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

Identificação de genes diferencialmente expressos em árvores de *Eucalyptus grandis* susceptíveis e resistentes à *Puccinia psidii*

Guillermo Rafael Salvatierra

**Tese apresentada, para obtenção do título de Doutor em
Agronomia. Área de Concentração: Genética e
Melhoramento de Plantas**

**Piracicaba
2006**

Guillermo Rafael Salvatierra
Licenciado em Genética

**Identificação de genes diferencialmente expressos em árvores de *Eucalyptus grandis*
susceptíveis e resistentes à *Puccinia psidii***

Orientador:
Prof. Dr. **CARLOS ALBERTO LABATE**

Tese apresentada para obtenção do título de Doutor em
Agronomia. Área de Concentração: Genética e Melhoramento
de Plantas

**Piracicaba
2006**

Dados Internacionais de Catalogação na Publicação (CIP)
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - ESALQ/USP

Salvatierra, Guillermo Rafael

Identificação de genes diferencialmente expressos em árvores de *Eucalyptus grandis* susceptíveis e resistentes à *Puccinia psidii* / Guillermo Rafael Salvatierra.
-- Piracicaba, 2006.

153 p. : il.

Tese (Doutorado) -- Escola Superior de Agricultura Luiz de Queiroz, 2006.
Bibliografia.

1. Eucalipto 2. Expressão gênica 3. Ferrugem -- Doença de planta 4. Fungo fitopatogênico 5. Genes 6. Resistência genética vegetal I. Título

CDD 634.9734

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

*Aos meus pais, Norma e Diego e
à minha irmã Stella,
que definem o significado do sacrifício
e a doação sem limites,
com amor*

OFEREÇO

*A minha esposa Daniela,
meu dia e minha noite,
e ao meu filho Diego,
razão do meu presente e futuro,
com amor*

DEDICO

AGRADECIMENTOS

Ao professor Dr. Carlos Alberto Labate, pela orientação, atenção, companherismo e pela minha formação desde a graduação.

À Mônica Teresa Veneziano Labate pelos ensinamentos, amizade, alegria constante e dedicação (e paciência com o meu português!!).

Ao Dr. David H. Moon, pelo apoio em momentos críticos, dinamismo, empenho constante (e paciência com o meu inglês!!).

Aos professores da ESALQ que em cada oportunidade deram o maior empenho para fazer de mim um pesquisador.

Aos amigos e colegas de pós-graduação do Departamento de Genética da ESALQ-USP, pelos momentos de amizade e companherismo. Não citarei nomes pois foram tantos que demandaria grande número de páginas.

Aos funcionários do Departamento de Genética da ESALQ-USP, em especial ao Juliano Bragatto.

À bibliotecária Gloria pelo auxílio e esforço.

Aos grandes amigos Alejandro Toro, Elizabeth, ao pequeno amigo Franco Toro, Esteban González e Alexander Andrade pelo carinho, amizade e ajuda incondicional.

À minha equipe de trabalho Danielle Gregório Gomes, Mayra Costa da Cruz Gallo de Carvalho, Raphael T Carneiro pelo constante apoio, alegria em cada passo, mão fraterna em momentos difíceis, palavras de conforto e encorajamento, em fim, por todos os momentos compartilhados sempre estarei agradecido obrigado e gracias!. Sem esquecer a Livia Franceschini pela ajuda constante.

Aos amigos do laboratório Max Feffer: Juliana Rodrigues, Gisele C.L. Boneli, Alex Vasconcelos de Souza, Simone Bragatto, Paola F. Celedon (muito obrigao por toda a ajuda prestada), Karem G.X. Meireles, Ana Letícia Bertolo, Fernanda Salvato, Daniela D. do

Nascimento, Eduardo L.O. Camargo, Felipe Boaretto, Karina Lima, Leandro F. Sturion, Lilia P. Barbosa, Mateus Bonatto (devem conhecer as gírias do bona!!), Luis de Andrade, Maria Inês Faraldo, Gabriela Conti, Renan Carange e Camila Silva, se esqueci de alguém saibam desculpar, a minha memória não é o meu ponto forte, mas com certeza estão aqui no meu peito.

Ao CNPq e à Companhia Suzano de Papel e Celulose pelo apoio financeiro e material biológico.

E a todos que, de alguma maneira, contribuíram para a realização do presente trabalho.

VERDADE

*A porta da verdade estava aberta,
mas só deixava passar
meia pessoa de cada vez*

*Assim não era possível atingir toda a verdade
porque a meia pessoa que entrava
só trazia o perfil de meia verdade.
E sua segunda metade
voltava igualmente com meio perfil.
E os meios perfis não coincidiam.*

*Arrebentaram a porta. Derrubaram a porta.
Chegaram ao lugar luminoso
onde a verdade esplendia seus fogos.
Era dividida em metades
Diferentes uma da outra.*

*Chegou-se a discutir qual a metade mais bela
Nenhuma das duas era totalmente bela.
E carecia optar. Cada um optou conforme
seu capricho, sua ilusão, sua miopia.*

Carlos Drummond de Andrade

SUMÁRIO

RESUMO	13
ABSTRACT	15
RESUMEN	17
LISTA DE FIGURAS	19
LISTA DE TABELAS	23
LISTA DE ABREVIATURAS.....	25
1 INTRODUÇÃO.....	29
2 REVISÃO BIBLIOGRÁFICA	33
2.1 As Florestas	33
2.2 Florestas no território Nacional	34
2.3 <i>Eucalyptus</i> spp.....	37
2.4 <i>Eucalyptus</i> no Brasil.....	38
2.5 Fungos patogênicos: A ferrugem.....	41
2.6 A ferrugem do eucalipto	44
2.7 Estratégias de controle.....	45
2.8 Interação entre planta e patógeno	46
2.9 Vias de sinalização	51
2.10 A Biotecnologia e a revolução da genômica funcional	53
2.11 Transcrissoma: Cartografia para entender a função gênica.....	54
2.12 SAGE.....	55
2.13 SAGE em plantas	57
3 MATERIAL E MÉTODOS.....	61
3.1 Material Vegetal:	61
3.2 Coleta das amostras:	61
3.3 Extração do RNA total:	62
3.4 Extração de RNA: Método de Salzman et al., (1999)	62
3.5 Extração de RNA: Método de Zeng e Yang (2002)	63
3.6 Determinação espectrofotométrica de pureza e concentração de RNA	63
3.7 Determinação da pureza e concentração de RNA	64
3.8 Isolamento do mRNA.....	64
3.9 PCR e Transcrição reversa	65
3.10 Construção de bibliotecas de SAGE:	65
3.11 A síntese do cDNA: primeira fita.....	65
3.12 A síntese do cDNA: segunda fita	66
3.13 A digestão do cDNA com a endonuclease <i>Nla</i> III.....	67
3.14 Ligação dos adaptadores ao cDNA	67
3.15 Preparação das etiquetas.....	67
3.16 Precipitação com etanol das etiquetas	68
3.17 Unindo os dímeros de etiquetas.....	68
3.18 Precipitação com etanol dos dímeros de etiqueta ligados	68
3.19 Ligação dos dímeros.....	69
3.20 Otimização das reações de PCR para os dímeros.....	69
3.21 Análise do produto do PCR por meio de eletroforese em gel de poliacrilamida	70
3.22 Amplificação por PCR	70
3.23 Purificação em gel dos dímeros de 100 pb	70

3.24	Eluição do DNA do gel.....	71
3.25	Precipitação com etanol dos dímeros eluídos do gel.....	71
3.26	Clivagem dos dímeros com a endonuclease <i>Nla</i> III.....	72
3.27	Precipitação dos fragmentos de digestão.....	72
3.28	Purificação dos dímeros de etiquetas sem adaptadores mediante eletroforese em gel de poliacrilamida.....	72
3.29	Extração dos dímeros a partir do gel de poliacrilamida.....	72
3.30	Precipitação com etanol.....	72
3.31	Reação de ligação.....	73
3.32	Eletroforese em gel.....	73
3.33	Extração dos concatâmeros do gel de poliacrilamida.....	73
3.34	Precipitação.....	73
3.35	Linearização do plasmídeo pZErO [®] -1.....	73
3.36	Ligação do concatâmero ao pZErO [®] -1.....	74
3.37	Preparo do meio de cultura LB seletivo.....	74
3.38	Transformação de <i>E. coli</i> One Shot [®] TOP10 Electrocompetente.....	74
3.39	Seqüenciamento.....	75
3.40	Análise das seqüências apropriadas com Phred-Phrap.....	75
3.41	Análise das seqüências por SAGE 2000 TM	78
3.42	Classificação das etiquetas mediante o teste estatístico Z.....	81
3.43	Comparação com a base de dados Genolyptus.....	81
3.44	Comparação entre as seqüências do Genolyptus e do NCBI.....	83
3.45	Anotação das ESTs identificadas com auxílio de outros bancos de dados públicos.....	85
4	RESULTADOS E DISCUSSÃO.....	89
4.1	SAGE.....	89
4.2	Distribuição de freqüências de expressão.....	90
4.3	Análise estatística.....	92
4.4	Análise das etiquetas com o auxílio do banco de dados do Genolyptus.....	94
4.4.1	Seqüências com expressão diferencial.....	94
4.4.2	Seqüências sem expressão diferencial.....	94
4.4.3	Discussão dos resultados do Genolyptus.....	95
4.5	Anotação com outros bancos de dados.....	95
4.5.1	Categorização funcional das etiquetas.....	97
4.5.1.1	Etiquetas preferencialmente expressas na biblioteca susceptível.....	97
4.5.1.2	Genes com expressão preferencialmente expressas na biblioteca resistente.....	98
4.5.1.3	Análise das 200 etiquetas com alta expressão, porém sem diferenças estatisticamente significativas entre as duas bibliotecas.....	99
4.5.1.4	Análise comparativa dos genes diferencialmente expressos nas duas bibliotecas.....	101
4.6	Descrição dos genes em cada categoria.....	102
4.6.1	Etiquetas com expressão preferencial na biblioteca susceptível.....	102
4.6.2	Etiquetas com expressão preferencial na biblioteca resistente.....	107
4.6.5	Etiquetas sem expressão preferencial.....	112
4.7	Discussão do envolvimento de genes anotados na resposta à infecção de <i>P. psidii</i>	117
4.7.1	Componentes Estruturais e a polarização celular.....	118
4.7.2	Explosão oxidativa, resposta hipersensível, resistência sistêmica adquirida e estresse abiótico.....	118
5	CONCLUSÕES.....	123

REFERÊNCIAS	125
ANEXOS.....	139

RESUMO

Identificação de genes diferencialmente expressos em árvores de *Eucalyptus grandis* susceptíveis e resistentes à *Puccinia psidii*

A atividade florestal apresenta-se como um vetor de desenvolvimento social, ambiental e econômico no Brasil. Em 2004 as exportações ligadas ao setor renderam US\$ 7 bilhões e contribuíram com US\$ 5,5 bilhões em impostos. No mesmo ano, no Brasil, dos 6 milhões de hectares de reflorestamento comercial, mais de 50% dos hectares são ocupados por eucalipto. Esta cultura pode ser utilizada para diversas funções, desde à produção de papel e celulose, até a produção de carvão vegetal. Mas a produtividade desta cultura atualmente tem uma série de fatores limitantes. Destes fatores, um dos mais severos é uma doença denominada ferrugem, causada pelo fungo *Puccinia psidii* Winter. Este patógeno entre tanto, não é conhecido nos centros de origem do gênero *Eucalyptus*, mas, utiliza como hospedeiros vários gêneros das Mirtáceas, infectando folhas, flores e frutos em desenvolvimento. O controle da doença é normalmente realizado mediante o uso de fungicidas. Porém, a utilização de plantas resistentes é o método mais aconselhável por diversos motivos, como o baixo custo, a praticidade e o menor impacto ambiental. Neste trabalho a metodologia do SAGE - Serial Analysis of Gene Expression foi utilizada, para a análise da expressão gênica uma vez que permite detectar e quantificar, de maneira global, a expressão de genes conhecidos e desconhecidos, envolvidos no processo de resistência à ferrugem. Neste trabalho analisou-se a expressão do conjunto de genes diferencialmente expressos e 200 genes sem diferenças no padrão de expressão, em *Eucalyptus grandis* infectado com *Puccinia psidii*. Utilizou-se para esta análise, indivíduos susceptíveis e resistentes à infecção por *Puccinia psidii* Winter. A população amostrada constituída por meios irmãos, apresentava segregação para o caráter de resistência à ferrugem. Os resultados produzidos permitiram identificar 421 genes com expressão diferencial ($p \leq 0,05$) nos dois fenótipos. Vários destes genes estavam relacionados a diversos mecanismos de defesa da planta, tais como genes envolvidos na formação de barreiras físicas e na resposta hipersensível, resistência sistêmica adquirida, bem como genes relacionados à polarização celular e resposta ao estresse oxidativo. Os dados em conjunto, corroboram a idéia de que são vários os mecanismos que atuam simultaneamente produzindo o fenótipo resistente. Estes genes apresentam-se para serem utilizados futuramente como marcadores moleculares na seleção de indivíduos resistentes, em cruzamentos naturais ou dirigidos e como fonte de genes para experimentos de transformação em estudos mais profundos de fitopatogênese.

Palavras-chave: SAGE; *Eucalyptus grandis*; *Puccinia psidii*; ferrugem; perfil transcricional

ABSTRACT

Identification of differentially expressed genes in susceptible and resistant *Eucalyptus grandis* trees exposed to *Puccinia psidii*

In Brazil the forest activities represent a vector for social, ambiental and economic development. In 2004 the exportations produced US\$ 7 billions and contributed with US\$ 5,5 billion in taxes. In the same year, in Brasil, of 6 million hectares of commercial reforestation, around 50% is occupied by eucalyptus. This crop can be used as the raw material for several industrial processes, from cellulose and paper production to vegetable charcoal production. However, eucalyptus productivity has several limiting factors. One of the most severe diseases is called rust, caused by the fungus *Puccinia psidii* Winter. This pathogen is unknown in the *Eucalyptus* center of origin; however, it has several Myrtaceae genera as its host, infecting leaves, flowers and developing fruits. Rust management normally involves the use of fungicides, but the use of resistant plants is a much better alternative due the lower costs and minimal ambient impact. In the current study SAGE - Serial Analysis of Gene Expression was used for gene expression analysis because it is a quantitative genome-wide method and that obtains expression profiles from known and unknown genes involved in the rust resistant process. This thesis analyzed the expression profile of the differentially expressed genes and 200 genes without significant statistical differences in their expression patterns, in *Eucalyptus grandis* infected with *Puccinia psidii*. For this analysis we used susceptible and resistant individuals to *Puccinia psidii* Winter infection taken from a population half siblings, which showed segregation for rust resistance. The results indicated 421 genes with differential expression ($p \leq 0.05$), 239 were preferentially expressed in the susceptible library and 232 in the resistant. Several of these genes were associated with plant defense mechanisms, such as genes involved in formation of physical barriers, hypersensitive response, systemic acquired resistance, genes related to cellular polarization and oxidative stress responses. All the data support the idea that not one but several mechanisms are acting at the same time producing the resistant phenotype. These genes are future candidates for molecular markers useful in the selection of resistant individuals in breeding programs and a source of genes for transformation experiments or as molecular tools to dissect the phytopathogenic process.

Key-words: SAGE; *Eucalyptus grandis*; *Puccinia psidii*; rust; transcriptional profile

RESUMEN

Identificación de los genes de expresión diferencial en árboles de *Eucalyptus grandis* susceptibles y resistentes a *Puccinia psidii*

La actividad forestal representa en Brasil un vector de desarrollo social, ambiental e económico. Las exportaciones produjeron US\$ 7 mil millones y contribuyeron con US\$ 5,5 mil millones en impuestos. Actualmente, en Brasil, más del 50% de los 6 millones de hectáreas de reforestamiento comercial, están siendo ocupadas por eucalipto. Este cultivo puede ser utilizado para múltiples funciones, desde la producción de papel y celulosa, hasta la producción de carbón vegetal. Pero actualmente la producción de este cultivo se encuentra afectada una serie de factores limitantes. De estos factores, uno de los más severos, es una enfermedad llamada roya que es una micosis es causada por el hongo *Puccinia psidii* Winter. Este patógeno, no es conocido en el centro de origen del género *Eucalyptus*. Sin embargo, este utiliza como hospederos varios géneros de la familia de las mirtáceas, infectando hojas, flores y frutos en crecimiento. El control de la micosis normalmente es realizado mediante el uso de fungicidas. Pero la utilización de plantas resistentes es el método más aconsejable, por varios motivos, como lo son el bajo costo, la practicidad y el menor impacto ambiental. En el presente estudio fue utilizada la metodología de SAGE - *Serial Analysis of Gene Expression*, para el análisis de la expresión génica, ya que éste permite detectar y cuantificar, de una manera global, la expresión de genes conocidos y desconocidos, que están involucrados en el proceso de resistencia a la roya. En la presente tesis se analizó el conjunto de genes que se expresan diferencialmente y 200 genes que no presentan diferencias en el padrón de expresión, de *Eucalyptus grandis* infectado con *Puccinia psidii*. Para el presente análisis se utilizaron individuos susceptibles y resistentes a la infección realizada por *Puccinia psidii* Winter. La población que fue muestreada estaba constituida por medios hermanos, los cuales presentaban el carácter de resistencia a la roya, segregando através de ellos. Los resultados obtenidos permitieron identificar 421 genes con una expresión diferencial en los dos genótipos ($p \leq 0,05$). Varios de estos genes se encontraban relacionados con diversos mecanismos de defensa de las plantas, tales como los genes involucrados en la formación de barreras físicas y en la respuesta hipersensible, resistencia sistemáticamente adquirida, así como genes relacionados con la polarización celular y los de respuesta al estrés oxidativo. Los datos en conjunto apoyan la idea de que son varios los mecanismos que actúan simultáneamente produciendo el fenotipo resistente. Estos genes se presentan como una buena opción para ser utilizados futuramente como marcadores moleculares en la selección de individuos resistentes, en cruzamientos naturales o dirigidos y también como una fuente de genes para experimentos de transformación, en estudios mas profundos de fitopatogénesis.

Palabras claves: SAGE; *Eucalyptus grandis*; *Puccinia psidii*; roya; perfil transcriptional

LISTA DE FIGURAS

- Figura 1 - Distribuição mundial das áreas plantadas por gênero de árvores (Fonte: FAO. Disponível em: <<http://www.fao.org>>).34
- Figura 2 – Distribuição das variedades florestais e áreas plantadas correspondentes no Brasil. Fonte: SBS (2005).36
- Figura 3 - Distribuição da família das Mirtáceas no mundo.....38
- Figura 4 – Representatividade das espécies de eucalipto mais plantadas no Brasil. (fonte: BRACELPA, 2004b).....39
- Figura 5 - Esquema do desenvolvimento da ferrugem no interior do tecido vegetal hospedeiro. (fonte: Kaiserslautern, 2005).....43
- Figura 6 – A área escura marca as regiões de alto risco para ocorrência da ferrugem causada pelo fungo *Puccinia psidii*-Winter no Neotrópico (BOOTH; JOVANOVIC, 2000).45
- Figura 7 – (a) e (b) Representação das interações possíveis entre os produtos gênicos dos genes de avirulência (Avr), nos patógenos e resistência (R), no hospedeiro: compatível, permite a ocorrência da doença e incompatível, que não permite o crescimento do patógeno. (1) Primeiro tipo de interação incompatível, que determina resistência pela interação direta dos produtos gênicos dos genes Avr e R. (2) Segundo tipo de interação incompatível que determina resistência pela interação indireta dos produtos gênicos dos genes Avr e R, denominada de “hipótese de guarda”, ou seja a molécula Avr atinge o seu alvo no hospedeiro, porém, a infecção não progride pela presença do gene R. Fonte: Hammond-Kosack e Parker (2003).50
- Figura 8 – Fórmula para cálculo da concentração de RNA. A_{260nm} absorvância a 260nm, Fc: fator de conversão, que indica a equivalência de 1 unidade de absorvância no comprimento de onda de 260nm correspondente a 40 μg RNA/ml. Fd é o fator de diluição da amostra. Neste caso 200 (porque são utilizados 4 μl da amostra em 800 μl de água). O

valor final é dividido por 1000, para expresarmos a unidade de concentração de RNA em $\mu\text{g}/\mu\text{l}$. Fonte: Sambrook et al. (1989b)	64
Figura 9 – Formula de cálculo da qualidade média - q . Onde p representa a probabilidade de erro estimado para as bases analisadas (EWING; GREEN, 1998)	76
Figura 10 – (a) Cromatograma representativo de uma seqüência que foi rejeitada. (b) Cromatograma representativo de uma seqüência com alta qualidade.....	77
Figura 11 – Arquivo de saída do software <i>phred</i> que permite identificar seqüências de DNA de alta qualidade.....	78
Figura 12 – Interface do SAGE2000 utilizada na comparação entre as duas bibliotecas de <i>Eucalyptus grandis</i> : susceptível e resistente à ferrugem.....	79
Figura 13 – Representação da distribuição das etiquetas geradas pelo <i>software</i> SAGE2000, o agrupamento por classes, os componentes que são agrupados e a abundância total desta classe.....	80
Figura 14 – Número de etiquetas diferentes que foram encontradas dentro da biblioteca resistente representando os genes que estão sendo expressos.....	80
Figura 15 – Interface gráfica da base de dados do Genolyptus, que inclui a ferramenta BLAST N de pesquisa de todas as bases presentes no Genolyptus, juntamente com as seqüências de corte da endonuclease e a seqüência da etiqueta de 10 bases.....	82
Figura 16 – Seqüência de um determinado contig pertencente ao Genolyptus que apresenta todas as características necessárias para ser selecionado: (amarelo) seqüência da etiqueta e de (azul) reconhecimento de corte da endonuclease em posição proximal à extremidade 3', cauda poli A e ausência de seqüências da endonuclease entre as últimas duas.....	83
Figura 17 – Interface gráfica da análise de BLAST, da base de dados do NCBI, de uma determinada seqüência resgatada do Genolyptus.....	84

Figura 18 – <i>ORF finder</i> : ferramenta de busca de quadros de leitura abertos codantes pertencente ao site do NCBI.....	84
Figura 19 – Outros bancos que foram pesquisados para a anotação das seqüências provenientes do Genolyptus. a) Base de dados do Pfam. b) Base de dados do COG. c) Base de dados do KEGG. d) Base de dados do CDD. e) Base de dados do SMART.....	87
Figura 20 – Representação de etiquetas seqüenciadas e como estas se encontram distribuídas nas duas bibliotecas.....	89
Figura 21 – Etiquetas totais por biblioteca e as etiquetas que foram identificadas como diferentes umas de outras pelo <i>software</i> SAGE2000 dentro de cada biblioteca.....	90
Figura 22 – Distribuição das etiquetas com expressão diferencial ($p \leq 0,05$) nas duas bibliotecas.....	93
Figura 23 –Número de contigs expressos em porcentagem resgatadas do banco de dados do Genolyptus a partir do total de etiquetas com expressão preferencial ou única na biblioteca susceptível, a 5% de probabilidade.....	94
Figura 24 – Número de contigs expressos em porcentagem resgatados do banco de dados do Genolyptus a partir do total de etiquetas com expressão preferencial ou única na biblioteca resistente, a 5% de probabilidade.....	94
Figura 25 – Número de contigs expressos em porcentagem resgatadas do total de 200 etiquetas com alta expressão nas duas bibliotecas, porém sem diferenças estatisticamente significativa.....	95
Figura 26 – Figura 26 – Distribuição das etiquetas em categorias funcionais (C1-C8) dos contigs para o grupo de etiquetas com expressão preferencial ou única na biblioteca susceptível, a $p \leq 0,05$ pelo teste estatístico Z.....	98

Figura 27 – Distribuição das etiquetas em categorias (C1-C8) funcionais pertencentes ao grupo de etiquetas com expressão preferencial ou única na biblioteca resistente, a $p \leq 0,05$ pelo teste estatístico Z.....99

Figura 28 – Distribuição em categorias (C1-C8) funcionais, dos 200 contigs mais frequentes pertencentes ao grupo de etiquetas com expressões não estatisticamente significativas, pelo teste Z.....100

LISTA DE TABELAS

Tabela 1 - Área nacional plantada, dividida por Estados no ano 2001 (fonte: disponível em http://www.manejoflorestal.org/noticia_print.cfm?id=132267 . Acesso em 12 dez. 2005).....	35
Tabela 2– Alíquotas das reações de PCR que foram utilizadas para a otimização das reações de amplificação.....	69
Tabela 3 – Resumo da distribuição de frequências das etiquetas geradas pela metodologia SAGE.....	91
Tabela 4 – Expressão diferencial das etiquetas nas duas bibliotecas (incluindo os “singletons”). Os dados foram analisados pelo teste Z ($p \leq 0,05$).....	93
Tabela 5 – Etiquetas com expressão preferencial na biblioteca susceptível à 5% de probabilidade.....	103
Tabela 6 – Etiquetas com expressão preferencial na biblioteca resistente à 5% de probabilidade.....	108
Tabela 7 —Etiquetas com expressão preferencial.....	113

LISTA DE ABREVIATURAS

- BLAST – Basic Local Alignment Search Tool. Algoritmo que executa comparações de sequências em pares procurando regiões de similaridade local.
- cDNA – Sequência de DNA gerada artificialmente por transcrição reversa do mRNA. Representa aproximadamente os componentes codificantes da região do DNA genômico que produziu o mRNA.
- °C – medida de temperatura, graus centígrados.
- CDD - Conserved Domain Database (Base de dados de domínios conservados).
- Contig – é um conjunto de *reads* cujas sequências se sobrepõem por similaridade de sequência.
- DEPC – dietilpirocarbonato.
- Ditag – Duas tags ou etiquetas, de aproximadamente 10 pb, unidas pelos extremos.
- DNA – ácido desoxiribonucleico.
- DNA genômico – Sequência de DNA produzida por sequenciamento genômico. Contém sequências de DNA codificantes e não-codificantes.
- dNTP – desoxi ribonucleotídeos (A = adenina, C = citosina, G = guanina e T = timidina)
- dT – desoxitimidina.
- EC – Número do código enzimático
- E. coli* – *Escherichia coli*.
- ESTs – expressed sequence tags. Sequências curtas de cDNA, completas ou não, que são sintetizada a partir de mRNA. As ESTs são utilizadas para identificação rápida de genes e não abrangem a sequência total de um gene.
- Exons – Parte de um determinado gene, expresso na proteína final desse.
- g* – Unidade de medida de velocidade de uma centrífuga.
- H₂O₂ – peróxido de hidrogênio.
- Introns – Sequência do interior de um determinado gene eucarioto que não é representada na proteína final desse.
- LiCl – cloreto de lítio.
- LTP – *Lipid Transfer Proteins* (Proteínas de transferência de lipídeos).
- μl – unidade de capacidade ou volume, correspondente a uma milionésima parte de um litro de solvente.

M – molar. Unidade de concentração que equivale ao n vezes o peso molecular do soluto em gramas num litro de solvente.

mg – miligrama. Medida de peso equivalente a milésima parte de uma grama.

MgCl₂ – cloreto de magnésio.

ml – medida de volume ou capacidade, correspondente a uma milésima parte de um litro.

mM – milimolar. Unidade de concentração que equivale a uma milésima parte de um molar.

MPC – “*Magnetic Particle Concentrator*” (Concentrador de partículas magnéticas).

mRNA – RNA mensageiro, é produzido pelo processo de transcrição, complementar à sequência de DNA genômico, salvo, as regiões que foram editadas pela célula para remover introns no caso dos organismos eucariotos.

NaCl – Cloreto de sódio.

NaOH – Hidróxido de sódio.

ng – nanograma. Medida de peso, equivalente a 1×10^{-9} grama.

nm – nanômetro. Medida de comprimento, equivalente a 1×10^{-9} metro.

ORF – *Open Reading Frame* ou quadro de leitura aberto.

ORF Finder - É uma ferramenta utilizada para localizar o quadro de leitura aberto de uma determinada sequência de DNA, e assim encontrar as regiões codificadas e a proteína ou domínio que codifica.

pb – pares de bases

PCR – “*Polymerase chain reaction*” (Reação em cadeia da DNA polimerase).

PIB - Produto Interno Bruto.

Primer – pequeno oligonucleotídeo que serve de iniciador da reação de síntese de uma determinada sequência de nucleotídeos.

p/v – medida de concentração onde a quantidade de soluto é medida em alguma unidade de peso e a de solvente em alguma unidade de volume.

Read – sequência de nucleotídeos obtida de um processo de sequenciamento, procedente de uma amostra de DNA (clone).

RNA – ácido ribonucléico.

RNA poli A – RNA com cauda poli A. A cauda poliadenilada é um processo de maturação das moléculas de RNA que se apresenta nas células eucarióticas.

RNases: Enzimas degradadoras de RNA

rpm – revoluções por minuto. Unidade de velocidade circular.

SAGE – “*Serial Analysis of Gene Expression*” (Análise serial de expressão gênica).

SDS –sulfato de sódio dodecil.

ton s/cc – Unidade de produção que expressa toneladas de madeira sem casca.

v/v - medida de concentração onde a quantidade de soluto e solvente são medidos em alguma unidade de volume.

Valor E – O valor E de um alinhamento indica o número de alinhamentos que pode ser encontrado com uma pontuação maior ou igual à pontuação de alinhamento observada.

1 INTRODUÇÃO

A produção florestal brasileira representa uma fonte de recursos naturais renováveis. No Brasil, com um Produto Interno Bruto – PIB no ano 2004 de US\$ 605 bilhões, a exploração florestal e sua cadeia de produção, industrialização e comercialização geraram, uma receita de mais de US\$ 28 bilhões, o que representou aproximadamente 4,5% do PIB nacional (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE, 2005). No mesmo ano, o setor florestal se constituiu no segundo item das exportações do agronegócio brasileiro com US\$ 7 bilhões e com uma arrecadação de US\$ 5,5 bilhões em impostos (LEITE, 2005; SOCIEDADE DE SILVICULTURA – SBS, 2005). Propiciou a geração de 1,6 milhões de empregos diretos e mais de 6,7 milhões diretos e indiretos (EMBRAPA, 2005).

Dos 850 milhões de hectares do território nacional, aproximadamente 550 milhões deles são cobertos por florestas nativas e 6 milhões, com florestas plantadas (EMBRAPA, 2005). O Eucalipto, como espécie exótica, começou a ser cultivado em larga escala no Brasil em 1903 (O EUCALIPTO..., 2001) e atualmente é plantado em mais de três milhões de hectares (EMBRAPA, 2005). Esta cultura pode ser utilizada para diversos fins, como na produção de carvão, extração de óleo vegetal, confecções de móveis, e na produção de papel e celulose (A MADEIRA..., 2001). Do volume total de madeira consumido no ano 2001, 19.128.924 ton s/cc - toneladas sem casca, 14.002.159 ton s/cc correspondiam a eucalipto, apresentando-se, portanto, como a cultura florestal mais importante para o Brasil (BRACELPA, 2003).

A cultura do eucalipto apresenta várias vantagens que a torna importante no reflorestamento comercial, como o seu rápido crescimento, a diversidade inerente ao gênero, adaptabilidade, alta produtividade e baixo custo (HO et al., 1998; ESPÉCIES..., 2001; MELHORAMENTO..., 2001). O melhoramento genético, dessa cultura permitiu o aumento da produtividade num menor período de tempo, levando à obtenção de espécies ou linhagens geneticamente puras (INSTITUTO DE PESQUISA E ESTUDOS FLORESTAIS – IPEF, 2005), exploração do vigor híbrido como resultado de cruzamentos interespecíficos e a seleção de árvores de características silviculturais superiores (BERTOLUCCI et al., 1995; MELHORAMENTO..., 2001).

No entanto, a produtividade desta cultura se vê restringida por uma doença, a ferrugem. Esta doença é causada pelo fungo *Puccinia psidii* Winter e apresenta-se como uma das doenças mais severas desta cultura, no Brasil (JUNGHANS; ALFENAS; MAFFIA, 2003). Em outros

ambientes subtropicais torna-se também um risco potencial para outras Mirtáceas que são cultivadas (RAYACHHETRY et al., 2001). Este patógeno é endêmico na América do Sul e Central e não é conhecido nos centros de origem do gênero *Eucalyptus*. Apareentemente este fungo apareceu originalmente nas Mirtáceas da América do Sul sendo altamente infeccioso em algumas espécies economicamente importantes como goiaba (*Psidium guajava*), e pimentão (*Pimento officinalis*) (COUTINHO et al., 1998; RAYACHHETRY et al., 2001). Porém, como este patógeno tem a capacidade de infectar Mirtáceas nativas e exóticas, adaptou-se também à cultura do eucalipto (TELECHEA et al., 2003; BOOTH et al., 2000; RAYACHHETRY et al., 2001; COUTINHO et al., 1998).

O controle da doença, até o momento, vem sendo realizado de várias formas, pela utilização de fungicidas, colheita de árvores para aproveitamento da rebrota em épocas desfavoráveis à ferrugem e pela utilização de plantas resistentes (AMBIENTEBRASIL, 2005). A utilização desta última medida é a mais aconselhável por várias razões, tais como baixo custo, praticidade e menor impacto ambiental, pela redução de fungicidas (DE CARVALHO et al., 1998; ALFENAS, 1989).

No eucalipto, a seleção de progênies ou clones resistentes pode ser realizada a partir de infecções naturais no campo, em áreas onde a doença é severa ou endêmica ou por meio de uma infecção artificial, em condições controladas (DE CARVALHO et al., 1998). A seleção destes indivíduos é importantíssima para os programas de melhoramento genético. Mas pouco se sabe sobre a genética da interação deste patógeno (*Puccinia psidii* Winter) e o gênero (*Eucalyptus* spp.), genes envolvidos, herança da resistência (JUNGHANS et al., 2003), vias de resposta à doenças ativadas durante o ataque deste patógeno e do ambiente genético no qual se desenvolve o processo (JUNGHANS, 2000).

A biotecnologia juntamente com as técnicas de genômica, transcrissômica, proteômica e bioinformática apresenta um imenso potencial no setor florestal e está revolucionando o conhecimento científico e o desenvolvimento econômico nesta área. Esta interação entre áreas permitirá a descoberta de novos genes, surgimento de novas perspectivas sobre a estrutura e função de genes e proteínas e a formulação de novas hipóteses sobre o funcionamento de respostas multigênicas complexas. Métodos globais para análises de expressão gênica como SAGE (“Serial Analysis of Gene Expression” ou Análise em Série de Expressão Gênica) e microarrays (de cDNA e de oligonucleotídeos), quantitativos como esse primeiro, representam

um avanço significativo no entendimento das bases moleculares envolvidas nas respostas do mecanismo de defesa das plantas a patógenos.

A metodologia SAGE tem a capacidade de detectar e quantificar a expressão de transcritos conhecidos e desconhecidos (VELCULESCU et al., 1995). Esta metodologia está baseada em dois princípios fundamentais. O primeiro está fundamentado na propriedade de que, seqüências curtas (etiquetas de 10 bases) de cDNA possuem informação suficiente para identificar os transcritos originais, e o segundo é que a concatenação destas seqüências curtas aumenta o número de mRNA representados por clone e portanto, a confiabilidade na identificação de transcritos. O pequeno tamanho das etiquetas diminui o custo do seqüenciamento e permite o seqüenciamento de etiquetas em larga escala. Em consequência, aumenta-se a confiabilidade estatística na identificação de transcritos (BOHELER; STERN, 2003).

Essa metodologia SAGE, portanto, foi utilizada para determinar, avaliar e quantificar o conjunto de genes envolvidos no processo de infecção entre o fungo *Puccinia psidii* e as plantas de *Eucalyptus grandis*. Para isto foram utilizadas amostras de tecido foliar de *Eucalyptus grandis* resistentes e susceptíveis à ferrugem, causada pelo fungo *Puccinia psidii*.

Para esta abordagem, considerando que as diferenças no número de etiquetas existentes entre as duas bibliotecas não são devidas ao acaso, formularam-se duas hipóteses:

1. A resistência é controlada por uma série de genes de resposta às doenças. Estes genes podem expressar-se nas plantas resistentes e não nas plantas susceptíveis ou podem expressar-se em ambas, porém, podem ter níveis de expressão diferentes, provavelmente por estarem regulados de forma diferente.
2. Este gene ou genes desencadeiam uma resposta de defesa na planta que provoca a ativação em cadeia, de vários genes de defesa.

Com o objetivo de quantificar o perfil transcricional de folhas de *Eucalyptus grandis* infectadas com o fungo patogênico *Puccinia psidii*, identificar genes diferencialmente expressos nos indivíduos resistentes e susceptíveis e genes potencialmente envolvidos na resposta de defesa, duas bibliotecas SAGE foram construídas. Uma de plantas resistentes e a outra de susceptíveis à ferrugem, infectadas com *Puccinia psidii* Winter. Nesse trabalho os dados do perfil transcricional de resposta dos indivíduos resistentes frente à ferrugem causada por *Puccinia psidii* Winter, estão apresentados. Além disso, os genes caracterizados são excelentes candidatos para

futuros experimentos de transformação, assim como bons candidatos para estudos de marcadores em futuros programas de melhoramento.

2 REVISÃO BIBLIOGRÁFICA

2.1 As Florestas

As florestas são áreas com alta densidade de árvores. Alguns dados demonstram que as florestas ocupam aproximadamente 30% da superfície terrestre (WIKIPEDIA, 2005). Segundo a FAO, as áreas florestais ocupam aproximadamente quatro bilhões de hectares no mundo todo. Os dez países com maiores áreas florestais concentram dois terços do total. A Rússia é o país onde ainda há mais florestas, com 809 milhões de hectares no ano 2005. O Brasil está em segundo lugar, com 478 milhões de hectares de áreas florestais (FAO, 2005). As florestas podem ser divididas em duas grandes categorias: naturais e artificiais. Estas últimas são aquelas implantadas com objetivos específicos e podem ser formadas por espécies nativas ou exóticas. Este é o tipo de floresta preferido para o uso em processos que se beneficiam da uniformidade da madeira produzida (WIKIPEDIA, 2005). Basicamente, deste último tipo de florestas, a atividade florestal mundial produz aproximadamente 300 bilhões de dólares ao ano (FAO, 2005).

A Silvicultura é a ciência que se ocupa do cuidado, aproveitamento e manutenção racional das florestas, naturais e implantadas, em função do interesse ecológico e científico, satisfazendo as necessidades do mercado (AMBIENTEBRASIL, 2005). A Silvicultura de vanguarda engloba tipicamente o uso de viveiros melhorados e a aplicação de técnicas como a preparação de solos, o uso controlado de fertilizantes e clones (FAO, 2005).

As florestas podem ser classificadas pelas suas características climáticas e ecológicas em florestas temperadas, tropicais e boreais (FAO, 2005), sendo as florestas temperadas e tropicais, as mais importantes para a produção de madeira. Sabe-se também que o solo e o clima são fatores que têm um elevado impacto no crescimento e produção das florestas. A umidade dos trópicos e os solos mais férteis permitem um crescimento mais rápido. Por exemplo, lugares de clima tropical como a Índia, freqüentemente apresentam uma produção de florestas de 4-8 m³ por hectare por ano. Isto devido parcialmente aos períodos de seca combinado com solos pobres. Espécies como *Gmelina arborea* e algumas de *Eucalyptus* são muito sensíveis ao solo onde são cultivadas. *Pinus* spp., por outro lado, tolera geralmente as condições adversas com mais facilidade e é mais flexível quanto ao tipo de solo (FAO, 2005).

Em geral, as principais espécies de crescimento rápido que possuem um curto período de rotação, encontram-se dentro dos gêneros *Eucalyptus* e *Acacia*. Neste caso, a idade de rotação é

inferior a 15 anos. Por outro lado, *Pinus* e outras coníferas são as principais espécies plantadas em regiões boreais e temperadas. Estas apresentam um período de rotação médio entre 20 e 30 anos. Entretanto, períodos de rotação menores podem ser aplicados para produção de polpa (Figura 1). A silvicultura e o melhoramento podem aumentar a produção e o rendimento de espécies de interesse como o *Eucalyptus grandis* e *E. urophylla* no Brasil e *Pinus radiata*, em alguns países do hemisfério Sul, são excelentes exemplos (FAO, 2005).

Figura 1 - Distribuição mundial das áreas plantadas por gênero de árvores (Fonte: FAO. Disponível em: <<http://www.fao.org>>)

2.2 Florestas no território Nacional

Dos 850 milhões de hectares do território nacional, aproximadamente 478 milhões de hectares são cobertos por florestas (FAO, 2005) e aproximadamente 6 milhões de hectares são florestas plantadas (Tabela 1). Desse total de florestas nativas, aproximadamente dois terços são

formados pela Floresta Amazônica, e o restante pelo Cerrado, Caatinga, Mata Atlântica e seus ecossistemas associados, totalizando assim, 14,5% da área florestal mundial (EMBRAPA, 2005). Das florestas plantadas, mais de 3 milhões de hectares pertencem à cultura do eucalipto, quase 2 milhões ao *Pinus*, e o restante, é subdividido entre outras espécies (seringueira, teca, etc) (VALVERDE, 2005) em detalhes na Figura 2.

Tabela 1 - Área nacional plantada, dividida por Estados no ano 2001 (fonte: disponível em http://www.manejoflorestal.org/noticia_print.cfm?id=132267. Acesso em 12 dez. 2005)

FLORESTAS PLANTADAS NO BRASIL (área em milhares de hectares)	
Minas	1.678
São Paulo	776
Paraná	672
Bahia	451
Santa Catarina	359
Rio Grande do Sul	252
Espírito Santo	152
Outros	990
TOTAL	5.000

Figura 2 – Distribuição das variedades florestais e áreas plantadas correspondentes no Brasil. Fonte: SBS (2005)

Segundo o Ministério do Meio Ambiente, o Brasil é o maior consumidor mundial de produtos de origem florestal. Setores estratégicos da economia brasileira, como a siderurgia, a indústria de papéis e embalagens, e a construção civil, são altamente dependentes do setor florestal (BRASIL - MINISTÉRIO DO MEIO AMBIENTE, 2005).

