

MAÍRA CRISTINA MENEZES FREIRE

FILOGEOGRAFIA MOLECULAR DE *Phakopsora pachyrhizi* COM BASE EM
DADOS DE SEQUENCIAMENTO DE REGIÕES GÊNICAS DO DNA
NUCLEAR

Tese apresentada à
Universidade Federal de Viçosa,
como parte das exigências do
Programa de Pós-Graduação em
Genética e Melhoramento, para
obtenção de título de *Doctor
Scientiae*.

VIÇOSA
MINAS GERAIS - BRASIL
2011

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

T

F866f
2011
Freire, Máira Cristina Menezes, 1983-
Filogeografia molecular de *Phakopsora pachyrhizi* com
base em dados de sequenciamento de regiões gênicas do
DNA nuclear / Máira Cristina Menezes Freire. – Viçosa, MG,
2011.

xii, 84f. : il. (algumas col.) ; 29cm.

Orientador: Luiz Orlando de Oliveira.

Tese (doutorado) - Universidade Federal de Viçosa.

Inclui bibliografia.

1. Ferrugem-da-soja. 2. Sequenciamento de nucleotídeo.
3. Soja - Doenças e pragas. 4. Variação (Genética).
5. Evolução (Biologia). I. Universidade Federal de Viçosa.
II. Título.

CDD 22. ed. 632.492

MAÍRA CRISTINA MENEZES FREIRE

FILOGEOGRAFIA MOLECULAR DE *Phakopsora pachyrhizi* COM BASE EM
DADOS DE SEQUENCIAMENTO DE REGIÕES GÊNICAS DO DNA
NUCLEAR

Tese apresentada à
Universidade Federal de Viçosa,
como parte das exigências do
Programa de Pós-Graduação em
Genética e Melhoramento, para
obtenção de título de *Doctor
Scientiae*.

APROVADA: 27 de julho de 2011.

Prof. Jorge Abdala Dergam dos
Santos

Prof^a. Valéria Monteze Guimarães

Pesq. Eunize Maciel Zambolim

Pesq. Trazilbo José de Paula Júnior

Prof. Luiz Orlando de Oliveira
(Orientador)

Dedico esse trabalho aos meus pais Ernani Campos Freire e Célia Lopes de Menezes Freire, verdadeiros responsáveis por cada vitória conquistada.

AGRADECIMENTOS

À Deus, meu eterno companheiro e guardião, que me concedeu muitas oportunidades e me apoiou em todos os momentos.

Aos meus pais, pela incessante torcida, companheirismo e amor, e por terem sempre depositado confiança em cada passo novo que eu precisei dar.

Aos meus irmãos Maria das Graças, Izabela Maria e Ernani Júnior. Meus amigos e companheiros para toda a vida. Obrigada por cada palavra, abraços, sorrisos - cada momento.

Às minhas avós que sempre estarão olhando por mim.

À Universidade Federal de Viçosa, em especial ao Programa de pós-graduação em Genética e Melhoramento, pela oportunidade de realização desse curso.

Ao Conselho Nacional de Pesquisa Científica e Tecnológica (CNPq), pelo apoio financeiro para a realização deste trabalho.

Ao meu orientador, professor Luiz Orlando de Oliveira, pela paciência, pelos ensinamentos e principalmente pela confiança.

Ao professor Eduardo Mizubuti, pelo incentivo nos momentos em que mais precisei.

Ao professor Gary Stacey, por ter permitido a realização dos meus experimentos em seu laboratório. Por cada um dos membros do laboratório pelo carinho que me receberam, e pela amizade conquistada a cada dia.

Às amigas do laboratório de Biologia Molecular e Filogeografia, por fazer cada dia se tornar especial. Obrigada pelo companheirismo e paciência necessários na convivência diária.

Às amigas que conquistei durante a realização do curso, em especial a Janaína Teixeira, Samuel Mazzinghy, Cassiana Severiano, Maria Andréia

Mendonça, Leonardo Almeida, Magali Garcia, e Ana Paula do Carmo, vocês são irmãos que ganhei nesses anos, mas que vou levar para a vida toda.

Às companheiras do time de handebol, especialmente Zélia, Rêmili, Valquíria e Nathália. Vocês, que se tornaram mais que colegas de time, obrigada por me fazerem estar sempre sorrindo. Até as coisas mais simples se tornam especiais com vocês.

Aos amigos de longa data, Letícia Figueiredo, Thiago Augusto, Natália Reis, Isabella Gomes, Thaísa Marques, Liliane Paulo, Rafaela Duarte, Hélio Guimarães, Rodolfo Ferreira, Salvador Lima, Rosane Nalon, Gláucia Costa, Suelen de Andrade, Vanessa Braga, Regina Antunes, Camila Morador, Jaqueline Ferreira, Luiz Menezes e Rafael Silva. Obrigada por entenderem meus momentos de ausência, mesmo de longe vocês foram essenciais.

Aos novos amigos que ganhei na minha temporada nos Estados Unidos, em especial a Cláudia Silva, Renata Dolazza, Chinonye Echebiri, Shawny Green, Elizabeth Oyelola e Andry. Vocês foram anjos que Deus colocou em minha vida.

Enfim, a todos que de alguma forma me ajudaram nessa etapa tão importante. Muito obrigada.

BIOGRAFIA

MAÍRA CRISTINE MENEZES FREIRE, filha de Ernani Campos Freire e Célia Lopes de Menezes Freire, nasceu em 27 de fevereiro de 1983, em Caratinga, Minas Gerais.

Em 2001, ingressou na Universidade Vale do Rio Doce (UNIVALE), Governador Valadares – MG, no curso de bacharelado em Ciências Biológicas, concluindo-o em dezembro de 2004.

Iniciou o curso de Mestrado em Genética e Melhoramento, na Universidade Federal de Viçosa, em março de 2005. Defendeu a dissertação no dia 02 de março de 2007, obtendo o título de *Magister Scientiae* em Genética e Melhoramento.

Em março de 2007 iniciou o curso de Doutorado em Genética e Melhoramento na UFV. Realizou o exame de qualificação em novembro de 2009. De março de 2010 a março de 2011 realizou o estágio do doutorado sanduíche na University of Missouri, Columbia, MO – EUA. Submeteu-se à defesa de tese em julho de 2011.

ÍNDICE

RESUMO.....	ix
ABSTRACT	xi
INTRODUÇÃO GERAL	1
REVISÃO DE LITERATURA	4
1 - A Cultura da soja	4
2 - <i>Phakospora pachrhizi</i> (ferrugem asiática da soja).....	5
2.1 - Histórico.....	5
2.2 - Ciclo de vida	7
2.3 - Formas de controle.....	7
2.4 - Variabilidade de <i>Phakopsora pachyrhizi</i>	9
3 - Filogeografia Molecular	11
REFERÊNCIA BIBLIOGRÁFICA.....	15
CAPÍTULO 1	21
<u>Nucleotide polymorphism in the 5.8S nrDNA gene and internal transcribed spacers in <i>Phakopsora pachyrhizi</i> viewed from structural models</u>	
Abstract.....	22
1 - Introduction	23
2 - Materials and methods.....	26
2.1 - Sampling strategy, DNA extraction, and PCR amplification	26

2.2 - Cloning, sequencing, and sequence alignment.....	26
2.3 - Structural analyses	27
3 - Results and discussion	28
3.1 - Predicted secondary structure of ITS1	28
3.2 - Predicted secondary structure of the 5.8S gene.....	30
3.3 - Predicted secondary structure of ITS2	31
3.4 - Co-evolution of the two spacers	32
3.5 - Usefulness of structural analyses in ITS.....	32
4 - References	34
CAPÍTULO 2	37
<u>Avaliação da variabilidade e da estrutura genética de populações</u>	
<u><i>Phakopsora pachyrhizi</i> com base em dados de sequências do DNA nuclear</u>	
Resumo	38
1 – Introdução	40
2 – Objetivos	42
3 - Metodologia.....	43
3.1 - Coleta de Amostras de Ferrugem de Soja	43
3.2 - Extração de DNA.....	43
3.3 - Reações de Amplificação via PCR.....	44
3.4 - Clonagem dos Produtos Amplificados.....	45
3.5 - Sequenciamento e Comparação das Sequências.....	47
3.6 - Análises Estatísticas.....	48
4 - Resultados	49
4.1 - Fator de ribosilação do ADP.....	49
4.1.1 - Comparação das Sequências.....	49
4.1.2 - Distribuição das sequências	50
4.1.3 – Relação genealógica entre os haplótipos	52
4.1.4 - Distribuição geográfica do haplótipos.....	53

4.1.5 - Fluxo gênico e diferenciação	56
4.2– Histona H4	57
4.2.1 - Comparação das Sequências.....	57
4.2.2 - Distribuição das sequências	58
4.2.3 - Relação genealógica entre os haplótipos.....	60
4.2.4 - Distribuição geográfica dos haplótipos	61
4.2.5 - Fluxo gênico e diferenciação	63
4.3 – ITS (Espaçador transcrito interno).....	64
4.3.1 - Comparação das sequências	64
4.3.2 – Distribuição das sequências	66
4.3.3 – Relação genealógica entre os ribótipos.....	69
4.3.4 - Distribuição geográfica dos ribótipos.....	70
4.3.5 - Fluxo gênico e diferenciação	72
5 – Discussão	75
5.1 – Variabilidade genética de <i>P. pachyrhizi</i>	75
5.2 - Estrutura genética das populações de <i>P. pachyrhizi</i>	76
5.3 – Efeito do vazio sanitário e relevância dos hospedeiros alternativos nas populações de <i>P. pachyrhizi</i>	77
5.4 – Rotas hipotéticas de dispersão e colonização de <i>P. pachyrhizi</i>	78
6 - Conclusões	81
7 – Referências Bibliográficas	82

RESUMO

FREIRE, Maíra Cristina Menezes, D. Sc., Universidade Federal de Viçosa, Julho de 2011. **Filogeografia molecular de *Phakopsora pachyrhizi* com base em dados de sequenciamento de regiões gênicas do DNA nuclear.** Orientador: Luiz Orlando de Oliveira. Co-orientadores: Maurilio Alves Moreira, Eduardo Seiti Gomide Mizubuti e Álvaro Manuel Rodrigues Almeida.

A ferrugem asiática da soja, causada por *Phakopsora pachyrhizi*, é uma importante doença em todas as áreas produtoras de soja, incluindo os dois países maiores produtores, Brasil e Estados Unidos. Genes de resistência (*Rpp*) à *P. pachyrhizi* vêm sendo constantemente pesquisados na soja e em espécies relacionadas, sendo que alguns deles já foram identificados. No entanto, devido à plasticidade genética de *P. pachyrhizi*, esses genes de resistência são rapidamente suplantados. Com o intuito de avaliar o nível de variabilidade genética de *P. pachyrhizi* nas principais áreas produtoras de soja no mundo e entender como essa variabilidade é distribuída foi feita uma amostragem de *P. pachyrhizi* em diferentes países e continentes. A maior parte da amostragem foi feita no Brasil e nos Estados Unidos. A variabilidade genética foi estimada a partir de dados moleculares. Três regiões gênicas nucleares (fator de ribosilação do ADP, histona H4 e ITS) foram amplificadas, clonadas e sequenciadas de amostras de *P. pachyrhizi* dos diferentes locais. Um banco de dados foi criado com todas as sequências para cada região gênica. As sequências foram comparadas, e os polimorfismos identificados permitiram a diferenciação dos clones. Entre as regiões nucleares, o ITS (espaçador transcrito interno) é a principal marca utilizada para estudos de diversidade genética de fungos, tendo como base o sequenciamento do DNA. As estruturas secundárias formadas pelas regiões de ITS1 e ITS2 de *P. pachyrhizi* foram utilizadas para: examinar o padrão de polimorfismos de ITS; identificar características estruturais das regiões ITS; investigar a presença de pseudogenes

ou artefatos de PCR e avaliar a utilidade do ITS em estudos de diversidade genética de *P. pachyrhizi*. Os resultados obtidos revelaram que as análises das estruturas secundárias permitiram avaliar a confiabilidade das mutações no ITS de *P. pachyrhizi*, possibilitando a utilização dessa região para estudos de diversidade genética. O sequenciamento das três regiões gênicas permitiu a diferenciação dos clones em haplótipos. Uma rede foi estimada para cada um dos bancos de dados para entender a relação genealógica entre os haplótipos. Os haplótipos foram plotados em mapas, nos respectivos locais de amostragem, para uma melhor compreensão da distribuição da variabilidade genética de *P. pachyrhizi*. Foram ainda estimados os índices de diversidade e de fluxo gênico para avaliar a estrutura genética de populações de *P. pachyrhizi*. Os resultados revelaram que *P. pachyrhizi* apresenta uma alta variabilidade genética e que grande parte dessa variabilidade está concentrada dentro das populações. Tais características podem ser explicadas pela alta taxa de mutação do fungo e pela alta capacidade de dispersão dos uredósporos, fazendo com que populações de diferentes países e continentes compartilhem os mesmos haplótipos. A comparação da variabilidade genética de *P. pachyrhizi* entre dois anos agrícolas brasileiros revelaram que o vazio sanitário não foi eficiente no estrangulamento genético de populações desse fungo. Ainda, a identificação dos mesmos haplótipos identificados na soja e no feijão kudzu, revela a relevância desse hospedeiro alternativo nos períodos de ausência de soja no campo. A distribuição geográfica dos haplótipos e as análises de fluxo gênico permitiram inferir as possíveis rotas de dispersão e colonização de *P. pachyrhizi*. Populações de *P. pachyrhizi* identificadas no Brasil apresentaram forte relação genealógica com as populações africanas, fortalecendo a hipótese desse continente como fonte de origem das populações brasileiras. Os Estados Unidos apresentaram populações de *P. pachyrhizi* distantemente relacionadas às identificadas nos outros países, dificultando a inferência do local de sua origem.

ABSTRACT

FREIRE, Máira Cristina Menezes, D. Sc., Universidade Federal de Viçosa, July, 2011. **Molecular phylogeography of *Phakopsora pachyrhizi* based on data from sequencing of nuclear DNA regions.** Adviser: Luiz Orlando de Oliveira. Co-advisers: Maurílio Alves Moreira, Eduardo Seiti Gomide Mizubuti and Álvaro Manuel Rodrigues Almeida.

Asian soybean rust caused by *Phakopsora pachyrhizi* is an important disease in all soybean producing areas, including the two largest producing countries Brazil and the United States. Resistance genes (Rpp) to *P. pachyrhizi* have been constantly studied in soybean and related species, and some genes have been identified. However, due to the genetic plasticity of *P. pachyrhizi* these resistance genes are quickly overcome. In order to assess the level of genetic variability of *P. pachyrhizi* in major soybean producing areas of the world and understand how this variability is distributed a broad sampling of *P. pachyrhizi* from different countries and continents was made. Most sampling was carried out in Brazil and the United States, the two largest producers. Following sampling genetic variability was estimated from molecular data. Three nuclear gene regions (ADP ribosylation factor, histone H4, and ITS) were amplified, cloned and sequenced from samples of *P. pachyrhizi* from different locations. A database was created with all the gene sequences for each region. The sequences were compared, and polymorphisms were identified allowing differentiation of the clones. Among the nuclear regions, the ITS (internal transcribed spacer) is the best marker based on the sequencing of DNA used in genetic diversity studies of fungi. The secondary structures formed by the ITS1 and ITS2 regions of *P. pachyrhizi* were used to examine the pattern of ITS polymorphisms, identify structural characteristics of the ITS regions, investigate the presence of pseudogenes or PCR artefacts and to evaluate the usefulness of ITS in studies of genetic diversity of *P. pachyrhizi*. The

data indicated that the analysis of secondary structures permits us to assess the reliability of mutations in ITS in *P.pachyrhizi*, allowing the use of this region for studies of genetic diversity. The sequencing of the three genetic regions allowed the differentiation of clones in haplotypes. A network of haplotypes was estimated for each of the databases to understand the genealogical relationship between the haplotypes. The haplotypes were plotted on maps in the respective sampling sites, to better understand the distribution of genetic variability of *P. pachyrhizi*. Diversity index and gene flow were estimated to assess the genetic structure of populations of *P. pachyrhizi*. The results showed that *P.pachyrhizi* has a high level of genetic variability, and that much of this variability is concentrated within populations. These characteristics can be explained by the high mutation rate of the fungus, and the high dispersal ability of spores, causing populations of different countries and continents share the same haplotype. The comparison of the genetic variability of *P. pachyrhizi* over two years revealed that the Brazilian agricultural field fallowing was not an effective genetic bottleneck of populations of this fungus. Still, the identification of these haplotypes observed in soybean, infecting the bean kudzu, reveals the importance of this alternate host during periods of absence of soybeans in the field. The geographical distribution of haplotypes and analysis of gene flow allowed us to infer the possible routes of dispersal and colonization of *P. pachyrhizi*. Populations of *P.pachyrhizi* identified in Brazil showed strong genealogical relationship with the African populations, strengthening the hypothesis that this continent was a source of origin of Brazilian populations. The United States had populations of *P. pachyrhizi* distantly related to those identified in other countries, complicating the inference of the origin of these populations.

INTRODUÇÃO GERAL

Fungos fitopatogênicos têm efeito substancial na agricultura mundial, podendo resultar em severos prejuízos econômicos (Nunney *et al.*, 2010). A dispersão a longa distância de esporos pela ação do vento é o principal meio de dispersão desses organismos, possibilitando a sua sobrevivência e colonização de novas áreas. Essa via de dispersão é especialmente relevante para fungos biotróficos obrigatórios, nos quais um processo de extinção e colonização pode estar constantemente ocorrendo (Brown e Hovmoller, 2002).

A capacidade de fungos patogênicos atingirem novas áreas depende da presença de hospedeiros suscetíveis. Grandes culturas, como a soja, foram redistribuídas a partir do seu centro de origem, sendo atualmente identificadas em praticamente todos os continentes. Essa ampla distribuição permite uma dispersão intercontinental dos patógenos, podendo alcançar uma escala global. O risco de propagação global da doença é ainda aumentado pela limitada diversidade genética de muitas culturas modernas (Brown e Hovmoller 2002).

Phakopsora pachyrhizi (Sydow), agente causador da ferrugem asiática da soja, é um patógeno biotrófico responsável por severos prejuízos econômicos à cultura da soja em todas as regiões produtoras. Identificado pela primeira vez no Japão em 1902, esse fungo se dispersou gradativamente para diferentes países e continentes alcançando também os dois maiores produtores de soja, Brasil e Estados Unidos (Schneider *et al.*, 2005). A hipótese mais aceita para essa ampla colonização de *P. pachyrhizi* é dispersão de uredósporos a longa distância por ação do vento (Isard *et al.*, 2005).

A principal forma de controle a *P. pachyrhizi* é com aplicação de fungicidas. No entanto, além dessa metodologia elevar muito os custos da produção, o nível de controle depende da pressão da doença e da identificação do momento exato da

aplicação de fungicidas (Mueller *et al.*, 2009). A melhor maneira de controlar a incidência de *P. pachyrhizi* seria a utilização de linhagens de soja resistentes ao fungo. Genes de resistência a *P. pachyrhizi* (*Rpp*) vêm sendo identificados em diferentes linhagens de soja, contudo essa resistência é rapidamente suplantada, devido à variabilidade genética do fungo (Garcia *et al.*, 2008; Hartman *et al.*, 2005; Goellner *et al.*, 2010).

Vários estudos vêm sendo realizados com o intuito de avaliar o nível de diversidade genética de populações de *P. pachyrhizi* (Freire *et al.*, 2008; Barnes *et al.*, 2008; Twizeyimana *et al.*, 2009; Twizeyimana *et al.*, 2011). Apesar de já existirem estudos da variabilidade genética desse fungo em diferentes países com resultados semelhantes, uma análise da variabilidade genética em nível global se torna necessária.

Estudos da diversidade genética de organismos tendo como base dados moleculares tem se tornado cada vez mais comuns. Entre as técnicas moleculares, o sequenciamento de regiões gênicas para identificação de mutações que diferenciam os organismos em nível intra-específicos tem se revelado promissor. A região espaçadora transcrita interna (ITS) do DNA ribossomal nuclear é uma das principais marcas para esse tipo de estudo (Hillis e Dixon, 1991; Alvarez e Wendel, 2003).

O ITS apesar de ser transcrito não é traduzido, sendo removido pelo processamento do rRNA. Uma vez que o ITS não é traduzido, acredita-se que essa região evolua de forma neutra, não existindo nenhum tipo de pressão evolutiva controlando sua evolução. Essa característica é de grande interesse para estudos de diversidade genética em nível intra-específico, pois se espera que o ITS acumule um grande número de mutações permitindo esse nível de diferenciação (Thines, 2007). No entanto, estudos da estrutura secundária formada pela região ITS na etapa de processamento do rRNA evidencia a existência de motivos conservados necessários a esse processamento, e que portanto, deve existir algum tipo de pressão para garantir a manutenção da estrutura secundária e dos motivos conservados (Hillis and Dixon, 1991; Nazar 2004; Mullineux and Hausner, 2009)

Tendo como base essas informações um estudo preliminar foi realizado com a região de ITS de *P. pachyrhizi*. Nesse estudo foram avaliados a sequência primária e as estruturas secundárias de ITS1 e ITS2 de *P. pachyrhizi*. Esse estudo teve como principal objetivo verificar a utilidade desses espaçadores para estudos de diversidade genética de *P. pachyrhizi*.

Para estudar a diversidade genética de *P. pachyrhizi* duas outras regiões gênicas nucleares, além do ITS, foram sequenciadas. Amostras foram obtidas de

diferentes partes do mundo, abrangendo as principais áreas produtoras de soja. As sequências de cada região gênica foram comparadas e diferenciadas em haplótipos. A diferenciação de organismos dentro de uma mesma espécie em haplótipos e a estimativa da rede de haplótipos pode fornecer dados das relações genealógicas entre os haplótipos (Hewitt, 2001).

Os haplótipos foram plotados nos mapas dos respectivos locais de amostragem. A análise da rede de haplótipos combinada com a visualização da distribuição geográfica permite não só inferir sobre a variabilidade genética do organismo, mas também como essa variabilidade se distribui e quais os possíveis processos evolutivos envolvidos na estruturação da população atual (Hewitt, 2004). Esses dados possibilitarão a inferência da história de dispersão de *P. pachyrhizi*, e fornecerão informações de populações desse fungo.

REVISÃO DE LITERATURA

1- A cultura da soja

A soja (*Glycine max.* L.) é economicamente uma das principais culturas em todo o mundo. É um grão versátil que dá origem a produtos e subprodutos muito usados pela agroindústria, indústria química e de alimentos. É valiosa como fonte de proteínas, óleos e vários nutrientes, incluindo antioxidantes e isoflovanóides (Soria-Guerra *et al.*, 2010). Apesar de ser muito importante na alimentação de animais e humana, na indústria de adesivos, na produção de adubos, na fabricação de fibras, entre outros, sua utilização mais conhecida é como óleo refinado. Ainda, recentemente, a soja vem crescendo também como fonte alternativa de combustível.

A soja é nativa da China, onde tem sido cultivada desde o século XI (Twizeyimana *et al.*, 2008), e onde ficou restrita por um longo tempo. No século XV a cultura foi introduzida na Europa, nos jardins botânicos da Inglaterra, França e Alemanha. Devido ao teor de óleo e proteína do grão, a soja começou a despertar interesse nas indústrias mundiais a partir da segunda década do século XX. Mas, foi apenas na década de 60 que a cultura chegou ao Brasil. Desde então, o país passou a investir em tecnologia para adaptação da cultura às condições climáticas brasileiras.

A produção de soja mundial é cerca de 260 milhões de toneladas por ano, com uma área plantada de mais de 102 milhões de hectares (Embrapa, 2011). Os Estados Unidos ocupam a primeira posição como maior produtor e exportador de soja (Soria-Guerra *et al.*, 2010). Na safra de 2009-2010, a produção de soja nos EUA chegou a alcançar mais de 91,4 milhões de toneladas (Embrapa, 2011). O

Brasil é o segundo maior produtor e tem um grande potencial para aumentar a produção no futuro dada a existência de áreas ainda não cultivadas e o aumento mundial da demanda de produtos oriundos da soja (Silva *et al.*, 2008). A produtividade média da soja brasileira é de 2941 kg/ha, chegando a alcançar cerca de 3036 kg/ha no estado do Mato Grosso, o maior produtor brasileiro da soja. O Brasil tem produzido, em cada safra, cerca de 60 milhões de toneladas de soja (Embrapa, 2011). Em 2010, o mercado da soja foi marcado pelo aumento da produção mundial, que permitiu a recomposição dos estoques mundiais do grão, e teve grande impacto na flutuação das cotações mundiais (Embrapa, 2011).

As doenças que incidem na cultura da soja têm assumido papel importante na definição da sua produtividade. As perdas anuais de produção devido as doenças são estimadas em cerca de 15 a 20% (Godoy *et al.*, 2006). Entre os problemas ameaçando a produção e a expansão da soja, a ferrugem asiática da soja causada pelo fungo *P. pachyrhizi*, é responsável por severos prejuízos econômicos (Silva *et al.*, 2008). Esses prejuízos comprometem a rentabilidade dos produtores e a economia do país onde a soja possui destaque dentro do agronegócio (Yorinori e Lazzarotto, 2004).

