

LEONARDO MAGALHÃES CRUZ



**CARACTERIZAÇÃO E ANÁLISE FILOGENÉTICA MOLECULAR
DE NOVOS ISOLADOS FIXADORES DE NITROGÊNIO**

Tese apresentada ao Curso de Pós-Graduação
em Bioquímica, Setor de Ciências Biológicas,
Universidade Federal do Paraná, como requeri-
mento parcial à obtenção do título de Doutor.

Orientador: Prof. Dr. Fábio de Oliveira Pedrosa

CURITIBA

2001

LEONARDO MAGALHÃES CRUZ

**CARACTERIZAÇÃO E ANÁLISE FILOGENÉTICA MOLECULAR
DE NOVOS ISOLADOS FIXADORES DE NITROGÊNIO**

Tese apresentada ao Curso de Pós-Graduação
em Bioquímica, Setor de Ciências Biológicas,
Universidade Federal do Paraná, como requeri-
mento parcial à obtenção do título de Doutor.

Orientador: Prof. Dr. Fábio de Oliveira Pedrosa

CURITIBA

2001

TERMO DE APROVAÇÃO

LEONARDO MAGALHÃES CRUZ

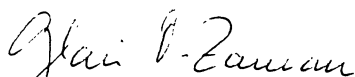
CARACTERIZAÇÃO E ANÁLISE FILOGENÉTICA MOLECULAR DE NOVOS ISOLADOS DE BACTÉRIAS FIXADORAS DE NITROGÊNIO

Tese aprovada como requisito parcial para obtenção do grau de Doutor no Curso de Pós-Graduação em Ciências-Bioquímica da Universidade Federal do Paraná, pela seguinte banca examinadora:

Orientador:



Prof. Dr. Fábio de Oliveira Pedrosa
Departamento de Bioquímica, UFPR



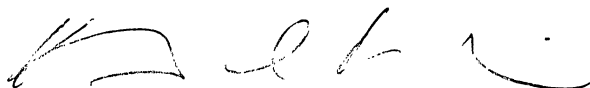
Dra.ª Glaci Zancan
Departamento de Bioquímica, UFPR



Prof.ª Dra.ª Maria Sueli Soares Felipe
Departamento de Biologia Celular, UnB



Prof. Dr. Sandro Luis Bonatto
Centro de Biologia Genômica e Molecular, PUC-RS



Prof. Dr. Humberto Maciel França Madeira
Centro de Ciências Agrárias e Ambientais, PUC-PR

A minha família, que sempre esteve presente. Em especial, aos meus pais, que sempre estiveram ao lado dos filhos, com especial carinho e atenção, não medindo esforços para que pudéssemos realizar nossos sonhos.

AGRADECIMENTOS

A Todos os colegas do Departamento, em especial, ao Humberto e Juliana Ramos, Guilherme Sasaki, Darlene Persuhn, Fabiane Rego, Cláudia Didonet e Elaine Benelli, pelas valiosas idéias, críticas, sugestões e discussões acadêmicas que tanto contribuíram para o meu trabalho e minha formação profissional.

Aos funcionários Valter Baura, Julieta Pie e Roseli Prado, insubstituíveis em suas tarefas e a tantos anos dedicados a um trabalho difícil e de fundamental importância ao laboratório, sempre preocupados com a qualidade de seus serviços e prontos a ajudar. E a quem estarei sempre em dívida e serei sempre grato pela participação direta no trabalho diário e pelo apoio dispensado.

Aos orientadores, Dr. Fábio Pedrosa e Dr. Emanuel de Souza pelo acompanhamento e revisão dos trabalhos ao longo destes anos.

Ao Dr. Geoffrey Yates e Dra. Glaci Zancan pela paciente e minuciosa avaliação deste trabalho.

A Olmar Weber, quem iniciou este trabalho e a quem tive o prazer de conhecer e consolidar uma grande amizade.

A todos os colegas da Embrapa/ CNPAB que me acompanharam no início desta jornada, em especial ao Dr. Ivo Baldani e à Dra. Kátia Teixeira, a quem devo o início do meu aprendizado.

A Darlene Persuhn e toda sua família, em especial ao Sr. Alcides e Sra. Rosilda Persuhn e Sônia Camati, pela amizade, apoio e carinho recebidos durante todos estes anos, fundamentais e especiais em todos os momentos, alegres ou difíceis.

A todos e em especial, a Lilian Noindorf, Luciano Huergo, Carolina Galvão, Mariana Araujo, Ana Cláudia Bonatto, Juliana Inaba, Rose Adele Monteiro e Gisele Klassen que, direta ou indiretamente, contribuíram para a realização deste trabalho.

SUMÁRIO

TERMO DE APROVAÇÃO	II
AGRADECIMENTOS	IV
SUMÁRIO	V
LISTA DE FIGURAS	VII
LISTA DE TABELAS	IX
LISTA DE SIGLAS	X
LISTA DE SÍMBOLOS	XII
RESUMO	XIII
ABSTRACT	XIV
1 INTRODUÇÃO	1
1.1 A IMPORTÂNCIA DO NITROGÊNIO PARA OS SERES VIVOS E A CONTRIBUIÇÃO DA FIXAÇÃO BIOLÓGICA DE NITROGÊNIO	1
1.2 BACTÉRIAS FIXADORAS DE NITROGÊNIO ASSOCIADAS A PLANTAS NÃO LEGUMINOSAS	7
1.2.1 DIAZOTRÓFICOS ENDOFÍTICOS	11
1.2.2 CONTRIBUIÇÃO DOS DIAZOTRÓFICOS ASSOCIATIVOS À FIXAÇÃO BIOLÓGICA DE NITROGÊNIO	15
1.2.3 OUTRAS CONTRIBUIÇÕES DOS DIAZOTRÓFICOS ASSOCIATIVOS	17
1.3 TAXONOMIA DE MICRORGANISMOS	18
1.4 INFERÊNCIA FILOGENÉTICA	23
1.4.1 OS MARCADORES FILOGENÉTICOS	25
1.4.1.1 O USO DO RNA RIBOSSOMAL PARA A RECONSTRUÇÃO FILOGENÉTICA	27
1.4.2 ALINHAMENTO DE SEQUÊNCIAS	30
1.4.3 DETERMINAÇÃO DO MODELO DE SUBSTITUIÇÃO	31
1.4.4 RECONSTRUÇÃO DA ÁRVORE FILOGENÉTICA	33
1.4.4.1 MÉTODO DA MÁXIMA VEROSSIMILHANÇA (<i>MAXIMUM LIKELIHOOD</i> - ML)	35
1.4.4.2 MÉTODO DA MÁXIMA PARCIMÔNIA (<i>MAXIMUM PARSIMONY</i> - MP)	36
1.4.4.3 MÉTODOS DE DISTÂNCIA (<i>DISTANCE METHODS</i>)	37
1.5 EVOLUÇÃO DA FIXAÇÃO BIOLÓGICA DE NITROGÊNIO E FILOGENIA DOS ORGANISMOS DIAZOTRÓFICOS	39
OBJETIVOS	43
2 MATERIAL E MÉTODOS	44
2.1 ESTIRPES USADAS	44
2.2 CONDIÇÕES DE CRESCIMENTO	46
2.3 EXTRAÇÃO DE DNA CROMOSSOMAL	46
2.4 INICIADORES USADOS NA AMPLIFICAÇÃO E SEQUENCIAMENTO DO 16S RDNA E REGIÃO INTERGÊNCIA 16S-23S RDNA	47
2.5 PREPARO DAS AMOSTRAS PARA REAÇÃO DA PCR A PARTIR DE CULTURAS	51
2.6 AMPLIFICAÇÃO DO 16S RDNA E REGIÃO INTERGÊNICA 16S-23S RDNA POR PCR	51
2.7 RESTRIÇÃO DO 16S RDNA E REGIÃO INTERGÊNICA 16S-23S RDNA	52
2.8 ANÁLISE DOS PERFIS DE RESTRIÇÃO	54
2.9 SEQUENCIAMENTO PARCIAL DO 16S RDNA	56
2.10 ANÁLISE DAS SEQUÊNCIAS	57
2.11 PROGRAMAS E MODELOS MATEMÁTICOS USADOS NA RECONSTRUÇÃO FILOGENÉTICA	62
3 RESULTADOS	71
3.1 AMPLIFICAÇÃO DO 16S RDNA E DA REGIÃO INTERGÊNICA 16S-23S RDNA	71
3.2 POLIMORFISMO DE FRAGMENTOS DE COMPRIMENTO ALEATÓRIO (RFLP) DA REGIÃO INTERGÊNICA 16S-23S RDNA	74
3.3 ANÁLISE DE RESTRIÇÃO DO DNA RIBOSSOMAL DE 16S AMPLIFICADO (ARDRA)	79

3.4 SEQÜENCIAMENTO DO 16S RDNA	94
3.5 ANÁLISE COMPARATIVA DE SEQÜÊNCIAS E PESQUISA DE HOMOLOGIA	97
3.6 RECONSTRUÇÃO FILOGENÉTICA	110
4 DISCUSSÃO	122
4.1 ANÁLISE DE RESTRIÇÃO DO 16S RDNA E REGIÃO INTERGÊNICA 16S-23S RDNA	122
4.2 ANÁLISE DAS SEQÜÊNCIAS E ESTRUTURA SECUNDÁRIA	125
4.3 ANÁLISE COMPARATIVA ENTRE OS MÉTODOS	131
5 CONCLUSÕES	136
6 REFERÊNCIAS	137
APÊNDICES	148

LISTA DE FIGURAS

FIGURA 1 - CICLO BIOGEOQUÍMICO DO NITROGÊNIO.....	3
FIGURA 2 - LOCALIZAÇÃO DOS INICIADORES USADOS PARA O SEQÜENCIAMENTO DA REGIÃO Y1-Y3 DO 16S RDNA.....	50
FIGURA 3 - MONTAGEM DA SEQÜÊNCIA CONTÍGUA DO ISOLADO DE BANANEIRA BA15 NO PROGRAMA CAP.....	58
FIGURA 4 - PRODUTO DE AMPLIFICAÇÃO POR PCR DA REGIÃO Y1-Y3 DO 16S RDNA DE ISOLADOS DIAZOTRÓFICOS DE BANANEIRA.....	72
FIGURA 5. A E B. PRODUTOS DE AMPLIFICAÇÃO DA REGIÃO INTERGÊNICA 16S- 23S RDNA DE ISOLADOS DIAZOTRÓFICOS E ESTIRPES DE <i>H.</i> <i>SEROPEDICAE</i> E <i>H. RUBRISUBALBICANS</i>	73
FIGURA 6 - A E B: PERFIL DE RESTRIÇÃO DOS AMPLIFICADOS DA REGIÃO INTERGÊNICA 16S-23S RDNA COM <i>ALUI</i>	76
FIGURA 8 - A E B: PERFIL DE RESTRIÇÃO DOS AMPLIFICADOS DA REGIÃO INTERGÊNICA 16S-23S RDNA COM <i>RS4I</i>	78
FIGURA 9 - DE A A G: PERFIS DE RESTRIÇÃO PARA ISOLADOS DE ABACAXIZEIRO E BANANEIRA E ESTIRPES DE <i>HERBASPIRILLUM</i> SP. E <i>BURKHOLDERIA</i> SP. COM A ENDONUCLEASE <i>ALUI</i> NA REGIÃO AMPLIFICADA Y1-Y3 DO 16S RDNA.....	83
FIGURA 10 - DE A A H: PERFIS DE RESTRIÇÃO PARA ISOLADOS DE ABACAXIZEIRO E BANANEIRA E ESTIRPES DE <i>HERBASPIRILLUM</i> SP. E <i>BURKHOLDERIA</i> SP. COM A ENDONUCLEASE <i>HAETIII</i> NA REGIÃO AMPLIFICADA Y1-Y3 DO 16S RDNA.....	86
FIGURA 11 - DE A A G: PERFIS DE RESTRIÇÃO PARA ISOLADOS DE ABACAXIZEIRO E BANANEIRA E ESTIRPES DE <i>HERBASPIRILLUM</i> SP. E <i>BURKHOLDERIA</i> SP. COM A ENDONUCLEASE <i>HINFI</i> NA REGIÃO AMPLIFICADA Y1-Y3 DO 16S RDNA.....	89
FIGURA 12 - DE A A F: PERFIS DE RESTRIÇÃO PARA ISOLADOS DE ABACAXIZEIRO E BANANEIRA E ESTIRPES DE <i>HERBASPIRILLUM</i> SP. E <i>BURKHOLDERIA</i> SP. COM A ENDONUCLEASE <i>RS4I</i> NA REGIÃO AMPLIFICADA Y1-Y3 DO 16S RDNA.....	91
FIGURE 13 - DENDROGRAMA INFERIDO A PARTIR DOS PERFIS DE RESTRIÇÃO OBTIDOS PARA A DIGESTÃO DOS PRODUTOS DE AMPLIFICAÇÃO DO 16S RDNA DE ISOLADOS DE ABACAXIZEIRO E BANANEIRA E ESTIRPES DE <i>H. SEROPEDICAE</i> , <i>H. RUBRISUBALBICANS</i> , " <i>B.</i> <i>BRASILENSIS</i> " E " <i>B. TROPICALIS</i> " COM <i>ALUI</i> , <i>HAETIII</i> E <i>HINFI</i> MOSTRADOS NA TABELA 9.....	93
FIGURA 14 - OBSERVAÇÃO DE LEITURA AMBÍGUA NA SEQÜÊNCIA DA REGIÃO Y1-Y2 DO GRUPO IX. COMPOSTO PELOS ISOLADOS AB48 E AB71.....	96
FIGURA 15 - MATRIZ MOSTRANDO A PROPORÇÃO DE IDENTIDADE (DIAGONAL SUPERIOR) E O NÚMEROS DE BASES DIFERENTES (DIAGONAL INFERIOR) ENTRE PARES DE SEQÜÊNCIAS DE 16S RDNA.....	99
FIGURA 16 - ÁRVORE FILOGENÉTICA RECONSTRUÍDA A PARTIR DE 219 SEQÜÊNCIAS DO 16S RDNA DE ISOLADOS DE BANANEIRA E ABACAXIZEIRO E SEQÜÊNCIAS HOMÓLOGAS, OBTIDAS POR PESQUISA COM O PROGRAMA BLASTN.....	114
FIGURA 17 - DETALHAMENTO DA FILOGENIA DO GÊNERO <i>BURKHOLDERIA</i> E ESPÉCIES RELACIONADAS APRESENTADA NA ÁRVORE FILOGENÉTICA DA FIGURA 21.....	115
FIGURA 18 - DETALHAMENTO DA FILOGENIA DO GÊNERO <i>HERBASPIRILLUM</i> E DOS ORGANISMOS RELACIONADOS APRESENTADA NA ÁRVORE FILOGENÉTICA DA FIGURA 21.....	116

FIGURA 19 - DETALHAMENTO DA FILOGENIA DA FAMÍLIA COMAMONADACEAE E DOS ORGANISMOS RELACIONADOS APRESENTADA NA ÁRVORE FILOGENÉTICA DA FIGURA 21	117
FIGURA 20 - ÁRVORE FILOGENÉTICA DA REGIÃO Y1-Y3 DO 16S RDNA DE ISOLADOS DE BANANEIRA E ABACAXIZEIRO E ESPÉCIES RELACIONADAS.....	118
FIGURA 21 - ÁRVORE FILOGENÉTICA DA REGIÃO Y1-Y3 DO 16S RDNA DE <i>BURKHOLDERIA</i> SP.....	119
FIGURA 22 - ÁRVORE FILOGENÉTICA DA REGIÃO Y1-Y3 DO 16S RDNA DO ISOLADO BA128 E ESPÉCIES DE COMAMONADACEAE RELACIONADAS.....	120
FIGURA 23 - ÁRVORE FILOGENÉTICA DA REGIÃO Y1-Y2 DO 16S RDNA MOSTRANDO O POSICIONAMENTO DOS ISOLADOS BA131 E ALA ENTRE ESPÉCIES DE ALFA PROTEOBACTÉRIAS.....	121
FIGURA 24 - ESTRUTURA SECUNDÁRIA DO 16S RRNA DE REPRESENTANTES DAS CINCO SUBDIVISÕES DAS PROTEOBACTÉRIAS (DE RIJK ET AL., 2000)	128
FIGURA 25 - ALINHAMENTO ENTRE OS GRUPOS DE SEQUÊNCIAS Y1-Y2 E SEQUÊNCIAS CONSENSO PARA AS SUBDIVISÕES DAS PROTEOBACTÉRIAS (INDICADO NO TÍTULO À ESQUERDA).....	129
FIGURA 26 -MAPEAMENTO DOS SÍTIOS DE RESTRIÇÃO PARA <i>ALU1</i> NA SEQUÊNCIA DA REGIÃO Y1-Y3 DE <i>H. SEROPEDICAE</i> E DOS ISOLADOS BA15, BA17, BA153 E X8.....	135
FIGURA 27 - REPRESENTAÇÃO DOS ELEMENTOS BÁSICOS DE UMA ÁRVORE.....	156

LISTA DE TABELAS

TABELA 1 - ALGUMAS ESPÉCIES DE RIZÓBIOS E SUAS LEGUMINOSAS HOSPEDEIRAS.....	6
TABELA 2 - ALGUMAS BACTÉRIAS FIXADORAS DE NITROGÊNIO ASSOCIATIVAS E SUAS PLANTAS HOSPEDEIRAS.....	10
TABELA 3 - BACTÉRIAS DIAZOTRÓFICAS ISOLADAS DE ABACAXIZEIRO (<i>ANANAS COMOSUS</i>) E BANANEIRA (<i>MUSA SPP.</i>) CULTIVADOS NOS ESTADOS DA BAHIA (BA) E RIO DE JANEIRO (RJ).....	45
TABELA 4 - PARÂMETROS FÍSICOS DOS INICIADORES UTILIZADOS NA AMPLIFICAÇÃO E SEQÜENCIAMENTO DO 16S RDNA E REGIÃO INTERGÊNICA 16S-23S RDNA.....	49
TABELA 5 - ENDONUCLEASES DE RESTRIÇÃO USADAS NAS ANÁLISES DE RFLP DO 16S RDNA (ARDRA) E 16S-23S RDNA.....	53
TABELA 6 - PROGRAMAS USADOS NAS ANÁLISES DE ARDRA E SEQÜENCIAMENTO DO 16S RDNA DE ISOLADOS DE BANANEIRA E ABACAXIZEIRO E ESPÉCIES RELACIONADAS.....	55
TABELA 7 - LISTA DOS ORGANISMOS PARA OS QUAIS A SEQÜÊNCIA DO 16S RDNA FORAM USADAS NAS ANÁLISES FILOGENÉTICAS.....	65
TABELA 8 - CARACTERIZAÇÃO DE BACTÉRIAS DIAZOTRÓFICAS POR RFLP DA REGIÃO INTERGÊNICA 16S-23S RDNA.....	75
TABELA 9 - CARACTERIZAÇÃO DE BACTÉRIAS DIAZOTRÓFICAS, POR ARDRA, DA REGIÃO Y1-Y3 DO 16S RDNA.....	81
TABELA 10 - CARACTERIZAÇÃO DE BACTÉRIAS DIAZOTRÓFICAS POR SEQÜENCIAMENTO DAS REGIÕES Y1-Y2 E Y1-Y3 DO 16S RDNA.....	95
TABELA 11 - ORGANISMOS QUE APRESENTARAM MAIOR HOMOLOGIA PARA A REGIÃO Y1-Y2 OU Y1-Y3 DA SEQÜÊNCIA DO 16S RDNA DE ISOLADOS DIAZOTRÓFICOS DE ABACAXIZEIRO E BANANEIRA, DE ACORDO COM PESQUISA REALIZADA COM O PROGRAMA BLASTN ^{r1)}	102

LISTA DE SIGLAS

16S rDNA - *16S ribosomal Deoxyribonucleic Acid*

23S rDNA - *23S ribosomal Deoxyribonucleic Acid*

16S rRNA - *16S ribosomal Ribonucleic Acid*

23S rRNA - *23S ribosomal Ribonucleic Acid*

AIA - *Ácido Indol Acético*

AFLP - *Amplified Fragment Length Polymorphism*

ARDRA - *Amplified 16S ribosomal DNA restriction analysis*

ATPase - *Adenosine Triphosphate synthase*

CAP - *Contig Assembly Program*

CNPAB - *Centro Nacional de Pesquisa de Agrobiologia*

dATP - *deoxiadenosina trifosfato*

dCTP - *deoxicitidina trifosfato*

dGTP - *deoxiguanosina trifosfato*

DNA - *Deoxyribonucleic Acid*

dTTP - *deoxitimidina trifosfato*

EDTA - *Ethylene Diamine Tetra-Acetate*

EF-Tu - *Elongation Factor Tu*

Embrapa - *Empresa Brasileira de Pesquisa Agropecuária*

ERIC - *Enterobacterial Repetitive Intergeneric Consensus*

GET - *Glicose/ EDTA/ Tris*

LS - *Least Square*

ME - *Minimum Evolution*

ML - *Maximum Likelihood*

MP - *Maximum Parsimony*

NJ - *Neighbor-Joining*

P. A. - *Puro para Análise*

Pb - pares de bases

PCR - *Poymerase Chain Reaction*

PFGE - *Pulsed Field Gel Electrophoresis*

PGR - *Plant Growth Regulators*

PHYLIP - *Phylogeny Inference Package*

PM - **Peso Molecular**

RAPD - *Random Amplified Polymorphic DNAs*

REP - *Repetitive Extragenic Palindromic*

RFLP - *Restriction Fragment Length Polymorphism*

RISA - *rRNA Intergenic Spacer Analysis*

RNA - *Ribonucleic Acid*

SDS - *Sodium Dodecyl Sulfate*

SSCP - *Single-Stranded Conformational Polymorphism*

T_m - *melting Temperature*

tRNA - *Transfer Ribonucleic Acid*

TSR - *Template Supression Reagent*

U - **Unidade enzimática**

UPGMA - *Unweighted Pair Group Method Using Arithmetic Averages*

UV - **Ultra Violeta**

LISTA DE SÍMBOLOS

A - adenosina

°C - graus Celsius

C - citidina

Ca - cálcio

Cl - cloro

Fe - ferro

g - grama

G - guanosina

H - hidrogênio

ha - hectare

K - potássio

L - litro

Mg - magnésio

μ - micro (10^{-6})

N- nitrogênio

Na - sódio

O - oxigênio

p - pico (10^{-12})

P - fósforo

S - enxofre

T - timidina

RESUMO

Bactérias diazotróficas isoladas de bananeira e abacaxizeiro, nos estados da Bahia e Rio de Janeiro, foram avaliados quanto ao grau de parentesco com estirpes referência de diversas bactérias fixadoras de nitrogênio. Análises prévias destes isolados, com base em dados morfológicos, fisiológicos e bioquímicos, permitiram separá-los em alguns grupos com semelhanças a espécies de *Herbaspirillum seropedicae*, *Herbaspirillum rubrisubalbicans*, *Herbaspirillum* sp., *Burkholderia* sp. e um grupo independente, sem semelhança aparente aos gêneros estudados. A caracterização molecular baseada na homologia do gene para o RNA ribossomal de 16S (16S rDNA) e região intergênica dos genes para os RNAs ribossomais de 16S e 23S (região intergênica 16S-23S rDNA) foi realizada no presente trabalho a partir das técnicas de Polimorfismo de Fragmentos de Comprimento Aleatório amplificados (PCR-RFLP), aplicada ao 16S rDNA e 16S-23S rDNA e seqüenciamento de DNA, aplicada ao 16S rDNA. Trinta e oito isolados foram divididos em quatro grupos identificados entre os gêneros *Herbaspirillum* e *Burkholderia* e oito grupos não identificados por PCR-RFLP. Análise de homologia e filogenética das seqüências de 16S rDNA, confirmaram a análise anterior e posicionou sete dos grupos não identificados entre os gêneros *Burkholderia*, *Azospirillum* e *Ochrobactrum* e um grupo entre a família das Comamonadaceas.

ABSTRACT

Diazotrophic bacteria isolated from banana trees and pineapple plants, from Bahia and Rio de Janeiro states were compared with reference strains of nitrogen-fixing bacteria. Previous analysis of these isolates, based on morphological, physiological, and biochemical data grouped them with *Herbaspirillum seropedicae*, *Herbaspirillum rubrisubalbicans*, *Herbaspirillum* sp., *Burkholderia* sp., and an unidentified group. Molecular characterization based on the homology of the 16S ribosomal RNA gene (16S rDNA) and the intergenic region of 16S and 23S ribosomal RNA genes (16S-23S rDNA intergenic region) was carried out in the present work by Restriction Fragment Length Polymorphism of amplified DNA (PCR-RFLP), applied to the 16S rDNA and 16S-23S rDNA intergenic region and DNA sequencing of the 16S rDNA. Thirty eight (38) isolates were divided into four groups identified in the *Herbaspirillum* and *Burkholderia* genus and eight unidentified groups by PCR-RFLP techniques. Homology and phylogenetic analysis from 16S rDNA sequences confirmed the previous results and placed seven (7) of the unidentified groups into *Burkholderia*, *Azospirillum* and *Ochrobactrum* genera and one (1) into the Comamonadaceae family.

1 INTRODUÇÃO

1.1 A IMPORTÂNCIA DO NITROGÊNIO PARA OS SERES VIVOS E A CONTRIBUIÇÃO DA FIXAÇÃO BIOLÓGICA DE NITROGÊNIO

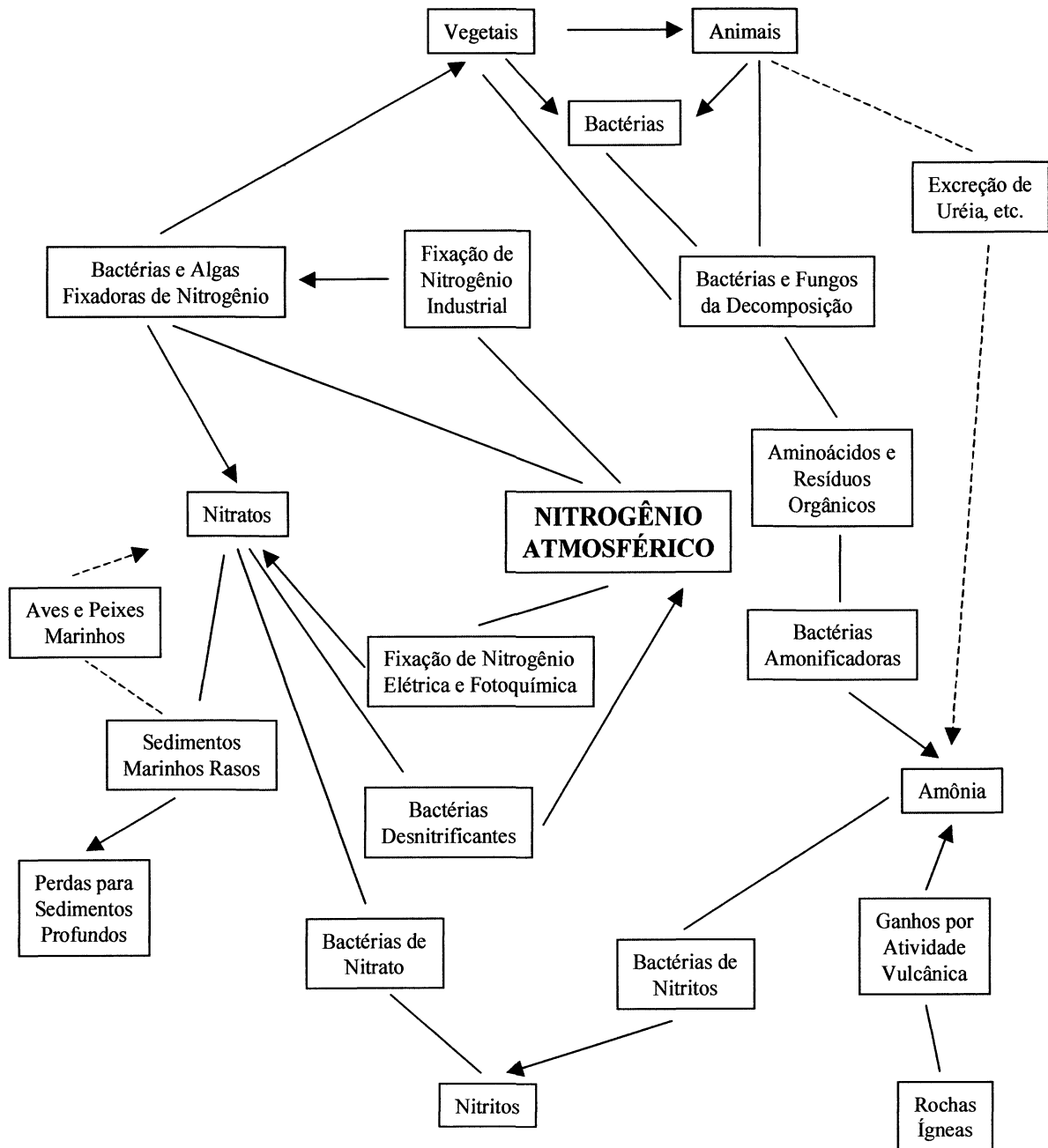
Comparando-se a composição química de uma célula com a composição química da Terra, é possível perceber que uma célula viva é uma estrutura altamente ordenada e não simplesmente uma mistura casual dos elementos químicos encontrados na Terra. A célula constitui um seletivo sistema químico composto primariamente de C, H, O, N, S e P, os principais elementos da vida. Ao mesmo tempo uma célula viva representa um sistema aberto, onde ocorre troca de massa e energia com o ambiente para manter a estrutura celular (MADIGAN *et al.*, 1997). Sendo assim, os elementos químicos essenciais à célula tendem a circular entre a biosfera e o ambiente em vias mais ou menos circulares, denominados CICLOS BIOGEOQUÍMICOS. Os ciclos, por sua vez, são compostos por um RESERVATÓRIO, geralmente não biológico, onde o movimento é lento e contém a maior parte do elemento e por um componente LÁBIL ou de CICLAGEM, composto por uma parcela menor, porém mais ativa que se permuta rapidamente entre os organismos e seu ambiente imediato. Dentre os seis elementos citados, o fósforo e o nitrogênio muitas vezes são os maiores responsáveis pelo limite e controle da abundância dos organismos em um ecossistema (ODUM, 1988) e por isso seus ciclos biogeoquímicos assumem um importante papel na manutenção da vida na Terra.

O fósforo apresenta um ciclo do TIPO SEDIMENTAR, por possuir seu reservatório localizado na crosta terrestre, relativamente simples e menos complexamente regulado. Por outro lado, o ciclo do nitrogênio é do TIPO GASOSO, com um reservatório localizado na atmosfera, muito mais complexo e bem tamponado (ODUM, 1988), figura 1. A atmosfera terrestre contém aproximadamente 78% do nitrogênio, armazenado na sua forma molecular (N_2). Entretanto, nesta forma o

nitrogênio é praticamente inerte, não podendo ser assimilado pela maioria dos seres vivos. Para sua incorporação ao componente lábil do seu ciclo é preciso que ele seja reduzido a compostos mais reativos (FIXADO), o que requer a quebra da ligação triplice entre os átomos de N, a um alto custo energético.

A fixação do nitrogênio produz amônia (NH_3), mas diversos outros compostos, como N_2O , NO , NO_2^- e NO_3^- também estão presentes e são constantemente ciclados nos ecossistemas. O nitrogênio pode ser fixado a partir de diversos processos naturais abióticos, tais como descargas elétricas, combustão e ação vulcânica. Entretanto estes processos contribuem com não mais do que 10% do nitrogênio fixado anualmente. Um outro processo, responsável por aproximadamente 25% da fixação total do nitrogênio anual, é a fixação artificial (NEWTON, 1999), feita industrialmente, principalmente para a produção de fertilizantes nitrogenados destinados à agricultura e com consumo de combustíveis fósseis. Um terceiro processo de fixação do nitrogênio é realizado por organismos vivos e por isso denominada FIXAÇÃO BIOLÓGICA DE NITROGÊNIO e contribui com cerca de 65% do total do nitrogênio fixado (NEWTON, 1999), sendo, portanto, o grande responsável pela disponibilização deste elemento para uso dos seres vivos. Apesar disto, o nitrogênio fixado e utilizado em todas as suas formas, compreende menos do que 0,0007% do nitrogênio na Terra e na sua atmosfera.