O setor florestal é o segundo item das exportações, produzindo US\$ 7 bilhões em exportações e US\$ 5,5 bilhões em impostos. Torna-se assim, a cultura arbórea economicamente mais importante neste país (LEITE, 2005; VALVERDE, 2005). O perfil das exportações no ano 2002 mostrou que a celulose, o papel e a madeira serrada foram os principais itens exportados (IBAMA - INSTITUTO BRASILEIRO DO MEIO AMBIENTE E DOS RECURSOS NATURAIS RENOVÁVEIS, 2002).

Em 2000, a produção brasileira de madeira roliça foi a terceira maior do mundo, com um volume de aproximadamente 300 milhões de m³. A produção nacional foi menor apenas que a dos Estados Unidos (da ordem de 400 milhões de m³) e da China (aproximadamente 320 milhões de m³). Observou-se também que o setor de papel e celulose consumiu aproximadamente 11% da madeira produzida naquele ano; para a produção de carvão vegetal 15%; madeira processada mecanicamente como serrados, laminados, compensados e painéis 20%; e energia industrial aproximadamente 10%. Finalmente, o segmento de lenha, para fins energéticos variados, como

consumo doméstico e secagem de grãos ficaram em torno de 44% do total de madeira produzida naquele ano (IBAMA, 2002).

Em outro âmbito, a produção brasileira de celulose chegou a 9 milhões de toneladas no ano de 2003, registrando um crescimento do setor de 12,2% sobre o ano anterior. Isto permitiu que a produção de papel chegasse na casa dos 7,9 milhões de toneladas no mesmo ano (BRACELPA, 2003). Já, segundo o informe da BRACELPA, de 2004, a produção de celulose foi da ordem de 9,6 milhões de toneladas e a de papel de 8,5 milhões de toneladas. Isto posiciona o Brasil como o sétimo produtor de celulose de todos os tipos, o primeiro produtor de celulose de fibra curta do mercado e décimo primeiro produtor de papel. Este setor da produção aportou, no ano de 2004, R\$ 2,1 bilhões em impostos, US\$ 3,5 bilhões em exportações, gerando 108 mil empregos diretos e atualmente representa 1,4% do PIB brasileiro (BRACELPA, 2004). Estes dados demonstram que este setor do mercado se encontra numa via de constante crescimento e superação.

As avaliações feitas revelam que o setor de papel e celulose tem hoje a totalidade do seu suprimento de madeira, originado de florestas plantadas, principalmente dos gêneros *Eucalyptus* e *Pinus*. Portanto, este é um setor que não exerce pressão sobre as florestas nativas remanescentes. No entanto, o setor de siderurgia e carvão vegetal tem 3/4 do seu suprimento de matéria-prima originado de florestas plantadas, e o restante é proveniente ainda, das florestas nativas remanescentes (IBAMA, 2002).

2.3 *Eucalyptus* spp.

O gênero *Eucalyptus* pertence à família das Mirtáceas. Esta é uma família muito extensa formada por um grande número de plantas lenhosas, abrange desde matas até grandes árvores. Apresenta folhas simples, geralmente opostas, com glândulas aromáticas, com consistência coreácea. É composta por aproximadamente 120 gêneros e 3.000 espécies originárias de regiões tropicais e subtropicais da Austrália principalmente, Ásia e América (Figura 3). Apresenta também uma grande importância econômica porque contém plantas de grande interesse e utilidade. Alguns dos gêneros mais destacados são *Agonis*, *Angophora*, *Callistemon*, *Eucalyptus*, *Eugenia*, *Feijoa*, *Lophomyrtus*, *Luma*, *Melaleuca*, *Metrosideros*, *Myrciaria*, *Psidium*, *Syncarpia*, *Syzygium*, *Tristania* (MYRTACEAE..., 2005).

Figura 3 - Distribuição da família das Mirtáceas no mundo

O gênero *Eucalyptus* envolve mais de 670 espécies. Algumas destas são utilizadas para a produção de madeira serrada, produção de carvão vegetal, fabricação de papel e produção de celulose, entre outros usos (ESPECIES..., 2001). No Brasil, o eucalipto, em função do seu rápido crescimento e das vantagens climáticas alcança uma produtividade aproximadamente 10 vezes superior a dos países líderes deste mercado. O crescimento das florestas na Finlândia, país tradicionalmente florestal, alcança, em média, rendimentos de 5 m³/ ha /ano; Portugal, 10; Estados Unidos, 15; África do Sul, 18; e Brasil, 45; embora possa alcançar só com a cultura do eucalipto produtividades de 60 a 70 m³ /ha / ano (VALVERDE, 2005; FAO, 2005).

Das espécies identificadas pelo Serviço Florestal da Austrália, apenas duas delas, *Eucalyptus urophylla* e *E. deglupta*, são de ocorrência natural fora do território australiano. Além do elevado número de espécies, existe também um número muito grande de variedades e de híbridos (O EUCALIPTO..., 2001).

2.4 *Eucalyptus* no Brasil

O *Eucalyptus*, como gênero exótico, foi introduzido no Brasil por volta do início do século XIX. As primeiras árvores foram plantadas no Jardim Botânico do Rio de Janeiro, no ano de 1825. Seu cultivo, em escala econômica, iniciou-se no ano de 1903, a partir do trabalho do agrônomo silvicultor Edmundo Navarro de Andrade, para atender a demanda da Companhia Paulista de Estradas de Ferro. Atualmente é a cultura florestal de maior importância econômica do Brasil (O EUCALIPTO..., 2001).

A cultura do eucalipto apresenta várias vantagens que a torna importante no reflorestamento comercial, como o seu rápido crescimento, a diversidade inerente ao gênero, adaptabilidade, alta produtividade e baixo custo (HO et al., 1998; AMBIENTEBRASIL, 2005; ESPÉCIES..., 2001). O melhoramento genético, desta cultura permitiu o aumento da produtividade em curto período de tempo e segue como tendência a utilização de espécies ou linhagens geneticamente puras, a exploração do vigor híbrido, obtidos mediante cruzamentos interespecíficos e a seleção de árvores de características silviculturais superiores (BERTOLUCCI et al., 1995; MELHORAMENTO..., 2001).

O eucalipto é considerado uma cultura muito importante na economia de mais de cem países. Atualmente, as dez espécies de eucalipto mais importantes nas plantações florestais no mundo, em termos de incremento médio anual de madeira, são: *Eucalyptus grandis*, *E. saligna*, *E. urophylla*, *E. camaldulensis*, *E. tereticornis*, *E. globulus*, *E. citriodora*, *E. robusta*, *E. esxerta* e *E. paniculata*. No Brasil, as espécies mais plantadas são o *Eucalyptus grandis*, *Eucalyptus saligna*, *Eucalyptus urophylla*, *Eucalyptus dunnii* (BRACELPA, 2004b). Entre outras espécies, destacam-se o *E. cloeziana* e o *E. citriodora* (ESPÉCIES..., 2001; O EUCALIPTO..., 2001). No entanto, da área total plantada com eucalipto no Brasil aproximadamente 50% é ocupada por algum tipo de híbrido (BRACELPA, 2004b).

Figura 4 – Representatividade das espécies de eucalipto mais plantadas no Brasil. (fonte: BRACELPA, 2004b)

As espécies mais representativas podem classificar-se segundo a sua utilidade (AMBIENTEBRASIL, 2005; ESPÉCIES..., 2001):

- Para lenha e carvão: *E. grandis*, *E. urophylla*, *E. torilliana*, por exemplo, oferecem grande quantidade de lenha em curto prazo.

- Para a produção de papel e celulose: *E grandis*, *E saligna*, *E urophylla* são as mais apropriadas por apresentarem cerne branco e macio.
- Para postes, moirões, dormentes e estacas: *E citriodora*, *E robusta*, *E globulus* são as mais apropriadas por apresentarem cerne duro e por resistir ao tempo.
- Para serrarias: *E dunnii*, *E viminalis*, *E grandis* se apresentam como as mais adequadas porque são espécies de madeira firme e não apresentam rachaduras.

Eucalyptus urophylla e *E. grandis* são espécies de maior interesse para a eucaliptocultura no Brasil. Sendo utilizados para a fabricação de celulose, chapas duras, produção de carvão e outros fins. O *Eucalyptus urophylla* possui uma ampla capacidade de adaptação a diversas condições de ambientes tropicais, aliada à tolerância ao patógeno *Cyphonectria cubensis*, causador do cancro do eucalipto (SANTOS et al., 1990). Já, o *E. grandis* está adaptado ao clima quente e úmido das regiões sudeste e nordeste da Austrália. A sua madeira é utilizada principalmente para produção de polpa (MURO-ABAD, 2001). Estas duas espécies foram utilizadas para a constituição do híbrido interespecífico (“Urograndis”) possibilitando a união das características mais desejáveis de ambas as espécies, tais como: boa produtividade, adaptação ecológica e resistência ao cancro. Assim, este híbrido passou a ser o mais cultivado no território nacional (BRACELPA, 2004b; BERTOLUCCI et al., 1995).

Considerando-se as espécies do gênero, *Eucalyptus grandis* é a espécie majoritariamente plantada no Brasil (Figura 4). Pode alcançar produções de 50-60 m³ por hectare, por ano, e em condições muito excepcionais a produtividade pode alcançar 100 m³ por hectare por ano. O *E. grandis* tem uma idade de colheita menor, dependendo do uso final da madeira. Para produção de polpa e celulose o tempo de colheita é de 5-6 anos. Já para postes de telefone a idade de colheita varia entre 7 e 8 anos e para uso como combustível industrial entre 10 e 12 anos (FAO, 2005).

Como cultura inserida no Brasil, o eucalipto é susceptível à uma série de doenças, na sua maioria, provocada por fungos patogênicos. A doença, como fenômeno biológico, interfere nos processos fisiológicos da planta, devido a elementos bióticos ou abióticos que determinam um desempenho anormal de suas funções vitais. Esta interferência prejudicial provoca uma redução da eficiência fisiológica nas plantas (KRUGNER, 1995).

As doenças bióticas, tradicionalmente, são vistas como uma interação entre dois organismos. A planta recebe o nome de hospedeiro e o agente causal recebe o nome de patógeno.

Ao colonizar uma planta o patógeno retira desta, os nutrientes necessários para o seu desenvolvimento. Em consequência o patógeno é considerado um parasita (KRUGNER, 1995).

Entre as principais doenças que atacam o eucalipto no Brasil podemos nomear as seguintes (PRINCIPAIS ..., 2001; KRUGNER; AUER, 1997):

- A ferrugem, causada pelo fungo *Puccinia psidii* Winter;
- O cancro, que é causado por várias espécies de fungos: *Cryphonectria cubensis*, *Valsa ceratosperma* na fase sexuada, *Cytospora* spp. na fase assexuada, *Botryosphaeria ribis*;
- O oídio, causado pelo patógeno *Oidium* sp.;
- O mofo cinzento que é causado pelo patógeno *Botrytis cinérea*;
- As manchas foliares de *Cilindrocladium*, provocadas por várias espécies deste patógeno;
- O tombamento de mudas, causado por algumas espécies de *Cilindrocladium*, *Rizhoctonia solani*, *Phytophthora* spp. e *Fusarium* spp.;
- A podridão da raiz, causada pela infecção das raízes pelos patógenos *Phytium* spp. e *Phytophthora* spp.;
- A podridão de estacas e miniestacas, causada pelos patógenos *Cilindrocladium* spp., *Rizhoctonia solani*, *Fusarium* spp., *Botryosphaeria ribis* e *Colletotrichum* spp.

Os problemas econômicos que resultam da ferrugem encontram-se relacionados aos plantios de campo, onde o tratamento com fungicidas é mais complicado. Plantas de *Eucalyptus grandis* infectadas com este fungo patogênico, sem tratamento, podem apresentar uma redução no desenvolvimento de cerca de 28% até 35%, se comparado a plantas que não sofreram o ataque (PRINCIPAIS..., 2001).

2.5 Fungos patogênicos: A ferrugem

As plantas encontram-se expostas a uma grande variedade de fatores ambientais. Entre estes fatores podemos relacionar os fatores abióticos, que podem ser variações de temperatura, incidência solar, etc. e os fatores bióticos, como uma ampla variedade de organismos patogênicos (DANGL; JONES, 2001).

Morfológicamente, os fungos podem apresentar dois tipos de estrutura básica: as estruturas vegetativas e as reprodutivas. As estruturas vegetativas podem ser unicelulares

(leveduriformes), simples (diminutas e arredondadas), plasmodiais ou miceliares. O talo miceliar é aquele característico da maioria dos fungos e está composto de filamentos tubulares, denominado hifa. Estas desempenham importantes funções no desenvolvimento do fungo. Por meio das hifas, o fungo coloniza o seu hospedeiro, absorve água e nutrientes. Os fungos, como parasitas das plantas, alteram a morfologia das suas hifas em resposta a características estruturais e fisiológicas das superfícies dos hospedeiros. Estas também mediam a união ao hospedeiro, reconhecimento, penetração, proliferação e nutrição (MENDGEN; HAHN, 2002). Estas hifas especializadas recebem diferentes nomes dependendo da especialização como (KRUGNER, 1995):

- Haustório: Estrutura, ramificada ou não, especializada na absorção de nutrientes a partir do citoplasma da célula hospedeira, no interior do qual este se desenvolve.
- Rizóide: Semelhante à estrutura de uma raiz de planta, ramificada, filamentosa. Atua na absorção de nutrientes.
- Hifopódio: Pequena ramificação de uma de duas células no micélio de alguns fungos epífitas. Funciona como âncora do fungo.
- Apressório: Estrutura achatada, formada pelo inchaço do tubo germinativo ou da hifa. Este se adere firmemente à superfície do hospedeiro para facilitar a penetração do fungo.

Por outro lado, os fungos como patógenos de plantas podem dividir-se em 2 grandes grupos (MENDGEN; HAHN, 2002):

- i. Os necrotróficos, matam rapidamente o hospedeiro, para posteriormente, alimentarem-se destes, como saprófitos e,
- ii. Os biotróficos, que precisam dos hospedeiros vivos, para completar o seu ciclo de vida levando, às vezes, à morte dos mesmos.

Os organismos biotróficos requerem a maquinaria celular do hospedeiro para poder alimentar-se e reproduzir-se. Bactérias e fungos podem participar de qualquer destes dois grupos, porém com a finalidade de ter acesso a nutrientes assim como um ambiente apropriado para o crescimento e a reprodução (LAM et al., 2001). O biotrofismo dos fungos apresenta uma série de características como: estruturas de infecção altamente desenvolvidas, atividade secretora limitada, supressão em longo prazo das defesas do hospedeiro e a presença de estruturas especializadas para a absorção de nutrientes (MENDGEN; HAHN, 2002).

A ordem Uredinales inclui os fungos causadores de ferrugens. Foram relatados aproximadamente 130 gêneros e 5000 espécies que afetam plantas, representando assim um dos maiores grupos de fungos fitopatogênicos. Atacam muitas espécies economicamente importantes como mirtáceas, leguminosas, cereais, cafeeiro, etc (KAISERSLAUTERN, 2005).

Os fungos causadores de ferrugem são patógenos biotróficos. As hifas especializadas destes, chamadas de haustórios, podem realizar interações estáveis com os hospedeiros, que podem durar até meses. O propósito deste patógeno é realizar uma infecção estável e causar um dano mínimo no seu hospedeiro (MENDGEN; HAHN, 2002). Depois da germinação dos uredósporos da ferrugem sobre a superfície foliar, a hifa penetra dentro da planta pelas aberturas estomatais. Uma vez dentro do tecido hospedeiro, um micélio intercelular prolifera e produz o que se denomina haustório dentro das células da planta (Figura 5) (KAISERSLAUTERN, 2005).

Figura 5 - Esquema do desenvolvimento da ferrugem no interior do tecido vegetal hospedeiro. (fonte: Kaiserslautern, 2005)

Até o momento, acredita-se que, o haustório tem uma função fundamental na absorção de nutrientes, nos tecidos das plantas. O mecanismo envolvido no processo de absorção de nutrientes dentro do haustório está relacionado a um transporte do tipo “*simport*”, que utiliza o potencial de gradiente dos prótons. O gradiente protônico é produzido pela atividade de uma ATPase protônica de membrana (KAISERSLAUTERN, 2005).

2.6 A ferrugem do eucalipto

Dentre os basidiomicetos pertencentes a ordem dos Uredinales, *Puccinia psidii* Winter apresenta uma distribuição geográfica ampla (Figura 6). Este patógeno é endêmico na América do Sul, América Central e no Caribe e é capaz de infectar diversas espécies de Mirtáceas nativas e exóticas. Já foram reportados como hospedeiros deste fungo 11 gêneros e 31 espécies (RAYACHHETRY et al., 2001).

Este fungo produz uma doença que atualmente afeta plantas jovens, viveiros e jardins clonais. A ferrugem do eucalipto só ataca plantas jovens com menos de dois anos de idade, sempre em órgãos tenros como primórdios foliares, terminais de galhos e haste principal (FIGUEREDO, 2001). Os sintomas do ataque são pequenas pontuações na parte inferior das folhas, levemente salientes de coloração verde clara ou vermelho-amarelada. Após duas semanas as pontuações se transformam em pústulas de uredósporos amarelos. As pústulas aumentam de tamanho e em seguida ocorre a típica esporulação uredospórica de coloração amarela forte, nos órgãos atacados, que ficam tomados pela infecção (RAYACHHETRY et al., 2001; PRINCIPAIS..., 2001).

A relação de susceptibilidade pode determinar três fenótipos gerais em relação à doença (RAYACHHETRY et al., 2001):

1. Sem a presença de sintomas o que é considerado imune;
2. Com pequenas pontuações que se desenvolvem em manchas cloróticas, como resistentes;
3. As que desenvolvem as pústulas, portanto, consideradas susceptíveis.

O desenvolvimento da doença causada por este fungo é dependente da temperatura e da umidade foliar, como mostrado por Tessmann, no primeiro estudo epidemiológico realizado numa Mirtácea no Brasil (TESSMANN et al., 2001). A ferrugem se tornou uma das mais severas doenças da cultura do eucalipto no Brasil, com um potencial de risco para outras regiões de clima semelhante como se vê na Figura 6 (COUTINHO et al., 1998).

Este patógeno não é conhecido nos centros de origem do gênero *Eucalyptus*. Aparentemente este fungo originou-se em Mirtáceas da América do Sul e tem a capacidade de adaptar-se para infectar diversos hospedeiros pertencentes a esta família (RAYACHHETRY et al., 2001). Já foi mostrado que este patógeno é altamente infectivo em algumas espécies economicamente importantes como goiaba (*Psidium guajava*), pimentão (*Pimento officinalis*)

(COUTINHO et al., 1998) e várias espécies de *Eucalyptus* são hospedeiras do fungo, entre elas *Eucalyptus citridora*, *Eucalyptus cloeziana*, *Eucalyptus grandis*, *Eucalyptus obliqua*, *Eucalyptus pilularis* e *Eucalyptus saligna* (KRUGNER; AUER, 1997).

Figura 6 – A área escura marca as regiões de alto risco para ocorrência da ferrugem causada pelo fungo *Puccinia psidii*-Winter no Neotrópico (BOOTH; JOVANOVIĆ, 2000)

2.7 Estratégias de controle

O controle da doença, até o presente, vem sendo realizado de várias formas como: utilização de fungicidas, colheita de árvores para aproveitamento da rebrota em épocas desfavoráveis à ferrugem e a utilização de plantas resistentes. A utilização desta última medida é a mais aconselhável por várias razões como baixo custo, praticidade e menor impacto ambiental, pela redução do uso de fungicidas (DE CARVALHO et al., 1998).

A resistência de espécies de eucalipto à ferrugem foi estudada no Brasil por Dianese et al., (1984) e posteriormente por De Carvalho et al., (1998) com auxílio de bases genéticas amplas.

Dianese et al. (1984) avaliaram 13 espécies de *Eucalyptus*, provenientes de 23 localidades onde a população foi cultivada, selecionada e adaptada, em relação a sua resistência à infecção por *Puccinia psidii*. *Eucalyptus microcorys* e *E. urophylla* apresentaram-se como altamente resistentes independente da localidade. Por outro lado, a resistência de *E. urophylla* dependia da procedência. Já as espécies de *E. cloeziana*, *E. citridora* e *E. saligna* apresentaram-se sensíveis à

infecção por este patógeno. *E. urophylla* S.T. Blake, *E. camaldulensis*, *E. tereticornis*, *E. pirocarpa*, *E. punctata*, *E. maculata*, *E. paniculata* e *E. saligna* se apresentaram moderadamente sensíveis. Os autores concluíram que a susceptibilidade a esta doença estava relacionada à procedência, assim como à espécie em questão (DIANESE et al., 1984).

Por outro lado, De Carvalho et al (1998) avaliaram o padrão de resistência de diferentes progênies das espécies *E. cloeziana*, *E. citridora*, *E. urophylla* e *E. pellita* distribuídas em distintas procedências. Estes detectaram variação dos níveis de resistências tanto entre procedências e progênies da mesma espécie, assim como entre espécies diferentes. Os autores identificaram como altamente resistentes todas as progênies de *E. citridora*, *E. urophylla* e *E. pellita*, independente da procedência e uma progênie de *E. cloeziana*, Ibitira X-28 arv. 4 procedente de Quadon-Austrália. As demais progênies de Quadon-Australia de *E. cloeziana* dividiam-se em resistentes e susceptíveis. Ibitira Z-28 arv. 7 e Ibitira X-33 arv. 2 apresentaram-se como resistentes e Ibitira X-29 arv.10 como susceptível. As outras progênies de *E. cloeziana* apresentaram-se como susceptíveis. Os autores concluíram que entre indivíduos de procedências e progênies suscetíveis podem-se detectar genótipos altamente resistentes (DE CARVALHO et al., 1998).

Portanto, a seleção de progênies ou clones resistentes pode ser realizada em *Eucalyptus*, a partir de infecções naturais no campo, em áreas onde a doença é severa ou em condições controladas por meio de uma infecção artificial (DE CARVALHO et al., 1998). A seleção destes indivíduos é importantíssima para os programas de melhoramento genético. Porém até o presente, pouco se sabe a respeito da genética da interação *Puccinia psidii* Winter-*Eucalyptus* spp., dos genes envolvidos, da herança da resistência ou das vias de resposta à doença ativadas durante o ataque deste patógeno (JUNGHANS, 2000).

2.8 Interação entre planta e patógeno

As plantas servem de fonte de alimento para os patógenos, pela elevada quantidade e qualidade dos materiais de reserva que acumulam e pela condição de imobilidade são alvos perfeitos para estes organismos, altamente especializados (LAM et al., 2001). Porém no curso de milhares de anos de evolução, tanto hospedeiros como patógenos desenvolveram inúmeras, eficientes e sutis técnicas de invasão e defesa. Para que os patógenos de plantas consigam infectá-las, primeiro precisam de sinais das plantas para induzir diferenciação e expressão de genes essenciais para a patogenicidade. Isto se torna um passo essencial para a ferrugem uma vez que a

diferenciação das hifas é ativada pela topografia da superfície das células da planta (THORDAL-CHRISTENSEN, 2003).

Como hospedeiras, as plantas apresentam vários níveis de defesa, que resolvem as duas principais limitações destas, a falta de um sistema de defesa circulatório e a imobilidade. Os mecanismos de defesa podem dividir-se em seis grandes grupos: defesa passiva ou barreiras pré-formadas, polarização celular e formação da papila, imunidade inata, defesa mediada por genes de resistência, resistência adquirida sistemicamente e resistência sistemicamente induzida.

A primeira linha de defesa, a defesa passiva, depende de capas epidérmicas densas e cutículas cerosas, parede celular, enzimas antimicrobianas e metabólitos secundários. O nível de sucesso dos patógenos em atravessar esta primeira barreira de defesa depende do grau de co-evolução da relação patógeno-hospedeiro (THORDAL-CHRISTENSEN, 2003). Os patógenos podem atravessar estas defesas de duas maneiras: sigilosamente, entrando pelas aberturas estomáticas, presentes no caule e folhas, ou com o uso da “força bruta”, mediante estruturas de penetração, tal o caso dos fungos, que introduzem uma hifa especializada em perfuração através das epidermes (HOLT et al., 2003). O fungo *Puccinia psidii* Winter geralmente entra sigilosamente através dos estômatos e realiza complexas interações biotróficas mediante hifas especializadas e os nutrientes são transportados através de sistemas de membrana especializados (THORDAL-CHRISTENSEN, 2003).

A polarização celular, temporalmente prévia à resposta hipersensível, é uma resposta que envolve uma série de mudanças morfológicas e uma rápida reprogramação metabólica da planta frente a fungos biotróficos. Esta apresenta um citoesqueleto metabolicamente muito ativo, um deslocamento do núcleo até o sítio da infecção através dos filamentos e microfilamentos do citoesqueleto, na procura de uma resposta rápida e localizada (GROSS et al., 1993) e a construção de uma barreira neste sítio reforçando a parede celular. Antes ainda da penetração fúngica detectam-se correntes citoplasmáticas indicando o início da resposta, correlacionada a uma reorganização profunda do citoesqueleto (KOBAYASHI et al., 1994), que são os componentes mais importantes (SCHLÜTER et al., 1997).

Acredita-se que as barreiras formadas têm a função de aprisionar ou excluir o patógeno do acesso aos nutrientes, determinando um engrossamento considerável da parede celular (SCHMELZER, 2002). Este engrossamento da parede celular chamado de papila é um fenômeno citológico conhecido (AIST, 1976) que apresenta características como: calose, aumento no

conteúdo de lignina, celulose, pectina, quitina, glicoproteínas ricas em hidroxiprolinas, peroxidases e lipídeos entre outros (SCHMELZER, 2002; HUECKELHOVEN et al., 1999).

A formação da papila é um processo altamente coordenado de síntese de compostos estruturais, transporte, secreção e reforço da parede altamente localizado e parece estar relacionado à resistência à infecção fúngica (SCHMELZER, 2002).

Uma via mais geral de defesa, para perceber organismos potencialmente patogênicos, é a imunidade inata. Esta é ativada pelo reconhecimento de PAMPs – *pathogen-associated molecular patterns* (padrões moleculares associados a patógenos) (ASAI et al., 2002) que são moléculas que apresentam características próprias de uma classe de organismo, seja este patogênico ou não (por exemplo, quitina e ergosterol são PAMPs detectados e característicos dos fungos patogênicos ou não). A resposta a estes elicitores é a produção de fluxos iônicos, compostos reativos de oxigênio, etileno, fortalecimento da parede celular e a indução de genes relacionados à defesa (NURNBERGER; BRUNNER, 2002; GÓMEZ-GÓMEZ; BOLLER, 2002) mediante cascatas de sinalização como as de MAP kinase (ASAI et al., 2002) porém, nunca uma resposta de morte celular característica da resposta hipersensível (GÓMEZ-GÓMEZ e BOLLER, 2002).

Por outro lado, a capacidade de certos patógenos de superar estes níveis de defesa levou ao desenvolvimento de mecanismos de identificação de patógenos e mecanismos interruptores que impedem que os patógenos completem seu ciclo de vida. A identificação é realizada através de interações clássicas 1gene-para-1gene (FLOR, 1971), envolvendo os genes de avirulência nos patógenos (Avr) e de resistência nos hospedeiros (R) de uma forma altamente específica. Esta defesa ativa tem o objetivo de impedir o progresso da infecção (Figura 7). (GRANT et al., 1998; DANGL; JONES, 2001).

Esta defesa é manifestada geralmente, como uma resposta hipersensível que se manifesta como uma morte celular programada, rápida e localizada no local da infecção, eliminando o organismo agressor (DANGL, 1998; LAM et al., 2001). Os produtos da expressão gênica dos genes Avr e R interagem diretamente ou indiretamente (“hipótese do guarda”) (Figura 7), produzindo uma resposta altamente específica. Nos primeiros 15 minutos se engatilha uma resposta que envolve fluxos iônicos através da membrana plasmática, produção de compostos reativos de oxigênio e óxido nítrico e uma reprogramação da expressão gênica por meio da ação de fatores de transcrição e kinases. Após este primeiro quarto de hora se produz a síntese de ácido

salicílico e etileno, fortalecimento da parede celular, processos de lignificação, síntese de compostos antimicrobianos, como por exemplo as fitoalexinas, culminando na morte celular programada, que caracteriza este tipo de resposta (DANGL; JONES, 2001). Exemplos deste tipo de resposta podem ser vistos em *Arabidopsis thaliana* Col-0 que é susceptível a *Pseudomonas syringae* DC3000. *Arabidopsis thaliana* pode apresentar resistência a este patógeno enquanto o patógeno expressar qualquer dos genes de avirulência *avrB* (STASKAWICZ et al., 1984), *avrRpm1* (DEBENER et al., 1991), *avrRpt2* (INNES et al., 1993) e *avrPph3* (JENNER et al., 1991). Por outro lado, mutantes de *Arabidopsis thaliana* para o gene *NDR1-1* apresentaram susceptibilidade a *Pseudomonas syringae* DC3000 expressando qualquer dos genes de avirulência, antes mencionados. Entretanto, a resposta hipersensível foi induzida em várias outras linhagens (CENTURY et al., 1995).

A mutação em *NDR1-1* em *Arabidopsis* também deixou as plantas sensíveis a vários isolados avirulentos de *Pseudomonas syringae* DC3000 e a um outro patógeno, *Perenospora parasitica* (CENTURY et al., 1995).

Outro exemplo dos genes R foi apresentado em soja. Nesse caso, o locus *Rpg1* (STEFFENSON, 1992) confere resistência à diferentes linhagens de *Pseudomonas syringae* pv *glycinea*, que expressem qualquer gene de avirulência *avrB* ou *avrRpm1* (GRANT et al., 1995; ASHFIELD et al., 2003).

Outro tipo de resposta de defesa pode ser a Resistência Sistêmica Adquirida, ou seja que imuniza o organismo após sucessivas infecções. A Resistência Sistêmica Adquirida é um estado de defesa aumentado, em partes não infeccionadas, ativado pela infecção primária de um patógeno (SCHENK et al., 2000; KUNKEL; BROOKS, 2002). Acredita-se que esta é induzida pela síntese de ácido salicílico, como resposta à presença do patógeno. Embora, no caso de sinalização a longas distâncias, este ácido não apresente as características de alta difusibilidade apropriadas e não é efetivo em baixas concentrações (GLAZEBROOK, 1999). No entanto, há indícios de que esta característica pode ser desenvolvida por moléculas derivadas de lipídeos que ativariam a síntese de ácido salicílico em sítios distanciados (MALDONADO et al., 2002). Esta resistência é manifestada mediante a expressão de genes que codificam proteínas relacionadas à patogênese (proteínas PR), mas na maioria dos casos, porém não é uma característica essencial para esta resposta (KIM; DELANEY, 2002). Como resultado disto, a planta ativa mecanismos de

defesa distantes do sítio de infecção (HUANG et al., 2002), temporariamente mais lentos que as respostas locais, tais como a resposta hipersensível (GLAZEBROOK, 1999).

A resistência sistemicamente induzida é um caso especial no qual a presença de um microrganismo na planta induz a resistência à infecção de outros patógenos (GLAZEBROOK, 1999). Esta via, porém, é independente do ácido salicílico (PIETERSE et al., 1996) e aparentemente dependente de ácido jasmônico e etileno (PIETERSE et al. 1998). Por exemplo, raízes de *Arabidopsis*, colonizadas por *Pseudomonas fluorescens* apresentavam um nível de resistência maior contra o fungo patogênico *Fusarium oxysporum* e também contra *Pseudomonas syringae* (PIETERSE et al., 1996).

Figura 7 – (a) e (b) Representação das interações possíveis entre os produtos gênicos dos genes de avirulência (Avr), nos patógenos e resistência (R), no hospedeiro: compatível, permite a ocorrência da doença e incompatível, que não permite o crescimento do patógeno. (1) Primeiro tipo de interação incompatível, que determina resistência pela interação direta dos produtos gênicos dos genes Avr e R. (2) Segundo tipo de interação incompatível que determina resistência pela interação indireta dos produtos gênicos dos genes Avr e R, denominada de “hipótese de guarda”, ou seja a molécula Avr atinge o seu alvo no hospedeiro, porém, a infecção não progride pela presença do gene R. Fonte: Hammond-Kosack e Parker (2003)

No caso de organismos biotróficos, as interações planta-patógeno pertencem geralmente ao terceiro grupo. Portanto, estas dependem de interações específicas entre o produto do gene

Avr (avirulência) do patógeno, e os produtos dos alelos dos genes de resistência (R) correspondentes das plantas (DANGL; JONES, 2001).

A enorme diversidade dos locus R basicamente pode ser dividida em 5 classes principais de proteínas (DANGL; JONES, 2001):

1. As proteínas intracelulares NB-LRR;
2. Proteínas do tipo Cf com um domínio transmembrana;
3. Proteínas Pto e Fen, kinases intracelulares;
4. Proteínas Xa21 e FLS2, do tipo kinases transmembrana;
5. As proteínas RPW8, que são proteínas de membrana com um domínio helicoidal e sem nenhuma semelhança com as classes anteriores.

Na atualidade, a principal classe de genes R é a classe NB-LRR “nucleotide-binding site-leucine-rich repeat” (VAN DER BIEZEN et al., 2002). Como o nome indica, cada uma destas proteínas contém um sítio de união a nucleotídeos essencial à união de ATP e GTP. Domínios LRR são encontrados em outras proteínas e funcionam como sítios de interação proteína-proteína, sítios de união peptídeo-ligante ou sítios de interação proteína-carboidrato, relacionados à especificidade da resistência (DANGL; JONES, 2001; KOBE; KAJAVA, 2001). Análises de mutantes ressaltam a importância destes domínios (DINESH-KUMAR; THAM; BAKER, 2000). Como exemplo, a ausência total e parcial do domínio NB não induz a ocorrência de resposta hipersensível, ressaltando a importância da função deste domínio na sinalização (ZHANG et al., 2004), união aos nucleotídeos trifosfatos e na hidrólise dos mesmos (TAMELING et al., 2002). Já no caso do domínio LRR e do número de suas repetições em diferentes alelos do locus L, há a hipótese de que este tem funções relacionadas ao controle da especificidade dos genes de resistência e seus correspondentes genes Avr (LAWRENCE et al., 1995).

Após o reconhecimento dos produtos dos genes Avr, por parte dos produtos gênicos R, uma série de eventos se desencadeiam, que determinarão a resposta à presença do patógeno (TORNERO et al., 2002; DANGL; JONES, 2001).

2.9 Vias de sinalização

A idéia geral é de que a resposta de defesa da planta é controlada por vias de sinalização interatuantes que envolvem moléculas sinalizadoras essenciais: ácido salicílico, jasmonato e etileno (SCHENK et al., 2000; KUNKEL; BROOKS, 2002; GLAZEBROOK, 1999). Estas moléculas parecem estar envolvidas nas duas principais vias de sinalização de defesa: a via de

sinalização dependente ou não de ácido salicílico. Nesta última estão envolvidas as outras duas moléculas, etileno e jasmonato (KUNKEL; BROOKS, 2002).

Schenk et al. (2000) com auxílio da tecnologia de microarrays analisaram o padrão de expressão gênica em plantas de *Arabidopsis* infectadas com o fungo *Alternaria brassicicola* e em presença das moléculas sinalizadoras: ácido salicílico, metil jasmonato e etileno. Neste trabalho são apresentadas evidências de que os genes envolvidos na resposta de defesa encontram-se regulados de forma coordenada e que estas vias de resposta encontram-se sobrepostas dado que vários genes aparecem funcionais em mais de uma via (SCHENK et al., 2000).

Após a infecção pelo patógeno, a concentração de ácido salicílico aumenta e a aplicação exógena de ácido salicílico determina um incremento na resistência para uma ampla faixa de patógenos. Esta molécula é importante para uma rápida ativação de respostas de defesa e para o estabelecimento Resistência Sistêmica Adquirida (KUNKEL; BROOKS, 2002).

Em tabaco e *Arabidopsis*, o bloqueio químico relacionado ao acúmulo de ácido salicílico assim como o bloqueio das vias de biossíntese de ácido salicílico, nos mutantes EDS16/SID2, afeta seriamente o mecanismo de defesa, aumentando a susceptibilidade a distintos patógenos. O que afeta a indução de Resistência Sistêmica Adquirida em tecidos distais (SHAH, 2003).

Sabe-se também que, o jasmonato é uma molécula derivada dos ácidos graxos. Esta molécula está envolvida no desenvolvimento do pólen, das sementes e no mecanismo de defesa contra os estresses mecânicos provocados por insetos e patógenos microbianos (KUNKEL; BROOKS, 2002). Plantas, nas quais a produção de jasmonato é inibida, apresentam uma sensibilidade incrementada à diferentes patógenos (STASWICK et al., 1998; CHEONG; YANG, 2003).

Por sua vez, a sinalização mediada por etileno é também importante para o mecanismo de defesa das plantas (DANGL; JONES, 2001). O etileno juntamente com o jasmonato são fundamentais na ativação da resistência sistêmica induzida (HAMMOND-KOSACK; PARKER, 2003). Este tipo de resistência sistêmica, por exemplo, é induzida pela bactéria colonizadora de raiz *Pseudomonas fluorescens* permitindo uma resistência incrementada contra outros patógenos. As sinalizações por jasmonato e etileno encontram-se intimamente relacionadas. Segundo Schenk et al. (2000), a metade dos genes que é ativada por jasmonato, é também induzida por etileno. Às vezes, estas duas moléculas estão presentes numa mesma via. Porém, este é um panorama

bastante simplista, uma vez que já foi mostrado que, as duas vias modulam-se de forma conjunta (KUNKEL; BROOKS, 2002).

As resistências sistêmicas adquiridas e induzidas podem ser ativadas simultaneamente. Isto aumentaria a proteção da planta contra o ataque de diversos tipos de patógenos (HAMMOND-KOSACK; PARKER, 2003).

O sentido evolutivo destas múltiplas vias de sinalização de defesa, provavelmente, é o de produzir uma fina modulação de resposta de defesa contra um determinado patógeno, o que leva a uma combinação de resposta contra patógenos específicos de acordo com as estratégias de virulência destes (KUNKEL; BROOKS, 2002).

Embora nos últimos anos, o entendimento, das vias acima mencionadas e os seus respectivos componentes que respondem a ubíquos patógenos, tenha aumentado consideravelmente, ainda muito, precisa ser esclarecido. Para isso, uma série de tecnologias de genômica funcional nos permitem analisar várias etapas da expressão gênica.

2.10 A Biotecnologia e a revolução da genômica funcional

A importância econômica mundial da cultura de eucalipto, a expansão das áreas afetadas pela ferrugem e o alto risco que esta representa para países com climas tropicais e sub-tropicais (COUTINHO et al., 1998; BOOTH et al., 2000), tornam necessária uma avaliação mais aprofundada das bases genéticas desta doença. A genômica funcional apresenta uma série de ferramentas que permitiram uma avaliação mais apurada de como se comportam os tecidos das plantas resistentes frente a esta doença.

Pode-se considerar que a genômica funcional, teve início em 1977, quando Sanger publicou a primeira seqüência de DNA completa de um organismo, o genoma do vírus ϕ X174, com 5386 pares de bases, codificando 9 proteínas (SANGER et al., 1977). A partir daí iniciou-se uma revolução no campo da genômica. Rápidos avanços em computação, robótica e o “shotgun”, uma técnica de enorme capacidade no processamento de seqüências que permitiu o seqüenciamento de genomas muito maiores. No ano de 1995, os genomas de dois organismos procarióticos, *Haemophilus influenzae* (FLEISCHMANN et al., 1995) e *Mycoplasma genitalium* (FRASER et al., 1995) foram seqüenciados, no ano seguinte foi também seqüenciado o primeiro organismo eucariótico, *Saccharomyces cerevisiae* (GOFFEAU et al., 1996). Atualmente, no banco de dados do “National Center for Biotechnology Information” (NCBI) podem-se encontrar o genoma de 100 microrganismos, mais de 1000 vírus, representantes de todos os domínios da

vida, onde arquea está representada por 16 genomas completos, as bactérias estão representadas por 121 genomas completos, o grupo dos eucariotos estão representados por 20 projetos genomas assim com também os genomas de mitocôndrias e cloroplastos (<http://www.ncbi.nlm.nih.gov/>).

Os esforços em genômica vegetal foram muitos, uma vez que o genoma de plantas pode ser muito variável em questão de tamanho. Genomas pequenos podem ser comparáveis a genomas de pequenos animais. Há também aqueles que têm tamanhos várias vezes maiores que o genoma humano. O seqüenciamento do primeiro genoma de plantas superiores, *Arabidopsis thaliana*, foi finalizado no ano de 2000 (THE ARABIDOPSIS GENOME INIATIVE, 2000). Em seguida outros genomas vegetais surgiram e novos programas de seqüenciamento vêm sendo disponibilizados: *Avena sativa*, *Glycine max*, *Hordeum vulgare*, *Lycopersicon esculentum*, *Oryza sativa*, *Triticum aestivum*, *Zea mays* (<http://www.ncbi.nlm.nih.gov/>).

A disponibilidade de seqüências genômicas completas, promoveram um grande avanço permitindo pela primeira vez, um estudo e análise do genoma como um todo. O principal objetivo deste tipo de estudos é construir uma ponte, sobre o espaço existente entre genótipo e fenótipo, reconhecer quais são os genes que são responsáveis por determinados processos tais como doenças, quais são os responsáveis por diferentes tipos de resposta à fatores bióticos e abióticos, como a resposta imune dos organismos à presença de patógenos, entre outros. Mas, para tudo isto, é necessário extrair o máximo de informação dos dados de seqüências mediante a análise de todo tipo de seqüências codificantes e não codificantes, tais como as seqüências promotoras, reguladoras e moduladoras. A análise das seqüências repetitivas, das seqüências nucleotídicas que sofrem modificações, entre outras, se tornam importantes para entender como este conjunto de informação se manifesta. Também, devem-se analisar profundamente as seqüências codificantes, onde e como se manifestam, como respondem à determinadas situações, suas dinâmicas e formas de ação (BORK; KOONIN, 1998).

