

CASLEY BORGES DE QUEIROZ

**USE OF THE IRAP MARKER TO STUDY GENETIC VARIABILITY IN
Pseudocercospora fijiensis POPULATIONS**

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

VIÇOSA
MINAS GERAIS- BRASIL
2013

Ficha catalográfica preparada pela Seção de Catalogação e
Classificação da Biblioteca Central da UFV

T

Q3u
2013

Queiroz, Casley Borges de, 1987-

Use of the irap marker to study genetc variability in
Pseudocercospora fijiensis populations / Casley Borges de
Queiroz. – Viçosa, MG, 2013.

v, 37f. : il. (algumas color.) ; 29cm.

Texto em inglês e português

Orientador: Marisa Vieira de Queiroz

Dissertação (mestrado) - Universidade Federal de Viçosa.

Inclui bibliografia.

1. *Mycosphaerella fijiensis*. 2. Polimorfismo (Genética).
3. Genética. 4. Sigatoka-negra. 5. Banana. I. Universidade
Federal de Viçosa. Departamento de Microbiologia.
Programa de Pós-Graduação em Microbiologia Agrícola.
II. Título.

CDD 22. ed. 632.46

CASLEY BORGES DE QUEIROZ

**USE OF THE IRAP MARKER TO STUDY GENETIC VARIABILITY IN
Pseudocercospora fijiensis POPULATIONS**

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

APROVADA: 22 de março de 2013

Andrea de Oliveira Barros Ribon

Eduardo Seiti G. Mizubuti
(Coorientador)

Marisa Vieira de Queiroz
(Orientadora)

AGRADECIMENTOS

Primeira mente agradeço a Deus pela saúde, força e determinação para superar os obstáculos e atingir meus objetivos.

À minha família pela compreensão, e todo apoio durante minha jornada.

À professora Marisa Vieira de Queiroz, por ter me orientado, por todo apoio, incentivo, ensinamentos e por toda a dedicação na pesquisa e na formação de seus alunos.

Ao Dr. Gilvan Ferreira da Silva, pelos os ensinamentos, por todo apoio e incentivo para que eu conseguisse carreira na pesquisa.

Ao professor Eduardo S. Gomide Mizubuti pelas valiosas sugestões na elaboração do artigo desse trabalho.

Ao colega de laboratório Mateus Santana pela amizade, ensinamentos e por toda a contribuição na execução deste trabalho.

Aos colegas de laboratório pela excelente convivência e amizade.

Ao Departamento de Microbiologia da Universidade Federal de Viçosa, pela oportunidade de realização do curso de Mestrado.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, pela concessão da bolsa de estudo de Mestrado.

À Empresa Brasileira de Pesquisa Agropecuária – Embrapa Amazônia Ocidental, pela concessão dos isolados.

A todos aqueles que me apoiaram e torceram por mim.

Muito obrigado!

SUMÁRIO

RESUMO	iv
ABSTRACT	v
INTRODUÇÃO GERAL	1
REFERÊNCIAS	7
ARTIGO.....	11
<i>Abstract</i>	12
Introduction	13
Materials and Methods	15
<i>Acquisition of Pseudocercospora fijiensis isolates</i>	15
<i>DNA extration</i>	15
<i>“In silico” analysis of Class I transposons present in the genome</i>	15
<i>IRAP markers</i>	16
<i>Data analysis</i>	17
Results.....	17
Discussion	19
Acknowledgements	21
References	22

RESUMO

QUEIROZ, Casley Borges, M.Sc., Universidade Federal de Viçosa, março de 2013. **Use of the IRAP marker to study genetic variability in *Pseudocercospora fijiensis* populations.** Orientadora: Marisa Vieira de Queiroz. Coorientadores: Elza Fernandes de Araujo e Eduardo Seiti Gomide Mizubuti.

Devido à ausência de estudo de caracterização da variabilidade genética das populações de *Pseudocercospora fijiensis* recentemente introduzidas no Brasil, o objetivo desse trabalho foi avaliar a adequabilidade do marcador IRAP para estudar variações genéticas entre indivíduos, bem como determinar a estrutura genética da população brasileira de *P. fijiensis* com base no *fingerprinting* gerado por amplificação de polimorfismo entre retrotransposons (IRAP). Um total de 22 locos foi amplificado, sendo 77.3 % polimórficos. A análise de agrupamento revelou dois principais grupos no Brasil. A diversidade gênica (H_E) foi de 0.22 e pela análise de variância molecular verificou-se que a maior variabilidade genética está dentro das populações. A Análise Discriminante de Componente Principal (DAPC) revelou que não há nenhuma estruturação relacionada com as origens geográficas e cultivo hospedeiro. O sistema de marcador baseado em retrotransposon IRAP é ferramenta apropriada para estudar a variabilidade genética em *P. fijiensis*.

ABSTRACT

QUEIROZ, Casley Borges, M.Sc., Universidade Federal de Viçosa, march 2013. **Use of the IRAP marker to study genetic variability in *Pseudocercospora fijiensis* populations.** Adviser: Marisa Vieira de Queiroz. Co-advisers: Elza Fernandes de Araujo and Eduardo Seiti Gomide Mizubuti.

Due to the lack of characterization study of genetic variability in populations of *Pseudocercospora fijiensis* recently introduced in Brazil, the objective of this study was to evaluate the suitability of IRAP marker for studying genetic variations between individuals, and to determine the genetic structure of the population of *P. fijiensis* based on fingerprinting generated by inter-retrotransposons amplified polymorphism (IRAP). A total of 22 loci were amplified and 77.3% showed a polymorphism. Cluster analysis revealed two major groups in Brazil. The observed genetic diversity (H_E) was 0.22, and through molecular analysis of variance, it was determined that the greatest genetic variability occurs within populations. The Discriminant Analysis of Principal Components (DAPC) revealed no structuring related to the geographical origin of culture of the host. The IRAP-based marker system is a suitable tool for the study of genetic variability in *P. fijiensis*.

INTRODUÇÃO GERAL

O fungo *Mycosphaerella fijiensis* M. Morelet (*Pseudocercospora fijiensis* (M. Morelet) Deighton), é o agente causal da sigatoka-negra, sendo esta considerada atualmente a doença mais destrutiva que ocorre em bananeiras no mundo (Churchill, 2011). Esse fungo produz seus estádios anamórfico (assexuado) e telomórfico (sexuado) sobre a mesma lesão. As estruturas anamórficas de *P. fijiensis* estão presentes durante a fase de estrias ou manchas jovens da doença, onde se observam frutificações anfígenas predominantemente na face abaxial. Os conidióforos surgem nos estádios iniciais da doença, emergindo dos estômatos, possuem coloração verde clara a marrom e são septados, podendo ser ramificados. Na fase sexual, os espermagônios são anfígenos e são formados logo após o coalescimento das lesões (Gasparotto et al., 2006; Churchill, 2011).

Os ascósporos são os principais esporos para a disseminação de patógenos a longas distâncias. Porém, por sua longa sobrevivência sobre as mais diversas superfícies, os conídios também podem ser inoculo capaz de dispersar a média e longa distâncias (Amil et al., 2007; Churchill, 2011), se medidas legislativas e/ou sanitárias não forem tomadas ou seguidas adequadamente em tempo hábil. O vento é o principal meio de dispersão dos esporos, principalmente os ascósporos (Churchill, 2011). A concentração de inóculo até quatro quilômetros de distância do bananal é idêntica à concentração de inóculo no ar próximo ao filoplano ou superfície das folhas. Ascósporos podem ser encontrados ou capturados até 60 quilômetros de distância do bananal (Gasparotto et al., 2006).

A doença foi descrita pela primeira vez, em 1963 no Vale da Sigatoka, na ilha de Vitu Levu em Fiji. Atualmente a doença se encontra disseminada em toda a América Central, grande parte da América do Sul, Ásia, Oceania e África. Na Oceania a doença foi relatada na Polinésia em 1964. Na Ásia, em Filipinas e Singapura, em 1964; em Zâmbia, África, em 1963; no continente americano foi constatada em Honduras em 1972, ocasionando uma epidemia. Em 1981, foi constatada no México e pela primeira vez na América do Sul, na Colômbia. Em 1991, na Venezuela; em 1994 no Peru e na Bolívia em 1997 (Carlier et al., 2000). No Brasil, foi constatada em fevereiro de 1998

nos municípios de Tabatinga e Benjamin Constant, no Estado do Amazonas (Pereira et al., 1998). Atualmente a sigatoka-negra está presente nos estados do Acre, Amazonas, Roraima, Amapá, Roraima, Pará, Tocantins, Mato Grosso, Mato grosso do sul, Minas Gerais, São Paulo, Paraná, Santa Catarina e Rio Grande do Sul (Ministério da Agricultura Pecuária e Abastecimento, 2012).