FIGURA 1 - CICLO BIOGEOQUÍMICO DO NITROGÊNIO.



Dada a importância do nitrogênio na composição química de elementos reguladores e estruturais, fundamentais à manutenção da vida e a escassez de formas prontamente absorvíveis pelas plantas, a Fixação Biológica de Nitrogênio é um processo de fundamental importância para a sustentabilidade da biosfera, sendo realizado exclusivamente por organismos procarióticos. Embora nenhum organismo eucariótico com a capacidade de fixar nitrogênio tenha sido encontrado, eles frequentemente se beneficiam de associações íntimas com procariontes diazotróficos (XIA *et al.*, 1994). Grande parte da Fixação Biológica de Nitrogênio é feita por bactérias simbióticas, o que permite a pronta disponibilização do nitrogênio para as plantas hospedeiras. Em última análise, o homem talvez seja o maior beneficiado dessas associações. Na agricultura, o nitrogênio é quantitativamente o elemento mais importante e o que mais frequentemente limita o crescimento vegetal, e associações entre plantas cultivadas com diazotróficos são frequentemente exploradas na agricultura.

A capacidade de fixar nitrogênio não é uma característica particular apresentada por organismos bem relacionados entre si e que formam um grupo taxonômico ou filogenético bem estruturado. Ao contrário, é possível observá-la nos mais distintos grupos taxonômicos entre os procariotos (YOUNG, 1992). Na maioria das vezes o grau de relacionamento entre organismos fixadores de nitrogênio, também denominados diazotróficos, muitas vezes, não vai além da sua capacidade fisiológica de crescer em ambientes pobres em nitrogênio. Portanto, qualquer tentativa de classificação dos diazotróficos leva em conta aspectos ecológicos. É comum sua divisão em organismos de vida livre, que vivem no solo ou em ambientes aquáticos; organismos associativos, que possuem uma interação íntima com plantas, vivendo na rizosfera ou no seu interior; e organismos simbióticos que, como era de se esperar, apresentam um complexo mecanismo de interação com suas plantas hospedeiras, culminando com o benefício de ambas as partes.

É justamente a partir dos diazotróficos simbióticos que o homem consegue beneficiar-se diretamente da Fixação Biológica de Nitrogênio. Dentre este grupo de organismos incluem-se aqueles genericamente denominados de rizóbios, que compreendem atualmente 35 espécies distribuídas entre os gêneros *Rhizobium*, *Sinorhizobium*, *Mesorhizobium*, *Bradyrhizobium* e *Azorhizobium* (YOUNG, 1996; DSMZ, 2001) e são capazes de formar nódulos em uma gama enorme de plantas leguminosas, inclusive aquelas com grande importância na agricultura (Tabela 1). Esta relação planta-bactéria constitui uma simbiose perfeita, com as plantas fornecendo energia, esqueleto carbônico e um ambiente restrito de oxigênio (nódulos), e com as bactérias convertendo o nitrogênio atmosférico a uma forma utilizável pela planta. A importância da simbiose entre *Bradyrhizobium japonicum* e as variedades de soja comerciais para a agricultura chega ao ponto de não haver necessidade de nenhum aporte de fertilizante nitrogenado em solos cultivados com variedades plantadas no Brasil e Argentina, estando o fornecimento de nitrogênio para a planta a cargo das bactérias simbiotas. Estima-se que esta contribuição represente cerca de 2,5 milhões de toneladas de nitrogênio vindo do ar, com uma economia para o país em torno de US\$ 1,8 bilhão anuais por ser desnecessário o uso de adubos nitrogenados (DÖBEREINER *et al.*, 1995).

TABELA 1 - ALGUMAS ESPÉCIES DE RIZÓBIOS E SUAS LEGUMINOSAS HOSPEDEIRAS.

espécie	hospedeiro
<i>Azorhizobium caulinodans</i>	<i>Sesbania rostrata</i>
<i>Bradyrhizobium elkani</i>	<i>Glycine max</i> , <i>Glycine soja</i> e <i>Macroptilium</i>
<i>Bradyrhizobium japonicum</i>	<i>Glycine max</i> , <i>Glycine soja</i> e <i>Macroptilium atropurpureum</i>
<i>Bradyrhizobium sp.</i>	<i>Vigna</i> , <i>Lupinus</i> , <i>Ornithopus</i> , <i>Arachis</i> , <i>Cicer</i> , <i>Sesbania</i> , <i>Leucaena</i> , <i>Desmodium</i> , <i>Mimosa</i> , <i>Lab lab</i> , <i>Acacia</i> , <i>Stylosanthes</i> , <i>Glycine</i> , <i>Macroptilium</i> , <i>Lotus</i> , <i>Phaseolus</i> e <i>Parasponia</i>
<i>Mesorhizobium ciceri</i>	<i>Cicer arietinum</i>
<i>Mesorhizobium huakuii</i>	<i>Astragalus sinicus</i>
<i>Mesorhizobium loti</i>	<i>Lotus spp.</i> , <i>Lupinus</i> , <i>Anthyllis</i> , <i>Cicer</i> , <i>Ornithopus</i> , <i>Caragna</i> , <i>Leucaena</i> , <i>Mimosa</i>
<i>Rhizobium etli</i>	<i>Phaseolus vulgaris</i>
<i>Rhizobium galegae</i>	<i>Galega officinalis</i> , <i>Galega orientalis</i>
<i>Rhizobium leguminosarum</i>	<i>Vicia fava</i> , <i>Pisum sativum</i> , <i>Trifolium</i> , <i>Lathyrus</i> , <i>Lens</i> , <i>Cicer</i> e <i>Phaseolus vulgaris</i>
<i>Rhizobium lupini</i>	<i>Lupinus</i> , <i>Ornithopus</i>
<i>Rhizobium tropici</i>	<i>Phaseolus vulgaris</i> , <i>Macroptilium</i> e <i>Leucaena</i>
<i>Sinorhizobium fredii</i>	<i>Glycine max</i> e <i>Glycine soja</i>
<i>Sinorhizobium meliloti</i>	<i>Medicago</i> , <i>Melilotus</i> , <i>Trigonella</i>
<i>Sinorhizobium saheli</i>	<i>Acacia</i> e <i>Sesbania</i>
<i>Sinorhizobium teranga</i>	<i>Acacia</i> e <i>Sesbania</i>

FONTE: MARTINEZ-ROMERO. E.: CABALLERO-MELLADO. J. *Rhizobium* phylogenies and bacterial genetic diversity. **Critical Reviews in Plant Sciences**, v. 15, n. 2, p. 113-140. 1996.

SIQUEIRA. J. O. & FRANCO. A. A. Fixação biológica de nitrogênio. In: _____. **Biotecnologia do solo - fundamentos e perspectivas**. Brasília: NAGY Ltda.. 1988. p. 179-216.

WATANABE. I. **Biological nitrogen fixation and its use in agriculture (outline)**. Disponível em: <<http://www.asahi-net.or.jp/~it6i-wtnb/BNF.HTM/>> Acesso em: 02 jun. 2000. Cantho Univ.

NOTA: Dados trabalhados pelo autor.

1.2 BACTÉRIAS FIXADORAS DE NITROGÊNIO ASSOCIADAS A PLANTAS NÃO LEGUMINOSAS

Outra importante contribuição da Fixação Biológica de Nitrogênio para a agricultura e o ambiente está na associação de bactérias diazotróficas com gramíneas. Apesar destes organismos estarem intimamente associados às plantas, não há formação de estruturas especializadas como os nódulos nas leguminosas e, até o presente momento, não foi possível mostrar de forma clara que estes diazotróficos são capazes de beneficiar suas plantas hospedeiras, transferindo diretamente o nitrogênio fixado para a planta.

O isolamento da bactéria associativa *Beijerinckia fluminensis* da rizosfera de plantas de cana-de-açúcar na década de 1950 (DÖBEREINER & RUSCHEL, 1958) abriu a perspectiva para o uso da Fixação Biológica de Nitrogênio em plantas não leguminosas. Mas, até a década de 1970 somente diazotróficos, na sua maioria de vida livre, capazes de crescer e fixar nitrogênio sob tensão atmosférica de oxigênio (devido a presença de mecanismos protetores da nitrogenase como a alta taxa de respiração e produção de muco) foram descobertos. Estes organismos pertencem aos gêneros *Azotobacter*, *Azomonas*, *Beijerinckia* e *Derrxia* (DÖBEREINER *et al.*, 1995). Foi somente após a redescoberta do gênero *Azospirillum* em 1975 por Döbereiner & Day que cientistas do mundo inteiro passaram a se interessar por bactérias diazotróficas associadas com gramíneas (BALDANI *et al.*, 1997). Esta bactéria, por não possuir os mecanismos necessários para a proteção da nitrogenase, não é capaz de crescimento diazotrófico sob tensões atmosféricas de O₂. A descoberta de organismos diazotróficos mais eficientes no uso de fontes de carbono para a fixação de N₂, que não possuem tais mecanismos, só foi possível com a elucidação dos mecanismos de funcionamento da nitrogenase e o emprego de novos métodos de isolamento (DÖBEREINER *et al.*,

1995), como por exemplo, o desenvolvimento de meio de cultura semi-sólido, onde a concentração de oxigênio não prejudica a atividade redutora de nitrogênio da enzima, permitindo o isolamento de diazotróficos microaerofílicos em meio livre de nitrogênio. Esta técnica permitiu o isolamento de novas espécies de bactérias fixadoras de nitrogênio, pertencentes aos gêneros *Azospirillum*, *Herbaspirillum*, *Azoarcus*, *Gluconacetobacter* e *Burkholderia* (KIRCHHOF *et al.*, 1997b). A descoberta deste novo grupo de bactérias despertou um grande interesse da comunidade científica mundial por dois principais motivos: estarem associados principalmente com gramíneas, especialmente aquelas largamente cultivadas nos trópicos como milho, sorgo, cana-de-açúcar, diversas forrageiras, etc. e serem, na sua maioria, organismos endofíticos, ocorrendo no interior dos tecidos das plantas hospedeiras.

Recentemente, um novo passo foi dado nas pesquisas em torno da Fixação Biológica de Nitrogênio com a descoberta de associações de diazotróficos com outros grupos de plantas de interesse econômico, ampliando o conjunto de plantas hospedeiras e aumentando a importância da Fixação Biológica de Nitrogênio na produção agrícola. JIMÉNEZ-SALGADO *et al.* (1997), isolaram a bactéria endofítica *Gluconacetobacter diazotrophicus* (GILLIS *et al.*, 1989) YAMADA *et al.*, 1997 de plantas de café (*Coffea arabica*), cultura de inegável importância em nossa economia; *G. diazotrophicus* e *Azospirillum* foram isolados de batata doce (*Ipomoea* spp.); *Burkholderia* e outras espécies de mandioca (WEBER, 1998); FERREIRA *et al.* (1995) isolaram bactérias dos gêneros *Azospirillum* e *Herbaspirillum* de dendezeiro (*Elaeis guineensis*) e pupunheira (*Bactris gasipaes*), palmeiras com potencial para produzir mais de 12 t. ha⁻¹ de óleo. Espécies de *Azospirillum* foram também isoladas de fruteiras, como abacaxizeiro (*Ananas comosus*), bananeira (*Musa* spp.) e outras (SUBBA RAO, 1983). Em abacaxizeiro foi isolada ainda a espécie *Gluconacetobacter diazotrophicus* (TAPIA-HERNANDEZ *et al.*, 2000). WEBER (1998) identificou também os gêneros *Herbaspirillum* e *Burkholderia*, além de *Azospirillum*, em plantas de abacaxizeiro e bananeira. Outras bactérias fixadoras de nitrogênio associadas foram

identificadas, mas provavelmente devido a seu baixo número, ou ocorrência restrita, não são bem exploradas (BALDANI *et al.*, 1997). O interesse na associação de diazotróficos com gramíneas reforça a importância do processo de Fixação Biológica de Nitrogênio para a agricultura sustentável no Brasil, onde baixas quantidades de fertilizantes nitrogenados são aplicados (DÖBEREINER, 1995).

TABELA 2 - ALGUMAS BACTÉRIAS FIXADORAS DE NITROGÊNIO ASSOCIATIVAS E SUAS PLANTAS HOSPEDEIRAS.

ESPÉCIE DE DIAZOTRÓFICO	HOSPEDEIRO	PARTE DO HOSPEDEIRO
<i>Azospirillum amazonense</i>	arroz, trigo, cana-de-açúcar, palmeira	raiz, caule, sementes, frutos
<i>Azospirillum brasilense</i>	arroz, trigo, <i>Panicum maximum</i> , <i>Digitaria</i> spp., <i>Brachiaria</i> spp., cana-de-açúcar, palmeira	raiz, caule, sementes, folhas
<i>Azospirillum irakense</i>	arroz	raiz
<i>Azospirillum lipoferum</i>	cereais, gramíneas forrageiras, cana-de-açúcar, palmeiras, tubérculos	raiz, caule, sementes, folhas, frutos, tubérculos
<i>Klebsiella</i> sp.	milho	caule
<i>Gluconacetobacter diazotrophicus</i>	cana-de-açúcar, <i>Pennisetum purpureum</i> (Cameroon), batata doce, café, chá, banana, insetos	raiz, caule, folhas, tubérculo, rizosfera, mecanismo de alimentação
<i>Alcaligenes faecalis</i>	arroz	raiz
<i>Azoarcus</i> spp.	capim <i>Kalar</i>	raiz, base do caule
<i>Burkholderia</i> spp.	cereais, Cassava, banana, batata doce	caule, folhas, raiz, tubérculo
<i>Herbaspirillum frisingense</i>	<i>Spartina pectinata</i> , <i>Miscanthus sinensis</i> , <i>Miscanthus sacchariflorus</i> , <i>Pennisetum purpureum</i>	raiz, caule, folhas
<i>Herbaspirillum rubrisubalbicans</i>	arroz, cana-de-açúcar	raiz, caule, folhas
<i>Herbaspirillum seropedicae</i>	<i>Zea mays</i> , <i>Oryza sativa</i> , <i>Sorghum bicolor</i> , <i>Brachiaria decumbens</i> , <i>Digitaria decumbens</i> , <i>Melinis minutiflora</i> , cereais, cana-de-açúcar, gramíneas forrageiras, palmeiras	raiz, caule, folhas, sementes

FONTE: BALDANI, J. I. *et al.* Emended description of *Herbaspirillum*; inclusion of [*Pseudomonas*] *rubrisubalbicans*, a mild plant pathogen, as *Herbaspirillum rubrisubalbicans* comb. nov.; and classification of a group of clinical isolates (EF Group I) as *Herbaspirillum* species 3. *Int. J. Syst. Bacteriol.* v. 46, n. 3, p. 802-810, 1996.

BALDANI, J. I. *et al.* Recent advances in BNF with non-legume plants. *Soil Biol. Biochem.* v. 29, n. 5/6, p. 911-922, 1997.

KIRCHHOF *et al.* *Herbaspirillum frisingense* sp. nov., a new nitrogen-fixing bacterial species that occurs in C4-fibre plants. *Int. J. Syst. Evol. Microbiol.* v. 51, p. 157-168, 2001.

SEVILLA, M.; KENNEDY, C. Genetic analysis of nitrogen fixation and plant-growth stimulating properties of *Acetobacter diazotrophicus*, an endophyte of sugarcane. In: TRIPLETT, E. **Prokaryotic nitrogen fixation**. Wyndham: Horizon Scientific Press, 2000. p. 737-760.

NOTA: Dados trabalhados pelo autor.

1.2.1 Diazotróficos Endofíticos

Bactérias endofíticas são onipresentes na maioria das espécies vegetais, colonizando os tecidos de forma localizada ou sistêmica. Várias definições foram propostas para bactérias endofíticas (DÖBEREINER, 1992; KLOPPER & BEAUCHAMP, 1992; HALLMAN *et al.*, 1997) e apesar de não serem totalmente concordantes, todas elas caracterizam os endofíticos como bactérias que colonizam o interior de plantas, sem necessariamente causar-lhes algum dano. Historicamente, tem-se acreditado que bactérias endofíticas são patógenos fracamente virulentos de plantas, mas recentemente foram descobertos vários efeitos benéficos nas plantas hospedeiras, tal como promoção de crescimento em plantas, aumento da resistência contra patógenos de plantas e parasitas (HALLMAN *et al.*, 1997) e a própria Fixação Biológica de Nitrogênio. Em relação aos endofíticos diazotróficos, foi proposta a sua divisão em endofíticos facultativos e obrigatórios (BALDANI *et al.*, 1997). Entretanto, esta divisão, apesar de fornecer uma maneira conveniente para agrupar estes organismos, é ainda controversa e o relacionamento levado em consideração neste sistema é artificial, não sendo possível classificar claramente algumas bactérias como facultativas ou obrigatórias. Além do mais, endofíticos obrigatórios podem ser mantidos em culturas puras (SEVILLA & KENNEDY, 2000).

Em geral, bactérias endofíticas originam-se a partir de comunidades bacterianas epifíticas da rizosfera e filoplano, bem como a partir de sementes infectadas com endofíticos ou material de plantio (HALLMAN *et al.*, 1997). Estes organismos dispersam aparentemente via sementes, propagação vegetativa, material de planta morto e pelo mecanismo de alimentação de insetos (BALDANI *et al.*, 1997; BAZZICALUPO & OKON, 1999). Bactérias endofíticas parecem penetrar ativamente nos tecidos de plantas usando enzimas hidrolíticas como celulase e pectinase, além de

entrar nas plantas através de aberturas naturais ou feridas. Considerando-se que estas enzimas são também produzidas por patógenos, um maior conhecimento na sua regulação e expressão será necessário para distinguir bactérias endofíticas de patógenos de plantas. Em geral, bactérias endofíticas ocorrem em baixas densidades populacionais em relação aos patógenos. Evolutivamente, endófitos parecem ser intermediários entre bactérias saprofíticas e patógenos de plantas, mas é possível apenas especular se elas são saprófitas evoluindo em direção a patógenos, ou são mais evoluídos do que patógenos de plantas e apenas conservam o ambiente protegido e suprimento de nutrientes por não matar seu hospedeiro (HALLMAN *et al.*, 1997). De maneira geral, a comunidade endofítica é dinâmica e influenciada por fatores bióticos e abióticos, com a própria planta constituindo um dos principais fatores de influência. Uma vez que bactérias endofíticas contam com o suprimento nutricional oferecido pela planta, qualquer parâmetro afetando o estado nutricional da planta pode, conseqüentemente, afetar a comunidade endofítica. (HALLMANN *et al.*, 1997).

Os diazotróficos que têm sido caracterizados como endofíticos estão distribuídos entre os gêneros *Azospirillum*, *Gluconacetobacter*, *Azoarcus*, *Herbaspirillum* e *Burkholderia* (BALDANI *et al.*, 1997). Os dois primeiros gêneros pertencem à subdivisão alfa das Proteobactérias e os demais à subdivisão beta. As alfa Proteobactérias abrigam também diazotróficos como os simbióticos que formam o grupo dos rizóbios, bactérias de vida livre, como *Beijerinckia*, fototróficos, como *Rhodopseudomonas*, metanotróficos, como *Methylocystis* e fitopatógenos como *Agrobacterium* (YOUNG, 1992). As beta Proteobactérias incluem diversos fitopatógenos e contrariamente à subdivisão alfa, possui relativamente poucas espécies capazes de fixar N₂, como nos gêneros *Derxia*, *Alcaligenes* e *Thiobacillus* (YOUNG, 1992), além daqueles já citados. A Tabela 2 lista os principais diazotróficos endofíticos e seus hospedeiros.

Os endofíticos são isolados do interior de plantas e, em geral, possuem uma baixa taxa de sobrevivência em solos naturais. Entretanto, embora espécies de

Azospirillum possam ser encontradas no interior dos tecidos de plantas, como nas raízes de arroz, estes diazotróficos são principalmente isolados da rizosfera. Por isso, algumas vezes eles têm sido classificados como endofíticos facultativos, embora esta classificação não esteja ainda completamente estabelecida, como mencionado anteriormente. O gênero *Azospirillum* compreende sete espécies: *Azospirillum brasilense*, *Azospirillum lipoferum*, *Azospirillum amazonense*, *Azospirillum Irakense*, *Azospirillum halopraeferans* (DÖBEREINER *et al.*, 1995b), *Azospirillum largomobile* (BEN DEKHIL *et al.*, 1997) e *Azospirillum doebereinae* (ECKERT *et al.*, 2001), todas com capacidade para fixar N₂. As espécies de *Azospirillum*, principalmente *A. brasilense* e *A. lipoferum*, têm sido isoladas de cereais (milho, trigo, arroz e sorgo), gramíneas forrageiras, cana-de-açúcar, palmeiras oleaginosas e plantas tuberosas, a partir de tecidos da raiz, caule, folhas, tubérculos, interior do xilema e em frutos e sementes (Tabela 2), principalmente em regiões tropicais (BALDANI *et al.*, 1997). A distribuição de *A. irakense* e *A. halopraeferans* parece ser mais restrita, tendo sido relatados apenas ocorrendo em plantas de arroz crescidas no Iraque e em gramínea *Kallar Grass* (*Leptochloa fusca*) crescida em solos salinos do Paquistão, respectivamente (DÖBEREINER *et al.*, 1995b).

O gênero *Herbaspirillum* é composto por apenas duas espécies. *Herbaspirillum seropedicae* foi isolada de um grande número de amostras de raízes, colmos e folhas de milho, sorgo, arroz, cana-de-açúcar, gramíneas forrageiras, palmeiras, abacaxizeiro e bananeira no Brasil (BALDANI *et al.*, 1997; WEBER, 1998 - Tabela 2). Há poucos relatos de sua ocorrência em outros países. O seu mecanismo de disseminação natural não está ainda completamente esclarecido, embora ela pareça ser transferida principalmente através das sementes (BALDANI *et al.*, 1997). *Herbaspirillum rubrisubalbicans*, que foi originalmente classificada como *Pseudomonas rubrisubalbicans* e posteriormente colocada no gênero *Herbaspirillum* por homologia de DNA:rRNA (BALDANI *et al.*, 1996), é considerado um patógeno que causa a doença da estria mosqueada em variedades de cana-de-açúcar plantadas

nos E. U. A. Entretanto, variedades plantadas comercialmente no Brasil mostram-se resistentes à doença. Esta espécie ocorre em plantas de cana-de-açúcar, arroz, palmeiras, abacaxizeiro e bananeira. Atividade da nitrogenase foi detectada nos vasos do xilema de folhas de cana-de-açúcar infectada com este endófito (BALDANI *et al.*, 1997). Recentemente foi sugerida a inclusão de uma nova espécie, *Herbaspirillum frisingense*, com base em análises fisiológicas e por homologia de DNA:DNA a partir de isolados de *Spartina pectinata*, *Miscanthus sinensis*, *Miscanthus sacchariflorus* e *Pennisetum purpureum* (KIRCHHOF *et al.*, 2001).

O gênero *Azoarcus* também é composto por duas espécies: *Azoarcus communis* e *Azoarcus indigens*. Sua ocorrência parece restrita a gramínea “Kallar Grass” (*Leptochloa fusca*), de onde a bactéria foi repetidamente isolada no Paquistão sem causar nenhum sintoma patogênico aparente (BALDANI *et al.*, 1997; DÖBEREINER *et al.*, 1995b).

Dentro do recém-criado gênero *Gluconacetobacter*, somente *Gluconacetobacter diazotrophicus* mostrou-se capaz de fixar N₂. Esta bactéria foi recentemente renomeada, sendo transferida a partir do gênero *Acetobacter* (YAMADA *et al.*, 1997). Esta espécie foi encontrada principalmente associada com plantas ricas em açúcar como cana-de-açúcar, batata doce e capim Cameroon (todos propagados vegetativamente) em números de 10⁵ a 10⁶ células.g⁻¹ (BALDANI *et al.*, 1997). Sua ocorrência restrita é confirmada pela completa ausência no solo e no tecido de plantas invasoras crescendo entre linhas em um campo de cana-de-açúcar (DÖBEREINER *et al.*, 1988). Entretanto, este diazotrófico foi isolado também de plantas de café (JIMÉNEZ-SALGADO *et al.* 1997).

O gênero *Burkholderia* inclui vários patógenos e foi criado a partir do grupo de homologia II de *Pseudomonas*, contendo originalmente sete espécies, sendo a espécie tipo *Burkholderia cepacia* (YABUUCHI *et al.*, 1992). Uma nova bactéria fixadora de nitrogênio isolada da rizosfera de arroz, cultivada no Vietnã, foi incluída no gênero devido a sua alta homologia DNA:DNA, DNA:rDNA e características

fenotípicas, tendo sido proposto o nome *Burkholderia vietnamiensis* (TRAN VAN *et al.*, 1994). Dois outros grupos de bactérias diazotróficas foram isoladas em plantas de arroz, batata doce e mandioca (raiz, tubérculo e caule) cultivadas no Brasil. Análises morfológicas, fisiológicas e seqüenciamento parcial do gene para o RNA ribossomal de 23S (23S rDNA) mostraram que estes isolados são relacionados ao gênero *Burkholderia*, mas com características bem diferentes de *B. vietnamiensis* (BALDANI *et al.*, 1997). Para duas estirpes destes isolados, M130 e Ppe8, foram propostos os nomes “*Burkholderia brasiliensis*” e “*Burkholderia tropicalis*”, respectivamente. Entretanto, o seqüenciamento do gene para o RNA ribossomal de 16S (16S rDNA) da estirpe M130 mostrou que há 100% de homologia com a seqüência de *Burkholderia kururiiensis* (J. I. BALDANI, comunicação pessoal), uma nova espécie proposta recentemente, isolada de águas poluídas e capaz de degradar tricloroetileno (ZHANG *et al.*, 2000). Um grande número destes diazotróficos tem sido encontrado no interior de raízes de arroz sem causar danos à célula (BALDANI *et al.*, 1997).

Outros diazotróficos capazes de colonizar o interior do tecido de plantas foram descritos: *Alcaligenes faecalis* em arroz, *Pantoea agglomerans* em aveia, *Bacillus* sp., *Klebsiella pneumoniae*, *Erwinia* sp. e *Enterobacter* sp. (WEBER, 1998).

1.2.2 Contribuição dos diazotróficos associativos à fixação biológica de nitrogênio

Embora não haja provas definitivas de que os organismos endofíticos fixadores de nitrogênio são capazes de fornecer nitrogênio combinado para a planta (KENNEDY, 1999), são muitos os experimentos que mostram o aumento de produção e/ou acúmulo de N em campos cultivados (OKON, 1999). Grande parte das evidências da contribuição da Fixação Biológica de Nitrogênio pelos endofíticos vem de estudos com balanço de N, diluição isotópica de ^{15}N e abundância natural de ^{15}N em plantas de cana-de-açúcar, arroz e “Kallar Grass”, que podem obter parte do nitrogênio de que necessitam através da Fixação Biológica de Nitrogênio (JAMES, 2000). Em arroz, o

potencial da contribuição da fixação biológica de nitrogênio para 4 a 6 cultivos consecutivos foi estimado em 18% do N acumulado na planta (APP *et al.*, 1980). Experimentos conduzidos em casa de vegetação no Brasil indicaram que 20% do nitrogênio incorporado pela planta foi proveniente da Fixação Biológica de Nitrogênio. As maiores respostas para esta cultura parecem estar associadas à inoculação com *Herbaspirillum seropedicae* e *Burkholderia* sp., sendo a resposta dependente da variedade da planta e estirpe da bactéria usadas (BALDANI, *et al.*, 1999).

A contribuição da fixação biológica de nitrogênio para a cultura da cana-de-açúcar também tem sido estudada e medições com técnica de diluição de ^{15}N sugerem que 60% do N total da planta é derivado deste processo (URQUIAGA *et al.*, 1992). Acredita-se que boa parte desta contribuição possa vir dos diazotróficos *H. seropedicae*, *Burkholderia* sp. e *G. diazotrophicus*, que ocorrem em 10^4 a 10^6 células g^{-1} de tecido fresco, além de *A. amazonense*. Em experimentos utilizando ^{15}N para plantas crescendo em potes de 50 kg, inoculação com diferentes combinações dos organismos citados mostraram que há uma ação sinérgica entre eles, variando o N acumulado pelas plantas de 5 a 33% (BALDANI *et al.*, 1999). Outros dados indicam o acúmulo de 10 a 80 kg de N. ha^{-1} . ano^{-1} em plantas de arroz e cana-de-açúcar, vindos da Fixação Biológica de Nitrogênio (BAZZICALUPO & OKON, 1999).

Em cálculos ainda muito debatidos, acredita-se que o balanço energético para a produção do álcool a partir de cana-de-açúcar é de cerca de 4 vezes, isto é, obtém-se 4 vezes mais energia com o álcool do que é usada para sua produção, podendo atingir o valor de 5, se for usada irrigação e aplicação mais adequada de P, K e Mo (DÖBEREINER, 1995).

Estudos recentes têm também demonstrado a expressão, pelos diazotróficos, dos genes que codificam para a nitrogenase, a expressão *in situ* das proteínas da nitrogenase e a atividade da nitrogenase (redução de acetileno), em associação com suas plantas hospedeiras. Usando fusões de genes *nif* com genes repórteres, tais como

lacZ e *gusA* e marcação da nitrogenase com ouro, foi mostrado que a Fixação Biológica de Nitrogênio é expressa no interior dos tecidos da planta (BAZZICALUPO & OKON, 1999). Entretanto, estes estudos não fornecem evidências conclusivas de que as plantas possuem dependência das bactérias associadas, uma vez que as técnicas não distinguem entre o N₂ fixado pelos endofíticos daquele fixado pelas bactérias heterotróficas de vida livre e cianobactérias da rizosfera e superfície das plantas. Outro problema surge quando se usa técnica de incorporação de ¹⁵N para mostrar a transferência de N fixado da bactéria para a planta. O longo tempo de exposição normalmente usado nestes experimentos pode indicar que o N fixado é transferido principalmente por células mortas e mineralização de diazotróficos assimbióticos (JAMES, 2000).

1.2.3 Outras contribuições dos diazotróficos associativos

Experimentos de inoculação com *Azospirillum* sp., conduzidos ao redor do mundo, têm mostrando um efeito positivo da bactéria na produção, mas ainda há um debate acerca do principal modo de ação pelo qual diazotróficos endofíticos contribuem para o nitrogênio acumulado nas plantas (BALDANI *et al.*, 1997). Inoculantes comerciais produzidos com *Azospirillum* têm mostrado efeitos positivos na utilização do nitrogênio e no aumento da produção de plantas de milho, trigo e sorgo em campos cultivados da Europa e África do Sul (BAZZICALUPO & OKON, 1999). Substâncias promotoras de crescimento, fixação de nitrogênio *per se* ou a habilidade da nitrato redutase bacteriana em ajudar na incorporação do nitrogênio assimilado a partir do solo pela planta (FERREIRA *et al.*, 1987) foram também demonstrados. A respeito destes diferentes efeitos exercidos por diazotróficos endofíticos em associação com gramíneas, aumentos variando de 5 a 30% na produção foram observados em vários experimentos de inoculação com *Azospirillum* (BALDANI *et al.*, 1997).

