



**UNIVERSIDADE DE BRASÍLIA**  
**INSTITUTO DE CIÊNCIAS BIOLÓGICAS**  
**DEPARTAMENTO DE BIOLOGIA CELULAR**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA MOLECULAR**

**ANÁLISE TRANSCRITÔMICA E PROTEÔMICA DA INTERAÇÃO *MUSA ACUMINATA* X  
*MYCOSPHAERELLA MUSICOLA***

**MARCO AURÉLIO NINÔMIA PASSOS**

**Brasília, Junho de 2014.**

**MARCO AURÉLIO NINÔMIA PASSOS**

**ANÁLISE TRANSCRITÔMICA E PROTEÔMICA DA INTERAÇÃO *MUSA ACUMINATA* X  
*MYCOSPHAERELLA MUSICOLA***

Tese apresentada ao Programa de Pós  
Graduação em Biologia Molecular da  
Universidade de Brasília como parte dos  
requisitos necessários para a obtenção do título  
de Doutor em Biologia Molecular

**Orientador**

Prof. Dr. Robert Neil Gerard Miller

**Brasília, Junho de 2014.**

**MARCO AURÉLIO NINÔMIA PASSOS**

**ANÁLISE TRANSCRITÔMICA E PROTEÔMICA DA INTERAÇÃO *MUSA ACUMINATA* X  
*MYCOSPHAERELLA MUSICOLA***

**Brasília, Junho de 2014.**

## AGRADECIMENTOS

Aos meus pais, Quiyomí Ninômia e Wilson de Andrade Passos, por serem o motivo dessa caminhada;

À minha irmã Thaís pelo grande apoio prestado, e pela torcida ao longo desses anos;

À Irene, minha segunda mãe, por todo apoio e dedicação;

À todos os amigos que sempre torceram e incentivaram para que eu chegasse a essa conquista;

Ao meu orientador, Dr. Robert Neil Gerard Miller, por todos os anos de aprendizado, e por todo apoio, respeito e confiança;

Aos pesquisadores Georgios Pappas, Roberto Togawa, Orzenil e Natália Martins por toda paciência, auxílio e ensinamentos, principalmente com as análises de bioinformática;

Ao grupo de Proteômica da Universidade Católica de Brasília, em especial aos Professores Octávio Franco e Simoni Campos, e aos colegas Nelson Júnior e Elizabeth Candido, por toda ajuda e aprendizado;

Um agradecimento especial à minha grande amiga Flávia Leonel, por estar sempre ao meu lado ao longo desses anos de estudo; agradeço pelo convívio, brigas, distrações, conselhos, incentivos, amizade e lealdade;

E por fim, à todos que contribuíram de alguma forma para a execução desta pesquisa.



## **DEDICATÓRIA**

*“A minha família, alicerce e base da minha força”*

## SUMÁRIO

LISTA DE TABELAS.....	9
LISTA DE FIGURAS .....	10
LISTA DE ABREVIATURAS .....	11
RESUMO .....	13
ABSTRACT .....	14
1. REVISÃO DE LITERATURA.....	15
1.1. Taxonomia de <i>Musa</i> .....	15
1.2. Importância social e econômica da banana.....	15
1.3. <i>Mycosphaerella musicola</i> .....	16
1.4. Melhoramento genético em <i>Musa</i> .....	21
1.5. Imunidade vegetal (ETI e PTI) .....	24
1.5.1. Imunidade disparada por PAMPs.....	24
1.5.2. Imunidade disparada por efetores .....	25
1.5.3. Sobreposições entre PTI e ETI.....	28
1.5.4. Mecanismos de sinalização.....	29
1.6. Tecnologias de sequenciamento de alto desempenho - NGS .....	31
1.6.1. Plataforma 454 .....	31
1.6.2. Plataforma Solexa .....	33
1.6.3. Aplicações.....	34
1.7. Proteômica.....	36
1.8. Transcritômica x Proteômica .....	37
2. JUSTIFICATIVA .....	40
3. OBJETIVOS .....	41
3.1. Objetivo geral: .....	41
3.2. Objetivos específicos:.....	41
4. METODOLOGIA.....	42
4.1. Coleção de isolados de <i>M. musicola</i> , identificação, crescimento e manutenção 42	
4.2. Bioensaio .....	45

4.3. Extração do RNA total de <i>M. acuminata</i> Calcutta 4 e <i>M. acuminata</i> Cavendish Grande Naine inoculados e não inoculados com <i>M. musicola</i> .....	45
4.4. Sequenciamento massal via 454 .....	46
4.5. Processamento de sequência e anotação .....	47
4.5.1. Pré-processamento .....	47
4.5.2. Montagem das sequências .....	47
4.5.3. Anotação das sequências.....	48
4.6. Análise Proteômica .....	48
4.6.1. Extração e quantificação de proteínas.....	48
4.6.2. Eletroforese bi-dimensional em gel de poliacrilamida (2DE-PAGE) .....	50
4.6.2.1. Focalização Isoelétrica (IEF).....	50
4.6.2.2. Separação em segunda dimensão: SDS-PAGE .....	50
4.6.2.3. Análise de imagem dos géis bidimensionais.....	51
5. RESULTADOS E DISCUSSÃO.....	52
5.1. Análise do transcritoma.....	52
5.1.1. Estatística e Montagem - 454.....	52
5.1.2. Anotação funcional e classificação.....	56
5.1.3. Distribuição de espécies.....	57
5.1.4. GO categorias.....	58
5.1.5. Expressão gênica em Calcutta 4 .....	63
5.1.6. Expressão gênica em Cavendish Grande Naine .....	65
5.1.7. Receptores de genes do hospedeiro .....	65
5.1.8. GO categorias.....	66
5.1.9. Mineração de dados em relação a modelos de genes de <i>M. fijiensis</i> .....	66
5.2. Análise proteômica da interação Banana x <i>M. musicola</i> .....	67
6. CONCLUSÕES .....	81
7. REFERÊNCIAS BIBLIOGRÁFICAS.....	82
8. ANEXOS.....	96

## LISTA DE TABELAS

Tabela 1 - Características da abordagem transcritômica e proteômica.....	38
Tabela 2 - Distribuição por tamanho de sequências de alta qualidade de <i>M. acuminata</i> Calcutta 4 e Cavendish Grande Naine, derivadas do sequenciamento massal.....	53

## LISTA DE FIGURAS

Figura 1 - Ciclo de <i>Mycosphaerella</i> spp em bananeira.....	17
Figura 2 - Escala de Meredith (1970), para diferentes graus de desenvolvimento das lesões de Sigatoka-amarela.....	18
Figura 3 – Folha de bananeira com sintomas característicos da Sigatoka Amarela .....	20
Figura 4 - Modelos de reconhecimento planta-patógeno .....	27
Figura 5 - Modelo “ziguezague” ilustrando as etapas do desenvolvimento do sistema imunológico da planta durante infecção.....	28
Figura 6 - Diagrama do processo do pirosequenciamento.....	32
Figura 7 - Visão geral da tecnologia de sequenciamento de 454 .....	32
Figura 8 - Câmara úmida de folhas de bananeira para emissão de esporódoquios.....	42
Figura 9 - Etapas do processo de multiplicação de <i>Mycosphaerella musicola</i> .....	43
Figura 10 - Sequência para liberação de conídios de <i>Mycosphaerella musicola</i> em cultura ...	44
Figura 11 - Distribuição conforme o tamanho dos <i>reads</i> de <i>M. acuminata</i> Calcutta 4 e Cavendish Grande Naine.....	54
Figura 12 - Resumo da montagem das sequências de <i>M. acuminata</i> Calcutta 4 e Cavendish Grande Naine.....	54
Figura 13 - Diagrama de Venn mostrando a sobreposição entre os conjuntos de dados do transcrito (contigs e singletons) para os genótipos de <i>M. acuminata</i> Calcutta 4 Cavendish e Grande Naine.....	56
Figura 14 - Exemplo de um contig de <i>M. acuminata</i> Calcutta 4 mapeando em um único modelo gênico de <i>M. acuminata</i> DH Pahang.....	56
Figura 15 - Distribuição das espécies dos unigenes de <i>M. acuminata</i> mostrados em porcentagem do total das sequências homólogas de plantas monocotiledôneas.....	58
Figura 16 - Apresentação do histograma do <i>Gene Ontology</i> (GO) após o processo de anotação dos unigenes de <i>M. acuminata</i> Calcutta 4 e Cavendish Grande Naine.....	60
Figura 17 - Representação dos unigenes de <i>M. acuminata</i> classificados em grupos funcionais..	62
Figura 18 – Mini-gel de poliacrilamida SDS-PAGE 12% mostrando o padrão proteico de duas concentrações do extrato protéico de <i>M. acuminata</i> .....	68
Figura 19 - Géis bidimensionais que foram analisados utilizando o software ImageMaster Platinum 7.0 .....	72
Figura 20 - Representação gráfica dos coeficientes de correlação obtidos entre as réplicas ...	73
Figura 21 – Mapas bidimensionais de folha de <i>Musa acuminata</i> utilizados na identificação de <i>spots</i> presentes no cultivar resistente Calcutta 4 (controle e tratado) e no cultivar suscetível Cavendish (controle e tratado).....	75
Figura 22 – Mapas bidimensionais de folha de <i>Musa acuminata</i> do cultivar resistente Calcutta 4.....	76
Figura 23 - Mapas bidimensionais de folha de <i>Musa acuminata</i> do cultivar suscetível Cavendish.....	77
Figura 24 – Mapas bidimensionais de folha de <i>Musa acuminata</i> indicando os <i>spots</i> em comum que apresentaram padrões de expressão para em cada tratamento em comparação ao controle.....	78

## LISTA DE ABREVIATURAS

°C – grau celsius  
Abs – absorbância  
APs – anilopirimidinas  
BSA – albumina do soro bovino  
C4 – Calcutta 4  
C4I – Calcutta 4 inoculado  
C4NI – Calcutta 4 não inoculado  
CAV – Cavendish  
CAVI – Cavendish inoculado  
CAVNI – Cavendish não inoculado  
cDNA – DNA complementar  
CHAPS – 3-[(3-Cholamidopropyl)dimethylammonio]-1 propanesulfonate  
CTAB – brometo de cetil-trimetilamônio  
DNA – ácido desoxirribonucleico  
DTT – ditioneitol  
EST – marcador de sequência expressa  
ET – etileno  
ETI – imunidade disparada por efetores  
g – grama  
Gb – gigabase  
GO – “*Gene Ontology*”  
HR – resposta de hipersensibilidade  
JA – ácido jasmônico  
KDa - quilodaltos  
LRR – repetição rica em leucina  
M – molar  
m/v – massa/volume  
MALDI – Ionização e dessorção a laser assistida por matriz  
MAMP – padrão molecular associado a micróbio  
Mb – megabase

MBCs – metil benzimidazoles carbonatos  
mg – miligrama  
mL – mililitro  
MM – massa molecular  
mM – milimoles  
MS – espectrômetro de massa  
NB – “nucleotide binding”  
NGS – “next generation sequencing”  
PAMP – padrão molecular associado a patógeno  
pb – pares de bases  
PCR – reação em cadeia da polimerase  
pH – potencial hidrogeniônico  
pI – ponto isoelétrico  
PPi – pirofosfato inorgânico  
PRs – genes relacionados a patogênese  
PTI – imunidade disparada por PAMP  
qRT-PCR – PCR quantitativo em tempo real  
R – coeficiente de correlação  
RGA – gene análogo de resistência  
RNA – ácido ribonucleico  
ROS – espécies reativas de oxigênio  
rpm – rotação por minuto  
rRNA – RNA ribossômico  
SA – ácido salicílico  
SAR – resposta sistêmica adquirida  
SNP – polimorfismo de nucleotídeo simples  
TCA – ácido tricloroacético  
TOF – tempo de voo  
v/v – volume/volume  
µg – micrograma  
µL – microlitro

## RESUMO

PASSOS, Marco Aurélio Ninomia. *Análise transcritômica e proteômica da interação *M. acuminata* x *M. musicola**. 2013. 95 Páginas. Tese de Doutorado em Biologia Molecular – Universidade de Brasília, Brasília, DF.

A Sigatoka Amarela em banana (*Musa* spp.), causada pelo fungo *Mycosphaerella musicola*, provoca desordem significativa de área foliar e amadurecimento prematuro do fruto. O desenvolvimento de genótipos resistentes a fungos patogênicos é de fundamental importância. A fim de desenvolver um recurso de genômica funcional para essa cultura oferecendo compreensão sobre os mecanismos moleculares das respostas de *Musa* a estresses bióticos, foi realizado um estudo de pirosequenciamento do transcrito de *Musa acuminata* utilizando genótipos contrastantes em resistência ao patógeno *M. musicola*. Amostras de RNA total foram preparadas a partir do material foliar dos genótipos Calcutta 4 (resistente) e Cavendish Grande Naine (suscetível), ambos não infectados (controles não inoculados) e artificialmente desafiados com o patógeno. O estudo gerou 978.133 seqüências de alta qualidade, com um comprimento médio de 334pb e totalizando 466Mb, das quatro bibliotecas de cDNA seqüenciadas usando o 454 GS-FLX Titanium. A montagem *de novo* gerou 36.384 e 35.269 seqüências unigenes de *M. acuminata* Calcutta 4 e Cavendish Grande Naine, respectivamente. Seqüências montadas foram funcionalmente mapeadas nos termos do Gene Ontology (GO). As funções do *unigenes* cobriram uma ampla gama de funções moleculares, processos biológicos e componentes celulares. Os genes oriundos de uma série de vias relacionadas com a defesa foram observadas em transcritos a partir de cada biblioteca de cDNA. Inúmeros fatores de transcrição foram identificados, que são tipicamente envolvidos na regulação do desenvolvimento vegetal, sinalização e resposta ao meio ambiente, entre outros papéis. Muitos deles são também conhecidas por estarem envolvidas na transdução de sinal e regulação da expressão de genes responsivos ao estresse. Mais de 99% dos unigenes mapearam em regiões exônicas no genoma referência de *M. acuminata* DH Pahang. Além disso, a extração de proteínas foi realizada a partir dos materiais foliares inoculados e não inoculados. A focalização isoeétrica foi realizada utilizando o sistema IPGphor. A segunda dimensão foi realizada utilizando um gel de SDS-PAGE a 12%, com géis corados com solução de azul de Coomassie G-250. Os géis foram digitalizados usando um scanner ImageScanner II (GE Healthcare), e as imagens foram analisadas utilizando o software Imagem Master 2D Platinum 7.0 (GeneBio). Foram feitas também análises para detecção de *spots* comuns ao mesmo cultivar (controle e tratado), bem como a identificação de *spots* com expressão diferencial, e detecção de *spots* presentes em apenas uma condição, denominados de *spots* exclusivos. No cultivar resistente Calcutta 4, foram observados 7 *spots* com uma expressão pelo menos 1,5 vezes maior no tratado quando comparado ao controle. O padrão de expressão foi diferente no cultivar suscetível, onde 1 *spot* apresentou diferença de expressão entre o controle e o tratado. Os resultados da análise proteômica desta interação poderá gerar informações importantes, tais como: os tipos de proteínas expressas, as mudanças pós-traducionais e respostas a diferentes condições ambientais, o que representa um elo entre a informação genotípica e fenotípica. Essa análise proteômica, em sobreposição com as informações transcritômicas geradas pelo seqüenciamento massal com NGS, podem vir a beneficiar futuros programas de melhoramento genético de *Musa* na obtenção de genótipos resistentes.

Palavras-chave: *M. acuminata*, *M. musicola*, transcritômica, proteômica.



## ABSTRACT

PASSOS, Marco Aurélio Ninomia. Transcriptomics and proteomics analysis of the interaction *M. acuminata* x *M. musicola*. 2013. 95 pages. Thesis (Doctorate in Molecular Biology) – Universidade de Brasília, Brasília, DF, Brazil

Sigatoka leaf spot disease in banana (*Musa* spp.), caused by *Mycosphaerella musicola*, causes significant reduction in functional leaf area and premature ripening of fruits. The development of genotypes resistant to fungal pathogens is of paramount importance. In order to develop a functional genomics resource for this crop which offers insights into molecular mechanisms of responses in *Musa* to biotic stresses, we performed a pyrosequencing study of the transcriptome in *Musa acuminata* genotypes contrasting in resistance to the fungal pathogen *M. musicola*. Total RNA samples were prepared from whole plant leaf material from genotypes Calcutta 4 (resistant) and Cavendish Grande Naine (susceptible), both uninfected (non-inoculated controls) and artificially challenged with the pathogen. The study generated 846,762 high quality sequence reads, with an average length of 334pb and totalling 283 MSb, from four full-length enriched cDNA libraries sequenced using a 454 GS-FLX system pyrosequencer with Titanium chemistry. *De novo* assembly generated 36,384 and 35,269 unigene sequences for *M. acuminata* Calcutta 4 and Cavendish Grande Naine, respectively. Assembled sequences were functionally mapped to Gene Ontology (GO) terms. Unigene functions covered a diverse range of molecular functions, biological processes and cellular components. Genes from a number of defense-related pathways were observed in transcripts from each cDNA library. Numerous transcription factors were also identified, which are typically involved in regulation of plant development, signaling and response to environment, amongst other roles. Many are also known to be involved in signal transduction and expression regulation of stress-responsive genes. Over 99% of contig unigenes mapped to exon regions in the reference *M. acuminata* DH Pahang whole genome sequence. Furthermore, protein extraction was conducted from the inoculated and non-inoculated leaf materials. Isoelectric focusing was performed using the IPGphor system. A second dimension was performed using an SDS-PAGE gel at 12%, with gels stained with Coomassie Blue solution G-250. Gels were scanned using a scanner ImageScanner II (GE Healthcare), with image analysis performed using the software Image Master 2D Platinum 7.0 (GeneBio). Assays to detect common spots were also made at the same cultivar (control and treated) as well as the identification of differentially expressed spots, and detection of spots present in only one condition, called exclusive spots. In the resistant cultivar Calcutta 4, 7 spots were observed at least 1,5 times higher expression in the treated compared with the control. The expression pattern was different in the susceptible cultivar, which showed 1 spot difference in expression between control and treated. The results of the proteomic analysis of this interaction can generate important information, such as: types of expressed proteins, post-translational changes and responses to different environmental conditions, representing a link between genotypic and phenotypic information. Proteomic analysis, overlapping with transcriptome information for this pathosystem, generated by NGS sequencing, can potentially benefit future programs for genetic improvement of *Musa* for development of resistant genotypes.

Keywords: *M. acuminata*, *M. musicola*, transcriptomics, proteomics

## 1. REVISÃO DE LITERATURA

### 1.1. Taxonomia de *Musa*

A bananeira (*Musa* spp.) é uma planta pertencente à classe *Liliopsida*, subclasse *Zingiberidae*, superordem *Lilianae*, ordem *Zingiberales*, família *Musaceae* (Cronquist, 1981; Belalcázar Carvajal, 1991). A família *Musaceae* é dividida em três subfamílias: *Heliconoideae*, *Strelitzoideae* e *Musoideae*. Dentro da subfamília *Musoideae* encontram-se os gêneros *Ensete* e *Musa*. As bananeiras com frutos comestíveis pertencem ao gênero *Musa*. A maioria dos cultivares é obtida a partir de duas espécies, *M. acuminata* (genoma A) e *M. balbisiana* (genoma B), sendo que a *M. acuminata* é a mais frequente das espécies de *EuMusa*. Mudanças cromossômicas estruturais que têm ocorrido espontaneamente, ou como um resultado de eventos de recombinação, resultaram no desenvolvimento de barreiras reprodutivas naturais no âmbito das espécies, causando divergência de subespécies e diversidade genética na espécie como um todo. Nove subespécies têm sido reconhecidas (Horry *et al.*, 1997).

O gênero *Musa* era dividido em quatro seções com base em seus números de cromossomos e características morfológicas: os que têm 11 pares de cromossomos (*EuMusa* e *Rhodochlamys*) e aqueles que têm 10 pares (*Australimusa* e *Callimusa*). A partir de 1976, foi criada a seção *Ingentimusa* para acomodar uma espécie solitária, *Musa ingens*, que tem sete pares de cromossomos. Recentemente, baseando-se em análises moleculares, este gênero possui duas seções apenas. A seção *Musa* que engloba os representantes das antigas *EuMusa* e *Rhodochlamys*, e a nova seção *Callimusa* que inclui as espécies das antigas seções *Australimusa* e *Callimusa*, além de também possuir o único representante da seção *Ingentimusa*, *M. ingens* (Häkkinen, 2013).

### 1.2. Importância social e econômica da banana

A banana é a mais importante fruticultura em termos de produção e consumo, sendo cultivada em regiões tropicais e sub-tropicais. Ela atua como um componente da dieta básica nos países em desenvolvimento, sendo considerada um alimento de alta energia, rico em minerais e vitaminas, com um curto prazo de validade (Trujillo *et al.*, 2007).

O Brasil é, atualmente, o quinto maior produtor de banana, contribuindo para, aproximadamente, 10% da produção mundial, com uma produção aproximada de 7,3 milhões de toneladas em uma área cultivada de 508 mil hectares. A produção anual global de mais de 120 países nos 5 continentes é estimada em cerca de 102 milhões de toneladas (FAOstat, 2012). A bananeira, todavia, é acometida por doenças provocadas principalmente por bactérias (*Ralstonia solanacearum*), nematóides (*Rhadopholus similis*), vírus (Banana streak virus), insetos (*Cosmopolites sordidus*) e fungos. As doenças que causam maior impacto na produtividade dessa cultura são causadas especialmente por fungos, dentre eles, *Fusarium oxysporum* f. sp. *cubense*, *Colletotrichum musae*, *Mycosphaerella fijiensis* e *Mycosphaerella musicola*.

### 1.3. *Mycosphaerella musicola*

A primeira descrição sucinta do fungo associado com a Sigatoka foi feita por Zimmermann (1902), como uma nova espécie de *Cercospora musae* Zimm. Durante quase quarenta anos após a sua descoberta, o fungo foi conhecido na sua forma imperfeita ou assexuada (conidial). Somente na década de 1930 Leach (1941), trabalhando na Jamaica, descobriu a forma perfeita (Teleomorfo) de *C. musae*, um fungo da classe dos Ascomycetes, para o qual a denominação de *Mycosphaerella musicola* foi atribuída.

O mal-de-Sigatoka ocasionado por *M. musicola*, três tipos de frutificações são produzidas nas manchas foliares ou manchas de Sigatoka em bananeiras: esporodóquios, espermogônio e peritécios (Stover, 1970). O processo sexuado no gênero *Mycosphaerella* envolve a formação de espermogônios, que produzem gametas masculinos, as espermásias, e o órgão sexual feminino, uma hifa espiralada, que é formada no interior de jovens ascocarpos, denominadas de tricogines (Wardlaw, 1961). Simmonds (1933) observou que espermogônios eram encontrados mais frequentemente por volta do final do ano em folhas manchadas e secas, ainda aderidas aos pseudocaulis. Em escala macroscópica, os espermogônios, de alguma forma, assemelham-se às pontuações negras formadas pelas frutificações conidiais, porém com um formato melhor delimitado de pontuação. Estas estruturas podem ser formadas em ambas as superfícies foliares, porém com maior predominância na abaxial. Sob microscópio de luz, os espermogônios, são pequenas frutificações negras em formato de frascos, imersas, que surgem no interior de uma base estromática de velhas frutificações conidiais ou independentemente. As espermásias, que são formadas em longas cadeias são bastante minúsculas, oblongas e hialinas, com formato semelhante às bactérias, e podem ser

visualizadas sendo expelidas a partir de um ostíolo ou poro no ápice dos espermogônios (Simmonds, 1933). O ciclo de vida de *Mycosphaerella* sp está ilustrado na Figura 1.

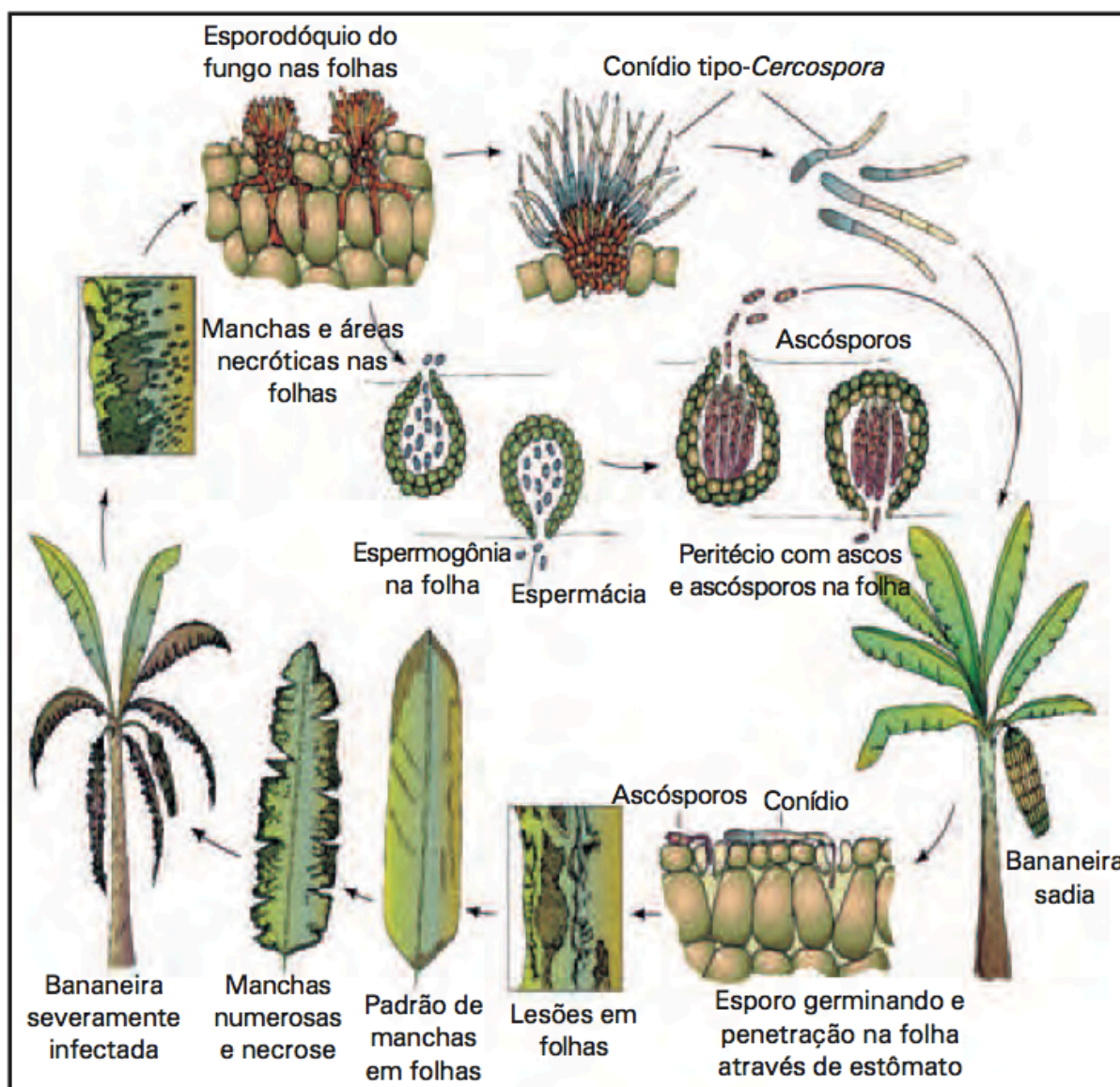


Figura 1 - Ciclo de *Mycosphaerella* spp em bananeira, mostrando as fases sexuada (telimórfica) e assexuada (anamórfica) do patógeno (Agrios, 2005).

Estão envolvidos, portanto, dois tipos de esporos no aparecimento da doença. O esporo sexuado que é o ascósporo, e o assexuado, que é o conídio. Estes são produzidos de forma contínua em climas úmidos, sendo disseminados pela água acumulada na superfície foliar durante as chuvas ou orvalho. Os ascósporos porém, produzidos nas mesmas lesões onde foram liberados os conídios anteriormente, surgem mais tardiamente, sendo ejetados a partir dos pseudotécios em períodos de alta umidade, ou mesmo em climas secos porém com ocorrências de orvalhos pesados (Simmonds, 1966).

No que concerne aos sintomas da Sigatoka-amarela, Meredith (1970), classificou o desenvolvimento das lesões em seis estádios: 1 - Estádio inicial de listra – A pontuação é perceptível ao olho nú apenas como uma leve mancha de descoloração verde-amarelada; 2- Segundo estágio de listra - A mancha aumenta de tamanho, principalmente em comprimento, ainda com aparência de descoloração verde-amarelada; 3 - Terceiro estágio de listra - A mancha começa a dilatar-se levemente na largura e principalmente no comprimento, tornando o centro da lesão com coloração levemente avermelhada enferrujada; 4 - Primeiro estágio de mancha - A listra ganha uma coloração marrom escuro, e simultaneamente, ou dentro de 24 horas, forma-se um halo com aspecto de encharcamento de cor marrom quando a folha encontra-se túrgida. Este halo é visível quando se observa a lesão contra a luz solar durante o período da manhã. A mancha aumenta consideravelmente de tamanho neste estágio. A listra chega a um estágio em que é facilmente reconhecida como uma lesão circular. 5 - Segundo estágio de mancha – A porção marrom escura da mancha torna-se enrugada e encolhida e o halo torna-se de coloração mais escurecida. 6 – Terceiro estágio de mancha – A mancha encontra-se plenamente desenvolvida, com a porção central abaulada e de coloração cinza, com o halo de coloração marrom escuro ou preto, formando um anel bem distinto ao redor da lesão. Este tipo de lesão permanece bem definida mesmo quando a folha está morta, com o anel escuro bem distinto circundando a lesão (Figura 2).

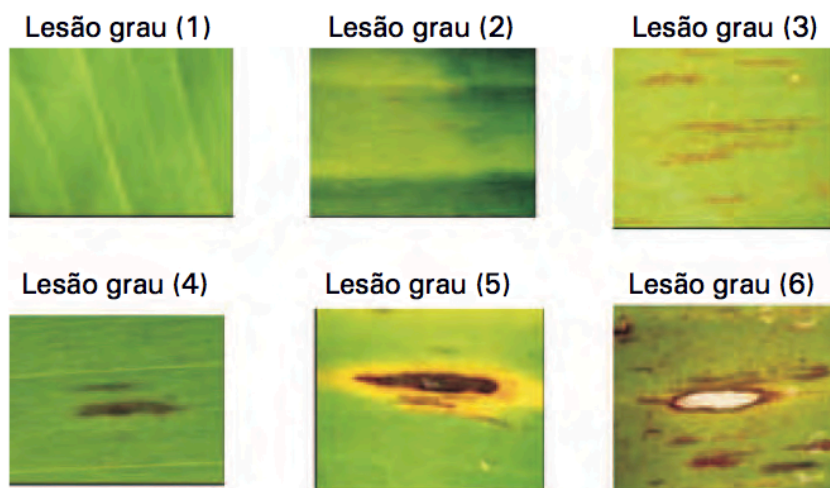


Figura 2 - Escala de Meredith (1970), para diferentes graus de desenvolvimento das lesões de Sigatoka-amarela. Foto: Aparecida Gomes de Araújo

Apesar dos severos danos ao limbo foliar, Leach (1946) afirma que a doença não afeta o desenvolvimento vegetativo em absoluto. Entretanto, Simmonds (1966) reporta a redução no tamanho dos cachos e dos frutos, presumivelmente pela redução da área fotossinteticamente ativa.

*M. musicola* é hemibiotrófico, heterotálico e pode apresentar reprodução sexuada, o que parece contribuir para o aumento da diversidade genética do patógeno (James *et al.*, 2010). Níveis elevados de variabilidade genética em populações de fungos, como de *M. musicola*, podem permitir a adaptação dos mesmos a diferentes ambientes e a genótipos de hospedeiras resistentes recém introduzidas (Churchill, 2011). Com isso, o controle do patógeno que, normalmente, é realizado com base na aplicação de fungicidas, é dificultado. Desta forma, conhecer a diversidade genética de populações de fungos e como os patógenos evoluem, bem como sua distribuição em uma escala regional e nacional contribui para o desenvolvimento de estratégias para o melhoramento de plantas e o manejo da resistência a doenças fornecendo, assim, informações para a busca de alternativas de controle para a doença. Contudo, ainda são escassos os estudos relacionados à diversidade genética de populações de *M. musicola*, sendo mais comuns em *M. fijiensis* (Condez-Ferráez *et al.*, 2007; Queiroz *et al.*, 2013).