Nas regiões onde há presença da sigatoka-negra, devido à maior agressividade de seu agente etiológico e também em decorrência da alta gama de cultivares de bananeiras suscetíveis, esta doença supera rápida e muito eficientemente a sigatoka-amarela (Vargas, 1996). As perdas devidas à doença podem atingir 100% da produção. O maior impacto da sigatoka negra ocorre nos países subtropicais, onde a banana e plátanos são a base alimentar de pessoas com menor poder aquisitivo sem recursos técnicos e financeiros para utilização dos fungicidas, uma vez que estes produtos químicos são de alto custo (Marín et al., 2003; Churchill, 2011). Apesar de a doença ser um sério problema, nos plantios comerciais que atendem os consumidores de maior poder aquisitivo, é controlada com aplicação de fungicidas. Para o seu controle efetivo, em regiões quentes e úmidas onde a doença já ocorre como na América Central, são necessárias 26 a 52 pulverizações por ano com fungicidas (Gasparoto et al., 2006). Dessa forma, os custos para o controle da doença por meio de fungicidas são incompatíveis com o cultivo dos plátanos em razão da produtividade ser muito baixa (Gasparoto et al., 2006). Apesar de a doença afetar a qualidade dos frutos, muitos produtores comercializam a banana no mercado interno. Assim o cultivo da banana por esses produtores torna-se uma prática de subsistência.

Em virtude da recente introdução de *P. fijiensis* no Brasil e as condições climáticas aqui encontradas, à análise molecular da variabilidade é necessária e crucial para a determinação da estrutura genética da população para direcionar as estratégias de melhoramento visando à resistência e até mesmo o manejo de fungicida em agroecossistemas (McDonald & Linde, 2002).

Atualmente a análise de elementos transponíveis tem contribuído para o conhecimento da organização de genomas de vários organismos como em plantas (Kalendar et al., 2011), insetos (Gomulski et al., 2004), animais (Marín, 2010; Tollis e Boissinot, 2011) e fungos (Bouvet et al., 2008). Baseado nas suas características

estruturais, os elementos transponíveis podem ser classificados em ordem hierárquica em nível de classe, subclasse, ordem, superfamília, família e subfamília. O nível mais alto (classe) divide os elementos transponíveis pela presença ou ausência de uma transposição intermediada por RNA. Assim, todos os elementos transponíveis pertencentes à Classe I transpõem via um RNA intermediário que é transcrito de uma cópia do transposon do genoma e um cDNA é feito via uma transcrição reversa realizada por uma proteína codificada pelo próprio elemento e chamada de transcriptase reversa. Isso significa que cada evento de transposição cria uma nova cópia do transposon, enquanto a cópia original permanece na região doadora. Consequentemente, os elementos da classe I são os maiores contribuidores da larga fração de DNA repetitivo nos genomas (Kumar & Bennetzen, 1999). Na Classe I estão presentes os retroelementos com LTR (do inglês *Long Terminal Repeats*) também chamados de retrotransposons por terem estrutura semelhante aos retrovírus. Os retrotransposons podem ser divididos em cinco ordens baseado nas suas características e mecanismo, organização e a filogenia da transcriptase reversa: retrotransposons LTR, elementos DIRS-like, elementos Penelope-like, LINEs e SINEs. Os retrotransposons LTR variam de algumas centenas de pares de base até excepcionalmente 22 kb (elemento *Ogre* encontrado na ervilha por Neumann et al., 2003). As LTRs dos elementos variam de umas poucas centenas de bases a mais de 5 kb e começam com 5'-TG-3' e terminam com 5'-CA-3'. Os retrotransposons também possuem regiões análogas à gag (que codifica as proteínas da cápsula viral) e pol (que codifica as enzimas transcriptase reversa, RNase H, integrase e protease). A protease (AP) cliva a poliproteína liberando pequenos produtos de proteínas; a transcriptase reversa (RT) produz um cDNA do retrotransposon usando um RNA como molde; a RNase H (RH) separa e degrada o RNA molde do híbrido DNA-RNA e a integrase (INT) com domínio DDE catalisa a inserção do cDNA no genoma hospedeiro. As duas superfamílias *Gypsy* e *Copia*, se diferenciam pelas posições da transcriptase reversa e da integrase na região (GAG, AP, RT, RH e INT) e (GAG, AP, INT, RT e RH), respectivamente. Os retrotransposons LTR apresentam um mecanismo de integração que leva a duplicação de uma pequena região de DNA dentro do sítio de integração. Devido a esse processo, os elementos transponíveis são flanqueados por duplicações do sítio alvo de 4-6 pb chamada de TSD

(*Target Site Duplications*) que possuem tamanhos conservados, sendo que, algumas vezes, a própria sequência do sítio alvo é conservada. Todas as outras superfamílias nas outras ordens de LTR usam um mecanismo de transposição similar (Wicker et al., 2007; Muszewska et al., 2011).

Os outros grupos da classe I foram descritos mais recentemente, tais como DIRS-like (Cappello et al., 1985) e Penelope-like (Evgen'ev et al., 1997). Os membros da ordem DIRS contêm um gene da tirosina recombinase no lugar de uma integrase. Embora esses elementos sejam flanqueados por repetições terminais, essas sequências são diferentes das LTRs típicas, sendo então repetições diretas denominadas de SDR (*Split Direct Repeats*) ou repetições invertidas. Essas características indicam um mecanismo de integração diferente dos elementos LTRs e LINEs. DIRS-like integram sem criar duplicações no sítio-alvo (TSD).

Os elementos da ordem *Penelope-like* codificam apenas duas proteínas, uma endonuclease e uma transcriptase reversa, sendo que esta última é mais relacionada filogeneticamente com a telomerase do que a transcriptase reversa dos retrotransposons LTR ou LINEs (Evgen'ev et al., 1997). Alguns membros dessa ordem podem ter sequências *LTR-like* que podem ter orientações diretas ou inversas.

Os membros da ordem LINE (*Longs Interspersed Nuclear Elements*) não possuem LTRs, seus tamanhos variam em algumas kilobases e são encontrados em todos eucariotos onde foram analisados. Os LINEs codificam pelo menos uma transcriptase reversa e uma nuclease para sua transposição, assim são também considerados elementos autônomos. Alguns elementos LINE possuem uma ORF *gag-like* na região 5' da poliproteína, mas sua função permanece desconhecida. Em suas extremidades 3', podem apresentar uma cauda poli (A), uma repetição em tandem ou uma região rica em adenina.

Os elementos da ordem SINE (*Shorts Enterspersed Nuclear Elements*) são menores que os outros retrotransposons, variando em tamanho de 80 a 500 pb. Na maioria dos casos, os SINE são flanqueados por curtas repetições diretas do DNA hospedeiro de 5 a 15 pb (TSDs). Como são retrotransposons não-autônomos, precisam completamente da maquinaria celular e das proteínas dos retrotransposons autônomos

para sua replicação; um promotor interno na extremidade 5' permite sua transcrição pela RNA polimerase III.

Elementos da Classe II possuem pequenas repetições terminais invertidas e codificam uma transposase. Normalmente aparecem em baixo número de cópias. São também encontrados em procariotos em formas simples chamadas de sequências de inserção (IS - *Insertion Sequences*) ou como parte de estruturas mais complexas. Os elementos da Classe II se movem para outra região do genoma via excisão (processo também chamado de “corte-cola”) do próprio elemento realizada pela transposase. No entanto, esses elementos podem aumentar seu número de cópias por transposição se durante a replicação do cromossomo o elemento já replicado for integrado em uma região ainda não replicada. Alternativamente, podem ser usados como molde para reparo de *gaps*, formando assim uma nova cópia (Nassif et al., 1994).