Diversas substâncias produzidas e excretadas por bactérias diazotróficas, genericamente denominadas de Reguladores do Crescimento de Plantas (PGR), tais como auxinas, giberelinas e citocininas, são capazes de exercer efeitos positivos nas plantas (BAZZICALUPO & OKON, 1999). A produção de ácido indol acético (AIA), um fitormônio da classe das auxinas, é freqüente entre bactérias que habitam a rizosfera de plantas e pode estar ligada a fitopatogênese, como em *Agrobacterium tumefaciens*, *Agrobacterium rhizogenes*, *Erwinia herbicola* e *Pseudomonas syringae* (PATTEN & GLICK, 1996). Entretanto, a produção de AIA foi também observada para bactérias benéficas, muitas delas diazotróficas como *Azospirillum* sp., *Alcaligenes faecalis*, *Klebsiella* sp., *Gluconacetobacter diazotrophicus*, *Herbaspirillum seropedicae*, entre outras (COSTACURTA & VANDERLEYDEN, 1995). AIA produzido por estes organismos poderia funcionar como promotor de crescimento nas plantas hospedeiras. Plantas de cana-de-açúcar inoculadas com a estirpe PA15 de *G. diazotrophicus* aumentaram sua massa e conteúdo de N quando crescidas sem N combinado. Aumento no crescimento também foi observado em plantas inoculadas com um mutante Nif⁻, na presença de N combinado (KENNEDY, 1999).

1.3 TAXONOMIA DE MICRORGANISMOS

Existem cerca de 2,0 milhões de espécies descritas entre todos os seres vivos (AMORIM, 1997). Entretanto, este número pode ser extremamente conservador em relação à diversidade existente, principalmente em relação à diversidade microbiana. É provável que a diversidade entre os microrganismos seja bem maior e que o volume atual de informação sobre as espécies que conhecemos não reflita a biodiversidade do planeta. Dessa forma, a expectativa quanto a espécies com novas funções ou novos produtos ainda é grande. Em relação à diversidade dos organismos fixadores de nitrogênio, surge ainda um outro problema. Apesar de somente os procariotos serem capazes de reduzir o nitrogênio molecular a amônia, esse processo é bem generalizado

entre este grupo de organismos. Entretanto, a identificação e classificação de microrganismos não têm como prática testar sua capacidade em fixar nitrogênio (YOUNG, 1992), o que, certamente ampliaria o número de espécies diazotróficas conhecidas.

A taxonomia é geralmente definida como a ciência da classificação (BUSSE & LUBITZ, 1996). Entretanto, ela pode ser dividida em três partes (VANDAMME *et al.*, 1996):

- a) a classificação, que trata do ordenamento dos organismos em grupos taxonômicos com base na sua similaridade;
- b) a nomenclatura, que refere-se ao processo de nomear as unidades classificadas e
- c) a identificação, onde é determinado se um organismo desconhecido pertence a uma ou outra unidade definida.

A taxonomia de bactérias é tradicionalmente baseada em análises morfológicas, fenotípicas e bioquímicas. Entretanto, é um consenso entre os autores que uma classificação deste tipo nem sempre é útil e muitas vezes torna-se errônea, já que as bactérias não possuem a diversidade de caracteres morfológicos presentes nos metazoários (OLSEN *et al.*, 1986; WOESE, 1987; YOUNG, 1992). Uma classificação baseada em relações evolutivas tem muitas vantagens sobre uma baseada em similaridades superficiais. Se esta é construída em dados suficientes, então ela será estável. Não dependerá de escolhas arbitrárias de propriedades “significantes” e permitirá que a evolução dos genes, de caminhos bioquímicos e de adaptações possam ser explicados (YOUNG, 1992). Acredita-se que uma vantagem da classificação tradicional é que ela seja “prática”: permite nomear um novo isolado com pouco esforço experimental. Entretanto, ela somente será útil se um taxon definido por um pequeno número de características for também relativamente homogêneo em outras propriedades. Este somente será o caso, se o taxon for filogeneticamente coerente (YOUNG, 1992). Recentemente tem-se dado grande importância a uma abordagem

“polifásica” (uso de técnicas diferentes) no delineamento de taxa (VANDAMME *et al.*, 1996).

A espécie é a unidade básica da taxonomia bacteriana e é definida como um grupo de estirpes, incluindo a estirpe tipo (VANDAMME *et al.*, 1996). Dois mecanismos diferentes contribuem para manter as espécies: o mecanismo de coesão genética, que envolve troca suficiente de genes entre membros de uma mesma espécie para manter um grupo comum de genes e o mecanismo de coesão ecológica que pode manter o isolamento de uma espécie, permitindo que seus membros compartilhem um nicho comum (YOUNG, 2000a). Entretanto, em eucariotos reproduzindo sexualmente, as espécies são uma unidade biológica natural, definida pela troca de genes intraespecíficos e barreiras interespecíficas, mas não está claro se espécies bacterianas são, ou podem ser, unidades naturais no mesmo sentido (YOUNG, 1992).

Um parâmetro fundamental, atualmente usado para estabelecer o limite entre espécies de bactérias é baseado no limite de reassociação de DNA:DNA entre dois microrganismos, que serão considerados como pertencendo à mesma espécie se este limite for superior a 70% e pertencendo a espécies diferentes em caso contrário (WAYNE *et al.*, 1987). Se esta mesma metodologia fosse aplicada aos primatas, estes pertenceriam todos a uma mesma espécie! O homem e o chimpanzé, por exemplo, mostram uma relação de 98,6% em relação a reassociação DNA:DNA (STALEY, 1999). Análise feita na molécula de RNA ribossomal é igualmente discrepante. *E. coli* possui sete operons de rRNAs e no mínimo 15 bases de diferença entre seus rRNAs de 16S, enquanto que entre o homem e o camundongo são observados apenas 16 bases de diferença entre o RNA de 18S (STALEY, 1999). Embora pareça ser unânime a crença na existência de espécies entre bactérias, parece que ainda é preciso um enorme esforço para que se estabeleça uma definição mais clara e que possam ser propostos métodos padronizados para a descrição desta entidade ecológica (SMITH *et al.*, 1991; YOUNG, 1992; STALEY, 1999).

Embora as bactérias sejam haplóides, tenham uma estrutura populacional

clonal e raramente sujeitam seus genomas a recombinação (SMITH *et al.*, 1991; LEVIN & BERGSTROM, 2000), genes individuais podem apresentar-se como mosaicos devido à troca local de gene entre espécies relacionadas. Bactérias diferindo em mais de 20% da seqüência do DNA podem fazer troca, normalmente localizada, de DNA cromossomal (SMITH *et al.*, 1991). LAWRENCE & OCHMAN (1998) mostraram que em *E. coli*, cerca de 18% da seqüência genômica provém de DNA estranho a esta espécie, perfazendo um total de 234 eventos de transferência lateral desde que este organismo divergiu de linhagens de *Salmonella*, a 100 milhões de anos atrás. Mas, apesar desta flexibilidade observada nos genomas bacterianos, é possível imaginarmos uma outra porção do genoma das bactérias, composta por genes essenciais sob quase todas as condições de crescimento e que é razoavelmente estável e predominantemente passado através das sucessivas gerações verticalmente. A definição de espécie a nível molecular deve, portanto ter como referência esta parte do genoma, a qual inclui os genes para o 16S rRNA (YOUNG, 2000a).

A classificação molecular de estirpes tem sido obtida por diferentes abordagens. Mas para uma classificação confiável, todos estes métodos devem reunir vários critérios para serem de grande utilidade (OLIVE & BEAN, 1999):

- a) Todo organismo dentro de uma espécie deve ser passível de tipagem pelo método usado. Para alguns métodos fenotípicos, tais como aqueles baseados em reação com um anticorpo ou a presença de um receptor para bacteriófago, a característica particular pode não estar presente em todos os membros da espécie.
- b) Qualquer método de subtipagem deve ter grande poder de diferenciação. Ele deve ser capaz de diferenciar claramente estirpes não relacionadas, tal qual aquelas que são geograficamente distintas dos organismos fonte e ao mesmo tempo demonstrar a relação de todos os organismos isolados de indivíduos infectados através da mesma fonte.
- c) Uma terceira preocupação para metodologia de subtipagem refere-se a

capacidade de uma técnica de reproduzir o mesmo resultado quando uma estirpe particular é testada repetidamente. Esta reprodutibilidade é especialmente importante para a construção de bancos de dados confiáveis, contendo todas as estirpes conhecidas dentro de uma espécie para a qual organismos desconhecidos podem ser comparados para classificação.

Muitas das técnicas moleculares usadas atualmente para tipagem contam com a separação eletroforética de fragmentos de DNA de diferentes comprimentos (OLIVE & BEAN, 1999). Estes métodos de caracterização e taxonomia molecular podem ser, grosseiramente, divididos em dois grupos: métodos baseados em um único *locus* e métodos baseados em múltiplos *loci*. Os métodos baseados em um único *locus*, normalmente caracterizam um grupo ou uma população de organismos com base na análise de um único gene (HAMPL *et al.*, 2001). Técnicas de seqüenciamento de DNA e SSCP (polimorfismo conformacional de fita única) são largamente empregadas. Entretanto, genes que codificam para proteínas ou regiões intergênicas podem não refletir a posição taxonômica destes organismos (HAMPL *et al.*, 2001). Por outro lado, mesmo em concordância com a posição taxonômica dos organismos, os genes não possuem a resolução necessária para se caracterizar organismos a nível intraespecífico, como ocorre, por exemplo, com o 16S rDNA. Os métodos baseados em *loci* múltiplos, baseiam-se na análise de todo o genoma dos organismos, permitindo a caracterização intraespecífica (HAMPL *et al.*, 2001).

A análise do genoma total por padrão de bandas do DNA tem sido facilitado pela descoberta de elementos repetitivos, incluindo DNAs Polimórficos Amplificados Aleatoriamente (RAPD), Consenso Intergenérico Repetitivo de Enterobactérias (ERIC) e Palíndromo Extragênico Repetitivo (REP) e análise de sítios de enzimas de restrição nos cromossomas bacterianos por Eletroforese em Gel de Campo Pulsado (PFGE) e Polimorfismo de Fragmentos Amplificados (AFLP). Todas estas técnicas reconhecem sítios aleatórios no genoma que não podem ser previstos sem o

seqüenciamento total do genoma (GÜRTLER & MAYALL, 2001). Estes métodos, normalmente, utilizam a técnica da reação em cadeia da polimerase (PCR) para amplificar os fragmentos em análise e o resultado é apresentado na forma de um padrão de bandas em um gel eletroforético (OLIVE & BEAN, 1999). Uma vez que estes padrões podem ser extremamente complexos (como nas técnicas de RAPD, ERIC/REP e AFLP), a facilidade com a qual os padrões são interpretados e relacionados é um fator na avaliação da utilidade de um método de tipagem particular. Juntamente com a facilidade de interpretação de um método particular está sua facilidade de uso. A dificuldade técnica, custo e tempo para obter um resultado deve ser também avaliado na determinação da utilidade de um método de tipagem particular.

A limitação dos métodos fenotípicos de tipagem tem levado ao desenvolvimento de métodos de tipagem baseados no genótipo microbiano ou seqüenciamento de DNA, que minimiza problemas com a “tipabilidade” e reprodutibilidade e, em alguns casos, permite o estabelecimento de grandes bancos de dados de organismos caracterizados (OLIVE & BEAN, 1999). O pivô neste processo de transição tem sido a molécula de RNA ribossomal de 16S (16S rRNA) ou, mais frequentemente, seu gene (16S rDNA), por vários motivos (ver adiante), dentre os quais a sua universalidade e conservação da sua função, permitindo construir um esboço da árvore evolutiva para as bactérias, a qual confirma certas categorias fisiológicas antigas, mas desmantelam muitas outras e remonta os componentes em novos grupos (YOUNG, 1992).

1.4 INFERÊNCIA FILOGENÉTICA

A inferência filogenética, ou mais especificamente, a reconstrução da evolução de um grupo de organismos, requer várias etapas, que culmina com a obtenção de uma árvore filogenética final que possa representar sua história evolutiva

de maneira confiável. De uma forma resumida, quatro passos são necessários: o ALINHAMENTO DAS SEQÜÊNCIAS, a DETERMINAÇÃO DO MODELO DE SUBSTITUIÇÃO, a CONSTRUÇÃO DA ÁRVORE e a AVALIAÇÃO DESTA ÁRVORE (HERSHKOVITZ & LEIPE, 1998). Entretanto, antes do desenvolvimento deste processo é preciso, em se tratando de filogenia molecular, que se tenha definido uma molécula que possa nos contar esta história evolutiva da melhor maneira possível, levando-se em conta as diferenças atuais entre os organismos. É preciso também obter os dados (as seqüências) dos organismos a serem estudados. Aqui a qualidade e precisão dos dados obtidos irão refletir diretamente na precisão da árvore filogenética inferida. Finalmente, é importante também ter, *a priori*, uma idéia grosseira da posição taxonômica dos organismos seqüenciados (se estes são desconhecidos). Uma vez que uma das grandes vantagens da análise de seqüências é a possibilidade de utilização de dados obtidos em diferentes laboratórios e em diferentes circunstâncias e sendo os bancos de dados atualmente de domínio público e com uma composição gigantesca, esta etapa também é de grande importância para o resultado final, devendo ser a filtragem dos dados criteriosa para evitar ambiguidades na análise final. Pode-se então seguir as etapas que fazem parte da inferência filogenética propriamente dita.

A análise filogenética de DNA ou proteína tem sido uma importante ferramenta para estudar a história evolutiva dos organismos, das bactérias aos homens. Uma vez que a taxa evolutiva das seqüências varia extensivamente com o gene ou segmento de DNA (WILSON *et al.*, 1977) por estar sob pressão de seleção, variando entre um gene e outro, entre diferentes partes do mesmo gene e entre uma linha de descendentes e outra. Isto significa que algumas seqüências mudam rápido o suficiente para mostrar variação significativa ao nível intraespecífico (pseudogenes, por exemplo, ou sítios silenciosos em códons), considerando que seqüências que são cruciais para a função são altamente conservadas e permanecem reconhecíveis em todos os organismos. Esta variação em taxa permite comparações de seqüências, se bem escolhidas, para serem úteis em todos os níveis de resolução taxonômicos, mas isto

também significa que análises cada vez mais sofisticadas têm que ser desenvolvidas para extrair a máxima informação filogenética (YOUNG, 1992). A análise filogenética é também importante para clarear o padrão evolutivo de famílias multigenes (ATCHLEY *et al.*, 1994), bem como para entender a evolução adaptativa ao nível molecular (JERMANN *et al.*, 1995; WISTOW, 1993). Esta técnica também permite uma visão aprofundada do mecanismo de manutenção de alelos polimórficos em populações (TAKAHATA, 1993).

Entretanto, há duas principais razões pelas quais a filogenia molecular pode falhar: a similaridade entre seqüências pode não ocorrer devido à ancestralidade e a filogenia de um gene pode não ser representativa da história evolutiva do genoma como um todo.

Evolução convergente é um problema bem menor em nível genotípico do que em nível fenotípico, uma vez que o mapa a partir do genótipo para o fenótipo é altamente redundante. Desta forma, a ação de forças seletivas similares pode perfeitamente causar a mesma capacidade metabólica evoluir independentemente em diferentes organismos, mas é virtualmente certo que as seqüências dos genes responsáveis irão reter diferenças características (YOUNG, 1992).

1.4.1 Os Marcadores Filogenéticos

Emile Zuckerkandl e Linus Pauling introduziram, em 1965, a idéia de que moléculas podem ser usadas como "cronômetros moleculares" (ZUCKERKANDL & PAULING, 1965), abrindo-se a possibilidade de se utilizar moléculas de DNA e/ou proteínas como "documentos da história evolutiva" dos seres vivos. Uma década mais tarde, Carl Woese e colaboradores identificaram o gene para o RNA ribossomal de 16S (16S rRNA) como um poderoso marcador filogenético, utilizando-o em trabalhos subsequentes para propor um terceiro domínio dos seres vivos, o domínio das Arqueobactérias (FOX *et al.*, 1980). Diversos outros marcadores como o 23S rRNA,

subunidade beta da F_1F_0 ATPase, EF-Tu (Fator de Elongação Tu, LUDWIG & SCHLEIFER, 1999), RecA (KARLIN *et al.*, 1995), citocromo c (WOESE, 1987), entre outros, têm sido utilizados. Atualmente, enquanto nós atravessamos a era genômica, novos marcadores estão sendo propostos como alternativa ao 16S rRNA. Entretanto, desde o início dos trabalhos de Woese e colaboradores com catálogos de oligonucleotídeos produzidos por digestão do gene para o 16S rRNA com a ribonuclease T_1 (FOX *et al.*, 1977) até os dias atuais, as informações sobre esta molécula, principalmente sua seqüência, têm-se acumulado dramaticamente. Isso se deveu, em parte, ao desenvolvimento de métodos rápidos e simples para o seqüenciamento do rRNA usando transcriptase reversa (LANE *et al.*, 1985) e mais tarde o uso da PCR (SAIKI *et al.*, 1988) para amplificação e seqüenciamento seletivo dos genes que codificam para os rRNAs, principalmente o de 16S (MEDLIN *et al.*, 1988; BÖTTGER, 1989; EMBLEY, 1991) o que tem levado a uma expansão explosiva dos bancos de dados de seqüências destas moléculas. Em consequência, mesmo aqueles marcadores que teoricamente contém maior informação filogenética, como é o caso da molécula de 23S rRNA, com características básicas semelhantes a do 16S rRNA, mas com aproximadamente o dobro do seu tamanho, ainda levarão um certo tempo para atingirem um volume de informação que permita usá-los em preferência ao 16S rRNA. Nesta perspectiva, o gene para o 16S rRNA tem sido a janela que permite olhar para um passado de milhões de anos. E parece que continuará sendo, já que cada vez mais créditos são depositados nesta molécula para exercer este papel (LUDWIG & SCHLEIFER, 1999).

O conteúdo informativo das moléculas marcadoras é crítico na análise filogenética e o esforço para reconstruir a história de um gene ou organismo analisando seqüências moleculares particulares está em função do número e caráter das mudanças detectáveis na seqüência. Assim, o conteúdo informativo máximo depende do número de nucleotídeos ou aminoácidos da molécula e dos estados potenciais assumidos pelos caracteres (4 para nucleotídeos e 20 para aminoácidos).

Entretanto, em geral, tais marcadores são razoavelmente conservados e somente alguns sítios são variáveis e, portanto, informativos. E, mesmo nestes sítios, o número de estados de caracteres permitidos é reduzido devido a barreiras funcionais e pressão seletiva. Em teoria, o mais informativo dos marcadores filogenéticos, a molécula do 23S rRNA, sob condições ótimas fornece uma resolução de um evento evolutivo por 0,2 - 0,3 milhão de anos. Portanto, a análise comparativa de moléculas marcadoras permite somente uma primeira aproximação do curso da evolução bacteriana (LUDWIG & SCHLEIFER, 1999).

1.4.1.1 O Uso do RNA Ribossomal para a Reconstrução Filogenética

Três abordagens têm sido usadas no estudo de relações bacterianas baseada em similaridade de rRNA: hibridização (DE SMEDT *et al.*, 1980), catálogo de oligonucleotídeos (WOESE, 1987) e seqüenciamento (WOESE, 1987). Na primeira, o rRNA marcado de uma estirpe é hibridizado com o DNA total de cada uma das outras e a temperatura de fusão (T_m) da molécula híbrida permite estimar a similaridade de cada par. Na segunda abordagem, o 16S rRNA purificado é clivado com ribonuclease T_1 e a seqüência de cada oligonucleotídeo resultante é determinada; o “catálogo” de tais oligonucleotídeos é característico para cada organismo e pode ser comparado com todos os catálogos previamente conhecidos. A terceira abordagem é o seqüenciamento completo ou parcial do rRNA ou do DNA que o codifica, seguido pelo alinhamento e comparação das seqüências. Catálogos têm sido determinados para um amplo espectro de organismos. Dados de hibridização, entretanto, estão disponíveis para um número muito maior de isolados e espécies e fornece uma visão mais detalhada, especialmente das Proteobactérias (YOUNG, 1992).

A hibridização depende da comparação par a par dos organismos. Toda nova estirpe tem que ser comparada experimentalmente contra cada uma de um grande (e crescente) número de outras antes que possa ser seguramente classificada. Em

contraste, dados de seqüenciamento são determinados somente uma vez para cada estirpe; todas as comparações podem então ser feitas por computador contra um banco de dados cumulativo de seqüências. A determinação de seqüências de rRNA ou do gene correspondente, entretanto, consumia muito tempo, mas avanços recentes têm tornado esta técnica muito fácil e rápida. Além disso, dados de seqüenciamento fornecem informações mais detalhadas e precisas sobre as moléculas de DNA (YOUNG, 1992).

A utilização da molécula do gene 16S rDNA no estudo da diversidade microbiana e análises filogenéticas de microrganismos tem sido largamente empregada devido a suas características estruturais e funcionais. As principais características que a tornam um excelente marcador filogenético são, segundo OLSEN *et al.* (1986) e WOESE (1987):

- a) Os rRNAs são moléculas universais, ocorrendo em todos os organismos, e diferentes posições em sua seqüência evoluem a diferentes taxas, permitindo medir relações filogenéticas em quase todos os níveis (entre organismos próximos e distantes).
- b) Os rRNAs são elementos chave da maquinaria de síntese de proteínas e são funcionalmente e evolutivamente homólogos em todos os organismos.
- c) Os rRNAs são moléculas anciãs e extremamente conservadas em toda a sua estrutura.
- d) As seqüências de nucleotídeos são também conservadas. Alguns trechos são invariáveis entre organismos evolutivamente distantes, enquanto outros variam. As seqüências e os elementos de estrutura secundária conservados permitem o alinhamento de seqüências variáveis, permitindo que somente nucleotídeos homólogos sejam empregados em análises filogenéticas.
- e) Os rRNAs fornecem informação de seqüência suficiente para permitir

comparações estatísticas significativas.

- f) Aparentemente os genes para os rRNAs não sofrem interferência de transferências laterais (transferência de genes por hibridização homóloga) entre os organismos.

A estas características, intrínsecas à molécula, pode-se acrescentar ainda o fato, já mencionado, da rápida expansão dos bancos de dados, que atualmente é da ordem de milhares de seqüências para o 16S rDNA e engloba representativamente os diversos ramos taxonômicos dos microrganismos.

Por outro lado, alguns problemas são observados na utilização dos RNAs ribossomais em geral para análises filogenéticas. Vários alelos do 16S rDNA são freqüentemente observados nos genomas bacterianos, variando de 1 a 15 cópias e não havendo nenhuma indício evolutivo aparente em relação ao número de operons observados (KLAPPENBACH *et al.*, 2001). Membros de famílias multigenes, tais como a família de genes dos RNAs ribossomais, têm uma tendência a possuir seqüências de DNA idênticas. Evidências recentes sugerem que o mecanismo molecular envolvido neste processo utiliza conversão gênica por recombinação para manter a homogeneidade das seqüências. Quando a taxa destes eventos de recombinação é alta, espera-se que as seqüências de DNA na família gênica sejam uniformes e quando a taxa é baixa, certa variabilidade é esperada (OLIVE & BEAN, 1999). Dessa forma, uma variabilidade entre 0,07% e 1,23% foi observada entre os operons de RNAs ribossomais a partir de seqüências de genomas completos (KLAPPENBACH *et al.*, 2001), embora variabilidade de até 6,5% tenha sido detectada (WANG *et al.*, 1997). Evidências para transferência lateral entre genes ribossomais para alguns organismos (ex., *Rhizobium etli* e *Rhizobium leguminosarum*) também foram relatadas (MARTÍNEZ-ROMERO & CABALLERO-MELLADO, 1996).

1.4.2 Alinhamento de Seqüências

Diversas ferramentas para o alinhamento entre seqüências foram desenvolvidas nos últimos 25 anos, podendo ser grosseiramente divididas em algoritmos GLOBAL e LOCAL, que por sua vez são separados em métodos de alinhamento EM PARES (duas seqüências) e MÚLTIPLO (mais de duas seqüências; ALTSCHUL, 1997b).

A maioria dos métodos de análise filogenética trabalha a partir de um grupo de seqüências alinhadas pelo método de alinhamento múltiplo, uma vez que se queira estabelecer as relações evolutivas de um grupo de organismos. Neste caso, as seqüências são dispostas cada uma em uma linha, de maneira que cada base (ou aminoácido), sempre que possível, é colocada em uma coluna (comumente referida como sítio) que contenha esta mesma base (ou aminoácido) em todas as seqüências, sem que para isso a estrutura primária da molécula seja alterada. Para isto, parte-se do pressuposto que a semelhança entre as seqüências reflete a homologia entre elas. A única "alteração" permitida é a introdução de espaços em algumas seqüências, denominadas falhas (*gaps*), que representam deleções de um ou mais nucleotídeos (ou aminoácidos) da seqüência, que se acumularam ao longo da evolução. Normalmente, como não se tem informações sobre a ancestralidade em relação aos organismos nas análises filogenéticas, é válido pensarmos também que as falhas possam representar inserções nas seqüências que não as possuem. Uma falha é difícil de se interpretar em termos evolutivos e não há métodos confiáveis que possam usar a informação contida em seus padrões (GOLDMAN, 1997).

Entre duas seqüências é possível encontrar diversas possibilidades para o alinhamento e este problema torna-se mais grave no alinhamento múltiplo. O número de comparações a serem feitas aumenta geometricamente em relação ao número de seqüências a serem alinhadas, o que torna os cálculos impraticáveis quando se trabalha com algumas dezenas de seqüências, mesmo em computadores velozes. Por isso a

maior parte dos programas para esta finalidade utiliza um conceito denominado *alinhamento progressivo* (SANKOFF, 1975). A idéia do alinhamento progressivo depende da existência de um relacionamento filogenético entre as seqüências, para que o alinhamento comece a ser feito entre as seqüências mais próximas e seja "progressivamente" conduzido, até que a seqüência mais distante seja incorporada. Este fato nos remete a um paradoxo: a análise filogenética de um grupo de seqüências depende do correto alinhamento entre estas seqüências e, da mesma forma, para se conseguir um alinhamento confiável é preciso ter um conhecimento prévio das relações filogenéticas entre as seqüências. Na prática, os programas mais usados, como o ClustalW (APÊNDICE 3), partem do alinhamento aos pares, montam uma matriz de similaridade entre todos os pares de seqüências e calculam uma "árvore filogenética" grosseira para guiar o alinhamento múltiplo propriamente dito. Este método não garante que tenhamos sempre o melhor alinhamento possível, mas torna-se tanto mais confiável quanto maior for o grau de conservação entre as seqüências e é bastante rápido, permitindo o alinhamento de centenas de seqüências em relativamente pouco tempo.

1.4.3 Determinação do Modelo de Substituição

Até o presente momento, dois elementos do modelo de substituição podem ser acessados para dados de nucleotídeos, mas não para dados de aminoácidos ou códon. Um deles é o modelo de substituição entre bases e o outro a taxa relativa de substituição total entre os diferentes sítios da seqüência. Variáveis mais complexas, como os modelos de substituição sítio-específicas ou linhagem-específicas, não podem ser determinados (HERSHKOVITZ & LEIPE, 1998).

Para se estabelecer o modelo de substituição entre bases é preciso considerar que dois tipos de substituições de bases, quimicamente distintos, são observados no DNA. A primeira é aquela onde uma base é substituída por outra quimicamente

semelhante, ou seja, uma purina sendo substituída por outra purina (ex., $A \rightarrow G$) ou uma pirimidina substituída por outra pirimidina (ex., $C \rightarrow T$). A este tipo de substituição denomina-se TRANSIÇÃO. O segundo tipo é aquele onde ocorre a substituição de bases quimicamente distintas, com uma purina sendo substituída por uma pirimidina (ex., $A \rightarrow C$) ou vice-versa. A este segundo tipo de substituição dá-se o nome de TRANSVERSÃO, que normalmente é menos freqüente que a transição (HERSHKOVITZ & LEIPE, 1998). Na prática, a especificação das taxas relativas de substituições entre resíduos particulares normalmente tem a forma de uma matriz quadrada, com o número de linhas e colunas sendo igual a 4 para as bases do DNA. Estas matrizes podem determinar, *a priori*, os "pesos" que serão aplicados às substituições pelos algoritmos de construção da árvore, como é feito quando se usa o método da Máxima Parcimônia (ver adiante). Neste caso, onde a matriz de "pesos" é aplicada, o método é referido como "Parcimônia Ponderada". Para os métodos de construção de árvores de distância e Máxima Verossimilhança, o "peso" dado às substituições pode ser derivado dos próprios dados, sendo a matemática envolvida mais complexa (HERSHKOVITZ & LEIPE, 1998).

O segundo elemento, a taxa relativa de substituição total entre diferentes sítios, também afeta os resultados da construção da árvore. O exemplo mais óbvio desta variação é aquela entre as três posições de códon, onde a terceira posição tende a ser mais variável que as duas primeiras. Por esta razão, em algumas análises feitas a partir de seqüências codificadoras de proteínas, a terceira posição é excluída (HERSHKOVITZ & LEIPE, 1998). Variação entre sítios também é observada no 16S rRNA, entre as posições mais ou menos conservadas (VAN DE PEER *et al.*, 1996). Entretanto, as abordagens a esta questão são matematicamente mais complexas, como no modelo não paramétrico (YANG *et al.*, 1996), no modelo invariante e no modelo de distribuição gama (SWOFFORD *et al.*, 1996). Estes modelos ainda não são rotineiramente empregados, mas o modelo de distribuição gama começa a ganhar força, à medida que novos algoritmos têm sido desenvolvidos e disponibilizados nos

pacotes de programas para análises filogenéticas.

1.4.4 Reconstrução da Árvore Filogenética

A reconstrução de árvores filogenéticas usando métodos estatísticos foi iniciada independentemente em taxonomia numérica para caracteres morfológicos (SOKAL & SNEATH, 1963) e em genética de populações para dados de frequência gênica (CAVALLI-SFORZA & EDWARDS, 1964). Alguns dos métodos estatísticos desenvolvidos para este propósito são ainda usados para análise filogenética de dados moleculares, mas recentemente novos métodos têm sido desenvolvidos.

As relações filogenéticas são tradicionalmente representadas em um diagrama denominado de "árvore" (APÊNDICE 2), assim denominado devido à maneira com que os ramos das diferentes linhagens analisadas são arranjados, lembrando os ramos de uma árvore. Embora esta representação não mostre todas as possíveis variações da maneira como os organismos evoluem (ex.: são incapazes de mostrar a transferência horizontal de informação genética), ela tem persistido, por ser, freqüentemente, uma boa aproximação e por ser a maneira mais facilmente tratada computacionalmente (GOLDMAN, 1997). O estudo de árvores forma uma parte da matemática conhecida como teoria gráfica e, embora haja uma considerável variação nos termos usados, alguns deles são aqui definidos e denominados no APÊNDICE 2 para um melhor entendimento das discussões posteriormente apresentadas acerca deste assunto.

Segundo NEI (1996), considera-se atualmente a reconstrução de uma árvore filogenética como uma inferência estatística de uma verdadeira árvore filogenética, entendendo-se como "verdadeira" a real história evolutiva do grupo de organismo em estudo, a qual é desconhecida. Há dois processos envolvidos nesta inferência: "estimativa" da topologia ou do padrão de ramos de uma árvore e estimativa dos comprimentos dos ramos para uma dada topologia de árvore. Quando a topologia é

conhecida, estimativas estatísticas do comprimento dos ramos são relativamente simples. O problema maior reside na estimativa ou reconstrução de uma topologia. O número de topologias a serem analisadas aumenta rapidamente em relação ao número de organismos ou seqüências sendo comparados. Desta forma, analisando-se 4 organismos, existem apenas três possíveis topologias (considerando-se árvores sem raiz e ramos divididos dicotomicamente); quando este número aumenta para 10 organismos, o número de topologias possíveis torna-se gigantesco, sendo igual a 2.027.025; se o número de organismos é igual a 20 o número de topologias possíveis atinge inimagináveis 8×10^{20} árvores (PENNY, 1991), o que seria impraticável de se calcular até mesmo em computadores com alta capacidade de processamento.

