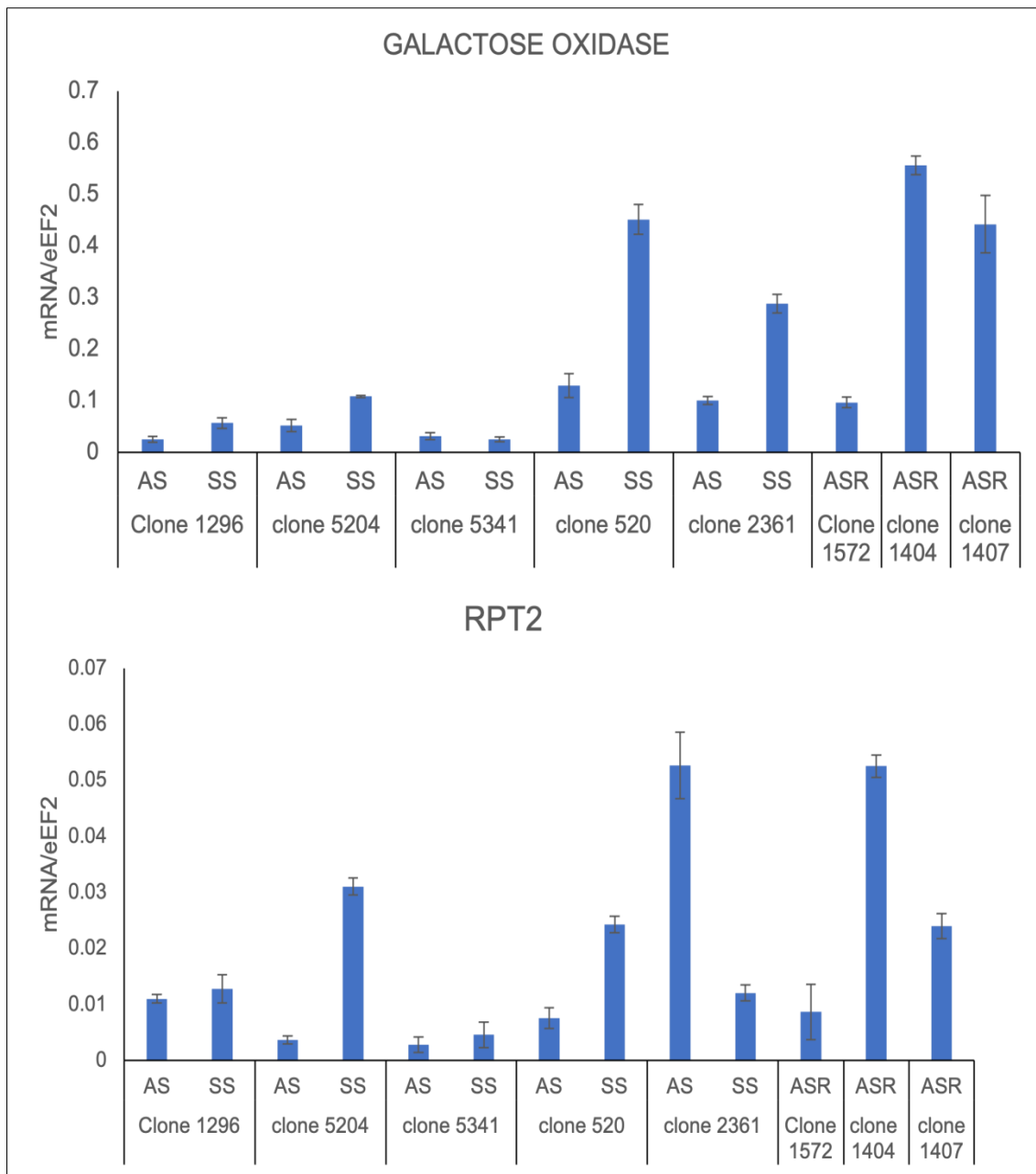


Figura 15

Petição 870220117662, de 