A Sigatoka-amarela em território brasileiro, foi constatada pela primeira vez em 1944, no estado do Amazonas (Kimati & Galli, 1980), alastrando-se, posteriormente, por todos os estados brasileiros. Já a Sigatoka-negra teve seu primeiro relato em 1998 também no Amazonas, com o patógeno, subseqüentemente, espalhando-se nos sete estados do norte do país (Gasparotto *et al.*, 2006). Em 2005 o patógeno foi relatado também em São Paulo, Minas Gerais, Paraná e Santa Catarina, e agora ameaça o região nordeste. De todas as doenças afetando a produção de *Musa* no Brasil, o maior impacto é atribuído ao patógeno *M. fijiensis*. Os sintomas da doença incluem amadurecimento precoce da fruta, lesões necróticas e decomposição foliar, assim como perdas na produção de até 100% em variedades suscetíveis Cavendish (*Musa cv. AAA*) e Prata (*Musa cv. AAB*) (Cordeiro & Matos, 2003). Os sintomas característicos da Sigatoka amarela podem ser vistos na Figura 3. Como consequência, a produção de Cavendish está ameaçada, e totalmente dependente do uso de pesticidas. Considerando que mais de 50 aplicações de fungicida podem ser necessárias para o controle da doença em áreas suscetíveis (Gasparotto *et al.*, 2001), aliado a altos custos (com estimativas de uso de agrotóxicos representando até 27% do custo de produção) (Stover & Simmonds, 1987), a produção pode se tornar economicamente inviável. Logo, a atual dependência de fungicidas revela a importância de implementação de estratégias integradas para evitar a seleção de isolados do patógeno resistentes a fungicidas.



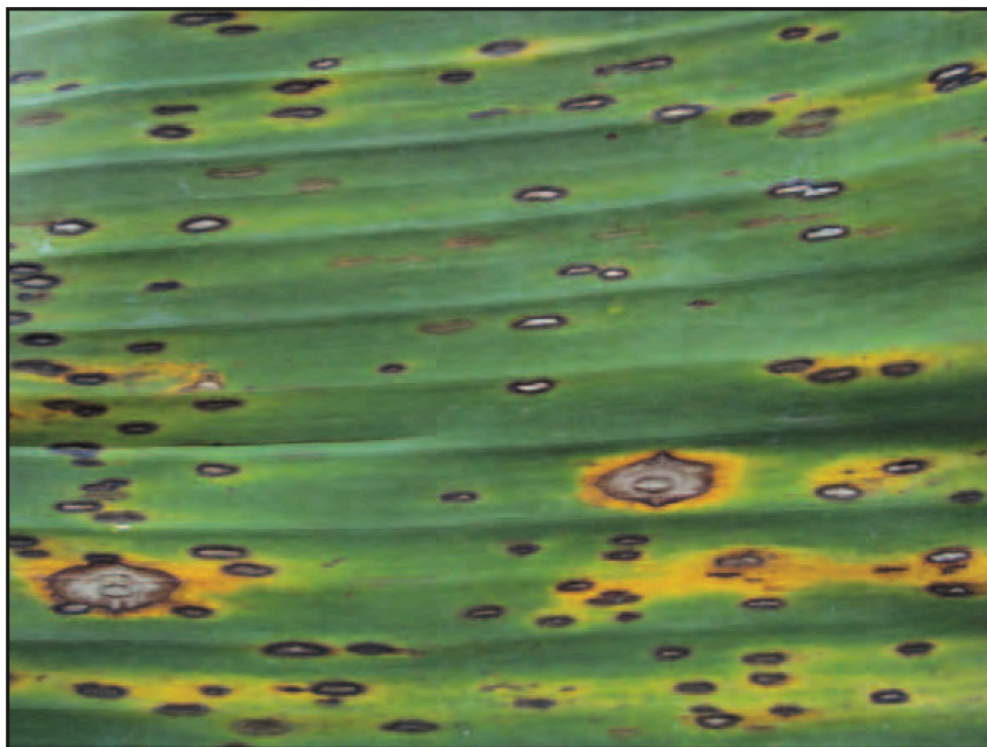


Figura 3 – Folha de bananeira com sintomas característicos da Sigatoka Amarela. Foto: Aparecida Gomes de Araújo

Em relação aos fungicidas, os que são adotados para o controle de Sigatoka-amarela são os mesmos utilizados para Sigatoka-negra. Sendo assim, as sete principais classes de fungicidas empregadas atualmente no controle de Sigatoka são os inibidores de demetilação (DMIs), aminas, *quinone outside inhibitors* (Qols; strobilurins), anilino-pirimidinas (APs), metil benzimidazoles carbonatos (MBCs) e guanidinas (Cañas-Gutiérrez *et al.*, 2009; Churchill, 2011). Contudo, os DMIs, um grande grupo de fungicidas denominado azole, estão entre os mais comumente usados. Esses, agem inibindo a esterol 14 $\alpha$ -demetilase, uma enzima da superfamília CYP51 envolvida na catalase do 14 $\alpha$ - demetilação do eburicol, um precursor da biossíntese do ergosterol. O ergosterol é importante para os fungos por regular o fluido na membrana e a permeabilidade das células fúngicas, sendo essencial para a sobrevivência da célula (Lepesheva & Waterman, 2004). Em contrapartida, a intensa aplicação deste fungicida causou uma pressão seletiva favorecendo o aparecimento de cepas resistentes (Cañas-Gutiérrez *et al.*, 2009).

O controle de Sigatoka-amarela também está relacionado a intensas aplicações de fungicidas DMIs. A baixa sensibilidade de *M. musicola* a propiconazole também já foi registrada (Peterson *et al.*, 2002). Entretanto, ainda não há relatos da relação da resistência do fungo com mutações no gene *CYP51*, tal como *M. fijiensis* e *M. graminicola*. Essa questão justifica um estudo mais detalhado a respeito da tolerância e resistência de *M. musicola* a fungicidas azole, já que, apesar dos severos danos provocados por *M. fijiensis*, no Brasil, *M.*

*musicola* ainda é a espécie de *Mycosphaerella* de maior ocorrência, provocando também, elevadas perdas nos cultivos de banana.

#### 1.4. Melhoramento genético em *Musa*

A maioria das espécies cultivadas e de interesse comercial encontra-se no subgênero *Eumusa* e compreendem diplóides ( $2n = 22$ ), triplóides ( $2n = 3x = 33$ ) e tetraplóides ( $2n = 4x = 44$ ) estéreis e que produzem frutos por partenocarpia e originaram-se de hibridações intra e interespecíficas entre duas espécies selvagens seminíferas, *M. acuminata* (genoma A) e *M. balbisiana* (genoma B), resultando em genótipos com grupo genômico AA, BB, AB, AAA, AAB, ABB, AAAA, AAAB, AABB e ABBB (Simmonds & Shepherd, 1955). Além dos grupos genômicos, foi estabelecido o termo subgrupo para denominar um complexo de cultivares com características comerciais de identidade específica, originárias por meio de mutações de uma única cultivar original (Shepherd *et al.*, 1984), como no caso o subgrupo Cavendish, grupo AAA e os subgrupos Prata e Terra do grupo AAB no Brasil. No entanto, em alguns híbridos estudados na Nova Guiné foi comprovada a participação de outras espécies, tais como *M. angustigemma* (genoma T) do subgênero *Australimusa* e de *M. schizocarpa* (genoma S) do subgênero *Rhodochlamys*, podendo ocorrer combinações dos tipos AS, AAS, ABBS, AAT e ABBT (D'Hont *et al.*, 2000).

Acredita-se que a evolução da bananeira selvagem seminífera para espécies cultivadas se deu provavelmente pela acumulação de esterilidade e aquisição de partenocarpia. Essas duas características foram selecionadas pelo homem dando origem as cultivares estéreis (Carreel *et al.*, 1994; Raboin *et al.*, 2005).

O surgimento da esterilidade, provavelmente se deve à origem dos diplóides cultivados que em sua maioria são híbridos entre subespécies distintas de *M. acuminata* com diferenças estruturais de seu genoma (Lanaud *et al.*, 1992; Perrier *et al.*, 2009). A seleção praticada pelos humanos deve ter ocorrido em cultivares que apresentava certa esterilidade e com desenvolvimento suficiente de polpa, que passaram a ser selecionadas e dispersas vegetativamente e ao longo do tempo cruzavam com outras formas. A origem das cultivares triplóides provavelmente envolve vários diplóides cultivados, parcialmente estéreis e que produziam gametas não reduzidos (Lanaud *et al.*, 1992) e que foram cruzando com diplóides fértil ou semi-fértil produzindo gametas haplóides (Raboin *et al.*, 2005).



O desenvolvimento de diplóides melhorados então, é considerado um importante componente no melhoramento genético. Sendo assim, o melhoramento convencional se inicia com a hibridação e seleção de recombinantes em nível diplóide, cujo objetivo é concentrar, em um mesmo genótipo (genitor masculino), um maior número de caracteres desejáveis, removendo eventuais defeitos inerentes dessas bananas não comerciais. Diplóides melhorados são então cruzados com cultivares triplóides, que apresentam fertilidade parcial, para produção de híbridos tetraplóides (Silva *et al.*, 1998). Nesse processo supõe-se que nos diplóides formam-se gametas reduzidos ( $n$ ) enquanto nos acessos triplóides formam gametas não reduzidos ( $3n$ ) dando origem aos híbridos tetraplóides primários. Os híbridos tetraplóides podem apresentar fertilidade masculina e feminina, reduzindo assim a qualidade da fruta devido à presença de sementes na polpa. Assim, uma outra possibilidade seria o desenvolvimento de híbridos triplóides secundários derivados do cruzamento entre acessos diplóides que produzem gametas reduzidos ( $n$ ) com cultivares ou híbridos tetraplóides com gametas ( $2n$ ) (Tenkouano *et al.*, 1998).

O objetivo do melhoramento da bananeira é, portanto, concentrar, em um mesmo genótipo, o maior número possível de características favoráveis como partenocarpia, elevado número de frutos e pencas, maior comprimento dos frutos, boa formação de cachos e resistência a doenças, para posteriormente tentar transferi-las às variedades triplóides comerciais, mediante a síntese de tetraplóides (Silva *et al.*, 1996; Silva *et al.*, 1997).

As dificuldades resultantes da baixa, e em alguns casos, ausência da produção de sementes em cruzamentos de bananeira, é potencialmente possível de se contornar, mediante o uso de tecnologias não convencionais de melhoramento, entre elas estão às técnicas de hibridação somática, duplicação de cromossomos, mutação e transformação genética.

A hibridação somática é uma estratégia promissora, que permite a introdução de resistência/tolerância a pragas e outras características de interesse encontradas em bananas diplóides (AA), nos cultivares triplóides, principalmente os estéreis. Matsumoto *et al* (2002) relataram a obtenção de híbridos somáticos entre o cultivar ‘Maçã’ e o diplóide ‘Lidi’ ressaltando o potencial desta técnica como ferramenta complementar para os programas de melhoramento.

Triplóides secundários podem ser produzidos numa pequena eficiência através do cruzamento de triplóides primários com diplóides melhorados. Uma alternativa do procedimento para a obtenção de triplóides secundários será a indução da duplicação de cromossomos de diplóides promissores, utilizando os agentes antimitóticos colchicina e orizalina (Ganga & Chezhiyan, 2002). Os autotetraplóides induzidos artificialmente podem

ser utilizados para o cruzamento com outros diplóides melhorados, gerando triplóides secundários que apresentem características de resistência às doenças e outras características agrônômicas superiores (Hamill *et al.*, 1992).

A mutagênese *in vitro* é uma técnica biotecnológica usada para alterar características agrônômicas governadas por um ou poucos genes, em genótipos de grande interesse, sendo considerada como um ajuste fino de uma variedade (Perez Ponce, 1998). Trabalhos já realizados utilizando a indução de mutação *in vitro* com o uso de raios gama em ápices caulinares de bananeira, indicaram um maior aumento na variabilidade genética, quando comparada com a taxa de variação somaclonal (Tulmann Neto & Domingues, 1990; Domingues *et al.*, 1994).

A transformação genética de bananeira via bombardeamento de partículas e por *Agrobacterium tumefaciens* foram conduzidos e resultados promissores têm sido obtidos, a exemplo Pei *et al.* (2005) obtiveram bananeiras transgênicas resistentes ao *Fusarium oxysporum* (FocR4) usando o gene da lisozima de humanos, que inibe o desenvolvimento desse fungo enquanto Kumar *et al.* (2005) obtiveram plantas transgênicas expressando o gene HBsAg (*Hepatitis B Surface Antigen*), revelando o potencial desta cultura para a produção de vacina contra a hepatite B. Embora essas técnicas até então não tenha gerado nenhuma variedade comercial mostra-se promissora para os casos onde o melhoramento convencional é difícil de se realizar.

Aliado a isso, a biotecnologia e a engenharia genética podem ajudar a superar os problemas no desenvolvimento de novos cultivares de banana (Heslop & Schwarzacher, 2007). Com isso, vários avanços têm sido alcançados. Estudos da citogenética molecular, identificação de marcadores moleculares, construção de mapas físico e genético, elucidação sobre os genes e sua expressão, bancos de dados de ESTs (*Expressed sequence tags*), e identificação de genes análogos de resistência (RGAs) vêm contribuindo para o entendimento do genoma da banana.

## 1.5. Imunidade vegetal (ETI e PTI)

O desenvolvimento de novas abordagens para a proteção das culturas exige avanços contínuos em nossa compreensão dos mecanismos moleculares que regulam a imunidade da planta. Ao longo das duas últimas décadas, as ferramentas moleculares e genômicas têm avançado consideravelmente, auxiliando nossa compreensão do sistema imune da planta.

### 1.5.1. Imunidade disparada por PAMPs

Um ramo do sistema imunológico da planta é conhecido como imunidade disparada por PAMPs (padrões moleculares associados a patógenos) (PTI - “PAMP-triggered immunity”), defesa inata ou defesa basal. Este ramo é baseado em interações entre a superfície do hospedeiro com PRRs (*pattern recognition receptors*) (Shiu & Bleeker, 2001; Altenbach & Robatzek, 2007), na matriz extracelular da planta ou da membrana plasmática (Boller & Felix, 2009), que reconhecem padrões moleculares associados aos micróbios ou patógenos (M/PAMPs) (Nürnberg & Kemmerling, 2009). Estes são epítomos microbianos, tais como flagelina bacteriana, fator de alongação bacteriano Tu, lipopolissacarídeos, peptídeoglicano,  $\beta$ -glucanases, ergosterol ou quitina de fungos (Boller & Felix, 2009). Tais epítomos indutores, são definidas como moléculas conservadas através de uma classe de um agente microbiano e consideradas essenciais na função ou na aptidão de sobrevivência do patógeno (Medzhitov & Janeway, 1997). Dado que as moléculas conservadas também podem ocorrer em microrganismos não patogênicos, também é utilizado o termo padrões moleculares associados aos micróbios (MAMPs) (Boller & Felix, 2009). Produtos de degradação da parede celular da planta hospedeira, também servem como indutores endógenos. Eles são libertados como resultado da presença de enzimas hidrolíticas do patógeno (Vorwerk *et al.*, 2004). Tais interações hospedeiro-patógeno ativa a PTI, que envolve a ativação de uma proteína associada ao mitógeno (MAP) quinase e de fatores de transcrição WRKY (TFS), impedindo ainda mais a colonização do hospedeiro, e confere resistência à maioria dos potenciais agentes patogênicos (Vivier & Malissen, 2005).

### 1.5.2. Imunidade disparada por efetores

Em um segundo ramo do sistema imunológico das plantas, patógenos podem superar a PTI. Após a secreção e a translocação para a célula hospedeira, estes genes suprimem ou interferem na sinalização da PTI, resultando em plantas que são suscetíveis à infecção por agentes patogênicos. A doença então, é o resultado desta susceptibilidade desencadeada por efetores (ETS) (Chisholm *et al.*, 2006; Da Cunha *et al.*, 2007; Block *et al.*, 2008; Gohre & Robatzek, 2008). Efetores diferem dos PAMPs, que são moléculas essenciais para a sobrevivência de agentes patogênicos microbianos. Neste caso, os patógenos visam especificamente os mecanismos de defesa do hospedeiro, transferindo efetores numa célula hospedeira suscetível. A partir disso, as plantas desenvolveram um mecanismo de defesa que reconhece os efetores através das proteínas NB-LRR, que possuem domínio de ligação a nucleotídeos (NB “nucleotide binding) e repetições ricas em leucina (LRR “leucine rich repeat”) (de tal forma que a proteína de virulência é agora classificada como uma proteína de avirulência). Estas proteínas ativam uma segunda linha de defesa conhecida como imunidade disparada por efetores (ETI) (Jones & Dangl, 2006), também conhecido anteriormente como gene-a-gene ou resistência raça-específica. Em contraste com a PTI, que ocorre em todos os membros de uma determinada espécie de planta, a ETI opera no nível intra-específico, com genótipos resistentes que possuem o polimorfismo necessário em genes de resistência (genes R) determinantes (Xiao *et al.*, 2008).

Quatro modelos importantes dentro da ETI têm surgido, que iniciou com o modelo de gene-a-gene e, posteriormente, o modelo de ziguezague, o modelo de guarda, e, mais recentemente, o modelo *decoy*. Estes modelos importantes podem ser resumidas como segue:

- a. No modelo de gene-a-gene (Flor, 1942), que foi originalmente descrito para explicar o patossistema *Melampsora lini* / linho, a reação de hipersensibilidade é o resultado de uma interação incompatível entre o gene de avirulência do patógeno e o gene de resistência do hospedeiro, Dentro deste modelo baseado na interação receptor-ligante (Keen, 1990), numerosos genes avr patogênicos de plantas e genes R já foram caracterizados (Figura 4).
- b. No modelo de guarda, uma proteína R não interage diretamente com a proteína AVR, mas guarda ou monitora o status de uma outra proteína do hospedeiro, que pode ser alvo do elicitor. (Van der Biezen & Jones, 1998; Dangl & Jones, 2001). Assim, não ocorre uma detecção direta da molécula elicitora do patógeno. O que acontece são alterações funcionais e/ou estruturais da célula hospedeira (Bent & Mackey, 2007) (Figura 4).

c. No modelo *decoy*, proteínas específicas, que são similares àquelas que podem ser alvos de elicitores patogênicos, são produzidas pela planta em algumas interações planta-patógeno. A função dessas proteínas *decoy* é se ligar aos elicitores e mediar as interações dos mesmos com proteínas R (van der Hoorn *et al.*, 2008). Por exemplo, uma proteína efectora AvrPto de *Pseudomonas syringae* infecta *Arabidopsis*, mas provoca uma resposta de resistência em plantas de tomateiro que transporta duas proteínas de resistência, Pto (que é um Ser / Thr quinase) e Prf (que é uma proteína de NBS-LRR) (Figura 4).

d. No modelo “zig-zague”, a amplitude final da resistência da doença ou suscetibilidade é proporcional à [PTI-ETS+ETI]. Na primeira fase, as plantas detectam a população microbiana (PAMPs), desencadeando uma resposta de imunidade (PTI). Na segunda fase, patógenos liberam efetores que interferem com a PTI, desencadeando a susceptibilidade por efetores (ETS). Na terceira fase, um efetor Avr é reconhecido pelas proteínas NBS-LRR, ativando a imunidade desencadeada por efetores (ETI), uma versão ampliada de PTI que muitas vezes induz a resposta de hipersensibilidade (HR). Na quarta fase, novos efetores surgem através do fluxo horizontal de genes, ajudando os patógenos a suprimir a ETI. A seleção por parte das plantas de novos alelos de proteínas NBS-LRR, que podem reconhecer um dos efetores recém adquiridos, dispara novamente a ETI (Figura 5) (Jones & Dangl, 2006).

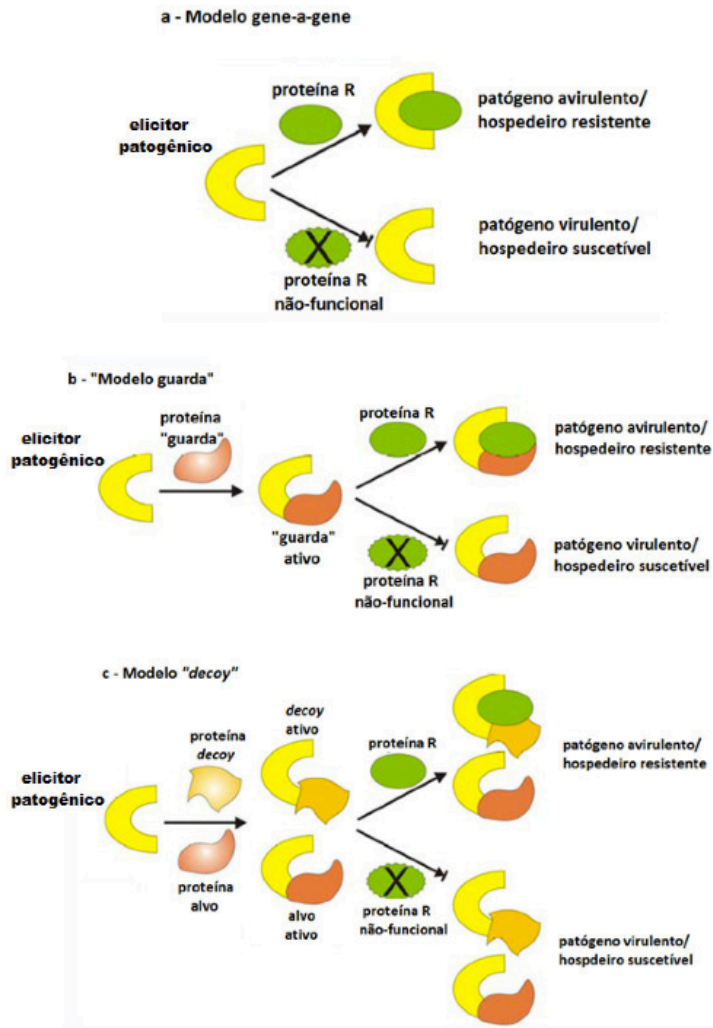


Figura 4 - Modelos de reconhecimento planta-patógeno. (a) Modelo de reconhecimento gene-a-gene, um reconhecimento específico direto do elicitor patogênico pelo receptor do hospedeiro desencadeando um processo de resistência. (b) Modelo de reconhecimento "guarda", em que uma proteína guarda coopera para que a resposta de resistência seja iniciada se ligando ao elicitor patogênico. (c) Modelo de reconhecimento "decoy". Uma proteína similar à proteína alvo do elicitor patogênico se liga ao mesmo e media sua interação com proteínas R. Fonte: Glowacki *et al.*, 2010.

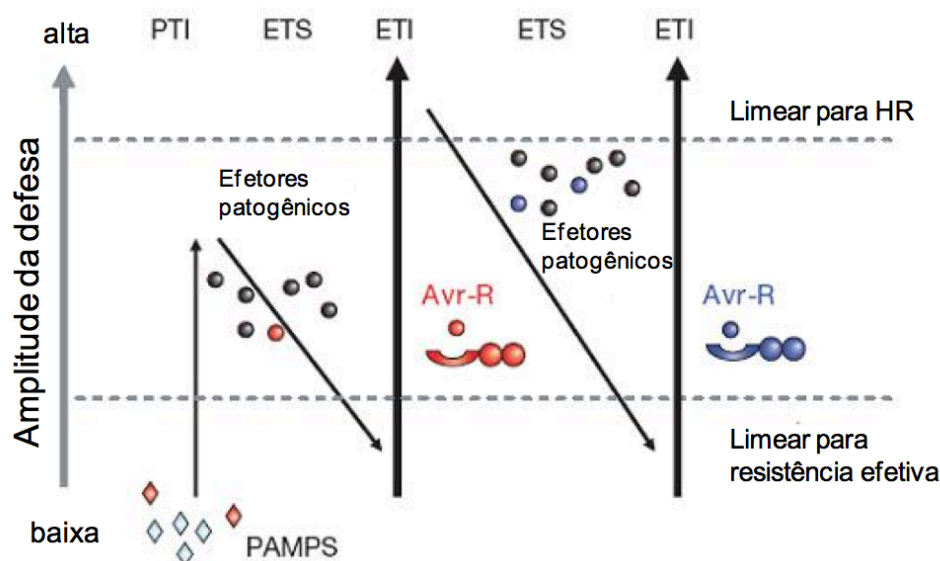


Figura 5 - Modelo “zigzague” ilustrando as etapas do desenvolvimento do sistema imunológico da planta durante infecção. Neste esquema, a amplitude final da resistência da doença ou susceptibilidade é proporcional à [PTI-ETS+ETI]. Na primeira fase, as plantas detectam a população microbiana (PAMPs, diamantes vermelhos), desencadeando uma resposta de imunidade (PTI). Na segunda fase, patógenos entregam com sucesso efetores que interferem com a PTI, desencadeando a susceptibilidade por efetores (ETS). Na terceira fase, um efector (indicado na figura em vermelho, Avr) é reconhecido pelas proteínas NB-LRR, ativando a imunidade desencadeada por efetores (ETI), uma versão ampliada de PTI que muitas vezes induz a resposta de hipersensibilidade (HR). Na quarta fase, novos efetores surgem através do fluxo horizontal de genes (em azul), ajudando os patógenos a suprimir a ETI. A seleção por parte das plantas de novos alelos de proteínas NB-LRR, que podem reconhecer um dos efetores recém adquiridos, dispara novamente a ETI. Fonte: Jones & Dangl, 2006.

### 1.5.3. Sobreposições entre PTI e ETI

Embora a PTI e a ETI sejam geralmente considerados em separado, as fronteiras não são claras (Thomma *et al.*, 2011). Por exemplo, têm sido caracterizados PAMPs com conservação estreita entre as classes de microrganismos, tais como Pep-13, que é um gene que codifica a proteína transglutaminase de parede celular, conservadas apenas entre os membros do gênero *Phytophthora* e envolvida na ativação de defesa da planta (Brunner *et al.*, 2002). Do mesmo modo, alguns efetores de agentes patogênicos foram caracterizados com uma ampla distribuição, tais como Nep1 (“necrosis and ethylene-inducing peptide1”), comuns em agentes patogênicos bacterianos, fúngicos e oomicetos (Gijzen & Nürnberger, 2006). Também contrariando as definições atuais, PAMPs podem desempenhar um papel na virulência. Um exemplo claro é a flagelina mutante de *Pseudomonas syringae* que possui mobilidade reduzida e, conseqüentemente, reduz virulência em plantas hospedeiras (Taguchi *et al.*, 2010).

Diferentes PAMPs, tais como o N terminal 22 de sequência de aminoácidos de flagelina flg22, peptidoglicano e quitina, foram relatados induzindo grandes respostas de transcrição na planta hospedeira, mostrando a expressão diferencial nas fases iniciais após o tratamento com PAMP (Gust *et al.*, 2007; Wan *et al.*, 2008), mas não durante as respostas tardias (Denoux *et al.*, 2005). Isto sugere sinalização comum no início da resposta, desencadeada por diferentes PAMPs, mas a expressão diferencial do gene em fases posteriores, talvez reflita diferentes respostas imunes necessárias para diferentes patógenos. Um estudo com flg22 e proteínas efetoras do patógeno também mostrou sobreposição na reprogramação transcricional induzida por ambos os ramos de imunidade da planta (Navarro *et al.*, 2004), com ETI evoluindo após PTI, com a maioria de seus mecanismos originados de mecanismos da PTI pré-existentes (Tsuda & Katagiri, 2010). Na verdade, muitos componentes de sinalização são compartilhados entre PTI e ETI, incluindo estresse oxidativo e alterações nos níveis hormonais de plantas. Entretanto, as respostas via ETI e PTI são tipicamente diferentes, com a resposta da ETI sendo geralmente mais rápida e duradoura do que a observada na PTI (Jones & Dangl, 2006; Tsuda & Katagiri, 2010).

#### **1.5.4. Mecanismos de sinalização**

A resposta de hipersensibilidade (HR), que é uma forma de morte celular programada, sendo uma resposta rápida, que restringe o alastramento do patógeno (Hammond-Kosack & Jones, 1997; Jones & Dangl, 2006). Embora associada a ETI, PAMPs como harpins e flagelinas também foram relatados como indutores de respostas de HR (por exemplo, Wei *et al.*, 1992; Taguchi *et al.*, 2003; Che *et al.*, 2000). Respostas de defesa celular na imunidade da planta geralmente iniciam com um estresse oxidativo através da produção de espécies reativas de oxigênio (ROS). Estas podem funcionar diretamente contra patógenos ou desempenham um papel na sinalização (Torres *et al.*, 2006). Na PTI, a produção de ROS é relatada como sendo rápida e apenas transitória (Zipfel *et al.*, 2006) e dependente da NADPH oxidase ArRbohD (Zhang *et al.*, 2007). Em contraste, a produção parece ser bifásica em ETI, com uma fase transiente inicial (possivelmente induzido por PAMPs), seguido por uma acumulação de maior dimensão e de maior duração, com ArRbohD também envolvidos na segunda fase (Torres *et al.*, 2006). Ocorrem em PTI e ETI, cascatas de sinalização celular mediada pela proteína quinase ativada por mitógeno (MAPK), embora variações na duração pode ser responsável por respostas diferentes nos dois ramos da imunidade (Tsuda & Katagiri,



2010). A resposta sistêmica adquirida (SAR) não está restrita apenas ao local da infecção, ela pode se ampliar, acumulando assim, compostos relacionados à defesa, em tecidos saudáveis, além disso, ela representa proteção por um determinado período do ciclo de vida da planta. A resposta do SAR é dependente da expressão de uma proteína reguladora NPR1 (Hammerschmidt, 2009). SAR é dependente do acúmulo de ácido salicílico (SA), levando a um aumento da expressão de genes relacionados a patogênese (PR) (Ward *et al.*, 1991). Embora o SA seja tipicamente importante na ETI contra organismos biotróficos e hemibiotróficos, a indução da SAR também foi relatada em *Arabidopsis* com PAMPs, tais como a flagelina e lipopolissacarídeos, com um típico aumento da expressão de ácido salicílico e proteínas relacionadas a patogênese (PR) (Mishina & Zeier, 2007). As vias de sinalização do ácido jasmônico (JA) e do etileno (ET) também estão envolvidas na defesa da planta, geralmente contra organismos necrotróficos (Glazebrook, 2005). O JA está envolvido na produção de proteínas PR basais e inibidores de proteases (Niki *et al.*, 1998). O ET também foi relatado por estar envolvido na produção de proteínas PR (van Loon *et al.*, 2006) e por possuir efeitos sinérgicos em relação a via de sinalização do JA (Lorenzo *et al.*, 2004). Embora as três vias são consideradas potencialmente antagônicas (Koornneef & Pieterse, 2008), também foi relatada evidência de sinergia entre os diferentes componentes dessa rede (Tsuda *et al.*, 2009). SA, JA e ET têm sido relatados tanto na PTI quanto na ETI, porém o modo como eles são empregados em cada ramo provavelmente diferem (Tsuda & Katagiri, 2010). A compreensão de todos os mecanismos moleculares que controlam a sinalização e a SAR está longe de ser resolvida, com sugestões de envolvimento de outras proteínas sinalizadoras, modificações pós-traducionais, expressão diferencial de alguns genes, e regulação positiva de fatores de transcrição (Van Der Ent *et al.*, 2009).

## 1.6. Tecnologias de sequenciamento de alto desempenho - NGS

O sequenciamento de DNA é uma ferramenta extremamente importante quando se trata de genômica. O sequenciamento de ESTs determina uma alternativa atraente para o sequenciamento de um genoma completo. Mesmo assim, como os genes são diferencialmente expressos, múltiplos tecidos devem ser amostrados, e, quando se utiliza o método tradicional de sequenciamento (Sanger), estes projetos exigem investimentos na construção e no sequenciamento de bibliotecas de cDNA. As novas tecnologias de sequenciamento, denominadas de tecnologias de sequenciamento de nova geração, começaram a ser comercializadas em 2005 e desde então, vêm evoluindo rapidamente. Todas essas tecnologias promovem o sequenciamento de DNA em plataformas capazes de gerar milhões de pares de bases em uma única corrida.

### 1.6.1. Plataforma 454

A primeira plataforma de sequenciamento de nova geração a ser comercializada foi o 454 *Life Sciences*, que desenvolveu um sistema de sequenciamento de DNA 100 vezes mais rápido do que os métodos normais, capazes de sequenciar mais de 1,000,000.000 fragmentos em 7 horas de corrida (Margulies *et al.*, 2005). O 454 utiliza o método de pirosequenciamento que se beneficia de uma técnica capaz de captar a emissão de luz causada pela adição de uma luciferase, acoplada à polimerização do DNA previamente fragmentado e aderido a microesferas, com o uso de seqüências adaptadoras (Figuras 6 e 7). Em média, o 454 (Titanium GS FLX 2009) lia, aproximadamente, 400 pb de comprimento. Atualmente, com esta tecnologia é possível obter 1.000.000 segmentos de leitura por corrida com tamanho aproximado de 1.000 pb (GS FLX). Esse aprimoramento das leituras advém de otimizações nas reações químicas do pirosequenciamento, as quais reduzem o ruído de fundo e aumentam o número de leituras por corrida, e do novo desenho do suporte de sequenciamento (*PicoTiterPlate*), o qual agregou duas mudanças principais: o uso de uma estrutura metálica, permitindo leituras mais acuradas, e esferas ainda menores, aumentando, tanto o tamanho das leituras, quanto o número de leituras por corrida (Roche, 2008). O maior tamanho das leituras e a grande capacidade de gerar informação tornam o processo de montagem mais fácil num projeto de sequenciamento *de novo* (sequenciamento de genomas desconhecidos) e permite trabalhar com coberturas genômicas mais amplas, favorecendo o processo de montagem.

Além disso, pelo fato de não envolver clonagem bacteriana, a representação do genoma é bem mais fiel, de forma que sequências difíceis de clonar e manter em bibliotecas genômicas podem ser acessadas.