Apesar dos fundamentos matemáticos em filogenia molecular não estarem bem estabelecidos, simulações por computador e dados empíricos indicam que os métodos atualmente usados, como Neighbor-Joining, Evolução Mínima, e Parcimônia produzem árvores filogenéticas razoavelmente boas quando um número suficientemente grande de nucleotídeos ou aminoácidos são usados. Entretanto, quando a taxa de evolução varia extensivamente de ramo a ramo, muitos métodos podem falhar em recuperar a verdadeira topologia (NEI, 1996).

Os métodos de inferência filogenética atualmente usados em filogenia molecular podem ser classificados em três grandes grupos: Métodos de Distância, Métodos de Máxima Verossimilhança e Métodos de Parcimônia. Alguns autores consideram também nesta classificação os Métodos de Análise de Grupos (*Cluster Analysis*), ainda usados atualmente na reconstrução das relações evolutivas entre organismos com base em técnicas de tipagem de DNA (*DNA fingerprint*) a partir de perfis eletroforéticos. Na maioria destes métodos o posicionamento dos organismos na árvore filogenética é baseado em modelos de evolução descrevendo regras assumidas para frequências de mudanças de bases, tipos de substituições e outros critérios semelhantes, sendo que nenhum destes índices reflete perfeitamente a história evolutiva real de um grupo de organismos. Em adição, os programas de computador

comumente usados para calcular árvores filogenéticas utilizam algoritmos heurísticos para conduzir os cálculos (LUDWIG & SCHLEIFER, 1999), o que significa que "tomam decisões", baseadas em critérios pré-estabelecidos, para a escolha da "melhor árvore", sem a necessidade de analisar todas as possibilidades.

1.4.4.1 Método da Máxima Verossimilhança (*Maximum Likelihood* - ML)

O princípio do método consiste em fazer o caminho contrário ao que é comumente feito nos demais métodos, virando o problema da filogenia às avessas. O método ML faz uma busca pelo modelo evolutivo, incluindo a própria árvore, que tem a maior probabilidade de produzir o dado observado (HERSHKOVITZ & LEIPE, 1998). Um grande problema encontrado no método é o tempo de processamento requerido se o número de seqüências é maior que dez (NEI, 1996).

Na prática, o método ML é derivado para cada posição (ou sítio) de base em um alinhamento. A "verossimilhança" é calculada em termos da probabilidade que o padrão de variação em um sítio produziria por um processo de substituição particular, dado uma árvore particular e a freqüência de bases observada. A "verossimilhança" torna-se a soma das probabilidades de cada possível reconstrução de substituições sob um processo de substituição particular. As "verossimilhanças" para todos os sítios são multiplicadas para dar uma "verossimilhança da árvore" (ex., a probabilidade dos dados fornecerem a árvore e o processo de substituição obtidos). Como se pode imaginar, para uma árvore particular, a "verossimilhança" dos dados é baixa em alguns sítios e alta em outros. Para uma boa árvore, muitos sítios terão alta "verossimilhança", então o produto das "verossimilhanças" é alto. Se não há sinal filogenético nos dados, todas as árvores obtidas ao acaso serão parecidas na "verossimilhança" (HERSHKOVITZ & LEIPE, 1998).

O modelo de substituição deve ser otimizado para ajustar-se ao dado observado. Entretanto, como o tempo computacional para desenvolver o método é alto,

normalmente é impraticável uma busca completa que simultaneamente otimize o modelo de substituição e a árvore para um dado grupo de dados (HERSHKOVITZ & LEIPE, 1998).

Quando seqüências de DNA são relativamente próximas, métodos de verossimilhança de DNA parecem funcionar bem. Entretanto, se elas são distantes e codificam seqüências de proteínas, diversas complicações podem surgir devido a taxa de substituições sinônimas ser geralmente muito maior que as substituições não sinônimas e existir a tendência de transição/transversão. A frequência relativa dos quatro nucleotídeos na terceira posição do códon também varia consideravelmente com a seqüência (NEI, 1996), sugerindo que o modelo estacionário (*stationary model*) de substituição de nucleotídeos não é apropriado. Em contraste, a mudança evolutiva de seqüências de proteína não sofre muito com estes problemas e mostra-se muito mais simples do que o de seqüências de DNA quando evolução em longo período é considerada. Analisando dados para genes mitocondriais de 11 espécies de vertebrados, RUSSO *et al.* (1996) mostrou que seqüências de proteína são mais confiáveis do que seqüências de DNA para obter a correta filogenia.

1.4.4.2 Método da Máxima Parcimônia (*Maximum Parsimony* - MP)

O uso do termo parcimônia vem do critério básico usado por este método para as estimativas da topologia da árvore filogenética, onde a árvore "escolhida" é aquela que requer o menor número de substituições (GOLDMAN, 1997). Na prática, a árvore de MP aquela com o menor número de mudanças, a qual, por definição é também aquela com o menor número de mudanças paralelas ou homoplasia (similaridade que evoluiu independentemente e não é indicativa de origem filogenética comum). O método pode ser aplicado a seqüências de DNA ou aminoácidos e só pode gerar árvores sem raiz. Para resolver este problema, freqüentemente se usa um grupo externo como raiz da árvore (APÊNDICE 2).

Um fator que faz o método da MP ineficiente é a tendência da transição/transversão e a heterogeneidade nas taxas de substituição dos diferentes sítios de nucleotídeos. Os nucleotídeos com mudança transicional e altas taxas de substituição, são pouco informativos para construções filogenéticas quando seqüências evolutivamente distantes são usadas. A razão para isto é que nestes sítios de substituições, substituições múltiplas são prováveis de terem ocorrido e isto irá introduzir “ruído” na inferência filogenética. Uma maneira de reduzir este “ruído” é dar altos “pesos” às mudanças de transversão ou sítios evoluindo lentamente e baixos “pesos” a mudanças de transição ou sítios evoluindo rapidamente (FARRIS, 1969). Neste caso, o comprimento da árvore dá uma estimativa aproximada do número mínimo de substituições de nucleotídeos e aumenta substancialmente a probabilidade de se obter a topologia correta (NEI *et al.*, 1995).

Se não há substituições múltiplas em cada sítio, espera-se que o método gere a topologia correta quanto maior for o número de sítios examinados. Na prática, seqüências de nucleotídeos estão freqüentemente sujeitas a substituições retrógradas e paralelas e isto introduz incertezas na filogenia inferida. O método é o único que pode facilmente levar em consideração as inserções e deleções de nucleotídeos ou aminoácidos, os quais algumas vezes fornecem importantes informações filogenéticas (NEI, 1996). Entretanto, normalmente consideram-se as falhas como nucleotídeos ou aminoácidos desconhecidos, não havendo nenhuma justificativa biológica ou teórica para isto, portanto não podendo esta prática ser recomendada (GOLDMAN, 1997).

1.4.4.3 Métodos de Distância (*Distance Methods*)

Nos métodos de distância uma árvore filogenética é determinada para uma matriz de

$$N(N - 1)/ 2$$

pares de distâncias evolutivas, onde "N" é o número de seqüências consideradas (VAN DE PEER, 1998). A distância entre cada par de seqüências é um valor único estimado para o número de substituições de nucleotídeos ou aminoácidos por sítio.

A árvore filogenética que é então construída a partir da matriz de distâncias, utiliza um critério pré-estabelecido para a escolha da "melhor árvore". O método de distância mais usado atualmente é o *Neighbor-Joining* (NJ - APÊNDICE 3). Este método usa uma versão simplificada do critério de EVOLUÇÃO MÍNIMA. Nos métodos de Evolução Mínima os comprimentos dos ramos de uma árvore são estimados por um certo algoritmo a partir dos dados de distâncias dos pares e a soma total dos comprimentos dos ramos é então computada para cada uma das possíveis topologias. A topologia que mostrar a menor soma será então escolhida como a árvore mais provável (NEI, 1996).

Devido ao fato de que algumas destas diferenças entre seqüências resultam de eventos múltiplos, estes valores representam uma sub-estimativa da verdadeira distância evolutiva entre os organismos (VAN DE PEER, 1998). Portanto, o número total de substituições entre qualquer par de seqüências é normalmente desconhecido e vários métodos diferentes para sua estimativa têm sido propostos (NEI, 1996). Exemplos são os métodos de KIMURA (1980) e HASEGAWA *et al.* (1985). Estes métodos são úteis para corrigir mutações paralelas e retrógradas e dão melhores estimativas do número de substituições por sítio do que a simples proporção de diferenças de nucleotídeos (distâncias p). Entretanto, estas estimativas geralmente têm uma maior variância em relação à distância não corrigida "p". Parcialmente por esta razão, a "distância p" ou uma medida de distância simples tal como a de JUKES & CANTOR (1969) tende a produzir a correta topologia mais freqüentemente do que medidas de distância sofisticadas, quando a taxa de substituição de nucleotídeos é mais ou menos constante em todas as linhagens evolucionárias e o número de nucleotídeos

examinados é muito grande (TAJIMA & TAKEZAKI, 1994). Neste caso, então, é preferível usar a "distância p" para a construção da topologia ao invés de um dos métodos para correção de substituições múltiplas. Entretanto, a vantagem da "distância p" diminui quando o número de nucleotídeos examinados é grande e a taxa evolutiva varia extensivamente com a linhagem evolutiva e neste caso é melhor usar medidas de distâncias imparciais (SAITOU & NEI, 1987). A eficiência da medida de distâncias na obtenção da árvore correta depende de no mínimo dois fatores: da relação linear com o número de substituições e o erro padrão ou coeficiente de variação das estimativas das medidas de distância (NEI, 1996).

Cada estágio na análise feita pelos métodos de distância, cálculo das distâncias e subsequente ajuste das distâncias em uma árvore, é muito mais rápida que as análises feitas pelos métodos de Máxima Verossimilhança ou Máxima Parcimônia, o que permite a análise de grande número de dados (GOLDMAN, 1997). Uma desvantagem do método reside no fato de que a conversão dos dados de seqüências, que são diretamente usados pelos outros dois métodos, em valores de distância, reduz o conteúdo filogenético informativo nelas contido originalmente e, além do mais, as propriedades estatísticas dos métodos de distância não são tão boas quanto aquelas dos métodos de Máxima Verossimilhança (GOLDMAN, 1997). Entretanto o método NJ é menos sensível a variação na taxa de substituição entre sítios (NEI, 1996).

1.5 EVOLUÇÃO DA FIXAÇÃO BIOLÓGICA DE NITROGÊNIO E FILOGENIA DOS ORGANISMOS DIAZOTRÓFICOS

A Fixação Biológica de Nitrogênio provavelmente evoluiu quando as reservas geoquímicas de nitrogênio fixado na biosfera estavam esgotadas. A considerável incerteza acerca de quando os procariotos adquiriram a capacidade para fixar N_2 é destacada pelas visões opostas da composição da paleoatmosfera pré-biótica. Ela pode ter sido fortemente redutora e contido grande quantidade de amônia,

metano, monóxido de carbono e sulfeto de hidrogênio (HART, 1979) ou menos redutora, com muito pouco destes três gases e contido N₂, dióxido de carbono e vapor d'água (LEVINE & AUGUSTSSON, 1983). Esta falta de concordância faz com que seja impossível estimar por quanto tempo as fontes de N fixado, as quais seriam acessíveis após o surgimento da biosfera, persistiram. Mesmo que o O₂ já fizesse parte da composição da atmosfera, gerado pela fotossíntese, se esta surgiu anteriormente, ou apenas pela fotólise da água, qualquer amônia de fonte geológica poderia reagir para formar N₂ e óxidos de nitrogênio. Eventualmente, a disponibilidade de óxidos de nitrogênio poderia limitar o crescimento de organismos, levando a uma pressão seletiva favorável ao aparecimento de organismos diazotróficos. De qualquer modo, a alta sensibilidade da nitrogenase à desnaturação por O₂ livre sugere que ela evoluiu na Terra sob condições anóxicas onde tal sensibilidade não seria um problema. Alternativamente, é claro, esta alta reatividade ao O₂ pode ser um pré-requisito inevitável para a fixação biológica de N₂ efetiva (NEWTON, 1999). Um outro ponto de vista é ainda sugerido por WÄCHTERSCHÄUSER (1988), envolvendo a existência de uma "nitrogenase" primitiva, na forma de um grupo Fe-S na forma de pirita, a qual produzia nitrogênio ligado a carbono ao invés de amônia. Este ponto de vista acredita na existência de uma "nitrogenase" capaz de fixar N₂ anterior a existência de enzimas e sugere que a Fixação Biológica de Nitrogênio é muito antiga. O grupo Fe-S na forma de pirita pôde então ter sido seqüestrado por uma proteína primitiva para dar origem a nitrogenase somente de Fe (CHISNELL *et al.*, 1988) que, por sua vez, poderia ter sido a precursora das três nitrogenases conhecidas atualmente. Esta visão contrasta com a sugestão de que a nitrogenase surgiu como um mecanismo assimilador de redução de cianeto para desintoxicar ambientes (POSTGATE, 1974), o que favoreceria ao aparecimento da nitrogenase de Mo em primeiro lugar, devido à sua superioridade em reduzir o cianeto. Uma última sugestão envolve uma nitrogenase originalmente ótima a partir da qual as três nitrogenases conhecidas evoluíram recentemente (POSTGATE & EADY, 1988). Esta teoria tem apoio nos seguintes argumentos: a) os genes

nifMBVUS suportam todas as três nitrogenases; b) os três grupos de genes estruturais foram provavelmente formados por interação gênica; c) o cofator metálico pode ter sido transferido entre as proteínas e; d) todas as três nitrogenases têm propriedades catalíticas similares (NEWTON, 1999).

Entre os procariotos, a Fixação Biológica de Nitrogênio está largamente distribuída entre eubactérias e arqueobactérias. Conseqüentemente, qualquer discussão da classificação dos organismos fixadores de nitrogênio deve cobrir a classificação das bactérias em geral. Entretanto, é bem possível que esta distribuição reflita melhor o ambiente explorado por estes organismos do que propriamente sua classificação taxonômica. Segundo YOUNG (1999b), existe uma distribuição desigual, entre os procariotos, da capacidade de fixar nitrogênio em consequência de uma história de duplicação e perda gênica, ao invés de múltiplas origens ou transferências laterais. Esta pode ser uma evidência da hipótese levantada acima, uma vez que a Fixação Biológica de Nitrogênio é muito cara para a célula em termos energéticos e requer um intrincado processo de regulação. Seria certamente desvantajoso manter um processo tão "caro" à célula em ambientes onde o nitrogênio necessário ao crescimento é facilmente acessível. Mas, até o presente, pouco é conhecido sobre a origem e evolução dos genes *nif*, bem como sobre os mecanismos moleculares envolvidos no desenvolvimento do processo de fixação de nitrogênio, e muito do que se propõem são hipóteses que envolvem inúmeras suposições arbitrárias feitas a partir de uns poucos genes envolvidos no processo (YOUNG, 1999b).

É notável que o conteúdo G+C dos genes *nifH* seja típico dos genomas nos quais eles foram encontrados (HENNECKE *et al.*, 1985). Este fato sugere que o gene é um componente estabelecido do genoma e que, portanto, não representa uma aquisição recente a partir de fontes externas. Esta observação é ainda reforçada pela análise filogenética feita com este gene que, aparentemente, concorda com aquela obtida com o gene para o 16S rRNA. Entretanto, em análise mais recente observa-se, na filogenia do gene *nifH*, uma divisão em quatro diferentes classes: A, B, C e D, sendo que apenas

a classe A é consistente com a filogenia dos organismos apresentada pelo 16S rRNA (YOUNG, 1999b). Esta diversidade no caráter do gene já havia sido observada pelo mesmo autor (YOUNG, 1992), que acredita que a distribuição do *nifH* não envolveu nenhum evento de transferência lateral e sim sucessivos eventos de duplicação gênica. Finalmente, um evento de duplicação do gene *nifH* recente entre as alfa Proteobactérias teria originado o gene *vnfH* associado com a vanádio dinitrogenase de *Azotobacter chroococcum*.

OBJETIVOS

Caracterizar novas bactérias diazotróficas isoladas de bananeira e abacaxizeiro, com base em técnicas moleculares de RFLP e seqüenciamento de DNA do gene que codifica para o RNA ribossomal de 16S e análises de grupo e filogenética.

2 MATERIAL E MÉTODOS

2.1 ESTIRPES USADAS

As bactérias diazotróficas caracterizadas no presente trabalho foram isoladas de caule, pseudocaule, folhas, raízes e frutos de abacaxizeiro e bananeira cultivadas nos estados da Bahia e Rio de Janeiro. O trabalho de isolamento e caracterização morfofisiológica e bioquímica dos isolados foi realizado por Olmar B. Weber no Centro Nacional de Pesquisa de Agrobiologia (CNPAB) da Embrapa - Seropédica, RJ (WEBER, 1998)¹. As estirpes de bactérias diazotróficas usadas como referência nos estudos comparativos foram fornecidas pelo CNPAB/Embrapa ou obtidas de coleção própria (Núcleo de Fixação Biológica de Nitrogênio - UFPR, PR). Foram elas: estirpes Z67, Z78 e M2 de *Herbaspirillum seropedicae* (BALDANI *et al.*, 1986), M4 de *Herbaspirillum rubrisubalbicans* (BALDANI *et al.*, 1996), M130 de "*Burkholderia brasilensis*" (nome proposto; HARTMANN *et al.*, 1995) e Ppe8 de "*Burkholderia tropicalis*" (nome proposto; BALDANI *et al.*, 1997). Todas as estirpes e isolados estão listados na Tabela 3.

¹ IMPORTANTE: Os isolados Ala, A2a, A3b, A8b e X8 foram isolados por FERREIRA *et al.* (1995) de palmeiras oleaginosas. Entretanto, por motivos maiores, dados específicos do processo de isolamento e caracterização realizados foram perdidos e, até o momento e apesar dos esforços dispensados, não foi possível recuperá-los. Sabe-se apenas que o isolado Ala é capaz de utilizar acetilglucosamina, arabinose e maltose como fonte de carbono e cresce em saís de NFb e JNFb (APÊNDICE 1); o isolado A2a utiliza arabinose como fonte de carbono e cresce em saís de JNFb; o isolado A3b utiliza arabinose e ramnose como fonte de carbono; o isolado A8b utiliza ramnose como fonte de carbono e cresce em saís de JNFb; e o isolado X8 utiliza acetilglucosamina, arabinose, ramnose e tartarato como fonte de carbono e cresce em saís de JMV, JMV pH 6,0, JNFb e NFb (J. I. BALDANI, comunicação pessoal).

TABELA 3 - BACTÉRIAS DIAZOTRÓFICAS ISOLADAS DE ABACAXIZEIRO (*Ananas comosus*) E BANANEIRA (*Musa spp.*) CULTIVADOS NOS ESTADOS DA BAHIA (BA) E RIO DE JANEIRO (RJ).

ESTIRPE	HOSPEDEIRO	CULTIVAR	TECIDO	ORIGEM
AB7	abacaxizeiro	Alenquer	folhas	Cruz das Almas/ BA
AB48	abacaxizeiro	Perolera	raiz	Cruz das Almas/ BA
AB71	abacaxizeiro	Perolera	caule	Cruz das Almas/ BA
AB98	abacaxizeiro	Pérola	fruto	Macaé/ RJ
AB117	abacaxizeiro	Smooth Cayenne	raiz	Quissamã/ RJ
AB119	abacaxizeiro	Pérola	folhas	Macaé/ RJ
AB120	abacaxizeiro	Pérola	caule	Macaé/ RJ
AB147	abacaxizeiro	Smooth Cayenne	caule	Quissamã/ RJ
BA10	bananeira	Butuhan	pseudocaule	Cruz das Almas/ BA
BA11	bananeira	Butuhan	folhas	Cruz das Almas/ BA
BA12	bananeira	Yangambi	raiz	Cruz das Almas/ BA
BA14	bananeira	Yangambi	pseudocaule	Cruz das Almas/ BA
BA15	bananeira	Prata Anã	raiz	Cruz das Almas/ BA
BA16	bananeira	Prata Anã	pseudocaule	Cruz das Almas/ BA
BA17	bananeira	Butuhan	pseudocaule	Cruz das Almas/ BA
BA22	bananeira	Prata Anã	folhas	Cruz das Almas/ BA
BA23	bananeira	Yangambi	pseudocaule	Cruz das Almas/ BA
BA25	bananeira	Prata Anã	pseudocaule	Cruz das Almas/ BA
BA27	bananeira	Yangambi	folhas	Cruz das Almas/ BA
BA88	bananeira	Maçã	folhas	Itaguaí/ RJ
BA104	bananeira	Prata Anã	pseudocaule	Itaguaí/ RJ
BA106	bananeira	Maçã	folhas	Itaguaí/ RJ
BA123	bananeira	Prata Manteiga	raiz	Macaé/ RJ
BA124	bananeira	Prata Manteiga	pseudocaule	Macaé/ RJ
BA126	bananeira	D'água	raiz	Itaguaí/ RJ
BA128	bananeira	Prata	fruto	Itaguaí/ RJ
BA131	bananeira	D'água	folhas	Itaguaí/ RJ
BA134	bananeira	Maçã	pseudocaule	Itaguaí/ RJ
BA136	bananeira	Prata	folhas	Itaguaí/ RJ
BA149	bananeira	Maçã	folhas	Itaguaí/ RJ
BA153	bananeira	Marmelo	fruto	Itaguaí/ RJ
BA161	bananeira	Maçã	raiz	Itaguaí/ RJ
Ala	n. d. ⁽¹⁾	n. d.	n. d.	n. d.
A2a	n. d.	n. d.	n. d.	n. d.
A3b	n. d.	n. d.	n. d.	n. d.
A8b	n. d.	n. d.	n. d.	n. d.
O1	bananeira	Yangambi	n. d.	Cruz das Almas/ BA
X8	n. d.	n. d.	n. d.	n. d.

(1) nenhum dado.

2.2 CONDIÇÕES DE CRESCIMENTO

Para os experimentos de amplificação do 16S rDNA e região intergênica 16S-23S rDNA, todas as bactérias foram crescidas durante a noite (~16h) em frascos de 25 mL contendo 5 mL de meio NFbHPN líquido contendo malato como fonte de carbono e 20 mmol/ L de NH₄Cl como fonte de nitrogênio (APÊNDICE 1), a 30 °C e 120 rpm. Para a extração de DNA cromossomal as bactérias foram crescidas nas mesmas condições, porém em frascos de 60 mL, contendo 20 mL de meio específico NFb HPN e malato como fonte de carbono para as estirpes Z67, Z78, M2 e M4 e para os isolados AB7, BA10, BA11, BA12, BA14, BA15, BA16, BA17, BA134, BA149, BA153 e BA161; NFbHPN e lactato como fonte de carbono para os isolados BA131, Ala, A2a, A3b e A8b; ou JMV (APÊNDICE 1) para as estirpes M130 e Ppe8 e para os isolados AB48, AB71, AB98, AB117, AB119, AB120, AB147, BA22, BA23, BA25, BA27, BA88, BA104, BA106, BA123, BA124, BA126, BA128, BA136 e O1 (todos também adicionados de 20 mmol/ L de NH₄Cl).

2.3 EXTRAÇÃO DE DNA CROMOSSOMAL

A massa celular de cerca de 20 mL de cultura cultivada por ~16 h foi coletada por centrifugação a 13.000 x g por 30 – 60 s ressuspendida em 500 µL de tampão GET (50 mmol/ L de Glucose, 25 mmol/ L de Tris HCl pH 8,0 e 10 mmol/ L de EDTA pH 8,0). Em seguida foi adicionado lisozima para uma concentração final de 300 µg/ mL e a mistura incubada a 37 °C por 30 min. Adicionou-se entre 100 - 200 µL de solução SDS (sódio dodecil sulfato) a 10%; pronase E e RNase para uma concentração final de 200 µg/ mL e 50 µg/ mL, respectivamente e a mistura foi mais uma vez incubada a 37 °C durante a noite. Cerca de 300 µL da mistura fenol: clorofórmio: álcool isoamílico (na proporção de 25:24:1) foi adicionado e os tubos gentilmente agitados por cerca de 15 min. A emulsão foi centrifugada a 13.000 x g por

10 min e a fase superior retirada com o auxílio de pipeta automática e transferida para tubos plásticos de 1,5 mL. Uma nova extração foi feita, desta vez com 1 volume de clorofórmio P. A.; a emulsão foi centrifugada a 13.000 x g por 10 min e a solução sobrenadante retirada com o auxílio de pipeta automática e transferida para um novo tubo plástico de 1,5 mL. Foram adicionados 2 volumes de etanol absoluto, o tubo foi gentilmente invertido por várias vezes até a completa precipitação do DNA, que foi então, retirado com a ajuda de uma alça de vidro capilar, lavado com etanol 80% e transferido para um novo tubo plástico de 1,5 mL. O excesso de etanol foi retirado em dessecador a vácuo. O DNA foi ressuspendido em água milli Q estéril, mantido a temperatura ambiente por 24 a 48 h, quantificado espectrofotometricamente em luz UV a 260 nm e estocado a 4 °C.

2.4 INICIADORES USADOS NA AMPLIFICAÇÃO E SEQÜENCIAMENTO DO 16S rDNA E REGIÃO INTERGÊNCIA 16S-23S rDNA

Os iniciadores Y1 e Y3, que anelam em regiões conservadas das pontas 5' e 3' da molécula do 16S rDNA, respectivamente, foram usados, permitindo a amplificação quase total do gene. O par de iniciadores permite a amplificação de 1.438 pb do 16S rDNA de *E. coli*, o que corresponde a 93,4% da molécula (excluindo as seqüências dos iniciadores). Em contrapartida, os iniciadores pHR e p23 foram usados na reação de PCR quando a intenção era amplificar o fragmento de DNA correspondendo à região intergênica 16S-23S rDNA (Tabela 4). O iniciador pHR anela em uma região conservada na ponta 3' do 16S rDNA, enquanto o iniciador p23 anela-se à ponta 5' do 23S rDNA. Desta forma, este par de iniciadores permite a amplificação completa da região intergênica.

O amplificado do 16S rDNA foi parcialmente sequenciado para 30 dos isolados e para as estirpes referência Z67, M4, M130 e Ppe8, utilizando-se os iniciadores Y1 e Y2, que delimitam uma região de aproximadamente 300 pb na ponta

5' da molécula. Para 19 isolados, representantes dos grupos obtidos nas caracterizações feitas, o amplificado do 16S rDNA foi também completamente sequenciado. Para tanto, os iniciadores Y1, 16S362f, 16S786f e 16S1203f foram usados no seqüenciamento da fita direta do DNA e os iniciadores Y3, 16S805r, 16S10110r e Y2 para o seqüenciamento da fita reversa (Tabela 4 e Figura 2). O prefixo 16S indica que os iniciadores foram desenhados para o 16S rDNA e os números que se seguem, a posição onde a ponta 3' do iniciador anela na seqüência direta de *E. coli*, de acordo com numeração de BROSIUS *et al.* (1981), no sentido direto (f - *forward*) ou reverso (r - *reverse*).

O seqüenciamento completo dos amplificados só foi possível a partir da construção de cinco dos oito iniciadores usados (16S362f, 16S786f, 16S1203f, 16S805r e 16S10110r). A construção destes iniciadores foi feita a partir do alinhamento de 65 seqüências do 16S rDNA de espécies das subdivisões alfa, beta e gama das Proteobactérias. As seqüências, obtidas no GenBank, foram inicialmente alinhadas no programa ClustalW, utilizando-se os parâmetros padrões. A partir do alinhamento obtido, foi feita uma busca visual por regiões conservadas, levando-se em consideração: temperaturas de fusão (T_m) o mais semelhante possível para os iniciadores, para que pudessem ser usados em um único ciclo de temperaturas para a PCR e escolha de um número mínimo de iniciadores, cobrindo regiões equidistantes, e capazes de permitir o seqüenciamento de toda a região do 16S rDNA com uma área de sobreposição entre elas suficiente para a montagem das seqüências contíguas com precisão (Figura 2).

Outros fatores relevantes na construção de iniciadores também foram levados em conta e, embora alguns apresentem características não desejáveis, como formação de dímeros e regiões de anelamento interno (dados não mostrados), na prática, não houve dificuldade para o seqüenciamento que pudesse ser atribuído a tais propriedades.

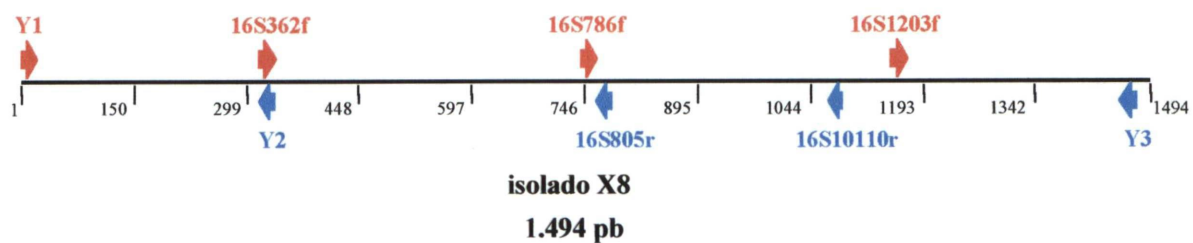
TABELA 4 - PARÂMETROS FÍSICOS DOS INICIADORES UTILIZADOS NA AMPLIFICAÇÃO E SEQUENCIAMENTO DO 16S rDNA E REGIÃO INTERGÊNICA 16S-23S rDNA.

NOME	GENE	SEQÜÊNCIA (DIREÇÃO 5' → 3')	PM	GC (%)	Tm (°C)
Y1	16S DNAr	TGGCTCAGAACGAACGCTGGCGGC	7465,87	66,7	80,56
16S362f	16S DNAr	CTCCTACGGGAGGCAGCAGTGGGG	7521,89	70,8	76,62
16S786f	16S DNAr	CGAAAGCGTGGGGAGCAAACAGG	7258,74	60,9	74,85
16S1203f	16S DNAr	GAGGTGGGGATGACGTCAAGTCCTC	7825,10	60,0	72,49
Y3	16S DNAr	TACCTTGTTACGACTTCACCCAGTC	7879,18	50,0	67,71
16S10110r	16S DNAr	TGCGCTCGTTGCGGGACTTAACC	7093,63	60,9	75,18
16S805r	16S DNAr	GACTACCAGGGTATCTAATCCTG	7085,65	47,8	58,37
Y2	16S DNAr	CCCCTGCTGCCTCCCGTAGGAGT	7327,79	66,7	74,62
	Região intergênica				
pHR	16S-23S DNAr	TGCGGCTGGATCACCTCCTT	6122,00	60,0	70,19
	Região intergênica				
p23	16S-23S DNAr	GGTTCTTTTCACCTTCCCTC	6327,15	47,6	61,54

NOME	COMPRIMENTO (pb)	POSIÇÃO EM <i>E. coli</i> ⁽¹⁾	REFERÊNCIA
Y1	24	20-43	Young, <i>et al.</i> 1991
16S362f	24	339-362	Esta tese
16S786f	23	764-786	Esta tese
16S1203f	25	1178-1203	Esta tese
Y3	26	1482-1507	Young, comunicação pessoal
16S10110r	23	1088-1110	Esta tese
16S805r	23	783-805	Esta tese
Y2	24	338-361	Young, <i>et al.</i> 1991
pHR	20	1522-1541	Fundação Tropical André Tossello
p23	21	465-485	Fundação Tropical André Tossello

(1) numeração de bases de acordo com BROSIUS *et al.* (1981); a numeração foi considerada para as moléculas de 16S e 23S rDNA individualmente.