2 - *Phakospora pachrhizi* (ferrugem asiática da soja)

2.1 – Histórico

O agente causador da ferrugem asiática da soja foi primeiramente registrado em 1902 no Japão. No entanto, foi apenas em 1914, quando foi identificado em Taiwan, que esse fungo foi nomeado como *Phakopsora pachyrhizi* Sydow. Em 1934, o fungo foi identificado pela primeira vez na Austrália. Até o ano de 1994 não havia registro da ferrugem asiática da soja no Novo Mundo, quando essa foi identificada pela primeira vez no Havaí. Esse fungo foi então identificado pela primeira vez na África no ano de 1997 e na América do Sul no ano de 2001.

No ano de 1976 a ferrugem da soja, *Phakopsora meibomia*, foi identificada pela primeira vez no continente sul americano. Taxonomistas dessa época não foram capazes de fazer distinção entre isolados asiáticos e isolados latino americanos, portanto, ambos isolados foram vistos como sendo da mesma espécie (Goellner *et al.*, 2010). Baseado na diferença de virulência, Bromfield (1984) propôs que isolados americanos poderiam representar uma *forma specialis* que seria menos agressiva e menos adaptada à soja. No entanto, em um estudo comparativo de isoenzimas realizado por Bonde (1988), um alto nível de polimorfismo foi

identificado entre os isolados asiáticos e americanos, levantando a possibilidade que essas poderiam ser, de fato, espécies distintas. Tal hipótese foi posteriormente confirmada por estudos morfológicos (Ono *et al.*, 1992), e por análise comparativa da sequência de nucleotídeos da região espaçadora interna (ITS) ribossomal (Frederick *et al.*, 2002). Hoje, já é conhecido que ferrugem da soja é causada por pelo menos duas espécies: ferrugem asiática da soja, causada pela *P. pachyrhizi*; e a ferrugem americana, causada pela *P. meibomiae* (Arthur), que é menos agressiva.

Após o ano de 2001 a ferrugem se estabeleceu em todas as áreas produtoras de soja do Brasil e nos outros países sul americanos. Na safra 2002-2003, a doença se espalhou por toda área de plantação de soja brasileira causando prejuízo estimado de US\$2 bilhões (Yorinori *et al.*, 2005). A dispersão a longas distâncias dos esporos é atribuída a tempestades de vento, mas alguns autores ainda acreditam na interferência humana (Goellner *et al.*, 2010).

A doença alcançou o continente norte americano apenas no ano de 2004 (Schneider *et al.*, 2005). No entanto, o alto nível de prejuízo esperado não ocorreu nos Estados Unidos, quando comparado aos prejuízos alcançados no Brasil. Tal fato se deve possivelmente a condições de clima não favoráveis para a dispersão do patógeno e/ou que a população do patógeno não se tornou completamente adaptada ao novo habitat (Christiano e Scherm, 2007). Uredósporos de *P. pachyrhizi* não são tolerantes a geadas, dessa forma, o patógeno deve procurar refúgios em hospedeiros alternativos no sul dos Estados Unidos (Goellner *et al.*, 2010). No verão, o patógeno pode se espalhar para as áreas de plantação de soja, no entanto, o fungo nem sempre é capaz de alcançar todas as áreas produtoras, principalmente aquelas localizadas ao norte do país (Hartman e Haudenschild, 2009).

Apesar do impacto da ferrugem asiática na produção de soja nos EUA ter sido menor do que o previsto, o estabelecimento de *P. pachyrhizi* nos EUA fornece a oportunidade para a evolução de populações adaptadas ambientalmente e mais virulentas (Fang *et al.*, 2010). Portanto, a presença da ferrugem asiática da soja no continente norte americano tem se tornado o principal foco das indústrias e produtores de soja, com o intuito de evitar um grande impacto na produção total da soja e nos custos dessa produção (Soria-Guerra *et al.*, 2010).

2.2 - Ciclo de vida

Após o contato com o hospedeiro, uredósporos rapidamente germinam levando ao desenvolvimento de um apressório de tamanho similar ao esporo (Hoefle *et al.*, 2009). A infecção de *P. pachyrhizi* se dá pela penetração direta das folhas, sem a utilização das aberturas estomatais. A hifa de penetração produzida a partir do apressório penetra diretamente na epiderme da planta hospedeira. Na penetração, as células da epiderme colapsam, tornam-se desorganizadas e mostram sinais de morte celular (Goellner *et al.*, 2010). O papel da morte das células epidérmicas para a resistência ainda não é conhecida (Hoefle *et al.*, 2009).

Após a penetração, a hifa cresce através das células da epiderme até alcançar o espaço intracelular. A hifa primária pode ramificar para formar a hifa secundária e o haustório em contato com as células do mesófilo. A formação do haustório ocorre de 24 a 48 horas após o contato inicial do uredósporo com o hospedeiro suscetível. Novas urédias são desenvolvidas a partir da urédia original, completando o ciclo assexuado de *P. pachyrhizi* (Goellner *et al.*, 2010). Uredósporos assexuais são formados dentro da urédia 5-8 dias após a inoculação nas folhas. Esses uredósporos são liberados através do ostíolo e dispersos pelo vento. Não existe nenhuma informação do ciclo sexual de *P. pachyrhizi*. Apesar de teliósporos terem sido observados em vários hospedeiros, inclusive na soja, a germinação deles não foi registrada na natureza (Bromfield, 1984; Goellner *et al.*, 2010).

Tanto a infecção quanto a prevalência de *P. pachyrhizi* dependem de condições favoráveis, como temperaturas noturnas quentes e presença de água na superfície da folha, as quais facilitam o desenvolvimento e a propagação do fungo (Mendes *et al.*, 2009). O fungo está adaptado a temperaturas que variam de 15°C a mais de 30°C e pode causar severas perdas da soja em todas as regiões onde ocorram períodos de molhamento da folha por mais de 10 horas (Yorinori e Lazzarotto, 2004). Os sintomas iniciais da doença são pequenas pústulas marrons escuras na superfície inferior da folha. Lesões também podem aparecer nos pecíolos, nas vagens e no caule (Mendes *et al.*, 2009).

2.3 - Formas de controle

A interação *P. pachyrhizi* - soja pode gerar três diferentes tipos de fenótipos: TAN, RB e imune. O fenótipo TAN, observado em cultivares suscetíveis, é caracterizado por lesões de cor castanho com urédias esporulando. Os cultivares resistentes desenvolvem lesões marrom avermelhadas (RB) com pouca ou

nenhuma esporulação. O fenótipo imune é uma interação incompatível sem nenhum sintoma visível da doença na folha hospedeira (Bromfield, 1984; Goellner *et al.*, 2010).

Sete genes de resistência já foram identificados em diferentes cultivares, apresentando, cada um, resistência para diferentes isolados de *P. pachyrhizi* (Goellner *et al.*, 2010). Esses sete locos são referidos como *Rpp1* (Cheng e Chan, 1968; Hidayat e Somaatmadja, 1977; McLean e Byth, 1980; Hartwig e Bromfield, 1983), *Rpp1b* (Ray *et al.* 2009), *Rpp2* (Hidayat e Somaatmadja, 1977; Bromfield, 1980), *Rpp3* (Bromfield e Hartwig, 1980), *Rpp4* (Hartwig, 1986), *Rpp5* (Garcia *et al.* 2008) and *Rpp?* (Monteros *et al.*, 2007). No entanto, ainda não há disponível uma variedade de soja com resistência a todos os isolados de população do patógeno (Hartman *et al.*, 2005). Entre esses genes apenas o *Rpp1* confere resposta imune a um particular isolado virulento, todos os outros genes conferem apenas resistência, dessa forma, lesões do tipo RB são observadas (Pandey *et al.*, 2010). A resistência mediada por esses genes têm sido constantemente suplantada por *P. pachyrhizi* devido à grande capacidade desse fungo de gerar novas raças (Bromfield, 1984).

Com base nesse fato, muitos pesquisadores estão se focando na tentativa de obter cultivares com resistência mais durável e de amplo espectro, por diferentes métodos. Análises de acessos de soja têm sido realizadas com o intuito de recrutar novas fontes de resistência ao fungo *P. pachyrhizi* (Garcia *et al.*, 2008; Hartman *et al.*, 2005; Goellner *et al.*, 2010). O rastreamento da resistência a esse fungo tem sido também expandido para espécies relacionadas à soja (Hartman *et al.*, 2005).

A ausência de cultivares resistentes faz com que o manejo da doença, por meio de aplicação de fungicidas, seja uma alternativa que viabiliza o cultivo da soja na presença da ferrugem (Godoy e Flausino, 2008). Entretanto, o nível de controle depende da pressão da doença, exercida pelos diferentes isolados; e do preciso tempo de aplicação de fungicida, ou seja, da rápida e precisa identificação da doença e do momento correto de se aplicar o fungicida (Goellner *et al.*, 2010). A detecção dos primeiros sintomas da doença é, portanto, de suma importância para que a decisão para o controle químico seja tomada no momento correto (Yorinori e Lazzarotto, 2004).

A falta de resistência durável da soja comercial e o elevado custo do controle utilizando fungicidas têm criado a necessidade para meios alternativos de controle da ferrugem. Fang *et al.* (2010) avaliaram a possibilidade de se utilizar abordagens farmacológicas nesse controle. Esses autores propuseram a utilização de peptídeos que se ligam ao tubo germinativo de *P. pachyrhizi* impedindo o seu desenvolvimento. O controle biológico também tem sido avaliado, e vários

resultados já vêm sendo alcançados em laboratório. No entanto, não se pode esperar que o controle biológico venha substituir completamente, ou significativamente, o controle químico da ferrugem, principalmente por causa das condições ambientais necessárias (Goellner *et al.*, 2010).

Uma combinação de estratégias tem sido a melhor forma de se reduzir o risco de danos provocados pela ferrugem asiática. Entre as estratégias recomendadas estão: a utilização de cultivares de ciclo precoce e semeaduras no início da época recomendada; a eliminação de plantas de soja voluntárias e a ausência de cultivo de soja na entressafra, por meio do vazio sanitário; o monitoramento da lavoura desde o início do desenvolvimento da cultura; a utilização de fungicidas na fase inicial de aparecimento dos sintomas; e a utilização de cultivares resistentes, quando disponíveis (Embrapa, 2011).

2.4 - Variabilidade de *Phakopsora pachyrhizi*

Fungos patogênicos de plantas possuem uma variedade de mecanismos pelos quais a variação genética pode ser gerada ou expandida, incluindo recombinação sexual, migração, mutação espontânea e hibridização somática (McDonald *et al.*, 1989). Uma vez que *P. pachyrhizi* é conhecido por se reproduzir apenas assexuadamente, era de se esperar que populações desse fungo apresentassem baixa variabilidade genética (Twizeyimana *et al.*, 2011). No entanto, a alta severidade da ferrugem e o fato de cultivares resistentes em um determinado ano terem sua resistência suplantada no ano seguinte são indícios de alta variabilidade genética do fungo (Yorinori e Lazzarotto, 2004). O conhecimento da variação genética de *P. pachyrhizi* e como essa variação se distribui entre populações é um importante fator para a compreensão da biologia de populações desse fungo, no entanto, poucas informações desse tipo estão disponíveis (Twizeyimana *et al.*, 2011).

Estudos da variação de populações de *P. pachyrhizi* têm se focado, na sua grande maioria, nas diferenças qualitativas da virulência. Essa natureza qualitativa é baseada nos três tipos diferentes de respostas da interação soja-*P. pachyrhizi* (Twizeyimana *et al.*, 2009). No entanto, devido a dificuldade em classificar reações intermediárias baseadas na designação TAN/RB, abordagens quantitativas já vêm sendo adotadas. Abordagens quantitativas são baseadas nas diferenças do desenvolvimento do patógeno em um dado hospedeiro, como por exemplo, número e tamanho de urédias produzidas.

Twizeyimana *et al.* (2009) avaliaram a diversidade de *P. pachyrhizi* na Nigéria baseados em dados de medidas quantitativas da virulência. Nesse estudo, uma considerável diversidade foi revelada dentro das populações de *P. pachyrhizi* ocorrendo nas áreas de produção. Usando técnicas multivariadas, os autores foram capazes de classificar os isolados do fungo em diferentes patótipos, utilizando apenas quatro genótipos de soja como plantas diferenciadoras (Twizeyimana *et al.*, 2009). No Japão, 18 raças foram identificadas em amostras coletadas em plantas de soja e hospedeiros selvagens (Yamaoka *et al.*, 2002).

Outra forma de se estudar a variabilidade genética de um organismo seria a partir de dados moleculares. Barnes *et al.* (2008) investigaram a variação genotípica de *P. pachyrhizi* a partir de esporos obtidos na chuva em toda a área produtora de soja nos Estados Unidos. Nesse estudo, a variação genotípica foi inferida a partir de dados moleculares. Amostras do fungo tiveram a região de ITS1 amplificadas e comparadas, com o intuito de identificar as diferenças entre eles. Para investigar a variação genotípica, duas marcas foram analisadas na região de ITS1, a partir de 98 sequências. Essas análises permitiram a identificação de seis diferentes genótipos. Os autores observaram que dois desses genótipos foram predominantes, e sugeriram que esses possam ser populações do patógeno predominante nos Estados Unidos e talvez na fonte de origem. Outro importante resultado obtido nesse trabalho foi a identificação da alta diversidade genotípica dentro de uma mesma localidade. Os autores concluíram que deve haver um genótipo predominante de *P. pachyrhizi* nos Estados Unidos, mas com baixo nível de outros genótipos misturados na população.

Em um estudo prévio, nosso grupo avaliou a variabilidade genética de isolados *P. pachyrhizi* no Brasil, baseado nos polimorfismos das regiões de ITS1 e ITS2 (Freire *et al.*, 2008). Dados similares aos observados por Barnes *et al.* (2008) foram encontrados para os isolados brasileiros: grande número de haplótipos para ambas regiões gênicas; prevalência de dois haplótipos em toda a região brasileira amostrada; alta variabilidade genética dentro de uma mesma localidade; e a identificação de possíveis haplótipos que migraram a partir da fonte de origem.

Ao compararmos as sequências da região de ITS1 de isolados brasileiros, obtidas no nosso trabalho, com as sequências dos isolados americanos, obtidos por Barnes *et al.* (2008), identificamos que os haplótipos de alta frequência no Brasil são os mesmos haplótipos de alta frequência nos Estados Unidos. Ainda, verificamos que os haplótipos de menor frequência foram específicos de cada país. Tais resultados fortaleceram a idéia da relação ancestral-derivado desses haplótipos (Freire *et al.*, 2010).

Ainda, Twizeyimana *et al.* (2011) utilizaram o marcador SSR (repetição de sequências únicas) para classificar 116 isolados de *P. pachyrhizi*, a partir da Nigéria. Oitenta e quatro genótipos foram identificados entre os 116 isolados analisados, revelando o alto grau de variação genética de *P. pachyrhizi* nesse país. Os dados da distribuição genética foram consistentes com os observados no Brasil por Freire *et al.* (2008), e nos Estados Unidos por Barnes *et al.* (2008) - a maior parte da diversidade genética localizada dentro das localidades; e baixa diferenciação genética entre as diferentes localidades.

3 - Filogeografia molecular

O termo filogeografia tem pouco mais de 20 anos (Avice *et al.*, 1987). É uma ciência recente e em rápido desenvolvimento que trata da distribuição geográfica de linhagens genealógicas (Avice, 1998). Estudos filogeográficos são realizados com o objetivo de entender os processos que influenciam na distribuição da variação genética dentro e entre espécies proximamente relacionadas. Tais estudos fornecem informações detalhadas de como eventos geológicos, influências ambientais, e fatores geográficos interagem com aspectos da ecologia da espécie e da história natural modelando sua evolução (Knowles, 2009). Esta abordagem tem sido empregada com sucesso para explorar associação entre rede de haplótipos e geografia em um grande número de espécies vegetais, animais e de micro-organismos (Templeton, 2004).

A análise da distribuição espacial da genealogia de genes forma a base da filogeografia, na qual os indivíduos são genotipados, distribuídos e são relacionados com padrões de distribuição geográfica. Essa ciência vem crescendo a partir da capacidade de genotipar os indivíduos por variações identificadas de sequências de DNA e, a partir destas, reconstruir a filogenia (Hewitt, 2004). Tal habilidade para obter sequências de DNA e deduzir sua história a partir da divergência representa uma grande perspectiva. Vários tipos de dados e inferências são combinados para elucidar muitas questões evolutivas, nos quais, genes podem ser vistos no tempo e no espaço (Hewitt, 2001).

As informações obtidas a partir dos polimorfismos das sequências de DNA são plotadas geograficamente para expor sua relação espacial e deduzir a origem evolutiva e a história da população, espécies ou subespécies (Hewitt, 2004). Estudar o padrão de variação genética no contexto de árvores ou rede de haplótipos tem contribuído consideravelmente para a compreensão de como os fatores têm influenciado na estrutura da população e divergência das espécies

(Avice, 1994; Knowles e Maddison, 2002). A representação em diagramas e figuras é mais esclarecedora, e transmite idéias complexas com explicações sucintas (Hewitt, 2001).

Templeton *et al.* (1992, 1995) evidenciaram uma explicação detalhada para a construção de redes de haplótipos usando o método da parcimônia por um modelo local finito da evolução do DNA. Este método, além de estimar uma rede da relação dos haplótipos, permite também estimar o número máximo de conexões parcimoniosas entre pares de sequências com 95% de probabilidade para todos os acoplamentos dos haplótipos dentro da rede. O “software” TCS (ver. 1.13) desenvolvidos por Clement *et al.* (2000) pode ser usado para a construção dessas redes de haplótipos.

Templeton *et al.* (1992) formularam algumas predições com base na teoria da coalescência em relação a redes de haplótipos. Uma rede de haplótipos apresentará multifurcações, pois se espera que haplótipos ancestrais sofram várias mutações, em tempos diferentes, podendo originar vários haplótipos descendentes. É esperado que haplótipos ancestrais e haplótipos descendentes estejam presentes em uma mesma rede, visto que seria extremamente improvável que todos os indivíduos do haplótipo ancestral tenham sofrido a mesma mutação. Na formação da rede, os haplótipos de origem mais recente devem ser encontrados nas pontas, enquanto os haplótipos ancestrais irão ser encontrados no interior da rede. De acordo com a distribuição geográfica é esperado que haplótipos ancestrais se apresentem geograficamente mais espalhados, enquanto haplótipos recentes tenham uma distribuição mais restrita, visto que não teriam tido tempo suficiente para uma maior dispersão. Esta análise combina um cladograma das relações genéticas obtidas entre haplótipos das populações em estudo com as suas respectivas distribuições geográficas, e desta forma reconstrói a história evolutiva das populações.

O mtDNA tem sido o principal foco de estudos filogeográficos em animais. No entanto, outras sequências são usadas, mais comumente o DNA cloroplastídeo em plantas, e regiões nucleares não codificantes em plantas e animais (Hewitt, 2004). Estudos filogeográficos usando dados de haplótipos nucleares tem sido um grande avanço para testar modelos históricos mais sofisticados (Hare, 2001). DNA nuclear não codificante, tipicamente, tem mais variação do que sequências codificantes adjacentes, por essa razão, íntrons de DNA têm sido alvo frequente de estudos em nível intra-específico e os exons relativamente conservados têm sido usados para desenhar primers (Hare, 2001). Independentemente da região a ser usada, uma explícita consideração dos limites de cada região deve ser levado em

consideração, para se fazer uma acurada inferência histórica, evitando uma super interpretação das análises.

Técnicas para obter informações de sequências de DNA permanecem avançando rapidamente, com sequências de genomas completos sendo produzidas em crescente número de organismos. Esse fato permite a identificação de sequências e o desenvolvimento de marcas para muitos tipos de investigações, sendo algumas úteis para estudos filogeográficos.

Inferências filogeográficas são usualmente derivadas por estudar a história genealógica reconstruída de genes individuais amostrados em diferentes populações. Entretanto, uma vez que muitos eventos podem ter ocorrido no passado (expansão da população, estrangulamento genético, vicariância e migração), e linhagens gênicas podem se perder, a história de uma espécie pode não ser facilmente inferida a partir da genealogia de um único gene (Knowles e Maddison, 2002). Dados genealógicos são requeridos a partir de vários locos independentes para fornecer uma completa e mais confiável história da espécie. O padrão filogeográfico observado por comparar locos evoluindo independentemente fornece pistas importantes para o esclarecimento do processo evolutivo (Hare, 2001). Polimorfismos únicos de nucleotídeos (SNPs) estão também se tornando disponíveis em todos os genomas, os quais irão produzir compreensivas medidas da diversidade genética e permitirá uma melhor construção da história da população (Hewitt, 2004).

Em estudos filogenéticos, sequências de diferentes locos são combinadas (concatenadas) e analisadas como um único gene, dessa forma o problema de diferentes árvores gênicas para locos diferentes é eliminado. Apenas uma árvore é formada pela concatenação. Em contraste, abordagens filogeográficas têm adotado metodologias que levam em conta as genealogias discordantes ao invés de forçar a formação de uma única história genealógica a partir de diferentes locos, o qual poderá ser uma violação dos padrões de evolução genética. De acordo com Knowles (2009) a variância entre locos fornece informações relevantes para entender a história evolutiva.

O avanço na tecnologia do DNA está produzindo dados ricos de indivíduos, populações e espécies, e existe um concomitante desenvolvimento nos métodos analíticos para predizer a história demográfica, a relação evolutiva, e para realizar testes pertinentes e informativos (Hewitt, 2004). Com o recente avanço da filogeografia estatística, as estimativas dos parâmetros de uma população genética de uma determinada espécie podem ser inferidas com impressionante precisão e acurácia (Knowles, 2009).

A escolha de um teste estatístico é um importante aspecto, já que ele pode determinar como os dados podem ser usados com maior eficiência (Knowles e Maddison, 2002). O conhecimento do sistema biológico em estudo é essencial para decidir qual é a abordagem ideal para as análises filogeográficas (Knowles, 2009). Uma vez que diferentes testes estatísticos podem extrair informações relevantes para diferentes questões, torna se evidente a necessidade de considerar cuidadosamente qual informação poderá recuperar o sinal dos diferentes processos aos quais a população foi submetida (Knowles e Maddison, 2002).

REFERÊNCIAS BIBLIOGRÁFICAS

- Alvarez I and Wendel JF (2003) Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution*, 29:417-434.
- Avise JC (1994) *Molecular Markers, Natural History and Evolution*. Chapman and Hall, New York.
- Avise JC (1998) The history and purview of Phylogeography: a personal reflection. *Molecular Ecology*, 7: 371-379.
- Avise JC, Arnold J and Ball RM (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematic: *Annual Review of Ecology and Systematics*, 18:489-522.
- Barnes CW, Szabo, LJ and Bowersox VC (2008) Identifying and quantifying *Phakopsora pachyrhizi* spores in rain. *Phytopathology*, 99:328-338.
- Bonde MR (1988) A comparasion of isoenzymes of *Phakopsora pachyrhizi* from the eastern hemisphere and the new world. *Phytopathology*, 78:1491-1494.
- Bromfield KR (1984) *Soybean Rust*. St. Paul, MN: American Phytopathological Society.
- Bromfield KR, Melching JS and Kingsolver CH (1980) Virulence and aggressiveness of *Phakopsora pachyrhizi* isolates causing soybean rust. *Phytopathology*, 70:17-21.
- Bromfield KR and Hartwig EE (1980) Resistance to soybean rust and mode of inheritance. *Crop Science*, 20:254-255.
- Brown JKM and Hovmeller MS (2002) Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science*, 297:537-541.

Cheng YW and Chan KL (1968) The breeding of rust resistant soybean Tainung 3. Journal Taiwan Agricultural Research, 17:30-34.

Christiano RSC and Scherm H (2007) Quantitative aspects of the spread of Asian soybean rust in the southeastern United States, 2005-2006. Phytopathology, 97:1428-1433.

Clement M, Posada D and Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Molecular Ecology, 9:1657-1659.

Embrapa Soja 2011. Dados econômicos: Soja em números (safra 2009/2010). http://www.cnpso.embrapa.br/index.php?op_page=294&cod_pai=17. Acessado em 06 de julho de 2011.

Fang ZD, Marois JJ, Stacey G, Schoelz JE, English JT and Schmidt FJ (2010) Combinatorially selected peptides for protection of soybean against *Phakopsora pachyrhizi*. Phytopathology, 100:1111-1117.

Frederick RD, Snyder, CL, Peterson GL and Bonde MR (2002) Polymerase chain reaction assays for the detection and discrimination of the soybean rust pathogens *Phakopsora pachyrhizi* and *P. meibomia*. Phytopathology, 92: 217-277.

Freire MCM, de Oliveria LO, de Almeida, AMR, Schuter I, Moreira MA, Liebenberg MM and Mienie CMS (2008) Evolutionary history of *Phakopsora pachyrhizi* (the Asian soybean rust) in Brazil based on nucleotide sequences of the internal transcribed spacer region of the nuclear ribosomal DNA. Genetics and Molecular Biology, 31:920-931.

Garcia A, Calvo, E Souza KR, Harada A, Hiromoto D and Vieira L (2008) Molecular mapping of soybean rust (*Phakopsora pachyrhizi*) resistance genes: discovery of a novel locus and alleles. Theor. Applied Genetics, 117:545-553.