Atualmente, os elementos transponíveis são utilizados como marcadores genéticos, agentes mutagênicos e vetores para transformação (Gertz et al., 2012; Kozeretska et al., 2011; Smýkal et al., 2011). Em fungos fitopatogênicos os elementos transponíveis têm sido amplamente explorados para traçar o perfil de populações de diferentes áreas geográficas, auxiliando na compreensão da estrutura populacional e epidemiologia de doenças (Pereira et al., 2006). Isso é possível em virtude de que cada evento de transposição gera uma inserção polimórfica que pode ser identificada usando várias técnicas moleculares. Os elementos transponíveis apresentam um grande número de cópias que podem gerar centenas ou até milhares de marcas. A presença ou a ausência do elemento transponível em um dado locus pode ser usado como um marca para *fingerprinting* e estudo de diversidade (Grzebelus, 2006). A estrutura e a estratégia de replicação dos retrotransposons atribuem a eles vantagens para serem usados como marcadores. Primeiro, os retrotransposons contêm longas sequências, definidas e conservadas que podem ser usadas para clonagem de marcas específicas. Segundo, a atividade de replicação dos retrotransposons pode produzir novas inserções no genoma aumentando o polimorfismo.

Uma das técnicas de análise de polimorfismo utilizada e baseada em retrotransposons é a S-SAP (*Sequence Specific Amplification Polymorphism*), proposta por Waugh et al. (1997). Nessa técnica é realizada a amplificação de uma região entre

um sítio próximo a uma extremidade de elemento transponível e um sítio de restrição. No primeiro passo da técnica, ocorre a digestão do DNA total com enzimas de restrição que não clivam o elemento transponível. A seguir, ocorre a ligação de adaptadores. Os *primers* utilizados para amplificação são pertencentes a região dos adaptadores e a região conservada do LTR do retrotransposon. O produto da amplificação pode ser então comparado entre os diferentes indivíduos, sendo que a presença de banda indica presença de elemento transponível e a ausência de bandas significa ausência de transposon num determinado locus (Waugh et al., 1997). A vantagem da sua utilização é o alto nível de polimorfismo apresentado.

Existe também a técnica RIBIP que foi desenvolvida usando o retrotransposon PDR1 (Flavell et al., 1998). Requer sequências 5' e 3' de regiões que flanqueiam a inserção do transposon. As amplificações podem ser analisadas usando a eletroforese em gel de agarose convencional, ou por hibridização. É uma técnica mais cara e complicada do que os outros métodos baseados em transposons.

Técnicas mais simples foram desenvolvidas utilizando sequências dos LTR de retrotransposons para a detecção de polimorfismo. IRAP (*Inter-Retrotransposon Amplified Polymorphism*) foi primeiro descrito por Kalendar et al. (1999). Esta técnica é baseada na amplificação de regiões entre dois retrotransposons. O polimorfismo é detectado por presença ou ausência de produtos da PCR (bandas). A ausência de uma banda indica a ausência do retrotransposon em um locus particular. As amplificações podem ser visualizadas por eletroforese em gel de agarose convencional ou em poliacrilamida. Essa técnica também possui a vantagem de amplificar um alto número de bandas devido a abundância de retroelementos e a capacidade de criarem novas cópias, e não requer o uso de enzimas de restrição e nem de adaptadores.

REFERÊNCIAS

- Amil AF, Heaney SP, Stanger C, Shaw MW (2008) Dynamics of QoI sensitivity in *Mycosphaerella fijiensis* in Costa Rica during 2000 to 2003. *Phytopathology* 97:1451–1457
- Bouvet GF, Jacobi V, Plourde K V, Bernier L (2008) Stress-induced mobility of OPHIO1 and OPHIO2, DNA transposons of the Dutch elm disease fungi. *Fungal Genet Biol* 45:565–578
- Cappello J, Handelsman K, Lodish H (1993) Sequence of *Dictyostelium DIRS-1*: an apparent retrotransposon with inverted terminal repeats and an internal circle copies of the putative R region. *Nucleic Acids Res* 21:2117–2123
- Carlier J, Fouré E, Gauhl F, Jones DR, Lepoivre P, Mourichon X, Pasberg-Gauhl C, Romero RA (2000) Black leaf streak. In: *Diseases of Banana, Abacá and Enset* (Jones, D.R., ed.), p. 37–79. New York: CABI Publishing
- Churchill ACL (2011) *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* 12:307–328.
- Evgen'ev M, Zelentsova H, ShostaK N, Kozitsina M, Barskyi V, Lankenau D-H, Corces VG (1997) Penelope, a new family of transposable elements and its possible role in hybrid dysgenesis in *Drosophila virilis*. *Proc Natl Acad Sci U S A* 94:196–201
- Flavell AJ, Knox MR, Pearce SR, Ellis THN (1998) Retrotransposon-based insertion polymorphisms (RBIP) for high throughput marker analysis. *The Plant Journal* 16:643–650
- Gasparotto L, Pereira JCR, Hanada RE, Montarroyos AVV (2006) *Sigatoka-negra da bananeira*, 1ª ed. Brasília, Embrapa

Gertz J, Varley KE, Davis NS, Baas BJ, Goryshin IY, Vaidyanathan R, Kuersten V, Myers RM (2012) Transposase mediated construction of RNA-seq libraries. *Genome Res* 22:134–141

Gomulski LM, Torti C, Murelli V, Bonizzoni M, Gasperi G, Malacrida AR (2004) Medfly transposable elements: diversity, evolution, genomic impact and possible applications. *Insect Biochem Mol Biol*.34:139–48

Grzebelus D (2006) Transposon insertion polymorphism as a new source of molecular markers. *J Fruit Ornam Plant Res* 14:21–29

Kalendar R, Flavell A, Ellis THN, Sjakste T, Moisy C, Schulman AH (2011) Analysis of plant diversity with retrotransposon-based molecular markers. *Heredity* 106:520–530.

Kalendar R, Grob T, Regina M, Suoniemi A, Schulman A (1999) IRAP and REMAP: two new retrotransposon-based DNA fingerprinting techniques. *Theor Appl Genet* 98:704–711

Kozeretska IA, Demydov SV, Ostapchenko LI (2011) Mobile genetic elements and cancer. From mutations to gene therapy. *Exp Oncol* 33:198–205

Kramerov D, Vassetzky N (2005) Short retroposons in eukaryotic genomes. *Int Rev Cytol* 247:165–221

Kumar A and Bennetzen JL (1999) Plant retrotransposons. *Annu Rev Genet* 33:479–532

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

Marín I (2010) GIN Transposons: Genetic elements linking retrotransposons. *Mol Biol Evol* 27:1903–1911

McDonald BA, Linde C (2002) Pathogen population genetics, evolutionary potential, and durable resistance, *Annu Rev Phytopathol* 40:349–79

- Muszewska A, Hoffman-Sommer M, Grynberg M (2011) LTR retrotransposons in fungi. *PLoS One* 6:1–10
- Nassif N, Penney J, Pal S, Engels W, Gloor G (1994) Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Mol Cell Biol* 14:1613–1625
- Neumann P, Pozarkova D, Macas J (2003) Highly abundant pea LTR retrotransposon Ogre is constitutively transcribed and partially spliced. *Plant Mol Biol* 53:399–410
- Pereira JCR, Gasparotto L, Coelho AFS, Urben AB (1998) Ocorrência da Sigatoka negra no Brasil. *Fitopatol Bras* 23:(abstract) 295
- Pereira JF, Araujo EF, Brommonschenkel SH, Queiroz MV (2006) Elementos transponíveis em fungos fitopatogênicos. *Rev Anu Patol Plant* 14:301–360
- Petrov DA, Hartl D L (1998) High rate of DNA loss in the *Drosophila melanogaster* and *Drosophila virilis* species groups. *Mol Biol Evol* 15: 293–302
- Ministério da Agricultura Pecuária e Abastecimento. Disponível em: <http://www.agricultura.gov.br/>. Acesso em: 24 maio, 2012
- Smýkal P, Bačová-Kerteszová N, Kalendar R, Corander J, Schulman AH, Pavelek M (2011) Genetic diversity of cultivated flax (*Linum usitatissimum* L.) germplasm assessed by retrotransposon-based markers. *Theor Appl Genet* 122:1385–1397
- Tollis M, Boissinot S (2011) The transposable element profile of the anolis genome: How a lizard can provide insights into the evolution of vertebrate genome size and structure. *Mob Genet Elements* 1:107-111
- Waugh R, McLean K, Flavell AJ, Pearce SR, Kumar A, Thomas BB, Powell W (1997) Genetic distribution of Bare-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). *Mol Gen Genet* 253:687-694

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P, Schulman AH (2007) A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* 8:973–82

Vargas VMM (1996) Prevencion y manejo de la sigatoka negra. Caldas, Colômbia: ICA, p 30

ARTIGO

USE OF THE IRAP MARKER TO STUDY GENETIC VARIABILITY IN *Pseudocercospora fijiensis* POPULATIONS

Queiroz, C.B., Santana, M.F., Silva, G.F, Mizubuti, E.S.G., Araújo, E.F., Queiroz, M.V.
Use of the IRAP marker to study genetic variability in *Pseudocercospora fijiensis*
populations. Current Microbiology, *under review*.