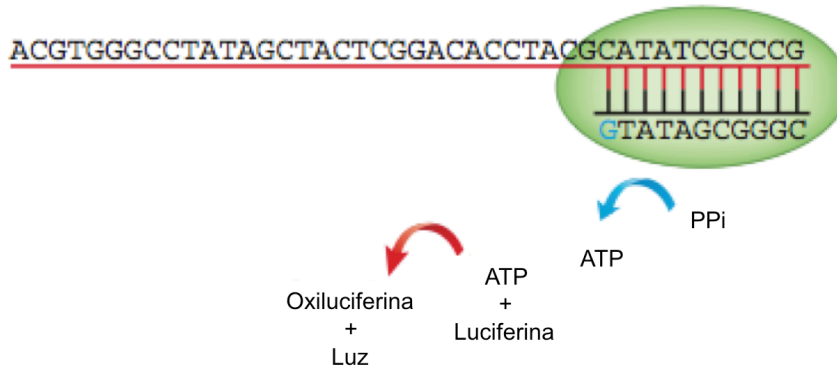


Figura 6 - Diagrama do processo do pirosequenciamento. A cadeia molde é representada em vermelho, o iniciador é mostrado em preto e a DNA polimerase é apresentada no círculo verde. A incorporação da base complementar (azul "G") gera um pirofosfato inorgânico (PPI), que é convertido em ATP pela enzima ATP sulfúrilase (seta azul). A luciferase (seta vermelha) utiliza o ATP para converter a luciferina em oxiluciferina, produzindo luz. Adaptado de Rothberg & Leamon, 2008.

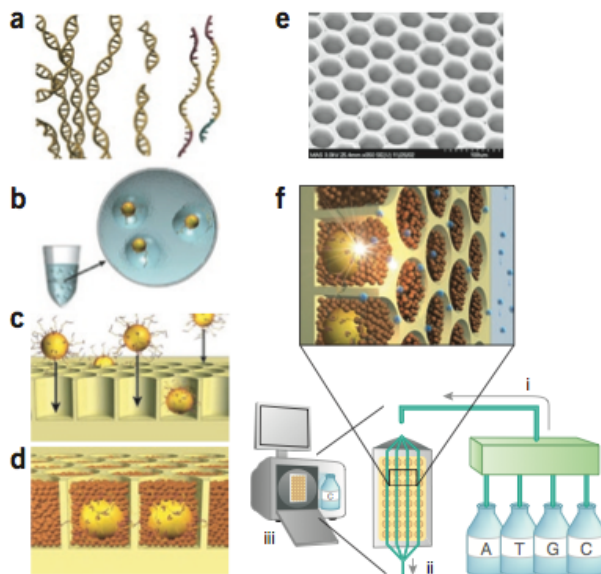


Figura 7 - Visão geral da tecnologia de sequenciamento de 454. (a) O DNA genômico é isolado, fragmentado, ligado a adaptadores e separados em cadeias simples. (b) Os fragmentos são ligados a esferas sob condições que favoreçam um fragmento por esfera. As esferas são isoladas e alocadas nas gotículas da reação da PCR em emulsão. A amplificação por PCR ocorre dentro de cada gotícula, o que resulta em dez milhões de cópias de um modelo de DNA original em cada esfera. (c) A emulsão é quebrada, as cadeias de DNA são desnaturadas, e as esferas que transportam os moldes de DNA de cadeia simples são depositadas em poços de uma lâmina de fibra óptica. (d) Esferas menores que transportam enzimas imobilizadas necessárias para uma reação de pirosequenciamento são depositadas em cada poço. (e) Microscopia eletrônica de uma parte da lâmina de fibra óptica, mostrando o revestimento de fibra óptica e os poços antes de deposição da esfera. (f) O aparelho do 454 consiste dos seguintes subsistemas principais: um conjunto de fluidos (objeto i), uma "flow cell" (objeto ii), sistema de detecção automático (objeto iii). Adaptado de Rothberg & Leamon, 2008.

Com relação às demais tecnologias de sequenciamento da segunda geração, a plataforma 454 é a que produz as sequências mais longas e por isso tem sido muito utilizada (Wicker *et al.*, 2006). Esta metodologia tem como principais desvantagens o alto custo dos reagentes e a taxa relativamente elevada de erro. Outra limitação importante da plataforma 454 é a baixa eficiência na determinação de homopolímeros. Como a intensidade do sinal de fluorescência relaciona-se ao número de vezes que um determinado nucleotídeo foi incorporado à sequência, a determinação precisa de sequências em que um único nucleotídeo é repetido mais de três vezes torna-se imprecisa. O custo do sequenciamento com essa plataforma é superior ao custo das plataformas Solexa e SOLiD, mas, nos casos em que a produção de leituras maiores é necessária, a plataforma 454 deve ser a melhor opção (Metzker, 2010), apesar de em muitos casos, não poder se considerar as bases localizadas no final dos *reads*. No entanto, devido ao grande número de dados gerados, esta tecnologia foi, e ainda é muito adotada. Esse fato levou a um novo desafio: o desenvolvimento de programas montadores de sequências para lidar com esse novo tipo de informação.

### **1.6.2. Plataforma Solexa**

O sequenciamento na plataforma Solexa, assim como o sequenciamento de Sanger, é realizado por síntese usando DNA polimerase e nucleotídeos terminadores marcados com diferentes fluoróforos. A inovação dessa plataforma consiste na clonagem *in vitro* dos fragmentos em uma plataforma sólida de vidro, processo também conhecido como PCR de fase sólida (Fedurco *et al.*, 2006; Turcatti *et al.*, 2008). A superfície de clonagem (*flow cells*) é dividida em oito linhas que podem ser utilizadas para o sequenciamento de até oito bibliotecas. Em cada linha, adaptadores são fixados à superfície pela extremidade 5', deixando a extremidade 3' livre para servir na iniciação da reação de sequenciamento dos fragmentos imobilizados no suporte por hibridização.

Os fragmentos de DNA da amostra são também ligados aos adaptadores em ambas as extremidades, o que permite sua fixação ao suporte de sequenciamento por hibridização a um dos adaptadores fixados. No primeiro ciclo de amplificação, nucleotídeos não marcados são fornecidos para que haja a síntese da segunda fita do fragmento imobilizado no suporte. A alta densidade de adaptadores no suporte facilita a hibridização do adaptador livre dos fragmentos imobilizados a sua sequência complementar fixa perto do clone inicial durante o ciclo de anelamento. Após o ciclo de anelamento, o fragmento forma uma estrutura em “ponte” na

superfície de sequenciamento e a extensão ocorre, formando a fita complementar também em “ponte”. No ciclo de desnaturação, as fitas são separadas e linearizadas. Esses ciclos são repetidos 35 vezes e assim as cerca de mil cópias geradas de cada fragmento nessa PCR de fase sólida permanecem próximas umas das outras, formando um *cluster* de sequenciamento. Etapas de desnaturação são necessárias para a separação dos duplex formados e, nos próximos ciclos de amplificação, nucleotídeos terminadores marcados são fornecidos para as reações de sequenciamento que ocorrem dentro de cada cluster. A alta densidade dos clusters de sequenciamento possibilita que o sinal de fluorescência gerado com a incorporação de cada um dos nucleotídeos terminadores tenha uma intensidade suficiente para garantir sua detecção exata. Até 50 milhões de *clusters* podem ser produzidos por linha, correspondendo a uma representação satisfatória da biblioteca. Após a incorporação de cada nucleotídeo no fragmento em síntese, a leitura do sinal de fluorescência é realizada. Em seguida, ocorre uma etapa de lavagem para remoção dos reagentes excedentes e remoção do terminal 3' bloqueado e do fluoróforo do nucleotídeo incorporado no ciclo anterior para que a reação de sequenciamento prossiga. A leitura das bases é feita pela análise sequencial das imagens capturadas em cada ciclo de sequenciamento. Em geral, leituras de 25-35 bases eram obtidas de cada *cluster* (Shendure & Ji, 2008), tornando-o menos adequado para a montagem *de novo* de genomas. Atualmente, esta tecnologia é capaz de gerar até 600 Gb com segmentos de 100 bases em média ([www.illumina.com](http://www.illumina.com)). O processamento e a análise estatística dos dados gerados resultam em novos desafios: atualmente um percentual dos *tags* são rotineiramente descartados devido a uma incapacidade para corresponder-lhes uma sequência de referência, reduzindo, assim, a efetiva capacidade da tecnologia (Rougemont *et al.*, 2008).

### 1.6.3. Aplicações

As novas plataformas de sequenciamento, tem revolucionado o conhecimento sobre o genoma das plantas, principalmente no que concerne ao estudo de variantes alélicas, SNPs, ao desenvolvimento de marcadores para seleção assistida e à clonagem baseada em mapeamento, representando uma importante ferramenta no melhoramento vegetal. Essa revolução será ainda maior com o avanço tecnológico das próprias plataformas de sequenciamento, produzindo leituras maiores, avanço das ferramentas de análise de leituras curtas e montagem de sequências.

As novas plataformas de sequenciamento apresentam a grande vantagem de permitir um

sequenciamento altamente representativo de genomas e/ou transcritomas em um único passo, o que é extremamente relevante, em razão da grande redução de custo alcançada com essas metodologias. Seu emprego tem revolucionado a transcriptômica com a geração de dados altamente reprodutíveis e informativos e com precisão na quantificação de transcritos. Em função do problema da montagem das leituras curtas produzidas por essas tecnologias, seu uso na genômica de plantas tem sido direcionado para o sequenciamento dos genomas plastidiais, sequências expressas, clones de interesse, ressequenciamento e detecção de variantes genóticas. Com os avanços mais recentes, a NGS pode auxiliar também, estudos de epigenômica, descoberta de RNAs não codificantes, elucidação da interação DNA-proteína, modificação de histonas e análise do degradoma.

Wicker *et al.* (2006) demonstraram que o pirosequenciamento via 454 permite um sequenciamento rápido, resultando em grandes porções de um genoma complexo. Vários estudos utilizando essa nova tecnologia têm sido descritos, envolvendo sequenciamento de ESTs para caracterização e análise do transcrito (Vera *et al.*, 2013; Emrich *et al.*, 2007). Além disso, o 454 foi utilizado na genômica comparativa de ervilha com os genomas da soja e da espécie *Mendicago trunculata* (Macas *et al.*, 2007) e na descoberta de SNPs em *Eucalyptus grandis* (Novaes *et al.*, 2008). Sequências derivadas do 454 foram utilizadas também na identificação de genes relacionados a resistência em feijão (Liu *et al.*, 2012) e na descoberta de genes diferencialmente expressos em *Prunus mume* (Shi *et al.*, 2012). O sequenciamento *de novo* é uma estratégia que está sendo amplamente adotada para a caracterização e análise do transcrito de diversas espécies de plantas (Zhang *et al.*, 2012; Yang *et al.*, 2013; Li *et al.*, 2013; Toledo-Silva *et al.*, 2013; Men *et al.*, 2013).

Em banana, o sequenciamento 454 foi utilizado na investigação de sequências repetitivas do genoma, fornecendo dados importantes para o isolamento de marcadores de moleculares para serem usados em estudos de diversidade genética e na seleção assistida por marcadores (Hřibová *et al.*, 2010). As estratégias adotadas neste presente trabalho são sustentadas por estudos recentes em *Musa*. Bai *et al.*, 2013 analisaram o perfil de expressão de raízes do cultivar Cavendish inoculado com *Fusarium oxysporum* f. sp. cubense tropical race 4, causador do Mal-do-Panamá. Os dados foram gerados pelo sequenciamento de alto desempenho utilizando o Illumina e validados por qRT-PCR. Os resultados apontam uma maior expressão de genes envolvidos em defesa no cultivar resistente, indicando que estes genes desempenham, de fato, papéis importantes na bananeira contra infecção a *F. oxysporum*. Esses resultados ajudam a explicar a resistência na cultivar resistente e proporciona mais conhecimentos na compreensão das interações compatíveis e incompatíveis

entre banana e este importante patógeno. Wang *et al.* (2012) compararam os perfis do transcrito em raízes de resistentes e suscetíveis do cultivar Cavendish também desafiado com *Fusarium oxysporum* f. sp. cubense raça tropical 4. Eles analisaram a expressão gênica 0, 2, 4 e 6 dias após a infecção com o patógeno. Os dados foram gerados utilizando o RNA-seq e montagem *de novo*. Os resultados indicaram que os mecanismos de defesa basais estão envolvidos no reconhecimento de PAMPs, e que os níveis elevados de transcritos relacionados com defesa podem contribuir para a resistência da bananeira (Li *et al.*, 2012). Assim como em nosso estudo, os resultados deste trabalho contribuem para a identificação de genes candidatos relacionados à resistência de plantas em um organismo não-modelo, e ajuda a melhorar a compreensão atual das interações patógeno-hospedeiro.

Dois genomas referência montados a partir do sequenciamento utilizando NGS foram descritos recentemente. D'Hont *et al.* (2012) descreveram a seqüência do genoma de 523MB de um genótipo duplo-haplóide de banana, *M. acuminata* ssp. malaccensis var. Pahang (DH Pahang). Davey *et al.* (*in press*) relataram o sequenciamento do genoma de *M. balbisiana* variedade ' Pisang Klutuk Wulung ' (PKW), que tem sido associada com uma melhor tolerância aos estresses bióticos e abióticos e é, também, um alvo para os programas de melhoramento de *Musa*. Eles alinharam o genoma da PKW (genoma B) contra o genoma A publicado anteriormente. O genoma B resultante representa 79 % do tamanho do genoma A, e contém 36.638 seqüências de genes preditos. Resultado este, quase idêntico aos 36.542 genes preditos no genoma A. Os dados gerados nesses estudos, tais como, genes preditos, listas de proteínas traduzidas, SSRs, SNPs, e famílias de microRNAs e seus possíveis alvos, poderão proporcionar um importante avanço para o melhoramento genético dessa cultura.

## **1.7. Proteômica**

O sequenciamento de genes em larga escala constitui hoje, uma importante ferramenta na dedução de seqüências de aminoácidos de suas proteínas. Porém, isto não informa de qual modo a informação está sendo expressa. Esta estratégia não revela informações sobre por exemplo, modificações pós-traducionais, regulação da expressão, funcionamento individual ou dedução do conjunto de proteínas de um organismo. Portanto, torna-se necessária uma abordagem proteômica no estudo de um conjunto de proteínas expressas pela célula em um dado momento e sob determinada condição, sendo possível identificar alterações e interações

que ocorrem dentro da mesma. Neste contexto, o uso de uma abordagem proteômica para estudar interações planta-patógeno tem se tornado mais frequente. A abordagem mais frequentemente utilizada nesses estudos tem sido o 2-DE seguido de espectrometria de massa (Quirino *et al.*, 2010).

O método clássico para a caracterização inicial de um proteoma é a eletroforese bidimensional de proteínas (2DE-PAGE). Essa técnica combina focalização isoelétrica (primeira dimensão) numa tira de gradient de pH imobilizado, com o objetivo de alocar as proteínas pelos seus pontos isoelétricos. Em seguida, é realizada a eletroforese em gel de poliacrilamida (segunda dimensão), que separa as proteínas pelas suas massas moleculares (MM) (Celis *et al.*, 2000). Desta forma, o ponto isoelétrico (pI) e a MM de cada spot do gel 2D podem ser determinados experimentalmente. As proteínas correspondentes aos *spots* são posteriormente excisadas do gel e digeridas utilizando uma protease (por exemplo, tripsina). Os peptídeos gerados são então analisados e validados por espectrometria de massa (Kuster *et al.*, 2001).

A proteômica de plantas tem conquistado progressos, porém ainda existem vários desafios analíticos que aguardam soluções experimentais e conceituais. Para plantas não-modelo, o gargalo em proteômica deve mudar de identificação de proteínas à determinação da função protéica. Particularmente, para o campo da interação planta-patógeno, existem características únicas para cada patossistema e trabalhar em comparação com organismos modelo, como o arroz não é o suficiente para revelar toda a “história” (Quirino *et al.*, 2010).

### **1.8. Transcritômica x Proteômica**

O aumento da frequência de anotação demonstra que o sucesso de uma abordagem transcritômica depende enormemente da disponibilidade de sequências depositadas em bancos de dados. Isto está em contraste com uma abordagem proteômica, principalmente quando uma base de dados pública de outras espécies de plantas é usada. Em ambas as abordagens, uma grande parte dos transcritos e proteínas ainda permanecem não identificados. A complementariedade dos estudos de transcritômica e proteômica ainda apresentam vantagens e desvantagem, como mostrado na Tabela 1.



Tabela 1 - Características da abordagem transcritômica e proteômica

Critérios de comparação	Transcritômica	Proteômica
Dependência de propriedades físico-químicas	Análise independente das propriedades físico-químicas da proteína codificada	Análise dependente das propriedades físico-químicas da proteína codificada (hidrofobicidade, peso molecular, pI)
Habilidade do operador	Independente do operador	Dependente do operador
Nível biológico de quantificação	Transcrito é um intermediário	Quantificação do produto final
Desempenho	Alto desempenho	Alto desempenho Viés para proteínas abundantes
Detecção de moléculas de baixa abundância	Transcritos pouco abundantes podem ser detectados	
Anotação	Anotação altamente dependente da caracterização das espécies	Poderosa anotação entre espécies

\*Adaptado de Carpentier *et al.*, 2008

Vários estudos recentes utilizam a integração das análises transcritômica e proteômica. Pesquisas com fungo, cianobactéria, bactéria, nematóide, planta e humanos têm sido relatadas (Barker *et al.*, 2012; Huang *et al.*, 2013; Liu *et al.*, 2013; McNulty *et al.*, 2013; Steiling *et al.*, 2009; Wilmes *et al.*, 2013). Em *Synechocystis* sp., a integração do transcrito e proteoma proporcionou um catálogo completo das modificações metabólicas, fornecendo novas hipóteses sobre o mecanismo de adaptação deste organismo em situação de estresse. Em *Arabidopsis* sp., a abordagem proteômica-transcritômica mostrou ser útil para estudar os padrões de proteínas e expressão de genes induzidos em resposta a diferentes estímulos. A integração dessas duas ferramentas de investigação tem o potencial de promover consideravelmente nosso entendimento sobre uma espécie, por facilitar a descoberta de genes, bem como verificar a expressão dos mesmos sob diversos tratamentos diferentes, e com isso possibilitar uma melhor compreensão da regulação da expressão gênica.

Em um estudo com o nematóide *Heligmosomoides polygyrus* foi empregado o sequenciamento de próxima geração e abordagens proteômicas. Essa abordagem permitiu identificar funções e proteínas associadas que podem desempenhar um papel na interação hospedeiro-parasita. Os métodos e as análises apresentados neste estudo são aplicáveis para estudar os aspectos bioquímicos e moleculares da relação hospedeiro-parasita em espécies

que possuem pouca informação de sequências em bancos de dados (Moreno *et al.*, 2011). Combinação das análises de sequenciamento do proteoma (LC-MS/MS) e do transcrito (Illumina) provou ser uma poderosa ferramenta para a descoberta de enzimas de degradação da parede celular de plantas (PCWDEs) ativos em uma espécie não-modelo. Os dados representam o ponto de partida de uma caracterização em profundidade funcional e evolutiva de famílias de genes PCWDE e sua contribuição para a adaptação destes herbívoros (Roy Kirsc *et al.*, 2012). Em arroz, a identificação de proteínas diferencialmente expressas e os respectivos mRNAs de coleótilo de arroz sob anoxia comparadas com coleótilo aeróbico foi realizada utilizando as ferramentas MALDI- TOF MS e qRT-PCR, onde a maioria das proteínas identificadas foram relacionados com respostas a estresse (Sadiq *et al.*, 2011).

## 2. JUSTIFICATIVA

Considerando as dificuldades para os melhoristas de banana, a genômica pode ser utilizada para o descobrimento de genes de resistência para o controle de Sigatoka Amarela. Genes envolvidos na resistência dessa cultura ao patógeno *M. musicola* poderão ser identificados neste estudo através das análises transcritômica e proteômica da interação patógeno-hospedeiro. Com isso, este estudo poderá contribuir para a elucidação dos mecanismos de defesa e resistência em *Musa*. Tecnologias que trabalham com perfis de expressão gênica facilitam a elucidação de tais mecanismos. Assim, o entendimento mais amplo das interações planta-praga, por meio da análise do transcritoma e da proteômica desta interação, poderá contribuir para a identificação eficiente de fatores genéticos associados à resistência não específica, estável e durável de plantas a esta praga. Pretende-se identificar genes envolvidos nestas interações que poderão futuramente ser introduzidas em variedades comerciais, visando obtenção de plantas resistentes a estas pragas.

### 3. OBJETIVOS

#### 3.1. Objetivo geral:

- Identificar e caracterizar de componentes genéticos de resistência e defesa ao patógeno *M. musicola* em *Musa* spp.

#### 3.2. Objetivos específicos:

- Caracterizar o transcrito, via piroseqüenciamento (454), de folhas de bananeira contrastantes infectadas *in vivo* com o patógeno *M. musicola*; sendo que *M. acuminata* ssp. *burmannicoides* var. Calcutta 4 representa a variedade resistente; e *M. acuminata* ssp. *burmannicoides* var. Grande Naine (Cavendish) representa o cultivar suscetível;
- Caracterizar e anotar genes de vias metabólicas implicadas na defesa via identificação de seqüências homólogas em bancos de dados de nucleotídeos e proteínas e mapeamento no genoma DH Pahang;
- Análisar *in silico* a expressão diferencial de genes gerados a partir do seqüenciamento massal nas variedades contrastantes;
- Construir mapas proteômicos do material foliar de *M. acuminata* Calcutta 4 e *M. acuminata* Cavendish inoculados e não-inoculados com o patógeno *M. musicola*;
- Identificar as proteínas expressas diferenciais entre as condições de tratamento (Calcutta 4 e Cavendish inoculados e não inoculados com *M. musicola*);

## 4. METODOLOGIA

### 4.1. Coleção de isolados de *M. musicola*, identificação, crescimento e manutenção

A coleção de *M. musicola* do presente trabalho possui 3 isolados, adquiridos de folhas de bananeira (*M. acuminata* Prata Anã, Prata Comum e Cavendish Grande Naine), pertencentes à Embrapa Mandioca e Fruticultura Tropical, em Cruz das Almas, BA.

Para o isolamento de colônias de *M. musicola*, fragmentos das folhas com as lesões características da Sigatoka amarela foram colocados em placas de petri com algodão umedecido e incubados por 2 a 3 dias em temperatura ambiente (Figura 8). A identificação do fungo foi realizada com auxílio de uma lupa e em seguida, o mesmo foi transferido para placas de petri com ágar-ágar, mantido em condições de luz constante com temperatura de 25°C. Após o surgimento de colônias, isolados de *M. musicola* foram transferidos, com o auxílio de um alça bisturi, para placas de petri com meio de cultura V8 (em 450 mL de água destilada, 1 g de CaCO<sub>3</sub>, 9g de Ágar, 50 mL de V8, 167 mg de Cloranfenicol). Após sete dias de crescimento a uma temperatura de 25°C e luz constante, as colônias presentes em cada placa de petri foram maceradas com auxílio de pistilo e cadinho de porcelana, ressuspensas em 1,2 mL de H<sub>2</sub>O milli-Q autoclavada. Alíquotas de 300 µL de suspensão foi transferido para novas placas de petri contendo meio de cultura V8, e em seguida foi espalhada empregando-se, para tal, uma alça de Drigalski (Figura 9). Após 10 dias a uma temperatura de 25°C e luz constante foi adicionado H<sub>2</sub>O milli-Q autoclavada em cada uma das placas. Com o auxílio de uma espátula, raspou-se as colônias para haver a liberação dos esporos. A suspensão obtida foi filtrada em gaze e sua concentração ajustada, em hemocitômetro, para  $2,0 \times 10^4$  esporos viáveis por mililitro (Figura 10). Algumas colônias foram transferidas para tubos Eppendorf de 2,0 mL com 500µL de glicerol 15% e mantidas a -80°C.

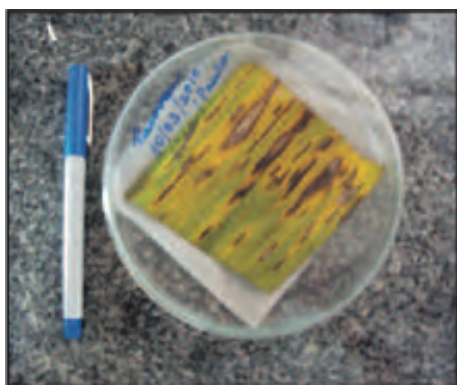


Figura 8 - Câmara úmida de folhas de bananeira para emissão de esporodóquios. Foto: Aparecida Gomes de Araújo



Figura 9 - Etapas do processo de multiplicação de *Mycosphaerella musicola* (A a F), para produção de inóculo. A e B – produção de um macerado a partir de colônias do fungo; C e D – transferência de uma alíquota do macerado para meio V8; E e F – espalhamento do macerado sobre o meio usando alça de Drigalski. Foto: Aparecida Gomes de Araújo

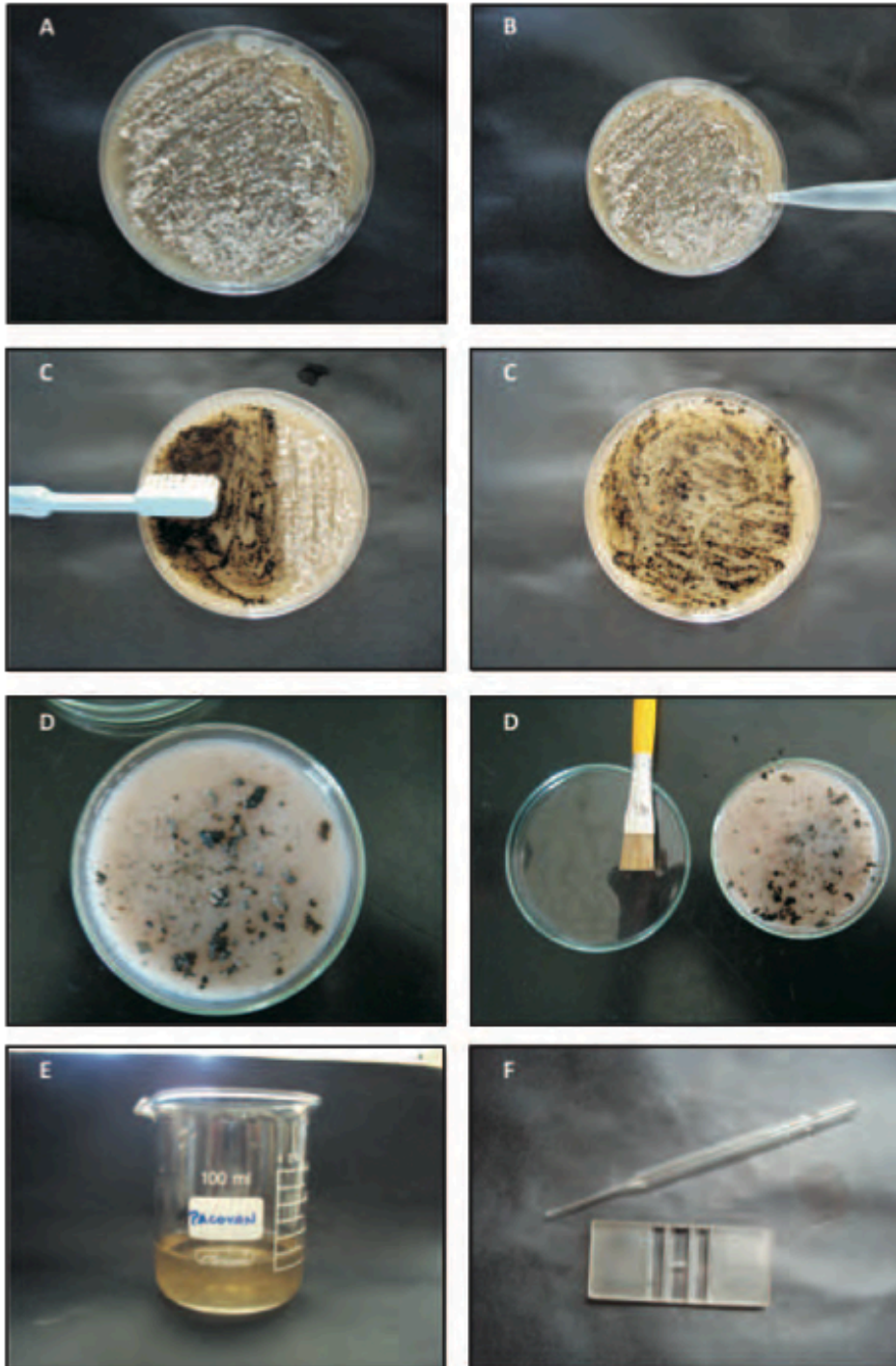


Figura 10 - Sequência para liberação de conídios de *Mycosphaerella musicola* em cultura: A – aspecto da colônia; B - adição de água; C – liberação de conídios com escova macia; D – liberação com pincel; E – suspensão obtida; F - contagem de esporos em câmara de Newbauer.  
Foto: Aparecida Gomes de Araújo

## 4.2. Bioensaio

Cultivares de *M. acuminata* Calcutta 4 e *M. acuminata* Cavendish Grande Naine foram propagadas *in vitro* e mantidas em casa de vegetação. As plantas foram mantidas em estufa sob um fotoperíodo de 12-h luz/12-h escuro a 25 ° C e umidade relativa de 85%. Plantas controle não infectadas foram mantidas sob condições idênticas. A inoculação foi realizada por pulverização da folha inteira, com uma suspensão de  $2,0 \times 10^4$  mL de conidiósporos e adição de agente tensoativo Tween 20 a 0,05%. borrifando as duas folhas mais jovens de nove plantas de *M. acuminata* Calcutta 4 e nove *M. acuminata* Cavendish. Outras 18 plantas, sendo nove de *M. acuminata* Calcutta 4 e nove *M. acuminata* Cavendish foram borrifadas apenas com H<sub>2</sub>O milli-Q autoclavada/Tween 20 a 0,05%. As folhas pulverizadas foram cobertas com sacos plástico transparentes para assegurar a uma alta umidade. Três réplicas independentes foram coletados para cada amostra.

Foram retirados 1g de material foliar da folha 1, sendo rapidamente transferidas para nitrogênio líquido e, posteriormente, armazenadas a -80°C até o momento da extração do RNA total. A folha 2 foi retirada e também, transferida para nitrogênio líquido e armazenada a uma temperatura de -80°C para posterior análise proteômica. Este procedimento foi realizado no terceiro dia após o processo de infecção, no sexto dia e no nono dia, de acordo com observações de penetração microbiano nos estômatos, conduzidas por microscopia eletrônica (dados nao mostrados).

## 4.3. Extração do RNA total de *M. acuminata* Calcutta 4 e *M. acuminata* Cavendish Grande Naine inoculados e não inoculados com *M. musicola*.

Foram testados quatro métodos de extração de RNA (Trizol/Invitrogen, CTAB, *Concert Plant RNA Reagent*/Invitrogen e *RNeasy mini Kit*/Qiagen). Nas extrações de RNA de folhas de *M. acuminata* Cavendish, obteve-se produto com os protocolos: CTAB (8ng/μL), CONCERT (10ng/μL), sendo este produto com boa integridade e pureza. As bandas originadas a partir da extração realizada utilizando esses protocolos apresentaram-se bem definidas, mostrando uma boa integridade do RNA extraído. Além disso, o gel não revelou nenhum tipo de impureza, já que não apresentou nenhum “rastros” nas extrações em que se



obteve sucesso. O método de extração de RNA selecionado foi o kit *Concert® Plant RNA Reagent* (Invitrogen).

A extração do RNA total de 36 amostras de folhas infectadas e não infectadas de *M. acuminata* Cavendish Grande Naine e *M. acuminata* Calcutta 4 foi conduzida para a análise do transcriptoma por pirosequenciamento e para o seu uso subsequente na validação da expressão gênica via qRT-PCR. A extração do RNA foi feita em tubos *falcon* de 15 mL. O RNA foi precipitado com isopropanol a -20°C e ressuspendido em H<sub>2</sub>O RNase free. O RNA aliquotado foi armazenado a uma temperatura de -80°C.

O RNA foi purificado utilizando o Kit Invisorb Spin Plant RNA Mini Kit (Invitex). A quantificação do RNA total extraído foi medida através da determinação das absorbâncias (Abs) a 260nm, 280nm e 230nm em espectrofotômetro ND-1000 (NanoDrop technologies), onde Abs<sub>260nm</sub> = 1 corresponde a 40µg de RNA em 1 mL de solução, quando a razão Abs<sub>260nm</sub>/Abs<sub>280nm</sub> = 1,9. A razão Abs<sub>260nm</sub>/Abs<sub>230nm</sub> é uma outra medida de pureza, com os valores de um ácido nucléico "puro" em geral na faixa de 1,8 - 2,2.