15/12/2022, pág. 82/89



**Figura 16**



**Figura 17**

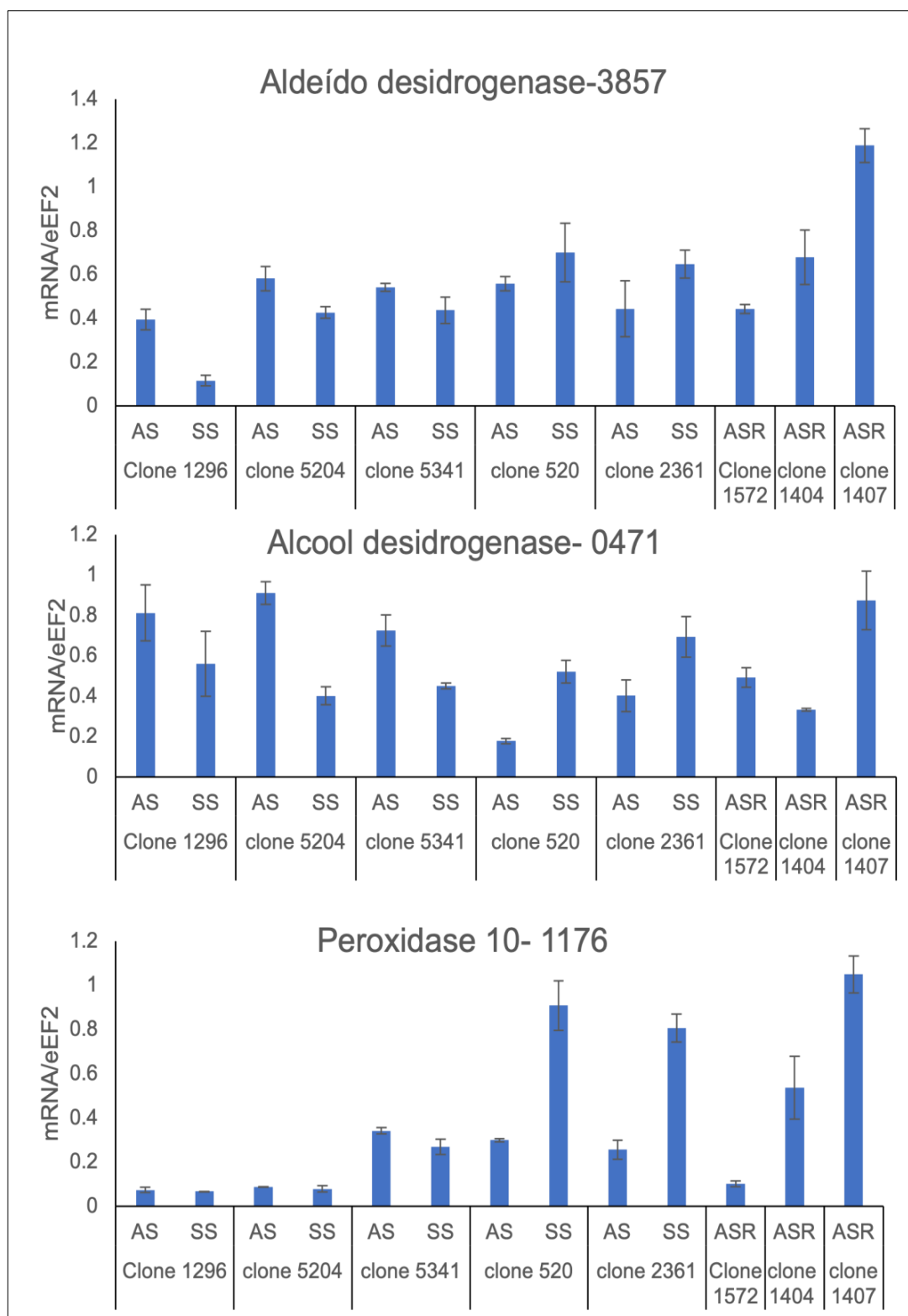