Use of the IRAP marker to study genetic variability in *Pseudocercospora fijiensis* populations

Casley Borges de Queiroz¹, Mateus Ferreira Santana¹, Gilvan Ferreira da Silva², Eduardo Seiti Gomide Mizubuti³, Elza Fernandes de Araújo¹ e Marisa Vieira de Queiroz¹

¹Departamento de Microbiologia - Universidade Federal de Viçosa, Minas Gerais, Brasil

² Laboratório de Genética de Micro-organismo - Embrapa Amazônia Ocidental, Amazonas, Brasil

³ Departamento de Fitopatologia - Universidade Federal de Viçosa, Minas Gerais, Brasil

Corresponding Author: Marisa Vieira de Queiroz

E-mail: mvqueiro@hotmail.com

Abstract

Pseudocercospora fijiensis is the etiological agent of black Sigatoka, which is currently considered one of the most destructive banana diseases in all locations where it occurs. It is estimated that a large portion of the *P. fijiensis* genome consists of transposable elements, which allows researchers to use transposon-based molecular markers in the analysis of genetic variability in populations of this pathogen. In this context, the inter-retrotransposon amplified polymorphism (IRAP) was used to study the genetic variability in *P. fijiensis* populations from different hosts and different geographical origins in Brazil. A total of 22 loci were amplified and 77.3% showed a polymorphism. Cluster analysis revealed two major groups in Brazil. The observed genetic diversity (H_E) was 0.22, and through molecular analysis of variance, it was determined that the greatest genetic variability occurs within populations. The Discriminant Analysis of Principal Components (DAPC) revealed no structuring related to the geographical origin of culture of the host. The IRAP-based marker system is a suitable tool for the study of genetic variability in *P. fijiensis*.

Keywords: black Sigatoka, inter-retrotransposon amplified polymorphism (IRAP), population genetics, structure

Introduction

Bananas are produced in more than 120 countries located in tropical and subtropical regions, especially in the rainy regions of Africa, Asia and Latin America [10]. In these regions, most of the production is used for subsistence. Brazil produces approximately 7 million tons of bananas and is the fourth largest banana producer in the world [11]. However, since the first report of the occurrence of black Sigatoka in the country in 1998 [34], banana production has decreased considerably [6]. Black Sigatoka is one of the most destructive fungal diseases on banana plantations, resulting in the loss of fruit quality due to early and non-uniform maturation and yielding fruit with no commercial value [4].

Black Sigatoka is caused by the ascomycete *Mycosphaerella fijiensis* M. Morelet, with *Pseudocercospora fijiensis* (M. Morelet) Deighton representing its anamorphic phase; the pathogen populations show high genetic variability [2, 37]. Several molecular markers have been used to assess the quantity and distribution of genetic variation in *P. fijiensis* populations and include random amplified polymorphic DNA (RAPD) [18], the restriction fragment length polymorphism (RFLP) [2, 17], single-nucleotide polymorphisms (SNPs) [48] and microsatellites [16, 37]. These studies have revealed that *P. fijiensis* populations display high genetic variability. Furthermore, it is known that the pathogen has a short life cycle, a mixed reproduction system and, apparently, a relatively high mutation rate [5]. These attributes allow the pathogens to be classified as having high evolutionary potential, as described by McDonald and Linde [30]. Transposable elements may play a crucial role in generating genetic variability in many species [38]. The presence of transposable elements in the genome can allow for rearrangements through recombination, and the mutational activity of these elements, excluding deleterious insertions, can promote beneficial genetic variations for the host [38, 24]. Due to their ubiquity, abundance, and genomic dispersion and the presence of conserved regions that allow for the design of specific primers, transposable elements have they been used as molecular markers [41, 22].

In many species, transposable element-based markers can be used successfully for the study of genetic diversity and variability, such as the sequence-specific amplified

polymorphism (SSAP) that has been used in barley cultivars [42] and in the fungus *Fusarium oxysporum* [33], the retrotransposon-based insertion polymorphism (RBIP) that has been used in rice [45], the retrotransposon-microsatellite amplified polymorphism (REMAP) that has been used in species of mushroom [26] and in the fungus *Magnaporthe grisea* [3] and the inter-retrotransposon amplified polymorphism (IRAP) that has been used in *Tricholoma matsutake* [31] and *Moniliophthora pernicious* [40]. Among the transposable element-based markers, IRAP stands out as a simple and efficient system, requiring only a simple PCR, followed by electrophoresis to resolve the products. The IRAP marker uses conserved retrotransposon sequences, termed long terminal repeats (LTRs), for detection of polymorphisms [22]. The IRAP method is based on the amplification of regions between two neighboring retrotransposons. The polymorphisms are detected by the presence or absence of loci. The absence of a band indicates the absence of the retrotransposon at a particular locus. The presence or absence of the transposable element at a particular locus can thus be used as a marker for fingerprinting, diversity studies and linkage maps [15].

Recently, the *P. fijiensis* genome has been sequenced by the Joint Genome Institute (JGI). By using "in silico" analysis, Clutterbuck [5] estimated that approximately 50% of the *P. fijiensis* genome is represented by repetitive DNA, making the use of transposon-based molecular markers a great potential tool for use in evolutionary genetic studies of the *P. fijiensis* population. Therefore, due to the estimated abundance of repetitive material in the genome of *P. fijiensis*, together with the lack of characterization studies of the genetic variability in *P. fijiensis* populations recently introduced in Brazil, two aims were established for the present study. The first was to evaluate the suitability of the IRAP marker for the study of genetic variation among individuals, and the second was to determine the genetic structure of the *P. fijiensis* population in Brazil based on the fingerprint generated by IRAP.

Materials and Methods

Acquisition of Pseudocercospora fijiensis isolates

P. fijiensis isolates were obtained from banana leaves showing symptoms of the disease. Samples were collected at 14 locations in Brazil in 2008 and 2009 (Table 1). A total of 69 monosporic isolates were obtained from conidia. Leaves with lesions were viewed under a stereomicroscope, and each conidium was transferred directly to a culture medium plate containing potato dextrose agar (PDA). The plates were maintained at 27°C in the dark.

DNA extraction

After obtaining pure monosporic cultures, mycelial fragments were removed from the colonies, crushed and transferred to 50 ml of liquid enriched PD medium (extract from 250 g of potatoes boiled in 500 ml of water, 10 g of dextrose, 2 g of peptone, 1.5 g of hydrolyzed casein and 2 g of yeast extract per liter of water). The isolates were cultured at 27°C for 7 days with continuous stirring at 120 rpm. The mycelia were recovered, washed, filtered and stored at -80°C until DNA extraction. DNA was obtained by grinding the mycelia using the cetyltrimethylammonium bromide (CTAB) method described by Doyle and Doyle [8]. DNA quality and quantity were determined with 0.8% agarose gels and a NanoDrop 2000 spectrophotometer, respectively.

"In silico" analysis of Class I transposons present in the genome

The genomic sequences of *P. fijiensis* Class I transposable elements were obtained by searching the fungus genome database (<http://genome.jgi-psf.org/Mycfi2/Mycfi2.home.html>) using the keyword search engines (transposon and reverse transcriptase) available on the previously mentioned website. Subsequently, the remaining copies of each element were obtained by BLAST (Basic Local Alignment

Search Tool) from each element previously identified against the *P. fijiensis* genome version 2.0. The long terminal repeats (LTRs) were identified with the aid of the *LTR_finder* [46] and by sequence alignment of the ends of the transposon using the MEGA program version 4.0 [43]. The found LTRs sequences were aligned in Clustal W [44]. The identification of domains related to transposition proteins was performed using tools available at NCBI (<http://www.ncbi.nlm.nih.gov>).