FIGURA 2 - LOCALIZAÇÃO DOS INICIADORES USADOS PARA O SEQÜENCIAMENTO DA REGIÃO Y1-Y3 DO 16S rDNA. AS POSIÇÕES, EM ESCALA, DOS INICIADORES SÃO MOSTRADAS DE ACORDO COM SUAS POSIÇÕES NA SEQÜÊNCIA DO ISOLADO X8.



2.5 PREPARO DAS AMOSTRAS PARA REAÇÃO DA PCR A PARTIR DE CULTURAS

O gene 16S rDNA e a região intergênica 16S-23S rDNA foram amplificados diretamente a partir de culturas crescidas em meio líquido. Aliquotas de 50 μ L das culturas foram diluídas 1:10 com água milliQ estéril, fervidas em banho maria por 5 a 10 min e imediatamente resfriadas em gelo. Após este tratamento as amostras foram usadas para reação de amplificação por PCR. Para as estirpes que tiveram o 16S rDNA completamente sequenciado na região Y1-Y3, partiu-se do DNA cromossomal extraído de acordo com o protocolo descrito no item acima “Extração de DNA cromossomal”.

2.6 AMPLIFICAÇÃO DO 16S rDNA E REGIÃO INTERGÊNICA 16S-23S rDNA POR PCR

Aliquotas de 10 μ L de culturas fervidas ou 50 a 100 ng de DNA cromossomal foram adicionadas a tampão de reação da *Taq* DNA polimerase (10 mmol/ L de Tris-HCl pH 9,0; 1,5 mmol/ L de MgCl₂; e 50 mmol/ L de KCl) diluído 10 vezes de acordo com recomendação do fabricante (GIBCO BRL), 200 μ mol/ L de dATP, dTTP, dCTP e dGTP, 0,1 pmol/ μ L de cada iniciador e 1 U de *Taq* DNA polimerase (GIBCO BRL). A solução foi completada para um volume final de 50 ou 100 μ L, com água milliQ estéril. A reação de amplificação foi realizada em termociclador OmniGene HYBAID ou PE Biosystem 9600 com os seguintes ciclos de temperatura: desnaturação inicial a 93 °C por 2 min; 34 ciclos de desnaturação a 93 °C por 30 - 45 s, anelamento a uma temperatura dependente do par de iniciadores usados (ver a seguir), por 30 - 45 s e extensão a 72 °C por 2 min; e uma extensão final a 72 °C por 5 min. Na amplificação do 16S rDNA o par de iniciadores Y1 e Y3 foi usado, com uma temperatura de anelamento de 62 °C e na amplificação da região intergênica 16S-23S rDNA o par pHR e p23 foi usado com uma temperatura de anelamento de 58 °C.

Após a reação de amplificação, as amostras eram submetidas a eletroforese em gel de agarose 0,7% (~6 V/ cm, por aproximadamente 40 min.), coradas com 0,5 µg/ mL de brometo de etídeo por 30 a 45 min e fotografadas sob luz UV para avaliação do amplificado com relação ao sucesso da reação, tamanho e concentração aproximada.

2.7 RESTRIÇÃO DO 16S rDNA E REGIÃO INTERGÊNICA 16S-23S rDNA

Alíquotas de 10 µL dos produtos de amplificação do 16S rDNA foram submetidos a digestão com endonucleases de restrição. Um excesso de enzima (5 U por reação), tampão específico (fornecido pelo fabricante; GIBCO BRL) diluído 10 vezes e água milli Q estéril para um volume final de 20 µL foram adicionados à reação. A mistura de reação foi incubada a 37 °C por 1,5 - 3,0 h e os fragmentos obtidos fracionados em eletroforese com gel de agarose 2,0% a 4,3 V/ cm durante 5 h, corados com 0,5 µg/ mL de brometo de etídeo por 45 min a 1 h e fotografados sob luz UV. As enzimas utilizadas para esta análise estão distribuídas entre as três classes estabelecidas por MOYER *et al.* (1996) em relação às frequências de tamanhos dos fragmentos produzidos pela digestão do DNA. Foram elas: *AluI*, *HaeIII* e *RsaI*, que reconhecem seqüências palindrômicas específicas de 4 pb, e *Hinfi*, que reconhece 5 pb, mas possui uma seqüência de reconhecimento degenerada (Tabela 5). Estas enzimas são usadas nas técnicas de RFLP em geral porque possuem uma alta frequência de corte, gerando fragmentos de DNA em número e tamanho satisfatórios para separação por eletroforese em gel de agarose e comparação dos perfis obtidos em diferentes estirpes (MOYER *et al.*, 1996).

Os experimentos de RFLP utilizando a região intergênica 16S-23S rDNA foram realizados nas mesmas condições descritas acima.

TABELA 5 - ENDONUCLEASES DE RESTRIÇÃO USADAS NAS ANÁLISES DE RFLP DO 16S rDNA (ARDRA) E 16S-23S rDNA.

ENZIMA	SEQUÊNCIA RECONHECIDA ⁽¹⁾	COMPRIMENTO (pb)	TEMPERATURA (°C)	FREQUÊNCIA DE CORTE NO 16S rDNA DE <i>E. coli</i>
<i>AluI</i>	AG ['] CT	4	37	8
<i>HaeIII</i>	GG ['] CC	4	37	9
<i>HinfI</i>	G ['] AnTC	5	37	6
<i>RsaI</i>	GT ['] AC	4	37	3

(1) a posição de corte é indicada por uma aspa simples ('); "n" indica a ocorrência de qualquer uma das bases.

2.8 ANÁLISE DOS PERFIS DE RESTRIÇÃO

Os perfis de restrição dos fragmentos obtidos da região Y1-Y3 do gene 16S rDNA foram comparados visualmente para o agrupamento dos isolados e estirpes referência. Os perfis de restrição estabelecidos para os grupos observados foram comparados, aos pares, e convertidos em uma matriz binária indicando a presença (1) ou ausência (0) de banda em uma dada posição ou massa molecular. A partir da matriz binária, um índice de similaridade ou divergência genética foi calculado para cada par de estirpes, usando-se a seguinte fórmula (NEI & LI, 1979):

$$D = 1 - [2N_{xy} / (N_x + N_y)],$$

onde D é a distância genética entre os perfis "x" e "y" analisados, N_{xy} é o número de fragmentos (bandas) compartilhadas por ambos os perfis de restrição, N_x é o número de fragmentos do perfil "x" e N_y o número de fragmentos do perfil "y". Com a matriz binária foi calculado um dendrograma pelo método de UPGMA (*unweighted pair group method using arithmetic averages*; APÊNDICE 3) para avaliar o grau de similaridade entre os grupos. Tanto os cálculos das distâncias genéticas, quanto o cálculo do dendrograma foram feitos com o auxílio do programa TreeCon (Tabela 6).

TABELA 6 - PROGRAMAS USADOS NAS ANÁLISES DE ARDRA E SEQÜENCIAMENTO DO 16S rDNA DE ISOLADOS DE BANANEIRA E ABACAXIZEIRO E ESPÉCIES RELACIONADAS.

PROGRAMA	VERSÃO	AUTOR	URL
BioEdit	4.8.8 (ou superior)	Tom Hall	http://jwbrown.mbio.ncsu.edu/RnaseP/info/programs/BIOEDIT/bioedit.html
BLAST	2.1	S. F. Altshul	http://www.ncbi.nlm.nih.gov/BLAST/
CAP	1.0	Xiaoqiu Huang	http://cs.mtu.edu/faculty/Huang.html
ClustalW	1.81	J. D. Thompson Jeroen Raes & Yves	ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX
ForCon	1.0	Van de Peer Karl B. Nicholas &	http://vulcan.rug.ac.be/~jerae/ForCon/index.html
GeneDoc	2.6	Hugh B. Nicholas Jr.	http://www.psc.edu/biomed/genedoc
Mega2	2.0	Sudhir Kumar	http://www.megasoftware.net/
Phylip/ DNAML	3.6 (alfa)	Joseph Felsenstein	http://evolution.genetics.washington.edu/phylip.html
TreeCon	1.3 (beta)	Yves Van de Peer	http://bioc-www.uia.ac.be/u/yvdp/treeconw.html
RnaViz	2.0	Peter de Rijk	http://rrna.uia.ac.be/rnaviz/

PROGRAMA	REFERÊNCIA	USO
BioEdit	HALL, 1997 - 2001	Edição dos cromatogramas; edição de sequências
BLAST	ALTSCHUL <i>et al.</i> , 1997	Pesquisa de homologia no GenBank
CAP	HUANG, 1992	Montagem das sequências contíguas
ClustalW	THOMPSON <i>et al.</i> , 1994	Alinhamento de sequências de DNA
ForCon	não está disponível	Interconversão entre formatos de arquivos de sequências
GeneDoc	NICHOLAS & NICHOLAS, 1997 - 2000	Edição de sequências
Mega2	KUMAR <i>et al.</i> , 1993 - 2000	Cálculo de distâncias evolutivas; reconstrução de árvore filogenética
Phylip/ DNAML	FELSENSTEIN, 1989 - 2000	Reconstrução de árvore filogenética pelo método de Máxima Verossimilhança
TreeCon	VAN DE PEER, 1994 - 1998	Cálculo de distâncias genéticas e construção do dendrograma na análise de ARDRA
RnaViz	DE RIJK, 1997 - 2001	Desenho das estruturas secundárias do 16S rDNA

2.9 SEQÜENCIAMENTO PARCIAL DO 16S rDNA

Produtos de amplificação foram purificados em colunas de gel permeação, *Nucleon QC clean-up spin column kit* (Amersham) ou de adsorção, QuiaKit (QuiaGen) de acordo com especificações do fabricante. Aliquotas de 1,0 a 5,0 μL dos produtos da PCR purificados, 10 a 15 pmol do iniciador Y1 ou Y2, 6 - 8 μL de *Ready Reaction Cycle Sequencing Kit* (PE Biosystem) e água milliQ estéril para um volume final de 20 μL formaram a mistura de reação de seqüenciamento. A reação foi conduzida em termociclador com a seguinte programação de temperatura: 25 ciclos de desnaturação a 96 °C por 30 s, anelamento a 62 °C por 10 s e extensão a 60 °C por 4 min. Posteriormente as amostras foram precipitadas com 3 volumes de etanol absoluto contendo 0,5 mol/ L de acetato de amônio, centrifugadas a 13.000 x g por 30 min, o precipitado lavado com etanol 70% e novamente centrifugado a 13 .000 x g por 5 min. O excesso de etanol foi retirado em dessecador a vácuo e as amostras foram ressuspensas em solução TSR (*Template Supression Reagent*– PE Biosystem) e mantidas a -20 °C até a leitura em sequenciador automático ABI PRISM 310 *Genetic Analyzer* (PE Biosystem).

O seqüenciamento completo da região amplificada Y1-Y3 foi realizado nas mesmas condições de reação de PCR feitas para a região Y1-Y2, com o seguinte programa de temperaturas: 25 ciclos de desnaturação a 96 °C por 15 s e extensão a 60 °C por 4 min. Os produtos da reação de seqüenciamento foram precipitados com 4 volumes de isopropanol 75%, centrifugados a 13.000 x g por 25 min, lavados uma única vez com etanol 70%, novamente centrifugados a 13.000 x g por 5 min, secos a vácuo e mantidos a -20 °C até leitura em sequenciador automático ABI PRISM 377 *Genetic Analyzer* (PE Biosystem).

2.10 ANÁLISE DAS SEQÜÊNCIAS

As seqüências da região Y1-Y3 foram obtidas com a montagem das seqüências parciais feitas a partir dos iniciadores Y1, 16S362f, 16S786f, 16S1203f para o seqüenciamento da fita direta e dos iniciadores Y2, 16S10110r, 16S805r e Y3 para o seqüenciamento da fita reversa do DNA (ver item “Iniciadores Usados na Amplificação e Seqüenciamento do 16S rDNA e Região Intergência 16S-23S rDNA”). A montagem foi possível devido às regiões de sobreposição entre as seqüências parciais e pela complementaridade entre as duas fitas seqüenciadas do DNA (Figura 3). Os cromatogramas de leitura para as seqüências parciais foram visualmente analisados e editados com o auxílio do programa BioEdit (Tabela 6), principalmente para retirada das pontas, onde ocorreram as maiores dificuldades na leitura das seqüências. Este processamento foi necessário devido a alta intensidade dos picos no início da seqüência, provavelmente pela presença de dideoxinucleotídeos livres, remanescentes da etapa de precipitação do produto de reação de seqüenciamento e pela baixa resolução dos picos observada no final da leitura. Posteriormente a montagem da seqüência contígua foi feita pela sobreposição das seqüências parciais, com uma média de $537,8 \pm 112,2$ pb por leitura, com exceção das seqüências dos iniciadores 16S1203f e Y2 que seqüenciam em direção às pontas do amplificado do 16S rDNA (Figura 2) e, por isso, ficaram com uma média de leitura de $292,9 \pm 14,8$ pb e $296,5 \pm 17,6$ pb, respectivamente. A sobreposição das seqüências parciais foi feita utilizando-se o programa CAP (*Contig Assembly Program*; APÊNDICE 3), incorporado no programa BioEdit. Finalmente, após a montagem, as seqüências contíguas foram visualmente analisadas para resolver eventuais posições com ambigüidade na leitura das bases entre as seqüências sobrepostas. A montagem da seqüência contígua, após edição, para o isolado BA15 é mostrado na Figura 3.

790 800 810 820 830 840 850 860 870 880 890 900 910

cap_BA15_Y2 (-)
cap_BA15_Y1 (+)
cap_BA15_760r (-)
cap_BA15_362f (+) TCCACGCCCTAAACGATGCTACTAGTTGTCGGGTCTTAATTGACTTGGTAACGCAGCTAACGCCGTAAGTAGACCCTGGGGAGTAC
cap_BA15_455r (-) TCCACGCCCTAAACGATGCTACTAGTTGTCGGGTCTTAATTGACTTGGTAACGCAGCTAACGCCGTAAGTAGACCCTGGGGAGTACGGTCGCAAGATTAAAACCTCAAAGGAATTGACGGGGACCCGC
cap_BA15_760r (-) TCCACGCCCTAAACGATGCTACTAGTTGTCGGGTCTTAATTGACTTGGTAACGCAGCTAACGCCGTAAGTAGACCCTGGGGAGTACGGTCGCAAGATTAAAACCTCAAAGGAATTGACGGGGACCCGC
cap_BA15_786f (+) CCCaCGCCCTAAACGATGCTACTAGTTGTCGGGTCTTAATTGACTTGGTAACGCAGCTAACGCCGTAAGTAGACCCTGGGGAGTACGGTCGCAAGATTAAAACCTCAAAGGAATTGACGGGGACCCGC
cap_BA15_Y3 (-)
cap_BA15_1203f (+) ..
Contig-0 TCCACGCCCTAAACGATGCTACTAGTTGTCGGGTCTTAATTGACTTGGTAACGCAGCTAACGCCGTAAGTAGACCCTGGGGAGTACGGTCGCAAGATTAAAACCTCAAAGGAATTGACGGGGACCCGC

920 930 940 950 960 970 980 990 1000 1010 1020 1030 1040

cap_BA15_Y2 (-)
cap_BA15_Y1 (+)
cap_BA15_760r (-)
cap_BA15_362f (+)
cap_BA15_455r (-) ACAAGCGGTGGATGATGTGGATTAATTCCGATGCAACCGGAAAAACCTTACCTACCCTTGACATGGATGGAATCCCGAAGAGATTTGGGAGTGCCTCGAAAGAGAACCATCACACAGGTGCTGCATGGCTGT
cap_BA15_760r (-) ACAAGCGGTGGATGATGTGGATTAATTCCGATGCAACCGGAAAAACCTTACCTACCCTTGACATGGATGGAATCCCGAAGAGATTTGGGAGTGCCTCGAAAGAGAACCATCACACAGGTGCTGCATGGCTGT
cap_BA15_786f (+) ACAAGCGGTGGATGATGTGGATTAATTCCGATGCAACCGGAAAAACCTTACCTACCCTTGACATGGATGGAATCCCGAAGAGATTTGGGAGTGCCTCGAAAGAGAACCATCACACAGGTGCTGCATGGCTGT
cap_BA15_Y3 (-)ATTATGTGNATTANTTCGATGCACCGGAAAAACCTTACCTACCCTTGACATGGATGGAATCCCGAAGAGATTTGGGAGTGCCTCGAAAGAGAACCATCACACAGGTGCTGCATGGCTGT
cap_BA15_1203f (+)
Contig-0 ACAAGCGGTGGATGATGTGGATTAATTCCGATGCAACCGGAAAAACCTTACCTACCCTTGACATGGATGGAATCCCGAAGAGATTTGGGAGTGCCTCGAAAGAGAACCATCACACAGGTGCTGCATGGCTGT

1050 1060 1070 1080 1090 1100 1110 1120 1130 1140 1150 1160 1170

cap_BA15_Y2 (-)
cap_BA15_Y1 (+)
cap_BA15_760r (-)
cap_BA15_362f (+)
cap_BA15_455r (-) CGTCAGCTcGtGTCGT
cap_BA15_760r (-) CGTCaGctCGTGT
cap_BA15_786f (+) CGTCAGCTCGTGTCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTTGCATTAGTTGCTACGAAAAGGGCACTCTAATGAGACTGCCGGTGACAAACCCGGAGGAAGGTGGGGATGACGTC AAG
cap_BA15_Y3 (-)16S10110r16S1203f
cap_BA15_1203f (+) CGTCAGCTCGTGTCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTTGCATTAGTTGCTACGAAAAGGGCACTCTAATGAGACTGCCGGTGACAAACCCGGAGGAAGGTGGGGATGACGTC AAG
Contig-0

1180 1190 1200 1210 1220 1230 1240 1250 1260 1270 1280 1290 1300

cap_BA15_Y2(-)
cap_BA15_Y1(+)
cap_BA15_760r(-)
cap_BA15_362f(+)
cap_BA15_455r(-)
cap_BA15_760r(-)
cap_BA15_786f(+)
cap_BA15_Y3(-)
cap_BA15_1203f(+)
Contig-0

T CCTCATGGCCCTTATGGGTAGGGCTTACACGTCATACAATGGTACATACAGAGGGCCGCCAACCCGCGAGGGGGAGCTAATCCCAGAAAAGTGATCGTAGTCCGGATTGGAGTCTGCAACTCGACTCC
T CCTCATGGCCCTTATGGGTAGGGCTTACACGTCATACAATGGTACATACAGAGGGCCGCCAACCCGCGAGGGGGAGCTAATCCCAGAAAAGTGATCGTAGTCCGGATTGGAGTCTGCAACTCGACTCC
..... tTATGGGTAGGGCTTACACGTCATACAATGGTACATACAGAGGGCCGCCAACCCGCGAGGGGGAGCTAATCCCAGAAAAGTGATCGTAGTCCGGATTGGAGTCTGCAACTCGACTCC
T CCTCATGGCCCTTATGGGTAGGGCTTACACGTCATACAATGGTACATACAGAGGGCCGCCAACCCGCGAGGGGGAGCTAATCCCAGAAAAGTGATCGTAGTCCGGATTGGAGTCTGCAACTCGACTCC

1310 1320 1330 1340 1350 1360 1370 1380 1390 1400 1410 1420 1430

cap_BA15_Y2(-)
cap_BA15_Y1(+)
cap_BA15_760r(-)
cap_BA15_362f(+)
cap_BA15_455r(-)
cap_BA15_760r(-)
cap_BA15_786f(+)
cap_BA15_Y3(-)
cap_BA15_1203f(+)
Contig-0

ATGAAGTTGGAATCGCTAGTAATCGCGGATCAACATGTCGC
ATGAAGTTGGAATCGCTAGTAATCGCGGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTGTTACACACCGCCCCGTACACCCATGGGAGCGGGTTTTACCAGAAAGTGGGTAGCCTAACCCGCAAGGAGG
ATGAAGTTGGAATCGCTAGTAATCGCGGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTGTTACACACCGCCCCGTACACCCATGGGAGCGGGTTTTACCAGAAAGTGGGTAGCCTAACCCGCAAGGAGG
ATGAAGTTGGAATCGCTAGTAATCGCGGATCAGCATGTCGCGGTGAATACGTTCCCGGGTCTTGTTACACACCGCCCCGTACACCCATGGGAGCGGGTTTTACCAGAAAGTGGGTAGCCTAACCCGCAAGGAGG

1440 1450 1460 1470 1480

cap_BA15_Y2(-)
cap_BA15_Y1(+)
cap_BA15_760r(-)
cap_BA15_362f(+)
cap_BA15_455r(-)
cap_BA15_760r(-)
cap_BA15_786f(+)
cap_BA15_Y3(-)
cap_BA15_1203f(+)
Contig-0

GCGCTCACGGCGGTAGgA
GCGCTCACCACGGTAGGATTTCGTGACTGGGGTGAAGTCGTAAACAAGGTAA
GCGCTCACCACGGTAGGATTTCGTGACTGGGGTGAAGTCGTAAACAAGGTAA

Y3

2.11 PROGRAMAS E MODELOS MATEMÁTICOS USADOS NA RECONSTRUÇÃO FILOGENÉTICA

As seqüências obtidas para os isolados diazotróficos de bananeira e abacaxizeiro foram usadas juntamente com seqüências da mesma região do 16S rDNA de estirpes referência das subdivisões alfa e beta das Proteobactérias. Todas as seqüências usadas para as análises filogenéticas não seqüenciadas no presente trabalho foram obtidas no banco de dados GenBank e encontram-se listadas na Tabela 7. Para cada uma das árvores filogenéticas inferidas, as seqüências foram inicialmente alinhadas no programa ClustalW (Tabela 6), usando-se os parâmetros padrões de análise.

Na etapa seguinte uma distância evolutiva foi calculada para cada par de seqüências do alinhamento, quando se pretendia inferir uma árvore filogenética pelo método Neighbor-Joining (APÊNDICE 3). Este método faz parte de um grupo de métodos genericamente denominados de Métodos de Distância. Eles utilizam uma matriz numérica, onde cada valor representa a estimativa da dissimilaridade (distância) entre duas seqüências. Por sua vez, as diferenças encontradas entre as seqüências são resultado de mutações ocorridas ao longo da evolução e a simples proporção de bases diferentes entre cada par de seqüências pode nos dar uma idéia da distância evolutiva entre os organismos. Entretanto, como já foi mencionado, devido ao fato de que algumas destas diferenças entre as seqüências são resultado de eventos múltiplos (duas ou mais mutações ocorrendo em um mesmo sítio) esta estimativa é atualmente uma sub-estimativa da distância evolutiva real que separa os organismos. Aplicando-se modelos evolutivos específicos, que fazem suposições à cerca da natureza da mudança evolutiva é possível fazer correções nestas distâncias, aproximando-as do real estado atual de divergência entre os organismos. Entretanto, uma vez que não temos um

registro histórico exato de eventos que tomaram lugar na evolução das seqüências, as correções da estimativa da distância evolutivas não são propriamente evidentes. A correção usada nas inferências filogenéticas feitas no presente trabalho foi o modelo de substituição de JUKES & CANTOR (1969). Este foi um dos primeiros modelos propostos e parte das seguintes pré-suposições:

- a) todas as substituições são independentes;
- b) todas as posições nas seqüências são igualmente sujeitas a mudanças;
- c) as substituições ocorrem casualmente entre os quatro tipos de nucleotídeos;
- d) inserções ou deleções não ocorreram.

A seguinte equação matemática é derivada deste modelo e foi empregada:

$$d_{AB} = -3/4 \ln (1 - 4/3 f_{AB}),$$

onde f_{AB} é a dissimilaridade (proporção de diferenças observadas) entre as seqüências “A” e “B” e d_{AB} é a estimativa da distância evolutiva entre as mesmas seqüências.

Após a construção da matriz de distâncias evolutiva, a topologia da árvore filogenética foi calculada pelo método Neighbor-Joining (APÊNDICE 3).

Árvores filogenéticas inferidas pelos métodos de Máxima Parcimônia e Máxima Verossimilhança também partiram dos alinhamentos, como feito para o método de *Neighbor-Joining*. Entretanto, estes dois métodos, usam os dados do próprio alinhamento para inferir a árvore filogenética, não havendo necessidade do cálculo das distâncias evolutivas (APÊNDICE 3).

As árvores construídas pelos métodos de Neighbor-Joining e Parcimônia (APÊNDICE 3) foram calculadas com o programa Mega2 e aquelas construídas pelo método da Máxima Verossimilhança foram calculadas com o programa DNAML do pacote PHYLIP. Os sítios contendo *gaps* em uma ou mais seqüências foram excluídos da análise. A análise de *bootstrap* (APÊNDICE 3) foi conduzida apenas para os

métodos de Neighbor-Joining e Parcimônia, também no programa Mega2, e suas especificações são indicadas nas respectivas figuras. Para o método de Máxima Verossimilhança, devido ao tempo de processamento ser proibitivo, não foram conduzidos testes de *bootstrap*. Todos os alinhamentos de seqüências de DNA necessários foram feitos com o programa ClustalW. O programa ForCon foi usado para a interconversão entre os formatos de arquivos de seqüências usados para conduzir as diferentes análises. Os programas BioEdit e GenDoc foram usados para edição das seqüências e alinhamentos. Todos os programas encontram-se listados na Tabela 6.

TABELA 7 - LISTA DOS ORGANISMOS PARA OS QUAIS A SEQÜÊNCIA DO 16S rDNA FORAM USADAS NAS ANÁLISES FILOGENÉTICAS.

ORGANISMO	No. DE ACESSO NO GenBank
2,4-D degrading bacterium TFD2	AF184927
2,4-D degrading bacterium TFD3	AF184929
<i>Acidovorax</i> sp. LW1	AJ130765
aniline degrading HY99	AF210313
<i>Aquaspirillum delicatum</i>	AF078756
<i>Aquaspirillum gracile</i>	AF078753
<i>Aquaspirillum psychrophilum</i>	AF078755
bacterium str 96444	AF227856
beta proteobacterium A0823	AF236011
beta proteobacterium A1020	AF236013
beta proteobacterium DSM 6920	Y11585
beta proteobacterium mz1L	AF110006
beta proteobacterium Wuba139	AF336363
beta proteobacterium Wuba68	AF336359
beta proteobacterium Wuba73	AF336362
<i>Burkholderia ambifaria</i>	AY028444
<i>Burkholderia brasiliensis</i>	AJ238360
<i>Burkholderia caribiensis</i>	Y17009
<i>Burkholderia caribiensis</i>	Y17010
<i>Burkholderia caribiensis</i>	Y17011
<i>Burkholderia caryophylli</i>	AB021423
<i>Burkholderia caryophylli</i>	U91570
<i>Burkholderia cepacia</i> KP24	AB015606
<i>Burkholderia cepacia</i>	AF097530
<i>Burkholderia cepacia</i>	AF097532
<i>Burkholderia cepacia</i>	AF097533
<i>Burkholderia cepacia</i>	AF175314
<i>Burkholderia cepacia</i>	AF311969
<i>Burkholderia cepacia</i>	AF311970
<i>Burkholderia cepacia</i>	AF311971
<i>Burkholderia cepacia</i>	AF311972
<i>Burkholderia cepacia</i>	U96927
<i>Burkholderia cepacia</i> DSM 50181	X87275
<i>Burkholderia cocovenenans</i>	AB021389
<i>Burkholderia gladioli</i>	AB012916
<i>Burkholderia gladioli</i>	U96934
<i>Burkholderia glathei</i>	U96935
<i>Burkholderia glumae</i>	U96931
<i>Burkholderia graminis</i>	U96939
<i>Burkholderia graminis</i>	U96940
<i>Burkholderia graminis</i>	U96941
<i>Burkholderia kururiensis</i>	AB024310
<i>Burkholderia mallei</i>	AF110187
<i>Burkholderia mallei</i>	AF110188
<i>Burkholderia phenazinium</i> ATCC 33666T	AB021394
<i>Burkholderia phenazinium</i>	U96936
<i>Burkholderia plantarii</i>	U96933
<i>Burkholderia pseudomallei</i>	AF093047
<i>Burkholderia pseudomallei</i>	AF093048
<i>Burkholderia pseudomallei</i>	AF093049
<i>Burkholderia pseudomallei</i>	AF093050

ORGANISMO	No. DE ACESSO NO GenBank
<i>Burkholderia pseudomallei</i>	AF093051
<i>Burkholderia pseudomallei</i>	AF093052
<i>Burkholderia pseudomallei</i>	AF093053
<i>Burkholderia pseudomallei</i>	AF093054
<i>Burkholderia pseudomallei</i>	AF093055
<i>Burkholderia pseudomallei</i>	AF093056
<i>Burkholderia pseudomallei</i>	AF093058
<i>Burkholderia pseudomallei</i>	AF093059
<i>Burkholderia pseudomallei</i>	AJ131790
<i>Burkholderia pseudomallei</i> 1026b	U91839
<i>Burkholderia pyrrocinia</i>	AB021369
<i>Burkholderia pyrrocinia</i>	U96930
<i>Burkholderia</i> sp. A6.33GD	AF247492
<i>Burkholderia</i> sp. AB101	AF219126
<i>Burkholderia</i> sp. AB2	AF219125
<i>Burkholderia</i> sp. C3B1M	U96938
<i>Burkholderia</i> sp. C4 7	AF247493
<i>Burkholderia</i> sp. CAB 02	AB041730
<i>Burkholderia</i> sp. DhA54	AJ011508
<i>Burkholderia</i> sp. EN B3	AF074711
<i>Burkholderia</i> sp. EN B9	AF074712
<i>Burkholderia</i> sp. IpA51	AJ011509
<i>Burkholderia</i> sp. JS150	AF262932
<i>Burkholderia</i> sp. LMG16225	AF215705
<i>Burkholderia</i> sp. LMG16307	AF215706
<i>Burkholderia</i> sp. LMG19076	AF215704
<i>Burkholderia</i> sp. MBIC3837	AB008002
<i>Burkholderia</i> sp. N2P5	U37342
<i>Burkholderia</i> sp. N2P6	U37343
<i>Burkholderia</i> sp. N3P2	U37344
<i>Burkholderia</i> sp. NF100DNA	AB025790
<i>Burkholderia</i> sp. S21	AF247494
<i>Burkholderia</i> sp. S411	AF247495
<i>Burkholderia stabilis</i>	AF148554
<i>Burkholderia thailandensis</i> E264	U91838
<i>Burkholderia vietnamiensis</i> AMMD	AF043302
<i>Burkholderia vietnamiensis</i> LMG 10929	AF097534
<i>Comamonas acidovorans</i>	AB021417
<i>Comamonas</i> sp. 12022	AF078773
<i>Comamonas</i> sp. 158	AJ002803
<i>Comamonas</i> sp. R7	AJ002810
<i>Delftia acidovorans</i>	AB020186
<i>Delftia acidovorans</i> ACM 489	AF078774
<i>Delftia acidovorans</i>	AF149849
<i>Duganella zoogloeoides</i> IAM 12670	D14256
<i>Herbaspirillum frisingense</i> 75B	AJ238357
<i>Herbaspirillum frisingense</i> GSF30	AJ238358
<i>Herbaspirillum frisingense</i> Mb11	AJ238359
<i>Herbaspirillum rubrisubalbicans</i>	AB021424
<i>Herbaspirillum rubrisubalbicans</i> ICMP 5777	AF137508
<i>Herbaspirillum rubrisubalbicans</i>	AJ238356
<i>Herbaspirillum seropedicae</i> Z67	AJ238361
<i>Herbaspirillum seropedicae</i>	Y10146
<i>Herbaspirillum</i> sp. B501	AB049133