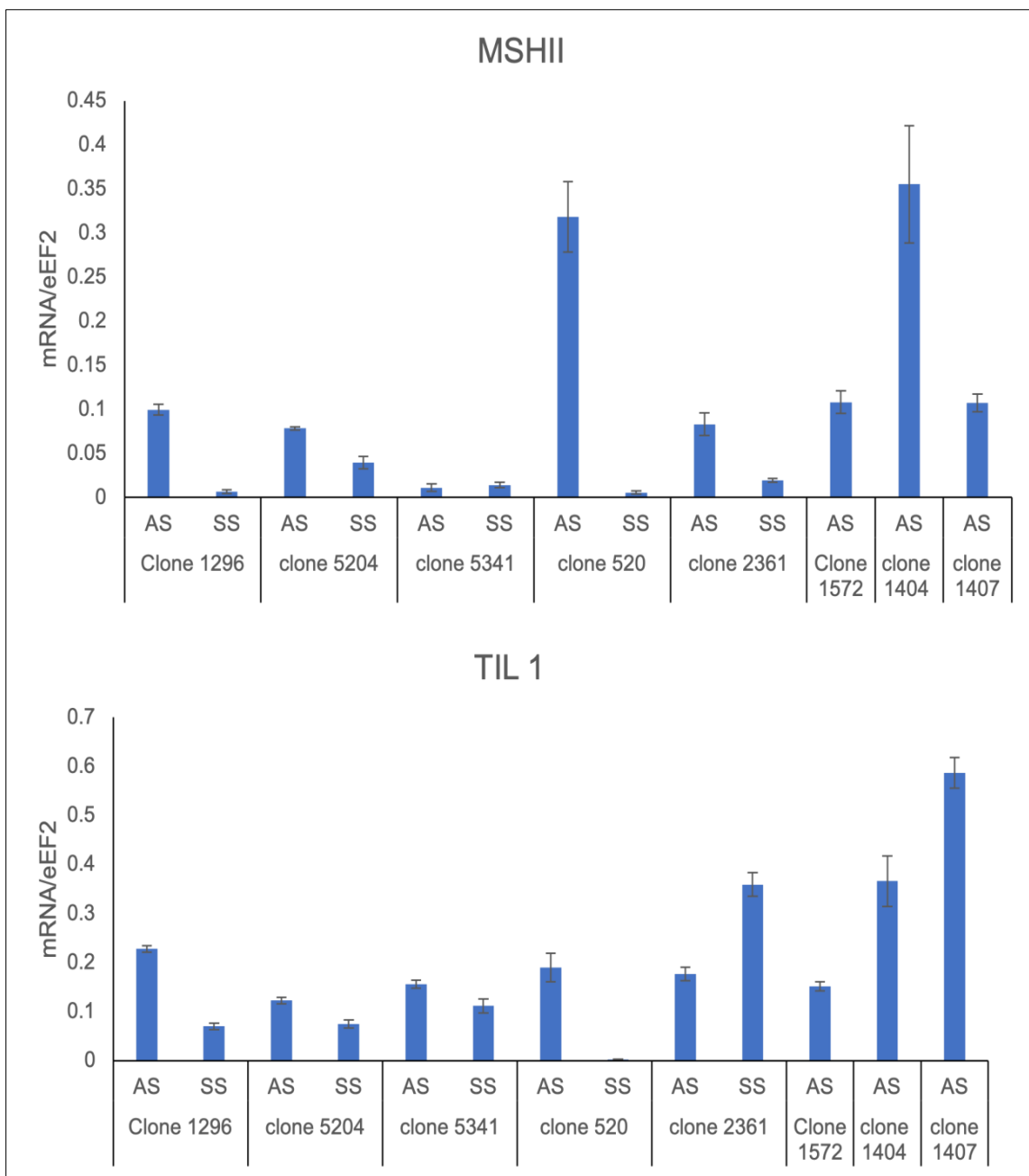
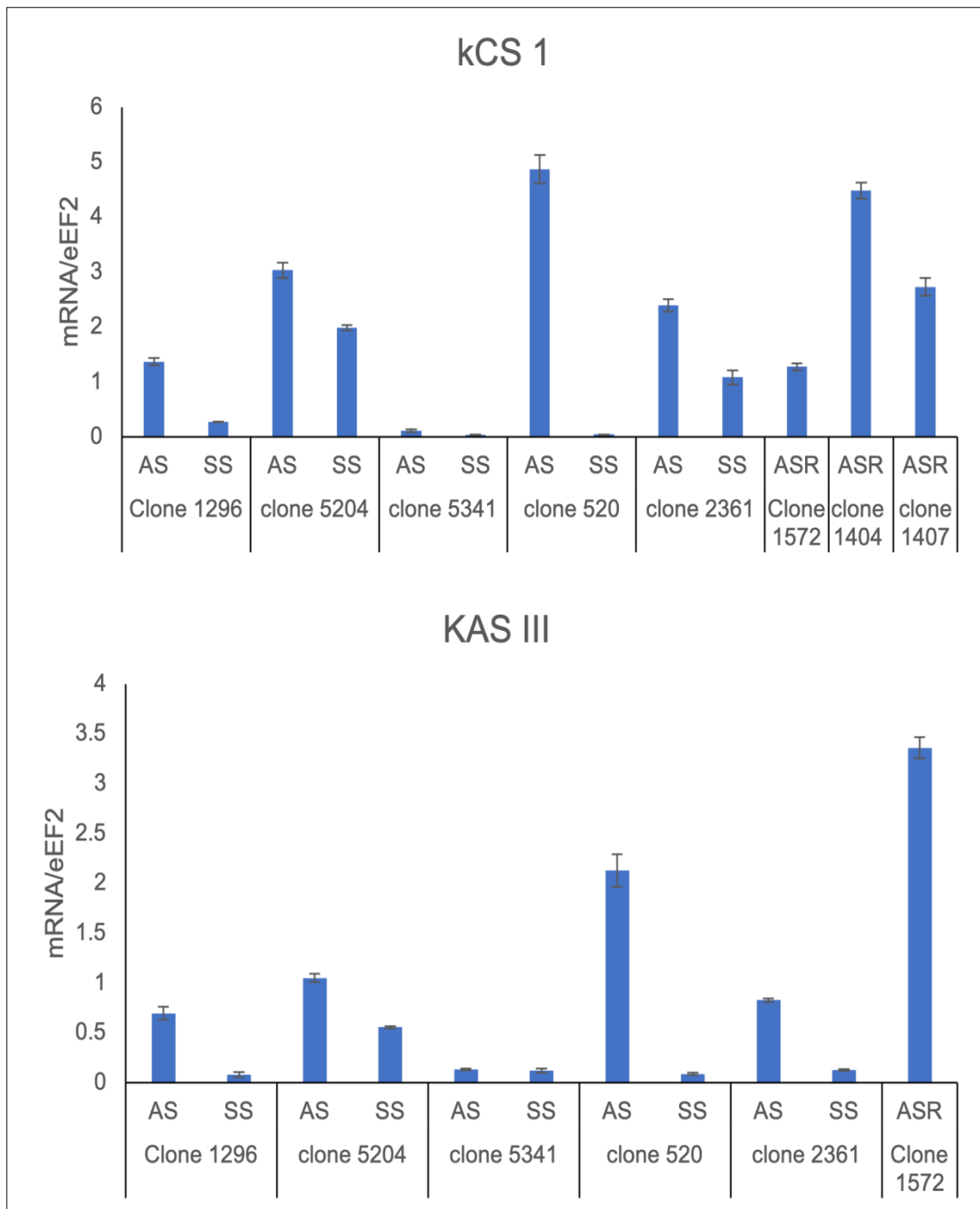