2.11 Transcrissoma: Cartografia para entender a função gênica

A informação dos projetos genoma, que agrupam seqüências de milhares de genes, apresenta a possibilidade de utilização das mesmas para identificar as suas respectivas seqüências ribonucleotídicas, os mRNAs. E em conseqüência, medir em paralelo e em larga escala, o nível de expressão dos mRNAs, em diferentes condições e organismos, para a caracterização da expressão gênica. Esta caracterização é um passo lógico para o entendimento da função biológica de um determinado gene. O desenvolvimento da metodologia de SAGE (serial analysis of gene

expression) (VELCULESCU et al., 1995), microarrays de cDNA (SCHENA et al., 1995) e microarrays de oligonucleotídeos (LOCKHART et al., 1996), permitiram que este passo fosse concretizado, numa perspectiva de análise ampla (SCHENA et al., 1996).

Portanto, desde que foi desenvolvida, e graças ao avanço da tecnologia de *chips* de DNA, a metodologia de microarrays de DNA (SCHENA et al., 1995), está se tornando uma ferramenta padrão para o monitoramento da expressão gênica. Esta metodologia foi aplicada a estudos de vários organismos modelos como: *Escherichia coli* (RICHMOND et al., 1999), *Saccharomyces cerevisiae* (KOERKAMP et al., 2002), animais como *Drosophila* (WHITE et al., 1999), em humanos (SCHENA et al., 1996; POLLACK et al., 1999), em culturas de células (ZIAUDDIN; SABATINI, 2001) e mais recentemente no estudo da expressão gênica de plantas (SCHENA et al., 1995; DESPREZ et al., 1998; RUAN; GILMORE; CONNER, 1998; YUE et al., 2001; REYMOND, 2001).

2.12 SAGE

Por outro lado, o desenvolvimento da metodologia SAGE (Serial Analysis of Gene Expression) permitiu a realização de análises da expressão gênica de um ponto de vista totalmente original. Esta metodologia tem a capacidade de identificar e quantificar a expressão de transcritos conhecidos e desconhecidos (VELCULESCU et al., 1995) e oferecer um perfil de expressão gênica da amostra processada.

Esta se encontra fundamentada em três princípios básicos (VELCULESCU et al., 1995; VELCULESCU et al., 2000):

1. Sequências curtas de 9-14 pares de bases, denominadas etiquetas ou *tags*, podem ser obtidas de regiões definidas de cada transcrito. Estas contêm a informação suficiente para identificar os respectivos transcritos originais.
2. Estas etiquetas podem ser unidas para formar uma molécula de DNA longa, que quando clonada em vetores de ampliação produzirá clones representativos de 30 e 50 mRNAs, diferente das outras metodologias ESTs, onde cada clone representa um único mRNA. O pequeno tamanho das mesmas, permite o seqüenciamento de um número maior de etiquetas, diminuindo o custo de seqüenciamento. Em consequência, isto possibilita o seqüenciamento de etiquetas em larga escala e o aumento da confiabilidade estatística na identificação dos transcritos (BOHELER; STERN, 2003). A confiança na identificação de uma etiqueta é um parâmetro

estatístico diretamente relacionado ao número de seqüências pertencentes ao mesmo transcrito (STOLLBERG et al., 2000).

3. O nível de expressão dos transcritos é então quantificado pelo número de vezes que uma etiqueta é observada, sendo os resultados apresentando em números inteiros.

Esta metodologia tem as suas vantagens, em relação aos outros métodos utilizados para a análise do transcissoma. Por exemplo: não é necessário ter-se um conhecimento prévio das unidades de transcrição para quantificá-las e analisá-las. Isto, devido ao fato que praticamente todos os mRNAs geram uma etiqueta, numa posição determinada (MADDEN; WANG; LANDES, 2000). Segundo Robinson et al. (2004) quase todos os mRNA poderiam gerar etiquetas para o SAGE, mediante o uso da endonuclease *Nla* III. No caso do organismo modelo por ele analisado, *Arabidopsis thaliana*, 98% dos genes puderam gerar etiquetas e 79 % das seqüências consenso (seqüência de reconhecimento da endonuclease – etiqueta no extremo 3', próximo da cauda poli A) puderam ser relacionados ao seu gene de origem (ROBINSON et al., 2004).

Essa técnica permite estudar seqüências de baixo número, uma vez que esse parâmetro depende do número de etiquetas seqüenciado (YAMAMOTO et al., 2001). Não é uma metodologia excludente, ou seja, pode complementar-se com resultados de análises de ESTs, por exemplo. E como característica de destaque, gera dados numéricos.

Apresenta-se como uma ótima estratégia para a identificação de genes em genomas de interesse (BOHELER; STERN, 2003; MADDEN; WANG; LANDES, 2000). Complementa os resultados obtidos com a metodologia de microarrays, no estudo de transcissomas e na identificação de transcritos novos. Por exemplo, no genoma humano Chen et al. (2002) utilizaram-se da técnica de SAGE com sucesso, em áreas complexas como a medicina cardiovascular (PATINO et al., 2003); na comparação entre os padrões de níveis de transcrição em distintas regiões do cérebro (DE CHALDÉE et al., 2003); no estudo de doenças como câncer (CERUTTI et al., 2003; WEERARATNA, 2003); no estudo do sistema imune de humanos, como por exemplo, na regulação das células T, que apresentam enormes perspectivas no campo dos transplantes de órgãos, mais especificamente, no estudo da tolerância a transplantes (COBBOLD et al., 2003). Esta tecnologia também foi utilizada em animais simples como *C. elegans*. Nesse caso, com auxílio do SAGE uma mudança no metabolismo deste organismo, foi observada como

decorrência de uma alteração do estágio larval desse nematóide para o estado anaeróbico (HOLT et al., 2003).

2.13 SAGE em plantas

Na área vegetal, muitos trabalhos de pesquisa vêm sendo realizados com auxílio da metodologia de SAGE em plantas modelo, como *Oryza sativa* (MATSUMURA; NIRASAWA; TERAUCHI, 1999; GIBBINGS et al. 2003; MATSUMURA et al., 2003), *Pinus taeda* (LORENZ; DEAN, 2002) e *Arabidopsis thaliana* (JUNG; LEE; LEE, 2003, LEE; LEE, 2003, ROBINSON et al., 2004).

Em *Oryza sativa*, foram realizadas análises globais do perfil transcricional de *seedlings* (MATSUMURA; NIRASAWA; TERAUCHI 1999), de folha madura e sementes imaturas (GIBBINGS et al., 2003), e culturas de células com e sem o estímulo de um elicitor do fungo patogênico de arroz (*Magnaphorte grisea*) (MATSUMURA et al., 2003), mediante a utilização desta metodologia. Estes experimentos apresentaram resultados surpreendentes. Um deles, por exemplo, mostrou que o transcrito mais abundante pertence a uma metalotionina (MATSUMURA et al., 2003). Resultado esse, corroborado por Gibbings et al. (2003) em análises de folhas e sementes. Segundo Gibbings et al. (2003) a metalotionina era do tipo 3. Uma segunda observação feita foi que, 70% dos transcritos de cópia única são comuns às duas bibliotecas, de semente e folha de arroz e uma quantidade significativa das etiquetas codificavam transcritos antisense, sugerindo, assim, um novo mecanismo de regulação gênica em plantas (GIBBINGS et al., 2003). Análises de expressão gênica desta mesma cultura, sob a ação de patógenos, também foram realizadas. Culturas de células de *Oryza sativa* foram postas em contato com o extrato de parede celular do fungo patogênico de arroz, *Magnaporthe grisea*. Este extrato provoca uma série de efeitos na cultura de células, como uma rápida produção de peróxido de hidrogênio, alcalinização do meio de cultura e morte celular. Neste caso, a metodologia de SAGE foi utilizada com o objetivo de identificar os genes envolvidos nesta morte celular, em culturas de células em suspensão, tratadas com elicitores. Entre os genes que apresentaram uma expressão reprimida, estava o gene que codifica o BI-1 (o inibidor Bax). As plantas transgênicas que superexpressam o gene BI-1, apresentaram uma resistência para desencadear a morte celular, em presença dos elicitores. Portanto, concluiu-se que este gene tem uma função na regulação da morte celular programada, pelo menos *in vitro*, em células de arroz (MATSUMURA et al., 2003).

O perfil transcricional de pinus (*Pinus taeda*) também foi realizado. Duas partes distintas da árvore foram utilizadas, o xilema lignificado isolado de duas regiões do tronco: parte superior e base, de árvores com 10 anos. Podê-se demonstrar com esse trabalho que há diferenças no nível de expressão gênica entre as duas partes do tronco. Este trabalho é bastante representativo por ser uma análise extensa realizada em árvores com uma descrição quantitativa do transcrissoma do xilema lignificado em pinus (LORENZ; DEAN, 2002).

Em *Arabidopsis* a transcrição em folhas foi analisada, estudando-se as variações que ocorriam no transcrissoma quando a planta se encontrava sob estresse de frio. Após o tratamento com frio, muitos genes foram superexpressados nas folhas. Entre estes genes, estavam os de resgate celular, defesa, morte celular, síntese protéica, de metabolismo, transporte, vários genes COR – *cold regulated* (genes regulados pelo frio), transportadores de lipídeos, álcool desidrogenase, β -amilase entre outros genes novos. Em contraposição, os genes que foram reprimidos, foram genes relacionados à fotossíntese, envolvidos na produção de energia. Estes resultados apresentam uma informação valiosa para o entendimento do mecanismo de tolerância ao frio (JUNG; LEE; LEE, 2003). Lee e Lee (2003) também analisaram o complemento de transcritos em pólen de *Arabidopsis*, utilizando esta tecnologia para analisar os transcritos próprios deste tipo de tecido e a expressão de transcritos diferenciais quando contrastados com os dados de folhas analisados por Jung, Lee e Lee (2003). Dessa forma, estes contrastes foram feitos em condições de temperatura normal e baixa e em tecidos de folhas e de pólen. A diferença dos experimentos de Jung, Lee e Lee (2003) é que os genes relacionados à síntese protéica, altamente expressos em folha, se encontravam pobremente expressos em pólen. A maioria dos transcritos não apresentava variações nas amostras contrastadas. Entre os que praticamente não modificaram a sua expressão, encontravam-se os que provavelmente participavam na aclimação dos vegetais como proteínas COR, transportadores de lipídeos e β -amilases (LEE; LEE, 2003).

Um dos pontos mais importantes desta técnica possivelmente é a produção de dados em número inteiros. Este tipo de dados permite que os mesmos possam ser utilizados posteriormente para comparações ao longo do espaço e do tempo, entre diversos experimentos. Esta característica foi explorada com *Arabidopsis thaliana*, analisando-se os resultados dos perfis de transcrição em folhas e pólen (LEE; LEE, 2003).

Os estudos dos perfis transcricionais chegaram até os organismos patogênicos de certas plantas, como é o caso da cevada. As folhas da cevada são alvo de um fungo patogênico, o

Blumeria graminis. Este se desenvolve sobre suas folhas e aí sofre uma série de alterações morfológicas que auxiliam no processo de infecção. Estas transformações incluem a formação do conídio, tubo germinal primário e a diferenciação das hifas em apressório. A metodologia SAGE foi utilizada com o objetivo de visualizar as alterações no padrão de transcrição, no percurso das transformações, produzindo um banco de dados para a caracterização das mudanças transcricionais, através do desenvolvimento deste patógeno durante o processo de infecção (THOMAS et al., 2002).

Os sistemas de defesa das plantas são diversos, versáteis e complexos (GLAZE BROOK, 1999). Porém, o entendimento destes sistemas será beneficiado com tecnologias como as de SAGE e microarrays. O SAGE surge como uma ferramenta poderosa para a análise dos complexos mecanismos que formam partes das respostas de defesa das plantas, frente a diferentes agressores. O SAGE permite uma avaliação global da expressão de transcritos. Portanto, a contribuição desta tecnologia será importantíssima para a determinação e o descobrimento de genes relacionados com o mecanismo de defesa das plantas. Isto abre a possibilidade da geração de novas perspectivas de seleção e de plantas resistentes, que se tornarão de vital importância para a agricultura.

3 MATERIAL E MÉTODOS

3.1 Material Vegetal:

Árvores jovens da espécie *Eucalyptus grandis* Hill ex Maiden, com aproximadamente 1 ano e meio de idade foram utilizadas para a coleta do material vegetal. Esta espécie é natural da Austrália, especificamente do norte do estado de New South Wales, sul de Queensland, com altitude de 300 a 900m (FERREIRA, 1979). Estas árvores foram obtidas a partir de sementes oriundas de Coff's Harbour, Austrália, cujas coordenadas geográficas são as seguintes latitude 30° 00', longitude 152° 55' e altitude de 91 metros. Essas sementes foram semeadas em Mogi Guaçu – SP. Indivíduos desta população foram selecionados quanto as características de crescimento e de acordo com os testes de progênie. As progênies destes indivíduos foram então semeadas no município de Itatinga – SP e submetidas a um novo ciclo de seleção. As melhores árvores oriundas das melhores famílias foram propagadas de forma vegetativa para a instalação de um pomar de sementes. Uma dessas árvores foi utilizada para a instalação de um pomar de sementes na fazenda São Miguel Arcanjo, no município de Itatinga, cuja progênie foi plantada para avaliação quanto à de resistência a ferrugem. Como estas árvores são uma população de meios irmãos, representam uma população de base genética restrita. Durante os períodos de infecção natural da ferrugem, essa população apresenta indivíduos com resistência e susceptibilidade à ferrugem. Como, nesta população fenótipos resistentes e susceptíveis à *Puccinia psidii* Winter são encontrados, amostras de folhas desse material foram fornecidos pela Suzano - Papel e Celulose, para a realização desse trabalho.

3.2 Coleta das amostras:

Amostras de tecido foliar, de dez indivíduos de cada fenótipo (resistente e susceptível) foram coletadas. Posteriormente, para cada fenótipo, realizou-se uma mistura das amostras de tecido foliar de vários indivíduos, de forma proporcional, com o objetivo de realizar uma amostragem mais representativa de cada fenótipo, ao invés de analisar somente um único indivíduo por fenótipo. A coleta foi realizada com luvas estéreis, todo o material utilizado foi tratado de forma a estar livre de RNase (como descrito abaixo). As amostras foram congeladas imediatamente em nitrogênio líquido para o transporte até o laboratório, e mantidas a -80°C até serem utilizadas nas análises.

As amostras foram posteriormente submetidas a um processo de extração de RNA total.

3.3 Extração do RNA total:

Todo material utilizado foi tratado a fim de estar livre de RNAses. O material de plástico foi submetido a uma série de lavagens com SDS – dodecil sulfato de sódio, a 2% (p/v) por pelo menos 2 horas, a 70°C, seguido de 20 minutos com álcool absoluto e 20 minutos com peróxido de hidrogênio – H₂O₂ 3% (p/v). O enxágüe final foi feito com água milli Q, previamente tratada com dietilpirocarbonato (DEPC), na concentração final de 0,01% (v/v), por mais 20 minutos, antes da autoclavagem (uma atm, 120°C) por 20 minutos.

Devido à dificuldade de obtenção de uma amostra de RNA total de boa qualidade, uma combinação de métodos de extração de RNA, foi utilizada. Os métodos de Zeng e Yang (2002) e de Salzman et al. (1999).

3.4 Extração de RNA: Método de Salzman et al., (1999)

A extração de RNA total realizada com auxílio do método desenvolvido por Salzman et al. (1999) foi realizada a partir de dois gramas de tecido foliar de *Eucalyptus grandis*, das amostras dos dois fenótipos. As amostras foram maceradas com auxílio de nitrogênio líquido, almofariz e pistilo. Ao macerado adicionou-se 10 ml de tampão de extração (ANEXO A) seguido de 10 ml de CIA (clorofórmio e álcool isoamílico), na proporção de 24:1, respectivamente. Em seguida agita-se (200 rpm), por 20 minutos. As amostras foram centrifugadas (16000 g) por 10 minutos, a 4°C. A fase superior aquosa foi transferida para tubo novo, repetindo-se a extração com CIA até a completa remoção da interfase. O sobrenadante final foi transferido para tubo novo, ao qual adicionou-se dois volumes de etanol absoluto e 0,1 volume de (1 ml) 5 M de NaCl - cloreto de sódio, para a precipitação do RNA que foi mantido a -20°C, por 16 horas. O RNA total foi recuperado por centrifugação (16000 g), por 10 minutos, a 4°C e ressuspendido em 10 ml de água milli Q tratada previamente com DEPC. Uma centrifugação adicional (20000 g), por 10 minutos, foi realizada para remoção do material insolúvel em água. O sobrenadante, contendo o RNA, foi transferido para um tubo novo, ao qual adicionou-se um volume igual (10 ml) de uma solução contendo CIA e fenol, na proporção de 1:1. Essa solução foi agitada (200 rpm), por 10 minutos à temperatura ambiente, e centrifugada (13000 g) por 10 minutos. A fase superior aquosa foi transferida para um tubo novo, repetindo mais uma vez a extração com CIA e fenol. À fase superior final foi transferida para um tubo novo, adicionou-se dois volumes de etanol absoluto e 0,1 volume de 5M NaCl para precipitação dos ácidos nucleicos e ribonucleicos. A amostra foi mantida a -20°C, por 16 horas, antes da centrifugação (16000 g), por 15 minutos, a 4°C. Os

ácidos nucléicos recuperados foram ressuspensos em 500 µl de água milli Q previamente tratada com DEPC. O volume foi ajustado para 1 ml, com a mesma água milli Q, ao qual adicionou-se 333 µl de 8M cloreto de lítio – LiCl para a precipitação seletiva (a 4°C) do RNA, por 3 horas. Após esse período, o “*pellet*” de RNA resultante da centrifugação (12000 g) por 20 minutos a 4°C, foi seco à temperatura ambiente, durante 10 minutos, e ressuspensionado em 100µl de água milli Q, previamente tratada com DEPC.

3.5 Extração de RNA: Método de Zeng e Yang (2002)

A extração de RNA total realizada pelo método desenvolvido por Zeng e Yang (2002) foi realizada a partir de um grama de tecido foliar de *Eucalyptus grandis*, das amostras de cada fenótipo. O material vegetal foi macerado com auxílio de nitrogênio líquido, almofariz e pistilo. Ao maceramento, adicionou-se 10 ml de tampão de extração (ANEXO B) pré-aquecido (65°C). A solução, após homogeneização foi mantida a 65°C, durante 15 minutos com agitação cada 5 minutos. A extração com clorofórmio foi feita pela adição de um volume de CIA (clorofórmio e álcool isoamílico, na proporção de 24:1, respectivamente), seguida de homogeneização e centrifugação (10000 g), durante 10 minutos, a 4°C a fase superior aquosa foi transferida para um tubo novo, repetindo-se a extração com CIA, antes da centrifugação (30000 g), durante 20 minutos, a 4°C. O sobrenadante foi transferido para um tubo novo e adicionou-se 0,25 volumes de 10M LiCl. Após homogeneizar as amostras foram mantidas por 16 horas, a 4°C, seguidas de centrifugação (30000 g), por 30 minutos, a 4°C. Os “*pellets*” resultantes foram lavados três vezes com etanol (75%), secos à temperatura ambiente durante 10 minutos, antes de serem ressuspensos em 110µl de água DEPC.

3.6 Determinação espectrofotométrica de pureza e concentração de RNA

As amostras de RNA foram avaliadas quanto à pureza e quantificadas por meio de espectrofotometria, em dois comprimentos de onda 260nm e 280nm. A razão, entre as absorbâncias 260/280 (A_{260nm}/A_{280nm}), é utilizada como medida padrão de qualidade das amostras. Valores superiores a 2.0 indicam que o RNA extraído é de boa qualidade. Já a absorbância em 260 nm (A_{260nm}) é utilizada para a quantificação do RNA total. A quantificação foi realizada utilizando-se a fórmula da Figura 8.

$$[\text{RNA}] = \frac{A_{260\text{nm}} \times F_c \times F_d}{1000}$$

Figura 8 – Fórmula para cálculo da concentração de RNA. $A_{260\text{nm}}$ absorvância a 260nm, F_c : fator de conversão, que indica a equivalência de 1 unidade de absorvância no comprimento de onda de 260nm correspondente a 40 μg RNA/ml. F_d é o fator de diluição da amostra. Neste caso 200 (porque são utilizados 4 μl da amostra em 800 μl de água). O valor final é dividido por 1000, para expressarmos a unidade de concentração de RNA em $\mu\text{g}/\mu\text{l}$. Fonte: Sambrook et al. (1989b)

3.7 Determinação da pureza e concentração de RNA

As amostras também foram avaliadas por eletroforese em gel de agarose (1%) em tampão TAE (SAMBROOK et al., 1989a) (ANEXO C e ANEXO D).

3.8 Isolamento do mRNA

O isolamento de RNA mensageiro foi realizado a partir de 75 μg das amostras de RNA total, com auxílio do kit *Dynabeads*[®] mRNA Purification Kit da Dynal, de acordo com as normas do fabricante. Este kit faz uso das *Dynabeads Oligo* (dT)₂₅, ou seja, oligos de desoxitimidina, que se unem por complementaridade à cauda poliA do mRNA.

A amostra de 50 μg de RNA total dissolvidos em 100 μl de água milli Q previamente tratada com DEPC, adicionou-se 1 volume do tampão de ligação (ANEXO E), seguidos de desnaturação por 2 minutos a 65°C. Nesse meio tempo, 200 μl de *Dynabeads Oligo* (dT)₂₅ foi lavado, com auxílio do *Dynal Magnetic Particle Concentrator* – MPC, (um suporte magnético que permitiu a concentração de partículas magnéticas) com 100 μl de tampão de ligação (ANEXO E), antes de serem ressuspensos em 100 μl desse mesmo tampão. Os oligo foram então adicionados à amostra de RNA total e homogeneizados, por rotação suave, por 5 minutos, à temperatura ambiente. A solução homogeneizada foi colocada no suporte magnético (MPC) por 30 segundos. O tubo foi retirado do MPC, o sobrenadante removido e o complexo *Dynabeads*-mRNA foi lavado duas vezes, com 200 μl de tampão de lavagem B (*Washing Buffer B*) (ANEXO F), sempre com auxílio do MPC.

Após lavagem, o complexo *Dynabeads*-RNA mensageiro foi ressuspendido em 20µl de 10 mM Tris-HCl. A solução foi submetida a aquecimento (65°C), por 2 minutos e imediatamente transferida para o MPC. O RNA mensageiro, no sobrenadante, foi então transferido para um tubo novo.

A qualidade do mRNA foi avaliada mediante a técnica de RT-PCR – “Reverse transcription-Polymerase chain reaction” com primers de genes de alta expressão (neste caso: UDP-glucose pirofosforilase) antes de se iniciar a construção das bibliotecas de SAGE, uma vez que a qualidade do RNA é imprescindível para esse processo.

3.9 PCR e Transcrição reversa

A síntese do cDNA foi realizada com auxílio da *SuperScript™ One-Step RT-PCR* da Invitrogen, de acordo com normas do fabricante, utilizando-se *primers* específicos (UDP-glucose pirofosforilase).

A reação num volume total de 25µl, está descrita no ANEXO G.

3.10 Construção de bibliotecas de SAGE:

A construção das bibliotecas foi realizada seguindo-se as instruções do método de I-SAGE™ *kit* da Invitrogen. Esta metodologia desenvolvida por Velculescu et al. (1995), produz bibliotecas apropriadas para a análise quantitativa da expressão gênica, de pequenas amostras, sem prévio conhecimento do transcissoma do indivíduo.

Uma alíquota de 50µg de RNA total, previamente avaliada quanto à qualidade, foi utilizada para a construção das bibliotecas de SAGE. Estas foram centrifugadas (16000 g), 40 minutos, a 4°C, o “*pellet*” lavado duas vezes com etanol 70% e novamente centrifugado (14000 g), por 5 minutos, a 4°C, antes de serem ressuspendidos em água milli Q, previamente tratada com DEPC.

3.11 A síntese do cDNA: primeira fita

Às partículas magnéticas poliT, do kit da Invitrogen, foram ressuspendidas, e uma quantidade de 100 µl foi transferida para um tubo estéril livre de RNases, que foi mantido em um suporte magnético MPC, por 2 minutos. Após descarte do sobrenadante, as partículas magnéticas foram lavadas por ressuspensão das mesmas em 500 µl de tampão de ligação e lise (*Lysis/binding buffer* ANEXO H), com auxílio do MPC.

Nesse meio tempo, ressuspendeu-se 50 µg de RNA total em 1 ml do mesmo tampão de ligação e lise (ANEXO H), e adicionou-se este 1 ml da solução sobre as partículas magnéticas equilibradas no tampão de ligação e lise. O mRNA associado às partículas de oligo dT foi suavemente agitado, por 30 minutos, à temperatura ambiente. O tubo contendo as amostras, foi transferido para o MPC, para realização de duas lavagens, com 1 ml de tampão de lavagem A (*Washing buffer A* ANEXO I), e uma lavagem posteriormente com 1 ml de tampão de lavagem B (*Washing buffer B* ANEXO J). Em seguida, realizaram-se quatro lavagens adicionais com o tampão de primeira fita 1X (1X *First Strand Buffer* ANEXO K).

Nesse meio tempo, preparou-se a solução para síntese da primeira fita de cDNA (*First Strand cDNA Synthesis* ANEXO L). As partículas magnéticas associadas ao mRNA foram ressuspendidas nessa solução, suavemente homogeneizadas e o tubo mantido a 37°C, por 2 minutos, ao qual adicionou-se 3 µl da transcriptase reversa *SuperScriptTM II* (ANEXO O). A reação de síntese da primeira fita foi realizada por uma hora, a 37°C. Após esse período, o tubo foi mantido em gelo por 2 minutos.

3.12 A síntese do cDNA: segunda fita

Para a síntese da segunda fita adicionou-se, ao tubo anterior, 465 µl de água DEPC, 150 µl de tampão para a segunda fita 5X (5X *Second strand buffer* ANEXO P), 15 µl de dNTP *mix* (ANEXO Q), 5 µl de *E. coli* DNA ligase (ANEXO R), 20 µl de *E. coli* DNA polimerase (ANEXO S) e 5 µl de *E. coli* RNase H (ANEXO T). Após leve agitação, o tubo foi mantido à 16°C, por 2 horas. Em seguida, o tubo foi transferido para gelo, ao qual adicionou-se 45 µl de 0,5M de EDTA para finalização da reação. O tubo foi então transferido ao MPC por 2 minutos. Após descarte do sobrenadante, adicionaram-se 750 µl de tampão de lavagem C (*Wash Buffer C* ANEXO U), pré-aquecido a 75°C. Após homogeneização, o material foi aquecido por 10 minutos, a 75°C para completar a inativação da polimerase. Após transferência do tubo para o MPC por 2 minutos e remoção do sobrenadante, os “microbeads” foram lavados novamente com tampão de lavagem C (ANEXO U), seguido de quatro lavagens com o tampão de lavagem D (*Wash Buffer D* ANEXO V). Após ressuspensão das partículas magnéticas em 200 µl de 1X tampão 4 (ANEXO W) a amostra foi transferida para tubo novo, que foi novamente lavado com outros 200 µl do mesmo tampão, com auxílio do MPC. Após essa série de quatro lavagens, o cDNA (dupla fita, ressuspendido em 1X tampão 4 ANEXO W) foi armazenado a 4°C, por 16 horas para posteriormente ser digerido pela endonuclease de sequência específica, a *Nla* III.

3.13 A digestão do cDNA com a endonuclease *Nla* III

Após descarte do tampão 4 (ANEXO W), a amostra de cDNA fita dupla foi ressuspensa em 172 µl do tampão LoTE (ANEXO X), 2 µl 100X BSA (ANEXO Y), 20 µl de 10X tampão 4 (ANEXO Z) e 6 µl de *Nla* III (ANEXO AA). A reação foi mantida a 37°C, por 1 hora. Após descarte do sobrenadante, com auxílio do MPC, a endonuclease foi inativada com duas lavagens da amostra com 750 µl de tampão de lavagem C (ANEXO AB), seguida de quatro lavagens com 750 µl do tampão D (ANEXO AC). Após descarte do sobrenadante, com auxílio do MPC, as partículas magnéticas associadas ao cDNA digerido, foram lavadas 2 vezes em 150 µl de 1X tampão para ligase (ANEXO AD). A amostra foi dividida em duas partes de igual volume (A e B). Ambas as amostras foram lavadas com o 1X tampão para ligase (ANEXO AD), ressuspensas no mesmo tampão e imediatamente submetidas à ligação com os adaptadores.

3.14 Ligação dos adaptadores ao cDNA

As amostras dos tubos A e tubo B foram transferidas para o MPC por 2 minutos, descartando-se cuidadosamente o sobrenadante. Os tubos foram então transferidos para gelo, ao qual se adicionaram os seguintes reagentes: 14 µl de LoTE (ANEXO X), 2 µl de 10X tampão para ligase (ANEXO AE). Aos tubos A acrescentou-se 1,5 µl do Adaptador A (40 ng/µl em LoTE), e aos tubos B 1,5 µl de adaptador B (40 ng/µl LoTE). Em seguida, as amostras foram aquecidas (50°C), por 2 minutos. Após resfriamento dos tubos à temperatura ambiente por 15 minutos, as amostras foram transferidas para gelo, à qual adicionaram-se 2,5 µl de ligase de DNA T4 (ANEXO AF). O material foi incubado por 2 horas, a 16°C, antes de serem lavados, quatro vezes, com 500 µl do tampão de lavagem D (ANEXO AC), e duas vezes com 200 µl de 1X do tampão 4 (ANEXO AG). As partículas magnéticas foram ressuspensas nesse último tampão e armazenadas a 4°C, por 16 horas, para o preparo dos dímeros de etiquetas.

3.15 Preparação das etiquetas

Para o preparo dos dímeros de etiquetas, os tubos contendo as amostras (tubo A e B), preparadas como descrito acima foram transferidas para o suporte magnético (MPC) e mantidos nesse por 2 minutos. Após descarte dos sobrenadantes, os tubos foram mantidos em gelo para o preparo da próxima reação. A cada tubo adicionaram-se os seguintes reagentes: 174 µl da solução LoTE (ANEXO X), 20 µl de 10X tampão 4 (ANEXO Z) e 4 µl de 100X BSA (ANEXO Y). Esta solução foi aquecida (65°C), por 2 minutos e em seguida, acrescentou-se à essa, 2 µl de *Bsm*F I (ANEXO AH). Os tubos foram então incubados por 1 hora, a 65°C. Após a digestão, com auxílio

do MPC, os sobrenadantes (compostos pelas seqüências de cDNA e adaptadores) foram transferidos para tubos novos.

3.16 Precipitação com etanol das etiquetas

Ao sobrenadante recuperado na etapa anterior, adicionou-se um volume igual de uma solução de fenol – clorofórmio (1:1) e foram centrifugadas (17968 g) por 5 minutos, à temperatura ambiente. Às fases aquosas transferidas para tubos novos, adicionaram-se 133 µl de 7,5 M de acetato de amônio, 3 µl de *mussel glycogen* (20 mg/ml), e 1000 µl de etanol 100%. Após homogeneização vigorosa, os tubos foram colocados em gelo seco por 10 minutos, antes de serem centrifugados (17968 g), por 50 minutos, a 4°C. Os sobrenadantes foram descartados cuidadosamente e as amostras, lavadas duas vezes com 1000 µl de etanol 70%. Após a remoção do etanol, as amostras foram deixadas à temperatura ambiente, por 5 a 10 minutos, para secagem antes de serem ressuspensas em 10 µl de LoTE (ANEXO X) e incubadas por 3 minutos, à 37°C.

3.17 Unindo os dímeros de etiquetas

A cada tubo (A e B) contendo os 10 µl de LoTE mais os adaptadores e as seqüências de cDNA adicionou-se 5 µl de 10X tampão do fragmento Klenow (ANEXO AI), 1 µl de 100X BSA (ANEXO Y), 2,5 µl de dNTP *mix* (ANEXO Q), 30,5 µl de água milli Q tratada com DEPC e 1 µl da polimerase Klenow (ANEXO AJ). Os tubos foram incubados por 30 minutos a 37°C. O conteúdo de cada tubo foi misturado transferindo-se 50 µl do tubo A para o tubo B. O remanescente do tubo A foi recuperado com 200 µl do LoTE (ANEXO X). Esse volume foi transferido para o tubo B, totalizando assim, 300 µl de amostra final.

3.18 Precipitação com etanol dos dímeros de etiqueta ligados

Ao tubo B contendo os dímeros de etiquetas e os adaptadores, mencionados acima, adicionaram-se um volume (300 µl) de uma solução de fenol – clorofórmio (1:1). Após homogeneização vigorosa, os tubos foram centrifugados (17968 g), à temperatura ambiente, por 5 minutos. A fase superior foi transferida para um tubo novo. Do total, 100 µl foram utilizados como controle negativo (sem ligase). A esse tubo contendo o controle negativo adicionaram-se 100 µl de água tratada com DEPC. Em seguida, às amostras, assim como ao controle negativo, adicionaram-se 133 µl de 7,5 M de acetato de amônia, 3 µl de *mussel glycogen* e 1000µl de etanol 100%. Os tubos foram então transferidos para gelo seco e mantidos assim por 20 minutos, antes de serem centrifugados (17968 g), por 40 minutos, a 4°C. Os precipitados resultantes foram

lavados cuidadosamente, com 1000 µl de etanol 70% gelado. Após secagem, à temperatura ambiente, por 10 minutos, foram ressuspensos em 1,5 µl de LoTE (ANEXO X) e incubados a 37°C, por 15 minutos.

3.19 Ligação dos dímeros

Após precipitação e ressuspensão em LoTE, adicionaram-se às amostras 1,5 µl de 2X *Ditag Reaction Mix* (ANEXO AK). Ao tubo contendo o controle negativo adicionou-se 1,5 µl da solução 2X concentrada do controle negativo sem ligase (*Negative Control Mix*) (ANEXO AM). Tanto as amostras como os controles, foram incubados a 16°C por 16 horas. Após esse período, adicionou-se 14 µl de LoTE (ANEXO X) para cada amostra, em câmara de fluxo laminar.

3.20 Otimização das reações de PCR para os dímeros

As amostras, como o controle, foram diluídas. O controle negativo (sem ligase) foi diluído na proporção de 1/20 e as amostras contendo o produto de ligação, foram diluídas nas proporções de 1/20, 1/40 e 1/80. Esta etapa foi realizada em câmara de fluxo laminar diferente da anterior, para eliminar qualquer contaminação da amostra do controle, uma vez que a reação de PCR é uma reação de amplificação muito sensível. Um controle fornecido pelo kit I-SAGE foi incorporado nas análises, juntamente com um controle, desprovido de cDNA.

Tabela 2– Aliquotas das reações de PCR que foram utilizadas para a otimização das reações de amplificação (continua)

<i>Reações de otimização de PCR</i>						
Reagente	1	2	3	4	5	6
10X Tampão BV (ANEXO AN)	5 µl	5 µl	5 µl	5 µl	5 µl	5 µl
DMSO	3 µl	3 µl	3 µl	3 µl	3 µl	3 µl
dNTP <i>mix</i> (ANEXO Q)	7,5 µl	7,5 µl	7,5 µl	7,5 µl	7,5 µl	7,5 µl
“ <i>Primer</i> ” para dímero de etiquetas- 1 (DTP-1) 175 ng/µl em água DEPC	2 µl	2 µl	2 µl	2 µl	2 µl	2 µl
“ <i>Primer</i> ” para dímero de etiquetas- 2 (DTP-2) 175 ng/µl em água DEPC	2 µl	2 µl	2 µl	2 µl	2 µl	2 µl
Água DEPC	29 µl	29 µl	29 µl	29 µl	30 µl	29 µl

(conclusão)

<i>Reações de otimização de PCR</i>						
Platinum [®] Taq DNA polymerase	0,5 µl	0,5 µl	0,5 µl	0,5 µl	0,5 µl	0,5 µl
Solução de dímeros diluída 1/20	1 µl	-	-	-	-	-
Solução de dímeros diluída 1/40	-	1 µl	-	-	-	-
Solução de dímeros diluída 1/80	-	-	1 µl	-	-	-
I-SAGE TM controle (kit)	-	-	-	1 µl	-	-
Controle negativo diluído 1/20	-	-	-	-	-	1 µl

As diluições dos dímeros foram preparadas numa câmara de fluxo diferente do local onde foram preparadas as reações de PCR. O programa utilizado nas reações de amplificação foi o seguinte: um ciclo de 2 minutos, a 95°C; 27 ciclos de 95°C por 30 segundos, 55°C por um minuto e 70°C por 1 minuto, seguido de uma extensão final a 70°C, por 5 minutos.

3.21 Análise do produto do PCR por meio de eletroforese em gel de poliacrilamida

Após a reação de amplificação, 10% (5 µl) das amostras foram carregadas em um gel de poliacrilamida (12%) para visualização dos produtos de PCR. Este gel foi realizado de acordo com as instruções do manual I-SAGETM (ANEXO AO). A diluição dos dímeros de etiquetas que produzia um padrão de bandas mais próximo do controle do kit I-SAGETM, foi escolhida para ser utilizada nas reações de PCR.

3.22 Amplificação por PCR

Duzentas reações de amplificação por PCR foram utilizadas. Para cada uma das 200 reações de PCR preparou-se uma solução de reagentes (ANEXO AR) de amplificação de 50 µl de volume total. O controle negativo foi preparado pela adição de 6 µl de água estéril na reação de PCR. Em outra câmara de fluxo, uma amostra de 6 µl dos dímeros diluídos foi adicionada à solução de reagentes de PCR. O programa de PCR utilizado foi o mesmo descrito anteriormente.

3.23 Purificação em gel dos dímeros de 100 pb

Os produtos de cada uma das 200 reações foram transferidos para tubos estéreis de 50 ml. A esses, adicionou-se um volume de fenol – clorofórmio (1:1) antes da centrifugação (2400 g), por 10 minutos, à temperatura ambiente. A fase aquosa (superior) de aproximadamente 12,5 ml foi transferida para tubo novo de 50 ml, a qual adicionaram-se 3,2 ml de 7,5 M de acetato de amônia, 72 µl de *mussel glycogen* e 22,8 ml de etanol absoluto. Após homogeneização, os tubos

foram transferidos para gelo seco e mantidos nesse por 20 minutos. Após centrifugação (14000 g) por 30 minutos a 4°C, o sobrenadante foi descartado e o “*pellet*” lavado, três vezes, com 25 ml de etanol 70% gelado. Após secagem, por 15 minutos, o “*pellet*” foi ressuscitado em 300 µl de LoTE (ANEXO X) e incubado por 10 minutos, a 37°C. Em seguida, adicionou-se à amostra 60 µl de 6X tampão de amostra TBE. Para cada amostra 40 µl foram aplicados em gel de poliacrilamida 12%, com 1,5 mm de espessura (ANEXO AO) e 120V até o corante azul chegar a 2 cm do fim do gel.

3.24 Eluição do DNA do gel

Após a eletroforese, o gel foi corado com brometo de etídio na concentração de 0,5-2 µg/ml e as bandas, visualizadas sobre luz UV. As bandas correspondentes a um produto de amplificação de 100 pb foram exisadas do gel, com auxílio de uma lâmina de bisturi. Cada banda foi subdividida em duas partes. As amostras foram centrifugadas (14000 g), por 3 minutos forçando o gel passar por um furo pequeno em um tubo de 0,7ml inserido em um tubo de 1,7ml, para o rompimento do gel em pequenas porções, facilitando assim a dissolução da poliacrilamida.

Os fragmentos de gel foram dissolvidos em 150 µl da solução de LoTE e acetato de amônia numa proporção de 125/25. Após homogeneização, as amostras foram incubadas por 2 horas, a 65°C. O conteúdo de 3 tubos foram combinados para passagem numa coluna S.N.A.PTM. Esse procedimento foi repetido para o restante dos tubos. Após centrifugação (14000 g) por 2 minutos, as frações eluídas foram combinadas e subdivididas em alíquotas de 300 µl em tubos novos e estéreis de 1,7 ml.

3.25 Precipitação com etanol dos dímeros eluídos do gel

Por cada alíquota de 300 µl dos dímeros eluídos do gel, adicionaram-se 133 µl de 7,5 M de acetato de amônia, 3 µl de *mussel glycogen* e 1000 µl de etanol 100%. As amostras foram então mantidas em gelo seco por 20 minutos, centrifugadas (14000 g) a 4°C por 30 minutos. Os precipitados resultantes foram lavados, duas vezes, com 500 µl de etanol 70% gelado, e após secagem à temperatura ambiente por 15 minutos, foram ressuscitados em 14 µl de LoTE (ANEXO X). Todas as alíquotas foram combinadas totalizando 126 µl e armazenadas à -20°C, por 16 horas.

3.26 Clivagem dos dímeros com a endonuclease *Nla* III

As amostras combinadas, como descrito acima, foram divididas em 3 partes iguais (~42 µl), em 3 tubos novos de 1,7ml. A essas amostras adicionaram-se 150 µl do “mix” de reagentes da endonuclease *Nla* III (ANEXO AS). Após a digestão, por duas horas a 37°C o volume final das amostras foi ajustado para 200 µl com LoTE (ANEXO X).

3.27 Precipitação dos fragmentos de digestão

Ao volume final das amostras adicionou-se um volume igual (200 µl) de solução fenol – clorofórmio (1:1). Após homogeneização, as amostras foram centrifugadas (14000 g) por 5 minutos à temperatura ambiente. A fase superior aquosa das amostras foi transferida para tubos novos às quais, adicionaram-se 90 µl de 7,5 M de acetato de amônia, 3 µl de *mussel glycogen* e 850 µl de etanol absoluto. Os tubos foram mantidos em gelo seco por 20 minutos e após centrifugação (12000 g) por 30 minutos a 4°C, os precipitados foram lavados, duas vezes, com 1000 µl de etanol 70%, secos por 10 minutos, à temperatura ambiente e ressuspensos em um volume total de 32 µl de LoTE (ANEXO X).