IRAP markers

The primers were designed based on conserved LTR regions of two different retrotransposons. A combination of the LTRMfF and LTRMfR primers was used to amplify the RetroMf1 element (scaffold 1, starting at 3,201,493 bp and ending at 3,208,625 bp), and a combination of the LTRG3F and LTRG3R primers was used to amplify the RetroMf2 element (scaffold 7, starting at 3,581,652 bp and ending at 3,588,100 bp). In our study, the combinations of the two primers amplify in the direction away from the element. The primers sequences are listed in Table 2.

The IRAP marker was amplified from the DNA samples based on the protocol by Kalendar and Schulman [23] with modifications; each 25 μ l reaction contained 0.5 μ M primers, 2.0 mM MgCl₂, 0.6 mM dNTP (equimolar mixture of dATP, dGTP, dCTP, and dTTP), 1x buffer and one unit of *goTaq* DNA polymerase (Promega). PCR was performed in a C1000 Bio-Rad thermocycler programmed to perform an initial denaturation at 95°C for 3 minutes, followed by 32 cycles at 95°C for 15 seconds, 60°C for 1 minute and 68°C for 2 minutes, and a final extension at 68°C for 5 minutes. The amplification products were separated by electrophoresis in a 1.5% agarose gel stained with 0.3 μ g/ml ethidium bromide and 1x TBE buffer (2 mM EDTA, 0.1 M Tris- HCl, and 0.1 M boric acid [pH 8.0]). The 1Kb Plus DNA Ladder (Invitrogen) was used to estimate the size of the amplicons.

Data analysis

Differences in the electrophoresis patterns among the isolates were visually analyzed in a 1.5% agarose gel. The bands for each primer combination used in the amplification were classified with the number one (presence) or zero (absence) among different isolates. The reproducibility of the band profiles was tested by repetition of the PCR with all the samples and selected primers. Only reproducible bands were considered for analysis. Bands common to all isolates were included in the analysis.

The number of amplified loci, polymorphism rate, and haplotype and singleton numbers were calculated using all the isolates. Subsequently, the isolates were grouped into 14 and 5 subpopulations based on the collection site and the host of origin, respectively. The genetic diversity (H_E) [32] and genotypic diversity (D_G) (*Shannon-Wiener* index) were calculated using the POPGENE software version 1.32 [47]. Arlequin software version 3.5 [9] was used to calculate the analysis of molecular variance (AMOVA). A dendrogram (bootstrap with 1,000 replicates) for *P. fijiensis* isolates grouped by geographical origin was constructed by the unweighted pair group method with arithmetic mean (UPGMA) using the R package pvclust software version 2.14 [35]. The Discriminant Analysis of Principal Component (DAPC) [20] was performed in the R package adegenet [19] to detect significant structuring within the dataset. DAPC uses *sequential K-means* and Bayesian Information Criterion (BIC) to determine the optimal number of clusters. To describe the groups identified, the DAPC was based on data transformation using the principal component analysis (PCA) to maximize separation among the groups.

Results

After searching for retrotransposons in the genome of *P. fijiensis*, we selected four elements containing conserved LTRs that were widely distributed in the genome. These four elements were used to develop four primer sets; however, only two sets were able to generate polymorphic products for the analysis of variability in *P. fijiensis*. Of

the two selected elements, the first was named *RetroMf1*, had a total size of 7,133 bp and contained 197 bp in each of its LTRs. The LTR analysis of this element via BLAST against the fungus genome resulted in 51 hits. This element is located in scaffold 1. The second element was named *RetroMf2* and had a total size of 6,448 bp. Of the analyzed elements, it was the only one that contained the complete ORFs of the *pol* region; however, its LTRs are broken at their ends (Figure 1). The LTR analysis of this element via BLAST against the *P. fijiensis* genome version 2.0 resulted in 41 hits. This element is located in scaffold 7.

A total of 22 loci were amplified. The combination of the LTRMfF and LTRMfR primers amplified 12 loci, whereas the combination of the LTRG3F and LTRG3R primers amplified 10 loci. The two primers sets demonstrated high reproducibility and were able to generate amplicons in all isolates used (Figure 2). Based on the IRAP electrophoretic patterns, it was determined that 77.3% of the amplified loci were polymorphic. The allelic frequency in the population ranged from 0.01 to 1.0. A total of 61 haplotypes were found, among which 56 were singletons. The combination of the LTRG3F and LTRG3R primers was able to detect a high rate of polymorphism, with 100% of bands amplified by this primer combination being polymorphic.

When the populations were subdivided by collection site and host of origin, the genetic diversity (H_E) and the genotypic diversity (D_G) were similar (Table 3). Most of the genetic variation was found within the subdivided populations (98.8% and 99.1%) according to collection site and host of origin, respectively (Table 4).

Cluster analysis based on the collection site revealed two major groups, group A and group B, with bootstrap support of 78 and 82, respectively. The first group includes populations from Eldorado-SP, Itacoatira-SP, Rio Branco-AC, Rio Preto da Eva and Caroebe-RR, while the second group includes populations from Iranduba-AM, Manaus-AM, Miracatu-SP, Careiro Castanho-AM, Cáceres-MT, Pariquera-Açu-SP and Manacapuru-AM. The isolates from Presidente Figueiredo-AM have a large genetic distance compared to other isolates from the Amazonas. The population from Porto Velho (Rondônia) has a significant genetic distance from the two main groups (Figure 3). DAPC was performed on all 69 subjects. The K value was 6, which represents a

good data summary. DAPC revealed the presence of six groups; however, this structuring is not correlated with geographical origin and origin of host (Figure 4).

Discussion

Transposable elements have been found in virtually all genomes, representing, for example, approximately 15% of the sequenced *Drosophila melanogaster* genome [1], 45% of the sequenced human genome [25], and more than 50% of the sequenced maize genome [39]. In filamentous fungi, transposable elements were found to represent 14% of the sequenced *Ascobolus* sp. genome [14], 10% of the sequenced *Neurospora crassa* genome [12], 9.7% of the sequenced *Magnaporthe grisea* genome [7], 21% of the sequenced *Laccaria bicolor* genome [28] and 58% of the sequenced *Tuber melanosporum* Vittad genome [29]. It has been estimated that approximately 50% of the *P. fijiensis* genome is composed of repetitive DNA; however, most of the elements are degraded [5].

Genetic variability studies in *P. fijiensis* have been conducted with RAPD [18], RFLP [2, 17], SNPs [48] and microsatellites [37, 36]. However, the use of transposable element-based markers has advantages over other molecular markers. IRAP markers are versatile, because can be used multiple primers that anneal to conserved regions of retrotransposons. Additionally, these specific primers yield highly reproducible results, unlike RAPD. It is a simple technique that uses inexpensive reagents when compared to AFLP. Furthermore, retrotransposon-based markers, such as IRAP, are advantageous in capturing large changes in the genome, unlike RFLP, SNPs, AFLP and microsatellites, which mainly detect single nucleotide changes with high frequencies of reversion. For example, microsatellite-based molecular markers usually detect a gain or loss of 20 nucleotides. Microsatellite alleles differ in their number of simple sequence repeats (SSRs) and, similar to single nucleotide changes, suffer from homoplasmy, as the number of SSRs can reversibly increase or decrease, making it impossible to distinguish the ancestral states of the subjects [23].

In the present study, the analysis of *P. fijiensis* isolates revealed a genetic diversity among populations in Brazil similar to that found by Robert et al [37] in Africa

($H_E = 0.22$). However, in that same study, the authors observed an H_E value of up to 0.65 in Southeast Asia, which is considered to be the center of origin of this species. The low H_E value found in the present study may be attributable to the recent introduction of *P. fijiensis* in Brazil. In fact, it has been shown that black Sigatoka emerged in Southeast Asia and later migrated to Oceania and Africa, and, simultaneously, from Asia and Oceania to America [37]. Since that migration, *P. fijiensis* has been found in the main banana-producing regions around the world. In Brazil, black Sigatoka was first observed in 1998 in the cities of Tabatinga and Benjamin Constant in the state of Amazonas [34] on the borders with Colombia and Peru. Therefore, our results are consistent with results from Robert et al [37], which demonstrated a trend of decreasing genetic diversity with increasing geographical distance from the center of origin.