A integridade do RNA total foi também avaliada pela intensidade de bandas referentes aos rRNA 28S e 18S, visualizada por meio de eletroforese utilizando gel de agarose 1% (14 V/cm) na presença de brometo de etídio (1µg/mL) e marcador molecular padrão *Low DNA Mass Ladder* (Invitrogen)

#### **4.4. Sequenciamento massal via 454**

Quatro '*pools*' de RNA total foram preparados, cada um a partir de nove amostras de RNA total, representando os diversos momentos durante o bioensaios. A síntese de bibliotecas de cDNA ('full-length enriched') e pirosequenciamento foram conduzidos por Eurofins MWG Operon (Ebersberg, Germany), a partir de 50µg de RNA total de cada pool. PCR em emulsão e sequenciamento via 454 foram conduzidos de acordo com os protocolos do fabricante (Roche Applied Science), utilizando GS FLX tecnologia e química da série Titanium, onde centenas de milhares de cDNA serão depositadas em poços de 44 µm de diâmetro em uma PicoTiterPlate e seqüenciados por síntese, em que a incorporação sucessiva de nucleotídeos será visualizada usando quimioluminescência.

## 4.5. Processamento de sequência e anotação

### 4.5.1. Pré-processamento

As etapas de processamento de sequências foram realizadas utilizando o software *est2assembly* (Papanicolaou *et al.*, 2009). Os arquivos das sequências em formato SFF foram convertidos para FASTA/QUAL (sequências reconhecidas por Java Scripts). Para a limpeza, sequências ligadas aos adaptadores foram mascaradas com SSAHA2 (Ning *et al.*, 2001), e as sequências indesejáveis identificadas (tais como o DNA mitocondrial, rDNA ou contaminantes) foram mascaradas utilizando BLAST2 (Altschul *et al.*, 1990). Sequências repetitivas foram identificadas e mascaradas utilizando o RepeatMasker versão open-3.3.0 (<http://repeatmasker.org>).

### 4.5.2. Montagem das sequências

A cobertura K-mer foi classificada utilizando o software Neson (Victorian Bioinformatics Consortium) e grupos de sequências (alta e baixa cobertura de K-mer) foram montados usando WcdEST versão 0.6.3 (Hazelhurst *et al.*, 2008). A montagem das sequências em contigs e singletons foi realizada utilizando o programa de montagem MIRA versão 3.1 (Chevreux *et al.*, 2004), com parâmetros de montagem específicos da tecnologia 454. Redundâncias em sequências montadas originárias de ambos os grupos K-mer foram inspecionadas usando BLAST2 (opção megablast).

A fim de avaliar a precisão na montagem das sequências montadas *de novo* para cada genótipo, foi realizado um alinhamento contra todos os modelos de genes anotados no genoma completo de *M. acuminata* ssp. *malaccensis* var. *Pahang* (DH Pahang) (baixado em <http://banana-genome.cirad.fr>), utilizando a ferramenta de mapeamento genômico e alinhamento GMAP (Wu & Watanabe, 2005). Critérios de alinhamento compostos por um mínimo de 90% da sequência em estudo mapeada no genoma referência e a identidade percentual da sequência em relação ao segmento mapeado superior a 95%. Validação de contigs unigenes que não mapearam modelos gênicos da referência foi realizado por meio da análise por TBLASTX contra o banco de dados GenBank (outros) (<http://www.ncbi.nlm.nih.gov>)

### 4.5.3. Anotação das sequências

A anotação funcional de proteínas/genes conhecidos para montagem de sequências únicas de cada grupo K-mer foi conduzida contra o banco de dados NCBI non-redundant protein (nr) (<http://www.ncbi.nlm.nih.gov>), usando o algoritmo BLASTX, with a typical E-value cut-off of  $10^{-5}$ . A anotação de unigenes sem similaridade significativa com as sequências dos bancos de dados foi posteriormente analisada para domínios protéicos com uma função celular específica, utilizando o software InterproScan version 4.8 (Zdobnov & Apweiler, 2001). A atribuição da categoria funcional para cada sequência única foi realizada usando o programa Blast2GO, classificando de acordo com termos GO dentro de funções moleculares, processos biológicos e componentes celulares (Conesa *et al.*, 2005; Gene Ontology Consortium, 2008).

Uma análise genômica comparativa entre o conjunto de dados de unigenes de Calcutta 4 e Cavendish Grande Naine foi realizada através da investigação de modelos gênicos comuns e distintos no genoma de referência de *Musa* para que os unigenes montados *de novo* fossem mapeados.

Para a identificação de potenciais transcritos patogênicos de fungos entre os unigenes diferencialmente expressos *in silico* derivados de folhas infectadas, cada unigene foi também alinhado usando BLASTN (opção megablast) contra o banco de dados *Mycosphaerella fijiensis* v2.0 All Gene Models (transcripts) usando BWA (Li & Durbin, 2009) com os parâmetros predefinidos. Contagens de acordo com as leituras mapeadas para modelos de genes de *M. fijiensis* foram calculadas usando GenomicRanges (<http://www.bioconductor.org>).

## 4.6. Análise Proteômica

### 4.6.1. Extração e quantificação de proteínas

Foi conduzida uma extração de proteínas a partir do material foliar inoculado e não-inoculado (armazenados a  $-80^{\circ}\text{C}$ ). 10g de folha foram macerados utilizando nitrogênio líquido e 2g foram colocados em tubos separados. Para retirar os pigmentos das folhas, as amostras foram ressuspendidas em 25 mL de acetona 100% gelada, agitadas e centrifugadas a 14.000 rpm a  $4^{\circ}\text{C}$  por 5 minutos. O sobrenadante foi descartado, e o procedimento repetido 3

vezes. Adicionou-se 10 mL de 10%TCA (m/v) em acetona gelada nas amostras. Estas foram então agitadas e centrifugadas a 14.000 rpm a 4°C por 5 minutos. O sobrenadante foi descartado, e o procedimento repetido 3 vezes. Esta etapa é realizada a fim de retirar pigmentos e precipitar proteínas. As amostras foram lavadas 3 vezes com 10%TCA em água gelada. As mesmas foram agitadas e centrifugadas a 14.000 rpm a 4°C por 5 minutos, e o sobrenadante foi descartado. Esta etapa retira a acetona e precipita proteínas hidrossolúveis. O precipitado foi lavado, a fim de retirar o TCA, com 10 mL de acetona 80% gelada, centrifugada por 5 minutos a 14.000 rpm. Para a extração fenólica, o precipitado foi ressuspenso em 10 mL de fenol tamponado com Tris pH= 10,5 e 10 mL de tampão de solubilização (30% Sacarose (m/v), 2% SDS (m/v), 0,1M Tris – Hcl pH=8,0 e 5% de  $\beta$ -mercaptanol (v/v)). Este tampão de solubilização auxilia na separação das fases na extração com fenol. O fenol solubiliza proteínas, incluindo de membrana, deixando substâncias solúveis em água (DNA, carboidratos) na fase aquosa. A mistura foi agitada e centrifugada a 14.000 rpm a 4°C por 5 minutos. Este procedimento foi repetido 2 vezes. A fase fenólica foi coletada, e adicionou-se a ela 3,7 volumes de 0,1M de acetato de amônio em metanol. As amostras foram acondicionadas em refrigerador -20°C até o dia seguinte. As amostras então foram centrifugadas a 14.000 rpm a 4°C por 5 minutos, descartando o sobrenadante. O precipitado foi lavado 2 vezes com 10 mL de acetato de amônio 0,1M em metanol e 2 vezes com 10 mL de acetona 100% gelada. Após descartar o sobrenadante, os tubos foram deixados abertos a temperatura ambiente até a evaporação total da acetona. As amostras foram estocadas em refrigerador -20°C. O material foi ressuspenso em 7M uréia, 2M tiouréia, 4% CHAPS, 1% IPG buffer ph 3-10, 1% DTT, Coquetel Inibidor de Protease (5ul/mg de tecido). A concentração total de proteína das amostras foi quantificada pelo método de Lowry, utilizando albumina de soro bovino (BSA) como padrão. Foi utilizado também utilizando o método Bradford, que é baseado na adsorção do reagente Coomassie Brilliant Blue G-250 (Bradford, 1976) (Tabela 3). Além disso, as amostras foram submetidas a um mini-gel de poliacrilamida para validar as quantificações realizadas, aplicando 50  $\mu$ g e 100  $\mu$ g de cada um dos quarto pools. (Figura 16).

## **4.6.2. Eletroforese bi-dimensional em gel de poliacrilamida (2DE-PAGE)**

### **4.6.2.1. Focalização Isoelétrica (IEF)**

As amostras foram submetidas à eletroforese bidimensional. Foram formados quatro *pools* (Calcutta 4 controle e tratado; Cavendish Grande Naine controle e tratado) a partir da mistura das replicas biológicas de cada amostra, retirada 9 dias após a infecção pelo patógeno. A re-hidratação da fita IPG para a focalização isoeletrica foi realizada durante 16 h em tampão contendo uréia 8 M, CHAPS 2% (m/v), azul de bromofenol (traços) e tampão IPG a 2% (v/v) (Gorg et al., 2009). A reidratação ocorreu por 16 horas.

A focalização isoeletrica foi realizada utilizando uma fita (*strip*) de gradiente entre 3-10 (GE Healthcare). Para efetivação do experimento, foi empregado um aparato para eletroforese bidimensional Ettan™ IPGphor™ 3 (GE Healthcare), para o acúmulo de 51.000 Volts.hora<sup>-1</sup>. Todas as corridas foram realizadas a 18°C. Foi adicionado óleo mineral sobre as tiras para evitar evaporação. Após a focalização isoeletrica, as tiras foram acondicionadas individualmente em tubos de vidro sem nenhuma solução, a -80°C, até sua utilização para a segunda dimensão.

### **4.6.2.2. Separação em segunda dimensão: SDS-PAGE**

Para a segunda dimensão, foi realizada uma SDS-PAGE utilizando um gel a 12% no sistema Ettan™ DALTsix (GE Healthcare). As tiras foram retiradas do *freezer* (-80°C) e foram equilibradas por quinze minutos com solução contendo Tris-HCl 50 mM (pH 8,8), uréia 6 M, glicerol 30%, SDS 2% e azul de bromofenol (traços) e 0,01% de ditioneitol (DTT)/e por quinze minutos contendo iodoacetoamida, sob agitação à temperatura ambiente. Posteriormente, foram colocadas sobre a superfície do gel de acrilamida (15cmx14cmx1mm) e fixadas com uma solução de agarose. As condições de corrida da eletroforese foram as seguintes: 1) 10mA (15') e 2) 20mA (5-7h50min). Os géis obtidos foram corados com solução de Coomassie Blue G-250. Foi utilizado o marcados molecular (MM) Fermentas® em todos os géis.

Quatro replicatas de cada amostra foram utilizadas para assegurar a reprodutibilidade do método, gerando quatro géis para cada condição.

#### 4.6.2.3. Análise de imagem dos géis bidimensionais

As quadruplicatas dos géis foram escaneados usando um scanner ImageScanner II, previamente calibrado de acordo com manual do fabricante (GE Healthcare). Um total de dezesseis géis foram selecionados e submetidos a análise *in silico* pelo *software* ImageMaster™ 2D Platinum v 7.0 (GE Healthcare Life Sciences). A detecção dos pontos proteicos foi realizada manualmente, circulando-se os *spots* e fornecendo-se o direcionamento para o *software* de massa e pI.

Posteriormente, as imagens foram intercruzadas para análise entre os *spots* proteicos. A análise das réplicas foi realizada utilizando-se alguns *spots* como ponto de referência, denominado “landmark”, bem como os marcadores de massa molecular. Análises de correlação estatística (R) foram aplicadas para avaliar a diferença entre as triplicatas dos géis. *Cut off* de correlação foi aplicado e os géis com  $R^2$  menor que 0,90 foram descartados. O volume relativo dos *spots* proteicos (dpi) foi usado para comparação de expressão proteica. Os *spots* comuns ao mesmo genótipo nos dois tratamentos realizados presentes em todos os 8 géis, e os *spots* exclusivos para cada um das quatro condições (genótipo resistente não inoculado, genótipo resistente inoculado, genótipo suscetível não inoculado, genótipo suscetível inoculado) presentes nos 4 géis, foram selecionados para uma futura etapa de identificação por espectrometria de massa. Além disso, foram identificados *spots* em comum aos tratamentos realizados com valor de diferença de expressão maior que 1,5 quando comparado com o controle.

Toda a análise proteômica deste estudo foi conduzida em colaboração com a Universidade Católica de Brasília (UCB).

## 5. RESULTADOS E DISCUSSÃO

### 5.1. Análise do transcrito

Esta parte do trabalho buscou gerar dados do transcrito de *M. acuminata* que inclui genes expressos na interação banana-*M.musicola*. utilizando os genótipos suscetíveis (Cavendish Grande Naine) e resistentes (Calcutta 4).

O genótipo Calcutta 4 é um diplóide selvagem fértil amplamente utilizado em programas de melhoramento de cultivares comerciais (por exemplo, Uma *et al.*, 2011). É considerado uma fonte importante de resistência a importantes fungos e nematoides, atuando como um doador de R-genes. Dada essa importância, tem sido adotado como um modelo para a genômica comparativa em arroz (Cheung & Town, 2007; Lescot *et al.*, 2008), com aplicações em genômica funcional (Santos *et al.*, 2005) e descoberta de candidatos a genes de resistência (Azhar & Heslop-Harrison, 2008; Miller *et al.*, 2008). Bananas do subgrupo Cavendish, tais como Grande Naine, pelo contrário, são triplóides estéreis, que, apesar de representar mais de 40% da produção mundial, não apresenta resistência a estresses bióticos, de tal modo que a aplicação de pesticidas regularmente é necessária para a produção comercial.

#### 5.1.1. Estatística e Montagem – 454

Cada biblioteca de cDNA foi sequenciada em um segmento de 1/4 de uma única placa de execução, gerando 978.133 *reads* para os dois genótipos, totalizando mais de 466 megabases de dados de sequências. A Tabela 2 e o Figura 11 apresentam um resumo da distribuição do tamanho dos *reads* nos dois genótipos estudados. Para ambos os conjuntos de dados, o comprimento mais comum dos *reads* ficou entre 301-500 pb.

Sequências de Calcutta 4 foram montadas *de novo* em 36.384 unigenes, incluindo 24.259 contigs e 12.125 singletons. Destes, um total de 25.381 unigenes foram representados no conjunto de dados do transcrito de folhas não-infectadas e 25.154 no conjunto de dados a partir de folhas desafiadas com o patógeno. No caso de sequências de Cavendish Grande Naine, foi realizada uma montagem semelhante, com um total de 35.269 unigenes, compostos de 23.729 contigs e 11.540 singletons. 18.611 unigenes foram representados no conjunto de

dados do transcrito de folhas não-infectadas e 29.223 no conjunto de dados a partir de folhas de infectadas pelo patógeno.

Singletons apresentaram um comprimento suficiente para permitir a anotação, com um tamanho médio de 332 pb em Calcutá 4 e 336 pb em Cavendish Grande Naine. O tamanho dos contigs variou de 150 a 3728 pb para Calcutta 4, com um tamanho médio de 552 pb. As estatísticas foram semelhantes para a montagem dos contigs de Cavendish Grande Naine, que apresentaram comprimentos que variaram de 151-4150 pb, e um tamanho médio de 548 pb. Distribuições de número de reads em cada unigene montado foram semelhantes em ambos genótipos (Figura 12A e 12B).

Tabela 2 - Distribuição por tamanho de sequências de alta qualidade de *M. acuminata* Calcutta 4 e Cavendish Grande Naine, derivadas do sequenciamento massal.

<b>M. acuminata Calcutta 4</b>				
	Min	Mean	Median	Max
Comprimento do EST (pb)	40	332	353	908
Distribuição do EST pelo comprimento (pb)	40 – 100	20472		
	101 – 200	39009		
	201 – 300	78901		
	301 – 400	152325		
	401 – 500	129795		
	>500	3737		
<b>M. acuminata Cavendish Grande Naine</b>				
	Min	Mean	Median	Max
Comprimento do EST (pb)	40	336	362	1161
Distribuição do EST pelo comprimento (pb)	40 – 100	22665		
	101 – 200	38635		
	201 – 300	71477		
	301 – 400	140099		
	401 – 500	143815		
	>500	5832		



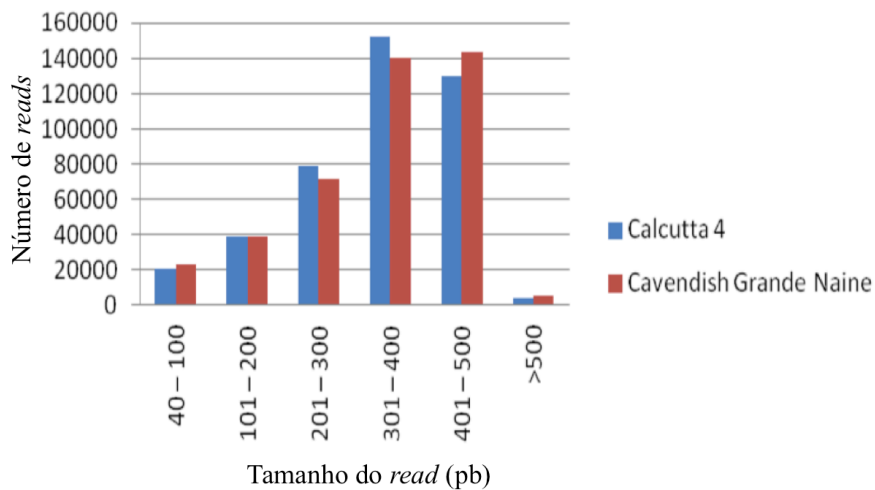


Figura 11 - Distribuição conforme o tamanho dos reads de *M. acuminata* Calcutta 4 e Cavendish Grande Naine.

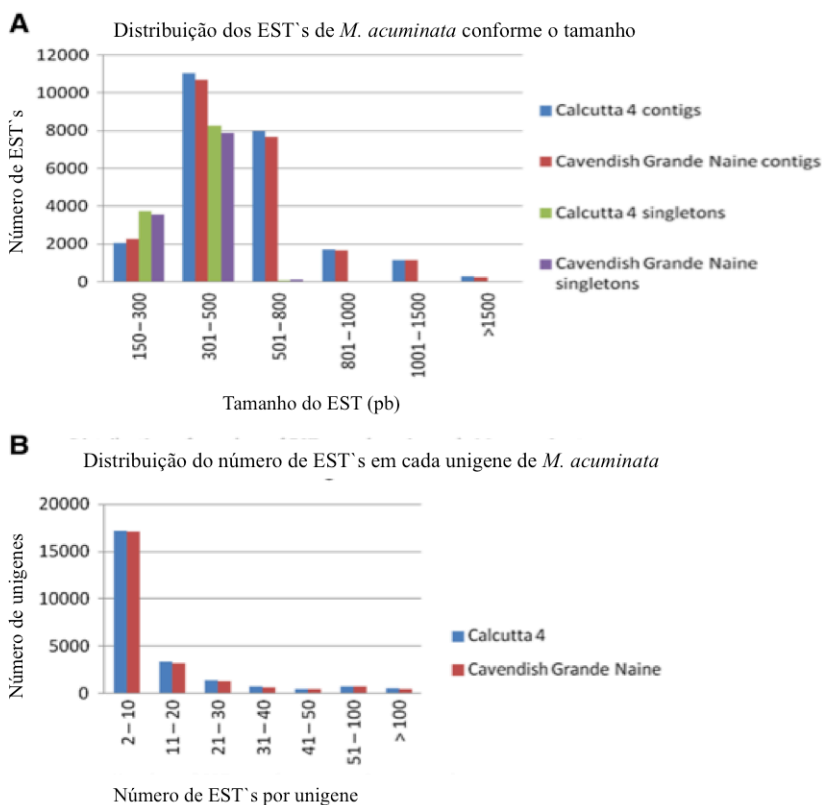
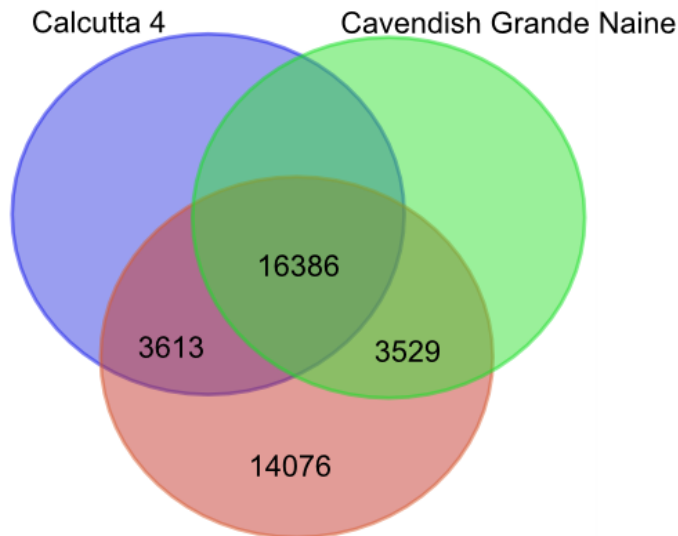


Figura 12 - Resumo da montagem das sequências de *M. acuminata* Calcutta 4 e Cavendish Grande Naine. **A)** Distribuição dos contigs e singletons de *M. acuminata* conforme o tamanho. **B)** Distribuição do número de EST's de *M. acuminata* por unigene.

A cobertura do genoma por genótipo correspondem bem com a mais recente estimativa de 36,542 modelos gênicos codificadores de proteínas na seqüência dos genomas de referência *M. acuminata* ssp. *malaccensis* var. *Pahang* (DH Pahang) (D'Hont *et al.*, 2012) e 36.638 de *M. balbisiana* variedade 'Pisang Klutuk Wulung' (PKW) (Davey *et al.*, *in press*), o que pode refletir uma análise rigorosa da qualidade e a utilização de parâmetros de montagem. Deve ser reconhecido, entretanto, que pode ocorrer uma superestimativa do números de genes como resultado de membros não sobrepostas da seqüência de leitura atuais por um único gene. O alinhamento dos contigs contra o genoma de referência revelou que 24.097 dos 24.259 contigs de Calcutta 4 (99,3%) mapearam nos 16.519 modelos gênicos, e de forma semelhante, 23.548 dos 23.729 contigs de Cavendish Grande Naine (99,2%) mapearam nos 16.402 modelos gênicos do genoma referência. Os altos percentuais de unigenes mapeados nos exons dos modelos gênicos valida tanto a montagem *de novo* dos contigs, quanto a anotação gênica do genoma de referência de *M. acuminata*, sobrepondo informações importantes em relação expressão do hospedeiro durante esta interação planta-patógeno (Figura 14). Através da análise por TBLASTX de contigs não mapeados, *hits* positivos para genes identificados no banco de dados NCBI EST (outros) revelou 162 novos potenciais unigenes em Calcutta 4 e 181 em Cavendish Grande Naine. Embora estas seqüências não alinhadas possam ser específicas para os genótipos de *M. acuminata*, estes dados podem indicar também outros genes que requerem uma anotação mais curada no genoma de referência. A fim de se obter uma idéia sobre grandes semelhanças e diferenças entre os conjuntos de dados do transcrito (contigs e singletons) para os dois genótipos de *M. acuminata* avaliados neste estudo, todos os modelos gênicos no genoma de referência foram usados como uma base para a identificação de genes comuns e distintos. Um diagrama de Venn ilustra a sobreposição entre os genótipos, com 16.386 modelos de genes comuns identificados (Figura 13). Este número representa 82,0 % dos modelos gênicos mapeados para cada genótipo. Embora o mapeamento específico para cada genótipo possa ser atribuído a diferentes distâncias evolutivas de *M. acuminata* ssp. *malaccensis* var. *Pahang*, a sobreposição entre os conjuntos de dados pode estar ligada com a cobertura das seqüências do 454.

Todos os dados das seqüências do presente estudo estão disponíveis para cada genótipo no "Sequence Read Archive" (SRA) do "National Center for Biotechnology Information" (NCBI) (submissão SRA055816).



### *M. acuminata* ssp. *malaccensis* var. Pahang (DH Pahang)

Figura 13 - Diagrama de Venn mostrando a sobreposição entre os conjuntos de dados do transcrito (contigs e singletons) para os genótipos de *M. acuminata* Calcutta 4 Cavendish e Grande Naine. Todos os modelos de genes de referência do genoma de *M. acuminata* DH Pahang foram usados como uma base para a identificação de genes comuns mapeados.

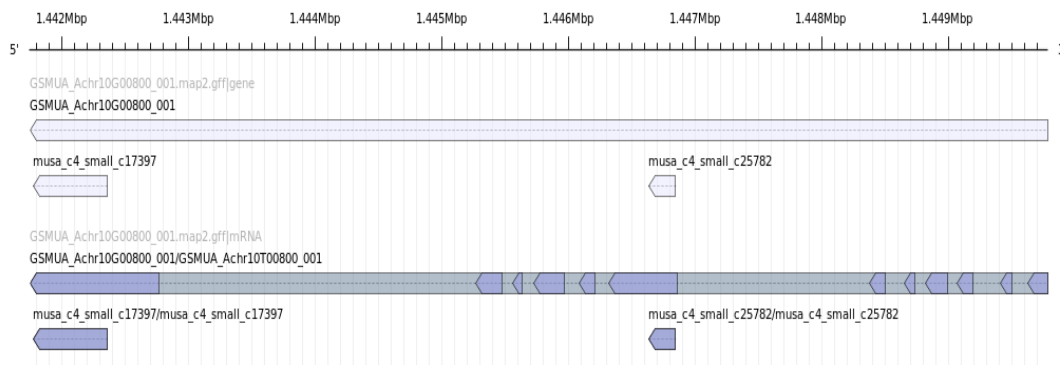


Figura 14 - Exemplo de um contig de *M. acuminata* Calcutta 4 mapeando em um único modelo gênico de *M. acuminata* DH Pahang

### 5.1.2. Anotação funcional e classificação

A anotação de sequências de unigenes montados foi realizada por meio da busca por similaridade de sequência com o banco de dados de sequências protéicas não redundantes do NCBI (nr). Os critérios do BLASTX adotados foram: comprimento do alinhamento deve ser maior do que 100 aminoácidos e  $E\text{-value}=10^{-5}$ . Do total de unigenes estimados para Calcutta 4, 10.080 apresentaram identidade significativa para genes que codificam proteínas com função conhecida, 1.633 para genes que codificam proteínas com função desconhecida, e 13.513 mostraram nenhuma identidade significativa para quaisquer sequência do banco de dados. Resultados semelhantes foram observados para Cavendish Grande Naine, com 10.645

unigenes mostrando identidade significativa para genes que codificam proteínas com função conhecida, 1.800 para genes que codificam proteínas com função desconhecida, e 11.971 mostrando nenhuma identidade significativa a quaisquer sequências do banco de dados.

A ferramenta de busca de domínios protéicos InterProScan (<http://www.ebi.ac.uk/Tools/InterProScan/>) também foi utilizada. Um total de 14.826 unigenes de Calcutta 4 continham domínios InterPro, com 3.949 domínios distintos representados no conjunto unigene. Das 13.513 sequências traduzidas sem “Blast hits” significantes, 192 mostraram domínios proteicos funcionais. Da mesma forma, 14.006 unigenes de Cavendish Grande Naine possuíam domínios Interpro, com um total de 3.974 domínios Interpro representados. Das 11.971 sequências traduzidas sem “Blast hits” significantes, 223 mostraram domínios proteicos funcionais.

Taxas de novidade baseadas nas análises de Blast e Interpro foi de 36,6%, no caso de Calcutta 4 e 33,3% para Cavendish Grande Naine.

### **5.1.3. Distribuição de espécies**

Sequências de unigenes com similaridade significativa à proteínas de plantas de bancos de dados públicos foram abundantes. Respectivamente, unigenes para Calcutta 4 e Cavendish Grande Naine, totalizam 8.518 e 8.948 genes encontrados em arroz (*Oryza sativa*), 5.718 e 6.078 genes encontrados em milho (*Zea mays*), 6.626 e 7.026 genes encontrados em sorgo (*Sorghum bicolor*), 9060 e 9579 genes correspondentes em uva (*Vitis vinifera*), 3771 e 3927 genes correspondentes em *Arabidopsis thaliana* (Figura 15). Distribuições semelhantes foram observados para famílias gênicas de *M. acuminata* DH-Pahang (D'Hont *et al.*, 2012). Para o conjunto de dados de Calcutta 4 e Cavendish Grande Naine, apenas 2,8% de “Blast hits” de seqüências homólogas de plantas estão publicamente disponíveis no banco de dados de proteínas nr *Musa* NCBI, indicando que nossos dados representam uma contribuição considerável para unigenes expressos para o gênero.

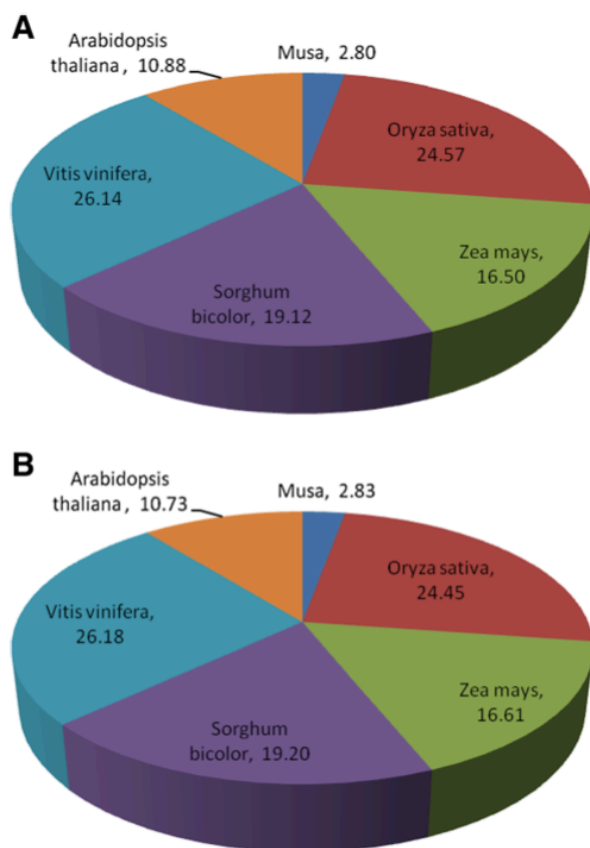


Figura 15 - Distribuição das espécies dos unigenes de *M. acuminata* mostrados em porcentagem do total das seqüências homólogas de plantas monocotiledôneas. Os melhores hits do Blast de cada seqüência foram analisados. A) unigenes de *M. acuminata* Calcutta 4; B) unigenes de *M. acuminata* Cavendish Grande Naine.

#### 5.1.4. GO categorias

Uma classificação funcional de conjuntos de unigenes foi realizada por meio do Gene Ontology (GO), utilizando Blast2GO. Os dados foram categorizados para cada genótipo nas três categorias principais de GO: processos biológicos, componente celular e função molecular. Unigenes de Calcutta 4 foram associados a um total de 341.244 anotações de GO, com 162.468 termos relacionados a processos biológicos, que representam 10 níveis de GO, 69.390 termos ligados a função molecular de 10 níveis e 109.386 termos de componentes celulares de 9 níveis. Atribuições semelhantes foram observados nos dados dos unigenes de Cavendish Grande Naine, com 351.220 termos GO distribuídos entre as três categorias principais, com um total de 168.015 termos de processos biológicos, 70.878 termos de função molecular e 112.327 termos de componentes celulares. Dentro da categoria de processo biológico, a maioria dos unigenes foram atribuídos a "processo celular" (5530 para C4, 5649 para CAV), "processo metabólico" (5.276 para C4, 5418 para CAV), "regulação biológica"

(1397 para C4, 1377 para CAV), "resposta ao estímulo" (1343 para C4, 1383 para CAV), e "localização" (1179 para C4, 1227 para CAV). Para a função molecular, os termos "ligação" (4732 para o C4, 4.863 para CAV), "atividade catalítica" (4659 para o C4, 4.771 para CAV) e "atividade transportador" (678 para o C4, 708 para CAV) foram os mais abundantes termos atribuídos para os conjuntos de dados de unigenes de ambos os genótipos. Em toda a categoria a função celular, os termos mais abundantes foram "componente celular" (7620 para o C4, 7.759 para CAV), "célula" (7.580 para o C4, 7.725 para CAV), "organela" (5589 para o C4, 5.735 para CAV) e "complexo macromolecular" (1.772 para o C4, 1.829 para CAV). Essas informações estão resumidas nas Figuras 16 e 17. Após essas análises, contigs relacionados com defesa da planta foram identificados, identificando-se também o número de leituras presentes em cada um dos tratamentos em ambos os cultivares. Esses dados podem ser visualizados nos Anexos 15 e 16.