Figura 18



**Figura 19**



**Figura 20**

## Resumo

### **Oligonucleotídeos para o processo de diagnóstico molecular do distúrbio fisiológico em Eucalipto (*Eucalyptus sp.*), processo de diagnóstico e usos.**

A presente invenção refere-se a oligonucleotídeos sintéticos para o diagnóstico molecular do distúrbio fisiológico em eucaliptos (*Eucalyptus sp.*), ao processo de diagnóstico, ao uso dos oligonucleotídeos para identificação de genes diferencialmente expressos pelo distúrbio fisiológico em caules, e produção de kits diagnósticos, e ao uso dos genes diferencialmente expressos para a identificação e diagnóstico do distúrbio fisiológico em eucaliptos. Os oligonucleotídeos utilizados neste processo biotecnológico permitem diagnosticar a síndrome do distúrbio fisiológico em clones de eucalipto a campo em diferentes estádios de desenvolvimento. A invenção pode ser aplicada no setor agroflorestal, nas áreas de genética vegetal, genética molecular e melhoramento genético.

Este anexo apresenta o código de controle da listagem de sequências biológicas.

### **Código de Controle**

#### **Campo 1**

50099D568C1F84FE

#### **Campo 2**

6CBC9E4590F0A82F

#### **Outras Informações:**

- Nome do Arquivo: Listagem de sequências biológicas.txt
- Data de Geração do Código: 15/12/2022
- Hora de Geração do Código: 09:56:22
- Código de Controle:
  - Campo 1: 50099D568C1F84FE
  - Campo 2: 6CBC9E4590F0A82F