Godoy CV and Flausino AM (2008) Eficiência de fungicidas para o controle da ferrugem asiática da soja em Londrina e Tamarana, PR, na safra 2007/08. Circular Técnica 57, Embrapa:1-4.

Godoy CV, Pimenta CB, Miguel-Wruck DS and Ramos Junior EU (2007) Eficiência de fungicidas para controle da ferrugem asiática da soja, *Phakopsora pachyrhizi*, na safra 2006/07. Resultados sumarizados dos ensaios em rede. Circular Técnica 42, Embrapa:1-8.

Goellner K, Loehrer M, Langenbach C, Conrath U, Koch E and Schaffrath U (2010) *Phakopsora pachyrhizi*, the causal agent of Asian soybean rust. Molecular Plant Pathology, 11:169-177.

- Hare MP (2001) Prospects for nuclear gene Phylogeography. *Trends in Ecology & Evolution*, 16:700-706.
- Hartman GL, Miles MR and Frederick RD (2005) Breeding for resistance to soybean rust. *Plant Disease*, 89:664-666.
- Hartman GL and Haudenshield JS (2009) Movement of *Phakopsora pachyrhizi* (soybean rust) spores by non-conventional means. *European Journal Plant Pathology*, 123:225-228.
- Hartwig EE (1986) Identification of a fourth major gene conferring to rust in soybeans. *Crop Science*, 26:1135-1136.
- Hartwig EE and Bromfield KR (1983) Relationships among 3 genes conferring specific resistance to rust in soybeans. *Crop Science*, 23:237-239.
- Hewitt GM (2001) Speciation, hybrid zones and Phylogeography – or seeing genes in space and time. *Molecular Ecology*, 10:537-549.
- Hewitt GM (2004) The structure of biodiversity – insights from molecular Phylogeography. *Frontiers in Zoology*: 1-4.
- Hillis DM and Dixon MJ (1991) Ribosomal DNA: molecular evolution and phylogenetic inference. *Q Rev Biol* 66:411-453.
- Hoefle C, Loehrer M, Schaffrath U, Frank M, Scultheiss H and Huckelhoven R (2009) Transgenic suppression of cell death limits penetration success of the soybean rust fungus *Phakopsora pachyrhizi* into epidermal cells of barley. *Phytopathology*, 99:220-226.
- Hidayat OO and Somaatmadja S (1977) Screening of soybean breeding lines for resistance to soybean rust (*Phakopsora pachyrhizi* Sydow). *Soybean Rust News*, 1:9-22.
- Isard SA, Gage SH, Comtois P and Russo JM (2005) Principles of the atmospheric pathway for invasive species applied to soybean rust. *BioScience* 55:851-861.
- Knowles LL (2009) Statistical Phylogeography. *Annual Review Ecology Evolution System*, 40: 593-612.
- Knowles LL and Maddison WP (2002) Statistical Phylogeography. *Molecular Ecology*, 11:2623-2635.

- McDonald BA, MsDermott JM, Allard RW and Webster RK (1989) Coevolution of host and pathogen populations in the *Hordeum vulgare-Rhynchosporium secalis* pathosystem. *Proceeding of the National Academy of Sciences, USA* 86:3924-3927.
- McLean RJ, Byth DE (1980) Inheritance of resistance to rust (*Phakopsora pachyrhizi*) in soybeans. *Aust journal Agriculture Research*, 31:951-956.
- Mendes RK, Carvalhal RF, Stach-Machado DR and Kubota LT (2009) Surface Plasmon resonance immunosensor for early diagnosis of Asian rust on soybean leaves. *Biosensors and Bioelectronics*, 24: 2483-2487.
- Monteros MJ, Missaoui AM, Phillips DV, Walker DR and Poerma HR (2007) Mapping and confirmation of the “Hyyuga” red-bron lesion resistance gene for Asian soybean rust. *Crop Science*, 47:829-836.
- Mueller TA, Miles MR, Morel W, Marois JJ, Wrihr DL, Kemerait RC, Levy C and Hartman GL (2009) Effect of fungicide and timing of application on soybean rust severity and yield. *Plant Disease*, 93:243-248.
- Mullineux T and Hausner G (2009) Evolution of rDNA ITS1 and ITS2 sequences and RNA secondary structures within members of the fungal genera *Grosmannia* and *Leptographium*. *Fungal Genetic Biology* 46:855-867.
- Nazar RN (2004) Ribosomal RNA processing and ribosome biogenesis in eukaryotes. *IUBMB Life* 56:457-465.
- Nunney L, Yuan X, Bromley R, Hartung J, Montero-Astúa M, Moreira L, Ortiz B and Stouthamer R (2010) Population genomic analysis of a bacterial plant pathogen: novel insight into the origin of pierce’s disease of Grapevine in the U.S. *PLoP ONE*, 5:1-9.
- Ono Y, Buritica P and Hennen JF (1992) Delimitation of *Phakopsora*, *Physopella* and *Cerotelium* and their species on *Leguminosae*. *Mycological Research*, 96:825-850.
- Pandey AK, Yang C, Zhang C, Graham MA, Horstamn HD, Lee Y, Zabortina OA, Hill JH, Pandley KF and Whitham SA (2011) Functional analysis of the Asian soybean rust resistance pathway mediated by *Rpp2*. *Molecular Plant-Microbe Interactions*, 24:194-206.
- Ray JD, Morel W, Smith JR, Frederick RD and Miles MR (2009) Genetics and mapping of adult plant rust resistance in soybean PI 587886 and PI 587880A. *Theoretical and Applied Genetics*, 119:271-280.

Schneider RW, Hollier CA, Whitam HK, Palm ME, McKemy, JM, Hernandez JR, Levy L and DeVries-Peterson, R (2005) First report of soybean rust caused by *Phakopsora pachyrhizi* in the continental United States. *Plant Disease*, 89:774-774.

Silva D, Yamanaka N, Brogin R, Arias C, Nepomuceno A, Di Mauro A, Pereira S, Nogueira L, Passianotto AL and Abdelnoor R (2008) Molecular mapping of two loci that confer resistance to Asian rust in soybean. *Theor Applied Genetics*, 117:57-63.

Soria-Guerra RE, Mendoza SR, Chang S, Haudenshield JS, Padmanaban A, Rodriguez-Zas S, Hartman GL, Ghabrial AS and Korban SS (2010) Transcriptome analysis of resistant and susceptible genotypes of *Glycine tomentella* during *Phakopsora pachyrhizi* infection reveals novel rust resistance genes. *Theor Applied Genetics*, 120:1315-1333.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonucleasas mapping and DNA sequence III. Cladogram estimation. *Genetics*, 132:619-633.

Templeton AR (2004) Statistical Phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology*, 13:789-809.

Templeton AR, Routman E and Phillips CA (1995) Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics*, 140:767-782.

Thines M (2007) characterization and phylogeny of repetitive elements giving rise to exceptional length of ITS2 in several downy medley genera (Peronosporaceae). *Fungal Genetics and Biology*, 44: 199-207.

Twizeyimana M, Ojiambo PS, Haudenshield JS, Caetano-Anollés G, Pedley KF, Bandyopadhyay R and Hartman GL (2011) Genetic structure and diversity of *Phakopsora pachyrhizi* isolates from soybean. *Plant Pathology*:1-11.

Twizeyimana M, Ojiambo PS, Ikotun T, Ladipo JL, Hartman GL and Bandyopadhyay R (2008) Evaluation of soybean germoplasma for resistance to soybean rust (*Phakopsora pachyrhizi*) in Nigeria. *Plant Disease*, 92:947-952.

Twizeyimana M, Ojiambo PS, Sonder K, Ikotun T, Hartman GL and Bandyopadhyay R (2009) Pathogenic variation of *Phakopsora pachyrhizi* infecting soybean in Nigeria. *Phytopathology*, 99:353-361.

Yamaoka Y, Fujiwara Y, Kakishima M, Katsuda K, Yamada K and Hagiwara H (2002) Pathogenic races of *Phakopsora pachyrhizi* on soybean and wild host plants collected in Japan. *Journal of Genetic Plant Pathology*, 68: 52-56.

Yorinori JT, Nunes Junior J and Lazzarotto JJ (2004) Ferrugem “asiática” da soja no Brasil: evolução, importância econômica e controle. Londrina: Embrapa Soja. 36p. (Documentos, 247).

Yorinori JT, Paiva WM, Frederick RD, Costamilan LM, Bertagnolli PF, Hartman GE, Godoy CV and Nunes J (2005) Epidemics of soybean rust (*Phakopsora pachyrhizi*) in Brazil and Paraguay from 2001 to 2003. *Plant Disease*, 89:675-677.

CAPÍTULO 1

**Nucleotide polymorphism in the 5.8S nrDNA gene and internal transcribed
spacers in *Phakopsora pachyrhizi* viewed from structural models**

Abstract

The assessment of nucleotide polymorphisms in environmental samples of obligate pathogens requires DNA amplification through the polymerase chain reaction (PCR) and bacterial cloning of PCR products prior to sequencing. The drawback of this strategy is that it can give rise to false polymorphisms owing to DNA polymerase misincorporation during PCR or bacterial cloning. We investigated patterns of nucleotide polymorphism in the internal transcribed spacer (ITS) region for *Phakopsora pachyrhizi*, an obligate biotrophic fungus that causes the Asian soybean rust. Field-collected samples of *P. pachyrhizi* were obtained from all major soybean production areas worldwide, including Brazil and the United States. Bacterially-cloned, PCR products using a high fidelity DNA polymerase yielded 370 sequences that were subjected to an array of complementary sequence analyses, which included analyses of secondary structure stability, the pattern of nucleotide polymorphisms, GC content, and the presence of conserved motifs. The sequences exhibited features of functional rRNAs. Overall, polymorphisms either took place within less conserved motives, such as loops and bulges, or gave rise to a non-canonical G-U pair within regions of double helix. We discuss the usefulness of structural analyses to filter out putative 'suspicious' bacterially cloned ITS sequences, thus keeping artificially-induced sequence variation to a minimum.

1 - Introduction

The internal transcribed spacer (ITS) region of the 18S–28S nuclear ribosomal DNA (nrDNA), which includes two spacers (ITS1 and ITS2) and the intervening 5.8S gene, constitutes a multicopy gene family. Due to concerted evolution, the tandem nrDNA repeats are more homogeneous within species and more distinct between species (Alvarez and Wendel, 2003). Transcription of the 18S–28S nrDNA cistron produces a single, long transcript from which, upon processing, the ribosomal RNAs are released with the spacers degraded to nucleotides (Coleman, 2009). Even though the spacers are not incorporated into the mature ribosomal RNA (rRNA), they are required to form functional secondary structures that contain binding sites for nuclear proteins required for ribosome biogenesis and correct processing of rRNA (Hillis and Dixon, 1991; Nazar 2004; Mullineux and Hausner, 2009). Sequence alignments of ITS have shown that polymorphic sites are distributed in a non-uniform manner across the ITS region. The 5.8S gene sequences are extremely conserved, the ITS1 and ITS2 sequences are highly polymorphic (Hillis and Dixon, 1991; Coleman, 2007; Nilsson et al., 2008). Given evolutionary constraints, both ITS1 and ITS2 sequences exhibit some conservation at the structural level, even though they can vary significantly at the sequence level (Hausner and Wang, 2005).

ITS sequences can be easily amplified from a variety of organisms with the use of universal primers. Compared to other nuclear genes, ITS sequences exhibit high levels of polymorphism even within species (White *et al.*, 1990). Thus, analysis of ITS sequences has become a key nuclear marker for phylogenetic and phylogeographic studies in eukaryotes (Hillis and Dixon, 1991; Alvarez and Wendel, 2003). Hence, the ITS region is one of the most widely sequenced DNA regions in fungi (Peay *et al.*, 2008).

Given the ease of DNA sequencing, a large database of ITS sequences is available and can be assembled easily at a population level. Such datasets are useful to investigate the distribution patterns of evolutionary constraints, either uniformly or discontinuously, across the ITS region within a given species or closely related species. Our rationale is that, within ITS1 and ITS2, there are motifs (such as double stranded helix and hairpin stems) that are less prone to accumulate changes given that their conformation is crucial for maintaining a functional secondary structure. In contrast, nucleotide interactions in other motifs (such as loops and bulges) may play a less important role in holding the secondary structure

together; as they are less conserved and accumulate higher number of polymorphisms over time. Therefore, an analysis of the predicted RNA secondary structure using ITS sequences may be helpful in distinguishing naturally-occurring polymorphisms from those that arose from experimental error. In phylogenetic or phylogeographic studies, for example, insertions or deletions (indels) that are bordered by mononucleotide repeats (e.g., polyA) or cloned singletons may be rejected as a source of information in statistical analyses. The aim is to reduce the effects of experimental error by removing these sites, which could arise from DNA polymerase misincorporation during PCR or bacterial cloning (Queiroz *et al.*, 2011) or evolutionary lability associated with that type of indel (Mast *et al.*, 2001). Structural analyses may also help to distinguish functional from potentially non-functional nrDNA copies, or pseudogenes (Queiroz *et al.*, 2011).

Phakopsora pachyrhizi (Syd. & P. Syd.) 1914 and *Phakopsora meibomiaae* (Arthur) Arthur 1917 are two closely related species and causing agents of soybean rust. *P. pachyrhizi*, the causal agent of the Asian soybean rust (ASR), is of Asian origin and historically confined to areas of Asia and Australia; whereas *P. meibomiaae* is native to the Americas (Ono *et al.*, 1992). ASR has the potential to cause yield losses up to 10% to 80% (Bromfield, 1984; Pandey *et al.*, 2011). Similar to other obligate biotrophic rust fungi, *P. pachyrhizi* depends on the presence of the host to survive (Cummins and Hiratsuka, 2004). Through asexual reproduction, *P. pachyrhizi* forms dikaryotic uredospores on soybean and on other 42 species in 19 genera of legumes (Bromfield, 1984; Ono *et al.*, 1992). Sexual stages have never been reported for *P. pachyrhizi*, which suggests that the species predominantly undergo asexual reproduction (Ono *et al.*, 1992).

Until the 2000/2001 season, the American continent was apparently free of ASR. *P. pachyrhizi* was reported in Paraguay (2001), Brazil and Argentina (2002), Bolivia (2003) as well as the United States (2004) (Schneider *et al.*, 2005). Worldwide emergence of ASR is a recent event and no geographic region is free of this disease (Freire *et al.*, 2008). Long-distance dispersal of the uredospores is attributed mainly to wind storms (Isard *et al.*, 2005).

Assessment of genetic diversity in obligate pathogens, such as *P. pachyrhizi*, is hampered by the inability to maintain the fungus outside its host. Therefore, for the purpose of sequence analyses, uredospores or other fungal tissue must be sampled directly from host tissues. A dikaryotic structure and an uncharacterized, potentially mixed fungal population from field samples are additional features that further hinder genetic studies in *P. pachyrhizi*. In fact, field samples of *P. pachyrhizi* typically consist of a pool of divergent isolates. In these circumstances, bacterial

cloning of PCR products prior to sequencing is mandatory given that direct sequencing of PCR products frequently yields chromatograms bearing unreadable results (Freire *et al.*, 2008).

In this report, we assembled a large dataset of ITS sequences derived from *P. pachyrhizi* and employed the utility of secondary structure analyses to: 1) examine patterns by which polymorphisms accumulate in ITS1 and ITS2; 2) identify ITS structural features, including conserved motifs and variable regions; 3) investigate the presence of either pseudogenes or putative PCR artifacts due to experimental errors; and 4) discuss the usefulness of secondary structure analyses as a quality control check for ITS sequence data in genetic diversity studies.

2 - Materials and methods

2.1 - Sampling strategy, DNA extraction, and PCR amplification

Samples of uredospores or soybean leaves bearing uredinia were obtained from 60 soybean fields in Brazil during 2005/2006 and 2008/2009 growing seasons. Additionally, a sample of uredospores was obtained from kudzu (*Pueraria lobata*) in Paraguay. Total genomic DNA was extracted as previously described (Freire *et al.*, 2008). For South African samples, infected leaves were collected from naturally infected highly susceptible local cultivars growing at the “Cedara Agricultural Research Station”, near Pietermaritzburg, during the 2005/2006 season. Genomic DNA was extracted at the ARC-Grain Crops Institute, Potchefstroom, following the protocol described below. After extraction, the genomic DNA was lyophilized and sent to our laboratory in Brazil for further analyses.

The entire ITS region (ITS1 - 5.8S gene - ITS2) was amplified using *P. pachyrhizi*-specific primers ITSP5 and ITSP3 and a PCR protocol described previously (Freire *et al.*, 2008). For all PCR amplifications, we used the *Phusion High-Fidelity DNA Polymerase* (New England BioLabs) and GeneAmp PCR System 9700 (Applied Biosystems).

2.2 - Cloning, sequencing, and sequence alignment

PCR products were cloned and sequences from one to four clones per sample were obtained. Cloning was carried out using the *pGEM-T Easy* (Promega) with DH5 α cells of *E. coli*, and plasmid DNA was extracted using a Wizard Minipreps DNA Purification System (Promega) following the manufacturer’s recommendations. Sequencing was performed by Macrogen Inc., USA (www.macrogen.com), using plasmid primer M13 F. A total of 227 sequences were produced and imported into the program Sequencher 4.8 (Gene Codes) for editing. Additionally, the dataset includes 37 sequences retrieved from GenBank (accession numbers EU436700-EU436723, and AF333488-AF333500) and 106 sequences from the Division of Plant Sciences, University of Missouri. Complete alignments of the 370 sequences were performed with the introduction of gaps to compensate for the presence of insertion/deletions (indels). At this point, any flanking regions of the 18S and 28S DNAs were trimmed off.

2.3 - Structural analyses

We defined the limits of the ITS1 and 5.8S gene sequences in *P. pachyrhizi* in accordance with published limits set for *P. pachyrhizi* and other Basidiomycetes in GenBank. The limits of the ITS2 sequences were set in accordance with The ITS2 Database II (Selig *et al.*, 2008). Finally, three datasets were assembled, one for each spacer (ITS1 and ITS2) and the 5.8S gene.

The DnaSP v5 software (Librado and Rozas, 2009) was used to define ribotypes and to estimate nucleotide and ribotype diversities within each of the three datasets. We used sequence alignments in Sequencher to identify unexpected nucleotide substitutions within known conserved motifs in both the 5.8S gene and the ITS2 region. In the 5.8S gene, we inspected the following three 5.8S conserved motifs: M1 (5'-CGAUGAAGAACGUAGC-3') and M3 (5'-UUUGAAYGCA-3') from Harpke and Peterson (2008), and M2 (5'-GAAUUGCAGAAUCC-3') from Jobes and Thien (1997).

The CG content of ribotypes with each of the three datasets was estimated with the online version of the program MFOLD 3.2 – RNAfolding (Mathews *et al.*, 1999; Zuker, 2003), with the default conditions for temperature (37°C) and ionic conditions. MFOLD also estimated the secondary structure and minimum free energy (ΔG) of ribotypes within both the ITS1 and 5.8S gene datasets. The secondary structures of the ITS2 ribotypes and associated minimum free energy (ΔG) were predicted by homology modeling using the secondary structure proposed for the ITS2 of *P. pachyrhizi* (accession 170674463) retrieved from the ITS2 Database II (Selig *et al.*, 2008). Homology modeling tested whether structural features were maintained in the ribotypes of ITS2 identified in this work. Subsequently, secondary structures were drawn and edited with VARNA (Darty *et al.*, 2009). Next, we inspected whether the secondary structures of ITS2 in *P. pachyrhizi* displayed the conserved features expected for an eukaryotic ITS2; that is, the full structure displays four helices, helix II exhibits a pyrimidine-pyrimidine mismatch (U-U), the 5' side of Helix III contains the motif 5'-UGGU-3' or any of its ribotypes (Schultz *et al.*, 2005; Coleman, 2007).

3 - Results and discussion

The three datasets each contained 370 sequences. After sequence alignments and the introduction of gaps to compensate for indels, ITS1 and ITS2 spanned 200 and 215 bases, respectively. The 5.8S gene was 154 bases in length. The concatenated dataset spanned 569 bases in length.

3.1 - Predicted secondary structure of ITS1

ITS1 exhibited six polymorphic sites (four substitutions and two indels), which defined 11 ribotypes (Fig. 1). Eight ribotypes were recovered multiple times, whereas three ribotypes were singletons (ribotypes that appeared only once). Ribotype A, with 261 occurrences, was the most frequent ribotype of ITS1. The ΔG required for formation of the secondary structures ranged from -107.426 to -118.294 KJ/mol, and the CG content of these sequences ranged from 25.5% to 26.5%.

A			B		
Rib	N	GenBank	Rib	N	GenBank
		11111112			
		14445890			
		62393590			
A	261	C--UUU-- EU436701	1	213	AC-----U EU436701
B	10AAU XXXXXXXX*	2	7	C..... XXXXXXXX*
C	26	.UU..AAU XXXXXXXX*	3	34	..UUU....G.. EU436703
D	11	.UU..... XXXXXXXX*	4	70	..UUU....G.A EU436700
E	1	UUU..A.. XXXXXXXX*	5	30A EU436709
F	6AU XXXXXXXX*	6	3	.U..... XXXXXXXX*
G	3	.UU...AU XXXXXXXX*	7	1U. XXXXXXXX*
H	11	.UU..A.. XXXXXXXX*	8	7	..UUU....GUA XXXXXXXX*
I	39A.. EU436700	9	1	..UUU....GU. XXXXXXXX*
J	1	...C.A.. EU436715	10	1G.. XXXXXXXX*
K	1AA.. XXXXXXXX*	11	1	..UUUUUU... XXXXXXXX*
			12	1G.A XXXXXXXX*
			13	1	..UUU.....A XXXXXXXX*

Figure 1. Sequence alignment of the variable sites in ITS1 (A) and ITS2 (B) of *Phakopsora pachyrhizi*. ITS1 and ITS2 span 200 and 215 bases, respectively. Dots indicate similarity to a reference sequence: Ribotype A (ITS1) or Ribotype 1 (ITS2); hyphens indicate gaps. Numbers on top indicate the nucleotide position in the reference sequence during alignment. The numbers of occurrences of the ribotypes (N) are as indicated. *GenBank accession numbers derived from this work.

Structural analyses indicated that none of the six polymorphic sites resulted in predicted changes in the secondary structure. The common secondary structure

of ITS1 consisted of a long, double helix that brought together the 5' and 3' termini (Fig. 2). At its central part, the double helix contained one large, internal loop; as well as, other structural motifs such as asymmetric internal loops of small sizes, bulges, and hairpins. The long double helix of ITS1 could operate as a 'biological spring' (Lalev and Nazar, 1998) and therefore bring to close proximity the termini of the 18S and 5.8S during the processing of the precursor pre-rRNA. Five (out of six) polymorphic sites occurred within loops and one base substitution (a transition C→U, at position 116) took place near the end of the long double helix and gave rise to a wobble G-U base pair. Non-canonical G-U pairing possesses great flexibility and, therefore, can be accommodated within rRNA secondary structures without imposing significant changes (Mullineux and Hausner, 2009). Apparently, indels that have an effect on the helix length or base substitutions that occur at loops do not affect the formation of mature rRNA and, therefore, those regions are prone to changes (Bridge *et al.*, 2008).

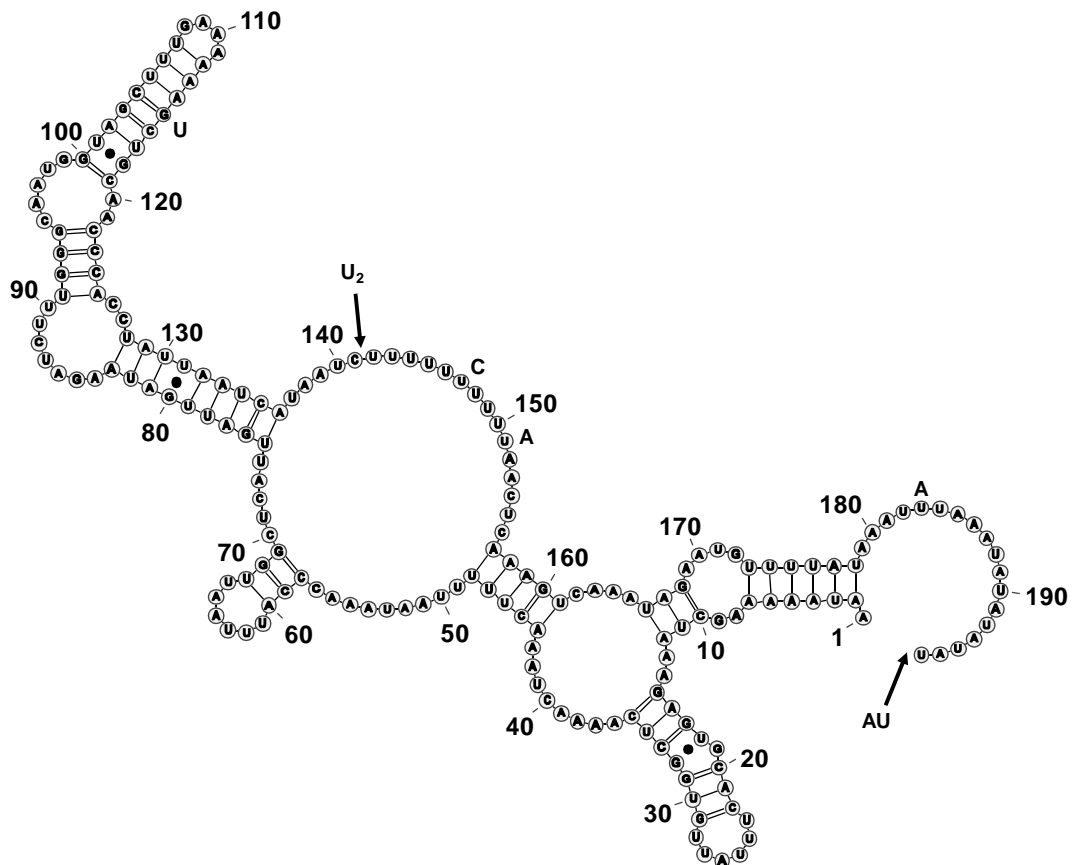


Figure 2. Secondary structure of ITS1 of *Phakopsora pachyrhizi*. The most frequent ribotype (Ribotype A) is used as the reference sequence. Positions are given every 10 nucleotides. Dot, non-canonical G-U pairings; arrow, insertion sites; substitutions as indicated.