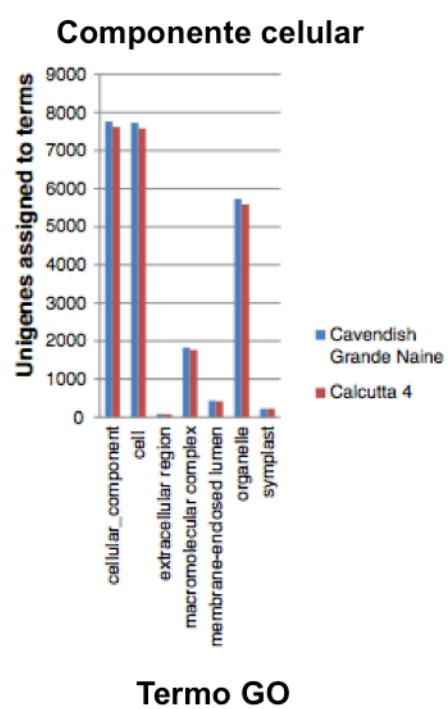
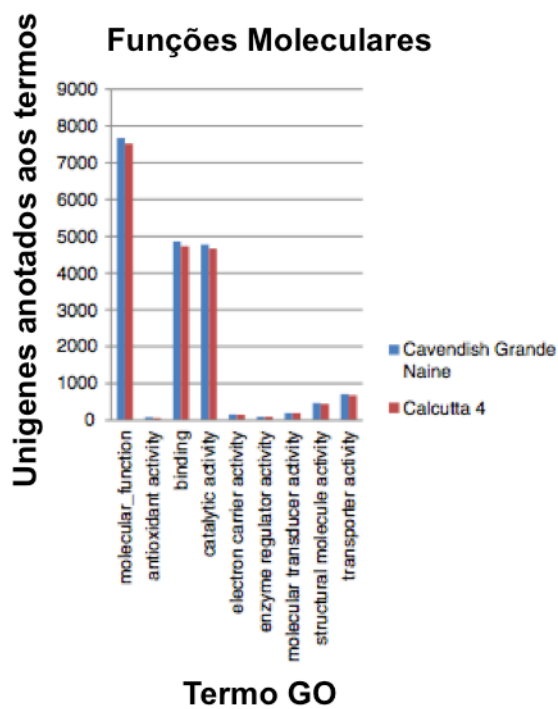
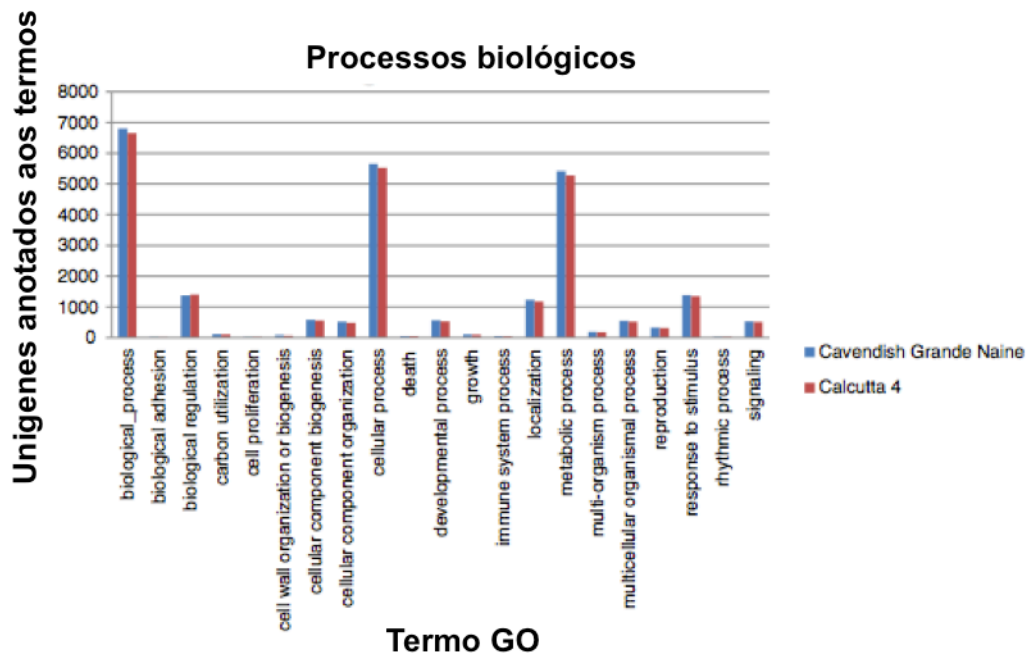
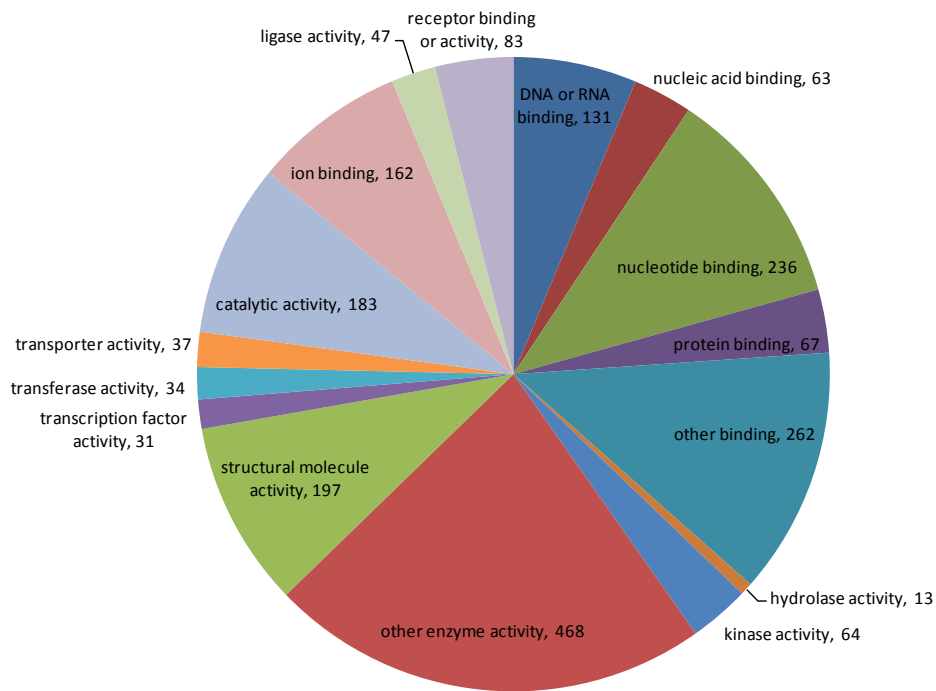
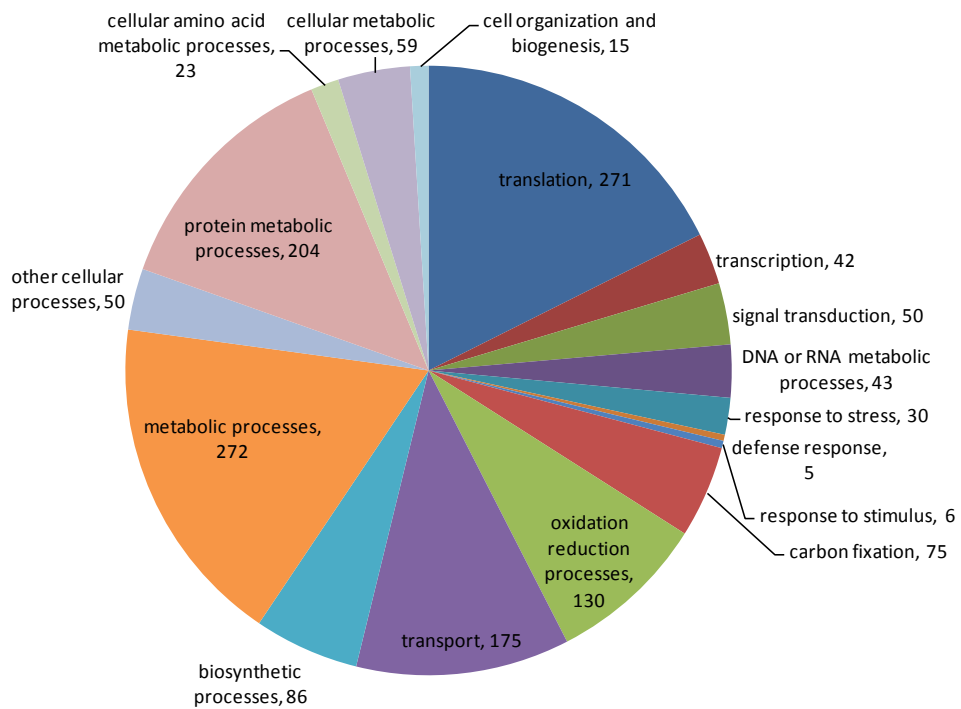


Figura 16 - Apresentação do histograma do *Gene Ontology* (GO) após o processo de anotação dos unigenes de *M. acuminata* Calcutta 4 e Cavendish Grande Naine. Os dados do GO são resumidos em três categorias principais: processo biológico, função molecular e componente celular.

**(a) Função Molecular**



**(b) Processos biológicos**





### (c) Componente celular

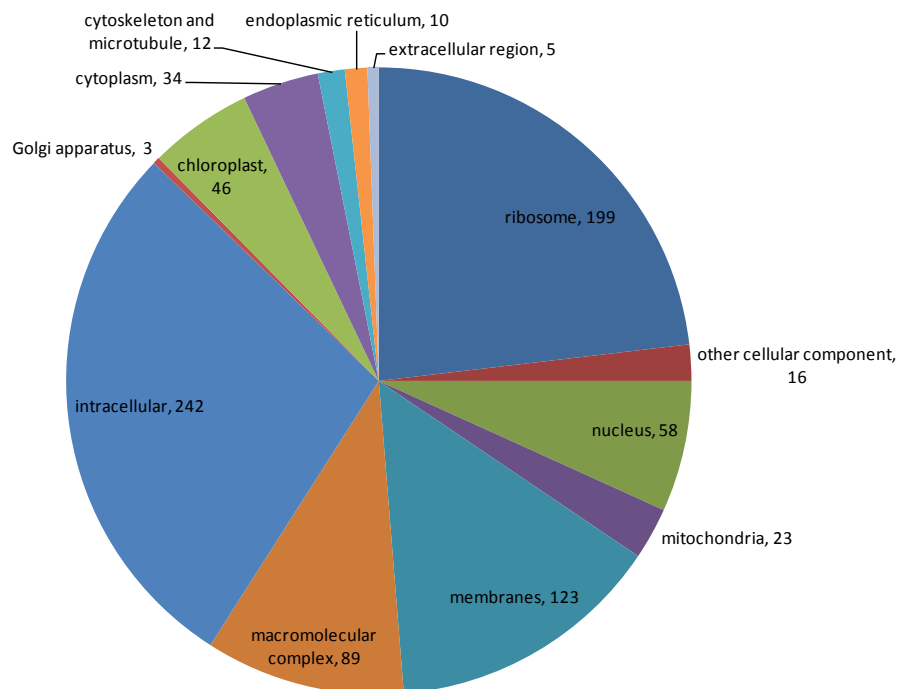


Figura 17 - Representação dos unigenes de *M. acuminata* classificados em grupos funcionais. Os hits GO foram atribuídos nas categorias (a) função molecular, (b) processos biológicos e (c) componente celular.

*M. musicola* é um agente patogênico hemibiotrófico que penetra através dos poros dos tecidos foliares, após um período de crescimento epifítico na superfície da folha. Uma vez dentro do hospedeiro, o patógeno coloniza o espaço intercelular dentro de camadas de tecido do mesófilo e do parênquima paliádico, sem formar haustórios ou infectar células hospedeiras. Esta fase biotrófica pode durar por algumas semanas antes do início do aparecimento de lesões necróticas nas células mesofílicas. O crescimento de conidióforos através de estômatos, então, permite o desenvolvimento do conidiósporo na superfície da folha. Diferenças notáveis no desenvolvimento da doença são observados em genótipos totalmente resistentes, com necrose precoce de células guarda estomática e morte de um número limitado de células hospedeiras no local da infecção. Não se observa esporulação de fungos em tais interações incompatíveis.

A análise da expressão de genes durante a interação banana-*Mycosphaerella* tem sido limitada até o momento (D'Hont *et al.*, 2012; Portal *et al.*, 2011; Passos *et al.*, 2012). Este estudo é o primeiro que visa a análise massal do transcrito no patossistema *M. acuminata* - *M. musicola*. A estratégia de sequenciamento para a identificação *in silico* da expressão

diferencial de transcritos a partir de organismos patogênicos e folhas desafiadas foi empregada, com a preparação de duas bibliotecas de cDNA por genótipo (infectados e não infectados). Estudos indicam que durante a fase inicial da infecção em banana com *M. fijiensis*, tubos germinativos penetram nos estômatos do 3<sup>o</sup> ao 6<sup>o</sup> dia (Beveraggi *et al.*, 1995). Neste trabalho, o material foliar foi coletado no 3<sup>o</sup>, 6<sup>o</sup>, e 9<sup>o</sup> dia após desafiado.

O uso de tecidos foliares destacados das plantas em bioensaios utilizando o patógeno de banana *Mycosphaerella*, foi relatado, porém apresentando resultados inconsistentes (por exemplo, Arzanlou *et al.*, 2007; Donzelli & Churchill, 2007), onde o desenvolvimento dos sintomas da doença nem sempre está relacionado com os observados em plantas intactas. Além disso, Arzanlou *et al.* (2007) sugeriu que os agentes patogênicos de banana exigem bananeiras saudáveis para o desenvolvimento da doença. Além disso, Churchill (2011), sugere que o estado fisiológico de folhas isoladas não são comparáveis com os de plantas inteiras que apresentam um sistema radicular intacto. Por estas razões, os bioensaios foram realizados usando folhas jovens intactas de plantas de seis meses de idade, com temperatura e umidade ideais.

Com base nos resultados da pesquisa de BlastX contra nr, anotação do Blast2GO e atribuição dos unigenes às categorias GO relacionadas com a defesa muitos unigenes potencialmente envolvidos na “imunidade ativada por efetores” (ETI) e na “imunidade ativada por PAMPs” (PTI) foram identificados em ambos os genótipos, além de unigenes do patógeno *M. musicola* (Anexo 14).

#### **5.1.5. Expressão gênica em Calcutta 4**

Numerosos fatores de transcrição foram up regulados. Envolvidos na regulação do desenvolvimento da planta, sinalização e resposta ao ambiente, entre outros papéis, muitos fatores de transcrição são conhecidos por estar envolvidos na transdução de sinal e regulação da expressão gênica em resposta ao estresse. Dentre os exemplos up-regulados observados em tecidos de folhas infectadas de Calcutta 4, incluiu-se uma proteína de domínio NAC e “ethylene insensitive-like protein 4”, um provável factor de transcrição que atua como um regulador positivo na via de resposta ao etileno.

O reconhecimento de plantas com resistência mediada por R-genes a um produto do gene patogênico está associada com a resposta de hipersensibilidade (HR), o que resulta em morte celular programada no local da infecção para limitar a disseminação do patógeno

(Morel & Dangl, 1997). Em plantas Calcutta 4 inoculadas, genes up-regulados, pode-se verificar a presença de metalotioneínas, que estão potencialmente envolvidas no processo de desintoxicação das plantas. Estes polipeptídeos de baixo peso molecular sequestram íons metálicos e estão associados com a regulação do potencial redox intracelular e com desintoxicação de oxigênio (Choi *et al.*, 1996; Hamer, 1986), que protege as células dos efeitos prejudiciais das espécies reativas de oxigênio (ROS). A geração de ROS pode indicar a atividade da Reação de Hipersensibilidade (RH), após infecção e reconhecimento. Quatro tipos distintos (MT1 a MT4) são conhecidos em plantas, de acordo com a distribuição de resíduos de cisteína. Foi relatado o isolamento dos tipos MT2 e MT3 em banana, com expressão influenciada, em resposta ao etileno e a metais (Liu *et al.*, 2002). Estudo mais recente apontou sua abundância em *M. acuminata* Calcutta 4 (Santos *et al.*, 2005). Este presente trabalho corrobora essa idéia, com a presença de sequências de MT2 e MT3.

A enzima Superóxido dismutase (SODs) foi igualmente up-regulada. Como as metalotioneínas, estas enzimas também podem atuar como antioxidantes, protegendo os componentes de células vegetais da oxidação por ROS. A proteína Glutathione S-transferase também foi representada. Esta pode estar envolvida em vias de sinalização celular, bem como na desintoxicação de produtos do estresse oxidativo durante a RH. A sua expressão no estágio final da interação *M. fijiensis*-*M. acuminata* foi recentemente relatada (Portal *et al.*, 2011). Acúmulo de H<sub>2</sub>O<sub>2</sub> e atividade da peroxidase em Calcutta 4 até o 10<sup>o</sup> dia após interação com *M. fijiensis* foi encontrada (Cavalcante *et al.*, 2011). Nossas observações vão de encontro com essas informações, onde houve um aumento da expressão de genes envolvidos na desintoxicação de ROS e na RH durante a interação com *M. musicola*.

Fenilpropanóides estão envolvidos numa série de respostas de defesa em plantas, incluindo a biossíntese de compostos antimicrobianos, tais como fitoalexinas e moléculas envolvidas em sinalização. O aumento de expressão da fenilalanina amônia-liase (PAL) foi observada em Calcutta 4 desafiada com o patógeno. Esta enzima catalisa uma reação na via biossintética dos fenilpropanóides, com a desaminação da fenilalanina em ácido cinâmico, reduzindo a síntese de fitoalexinas, assim como a produção de ácido salicílico, uma molécula sinal envolvida na Resistência Sistêmica Adquirida (SAR). 4-cumarato-CoA-ligase, uma outra enzima importante na biossíntese dos flavonóides e isoflavonóides, também foi up-regulada.

### 5.1.6. Expressão gênica em Cavendish Grande Naine

Um total de 1.371 unigenes em Cavendish Grande Naine mostraram-se com expressão diferencial estatisticamente significativa entre as plantas não inoculadas e inoculadas. A anotação revelou uma abundância de transcritos com similaridade pobremente caracterizada no GenBank. No entanto, quando as descrições foram possíveis, um número de unigenes regulados positivamente em folhas infectadas foram identificados como potencialmente envolvidos na resposta das plantas. Estes incluem fatores de transcrição, tais como as metalotioneínas e superóxido dismutase (desintoxicação de plantas), “4-coumarate:coA ligase 2”, ácido cinâmico 4- hidroxilase e “isoflavone reductase-like protein” (via fenilpropanóide), proteína F-box relacionada a doença e calmodulina (resposta de defesa). Em contraste, um número maior de unigenes apresentaram-se regulados negativamente em folhas infectadas na interação compatível, quando comparados com os tecidos não infectados. Dentre estes estão incluídos o fator de transcrição WRKY 17 e MAP quinase BIMK1 (sinalizadoras de defesa), “putative callose synthase 1 catalytic subunit” (síntese de calose), xiloglucano endotransglicosilase (XET), endo-quitinases e putativos quitinases (degradação da parede celular dos fungos), “pathogenesis-related protein 1”, F-box, “wd40 domain protein” e “hypersensitive-induced response protein” (defesa da planta), glutathione S-transferase 1 (desintoxicação da planta), “type III polyketide synthase 4”, cinamato 4-hidroxilase (via fenilpropanóide), proteína quinase CTR1, “ethylene receptor-like protein”, “ethylene response factor 11” e “ethylene-responsive transcriptional coactivator” (sinalização), oxidases responsáveis pela “explosão respiratória” (sinalização de ROS, transdução de sinal e morte celular), taumatina (degradação das paredes celulares dos fungos), serina/treonina-quinase com LRRs, Ser/Thr quinase e proteínas de resistência NBS-LRR (genes receptores).

### 5.1.7. Receptores de genes do hospedeiro

Um número de diferentes famílias de R-genes são reconhecidas como envolvidas em ETI e PTI, com base em domínios protéicos e função celular. A mais abundante é a de proteínas receptoras citoplasmáticas com sítios de ligação a nucleotídeos e proteínas ricas em leucina (NBS-LRR) (Hammond-Kosack & Jones, 1997). Em arroz, a cerca de 400 genes NBS-LRR foram caracterizados, com 150 presentes no genoma de *Arabidopsis* (McHale *et al.*, 2006), e 89 identificados no genoma de *M. acuminata* DH-Pahang (D’Hont *et al.*, 2012). No presente

estudo, 14 genes expressos de NBS-LRR foram identificados através do Blast de ambos tecidos foliares infectados e não-infectados de Calcutta 4 e 25 em Cavendish Grande Naine. O mapeamento de contigs nos modelos gênicos contendo domínio NB-ARC no genoma de referência de *M. acuminata* identificou 38 contigs mapeados em 40 modelos gênicos no caso dos dados do transcrito de Calcutta 4, e 43 contigs de Cavendish Grande Naine mapeados em 40 modelos gênicos (Anexo 12).

Outras classes conhecidas de R-genes de plantas incluem LRRs extracelulares ancorados por domínios transmembranas (proteínas receptoras), LRRs extracelulares ligadas a domínios citoplasmáticos de serina/treonina-quinases, serina/treonina-quinases intracelulares, e as proteínas com um domínio em espiral (*coiled-coil*) ancoradas à membrana celular.

#### **5.1.8. GO categorias**

A atribuição a categorias GO relacionadas com a defesa forneceram um nível adicional de genes candidatos envolvidos na resposta de defesa ao hospedeiro, complementando a anotação feita por Blast-nr. Por exemplo, foram identificados interessantes unigenes regulados positivamente em tecidos foliares de Calcutta 4 inoculados com o patógeno. Estes incluem componentes da síntese de glucano, que são associados com a deposição de calose nas paredes de células hospedeiras durante a resposta de defesa, “Rpm1 interacting protein 4”, relatado na defesa de plantas envolvendo “Mac perforin domain-containing proteins”, que estão associadas com o ácido salicílico (SA) - via da morte celular programada, “brassinosteroid insensitive 1- associated receptor kinase 1”, a qual está envolvida na PTI e na morte celular programada, e “mlo-like protein 1”, conhecida por estar envolvida na defesa de plantas e fortalecimento da parede celular.

#### **5.1.9. Mineração de dados em relação a modelos de genes de *M. fijiensis***

Uma mineração de dados dos transcritos do patógeno *M. musicola* entre os dados das sequências do 454 pré-processadas derivadas de folhas infectadas foi realizada por meio de mapeamento e alinhamento no genoma usando o bando de dados *Mycosphaerella fijiensis* v2.0 All Gene Models (transcripts). Com base na análise da sequência ITS do rDNA, esta espécie forma um grupo monofilético com *M. musicola* ( Goodwin *et al.*, 2001). Embora a abundância de transcritos do patógeno seja dependente do sequenciamento 454, um total de

10 unigenes de transcritos em Calcutta 4 infectado com o patógeno e 12 transcritos em Cavendish Grande Naine mapearam em modelos de genes de *M. fijiensis*. Além de proteínas hipotéticas e *no hits* que foram identificadas na interação com C4 e CAV, potenciais genes patogênicos foram identificados nos dados do transcriptoma dados de CAVI, incluindo uma “*SAP family cell cycle dependent phosphatase-associated protein*”, duas proteínas da família Hsp70, uma proteína de domínio de ligação FAD, e um canal de cálcio. No caso dos dados de C4I, entre os transcritos mapeados positivamente nos transcritos de *M. fijiensis*, incluem uma proteína extracelular 6 (Ecp6), e duas proteínas ribossomais 60S. Um total de sete Ecps foram identificados em *Cladosporium fulvum* (syn. *Passalora fulva*), que são secretadas durante a interação com tomate (Bolton *et al.*, 2008). A proteína efetora Ecp6 de *C. fulvum*, crucial para a virulência, contém domínios Lisina de ligação a carboidratos (LysM) que podem estar envolvidos em ligação a quitina liberada das paredes celulares dos fungos durante o processo de infecção (de Jonge *et al.*, 2010). Tal ligação pode impedir a indução de respostas de defesa basal da planta. Ortólogos de Ecp6 foram identificados em sequências de EST e dados de genoma num grande número de gêneros de fungos. No gênero *Mycosphaerella*, LysM contendo proteínas Ecp6 semelhantes foram identificadas em *M. fijiensis* (Bolton *et al.*, 2008) e *Mycosphaerella graminicola* (Bolton *et al.*, 2008; Marshall *et al.*, 2011), indicando a sua provável presença em outros membros do gênero. Enquanto transcritos homólogos de efetores Ecp2 de *C. fulvum* foram descritos em *M. fijiensis* cultivados *in vitro* (Cho *et al.*, 2008), até o presente momento não houve relatos de expressão *in vivo* de um homólogo do efector Ecp para o organismo patogênicos da bananeira, *Mycosphaerella* sp., durante a interação patógeno-hospedeiro. Nossa identificação de transcritos homólogos a proteína efetora Ecp6 em *M. musicola* pode contribuir para a identificação de efetores do patógeno e potenciais genes de resistência a doenças em *Musa*.

## **5.2. Análise proteômica da interação Banana x *M. musicola***

Os resultados da quantificação realizada utilizando o método de Lowry não foram considerados. Tais resultados não foram considerados confiáveis entre as réplicas técnicas. As quantidades de amostra aplicadas no mini-gel de poliacrilamida (Figura 18) foram calculadas baseadas na quantificação feita pelo método de Bradford, utilizando uma curva de concentrações (0, 0.03, 0.06, 0.125, 0.3 e 0.5 mg/mL) das amostras de *M. acuminata* Calcutta 4 não inoculadas e inoculadas com o patógeno (C4NI e C4I) e Cavendish Grande Naine não

inoculadas e inoculadas com o patógeno (CAVNI e CAVI). A concentração final de C4NI foi de 1,12mg/mL, de C4I 1,35mg/mL, de CAVNI 1,18 mg/mL e de CAVI 1,13mg/mL.

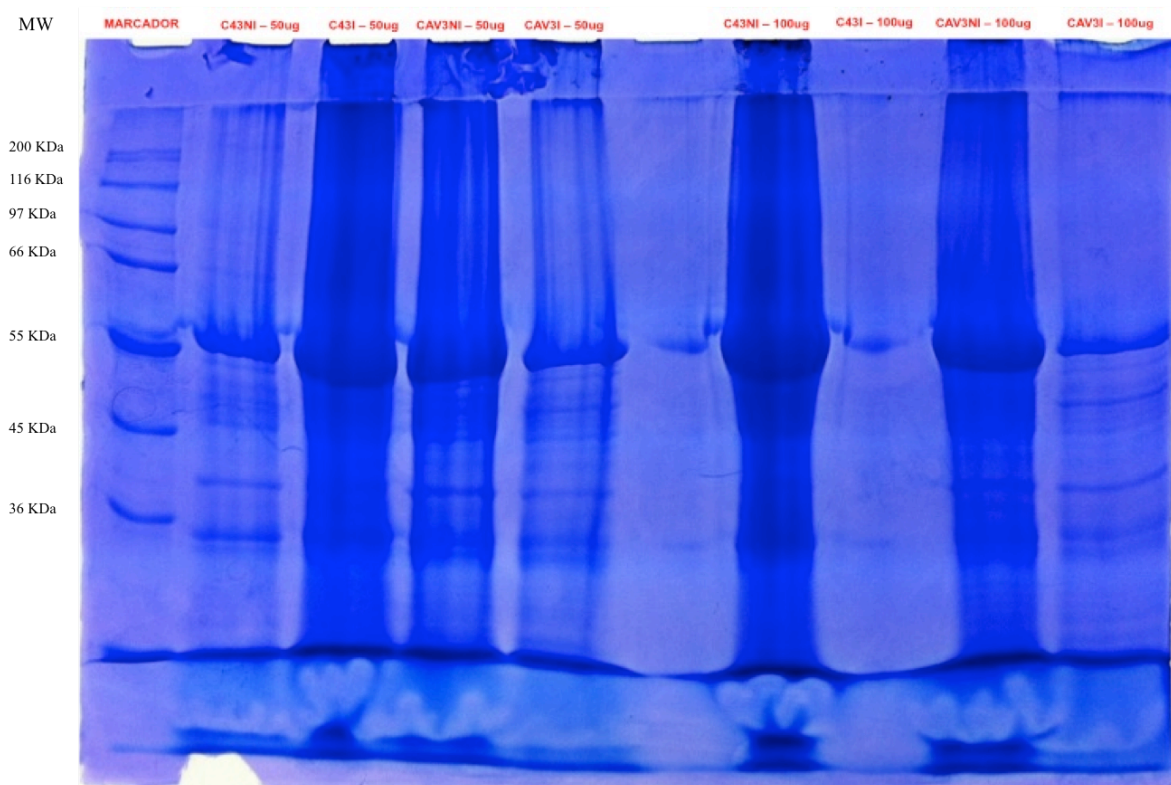
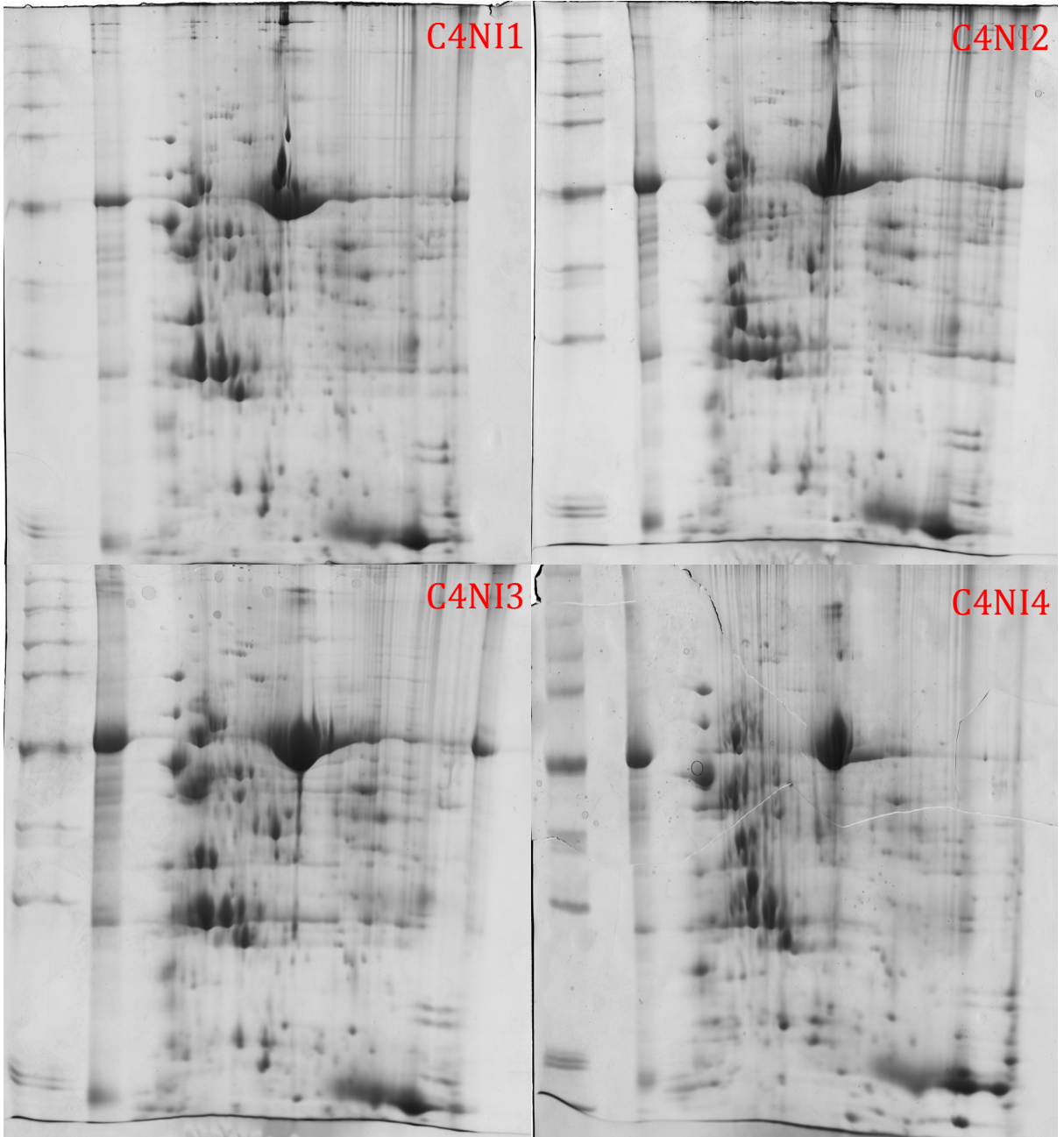


Figura 18 – Mini-gel de poliacrilamida SDS-PAGE 12% mostrando o padrão proteico de duas concentrações do extrato protéico de *M. acuminata*. Ordem de aplicação no gel: Marcador molecular. **C43NI** - *M. acuminata* Calcutta 4 não inoculada com o patógeno; **C43I** - *M. acuminata* Calcutta 4 inoculada com o patógeno; **CAV3NI** - *M. acuminata* Cavendish Grande Naine não inoculada com o patógeno; **CAV3I** - *M. acuminata* Cavendish Grande Naine inoculada com o patógeno. 50 e 100  $\mu\text{g}\cdot\text{mL}^{-1}$ . Concentração do extrato proteico utilizando. O gel foi corado com azul de comassie G-250.

Baseado na quantificação realizada, optou-se por fazer um gel bidimensional aplicando 250 $\mu\text{g}$  de cada pool (224 $\mu\text{l}$  de C4NI, 92.5 $\mu\text{l}$  de C4I, 105.75 $\mu\text{l}$  de CAVNI, 222 $\mu\text{l}$  de CAVI). A concentração de 250 $\mu\text{g}$  foi uma estimativa. Essa dificuldade na quantificação pode ter acontecido devido ao tampão de ressuspensão utilizado nas amostras. Os componentes deste tampão podem ter interferido nos métodos de quantificação, assim como na corrida do mini-gel.