ORGANISMO	No. DE ACESSO NO GenBank
<i>Herbaspirillum</i> sp. B512	AB049104
<i>Herbaspirillum</i> sp. B65	AB049103
<i>Hydrogenophaga intermedia</i> S1	AF019037
<i>Janthinobacterium agaricidamnosum</i>	Y08845
<i>Janthinobacterium lividum</i>	AF174648
<i>Janthinobacterium lividum</i> DSM 1522T	Y08846
<i>Massilia timonae</i>	U54470
<i>Matsuebacter</i> sp. 9	AB024305
nitrogen fixing bacterium COL	AF214642
<i>Oxalobacter formigenes</i>	U49749
<i>Oxalobacter formigenes</i>	U49751
<i>Oxalobacter formigenes</i>	U49753
<i>Oxalobacter formigenes</i>	U49754
<i>Oxalobacter formigenes</i>	U49755
<i>Oxalobacter formigenes</i>	U49756
<i>Oxalobacter formigenes</i>	U49757
<i>Oxalobacter formigenes</i>	U49758
<i>Pandoraea apista</i>	AF139172
<i>Pandoraea apista</i>	AF139173
<i>Pandoraea pnomenusa</i>	AF139174
<i>Pandoraea pulmonicola</i>	AF139175
<i>Pandoraea</i> sp. G3307	AF247699
<i>Pandoraea</i> sp. G3308	AF247691
<i>Pandoraea</i> sp. G5056	AF247692
<i>Pandoraea</i> sp. G7835	AF247694
<i>Pandoraea</i> sp. G8107	AF247695
<i>Pandoraea</i> sp. G9805	AF247697
<i>Paucimonas lemoignei</i> LMG 16480	X92554
<i>Paucimonas lemoignei</i>	X92555
<i>Pseudomonas huttiensis</i>	AB021366
<i>Pseudomonas lanceolata</i>	AB021390
<i>Pseudomonas lemoignei</i>	AB021375
<i>Pseudomonas mephitica</i>	AB021388
<i>Pseudomonas</i> sp. LB400	U86373
rape rhizosp. here tsb068	AJ295450
rape rhizosp. here tsb077	AJ295454
<i>Rhodoferax fermentans</i> FR2	D16211
ultramicrobacterium D-6	AB008504
ultramicrobacterium D-7	AB008505
ultramicrobacterium MY14	AB008503
ultramicrobacterium ND5	AB008506
uncultured bacterium	AF358019
uncultured bacterium FukuN55	AJ289999
uncultured bacterium GKS2 122	AJ290026
uncultured bacterium SJA 62	AJ009470
uncultured <i>Duganella</i> CTHB-18	AF067655
uncultured eubacterium WD202	AJ292637
uncultured eubacterium WD2100	AJ292623
uncultured eubacterium WD2114	AJ292647
uncultured eubacterium WD2115	AJ292627
uncultured eubacterium WD2116	AJ292648
uncultured eubacterium WD227	AJ292639
uncultured eubacterium WD240	AJ292622
uncultured eubacterium WD258	AJ292640

ORGANISMO	No. DE ACESSO NO GenBank
uncultured eubacterium WD263	AJ292641
uncultured eubacterium WD266	AJ292642
uncultured eubacterium WD268	AJ292643
uncultured eubacterium WD285	AJ292644
uncultured eubacterium WD289	AJ292645
uncultured eubacterium WD291	AJ292624
uncultured eubacterium WD293	AJ292625
uncultured hydroc seep BPC087	AF154097
uncultured rape rhizosp.h wr0040	AJ295498
uncultured sludge A6b	AF234720
uncultured sludge H21	AF234687
unidentified beta proteobacterium 1040	U34035
Unidentified bacteria HW1	D89026
unidentified bacterium	X95838
unidentified bacterium 19523	AF097797
unidentified bacterium rA10	AB021361
unidentified bacterium rJ1	AB021320
unidentified bacterium rJ16	AB021334
unidentified bacterium rM5	AB021340
unidentified bacterium rM6	AB021341
unidentified bacterium T33	Z93960
unidentified bacterium T41	Z93973
unidentified bacterium T47	Z93977
unidentified bacterium T70	Z93974
unidentified bacterium T71	Z93968
unidentified bacterium T83	Z93969
unidentified BD1 33	AB015530
unidentified BD5 11	AB015567
unidentified HTB091	AB010869
unidentified HTC018	AB010865
unidentified beta proteobacterium MC 13289	AB006750
unidentified proteobacterium K1	AJ001384
<i>Variovorax paradoxus</i>	AB008000
<i>Variovorax paradoxus</i>	AF209469
<i>Variovorax paradoxus</i>	AF250030
<i>Variovorax paradoxus</i> IAM 12373	D30793
<i>Variovorax paradoxus</i>	D88006
<i>Vestimentiferan symbiont</i> TW5	AB042420
<i>Xylophilus ampelinus</i> ATCC 33914	AF078758
<i>Zoogloea ramigera</i> ATCC 25935	X74914
<i>Burkholderia norimbergensis</i>	Y09879
<i>Burkholderia cepacia</i> ATCC 25416	M22518
<i>Burkholderia</i> sp. CRE 7	U37340
<i>Burkholderia</i> sp. JB1	X92188
<i>Burkholderia glathei</i> ATCC 29195T	AB021374
<i>Burkholderia multivorans</i> LMG 14293	AF097531
<i>Alcaligenes faecalis</i> IAM 12369	D88008
<i>Herbaspirillum</i> sp. G8A1	AJ012069
uncultured proteobacterium OCS7	AF001645
<i>Oxalobacter formigenes</i> BLISS	U49750
<i>Acidovorax facilis</i> CCUG 2113	AF078765
<i>Acidovorax temperans</i> CCUG 11779	AF078766
<i>Acidovorax delafieldii</i> ATCC 17505	AF078764
<i>Acidovorax</i> sp. 7078	AF078767

ORGANISMO	No. DE ACESSO NO GenBank
<i>Acidovorax avenae</i> subsp. <i>avenae</i> ATCC 19860	AF078759
<i>Acidovorax</i> sp. IMI 357678	AF078763
<i>Acidovorax koniaki</i> ATCC 33996	AF078760
<i>Comamonas terigena</i> IMI 359870	AF078772
<i>Comamonas testosteroni</i> MBIC 3841	AB007997
<i>Hydrogenophaga flava</i> CCUG 1658	AF078771
<i>Hydrogenophaga pseudoflava</i> ATCC 33668	AF078770
<i>Hydrogenophaga palleronii</i> DSM 63	AF019073
<i>Hydrogenophaga taeniospiralis</i> ATCC 49743	AF078768
<i>Matsuebacter chitosanotabidus</i>	AB006851
<i>Ideonella dechloratans</i>	X72724
<i>Rubrivivax gelatinosus</i> IC 144	AB016167
<i>Leptothrix discophora</i> SS-1	L33975
<i>Leptothrix cholodnii</i> CCM 1827	X97070
<i>Leptothrix mobilis</i> Feox-1	X97071
<i>Escherichia coli</i>	J01859
<i>Acetobacter aceti</i>	X74066
<i>Azospirillum amazonense</i> Y2	X79742
<i>Azospirillum brasilense</i> Sp7	X79739
<i>Azospirillum lipoferum</i> ATCC 29708	X79729
<i>Azospirillum lipoferum</i> F	X79736
<i>Azospirillum lipoferum</i> WO3	X79741
<i>Azospirillum doebereineriae</i>	AJ238567
<i>Bradyrhizobium japonicum</i>	X87272
<i>Beijerinckia indica</i>	M59060
<i>Bradyrhizobium elkanii</i>	U35000
<i>Caulobacter bacteroides</i>	M83796
<i>Caulobacter fusiformis</i>	AJ007803
<i>Caulobacter henricii</i>	AB016846
<i>Caulobacter intermedius</i>	AJ007802
<i>Caulobacter variabilis</i>	AJ007808
<i>Methylobacterium organophilum</i>	D32226
<i>Methylobacterium rhodesianum</i>	D32228
<i>Methylobacterium zatmanii</i>	D32230
<i>Nitrobacter hamburgensis</i>	
<i>Nitrobacter winogradskyi</i>	L11661
<i>Ochrobactrum anthropi</i>	U88440
<i>Ochrobactrum intermedium</i>	AJ242583
<i>Ochrobactrum tritici</i>	AJ242579
<i>Octadecabacter arcticus</i>	U73725
<i>Pseudomonas denitrificans</i>	X69159
<i>Paracraurococcus ruber</i>	D85827
<i>Rhizobium leguminosarum</i> LMG 8819	X77122
<i>Rhizobium litoralis</i> ATCC 49566	X78312
<i>Rhizobium thiosulfatophilus</i>	X72908
<i>Rhizobium etli</i>	U47303
<i>Rhizobium fredii</i>	D12792
<i>Rhizobium gallicum</i>	AF008126
<i>Rhizobium hainanensis</i>	U71078
<i>Rhizobium leguminosarum</i>	M55235
<i>Rhizobium loti</i>	D14514
<i>Rhizobium mongolense</i>	U89820
<i>Rhizobium tropici</i>	AF260294
<i>Rhodobium orientis</i>	D30792

conclusão

ORGANISMO	No. DE ACESSO NO GenBank
<i>Rhodopila globiformis</i>	D86513
<i>Rhodoplanes roseus</i>	D25313
<i>Rhodopseudomonas palustris</i>	D25312
<i>Rhodospirillum fulvum</i> NCIMB 11762	D14433
<i>Rhodospirillum rubrum</i>	D30778
<i>Rickettsia prowazekii</i>	M21789
<i>Shinorhizobium xinjiangensis</i>	D12796
<i>Sphingomonas adhaesiva</i>	D16146
<i>Sphingomonas agrestis</i>	Y12803
<i>Sphingomonas echinoides</i>	AB033944
<i>Sphingomonas paucimobilis</i>	D16144
<i>Sulfitobacter pontiacus</i>	Y13155
<i>Xanthobacter autotrophicus</i>	X94203

3 RESULTADOS

3.1 AMPLIFICAÇÃO DO 16S rDNA E DA REGIÃO INTERGÊNICA 16S-23S rDNA

A amplificação do 16S rDNA, pelos iniciadores Y1 e Y3, de 38 isolados diazotróficos de abacaxizeiro e bananeira e das 6 estirpes de espécies diazotróficas referência, produziu um único fragmento de tamanho molecular entre 1.018 e 1.636 pb (Figura 4). A partir da seqüência de *Herbaspirillum seropedicae* (número de acesso no GenBank Y10146) e *Escherichia coli* (BROSIUS *et al.*, 1981) foi verificado que os tamanhos esperados para os produtos da PCR com estes iniciadores seriam de 1.479 e 1.487 pb, respectivamente. Portanto, a migração dos fragmentos amplificados ficou dentro da faixa esperada.

Para a região intergênica 16S-23S rDNA, também foi encontrado um único fragmento como produto de amplificação para a maioria dos isolados, com um tamanho molecular aproximado de 1.100 pb (Figura 5). Entretanto, três outros padrões foram observados: o primeiro contendo um fragmento único com um peso molecular de cerca de 1.000 pb, observado para o isolado BA131; o segundo apresentando também uma banda única, correspondendo a um fragmento com peso molecular entre 700 a 800 pb, observado para o isolado Ala; e um terceiro padrão, detectado nos isolados A2a e A3b, contendo dois fragmentos com pesos moleculares correspondentes aos padrões anteriores (Figura 5).

FIGURA 4 - PRODUTO DE AMPLIFICAÇÃO POR PCR DA REGIÃO Y1-Y3 DO 16S rDNA DE ISOLADOS DIAZOTRÓFICOS DE BANANEIRA. ELETROFORESE EM GEL DE AGAR 1% (60 V POR 40 MIN). 1. 1 KB DNA LADDER (GIBCO BRL), 2. BA14 E 3. BA153.

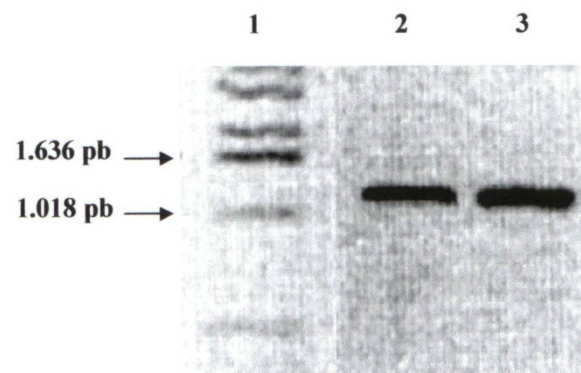
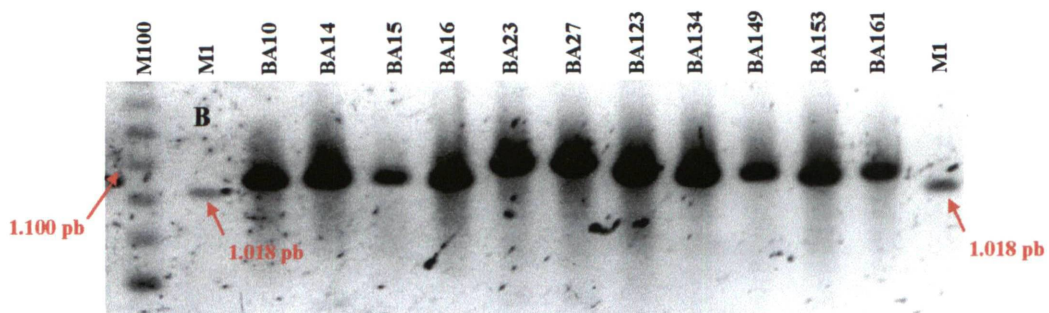
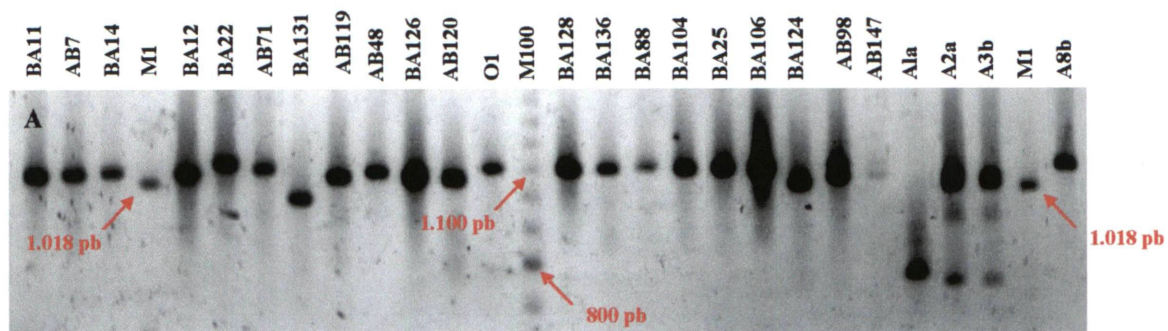


FIGURA 5. A E B. PRODUTOS DE AMPLIFICAÇÃO DA REGIÃO INTERGÊNICA 16S-23S rDNA DE ISOLADOS DIAZOTRÓFICOS E ESTIRPES DE *H. seropedicae* E *H. rubrisubalbicans*; M1 = 1 KB DNA LADDER (GIBCO BRL) E M100 - 100 BP DNA LADDER (GIBCO BRL).

3.2



3.2 POLIMORFISMO DE FRAGMENTOS DE COMPRIMENTO ALEATÓRIO (RFLP) DA REGIÃO INTERGÊNICA 16S-23S rDNA

A restrição dos produtos de amplificação da região intergênica 16S-23S rDNA dos isolados diazotróficos com as endonucleases de restrição *AluI*, *HaeIII* e *RsaI* produziram diversos fragmentos que foram separados, por sua massa molecular, em eletroforese em gel de agarose. Os fragmentos produzidos por uma endonuclease específica apresentaram um padrão de bandas, que foi denominado perfil de restrição, representado por letras minúsculas na Tabela 8. Cada perfil de restrição é compartilhado somente por aqueles organismos que possuem os mesmos sítios de restrição conservados, ou seja, os organismos que apresentaram o mesmo perfil de restrição para uma determinada enzima possuem a mesma letra na Tabela 8. Para os isolados diazotróficos a endonuclease *RsaI* mostrou-se com maior poder de resolução, tendo sido capaz de revelar 13 diferentes perfis de restrição. A endonuclease *HaeIII* revelou 11 perfis e *AluI*, 10 perfis. Analisando-se os perfis obtidos com as três endonucleases conjuntamente, uma maior resolução foi obtida no agrupamento dos isolados, sendo detectados 15 diferentes padrões, denominados Grupos 16S-23S e representados por letras maiúsculas na Tabela 8, entre os 34 isolados analisados (Figuras 5, 6 e 7). A maior parte dos Grupos 16S-23S obtidos (B, C, F, H, I, J, K, L e N) foram compostos por um único isolado representante. Três outros Grupos 16S-23S (D, G e M) foram compostos por dois isolados cada um e somente dois Grupos 16S-23S (A e E) foram compostos por oito ou mais isolados.

TABELA 8 - CARACTERIZAÇÃO DE BACTÉRIAS DIAZOTRÓFICAS POR RFLP DA REGIÃO INTERGÊNICA 16S-23S rDNA.

ESTIRPE	PERFIL DE RESTRIÇÃO ⁽¹⁾			GRUPOS 16S-23S ⁽²⁾
	<i>AluI</i>	<i>HaeIII</i>	<i>RsaI</i>	
BA10	a	a	a	A
BA11	a	a	a	A
BA12	a	a	a	A
BA14	a	a	a	A
BA15	a	a	a	A
BA16	a	a	a	A
BA134	a	a	a	A
BA149	a	a	a	A
BA153	a	a	a	A
BA161	a	a	a	A
X8	a	a	a	A
AB7	b	b	b	B
BA124	c	c	c	C
AB98	d	d	d	D
AB147	d	d	d	D
O1	e	e	e	E
BA22	e	e	e	E
BA23	e	e	e	E
BA25	e	e	e	E
BA88	e	e	e	E
BA104	e	e	e	E
BA106	e	e	e	E
BA128	e	e	e	E
BA136	e	e	e	E
AB120	d	f	f	F
BA123	f	g	f	G
BA126	f	g	f	G
AB48	d	a	g	H
AB71	d	a	f	I
AB119	g	h	h	J
Ala	h	i	i	K
BA131	i	j	j	L
A2a	j	k	k	M
A3b	j	k	k	M
A8b	e	e	l	N
BA27	g	n. d. ⁽³⁾	m	O

(1) Padrão de bandas obtido na eletroforese em gel de agarose após restrição dos produtos de amplificação para a região intergênica 16S-23S rDNA com as endonucleases indicadas (mostrados nas Figuras de 6 a 8).

(2) Padrão de bandas obtido para o conjunto das três endonucleases usadas na análise.

(3) nenhum dado.

FIGURA 6 - A E B: PERFIL DE RESTRIÇÃO DOS AMPLIFICADOS DA REGIÃO INTERGÊNICA 16S-23S rDNA COM *AluI*; AS LETRAS EM CAIXA BAIXA REPRESENTAM OS PERFIS DE RESTRIÇÃO DA TABELA 8; MARCADOR: M100 = 100 BP DNA LADDER (GIBCO BRL).

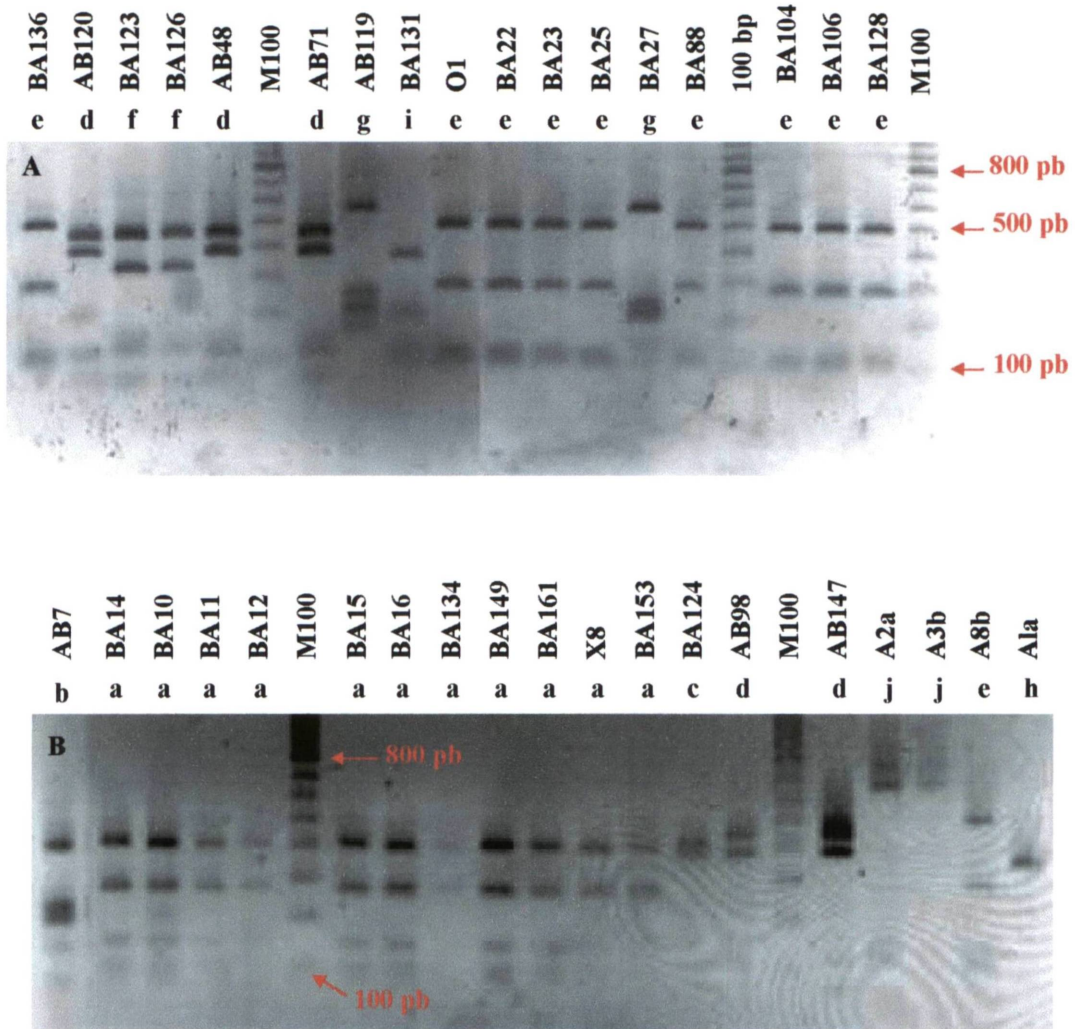


Figura 7 - A e B: perfil de restrição dos amplificadas da região intergênica 16S-23S rDNA com *Hae*III; as letras em caixa baixa representam os perfis de restrição da tabela 8; marcador: M100 = 100 bp DNA ladder (GIBCO BRL).

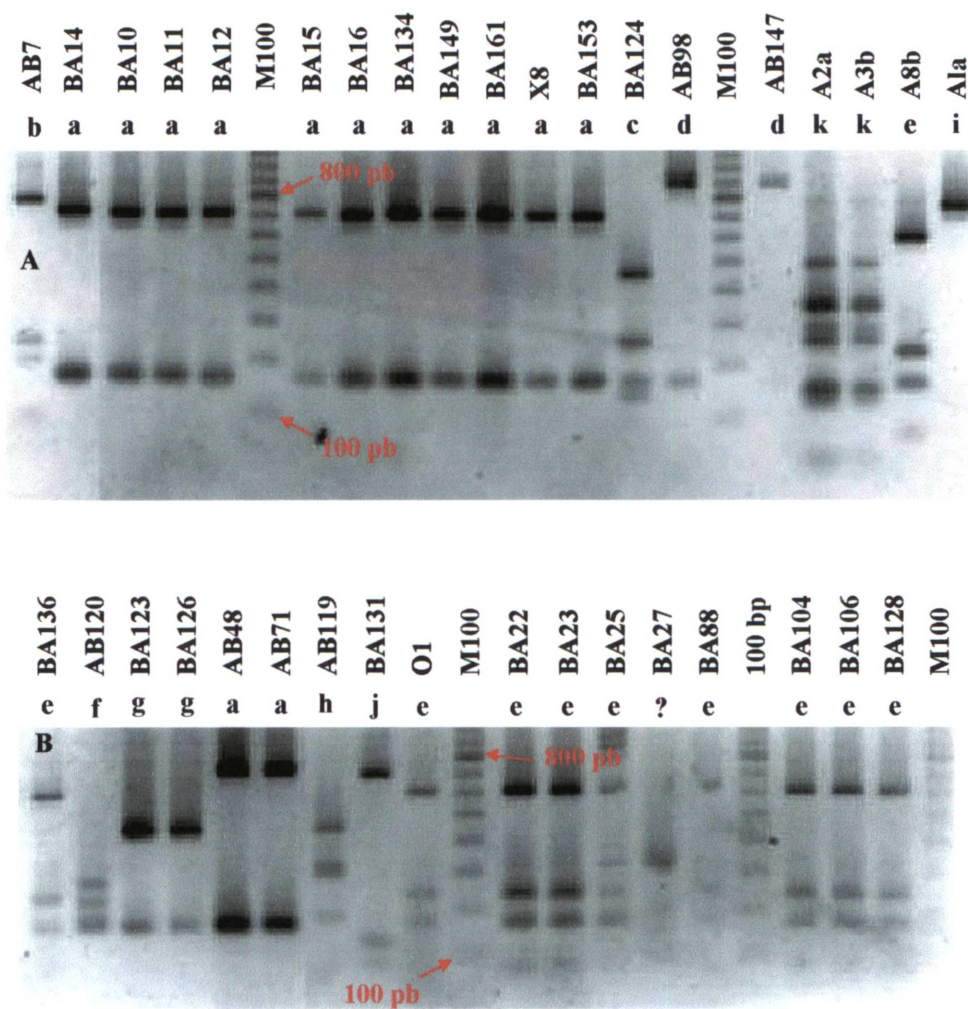
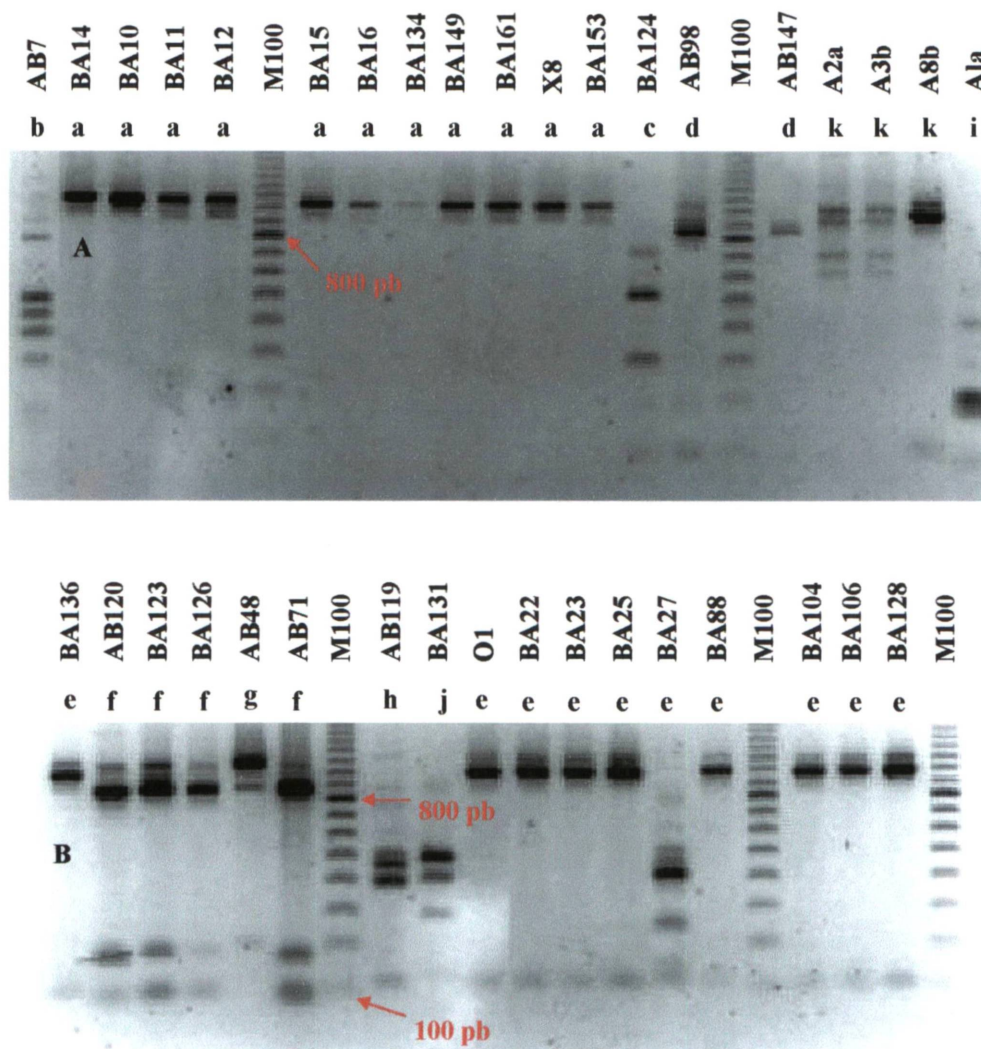


FIGURA 8 - A E B: PERFIL DE RESTRIÇÃO DOS AMPLIFICADOS DA REGIÃO INTERGÊNICA 16S-23S rDNA COM *Rsa*I; AS LETRAS EM CAIXA BAIXA REPRESENTAM OS PERFIS DE RESTRIÇÃO DA TABELA 8; MARCADOR: M100 = 100 BP DNA LADDER (GIBCO BRL).



3.3 ANÁLISE DE RESTRIÇÃO DO DNA RIBOSSOMAL DE 16S AMPLIFICADO (ARDRA)

Os fragmentos amplificados com os primers Y1 e Y3 foram tratados com *AluI*, *HaeIII*, *HinfI* e *RsaI*, produzindo de 3 a 7 fragmentos detectados pela eletroforese em gel de agarose (Figuras de 8 a 12). Entre 5 a 10 padrões únicos foram obtidos para cada enzima individualmente e estão representados por letras minúsculas na Tabela 9. A enzima *HaeIII* sozinha foi capaz de resolver a maioria dos grupos (dez), seguido pela enzima *AluI* (sete) e *HinfI/RsaI* (cinco; Tabela 9).

Doze diferentes padrões de restrição foram obtidos pela análise conjunta das quatro endonucleases de restrição, representando 12 Grupos ARDRA diferentes que foram numerados em arábico na Tabela 9. Quatro grupos foram idênticos àqueles obtidos para as estirpes usadas como referência e puderam ser especificamente identificados. Dois isolados (BA153 e X8) apresentaram perfil semelhante ao obtido para as estirpes Z67, Z78 e M2 de *Herbaspirillum seropedicae*; onze isolados (AB7, BA10, BA11, BA12, BA14, BA15, BA16, BA17, BA134, BA149 e BA161) tiveram o mesmo padrão obtido para a estirpe M4 de *Herbaspirillum rubrisubalbicans*; um isolado (BA124) teve o mesmo padrão obtido para a estirpe M130 de *Burkholderia brasilensis*; e, finalmente, dois isolados (AB98 e AB147) tiveram o mesmo padrão obtido para a estirpe Ppe8 de *Burkholderia tropicalis*. Entretanto, 22 isolados produziram oito novos Grupos ARDRA, grupos de 5 a 12 (Tabela 9).

Vários Grupos ARDRA apresentaram-se muito próximos, sendo diferenciados por um único padrão enzimático ou por um único fragmento de restrição. Os Grupos ARDRA 1 (*H. seropedicae*) e 2 (*H. rubrisubalbicans*) compartilharam todos os fragmentos de restrição, exceto um único fragmento obtido com *AluI* (Figura 9A). Entre os Grupos ARDRA 10 e 12, contendo isolados não identificados, a diferenciação ocorreu por dois fragmentos obtidos com *AluI* (ex., A8b e Ala - Figura 9B). Por outro lado, os Grupos ARDRA 6 e 8, também contendo

espécies não identificadas, foram somente diferenciados pelo perfil de restrição obtido com *HaeIII* (ex., AB71 e AB117 - figura 10B).