3.2 - Predicted secondary structure of the 5.8S gene

The 5.8S gene contained a single polymorphic site, which defined two 5.8S ribotypes. Both ribotypes exhibited the predicted secondary structure for a functional 5.8S gene, with two large central loops from which four helices emerge (Fig. 3). The single polymorphic site consisted of an extra A in a stretch of a polyA that appeared within the loop in Helix II. The ΔG required for formation of the secondary structure of the 5.8S gene was -110.77 KJ/mol, and the CG content was 35.9%.

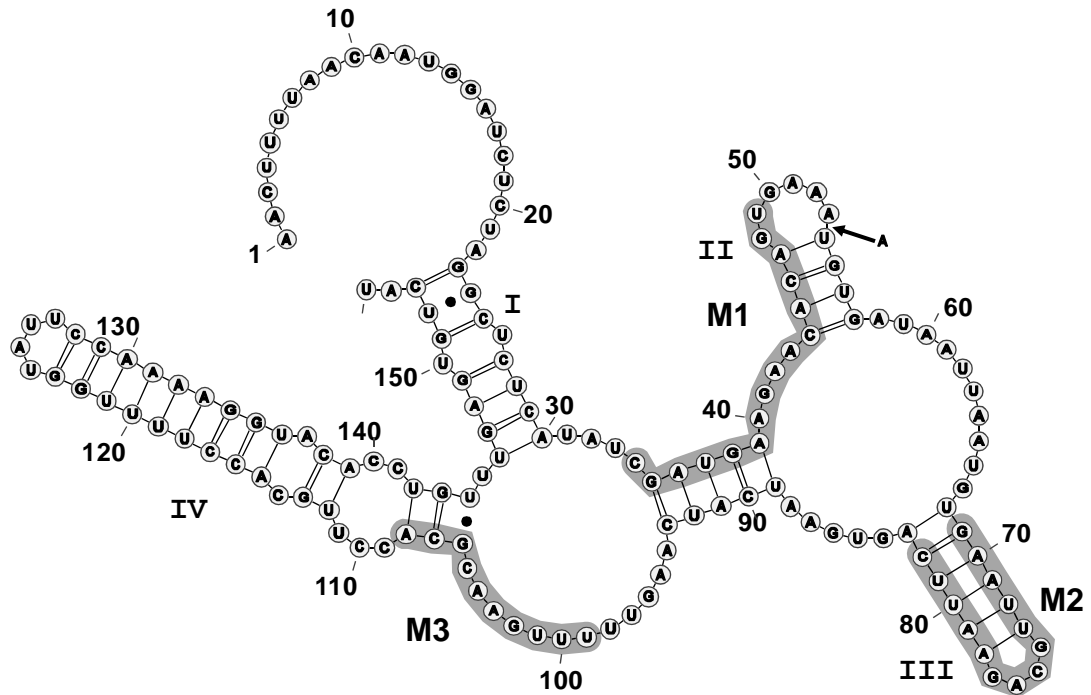


Figure 3. Secondary structure of the 5.8S gene of *Phakopsora pachyrhizi*. The most common ribotype of 5.8S gene is used as the reference sequence. Helices were numbered I to IV; positions are given every 10 nucleotides. Dot, non-canonical G-U pairing; arrow, insertion site; gray shadow, conserved motifs M1, M2 and M3.

Conserved motifs for the 5.8S gene are poorly described in fungi. However, at least three motifs of the 5.8S gene are conserved among angiosperms: M1 (5'-CGAUGAAGAACGUAGC-3') and M3 (5'-UUUGAACGCA-3') (Harpke and Peterson, 2008), and M2 (5'-GAAUUGCAGAAUCC-3') (Jobes and Thien, 1997). In *P. pachyrhizi*, those three motifs were found as (base substitutions are shown in bold): M1 (5'-CGAUGAAGAAC**AC**AGU-3'), M2 (5'-GAAUUGCAGAA**UUC**-3'), and M3 (5'-UUUGAACGCA-3'). At the DNA level, the motif M2 harbors an EcoRI restriction site (underlined), which is highly conserved in fungi and distinguishes between fungal and angiosperm M2 motifs (Jobes and Thien, 1997). Sequences in motifs M1, M2,

and M3 of *P. meibomia* are identical to those of *P. pachyrhizi* (data not shown), which suggest that those motifs play an important biological role in rRNA function.

3.3 - Predicted secondary structure of ITS2

ITS2 contained seven polymorphic sites (three substitutions and four indels), which defined 13 ribotypes (Fig. 1). Seven of those 13 ribotypes were shared, while six were singletons. Ribotype 1, with 213 occurrences, was the most frequent ITS2 ribotype. The ΔG required for formation of the secondary structures ranged from -96.14 to -125.818 KJ/mol, and the CG content of these sequences varied from 18.7% to 19.9%. Homology modeling analyses provided high quality models for the 13 ribotypes and recovered a common secondary structure for the 13 ribotypes, consisting of four helices radiating from a central loop (Fig. 4). This structure of ITS2 is conserved in the vast majority of eukaryotes (Schultz *et al.*, 2005; Coleman, 2007). For every ribotype, percentages of helix transfer were 100% for any of the four helices.

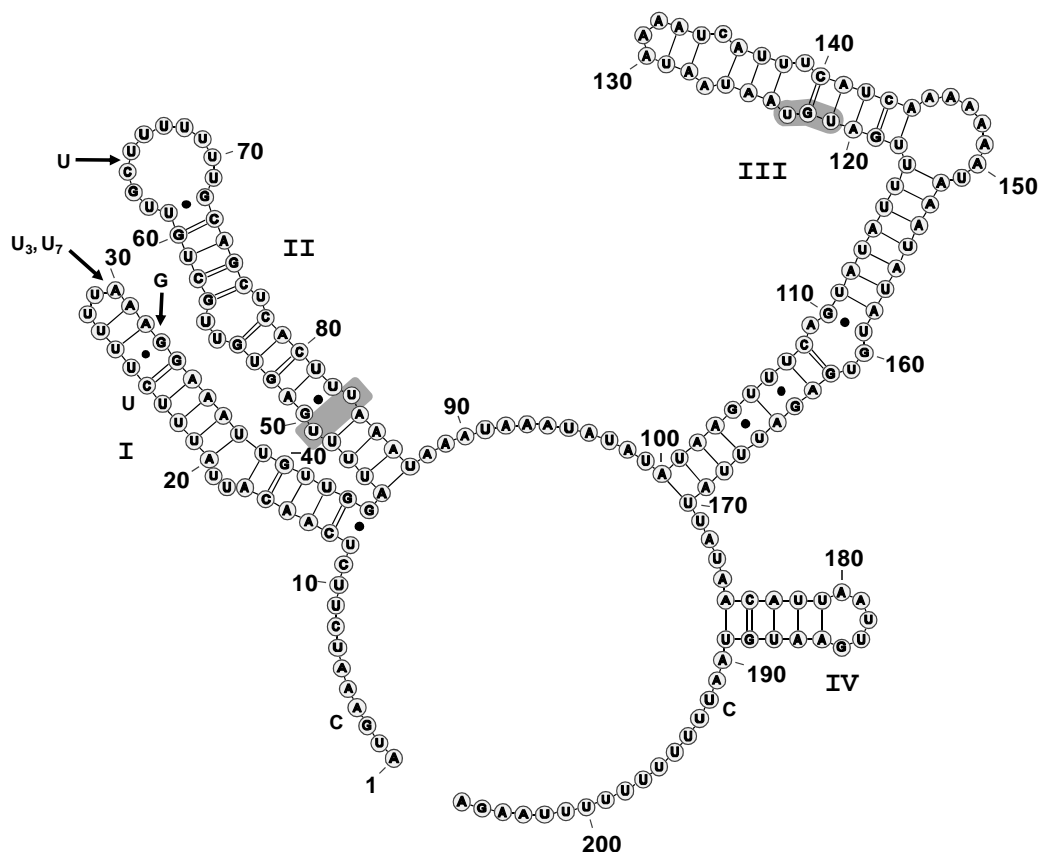


Figure 4. Secondary structure of ITS2 of *Phakopsora pachyrhizi*. Accession 170674463 retrieved from The ITS2 Database II (Selig *et al.*, 2008) is used as the reference sequence. Helices (I to IV) are numbered according to Coleman (2007); positions are given every 10 nucleotides. Dot, non-canonical G-U pairing; arrow, insertion sites; substitutions as indicated. Gray shadow: in helix II, the characteristic pyrimidine–pyrimidine bulge; in Helix III, the conserved sequence 5'-UGU-3'.

Among the four helices of ITS2, Helix I showed the greatest level of variation. It contained one base substitution (that gave rise to a non-canonical G-U pair within the double helix), one 1bp-indel, and one naturally-polymorphic poly-U tract within the loop (size ranging from U₂ to U₉). The evolution of such mononucleotide tracts may arise by such mechanisms as slipped-strand mispairing (Levinson and Gutman, 1987), unequal crossing over, and biased gene conversion (Hillis and Dixon, 1991). All of which may contribute to increases in the level of naturally-occurring ITS ribotypes. The flexibility of the G-U pairing allows secondary structures to remain unchanged (Mullineux and Hausner, 2009). Helix II exhibited a high level of sequence conservation, including two bulges with a pyrimidine-pyrimidine mismatch (U-U) that are highly conserved among eukaryotes (Schultz *et al.*, 2005; Coleman, 2007). In ribotypes 7, 8 and 9, Helix II contained an additional U that extended a polyU tract within its hairpin loop. Helix III was the longest among the four helices of ITS2. In *P. pachyrhizi*, Helix III contained the motif 5'-UGU-3', which may have replaced the conserved motif 5'-UGGU-3' (Schultz *et al.*, 2005). Helix IV displayed a hairpin of only five base pairs and a short terminal loop. Helix IV of *P. pachyrhizi* was similar in size to the Helix IV found in other eukaryotes (Coleman, 2007).

3.4 - Co-evolution of the two spacers

There is evidence that ITS1 and ITS2 do not evolve independently. In fungi, the overall variation in ITS1 and ITS2 was found to be highly correlated (0.87) in about 66% of the fungal species. The observed variability of ITS1 was higher than that of ITS2 (Nilsson *et al.*, 2008). In *P. pachyrhizi*, nucleotide and ribotype diversities for ITS1 and ITS2 were 0.00191 and 0.00223, and 0.4848 and 0.6186, respectively. Therefore, nucleotide diversity and the number of ribotypes were similar (11 ribotypes of ITS1 against 13 ribotypes of ITS2).

3.5 - Usefulness of structural analyses in ITS

Sequencing of bacterially cloned PCR products can recover sequence errors such as introduction of single-base substitutions (Kobayashi *et al.*, 1999). If neglected, those PCR artifacts become 'false' polymorphisms, which can lead to an overestimation of the true levels of genetic variation. The usefulness of ITS as a marker for intraspecific studies in *P. pachyrhizi* is the result of the polymorphisms that accumulate. Our results suggested that we can rely upon structural analyses to

filter out putative 'suspicious' bacterially cloned ITS sequences; thus, keeping artificially-induced sequence variation to a minimum. We attribute the absence of suspicious sequences from our dataset to the extreme fidelity of the *Phusion High-Fidelity DNA Polymerase* (about 50 times greater than *Taq DNA polymerase* – Frey and Suppmann, 1995).

Relying upon structural analyses to explore natural variation is critical given that sequencing of bacterially cloned PCR products is prone to display 'false' polymorphisms from misincorporation during PCR or bacterial cloning. Our structural analyses suggested that the DNA polymorphisms found are likely natural sequence variations in *P. pachyrhizi*, which allowed ITS1 and ITS2 to exhibit some degree of structural changes without compromising functionality.

4 - References

- Alvarez, I., Wendel, J.F., 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Mol Phylo Evo.* 29, 417-434.
- Bridge, P.D., Schlitt, T., Cannon, P.F., Buddie, A.G., Baker, M., Borman, A.M., 2008. Domain II hairpin structure in ITS1 sequences as an aid in differentiating recently evolved animal and plant pathogenic fungi. *Mycopathology.* 166, 1-16.
- Bromfield, K.R., 1984. Soybean rust. Saint Paul: American Phytopathological Society.
- Coleman, A.W., 2007. Pan-eukaryote ITS2 homologies revealed by RNA secondary structure. *Nuc. Acids Res.* 10, 3322-3329.
- Coleman, A.W., 2009. Is there a molecular key to the level of “biological species” in eukaryotes? A DNA guide. *Mol. Phylo. Evol.* 50, 197–203.
- Cummins, G.B., Hiratsuka, Y., 2004. Illustrated Genera of Rust Fungi (APS Press, St. Paul)
- Darty, K., Denise, A., Ponty, Y., 2009. Varna: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics.* 25, 1974-1975.
- Freire, M.C.M., Oliveira, L.O., Almeida, A.M.R., Schuster, I., Moreira, M.A., Liebenberg, M.M., Mienie, C.M.S., 2008. Evolutionary history of *Phakopsora pachyrhizi* (the Asian soybean rust) in Brazil based on nucleotide sequences of the internal transcribed spacer region of the nuclear ribosomal DNA. *Genet Mol Biol.* 31, 920-931.
- Frey, B., Suppmann, B., 1995. Demonstration of the Expand PCR System's greater fidelity and higher yields with a *lacI*-based PCR fidelity assay. *Biochemica.* 2, 34-35.
- Harpke, M.P., Peterse, A., 2008. 5.8S motifs for the identification of pseudogenetic ITS regions. *Botany.* 86, 300-305.
- Hausner, G., Wang, X., 2005. Unusual compact rDNA gene arrangements within some members of the Ascomycota: evidence for molecular co-evolution between ITS1 and ITS2. *Genome.* 48, 648-660.
- Hillis, D.M., Dixon, M.J., 1991. Ribosomal DNA: molecular evolution and phylogenetic inference. *Q Rev Biol.* 66, 411-453.

- Isard, S.A., Gage, S.H., Comtois, P., Russo, J.M., 2005. Principles of the atmospheric pathway for invasive species applied to soybean rust. *BioScience*. 55, 851-861.
- Jobes, D.V., Thien, L.B., 1997. A conserved motif in the 5.8S ribosomal RNA (rRNA) gene is a useful diagnostic marker for plant internal transcribed spacer (ITS) sequence. *Plant Mol Biol Rep*. 15, 326-334.
- Kobayashi, N., Tamura, K., Aotsuka, T., 1999. PCR error and molecular population genetics. *Bioch Genet*. 37, 317-321.
- Lalev, A.I., Nazar, R.N., 1998. Conserved core structure in the internal transcribed spacer 1 of the *Schizosaccharomyces pombe* precursor ribosomal RNA. *J Mol Biol*. 284, 1341-1351.
- Levinson, G., Gutman, G.A., 1987. Slipped-strand mispairing: A major mechanism for DNA sequence evolution. *Mol Biol Evol*. 4, 203-221.
- Librado, P., Rozas, J., 2009. DNAsp V5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. 25, 1451-1452.
- Mast, A.R., Kelso, S., Richards, A.J., Lang, D.J., Feller, D.M.S., Conti, E., 2001. Phylogenetic relationship in *Primula* L. and related genera (Primulaceae) based on noncoding chloroplast DNA. *J. Plan Sci*. 162, 1381-1400.
- Mathews, D.H., Sabina, J., Zuker, M., Turner, D.H., 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structures. *J Mol Evol*. 288, 911- 940.
- Mullineux, T., Hausner, G., 2009. Evolution of rDNA ITS1 and ITS2 sequences and RNA secondary structures within members of the fungal genera *Grosmannia* and *Leptographium*. *Fungal Genet Biol*. 46, 855-867.
- Nazar, R.N., 2004. Ribosomal RNA processing and ribosome biogenesis in eukaryotes. *IUBMB Life*. 56, 457-465.
- Nilsson, R.H., Kritiansson, E., Ryberg, M., Hallenberg, N., Larsson, K.H., 2008. Intraspecific ITS variability in the kingdom Fungi as expressed in the international sequences database and its implications for molecular species identification. *Evol Bioinform*, 4, 193-201.
- Ono, Y., Buritica, P., Hennen, J.F., 1992. Delimitation of *Phakospora*, *Physopella* and *Cerotelium* and their species on Leguminosae. *Myc Res*. 96, 825-850.

- Pandey, A.K., Yang, C., Zhang, C., Graham, M.A., Horstman, H.D., Lee, Y., Zabolina, O.A., Hill, J.H., Pedley, K.F., Whitham, S.A., 2011. Functional analysis of the asian soybean rust resistance pathway mediated by *Rpp2*. *Mol Plan-Micro Interac.* 24, 194-206
- Peay, K.G., Kennedy, P.G., Bruns, T.D., 2008. Fungal community ecology: a hybrid beast with a molecular master. *BioScience.* 58,799-810.
- Queiroz, C.S., Batista, F.R.C., Oliveira, L.O., 2011. Evolution of the 5.8S nrDNA gene and internal transcribed spacers in *Carapichea ipecacuanha* (Rubiaceae) within a phylogeographic context. *Mol Phy Evol.* 59, 293-302.
- Schneider, R.W., Hollier, C.A., Whitman, H.K., Palm, M.E., McKemy, J.M., Hernandez, J.R., Levy, L., DeVries-Paterson, R., 2005. First report of soybean rust caused by *Phakopsora pachyrhizi* in the continental United States. *Plan Dis.* 89, 773-779.
- Schultz, J., Maisel, S., Gerlach, D., Muller, T., Wolf M., 2005. A common core of secondary structure on the internal transcribed spacer 2 (ITS2) throughout the Euckaryota. *RNA.* 11, 361-364.
- Selig, C., Wolf, M., Muller, T., Dandekar, T., Schultz, J., 2008. The ITS2 Database II: homology modeling RNA structure for molecular systematic. *Nucleic Acids Res.* 36, D377-D380.
- White, T.J., Bruns, T., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ and White TJ (eds) *PCR protocols: a guide to methods and applications.* Academic Press, London, pp. 315-322.
- Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31, 3406-3415.

CAPÍTULO 2

Avaliação da variabilidade e da estrutura genética de populações *Phakopsora pachyrhizi* com base em dados de sequências do DNA nuclear

Resumo

A ferrugem asiática da soja, causada pelo fungo *Phakopsora pachyrhizi*, é uma das principais doenças da agricultura mundial. Esse fungo biotrófico infecta as plantações de soja levando ao desfolhamento precoce e redução da produção e do peso final do grão. A principal forma de dispersão desse fungo é a dispersão dos uredósporos pelo vento, que podem infectar plantações próximas, ou podem alcançar longas distâncias. Como resultado dessa ampla dispersão, *P. pachyrhizi* é atualmente identificado em praticamente todas as áreas produtoras de soja, causando severos prejuízos econômicos. Genes de resistência (*Rpp*) à *P. pachyrhizi* vêm sendo constantemente pesquisados na soja e em espécies relacionadas, e alguns genes já foram identificados. No entanto, devido à plasticidade genética de *P. pachyrhizi* esses genes de resistência são rapidamente suplantados. Esse estudo teve como principais objetivos avaliar o nível de variabilidade genética de *P. pachyrhizi*, e entender como essa variabilidade está distribuída entre as principais regiões produtoras de soja em todo o mundo. Foi realizada uma ampla amostragem de *P. pachyrhizi* em dois anos agrícolas brasileiros, em um local da África do Sul, um local no Paraguai, e uma amostra foi proveniente de feijão kudzu também do Paraguai. O DNA genômico foi extraído e três regiões nucleares (fator de ribosilação do ADP, histona H4, e ITS) foram amplificadas e clonadas. Foram sequenciados cerca de quatro clones para cada região gênica para cada amostra. Sequências das três regiões gênicas de *P. pachyrhizi* foram obtidas no Genbank e no banco de dados da Division of Plant Sciences, Life Sciences Center (University of Missouri). Foi criado um banco de dados para cada uma das regiões gênicas, as sequências foram alinhadas e comparadas entre si. O programa DnaSP forneceu informações das características das mutações das sequências, diferenciou as sequências em haplótipos e calculou

a diversidade haplotípica de cada região gênica. Para avaliar a relação genealógica entre os haplótipos o programa TCS foi utilizado para estimar uma rede de haplótipos. Os haplótipos foram plotados nos mapas dos respectivos locais de amostragem. O programa Arlequin foi ainda utilizado para estimar índices de diversidade e para avaliar como as populações de *P. pachyrhizi* estão estruturadas. As análises dos três bancos de dados das diferentes regiões gênicas nucleares forneceram informações complementares e compatíveis. Foi identificada uma alta diversidade de haplótipos para as três regiões gênicas, revelando a alta variabilidade genética de *P. pachyrhizi*. Grande parte dessa variabilidade genética (>90%) se concentra dentro das populações, indicando que populações locais são altamente diversas, e que as populações de diferentes países e continentes são uniformes. Esses resultados podem ser explicados pela alta taxa de mutação do fungo e pela alta capacidade de dispersão dos esporos. Os dados ainda revelaram que o vazio sanitário, adotado no Brasil, não promoveu o estrangulamento genético de populações de *P. pachyrhizi* entre os dois anos agrícolas avaliados, e que os haplótipos de *P. pachyrhizi* identificados no feijão kudzu são os mesmos identificados na soja, revelando a importância dessa cultura como hospedeiro alternativo. A análise da distribuição geográfica dos haplótipos de *P. pachyrhizi* nas principais áreas produtoras de soja possibilitou a inferência das possíveis rotas de dispersão do fungo. Os dados obtidos no presente trabalho reforçam a hipótese de dispersão a longa distância de esporos de *P. pachyrhizi*. Esse fungo identificado pela primeira vez nos países orientais se dispersou alcançando a Austrália e mais tarde o continente africano. Uredósporos de populações de *P. pachyrhizi* dispersaram da África alcançando a América do Sul, inclusive o Brasil, onde esse fungo foi amplamente disperso. O nível de compartilhamento de haplótipos dos Estados Unidos com os demais países e continentes revelou a possibilidade dos Estados Unidos ter sofrido introdução de mais de um local (Brasil, Havaí e África), sendo a rota pelo Havaí a mais evidenciada.

1 - Introdução

No gênero *Phakopsora* duas espécies (*P. pachyrhizi* e *P. meibomiae*) infectam a soja causando a ferrugem. Apesar dessas duas espécies serem proximamente relacionadas, *P. pachyrhizi* é mais agressiva (Hoefle *et al.*, 2009). Essa doença, também conhecida como ferrugem asiática da soja, tem potencial para causar sérios prejuízos econômicos na produção da soja (Sinclair e Hartman, 1996). Atualmente, a ferrugem da soja é um dos mais importantes problemas econômicos para a produção da soja (Goellner *et al.*, 2010), podendo causar perdas de 10% a 80% na produção (Bromfield, 1984; Pandey *et al.*, 2010). Plantas severamente infectadas mostram desfolhação precoce, que reduz a formação das vagens e peso final do grão, conseqüentemente reduz o valor e a qualidade do produto (Mendes *et al.*, 2009).

P. pachyrhizi é um patógeno obrigatório, propagado principalmente pela dispersão aérea dos uredósporos, que é tipicamente seu único tipo de esporo infectivo. Sua forma de penetração se diferencia da observada nas outras ferrugens. A penetração é de forma direta através da cutícula utilizando enzimas de digestão e força física (Luster *et al.*, 2010). A taxa de progressão da doença é dependente da temperatura e condições de umidade favoráveis (Luster *et al.*, 2010). Outra característica que diferencia a ferrugem asiática da soja de outras ferrugens é sua capacidade de infectar uma ampla variedade de hospedeiros (Luster *et al.*, 2010). Esse fungo é capaz de esporular em 95 espécies de mais de 42 gêneros de plantas (Bromfield, 1984; Ono *et al.*, 1992).