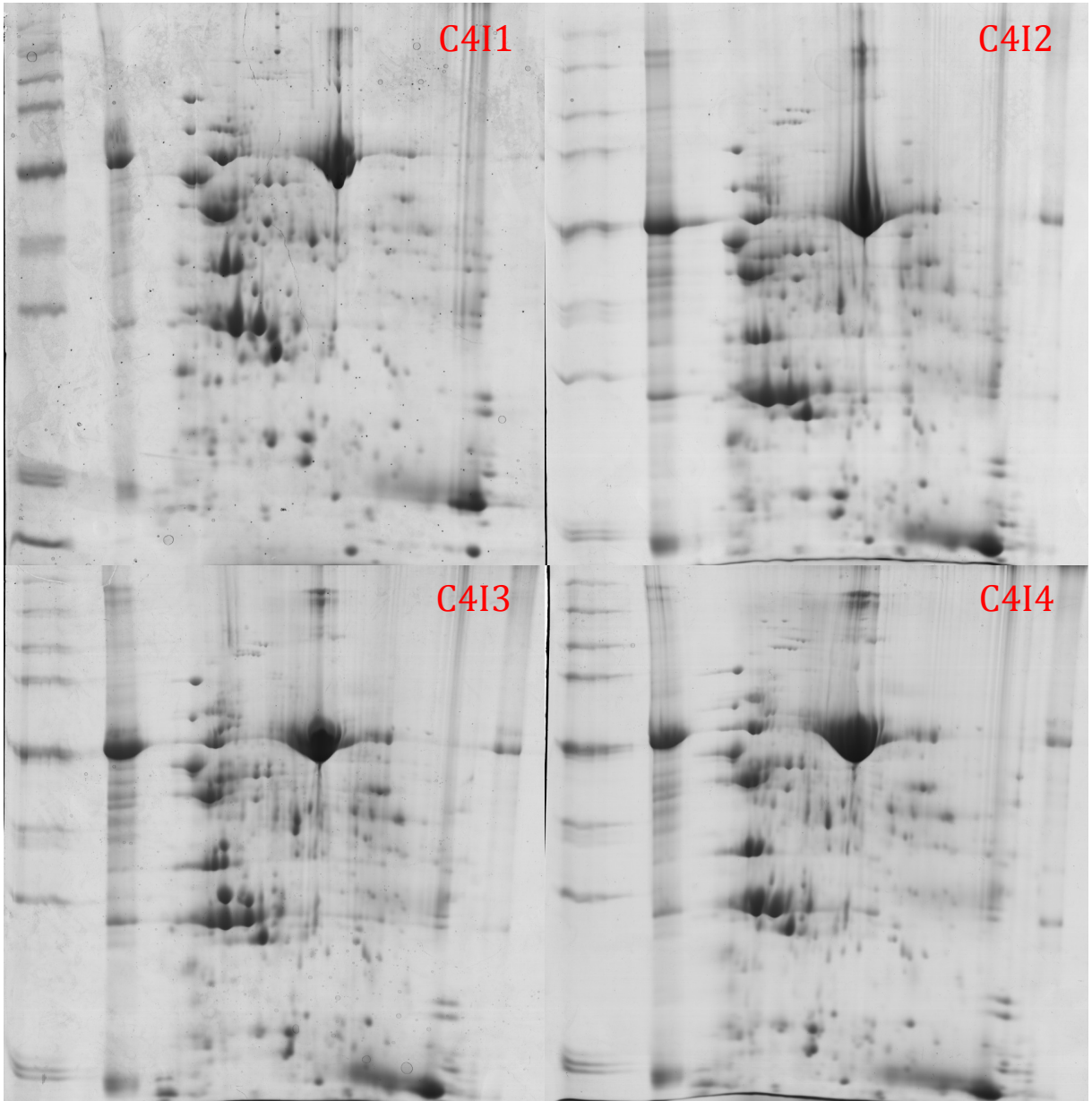
Os géis bidimensionais mostraram uma distribuição uniforme indicando proteínas em todos os pesos moleculares bem como em todos os pIs. As imagens das réplicas que foram selecionadas e alinhadas estão representadas na Figura 19.



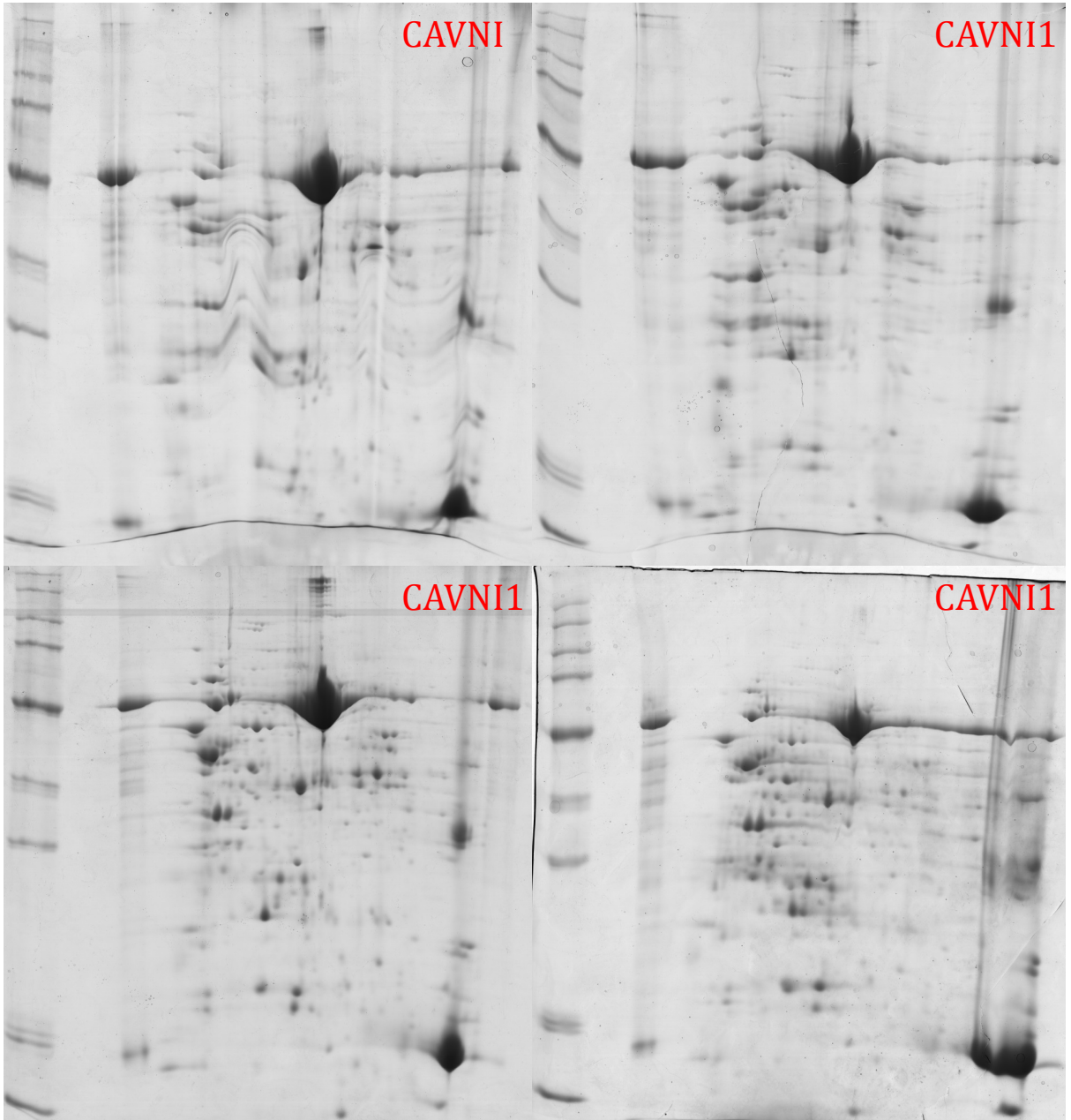


(a)

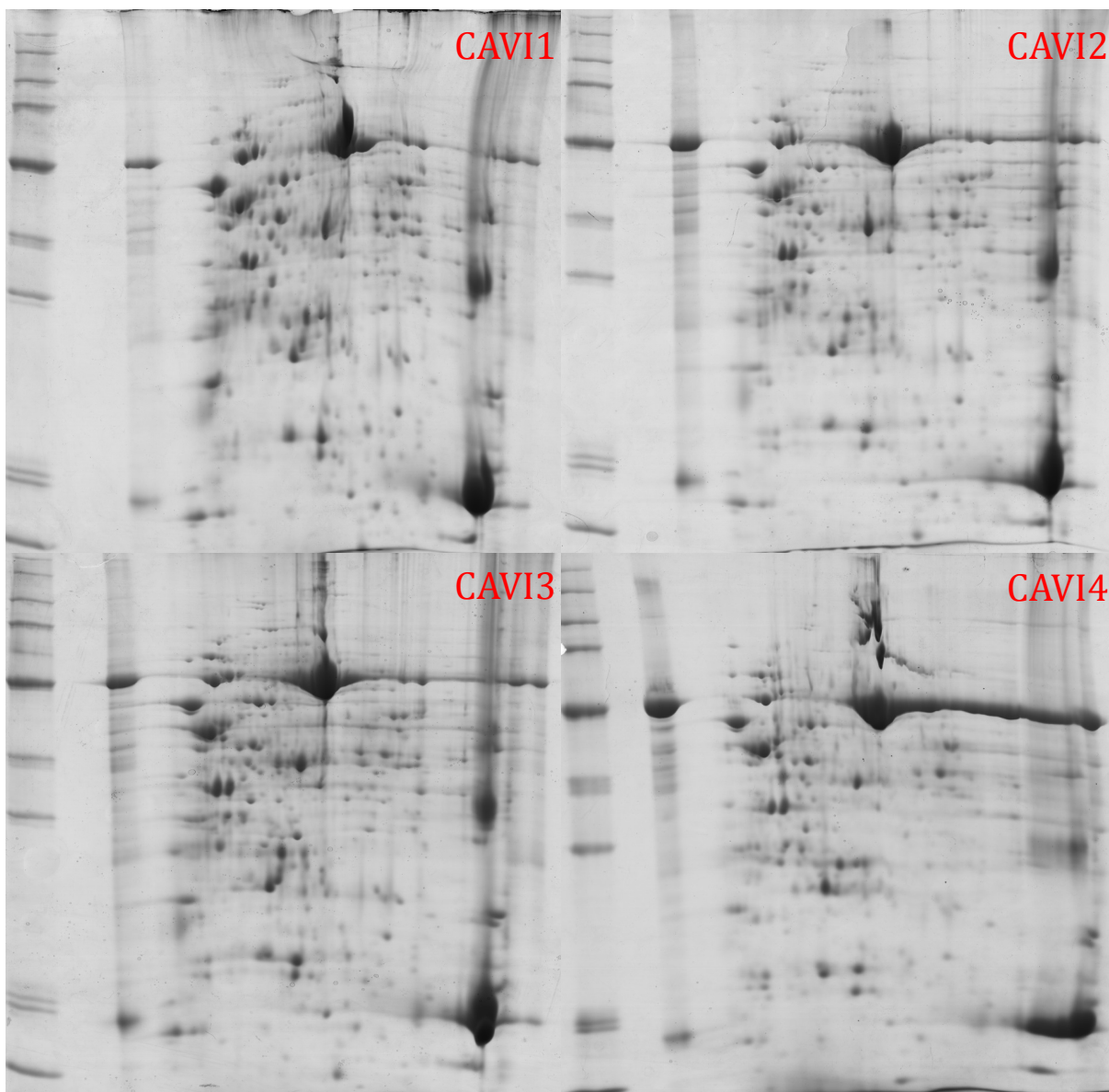




(b)



(c)



(d)

Figura 19 - Géis bidimensionais que foram analisados utilizando o software ImageMaster Platinum 7.0. (a) – Quadruplicatas do genótipo Calcutta 4 controle (C4NI1, C4NI2, C4NI3 e C4NI4); (b) – Quadruplicatas do genótipo Calcutta 4 inoculado com o patógeno *M. musicola* (C4I1, C4I2, C4I3 e C4I4); (c) - Quadruplicatas do genótipo Cavendish controle (CAVNI1, CAVNI2, CAVNI3 e CAVNI4); (d) - Quadruplicatas do genótipo Cavendish Grande Naine inoculado com o patógeno *M. musicola* (CAVI1, CAVI2, CAVI3 e CAVI4).

Os mapas protéicos de cada um dos tratamentos (cultivar resistente não inoculado e inoculado/cultivar suscetível não inoculado e inoculado) foram realizados em quadruplicatas e analisados utilizando o “software *Image master 2D Platinum 7*”. Um gel referente ao cultivar Cavendish não inoculado foi descartado das análises por não ter obedecido os parâmetros inicialmente propostos. As réplicas foram comparadas obtendo-se coeficientes de correlação com valores entre 0,91 e 0,99, considerados como aceitáveis para as análises seguintes de sobreposição dos géis para identificação de *spots* diferencialmente expressos (Figura 20). A correlação entre as repetições garante que as diferenças nas análises entre as condições não sejam atribuídas a limitações da técnica de extração, e sim, a variações no mapa protéico dentre as 4 condições testadas (C4NI, C4I, CAVNI e CAVI).

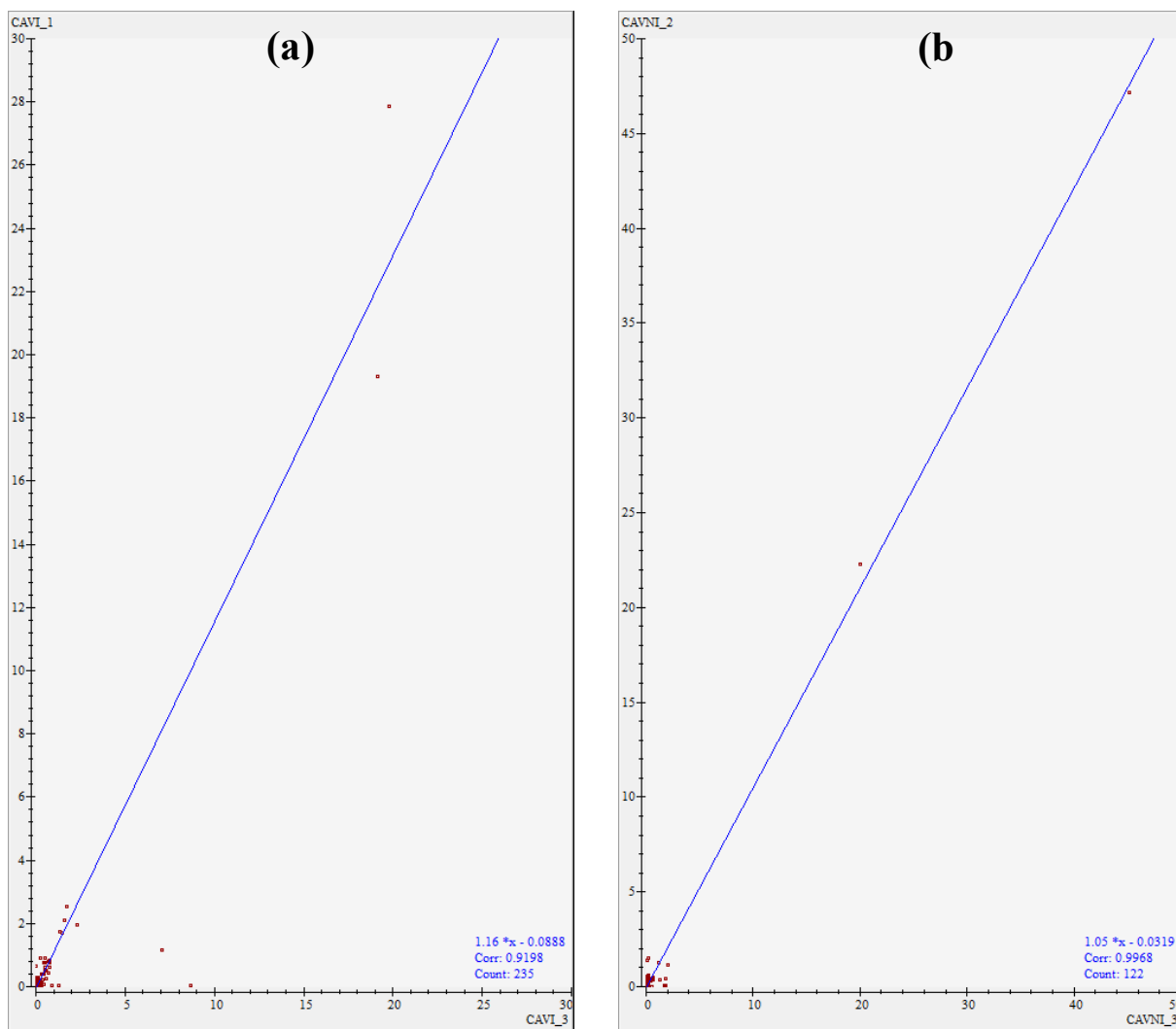


Figura 20 - Representação gráfica dos coeficientes de correlação obtidos entre as réplicas, demonstrando-se em **a)** menor valor de correlação obtido (0,9198) e em **b)** o maior (0,9968). Estes valores foram considerados como aceitáveis para as análises seguintes de sobreposição dos géis para identificação de *spots* diferencialmente expressos

Os mapas proteômicos de banana apresentaram *spots* bem definidos e com boa separação das proteínas abundantes neste tecido (Figura 19). A interferência de proteínas abundantes em folhas já foi descrita como uma das limitações para a obtenção de mapas proteômicos bem definidos (Wilson *et al.*, 2002). O número de *spots* detectados nos mapas proteômicos foi diferente entre as condições testadas, com um total de 242 *spots* para o tratamento C4NI, 230 *spots* para C4I, 155 para CAVNI e 323 para CAVI.

Foram feitas também análises para detecção de *spots* comuns ao mesmo cultivar (controle e tratado), bem como a identificação de *spots* com expressão diferencial, e detecção de *spots* presentes em apenas uma condição, denominados de *spots* exclusivos.

Foram identificados 68 *spots* comuns em Calcutta 4 não inoculado (controle) e inoculado com o fungo (C4NI e C4I). Em Cavendish, 48 *spots* estavam presentes em todos os géis do controle e do tratado. 23 *spots* estavam presentes em todos os 8 géis de Calcutta 4 controle e tratado, e 6 em Cavendish (Figura 21). A presença em todos os géis foi um parâmetro de seleção dos *spots* que serão identificados posteriormente por espectrometria de massa.

Os *spots* exclusivos foram detectados em todos os tratamentos. Para Calcutta 4, foram observados 179 *spots* exclusivos do controle e 170 exclusivos do inoculado com o fungo. Destes, 52 e 57 estavam presentes em todas as réplicas do controle e do tratado, respectivamente (Figura 22). Em Cavendish, 124 *spots* exclusivos foi identificado tanto no controle quanto no tratado. 53 e 74 *spots* foram identificados em todas as réplicas válidas para a análise no controle e no tratado respectivamente (Figura 23). Os *spots* em comum que apresentaram padrões de expressão também foram descritos para cada tratamento em comparação ao controle. No cultivar resistente Calcutta 4, foram observados 7 *spots* com uma expressão pelo menos 1,5 vezes maior no tratado quando comparado ao controle. O padrão de expressão foi diferente no cultivar suscetível, onde 1 *spot* apresentou diferença de expressão entre o controle e o tratado (Figura 24).

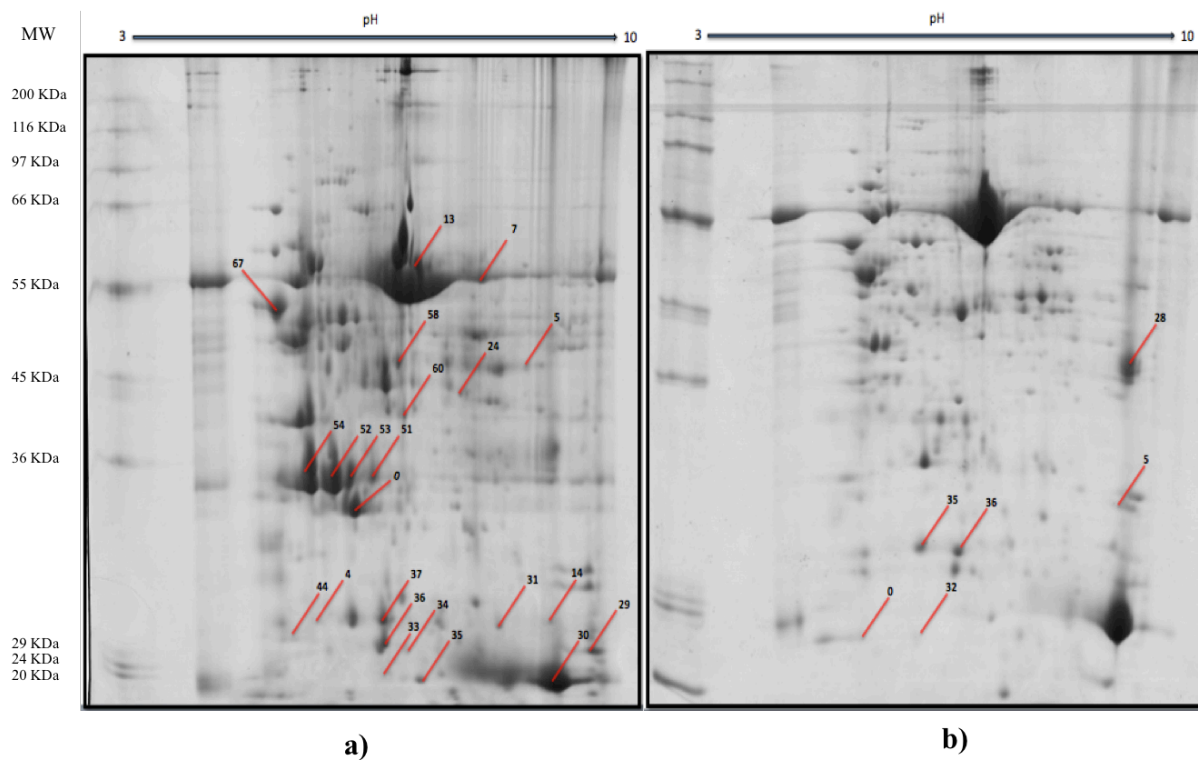


Figura 21 – Mapas bidimensionais de folha de *Musa acuminata* utilizados na identificação de *spots* presentes no cultivar resistente Calcutta 4 (controle e tratado) e no cultivar suscetível Cavendish (controle e tratado). **a)** os 23 *spots* presentes em todos os 8 géis de Calcutta 4. **b)** 6 *spots* em comum em ambos os tratamentos realizados em Cavendish. A presença em todos os géis foi um parâmetro de seleção dos *spots* que serão identificados posteriormente por espectrometria de massa. Os *spots* estão representados pelos números de identificação gerados pelo *software* utilizando nas análises dos géis.



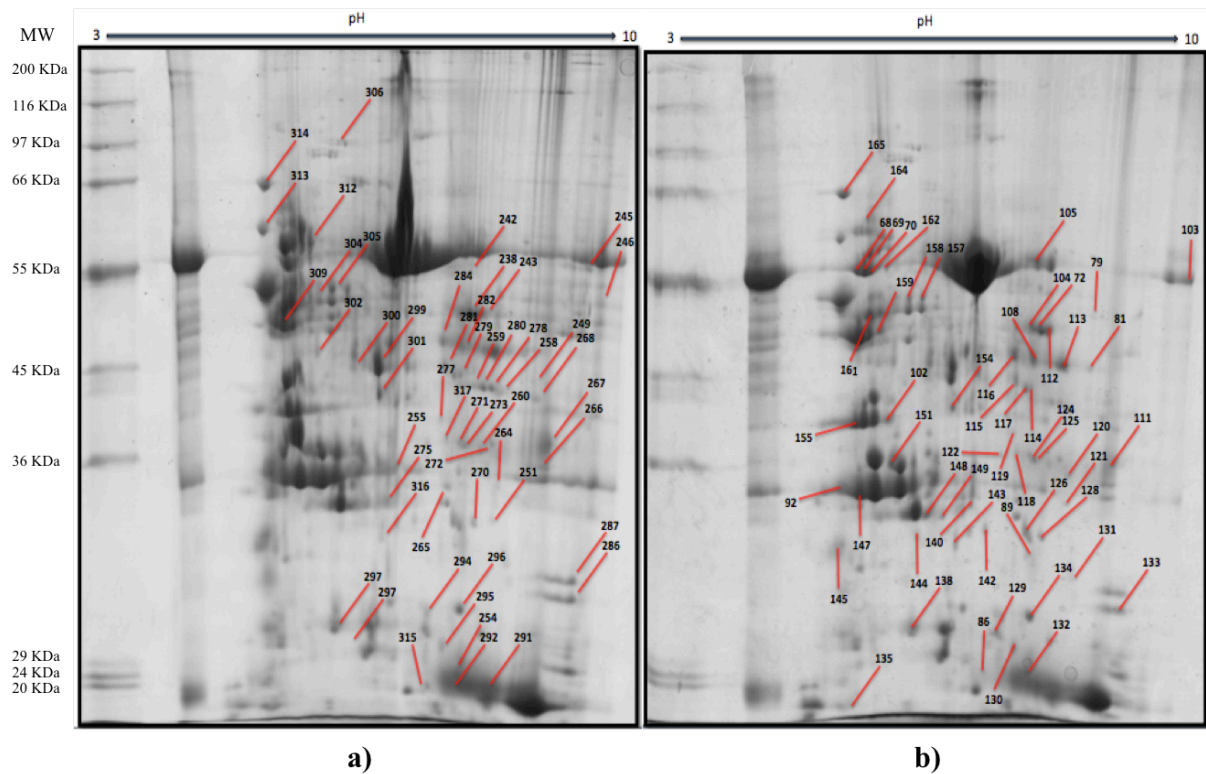


Figura 22 – Mapas bidimensionais de folha de *Musa acuminata* do cultivar resistente Calcutta 4. **a)** as 52 proteínas identificadas exclusivamente no controle, e presentes em todas as réplicas técnicas. **b)** as 57 proteínas que foram identificadas apenas no tratado, e em todas as réplicas. Os spots estão representados pelos números de identificação gerados pelo software utilizado nas análises dos géis.

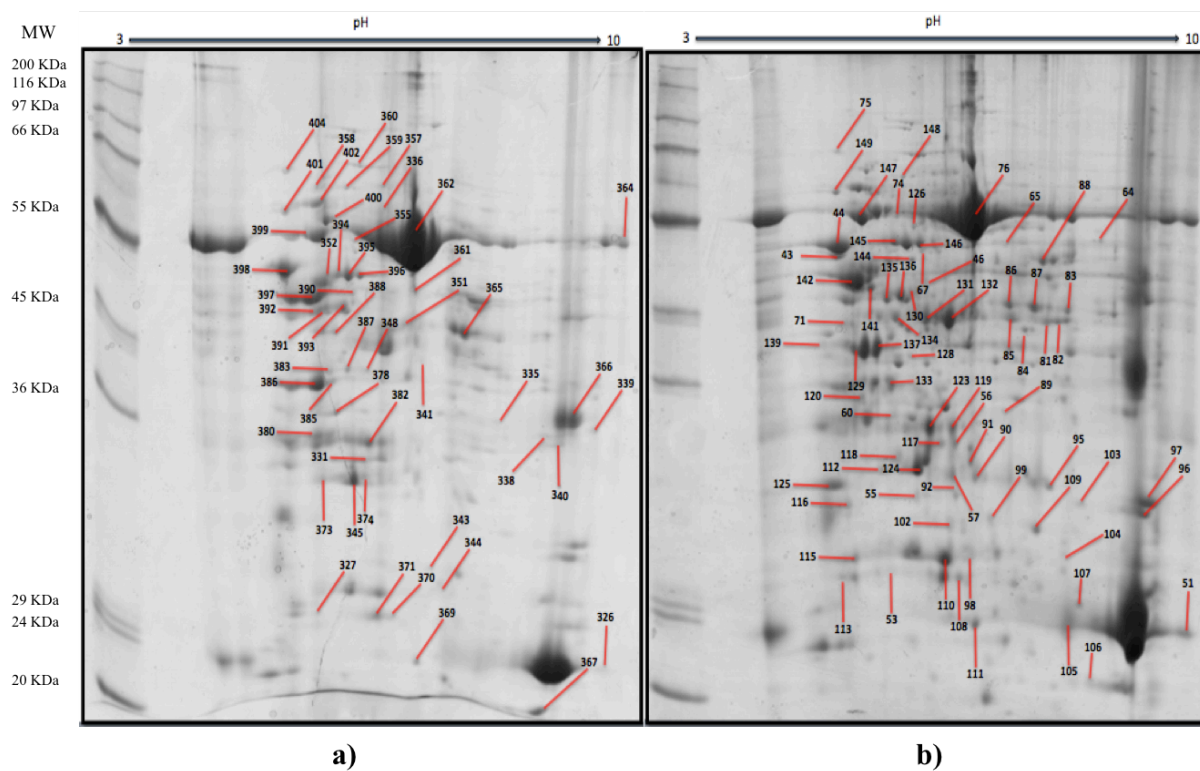


Figura 23 - Mapas bidimensionais de folha de *Musa acuminata* do cultivar suscetível Cavendish. **a)** as 53 proteínas identificadas exclusivamente no controle não inoculado com o patógeno, e presentes em todas as réplicas técnicas. **b)** as 74 proteínas que foram identificadas apenas no tratado, e em todas as réplicas. Os *spots* estão representados pelos números de identificação gerados pelo *software* utilizando nas análises dos géis.



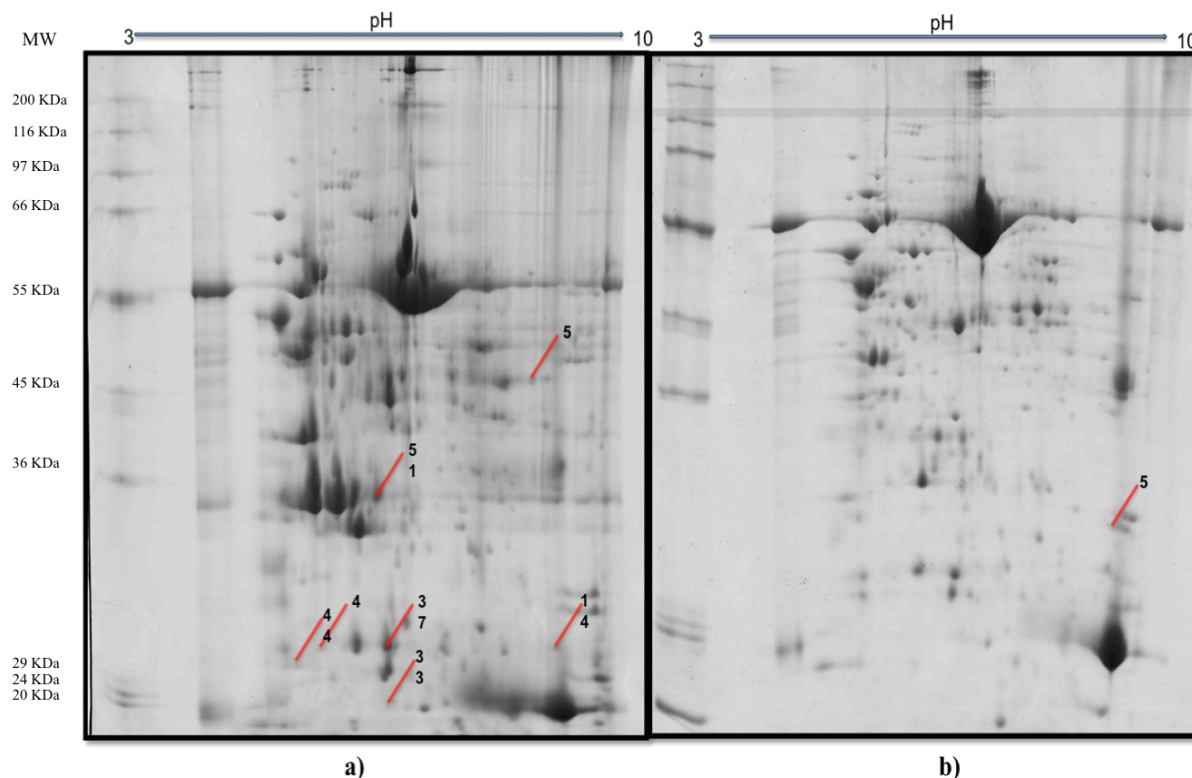


Figura 24 – Mapas bidimensionais de folha de *Musa acuminata* indicando os *spots* em comum que apresentaram padrões de expressão para cada tratamento em comparação ao controle. **a)** 7 *spots* observados no cultivar resistente Calcutta 4 com uma expressão pelo menos 1,5 vezes maior no tratado quando comparado ao controle. **b)** O único *spot* que apresentou diferença de expressão entre o controle e o tratado no cultivar suscetível Cavendish.