3.28 Purificação dos dímeros de etiquetas sem adaptadores mediante eletroforese em gel de poliacrilamida

Às amostras ressuspensas 32 µl de LoTE adicionaram-se 7µl de 6X TBE. Do total de 39 µl, aplicou-se 13 µl no gel de poliacrilamida 12% (ANEXO AO). Após a corrida de 120v até o corante azul chegar a 2 cm do fim do gel, o mesmo foi corado com brometo de etídio (0,5 µg / ml) e a visualização das bandas foi feita sobre luz UV.

3.29 Extração dos dímeros a partir do gel de poliacrilamida

Após a visualização do gel, as bandas de tamanho esperado foram retiradas do gel para eluição do DNA da mesma forma como descrito acima. Após centrifugação (14000 g) por 2 minutos, as frações eluídas foram combinadas e subdivididas em alíquotas de 150 µl e transferidas para tubos novos e estéreis de 1,5 ml.

3.30 Precipitação com etanol

As amostras de DNA foram precipitadas da mesma maneira como descrito anteriormente e ressuspensas em 7,75 µl de LoTE (ANEXO24) e mantidas a 4°C.

3.31 Reação de ligação

Às amostras ressuspensas em 7,75 µl de LoTE, adicionaram-se 1 µl de 10X tampão de ligase (ANEXO AL) e 1,25 µl da ligase de DNA T4 (ANEXO AF). Os tubos foram mantidos por 3 horas a 16°C. Em seguida, adicionaram-se 2 µl de 6X de tampão de amostra TBE, e os tubos foram mantidos por 10 minutos, a 65°C.

3.32 Eletroforese em gel

As reações de ligação foram centrifugadas (14000 g) brevemente e aplicadas em gel de poliacrilamida 8% (ANEXO AT). Após a corrida a 120v até o corante azul chegar a 2cm do fim do gel, o mesmo foi corado com brometo de etídio (0,5 µg/ml) e as bandas visualizadas sobre luz UV.

3.33 Extração dos concatâmeros do gel de poliacrilamida

As bandas correspondentes a 400-1000 pb foram excisadas do gel. Cada banda foi subdividida em 2 tubos. As amostras foram centrifugadas (14000 g) por 6 minutos, antes da adição de 200 µl de uma solução de LoTE (ANEXO X) e acetato de amônia, 5:1 (v/v) e incubadas por 2 horas. O conteúdo dos dois tubos foi combinados, antes da passagem em colunas S.N.A.PTM. As colunas contendo as amostras foram centrifugadas (14000 g) por 2 minutos e as amostras, subdivididas em alíquotas de 200 µl.

3.34 Precipitação

As amostras foram precipitadas seguindo-se o mesmo procedimento da Precipitação com etanol, posterior à clivagem com a endonuclease *Nla* III, citado anteriormente. Após à precipitação, os “*pellets*”, após secagem à temperatura ambiente, por 20 minutos, foram ressuspensos em 6 µl de LoTE (ANEXO X).

3.35 Linearização do plasmídeo pZErO[®]-1

Para clonar o concatâmero foi necessário linearizar o plasmídeo, assim, o pZErO[®]-1 (ANEXO AU) foram digeridos com a nuclease *Sph* I (ANEXO AW). À amostra, de 2 µl de pZErO[®] adicionou-se 1, 2,5 µl de 10X tampão 2 (ANEXO AX), 1,4 µl da endonuclease *Sph* I e 19,1 µl de água estéril. A reação foi mantida por 25 minutos, a 37°C. Em seguida, adicionou-se 175 µl de LoTE (ANEXO X) e um volume igual de uma solução de fenol:clorofórmio (1v:1v). As amostras foram centrifugadas (14000 g) à temperatura ambiente por 2 minutos. As fases aquosas foram transferidas para tubos novos às quais adicionaram-se 65 µl de acetato de amônia

(7,5 M) e 600 µl de etanol absoluto. Os tubos foram então mantidos em gelo seco, por 15 minutos antes da centrifugação (14000 g) por 30 minutos a 4°C. Os “*pellets*” foram lavados, duas vezes, com etanol 70% gelado e após secagem à temperatura ambiente por 10 minutos, foram ressuspensos em 50 µl de LoTE (ANEXO X).

3.36 Ligação do concatâmero ao pZErO[®]-1

Todas as reações de ligação foram preparadas em gelo. A 6 µl da solução de concatâmeros adicionou-se 1 µl de pZErO[®]-1 linearizado, como descrito acima, 1 µl de 10X tampão de ligação, 2 µl de ligase DNA T4. A reação foi mantida por 3 horas, a 16°C. Em seguida, o volume foi ajustado para 200 µl com LoTE (ANEXO X), ao qual adicionou-se um volume igual de uma solução de fenol:clorofórmio (1v:1v). Após a centrifugação (14000 g) a 4°C por 30 minutos, os “*pellets*” resultantes foram lavados, duas vezes, com etanol 70% gelado, secos à temperatura ambiente por 15 minutos e ressuspensos em 12 µl de LoTE.

3.37 Preparo do meio de cultura LB seletivo

Esse meio de cultura, com baixa concentração de sal foi preparado com triptona 1%, extrato de levedura 0,5%, NaCl 0,5% e agar 1,5% e o pH foi ajustado para 7,5 com NaOH (1N). Após autoclavagem (1 atm, 121°C) por 20 minutos, o meio foi mantido a 55°C, por 20 minutos antes da adição do antibiótico seletivo, ZeocinaTM na concentração final de 50 µg/ml, a ser utilizado na seleção dos clones recombinantes.

3.38 Transformação de *E. coli* One Shot[®] TOP10 Electrocompetente

Células (50 µl) da linhagem de *E. coli* One Shot TOP10 eletrocompetentes (Invitrogen) foram utilizadas para os experimentos de transformação, por eletroporação. A essas adicionaram-se 1-2 µl da ligação, concatâmeros clonados no plasmídeo pZErO[®]-1. A combinação das células mais o plasmídeo recombinante foi transferida para uma cubeta de eletroporação 0,1 cm, gelada, que foi submetida a um pulso elétrico (2,5kV, 4 milli-segundos). Após a eletroporação, adicionou-se às culturas 250 µl de meio SOC (ANEXO AY), à temperatura ambiente e foram mantidas sob agitação (200 rpm), à 37°C, por 1 hora. Em seguida, adicionou-se a essas, 750µl de meio SOC, antes de plaquear a alíquota de 100 µl em placas de meio LB seletivo, acrescido de ZeocinaTM, como descrito anteriormente. As placas foram mantidas a 37°C, por 18 horas.

Os clones de *E. coli* TOP10 recombinante, contendo os concatâmeros, resistentes ao antibiótico Zeocina foram selecionados. Os plasmídeos desses clones foram isolados para seqüenciamento dos concatâmeros correspondentes.

3.39 Seqüenciamento

O seqüenciamento dos plasmídeos recombinantes, foi realizado em colaboração com o Laboratório de Biologia Celular e Molecular do CENA-USP sob a supervisão da Dra. Siu Mui Tsai, com auxílio do ABI PRISM 3100 “*genetic analyser*”, que consiste de uma plataforma automatizada com capacidade de analisar 16 amostras simultaneamente por eletroforese capilar. As amostras de DNA foram seqüenciadas utilizando BigDye® Terminator v3.1, desnaturadas a 95°C, por 5 minutos em termociclador “Gene Amp 9700” e mantidas em gelo por 2 minutos, antes de serem transferidas para o seqüenciador 3100 (Applied Biosystems).

3.40 Análise das seqüências apropriadas com Phred-Phrap

Softwares como Phred, Phrap e Consed (EWING; GREEN, 1998; EWING et al., 1998; GORDON; ABAJIAN; GREEN, 1998) para a montagem de seqüências são uma necessidade básica em projetos de seqüenciamento, uma vez que as tecnologias de seqüenciamento, geralmente geram moléculas de aproximadamente 400-500 pb de cada amostra de DNA. Portanto, para a obtenção de moléculas maiores, ou seja, uma seqüência completa, é necessário fazermos a montagem desses fragmentos, gerando assim, o que se denomina um *contig*. Cada *contig* é formado por seqüências de *reads* sobrepostas. Estes “*softwares*” de montagem juntam estes *reads*, por identidade entre seqüências, as agrupam em *contigs*, determinando assim uma seqüência concenso, que representa uma molécula de DNA original, fragmentada.

Entre os “*softwares*” mais comumente utilizados, o programa Phred (EWING; GREEN, 1998; EWING et al., 1998) analisa os arquivos de seqüências gerados pelo seqüenciador automático produzindo os cromatogramas. Este programa lê dados de sinais de fluorescência oriundos de seqüenciadores de DNA, define as bases e atribui a essas, valores de qualidade das bases da seqüência de DNA (EWING et al, 1998).

Os arquivos do programa utilizado geraram cromatogramas no formato ABI (Applied Biosystems) do seqüenciamento dos concatâmeros. O programa Phred determinou as posições dos picos ideais para os sinais das 4 diferentes fluorescências (uma para cada base A, T, C e G). Os fatores mais importantes que ele mostrou foram o ponto de informação do conjunto de dados

e o tempo em que este ponto de informação foi gerado pelo seqüenciador, o que permite prever uma série de localizações homogêneas espaçadas no tempo. Na segunda fase, o programa analisa cada sinal para determinar o centro de cada pico de fluorescência, real, observado. Esses picos são detectados de forma independente, podendo assim, gerar uma sobreposição de alguns desses. Na terceira fase, os picos detectados de forma independente são associados às localizações ideais previstas no primeiro passo. Na fase final, ele avalia todos os parâmetros de qualidade para quantificar o sinal (EWING et al., 1998; EWING; GREEN, 1998).

As seqüências geradas pelo seqüenciador são examinadas pelo programa, phred (EWING et al., 1998; EWING; GREEN, 1998), isoladamente. Como padrão se estabeleceu que as seqüências a serem utilizadas para análise, mediante o software ou programa SAGE 2000TM, deviam ser maiores de 300 bases contínuas, de alta qualidade com um $q=20$, o que representa uma precisão de 99%. Assim, admitiu-se no seqüenciamento, um erro a cada 100 bases seqüenciadas (Figura 9) (MADEIRA, 2004).

$$q = -10 \times \log_{10} (p)$$

Figura 9 – Formula de cálculo da qualidade média - q . Onde p representa a probabilidade de erro estimado para as bases analisadas (EWING; GREEN, 1998)

Isto permitiu a separação de uma série de cromatogramas para re-seqüenciamento, como exemplificado na Figura 10. O padrão de qualidade está determinado pela fórmula acima detalhada.

Figura 10 – (a) Cromatograma representativo de uma sequência que foi rejeitada. (b) Cromatograma representativo de uma sequência com alta qualidade

Com auxílio do sistema operativo *linux* e Interface gráfica *Konqueror*, o programa *phred* foi executado para cada sequência isoladamente. Isto se realiza mediante a linha de comando (MADEIRA, 2004) em detalhes, passo a passo, abaixo:

<phred>

Após esse comando, escreve-se o nome do arquivo ou arquivos que o programa vai analisar. Os arquivos devem estar num diretório denominado *chromat_dir*, que se encontrava dentro de outro diretório com o nome do usuário, dentro de *home*. Para a execução do programa *phred* outros diretórios foram criados dentro do diretório *analysis*. Estes são *edit_dir* e *phd_dir*, e o programa foi executado dentro do diretório *edit_dir*.

Dentro do diretório *edit_dir* digita-se:

<phred -id ../chromat_dir -trim_alt "" -trim-cutoff 0.01 -sa trimteste>

Estes parâmetros que são introduzidos determinam o seguinte:

- id ../chromat_dir: Define o diretório onde estão os arquivos que serão processados pelo comando *phred* (por exemplo: *home/guillermo/analysis/chromat_dir*).
- trim_alt "": Realiza um corte para localizar a região de alta qualidade de cada read. Utiliza para isto o parâmetro de qualidade definido pelo *-trim_cutoff*.
- trim_cutoff: Define a probabilidade de erro (da equação de qualidade do *phred*) e vai ser utilizada pela opção *-trim_alt*. O valor default é 0,05.

-as: Grava o resultado do processamento do phred de todos os reads em um arquivo com o nome “trimteste”.

Para checar os dados da etapa anterior digita-se

<more trimteste>

```
>HS03-FP-00-000-009-G08.g2      324      234      89   ESD
TCGCGGCAGCATCGCTGGAGTCCGCTGTCGGTCTGGCAGACGGCTCGGCA
CCCAGAGCACATGAACATGAACGAAGTGCTGGCAACCGCGCCTCGGAACT
CTTGGGCGGGTGCGTGGCGAGGATCGCTGATCCAACCCAATGACGACGTC
AACCGTGGCCAGTCCTCCAACGATATCTTCCCCACCGGCATGCACGTGGC
CGCCTGCATGGCCGTGGCGACCCACCTGATCCCGGCGCTGCACACGCTGC
GCGGCACGCTGGCGGAGAAATCCAGCCAGTTCCGCCGACATCGTCAAGATC
GGTCGCACCCACCTGCAAGACGCT
```

Figura 11 – Arquivo de saída do software *phred* que permite identificar seqüências de DNA de alta qualidade

3.41 Análise das seqüências por SAGE 2000™

As seqüências que não foram aceitas pelo programa Phred foram re-seqüenciadas e novamente analisadas até conseguir-se seqüências com a qualidade desejada. Toda arquivo (abi) que passou pela análise do programa *phred* o arquivo correspondente de texto (.seq) com a seqüência em bases foi analisada com o programa SAGE 2000™.

O software SAGE 2000™ versão 4.5 é um programa apropriado para análise dos dados de seqüenciamento provenientes dos concatâmeros inseridos nos plasmídeos. O programa extrai as etiquetas SAGE™ de cada clone seqüenciado. Para isto, o programa importa arquivos de seqüência com a extensão (.seq), identifica e contabiliza a ocorrência de cada etiqueta com esses dados, no final, gera um arquivo de saída do conjunto das etiquetas e suas respectivas abundâncias, utilizando a seqüência de reconhecimento da endonuclease *Nla* III (CATG), como referência (Figura 12). A freqüência com que cada etiqueta aparece determina o nível de expressão da mesma, refletindo assim, a abundância do mRNA original no tecido vegetal amostrado.

Um dos problemas dos dados produzidos pela metodologia SAGE é a análise das etiquetas de cópia única que também são chamadas de “*singletons*”. Este tipo de etiquetas pode representar um gene de baixa expressão ou simplesmente um artefato produzido pela amplificação por PCR ou também um erro de seqüenciamento. Este problema se complica um

pouco, quando não se pode encontrar uma sequência que possua uma determinada etiqueta em nenhuma base de dados. Nas análises das bibliotecas de *Eucalyptus grandis*, resistente e susceptível à ferrugem, as frequências dos singletons, apesar de não terem sido ignoradas, não fizeram parte da caracterização das etiquetas.

Figura 12 – Interface do SAGE2000 utilizada na comparação entre as duas bibliotecas de *Eucalyptus grandis*: susceptível e resistente à ferrugem

O *software* SAGE2000 apresenta uma série de dados após as análises das sequências, que permitiram identificar cada uma das etiquetas, a abundância das mesmas, as classes em que elas se encontram, a abundância das respectivas classes (Figura 13) e o número total de etiquetas diferentes dentro de cada biblioteca (Figura 14).

Figura 13 – Representação da distribuição das etiquetas geradas pelo *software* SAGE2000, o agrupamento por classes, os componentes que são agrupados e a abundância total desta classe

Figura 14 – Número de etiquetas diferentes que foram encontradas dentro da biblioteca resistente representando os genes que estão sendo expressos

3.42 Classificação das etiquetas mediante o teste estatístico Z

O arquivo de saída gerado com o programa SAGE 2000 foi organizado pela classificação das etiquetas. O critério utilizado nessa classificação foi a abundância das etiquetas, ou seja, as diferenças no nível de expressão dessas nas duas bibliotecas (susceptível e resistente à ferrugem).

O teste escolhido para a análise estatística desses dados foi o proposto por Kal et al. (1999), também chamado de teste Z, utiliza as proporções de etiquetas específicas em cada biblioteca. Estas proporções de etiquetas seguiriam uma distribuição binomial. O mesmo é calculado como a diferença observada entre as proporções de etiquetas específicas, nas duas bibliotecas, dividido pelo erro padrão desta diferença. Quando a H_0 é verdadeira: O valor estatístico Z calculado com as proporções de uma determinada etiqueta, vai cair dentro dos valores de aceitação. Por outro lado quando a H_0 é falsa o valor Z calculado vai cair dentro dos valores críticos ou níveis de confiança ou significância bilateral (RUIJTER et al., 2002).

Man et al. (2000) comparou os teste de χ^2 , o teste proposto por Audic e Claverie (1997) e teste exato de Fisher. Para esta comparação ele utilizou o método de simulação de Monte Carlo para produzir possíveis distribuições de frequências de bibliotecas SAGE. Ele avaliou os três métodos em especificidade, poder e robustês. Esta comparação mostrou que o teste de χ^2 era consistentemente mais poderoso e muito mais robusto que os outros métodos. Estas diferenças eram mais marcadas quando o nível de expressão era baixo, menor que 15 cópias num total de 50000 etiquetas por biblioteca. Este teste χ^2 quando tomado de forma bilateral é equivalente do teste Z. Portanto este método é o mais indicado para a avaliação dos experimentos de SAGE (MAN et al., 2000) e foi o escolhido para as análises de nossas bibliotecas.

A fórmula de cálculo do mesmo se encontra detalhado na eq. (1):

Eq. 1

$$Z = \frac{p_1 - p_2}{\sqrt{p_0(1-p_0) \left(\frac{1}{N_1} + \frac{1}{N_2} \right)}}$$

3.43 Comparação com a base de dados Genolyptus

As seqüências das etiquetas, categorizadas com o teste Z foram procuradas na base de dados do Genolyptus. Esta é uma base de dados de ESTs de diferentes espécies de *Eucalyptus*: *Eucalyptus globulus*, *E. grandis*, *E. pellita* e *E. urophila*. Os 70893 ESTs presentes neste banco,

distribuídos em 23292 contigs têm procedência de diferentes tecidos vegetais, de diferentes idades.

O contraste das seqüências das etiquetas e as seqüências de reconhecimento da endonuclease *Nla* III contra as seqüências do Genolyptus foi realizado com auxílio da ferramenta de bioinformática BLAST-N (ALTSCHUL et al, 1997) (Figura 15). Estas seqüências deviam ser encontradas nos contigs pertencentes ao Genolyptus juntas as seqüências de corte (CATG) e devia ser próxima a ultima seqüência de corte (CATG) antes da cauda poli A, na posição 3'. A presença de alguma seqüência de corte intermediária entre a seqüência da etiqueta e a cauda poliA, ou a falta total ou parcial da seqüência de corte ou da etiqueta foram características desclassificadoras dos contigs (Figura 16).

Figura 15 – Interface gráfica da base de dados do Genolyptus, que inclui a ferramenta BLAST N de pesquisa de todas as bases presentes no Genolyptus, juntamente com as seqüências de corte da endonuclease e a seqüência da etiqueta de 10 bases

```

>Contig1798ur-xy·(whole·contig)¶
CTTCAAGTACCACTTCTTCACTGTCTCTCTTCAATCCCCCTCTTACCAAGAGCAAAGCAAAA
AACCATGATGAAGAGACTGTCAATTTCTGCTCCTACTGGTCTGCTCTTCCAATGCTCTACCACTT
GGCTCAGCCTGCGGCCGCCAGCTCGCCTGTGATAGCCCCGGCTGCACCTGCTACGCTGCCT
TAGGCCCGGCTCCTCTCTTAGGCCAGCTCCTGCAGGCCAACCGACATCAGGAAGGTCCCTC
AAGAAGGCGAGCCAAATTTACGGTGTCTCAAGCTCATGAGGGCCACCTCAGTCGACAAACAGAT
CAACGCGCAACTTAACAACCTGAACAATGGCCTAACCATCTTCGCCCCGAATGACAATGCCCTTCT
CGAGCCTCAAGTCTGGGACTTTGAACAGCTCTCGTCTGAGCAGCAGGTTGAGCTCGTGCAGTTC
CATGTAGTCCCGACATACTTGGCTGTGCCCAATTCAAACAGTCAGCAATCCGCTGAGGACACA
AGCAGGCGGGAGCGGGAAGTTCGAGTTCCTCACTGACTTTGACTACATCCGGTAGCAGTGTGAACA
TATCGACTGGTGTCAACCAATGCAACAGTGGACCAACAGTATATAACGATGGCCAGTTGGCTGTG
TATATGGTCGACAAGTTCTGCTTCCAATGAGCATATTCGGCCCGAAGCCACCTGCGCCGGCGCC
TGGCCAGCAAGCCAAAGAAGAAGCCTGCCTTGGCATCAAGCCCGGTAGCCGACAGCGTCCCGG
TGGACACATCTGGTGAAGGGAATGATGGTCCCTGGAATGATGGTGTTCATTGGATTGAGTGT
GTTGGGACATTTCTACTCGTGACGAGGAACATTTGGGACATCACGAGAATGATCAGCTTGAACAC
CGAGGTGGGGTCACTTTTGGGTCTCCTATGAGCTATTATTTGGTGAATTTGGTTAGAGGGCAG
TGAGAGATGTGGTCAGTGATTTGAAATAAATTCGAGTCCGATGGTAGAACACGCCAGTGATGAG
AGAAAAAGAGAAAAAGAGAGTCAAGAGTCTGTTCATTTTCCTTTTCCTTTCCCTTTCCCTTCCCTG
TGTTCAGTCTTTCAGTTTTCATGTACTTATTGTTGTGTCTTGTGAGGAAGTGTGTATGTT
CCTGGATTGTGAATTGATCTTCTTCTTCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA¶

```

Figura 16 – Sequência de um determinado contig pertencente ao Genolyptus que apresenta todas as características necessárias para ser selecionado: (amarelo) sequência da etiqueta de (azul) reconhecimento de corte da endonuclease em posição proximal à extremidade 3', cauda poli A e ausência de sequências da endonuclease entre as últimas duas

3.44 Comparação entre as sequências do Genolyptus e do NCBI

As sequências dos contigs resgatadas do Genolyptus que apresentavam as sequências de reconhecimento de endonucleases e etiquetas nas posições corretas foram comparadas com a base de dados do NCBI – National Center for Biotechnology Information, com auxílio das diferentes ferramentas de bioinformática: Basic Local Alignment Search Tool – BLAST, em diferentes variantes BLAST N, BLAST X e “ORF finder” pertencentes ao banco de dados públicos do NCBI (www.ncbi.nlm.nih.gov/SAGE) (Figura 17).

A análise de BLAST possibilita comparações entre sequências, por regiões de similaridade dentro dos bancos de dados (ALTSCHUL et al, 1997). Uma pesquisa do BLAST num banco de dados de sequências pode produzir dezenas de alinhamentos. As informações que oferece primariamente são pontuações brutas, pontuações de “bits” e valores *E*. As pontuações brutas de um alinhamento local, muitas vezes, não podem ser comparadas diretamente porque as matrizes de pontuação podem ser diferentes. As pontuações de “bits” são pontuações brutas convertidas a partir da base de registro da matriz de pontuação que criou o alinhamento. Isto permite que as pontuações de “bit” sejam comparadas entre diferentes alinhamentos. Por outro lado, os valores *E* fornecem informações sobre probabilidades. Indicam o número de alinhamentos que podem ser encontrados com uma pontuação maior ou igual à pontuação de alinhamento observada. Valores *E* altos indicam que os alinhamentos são produzidos pelo acaso. Valores de 0,1 ou 0,05 são usados como limites em pesquisas em bancos de dados (GIBAS; JAMBECK, 2001). *E values* em torno de $x=E^{-20}$ foram tomados como padrão. Assim, *E values*

inferiores ($\leq E^{-20}$) foram tomados como padrão de seleção dos dados para anotação por homologia. Valores superiores foram desconsiderados.

The screenshot shows the NCBI BLAST interface. At the top, there's a navigation bar with 'Nucleotide', 'Protein', 'Translations', and 'Retrieve results for an RID'. The 'Nucleotide' tab is selected. Below the navigation bar, there's a text input field containing a nucleotide sequence: '>Contig1798ur-xy (whole contig) CTTCAGTACCACCTTCTTCACTGCTCTCTTCACAATCCCCCTCTTACCAAGAGCAAA GCAAAAACATGATGAAGAGACTGTCATTTCTGCTCCTACTGGTCTGCTCTTCCAATG CTCTACCACCTTGGCTCAGCCTGCGGCCGCCCAAGCTCCGCTGTGATAGCCCCGGCTG CACCTGCTACGCTGCCTTAGGCCGGCTCCTCTGTCTTAGGCCAGCTCCTGCAGGC'. Below the sequence, there's a 'Set subsequence' section with 'From:' and 'To:' input fields. There's also a 'Choose database' dropdown menu set to 'nr'. At the bottom, there's a 'Now: BLAST!' button, along with 'Reset query' and 'Reset all' buttons.

Figura 17 – Interface gráfica da análise de BLAST, da base de dados do NCBI, de uma determinada sequência resgatada do Genolyptus

Outra ferramenta utilizada foi a de *ORF – Finder* (“*Open Reading Frame*” - Quadro de leitura aberto), que localiza quadros de leitura abertos em uma determinada sequência de DNA. Esta ferramenta traduz seqüências nos seis quadros de leitura possíveis. Quando o quadro não é o apropriado aparecem com frequência códons de interrupção. Seqüências longas interrompidas, sem estes códons são consideradas regiões com alta probabilidade de serem codantes (GEBAS; JAMBECK, 2002) (Figura 18).

The screenshot shows the NCBI ORF Finder (Open Reading Frame Finder) interface. The page has a blue header with the NCBI logo and the title 'ORF Finder (Open Reading Frame Finder)'. Below the header, there's a navigation bar with 'Published', 'Criteria', 'BLAST', 'ORF', 'Translation', and 'Statistics'. The 'Criteria' tab is selected. The main content area contains a description of the tool: 'The ORF Finder (Open Reading Frame Finder) is a graphical analysis tool which finds all open reading frames of a selectable minimum size in a user's sequence or in a sequence already in the database. This tool identifies all open reading frames using the standard or alternative genetic codes. The deduced amino acid sequences can be saved in various formats and searched against the sequence database using the WWW BLAST server. The ORF Finder should be helpful in preparing complete and accurate sequence submissions. It is also packaged with the Seqin sequence submission software.' Below the description, there's a form with two input fields: 'Enter GI or ACCESSION' and 'or Sequence in FASTA format'. There are also 'Go' and 'Clear' buttons.

Figura 18 – *ORF finder*: ferramenta de busca de quadros de leitura abertos codantes pertencente ao site do NCBI

3.45 Anotação das ESTs identificadas com auxílio de outros bancos de dados públicos

Só as funções de uma pequena parte das proteínas conhecidas foi determinada por meio de experimentação. Portanto, o uso de ferramentas computacionais de análises de sequência é essencial para a anotação ou atribuição de funções a novos genes ou genomas (SCHULTZ et al., 1998).

Portanto, a posterior caracterização das seqüências procedentes do Genolyptus foram utilizadas várias bases de dados públicas e suas respectivas ferramentas computacionais de análise de seqüência tais como (Figura 19):

- CDD - *Conserved Domain Database* (Base de dados de domínios conservados) (MARCHLER-BAUER et al., 2003; MARCHLER-BAUER et al., 2005);
- Pfam (BATEMAN et al, 2004; BATEMAN et al, 2002; BATEMAN et al., 2000; BATEMAN et al., 1999; SONNHAMMER, et al., 1998);
- COG - *Clusters of Orthologous Groups of proteins* (TATUSOV; KOONIN e LIPMAN, 1997)
- SMART - *Simple Modular Architecture Research Tool* (Ferramenta de pesquisa de arquitetura modular simples) (SCHULTZ et al., 1998) e o,
- KEGG - *Kyoto Encyclopedia of Genes and Genomes* (Enciclopédia de genes e genomas de Kioto) (KANEHISA; GOTO, 2000; GIBAS; JAMBECK, 2001).

Todas são bases de dados que armazenam informações de seqüências protéicas. Para realizar a busca por similaridade, as seqüências de nucleotídeos tiveram que ser traduzidas com o auxílio da ferramenta BLAST do NCBI. Este permite traduzir as seqüências e procura seqüências com similaridade nas bases acima mencionadas.

As proteínas apresentam com freqüência uma estrutura modular e estes são denominados como domínios, cada um com diferente função e origem evolutiva. A bioinformática determina os domínios conservados mediante a presença de padrões de aminoácidos. O Conserved Domain Database (CDD) é uma coleção de alinhamentos de múltiplas seqüências para proteínas completas e domínios ancestrais. É utilizado para identificar os domínios conservados presentes nas seqüências traduzidas dos “contigs” pertencentes ao Genolyptus (MARCHLER-BAUER et al., 2003; MARCHLER-BAUER et al., 2005).

Já, o Pfam é também um banco de dados de alinhamentos de famílias de domínios protéicos, que agrupa alinhamentos múltiplos de proteínas e perfis to tipo “Hidden Markov

Model” – HMM dessas famílias protéicas. É formado por dois bancos de dados: Pfam-A e Pfam-B. O Pfam-A é um banco de dados composto por mais de 2700 perfis. A maior parte abrange domínios inteiros de proteínas e está baseado em pesquisas HMM. De forma semelhante que a ferramenta BLAST do NCBI, ele apresenta *E-values*, que representam a probabilidade de se obter uma sequência semelhante pelo acaso. Um E-value muito bom deve ter um valor muito menor que 1. Já, um *E-value* com valor próximo de 1 indica que a semelhança entre as sequências pode ser devida ao acaso. Por outro lado, o Pfam-B agrupa, automaticamente, os perfis das sequências que restaram do Pfam-A (BATEMAN et al, 2004; BATEMAN et al, 2002; BATEMAN et al., 2000; BATEMAN et al., 1999; SONNHAMMER, et al., 1998).

O KEGG é um recurso que fornece uma informação global relativa ao metabolismo com ilustrações em mapas de vias metabólicas e não somente texto. Também fornece listas de números EC (código enzimático) e de suas enzimas correspondentes, combinando assim informações de genômica e de ordem funcional (KANEHISA; GOTO, 2000; GIBAS; JAMBECK, 2001).

O COG como base, reúne informações de genômica comparativa e classificação protéica, em uma perspectiva do sistema natural. Agrupa genes conservados a partir de clados filogeneticamente distantes, mas com funções/sequências conservadas, o que permite, ao classificarmos uma proteína dentro de um grupo, fazermos umas predições funcionais automática, que corresponderia a regiões ancestrais conservadas (TATUSOV; KOONIN; LIPMAN, 1997).

O SMART é uma base de dados de alinhamentos múltiplos pertencentes a 86 domínios, que são encontrados nas proteínas sinalizadoras. Este permite uma rápida identificação e anotação de domínios de sequências sinalizadoras e é capaz de determinar a arquitetura modular de sequências simples ou genômicas (SCHULTZ et al., 1998).

Figura 19 – Outros bancos que foram pesquisados para a anotação das seqüências provenientes do Genolyptus. a) Base de dados do Pfam. b) Base de dados do COG. c) Base de dados do KEGG. d) Base de dados do CDD. e) Base de dados do SMART

4 RESULTADOS E DISCUSSÃO

4.1 SAGE

A metodologia SAGE permitiu que se sequenciassem um total de 1154 e 1306 clones para a biblioteca resistente e susceptível, respectivamente. Este sequenciamento produziu um total de 33263 e 42256 etiquetas para as bibliotecas susceptível e resistente, respectivamente. Para evitar a contagem de etiquetas geradas da amplificação por PCR tendenciosa, os dímeros de etiquetas repetidos em cada biblioteca foram eliminados. Produzindo em consequência um total de 71609 etiquetas para os tecidos de folha. Portanto, o perfil transcricional de folhas de *Eucalyptus grandis* encontra-se dividido em 2 bibliotecas, de 31645 e 39964 etiquetas pertencentes às bibliotecas susceptível e resistente, respectivamente (Figura 20). Estas etiquetas foram distribuídas em 17366 etiquetas diferentes ou possíveis genes para a biblioteca de folhas resistentes e 13915 etiquetas diferentes ou possíveis genes para a biblioteca de folhas susceptíveis (Figura 21). Apresentando assim uma eficácia final de 24,2 e 34,6 etiquetas por clone, para a biblioteca susceptível e resistente, respectivamente.

Figura 20 – Representação de etiquetas sequenciadas e como estas se encontram distribuídas nas duas bibliotecas

Figura 21 – Etiquetas totais por biblioteca e as etiquetas que foram identificadas como diferentes umas de outras pelo *software* SAGE2000 dentro de cada biblioteca

4.2 Distribuição de frequências de expressão

Os genes foram agrupados pelas etiquetas que geraram e a frequência com que estas etiquetas se apresentaram nas bibliotecas. O *software* SAGE2000 agrupou as etiquetas em categorias denominadas classes. Estas classes foram formadas por diferentes etiquetas que apresentavam a mesma abundância. Por exemplo, o conjunto de etiquetas que apresentava uma abundância de 25 cópias na biblioteca resistente formou uma classe. Esta classe agrupou, portanto seis etiquetas diferentes com uma abundância de 25 cópias cada uma. A abundância total desta classe foi de 150 etiquetas. As etiquetas com seqüências diferentes representam putativos ou possíveis genes. O SAGE2000 fornece também o número de etiquetas diferentes identificadas dentro de cada biblioteca. Estas que representariam os possíveis genes que estariam se expressando no tecido no momento da infecção.

Os conjuntos de dados referentes a ambas bibliotecas foram distribuídos em 4 categorias segundo a frequência de ocorrência das etiquetas. Seguindo as distribuições de Lorenz e Dean (2002), estas categorias constituem agrupamentos de várias das classes geradas pelo SAGE2000, que denominaremos categorias. Pode-se encontrar a categoria das etiquetas de alta expressão (etiquetas com 20 ou mais cópias), de expressão intermediária (etiquetas com abundâncias entre 5 e 19 cópias), de baixa expressão (etiquetas com abundâncias entre 4 e 2 cópias) e as etiquetas de cópia única ou *singletons*. As frequências relativas destas categorias foram representadas pelas

porcentagens dos números de etiquetas diferentes dentro de cada categoria em relação com as etiquetas diferentes identificadas dentro da biblioteca respectiva. As abundâncias das categorias foram calculadas como a somatória de cópias de cada etiqueta pertencente à categoria, estando também representada pelas suas porcentagens. Isso dá uma idéia da representatividade de cada categoria no número total de tags sequenciadas. A distribuição das etiquetas se encontra resumida na tabela 3.

Tabela 3 – Resumo da distribuição de freqüências das etiquetas geradas pela metodologia SAGE

<u>Freqüência</u>	≥ 20	5 até 19	2 até 4	=1	Total
<u>Folhas resistentes</u>					
<i>Etiquetas diferentes</i>	156 (0,88%)	1186 (6,83%)	3871 (22,3%)	12153 (69,97%)	17366 (100%)
<i>Abundância da categoria</i>	8467 (21,19%)	9645 (24,13%)	9699 (24,27%)	12153 (30,41%)	39964 (100%)
<u>Folhas susceptíveis</u>					
<i>Etiquetas diferentes</i>	135 (0,97)	906 (6,51%)	3054 (21,95%)	9820 (70,57)	13915 (100%)
<i>Abundância da categoria</i>	6966 (22,01%)	7235 (22,86%)	7624 (24,09%)	9820 (31,03%)	31645 (100%)

A categoria de etiquetas de alta expressão (expressão ≥ 20) está representada por 0,88% e 0,97% do total de etiquetas diferentes identificadas, nas bibliotecas de folhas resistentes e susceptível, respectivamente. As etiquetas com expressão intermediária (5-19 cópias) apresentam uma freqüência relativa de 6,83% na biblioteca de folhas resistentes e 6,51% na biblioteca susceptível. As etiquetas de baixa expressão (2-4 cópias) apresentam uma freqüência relativa de 22,3% e 21,95% das etiquetas identificadas, nas bibliotecas resistente e susceptível, respectivamente. Por último, a categoria das etiquetas representadas por 1 cópia se encontra representada por uma freqüência relativa de 69,97% na biblioteca resistente e 70,57% na susceptível.

Com relação à abundância, a categoria de etiquetas de alta freqüência (≥ 20) agrupa 21,19% e 22,01% do total de etiquetas (39964 e 31645 etiquetas), nas bibliotecas de folhas resistentes e susceptíveis, respectivamente. A categoria de etiquetas de expressão intermediária

(19-5 cópias), representa 24,13% e 22,86% do total das etiquetas, nas bibliotecas resistente e susceptível, respectivamente. Já a categoria das etiquetas de baixa expressão (4-2 cópias) agrupa 24,27% e 24,09% do total de etiquetas, nas bibliotecas resistente e susceptível, respectivamente, e a categoria das etiquetas representadas por uma única cópia representa, respectivamente, 30,41% e 31,03% do total de etiquetas das bibliotecas resistente e susceptível.

Os valores relativos produzidos por ambas bibliotecas foram muito próximos como pode se ver nos dados da Tabela anterior, mostrando uma distribuição coerente das etiquetas dentro das categorias em ambas bibliotecas.

4.3 Análise estatística

O objetivo geral da maioria dos estudos de genômica funcional é identificar genes de interesse comparando a expressão de uma determinada etiqueta entre duas bibliotecas. Para isto, a primeira suposição que devemos aceitar é que cada cópia de mRNA, tem a mesma probabilidade de gerar uma etiqueta, na biblioteca do tecido. A segunda suposição é que as diferenças que são observadas não são produzidas por um erro de amostragem. Em termos estatísticos, as etiquetas de interesse serão as que consigam rejeitar a hipótese de nulidade (H_0). A H_0 estabelece que não existe diferença na abundância das etiquetas nas duas bibliotecas (RUIJTER et al., 2002).

O teste Z outorgou diferentes valores Z que permitiram localizar cada uma das comparações de frequências das etiquetas entre as duas bibliotecas nas regiões de aceitação ou rejeição da H_0 , ou seja, permitia determinar se as diferenças de expressões em ambas as bibliotecas eram estatisticamente significativas. Isto permitiu classificar as etiquetas geradas pela metodologia SAGE, estabelecendo-se como limite de probabilidade o valor de $p \leq 0,05$. Isto permitiu agrupar as diferenças estatísticas em 417 etiquetas nesse limite de probabilidade (Tabela 4), mostrando que a vasta maioria das etiquetas e as funções biológicas que elas representaram se mostravam com expressões semelhantes nas duas amostras de indivíduos susceptíveis e resistentes. Portanto, a resposta diferencial das plantas a este patógeno deve se encontrar entre as funções biológicas que estas 471 etiquetas representam.

Tabela 4 – Expressão diferencial das etiquetas nas duas bibliotecas (incluindo os “singletons”). Os dados foram analisados pelo teste Z ($p \leq 0,05$)

Etiquetas com expressão diferencial ($p \leq 0,05$)	Etiquetas sem expressão diferencial	Etiquetas Totais (das duas bibliotecas)
471	25107	25578

As 471 etiquetas identificadas com uma probabilidade de $p \leq 0,05$ foram divididas em 2 grupos. Um grupo com expressão preferencial na biblioteca susceptível (239 etiquetas diferentes) e outro conjunto com expressão preferencial na biblioteca resistente (232 etiquetas diferentes), como ilustrado na Figura 22.

Figura 22 – Distribuição das etiquetas com expressão diferencial ($p \leq 0,05$) nas duas bibliotecas

As etiquetas que não apresentaram diferenças estatisticamente significativas pela alteração do nível de expressão foram ordenadas segundo o critério de abundância, de acordo com a ordem decrescente de expressão das mesmas. As 200 etiquetas com maior expressão, de acordo com esse critério, foram mantidas, para serem posteriormente avaliadas quanto ao padrão metabólico das plantas resistentes e susceptíveis.

4.4 Análise das etiquetas com o auxílio do banco de dados do Genolyptus

4.4.1 Sequências com expressão diferencial

Com auxílio da ferramenta BLAST (ALTSCHUL et al., 1997) as seqüências das etiquetas foram buscadas nas seqüências dos *contigs* da base de dados do Genolyptus. As análises geraram 119 e 148 seqüências de contigs, a partir das etiquetas com expressão preferencialmente ou única na biblioteca susceptível e resistente, respectivamente. Esses números representam, respectivamente, 50% e 64% das etiquetas da biblioteca susceptível (Figura 23) e resistente (Figura 24) a 5% de probabilidade de erro.

Figura 23 – Número de contigs expressos em porcentagem resgatadas do banco de dados do Genolyptus a partir do total de etiquetas com expressão preferencial ou única na biblioteca susceptível, a 5% de probabilidade

Figura 24 – Número de contigs expressos em porcentagem resgatados do banco de dados do Genolyptus a partir do total de etiquetas com expressão preferencial ou única na biblioteca resistente, a 5% de probabilidade

4.4.2 Sequências sem expressão diferencial

As seqüências das etiquetas que não apresentaram diferenças estatisticamente significativas quanto à expressão diferencial, também foram contrastadas com as seqüências dos contigs do banco de dados do Genolyptus. Por essa análise pode-se resgatar 151 contigs que representam 75% de um total de 200 etiquetas desse grupo de ambas bibliotecas (Figura 25).

Figura 25 – Número de contigs expressos em porcentagem resgatadas do total de 200 etiquetas com alta expressão nas duas bibliotecas, porém sem diferenças estatisticamente significativa

4.4.3 Discussão dos resultados do Genolyptus

Quando analisam-se as proporções etiquetas que não conseguiram resgatar um contig do banco de dados Genolyptus, pode-se observar uma tendência de aumento de etiquetas sem contigs: das 200 etiquetas sem expressão diferencial, 25% não apresentam contigs; das 232 etiquetas preferencialmente expressas na biblioteca resistente, 36% se encontram sem contigs; e das 239 etiquetas preferencialmente expressas na biblioteca susceptível, 50% das etiquetas não resgataram nenhum contig.