O patógeno dispersou do seu ponto origem a Ásia para Austrália, Havaí, África, América do Sul e, mais recentemente, para os Estados Unidos. O único método de controle utilizado da ferrugem asiática é a aplicação de fungicidas, o que aumenta substancialmente os custos da produção e é uma contínua fonte de contaminação ambiental (Silva *et al.*, 2008). Apesar do uso de fungicidas ser eficiente no controle da ferrugem asiática, o desenvolvimento de cultivares resistentes de soja pode fornecer um método durável e sustentável de proteção.

A resistência da soja à ferrugem foi registrada pela primeira vez na década de 60 em Taiwan. Desde então, a resistência a *P. pachyrhizi* tem sido investigada na soja e em espécies relacionadas. Sete genes conferindo resistência já foram identificados. No entanto, devido à variabilidade genética de populações de *P. pachyrhizi*, cinco desses sete genes já tiveram sua resistência suplantada (Soria-Guerra *et al.*, 2010).

Pesquisas da ferrugem asiática da soja têm como principais objetivos os estudos do hospedeiro, na epidemiologia e na avaliação na perda da produção e meios de controle. No entanto, como ainda não existem cultivares de soja com resistência para todos os isolados de *P. pachyrhizi* e é questionável a possibilidade de se obter esse cultivar, pesquisas mais básicas são também necessárias em relação ao patógeno (Goellner *et al.*, 2010). Pouco é conhecido sobre a interação soja-ferrugem em nível molecular (Luster *et al.*, 2010). Detalhes sobre a epidemiologia e a variabilidade genética do fungo também devem ser pesquisados (Yorinori e Lazzarotto, 2004). De acordo com Twizeyimana *et al.* (2011), uma vez que a durabilidade da resistência da planta hospedeira é um processo evolutivo, o desenvolvimento e o melhoramento da resistência durável pode ser mais efetivo no controle da ferrugem da soja se a estrutura genética da população de *P. pachyrhizi* for conhecida.

2 – Objetivos

Esse estudo teve como principal objetivo avaliar o nível de variabilidade genética de *P. pachyrhizi*. Para essa avaliação foram feitas amostragens do fungo *P. pachyrhizi* em diferentes países e continentes envolvendo os principais locais de produção de soja. A variabilidade genética das populações de *P. pachyrhizi* foi avaliada a partir de dados moleculares. Foi utilizado o sequenciamento e a comparação de diferentes regiões gênicas nucleares para diferenciar os organismos. Essas análises forneceram informações para responder às seguintes questões:

- a) Qual o nível de variabilidade genética de *P. pachyrhizi* e como essa variabilidade se distribui entre os diferentes países e continentes? A maior parte da variabilidade genética estaria concentrada em determinados locais ou estaria amplamente distribuída?
- b) Quais são os possíveis fatores envolvidos na estruturação das populações de *P. pachyrhizi*, e quais os efeitos de cada fator?
- c) O vazio sanitário, adotado no Brasil desde 2006, além de eficiente na redução de inóculo inicial para cada ano agrícola, é também eficiente na redução da variabilidade genética de *P. pachyrhizi* entre os diferentes anos agrícolas brasileiros?
- d) Populações de *P. pachyrhizi* identificadas no feijão kudzu, são as mesmas identificadas na soja? E existe algum tipo de diferenciação do fungo para a infecção dos diferentes hospedeiros?
- e) Quais são os centros de origem das populações de *P. pachyrhizi* dos dois maiores produtores de soja, Brasil e Estados Unidos? Com as informações obtidas, é possível explicar a rota de dispersão de *P. pachyrhizi*?

3 - Metodologia

3.1 - Coleta de amostras

Amostras de folhas de soja infectadas com *P. pachyrhizi* foram coletadas em campos de soja no Brasil, em dois anos agrícolas 2005-2006 (29 amostras) e 2008-2009 (28 amostras). No continente africano, amostras foram coletadas na África do Sul em colaboração com o ARC-Grain Crops Institute, Potchefstroom. No Paraguai, uma amostra de soja e uma amostra de kudzu infectados foram coletadas com o apoio da Embrapa-Soja. Para as amostras de origem brasileira, as folhas contaminadas foram embaladas, devidamente identificadas e enviadas ao Laboratório de Biologia Molecular e Filogeografia do Instituto de Biotecnologia Aplicada à Agropecuária (BIOAGRO) da Universidade Federal de Viçosa (UFV), Viçosa-MG. Após o recebimento das amostras, os esporos das folhas mais contaminadas foram removidos com auxílio de pincel. Das amostras que apresentaram pouca quantidade de esporos, pedaços das folhas com maior intensidade de infecção foram cortados para serem utilizadas. Os esporos e os pedaços de folhas com esporos foram colocados em microtubos, liofilizados e armazenados a -80°C. Para as amostras provenientes de outros países, o DNA foi previamente extraído por laboratórios colaboradores, seguindo metodologia pré-estabelecida e, em seguida, enviada ao BIOAGRO.

3.2 - Extração de DNA

O DNA genômico total foi extraído de aproximadamente 30mg de esporos e/ou folhas contaminadas, usando o protocolo de CTAB descrito por Doyle e Doyle (1987), com modificações, como descrito por Freire *et al.* (2008). As extrações de DNA foram feitas no Laboratório de Biologia Molecular e Filogeografia da UFRV.

O material foi macerado em microtubos de capacidade de 1,5mL na presença de 600µL de tampão de extração CTAB: NaCl; Tris -HCl 1M; EDTA 0,5M; CTAB 5%; β-mercaptoetanol. O material foi homogeneizado e incubado a 65°C durante 15 minutos. Foram adicionados 400 µL de clorofórmio álcool-isoamílico 24:1, em seguida os tubos foram centrifugados a 13.000rpm por 10 minutos, para separar a fase orgânica da aquosa. Após a centrifugação, o sobrenadante foi transferido para novos microtubos, onde foram adicionados 500 µL de isopropanol gelado. Os tubos foram agitados e centrifugados para sedimentar o "pellet". O

precipitado foi lavado com etanol 70% e em seguida colocado para secar em câmara de fluxo laminar. O material foi então ressuspenso em TE (Tris-HCl 10mM; EDTA 1mM, pH 8,0) e RNase 40 µg/mL, e finalmente armazenado a -20°C.

Foi extraído DNA de 57 amostras brasileiras, sendo 29 referentes ao ano agrícola 2005-2006, e 28 do ano agrícola 2008-2009. As amostras de DNA foram enviadas para o Life Science Center da University of Missouri, Columbia-MO, onde foram realizadas as amplificações e clonagens. Para a exportação e importação de amostras de DNA de *P. pachyrhizi*, foi obtida autorização do Ministério da Agricultura (ANVISA). Esse documento foi necessário para garantir a legalidade da saída e da entrada das amostras de DNA no país, comprovando que essas amostras estariam sendo exportadas com a única finalidade de pesquisa. Documento para a permissão da entrada dessas amostras nos Estados Unidos também foi concedido pela APHIS -USDA (Animal and Plant Health Inspection Service).

3.3 - Reações de amplificação via PCR

Foram realizados testes de amplificação de regiões do genoma nuclear e do genoma mitocondrial, para algumas amostras. Para a seleção dos genes a serem amplificados e analisados para todas as amostras, foram avaliados os genes mais viáveis a amplificação, e de maior nível de polimorfismo, entre as amostras selecionadas. Três regiões do genoma nuclear foram selecionadas: (1) o gene que codifica para o fator de ribosilação do ADP; (2) o gene que codifica para a histona H4 (HIS); (3) e as regiões espaçadoras internas (ITS) mais a região de 5.8S do DNA ribossomal nuclear. As amplificações foram conduzidas em termociclador GeneAmp PCR System 9700 (Applied Biosystems).

A região espaçadora interna (ITS) foi amplificada com os primers específicos para *P. pachyrhizi*, previamente desenhados, ITSP5 e ITSP3 (Freire *et al.*, 2008). Testes de amplificação foram realizados com os primers universais de ITS, ITS4 e ITS5 de White *et al.* (1990), no entanto, esses não apresentaram boa repetibilidade. As amplificações do gene histona H4 e fator de ribosilação do ADP foram conduzidas com primers obtidos na Division of Plant Sciences, Life Sciences Center (University of Missouri). A combinação de primers DN739857F2 e DN739857R5b foram utilizados para amplificar os genes de fator de ribosilação ADP, e os primers HisH4F3EcoR e HisH4R3Pst para amplificar os genes de histona H4 (Tabela 1).

As amplificações foram conduzidas com o mesmo programa para as três regiões gênicas. As amplificações foram realizadas com um ciclo inicial de desnaturação a 94°C por cinco minutos; seguidos por 35 ciclos de 94°C por 30 segundos, 58°C por 30 segundos para anelamento dos primers, 72°C por 45 segundos para extensão; e um ciclo de extensão final de 72°C por cinco minutos. As condições de amplificação foram praticamente às mesmas para os três genes, para um volume final de 20µl de reação: 4µl do tampão 5x Phusion HF Buffer (1,5 mM de MgCl₂); 1µl de cada primer a 10µM; 0,2µl de dNTPs a 10mM; 2,0µl de DNA; 0,2µl de *Phusion High-Fidelity DNA Polymerase*; e água mili-Q autoclavada para completar o volume. Para a amplificação da região de ITS foi adicionado DMSO à reação.

Após cada amplificação, foram adicionados 0,2µl da enzima *Taq* polimerase em cada reação e essas foram incubadas a 37°C por 30 minutos. A adição de *Taq* teve por finalidade a adição de adenina no final de cada fita amplificada, com o intuito de facilitar a reação de ligação no plasmídeo pGEM-T Easy (PROMEGA). Os fragmentos amplificados foram separados por meio de eletroforese em gel de agarose (0,8%), visualizados por meio de luz ultravioleta, e suas imagens digitalizadas em um sistema captura de imagens.

3.4 - Clonagem dos produtos amplificados

Para a clonagem dos produtos de amplificação, foi feita a reação de ligação utilizando o vetor pGEM-T Easy (Promega), de acordo com as recomendações do fabricante. Em seguida, foi realizada a transformação por choque térmico, utilizando células DH5α ultracompetentes de *Escherichia coli*. Em 100µl de células competentes foram adicionados 3µl de reação de ligação, e mantidos no gelo por 30 minutos. Após os 30 minutos, os tubos foram incubados em banho-maria a 42°C durante dois minutos e novamente colocados em gelo por dois minutos. Para a recuperação das células, foram adicionados 900µl de meio SOC (0,5% de extrato de levedura, 2% de triptona, NaCl 10mM, KCl 2,5mM, MgCl₂ 10mM, MgSO₄ 10mM) e esses foram incubados em agitador a 37°C/150rpm durante uma hora. Transcorrido o tempo de crescimento, os microtubos foram centrifugados a 10.000rpm por dois minutos (Sambrook *et al.*, 1989). O precipitado foi ressuspenso em 150µl de meio SOC, e desses 100µl foram plaqueados em meio LB sólido contendo ampicilina 50µg/ml e X-Gal 0,025mg/ml. As placas foram incubadas a 37°C por cerca de 20 horas.

Para as análises dos transformantes, colônias brancas isoladas foram repicadas em meio sólido e analisadas por PCR de colônia, para avaliar quais colônias apresentavam o inserto de interesse. As reações foram preparadas em um volume final de 30µl: 6,0µl do tampão 5x Green GoTaq Reaction Buffer; 1,5µl de cada primer a 10µM; 0,2µl de dNTPS a 10mM; 0,2µl de *Taq* DNA Polimerase; e água mili-Q autoclavada para completar o volume. As reações de PCR de colônia foram conduzidas com primers específicos. Os primers foram desenhados com o auxílio do “Primer3 Input Program” (<http://frodo.wi.mit.edu/primer3/>), com base nas sequências de *P. pachyrhizi*. O ITS foi amplificado com os primers ITSP5 e ITS6, o gene de histona H4 com os primers HIS2F e HIS2R, e o gene de fator de ribosilação do ADP com os primers ADP3F e ADP2R. As sequências de cada primer, assim como a temperatura de anelamento, encontram-se resumidas na Tabela 1. As reações foram conduzidas nas mesmas condições da primeira amplificação. Em seguida, os fragmentos foram visualizados por meio de luz ultravioleta e suas imagens digitalizadas em um sistema de captura de imagens.

Tabela 1: Definição dos primers utilizados para a amplificação das três regiões gênicas de *P. pachyrhizi*.

	Primer	Sequencia	Tm
ITS	ITSP5	5' GCAACGGCACTTTACTGGCTC 3'	59,1°C
	ITSP3	5' GTTCAGTGGGTAGTCTCACCTGATT 3'	58°C
	ITS6	5' TACCAAAAGGTGCAAGGTGCG 3'	58,4°C
ADP	DN739857F2	5' AGCAGACCCAACCCATCGATAT 3'	58,2°C
	DN73985R5b	5' AGTCCACCACAAAGATGATTCCCT 3'	58,2°C
	ADP3F	5' GATATCTCCCTCTCCCTCTC 3'	52,8°C
	ADP2R	5' ATGATCCCTGAGTGTTTTG 3'	50,4°C
HIS	HisH4F3EcoR	5' CTTCTGTGAATCCGGAGAACGTTA 3'	60,3°C
	HisH4R3Pst	5' TCAGAACTGCAGTATCTCCTATCCTA 3'	59,3°C
	HIS2F	5' ATCAAAAGATGTCTGGACGTG 3'	52,8°C
	HIS2R	5' TTCCGGAGAACGTTATTTGT 3'	51,8°C

Os produtos de PCR de colônia foram digeridos com enzimas de restrição específicas para cada região gênica. A digestão com enzimas permitiu a diferenciação de amplificações de *P. pachyrhizi* e de contaminantes. Para o ITS e Histona H4 foi utilizada a enzima Bgl II. A enzima Sac I foi utilizada para digerir o fator de ribosilação do ADP. Os produtos gerados após a digestão com as enzimas de restrição foram separados por meio de eletroforese em gel de agarose (1,5%), visualizados por meio de luz ultravioleta e suas imagens digitalizadas em um sistema de captura de imagens. As colônias que apresentaram as melhores amplificações foram selecionadas para fazer a extração de DNA plasmidial. As

colônias selecionadas foram inoculadas em tubos de ensaio contendo 5ml de meio LB líquido com 2,5µl de ampicilina 50µg/ml. Os tubos contendo meio LB e bactérias foram incubados por cerca de 14 horas a 37°C sob agitação de 180rpm. A extração de DNA plasmidial com o Kit Wizard Mini Prep (Promega), de acordo com as recomendações do fabricante. A quantificação do DNA foi feita com o auxílio do Nano-Drop (Thermo-Scientific).

A extração de DNA plasmidial não funcionou bem para algumas amostras dos genes de Histona H4 e do fator de ribosilação do ADP. Para essas amostras, o produto de PCR de colônia foi purificado (para a remoção de primers e dNTPs não incorporados na reação) e enviado para o sequenciamento. Para a purificação do produto de PCR de colônia foram utilizadas as enzimas Exonuclease I (Exo) e Alcalina Fosfatase (SAP). Foi adicionado 2,0µl de SAP e 0,2µl de Exo em 10,0µl de produto de PCR, esses foram incubados em 37°C por 60 minutos para a ação das enzimas e logo em seguida a 75°C por 15 minutos para a inativação das enzimas.

3.5 - Sequenciamento e comparação das sequências

Os sequenciamentos foram realizados pela equipe da MacroGen-USA. Sequenciamento do DNA plasmidial foi conduzido com o primer universal M13pUCF. Sequenciamento dos produtos de PCR de colônia purificados foram realizados com primers específicos de cada região gênica (utilizados no PCR de colônia). Para o ITS foi sequenciada a região completa (18S, 5.8S, 28S). Para os genes de histona H4 e fator de ribosilação do ADP foram sequenciadas apenas parte dos genes, envolvendo regiões de introns e exons.

As sequências foram importadas para o aplicativo computacional SEQUENCHER versão 4.1.4. Sequências de DNA de *P. pachyrhizi*, referentes às três regiões amplificadas, foram obtidas no banco de dados da Division of Plant Sciences, Life Sciences Center (University of Missouri), e no Genbank (números de acesso: EU436700-EU436723 e AF333488-AF333500). Foram montados três bancos de dados, um para cada região gênica analisada. As sequências foram alinhadas e submetidas a correções e edições manuais. Os primers foram também alinhados às regiões gênicas correspondentes e utilizados como referência para a eliminação das extremidades das sequências. No alinhamento foi necessária a introdução de “gaps” para compensar as inserções ou deleções de bases. A comparação das sequências, e a análise do cromatograma de cada uma delas permitiu a realização das correções necessárias e a identificação das mutações.

3.6 - Análises estatísticas

Foram preparados três bancos de dados, um para cada região gênica, e esses foram analisados no programa DNAsp (Librado e Rozas, 2009). Esse programa forneceu dados básicos de cada alinhamento, como: número de sequências analisadas, tamanho das sequências, número de gaps e a posição deles na sequência, número de sítios polimórficos e posição na sequência, número de sítios polimórficos informativos, números de singltons, conteúdo GC, número de haplótipos, diversidade nucleotídica e diversidade haplotípica. Ainda no programa DNAsp foi possível avaliar o nível de fluxo gênico e de diferenciação entre as populações de *P. pachyrhizi* provenientes de diferentes países e/ou continentes, ao calcular o F_{ST} . Para análises a nível populacional foram feitos agrupamentos referentes ao país de origem de cada amostra.

Os dados haplotípicos e os dados de agrupamento obtidos pelo programa DNAsp foram exportados para o programa ARLEQUIN versão 3.1 (Excoffier, 2006), como qual foram realizados vários testes a nível intra e interpopulacional. Foram calculados índices da diversidade molecular; testes da estrutura genética da população: análise da variância molecular (AMOVA); e análise de diferenciação das populações.

O programa TCS (versão 1.13), desenvolvido por Clement *et al.* (2000), foi utilizado para evidenciar a relação genealógica entre os haplótipos obtidos. Foi estimada uma rede de haplótipos para cada uma das regiões gênicas. Esse programa aplica a estatística da parcimônia pelo algoritmo desenvolvido por Templeton *et al.* (1992). Para estimar a rede de haplótipos, cada uma das bases (A, G, C e T) foi considerada como um caráter, e as inserções ou deleções (indels) foram consideradas como um quinto caráter, independente do seu tamanho. O número máximo de conexões parcimoniosas entre as sequências foi estimado com 95% de probabilidade para todos os acoplamentos dos haplótipos dentro da rede.

Os haplótipos foram plotados em mapas das localidades amostradas e comparados com a rede de haplótipos obtida. Esse processo foi realizado separadamente para cada uma das três regiões gênicas estudadas. Os dados obtidos para cada uma das regiões gênicas foram comparados com o intuito de se avaliar o nível de confiabilidade de cada um dos dados, e a partir daí tentar se aproximar ao máximo: do real nível de variabilidade genética; da distribuição dessa variabilidade; e da história evolutiva e de migração desse fungo.

4 - Resultados

4.1 - Fator de ribosilação do ADP

4.1.1 - Comparação das sequências

Para o gene do fator de ribosilação do ADP foram sequenciados 154 clones provenientes de amostras do Brasil, 67 referentes ao ano agrícola 2005-2006 e 87 do ano agrícola 2008-2009. Foram ainda sequenciados 14 clones provenientes de amostras da África do Sul, nove clones do Paraguai, e uma amostra proveniente de feijão kudzu infectado com *P. pachyrhizi*, também do Paraguai. Foram incorporadas ao estudo 69 sequências do banco de dados da Division of Plant Sciences, Life Sciences Center (University of Missouri). Foram analisadas 247 sequências do fator de ribosilação do ADP.

As 247 sequências do fator de ribosilação do ADP foram alinhadas. O tamanho da sequência consenso foi de 728 bases. Análises das sequências permitiram a identificação de três indels (um de seis bases e dois de cinco bases cada), e cinco substituições de bases, todas parsimoniosamente informativas (Figura 1). O conteúdo GC das sequências do fator de ribosilação do ADP foi de 41,4%.

As oito mutações do fator de ribosilação do ADP permitiram a diferenciação de 19 haplótipos, com diferentes frequências, resultando em uma diversidade haplotípica de 0,8170, a qual corresponde à diversidade gênica média entre todos os locais amostrados.

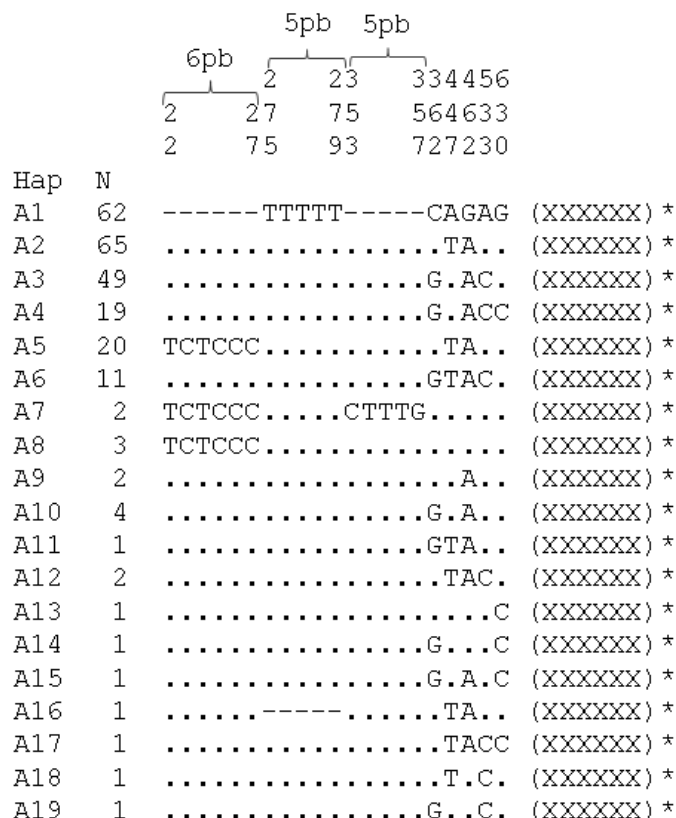


Figura 1: Alinhamento dos sítios variáveis no fator de ribosilação do ADP definindo os 19 haplótipos de *P. pachyrhizi*. Pontos indicam similaridades com o haplótipo A1 (tomado como sequência de referência) e hífens indicam gaps. Números ao alto indicam a posição do nucleotídeo na sequência de referência durante o alinhamento.*Números de acesso no GenBank derivados do presente estudo.

4.1.2 - Distribuição das sequências

As 247 sequências do fator de ribosilação do ADP foram representadas na Tabela 2, nos respectivos locais de amostragem. Os haplótipos A1 e A2 foram os haplótipos mais frequentes, os quais foram identificados em 62 e 65 sequências, respectivamente. Ambos haplótipos foram os mais frequentes no Brasil (nos dois anos agrícolas) e nos outros países analisados, com a exceção dos Estados Unidos. Os haplótipos A3, A4 e A5 (49, 19 e 20 sequências, respectivamente) apresentaram distribuição mais restrita em determinados locais. Os haplótipos A3 e A4 foram praticamente restritos ao Brasil. O haplótipo A5, haplótipo mais frequente nos Estados Unidos (20 sequências), foi restrito a esse país. Entre os haplótipos com menor frequência, três haplótipos (A6, A9 e A10) foram compartilhados com diferentes países, e 11 haplótipos (A7, A8, A11-A19) foram específicos de um determinado país. Desses 11 haplótipos restritos, nove foram singletons (identificado uma única vez).

Foi calculada a diversidade haplotípica para cada local. Para esse cálculo foi avaliado o número de sequências analisadas em cada país, o número de haplótipos identificados e a frequência de cada haplótipo. A Tabela 3 sumariza os dados de diversidade haplotípica para o fator de ribosilação do ADP. Entre as 67 sequências do primeiro ano agrícola brasileiro (2005-2006) foi possível identificar oito haplótipos, resultando em uma diversidade haplotípica 0,75079. Para o segundo ano agrícola brasileiro (2008-2009) foram identificados 12 haplótipos entre as 87 sequências analisadas, resultando em uma diversidade haplotípica de 0,81743. Altos valores de diversidade haplotípica foram identificados em todos os locais. O maior valor de diversidade haplotípica foi identificado no Brasil (0,81743), e o menor foi identificado nos Estados Unidos (0,62731).

Tabela 3: Número de sequências do fator de ribosilação do ADP para cada localidade, com os respectivos números de haplótipos diferenciados e a valor da diversidade haplotípica estimada.

	Sequências	Haplótipos	Hd
Brasil 2006	67	8	0,75079
Brasil 2009	87	12	0,81743
África	21	6	0,72857
América do Sul	14	5	0,75824
Estados Unidos	42	4	0,62731
Outros	12	4	0,71212

4.1.3 – Relação genealógica entre os haplótipos

Uma rede de haplótipos foi estimada para entender quais as relações genealógicas entre os haplótipos do fator de ribosilação do ADP de *P. pachyrhizi*. O valor máximo estimado de conexões parcimoniosas para o fator de ribosilação do ADP foi de 11 eventos de mutação. Uma única rede de haplótipos foi obtida abrangendo os 19 haplótipos. A rede de haplótipos do fator de ribosilação do ADP foi representada na figura 2. A rede de haplótipos obtida para a região do fator de ribosilação do ADP apresentou cinco “loops” (haplótipos com mais de uma origem possível).