The fact that low genetic variability was found among populations subdivided by geographical origin and origin of host suggests that many alleles are shared among different populations. AMOVA values for populations subdivided by geographic origin show that the variability of isolates was greater within the collection site than among them, with an estimated 98.8% of total variability found to occur within each collection site. Likewise, AMOVA values for populations subdivided by host of origin show that the variability of isolates was greater within the host of origin than among them, with an estimated 99.1% of total variability found to occur within each host of origin. This result is mainly due to the existence of a large number of heterogeneous haplotypes within each population.

The dendrogram obtained through genetic distance analysis showed no correlation between geographical distance and genetic distance, as it was observed that samples from geographically distant regions have high genetic similarity. For example, Itacoatiara-AM and Eldorado-SP have a straight-line geographical distance of more than 2,500 kilometers but were shown to be genetically close in the generated dendrogram. In contrast, Presidente Figueiredo-AM and Rio Preto da Eva-AM are geographically close regions separated by a straight-line geographical distance of 82 kilometers but these populations were allocated to the different cluster in the dendrogram. Likewise, although the DAPC revealed structuring in the *P. fijiensis* population, the genetic diversity of the isolates is independently related as to the geographical origin and the host cultivar from

which they were collected. This result is possibly due to the type of spore dispersal and the anthropogenic activity. Wind dispersal of ascospores and conidia is considered the most important source of inoculum for long distance dissemination of the fungus [4, 21]. Moreover, anthropogenic activity can also accelerate dissemination of the disease through the transport of infected material, such as banana leaves [37]. Ascospore dispersal can lead to the exchange of genetic material among fungi populations [4], potentially increasing the genetic variability.

The present study is the first report of the genetic variability of *P. fijiensis* populations in Brazil and provides a better understanding of the distribution of this pathogen in the world. Furthermore, we have demonstrated that the retrotransposon-based marker system IRAP is a useful tool for the analysis of genetic variability and characterization of *P. fijiensis* population structure. We also believe our results are important in the efforts to control the disease and in the successful development of genetic improvement programs to produce banana varieties resistant to black Sigatoka.

Acknowledgments

This work was financially supported by the Brazilian Agency CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico).

References

1. Biémont C, Cizeron G (1999) Distribution of transposable elements in *Drosophila* Species. *Genetica* 105:43-62
2. Carlier J, Lebrun MH, Zapater MF, Dubois C, Mourichon X (1996) Genetic structure of the global population of banana black leaf streak fungus, *Mycosphaerella fijiensis*. *Mol Ecol* 5:499-510
3. Chadha S, Gopalakrishna T (2005) Retrotransposon-microsatellite amplified polymorphism (REMAP) markers for genetic diversity assessment of the rice blast pathogen (*Magnaporthe grisea*). *Genome* 48: 943-945
4. Churchill ACL (2011) *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* 12:307–328
5. Clutterbuck AJ (2011) Genomic evidence of repeat-induced point mutation (RIP) in filamentous ascomycetes. *Fungal Genet Biol* 48:306–326
6. Cordeiro ZJM, Matos AP (2002) Impact of *Mycosphaerella* spp in Brazil. In: *Mycosphaerella* leaf spot diseases of bananas: present status and outlook. Proceedings of the 2nd International Workshop on *Mycosphaerella* leaf spot diseases held in San Jose. INIBAP Costa Rica, pp 91-97
7. Dean RA, Talbot NJ, Ebbole DJ et al (2005) The genome sequence of the rice blast fungus *Magnaporthe grisea*. *Nature* 434:980-986
8. Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 19:11-15
9. Excoffier L, Laval G, Schneider S (2005) ARLEQUIN ver. 3.1: an integrated software package for population genetics data analysis. *Evol Bioinform Online* 1:47-50
10. FAO (2009). Major diseases of banana and plantain: an update of their spread, impact and response strategies. Joint Meeting of the 4th Session of the Sub-Group on Bananas and the 5th Session of the Sub-Group on Tropical Fruits, Committee on Commodity Problems. Rome: Food and Agriculture Organization of the United Nations.
11. FAO (2010). Food and Agricultural Organization. FAOSTAT 2010. Available at <http://faostat.fao.org/site/567/default.aspx>. Accessed 3 July 2012

12. Galagan JE, Calvo SE, Borkovich KA et al (2003) The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* 422:859-868
13. Gilbert C, Schaack S, Pace JKII, Brindley PJ, Feschotte C (2010) A role for host-parasite interactions in the horizontal transfer of transposons across phyla. *Nature* 464:1347–1350
14. Goyon C, Rossignol JL, Faugeron G (1996) Native DNA repeats and methylation in *Ascobolus*. *Nucleic Acids Res* 24:3348-3356
15. Grzebelus D (2006) Transposon insertion polymorphism as a new source of molecular markers. *J Fruit and Ornam Plant Res* 14:21-29
16. Halkett F, Coste D, Rivas Platero GG, Zapater MF, Abadie C, Carlier J (2010) Genetic discontinuities and disequilibria in recently established populations of the plant pathogenic fungus *Mycosphaerella fijiensis*. *Mol Ecol* 19:3909-3923
17. Hayden HL, Carlier J, Aitken EAB (2003) Genetic structure of *Mycosphaerella fijiensis* populations from Australia, Papua New Guinea and the Pacific Islands. *Plant Pathol* 52:703-712
18. Johanson A, Crowhurst RN, Rikkerink EHA, Fullerton RA, Templeton MD (1994) The use of species-specific DNA probes for the identification of *Mycosphaerella fijiensis* and *M. musicola*, the causal agents of Sigatoka disease of banana. *Plant Pathol* 43:701-707
19. Jombart T (2008) Adegnet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403-1405
20. Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* doi:10.1186/1471-2156-11-94
21. Jones DR (2000) Introduction to banana, abacá and enset. In: Jones DR (ed) *Diseases of Banana, Abacá and Enset*. CABI Publishing, Wallingford, pp 37-79
22. Kalendar R, Grob T, Regina M, Suoniemi A, Schulman A (1999) IRAP and REMAP: two new retrotransposon-based DNA fingerprinting techniques. *Thor Appl Genet* 98:704-711
23. Kalendar R, Schulman AH (2006) IRAP and REMAP for retrotransposon-based genotyping and fingerprinting. *Nat Protoc* 1:2478-2484

24. Kraaijeveld K, Zwanenburg B, Hubert B, Vieira C, De Pater S, Van Alphen JJM, Den Dunnen JT, De Knijff P (2012), Transposon proliferation in an asexual parasitoid. *Mol Ecol*. doi: 10.1111/j.1365-294X.2012.5582
25. Lander ES, Linton LM, Birren B et al (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860-921
26. Le QV, Won H-K, Lee T-S, Lee C-Y, Lee H-S, Ro H-S (2008) Retrotransposon microsatellite amplified polymorphism strain fingerprinting markers applicable to various mushroom species. *Mycobiology* 36:161-166
27. Lewontin RC (1972) The apportionment of human diversity. *Evol Biol* 6: 381-398
28. Martin F, Aerts A, Ahrén D et al (2008) The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature* 452:88-92
29. Martin F, Kohler A, Murat C et al. (2010) Périgord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. *Nature* 464:1033-1038
30. McDonald BA, Linde C (2002) Pathogen population genetics, evolutionary potential, and durable resistance. *Annu. Rev. Phytopathol* 40:349-379
31. Murata H, Babasaki K, Saegusa T, Takemoto K, Yamada A, Ohta A (2008) Traceability of Asian Matsutake, specialty mushrooms produced by the ectomycorrhizal basidiomycetes *Tricholoma matsutake*, on the basis of retroelement-based DNA markers. *Appl Environ Microb* 74:2023-2031
32. Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci USA* 70:3321-3323
33. Pasquali M, Saravanakumar D, Gullino ML, Garibaldi A (2008) Sequence-specific amplified polymorphism (SSAP) technique to analyse *Fusarium oxysporum* f. sp. *lactucae* VCG 0300 isolate from lettuce. *J Plant Pathol* 90:527-535
34. Pereira JCR, Gasparotto L, Coelho AFS, Urban AB (1998) Ocorrência da Sigatoka negra no Brasil. *Fitopatol Bras* 23:(abstract) 295
35. R development Core Team (2007) R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing
36. Rieux A, De Lapeyre De Bellaire L, Zapater MF, Ravigne V, Carlier J (2012) Recent range expansion and agricultural landscape heterogeneity have only minimal effect on