Provavelmente a maior porcentagem dessas etiquetas sem contig na biblioteca susceptível, pode ser devido ao mRNA do patógeno, outra parte das etiquetas não representadas no banco de dados e a ultima parte dos genes somente expressos em momentos de estresse severo, ainda desconhecidos. Por outro lado a menor porcentagem no grupo das 200 etiquetas sem diferenças estatisticamente significativas, deve-se ao fato de serem transcritos de elevada expressão, em consequência mais fáceis de serem isolados e com maior probabilidade de serem encontrados na base de dados do Genolyptus

4.5 Anotação com outros bancos de dados

Após a seleção dos melhores contigs, realizou-se uma análise comparativa desses com as bases de dados do National Center for Biotechnology Information (NCBI, 2005), pfam (BATEMAN et al, 2004; BATEMAN et al, 2002; BATEMAN et al., 2000; BATEMAN et al., 1999; SONNHAMMER, et al., 1998), COG - *Clusters of Orthologous Groups of proteins* (TATUSOV; KOONIN e LIPMAN, 1997), SMART - *Simple Modular Architecture Research Tool* (Ferramenta de pesquisa de arquitetura modular simples) (SCHULTZ et al., 1998), KEGG KEGG - *Kyoto Encyclopedia of Genes and Genomes* (Enciclopédia de genes e genomas de

Kioto) (KANEHISA; GOTO, 2000; GIBAS; JAMBECK, 2001) e o CDD - *Conserved Domain Database* (Base de dados de domínios conservados) (MARCHLER-BAUER et al., 2003; MARCHLER-BAUER et al., 2005).

Como todas as bases de dados acima descritas contribuíram com à caracterização das seqüências obtidas do Genolyptus foi possível categorizar todas as etiquetas de acordo com as anotações funcionais da seguinte forma:

- *Categoria 1* – Metabolismo: inclui os genes com funções relacionadas ao metabolismo de aminoácidos, nitrogênio, nucleotídeos, carboidratos, lipídeos, cofatores e metabolismo secundário.
- *Categoria 2* – Energia: engloba os genes com funções relacionadas à fotossíntese, ao transporte de elétrons, associados ao ATP, ao metabolismo da glicose e ao metabolismo C1.
- *Categoria 3* – Processos celulares: agrupa os genes relacionados à regulação intracelular, crescimento e desenvolvimento, estresse, homeostase celular e ritmo circadiano.
- *Categoria 4* – Transporte: inclui os genes relacionados ao transporte de macromoléculas (tais como proteínas, DNA, RNA, carboidratos e lipídeos), água, pequenas moléculas, vesículas, fatores de transporte e quinesinas.
- *Categoria 5* – Componentes estruturais: encontram-se os genes relacionados à membrana plasmática, parede celular (polissacarídeos, proteínas, lignina, etc.), e citoesqueleto (actina, tubulina e fibras).
- *Categoria 6* – Metabolismo macromolecular: inclui os genes relacionados ao metabolismo de DNA, RNA, Proteínas e lipídeos.
- *Categoria 7* – Seqüências expressas: abrange os contigs para os quais não há seqüências homólogas dentro dos bancos de dados pesquisados.
- *Categoria 8* – Sem anotação: agrupa as etiquetas para as quais não há contigs com alguma homologia dentro do banco de dados do Genolyptus.

4.5.1 Categorização funcional das etiquetas

4.5.1.1 Etiquetas preferencialmente expressas na biblioteca susceptível

A Figura 26 mostra a distribuição das 239 etiquetas com expressão preferencial na biblioteca susceptível, a 5% de probabilidade, podendo-se observar que 66% dessas etiquetas (*categorias 7 e 8*) apresentam função biológica desconhecida.

Por outro lado, as *categorias 1, 2 e 4* que agrupam genes relacionados ao metabolismo basal, a produção de energia e ao transporte inter e intracelular, respectivamente, estão representadas por 3%, 2% e 4%, das etiquetas com expressão preferencial na biblioteca susceptível. Já a *categoria 3* dos processos celulares se encontra altamente representada, agrupa 15% das etiquetas preferencialmente expressas na biblioteca dos indivíduos susceptíveis. A *categoria 5* está representada por apenas 1 etiqueta (0,42%) relacionada aos componentes estruturais. Por último, a *categoria 6*, relacionada ao metabolismo macromolecular, se encontra representada por 9% das etiquetas preferencialmente expressas na biblioteca susceptível.

Figura 26 – Distribuição das etiquetas em categorias funcionais (C1-C8) dos contigs para o grupo de etiquetas com expressão preferencial ou única na biblioteca susceptível, a $p \leq 0,05$ pelo teste estatístico Z

4.5.1.2 Genes com expressão preferencialmente expressas na biblioteca resistente

A Figura 27, mostra a distribuição das 232 etiquetas com expressão preferencial na biblioteca resistente. Este conjunto de etiquetas foi anotado e distribuído em 8 categorias (seis funcionais e duas sem função determinada), podendo-se observar que 55% das etiquetas apresentam função biológica desconhecida (*categorias 7 e 8*).

As *categorias 1 e 2*, que compreendem, respectivamente, às etiquetas relacionadas ao metabolismo e energia, representam praticamente o dobro das etiquetas das correspondentes categorias na biblioteca susceptível (5 e 6%, respectivamente). A *categoria 3* que agrupa as etiquetas relacionadas aos processos celulares tem uma representação levemente menor que na biblioteca susceptível, equivalente a 14% das etiquetas. A *categoria 4*, que agrupa os genes relacionados ao transporte, representa 3% das etiquetas, apresentando uma leve diferença com a

categoria 4 da biblioteca susceptível. Por outro lado, uma grande diferença pode ser observada na *categoria 5*, que agrupa genes relacionados aos componentes estruturais. Esta categoria se encontra fortemente representada na biblioteca resistente por 9% das etiquetas, em contraste com a biblioteca susceptível, onde esta categoria compreende apenas 0,42% das etiquetas. Por último, a *categoria 6* que agrupa genes relacionados ao metabolismo macromolecular está representada por 8% das etiquetas, ou seja, encontra-se ligeiramente menos representada nessa biblioteca.

Figura 27 – Distribuição das etiquetas em categorias (C1-C8) funcionais pertencentes ao grupo de etiquetas com expressão preferencial ou única na biblioteca resistente, a $p \leq 0,05$ pelo teste estatístico Z

4.5.1.3 Análise das 200 etiquetas com alta expressão, porém sem diferenças estatisticamente significativas entre as duas bibliotecas

Como só uma pequena parte do total das etiquetas identificadas pela metodologia SAGE apresenta diferenças de expressão estatisticamente significativas entre as duas bibliotecas, realizou-se também uma divisão por categorias de 200 etiquetas que apresentam alta expressão.

Portanto, a Figura 28 mostra a distribuição das 200 etiquetas sem expressão preferencial, podendo-se observar que 42% dessas etiquetas (*categorias 7 e 8*) apresenta função biológica desconhecida.

Entre as categorias com função conhecida, a *categoria 1* e 2, que agrupam respectivamente genes relacionados ao metabolismo e à produção de energia, estão representadas por 9% e 12% das etiquetas. Já a *categoria 3* relacionada aos processos celulares está representada por 7% das etiquetas. As *categorias 4* e 5, relacionadas ao transporte e aos componentes estruturais, estão respectivamente representadas por 3% e 4% das etiquetas analisadas, apresentando a *categoria 3* um valor semelhante aos observados nas distribuições anteriores. Em contraste a *categoria 6* relacionada ao metabolismo macromolecular está fortemente representada por 23% do total das etiquetas analisadas.

Figura 28 – Distribuição em categorias (C1-C8) funcionais, dos 200 contigs mais frequentes pertencentes ao grupo de etiquetas com expressões não estatisticamente significativas, pelo teste estatístico Z

4.5.1.4 Análise comparativa dos genes diferencialmente expressos nas duas bibliotecas

Quando a distribuição das categorias dos genes preferencialmente expressos nas bibliotecas susceptível e resistente é comparada surgem uma série de diferenças interessantes, que fazem parte importante da resposta diferencial dos indivíduos resistentes. Estas diferenças são encontradas na categoria de energia e componentes estruturais (*categoria 2 e 5*) e em menor grau a *categoria 1* do metabolismo de pequenas moléculas, sendo mais representadas nos indivíduos resistentes.

A categoria do metabolismo de pequenas moléculas (*categoria 1*) é a categoria que agrupa genes relacionados com a produção de energia (*categoria 2*) mostra um panorama onde os indivíduos resistentes, se apresentam metabolicamente mais ativos, produzindo provavelmente um adicional de energia para as funções celulares e respondendo assim de maneira mais eficiente; ou um panorama onde a planta susceptível poderia encontrar-se fotossinteticamente comprometida.

Por outro lado, a categoria dos componentes estruturais apresenta a maior diferença, aproximadamente 20 vezes maior quando comparada com os indivíduos susceptíveis, apoiando a idéia de que os componentes da parede celular, membrana plasmática e citoesqueleto participam ativamente na defesa contra patógenos.

Interessante, a categoria dos processos celulares (*categoria 3*) que agrupa, entre outros, os genes relacionados ao estresse e à homeostase quase não apresenta diferenças entre as bibliotecas. No entanto, esta é a categoria funcional mais representada em ambas as bibliotecas, sendo provavelmente parte importante da resposta diferencial ao patógeno, tanto dos indivíduos susceptíveis e resistentes, porém com genes diferencialmente expressos em cada fenotipo.

Outra observação, é que as diferenças de representatividade nas *categorias 1 e 2*, que agrupam genes relacionados ao metabolismo e à produção de energia, podem ser devidas a uma queda no metabolismo dos indivíduos susceptíveis, causada pela atividade do patógeno. No entanto, quando analisa-se a categorização funcional das 200 etiquetas de maior expressão, vemos que as categorias de metabolismo, metabolismo macromolecular e energia (*categorias 1, 6, e 2*) se encontram bem representadas tanto nos indivíduos resistentes como susceptíveis indicando que tanto susceptíveis como resistentes, encontram-se metabolicamente ativos.

4.6 Descrição dos genes em cada categoria

4.6.1 Etiquetas com expressão preferencial na biblioteca susceptível

Na Tabela 5, pode-se encontrar a relação das etiquetas com expressão preferencial na biblioteca susceptível. Dentro da *categoria 6*, cinco etiquetas (de 1-5) se encontram representadas por contigs relacionados à degradação de proteínas e onze etiquetas se encontram representadas por contigs envolvidos em funções ribossômicas (de 6-16). A atividade de degradação de proteínas, em conjunto totaliza 148 e 96 etiquetas para a biblioteca susceptível e resistente, respectivamente. A atividade ribossomal, em conjunto, totaliza 140 e 49 etiquetas nas bibliotecas susceptível e resistente, respectivamente. Ainda relacionada à esta categoria, duas etiquetas se encontram representadas por contigs relacionados funcionalmente à modificações pós-traducionais (de 17-18), duas etiquetas apresentam contigs associados ao *splicing* (de 19-20) e uma etiqueta associada provavelmente ao metabolismo de DNA (etiqueta 21).

No relacionado aos componentes estruturais (*categoria 5*) só foi contabilizada uma etiqueta que tem alta similaridade com uma molécula associada à actina (etiqueta 22), cuja expressão foi 12 e 5, na biblioteca resistente e susceptível, respectivamente.

Na *categoria 4* relacionada ao transporte, foram contabilizadas 8 etiquetas (de 23-30). Entre elas pode-se encontrar uma etiqueta com alta semelhança à fatores de transporte (etiqueta 23), outra relacionada ao transporte vesicular (etiqueta 24), cujas expressões respectivamente são 29 e 15 cópias e 11 e 4 cópias, nas bibliotecas susceptível e resistente, respectivamente. Também foram contabilizadas três etiquetas relacionadas ao transporte de pequenas moléculas (de 25-27) e três relacionadas ao transporte de água (de 28-30), cujas expressões relativas são 21 e 4 cópias e 35 e 11 cópias nas bibliotecas susceptível e resistente. Ainda nesta categoria foi detectada uma etiqueta relacionada ao transporte de proteínas (etiqueta 31) representada exclusivamente nesta biblioteca com 14 cópias.

Na *categoria 3* relacionada aos processos celulares (de 32-66) foram contabilizadas duas etiquetas com funções relacionadas à homeotase redox (de 32-33), cuja expressão totaliza respectivamente 17 e 3 cópias, na biblioteca susceptível e resistente. Dentro desta categoria, no relacionado ao estresse (de 34-54) foram contabilizadas sete etiquetas relacionadas ao estresse abiótico (de 34-40), doze ao biótico (de 41-52) e duas ao estresse geral (53-54). Das etiquetas relacionadas ao estresse abiótico foram relacionadas quatro etiquetas relacionadas à produção de glutathionas (etiquetas 35, 37 e 39-40) e uma metalotionina (etiqueta 38) cujas expressões

totalizam respectivamente 253 e 66 cópias; e 348 e 209 cópias, na biblioteca susceptível e resistente. E duas etiquetas relacionadas ao amadurecimento e à desidratação (etiquetas 34 e 36, respectivamente). Na resposta ao estresse biótico foram contabilizadas três etiquetas com semelhança à proteínas de resposta à patogênese (de 45-46 e 49) e três com função de catalase (etiquetas 48 e 50-51) cujas expressões totalizam respectivamente 22 e 4 cópias e 46 e 23 cópias, nas bibliotecas susceptível e resistente. Também foram encontradas duas taumatinas (etiquetas 43 e 52) exclusivamente expressas nesta biblioteca, totalizando 10 cópias. Foi detectada também uma glucanase (etiqueta 41) e uma defensina (etiqueta 42) cujas expressões são respectivamente 40 e 3 cópias e 50 e 12 cópias, na biblioteca susceptível e resistente, respectivamente. Ainda foram detectados um inibidor de proteases (etiqueta 44) e uma quitinase (etiqueta 47) de baixa expressão. Relacionadas ao estresse geral foram detectadas uma chaperona e uma *USP-universal stress protein* (etiqueta 53 e 54, respectivamente). Ainda relacionadas aos processos celulares e expressão gênica foram encontradas cinco etiquetas relacionadas ao controle transcricional (de 62-66) e uma fosfatase (etiqueta 58) de baixa expressão, entre outras.

Na categoria de genes relacionados à produção de energia (*categoria 2*) foram contabilizadas 5 etiquetas (de 67-71), entre elas três associadas à fotossíntese (68-70) e uma associada à cadeia menor da rubisco (etiqueta 71).

Na *categoria 1*, foram contabilizadas foram 8 etiquetas (de 72-79).

Desse grupo de etiquetas com expressão diferencial, as mais abundantes foram a cadeia menor da ribulose bifosfato carboxilase (etiqueta 71) com 321 cópias, o que é bastante plausível, uma vez que esta é uma das proteínas de maior expressão, nos tecidos foliares e uma metalotionina (etiqueta 38) cuja expressão é de 348 cópias, essa provavelmente tem uma importante função relacionada à resposta ao estresse de metais pesados. A expressão desta e a sua função putativa concorda com os dados de Matsumura; Nirasawa e Terauchi (1999) referente a análise da expressão gênica de *seedlings* de arroz.

Tabela 5 - Etiquetas com expressão preferencial na biblioteca susceptível a 5% de probabilidade

							(continua)
N	Etiqueta	S	R	contig (Genolyptus)	NCBI e Outras bases		Categorias
1	TGGTGAAGGG	12	2	Contig806gr-pu	gi 78707589 gb ABB46564.1 pfam00450, Peptidase_S10, Serine carboxypeptidase	Serine carboxypeptidase	6.3.4 - PROTEIN DEGRADATION
2	GAACGCTATC	117	92	Contig4188g1-xy	gi 2677828 gb AAB97142.1 Peptidase_C1A, Peptidase C1A subfamily	cysteine protease cd02248,	6.3.4 - PROTEIN DEGRADATION

N	Etiqueta	S	R	contig (Genolyptus)	NCBI e Outras bases	Categorias
3	TGAGAGTCCA	10	2	Contig4284gl-xy	gi 15235544 ref NP_194633.1 RPT2A; ATPase gnl CDD 28893 cd00009, AAA, AAA-superfamily of ATPases associated with a wide variety of cellular activities, including membrane fusion, proteolysis, and DNA replication	6.3.4 - PROTEIN DEGRADATION
4	ATCATCTGCT	5	0	Contig1785sp-fx	gi 6066285 gb AAF03236.1 ubiquitin carrier protein 4 cd00195.3 - Ubiquitin-conjugating enzyme E2, catalytic (UBCc) domain.	6.3.4 - PROTEIN DEGRADATION
5	GAAGAACTGA	4	0	Contig923ur-xy	gi 71040675 gb AAZ20286.1 ubiquitin-conjugating enzyme 1 gnl CDD 29157 cd00195, UBCC, Ubiquitin-conjugating enzyme E2, catalytic (UBCc) domain.	6.3.4 - PROTEIN DEGRADATION
6	ATGAGCAGAA	12	2	Contig3012sp-fx	gi 40287526 gb AAR83877.1 60S ribosomal protein L19 cd01417, Ribosomal_L19e_E, Ribosomal protein L19e	6.3.3 - RIBOSOME
7	AAACAGTTAT	23	11	Contig2186gl-xy	gi 7440236 pir T04082 probable ribosomal protein S8 pfam01201, Ribosomal protein S8e	6.3.3 - RIBOSOME
8	CGGTCTCTGT	21	11	Contig521pe-xy	gi 24473796 gb AAL91663.1 60s acidic ribosomal protein pfam00428, Ribosomal 60s	6.3.3 - RIBOSOME
9	AGAATTTTGT	7	1	Contig913ur-xy	gi 53792736 dbj BAD53772.1 putative 60S acidic ribosomal 110 4e-23	6.3.3 - RIBOSOME
10	CTAGGAGGGA	7	1	Contig651ur-xy	gi 40287526 gb AAR83877.1 60S ribosomal protein L19 pfam01280, Ribosomal protein L19e.	6.3.3 - RIBOSOME
11	ATTGAGCATA	12	5	Contig1266gr-se	gi 15220003 ref NP_178103.1 RPS17; structural constituent of ribosome pfam00366, Ribosomal_S17	6.3.3 - RIBOSOME
12	ATCAAAGAGA	4	0	Contig827ur-xy	gnl CDD 25719 pfam01248, pfam01248.11- Ribosomal protein L7Ae/L30e/S12e/Gadd45 family	6.3.3 - RIBOSOME
13	GAGATTATTC	4	0	Contig2158gr-ts	gi 24817254 emb CAD56219.1 ribosomal protein S3a [Cicer arietinum] 441 3e-122 gnl CDD 1566 pfam01015, Ribosomal_S3Ae, Ribosomal S3Ae family	6.3.3 - RIBOSOME
14	TATTGAATTA	4	0	Contig3610sp-fx	COG1471.1 - Ribosomal protein S4E	6.3.3 - RIBOSOME
15	ATATTCCATC	36	16	Contig3652sp-fx	gi 2500380 sp Q96499 RL44_GOSHI 60S ribosomal protein L44 (Gossypium hirsutum) 224 7e-58 pfam00935, Ribosomal protein L44	6.3.3 - RIBOSOME
16	TTATTTTGAA	10	2	Contig330pe-xy	gi 44662866 gb AAS47512.1 ribosomal protein L37 [Glycine max] 183 3e-45 pfam01907, Ribosomal_L37e, Ribosomal protein L37e	6.3.3 - RIBOSOME
17	GTGATGGTAT	39	22	Contig525ur-xy	gi 50660327 gb AAT80888.1 chloroplast chaperonin 21 [Vitis... 409 e-113 pfam00166, Chaperonin 10 Kd subunit	6.3.2 - POST- TRANSLATIONAL MODIFICATION
18	GTAAGAGAAT	4	0	Contig3284sp-fx	gi 79315364 ref NP_001030874.1 protein binding / ubiquitin-protein ligase/ zinc ion binding [Arabidopsis thaliana] 103 3e-21 gnl CDD 14624 COG5540, COG5540, RING-finger-containing ubiquitin ligase [Posttranslational modification, protein turnover, chaperones].	6.3.2 - POST- TRANSLATIONAL MODIFICATION
19	ATCAGTGTAA	35	18	Contig2166gr-se	gi 19754 emb CAA43427.1 29kD 29kD A ribonucleoprotein [Nicotiana glauca] 259 1e-67 pfam00076, RNA recognition motif	6.2.1 - SPLICING
20	TTATGGTTTC	14	4	Contig1771ur-xy	gi 34851124 gb AAL13082.1 putative glycine-rich RNA- binding protein [Prunus avium] 158 2e-37 pfam00076, RNA recognition motif	6.2.1 - SPLICING
21	TAGTTGTTTG	38	17	Contig3692sp-fx	gi 63020536 gb AAAY26151.1 high mobility group protein [Cucumis sativus] 140 4e-32 cd01390.1 HMGB- UBF_HMG-box HMGB-UBF_HMG-box, class II and III members of the HMG-box superfamily of DNA-binding proteins	6.1 - DNA METABOLISM
22	TAATTTGAAA	12	5	Contig1190ur-xy	gi 6979167 gb AAF34341.1 latex profilin Hev b 8 [Hevea brasiliensis] 215 1e-54 cd00148, PROF, Profilin binds actin monomers	5.3.1 - ACTIN
23	TAGAGATGGA	29	15	Contig1071gr-pu	gi 21555203 gb AAM63803.1 nuclear transport factor 2, putative [Arabidopsis thaliana]. 95.1 2e-18 gnl CDD 29630 cd00780, NTF2, Nuclear transport factor 2 (NTF2)	4.5 - TRANSPORT FACTORS
24	AATAAGTAT	11	4	Contig1207ur-xy	gi 34909214 ref NP_915954.1 putative ADP-ribosylation factor [Oryza sativa (japonica cultivar-group)] 372 1e-101 gnl CDD 16586 pfam00025, Arf, ADP- ribosylation factor family	4.4 - VESICULAR TRANSPORT
25	AATAAGCAAG	24	11	Contig1700pe-xy	gb AAF69010.1 H+-pyrophosphatase [Vitis vinifera] 227 9e-59 pfam03030, Inorganic H+ pyrophosphatase. Ion transport	4.3 - SMALL MOLECULE TRANSPORT

N	Etiqueta	S	R	contig (Genolyptus)	NCBI e Outras bases	Categorias
26	TTGTCTTGTG	6	0	Contig1628gr-ts	gi 62912516 gb AA21813.1 tyrosine aminotransferase [Glycine max]. 319 1e-85 COG0436, COG0436, Aspartate/tyrosine/aromatic aminotransferase [Amino acid transport and metabolism].	4.3 - SMALL MOLECULE TRANSPORT
27	AATATATGAA	5	0	Contig477gr-ts	gi 24817260 emb CAD56222.1 aquaporin-like water channel protein [Cicer arietinum] 195 9e-49 cd00333, MIP, Major intrinsic protein (MIP) superfamily	4.3 - SMALL MOLECULE TRANSPORT
28	GATTCAATGA	4	0	Contig189sp-fx	gi 28395418 gb AA039007.1 plasma intrinsic protein 2,1 [Juglans regia] 259 4e-68 cd00333, MIP, Major intrinsic protein (MIP) superfamily	4.2 - WATER TRANSPORT
29	GTGGTGGGGG	11	3	Contig856gr-pu	gi 3158476 gb AAC17529.1 aquaporin 2 [Samanea saman] 506 5e-142 cd00333, MIP, Major intrinsic protein (MIP) superfamily.	4.2 - WATER TRANSPORT
30	ACCTGTGGAA	6	1	Contig2209gr-se	gi 38198152 emb CAE53882.1 aquaporin [Ricinus communis] 528 2e-148	4.2 - WATER TRANSPORT
31	TGAATGATGA	14	0	Contig722sp-fx	gi 13752560 gb AAK38726.1 importin alpha 1 [Capsicum annuum] 417 2e-115 gnl CDD 14193 COG5064, SRP1, Karyopherin (importin) alpha [Intracellular trafficking and secretion].	4.1.1 - PROTEIN
32	GATTATGAAT	7	1	Contig1348gr-ts	gi 53988164 gb AAV28174.1 aldo/keto reductase [Fragaria x ananassa]. 244 3e-63 pfam00248.11 - Aldo/keto reductase family	3.4.1 - REDOX
33	TGACATTCAA	10	2	Contig1399gr-ts	gi 62526573 gb AAX84672.1 aldo/keto reductase AKR [Manihot esculenta]. 477 4e-133 COG0667, Tas, Predicted oxidoreductases (related to aryl-alcohol dehydrogenases) [Energy production and conversion]	3.4.1 - HOMEOSTASIS REDOX
34	GCAAGAAGCA	5	0	Contig1604ur-xy	gi 16588758 gb AAL26889.1 abscisic stress ripening-like protein 48.9 8e-05	3.3.2 - ABIOTIC STRESS
35	TCTGATATTT	6	1	Contig776gr-pu	gi 21593459 gb AAM65426.1 lactoylglutathione lyase-like prot... 495 1e-138	3.3.2 - ABIOTIC STRESS
36	TATTTTCCTT	6	1	Contig1704sp-fx	gi 8489786 gb AAF75749.1 dehydration-induced protein ERD15 [Lycopersicon esculentum] 96.3 8e-19	3.3.2 - ABIOTIC STRESS
37	ATGAGCAACA	68	60	Contig896gr-pu	gi 2920666 gb AAC18566.1 2,4-D inducible glutathione S-transferase ... 358 2e-97	3.3.2 - ABIOTIC STRESS
38	TAATACCATC	348	209	Contig2102gr-se	gi 1566700 emb CAA69624.1 metallothionein-like protein [Cari... 90.5 2e-17	3.3.2 - ABIOTIC STRESS
39	TTGGGGCTTA	28	4	Contig3740sp-fx	gnl CDD 24651 pfam02798, GST_N, Glutathione S-transferase, N-terminal domain.	3.3.2 - ABIOTIC STRESS
40	AATTCTGGTT	9	2	Contig1477gr-ts	gi 2920666 gb AAC18566.1 2,4-D inducible glutathione - transferase [Glycine max] 311 2e-83 pfam00043.11 - Glutathione S-transferase, C-terminal domain	3.3.2 - ABIOTIC STRESS
41	CAGAATCAGG	40	3	Contig2952gr-se	gi 16903144 gb AAL30426.1 beta-1,3-glucanase [Prunus persica] 207 9e-53	3.3.1 - BIOTIC STRESS
42	GTGCTTCTG	50	12	Contig3314gr-se	gi 28624546 gb AAL85480.1 defensin protein 1 [Prunus persica] 75.5 7e-13 pfam00304, Gamma-thionins family	3.3.1 - BIOTIC STRESS
43	AGTGTATTGC	6	0	Contig238gr-pu	gi 54609347 gb AAV34889.1 osmotin-like [Theobroma cacao] 332 7e-90 smart00205, THN, Thaumatin family	3.3.1 - BIOTIC STRESS
44	TCAAATAAAT	5	0	Contig2634gr-ts	gi 1762933 gb AAC49969.1 tumor-related protein [Nicotiana tabacum]. 201 2e-50 cd00178, STI, Soybean trypsin inhibitor (Kunitz) family of protease inhibitors	3.3.1 - BIOTIC STRESS
45	CTGGGCACCA	7	1	Contig2632pe-xy	gi 11182126 emb CAC16166.1 pathogenesis-related protein 10 [Vitis vinifera]. 218 2e-55 pfam00407, Bet_v_I, Pathogenesis-related protein Bet v I family	3.3.1 - BIOTIC STRESS
46	CTGTGCCCCA	4	0	Contig3542gr-ts	gi 3511147 gb AAC33732.1 PR-4 type protein [Vitis vinifera] 199 5e-50 "Barwin family" PF00967 - pathogenesis-related protein 4 from tobacco	3.3.1 - BIOTIC STRESS
47	TGCAGCTAGA	4	0	Contig2337gr-ts (whole contig)	28 0.12 pfam00182, Glyco_hydro_19, Chitinase class I.	3.3.1 - BIOTIC STRESS
48	GGTGTGAAGA	4	0	Contig2883sp-fx	gi 3982596 gb AAC83463.1 cationic peroxidase 2 [Glycine max] 558 2e-157 cd00693, secretory_peroxidase, Secretory peroxidases	3.3.1 - BIOTIC STRESS

N	Etiqueta	S	R	contig (Genolyptus)	NCBI e Outras bases	Categorias
49	CTAAATAAA	11	3	Contig2375gr-ts	gi 44409451 gb AAS47035.1 major cherry allergen Pru av 1.0201[Prunus avium] 177 4e-43 pfam00407.11 - Pathogenesis-related protein Bet v I family	3.3.1 - BIOTIC STRESS
50	CGTCTTGCCT	5	0	Contig3754gr-ts	gi 77551234 gb ABA94031.1 Rubredoxin, putative [Oryza sativa (j 149 8e-35 PF00301 - cd00730.2 - Rubredoxin; nonheme iron binding domains containing a [Fe(SCys)4] center	3.3.1 - BIOTIC STRESS
51	TGAATTGCTT	37	23	Contig233ur-xy	gi 6006609 emb CAB56850.1 catalase [Prunus persica] 189 8e-47 pfam00199, Catalase	3.3.1 - BIOTIC STRESS
52	TAATCTGCCG	4	0	Contig2505gr-ts	gi 25091401 sp O80327 TLP1_PYRPY Thaumatin-like protein 1 precursor 306 8e-82 gnl CDD 14921 smart00205, THN, Thaumatin family; The thaumatin family gathers proteins related to plant pathogenesis	3.3.1 - BIOTIC STRESS
53	ACATTGTGTA	10	0	Contig570gr-ts	gi 11762216 gb AAG40386.1 AT5g02160 [Arabidopsis thaliana].. 108 9e-23 gnl CDD 10357 COG0484, DnaJ, DnaJ-class molecular chaperone with C-terminal Zn finger domain [Posttranslational modification, protein turnover, chaperones].	3.3 - STRESS
54	TCATTCTTCA	14	6	Contig1508pe-xy	gi 60100214 gb AA13276.1 USP family protein [Triticum aestivum]. 202 1e-50 gnl CDD 30165 cd00293, USP_Like, Usp: Universal stress protein family	3.3 - STRESS
55	AGAGGTTTGA	6	0	Contig1621gr-ts	gi 23466389 tpg DAA00285.1 TPA: putative phytosulfokine peptide precursor [Lycopersicon esculentum]. 49.7 3e-05 pfam06404, PSK, Phytosulfokine precursor protein (PSK).	3.2.5 - CELL DIVISION
56	AGGCTCTTTA	16	7	Contig2565gl-xy	gi 32470635 gb AAP45162.1 putative Mob1/phocein family protein 428 1e-118 pfam03637, Mob1_phocein, Mob1/phocein family.	3.2.5 - CELL DIVISION
57	AGTGICTTAT	4	0	Contig1956gr-se	gi 50943465 ref XP_481260.1 senescence-associated protein-like [Oryza sativa (japonica cultivar-group)] 204 8e-51 pfam00335.11 - Tetraspanin family	3.2.4 - INTERCELLULAR SIGNALLING
58	GGGGACCACT	4	0	Contig927gr-pu	gi 3341443 emb CAA11075.1 acid phosphatase [Glycine max] 186 1e-45 pfam03767, Acid_phosphat_B, HAD superfamily	3.2.4 - INTERCELLULAR SIGNALLING
59	AGAGGCTGGT	48	16	Contig3413gr-ts	gi 62526583 gb AAX84677.1 auxin-repressed protein-like protein 147 2e-34 pfam05564, Auxin_repressed, Dormancy/auxin associated protein	3.2.1 - HORMONE RESPONSIVE
60	GTAACCTTGG	19	6	Contig1597gl-xy	gi 4193388 gb AAD10032.1 translationally controlled tumor pr... 179 1e-43 CDD 4349 pfam00838, TCTP, Translationally controlled tumour protein	3.2 - GROWTH AND DEVELOPMENT
61	GTCTGAAGGA	14	3	Contig2241pe-xy	gi 9979196 sp Q9ZSW9 TCTP_HEVBR Translationally controlled tumor protein homolog (TCTP)...256 6e-67 pfam00838, Translationally controlled tumour protein	3.1.2 - TRANSLATIONAL CONTROL
62	ATGAAGTTGA	52	18	Contig564pe-xy	gi 30698145 ref NP_569017.2 IAA9; transcription factor [Arabido 172 1e-41 pfam02309.11 AUX/IAA family	3.1.1 - TRANSCRIPTIONAL CONTROL
63	TTGGGTTTGG	20	9	Contig2688pe-xy	gi 15242096 ref NP_197590.1 unknown protein [Arabidopsis thaliana] 123 2e-26 gnl CDD 25073 pfam06200, Zim, ZIM motif	3.1.1 - TRANSCRIPTIONAL CONTROL
64	AGTGATAAAT	11	3	Contig2902gl-xy	gi 40806814 gb AAR92206.1 MADS box transcription factor [Populus tomentosa]	3.1.1 - TRANSCRIPTIONAL CONTROL
65	TCAACGTAAA	20	10	Contig1424gr-ts	gi 23397145 gb AAN31856.1 unknown protein [Arabidopsis thali... 117 4e-25 pfam00098, zf-CCHC, Zinc knuckle	3.1.1 - TRANSCRIPTIONAL CONTROL
66	ACTTATGATG	10	3	Contig260ur-xy	gi 21592325 gb AAM64276.1 bHLH transcription factor, putative [Arabidopsis thaliana] 112 5e-24	3.1.1 - TRANSCRIPTIONAL CONTROL
67	AGGTTTAGTG	19	10	Contig1641gl-xy	gi 732815 emb CAA52445.1 Mg-dependent ATPase 1 [Lycopersicon esculentum] 635 2e-180 gnl CDD 28893 cd00009, AAA, AAA-superfamily of ATPases associated with a wide variety of cellular activities, including membrane fusion, proteolysis, and DNA replication	2.4 - ATP ASSOCIATED
68	GCCACCGCAC	13	0	Contig1507gr-ts	gb AAN60205.1 photosystem II 10 kDa protein [Xerophyta humilis] 224 6e-58 pfam04725, PsbR, Photosystem II 10 kDa polypeptide PsbR	2.2 - PHOTOSYNTHESIS
69	AGGGAAGCAT	102	96	Contig2475gr-ts	gi 19911 emb CAA39039.1 photosystem II 23kDa polypeptide ... 387 e-106	2.2 - PHOTOSYNTHESIS
70	GTGGTCGTTT	15	8	Contig936gr-ml	gi 25990286 gb AAK49535.2 chloroplast protein 12 [Nicotiana ... 137 2e-31	2.2 - PHOTOSYNTHESIS

N	Etiqueta	S	R	contig (Genolyptus)	NCBI e Outras bases	Categorias
71	GTGGATCGTG	321	274	Contig1075gr-pu	gi 77157637 dbj BAE46384.1 ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit [Panax ginseng]. 308 2e-82 pfam00101, Ribulose bisphosphate carboxylase, small chain	2.1 - METABOLISM C1
72	TGTAACAAAA	5	0	Contig1560pe-xy	gi 1296816 emb CAA64452.1 naringenin-chalcone synthase [Juglans nigra x Juglans regia]. 683 0.0 cd00831, CHS_like, Chalcone and stilbene synthases	1.7 - SECONDARY METABOLISM
73	ACCCATCATC	4	0	Contig785pe-xy (whole contig) 28 0.12	gi 21593078 gb AAM65027.1 unknown [Arabidopsis thaliana] 199 5e-50 pfam00903.11 - Glyoxalase/Bleomycin resistance protein/Dioxygenase superfamily	1.7 - SECONDARY METABOLISM
74	GCCCTGGTGG	91	85	Contig4207gl-xy	gi 1838961 emb CAA71885.1 [Casuarina ... 146 1e-33 acyl carrier protein	1.7 - SECONDARY METABOLISM
75	ATGGCCAAGT	13	4	Contig1236ur-xy	gi 21913285 gb AAM81202.1 calmodulin 1 [Medicago truncatula] 298 2e-79 cd00051, EFh, EF-hand, calcium binding motif	1.6 - COFACTORS AND SMALL MOLECULE METABOLISM
76	AGCCACGTGA	85	78	Contig799gr-se	gi 4588376 gb AAD26116.1 galactinol synthase [Brassica napus] 509 e-143	1.4 - CARBOHYDRATE METABOLISM
77	CAATAAGAAT	10	4	Contig1041gr-ts	gi 3021357 emb CAA06339.1 UDP-galactose 4-epimerase [Cyanops... 362 7e-99	1.3 - NUCLEOTIDE METABOLISM
78	TAGAAACATA	25	16	Contig2958gr-ts	gi 18032028 gb AAL47679.1 aminotransferase 1 [Cucumis melo] 238 7e-62	1.1 - AMINO ACID METABOLISM
79	GGGAATGGTT	108	103	Contig2889gr-se	gi 1773330 gb AAB40396.1 glycolate oxidase [Mesembryanthem... 643 0.0	1.1 - AMINO ACID METABOLISM

Notas: Sinais convencionais utilizados

N número da etiqueta

S expressão na biblioteca susceptível

R expressão na biblioteca resistente

4.6.2 Etiquetas com expressão preferencial na biblioteca resistente

Na Tabela 6, pode-se encontrar a relação das etiquetas com expressão preferencial na biblioteca resistente. Dentro da *categoria 6* (de 1-18), foram contabilizadas 5 etiquetas (de 1-5) com funções relacionadas à degradação de proteínas e duas à biossíntese de proteínas (de 14-15), cujas expressões respectivas são 8 e 47 cópias e 1 e 17 cópias, nas bibliotecas susceptível e resistente. Também foram contabilizadas quatro etiquetas relacionadas à atividade ribossomal (de 6-9) e quatro à modificações pós-traducionais (de 10-13), cujas respectivas expressões totalizam 6 e 42 cópias e 3 e 31 cópias, nas bibliotecas susceptível e resistente. Ainda foram encontradas duas etiquetas relacionadas à transcrição (de 16-17) e uma histona (etiqueta 18).

No relacionado à *categoria 5* dos componentes estruturais (de 19-39) foram contabilizadas quatro etiquetas com funções de tubulina (de 19-22), uma actina (etiqueta 24) e uma com função associada à actina (etiqueta 23), cujas expressões totalizam 8 e 79 cópias, 1 e 14 cópias e 1 e 13 cópias, nas bibliotecas susceptível e resistente. Dentro desta categoria, relacionas a parede celular foram contabilizadas duas etiquetas relacionadas ao metabolismo de glicoproteínas (de 25-26) e sete etiquetas relacionadas à síntese de lignina (de 27-33), cujas expressões respectivas são 3 e 36 cópias e 3 e 65 cópias, nas bibliotecas susceptível e resistente. Também foram contabilizadas quatro etiquetas relacionadas ao metabolismo de arabinogalactanos (de 34-37), uma ao

metabolismo de pectinas (etiqueta 38) e uma ao metabolismo de celulose (etiqueta 39), expressas exclusivamente nesta biblioteca e cujas expressões totalizam 44 cópias, 5 cópias e 8 cópias respectivamente.

A *categoria 4* relacionada ao transporte (de 40-46), foram encontradas duas etiquetas relacionadas ao transporte vesicular (de 40-41) e três ao transporte de água (de 44-46) cujas respectivas expressões foram 7 e 28 cópias e 11 e 64 cópias, nas bibliotecas susceptível e resistente. Também foram contabilizadas duas etiquetas relacionadas ao transporte de pequenas moléculas (de 42-43) exclusivamente expressas nesta biblioteca com 11 cópias.

Na *categoria 3* (de 47-79), foram contabilizadas 3 etiquetas relacionadas ao estresse abiótico (de 47-49). Entre elas se contabilizou uma glutathione (etiqueta 47) e duas isoformas de metalotionina (de 48-49). No relacionado ao estresse biótico (de 41-52) foram contabilizadas duas quitinases (etiquetas 51 e 55) exclusivamente expressas nesta biblioteca totalizando 57 cópias. Também foram contabilizadas duas isoflavonas redutases (etiquetas 57 e 61) (precursor de medicarpin, uma fitoalexina) e duas isoformas de proteínas LTP (de 52-53) cujas expressões totalizam 3 e 99 cópias e 140 e 259 cópias, respectivamente. Foram também caracterizadas uma superóxido dismutase (etiqueta 58) e três catalases (etiquetas 54 e 59-60) cujas respectivas expressões foram 0 e 9 cópias e 11 e 50 cópias, nas bibliotecas susceptível e resistente. Ainda relacionadas à este estresse foram identificadas duas etiquetas relacionadas com a resposta a doenças (etiquetas 50 e 56), sendo a última um gene harpin (etiqueta 56) e duas fitoquelatinas (etiquetas 62-63). No relacionado a ativação de cascatas de sinalização foram identificadas seis quinases (de 71-76) e duas fosfatases (de 77-78) cujas expressões totalizam 14 e 74 cópias e 3 e 22 cópias, nas bibliotecas susceptível e resistente, respectivamente.

Na *categoria 2*, foram anotadas três etiquetas com funções no metabolismo de glucose (de 80-82), seis no transporte de elétrons (de 83-88), três na fotossíntese (de 89-91).

Na *categoria 1*, onze etiquetas (de 93-103) foram contabilizadas com funções relacionadas ao metabolismo basal.