Três haplótipos (A1, A2 e A3) ocorreram em maior frequência. Esses haplótipos ocuparam pontos centrais da rede e apresentaram multifurcações para haplótipos de menor frequência. Essas características são indícios da ancestralidade destes três haplótipos. Os haplótipos A4 e A6 se diferenciam do haplótipo A3 por um único passo

de mutação cada. Os haplótipos A9 e A10 ocuparam posição central na rede, sendo o haplótipo A10 mais proximamente relacionado ao haplótipo A3, e o haplótipos A9 aos haplótipos A1 e A2. Nas extremidades da rede foi identificada a presença de quatro haplótipos singletons (A16, A17, A18 e A19), e dos haplótipos A5 e A7. A posição ocupada por esses haplótipos servem como indicativo que esses são derivados dos haplótipos que ocupam posições centrais na rede.

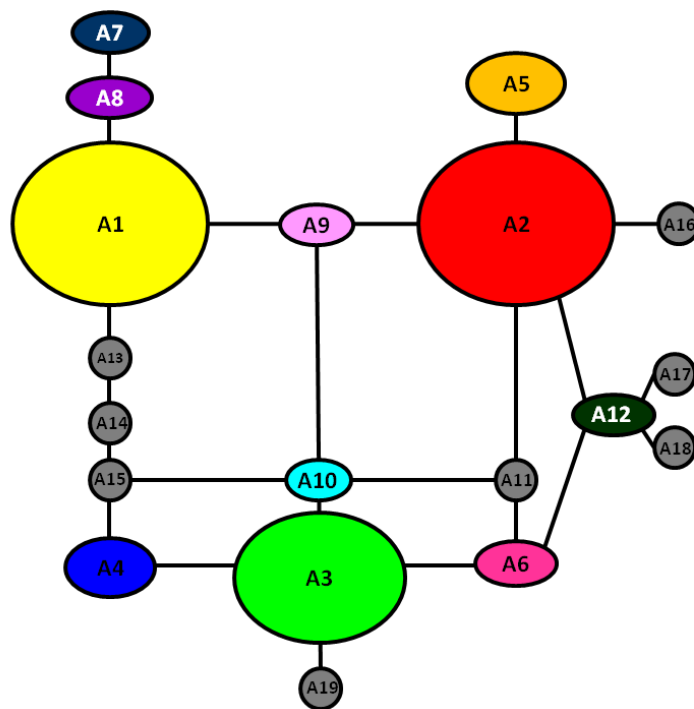


Figura 2: Rede de haplótipos para os haplótipos do fator de ribosilação do ADP de *P. pachyrhizi*. Cada círculo representa um haplótipo; número dentro do círculo indica o haplótipo; o tamanho dos círculos tem relação com a frequência dos haplótipos; círculos coloridos são haplótipos frequentes ($N > 2$); círculos cinza são singletons ($N = 1$). Cada linha representa um passo de mutação.

4.1.4 - Distribuição geográfica dos haplótipos

Os 19 haplótipos identificados para o fator de ribosilação do ADP de *P. pachyrhizi* foram plotados em mapas referentes aos locais de amostragem, para melhor visualização da distribuição dos haplótipos (Figura 3).

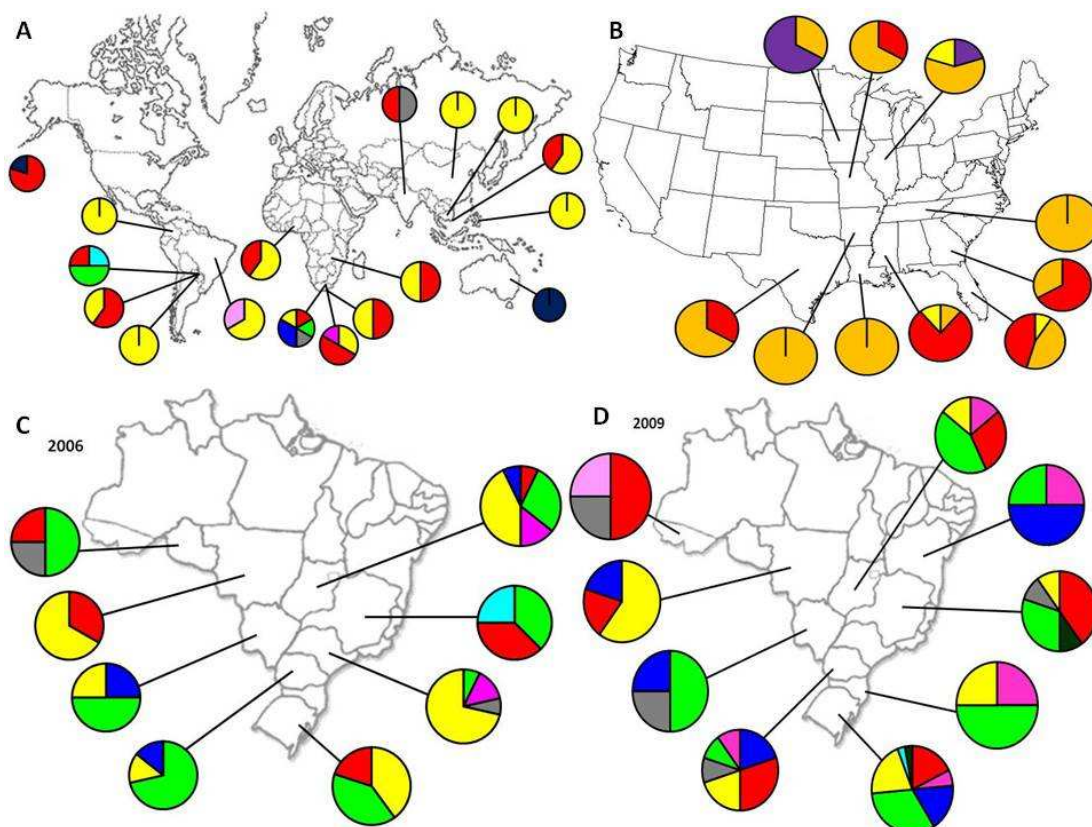


Figura 3: Distribuição geográfica dos 19 haplótipos do fator de ribosilação do ADP de *P. pachyrhizi*. (A) Ásia, Austrália, África, América do Sul e Havaí. (B) Estados Unidos. (C) Brasil, ano agrícola 2005-2006. (D) Brasil, ano agrícola 2008-2009. Círculos representam o conjunto de haplótipos identificados em cada local, e as divisões de cada círculo são proporcionais a frequência de cada haplótipo identificado.

Nos países orientais (China, Índia, Filipinas, Tailândia e Vietnã) houve uma grande prevalência dos haplótipos A1 e A2, reforçando a hipótese de ancestralidade desses haplótipos. Na Austrália foi identificado o haplótipo A7, o qual foi identificado apenas nesse país e no Havaí. No Havaí além do haplótipo A7 foi também identificado o haplótipo A2.

No continente africano também houve prevalência dos haplótipos A1 e A2, haplótipos identificados na Nigéria e no Zimbábue. Na África do Sul além dos haplótipos A1 e A2 quatro novos haplótipos foram identificados (A3, A4, A6 e A15). O haplótipo A3 foi também identificado no Paraguai e amplamente distribuído no Brasil. Os haplótipos A4 e A6 se diferenciam do haplótipo A3 por uma única mutação cada, e também foram amostrados no Brasil.

Na América do Sul, dois países além do Brasil foram amostrados, Colômbia e Paraguai. Na Colômbia foi identificado o haplótipo A1. No Paraguai quatro haplótipos

foram identificados (A1, A2, A3 e A10). A sequência proveniente de feijão kudzu infectado foi referente ao haplótipo A2.

No Brasil, no ano agrícola 2005-2006, foi identificada a presença de oito haplótipos. Os haplótipos A1, A2 e A3 foram os mais frequentes e mais amplamente distribuídos, sendo identificados em quase todos os estados amostrados. Os haplótipos A4 e A6, que já haviam sido identificados na África do Sul, foram identificados em mais de um estado brasileiro. O haplótipo A10, que foi identificado no Paraguai, foi também identificado em um estado brasileiro. Foram identificados ainda dois novos haplótipos, A11 e A14, sendo ambos singletons. No ano agrícola 2008-2009 foram amostrados 12 dos 19 haplótipos. Nesse ano, os haplótipos A1, A2, A3 e A4 foram os mais frequentes e mais amplamente distribuídos. O haplótipos A6 e A10 foram novamente amostrados, e seis novos haplótipos foram identificados, sendo esses específicos desse ano agrícola do Brasil.

Nos Estados Unidos quatro haplótipos foram identificados (A1, A2, A5 e A8). O haplótipo A5, específico dos Estados Unidos, foi o mais frequente e mais amplamente distribuído, sendo identificado em todos os estados amostrados. A posição do haplótipo A5 na rede é um indício que é um haplótipo derivado, mas a sua ampla distribuição nos Estados Unidos indica que sua origem não é recente. Os haplótipos A1 e A2 estão presentes em diferentes frequências nos Estados Unidos, sendo esses os únicos haplótipos compartilhados com as demais localidades. Essa diferença na frequência revela a possibilidade desses dois haplótipos terem chegado aos Estados Unidos em épocas diferentes. Um haplótipo específico (haplótipo A8) foi identificado em dois estados norte-americanos. O haplótipo A8 parece ter sido originado a partir de uma mutação no haplótipo A1, e também se diferencia do haplótipo A7 por uma única mutação, que foi identificada apenas no Havaí e na Austrália.

Com base na rede de haplótipos e na distribuição geográfica é possível inferir que os haplótipos A1 e A2 são os haplótipos ancestrais. Esses dois haplótipos foram os mais frequentes, apresentaram multifurcações na rede de haplótipos dando origem aos haplótipos derivados, tiveram ampla distribuição geográfica, e foram os mais frequentes nos possíveis pontos de origem de *P. pachyrhizi*. Pode se inferir que populações de *P. pachyrhizi* portando os haplótipos A1 e A2 dispersaram dos países orientais alcançando a África. Essas populações se dispersaram por esse continente, e sofreram mutações que deram origem a novos haplótipos (A3, A4 e A6). Populações apresentando os haplótipos A1, A2, A3, A4 e A6 se dispersaram da África alcançando

o Brasil e outros países da América do Sul. No Brasil esses haplótipos apresentaram diferentes frequências, evidenciando a possibilidade dessa dispersão ter ocorrido mais de uma vez. A identificação da fonte de origem dos haplótipos norte americanos ficou dificultada. Os Estados Unidos apresentaram apenas os haplótipos A1 e A2 compartilhados com os outros locais, e esses são frequentes em todos os outros países amostrados. Os outros haplótipos presentes nos Estados Unidos (A5 e A8) foram específicos. O haplótipo A8 tem próxima relação genealógica com o haplótipo A7, que foi identificado apenas no Havaí e na Austrália. No entanto, não é possível inferir, com base nessas informações, qual foi a direção da dispersão.

4.1.5 - Fluxo gênico e diferenciação

Para avaliar o nível de fluxo gênico e diferenciação das populações desse fungo entre diferentes países, foi estimado o valor F_{ST} . A Tabela 4 sumariza os valores de F_{ST} estimados pela comparação de todos os pares populações. O valor de F_{ST} pode variar de 0 a 1, onde 0 significa que as duas populações comparadas estão agindo como uma população única (alto fluxo gênico), e o 1 significa isolamento genético total.

Entre os dois anos agrícolas brasileiros (2005-2006; 2008-2009) foi identificado um baixo valor de F_{ST} (0,02600). Esse baixo valor de F_{ST} é reflexo do alto nível de compartilhamento de haplótipos entre esses dois anos agrícolas. Entre os oito haplótipos identificados no ano agrícola 2005-2006, seis foram também identificados no ano agrícola 2008-2009. Esses resultados indicam que praticamente não existe diferença na variabilidade genética entre esses dois anos agrícolas no Brasil.

Tabela 4: Valores de F_{ST} entre pares de populações, tendo como base o nível de compartilhamento de haplótipos.

F_{ST}	África	América do Sul	Estados Unidos	Outros
Brasil	0,07119	0,09413	0,49163	0,18059
África		0,03359	0,43450	0,07386
América do Sul			0,46530	0,05468
Estados Unidos				0,24181

Ao comparar os dois anos agrícolas do Brasil com as demais localidades amostradas, o menor valor de F_{ST} foi identificado para a África (0,07119). Esses

valores indicam que o Brasil tem um alto nível de fluxo gênico com esse continente, e que esse fluxo é maior do que com qualquer outra localidade amostrada. Entre os haplótipos identificados no continente africano, apenas um não foi identificado no Brasil.

Na comparação dos Estados Unidos com os demais países amostrados, um alto valor de F_{st} foi identificado em todos os casos, sendo para a maior parte, maior que 0,4. Apenas ao comparar os Estados Unidos com os países do oriente o valor alcançado foi um pouco menor (0,24181). Esses altos valores são reflexos de apenas dois haplótipos norte-americanos serem compartilhados com as demais localidades, e dos dois outros haplótipos serem específicos desse país, sendo um deles (haplótipo A5) o mais frequente. Valores de F_{ST} inferiores a 0,15 foram identificados em todas as outras comparações. Esses valores indicam um fluxo gênico de moderado a alto entre todos os locais amostrados, com exceção dos Estados Unidos.

4.2– Histona H4

4.2.1 - Comparação das sequências

Para o gene histona H4 foi amplificada uma região de 660 bases. No total 116 clones foram sequenciados: 92 provenientes do Brasil (59 do ano agrícola 2005-2006, 32 do ano agrícola 2008-2009, e um do banco de dados da University of Missouri); oito provenientes dos Estados Unidos; sete da África; três da China; e seis do Paraguai.

A comparação das 116 sequências permitiu a identificação de 19 mutações dentro da extensão de 660 bases. Três indel (um de uma base, um de duas bases e um de três bases), e 16 substituições de base, sendo 15 identificadas em mais de uma sequência. As posições e as características das mutações podem ser observadas na figura 4. Essas mutações acarretaram na diversidade nucleotídica de 0,00647.

A análise das sequências permitiu a estimativa do conteúdo GC em 41,4%. As 19 mutações permitiram a diferenciação das 116 sequências do gene de histona H4 em 23 haplótipos, com diferentes frequências, resultando em uma diversidade haplotípica de 0,8594.

		3pb	
	2pb	1111	1112222233333333
	5580345	578568990011136	
	0162940	200059151912710	
Hap	N		
H1	32	--CATA---AAATCTATTAA-AA	(XXXXXXX) *
H2	24T	(XXXXXXX) *
H3	14	TT..G.AAA...AGGGCAG.AGT	(XXXXXXX) *
H4	10GG.....G.....T	(XXXXXXX) *
H5	5	TT..GG.....G.....T	(XXXXXXX) *
H6	4	TT.....	(XXXXXXX) *
H7	4GG....C..G....C..T	(XXXXXXX) *
H8	3GG...CC..G....C..T	(XXXXXXX) *
H9	2	TT.....T	(XXXXXXX) *
H10	2	TTA.GG.....G.....T	(XXXXXXX) *
H11	2	..ATGG.....G.....T	(XXXXXXX) *
H12	2G.AAA....G.....T	(XXXXXXX) *
H13	2G.AAA...AGGGCAG.AGT	(XXXXXXX) *
H14	1	TT...G.....T	(XXXXXXX) *
H15	1G.....T	(XXXXXXX) *
H16	1	..A.GG.....G.....T	(XXXXXXX) *
H17	1GG....C..G....C.GT	(XXXXXXX) *
H18	1GG....CG.G....C..T	(XXXXXXX) *
H19	1GG....CG.G.....T	(XXXXXXX) *
H20	1	TT..GG...CC..G....C..T	(XXXXXXX) *
H21	1G.AAA...AGGG.A..A.T	(XXXXXXX) *
H22	1	TT..G.AAA...AGGGCAG.A.T	(XXXXXXX) *
H23	1	TT..GG.....GGGCAG.A.T	(XXXXXXX) *

Figura 4: Alinhamento dos sítios variáveis na histona H4 definindo os 23 haplótipos de *P. pachyrhizi*. Pontos indicam similaridades com o haplótipo H1 (tomado como sequência de referencia) e hífens indicam gaps. Números ao alto indicam a posição do nucleotídeo na sequência de referencia durante o alinhamento. *Números de acesso no GenBank derivados do presente estudo.

4.2.2 - Distribuição das sequências

As 116 sequências analisadas para o gene de histona H4 foram representadas em uma tabela de acordo com o local de amostragem (Tabela 5). O haplótipos H1 (com 32 ocorrências) e o haplótipo H2 (com 24 ocorrências), foram os haplótipos de histona H4 mais frequentes. Esses haplótipos foram praticamente restritos aos dois anos agrícolas brasileiros, sendo compartilhados apenas com África do Sul e Paraguai. O haplótipo H1 foi mais frequente no ano agrícola 2005-2006 (21 sequências do total de 32), e o haplótipo H2 foi mais frequente no ano agrícola 2008-2009 (13 sequências no total de 24).

O haplótipo H3 foi identificado em 14 sequências, sendo 11 no Brasil e as outras na Nigéria e China. O haplótipos H4, 10 sequências, foi também identificado apenas no Brasil e na África do Sul. Os haplótipos de H5 a H23 foram identificados em baixa frequência, de uma a cinco sequências.

Foi estimado o valor de diversidade haplotípica em cada local amostrado. Os valores são sumarizados na tabela 6. Nos dois anos agrícolas brasileiros foi identificada uma diversidade haplotípica maior que 80%, reflexo da grande quantidade de número de haplótipos identificados nesse país. O maior valor de diversidade haplotípica foi observado na África, onde as sete sequências foram diferenciadas em seis haplótipos, e o menor valor foi identificado no Paraguai e China, provavelmente devido à pequena amostragem.

Tabela 6: Número de sequências de histona H4 para cada localidade, com os respectivos números de haplótipos diferenciados e a valor da diversidade haplotípica estimada.

	# sequências	Hap	Hd
Bra 2006	59	17	0,83635
Bra 2009	33	10	0,80492
África	7	6	0,95238
Estados Unidos	8	4	0,75000
Paraguai	6	2	0,60000
China	3	2	0,66667

4.2.3 - Relação genealógica entre os haplótipos

Uma rede de haplótipos foi estimada para revelar a relação genealógica entre os haplótipos de histona H4. Com base no tamanho das sequências (660pb) o número máximo de conexões parcimoniosas foi de 11 eventos mutação. Uma única rede foi obtida abrangendo os 23 de haplótipos (Figura 5). Na rede de haplótipos foi identificada a presença cinco “loops”, e a presença de “nós” ligando haplótipos. Os nós indicam haplótipos não amostrados, reflexo da pequena amostragem principalmente nos países orientais.

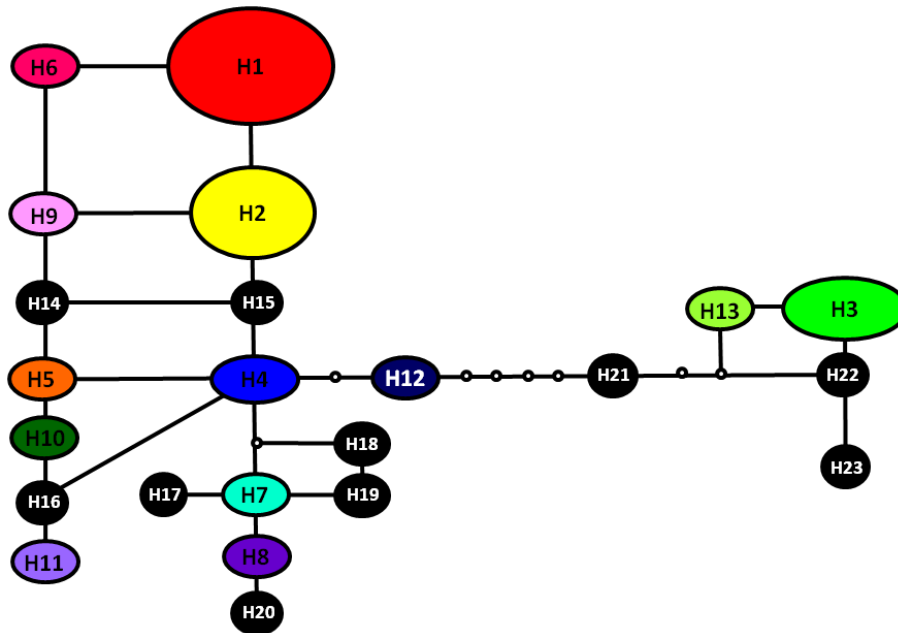


Figura 5: Rede de haplótipos para os haplótipos de histona H4 de *P. pachyrhizi*. Cada círculo representa um haplótipo; número dentro do círculo indica o haplótipo; o tamanho dos círculos tem relação com a frequência dos haplótipos; círculos coloridos são haplótipos frequentes ($N > 2$); círculos pretos são singletons ($N = 1$). Cada linha representa um passo de mutação. Círculos vazios são haplótipos não amostrados.

A rede de haplótipos de histona H4 apresentou dois subgrupos, definidos pelo grande número de mutações que os diferenciam. Um dos subgrupos (A) apresentou três haplótipos singletons, o haplótipo H13 identificado duas vezes, e o haplótipo H3, de maior frequência desse subgrupo. O outro subgrupo (B) comportou os outros 18 haplótipos. No subgrupo B foram identificados três haplótipos de grande frequência (H1, H2 e H4). Apesar dos haplótipos H1 e H2 terem sido observados em maior frequência, o haplótipo H4 ocupou o centro do subgrupo B. Haplótipos de baixa frequência (H8 e H11), e haplótipos singletons ocuparam a extremidades da rede, reforçando a idéia de serem haplótipos derivados. A grande quantidade de loops na rede também dificultou a inferência da relação ancestral/derivado de cada haplótipo.

4.2.4 - Distribuição geográfica dos haplótipos

Os 23 haplótipos de histona H4 foram plotados em mapas nos respectivos locais de amostragem (Figura 6). Na China foram identificados dois haplótipos (H3 e H22), ambos haplótipos identificados no subgrupo A da rede de haplótipos. O haplótipo H3 foi identificado em mais 13 localidades, incluindo África e Brasil. Na África seis

haplótipos foram identificados, três na África do Sul (H1, H4 e H6), e três na Nigéria (haplótipos H3, H5 e H23). O haplótipo H3 foi o único compartilhado com a China. O haplótipo H5, identificado na Nigéria, foi compartilhado apenas com os Estados Unidos. Os haplótipos da África do Sul (H1, H4 e H6) foram todos compartilhados com o Brasil.

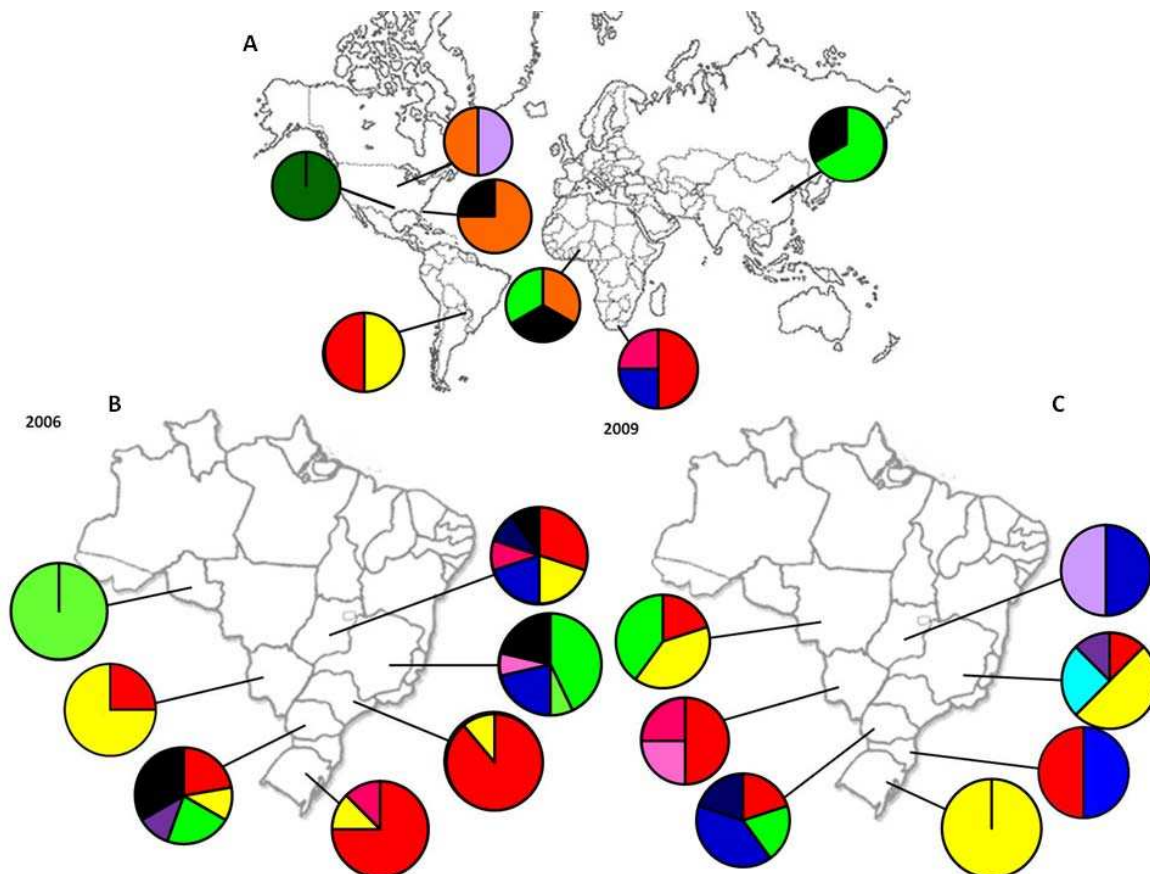


Figura 6: Distribuição geográfica dos 23 haplótipos de histona H4 de *P. pachyrhizi*. (A) China, África, Paraguai e Estados Unidos. (B) Brasil, ano agrícola 2005-2006. (C) Brasil, ano agrícola 2008-2009. Círculos representam o conjunto de haplótipos identificados em cada local, e as divisões de cada círculo são proporcionais a frequência de cada haplótipo identificado.