Esta relação próxima entre as estirpes testadas é observada na topologia obtida para um dendograma calculado pelo método UPGMA (APÊNDICE 3), a partir de distâncias determinadas pelo método de NEI & LI (1979), o qual permite estimar o número de substituições de nucleotídeos entre duas populações ou espécies através da mudança evolutiva de sítios de restrição (Figura 13). Os resultados obtidos para a endonuclease *RsaI* foram excluídos desta análise porque o isolado BA131 não foi testado com esta enzima (Tabela 9). O dendograma produzido mostrou três grupos principais. O primeiro é formado pelos Grupos ARDRA 1 e 2 e inclui as estirpes de *H. seropedicae* e *H. rubrisubalbicans* usadas como referência; o segundo é formado pelos Grupos ARDRA de 3 a 9 e inclui as estirpe M130 de *B. brasiliensis* (Grupo ARDRA 3) e Ppe8 de *B. tropicalis* (Grupo ARDRA 4); e o terceiro grupo, formado pelos Grupos ARDRA de 10 a 12, os quais se mostraram muito distantes dos demais. O grupo ARDRA 5 separa os grupos "*Burkholderia*" e "*Herbaspirillum*".

TABELA 9 - CARACTERIZAÇÃO DE BACTÉRIAS DIAZOTRÓFICAS, POR ARDRA, DA REGIÃO Y1-Y3 DO 16S RDNA.

continua

ESTIRPE ⁽¹⁾	HOSPEDEIRO	PERFIS DE RESTRIÇÃO ⁽²⁾				GRUPOS ARDRA ⁽³⁾
		<i>AluI</i>	<i>HaeIII</i>	<i>HinfI</i>	<i>RsaI</i>	
<i>Herbaspirillum seropedicae</i>						
Z67		a	a	a	a	1
Z78		a	a	a	a	1
M2		a	a	a	a	1
BA153	banana	a	a	a	a	1
X8	n. d. ⁽⁴⁾	a	a	a	a	1
<i>Herbaspirillum rubrisubalbicans</i>						
M4		b	a	a	a	2
AB7	abacaxizeiro	b	a	a	a	2
BA10	banana	b	a	a	a	2
BA11	banana	b	a	a	a	2
BA12	banana	b	a	a	a	2
BA14	banana	b	a	a	a	2
BA15	banana	b	a	a	a	2
BA16	banana	b	a	a	a	2
BA17	banana	b	a	a	a	2
BA134	banana	b	a	a	a	2
BA149	banana	b	a	a	a	2
BA161	banana	b	a	a	a	2
<i>Burkholderia brasilensis</i>						
M130		c	b	b	b	3
BA124	banana	c	b	b	b	3
<i>Burkholderia tropicalis</i>						
Ppe8		d	c	c	b	4
AB98	abacaxizeiro	d	c	c	b	4
AB147	abacaxizeiro	d	c	c	b	4
Unknown						
O1	banana	g	e	e	d	5
BA22	banana	g	e	e	d	5
BA23	banana	g	e	e	d	5
BA25	banana	g	e	e	d	5
BA27	banana	g	e	e	d	5
BA88	banana	g	e	e	d	5
BA104	banana	g	e	e	d	5
BA106	banana	g	e	e	d	5
BA128	banana	g	e	e	d	5
BA136	banana	g	e	e	d	5
AB117	abacaxizeiro	g	f	c	c	6
AB120	abacaxizeiro	g	f	c	c	6
BA123	banana	d	g	c	e	7
BA126	banana	d	g	c	e	7

ESTIRPE ⁽¹⁾	HOSPEDEIRO	Perfis de Restrição ⁽²⁾				GRUPOS ARDRA ⁽³⁾
		<i>AluI</i>	<i>HaeIII</i>	<i>HinfI</i>	<i>RsaI</i>	
AB48	abacaxizeiro	g	h	c	c	8
AB71	abacaxizeiro	g	h	c	c	8
AB119	abacaxizeiro	d	l	b	b	9
Ala	n. d. ⁽⁴⁾	f	d	d	c	10
BA131	banana	b	i	f	n. d.	11
A2a	n. d.	e	d	d	c	12
A3b	n. d.	e	d	d	c	12
A8b	n. d.	e	d	d	c	12

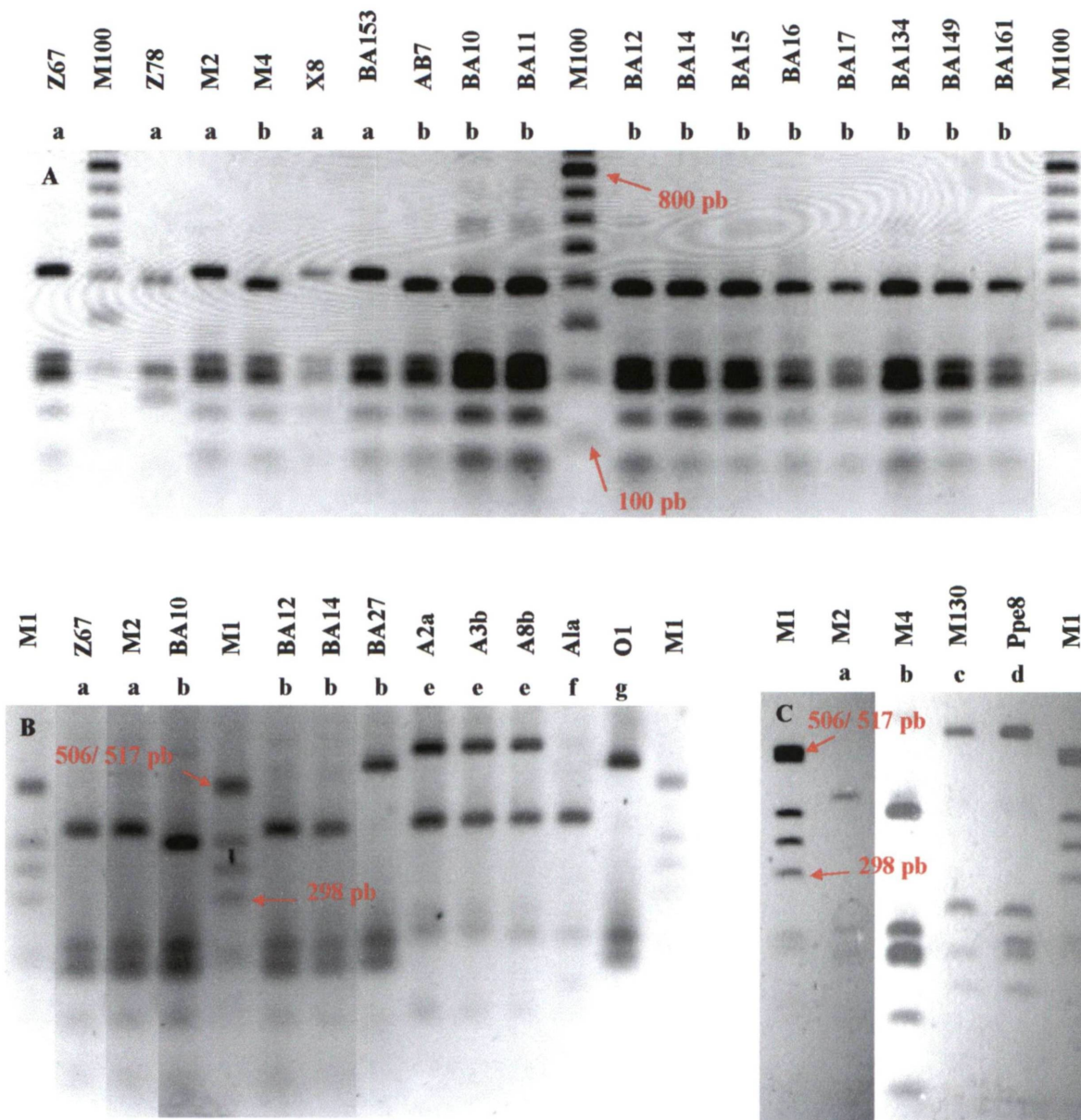
(1) As estirpes usadas como referência foram: Z67, Z78 e M2 de *Herbaspirillum seropedicae*, M4 de *Herbaspirillum rubrisubalbicans*, Ppe8 de "*Burkholderia tropicalis*", e M130 de "*Burkholderia brasilensis*".

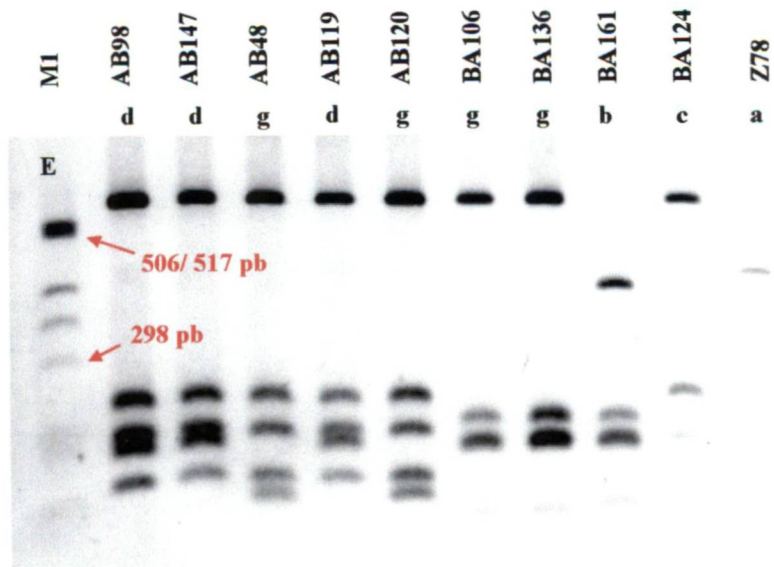
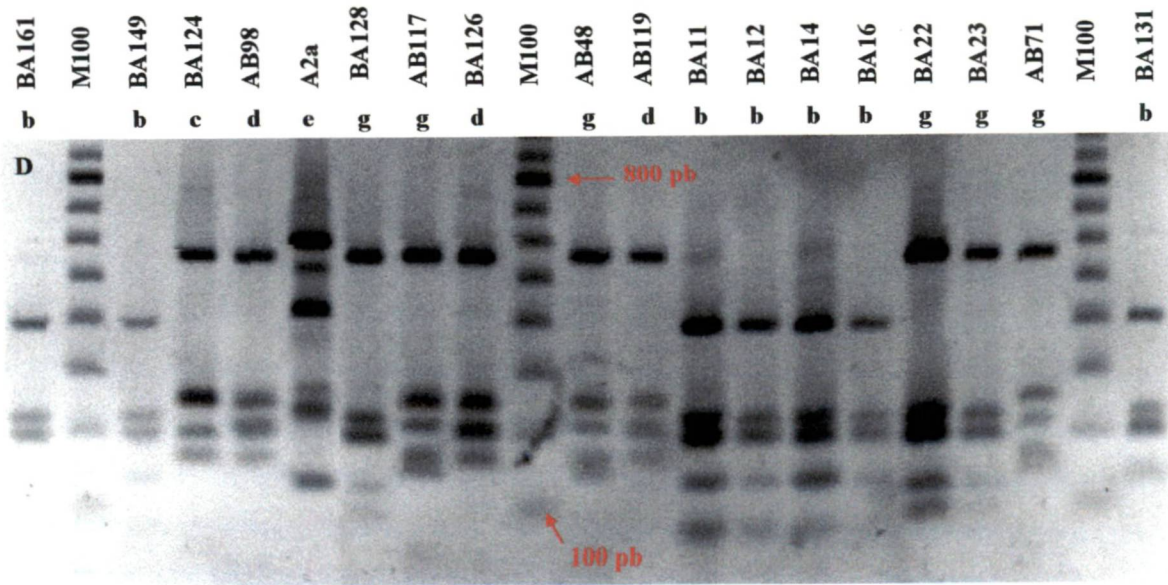
(2) Perfil de mobilidade dos fragmentos de DNA em gel de agarose após restrição dos produtos de amplificação para a região Y1-Y3 do 16S rDNA com as endonucleases indicadas.

(3) Padrão obtido para o conjunto das quatro endonucleases usadas na análise.

(4) nenhum dado.

FIGURA 9 - DE A A G: PERFIS DE RESTRIÇÃO PARA ISOLADOS DE ABACAXIZEIRO E BANANEIRA E ESTIRPES DE *Herbaspirillum* sp. E *Burkholderia* sp. COM A ENDONUCLEASE *AluI* NA REGIÃO AMPLIFICADA Y1-Y3 DO 16S rDNA; LETRAS EM CAIXA BAIXA REPRESENTAM OS PERFIS DE RESTRIÇÃO DA TABELA 9; MARCADORES: M1 = 1 KB DNA LADDER (GIBCO BRL) E M100 = 100 BP DNA LADDER (GIBCO BRL).





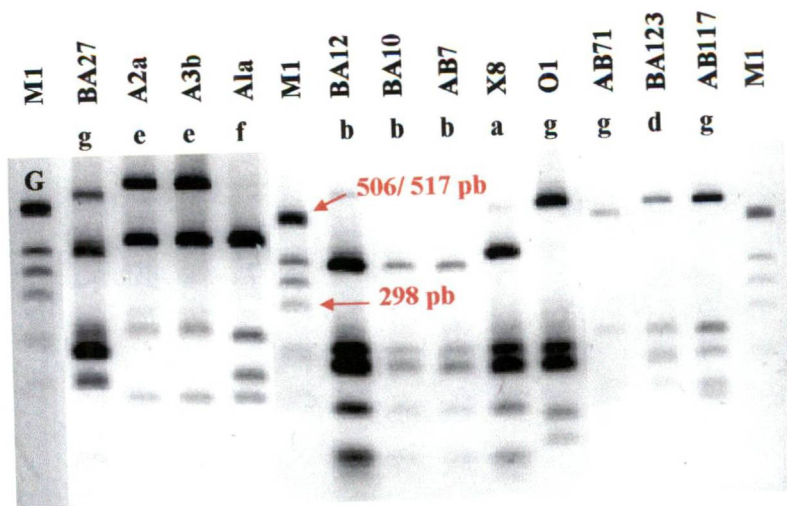
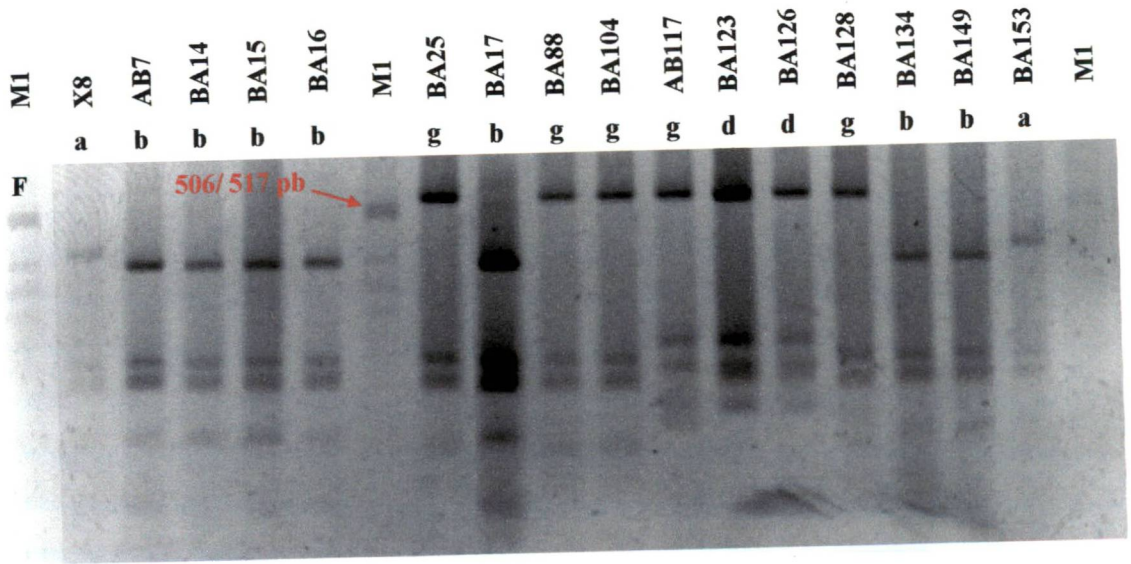
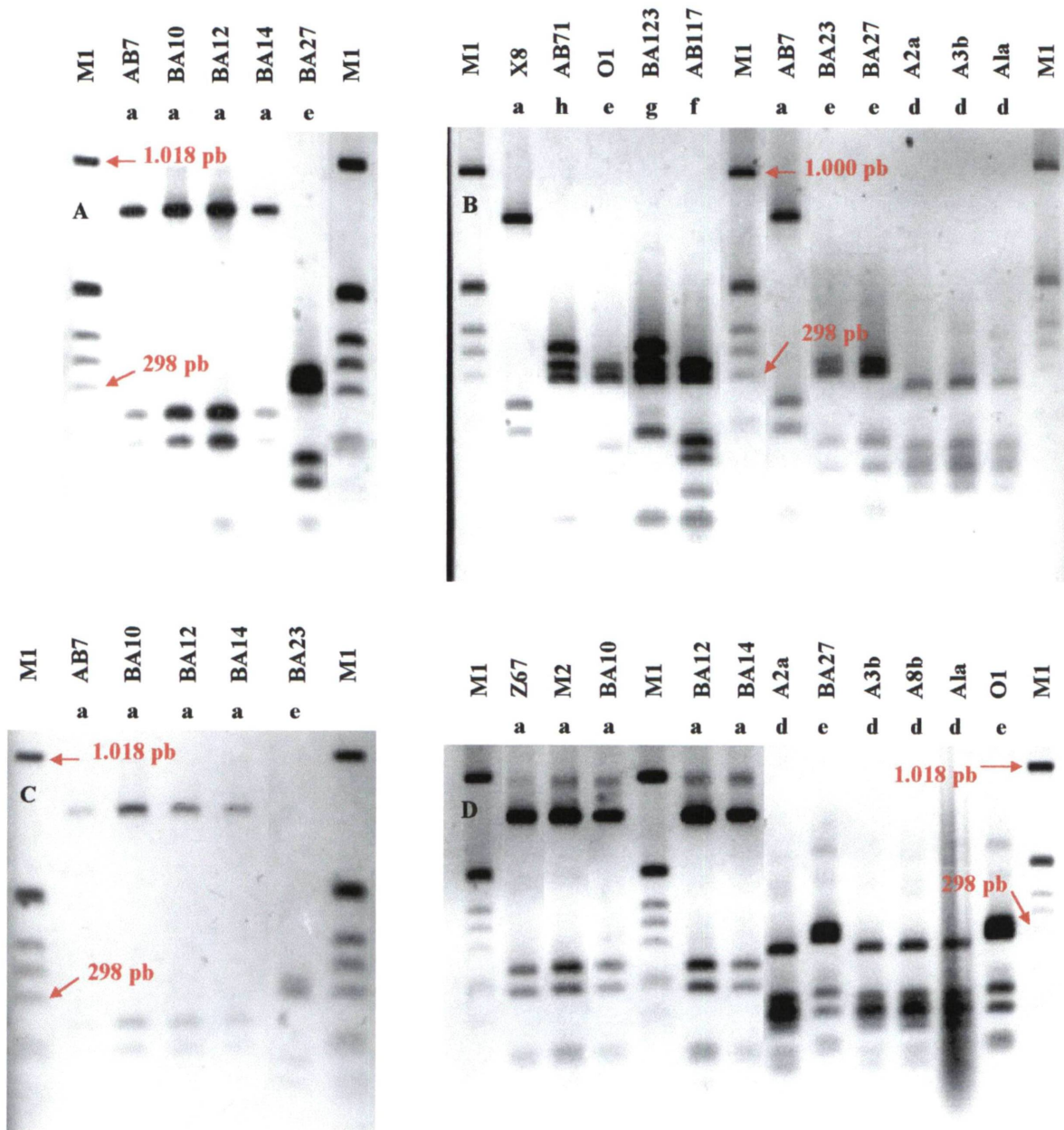
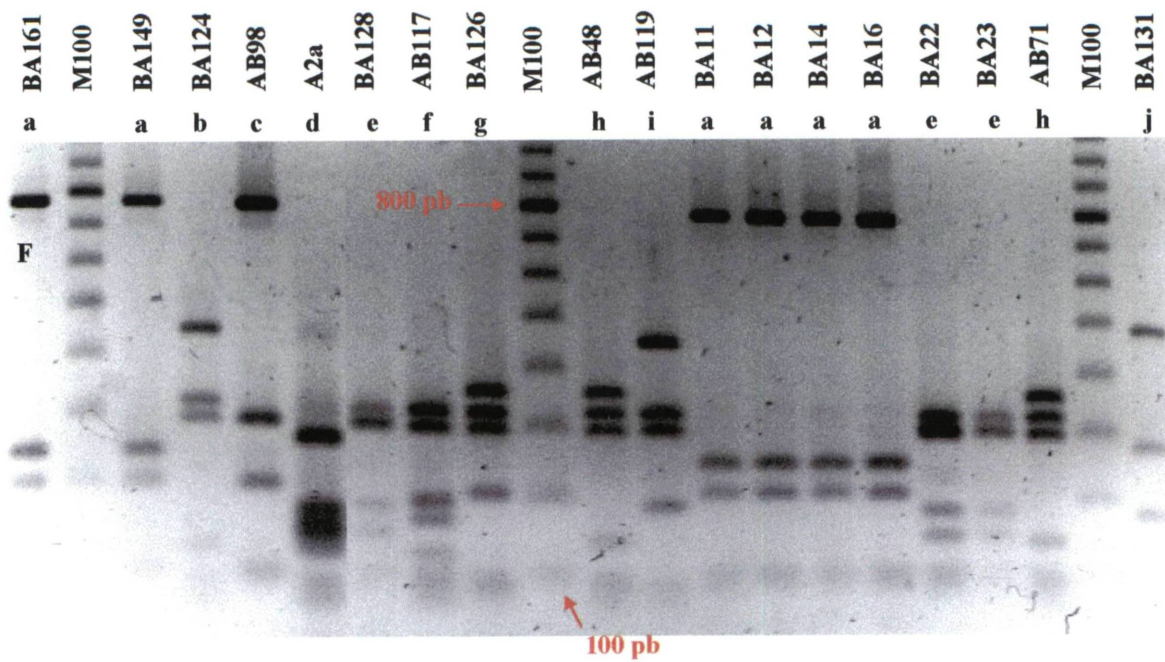
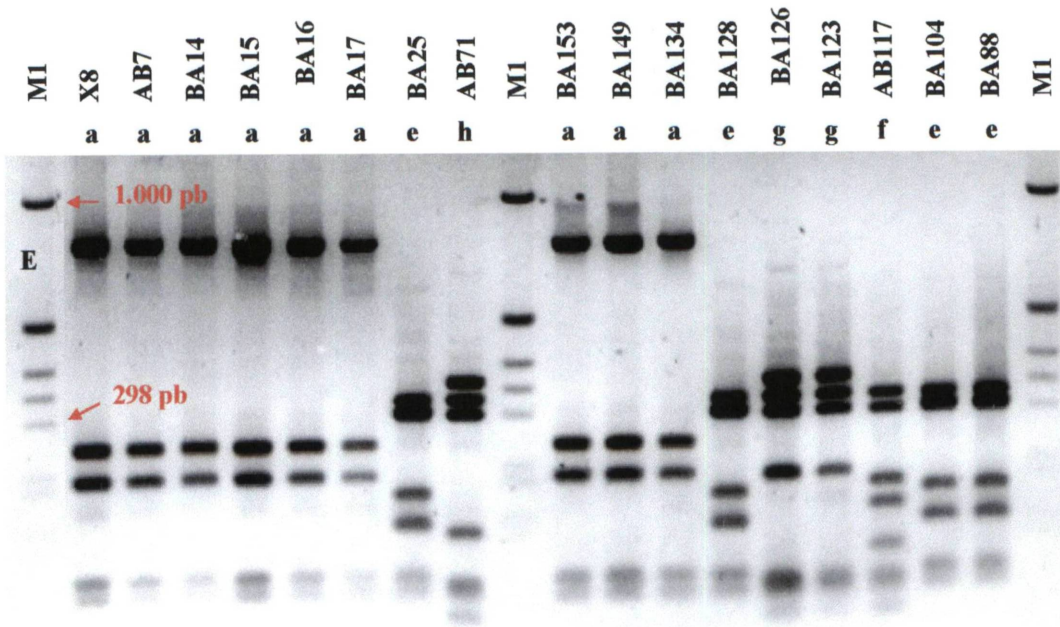


FIGURA 10 - DE A A H: PERFIS DE RESTRIÇÃO PARA ISOLADOS DE ABACAXIZEIRO E BANANEIRA E ESTIRPES DE *Herbaspirillum* sp. E *Burkholderia* sp. COM A ENDONUCLEASE *Hae*III NA REGIÃO AMPLIFICADA Y1-Y3 DO 16S rDNA; LETRAS EM CAIXA BAIXA REPRESENTAM OS PERFIS DE RESTRIÇÃO DA TABELA 9; MARCADORES: M1 = 1 KB DNA LADDER (GIBCO BRL) E M100 = 100 BP DNA LADDER (GIBCO BRL).





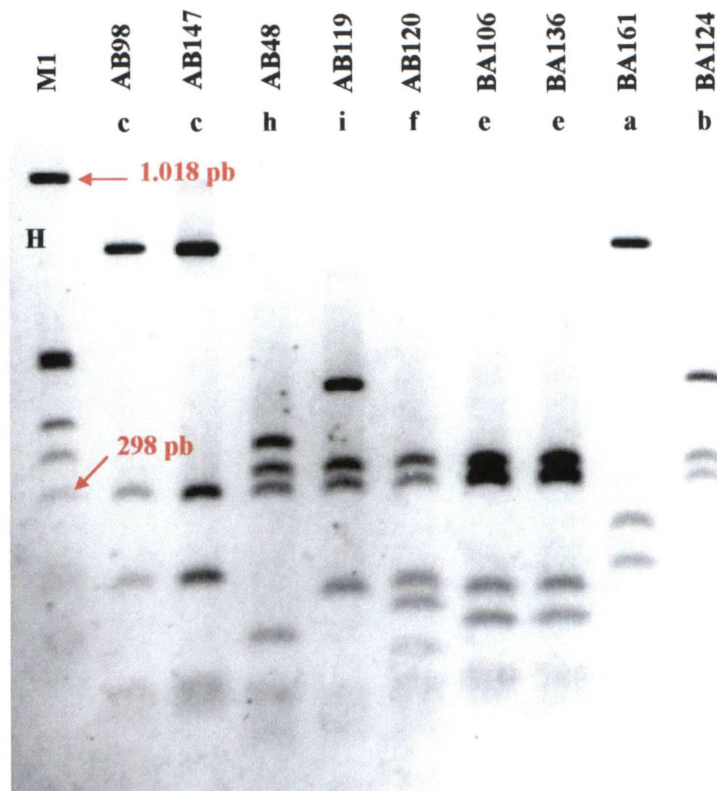
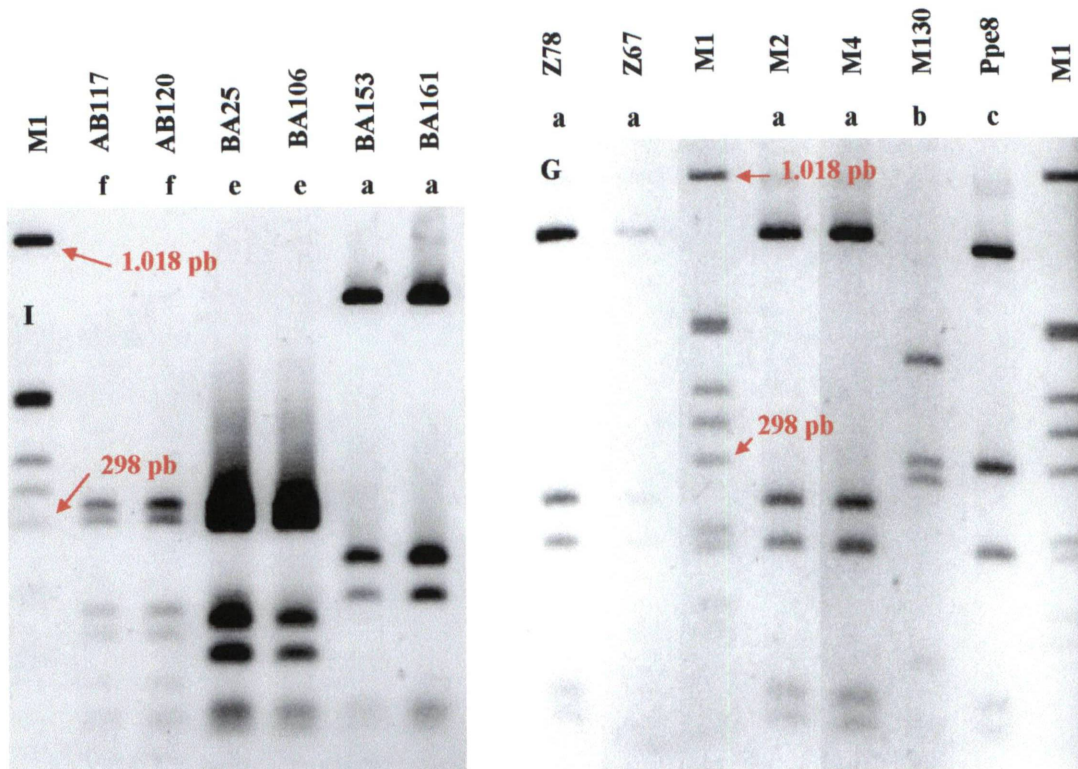
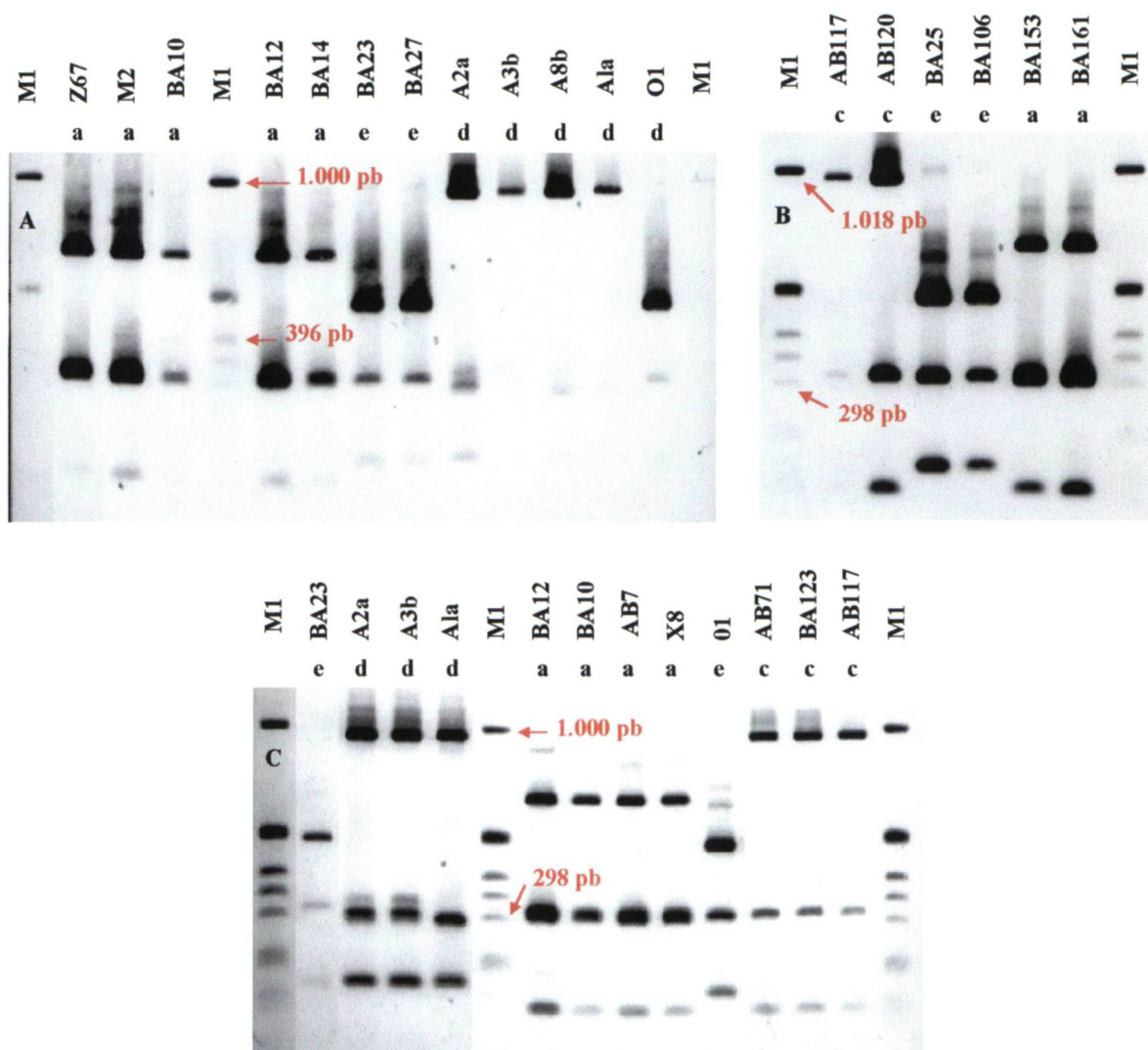


FIGURA 11 - DE A A G: PERFIS DE RESTRIÇÃO PARA ISOLADOS DE ABACAXIZEIRO E BANANEIRA E ESTIRPES DE *Herbaspirillum* sp. E *Burkholderia* sp. COM A ENDONUCLEASE *Hinf*I NA REGIÃO AMPLIFICADA Y1-Y3 DO 16S rDNA; AS LETRAS EM CAIXA BAIXA REPRESENTAM OS PERFIS DE RESTRIÇÃO DA TABELA 9; MARCADORES: M1 = 1 KB DNA LADDER (GIBCO BRL) E M100 = 100 BP DNA LADDER (GIBCO BRL).