Em bananeira, já foi relatada a identificação com sucesso de proteínas da membrana plasmática utilizando MALDI-TOF/TOF. O grande desafio é realizar estudos quantitativos confiáveis (Vertommen *et al.*, 2011). Carpentier *et al.* (2011) caracterizaram polimorfismos protéicos inter e intra cultivares em banana utilizando gel 2DE e sequenciamento *de novo* de proteínas. Nesse mesmo estudo, indica-se que observações da proteômica fornecem boas indicações para a estrutura do genoma complexo e rearranjo genômico em algumas variedades de banana. Análises de proteômica quantitativa revelam que mecanismos de antioxidação contribuem para a tolerância ao frio em *Musa paradisiaca* (Yang *et al.*, 2012). Bananas (*Musa* spp.) são um excelente modelo para o estudo do proteoma do meristema. Utilizando um protocolo de 2-DE desenvolvido para pequenas quantidades de tecido e identificação de polipeptídeos por MS, os autores concluíram que uma análise do proteoma pode ser usada com sucesso para realizar uma análise quantitativa e investigar a variabilidade genética entre cultivares distintos (Carpentier *et al.*, 2007). Samyn *et al.* (2006) realizaram uma análise com sucesso do proteoma funcional de banana, uma planta não-modelo, utilizando MALDI-MS/MS e sequenciamento *de novo*. Zhang *et al.* (2012) demonstraram que

a extração fenólica gera proteínas em maior concentração e qualidade, em comparação com precipitação com TCA e precipitação TCA/Acetona. Utilizando MALDI-TOF MS e MALDI-TOF/TOF MS eles identificaram proteínas diferencialmente expressas em bananas induzidas a estresse abiótico quando comparadas ao controle. Eles discutem que o cDNA correspondente pode ser obtido, e através de RNAi ou geração de plantas transgênicas, poderá ser possível elucidar funções de novas proteínas. Vanhove *et al.* (2012) demonstraram a utilização da proteômica (2DE-MS) como ferramenta para descoberta de variedades de banana tolerantes a seca. Em uma análise recente da proteômica de bananeira, revelou 1.131 proteínas fornecendo uma base para os futuras pesquisas que visam criar organismos geneticamente modificados resistentes contra agentes patogênicos ou que expressem proteínas com um elevado valor nutricional (Esteve *et al.*, 2013).

A análise proteômica combinada com NGS, com a finalidade de buscar proteínas diferencialmente expressas tem sido relatada em diversos estudos e com vários organismos, como anelídeos, nematóides, insetos, fungos e plantas (Holder *et al.*, 2013; Moreno *et al.*, 2011; Kirsc *et al.*, 2012; Yin *et al.*, 2012; Sadiq *et al.*, 2011). Em um trabalho recente, transcritômica e proteômica foram utilizadas como ferramentas para fornecer um recurso para o estudo do fungo *Agrocybe aegerita* (Wang *et al.*, 2013). Huang *et al.* (2013) aliaram um conjunto de dados de EST com o perfil protéico utilizando 2-DE em *Tritrichomonas foetus*, a fim de fornecer uma base para o futuro sequenciamento do genoma, além de estabelecer alvos terapêuticos potenciais contra a infecção por esse parasita.

Em humanos, foi realizada uma análise proteômica do líquido sinovial do joelho de pacientes com osteoartrite. Os dados foram comparados com as análises do transcriptoma de tecidos comuns, identificando proteínas diferencialmente expressas que podem contribuir para a lesão articular. (Ritter *et al.*, 2013).

Zeng *et al.* (2013) integraram dados de transcriptoma, proteômica e metabolômica no estudo da via de biosíntese de alcalóides em plantas do gênero *Macleaya* spp, contribuindo para novas pesquisas de plantas não-modelo, utilizando a integração de diferentes tecnologias de alta capacidade.

Até recentemente, as investigações proteômicas em larga escala só foram possíveis para algumas espécies modelo, para as quais todas as seqüências do genoma já havia sido completamente determinadas. Por outro lado, para muitas outras espécies de interesse econômico, como fontes de alimentação humana e animal, pouco se sabia sobre a sua seqüência do genoma. Em consequência, a identificação do proteoma nestas espécies ainda era considerada um desafio. No entanto, o progresso tem sido feito como resultado de vários

avanços recentes nas ferramentas proteômicas, por exemplo, a tecnologia de espectrometria de massa, a crescente disponibilidade de seqüências genômicas e de cDNA de várias espécies. Além disso, as tecnologias de sequenciamento de última geração tornam agora possível determinar rapidamente, a um custo razoável, a seqüência do genoma ou do transcrito de espécies que atualmente não são consideradas como modelos, expandindo assim, os bancos de dados de seqüência de plantas. Desta forma, estes avanços permitem identificar um grande conjunto de proteínas, mesmo para espécies consideradas não-modelos (Champagne & Boutry, 2012).

## 6. CONCLUSÕES

Esta investigação NGS baseada no transcriptoma da interação *M. acuminata* - *M. musicola* pôde proporcionar dados úteis da expressão gênica durante as respostas imunológicas da planta neste patossistema, com 36.384 e 35.269 unigenes foram identificados, respectivamente, para os genótipos contrastantes *M. acuminata* Calcutta 4 Cavendish e Grande Naine. Os dados representam um recurso para o estudo desta planta não-modelo, com a identificação de genes candidatos regulados pela infecção, contribuindo para a nossa compreensão dos mecanismos de defesa do hospedeiro contra esse importante patógeno. O genoma de referência publicado recentemente representa aproximadamente 90% do tamanho do genoma estimado. As atualizações de anotação do genoma serão facilitadas pela disponibilidade crescente de um conjunto abrangente de sequências de genes expressos para esta espécie.

Na análise proteômica, foram feitas análises para detecção de proteínas expressas comuns ao mesmo cultivar (controle e tratado), bem como a identificação de proteínas expressas com expressão diferencial, e detecção de proteínas expressas presentes em apenas uma condição. No cultivar resistente Calcutta 4, foram observados 7 proteínas com uma expressão pelo menos 1,5 vezes maior no tratado quando comparado ao controle. O padrão de expressão foi diferente no cultivar suscetível, onde apenas 1 proteína apresentou diferença de expressão entre o controle e o tratado. As proteínas exclusivas a um genótipo, além das que apresentaram um padrão de expressão diferencial serão submetidas a uma digestão enzimática para as análises através de espectrometria de massa MALDI-ToF (*Matrix-Assisted Laser Desorption Ionization – Time of Flight*) para determinação da massa proteica.

Na era pós-genômica, a proteômica pode gerar informações importantes como: os tipos de proteínas expressas, modificações pós-traducional e respostas a diferentes condições ambientais, representando um elo entre a informação genotípica e fenotípica. Estes dados, em conjunto com as informações do transcriptoma, poderão facilitar a gestão de doenças em banana, baseando-se em melhoramento genético, através de transformação de plantas e seleção assistida por marcadores.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

- Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature*. 2003;422(6928):198–207.
- Altenbach, D. and Robatzek, S. Pattern recognition receptors: from the cell surface to intracellular dynamics. *Molecular Plant-Microbe Interactions*. 2007;20:1031-1039.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403–10.
- Arzanlou M, Abeln ECA, Kema GHJ, Waalwijk C, Carlier J, de Vries I, Guzman M, Crous PW: Molecular diagnostics for the Sigatoka disease complex of banana. *Phytopathology*. 2007;97:1112–1118.
- Audic S, Claverie JM. The significance of digital gene expression profiles. *Genome Res*. 1997;7(10):986–95.
- Azhar M, Heslop-Harrison JS: Genomes, diversity and resistance gene analogues in *Musa* species. *Cytogenet Genome Res*. 2008; 121:59–66.
- Bai TT, Xie WB, Zhou PP, Wu ZL, Xiao WC, Zhou L, *et al*. Transcriptome and Expression Profile Analysis of Highly Resistant and Susceptible Banana Roots Challenged with *Fusarium oxysporum* f. sp. cubense Tropical Race 4. *PLoS ONE*. 2013;8(9):739-745.
- Baker B, Zambryski P, Staskawicz B, Dinesh-Kumar SP. Signaling in plant-microbe interactions. *Science*. 1997;276:726-733.
- Barker BM, Kroll K, Vödisch M, Mazurie A, Kniemeyer O, Cramer RA. Transcriptomic and proteomic analyses of the *Aspergillus fumigatus* hypoxia response using an oxygen-controlled fermenter. *BMC Genomics*. 2012;13:62.
- Belalcázar Carvajal, SL. El cultivo del plátano en el trópico. Cali: Impresora Feriva, 1991;367
- Bent, A.F. and Mackey, D. Elicitors, effectors, and R genes: the new paradigm and a lifetime supply of questions. *Annu. Rev. Phytopathol*. 2007;45:399-436.
- Beveraggi A, Mourichon X, Salle G: Etude des interactions hôte-parasite chez des bananiers sensibles et résistants inoculés par *Cercospora fijiensis* (*Mycosphaerella fijiensis*) responsable de la maladie des raies noires. *Can J Bot*. 1995;73:1328–1337.
- Block A, Li G, Fu ZQ, Alfano JR. Phytopathogen type III effector weaponry and their plant targets. *Curr. Opin. Plant Biol*. 2008;11:396-403.
- Boller T, Felix G. A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annu. Rev. Plant Biol*. 2009;60:379-406.

Bolton MD, van Esse HP, Vossen JH, de Jonge R, Stergiopoulos I, Stulemeijer IJE, *et al.* The novel *Cladosporium fulvum* lysin motif effector Ecp6 is a virulence factor with orthologues in other fungal species. *Mol Microbiol.* 2008;69(1):119–36.

Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem.* 1976;72:248–54.

Brunner F, Rosahl S, Lee J, Rudd JJ, Geiler C, Kauppinen S, Rasmussen G, Scheel D, Nürnberger T. Pep-13, a plant defense-inducing pathogen-associated pattern from *Phytophthora* transglutaminases. *EMBO J.* 2002;21:6681-6688.

Cañas-Gutiérrez GP, Angarita-Velásquez MJ, Restrepo-Flórez JM, Rodríguez P, Moreno CX, Arango R. Analysis of the CYP51 gene and encoded protein in propiconazole-resistant isolates of *Mycosphaerella fijiensis*. *Pest Manag Sci.* 2009;65(8):892–9.

Carlier J, Zapater MF, Lapeyre F, Jones DR, Mourichon X. Septoria Leaf Spot of Banana: A Newly Discovered Disease Caused by *Mycosphaerella eumusae* (Anamorph *Septoria eumusae*). *Phytopathology.* 2000;90(8):884–90.

Carpentier SC, Panis B, Renaut J, Samyn B, Vertommen A, Vanhove A-C, *et al.*, The use of 2D-electrophoresis and de novo sequencing to characterize inter- and intra-cultivar protein polymorphisms in an allopolyploid crop. *Phytochemistry.* 2011 Jul;72(10):1243–50.

Carpentier SC, Witters E, Laukens K, Van Onckelen H, Swennen R, Panis B. Banana (*Musa* spp.) as a model to study the meristem proteome: acclimation to osmotic stress. *Proteomics.* 2007;7(1):92–105.

Carreel F, Faure S, Gonzalez D, Lagoda PJJ, Perrier X, Bakry F, Montcel HT, Lanaud C, Horry, JP. Evaluation de la diversité génétique chez les bananiers diploïdes (*Musa* spp.). *Genetics Selection Evolution.* 1994; 26(1): 125-136.

Cavalcante MJB, Escoute J, Madeira JP, Romero RE, Nicole M, Oliveira LC, Hamelin C, Lartaud M, Verdeil JL: Reactive oxygen species and cellular interactions between *Mycosphaerella fijiensis* and banana. *Trop Plant Biol.* 2011;4:134–143.

Celis JE, Kruhøffer M, Gromova I, Frederiksen C, Ostergaard M, Thykjaer T, *et al.* Gene expression profiling: monitoring transcription and translation products using DNA microarrays and proteomics. *FEBS Lett.* 2000;480(1):2–16.

Champagne A, Boutry M. Proteomics of nonmodel plant species. *Proteomics.* 2013;13(3-4):663–73.

Che, F.S., Nakajima, Y., Tanaka, N., Iwano, M., Yoshida, T., Takayama, S., Kadota, I, Isogai, A. Flagellin from an incompatible strain of *Pseudomonas avenae* induces a resistance response in cultured rice cells. *J. Biol. Chem.* 2000;275:32347-32356.

Cheung F, Town CD: A BAC end view of the *Musa acuminata* genome. *BMC Plant Biol.* 2007;7:29.

Chevreur B, Pfisterer T, Drescher B, Driesel AJ, Müller WEG, Wetter T, *et al.* Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection

in sequenced ESTs. *Genome Res.* 2004;14(6):1147–59.

Chisholm ST, Coaker G, Day B, Staskawicz BJ. Host-microbe interactions: shaping the evolution of the plant immune response. *Cell.* 2006;124:803-814.

Cho Y, Hou S, Zhong S. Analysis of expressed sequence tags from the fungal banana pathogen *Mycosphaerella fijiensis*. *Open Mycol J.* 2008;2:61–73.

Choi D, Kim HM, Yun HK, Park JA, Kim WT, Bok SH. Molecular cloning of a metallothionein-like gene from *Nicotiana glutinosa* L. and its induction by wounding and tobacco mosaic virus infection. *Plant Physiol.* 1996;112:353–359.

Churchill ACL. *Mycosphaerella fijiensis*, the black leaf streak pathogen of banana: progress towards understanding pathogen biology and detection, disease development, and the challenges of control. *Mol Plant Pathol.* 2011;12:307–328.

Conde-Ferrández L, Waalwijk C, Canto-Canché BB, Kema GHJ, Crous PW, James AC, *et al.* Isolation and characterization of the mating type locus of *Mycosphaerella fijiensis*, the causal agent of black leaf streak disease of banana. *Mol Plant Pathol.* 2007;8(1):111–20.

Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics.* 2005;21(18):3674–6.

Cordeiro MZJ, Matos PA. Impact of *Mycosphaerella* spp. in Brazil. In: *Mycosphaerella* Leaf Spot Diseases of Bananas: Present Status and Outlook. Proceedings of the Workshop on *Mycosphaerella* Leaf Spot Diseases, San José, Costa Rica. Montpellier: The International Network for the Improvement of Banana and Plantain. 2003;91-97.

Cordeiro ZJM. Panorama Nacional das Principais Doenças da Bananeira. Manejo integrado de doenças de fruteiras. Universidade Federal de Lavras: Sociedade Brasileira de Fitopatologia. 2007;165-184.

Cronquist A. An integrated system of classification of flowering plants. New York: Columbia University. 1981;805-8.

Crouch HK, Crouch JH, Jarret RL, Cregan PB, Ortiz R. Segregation of microsatellite loci from haploid and diploid gametes in *Musa*. *Crop Science.* 1998;211–217.

Crous PW, Kang JC, Braun U. A phylogenetic redefinition of anamorph genera in *Mycosphaerella* based on ITS rDNA sequence and morphology. *Mycologia.* 2001; 93:1081–1101.

Crous PW, Schoch CL, Hyde KD, Wood AR, Gueidan C, Gasparotto L, Pereira JCR, Costa MM, Pereira MCN. Fungicidas para o controle da Sigatoka negra da bananeira. *Fitopatologia Brasileira.* 2009;26:434. (Abstract 636).

D’Hont A, Denoeud F, Aury JM, Baurens FC, Carreel F, Garsmeur O, Noel B, Bocs S, Droc

G, Rouard M, Da Silva C, Jabbari K, Cardi C, Poulain J, Souquet M, Labadie K, Jourda C, Lengellé J, Rodier-Goud M, Alberti A, Bernard M, Correa M, Ayyampalayam S, Mckain MR, Leebens-Mack J, *et al.* The banana (*M. acuminata*) genome and the evolution of monocotyledonous plants. *Nature* 2012, doi:10.1038/nature11241.

D'Hont A, Paget-Goy A, Escoute J, Carreel F. The interspecific genome structure of cultivated banana, *Musa spp.* revealed by genomic DNA *in situ* hybridization. *Theoretical and Applied Genetics*. 2000;100(2):177-183.

da Cunha L, Sreerekha MV, Mackey D. Defense suppression by virulence effectors of bacterial phytopathogens. *Curr. Opin. Plant Biol.* 2007;10:349-357.

Dangl JL, Jones JD. Plant pathogens and integrated defence responses to infection. *Nature*. 2001;411:826-33.

Davey, MW, Gudimella R, Harikrishna JA, Sin LW, Khalid N, Keulemans W. *BMC Genomics*. 2013. *In press*.

De Jonge R, van Esse HP, Kombrink A, Shinya T, Desaki Y, Bours R, *et al.* Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants. *Science*. 2010;329(5994):953-5.

Denoux C, Galletti R, Mammarella N, Gopalan S, Werck D, De Lorenzo G, *et al.* Activation of defense response pathways by OGs and Flg22 elicitors in *Arabidopsis* seedlings. *Mol. Plant*. 2008;1:423-445.

Domingues ET, Tulmann Neto A, Mendes BMJ, Ando A. Efeitos de doses de raios-gama em ápices caulinares de bananeira (*Musa sp*) desenvolvidos *in vitro* visando a indução de mutação. *Pesquisa Agropecuária Brasileira*. 1994;29(7):1091-109.

Donzelli BGG, Churchill ACL: A quantitative assay using mycelial fragments to assess virulence of *Mycosphaerella fijiensis*. *Phytopathology*. 2007;97:916-929.

Emrich SJ, Barbazuk WB, Li L, Schnable PS. Gene discovery and annotation using LCM-454 transcriptome sequencing. *Genome Res*. 2007;17(1):69-73.

Esteve C, D'Amato A, Marina ML, García MC, Righetti PG. In-depth proteomic analysis of banana (*Musa spp.*) fruit with combinatorial peptide ligand libraries. *Electrophoresis*. 2013;34(2):207-14.

FAOSTAT. Food and agricultural commodities production. Food and Agricultural organization Statistics. 2012. <http://faostat.fao.org/>

Fedurco M, Romieu A, Williams S, Lawrence I, Turcatti G. BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies. *Nucleic acids research*. 2006;34: 22.

Flor HH. Inheritance of pathogenicity in *Melampsora lini*. *Phytopath*. 1942;32:653-669.



- Ganga M, Chezhiyan N. Influence of the antimetabolic agents colchicine and oryzalin on in vitro regeneration and chromosome doubling of diploid bananas (*Musa* spp.). *Journal of Horticultural Science & Biotechnology*. 2002;77(5): 572-575.
- Gasparotto L, Pereira JCR, Costa MM, Pereira MCN. Fungicidas para o controle da Sigatoka Negra da bananeira. *Fitopatologia Brasileira. Euphytica*. 2000; 113: 163–185.
- Gasparotto L, Pereira JCR, Hanada RE, Montarroyos AVV. Sigatoka-negra da bananeira. Brasília, Embrapa. 2006; 177p.
- Gene Ontology Consortium: The Gene Ontology project in 2008. *Nucleic Acids Res* 2008;36:440–444.
- Gijzen M, Nürnberger T. Nep1-like proteins from plant pathogens: Recruitment and diversification of the NPP1 domain across taxa. *Phytochemistry*. 2006;67:1800-1807.
- Glazebrook J. Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annu. Rev. Phytopathol.* 2005;43:205-227.
- Glowacki S, Macioszek VK, Kononowicz AK. R proteins as fundamental of plant innate immune system. *Cellular and Molecular Biology Letters*. 2010;16: 1-24.
- Gohre V, Robatzek S. Breaking the barriers: microbial effector molecules subvert plant immunity. *Annu. Rev. Phytopathol.* 2008;46:189-215.
- Goodwin SB, Dunkle LD, Zismann VL. Phylogenetic analysis of cercospora and mycosphaerella based on the internal transcribed spacer region of ribosomal DNA. *Phytopathology*. 2001;91(7):648–58.
- Görg A, Drews O, Lück C, Weiland F, Weiss W. 2-DE with IPGs. *Electrophoresis*. 2009;30 Suppl 1:122–132.
- Gust AA, Biswas R, Lenz HD, Rauhut T, Ranf S, Kemmerling B. Bacteria-derived peptidoglycans constitute pathogen-associated molecular patterns triggering innate immunity in *Arabidopsis*. *J. Biol. Chem.* 2007;82:32338-32348.
- Häkkinen M. Reappraisal of sectional taxonomy in *Musa* (*Musaceae*). *International Association for Plant Taxonomy*. 2013;62(4):809-813.
- Hamer DH: Metallothionein. *Annu Rev Biochem.* 1986;55:913–951.
- Hamill SD, Smith MK, Dodd WA. In vitro induction of banana autotetraploids by colchicine treatment of micropropagated diploids. *Australian Journal of Botany*. 1992;40: 887-96.
- Hammerschmidt R. Systemic acquired resistance. *Advances in Botanical Research*. 2009;51:173-222.
- Hammond-Kosack KE, Jones JDG: Plant disease resistance genes. *Annu Rev Plant Physiol Plant Mol Biol*. 1997;48:573–605.

- Hammond-Kosack KE, Jones JDG. Plant disease resistance genes. *Annual Review of Plant Physiology and Plant Molecular Biology*. 1997;48:573-605.
- Hazelhurst S, Hide W, Lipták Z, Nogueira R, Starfield R. An overview of the wcd EST clustering tool. *Bioinformatics*. 2008;24(13):1542–6.
- Heslop-Harrison JS, Schwarzacher T. Domestication, genomics and the future for banana. *Ann Bot (Lond)*. 2007;100(5):1073-84.
- Holder T, Basquin C, Ebert J, Randel N, Jollivet D, Conti E, *et al*. Deep transcriptome-sequencing and proteome analysis of the hydrothermal vent annelid *Alvinella pompejana* identifies the CvP-bias as a robust measure of eukaryotic thermostability. *Biol Direct*. 2013;8:2.
- Horry JP, Ortiz R, Arnaud E, Crouch JH, Ferris RSB, Jones DR, Mateo N, Picq C, Vuylsteke D. Banana and Plantain. 1997;67-81.
- Hribová E, Neumann P, Matsumoto T, Roux N, Macas J, Dolezel J. Repetitive part of the banana (*M. acuminata*) genome investigated by low-depth 454 sequencing. *BMC Plant Biol*. 2010;10:204.
- Huang KY, Shin JW, Huang PJ, Ku FM, Lin WC, Lin R, *et al*. Functional profiling of the *Tritrichomonas foetus* transcriptome and proteome. *Mol Biochem Parasitol*. 2013;187(1):60–71.
- James AC, Arzanlou M, Canche BC, Ramirez JH, Conde Ferraez L, Echeverria SP. Fungal Disease of Banana. *Bananas: Nutrition, Diseases and Trade Issues*. 2010; 3: 65-122.
- Jones JDG, Dangl JL. The plant immune system. *Nature*. 2006;444:323-329.
- Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem*. 1988;60(20):2299–301.
- Keen NT. Gene-for-gene complementarity in plant-pathogen interactions. *Annu. Rev. Genet*. 1990;24:447-463.
- Kimati H, Galli F. Doenças da bananeira *Musa* spp. *Manual de Fitopatologia; doenças das plantas cultivadas*. Ed. Agronômica Ceres. 1980;2: 87-101.
- Kirsch R, Wielsch N, Vogel H, Svatoš A, Heckel DG, Pauchet Y. Combining proteomics and transcriptome sequencing to identify active plant-cell-wall-degrading enzymes in a leaf beetle. *BMC Genomics*. 2012;13:587.
- Koornneef A, Pieterse CMJ. Cross talk in defense signaling. *Plant Physiology*. 2008;146:839-844.
- Kumar SGB, Ganapathi TR, Revathi CJ, Srivivas L, Bapat VA. Expression of hepatitis B surface antigen in transgenic banana plants. *Planta*. 2005;222(3): 484-493.
- Küster B, Mortensen P, Andersen JS, Mann M. Mass spectrometry allows direct identification of proteins in large genomes. *Proteomics*. 2001;1(5):641–50.

Lanaud C, Montcel HT, Jolivot MP, Glaszmann JC, Gonzalez D. Variation of ribosomal gene spacer length among wild and cultivated banana. *Heredity*. 1992;68(2):147-156.

Leach R. Banana leaf spot (*Mycosphaerella musicola*) on the Gros Michel variety in Jamaica. Gov. Printer. 1946; 118p.

Lepesheva GI, Waterman MR. CYP51--the omnipotent P450. *Mol Cell Endocrinol*. 2004;215(1-2):165-70.

Lery LMS, Coelho A, von Kruger WMA, Gonçalves MSM, Santos MF, Valente RH, *et al*. Protein expression profile of *Gluconacetobacter diazotrophicus* PAL5, a sugarcane endophytic plant growth-promoting bacterium. *Proteomics*. 2008;8(8):1631-44.

Lescot M, Piffanelli P, Ciampi AY, Ruiz M, Blanc G, Leebens-Mack J, da Silva FR, Santos CM, D'Hont A, Garsmeur O, Vilarinhos AD, Kanamori H, Matsumoto T, Ronning CM, Cheung F, Haas BJ, Althoff R, Arbogast T, Hine E, Pappas GJ Jr, Sasaki MT Jr, Souza MT Jr, Miller RNG, Glaszmann JC, Town CD: Insights into the *Musa* genome: syntenic relationships to rice and between *Musa* species. *BMC Genomics* 2008;9:58.

Li CY, Deng GM, Yang J, Viljoen A, Jin Y, Kuang RB, *et al*. Transcriptome profiling of resistant and susceptible Cavendish banana roots following inoculation with *Fusarium oxysporum* f. sp. *cubense* tropical race 4. *BMC Genomics*. 2012;13:374.

Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-60.

Liu H, Luo Y, Han J, Wu J, Wu Z, Feng D, *et al*. Proteome reference map of *Haloarcula hispanica* and comparative proteomic and transcriptomic analysis of polyhydroxyalkanoate biosynthesis under genetic and environmental perturbations. *J Proteome Res*. 2013;12(3):1300-15.

Liu P, Goh CJ, Loh CS, Pua EC. Differential expression and characterization of three metallothionein-like genes in Cavendish banana (*M. acuminata*). *Physiol Plant*. 2002;114:241-250.

Liu Z, Crampton M, Todd A, Kalavacharla V. Identification of expressed resistance gene-like sequences by data mining in 454-derived transcriptomic sequences of common bean (*Phaseolus vulgaris* L.). *BMC Plant Biol*. 2012;12:42.

Lorenzo O, Piqueras R, Sánchez-Serrano JJ, Solano R. Ethylene response factor1 integrates signaling from ethylene and jasmonate pathways in plant defense. *Plant Cell*. 2003;15:165-178.

Macas J, Neumann P, Navrátilová A. Repetitive DNA in the pea (*Pisum sativum* L.) genome: comprehensive characterization using 454 sequencing and comparison to soybean and *Medicago truncatula*. *BMC Genomics*. 2007;8:427.

Manrique-Trujillo SM, Ramírez-López AC, Ibarra-Laclette E, Gómez-Lim MA. Identification of genes differentially expressed during ripening of banana. *J Plant Physiol*. 2007;14:1037-

1050.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature*. 2005;437(7057):376–80.

Marín DH, Romero RA, Guzmán, M. Sutton TB. Black Sigatoka: an increasing threat to banana cultivation. *Plant Dis*. 2003; 87: 208–222.

Marshall R, Kombrink A, Motteram J, Loza-Reyes E, Lucas J, Hammond-Kosack KE, *et al.* Analysis of two in planta expressed LysM effector homologs from the fungus *Mycosphaerella graminicola* reveals novel functional properties and varying contributions to virulence on wheat. *Plant Physiol*. 2011;156(2):756–69.

Matsumoto K, Vilarinhos AD, Oka S. Somatic hybridization by electrofusion of banana protoplasts. *Euphytica*. 2002;125(3): 317-324.

McHale L, Tan X, Koehl P, Michelmore RW: Plant NBS-LRR proteins: adaptable guards. *Genome Biol*. 2006;7:212.

McNulty SN, Abubucker S, Simon GM, Mitreva M, McNulty NP, Fischer K, *et al.* Transcriptomic and proteomic analyses of a Wolbachia-free filarial parasite provide evidence of trans-kingdom horizontal gene transfer. *PLoS ONE*. 2012;7(9):45777.

Medzhitov, R. and Janeway, C.A., Jr. 1997. Innate immunity: the virtues of a nonclonal system of recognition. *Cell* 91:295-298.

Men L, Yan S, Liu G. De novo characterization of *Larix gmelinii* (Rupr.) Rupr. transcriptome and analysis of its gene expression induced by jasmonates. *BMC Genomics*. 2013;14(1):548.

Meredith DS, Lawrence JS. Black leaf streak disease of bananas (*Mycosphaerella fijiensis*): symptoms of disease in Hawaii, and notes on the conidial state of the causal fungus. *Transactions of the British Mycological Society*. 1970; 52: 459-467.

Meredith DS. Banana leaf spot disease (Sigatoka) caused by *Mycosphaerella musicola*. *Phytopathology* 1970; 147p.

Metzker ML. Sequencing technologies - the next generation. *Nat Rev Genet*. 2010 Jan;11(1):31–46.

Metzker ML. Sequencing technologies: the next generation. *Nature reviews Genetics*. 2010;11:31-46.

Miller RNG, Bertioli DJ, Baurens FC, Santos CM, Alves PC, Martins NF, Togawa RC, Souza MT Junior, Pappas GJ Junior. Analysis of non-TIR NBS-LRR resistance gene analogs in *M. acuminata* Colla: Isolation, RFLP marker development, and physical mapping.. *BMC Plant Biology*. 2008;8:15.

Miller RNG, Bertioli DJ, Baurens FC, Quirino BF, Ciampi AY, Santos CMR, *et al.*

Understanding plant responses to biotic stress: ongoing research in *Musa*. *Acta Hort.* 2009;828:255-272.

Mishina TE, Zeier J. Pathogen-associated molecular pattern recognition rather than development of tissue necrosis contributes to bacterial induction of systemic acquired resistance in *Arabidopsis*. *Plant J.* 2007;50:500-513.

Morel JB, Dangl JL: The hypersensitive response and the induction of cell death in plants. *Cell Death Differ.* 1997;4:671–683.

Moreno Y, Gros P-P, Tam M, Segura M, Valanparambil R, Geary TG, *et al.* Proteomic Analysis of Excretory-Secretory Products of *Heligmosomoides polygyrus* Assessed with Next-Generation Sequencing Transcriptomic Information. *PLoS Negl Trop Dis.* 2011;5(10).

Navarro L, Zipfel C, Rowland O, Keller I, Robatzek S, Boller T, Jones JD. The transcriptional innate immune response to flg22, Interplay and overlap with Avr gene-dependent defense responses and bacterial pathogenesis. *Plant Physiol.* 2004;135:1113-1128.

Niki T, Mitsuhashi I, Seo S, Ohtsubo N, Ohashi Y. Antagonistic effect of salicylic acid and jasmonic acid on the expression of pathogenesis-related (PR) protein genes in wounded mature tobacco leaves. *Plant and Cell Physiology.* 1998;39:500- 507.

Ning Z, Cox AJ, Mullikin JC. SSAHA: a fast search method for large DNA databases. *Genome Res.* 2001;11(10):1725–9.

Novaes E, Drost DR, Farmerie WG, Pappas GJ Jr, Grattapaglia D, Sederoff RR, *et al.*, High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome. *BMC Genomics.* 2008;9:312.

Nürnberg T, Kemmerling B. PAMP-triggered basal immunity in plants. *Advances in Botanical Research.* 2009;51:1-38.

Ortiz R. Secondary polyploids, heterosis, and evolutionary crop breeding for further improvement of the plantain and banana (*Musa* spp. L.) genome. *Theor. Appl. Genet.* 1997;94: 1113–1120.

Papanicolaou A, Stierli R, French-Constant RH, Heckel DG. Next generation transcriptomes for next generation genomes using est2assembly. *BMC Bioinformatics.* 2009;10:447.

Passos MAN, de Oliveira Cruz V, Emediato FL, de Camargo Teixeira C, Souza MT Jr, Matsumoto T, Rennó Azevedo VC, Ferreira CF, Amorim EP, de Alencar Figueiredo LF, Martins NF, de Jesus Barbosa Cavalcante M, Baurens FC, da Bonfim Silva O Jr, Pappas GJ Jr, Pignolet L, Abadie C, Ciampi AY, Piffanelli P, Miller RNG: Development of expressed sequence tag and EST-SSR marker resources for *M. acuminata*. *AoB Plants.* 2012, doi:10.1093/aopbla/pls030. pls030.

Pei XW, Chen SK, Wen RM, Ye S, Huang JQ, Zhang YQ, Wang BS, Wang ZX, Jia SR. Creation of transgenic bananas expressing human lysozyme gene for panama wilt resistance. *Journal of Integrative Plant Biology.* 2005;47(8): 971-977.

Pérez Ponce JN. Mutagêneses in vitro. Propagación y mejora genética de plantas por biotecnología. Santa Clara: Instituto de Biotecnología de las plantas. 1998;297-326.

Perrier X, Bakry F, Carreel F, Jenny C, Horry JP, Lebot V, Hippolyte I. Combining biological approaches to shed light on evolution of edible bananas. *Ethnobotany Research and Applications*. 2009;7(1):199-216.

Peterson R, Grice K, De La Rue S. Management of *Mycosphaerella* leaf spot diseases in Australia. 2003. *Mycosphaerella* Leaf Spot of Bananas: Present Status and Outlook, Proceedings of the 2nd International Workshop on *Mycosphaerella* Leaf Spot Diseases, 20-23 May 2002, San José, Costa Rica.