- the spatial genetic structure of the plant pathogenic fungus *Mycosphaerella fijiensis*. *Heredity*. doi: 10.1038/hdy.2012.55
37. Robert S, Ravigne V, Zapater MF, Abadie C, Carlier J (2012) Contrasting introduction scenarios among continents in the worldwide invasion of the banana fungal pathogen *Mycosphaerella fijiensis*. *Mol Ecol* 21:1098-1114
 38. Rouzic AL, Boutin TS, Capy P (2007) Long-term evolution of transposable elements. *Evolution* 104:9375-19380
 39. SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL (1998) The paleontology of intergene retrotransposons of maize. *Nat Genet* 20:43-45
 40. Santana MF, Araújo EF, Souza JT, Mizubuti ESG, Queiroz MV (2012) Development of molecular markers based on retrotransposons for the analysis of genetic variability in *Moniliophthora perniciosa*. *Eur J Plant Pathol*. doi: 10.1007/s10658-012-0031-4
 41. Schulman AH, Flavell AJ, Ellis TH (2004) The application of LTR retrotransposons as molecular markers in plants. *Methods Mol Biol* 260:145–173
 42. Soleimani VD, Baum BR, Johnson DA (2005) Genetic diversity among barley cultivars assessed by sequence-specific amplification polymorphism. *Theor Appl Genet* 110:1290-300
 43. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596-9
 44. Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W: improving the sensibility of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673-4680
 45. Vitte C, Ishii T, Lamy F, Brar D, Panaud O (2004) Genomic paleontology provides evidence for two distinct origins of Asian rice (*Oryza sativa* L). *Mol Genet Genomics* 272:504-511
 46. Xu Z, Wang H (2007) LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res* 35:265-268
 47. Yeh FC, Yang R, Boyle TJ et al (1999) PopGene32, Microsoft Windows-Based Freeware for Population Genetic Analysis, Version 1.32. Molecular Biology and Biotechnology Centre, University of Alberta, Edmonton

48. Zandjanakou-Tachin M, Vroh-Bi I, Ojiambo PS, Tenkouano, A, Gumedzoe YM, Bandyopadhyay R (2009) Identification and genetic diversity of *Mycosphaerella* species on banana and plantain in Nigeria. *Plant Pathol* 58:536-546

Table 1. *Pseudocercospora fijiensis* isolates used in this study.

Isolates	Collection site	Variety
2Mf	Presidente Figueiredo – AM	THAP MAEO
3Mf	Presidente Figueiredo – AM	FHIA 18
5Mf	Presidente Figueiredo – AM	Pacovan
105Mf	Presidente Figueiredo – AM	Costela de vaca
113Mf	Presidente Figueiredo – AM	Prata
123Mf	Presidente Figueiredo – AM	Pacovan
127Mf	Presidente Figueiredo – AM	Caru roxa
7Mf	Manacapuru – AM	Pacovan
9Mf	Manacapuru – AM	Pacovan
10Mf	Manacapuru – AM	Maçã
23Mf	Manacapuru – AM	Maçã
62Mf	Manacapuru – AM	Maçã
63Mf	Manacapuru – AM	Prata
68Mf	Manacapuru – AM	Pacovan
20Mf	Rio Preto da Eva – AM	Maçã
24Mf	Rio Preto da Eva – AM	Prata
32Mf	Rio Preto da Eva – AM	Prata
37Mf	Rio Preto da Eva – AM	Maçã
40Mf	Rio Preto da Eva – AM	Pacovan
41Mf	Rio Preto da Eva – AM	Maçã
44Mf	Manaus – AM	Prata
46Mf	Manaus – AM	Maçã
47Mf	Manaus – AM	Maçã
49Mf	Manaus – AM	Pacovan
52Mf	Manaus – AM	Maçã
54Mf	Manaus – AM	prata
97Mf	Irاندuba – AM	Maçã
99Mf	Irاندuba – AM	Pacovan
102Mf	Irاندuba – AM	Maçã

103Mf	Irاندوبا – AM	Pacovan
223Mf	Irاندوبا – AM	Urucuri
224Mf	Irاندوبا – AM	Nanica
225Mf	Irاندوبا – AM	Nanica
226Mf	Irاندوبا – AM	Prata
160Mf	Careiro Castanho – AM	Nanica
167Mf	Careiro Castanho – AM	Prata
173Mf	Careiro Castanho – AM	Pacovan
174Mf	Careiro Castanho – AM	Maçã
182Mf	Itacoatiara – AM	FHIA 18
188Mf	Itacoatiara – AM	Maçã
192Mf	Itacoatiara – AM	Nanica
195Mf	Itacoatiara – AM	Pacovan
82Mf	Cáceres – MT	IAC 2001
83Mf	Cáceres – MT	D'angola
87Mf	Cáceres – MT	Grand naine
171Mf	Rio Branco – AC	D'angola
177Mf	Rio Branco – AC	Carrú cinza
185Mf	Rio Branco – AC	D'angola
196Mf	Rio Branco – AC	ST 1231
106Mf	Caroebe – RR	Maçã
118Mf	Caroebe – RR	Maçã
119Mf	Caroebe-RR	Prata
120Mf	Caroebe – RR	Maçã
125Mf	Caroebe – RR	Prata
130Mf	Caroebe – RR	Pacovan
131Mf	Caroebe-RR	Pacovan
169Mf	Porto Velho – RO	Prata
170Mf	Porto Velho – RO	Caru roxa
175Mf	Porto Velho – RO	Caru roxa
198Mf	Miracatu – SP	Prata

199Mf	Miracatu – SP	Prata
204Mf	Miracatu – SP	Prata
205Mf	Miracatu – SP	Prata
208Mf	Eldorado – SP	Nanicão
210Mf	Eldorado – SP	Prata
212Mf	Eldorado – SP	Nanica
219Mf	Pariquera-açu – SP	Maçã
220Mf	Pariquera-açu – SP	Naniquinha
222Mf	Pariquera-açu – SP	Figo

Brazilian states: AM (Amazonas), AC (Acre), RO (Rondônia), RR (Roraima), MT (Mato Grosso), SP (São Paulo).

Table 2. Primers used in IRAP analysis of *Pseudocercospora fijiensis*.

Identification	Sequence 5' – 3'
LTRMf F	GCGCTTAGCGTTAGGCTAACT
LTRMfR	CGTG TAGCCTCTTTGGCCCTA
LTRG3F	CGAGTAGTAGGAAGGAACCGG
LTRG3R	GGCGGCTAGCTTATAGGACTT

Table 3. Estimates of the genetic diversity of *Pseudocercospora fijiensis* populations

Hierarchical level	N	H_E (SE)	D_G (SD)
Collection site	69	0.22(0.17)	0.34(0.25)
Host	58	0.21(0.17)	0.33(0.25)

N , N , individuals number; H_E , gene diversity (Nei, 1978);

D_G , genotypic diversity; SD, standard deviation.

Table 4. Analysis of Molecular Variance (AMOVA): variation attributed to differences among and within populations subdivided by different geographical origins (collection sites) and origin host.

Variation sources	Degrees of freedom	Sum of squares	Variance components	Variance (%)
Among populations of cultivars	4	10.790	0.02	0.84
Within populations of cultivars	53	130.969	2.47	99.16
Among populations of collecting sites	13	33.074	0.03	1.16
Within populations of collecting sites	55	132.375	2.40	98.84

Figure legends

Figure 1. Basic structure of the *RetroMf2* retrotransposon. The ORF of the gag region overlaps with the first ORF of the *pol* region, the protease (PR). Other protein-encoding regions found in the *pol* region are the reverse transcriptase (RT), RNase H (RH), and integrase (IN). Arrows indicate the regions where the primers anneal and the direction of amplification.

Figure 2. Amplification profile of IRAP. A, IRAP banding pattern generated by the combination of the LTRMfF and LTRMfR primers. B, IRAP banding pattern generated by the combination of the LTRG3F and LTRG3R primers.

Figure 3. Dendrogram of genetic distances of 14 *Pseudocercospora fijiensis* populations in Brazil. The genetic distance was based on loci obtained by IRAP. Bootstrap values were obtained from 1,000 replicates; only values equal to or greater than 70 were shown. The dendrogram was constructed by the UPGMA method.

Figure 4. Scatter plot of the Discriminant Analysis of Principal Component (DAPC) performed on the data from the 69 subjects ($K = 6$). The individuals are represented by dots, and the groups are represented by ovals and colors. The representative percentage of individuals in the groups related to the geographical origin is shown on the map.