No Paraguai foram identificados os haplótipos H1 e H2, na mesma frequência. Esses dois haplótipos se diferenciam por uma única mutação e foram os haplótipos mais frequentes. Nos dois anos agrícolas brasileiros foram identificados 18 haplótipos, sendo nove deles observados nos dois anos. Os H1 e H2 foram os mais frequentes, sendo identificados em praticamente todos os estados brasileiros amostrados. O haplótipo H3, que já havia sido identificado na China e Nigéria, também foi identificado nos dois anos agrícolas em diferentes estados. Os haplótipos H4 e H6, identificados na

África do Sul, foram identificados em baixa frequência nos dois anos agrícolas. Os outros haplótipos foram privativos do Brasil.

Nos Estados Unidos quatro haplótipos foram identificados. O haplótipo de maior frequência, haplótipo 5, foi identificado em quatro sequências em dois diferentes estados. Esse haplótipo foi compartilhado apenas com a Nigéria. Os haplótipos 10 e 16 foram privativos desse país, sendo identificados duas e uma vez, respectivamente. O haplótipo 11 foi identificado uma vez nesse país e uma vez no Brasil. Os quatro haplótipos presentes nos Estados Unidos (5, 10, 11 e 16) são proximamente relacionados na rede de haplótipos, se diferenciam uns dos outros por uma mutação, e parecem ser derivados do haplótipo 4.

As informações obtidas a partir da rede de haplótipos de histona H4 e das respectivas distribuições geográficas indicam que, haplótipos ocupando o subgrupo A da rede de haplótipos parecem ser os haplótipos ancestrais. Esses haplótipos ancestrais sofreram mutações, originando novos haplótipos que foram identificados uma única vez, ou não foram amostrados, reflexo da baixa amostragem nos países orientais. A presença do haplótipo H3 na China e na África, permitiu inferir que populações portando esse haplótipo dispersaram para a África, onde sofreu mutações e deu origem a novos haplótipos, tais como: H1, H4, H5 e H6. Pela posição ocupada na rede, o haplótipo H4 parece ter sido um dos primeiros haplótipos originado, o qual sofreu mutação dando origem aos outros haplótipos. O nível de compartilhamento de haplótipos entre África e Brasil possibilita inferir que populações africanas dispersaram e colonizaram produções de soja no Brasil. A diferença na frequência dos haplótipos compartilhados entre Brasil e África indica que esses haplótipos devem ter chegado ao Brasil em épocas diferentes. Já o nível de compartilhamento de haplótipos americanos com as demais localidades, não permitiu inferir, com segurança, a origem dos haplótipos presentes nesse país.

4.2.5 - Fluxo gênico e diferenciação

O nível de fluxo gênico entre os diferentes países foi estimado ao calcular o valor de F_{ST} . A Tabela 7 sumariza os valores da comparação entre pares de populações. Ao comparar os dois anos agrícolas brasileiros, um baixo valor de F_{ST} foi identificado (0,00247), revelando que não existe diferença no nível de variabilidade genética entre os dois diferentes anos.

Tabela 7: Valores de F_{ST} entre pares de populações, tendo como base o nível de compartilhamento de haplótipos de histona H4.

F_{ST}	África	Estados Unidos	Paraguai	China
Brasil	0,023825	0,39654	0,16916	0,752245
África		0,19572	0,26329	0,61572
Estados Unidos			0,79746	0,90833
Paraguai				0,95824

O menor valor de F_{ST} foi identificado entre Brasil e África. Esse baixo valor evidencia o alto nível de fluxo gênico entre essas duas localidades. Altos valores de F_{ST} foram identificados ao comparar os Estados Unidos com as outras localidades, indicando o alto nível de diferenciação de populações do fungo presentes nesse país em relação às populações presentes nos outros e continentes.

4.3 – ITS (Espaçador transcrito interno)

4.3.1 - Comparação das sequências

Para a região espaçadora interna (ITS) do DNA ribossomal nuclear foram sequenciados 227 clones. 112 provenientes do Brasil (2005-2006), 102 do Brasil (2008-2009), 10 da África do Sul, e três de feijão kudzu infectados com *P. pachyrhizi*, provenientes do Paraguai. Para as análises dessas sequências, foram incorporadas 106 sequências obtidas no banco de dados da Division of Plant Sciences, Life Sciences Center (University of Missouri) e 37 sequências obtidas no Genbank. Das sequências obtidas no banco de dados da Division of Plant Sciences, 61 foram obtidas em 11 diferentes estados americanos, e 45 foram provenientes de outros países. Entre as sequências obtidas no Ganbank 13 foram obtidas por Frederick *et al.* (2002), principalmente amostras de países orientais e África, e 24 foram sequências obtidas por Barnes *et al.* (2008), todas proveniente de Flórida (EUA). Foram analisadas 370 sequências de ITS de *P. pachyrhizi*.

As 370 sequências de ITS de *P. pachyrhizi* foram alinhadas, resultando em uma sequência consenso de 646 bases. A comparação das 370 sequências permitiu a identificação de sete sítios polimórficos e sete indels, totalizando em uma diversidade nucleotídica de 0,00135 (Figura 7).

Hap	N		
			2pb 2pb 3pb 4pb 1112222223444 44 4446 799003550133 44 4580 067379348228 01 4019
R1	187	C--TTT---AC-----T	EU436701
R2	7C.....T	(XXXXXX)*
R3	17TTT....G.T	(XXXXXX)*
R4	6AAT...TTT....G.A	(XXXXXX)*
R5	18	.TT..AAT...TTT....G.A	(XXXXXX)*
R6	6	.TT..AAT...TTT....G.T	(XXXXXX)*
R7	5	.TT.....T	(XXXXXX)*
R8	3	.TT.....A	(XXXXXX)*
R9	1	.TT.....TTT....G.T	(XXXXXX)*
R10	23A	EU436709
R11	1	TTT..A.....T	(XXXXXX)*
R12	3T.....T	(XXXXXX)*
R13	1	.TT..AAT.....A	(XXXXXX)*
R14	3AAT.....T	(XXXXXX)*
R15	1AT...TTT....G.T	(XXXXXX)*
R16	1	.TT..AT...TTT....G.T	(XXXXXX)*
R17	1TT	(XXXXXX)*
R18	7	.TT..A.....T	(XXXXXX)*
R19	1AT.....T	(XXXXXX)*
R20	1	.TT..AT.....T	(XXXXXX)*
R21	1	.TT..A.....A	(XXXXXX)*
R22	1	.TT..A.....TTT....G.T	(XXXXXX)*
R23	14TTT....G.A	EU436708
R24	3AT...TTT....G.A	(XXXXXX)*
R25	23A.....TTT....G.A	EU436700
R26	1	.TT..AT...TTT....G.A	(XXXXXX)*
R27	7A.....TTT....G.T	EU436703
R28	1	...C.A.....TTT....G.A	EU436715
R29	6TTT....GTA	(XXXXXX)*
R30	1	.TT.....TTT....G.A	(XXXXXX)*
R31	2	.TT..A.....TTT....G.A	(XXXXXX)*
R32	1A.....TTT....GTA	(XXXXXX)*
R33	1	.TT.....TTT....GTT	(XXXXXX)*
R34	1G.T	(XXXXXX)*
R35	1	.TT..AAT.....T	(XXXXXX)*
R36	1TTTTTTT..T	(XXXXXX)*
R37	1AAT.....G.A	(XXXXXX)*
R38	2A.....A	(XXXXXX)*
R39	6A.....T	(XXXXXX)*
R40	1AA.....TTT....G.A	(XXXXXX)*
R41	1ATA.....T	(XXXXXX)*
R42	1TTT.....A	(XXXXXX)*

Figura 7: Alinhamento dos sítios variáveis no ITS definindo os 42 ribótipos de *P. pachyrhizi*. Pontos indicam similaridades ao ribótipo R1 (tomado como sequência de referência) e hífens indicam gaps. Números ao alto indicam a posição do nucleotídeo na sequência de referência durante o alinhamento.*Números de acesso no GenBank derivados do presente estudo.

Das sete substituições de base identificadas em ITS, quatro foram identificadas mais de uma vez e três foram singltons. Dos indels, três foram de uma única base, dois foram de duas bases cada, um foi de três bases, e um foi de quatro bases. As 14 mutações identificadas no ITS permitiu a diferenciação de 42 ribótipos, resultando em uma diversidade haplotípica de 0,7301.

4.3.2 – Distribuição das sequências

As 370 sequências de ITS de *P. pachyrhizi* foram distribuídas em uma tabela nos respectivos locais de amostragem (Tabela 8).

O ribótipo R1 apresentou uma elevada frequência, 187 sequências das 370 analisadas. Esse ribótipo foi o mais frequente em todas as localidades amostradas, sendo essa frequência mais ressaltada nos dois anos agrícolas brasileiros. Cinco ribótipos (R3, R5, R10, R23 e R25) foram frequentes em locais específicos. O ribótipo R25 foi praticamente restrito aos Estados Unidos, sendo observado apenas uma vez no feijão kudzu do Paraguai, e uma vez no Havaí. Os outros 36 ribótipos foram identificados em baixa frequência (1 a 7 sequências). Dentre esses ribótipos, 14 foram privativos ao Brasil, cinco foram privativos dos Estados Unidos e os outros foram distribuídos entre os outros países.

A diversidade haplotípica de cada país ou continente foi estimada com base no número e frequência de ribótipos (Tabela 9). Para o primeiro ano agrícola brasileiro, 112 sequências foram analisadas e 18 ribótipos foram identificados. Para o segundo ano agrícola brasileiro, as 102 sequências foram diferenciadas em 16 haplótipos. Esses valores permitiram a inferência de uma diversidade haplotípica de cerca de 57% para esse país. Esse valor foi muito inferior ao observado nas outras localidades, e ao valor de diversidade haplotípica observada no Brasil para os outros genes. Esse valor de diversidade haplotípica no Brasil é reflexo da alta frequência do ribótipo R1 nesse país. Os outros países amostrados apresentaram alto valor de diversidade haplotípica, principalmente na América do Sul e nos países orientais, onde esse valor passou de 0,9.

Tabela 8: Distribuição das 370 sequências do ITS de *P. pachyrhizi* entre as localidades amostradas e os 42 ribótipos correspondentes, contendo o número total de sequências analisadas (#) e o número de haplótipos identificados em cada localidade (Nhap).

	#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
Bra 2006	112	73	4	5	1	7	2	3	1	1	7	1	1	1	1	1	1	1	1											
Bra 2009	102	66	3	5	5	4	1	1			5		2		1				4	1	1	1	1	1						
EUA	85	28		6							6													7		21		6	1	
Africa Sul	13	9				1		1			1																			
Paraguai	3	2																												
Kudzu	3																								1	1	1			
Colômbia	3					1																								2
Brasil	4	1					1																							1
Nigéria	6	2				1			1																					
Zimbabue	4					1													1					1						
Hawai	3			1																					1	1				
Austrália	2																								1			1		
China	6	2				1			1						1															
Índia	3																								2					
Filipinas	4					1																								1
Vietnã	7	2					2				2														1					
Tailândia	4	1				1					1													1						
Taiwan	6	1									1								1					1						2
Total	370	187	7	17	6	18	6	5	3	1	23	1	3	1	3	1	1	1	7	1	1	1	1	14	3	23	1	7	1	6

	30	31	32	33	34	35	36	37	38	39	40	41	42	NHap
Bra 2006														18
Bra 2009														16
EUA									1	6	1	1	1	12
Africa do Sul						1								5
Paraguai			1											2
Kudzu														3
Colômbia														2
Brasil									1					4
Nigéria				1	1									5
Zimbabue								1						4
Hawai														3
Austrália														2
China		1												5
India						1								2
Filipinas	1	1												4
Vietnã														4
Tailândia														4
Taiwan														5
Total	1	2	1	1	1	1	1	1	2	6	1	1	1	

Continuação da tabela 8.

Tabela 9: Número de sequências de ITS para cada localidade, com os respectivos números de ribótipos diferenciados e a valor da diversidade haplotípica estimada.

	# sequências	hap	Hd
Brasil 2006	112	18	0,56725
Brasil 2009	102	16	0,57860
África	23	11	0,77075
EUA	85	12	0,80975
América do Sul	13	9	0,92308
Outros	35	16	0,93277

4.3.3 – Relação genealógica entre os ribótipos

Uma única de rede ribótipos foi estimada revelando a relação genealógica entre os 42 ribótipos de ITS de *P. pachyrhizi* (Figura 8). Para a obtenção dessa rede o número máximo possível calculado de mutações entre dois vizinhos mais próximos foi de 10 eventos de mutação. A rede de haplótipos obtida para a região de ITS revelou a presença de muitos “loops”, dificultando a identificação da origem dos ribótipos. Apenas um “nó” foi identificado na rede, o qual corresponde um ribótipo já extinto ou não amostrado.

O ribótipo R1, ribótipo mais frequente, ocupou posição central na rede de ribótipos, e apresentou multifurcações para outros ribótipos. Essas características são indícios da ancestralidade deste ribótipo. Outros ribótipos de alta frequência (R3, R5, R10, R23 e R25), também ocuparam pontos centrais da rede, e apresentaram multifurcações que originaram outros ribótipos. Essas características são indícios que os ribótipos R3, R5, R10, R23 e R25 são derivados em relação ao ribótipo R1, mas que essa origem se deu a tempo suficiente para que esses ribótipos tivessem elevada frequência, e que sofressem mutações para originar novos ribótipos. Os outros ribótipos ocuparam as extremidades da rede e/ou foram identificados em baixa frequência, indicando serem ribótipos derivados e de origem recente.

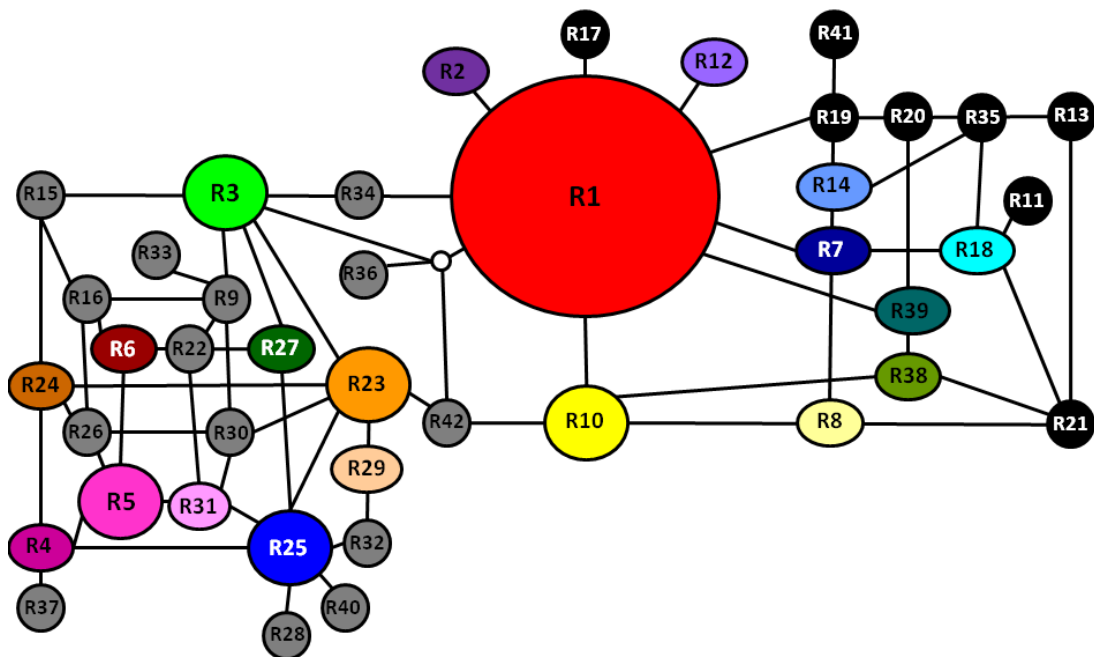


Figura 8: Rede de ribótipos para ITS de *P. pachyrhizi*. Cada círculo representa um ribótipo; número dentro do círculo ou ao lado indica o haplótipo; o tamanho dos círculos tem relação com a frequência dos ribótipos; círculos coloridos são haplótipos frequentes (N>2); círculos cinza ou pretos são singletons (N=1). Cada linha representa um passo de mutação. Círculos vazios são haplótipos não amostrados.

4.3.4 - Distribuição geográfica dos ribótipos

Os 42 ribótipos de ITS de *P. pachyrhizi* foram plotados nos mapas referentes aos locais de amostragem (Figura 9). Entre os países orientais, possível local de origem de *P. pachyrhizi*, foi observada uma grande diversidade de ribótipos. Nesses países não houve prevalência de nenhum ribótipo específico. Dentre os 12 ribótipos observados, foram identificados os haplótipos R1, R5, R10 e R23. Esses ribótipos se apresentaram em alta frequência nas outras localidades, e ocuparam pontos centrais na rede.

Na África, o ribótipo R1 foi o mais frequente, principalmente na África do Sul. Dos ribótipos de alta frequência, os haplótipos R1, R5, R10 e R23 foram identificados na África e compartilhados com os países orientais. Ainda foram compartilhados entre a África e os países orientais, os ribótipos R8 e R18. Foram identificados nesse continente o ribótipo R7 que posteriormente foi identificado também no Brasil, e quatro ribótipos singletons.

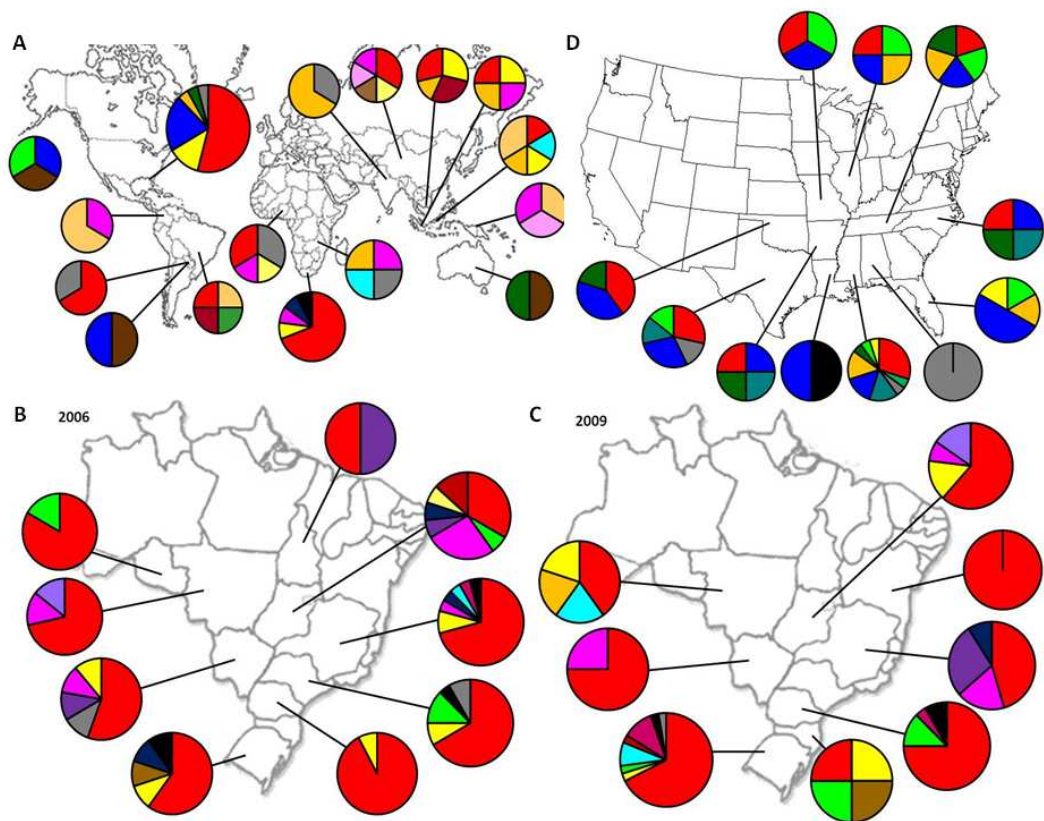


Figura 9: Distribuição geográfica dos 42 ribótipos de ITS de *P. pachyrhizi*. (A) Ásia, Austrália, África, América do Sul, Havaí e Florida (EUA). (B) Estados Unidos. (C) Brasil, ano agrícola 2005-2006. (D) Brasil, ano agrícola 2008-2009. Círculos representam o conjunto de ribótipos identificados em cada local, e as divisões de cada círculo são proporcionais a frequência de cada ribótipo identificado.

Entre os países sul-americanos (Colômbia e Paraguai) seis ribótipos foram identificados. Na Colômbia foram identificados os ribótipos R5 e R29. O ribótipo R5, ribótipo de alta frequência, já havia sido identificado nos países orientais e África, e o ribótipo R29 compartilhado apenas com os países orientais. No Paraguai dois ribótipos (R1 e R32) foram identificados em amostras de soja, e dois ribótipos (R24 e R25) foram identificados em amostras de feijão kudzu. O ribótipo R24 foi identificado apenas nesse país, no Havaí e na Austrália, e o ribótipo R25 foi também identificado no Havaí e com alta frequência nos Estados Unidos.

Nos dois anos agrícolas brasileiros (2005-2006; 2008-2009), o ribótipo R1 foi o mais frequente, sendo identificado em todos os estados amostrados. Além do ribótipo R1, 22 outros ribótipos foram identificados no Brasil, sendo 13 específicos desse país. Com exceção dos ribótipos singletons, todos os ribótipos identificados na África foram também identificados no Brasil.

Nos Estados Unidos 12 ribótipos foram amostrados, com prevalência dos ribótipos R1 e R25. Dentre esses ribótipos, o ribótipo R1 foi identificado em todos

os outros locais amostrados, e o ribótipo R25 foi identificado apenas no Havaí e na amostra de feijão kudzu, do Paraguai. O ribótipo R3, identificado em seis sequências nesse país, foi identificado nos dois anos agrícolas brasileiros, e no Havaí. Ainda, cinco outros ribótipos foram específicos dos Estados Unidos.

A análise da frequência dos ribótipos, da rede com as respectivas relações genealógicas, e da distribuição geográfica dos ribótipos de ITS de *P. pachyrhizi*, permitiram inferir a possível história evolutiva de *P. pachyrhizi*. A frequência dos ribótipos R1, R5, R10 e R23, os seus posicionamentos centrais na rede, e a identificação desses ribótipos no possível ponto de origem desse fungo, fornecem indícios que esses são ribótipos ancestrais. A ocorrência desses ribótipos (R1, R5, R10 e R23) na África evidencia que populações desse fungo se dispersaram dos países orientais alcançando a África, onde sofreram mutação e deram origem a novos ribótipos. O nível de compartilhamento de haplótipos de ITS entre América do Sul e África reforça a hipótese da dispersão de esporos de fungo entre esses dois continentes. No Paraguai foi identificado pela primeira vez o ribótipo R25, que foi identificado também no Havaí e teve ampla distribuição nos Estados Unidos. No Brasil, os ribótipos provenientes da África tiveram ampla distribuição, principalmente o ribótipo R1. Novos ribótipos foram identificados no Brasil, entre eles o ribótipo R3, também identificado no Havaí e de alta frequência nos Estados Unidos. Não foi possível inferir, com segurança, a origem dos ribótipos norte americanos. Com base na alta frequência dos ribótipos R1, R3 e R25 nesse país, existem a possibilidade do ponto de origem ser o Havaí e/ou a América do Sul.

4.3.5 - Fluxo gênico e diferenciação

Tendo como base o nível de compartilhamento de ribótipos de ITS de *P. pachyrhizi* entre os diferentes países e continentes, foi possível calcular o valor de F_{ST} entre pares de populações (Tabela 10). Esses valores indicam o nível de fluxo gênico entre as diferentes localidades, permitindo inferir sobre o grau de diferenciação de populações. Baixo valor de F_{ST} foi identificado ao comparar os dois anos agrícolas brasileiros (0,00727). Esse baixo valor indica que não existe diferença na variabilidade genética nas populações desse fungo nos dois diferentes anos no Brasil.

Tabela 10: Valores de F_{ST} entre pares de populações, tendo como base o nível de compartilhamento de haplótipos de ITS.