Tabela 6 - Etiquetas com expressão preferencial na biblioteca resistente a 5% de probabilidade

(continua)

	Etiqueta	S	R	contig (Genolyptus)	NCBI e Outras bases	Categorias		
1	CAGATCTTCG	2	10	Contig1244ur-xy	gi 288114 emb CAA45622.1 [Petroselinum crispum]. 861	polyubiquitin 0.0	6.3.4 - DEGRADATION	PROTEIN
2	TCTGTGTTGT	2	10	Contig3763gr-ts	ubiquitinating enzyme		6.3.4 - DEGRADATION	PROTEIN
3	GCCAACCACT	0	5	Contig3020gr-ts	ubiquitin extension protein		6.3.4 - DEGRADATION	PROTEIN
4	GTCTTCGTGA	3	13	Contig1306gr-ts	gi 22795037 gb AAN05420.1 protein [Populus x cane 296	putative RING 6e-79	6.3.4 - DEGRADATION	PROTEIN

(continuação)

	Etiqueta	S	R	contig (Genolytus)	NCBI e Outras bases	Categorias
5	GGATGGAGGT	1	9	Contig1946sp-fx	ubiquitin carrier protein [Populus tomentosa].	6.3.4 - PROTEIN DEGRADATION
6	GTGTCTCCAG	1	15	Contig3829gl-xy	18S ribosomal RNA gene	6.3.3 - RIBOSOME
7	CGATAATTAA	0	5	Contig3460gr-se	chloroplast 23S ribosomal RNA	6.3.3 - RIBOSOME
8	CATCGTTTGG	5	17	Contig1381gl-xy	gi 4567232 gb AAD23647.1 40S ribosomal protein S25 [Arabidop... 132 5e-30	6.3.3 - RIBOSOME
9	TTTCTCAGCA	0	5	Contig3200gr-ts	ribosomal protein L44 isoform b	6.3.3 - RIBOSOME
10	AAGCTTCTTC	0	5	Contig1316pe-xy	ARG1-like protein 1	6.3.2 - POST-TRANSLATIONAL MODIFICATION
11	AGGAGGTTCA	0	5	Contig1780ur-xy	pfam00011, Hsp20/alpha crystallin family	6.3.2 - POST-TRANSLATIONAL MODIFICATION
12	CTTCCTAGTT	0	5	Contig2533gl-xy	gi 2104959 gb AAB57848.1 immunophilin [Vicia faba] >gi 74373... 212 1e-53	6.3.2 - POST-TRANSLATIONAL MODIFICATION
13	GGAGGTATGA	3	16	Contig1935gl-xy	putative O-linked GlcNAc transferase	6.3.2 - POST-TRANSLATIONAL MODIFICATION
14	GATGATTGCT	1	9	Contig708gr-se	ubiquinol--cytochrome c reductase	6.3.1 - PROTEIN BIOSYNTHESIS
15	GGCGAAATAT	0	8	Contig105sp-fx	ubiquinol--cytochrome c reductase	6.3.1 - PROTEIN BIOSYNTHESIS
16	CTTCAAAAAA	1	9	Contig1921gr-ts	putative TIP120 protein	6.2.1 - TRANSCRIPTION
17	TGGAAGAACC	0	6	Contig791pe-xy	splicing factor	6.2.1 - TRANSCRIPTION
18	TGACAAATCG	3	14	Contig2913sp-fx	histone H2A.F/Z	6.1.2 - HISTONES
19	ATCTCTATTG	0	13	Contig1342ur-xy	gi 56481497 gb AAV92379.1 alpha tubulin 1 [Pseudotsuga menziesii var. menziesii]. 259 4e-68 gnl CDD 30217 cd02186, alpha_tubulin, The tubulin superfamily includes five distinct families, the alpha-, beta-, gamma-, delta-, and epsilon-tubulins and a sixth family (zeta-tubulin) which is present only in kinetoplastid protozoa	5.3.2 - TUBULIN
20	CTTGCTTGCG	3	25	Contig2604gr-ts	gi 8928425 sp Q9ZPN9 TBB2_ELEIN Tubulin beta-2 chain (Beta-2 tubulin) 773 0.0 cd02187, beta_tubulin	5.3.2 - TUBULIN
21	TTGTGGATTG	0	22	Contig2527gl-xy	gi 23452315 gb AAN32988.1 beta-tubulin 1 [Gossypium hirsutum] 712 0.0 pfam03953, Tubulin_C, Tubulin/FtsZ family, C-terminal domain	5.3.2 - TUBULIN
22	AATGTCTATC	5	19	Contig2684gl-xy	alpha tubulin 1	5.3.2 - TUBULIN
23	TTTGGAATTT	1	13	Contig2708pe-xy	gi 38016521 gb AAR07596.1 fiber protein Fb34 [Gossypium barbadense] 270 7e-71	5.3.1 - ACTIN
24	CTGTGCCAGG	1	14	Contig3822sp-fx	gi 23955912 gb AAN40685.1 actin [Stevia rebaudiana] 538 1e-151 gnl CDD 28896 cd00012, ACTIN, Actin	5.3.1 - ACTIN
25	CTCTTCCGGG	2	25	Contig2525gl-xy	gi 886471 emb CAA58474.1 methionine synthase [Catharanthus roseus] 957 0.0 pfam01717, Methionine_synt, Methionine synthase	5.2.5 - GLYCOPROTEIN METABOLISM
26	TTGTTTTTCT	1	11	Contig3850gl-xy	hydroxymethyltransferase	5.2.5 - GLYCOPROTEIN METABOLISM
27	GTTCCCGCT	0	9	Contig3021sp-fx	gi 13540318 gb AAK29410.1 S-adenosyl-L-methionine synthetase [Elaeagnus umbellata] 752 0.0 pfam02772, S-AdoMet_synt_M, S-adenosylmethionine synthetase, central domain	5.2.3 - LIGNIN METABOLISM
28	GTGATGCTGG	0	6	Contig2671gr-ts	S-adenosyl-L-methionine synthetase	5.2.3 - LIGNIN METABOLISM
29	AACCTGAAAG	0	11	Contig1308ur-xy	gi 84514155 gb ABC59086.1 cytochrome P450 monooxygenase CYP98A [Medicago truncatula] 784 0.0 pfam00067, p450, Cytochrome p450.	5.2.3 - LIGNIN METABOLISM
30	AATGCGGTGCG	0	11	Contig2283gl-xy	gi 1621467 gb AAB17194.1 laccase [Liriodendron tulipifera] 611 2e-173 gnl CDD 11840 COG2132, SufI, Putative multicopper oxidases [Secondary metabolites biosynthesis, transport, and catabolism].	5.2.3 - LIGNIN METABOLISM
31	AACCTAGAAA	3	15	Contig2163pe-xy	caffeoyl-CoA O-methyltransferase	5.2.3 - LIGNIN METABOLISM
32	GTTTGCCGCT	0	8	Contig1607pe-xy	S-adenosyl-L-methionine synthetase	5.2.3 - LIGNIN METABOLISM

Etiqueta	S	R	contig (Genolyptus)	NCBI e Outras bases	Categorias
33 GTTCTTGGAT	0	5	Contig3526gl-xy	gi 437777 emb CAA52814.1 Methyltransferase [Eucalyptus g... 0.0	0- 701 5.2.3 - LIGNIN METABOLISM
34 TACTTATTG	0	22	Contig1798ur-xy	gi 30841338 gb AAO92753.1 arabinogalactan protein [Gossypium hirsutum] 211 5e-53 gnl CDD 24301 smart00554, FAS1, Four repeated domains in the Fasciclin I family of proteins	5.2.2.1 - ARABINOALACTAN METABOLISM
35 CTTTGAAAAG	0	6	Contig2366gl-xy	arabinogalactan protein-like	5.2.2.1 - ARABINOALACTAN METABOLISM
36 AAAACTGAAA	0	8	Contig3943gl-xy	gi 30841338 gb AAO92753.1 arabinogalactan protein [Gossypium hi 211 4e-53	5.2.2.1 - ARABINOALACTAN METABOLISM
37 ACCAGGCCG	0	8	Contig1332ur-xy // Contig2700pe-xy	fasciclin-like AGP 13	5.2.2.1 - ARABINOALACTAN METABOLISM
38 TGGCGCGGG	0	5	Contig941pe-xy	beta xylosidase	5.2.1.2 - PECTATE AND PECTIN METABOLISM
39 GAATCAAAAT	0	8	Contig1884pe-xy	cellulose synthase Cesa-1	5.2.1.1 - CELLULOSE METABOLISM
40 CGTAACGCA	0	5	Contig2639gl-xy	gi 47026845 gb AAT08648.1 factor [Hyacinthus o 363 ADP-ribosylation 6e-99	4.4 - VESICULAR TRANSPORT
41 GGGTTAACTG	7	23	>Contig4062gl-xy	gi 21553518 gb AAM62611.1 factor-like prote... 260 ADP-ribosylation 3e-68	4.4 - VESICULAR TRANSPORT
42 CAAAAATTGC	0	6	Contig4154gl-xy	gi 9759492 dbj BAB10497.1 nuclear pore protein-like [Arabi... 130 1e-58	4.3 - SMALL MOLECULE TRANSPORT
43 TGGCGTGT	0	5	Contig1129sp-fx	lamin	4.3 - SMALL MOLECULE TRANSPORT
44 ATTAGTGGTG	11	52	Contig1253ur-xy	gi 60498693 dbj BAD90701.1 plasma membrane intrinsic protein 2;5 [Mimosa pudica]. 532 7e-150 cd00333, MIP, Major intrinsic protein (MIP) superfamily	4.2 - WATER TRANSPORT
45 CCTCCTTAA	0	7	EGEQBK1002A07.g	aquaporin	4.2 - WATER TRANSPORT
46 GGATGATCAG	0	5	Contig1648ur-xy	gi 62945923 gb AAY22204.1 putative aquaporin [Phaseolus vulgaris] 73.2 3e-12	4.2 - WATER TRANSPORT
47 ACAGTGTCTG	0	6	Contig2557sp-fx	glutathione S-transferase GST 22	3.3.2 - ABIOTIC STRESS
48 CACATCCCTT	113	273	Contig2769gr-ts	pfam01439, Metallothio_2, Metallothionein	3.3.2 - ABIOTIC STRESS
49 GTTATGAAGT	15	84	Contig2372sp-fx	pfam01439, Metallothio_2, Metallothionein	3.3.2 - ABIOTIC STRESS
50 CGGATGACCC	5	18	Contig4175gl-xy	gi 15218836 ref NP_176762.1 disease resistance-responsive family protein [Arabidopsis thaliana]. ... 150 5e-35	3.3.1 - BIOTIC
51 CAGTTAAAAG	0	50	Contig3759gl-xy	gi 34016875 gb AAQ56598.1 chitinase-like protein [Gossypium hirsutum] 510 5e-143 gnl CDD 29557 cd00325, chitinase_glyco_hydro_19, Glycoside hydrolase family 19 chitinase domain	3.3.1 - BIOTIC
52 TCGCGTACT	139	238	Contig3342gr-ts	gi 11095210 gb AAG29777.1 lipid transfer protein 3 precursor [Gossypium hirsutum] 149 9e-35 gnl CDD 29335 cd00262, AAI_plantLTP, Plant lipid transfer protein (LTP) domain, Alpha Amylase Inhibitor (AAI) subgroup	3.3.1 - BIOTIC
53 CTTTGTGT	1	21	Contig889gr-pu	gi 28194084 gb AAO33393.1 lipid transfer protein isoform 1 [Vitis vinifera]. 119 5e-26 gnl CDD 29335 cd00262, AAI_plantLTP, Plant lipid transfer protein (LTP) domain, Alpha Amylase Inhibitor (AAI) subgroup	3.3.1 - BIOTIC
54 CAATCCAGAC	3	13	Contig834gr-ML	peroxiredoxin Q	3.3.1 - BIOTIC
55 GTCTCCCACT	0	7	Contig131gr-xy	chitinase-like protein	3.3.1 - BIOTIC
56 GAAGAATCGA	2	13	Contig1250ur-xy	Harpin-induced protein 1 (Hin1)	3.3.1 - BIOTIC
57 AATTTCACAG	1	88	Contig1719ur-xy	gi 10764491 gb AAG22740.1 allergenic isoflavone reductase-like protein Bet v 6.0102 [Betula pendula]. 105 6e-22	3.3.1 - BIOTIC
58 CTTGAAAAA	0	9	Contig798gr-ml	gi 38228697 emb CAE54085.1 superoxide dismutase [Fagus sylvatica] 274 4e-72 gnl CDD 25385 pfam00080, Sod_Cu, Copper/zinc superoxide dismutase (SODC)	3.3.1 - BIOTIC

	Etiqueta	S	R	contig (Genolyptus)	NCBI e Outras bases	Categorias
59	TGCTGATCTT	6	26	Contig4289gl-xy	gi 32526566 emb CAD42908.1 catalase [Prunus persica] 981 0.0 pfam00199, Catalase	3.3.1 - BIOTIC
60	CAGCAAATTA	2	11	Contig2025sp-fx	peroxidase	3.3.1 - BIOTIC
61	TACAACCCCC	2	11	Contig21gr-yl	isoflavone reductase related protein	3.3.1 - BIOTIC
62	ATTCTTGAC	0	10	Contig377ur-xy	gi 9755608 emb CAC01762.1 putative phytochelatin synthetase [Arabidopsis thaliana] 100 4e-20	3.3 - STRESS
63	CAATACTGAG	0	5	Contig2560gl-xy	putative phytochelatin synthetase	3.3 - STRESS
64	TTAATCGTGC	0	5	Contig2109gr-ts	Skp1	3.2.5 - CELL DIVISION
65	ACTAGCCCTG	1	8	Contig2067gr-ts	EIN3-like protein	3.2.1 - HORMONE RESPONSIVE
66	ATCCGATATT	3	19	Contig1762gr-ts	gi 47575681 gb AAT35818.1 abscisic stress ripening protein-like protein [Musa acuminata]. 62.0 2e-08	3.2 - GROWTH AND DEVELOPMENT
67	TAAGTACCGG	0	11	Contig1176ur-xy	gi 7406663 emb CAB85625.1 putative ripening-related protein [Vitis vinifera]. 210 5e-53 pfam04043, Plant invertase/ pectin methylesterase inhibitor	3.2 - GROWTH AND DEVELOPMENT
68	AAGAAGCAGT	0	5	Contig2306pe-xy	germin-like protein (GLP10)	3.2 - GROWTH AND DEVELOPMENT
69	TAAAGACGCT	0	5	Contig1513ur-xy	gi 37695573 gb AAR00326.1 rapid alkalinization factor 2 [Solanu 120 3e- 26	3.2 - GROWTH AND DEVELOPMENT
70	TGCTGAAAAA	0	8	Contig4288gl-xy	gi 20465441 gb AAM20180.1 unknown protein [Arabidopsis thaliana] 563 e-159	3.2 - GROWTH AND DEVELOPMENT
71	CTGCGTATAC	1	8	Contig4184gl-xy	gi 21553971 gb AAM63052.1 calcium- dependent protein kinase... 261 1e-68	3.1.4 - KINASES
72	TTTGGCGCCA	1	8	Contig4316gl-xy	gi 6967109 emb CAB72463.1 protein kinase- like protein 176 2e-50	3.1.4 - KINASES
73	AGACGGAATA	10	30	Contig4143gl-xy	gi 9757781 dbj BAB08390.1 adenosine kinase [Arabidopsis th... 568 e-161	3.1.4 - KINASES
74	TGATGGGATA	2	10	Contig2676pe-xy	ref XP_468084.1 putative UMP/CMP kinase a [Oryza sativa (japonica cultivar-group)] 324 7e-88	3.1.4 - KINASES
75	TGTACATTCG	0	5	Contig3243gl-xy	serine/threonine-protein kinase	3.1.4 - KINASES
76	TGTGCTGTAG	0	13	Contig3238gl-xy	gi 3482933 gb AAC33218.1 Similar to cdc2 protein kinases [Arabidopsis thaliana]. 151 8e-36 pfam04669, DUF579, Protein of unknown function (DUF579)	3.1.4 - KINASE
77	TGACGTAGCG	2	13	Contig975gr-ML	phosphatase 2C, putative / PP2C	3.1.3 - INTRACELLULAR SIGNALING
78	TACTGTTGTT	1	9	Contig2232gl-xy	putative protein phosphatase	3.1.3 - INTRACELLULAR SIGNALING
79	AGATAAAAAA	0	5	Contig2889gr-ts	WD-40 repeat family protein	3.1 - INTRACELLULAR REGULATION
80	TGCTCTGCAG	9	24	Contig4287gl-xy	gi 1086252 pir S51376 sucrose cleavage protein - potato... 211 2e-53	2.5 - GLUCOSE METABOLISM
81	ATGTTTGAGT	4	16	Contig1167gr-pu (whole contig) 26 0.49		2.5 - GLUCOSE METABOLISM
82	TGACTAACCT	0	7	Contig1599pe-xy	sucrose synthase	2.5 - GLUCOSE METABOLISM
83	ATCCCCGTAC	1	8	Contig1187gr-ml	Plastocyanin	2.3 - ELECTRON TRANSPORT
84	ATTGCTGTGT	0	11	Contig181gr-xy	gi 21553614 gb AAM62707.1 blue copper protein, putative [Arabid 164 6e-39 pfam02298, Cu_bind_like, Plastocyanin-like domain	2.3 - ELECTRON TRANSPORT
85	TACGTTGTAA	0	5	Contig1285ur-xy	NADH:cytochrome b5 reductase	2.3 - ELECTRON TRANSPORT
86	AGGTTTCTTG	0	7	Contig1428pe-xy	gi 7268000 emb CAB78340.1 putative putative protein [Arabidopsis thaliana]. 373 8e-102	2.3 - ELECTRON TRANSPORT
87	CTACTTTGGG	0	7	Contig1737pe-xy	thioredoxin	2.3 - ELECTRON TRANSPORT
88	GTCATCTCAT	0	8	Contig1775ur-xy	blue copper protein	2.3 - ELECTRON TRANSPORT
89	CAAAAATTAC	194	338	Contig1310gr-ts	gi 3417451 dbj BAA32346.1 light-harvesting chlorophyll a/b-binding protein of photosystem II [Cryptomeria japonica] 299 1e-79 pfam00504, Chloroa_b-bind, Chlorophyll A-B binding protein	2.2- PHOTOSYNTHESIS

(conclusão)

	Etiqueta	S	R	contig (Genolyptus)	NCBI e Outras bases	Categorias
90	TACCTGCACT	18	58	Contig2595gr-ts	gi 3337435 gb AAC95000.1 PsbY precursor; putative photosystem II peptide [Spinacia oleracea] 102 2e-20	2.2- PHOTOSYNTHESIS
91	TACCTGATCT	8	37	Contig1226gr-pu	gi 28629385 gb AAO49652.1 photosystem I-N subunit [Phaseolus vulgaris]. 206 6e-52 pfam05479, PsaN, Photosystem I reaction centre subunit N	2.2 - PHOTOSYNTHESIS
92	GAATATCCTG	3	14	Contig2625gl-xy	light-inducible protein ATLS1	2 - CELLULAR REGULATION
93	CTATTGCGGG	4	20	Contig333gr-PU	gi 14331015 emb CAC41012.1 putative chloroplast terpene synthase [Quercus ilex]. 133 4e-30 cd00684, Terpene_cyclase_plant_C1, Plant Terpene Cyclases	1.7 - SECONDARY METABOLISM
94	GCCATCATAT	0	5	Contig1545ur-xy	hydroxycinnamoyl transferase	1.7 - SECONDARY METABOLISM
95	TGTCCAGCA	2	10	Contig3872sp-fx	acyl-CoA-binding protein	1.5 - LIPID METABOLISM
96	TGTCCGGCA	2	13	Contig1239ur-xy	acyl-CoA-binding protein	1.5 - LIPID METABOLISM
97	GTCTAGTTTT	47	117	Contig947gr-ml	gi 5608497 emb CAB51533.1 galactinol synthase, isoform GolS-1 [Ajuga reptans] 543 9e-153 gnl CDD 25768 pfam01501, Glyco_transf_8, Glycosyl transferase family 8	1.4 - CARBOHYDRATE METABOLISM
98	GTGAAGTCAG	1	8	Contig1327ur-xy	gi 1518540 gb AAB58398.1 UDP-glucose dehydrogenase [Glycine ... 789 0.0	1.4 - CARBOHYDRATE METABOLISM
99	TCGGTCTTAT	4	16	Contig2898sp-fx	UDP-glucose pyrophosphorylase	1.4 - CARBOHYDRATE METABOLISM
100	CCTGACCCCG	13	35	Contig957gr-ml	gi 5608497 emb CAB51533.1 galactinol synthase, isoform GolS-1 [Ajuga reptans]. 541 2e-152	1.4 - CARBOHYDRATE METABOLISM
101	CCATCGGTTG	0	9	Contig2320gr-ts	gi 7268039 emb CAB78378.1 3-hydroxyisobutyryl-coenzyme A hydrolase-like protein [Arabidopsis thaliana]... 67.4 2e-10	1.1 - AMINO ACID METABOLISM
102	AACCATTTTG	1	15	Contig2252gl-xy	gi 3452497 emb CAA76854.1 ketol-acid reductoisomerase [Pisum sativum].... 656 0.0 pfam01450, IlvC, Acetohydroxy acid isomeroreductase, catalytic domain	1.1 - AMINO ACID METABOLISM
103	GTTGCCGCCG	0	8	Contig3139sp-fx	5-methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase	1.1 - AMINO ACID METABOLISM

Notas: Sinais convencionais utilizados

N número da etiqueta

S expressão na biblioteca susceptível

R expressão na biblioteca resistente

4.6.5 Etiquetas sem expressão preferencial

Na Tabela 7, pode-se encontrar a relação das etiquetas sem expressão preferencial. Dentro da *categoria 6* relacionada ao metabolismo macromolecular, sete etiquetas foram anotadas como relacionadas à degradação de proteínas (de 1-7) e vinte relacionadas com a atividade ribossomal (de 8-27), cujas expressões respectivas totalizam 180 e 235 cópias e 365 e 390 cópias nas bibliotecas susceptível e resistente. Também foram contabilizadas sete etiquetas com funções relacionadas a modificações pós-traducionais (de 28-34) e oito relacionadas à biossíntese de proteínas (de 35-42) com expressões de 115 e 150 cópias e 167 e 199 cópias, nas bibliotecas susceptível e resistente, respectivamente. Por último, foram detectadas quatro etiquetas relacionadas ao metabolismo do RNA cuja expressão totaliza 66 e 72 cópias, nas bibliotecas susceptível e resistente, respectivamente.

No referente à *categoria 5*, dos componentes estruturais foram anotadas uma isoforma da tubulina (etiqueta 48) e duas proteínas associadas ao citoesqueleto (de 49-50), cujas expressões respectivas totalizam 18 e 21 cópias e 21 e 27 cópias) nas bibliotecas susceptível e resistente . Ainda dentro deste grupo foram identificadas quatro proteínas relacionadas á parede celular (de 51 à 54) e uma relacionada à membrana plasmática (etiqueta 55).

Na categoria do transporte (*categoria 4*), foram anotadas quatro etiquetas relacionadas ao transporte vesicular (de 56-59) e duas relacionadas ao transporte de pequenas moléculas (de 60-61).

Já na categoria dos processos celulares (de 62-75) foram contabilizadas quatro etiquetas relacionadas ao estresse biótico (de 63-66). Uma delas está anotada como induzida por resposta hipersensível, uma quitinase , uma metalotionina, uma uma superóxido dismutase (etiquetas 63, 64, 65 e 66, respectivamente) cujas respectivas expressões são 18 e 14 cópias, 66 e 64 cópias, 11 e 14 cópias e 17 e 11, nas bibliotecas susceptível e resistente. Também foram detectadas duas relacionadas com a sinalização intracelular (de 71-72), três com o controle transcricional (de 73-75) e uma produtora de etileno (etiqueta 67), entre outras.

No relacionado à *categoria 2* (de 76-98), foram identificadas 3 etiquetas relacionadas com metabolismo de glucose (de 76-78), quatro associadas ao ATP (de 79-82), oito relacionadas ao transporte de elétrons (de 83-90) e oito relacionadas à fotossíntese (de 91-98), apresentando um nível de produção de energia alto nas duas bibliotecas.

Na categoria do metabolismo, *categoria 1* (de 99-115), foram contabilizadas dezessete etiquetas. Duas relacionadas ao metabolismo de pequenas moléculas (de 99-100), duas ao metabolismo lipídico (de 101-102), quatro ao metabolismo de carboidratos (de 103-106), cinco ao metabolismo de nitrogênio (107-111), três ao metabolismo de aminoácidos (de 112-114) e uma relacionada ao metabolismo de glucose (etiqueta 115), também apresentando um nível metabólico forte.

Tabela 7 - Etiquetas sem expressão diferencial

(continua)

	Etiqueta	S	R	contig (Genolyptus)	NCBI e Outras bases	Categorias
1	CGGGCCGATC	44	62	Contig2714gl-xy	gi 13359451 dbj BAB33421.1 putative senescence-associated protein [Pisum sativum]. 310 1e-82	6.3.4 - PROTEIN DEGRADATION
2	TAAGTGGCTT	37	51	Contig1254gl-xy	gi 12231174 dbj BAB20970.1 aspartic proteinase 2 [Nepenthe... 759 0.0	6.3.4 - PROTEIN DEGRADATION
3	TATGCTAACT	32	51	Contig202gr-ts	gi 18086 emb CAA27609.1 pot. cysteine proteinase [Carica papaya]. 128 6e-29	6.3.4 - PROTEIN DEGRADATION
4	TTGGTAGTTT	18	21	Contig213lgr-ts	gi 322750 pir S28420 ubiquitin / ribosomal protein CEP52 -... 258 9e-68	6.3.4 - PROTEIN DEGRADATION

(continuação)

	Etiqueta	S	R	contig (Genolyptus)	NCBI e Outras bases	Categorias
5	TACATAAGGC	18	17	Contig1496ur-xy	gi 7381219 gb AAF61440.1 papain-like cysteine proteinase isoform I... 584 e-165	6.3.4 - PROTEIN DEGRADATION
6	ATAGGATTGT	20	14	Contig1763gr-ts	gi 51534981 dbj BAD38105.1 polyubiquitin 2 [Oryza sativa (japon 262 1e-68	6.3.4 - PROTEIN DEGRADATION
7	GAAAACGCTC	11	19	Contig2366gr-se	gi 12484207 gb AAG54002.1 unknown protein [Arabidopsis thali... 160 2e-38	6.3.4 - PROTEIN DEGRADATION
8	GAGAGAAGGA	33	26	Contig4198gl-xy	gi 7582403 gb AAF64312.1 plastid ribosomal protein L19 precursor ... 173 7e-42	6.3.3 - RIBOSOME
9	GATCCAGTTG	27	23	Contig3394gr-se	gi 6984222 gb AAF34799.1 40S ribosomal protein S16 [Euphorbia e 206 2e-52	6.3.3 - RIBOSOME
10	AGATGTGGTG	18	27	Contig3022sp-fx	gi 7270565 emb CAB81522.1 putative ribosomal protein L8 [A... 507 e-142	6.3.3 - RIBOSOME
11	TTGATTGTTG	21	23	Contig1215pe-xy	gi 15241026 ref NP_198122.1 40S ribosomal protein S21 (RPS21C) 142 4e-33	6.3.3 - RIBOSOME
12	TTGAGGTTTT	16	28	Contig1432gr-se	gi 55773694 dbj BAD72277.1 40S ribosomal protein S30-like [O... 106 3e-22	6.3.3 - RIBOSOME
13	ATGACAGATG	16	27	Contig3166gr-ts	gi 21553368 gb AAM62461.1 putative ribosomal protein [Arab... 44 9e-11	6.3.3 - RIBOSOME
14	TGAAATCAGA	15	17	Contig2630sp-fx	gi 50948455 ref XP_483755.1 putative 60S ribosomal protein L... 353 5e-96	6.3.3 - RIBOSOME
15	AGGTATTTGC	12	19	Contig1162gr-ts	gi 166867 gb AAA32866.1 ribosomal protein S11 (probable start c 250 4e-65	6.3.3 - RIBOSOME
16	AATGGATTGT	15	14	Contig2306gl-xy	gi 166867 gb AAA32866.1 ribosomal protein S11 (probable start c 250 3e-65	6.3.3 - RIBOSOME
17	CGGAACACAA	8	18	Contig4173gl-xy	gi 20465445 gb AAM20182.1 putative chloroplast ribosomal protein... 138 8e-32	6.3.3 - RIBOSOME
18	TAGACTACCT	16	9	Contig972gr-ts	gi 21592421 gb AAM64372.1 ribosomal protein L11, cytosolic... 312 6e-84	6.3.3 - RIBOSOME
19	GTTTATTGTA	15	9	Contig1626pe-xy	gi 31745220 gb AAP68880.1 putative ribosomal protein S29 [... 126 3e-28	6.3.3 - RIBOSOME
20	TCAGCTCTAG	11	13	Contig1456pe-xy	gi 21617950 gb AAM67000.1 putative ribosomal protein S10 [Ar... 180 5e-44	6.3.3 - RIBOSOME
21	GTGAATTCGT	28	28	Contig874gr-pu	gi 166867 gb AAA32866.1 ribosomal protein S11 (probable start c 249 8e-65	6.3.3 - RIBOSOME
22	GGGGGCCTGA	21	30	Contig447gr-pu	gi 248303 gb AAB21989.1 ribosomal protein L12; CL12 [Nicotia... 119 9e-26	6.3.3 - RIBOSOME
23	AGTTTCTCCC	27	21	Contig3644sp-fx	gi 7671467 emb CAB89407.1 40S ribosomal protein S6 [Arabido... 390 e-107	6.3.3 - RIBOSOME
24	AATGAATTGT	21	16	Contig512ur-xy	gi 21902507 gb AAM78552.1 ribosomal protein small subunit ... 99 4e-20	6.3.3 - RIBOSOME
25	ATCCACTTAC	18	14	Contig2514gr-ts	gi 1362041 pir S56673 ribosomal protein S23.e, cytosolic (... 285 7e-76	6.3.3 - RIBOSOME
26	TGACTTCAAG	17	13	Contig3228gr-ts	gi 49524072 emb CAG47084.1 40S ribosomal protein S9 [Catha... 356 5e-97	6.3.3 - RIBOSOME
27	AGAATTGTTG	10	15	Contig2151gr-ts	gi 7440901 pir T01744 ribosomal protein L17, chloroplast - common tobacco 259 5e-68	6.3.3 - RIBOSOME
28	GATGGCACTT	24	36	Contig2389gl-xy	gi 47176684 gb AAT12488.1 copper chaperone [Populus alba x... 132 9e-30	6.3.2 - POST- TRANSLATIONAL MODIFICATION
29	CGGATACCGA	26	41	Contig4174gl-xy	gi 1143427 emb CAA52149.1 heat shock protein 70 [Cucumis s... 1091 0.0	6.3.2 - POST- TRANSLATIONAL MODIFICATION
30	TGGGAAGAGT	14	21	Contig2500gr-ts	gi 33772193 gb AAQ54533.1 putative DnaJ protein [Malus x domest 139 2e-31	6.3.2 - POST- TRANSLATIONAL MODIFICATION
31	TGTTATTGAT	19	14	Contig438ur-xy	gi 17473863 gb AAL38353.1 putative heat-shock protein [Arabidop 210 5e-53	6.3.2 - POST- TRANSLATIONAL MODIFICATION
32	ACCTCACGCC	11	13	Contig1475pe-xy	gi 38325815 gb AAR17080.1 heat shock protein 70-3 [Nicotiana ta 824 0.0	6.3.2 - POST- TRANSLATIONAL MODIFICATION
33	GGTGTTC AAT	11	12	Contig2182gl-xy	gi 64213910 gb AAY41234.1 putative peptidyl- prolyl cis-trans... 240 5e-62	6.3.2 - POST- TRANSLATIONAL MODIFICATION
34	GTGTTTTAGG	10	13	Contig1109ur-xy	gi 21886603 emb CAC84116.1 peptidylprolyl isomerase (cyclophilin). 305 1e-81	6.3.2 - POST- TRANSLATIONAL MODIFICATION
35	ATAAGAGTTT	36	33	Contig2666gl-xy	gi 4585973 gb AAD25609.1 translation initiation factor [Ar... 215 7e-55	6.3.1 - PROTEIN BIOSYNTHESIS
36	TTGAGGGTGT	27	27	Contig802gr-se	gi 6561948 emb CAB62452.1 alpha NAC-like protein [Arabidopsi... 183 6e-45	6.3.1 - PROTEIN BIOSYNTHESIS

	Etiqueta	S	R	contig (Genolyptus)	NCBI e Outras bases	Categorias
37	AAAGCTTGCG	25	27	Contig1744sp-fx	gi 475216 emb CAA55641.1 translation initiation factor (eIF-... 378 1e-103	6.3.1 - PROTEIN BIOSYNTHESIS
38	GCCGGACCTT	18	28	Contig2620pe-xy	gi 6561948 emb CAB62452.1 alpha NAC-like protein [Arabidopsi... 141 3e-32	6.3.1 - PROTEIN BIOSYNTHESIS
39	TGATAGGCGG	21	22	Contig1547pe-xy	gi 6015059 sp O64937 EF1A_ORYSA Elongation factor 1-alpha (... 868 0.0	6.3.1 - PROTEIN BIOSYNTHESIS
40	ACTCGAGGTA	15	25	Contig1152sp-fx	gi 21554974 gb AAM63744.1 40S ribosomal protein S21 homolog ... 82 4e-15	6.3.1 - PROTEIN BIOSYNTHESIS
41	GTGTCGGGTT	14	23	Contig3586sp-fx	gi 19702 emb CAA45104.1 eukaryotic initiation factor 5A (2... 309 9e-83	6.3.1 - PROTEIN BIOSYNTHESIS
42	GAAATGGTTT	11	14	Contig1242pe-xy	gi 33325119 gb AAQ08192.1 eukaryotic translation initiation factor 5A... 307 3e-82	6.3.1 - PROTEIN BIOSYNTHESIS
43	TATTTTGTCA	20	28	Contig333ur-xy	gi 4960152 gb AAD34614.1 DNA-directed RNA polymerase IIb [Ni... 119 4e-26	6.2.1 - TRANSCRIPTION
44	TGGATCAGAG	12	15	Contig4291gl-xy	gi 22331232 ref NP_188747.2 RNA recognition motif (RRM)-co... 269 3e-70	6.2.1 - SPLICING
45	TGTGGTCGTA	12	13	Contig2425gr-ts	gi 50946241 ref XP_482648.1 splicing factor 3B subunit 5-lik... 177 2e-43	6.2.1 - SPLICING
46	CCGTCACAG	22	16	Contig1705gr-ts	gi 12642936 gb AAK00410.1 putative transcriptional coactivator protein ... 221 2e-56	6.2 - RNA METABOLISM
47	TTCTGGTCCA	8	19	Contig1407pe-xy	gi 42572635 ref NP_974413.1 high mobility group protein alpha (HMGalpha)... 118 3e-25	6.1.2 - HISTONES
48	GAAGTCGAAA	18	21	Contig4192gl-xy	gi 11558252 emb CAC17794.1 microtubule-associated protein ... 840 0.0	5.3.2 - TUBULIN
49	GGAACCTCCG	7	19	Contig1633ur-xy	gi 34851178 gb AAP15201.1 profilin-like protein [Ambrosia ... 250 3e-65	5.3.1 - ACTIN
50	TAAGAGTGTT	14	8	Contig1300pe-xy	gi 4731573 gb AAD28506.1 remorin 1 [Lycopersicon esculentum] 124 4e-27	5.3 - CYTOSKELETON
51	AAGGTTTGGT	11	14	Contig1128pe-xy	gi 1431629 emb CAA67728.1 pectinacetylesterase precursor [Vi... 248 2e-64	5.2.1.2 - PECTATE AND PECTIN METABOLISM
52	AGTTTATTT	10	15	Contig4147gl-xy	gi 37781495 gb AAP40636.1 cellulose synthase 6 [Populus tr... 1046 0.0	5.2.1.1 - CELLULOSE METABOLISM
53	GGAGGATACC	15	29	Contig1864gr-ts	gi 421806 pir A46130 fiber protein E6 (clone CKE6-1A) - upla... 55.1 1e-06	5.2 - CELL WALL
54	TTGAAGTTGA	20	16	Contig740gr-pu	gi 30841448 gb AAP34362.1 fiber protein Fb15 [Gossypium barbade 148 9e-35	5.2 - CELL WALL
55	TACTTGCAAG	15	28	Contig1289gr-ts	gi 23197608 gb AAN15331.1 membrane related protein CP5, putative [Arabidopsis thaliana]... 502 e-140	5.1 - MEMBRANE
56	GGATTGCATA	45	74	Contig1767ur-xy	gi 56311394 emb CAI29265.1 ADP-ribosylation factor 1 [Medicago truncatula]... 365 8e-100	4.4 - VESICULAR TRANSPORT
57	TTAGCTGAGA	53	64	Contig870gr-ml	gi 438247 emb CAA81082.1 glycine hydroxymethyltransferase [Solanum tuberosum]... 658 0.0	4.4 - VESICULAR TRANSPORT
58	GGTGTTTGGA	22	27	Contig1207pe-xy	gi 23397041 gb AAN31806.1 putative RAN2 small Ras GTP-binding nuclear protein (Ran-2)... 415 e-114	4.4 - VESICULAR TRANSPORT
59	TTTGAGCTTC	11	12	Contig2540gl-xy	gi 20149262 gb AAM12880.1 GTP-binding protein [Helianthus ... 432 e-120	4.4 - VESICULAR TRANSPORT
60	TGCCCTCCT	18	15	Contig1780sp-fx	gi 732893 emb CAA59694.1 tobacco calretulin [Nicotiana tab... 119 1e-35	4.3 - SMALL MOLECULE TRANSPORT
61	CGGTTGTACT	7	19	Contig723ur-xy	gi 3334138 sp Q39817 CALX_SOYBN Calnexin homolog precursor >gi 6 189 7e-47	4.3 - SMALL MOLECULE TRANSPORT
62	CGCTGTGATT	9	18	Contig3060sp-fx	gb AAP76396.1 glyoxalase I [Zea mays] 473 4e-132	3.3.2 - ABIOTIC
63	TGAATTGCT	18	14	Contig3013gr-ts	gi 21554781 gb AAM63689.1 hypersensitive-induced response protein ... 473 e-132	3.3.1 - STRESS BIOTIC
64	ACAAGGTTTC	66	64	Contig3401gr-ts	gi 3126965 gb AAC16011.1 basic chitinase [Elaeagnus umbellata] 165 9e-40	3.3.1 - BIOTIC STRESS
65	CAACTGCTGA	11	14	Contig3470gr-ts	gi 1223652 emb CAA65338.1 metallothionein [Casuarina glauca]... 51.2 2e-05	3.3.1 - BIOTIC STRESS
66	ACAACAGATA	17	11	Contig1469ur-xy	gi 38228697 emb CAE54085.1 superoxide dismutase [Fagus syl... 271 2e-71	3.3.1 - BIOTIC STRESS

Etiqueta	S	R	contig (Genolyptus)	NCBI e Outras bases	Categorias
67 TATGTGTATT	14	12	Contig2738gr-ts	gi 398994 sp P31237 ACCO_ACTCH 1-aminocyclopropane-1-carboxylate oxidase (ACC oxidase) (Ethylene-forming enzyme) (EFE). 524 3e-147	3.3 - STRESS
68 GTGACTTTGG	10	13	Contig614sp-fx	gi 36963062 gb AAQ87663.1 translationally controlled tumor prot 288 2e-76	3.2.5 - CELL DIVISION
69 TACGGTGCTT	23	31	Contig2315sp-fx	gi 114734 sp P13089 AUX28_SOYBN Auxin-induced protein AUX28 >... 253 7e-66	3.2.1 - HORMONE RESPONSIVE
70 AGAGGCTCCT	10	16	Contig3211gr-ts	gi 13430203 gb AAK25768.1 auxin-repressed protein like-protein... 138 2e-31	3.2.1 - HORMONE RESPONSIVE
71 TAAACCAGAC	10	18	Contig1308gr-ts	gi 50725153 dbj BAD33770.1 unknown protein [Oryza sativa (japon 182 1e-44	3.1.3 - INTRACELLULAR SIGNALING
72 TCACTTGGA	11	12	Contig334gr-ml	gi 21592643 gb AAM64592.1 unknown [Arabidopsis thaliana] >gi... 231 2e-59	3.1.3 - INTRACELLULAR SIGNALING
73 AGATTGATCG	19	16	Contig2114gr-ts	gi 37991674 dbj BAD00048.1 perchloric acid soluble translation inhibitor protein homolog... 247 4e-64	3.1.2 - TRANSLATIONAL CONTROL
74 GGTTAGTCGA	28	27	Contig3486gl-xy	gi 13400109 gb AAK21974.1 rRNA promoter binding protein [Rat... 79.7 9e-14	3.1.1 - TRANSCRIPTIONAL CONTROL
75 TGAAGTTGTT	13	10	Contig2695gr-ts	gi 38503523 gb AAR20445.2 putative leucine zipper protein [Goss 670 0.0	3.1.1 - TRANSCRIPTIONAL CONTROL
76 GTGACTTTTT	45	66	Contig954gr-ml	gi 21311559 gb AAM46780.1 latex plastidic aldolase-like protein [Hevea brasiliensis] 714 0.0	2.5 - GLUCOSE METABOLISM
77 GTATTCGAGT	31	54	Contig3370sp-fx	gi 4539543 emb CAB39974.1 glyceraldehyde-3-phosphate dehydrogenase .. 599 e-170	2.5 - GLUCOSE METABOLISM
78 CTGAGGCTTT	28	24	Contig4183gl-xy	gi 4539543 emb CAB39974.1 glyceraldehyde-3-phosphate dehydrogenase ... 580 e-164	2.5 - GLUCOSE METABOLISM
79 CTGGTGCAGG	11	15	Contig2421gr-ts	gi 415852 emb CAA79857.1 ribulose-1,5-bisphosphate carboxylase/oxygenase activase [Malus x domestica].. 357 3e-97	2.4 - ATP ASSOCIATED (CITRIC ACID CYCLES)
80 ATGTCCCGTT	25	38	Contig938gr-ml	gi 12620883 gb AAG61121.1 ribulose-1,5-bisphosphate carboxyl... 578 3e-163	2.4 - ATP ASSOCIATED (CITRIC ACID CYCLES)
81 TGAATCAATA	30	25	Contig1592sp-fx	gi 48209968 gb AAT40531.1 putative mitochondrial ATP synthase ... 300 2e-80	2.4 - ATP ASSOCIATED
82 TCTATGTCGG	8	21	Contig2697gr-se	gi 19787 emb CAA45153.1 chloroplast ATP synthase (delta subunit) [Nicotiana tabacum]. 273 6e-72	2.4 - ATP ASSOCIATED
83 TTTGTTAGTT	12	11	Contig73gr-yl	gi 3023743 sp 004683 FER1_MESCR Ferredoxin-1, chloroplast pre... 182 1e-44	2.3 - ELECTRON TRANSPORT
84 CTCTATTCT	86	129	Contig3372gr-ts	gi 3023743 sp 004683 FER1_MESCR Ferredoxin-1, chloroplast precursor (Ferredoxin I).... 180 4e-44	2.3 - ELECTRON TRANSPORT
85 TTGTATGGTT	24	38	Contig1195ur-xy	gi 18654477 gb AAL35363.2 thioredoxin peroxidase [Capsicum... 292 6e-78	2.3 - ELECTRON TRANSPORT
86 AACATACTGC	26	35	Contig4133gl-xy	gi 431957 emb CAA53900.1 thioredoxin m [Pisum sativum] >gi... 191 1e-47	2.3 - ELECTRON TRANSPORT
87 ATAACGTAC	27	33	Contig2696gr-ts	gi 20832 emb CAA45151.1 chloroplast Rieske FeS protein [Pisu... 327 3e-88	2.3 - ELECTRON TRANSPORT
88 AACACTAGGA	16	35	Contig2587gr-ts	gi 21592405 gb AAM64356.1 plastocyanin, putative [Arabidopsis t 186 5e-46	2.3 - ELECTRON TRANSPORT
89 GATGATGATT	14	17	Contig438gr-ts	gi 11135152 sp O81332 TRXF_MESCR Thioredoxin F-type, chloropl... 174 3e-42	2.3 - ELECTRON TRANSPORT
90 GTCGGAATGA	9	22	Contig627pe-xy	gi 21536569 gb AAM60901.1 NADH:ubiquinone oxidoreductase-lik... 196 4e-49	2.3 - ELECTRON TRANSPORT
91 ATCATCAGAT	15	11	Contig3179gr-ts	gi 33438474 emb CAE30280.1 chlorophyll a /b binding protein [Beta vulgaris].... 419 8e-116	2.2- PHOTOSYNTHESIS
92 TATTGCTGCT	103	124	Contig2699gr-ts	gi 3036945 dbj BAA25390.1 light harvesting chlorophyll a/b-b... 496 6e-139	2.2 - PHOTOSYNTHESIS
93 CTAATTGTCT	91	96	Contig908gr-ml	gi 170388 gb AAA34140.1 chlorophyll a/b-binding protein >gi ... 413 4e-114	2.2 - PHOTOSYNTHESIS