Figure 1

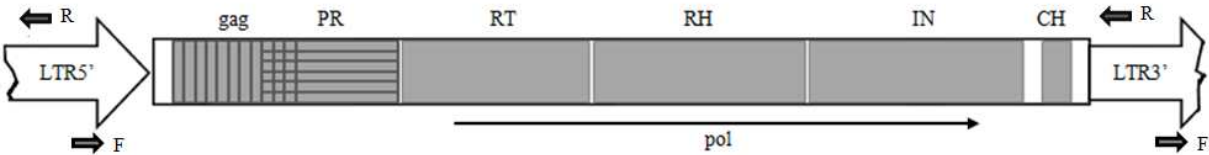


Figure 2

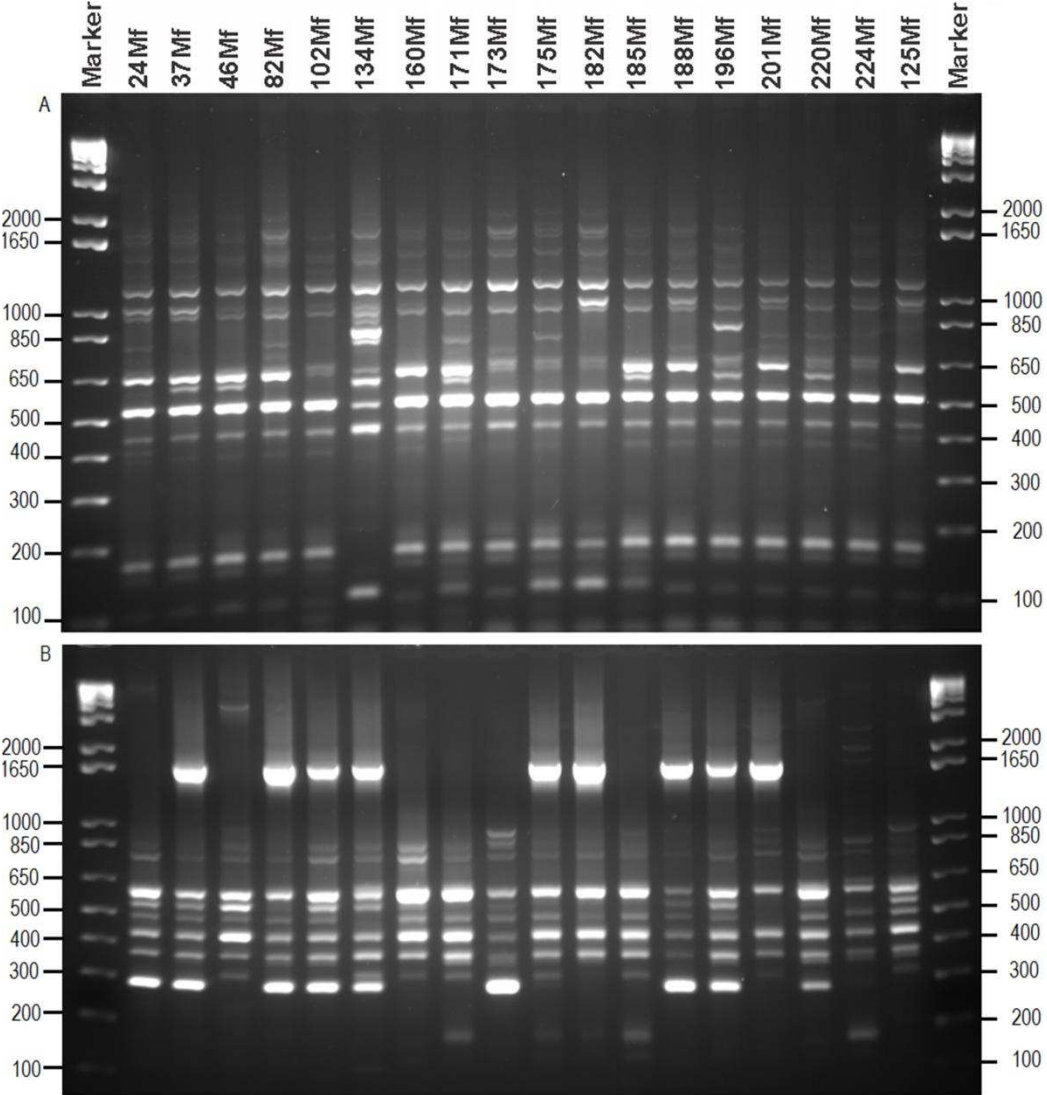


Figure 3

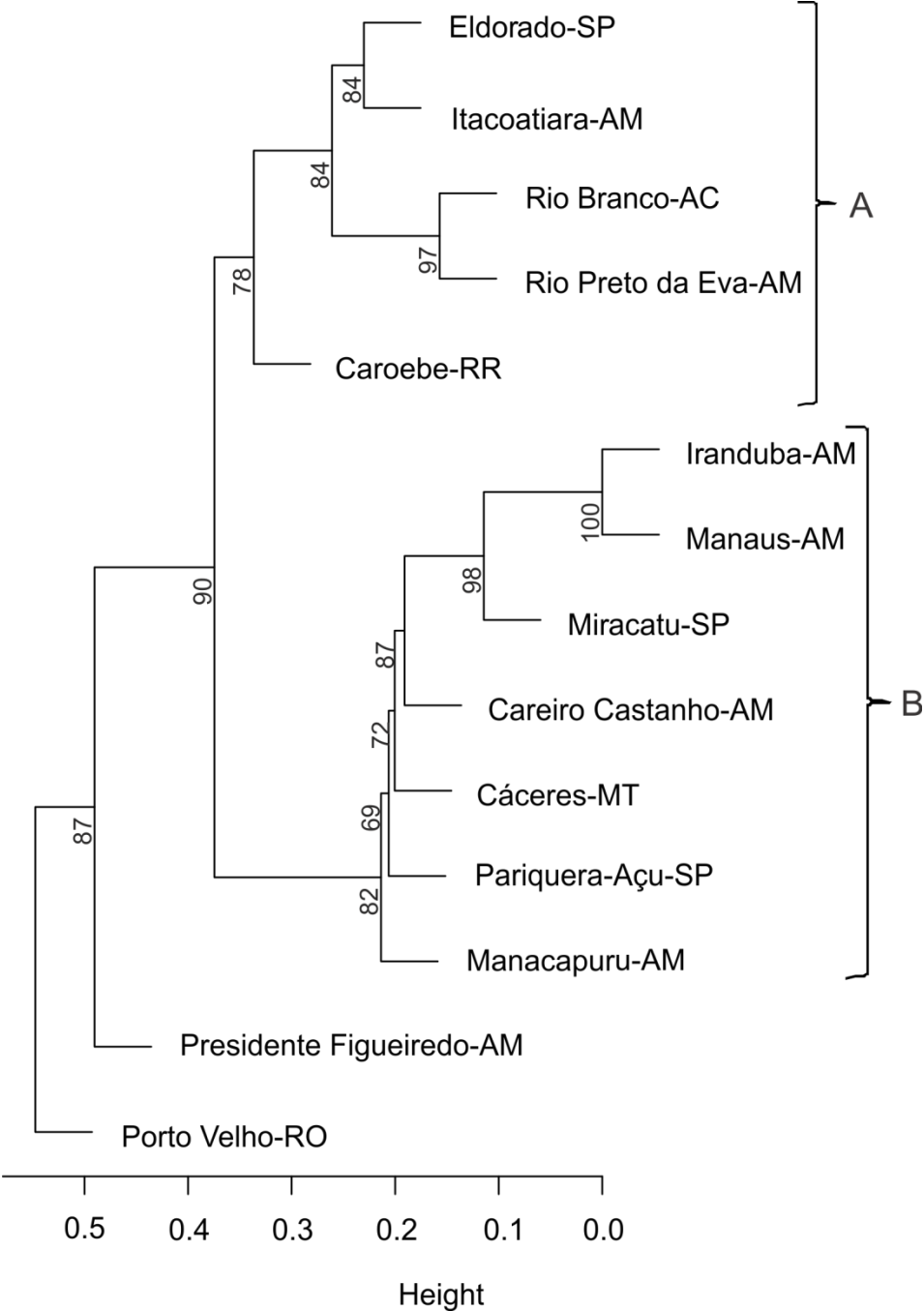


Figure 4