Portal O, Izquierdo Y, De Vleeschauwer D, Sánchez-Rodríguez A, Mendoza-Rodríguez M, Acosta-Suárez M, Ocaña B, Jiménez E, Höfte M. Analysis of expressed sequence tags derived from a compatible *Mycosphaerella fijiensis*-banana interaction. *Plant Cell Rep*. 2011;30:913–928.

Queiroz CB, Miranda EC, Hanada RE, Sousa NR, Gasparotto L, Soares MA, *et al*. Distribution of mating-type alleles and M13 PCR markers in the black leaf spot fungus *Mycosphaerella fijiensis* of bananas in Brazil. *Genet Mol Res*. 2013;12(1):443–52.

Quirino BF, Candido ES, Campos PF, Franco OL, Krüger RH. Proteomic approaches to study plant-pathogen interactions. *Phytochemistry*. 2010;71:351–62.

Raboin LM, Carreel F, Noyer JL, Baurens FC, Horry JP, Bakry F, Montcel HT, Ganry J, Lanaud C, Lagoda P. Diploid ancestors of triploid export banana cultivars: molecular identification of  $2n$  restitution gamete donors and  $n$  gamete donors. *Molecular Breeding*. 2005;16(4): 333-341.

Ritter SY, Subbaiah R, Bebek G, Crish J, Scanzello CR, Krastins B, *et al*. Proteomic analysis of synovial fluid from the osteoarthritic knee: comparison with transcriptome analyses of joint tissues. *Arthritis Rheum*. 2013;65(4):981–92.

Roche Applied Science. 2008. Genome Sequencer Data Analysis Software Manual. Software Version 2.0.00, Roche Diagnostics GmbH.

Romualdi C, Bortoluzzi S, D'Alessi F, Danieli GA. IDEG6: a web tool for detection of differentially expressed genes in multiple tag sampling experiments. *Physiol Genomics*. 2003;12(2):159–62.

Rothberg JM, Leamon JH. The development and impact of 454 sequencing. *Nature Biotechnology*. 2008;10: 1117-1124.

Rougemont J, Amzallag A, Iseli C, Farinelli L, Xenarios I, Naef F. Probabilistic base calling of Solexa sequencing data. *BMC Bioinformatics*. 2008;9:431.

Sadiq I, Fanucchi F, Paparelli E, Alpi E, Bachi A, Alpi A, *et al*. Proteomic identification of differentially expressed proteins in the anoxic rice coleoptile. *J Plant Physiol*. 2011;168(18):2234–43.



Samyn B, Sergeant K, Carpentier S, Debyser G, Panis B, Swennen R, *et al.*, Functional proteome analysis of the banana plant (*Musa* spp.) using de novo sequence analysis of derivatized peptides. *J Proteome Res.* 2007 Jan;6(1):70–80.

Santos CMR, Martins NF, Horberg HM, de Almeida ERP, Coelho MCF, Togawa RC, da Silva FR, Caetano AR, Miller RNG, Souza MT Jr. Analysis of expressed sequence tags from *M. acuminata* ssp. burmannicoides, var. Calcutta 4 (AA) leaves submitted to temperature stresses. *Theor Appl Genet* 2005;110:1517–1522.

Shendure J, Ji H. Next-generation DNA sequencing. *Nature biotechnology.* 2008;26:1135-1145.

Shepherd K. Taxonomia e caracterização de cultivares de banana. Cruz das Almas: Embrapa CNPMF, 1984.

Shi T, Gao Z, Wang L, Zhang Z, Zhuang W, Sun H, *et al.* Identification of differentially-expressed genes associated with pistil abortion in Japanese apricot by genome-wide transcriptional analysis. *PLoS ONE.* 2012;7(10):47810.

Shiu SH, Blecker AB. Receptor-like kinases from *Arabidopsis* form a monophyletic gene family related to animal receptor kinases. *Proc. Natl Acad. Sci. USA.* 2001;98:10763-10768.

Silva SO, Matos AP, Alves EJ. Melhoramento genético de bananeira. *Pesquisa Agropecuária Brasileira.* 1998;33(5):693-703.

Silva SO, Shepherd K, Dantas JLL, Alves EJ, Borges AL, Fancelli M, Oliveira SL, Almeida MA. Avanços do Recursos Genéticos e Melhoramento de Plantas para o Nordeste Brasileiro programa de pesquisa em *Musa* no CNPMF, Embrapa, Brasil. Cruz das Almas: Embrapa-CNPMF. 1996;37p.

Silva SO, Shepherd K, Dantas JLL, Souza AS, Carneiro MS. Germoplasma de banana. A cultura da banana. Aspectos técnicos socioeconômicos e agroindustriais. 1997;61-84.

Simmonds JH. Banana Leaf Spot. Queensland: Department of Agriculture and Stock Division of Entomology and Plant Pathology. 1933.

Simmonds NW, Shepherd K. The taxonomy and origin of the cultivated bananas. *The Botany Journal of the Linnean Society of London.* 1955;55:302-312.

Simmonds NW, Shepherd K. The taxonomy and origins of the cultivated bananas. *Journal of the Linnean Society of London (Botany).* 1955;55:302-312.

Simmonds NW. *Bananas Diseases* Longmans. 1966;2:366-408.

Souza SAC. Avaliação da variabilidade genética em *Musa* ssp. utilizando marcadores microsatélites. Piracicaba. 2002;86p

Steiling K, Kadar AY, Bergerat A, Flanigon J, Sridhar S, Shah V, *et al.* Comparison of proteomic and transcriptomic profiles in the bronchial airway epithelium of current and never smokers. *PLoS ONE.* 2009;4(4):5043.

- Stover RH, Simmonds NW. Bananas. Longmans. (Tropical Agriculture Series). 1987;468 p.
- Stover RH. Leaf spot of bananas caused by *M. musicola*: role of conidia in epidemiology. *Phytopathology*. 1970;60:856-860.
- Taguchi F, Yamamoto M, Ohnishi-Kameyama M, Iwaki M, Yoshida M, Ishii *et al.* Defects in flagellin glycosylation affect the virulence of *Pseudomonas syringae* pv. *tabaci* 6605. *Microbiology*. 2010;156:72-80.
- Tenkouano A, Ortiz R, Vuylsteke D. Combining ability for yield and plant phenology in plantain-derived populations. *Euphytica*. 1998;104(3):151-158.
- Thomma PBHJ, Nürnberger T, Joosten MHAJ. Of PAMPs and effectors: the blurred PTI-ETI dichotomy. *The Plant Cell*. 2011;23:4-15.
- Toledo-Silva G, Cardoso-Silva CB, Jank L, Souza AP. De novo transcriptome assembly for the tropical grass *Panicum maximum* Jacq. *PLoS ONE*. 2013;8(7):70781.
- Torres MA, Jones JD, Dangl JL. Reactive oxygen species signaling in response to pathogens. *Plant Physiol*. 2006;141:373-378.
- Trujillo IL, Hermoso Y, García E. Caracterización estructural de clones de banano: resistentes y no resistentes a la sigatoka amarilla. *Anales de Botánica Agrícola*. 1997; 4:59-62.
- Tsuda K, Katagiri F. Comparing signaling mechanisms engaged in pattern- triggered and effector-triggered immunity. *Curr. Opin. Plant Biol*. 2010;13:459-465.
- Tsuda K, Sato M, Stoddard T, Glazebrook J, Katagiri F. Network properties of robust immunity in plants. *PLoS Genetics*. 2009;5:12.
- Tulmann Neto A, Domingues ET. In vitro mutation breeding methodology for Fusarium wilt resistance in banana. *Mutation Breeding Newsletter*. 1990;36:12-13.
- Turcatti G, Romieu A, Fedurco M, Tairi AP. A new class of cleavable fluorescent nucleotides: synthesis and optimization as reversible terminators for DNA sequencing by synthesis. *Nucleic acids research*. 2008;36: 25.
- Uma S, Mustaffa MM, Saraswathi MS, Durai P. Exploitation of diploids in Indian breeding programmes. *Acta Hort*. 2011; 897:215–223.
- Van der Biezen E, Jones JG. Plant disease-resistance proteins and the gene- for-gene concept. *Trends in Biochemical Sciences*. 1998;23:454-456.
- Van Der Ent S, Van Hulten M, Pozo MJ, Czechowski T, Udvardi MK, Pieterse CMJ, Ton J. Priming of plant innate immunity by rhizobacteria and beta- aminobutyric acid: differences and similarities in regulation. *New Phytologist*. 2009;183:419-431.
- Van der Hoorn RA, Kamoun S. From guard to decoy: a new model for perception of plant pathogen effectors. *Plant Cell*. 2008;20:2009-2017.



Van Loon LC, Rep M, Pietersen CMJ. Significance of inducible defense- related proteins in infected plants. *Annual Review of Phytopathology*. 2006;44:135-162.

Vanhove AC, Vermaelen W, Panis B, Swennen R, Carpentier SC. Screening the banana biodiversity for drought tolerance: can an in vitro growth model and proteomics be used as a tool to discover tolerant varieties and understand homeostasis. *Front Plant Sci*. 2012;3:176.

Vera M, Alvarez-Dios JA, Fernandez C, Bouza C, Vilas R, Martinez P. Development and Validation of Single Nucleotide Polymorphisms (SNPs) Markers from Two Transcriptome 454-Runs of Turbot (*Scophthalmus maximus*) Using High-Throughput Genotyping. *Int J Mol Sci*. 2013;14(3):5694–711.

Vertommen A, Møller ALB, Cordewener JHG, Swennen R, Panis B, Finnie C, *et al*. A workflow for peptide-based proteomics in a poorly sequenced plant: a case study on the plasma membrane proteome of banana. *J Proteomics*. 2011;74(8):1218–29.

Vivier E, Malissen B. Innate and adaptive immunity: specificities and signaling hierarchies revisited. *Nat Immunol*. 2005;6(1):17–21.

Vorwerk S, Somerville S, Somerville C. The role of plant cell wall polysaccharide composition in disease resistance. *Trends Plant Sci*. 2004;9:203-209.

Wan J, Zhang XC, Neece D, Ramonell KM, Clough S, Kim SY, Stacey MG, Stacey G. A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in *Arabidopsis*. *Plant Cell*. 2008;20:471-481.

Wang M, Gu B, Huang J, Jiang S, Chen Y, Yin Y, *et al*. Transcriptome and proteome exploration to provide a resource for the study of *Agrocybe aegerita*. *PLoS ONE*. 2013;8(2):566-86.

Ward ER, Uknes SJ, Williams SC, Dincher SS, Wiederhold DL, Alexander DC, *et al*. Coordinate gene activity in response to agents that induce systemic acquired resistance. *Plant Cell*. 1991;3:1085-1094.

Wardlaw CW. Leaf Spot. (Sigatoka Disease). *Banana diseases: Including Plantains and Abaca*. 1961;11:314-341.

Wei ZM, Laby RJ, Zumoff CH, Bauer DW, He SY, Collmer A, Beer SV. Harpin, elicitor of the hypersensitive response produced by the plant pathogen *Erwinia amylovora*. *Science*. 1992;257:85-88.

Wicker T, Schlagenhauf E, Graner A, Close TJ, Keller B, Stein N. 454 sequencing put to the test using the complex genome of barley. *BMC Genomics*. 2006;7:275.

Wicker, T, Schlagenhauf E, Graner A, Close TJ, Keller B, Stein N. 454 sequencing put to the test using the complex genome of barley. *BMC genomics*. 2006;7: 275.

Wilmes A, Limonciel A, Aschauer L, Moenks K, Bielow C, Leonard MO, *et al*. Application of integrated transcriptomic, proteomic and metabolomic profiling for the delineation of mechanisms of drug induced cell stress. *J Proteomics*. 2013;79:180–94.

Wilson KA, Mcmanus MT, Gordon ME, Jordan TW. The proteomics of senescence in leaves of white clover, *Trifolium repens* (L.). *Proteomics*. 2002; 9:1114-22.

Wu TD, Watanabe CK. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics*. 2005;21(9):1859–75.

Xiao S. 2006 Current perspectives on molecular mechanisms of plant disease resistance, p.317-333 in *Floriculture, Ornamental and Plant Biotechnology: Advances and Topical Issues*, Vol. III, edited by J. Teixeira da Silva. Global Science Books, London.

Xiao S, Wang W, Yang X. Evolution of resistance genes in plants. Innate immunity of plants, animals, and humans. *Nucleic Acids and Molecular Biology*. 2008;21:1- 25.

Yang QS, Wu JH, Li CY, Wei YR, Sheng O, Hu CH, *et al.* Quantitative proteomic analysis reveals that antioxidation mechanisms contribute to cold tolerance in plantain (*Musa paradisiaca* L.; ABB Group) seedlings. *Mol Cell Proteomics*. 2012;11(12):1853–69.

Yin Y, Yu G, Chen Y, Jiang S, Wang M, Jin Y, *et al.* Genome-wide transcriptome and proteome analysis on different developmental stages of *Cordyceps militaris*. *PLoS ONE*. 2012;7(12):51853.

Zdobnov EM, Apweiler R. InterProScan: an integration platform for the signature-recognition methods in InterPro. *Bioinformatics*. 2001;17(9):847–8.

Zeng J, Liu Y, Liu W, Liu X, Liu F, Huang P, *et al.* Integration of transcriptome, proteome and metabolism data reveals the alkaloids biosynthesis in *Macleaya cordata* and *Macleaya microcarpa*. *PLoS ONE*. 2013;8(1):e53409.

Zhang LL, Feng RJ, Zhang YD. Evaluation of different methods of protein extraction and identification of differentially expressed proteins upon ethylene-induced early-ripening in banana peels. *J Sci Food Agric*. 2012;92(10):2106–15.

Zhang J, Shao F, Li Y, Cui H, Chen L, Li H, *et al.* A *Pseudomonas syringae* effector inactivates MAPKs to suppress PAMP-induced immunity in plants. *Cell Host Microbe*. 2007;1:175-185.

Zimmerman A. Über einige tropischer Kulturpflanzen beobachtete Pilze. *Zentralblatt für Bakteriologie, Parasitenkunde, Infektionskrankheiten und Hygiene*. 1902;8:219.

Zipfel C, Kunze G, Chinchilla D, Caniard A, Jones JD, Boller T, Felix G. Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts *Agrobacterium*-mediated transformation. *Cell*. 2006;125:749-760.

## 8. ANEXOS

**Anexo 1:** Passos MAN, de Cruz VO, Emediato FL, de Teixeira CC, Azevedo VCR, Brasileiro ACM, *et al.* Analysis of the leaf transcriptome of *Musa acuminata* during interaction with *Mycosphaerella musicola*: gene assembly, annotation and marker development. *BMC Genomics*. 2013;14:78.

Passos *et al.* *BMC Genomics* 2013, **14**:78  
<http://www.biomedcentral.com/1471-2164/14/78>



**RESEARCH ARTICLE**

**Open Access**

# Analysis of the leaf transcriptome of *Musa acuminata* during interaction with *Mycosphaerella musicola*: gene assembly, annotation and marker development

Marco A N Passos<sup>1</sup>, Viviane Oliveira de Cruz<sup>1</sup>, Flavia L Emediato<sup>1</sup>, Cristiane Camargo de Teixeira<sup>2</sup>, Vânia C Rennó Azevedo<sup>3</sup>, Ana C M Brasileiro<sup>3</sup>, Edson P Amorim<sup>4</sup>, Claudia F Ferreira<sup>4</sup>, Natalia F Martins<sup>3</sup>, Roberto C Togawa<sup>3</sup>, Georgios J Pappas Júnior<sup>1</sup>, Orzenil Bonfim da Silva Jr<sup>3</sup> and Robert NG Miller<sup>1\*</sup>

**Anexo 2:** Passos MAN, de Oliveira Cruz V, Emediato FL, de Camargo Teixeira C, Souza MT Jr, Matsumoto T, *et al.* Development of expressed sequence tag and expressed sequence tag-simple sequence repeat marker resources for *Musa acuminata*. *AoB Plants*. 2012;2012:pls030.



**AoB PLANTS**

<http://aobplants.oxfordjournals.org/>

**Open access – Research article**

**THIS ARTICLE IS PART OF A SPECIAL ISSUE ENTITLED  
'MOLECULAR TECHNOLOGIES TO IMPROVE  
VEGETATIVELY PROPAGATED BANANA AND CASSAVA'**

## **Development of expressed sequence tag and expressed sequence tag-simple sequence repeat marker resources for *Musa acuminata***

Marco A. N. Passos<sup>1</sup>, Viviane de Oliveira Cruz<sup>1</sup>, Flavia L. Emediato<sup>1</sup>, Cristiane de Camargo Teixeira<sup>2</sup>, Manoel T. Souza Jr<sup>3</sup>, Takashi Matsumoto<sup>4</sup>, Vânia C. Rennó Azevedo<sup>3</sup>, Claudia F. Ferreira<sup>5</sup>, Edson P. Amorim<sup>5</sup>, Lucio Flavio de Alencar Figueiredo<sup>1</sup>, Natalia F. Martins<sup>3</sup>, Maria de Jesus Barbosa Cavalcante<sup>6</sup>, Franc-Christophe Baurens<sup>7</sup>, Orzenil Bonfim da Silva Jr<sup>3</sup>, Georgios J. Pappas Jr<sup>1,3</sup>, Luc Pignolet<sup>8</sup>, Catherine Abadie<sup>8</sup>, Ana Y. Ciampi<sup>3</sup>, Pietro Piffanelli<sup>7,9</sup> and Robert N. G. Miller<sup>1\*</sup>

**Anexo 3:** Passos, MAN, Emediato, FL, Cruz, VO, Teixeira, CC, de Alencar Figueiredo, LF, Martins, NF, Togawa, RC, Costa, MMC, Silva Jr, O, Pappas Jr, GJ and Miller, RNG. 2013. Understanding plant immunity: transcriptome profiling in *Musa*-pathogen interactions using next generation sequencing. Acta Hort. (ISHS) 986:227-240.

### **Understanding Plant Immunity: Transcriptome Profiling in *Musa*-Pathogen Interactions Using Next Generation Sequencing**

M.A.N. Passos<sup>1</sup>, F.L. Emediato<sup>1</sup>, V.O. Cruz<sup>1</sup>, C. de Camargo Teixeira<sup>2</sup>, L.F. de Alencar Figueiredo<sup>1</sup>, N.F. Martins<sup>3</sup>, R.C. Togawa<sup>3</sup>, M.M.C. Costa<sup>3</sup>, O. Silva Jr.<sup>3</sup>, G.J. Pappas Jr.<sup>2,3</sup> and R.N.G. Miller<sup>1</sup>

**Anexo 4:** Emediato, FL, Passos, MAN, Teixeira, CC, Pappas Jr., GJ and Miller, RNG. 2013. Analysis of expression of nbs-lrr resistance gene analogs in *Musa acuminata* during compatible and incompatible interactions with *Mycosphaerella musicola*. Acta Hort. (ISHS) 986:255-258

**Analysis of Expression of NBS-LRR Resistance Gene Analogs in *Musa acuminata* during Compatible and Incompatible Interactions with *Mycosphaerella musicola***

F.L. Emediato<sup>1</sup>, M.A.N. Passos<sup>1</sup>, C. de Camargo Teixeira<sup>2</sup>,  
G.J. Pappas Jr.<sup>2,3</sup> and R.N.G. Miller<sup>1</sup>

**Anexo 5:** Passos, MAN; Pappas Júnior, Georgios J; Emediato, FL; Teixeira, CC ; Miller, RNG. Identificação e caracterização de componentes genéticos de resistência e defesa aos patógenos *Mycosphaerella fijiensis* e *Mycosphaerella musicola* em *Musa* spp.. In: Congresso Brasileiro de Fitopatologia, 2010, Cuiabá. 43º Congresso Brasileiro de Fitopatologia, 2010.

## 07.040

**Identificação e caracterização de componentes genéticos de resistência e defesa aos patógenos *Mycosphaerella fijiensis* e *Mycosphaerella musicola* em *Musa* spp.** Passos, M. A. N.<sup>1</sup>; Pappas Jr., G. J.<sup>2</sup>; Emediato, F. L.<sup>1</sup>; Teixeira, C. C.<sup>3</sup>; Miller, R. N. G.<sup>1</sup> - <sup>1</sup>Universidade de Brasília - Biologia Celular; <sup>2</sup>Embrapa Recursos Genéticos e Biotecnologia - Laboratório de Bioinformática; <sup>3</sup>Universidade Católica de Brasília - Ciências Genômicas e Biotecnologia. *E-mail: marconinomia@gmail.com*. Identification and characterization of genetic components of resistance and defense against pathogens *Mycosphaerella fijiensis* e *Mycosphaerella musicola* in *Musa* spp.

O desenvolvimento de programas de melhoramento genético possibilita a obtenção de genótipos resistentes a fitopatógenos. Nesse contexto, foi conduzida a análise da expressão gênica diferencial em folhas de cultivares contrastantes de *Musa acuminata* Calcutta 4 e *M. acuminata* Cavendish inoculadas in vitro com *Mycosphaerella fijiensis*, causador da Sigatoka Negra. O seqüenciamento foi realizado em 8000 clones. A análise diferencial in silico demonstrou que 220 genes foram diferencialmente expressos, sendo que mais de 20 deles estão relacionados à resistência e defesa, incluindo metalotioneína, germin-like proteínas, ferredoxinas, glutatona S-transferases, proteínas relacionadas a patogênese, tioredoxina, glicolato oxidase, proteínas responsivas a etileno, superoxidases dismutases, proteínas responsivas a stress e proteínas responsivas a auxina. Bioensaios foram conduzidos in planta inoculados com *M. fijiensis* e *M. musicola*. A utilização de seqüenciamento massal está sendo realizada visando o aumento da resolução da análise de genes diferencialmente expressos. **Apoio Financeiro:** IAEA, CNPq, UCB, FINEP.

*Tropical Plant Pathology 35 (Suplemento), agosto 2010*



**Anexo 6:** Passos, MAN; Emediato, FL; Teixeira, CC; Franco, OC ; Miller, RNG. A protein extraction method compatible with proteomic analysis of the interaction between *Musa acuminata* and *Mycosphaerella musicola*. In: International ISHS/ProMusa Symposium - Bananas and Plantains: Toward sustainable global production and improved uses, 2011, Salvador. International ISHS/ProMusa Symposium - Bananas and Plantains: Toward sustainable global production and improved uses. Leuven, 2011.

## **A Protein Extraction Method Compatible with Proteomic Analysis of the Interaction between *Musa acuminata* and *Mycosphaerella musicola***

M.A.N. Passos<sup>1</sup>, F.L. Emediato<sup>1</sup>, C.C. Teixeira<sup>2</sup>, O.C. Franco<sup>2</sup> and R.N.G. Miller<sup>1</sup>

<sup>1</sup>Universidade de Brasília/ IB/BIOCEL/ CEP 70910-900, Brasília DF, Brazil;

<sup>2</sup>Universidade Católica de Brasília/ Pós-graduação em Ciências Genômicas e Biotecnologia/ SGAN 916, Módulo B/ CEP 70790-160, Brasília, DF, Brazil

**Keywords:** Biotic stress, Yellow Sigatoka, disease resistance, proteomics

Sigatoka leaf spot, caused by *Mycosphaerella musicola*, reduces functional leaf area and causes premature fruit ripening. The wild diploid cultivar *Musa acuminata* ssp. *burmannicoides* 'Calcutta-4' is used as a source of resistance in breeding programs and serves as a model for comparative genomics, functional genomics and candidate resistance gene discovery. In this context, a proteomics study was undertaken in order to identify protein expression in the *M. acuminata* – *M. musicola* interaction. Total proteins were extracted from leaf material for 'Calcutta-4' (resistant) and Cavendish cultivar 'Grande Naine' (AAA, susceptible), both uninoculated and inoculated with *M. musicola* conidiospores. The extraction methodology was optimized, using acetone to remove leaf pigments, followed by 10% TCA/acetone to remove further pigments and precipitate total proteins, plus a second exposure to 10% TCA/acetone to remove water and precipitate soluble proteins. A solubilization buffer was used to assist in the separation of phases in the extraction with phenol. After dissolving the material, proteins were precipitated in 75% TCA. Isoelectric focusing (1st dimension) was performed using the IPGphor system, in which proteins are separated by electrophoretic migration in a pH gradient. SDS-PAGE (2nd dimension) was performed to separate proteins by size. Proteomic analyses of this interaction are currently ongoing. Results will provide important information on the profiles of expressed proteins and post-



**Anexo 7:** Passos, MAN; Pappas Júnior, Georgios J; Emediato, FL; Teixeira, CC ; Miller, RNG. Identificação e caracterização de componentes genéticos de resistência a *Mycosphaerella musicola* em *Musa acuminata* In: Congresso Brasileiro de Fitopatologia, 2012, Manaus. 45º Congresso Brasileiro de Fitopatologia, 2012.



*Tropical Plant Pathology* 38 (Suplemento), agosto 2012  
45º Congresso Brasileiro de Fitopatologia - Manaus, AM  
Copyright the Brazilian Phytopathological Society <http://www.sbfito.com>

## MELHORAMENTO GENÉTICO

470

### **Identificação e caracterização de componentes genéticos de resistência a *Mycosphaerella musicola* em *Musa acuminata*.**

(Identification and characterization of genetic components of resistance to *Mycosphaerella musicola* in *Musa acuminata*.)

**Passos, M.A.N.<sup>1</sup>, Pappas J.R., Silva J.R., O.<sup>2</sup>, Emediato, F.L.<sup>1</sup>, Teixeira, C.C.<sup>3</sup>, Miller, R.N.G.<sup>1</sup>**

<sup>1</sup>Universidade de Brasília, – <sup>2</sup>EMBRAPA Recursos Genéticos e Biotecnologia. <sup>3</sup> Universidade Católica de Brasília *E-mail*: marconinomia@gmail.com

**Anexo 8:** Miller, RNG; Passos, MAN; Emediato, FL; Teixeira, CC, Silva Jr, O; Pappas Jr, Georgios J. Gene expression analysis in *Musa acuminata*- a *Mycosphaerella* interactions. In: Congresso Brasileiro de Fitopatologia, 2012, Manaus. 45º Congresso Brasileiro de Fitopatologia, 2012.



*Tropical Plant Pathology* 37 (Suplemento), agosto 2012  
45º Congresso Brasileiro de Fitopatologia - Manaus, AM  
Copyright the Brazilian Phytopathological Society <http://www.sbfito.com.br>

**Gene expression analysis in *Musa acuminata*-*Mycosphaerella* interactions**  
(Análise de expressão genética durante as interações *Musa acuminata*-*Mycosphaerella*)

**Miller, R.N.G.<sup>1</sup>, Passos, M.A.N.<sup>1</sup>, Emediato, F.L.<sup>1</sup>, Teixeira, C.C.<sup>2</sup>, Silva Jr, O<sup>3</sup>; Pappas Jr, G.J.<sup>3</sup>.**

<sup>1</sup>Universidade de Brasília, <sup>2</sup>Universidade Católica de Brasília, <sup>3</sup>EMBRAPA Recursos Genéticos e Biotecnologia. *E-mail*: robertmiller@unb.br.

**Anexo 9:** Passos, MAN; Pappas Jr, Georgios J; Silva Jr, O; Emediato, FL; Teixeira, CC; Franco, OL; Miller, RNG. Characterization of genetic components of resistance to *Mycosphaerella musicola* in *Musa acuminata*. In: Simpósio Brasileiro de Genética Molecular de Plantas, 2013, Bento Gonçalves. IV Simpósio Brasileiro de Genética Molecular de Plantas, 2013.



## CHARACTERIZATION OF GENETIC COMPONENTS OF RESISTANCE TO *MYCOSPHAERELLA MUSICOLA* IN *MUSA ACUMINATA*

**Passos, MAN<sup>1</sup>; Pappas Jr, GJ<sup>1</sup>; Silva JR, O<sup>2</sup>; Emediato, FL<sup>1</sup>; Teixeira, CC<sup>3</sup>; Franco, OL<sup>3</sup>; Miller, RNG<sup>1</sup>.**

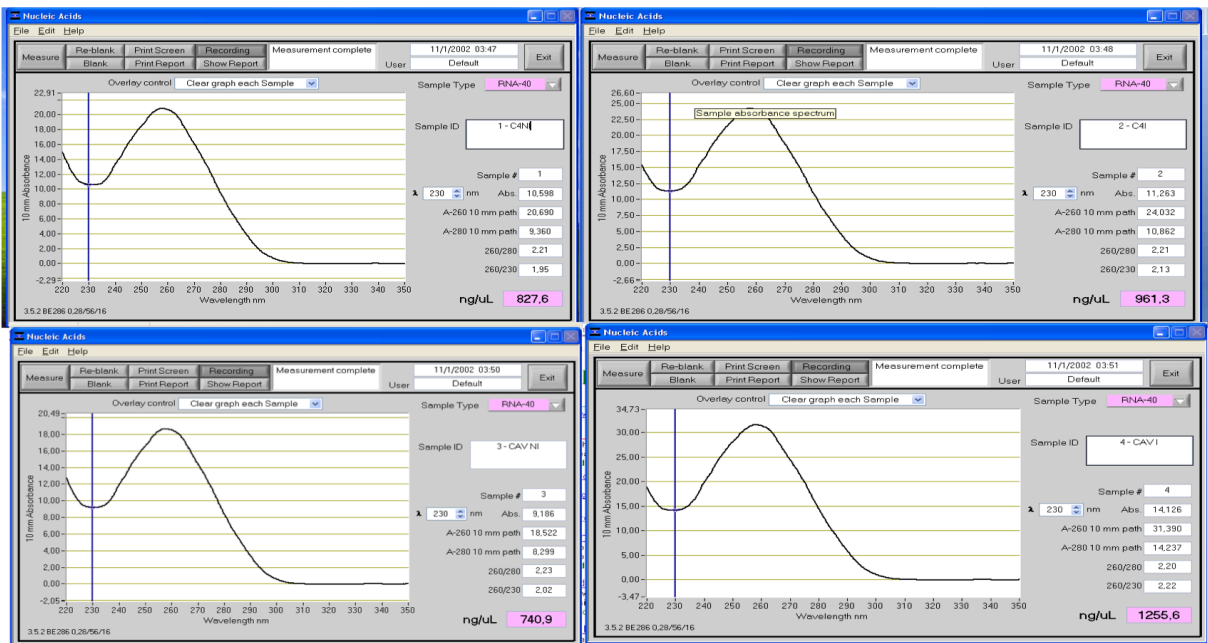
<sup>1</sup>Universidade de Brasília/ IB/BIOCEL/ CEP 70910-900, Brasília DF, Brasil. <sup>2</sup>EMBRAPA Recursos Genéticos e Biotecnologia, Parque Estação Biológica, CP 02372, CEP 70.770-900, Brasília, DF, Brazil. <sup>3</sup>Universidade Católica de Brasília/ Pós-graduação em Ciências Genômicas e Biotecnologia/ SGAN 916, Módulo B/ CEP 70790-160, Brasília, DF, Brasil.

\*E-mail: marconinomia@gmail.com.

**Anexo 10:** Análise da qualidade e a quantificação do RNA verificadas pelo NanoDrop, analisando as razões de absorbância 260/280 e 260/230 nos quatro *pools*: Calcutta 4 não-inoculado (C4NI), Calcutta 4 inoculado (C4I), Cavendish não inoculado (CAVNI) e Cavendish inoculado (CAVI).

Sample ID	Date	ng/ul	A260	A280	260/280	260/230	Constant	Cursor Pos.	Cursor abs.	340 raw
1 - C4NI	05/24/10	827,6	20,69	9,36	2,21	1,95	40,00	230,00	10,60	-0,02
2 - C4I	05/24/10	961,2	24,03	10,86	2,21	2,13	40,00	230,00	11,26	-0,02
3 - CAVNI	05/24/10	740,8	18,52	8,30	2,23	2,02	40,00	230,00	9,19	0,01
4 - CAVI	05/24/10	1.255	31,39	14,24	2,20	2,22	40,00	230,00	14,13	0,05

Sample ID	$\mu$ l	total ( $\mu$ g)
1 - C4NI	72,50 $\mu$ l	60
2 - C4I	62,40 $\mu$ l	60
3 - CAVNI	85,00 $\mu$ l	60
4 - CAVI	47,80 $\mu$ l	60



**Anexo 11:** Dados do mapeamento de *M. acuminata* Calcutta 4 e Cavendish Grande Naine Unigene alinhados com o genoma referência *M. acuminata* DH Pahang.

<http://www.biomedcentral.com/content/supplementary/1471-2164-14-78-S1.xlsx>

**Anexo 12:** Seqüências consenso e análise BLASTX de contigs unigenes montados de *M. acuminata* Calcutta 4 e Cavendish Grande Naine.

<http://www.biomedcentral.com/content/supplementary/1471-2164-14-78-S4.xlsx>

**Anexo 13:** Análise de unigene montados de *M. acuminata* Calcutta 4 e Cavendish Grande Naine pelo Gene Ontology (GO). Contigs montados de *M. acuminata* Calcutta 4 e Cavendish Grande Naine foram anotados para os termos GO relacionados com a defesa.

<http://www.biomedcentral.com/content/supplementary/1471-2164-14-78-S6.xlsx>