F_{ST}	África	Estados Unidos	América do Sul	Outros
Brasil	0,02178	0,18156	0,32690	0,21959
África		0,10214	0,18920	0,09115
Estados Unidos			0,10443	0,04693
América do Sul				0,01971

Ao comparar o Brasil com os demais países e continentes, baixo valor de F_{ST} foi identificado na comparação com a África (0,02178). Esse valor reflete o alto nível de compartilhamento de ribótipos entre as duas localidades, revela o baixo nível de diferenciação das populações desses dois locais, indicando a possibilidade dos ribótipos que chegaram ao Brasil serem originários da África.

Ao comparar os Estados Unidos com os demais países e continentes, um valor de F_{ST} de 0,04693 foi identificado na comparação com os países denominados “outros”. Esse valor indica um baixo nível de diferenciação. Entre os países denominados “outros” está o Havaí. Essas características evidenciam alto nível de fluxo gênico entre Havaí e Estados Unidos, e que ribótipos norte americanos podem ser originários do Havaí.

O programa Arlequin versão 3.1 (Excoffier *et al.*, 2005) foi utilizado para fazer a análise de variância molecular das populações desse fungo baseado nos dados das sequências de ITS (Tabela 11). Foi calculada a porcentagem de variação em três níveis: entre grupos (entre países e/ou continentes); entre populações dentro do grupo (entre os estados de um mesmo país); e dentro de populações (entre as diferentes amostras de um mesmo estado).

A maior parte da variação foi observada dentro das populações, 91,97%. Essa alta porcentagem de variação dentro das populações, e baixa porcentagem entre grupos, refletem o alto nível de compartilhamento dos ribótipos mais frequentes em praticamente todas as localidades amostradas, e da presença de ribótipos de menores frequência específicos de cada local. Esses valores indicam que existe um alto nível de fluxo gênico entre todas as populações de *P. pachyrhizi*, e que essas populações podem agir como uma única população.

Tabela 11: Cálculo da análise da variância molecular (AMOVA) de populações de *P. pachyrhizi* com base no ITS, feito pelo programa Arlequin versão 3.1 (Excoffier *et al.*, 2005).

Fonte de Variação	GL	Soma dos quadrados	Componente de Variância	Porcentagem de Variação
Entre Grupos	4	9,505	0,00825 va	2,21
Entre pop. dentro do grupo	1	0,705	0,02178 vb	5,83
Dentro de populações	364	125,117	0.34373 vc	91,97
Total	369	135,327	0.37375	

5 – Discussão

5.1 – Variabilidade genética de *P. pachyrhizi*

Um dos objetivos deste estudo foi determinar o nível de variabilidade genética intra-específica do fungo causador da ferrugem da soja, *P. pachyrhizi*. Para essa determinação, foi examinada a presença de polimorfismos em três regiões gênicas do DNA nuclear desse organismo. A escolha de mais de uma região gênica teve como objetivo garantir que genes com diferentes níveis de evolução fossem analisados.

Para as três regiões gênicas analisadas, foram identificados valores de diversidade haplotípica superiores a 70%. O valor de diversidade haplotípica corresponde à diversidade gênica média, a qual é uma medida da variabilidade genética da população em estudo. Os valores de diversidade haplotípica identificados para populações de *P. pachyrhizi* indicam que esse fungo tem uma alta variabilidade genética, principalmente por ser um fungo que parece se reproduzir apenas assexuadamente. Uma alta diversidade genética de *P. pachyrhizi* já foi estimada por estudos baseados em: diferenciação de raças (Yamaoka *et al.* 2002) ; sequenciamento de regiões gênicas (Freire *et al.* 2008; Barnes *et al.* 2009); na medição quantitativa da virulência (Twizeyimana *et al.* 2009); e marcadores SSR (Twizeyimana *et al.* 2011).

Duas características de populações de *P. pachyrhizi* podem contribuir para o alto nível de variabilidade desse fungo: alta taxa de mutação e alta taxa de dispersão. Genes de resistência (*Rpp*) conhecidos por serem eficientes a isolados do fungo têm sua resistência rapidamente suplantada. Ainda, alta diversidade do fungo tem sido identificada no Brasil e nos Estados Unidos, países onde o fungo foi introduzido recentemente. Tais características indicam que populações de *P. pachyrhizi* têm potencial para evoluir e se adaptar rapidamente a mudanças ambientais. Essa alta taxa de mutação aliada à alta taxa de dispersão permite que populações de diferentes locais e geneticamente diversas, estejam sendo constantemente trocadas, acentuando a variabilidade genética de *P. pachyrhizi*.

Fungos patogênicos de plantas possuem uma variedade de mecanismos que podem resultar no aumento da diversidade genética, entre eles a recombinação. No entanto, *P. pachyrhizi* parece se reproduzir apenas assexuadamente (o estágio sexual ainda não foi identificado na natureza) o que limitaria a variação genética desse fungo. A ocorrência de recombinação sexual em

populações desse fungo já foi avaliada e não foi identificada (Twizeyimana *et al.*, 2011). No entanto, os dados do nível de variação genética desse fungo indicam a possibilidade da existência de outros fatores genéticos envolvidos no alto grau de variação desse organismo, que não poderia ser explicado apenas pelo acúmulo de mutação de ponto (Freire *et al.*, 2008).

Características das mutações identificadas no presente trabalho indicam a possibilidade de eventos de recombinação interna nas regiões gênicas analisadas. As redes de haplótipos obtida para cada uma das regiões gênicas apresentaram loops internos. A presença desses loops na rede de haplótipos indica uma relação reticulada dos haplótipos (Hare, 2001). A recombinação é o mecanismo responsável por esse padrão reticulado de evolução (Nunney *et al.*, 2010). Essas observações sugerem que as regiões gênicas de *P. pachyrhizi* poderão se diversificar por pelo menos duas vias: alta taxa de mutações de ponto; e, possivelmente, por recombinação somática.

5.2 - Estrutura genética das populações de *P. pachyrhizi*

O conhecimento do nível da variação genética de uma espécie e como essa variação é distribuída entre as populações é um importante fator para o entendimento na biologia de populações de fungos patogênicos. A informação da estrutura genética permitiu inferir quais forças evolutivas estão influenciando na evolução das populações do patógeno (Twizeyimana *et al.*, 2011).

A análise da distribuição da diversidade de *P. pachyrhizi* permitiu identificar que apenas um ou dois haplótipos, de cada região gênica, apresentaram alta frequência e foram amplamente distribuídos nos países amostrados. A maioria dos haplótipos foram identificados em baixa frequência, e tiveram a distribuição restrita a um determinado local ou país. Esse tipo de distribuição resultou em uma estrutura genética em que a maior parte da diversidade genética de *P. pachyrhizi* se concentra dentro das populações (>90%). Esse tipo de estrutura da população sugere a possibilidade de rápida adaptação do fungo a mudanças ambientais.

Essa estrutura genética para populações de *P. pachyrhizi* já vêm sendo identificada por outros autores em nível de país. Freire *et al.* (2008) identificaram esse padrão de distribuição ao analisar populações desse fungo no Brasil, Barnes *et al.* (2008) nos Estados Unidos, e Twizeyimana *et al.* (2011) na Nigéria. Ainda, os baixos valores de F_{ST} identificados ao comparar os diferentes países e continentes indicam um alto nível de fluxo gênico das populações do fungo. A alta taxa de dispersão, e conseqüentemente fluxo gênico, podem ser responsáveis pela limitada

variação genética entre as diferentes localidades. Esses dados permitem assumir que a dispersão do fungo a longas distâncias pela ação do vento, e a alta taxa de mutação são uma das principais forças agindo na estrutura da população de *P. pachyrhizi*.

5.3 – Efeito do vazio sanitário e relevância dos hospedeiros alternativos nas populações de *P. pachyrhizi*

No Brasil, o vazio sanitário foi implantado desde 2006, como estratégia de controle de *P. pachyrhizi*. O vazio sanitário corresponde ao período de 90 dias sem o cultivo da soja durante a entressafra. Essa estratégia tem como objetivo reduzir o inóculo de *P. pachyrhizi* nos primeiros plantios, e dessa forma diminuir a incidência da doença. A adoção do vazio sanitário tem se mostrado eficiente, resultando na redução do número médio de aplicação de fungicidas (Godoy 2009). Essa redução no número de inóculo pode resultar no estrangulamento genético (redução da diversidade gênica) das populações de *P. pachyrhizi* no Brasil.

Dois anos agrícolas brasileiros foram amostrados, com diferença de dois anos entre eles. Os dados obtidos a partir dessa amostragem permitiram inferir que não existe diferença no nível de variabilidade genética nas populações de *P. pachyrhizi* entre os dois anos. Essa observação indica duas possibilidades: o vazio sanitário não está sendo eficiente como um gargalo genético nas populações de *P. pachyrhizi* no Brasil, ou o vazio sanitário está sendo compensado pela alta taxa de dispersão, que pode levar a uma alta diversidade genética, mesmo em populações que sofreram estrangulamento.

O fungo *P. pachyrhizi* é capaz de infectar uma variedade de hospedeiros, entre eles o feijão kudzu. Na ausência de soja no campo, (vazio sanitário no Brasil e época de geadas nos Estados Unidos) esses hospedeiros parecem agir como fonte de inóculo para o próximo ano agrícola. A sobrevivência de *P. pachyrhizi* nos Estados Unidos no período de inverno é dependente da presença de outros hospedeiros, como o feijão kudzu, localizados principalmente nos estados da Flórida e Texas (Christiano e Scerm, 2007). Esses locais são as principais áreas de refúgio do fungo *P. pachyrhizi* no inverno.

A amostragem de *P. pachyrhizi* a partir do feijão kudzu, realizada no presente trabalho, permitiu concluir que os haplótipos do fungo identificados nesse hospedeiro alternativo são os mesmos haplótipos também identificados na soja. Portanto, para as três regiões gênicas analisadas, não foi identificada nenhuma forma de especialização desse fungo para a infecção dos diferentes hospedeiros.

Esses resultados reforçam a hipótese da importância dos hospedeiros alternativos como fonte de inóculo de *P. pachyrhizi* no início da safra.

5.4 – Rotas hipotéticas de dispersão e colonização de *P. pachyrhizi*

As análises das redes de haplótipos e das respectivas distribuições geográficas dos haplótipos forneceram bases para inferências das possíveis rotas de dispersão e colonização de *P. pachyrhizi* (Figura 10).

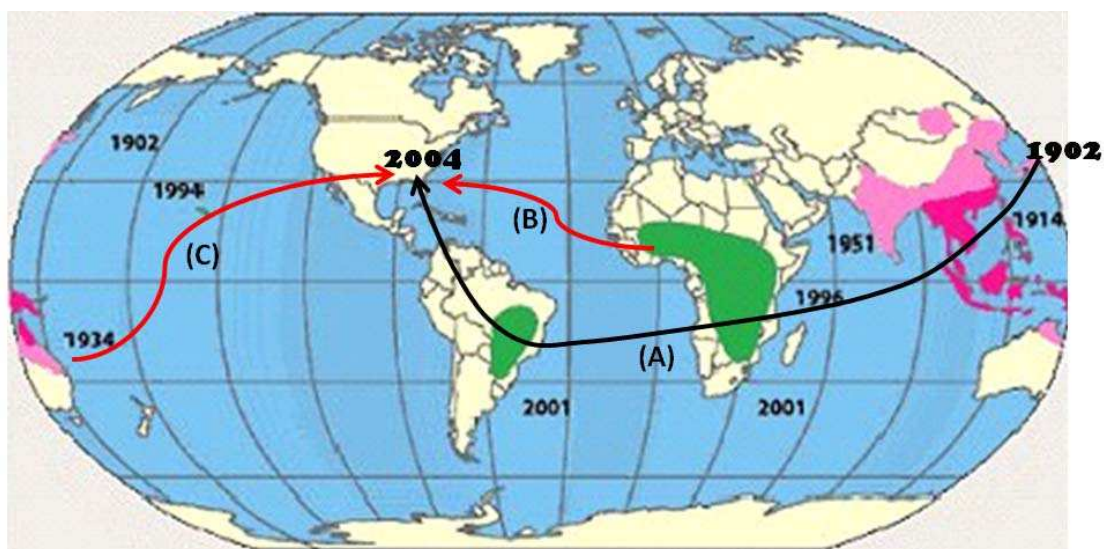


Figura 10: Possíveis rotas de dispersão de *Phakopsora pachyrhizi*. Datas próximas aos continentes correspondem ao ano de primeira identificação de *P. pachyrhizi*. Setas indicam possíveis rotas de dispersão de *P. pachyrhizi*. As três rotas são coincidentes até a ocorrência no Brasil. (A) Rota de dispersão mais aceita, América do Sul como fonte de dispersão para os Estados Unidos. (B) África como fonte de dispersão para os Estados Unidos. (C) Dispersão da Austrália para o Havaí e do Havaí para os Estados Unidos.

Os esporos do fungo *P. pachyrhizi* são capazes de se dispersar em longas distâncias pela ação do vento. A hipótese mais aceita para a introdução desse fungo na África é a dispersão, pelo vento, proveniente de países orientais. Os níveis de compartilhamento de haplótipos africanos e dos países orientais reforçam a hipótese dos países orientais como ponto de origem dos haplótipos africanos. O compartilhamento de mais de um haplótipo entre esses locais, e a diferença na frequência desses haplótipos na África sugerem que a introdução inicial foi de mais de um haplótipo e/ou, que a diversidade observada é fruto de múltiplas introduções. Ao analisar a população de *P. pachyrhizi* da Nigéria, Twizeyimana *et al.* (2011) inferiram que a atual diversidade genética seria resultado de múltiplas introduções, e que essas seriam provenientes de isolados de Taiwan.

Na continente africano ainda foram observados haplótipos que ainda não haviam sido identificados nos países orientais. Foram identificados haplótipos que foram compartilhados com o continente americano, e haplótipos exclusivos da África. Os haplótipos identificados pela primeira vez na África e compartilhados com o continente africano ocuparam posições intermediárias na rede, ligando os haplótipos centrais e haplótipos periféricos. Esses haplótipos parecem ter origem mais recente, mas que essa origem já ocorreu a tempo suficiente para que populações portando esses haplótipos sofressem dispersão para o continente americano, e que apresentassem uma ampla distribuição, principalmente no Brasil. Os haplótipos privativos da África ocuparam as extremidades das redes de haplótipos. Esses haplótipos apresentam característica de haplótipos derivados, e de origem recente.

Os continentes africanos e sul americanos estão distantes por cerca de 4000 km, separados pelo oceano Atlântico. Já existe relato de patógenos de plantas que se dispersaram entre esse dois continentes, por exemplo, esporos do fungo *Hemelia vastatrix*, fungo causador da ferrugem do café, que se dispersou a partir da Angola (África) alcançando o Brasil (Nagarajan e Singh, 1990). Freire *et al.* (2008) ao avaliarem a diversidade de haplótipos de populações brasileiras de *P. pachyrhizi*, concluíram que os haplótipos identificados no Brasil são proximamente relacionados genealogicamente aos haplótipos identificados no continente africano.

No presente trabalho um alto nível de fluxo gênico foi identificado entre África e América do Sul, principalmente África e Brasil. Esse alto nível de fluxo gênico foi refletido pelos baixos valores de F_{ST} identificados ao comparar esses dois locais. Para as três regiões gênicas praticamente todos os haplótipos identificados no continente africano, que não foram restritos a esse continente, foram também identificados no Brasil. Essas características reforçam a hipótese de dispersão de isolados africanos de *P. pachyrhizi* para América do Sul, inclusive Brasil. Os resultados ainda permitem inferir que a introdução de *P. pachyrhizi* no Brasil e nos outros países da América do Sul foi haploticamente diversa, e esse processo parece ter ocorrido mais de uma vez. Com base no nível de compartilhamento de haplótipos entre Brasil e África é possível ainda inferir que o processo de dispersão entre esses dois países é um processo contínuo.

O Brasil ainda apresentou uma variedade de haplótipos que não haviam sido identificados nos países orientais ou África. Alguns desses haplótipos tiveram ampla distribuição nesse país, enquanto outros tiveram distribuição restrita, sendo alguns singltons. Os haplótipos de ampla distribuição parecem ter sido originados

a tempo suficiente para se dispersarem nesse país. E os haplótipos de distribuição restrita seriam haplótipos derivados e de origem recente.

Três hipóteses são utilizadas para explicar a introdução de *P. pachyrhizi* nos Estados Unidos, todas essas tendo como base a dispersão em longa distância. Essas hipóteses se diferenciam pelo possível ponto de origem da dispersão desses fungos. Os três possíveis focos de dispersão de *P. pachyrhizi* para os Estados Unidos são: América do Sul; África (Isard *et al.*, 2005); e Havaí (Zhang, informação pessoal). Isard *et al.*, (2005) ao avaliarem a produção de esporos na África e na América do Sul concluíram que a América do Sul é a mais provável fonte de esporos para os Estados Unidos, e que essa dispersão teria sido facilitada por ação de tornados. Recentemente, Zhang (informação pessoal) avaliou o nível de compartilhamento de haplótipos de *P. pachyrhizi* entre os diferentes países e inferiu uma nova rota de dispersão a partir do Oeste da Ásia. De acordo com dados obtidos, populações de *P. pachyrhizi* migrariam da Ásia para Austrália, ou para o Havaí diretamente; da Austrália para o Havaí; e do Havaí para os Estados Unidos.

Para o gene de histona H4 os haplótipos dos Estados Unidos tiveram relações genealógicas mais evidenciadas com haplótipos do continente africano. No entanto, a pequena amostragem norte americana para esse gene, não garante confiabilidade para essa estimativa. As análises da distribuição geográfica e da rede de haplótipos dos genes de ITS e do fator de ribosilação do ADP evidenciaram uma forte relação genealógica entre haplótipos da Austrália com o Havaí, e do Havaí com os Estados Unidos. Esses dados reforçaram a hipótese evidenciada por Xuencheng (dados não publicados). Para a região de ITS ainda foram identificados haplótipos compartilhados apenas entre Brasil e Estados Unidos, revelando a possibilidade desse local como fonte de origem dos haplótipos norte americanos. Esses resultados não permitem a exclusão de nenhuma das possíveis fontes de origem do haplótipos identificados nos Estados Unidos, e ainda indicam a possibilidade desse país ter sofrido a introdução de mais de um local.

Assim como para o continente africano e o continente sul americano, populações de *P. pachyrhizi* do continente norte americano parecem ter origem genealogicamente diversa, e/ou que a diversidade genética atual nesse país é resultado de múltiplas introduções, como já havia sido evidenciado por Barnes *et al.*, (2008).

6 - Conclusões

- O fungo *Phakopsora pachyrhizi*, apesar de ser um fungo que aparentemente se reproduz apenas assexuadamente, apresenta uma alta variabilidade genética.
- Análises de populações de *P. pachyrhizi* revelaram que populações locais são altamente diversas.
- Populações de *P. pachyrhizi* de diferentes países e continentes são altamente similares.
- O vazio sanitário, adotado no Brasil, apesar de eficiente na redução de inóculo inicial em cada ano agrícola, não se revelou eficiente na redução da variabilidade genética de *P. pachyrhizi* entre os dois anos agrícolas analisados.
- Os dados obtidos reforçaram a hipótese do continente africano como fonte de origem de populações de *P. pachyrhizi* que infectam plantações de soja no Brasil.
- Os Estados Unidos possivelmente sofreram, ou sofrem, introduções de uredósporos de *P. pachyrhizi* de diferentes fontes de origem.

7 – Referências Bibliográficas

- Barnes CW, Szabo, LJ and Bowersox VC (2008) Identifying and quantifying *Phakopsora pachyrhizi* spores in rain. *Phytopathology*, 99:328-338.
- Bromfield KR (1984) Soybean Rust. St. Paul, MN: American Phytopathological Society.
- Christiano RSC and Scherm H (2007) Quantitative aspects of the spread of Asian soybean rust in the southeastern United States, 2005-2006. *Phytopathology*, 97:1428-1433.
- Clement M, Posada D and Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, 9:1657-1659.
- Doyle JJ and Doyle JL (1987) A rapid DNA isolation procedure for small amounts of fresh leaf tissue. *Phytochemical Bulletin*, 19:11-15.
- Excoffier L (2006) Arlequin ver 3.1: Computational and Molecular Population Genetics Lab CMPG. Zoological Institute, University of Berne.
- Freire MCM, de Oliveria LO, de Almeida, AMR, Schuter I, Moreira MA, Liebenberg MM and Mienie CMS (2008) Evolutionary history of *Phakopsora pachyrhizi* (the Asian soybean rust) in Brazil based on nucleotide sequences of the internal transcribed spacer region of the nuclear ribosomal DNA. *Genetics and Molecular Biology*, 31:920-931.
- Godoy C (2009) Começa o período de vazio sanitário da soja. <http://noticias.ambientebrasil.com.br/clipping/2009/06/15/46257-comeca-o-periodo-de-vazio-sanitario-da-soja.html> Acessado: 06 de julho de 2011.
- Goellner K, Loehrer M, Langenbach C, Conrath U, Koch E and Schaffrath U (2010) *Phakopsora pachyrhizi*, the causal agent of Asian soybean rust. *Molecular Plant Pathology*, 11:169-177.
- Hare MP (2001) Prospects for nuclear gene Phylogeography. *Trends in Ecology & Evolution*, 16:700-706.
- Hoefle C, Loehrer M, Schaffrath U, Frank M, Scultheiss H and Huckelhoven R (2009) Transgenic suppression of cell death limits penetration success of the soybean rust fungus *Phakopsora pachyrhizi* into epidermal cells of barley. *Phytopathology*, 99:220-226.

- Isard SA, Gage SH, Comtois P and Russo JM (2005) Principles of the atmospheric pathway for invasive species applied to soybean rust. *BioScience* 55:851-861.
- Librado P and Rozas J (2009) DNAsp V5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451-1452.
- Luster DG, McMahon MB, Carter ML, Fortis LL and Nuñez A (2010) Proteomic analysis of germinating urediniospores of *Phakopsora pachyrhizi*, causal agent of Asian soybean rust. *Proteomics*, 10:3549-3557.
- Mendes RK, Carvalhal RF, Stach-Machado DR and Kubota LT (2009) Surface Plasmon resonance immunosensor for early diagnosis of Asian rust on soybean leaves. *Biosensors and Bioelectronics*, 24: 2483-2487.
- Nagarajan S and Singh (1990) Long-distance dispersion of rust pathogens. *Annual Review of Phytopathology*, 28:139-153.
- Nunney L, Yuan X, Bromley R, Hartung J, Montero-Astúa M, Moreira L, Ortiz B and Stouthamer R (2010) Population genomic analysis of a bacterial plant pathogen: novel insight into the origin of pierce's disease of Grapevine in the U.S. *PLoS ONE*, 5:1-9.
- Ono Y, Buritica P and Hennen JF (1992) Delimitation of *Phakopsora*, *Physopella* and *Cerotelium* and their species on *Leguminosae*. *Mycological Research*, 96:825-850.
- Pandey AK, Yang C, Zhang C, Graham MA, Horstamn HD, Lee Y, Zabortina OA, Hill JH, Pandley KF and Whitham SA (2011) Functional analysis of the Asian soybean rust resistance pathway mediated by *Rpp2*. *Molecular Plant-Microbe Interactions*, 24:194-206.
- Sambrook J, Fritsch EF and Maniatis T (1989) *Molecular cloning, a laboratory manual*. 2nd ed. USA: Cold Spring Harbor Laboratory.
- Silva D, Yamanaka N, Brogin R, Arias C, Nepomuceno A, Di Mauro A, Pereira S, Nogueira L, Passianotto AL and Abdelnoor R (2008) Molecular mapping of two loci that confer resistance to Asian rust in soybean. *Theor Applied Genetics*, 117:57-63.
- Sinclair JB and Hartman GL (1996) *Soybean Rust Workshpo*. Urbana, Illinois. Proceedings.
- Soria-Guerra RE, Mendoza SR, Chang S, Haudenschild JS, Padmanaban A, Rodriguez-Zas S, Hartman GL, Ghabrial AS and Korban SS (2010) Transcriptome

analysis of resistant and susceptible genotypes of *Glycine tomentella* during *Phakopsora pachyrhizi* infection reveals novel rust resistance genes, 120:1315-1333.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonucleases mapping and DNA sequence III. Cladogram estimation. *Genetics*, 132:619-633.

Twizeyimana M, Ojiambo PS, Haudenshield JS, Caetano-Anollés G, Pedley KF, Bandyopadhyay R and Hartman GL (2011) Genetic structure and diversity of *Phakopsora pachyrhizi* isolates from soybean. *Plant Pathology*:1-11.

Twizeyimana M, Ojiambo PS, Sonder K, Ikotun T, Hartman GL and Bandyopadhyay R (2009) Pathogenic variation of *Phakopsora pachyrhizi* infecting soybean in Nigeria. *Phytopathology*, 99:353-361.

White TJ, Bruns T, Lee S and Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: M.A. *et al.* PCR protocols, Academic press, San Diego, CA.

Yamaoka Y, Fujiwara Y, Kakishima M, Katsuda K, Yamada K and Hagiwara H (2002) Pathogenic races of *Phakopsora pachyrhizi* on soybean and wild host plants collected in Japan. *Journal of Genetic Plant Pathology*, 68: 52-56.

Yorinori JT, Nunes Junior J and Lazzarotto JJ (2004) Ferrugem “asiática” da soja no Brasil: evolução, importância econômica e controle. Londrina: Embrapa Soja. 36p. (Documentos, 247).