	Etiqueta	S	R	contig (Genolyptus)	NCBI e Outras bases	Categorias	
94	AACTACCTTT	55	54	Contig2750gr-ts	gi 728716 emb CAA59409.1 protein of photosystem II [Spinacia oleracea]. 100 3e-20	2.2 - PHOTOSYNTHESIS	-
95	TGAAAAGAGG	38	44	Contig1286gr-ml	gi 82080 pir S04125 chlorophyll a/b-binding protein type III precursor - tomato. 446 1e-123	2.2 - PHOTOSYNTHESIS	-
96	CAAGGTATT	32	30	Contig774gr-se	gi 37927438 gb AAP69815.1 chlorophyll a/b-binding protein [Viti 67.4 1e-10	2.2 - PHOTOSYNTHESIS	-
97	TCATTTAGCT	17	36	Contig3594gr-ts	gi 1181599 dbj BAA11677.1 subunit of photosystem I [Cucumis sat 306 7e-82	2.2 - PHOTOSYNTHESIS	-
98	TTTTTGCTTC	12	19	Contig2764gr-ts	gi 7443236 pir T09838 chlorophyll a/b binding protein prec... 466 e-130	2.2 - PHOTOSYNTHESIS	-
99	TTAAGTGTG	16	11	Contig2161pe-xy	gi 71040683 gb AAZ20290.1 type 2 metallothionein [Arachis hypog 65.5 2e-09	1.6 - COFACTORS AND SMALL MOLECULE METABOLISM	-
100	CTGTTGGTGA	14	9	Contig1340gr-pu	gi 22136156 gb AAM91156.1 putative thiamin biosynthesis protein... 120 2e-26	1.6 - COFACTORS AND SMALL MOLECULE METABOLISM	-
101	AAGTCCACA	17	29	Contig4137gl-xy	gi 1617270 emb CAA64327.1 acyl-CoA synthetase [Brassica na... 1079 0.0	1.5 - LIPID METABOLISM	-
102	CCTATTGAC	14	19	Contig63sp-fx	gi 21537410 gb AAM61751.1 putative 3-beta hydroxysteroid deh... 132 7e-30	1.5 - LIPID METABOLISM	-
103	GGATCCTGTG	25	34	Contig2449gr-ts	gi 51705411 gb AAU09443.1 UDP-glucose glucosyltransferase [Frag 635 0.0	1.4 - CARBOHYDRATE METABOLISM	-
104	TTAACTGGGT	24	26	Contig3101sp-fx	gi 30692244 ref NP_564372.3 UDP-N-acetylglucosamine pyrophos... 363 5e-99	1.4 - CARBOHYDRATE METABOLISM	-
105	GGATCCCGTG	20	14	Contig1153pe-xy	gi 51705411 gb AAU09443.1 UDP-glucose glucosyltransferase [Frag 255 9e-67	1.4 - CARBOHYDRATE METABOLISM	-
106	GTAATGAGAA	13	15	Contig4224gl-xy	gi 30699056 ref NP_849887.1 ADP-glucose pyrophosphorylase ... 657 0.0	1.4 - CARBOHYDRATE METABOLISM	-
107	TACTCGGATG	12	22	Contig2112gr-ts	gi 1750380 gb AAB39248.1 NADP-isocitrate dehydrogenase [Eu... 842 0.0	1.2 - NITROGEN METABOLISM	-
108	TATATCGTCG	25	34	Contig1701ur-xy	gi 99699 pir S18601 glutamate-ammonia ligase (EC 6.3.1.2)... 105 1e-21	1.2 - NITROGEN METABOLISM	-
109	GACATCTTCC	25	24	Contig2078gr-ts	gi 5733730 gb AAD49734.1 glutamine synthetase precursor [Juglan 766 0.0	1.2 - NITROGEN METABOLISM	-
110	AACTGGGCTA	18	16	Contig600gr-ml	gi 20259533 gb AAM13886.1 putative carbonic anhydrase, chlor... 211 1e-53	1.2 - NITROGEN METABOLISM	-
111	TTTAATAATG	16	13	Contig2209gr-ts	gi 13277342 emb CAC34417.1 Germin-like protein [Pisum sativum] 291 1e-77	1.2 - NITROGEN METABOLISM	-
112	TAATTAAGTC	30	36	Contig2714gr-ts	gi 34420138 gb AAQ67414.1 glycine decarboxylase complex H-protein ... 282 7e-75	1.1 - AMINO ACID METABOLISM	-
113	GTGCGGATGT	28	34	Contig2645gr-ts	gi 3913439 sp Q96471 DCAM_IPONI S-adenosylmethionine decarboxylase proenzyme ... 525 e-147	1.1 - AMINO ACID METABOLISM	-
114	GTCGATTCCT	8	21	Contig4232gl-xy	gi 30682078 ref NP_192966.2 copper amine oxidase, putative... 834 0.0	1.1 - AMINO ACID METABOLISM	-
115	AACTTGTGTC	6	18	Contig4135gl-xy	gi 1161600 emb CAA88841.1 phosphoglycerate kinase [Nicotia... 732 0.0	1 - GLUCOSE METABOLISM	-

Notas: Sinais convencionais utilizados

N número da etiqueta

S expressão na biblioteca susceptível

R expressão na biblioteca resistente

4.7 Discussão do envolvimento de genes anotados na resposta à infecção de *P. psidii*

No presente trabalho, foram anotadas várias etiquetas com funções relacionadas ao fortalecimento da parede celular, polarização celular, com a resposta hipersensível, de resistência sistêmica adquirida e estresse abiótico, confirmando um panorama onde tanto os indivíduos susceptíveis como resistentes respondem ao patógeno, porém diferencialmente. Estes mecanismo podem ser analisados da seguinte ótica:

4.7.1 Componentes Estruturais e a polarização celular

No referente a polarização celular e fortalecimento da parede celular, uma vez que o fungo patógeno contactou a superfície da planta, a polarização celular leva a um profundo rearranjo do citoesqueleto, tanto dos componentes como das proteínas associadas, são polimerizados microtúbulos e filamentos de actina até o local da infecção, levando vesículas e o núcleo até o local (SCHMELZER, 2002).

Nos dados analisados foram detectadas cinco etiquetas com alta identidade às isoformas de tubulina, (etiquetas de 19-22 da tabela 6 e a etiqueta 48 da tabela 7), outra com identidade à actina (etiqueta 24 da tabela 6), e quatro com identidade à proteínas associadas ao citoesqueleto (etiqueta 23 da tabela 6, etiquetas 49-50 da tabela 7 e a etiqueta 22 da tabela 5) cujas respectivas expressões totalizam 26 e 100 cópias, 1 e 14 cópias e 34 e 45, para a biblioteca susceptível e resistente, respectivamente.

Após o rearranjo do citoesqueleto acontece um depósito de material e um robustecimento considerável da parede celular chamado de papila. A composição desta é extremadamente variável, lignina, celulose, pectina, calose, quitina, suberina, lipídeos, proteínas ricas em hidroxiprolina, peroxidases, etc (SCHMELZER, 2002).

Os dados apresentaram unicamente no grupo de genes preferencialmente expressos na biblioteca resistente, duas etiquetas relacionadas ao metabolismo de glicoproteínas (de 25-26 tabela 6), este proteoglicano, de matriz extracelular foi relacionado com a diminuição da eficiência de infecção (GASPAR et al., 2004). Foram também detectadas sete etiquetas relacionadas ao metabolismo de lignina (de 27-33 tabela 6). Ainda foram contabilizadas quatro etiquetas relacionadas ao metabolismo de arabinogalactanos (de 34-37 tabela 6), uma relacionada ao metabolismo de pectinas (etiqueta 38 tabela 6) e uma relacionada ao metabolismo de celulose unicamente encontradas na biblioteca resistente (etiqueta 39).

Os dados mostram uma resposta de fortalecimento da parede celular e de polarização celular importantes, apoiando a hipótese de que estes são componentes da resposta a este patógeno.

4.7.2 Explosão oxidativa, resposta hipersensível, resistência sistêmica adquirida e estresse abiótico

A explosão oxidativa é uma resposta da planta à presença do patógeno quando forma parte do que se conhece como resposta hipersensível. Esta resposta produzirá diversos compostos

denominados espécies reativas de oxigênio ou também chamadas de ROS - *Reactive Oxygen Species*. Estas espécies de oxigênio são o peróxido de hidrogênio (H_2O_2), outros peróxidos, o ânion superóxido (O_2^-) e radicais hidroxilas livres ($OH^\cdot$) (RESENDE; SALGADO; CHAVES, 2003). Embora reativas por natureza própria podem reagir com íons metálicos intracelulares e converte-se em espécies ainda mais reativas (GIJZEN et al., 1999).

Portanto, no percurso da evolução, os vegetais desenvolveram enzimas reguladoras que podem detoxificar eficientemente a célula destas espécies ativas de oxigênio. Tais como: a Superóxido Dismutase que catalisa a dismutação do O_2^- em H_2O_2 , a Catalase que converte H_2O_2 e O_2 (RESENDE; SALGADO; CHAVES, 2003), a peroxiredoxina Q que é uma peroxidase ubíqua dependente de tioredoxina ou glutaredoxina e cuja função é a destruição de peróxidos (ROUHIER et al., 2004). Devido a vida media curta que apresentam estes compostos, a presença deles pode ser deduzida pela presença das moléculas que os neutralizam.

No presente trabalho, encontramos diversas etiquetas com funções relacionadas com várias destas enzimas duas superóxido dismutases, uma expressa nas duas bibliotecas (etiqueta 66 tabela 7), porém a outra exclusivamente expressa na biblioteca resistente (etiqueta 58 tabela 6). Ainda foram detectadas três catalases (etiquetas 54 e 59-60 tabela 6) preferencialmente expressas na biblioteca resistente, cujas expressões totalizam 49 e 60 cópias, nas bibliotecas susceptível e resistente, respectivamente. Também foram encontradas duas catalase preferencialmente expressas na biblioteca susceptível (etiqueta 48 e 51 tabela 5), totalizando uma expressão de 41 e 23 cópias, na biblioteca susceptível e resistente, respectivamente. Isso apresenta uma resposta ao estresse oxidativo em ambas bibliotecas, porém um pouco maior nos indivíduos resistentes.

Após essa explosão oxidativa a célula sofre fluxos iônicos, síntese de óxido nítrico, uma reprogramação transcricional, síntese de quinases, fosfatases e fatores de transcrição (DANGL; JONES, 2001). No presente trabalho também foram detectados desses genes preferencialmente expressos na biblioteca resistente, entre eles seis quinases (de 71-76 tabela 6) e duas fosfatases (de 77-78 tabela 6), cuja expressão totalizava 14 e 74 cópias para as quinases e 3 e 22 cópias para as fosfatases, nas bibliotecas susceptível e resistente respectivamente, demonstrando uma forte reprogramação metabólica na célula.

Após a reprogramação ocorre a síntese de ácido salicílico e etileno, o fortalecimento da parede celular, síntese de componentes antimicrobianos e por último a resposta hipersensível

(DANGL; JONES, 2001). No presente trabalho foi detectada um enzima produtora de etileno (etiqueta 67 tabela 7) cuja expressão foi de 14 e 12 cópias para a biblioteca susceptível e resistente, respectivamente. Os genes de fortalecimento da parede celular que já foram acima descritos (em fortalecimento da parede e polarização celular) se expressam fortemente na biblioteca resistente. Por outro lado, na síntese de genes PR surgem uma importante diferença entre ambas bibliotecas. Os indivíduos susceptíveis produzem preferencialmente glucanases (etiqueta 41 tabela 5), contabilizando 40 cópias e 3 nos indivíduos susceptíveis e resistentes, respectivamente, sendo estes compostos degradadores das paredes fúngicas e liberadores de elicitores. Também produzem defensinas (etiqueta 42 tabela 5) que interagem com lipídeos da membrana plasmática e produzem a formação de poros e a saída de componentes celulares e desestabilizam o potencial de membrana (THEIS; STAHL, 2004; BROEKAERT et al., 1995). No presente trabalho, uma etiqueta foi anotada com função de defensina (etiqueta 42 tabela 5), com 50 cópias na biblioteca susceptível e 12 na resistente. Foram também detectadas taumatinas (etiquetas 43 e 52 tabela 5) que apresentam atividades β 1,3-glucanase. Ainda foram detectadas três genes de proteínas relacionadas com a patogênese (de 45-46 e 49 tabela 5) com 22 e 4 cópias, na biblioteca susceptível e resistente, respectivamente. No caso de taumatinas e defensinas preferencialmente expressos no susceptível poderia indicar que estas respostas não são suficientes para conter o avanço desse patógeno.

Por outro lado, os indivíduos resistentes expressam uma bateria de genes diferentes. A resposta desses é baseada em quitinases, hidrolizando a quitina e debilitando a parede celular, conseqüentemente produz lise celular (THEIS; STAHL, 2004; FUKUDA, 1997). No presente trabalho foram anotadas quatro etiquetas com funções de quitinase, duas delas com expressão exclusiva na biblioteca resistente (etiquetas 51 e 55 tabela 6), totalizando 57 cópias. Interessante, existe outra etiqueta também codificando outra quitinase que unicamente se expressa na biblioteca susceptível (etiqueta 47 tabela 5) porém a expressão é baixa, quatro cópias. Por outro lado também detectada outra quitinase, cuja expressão não é diferencial (etiqueta 64 tabela 7), com 66 e 64 cópias na biblioteca susceptível e resistente respectivamente. Provavelmente em *Eucalyptus* existam várias isoformas de quitinases, e os indivíduos resistentes conseguem expressar três das quatro detectadas. Estas isoformas de quitinases podem ter a função de gerar elicitores ou atacar diretamente a parede celular do fungo.

Também a produção de uma fitoalexina é indicada pela detecção de isoflavona redutase preferencialmente expressos na biblioteca resistente. Já foi demonstrado *in vitro* a relação da presença de fitoalexinas e a tolerância das plantas a diversos fungos fitopatogênicos (VAN ETTEN; MATTHEWS; MATTHEWS, 1989; WU; VAN ETTEN, 2004). Duas etiquetas com alta similaridade com isoflavona redutases apresentam uma expressão preferencial de 99 cópias no resistente e 3 no susceptível (etiqueta 57 e 61 tabela 6), indicando provavelmente a via biossintética de fitoalexinas como elemento importante na resposta dos indivíduos resistentes.

Por último, os indivíduos resistentes expressam diferencialmente genes que codificam proteínas LTP – *Lipid Transfer Proteins* (Proteínas de transferência de lipídeos), as que apresentam forte atividade antimicrobiana (CAMMUE et al., 1995). No presente trabalho duas etiquetas foram anotada com a função LTP (de 52-53 tabela 6). A expressão das mesmas foi de 140 e 259, para a biblioteca susceptível e resistente, respectivamente.

Normalmente a indução de uma resposta hipersensível requer uma interação gene a gene (FLOR; 1971) sendo um dos genes da planta infectada e o outro do patógeno. Tradicionalmente os genes que interagem com o gene avr do patógeno pertencem a família NB-LRR, chamados de genes R. Infelizmente, nenhum gene dessa família foi encontrado no material resistente provavelmente este gene não esteja representado no banco de dados Genolyptus, porém, um gene com similaridade aos genes Harpin foi encontrado (etiqueta 56 tabela 6). Os produtos dos genes Harpin também são conhecidos como indutores da resposta hipersensível em plantas (NEWMAN et al., 2001). Portanto, este gene poderia substituir a função dos genes R na indução da resposta hipersensível em eucalipto.

Por outro lado, o peróxido de hidrogênio também é responsável da ativação da hidrolase do ácido benzóico e esta é responsável pela produção do ácido salicílico. O que permitiria uma resposta de defesa dependente do ácido salicílico, permitindo explosões oxidativas no sítio da infecção e explosões secundárias longe do sítio de infecção, induzindo a expressão de proteínas relacionadas à patogênese o que se conhece como Resistência Sistêmica Adquirida (SAR – *Systemic Acquired Resistance*) (SHAH, 2003), por exemplo, permitiria a síntese de quitinases, glucanases, thaumatinas, inibidores de protease, etc. em longas distâncias.

A uma resposta mais generalizada contra o estresse foi também induzida nos dois fenótipos, respondendo aos estresses oxidativos, por metais pesados e outros compostos tóxicos, e ambientais, exemplificado pela presença de genes relacionados com a produção de glutathionas

(etiquetas 35 e 39 tabela 5; etiqueta 47 tabela 6), chaperonas (etiqueta 53 tabela 5), USP-*universal stress protein* (etiqueta 54 tabela 5) e metalotioninas (etiqueta 38 tabela 5; de 48-49 tabela 6).

Os dados em conjunto apóiam a idéia de que a explosão oxidativa, resposta hipersensível, a ativação diferencial de compostos antimicrobianos e genes relacionados a patogênese (PR) são parte fundamental da resposta à *Puccinia psidii*.

As presenças de diferentes genes expressados de maneira diferencial nas duas bibliotecas poderiam apoiar as observações de Casteels et al. (1993) de que muitas proteínas antifúngicas podem atuar sinergicamente com outras proteínas antifúngicas. Estes efeitos sinérgicos podem alterar a atividade ou ainda a especificidade das respostas (CASTEELS, et al., 1993). No entanto, percebe-se a presença de uma resposta multifatorial com diferentes níveis no espaço e no tempo.

Provavelmente quando as células percebem a presença dos patógenos, estes já se encontraram com o primeiro impedimento das barreiras físicas, já existentes ou rapidamente induzidas, seguido de complexos passos metabólicos que são postos em marcha. Uma polarização celular, cujo objetivo é a de maximizar a eficiência da resposta contra o patógeno e o fortalecimento da parede, é acionada. Espécies reativas de oxigênio são enviadas contra o patógeno e acontece a explosão oxidativa. Ao mesmo tempo uma eficiente resposta antioxidativa é induzida para conter os danos causados pelas espécies reativas de oxigênio. Por sua vez, estes compostos reativos de oxigênio induzem a expressão de uma série de proteínas com atividades fúngicas (PR: proteínas relacionadas com a patogênese), e a ativação da resposta dependente de ácido salicílico, que permitira o estabelecimento da Resistência Sistêmica Adquirida. Se houver uma interação gene a gene, gene avr do fungo com o gene R da planta, esta explosão oxidativa se transformará na mais severa medida de defesa que toma a planta a Resposta Hipersensível.

5 CONCLUSÕES

A metodologia SAGE se apresentou eficiente na geração das bibliotecas resistentes e susceptíveis, para a obtenção de perfis de expressão, e ofereceu um meio para analisar diferenças e semelhanças na expressão gênica entre dois fenótipos (resistente e susceptível). Provando que é uma ferramenta poderosa para acrescentar o entendimento dos processos metabólicos envolvidos na resposta de frente à ferrugem.

A grande maioria dos genes apresenta expressões semelhantes nos dois fenótipos em presença do fungo *Puccinia psidii*. So uma pequena porção desses genes apresenta expressão diferencial e unicamente 232 etiquetas apresentam uma expressão preferencial na biblioteca resistente.

As barreira físicas, a polarização celular apresentam-se como importantes vias de resposta ao patógeno, que utilizam os indivíduos resistentes.

A resposta ao estresse oxidativo e ao estresse biótico são componentes importantes neste processo de resposta ao patógeno. Esses são expressos de maneira diferencial nos indivíduos resistente e susceptíveis.

Tanto indivíduos resistentes e susceptíveis respondem à esta doença. No entanto a resposta é diferente em ambos fenótipos. As baterias de genes de resposta a doenças é diferencial em ambos fenótipos. A resposta hipersensível e a resistência sistêmica adquirida são utilizadas e são diferencialmente expressas entre estes fenótipos.

Todos os dados em conjunto suportam a idéia de que não há um mecanismo, porém vários mecanismos atuam ao mesmo tempo de forma concertada produzindo o fenotipo resistente, entre eles, resposta eficiente ao estresse oxidativo, robustecimento da parede celular, polarização celular, resposta hipersensível e resistência sistêmica adquirida.

REFERÊNCIAS

- A MADEIRA de eucalipto. **Revista da Madeira**, São Paulo, n. 59, p. 6-8, set. 2001.
- AIST, J.R. Papillae and related wound plugs of plant cells. **Annual Review of Phytopathology**, Colorado, v.14, p. 145-163, 1976.
- ALFENAS, A.C.; DEMUNER, N.L.; BARBOSA, M.M.A. Ferrugem e as opções de controle. **Correio Agrícola**, v. 1, p. 18-20, 1989.
- ALTSCHUL, S.F.; MADDEN, T.L.; SCHAFER, A.A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D.J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Research**, London, v. 25, n. 17, p.3389-402, 1997.
- AMBIENTEBRASIL. Disponível em:
<<http://www.ambientebrasil.com.br/composer.php3?base=/florestal/index.html&conteudo=/florestal/eucalipto.html>>. Acesso em 13 dez. 2005.
- ASAI, T.; TENA, G.; PLOTNIKOVA, J.; WILLMANN, M.R.; CHIU, W.-L.; GOMEZ-GOMEZ, L.; BOLLER, T.; AUSUBEL, F.M.; SHEEN, J. MAP kinase signalling cascade in *Arabidopsis* innate immunity. **Nature**, London, v.415, p.977-983, 2002.
- ASHFIELD, T.; BOCIAN, A.; HELD, D.; HENK, A.D.; MAREK, L.F.; DANESH, D.; PENUELA, S.; MEKSEM, K.; LIGHTFOOT, D.A.; YOUNG, N.D.; SHOEMAKER, R.C.; INNES, R.W. Genetic and physical localization of the soybean Rpg1-b disease resistance gene reveals a complex locus containing several tightly linked families of NBS-LRR genes. **Molecular Plant Microbe Interact**, St Paul, v.16, n.9, p.817-826, 2003.
- AUDIC, S.; CLAVERIE, J.M. The Significance of Digital Gene Expression Profiles. **Genome Research**, Woodbury, v.7, n.10, p.986-995, 1997.
- BATEMAN, A.; COIN, L.; DURBIN, R.; FINN, R.D.; HOLLICH, V.; GRIFFITHS-JONES, S.; KHANNA, A.; MARSHALL, M.; MOXON, S.; SONNHAMMER, E.L.L.; STUDHOLME, D.J.; YEATS, C.; EDDY, S.R. The Pfam protein families database. **Nucleic Acids Research**, London, v. Database Issues, n. 32, p.138-141, 2004.
- BATEMAN, A.; BIRNEY, E.; CERRUTI, L.; DURBIN, R.; ETWILLER, L.; EDDY, S.R.; GRIFFITHS-JONES, S.; HOWW, K.L.; MARSHALL, M.; SONNHAMMER, E.L.L.; STUDHOLME, D.J.; YEATS, C. The Pfam protein families database. **Nucleic Acids Research**, London, v. 30, n. 1, p.276-280, 2002.
- BATEMAN, A.; BIRNEY, E.; DURBIN, R.; EDDY, S.R.; HOWE, K.L.; SONNHAMMER, E.L.L. The Pfam protein families database. **Nucleic Acids Research**, London, v. 28, p.263-266, 2000.
- BATEMAN, A.; BIRNEY, E.; DURBIN, R.; EDDY, S.R.; FINN, R.D.; SONNHAMMER, E.L.L. Pfam 3.1: 1313 multiple alignments match the majority of proteins. **Nucleic Acids Research**, London, v. 27, p.260-262, 1999.

- BERTOLUCCI, F.; REZENDE, G.; PENCHEL, R. Produção e utilização de híbridos de eucalipto. **Silvicultura**, São Paulo, v.51, p.12-16, 1995.
- BOHELER, K.R.; STERN, M.D. The New Role of Sage in Gene Discovery. **Trends in Biotechnology**, Cambridge, v.21, n.2, p.55-57, 2003.
- BOOTH, T.H.; OLD, K.M.; JOVANOVIĆ, T. A Preliminary Assessment of High Risk Areas for *Puccinia psidii* (*Eucalyptus* Rust) in the Neotropics and Australia. **Agriculture Ecosystems & Environment**, Amsterdam, v.82, n.1-3, p.295-301, 2000.
- BORK, P.; KOONIN, E.V. Predicting functions from protein sequences - where are the bottlenecks?. **Nature Genetics**, New York, v. 18, n. 4, p. 313-318, 1998.
- BRASIL - MINISTÉRIO DO MEIO AMBIENTE. **Programas Nacional de Florestas**. Brasília, 2005. Disponível em: <http://www.mma.gov.br/index.cfm?id_estrutura=5&id_menu=1194>. Acesso em 11 nov. 2005.
- BRASILEIRA DE CELULOSE E PAPEL, ASSOCIAÇÃO. **Desempenho do setor em 2003**. São Paulo, 2003. 4p.
- BRASILEIRA DE CELULOSE E PAPEL, ASSOCIAÇÃO. **O setor brasileiro de celulose e papel: perfil do setor**. São Paulo, 2004. 6p.
- BRASILEIRA DE CELULOSE E PAPEL, ASSOCIAÇÃO. **Espécies mais plantadas existentes em 31/12/2004: folhosas**. São Paulo, 2004b. 1p.
- BRASILEIRA DE CELULOSE E PAPEL, ASSOCIAÇÃO. **Utilidades do eucalipto**. São Paulo, 2005. 2p.
- BROEKAERT, W.R.; TERRAS, F.R.; CMMUE, B.P.; OSBORN, R.W. Plant defensins: novel antimicrobial peptides as components of the host defense system. **Plant Physiology**, Rockville, v.108, p. 1353-1358, 1995.
- CAMMUE, B.P.A.; THEVISSSEN, K.; HENDRIKS, M.; EGGERMONT, K.; GODERIS, I.J.; PROOST, P.; VAN DAMME, J.; OSBORN, R.W.; GUERBETTE, F.; KADER, J.C.; BROEKAERT, W.F. A potent antimicrobial protein from onion seeds showing sequence homology to plant lipid transfer proteins. **Plant Physiology**, Rockville, v. 109, p. 445-455, 1995.
- CASTEELS, P.; AMPE, C.; JACOBS, F.; TEMPST, P. Functional and chemical characterization of hymenoptaecin, an antimicrobial polypeptide that is infection-inducible in the honeybee (*Apis mellifera*). **The Journal of Biological Chemistry**, Baltimore, v.268, p.7044-7054, 1993.
- CENTURY, K.S.; HOLUB, E.B.; STASKAWICZ, B.J. Ndr1, a Locus of Arabidopsis-Thaliana That Is Required for Disease Resistance to Both a Bacterial and a Fungal Pathogen. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v.92, n.14, p.6597-6601, 1995.

- CERUTTI, J.M.; RIGGINS, G.J.; de SOUZA, S.J. What can digital transcript profiling reveal about human cancers?. **Brazilian Journal of Medical and Biological Research**, Ribeirão Preto, v. 36, p. 975-985, 2003.
- DE CHALDEE, M.; GAILLARD, M.C.; BIZAT, N.; BUHLER, J.M.; MANZONI, O.; BOCKAERT, J.; HANTRAYE, P.; BROUILLET, E.; ELALOUF, J.M. Quantitative assessment of transcriptome differences between brain territories. **Genome Research**, Woodbury, v. 13, n. 7, p. 1646-1653, 2003.
- CHEN, J.J.; SUN, M.; LEE, S.G.; ZHOU, G.L.; ROWLEY, J.D.; WANG, S. Identifying novel transcripts and novel genes in the human genome by using novel SAGE tags. **Proceedings Of The National Academy Of Sciences Of The United States Of America**, Washington, v. 99, n. 19, p. 12257-12262, 2002
- CHEONG, J.-J.; YANG, D.C. Methyl jasmonate as a vital substance in plants. **Trends in Genetics**, Cambridge, v.19, n.7, p.409-413, 2003.
- CHOI, D.; KIM, H.M.; YUN, H.K.; PARK, J.A.; KIM, W.T.; BOK, S.H. Molecular cloning of a metallothionein-like gene from *Nicotiana glutinosa* L. and its induction by wounding and tobacco mosaic virus infection. **Plant Physiology**, Rockville, v. 112, p. 353-359, 1996.
- COBBOLD, S.P.; NOLAN, K.F.; GRACA, L.; CASTEJON, R.; LE MOINE, A.; FREWIN, M.; HUMM, S.; ADAMS, E.; THOMPSON, S.; ZELENKA, D.; PATERSON, A.; YATES, S.; FAIRCHILD, P.J.; WALDMANN, H. Regulatory T cells and dendritic cells in transplantation tolerance: molecular markers and mechanisms. **Immunological Reviews**, Copenhagen, v. 196, n. 1, p. 109-124, 2003.
- COUTINHO, T.A.; WINGFIELD, M.J.; ALFENAS, A.C.; CROUS, P.W. Eucalyptus Rust: a Disease With the Potential for Serious International Implications. **Plant Disease**, St Paul, v.82, n.7, p.819-825, 1998.
- DANGL, J. Innate Immunity - Plants Just Say No to Pathogens. **Nature**, London, v.394, n.6693, p.525-+, 1998.
- DANGL, J.L.; JONES, J.D.G. Plant Pathogens and Integrated Defence Responses to Infection. **Nature**, London, v.411, n.6839, p.826-833, 2001.
- DE CARVALHO, A.D.O.; ALFENAS, A.C.; MAFFIA, L.A.; DO CARMO, M.G.F. Resistance of *Eucalyptus* Species, Progenies and Provenances to *Puccinia psidii* Winter. **Pesquisa Agropecuaria Brasileira**, Rio de Janeiro, v.33, n.2, p.139-147, 1998.
- DE MIRANDA, J.R.; THOMAS, M.A.; THURMAN, D.A.; TOMSETT, A.B. Metallothionein gene from the flowering plant *Mimulus guttatus*. **FEBS Letters**, Amsterdam, v. 260, p. 277-280, 1990.
- DEBENER, T.; LEHNACKERS, H.; ARNOLD, M.; DANGL, J.L. Identification and Molecular Mapping of a Single Arabidopsis- Thaliana Locus Determining Resistance to a Phytopathogenic Pseudomonas-Syringae Isolate. **The Plant Journal**, Oxford, v.1, n.3, p.289-

302, 1991.

DESPREZ, T.; AMSELEM, J.; CABOCHE, M.; HÖFTE, H. Differential gene expression in *Arabidopsis* monitored using cDNA arrays. **The Plant Journal**, Oxford, v. 14, n. 5, p.643-652, 1998.

DIANESE, J.C.; MORAES, T.S.D.; SILVA, A.R. Response of Eucalyptus Species to Field Infection by Puccinia- Psidii. **Plant Disease**, St Paul, v.68, n.4, p.314-316, 1984.

DINESH-KUMAR, S.P.; THAM, W.H.; BAKER, B.J. Structure-function analysis of the tobacco mosaic virus resistance gene N. **Proceedings Of The National Academy Of Sciences Of The United States Of America**, Washington, v. 97, n. 26, p. 14789-14794, 2000.

EMBRAPA. Disponível em:
<http://www21.sede.embrapa.br/linhas_de_acao/temas_basicos/florestas/index_html>. Acesso em: 17 nov. 2005.

ESPÉCIES de eucalipto. **Revista da Madeira**, São Paulo, n. 59, p. 16-22, set. 2001

EWING, B.; HILLIER, L.D.; WENDL, M.C.; GREEN, P. Base-calling of automated sequencer traces using *Phred*. I. Accuracy assessment. **Genome Research**, Woodbury, v. 8, p. 175-185, 1998.

EWING, B.; GREEN, P. Base-calling of automated sequencer traces using *Phred*. II. Error probabilities. **Genome Research**, Woodbury, v. 8, p. 186-194, 1998.

FAO. Disponível em:
<http://www.fao.org/documents/show_cdr_head.asp?url_file=?docrep/004/y1997e/y1997e08.htm>. Acesso em: 10 dez. 2005.

FERREIRA, M. Escolha de espécies de eucalipto. **Circular técnica IPEF**, São Paulo, v.47, p.1-30, 1979.

FIGUEIREDO, M.B. Doenças fúngicas emergentes em grandes culturas. **Biológico**, São Paulo, v.63, n.1, p.29-32, 2001.

FIZAMES, C.; MUÑOS, S.; CAZETTES, C.; NACRY, P.; BOUCHEREZ, J.; GAYMARD, F.; PIQUEMAL, D.; DELORME, V.; COMMES, T.; DOUMAS, P.; COOKE, R.; MARTI, J.; SENTENAC, H.; GOJON, A. The *Arabidopsis* root transcriptome by serial analysis of gene expression. Gene identification using the genome sequence. **Plant Physiology**, Rockville, v. 134, p. 67-80, jan 2004.

- FLEISCHMANN, R.D.; ADAMS, M.D.; WHITE, O.; et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* RD. **Science**, Washington, v. 269, n. 5223, p. 496-512, 1995.
- FLOR, H.H. Current status of the gene-for-gene concept. **Annual Review of Phytopathology**, Colorado, v. 9, p. 275-298, 1971.
- FRASER, C.M.; GOCAYNE, J.D.; WHITE, O.; ADAMS, M.D.; CLAYTON, R.A.; FLEISCHMANN, R.D.; BULT, C.J.; KERLAVAGE, A.R.; SUTTON, G.; KELLEY, J.M.; FRITCHMAN, J.L.; WEIDMAN, J.F.; SMALL, K.V.; SANDUSKY, M.; FUHRMANN, J.; NGUYEN, D.; UTTERBACK, T.R.; SAUDEK, D.M.; PHILLIPS, C.A.; MERRICK, J.M.; TOMB, J.F.; DOUGHERTY, B.A.; BOTT, K.F.; HU, P.C.; LUCIER, T.S.; PETERSON, S.N.; SMITH, H.O.; HUTCHISON, C.A.; VENTER, J.C. The minimal gene complement of *Mycoplasma genitalium*. **Science**, Washington, v. 270, n. 5235, p. 397-403, 1995.
- FREGENE, M.; MATSUMURA, H.; AKANO, A.; DIXON, A.; TERAUCHI, R. Serial analysis of gene expression (SAGE) of host-plant resistance to the cassava mosaic disease (CMD). **Plant Molecular Biology**, Dordrecht, v. 56, p. 563-571, 2004.
- FUKUDA, Y. Interaction of tobacco nuclear protein with a elicitor-responsive element in the promoter of a basic class I chitinase gene. **Plant Molecular Biology**, Dordrecht, v. 34, p. 81-87, 1997.
- GASPAR, Y.M.; NAM, J.; SCHULTZ, C.J.; LEE, L.Y.; GILSON, P.R.; GELVIN, S.B.; BACIC, A. Characterization of the *Arabidopsis* lysine-rich arabinogalactan-protein *AtAGP17* mutant (*rat1*) that results in a decreased efficiency of *Agrobacterium* transformation. **Plant Physiology**, Rockville, v. 135, p. 2162-2171.
- GENOLIPTUS. Disponível em <<http://genoma.embrapa.br/genoma/>>. Acesso em: jan. 2005.
- GIBAS, C.; JAMBECK, P. **Desenvolvendo bioinformática**. Tradução de C.A Machado. Rio de Janeiro: Campus, 2001. 440p.
- GIBBINGS, J.G.; COOK, B.P.; DUFAULT, M.R.; MADDEN, S.L.; KHURI, S.; TURNBULL, C.J.; DUNWELL, J.M. Global transcript analysis of rice leaf and seed using SAGE technology. **Plant Biotechnology**, New York, v.1, p.271-285, 2003.
- GIJZEN, M.; MILLER, S.S.; BOWMAN, L.A.; BATCHELOR, A.K.; BOUTILIER, K.; MIKI, B.L.A. Localization of peroxidase mRNAs in soybean seeds by *in situ* hybridization. **Plant Molecular Biology**, Dordrecht, v. 41, p. 57-63, 1999.
- GLAZEBROOK, J. Genes Controlling Expression of Defense Responses in Arabidopsis. **Current Opinion in Plant Biology**, London, v.2, n.4, p.280-286, 1999.

- GOFFEAU, A.; BARRELL, B.G.; BUSSEY, H.; DAVIS, R.W.; DUJON, B.; FELDMANN, H.; GALIBERT, F.; HOHEISEL, J.D.; JACQ, C.; JOHNSTON, M.; LOUIS, E.J.; MEWES, H.W.; MURAKAMI, Y.; PHILIPPSSEN, P.; TETTELIN, H.; OLIVER, S.G. Life with 6000 genes. **Science**, Washington, v. 274, n. 5287, p. 546-567, 1996.
- GOMEZ-GOMEZ, L.; BOLLER, T. Flagellin Perception: a Paradigm for Innate Immunity. **Trends in Plant Science**, Kidlington, v. 7, n. 6, p. 251-256, 2002.
- GORDON, D.; ABAJIAN, C.; GREEN, P. A graphical tool for sequence finishing. **Genome Research**, Woodbury, v. 8, n. 3, p. 195-202, 1998.
- GRANT, M.R.; GODIARD, L.; STRAUBE, E.; ASHFIELD, T.; LEWALD, J.; SATTLER, A.; INNES, R.W.; DANGL, J.L. Structure of the Arabidopsis Rpm1 Gene Enabling Dual-Specificity Disease Resistance. **Science**, Washington, v.269, n.5225, p.843-846, 1995.
- GRANT, M.R.; MCDOWELL, J.M.; SHARPE, A.G.; ZABALA, M.D.T.; LYDIATE, D.J.; DANGL, J.L. Independent Deletions of a Pathogen-Resistance Gene in Brassica and Arabidopsis. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v.95, n.26, p.15843-15848, 1998.
- GROSS, P.; JULIUS, C.; SCHMELZER, E.; HAHLBROCK, K. Translocation of cytoplasm and nucleus to fungal penetration sites is associated with depolymerization of microtubules and defense gene activation in infected, cultured parsley cells. **EMBO Journal**, Oxford, v. 12, n. 5, p. 1735-1744, 1993.
- HAMER, D.H. Metallothionein. **Annual Review of Biochemistry**, Palo Alto, v. 55, p. 913-951, 1986.
- HAMMOND-KOSACK, K.E.; PARKER, J.E. Deciphering Plant-Pathogen Communication: Fresh Perspectives for Molecular Resistance Breeding. **Current Opinion in Biotechnology**, Philadelphia, v. 14, n. 2, p. 177-193, 2003.
- HO, C.K.; CHANG, S.H.; TSAY, J.Y.; TSAI, C.J.; CHIANG, V.L.; CHEN, Z.Z. Agrobacterium Tumefaciens-Mediated Transformation of Eucalyptus Camaldulensis and Production of Transgenic Plants. **Plant Cell Reports**, New York, v. 17, n. 9, p. 675-680, 1998.
- HOLT III, B.F.; HUBERT, D.A.; DANGL, J.L. Resistance Gene Signaling in Plants - Complex Similarities to Animal Innate Immunity. **Current Opinion in Immunology**, London, v. 15, n. 1, p. 20-25, 2003.
- HUANG, X.; KIEFER, E.; VON RAD, U.; ERNST, D.; FOISSNER, I.; DURNER, J. Nitric Oxide Burst and Nitric Oxide-Dependent Gene Induction in Plants. **Plant Physiology and Biochemistry**, Paris, v. 40, n. 6-8, p. 625-631, 2002.
- HUCKELHOVEN, R.; et al. Hypersensitive cell death and papilla formation in barley attacked by the powdery mildew fungus are associated with hydrogen peroxide but not with salicylic acid accumulation. **Plant Physiology**, Rockville, v. 119, n. 4, p. 1251-1260, 1999.

- IBAMA - INSTITUTO BRASILEIRO DO MEIO AMBIENTE E DOS RECURSOS NATURAIS RENOVÁVEIS. **Relatório perspectivas do meio ambiente do Brasil**. Brasília, 2002. 450 p.
- INNES, R.W.; BENT, A.F.; KUNKEL, B.N.; BISGROVE, S.R.; STASKAWICZ, B.J. Molecular Analysis of Avirulence Gene Avrpt2 and Identification of a Putative Regulatory Sequence Common to All Known Pseudomonas-Syringae Avirulence Genes. **Journal of Bacteriology**, Baltimore, v. 175, n. 15, p. 4859-4869, 1993.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. **Produção da extração vegetal e da silvicultura – 2004**. Brasília, 2005. 6 p.
- INSTITUTO DE PESQUISA E ESTUDOS FLORESTAIS. **MCT/IPEF – Pesquisa Florestal no Brasil: Melhoramento e conservação genética**. São Paulo, 2005. 10 p.
- JENNER, C.; HITCHIN, E.; MANSFIELD, J.; WALTERS, K.; BETTERIDGE, P.; TEVERSON, D.; TAYLOR, J. Gene-for-Gene Interactions Between Pseudomonas-Syringae Pv Phaseolicola and Phaseolus. **Molecular Plant-Microbe Interactions**, St Paul, v.4, n.6, p.553-562, 1991.
- JUNG, S.H.; LEE, J.Y.; LEE, D.H. Use of Sage Technology to Reveal Changes in Gene Expression in Arabidopsis Leaves Undergoing Cold Stress. **Plant Molecular Biology**, Dordrecht, v. 52, n. 3, p. 553-567, 2003.
- JUNGHANS, D.T. **Quantificação da severidade, herança da resistência e identificação de marcadores RAPD ligados à resistência à ferrugem (*Puccinia psidii*) em *Eucalyptus grandis***. 2000. 53 p. Tese (Doctor Scientiae) - Universidade Federal de Viçosa, Viçosa, 2000.
- JUNGHANS, D.T; ALFENAS, A.C.; BROMMONSCHENKEL, S.H.; ODA, S.; MELLO, E.J.; GRATAPAGLIA, D. Resistance to rust (*Puccinia psidii* Winter) in *Eucalyptus*: mode of inheritance and mapping of a major gene with RAPD markers. **Theoretical and Applied Genetics**, New York, v. 108, n. 1, p. 175-180, dec 2003.
- JUNGHANS, D.T; ALFENAS, A.C.; MAFFIA, L.A. Escala de notas para quantificação da ferrugem em *Eucalyptus*. **Fitopatologia brasileira**, Brasília, v. 28, n. 2, p. 184-188, 2003.
- KAL, A.J; VAN ZONNEVELD, A.J; BENES, V.; VAN DEN BERG, M.; KOERKAMP, M.G.; ALBERMANN, K.; STRACK, N.; RUIJTER, J.M.; RICHTER, A.; DUJON, B.; ANSORGE, W.; TABAK, H. Dynamics of gene expression revealed by comparison of serial analysis of gene expression transcript profiles from yeast grown on two different carbon sources. **Molecular Biology of the Cell**, Bethesda, v. 10, n. 6, p. 1859-1872, 1999.
- KANEHISA, M.; GOTO, S. KEGG: Kyoto encyclopedia of genes and genomes. **Nucleic Acids Research**, London, v. 28, n. 1, p. 27-30, 2000.
- KAISERSLAUTERN TECHNISCHE UNIVERSITÄT. Disponível em: <<http://www.uni-kl.de/FB-Biologie/AG-Hahn/Research/rustfungi.html>>. Acesso em: 10 out. 2005.

- KENZELMANN, M.; MÜHLEMANN, K. Transcriptome analysis of fibroblast cells immediately after human cytomegalovirus infection. **Journal of Molecular Biology**, London, v. 304, n. 5, p.741-751, 2000.
- KIM, H.S.; DELANEY, T.P. *Arabidopsis* SON1 is an F-box protein that regulates a novel induced defense response independent of both salicylic acid and systemic acquired resistance. **The Plant Cell**, Baltimore, v. 14, n. 7, p. 1469-1482, 2002.
- KOBAYASHI, I.; et al. Dynamic reorganization of microtubules and microfilaments in flax cells during the resistance response to flax rust infection. **Planta**, Berlin, v. 195, n. 2, p. 237-247, 1994.
- KOBE, B.; KAJAVA, A.V. The leucine-rich repeat as a protein recognition motif. **Current Opinion in Structural Biology**, Philadelphia, v. 11, n. 6, p. 725-732, 2001.
- KOERKAMP, M.G.; REP, M.; BUSSEMAKER, H.J.; HARDY, G.P.M.A.; MUL, A.; PIEKARSKA, K.; SZIGYARTO, C.A.; de MATTOS, J.M.T.; TABAK, H.F. Dissection of transient oxidative stress response in *Saccharomyces cerevisiae* by using DNA microarrays. **Molecular Biology of the Cell**, Bethesda, v. 13, n. 8, p. 2783-2794, 2002.
- KUNKEL, B.N.; BROOKS, D.M. Cross talk between signaling pathways in pathogen defense. **Current Opinion in Plant Biology**, London, v. 5, n. 4, p. 325-331, 2002.
- KRUGNER, T.L.; AUER, C.G. Doenças de eucalipto. In: KIMATI, M.; AMORIM, L.; BERGAMIN FILHO, A.; CAMARGO, L.E.A.; REZENDE, J.A.M (Ed.). **Manual de Fitopatologia**. São Paulo: Agronômica Ceres, 1995-1997. chap. 33, p. 358-375. 2 v.
- KRUGNER. A natureza da doença. In: BERGAMIN, A. F.; KITAMI, H.; AMORIM, L. (Ed.) **Manual de Fitopatologia**. São Paulo: Agronômica Ceres, 1995. cap.3.
- LAM, E.; KATO, N.; LAWTON, M. Programmed Cell Death, Mitochondria and the Plant Hypersensitive Response. **Nature**, London, v. 411, n. 6839, p. 848-853, 2001.
- LAWRENCE, G.J.; FINNEGAN, E.J.; AYLIFFE, M.A.; ELLIS, J.G. The *L6* gene for flax rust resistance is related to the *Arabidopsis* bacterial resistance gene *RPS2* and the tobacco viral resistance gene. **The Plant Cell**, Baltimore, v. 7, n. 8, p. 1195-1206, 1995.
- LEE, J.Y.; LEE, D.H. Use of serial analysis of gene expression technology to reveal changes in gene expression in *Arabidopsis* pollen undergoing cold stress. **Plant Physiology**, Rockville, v.132, n. 2, p.517-529, 2003.
- LEITE, N.B. Avancos da silvicultura brasileira são significativos. **Visão Agrícola**, São Paulo, n. 4, p. 58-61, jul-dez, 2005.
- LOCKHART, D.J.; DONG, H.L.; BYRNE, M.C.; FOLLETTIE, M.T.; GALLO, M.V.; CHEE, M.S.; MITTMANN, M.; WANG, C.W.; KOBAYASHI, M.; HORTON, H.; BROWN, E.L. Expression Monitoring by Hybridization to High-Density Oligonucleotide Arrays. **Nature Biotechnology**, New York, v. 14, n. 13, p. 1675-1680, 1996.

- LORENZ, W.W.; DEAN, J.F. SAGE Profiling and demonstration of differential gene expression along the axial developmental gradient of lignifying xylem in loblolly pine (*Pinus taeda*). **Tree Physiology**, Victoria, v.22, n.5, p.301-10, 2002.
- MADDEN, S.L.; WANG, C.J.; LANDES, G. Serial analysis of gene expression: from gene discovery to target identification. **Drug Discovery Today**, Oxford, v. 5, n. 9, p. 415-425, 2000.
- MADEIRA, H.M.F. **Tutorial phred/ phrap/ consed**. São José dos Pinhais: I JOBIP – Jornada de Bioinformática do Paraná, 2004. 28p.
- MALDONADO, A.M.; DOERNER, P.; DIXON, R.A.; LAMB, C.J.; CAMERON, R.K. A putative lipid transfer protein involved in systemic resistance signalling in *Arabidopsis*. **Nature**, London, v. 419, n. 6905, p. 399-403, 2002.
- MAN, MZ; WANG, X.; WANG, Y. POWER_SAGE: comparing statistical tests for SAGE experiments. **Bioinformatics**, Oxford, v. 16, n. 11, p. 953-959, 2000.
- MARCHLER-BAUER, A.; ANDERSON, J.B.; DeWEESE-SCOTT, C.; FEDOROVA, N.D.; GEER, L.Y.; HE, S.; HURWITZ, D.I.; JACKSON, J.D.; JACOBS, A.R.; LANCZYCKI, C.J.; LIEBERT, C.A.; LIU, C.; MADEJ, T.; MARCHLER, G.H.; MAZUMDER, R.; NIKOLSKAYA, A.N.; PANCHENKO, A.R.; RAO, B.S.; SHOEMAKER, B.A.; SIMONYAN, V.; SONG, J.S.; THIESSEN, P.A.; VASUDEVAN, S.; WANG, Y.; YAMASHITA, R.A.; YIN, J.J.; BRYANT, S.H. CDD: a curated Entrez database of conserved domain alignments. **Nucleic Acids Research**, London, v. 31, n. 1, p. 383-387, 2003
- MARCHLER-BAUER, A.; ANDERSON, J.B.; CHERUKURI, P.F.; DeWEESE-SCOTT, C.; GEER, L.Y.; GWADZ, M.; HE, S.; HURWITZ, D.I.; JACKSON, J.D.; KE, Z.; LANCZYCKI, C.J.; LIEBERT, C.A.; LIU, C.; LU, F.; MARCHLER, G.H.; MULLOKANDOV, M.; SHOEMAKER, B.A.; SIMONYAN, V.; SONG, J.S.; THIESSEN, P.A.; YAMASHITA, R.A.; YIN, J.J.; ZHANG, D.; BRYANT, S.H. CDD: a Conserved Domain Database for protein classification. **Nucleic Acids Research**, London, v. 33, p. D192-D196, 2005.
- MATSUMURA, H.; NIRASAWA, S.; TERAUCHI, R. Transcript profiling in rice (*Oryza sativa* L.) seedlings using serial analysis of gene expression (SAGE). **The Plant Journal**, Oxford, v. 20, n. 6, p. 719-726, 1999.
- MATSUMURA, H.; NIRASAWA, S.; KIBA, A.; URASAKI, N.; SAITOH, H.; ITO, M.; KAWAI-YAMADA, M.; UCHIMIYA, H.; TERAUCHI, R. Overexpression of Bax Inhibitor Suppresses the Fungal Elicitor-Induced Cell Death in Rice (*Oryza Sativa* L.) Cells. **The Plant Journal**, Oxford, v. 33, n. 3, p. 425-434, 2003.
- MELHORAMENTO genético para a qualidade da madeira. **Revista da Madeira**, São Paulo, n. 59, p. 48-54, set. 2001.
- MENDGEN, K.; HAHN, M. Plant Infection and the Establishment of Fungal Biotrophy. **Trends in Plant Science**, Kidlington, v. 7, n. 8, p. 352-356, 2002.

- MURO-ABAD, J.I.; GOMES, E.A.; CANCIO, O.N.; ARAÚJO, E.F. Genetic analysis of *Eucalyptus urophylla* and *E. grandis* clones selected in commercial crops from the Brazilian Amazon by RAPD markers. **Silvae Genetica**, Frankfurt, v. 50, n. 5-6, p. 177-181, 2001.
- MYRTACEAE. Disponível em: <<http://www.arbolesornamentales.com/Myrtaceae.htm>>. Acesso em: 12 dez. 2005.
- NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION. Disponível em: <<http://www.ncbi.nlm.nih.gov/>>. Acesso em: 20 jan 2006.
- NURNBERGER, T.; BRUNNER, F. Innate Immunity in Plants and Animals: Emerging Parallels Between the Recognition of General Elicitors and Pathogen- Associated Molecular Patterns. **Current Opinion in Plant Biology**, London, v.5, n.4, p.318-324, 2002.
- O EUCALIPTO e suas origens. **Revista da Madeira**, São Paulo, n. 59, p. 10-12, set. 2001.
- PATINO, W.D.; MIAN, O.Y.; SHIZUKUDA, Y.; HWANG, P.M. Current and future applications of SAGE to cardiovascular medicine. **Trends in Cardiovascular Medicine**, London, v. 13, n. 4, p. 163-168, may 2003.
- PIETERSE, C.M.J.; VAN WEES, S.C.M.; HOFFLAND, E.; VAN PELT, J.A.; VAN LOON, L.C. Systemic resistance in *Arabidopsis* induced by biocontrol bacteria is independent of salicylic acid accumulation and pathogenesis-related gene expression. **The Plant Cell**, Baltimore, v. 8, n. 8, p. 1225-1237, 1996.
- PIETERSE, C.M.J.; VAN WESS, S.C.M.; VAN PELT, J.A.; KNOESTER, M.; LAAN, R.; GERRITS, H.; WEISBEEK, P.J.; VAN LOON, L.C. A novel signaling pathway controlling induced systemic resistance in *Arabidopsis*. **The Plant Cell**, Baltimore, v. 10, n. 9, p. 1571-1580, 1998.
- POLLACK, J.R.; PEROU, C.M.; ALIZADEH, A.A.; EISEN, M.B.; PERGAMENSCHIKOV, A.; WILLIAMS, C.F.; JEFFREY, S.S.; BOTSTEIN, D.; BROWN, P.O. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. **Nature Genetics**, New York, v. 23, n. 1, p. 41-46, 1999.
- PRINCIPAIS doenças na cultura do *Eucalyptus*. **Informações Agronômicas**, Piracicaba, n. 93, p. 26-32, Março 2001.
- RAYACHHETRY, M.B.; VAN, T.K.; CENTER, T.D.; ELLIOTT, M.L. Host Range of Puccinia Psidii, a Potential Biological Control Agent of Melaleuca Quinquenervia in Florida. **Biological Control**, Orlando, v. 22, n. 1, p. 38-45, 2001.
- RESENDDE, M.L.V.; SALGADO, S.M.L.; CHAVES, Z.M. Espécies ativas de oxigênio na resposta de defesa de plantas a patógenos. **Fitopatologia brasileira**, Brasília, v. 28, n. 2, p. 123-130, 2003.
- REYMOND, P. DNA microarrays and plant defence. **Plant Physiology and Biochemistry**, Paris, v. 39, n. 3-4, p. 313-321, 2001.

- RICHMOND, S.C.; GLASNER, J.D.; MAU, R.; JIN, H.; BLATTNER, F.R. Genome-wide expression profiling in *Escherichia coli* K-12. **Nucleic Acids Research**, London, v. 27, n. 19, p. 3821-3835, 1999.
- ROBINSON, S.J.; CRAM, D.J.; LEWIS, C.T.; PARKIN, I.A.P. Maximizing the efficacy of SAGE analysis identifies novel transcripts in *Arabidopsis*. **Plant Physiology**, Rockville, v. 136, n. 2, p. 3223-3233, October 2004.
- ROUHIER, N.; GELHAYE, E.; GUALBERTO, J.M.; JORDY, M.N.; DE FAY, E.; HIRASAWA, M.; DUPLESSIS, S.; LEMAIRE, S.D.; FREY, P.; MARTIN, F.; MANIERI, W.; KNAFF, D.B.; JACQUOT, J.P. Poplar peroxiredoxin Q. A thioredoxin-linked chloroplast antioxidant functional in pathogen defense. **Plant physiology**, Rockville, v. 134, n. 3, p. 1027-1038, 2004.
- RUAN, Y.; GILMORE, J.; CONNER, T. Towards *Arabidopsis* genome analysis: monitoring expression profiles of 1400 genes using cDNA microarrays. **The Plant Journal**, Oxford, v. 15, n. 6, p. 821-833, 1998.
- RUIJTER, J.M.; VAN KAMPEN, A.H.C.; BAAS, F. Statistical evaluation of SAGE libraries: consequences for experimental design. **Physiological Genomics**, Bethesda, v. 11, n. 2, p. 37-44, 2002.
- SALZMAN, R.A.; FUJITA, T.; ZHU-SALZMAN, K.; HASEGAWA, P.M.; BRESSAN, R.A. An Improved Rna Isolation Method for Plant Tissues Containing High Levels of Phenolic Compounds or Carbohydrates. **Plant Molecular Biology Reporter**, Piscataway, v. 17, n. 1, p. 11-17, 1999.
- SAMBROOK, J.; FRITSCH, E.F.; MANIATIS, T. Preparation of reagents and buffers used in molecular cloning. In: _____. **Molecular Cloning: a laboratory manual**. New York: Cold Spring Harbor Laboratory Press, 1989a. chap. 6, p. 6.3-6.62.
- SAMBROOK, J.; FRITSCH, E.F.; MANIATIS, T. Preparation of reagents and buffers used in molecular cloning. In: _____. **Molecular Cloning: a laboratory manual**. New York: Cold Spring Harbor Laboratory Press, 1989b. chap. Appendix E, p. E.1-E.39.
- SANGER, F.; AIR, G.M.; BARRELL, B.G.; BROWN, N.L.; COULSON, A.R.; FIDDES, J.C.; HUTCHISON, C.A.; SLOCOMBE, P.M.; SMITH, M. Nucleotide-sequence of bacteriophage Φ X 174 DNA. **Nature**, London, v. 265, n. 5596, p. 687-695, 1977.
- SANTOS, P.E.T.; GARCIA, C.H.; MORI, E.S.; MORAES, M.L.T. Potencial para programas de melhoramento, estimativas de parâmetros genéticos e interação progênies x locais em populações de *Eucalyptus urophylla* S. T. Blake. **IPEF**, Piracicaba, n. 43/44, p. 1119, 1990.
- SCHENA, M.; SHALON, D.; DAVIS, R.W.; BROWN, P.O. Quantitative Monitoring of Gene-Expression Patterns With a Complementary-Dna Microarray. **Science**, Washington, v. 270, n. 5235, p. 467-470, 1995.

- SCHENA, M.; SHALON, D.; HELLER, R.; CHAI, A.; BROWN, P.O.; DAVIS, R.W. Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 93, n. 20, p. 10614-9, 1996.
- SCHENK, P.M.; KAZAN, K.; WILSON, I.; ANDERSON, J.P.; RICHMOND, T.; SOMERVILLE, S.C.; MANNERS, J.M. Coordinated Plant Defense Responses in Arabidopsis Revealed by Microarray Analysis. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 97, n. 21, p. 11655-11660, 2000.
- SCHLÜTER, K. et al. Profilins as regulators of actin dynamics. **Biochimica et Biophysica Acta- Gene Structure And Expression**, Amsterdam, v. 1359, n. 2, p. 97-109, 1997.
- SCHMELZER, E. Cell polarization, a crucial process in fungal defence. **Trends in Plant Science**, Kidlington, v. 7, n. 9, p. 411-415, 2002.
- SCHULTZ, J.; MILPETZ, F.; BORK, P.; PONTING, C.P. SMART, a simple modular architecture research tool: identification of signaling domains. **Proceedings of the National Academy of Sciences of the United States of América**, Washington, v. 95, n. 11, p. 5857-5864, may 1998.
- SHAH, J. The salicylic acid loop in plant defense. **Current Opinion in Plant Biology**, London, v.6, n. 4, p.365-371, 2003.
- SOCIEDADE DE SILVICULTURA BRASILEIRA. Disponível em: <<http://www.sbs.org.br>>. Acesso em: 12 dezembro 2005.
- SONNHAMMER, E.L.L.; EDDY, S.R.; BIRNEY, E.; BATEMAN, A.; DURBIN, R. Pfam: multiple sequence alignments and HMM-profiles of proteins domains. **Nucleic Acids Research**, London, v. 26, n. 1, p.320-322, 1998.
- STASKAWICZ, B.J.; DAHLBECK, D.; KEEN, N.T. Cloned Avirulence Gene of Pseudomonas-Syringae Pv Glycinea Determines Race-Specific Incompatibility on Glycine-Max (L) Merr. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, v. 81, n. 19, p. 6024-6028, 1984.
- STASWICK, P.E.; YUEN, G.Y.; LEHMAN, C.C. Jasmonate Signaling Mutants of Arabidopsis Are Susceptible to the Soil Fungus Pythium Irregulare. **The Plant Journal**, Oxford, v.15, n. 6, p.747-754, 1998.
- STEFFENSON, B.J. Analysis of Durable Resistance to Stem Rust in Barley. **Euphytica**, Wageningen, v. 63, n. 1-2, p. 153-167, 1992.
- STOLLBERG, J.; URSCHITZ, J.; URBAN, Z.; BOYD, C.D. A quantitative evaluation of SAGE. **Genome Research**, Woodbury, v. 10, n. 8, p. 1241-1248, 2000.

- TAMELING, W.I.L.; ELZINGA, S.D.J.; DARMIN, P.S.; VOSSSEN, J.H.; TAKKEN, F.L.W.; 'HARING, M.A.; CORNELISSEN, V.J.C. The tomato *R* gene products I-2 and Mi-1 are functional ATP binding proteins with ATPase activity. **The Plant Cell**, Baltimore, v. 14, n. 11, p. 2929-2939, 2002.
- KRUGER; BACCHI, L. M. A. In: BERGAMIN, A. F.; KITAMI, H.; AMORIM, L. (Ed.) **Manual de Fitopatologia**. São Paulo: Agronômica Ceres, 1995. cap.4.
- TATUSOV, R.L.; KOONIN, E.V.; LIPMAN, D.J. A genomic perspective on protein families. **Science**, Washington, v. 278, n. 5338, p. 631-637, 1997.
- TELECHEA, N.; ROLFO, M.; COUTINHO, T.A.; WINGFIELD, M.J. *Puccinia psidii* on *Eucalyptus globulus* in Uruguay. **Plant Pathology**, London, v. 52, n. 3, p. 427, 2003.
- TESSMANN, D.J.; DIANESE, J.C.; MIRANDA, A.C.; CASTRO, L.H.R. Epidemiology of a Neotropical Rust (*Puccinia Psidii*): Periodical Analysis of the Temporal Progress in a Perennial Host (*Syzygium Jambos*). **Plant Pathology**, London, v.50, n.6, p.725-731, 2001.
- THE ARABIDOPSIS GENOME INIATIVE. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. **Nature**, London, v. 408, n. 6814, p. 796-815, 2000.
- THEIS, T.; STHAL, U. Antifungal proteins: targets, mechanisms and prospective applications. **CMLS Cellular and Molecular Life Sciences**, Basel, v. 61, n. 4, p.437-455, 2004
- THOMAS, S.W.; GLARING, M.A.; RASMUSSEN, S.W.; KINANE, J.T.; OLIVER, R.P. Transcript profiling in the barley mildew pathogen *Blumeria graminis* by serial analysis of gene expression (SAGE). **Molecular Plant-Microbe Interactions**, St Paul, v. 15, n. 8, p. 847-856, 2002.
- THORDAL-CHRISTENSEN, H. Fresh insights into processes of nonhost resistance. **Current Opinion in Plant Biology**, London, v. 6, n. 4, p.351-357, 2003.
- TORNERO, P.; MERRITT, P.; SADANANDOM, A.; SHIRASU, K.; INNES, R.W.; DANGL, J.L. Rar1 and Ndr1 Contribute Quantitatively to Disease Resistance in Arabidopsis, and Their Relative Contributions Are Dependent on the R Gene Assayed. **The Plant Cell**, Baltimore, v. 14, n. 5, p. 1005-1015, 2002.
- VALVERDE, S.R. **As plantações de eucalipto no Brasil**. Disponível em: http://www.sbs.org.br/destaque_plantacoesnobrasil.htm> Acesso em: 21 dez. 2005.
- VAN DER BIEZEN, E.A.; FREDDIE, C.T.; KAHN, K.; PARKER, J.E.; JONES, J.D.G. Arabidopsis Rpp4 Is a Member of the Rpp5 Multigene Family of Tir-Nb-Lrr Genes and Confers Downy Mildew Resistance Through Multiple Signalling Components. **The Plant Journal**, Oxford, v.29, n. 4, p. 439-451, 2002.
- VAN ETEN, H.D.; MATTHEWS, D.E.; MATTHEWS, P.S. Phytoalexin detoxification: importance for pathogenicity and practical implications. **Annual Review of Phytopathology**, Colorado, v. 27, p. 143-164, 1989.

- VELCULESCU, V.E.; ZHANG, L.; VOGELSTEIN, B.; KINZLER, K.W. Serial Analysis of Gene-Expression. **Science**, Washington, v.270, n.5235, p.484-487, 1995.
- VELCULESCU, V.E.; VOGELSTEIN, B.; KINZLER, K.W. Analysing Uncharted Transcriptomes With Sage. **Trends in Genetics**, Cambridge, v.16, n.10, p.423-425, 2000.
- WEERARATNA, A.T. Serial analysis of gene expression (SAGE): Advances, analysis and applications to pigment cell research. **Pigment Cell Research**, Copenhagen, v. 16, n. 3, p. 183-189, 2003.
- WHITE, K.P.; RIFKIN, S.A.; HURBAN, P.; HOGNESS, D.S. Microarray analysis of *Drosophila* development during metamorphosis. **Science**, Washington, v. 286, n. 5447, p. 2179-2184, 1999.
- WIKIPEDIA. **Florestas**. Disponível em: <<http://pt.wikipedia.org/wiki/Floresta>>. Acesso em: 17 nov. 2005.
- WU, Q.; VAN ETEN, H.D. Introduction of plant and fungal genes into pea (*Pisum sativum* L.) hairy roots reduces their ability to produce pisatin and affects their response to a fungal pathogen. **Molecular Plant Microbe Interact**, St Paul, v. 17, n. 7, p. 798-804, 2004.
- YAMAMOTO, M.; WADATSUKI, T.; HADA, A.; RYO, A. Use of serial analysis of gene expression (SAGE) technology. **Journal of Immunological Methods**, Amsterdam, v. 250, n. 1-2, p. 45-66, 2001.
- YUE, H.; EASTMAN, P.S.; WANG, B.B.; MINOR, J.; DOCTOLERO, M.H.; NUTTALL, R.L.; STACK, R.; BECKER, J.W.; MONTGOMERY, J.R. VAINER, M.; JOHNSTON, R. An evaluation of the performance of cDNA microarrays for detecting changes in global mRNA expression. **Nucleic Acids Research**, London, v. 29, n. 8, p. e41-e50, 2001.
- ZHANG, Y.; DOREY, S.; SWIDERSKY, M.; JONES, J.D.G. Expression of *RPS4* in tobacco induces an AvrRps4-independent HR that requires EDS1, SGT1 and HSP90. **The Plant Journal**, Oxford, v. 40, n. 2, p. 213-224, 2004.
- ZENG, Y.; YANG, T. RNA isolation from highly viscous samples rich in polyphenols and polysaccharides. **Plant Molecular Biology Reporter**, Canada, v. 20, p. 417a-417c, 2002.
- ZIAUDDIN, J.; SABATINI, D.M. Microarrays of cells expressing defined cDNAs. **Nature**, London, v. 411, n. 68433, p. 107-110, 2001.

ANEXOS

ANEXO A

Tampão de extração de RNA (SALZMAN et al., 1999)

Guanidina tiocianato	4 M
Tris-HCl, pH 8	100 mM
Citrato de sódio, pH 8	25 mM
N-lauril sarcosina	0,5%

ANEXO B

Tampão de extração (ZENG; YANG, 2002)

CTAB	2%
PVP – polivinilpirrolidona	2% (peso molecular 25)
Tris-HCl, pH 8	100 mM
EDTA	25 mM
NaCl	2 M
Espermidina trihidrocloridra	0,05%
β -mercaptoetanol	2%

ANEXO C

Mix de RNA aplicado no gel

MOPS 10X	1,5 μ l
Formaldeído 37%	3 μ l
Formamida deionizada	7,5 μ l
Tampão Dye III	2 μ l
Brometo de Etídeo	1 μ l
RNA	5 μ g

ANEXO D

Tampão TAE

Tris-acetato	0,04 M
EDTA	0,001 M

ANEXO E

<i>Tampão de ligação (Binding buffer) (Dynabeads[®])</i>	
Tris-HCl, pH 7,5	20 mM
LiCl	1 M
EDTA	2 mM
Duodecil-sulfato de lítio	1%
DTT – DI-TIOTREITOL (EM ÁGUA DEPC)	5mM

ANEXO F

<i>Tampão e lavagem B (Washing Buffer B) (Dynabeads[®])</i>	
Tris-HCl, pH 7,5	10 mM
LiCl	0,15 M
EDTA	1 mM

ANEXO G

2X reaction *mix* PCR-RT: Tampão contendo 0,4 mM de cada dNTP e 24 mM de SO₄Mg.

ANEXO H

<i>Tampão de ligação e lise (Lysis/binding buffer) (I-SAGETM)</i>	
Tris-HCl, pH 7,5	100 mM
LiCl	500 mM
EDTA	10 mM
Duodecil sulfato de lítio	1%
DTT – DI-TIOTREITOL (EM ÁGUA DEPC)	5 mM

ANEXO I

<i>Tampão de lavagem A (Wash buffer A) (I-SAGETM)</i>	
Tris-HCl, pH 7,5	10 mM
LiCl	0,15 M
EDTA	1 mM
Duodecil-sulfato de lítio	0,1%
glicogênio	10 µl/ml

ANEXO J

<i>Tampão de lavagem B (Washing buffer B) (I-SAGETM)</i>	
Tris-HCl, pH 7,5	10 mM
LiCl	150 mM
EDTA	1 mM
glicogênio	10 µl/ml

ANEXO K

<i>Tampão de primeira fita 1X (1X First Strand Buffer) (I-SAGETM)</i>	
Tris-HCl, pH 8,3	50 mM
KCl	75 mM
MgCl ₂	3 mM
<i>Mussel glycogen</i>	10 µl/ml

ANEXO L

<i>Solução para síntese da primeira fita de cDNA (First Strand cDNA Synthesis) (I-SAGETM)</i>	
5X Tampão da primeira fita (5X <i>First strand buffer</i>) (ANEXO M)	18 µl
RNaseOUT TM (ANEXO N)	1 µl
Água DEPC	54,5 µl
0,1 M DTT – di-tiotreitol (em água DEPC)	9 µl
dNTP <i>mix</i> (10 mM de cada nucleotídeo)	4,5 µl

ANEXO M

<i>Tampão da primeira fita 5X (1X First Strand Buffer) (I-SAGETM)</i>	
Tris-HCl, pH 8,3	250 mM
KCl	375 mM
MgCl ₂	15 mM

ANEXO N

<i>RNaseOUTTM (I-SAGETM)</i>	
RNaseOUT	40 U/μl
Tris-HCl, pH 8	20 mM
KCl	50 mM
EDTA	0,5 mM
DTT – DI-TIOTREITOL (EM ÁGUA DEPC)	8 mM
Glicerol	50% (peso/volume)

ANEXO O

<i>Transcritase reversa SuperScriptTM II</i>	
Transcritase reversa <i>SuperScriptTM II</i>	200 U/μl
Tris-HCl, pH 7,5	20 mM
NaCl	100 mM
EDTA	0,1 mM
DTT – DI-TIOTREITOL (EM ÁGUA DEPC)	1 mM
Nonidet-40	0,01% (volume/volume)
Glicerol	50% (peso/volume)

ANEXO P

Tampão da a segunda fita 5X (5X Second strand buffer) (I-SAGETM)

Tris-HCl, pH 6,9	100 mM
KCl	450 mM
MgCl ₂	23 mM
B-NAD ⁺	0,075 mM
Sulfato de amônio	50 mM

ANEXO Q

dNTP mix (I-SAGETM)

dATP	10 mM
dGTP	10 mM
dCTP	10 mM
dTTP	10 mM
Água DEPC	

ANEXO R

Ligase de DNA de Escherichia coli (I-SAGETM)

Ligase de DNA de <i>Escherichia coli</i>	10 U/μl
Tris-HCl, pH 7,4	10 mM
KCl	50 mM
EDTA	0,1 mM
DTT – DI-TIOTREITOL (EM ÁGUA DEPC)	1 mM
BSA	200 μg/ml
Glicerol	50% (peso/volume)
Triton [®] X-100	0,1%

ANEXO S

<i>DNA polimerase de Escherichia coli (I-SAGETM)</i>	
DNA polimerase de <i>Escherichia coli</i>	10 U/μl
Potasio fosfato, pH 7,4	50 mM
KCl	100 mM
DTT – DI-TIOTREITOL (EM ÁGUA DEPC)	1 mM
Glicerol	50% (peso/volume)

ANEXO T

<i>RNase H de Escherichia coli (I-SAGETM)</i>	
RNase H de <i>Escherichia coli</i>	2 U/μl
Tris-HCl, pH 7,5	20 mM
KCl	100 mM
MgCl ₂	10 mM
EDTA	0,1 mM
DTT – DI-TIOTREITOL (EM ÁGUA DEPC)	0,1 mM
BSA	50 μg/ml
Glicerol	50% (peso/volume)

ANEXO U

<i>Tampão de lavagem C (Wash Buffer C) (I-SAGETM)</i>	
Tris-HCl, pH 7,5	5 mM
EDTA	0,5 mM
NaCl	1 M
SDS – duodecil sulfato sódico	1%
<i>Mussel glycogen</i>	10 μl/ml

ANEXO V

Tampão de lavagem D (Wash Buffer D) (I-SAGETM)

Tris-HCl, pH 7,5	5 mM
EDTA	0,5 mM
NaCl	1 M
BSA – albumina de soro bovino	200 µl/ml

ANEXO W

1X tampão 4 (I-SAGETM)

Tris-acetato, pH 7,9	20 mM
Acetato de magnésio	10 mM
Acetato de potássio	50 mM
DTT – DI-TIOTREITOL (EM ÁGUA DEPC)	1 mM
BSA	200 µg/ml

ANEXO X

LoTE (I-SAGETM)

Tris-HCl, pH 7,5	3 mM
EDTA, pH 7,5	0,2 mM

ANEXO Y

100X BSA (I-SAGETM)

BSA	10 mg/ml
Fosfato potássio, pH 7	20 mM
NaCl	50 mM
EDTA	0,1 mM
Glicerol	5% (p/v)

ANEXO Z

<i>10X tampão 4 (I-SAGETM)</i>	
Tris-acetato, pH 7,9	200 mM
Acetato de magnésio	100 mM
Acetato de potásio	500 mM
DTT – DI-TIOTREITOL (EM ÁGUA DEPC)	10 mM

ANEXO AA

<i>Nla III (I-SAGETM)</i>	
<i>Nla</i> III	10 U/μl
KCl	200 mM
Tris-HCl, pH 7,5	10 mM
EDTA	0,1 mM
DTT – DI-TIOTREITOL (EM ÁGUA DEPC)	1 mM
BSA (ANEXO Y)	500 μg/ml
Glicerol	50% (peso/volume)

ANEXO AB

<i>Tampão de lavagem C (Wash buffer C) (I-SAGETM)</i>	
Tris, pH 7,5	5 mM
EDTA	0,5 mM
NaCl	1 M
SDS – duodecil sulfato de sódio	1% (volume/volume)
<i>Mussel glycogen</i>	10 μg/ml

ANEXO AC

<i>Tampão de lavagem D (Wash buffer D) (I-SAGETM)</i>	
Tris, pH 7,5	5 mM
EDTA	0,5 mM
NaCl	1 M

BSA (ANEXO Y)	200 µg/ml
ANEXO AD	
<i>1X tampão para ligase (I-SAGETM)</i>	
10X tampão para ligase (ANEXO AE)	Diluído para 1X em água DEPC
ANEXO AE	
<i>10X tampão para ligase (I-SAGETM)</i>	
Tris-HCl, pH 7,5	60 mM
MgCl ₂	60 mM
NaCl	50 mM
BSA	1 mg/ml
β-mercaptoetanol	70 mM
ATP	1 mM
DTT – DI-TIOTREITOL (EM ÁGUA DEPC)	20 mM
Espermidina	10 mM
ANEXO AF	
<i>Ligase de DNA T4 (T4 DNA ligase)</i>	
Ligase de DNA T4	4 unidades de Weiss/µl
Tris-HCl, pH 7,5	10 mM
Cloreto de potássio	50 mM
DTT – DI-TIOTREITOL (EM ÁGUA DEPC)	1 mM
Glicerol	50% (peso/volume)
ANEXO AG	
<i>1X tampão 4 (1X buffer 4)</i>	
10X tampão 4 (ANEXO Z)	Diluído para 1X em água DEPC

ANEXO AH

<i>BsmF I</i>	
<i>BsmF I</i>	2 unidades/ml
NaCl	50 mM
Tris-HCl, pH 7,4	10 mM
EDTA	0,1 mM
DTT – DI-TIOTREITOL (EM ÁGUA DEPC)	1 mM
BSA	200 µg/ml
Glicerol	50% (peso/volume)

ANEXO AI

<i>10X Tampão do fragmento Klenow</i>	
Tris-HCl, pH 8	500 mM
MgCl ₂	100 mM
NaCl	500 mM

ANEXO AJ

<i>Polimerase Klenow</i>	
Polimerase Klenow	3-9 U/µl
Fosfato de potásio, pH 7	50 mM
KCl	100 mM
DTT – DI-TIOTREITOL (EM ÁGUA DEPC)	1 mM
Glicerol	50% (peso/volume)

ANEXO AK

<i>2X solução de reação de dímeros de etiquetas (Ditag reaction mix)</i>		
Tris-HCl, pH 7,5	3 mM	1,25 µl
10X Tampão da ligase (ANEXO AL)		0,75 µl
Água DEPC		0,75 µl
Ligase de DNA T4 (ANEXO AF)		1 µl

ANEXO AL

10X tampão da Ligase (10X Ligase buffer)

Tris-HCl, pH 7,5	60 mM
MgCl ₂	60 mM
NaCl	50 mM
BSA (ANEXO Y)	1 mg/ml
β-mercaptoetanol	70 mM
ATP	1 mM
DTT - DI-TIOTREITOL (EM ÁGUA DEPC)	20 mM
Espermidina	10 mM

ANEXO AM

2X da solução do controle negativo (Negative Control Mix)

Tris-HCl, pH 7,5	3 mM	2,25 µl
10X Tampão da ligase (ANEXO AL)		0,75 µl
Água DEPC		0,75 µl
Ligase de DNA T4 (ANEXO AF)		

ANEXO AN

10X tampão BV

Sulfato de amônio	166 mM
Tris-HCl, pH 8,8	670 mM
MgCl ₂	67 mM
β-mercaptoetanol	100 mM

ANEXO AO

Gel de poliacrilamida 12%

40% de acrilamida/bisacrilamida	12,8 ml
Tampão de corrida 5X TBE (ANEXO AP)	8 ml
TEMED	16 µl
Água deionizada	18,8 ml

Adicionou-se 0,4 ml de persulfato de amônio 10%. Se misturou tudo e depositou na plataforma para confecção do gel com ajuda de uma seringa. Inseriu-se o pente na parte superior e foi deixado polimerizando 60 minutos a temperatura ambiente. Preparou-se 1X tampão de corrida TBE (ANEXO AP). Preparouse as amostras com 5X tampão de amostra (ANEXO AQ) e se carregaram as amostras no gel.

ANEXO AP

Tampão de corrida 5X TBE

Tris	89 mM
Ácido bórico	89 mM
EDTA	2 mM
pH	8,3

Para 1X TBE, diluiu-se com água deionizada.

ANEXO AQ

Tampão de amostra 5X TBE

Tris base	18 mM
Ácido bórico	18 mM
EDTA (livre de ácidos)	0,4 mM
Ficoll™ tipo 400	3%
Azul de bromofenol	0,02%
Xileno cianol	0,02%

ANEXO AR

Solução para amplificação por PCR (X 200 reações)

Tampão 10X BV	5 µl
DMSO	3 µl
dNTP (10 mM de cada um) (ANEXO Q)	7,5 µl
DTP-1 (<i>ditag pirmer-1</i>)	2 µl
DTP-2 (<i>ditag primer-2</i>)	2 µl
Água DEPC	24 µl
DNA polimerase Platinum [®] <i>Taq</i>	0,5 µl

ANEXO AS

Mix para a clivagem com Nla III

Dímero de etiquetas (100 pb)	42 µl
10X Tampão 4 (ANEXO Z)	15 µl
100X BSA (ANEXO Y)	2 µl
<i>Nla</i> III (ANEXO AA)	12 µl
Água DEPC	79 µl

ANEXO AT

Gel de poliacrilamida 8%

40% de acrilamida/bisacrilamida	8,2 ml
Tampão de corrida 5X TBE (ANEXO AP)	8 ml
TEMED	16 µl
Água deionizada	22,8 ml

Adicionou-se 0,4 ml de persulfato de amônio 10%. Se misturou tudo e depositou na plataforma para confecção do gel com ajuda de uma seringa. Inseriu-se o pente na parte superior e foi deixado polimerizando 60 minutos a temperatura ambiente. Preparou-se 1X tampão de corrida TBE (ANEXO AP). Preparouse as amostras com 5X tampão de amostra (ANEXO AQ) e se carregaram as amostras no gel.

ANEXO AU

<i>pZErO[®]-1</i>	
pZErO [®] -1	1 µg/µl em tampão TE (ANEXO AV), pH 7,5

ANEXO AV

<i>Tampão TE</i>	
Tris-HCl, pH 7,5	10 mM
EDTA	1 mM

ANEXO AW

<i>nuclease Sph I</i>	
<i>Sph I</i>	5 U/µl
NaCl	100 mM
Tris-HCl, pH 7,4	10 mM
EDTA	0,1 mM
DTT - DI-TIOTREITOL (EM ÁGUA DEPC)	1 mM
BSA (ANEXO Y)	100 µg/ml
Glicerol	50% (peso/volume)

ANEXO AX

<i>10X tampão 2</i>	
Tris-HCl, pH 7,9	100 mM
MgCl ₂	10 mM
NaCl	50 mM
DTT - DI-TIOTREITOL (EM ÁGUA DEPC)	1 mM

ANEXO AY

MEIO SOC

Bacto triptona	2% (p/v)
Bacto extrato de levedura	0,5% (p/v)
NaCl	10 mM
KCl	2,5 mM
Mg Cl ₂ [*]	10 mM
MgSO ₄ [*]	10 mM
Glicose [*]	20 mM
Água Milli Q	q.s.q 1 l
pH	7,0
Esterilização por autoclavagem	

^{*}Preparados à parte, esterilizados por filtração (Millipore 0,2 µm) e adicionados após autoclavagem.