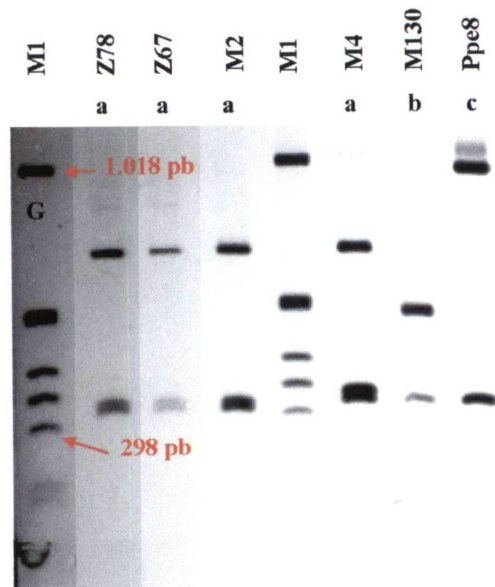
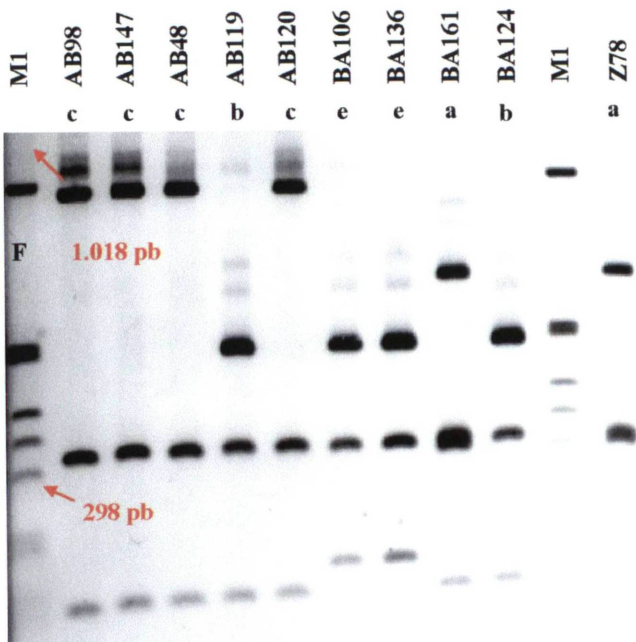
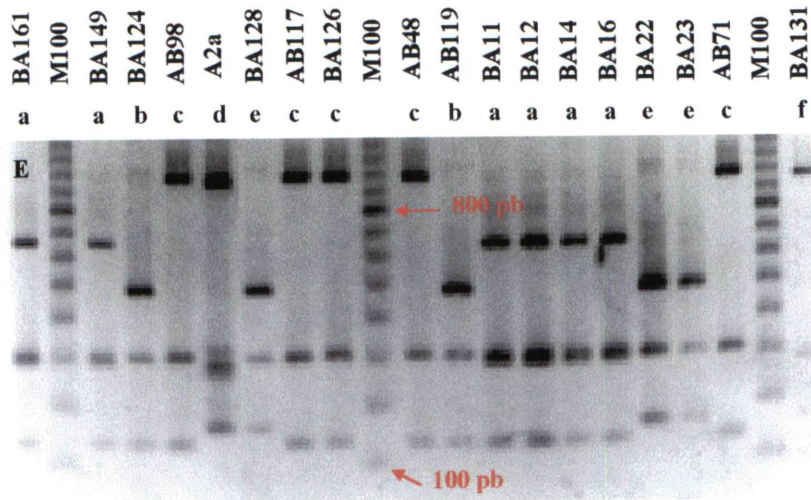
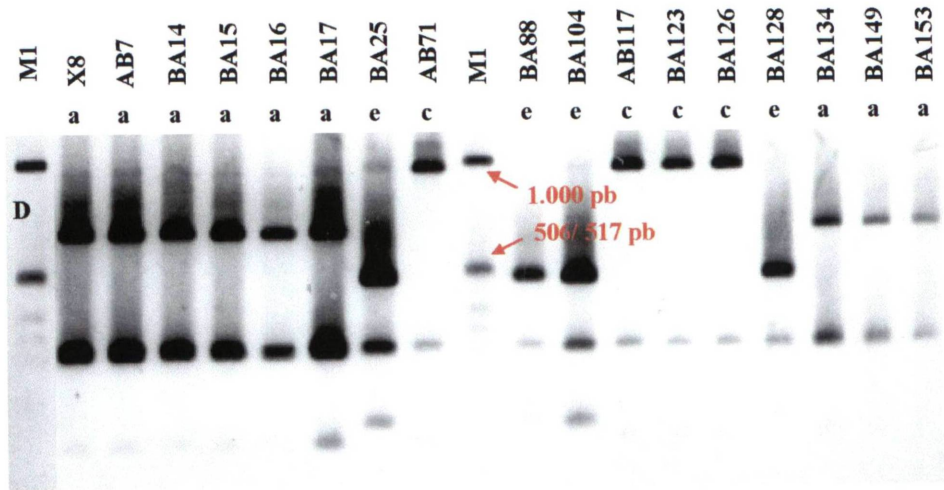
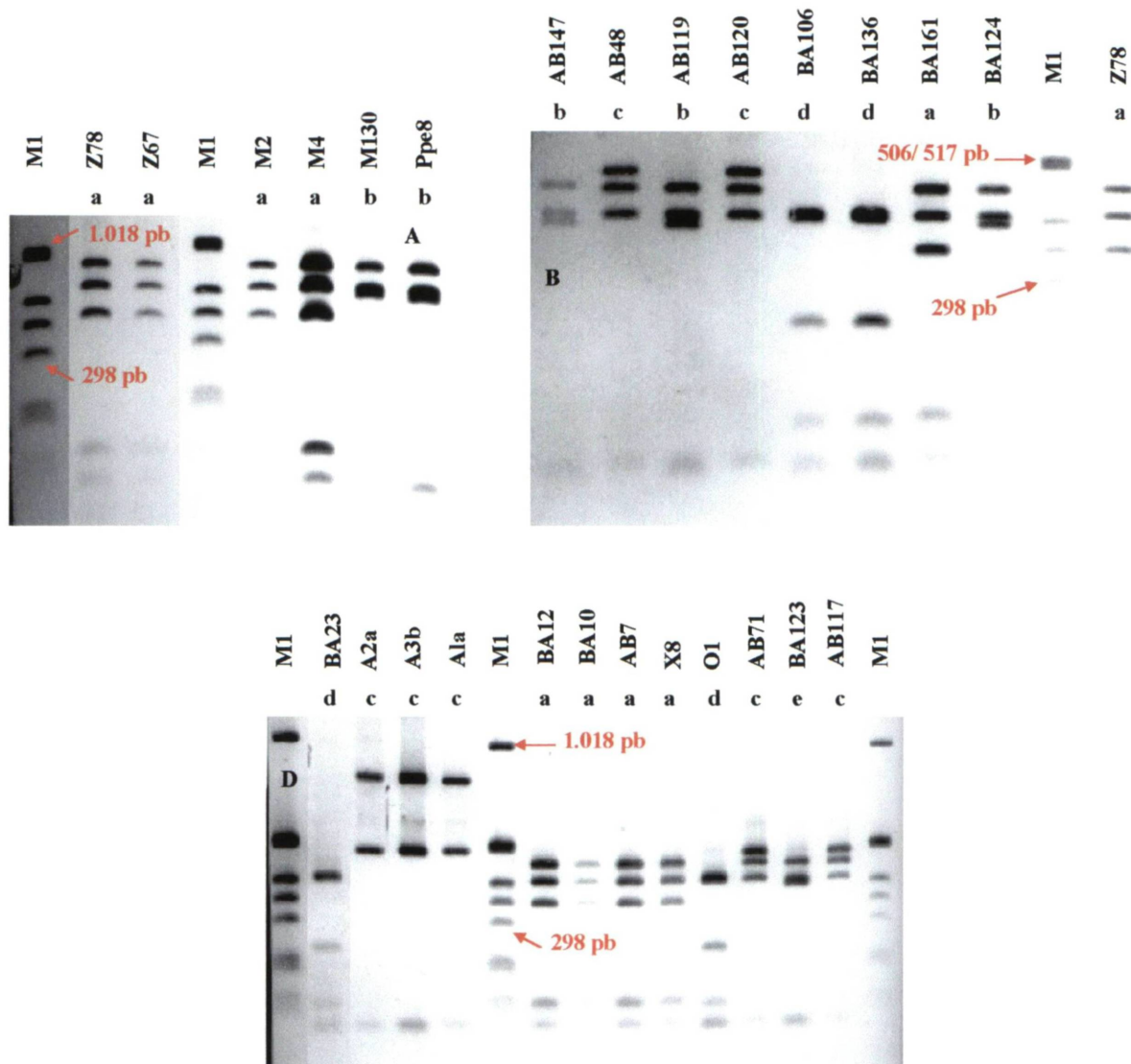


FIGURA 12 - DE A A F: PERFIS DE RESTRIÇÃO PARA ISOLADOS DE ABACAXIZEIRO E BANANEIRA E ESTIRPES DE *Herbaspirillum* sp. E *Burkholderia* sp. COM A ENDONUCLEASE *RsaI* NA REGIÃO AMPLIFICADA Y1-Y3 DO 16S rDNA; AS LETRAS EM CAIXA BAIXA REPRESENTAM OS PERFIS DE RESTRIÇÃO DA TABELA 9; MARCADORES: M1 = 1 KB DNA LADDER (GIBCO BRL) E M100 = 100 BP DNA LADDER (GIBCO BRL).



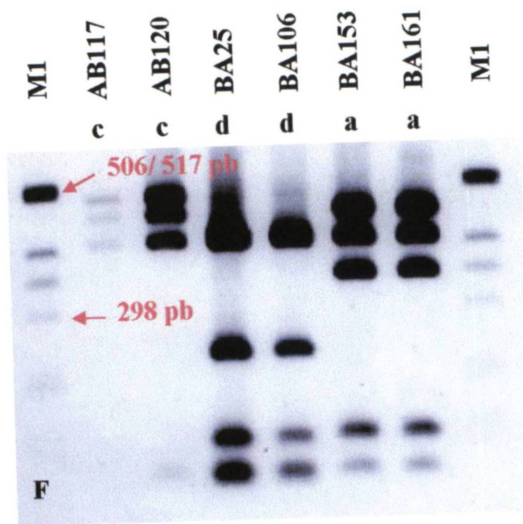
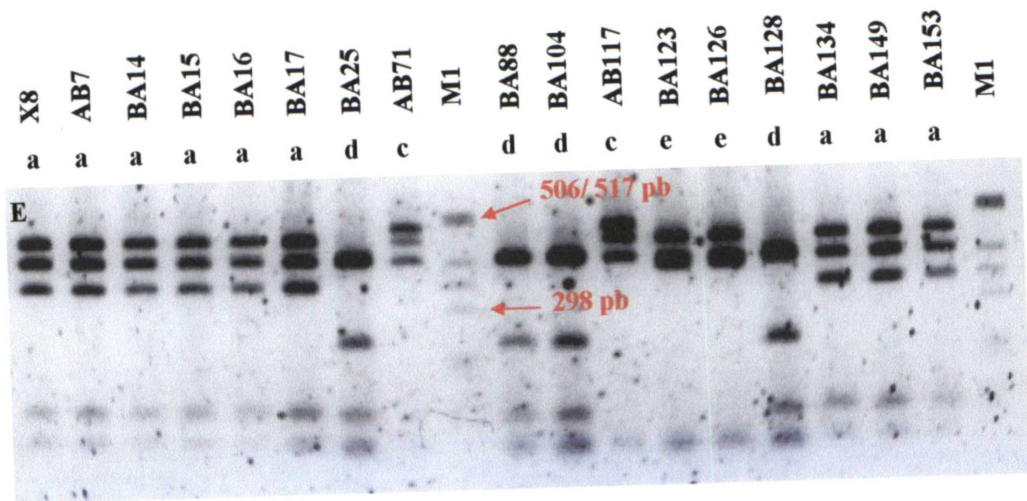
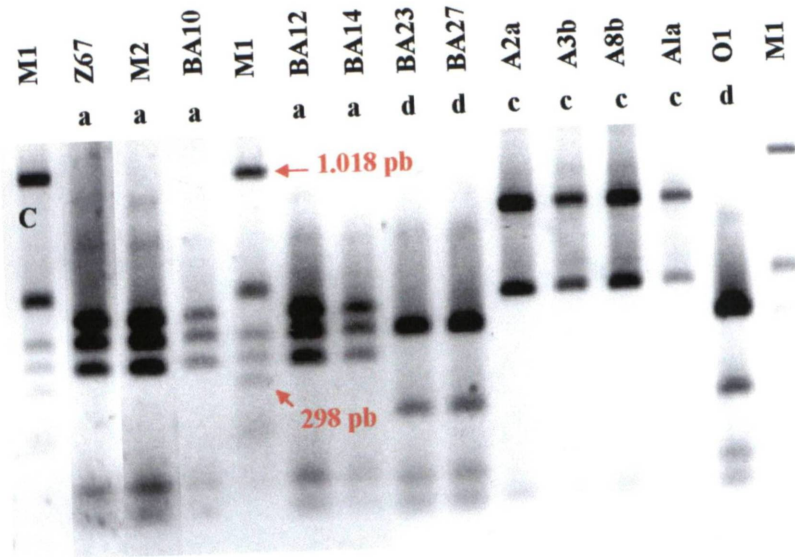
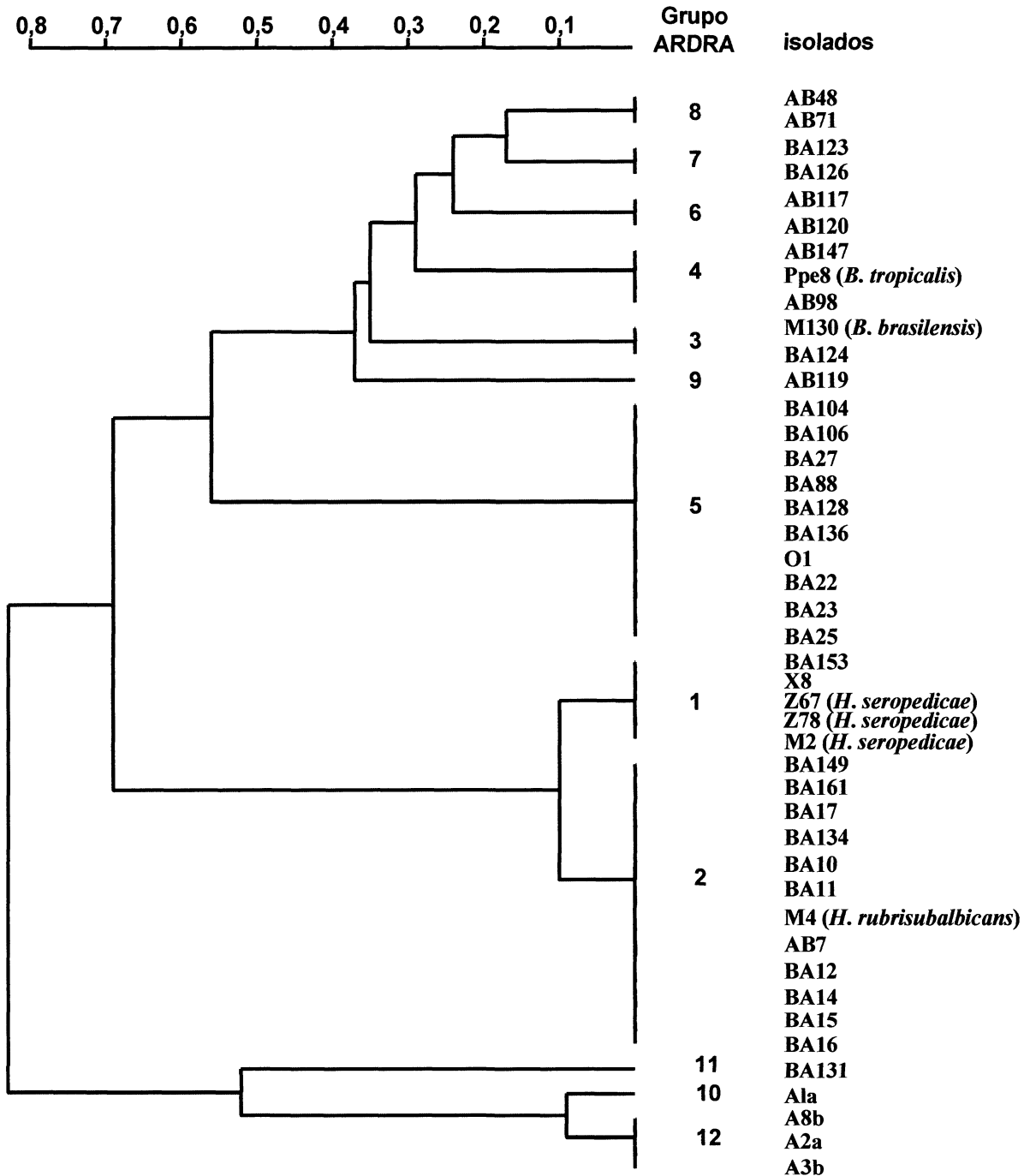


FIGURE 13 - DENDROGRAMA INFERIDO A PARTIR DOS PERFIS DE RESTRIÇÃO OBTIDOS PARA A DIGESTÃO DOS PRODUTOS DE AMPLIFICAÇÃO DO 16S rDNA DE ISOLADOS DE ABACAXIZEIRO E BANANEIRA E ESTIRPES DE *H. seropedicae*, *H. rubrisubalbicans*, "*B. brasiliensis*" E "*B. tropicalis*" COM *AluI*, *HaeIII* E *HinfI* MOSTRADOS NA TABELA 9. TOPOLOGIA OBTIDA PELO MÉTODO DE UPGMA (SNEATH & SOKAL, 1973) A PARTIR DE DISTÂNCIAS CALCULADAS PELO MÉTODO DE NEI & LI (1979).



3.4 SEQUENCIAMENTO DO 16S rDNA

As seqüências da região Y1-Y2 do 16S rDNA de 34 estirpes bacterianas foram determinadas, tendo sido encontradas 12 diferentes seqüências, representadas pelos Grupos de seqüência Y1-Y2 de I a XII (Tabela 10). Todas as seqüências foram obtidas diretamente a partir dos produtos de PCR das culturas bacterianas com o par de iniciadores Y1 e Y3, sem clonagem dos fragmentos. As seqüências foram obtidas em ambas as fitas do DNA.

O seqüenciamento direto dos produtos de PCR para os isolados AB48 e AB71, que pertencem ao Grupo de seqüência Y1-Y2 IX, apresentaram uma leitura de base ambígua C ou T (Y) na posição 157 (Figuras 13). De maneira análoga, foram observados picos para bases G ou A na leitura da seqüência da fita complementar, sugerindo que estas estirpes têm, no mínimo, dois operons 16S rDNA distintos.

Dezenove isolados, representantes dos Grupos de seqüências Y1-Y2 I, II, V, VI, VII, VIII, IX e XII foram também seqüenciados na região Y1-Y3 do 16S rDNA (Tabela 10). O seqüenciamento foi feito em ambas as fitas do DNA utilizando-se os oito iniciadores descritos na seção “Iniciadores Usados na Amplificação e Seqüenciamento do 16S rDNA e Região Intergência 16S-23S rDNA” do Material e Métodos. A região seqüenciada corresponde a mais de 90% da molécula o que permite uma análise filogenética mais precisa e confiável (ver adiante). O comprimento das seqüências variou de 1.425 a 1.430 pb (excluindo as seqüências para os iniciadores). Variação foi também encontrada entre seqüências, da mesma região, de espécies representantes da subdivisão beta das Proteobactérias.

TABELA 10 - CARACTERIZAÇÃO DE BACTÉRIAS DIAZOTRÓFICAS POR SEQÜENCIAMENTO DAS REGIÕES Y1-Y2 E Y1-Y3 DO 16S rDNA.

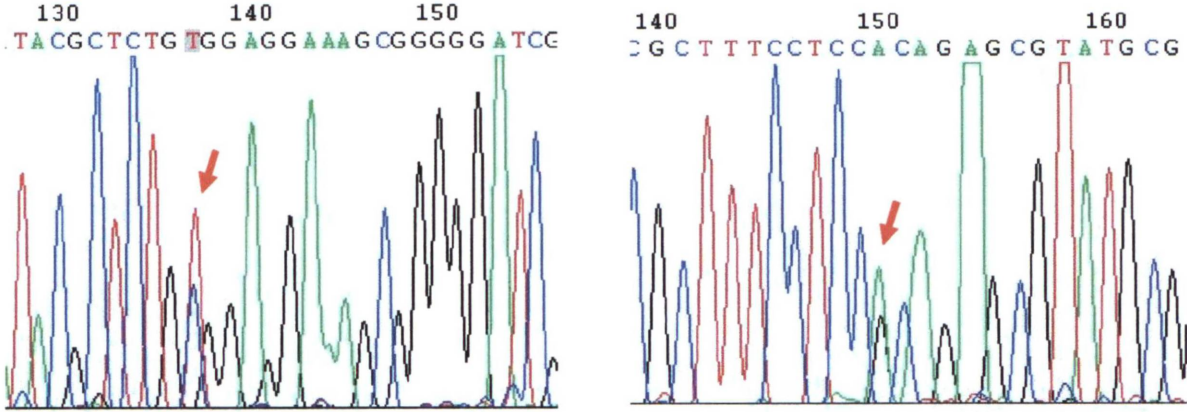
ESTIRPE ⁽¹⁾	ACESSO ⁽²⁾	SEQÜÊNCIA Y1-Y2		SEQÜÊNCIA Y1-Y3 ⁽³⁾
		Seqüência ⁽³⁾	Grupos	
<i>Herbaspirillum seropedicae</i>				
Z67		290 pb	I	
BA153	AF164062	290 pb	I	1 429 pb
X8	AF164065	290 pb	II	1 429 pb
<i>Herbaspirillum rubrisubalbicans</i>				
M4		290 pb	III	
AB7	AF164042	290 pb	I	1 428 pb
BA10	AF164049	290 pb	I	1 428 pb
BA11	AF164050	290 pb	I	1 429 pb
BA12	AF164051	290 pb	I	1 428 pb
BA14	AF164052	290 pb	I	1 429 pb
BA15	AF164053	290 pb	I	1 429 pb
BA16	AF164054	290 pb	I	1 429 pb
BA17	AF364861	290 pb	I	1 429 pb
BA134	AF164060	290 pb	I	1 429 pb
BA149		290 pb	I	1 429 pb
BA161	AF164063	290 pb	I	1 429 pb
<i>Burkholderia brasilensis</i>				
M130		288 pb	IV	
<i>Burkholderia tropicalis</i>				
Ppe8		290 pb	V	
AB98	AF164045	290 pb	V	1 429 pb
AB147	AF164047	290 pb	V	
Desconhecido				
O1	AF164064	286 pb	VI	
BA22	AF164055	286 pb	VI	
BA23	AF164056	286 pb	VI	
BA27	AF164057	286 pb	VI	
BA128	AF364862	286 pb	VI	1 425 pb
BA136	AF164061	286 pb	VI	
AB117	AF164046	288 pb	VII	
AB120	AF364860	288 pb	VII	1 428 pb
BA123	AF164058	288 pb	VIII	
BA126	AF164059	288 pb	VIII	1 427 pb
AB48	AF164043	288 pb	IX	1 427 pb
AB71	AF164044	288 pb	IX	
Ala	AF164048	260 pb	X	
BA131	AF213248	259 pb	XI	
AB119	AF364859	290 pb	XII	1 430 pb

(1) Z67 - *Herbaspirillum seropedicae*, M4 - *Herbaspirillum rubrisubalbicans*, Ppe8 - "*Burkholderia tropicalis*" e M130 - "*Burkholderia brasilensis*".

(2) Número de acesso das seqüências obtidas no presente trabalho e depositadas no banco de dados GenBank (<http://www.ncbi.nlm.nih.gov>).

(3) Não incluem as seqüências dos iniciadores.

FIGURA 14 - OBSERVAÇÃO DE LEITURA AMBÍGUA NA SEQÜÊNCIA DA REGIÃO Y1-Y2 DO GRUPO IX, COMPOSTO PELOS ISOLADOS AB48 E AB71. AS SETAS INDICAM A OCORRÊNCIA DAS BASES T E C NA SEQÜÊNCIA FEITA COM O INICIADOR Y1 (ESQUERDA) E DAS BASES A E G NA SEQÜÊNCIA DA FITA COMPLEMENTAR FEITA COM O INICIADOR Y2 (DIREITA).



3.5 ANÁLISE COMPARATIVA DE SEQUÊNCIAS E PESQUISA DE HOMOLOGIA

A identidade entre cada par de seqüências dos Grupos de seqüências Y1-Y2 do 16S rDNA é mostrada na matriz da Figura 15A. A identidade foi calculada após o alinhamento das seqüências no programa ClustalW (APÊNDICE 3), usando-se os parâmetros padrões. Um valor de identidade acima de 98% é compartilhado entre os isolados dos Grupos I, II e III. Valores de identidade acima de 90% foram também observados entre os isolados dos Grupos IV, V, VII, VIII, IX e XII. Os isolados dos Grupos VI, X e XI mostraram-se mais distantes dos demais. O Grupo VI apresentou valores de identidade mais ou menos constantes em relação a todos os outros grupos, com um valor médio de 83,6%. O mesmo foi observado para os Grupos X e XI, mas com uma identidade média de 66,1%. Entre estes dois últimos Grupos, a identidade foi de 79,3%. Estes dados permitiram a separação dos organismos analisados em quatro grandes grupos relacionados entre si. O primeiro formado pelos Grupos de seqüências Y1-Y2 I, II e III, semelhantes a *Herbaspirillum*; o segundo composto pelos Grupos de seqüências Y1-Y2 IV, V, VII, VIII, IX e XII e semelhantes a *Burkholderia*; o terceiro composto apenas pelo Grupo de seqüência Y1-Y2 VI; e o quarto composto pelos Grupos de seqüências Y1-Y2 X e XI. Para estes dois últimos Grupos, não foi observada proximidade a nenhuma das estirpes usadas como referência. As identidades obtidas também para a região Y1-Y3 do 16S rDNA concordam com aquelas apresentadas acima (Figura 15B). Entretanto, os valores foram, em geral, ligeiramente superiores. Para esta região, todos os isolados do Grupo de seqüência Y1-Y2 I e também o isolado do Grupo II foram seqüenciados. Estes mostraram um alto grau de identidade (acima de 98,6%). Entretanto, uma identidade de 100% só foi observada entre os isolados:

- a) BA11, BA16 e BA17;
- b) BA14, BA15 e BA149;

c) BA12 e BA134.

Mostrando haver heterogeneidade no Grupo de seqüência Y1-Y2 I, o que confirmou a ambigüidade apresentada pelas análises de ARDRA e seqüenciamento da região Y1-Y2.

FIGURA 15 - MATRIZ MOSTRANDO A PROPORÇÃO DE IDENTIDADE (DIAGONAL SUPERIOR) E O NÚMEROS DE BASES DIFERENTES (DIAGONAL INFERIOR) ENTRE PARES DE SEQÜÊNCIAS DE 16S rDNA. A) CALCULADOS A PARTIR DO ALINHAMENTO ENTRE AS SEQÜÊNCIAS NA REGIÃO Y1-Y2 DOS GRUPOS OBTIDOS; VALORES DE IDENTIDADE ACIMA DE 90% (0,900) SÃO MOSTRADOS EM NEGRITO; B) CALCULADOS A PARTIR DO ALINHAMENTO ENTRE SEQÜÊNCIAS DA REGIÃO Y1-Y3 DE DIVERSOS ISOLADOS; VALORES DE IDENTIDADE ACIMA DE 98% (0,980) DE IDENTIDADE SÃO MOSTRADAS EM NEGRITO.

A												
Seq->	I	II	III	IV	V	VII	VIII	IX	XII	VI	X	XI
GRUPO I		0,993	0,993	0,834	0,868	0,831	0,848	0,841	0,855	0,865	0,662	0,658
GRUPO II	2		0,986	0,827	0,875	0,824	0,841	0,834	0,855	0,865	0,668	0,665
GRUPO III	48	2	4	0,841	0,868	0,834	0,851	0,844	0,858	0,858	0,662	0,651
GRUPO IV	36	38	21	46	0,920	0,947	0,909	0,934	0,920	0,831	0,673	0,666
GRUPO V	49	51	48	21	15	0,927	0,913	0,913	0,965	0,834	0,672	0,679
GRUPO VII	44	46	43	26	25	15	0,954	0,931	0,813	0,817	0,652	0,645
GRUPO VIII	46	48	45	19	25	9	13	0,927	0,810	0,656	0,652	
GRUPO IX	42	42	41	21	10	10	18	19	0,831	0,662	0,658	
GRUPO XII	39	39	41	47	48	52	51	53	49	0,678	0,657	
GRUPO VI	96	98	94	95	100	99	99	98	92		0,793	
GRUPO X	99	97	101	96	93	102	101	100	99	98	52	
GRUPO XI												

B																			
Seq->	BA153	X8	AB7	BA10	BA11	BA12	BA14	BA15	BA16	BA17	BA134	BA149	BA161	AB98	BA128	AB120	BA126	AB48	AB119
BA153		0,992	0,989	0,990	0,991	0,990	0,990	0,990	0,991	0,991	0,990	0,990	0,989	0,906	0,894	0,898	0,900	0,900	0,904
X8	11		0,986	0,986	0,988	0,986	0,987	0,987	0,988	0,988	0,987	0,987	0,986	0,906	0,890	0,897	0,898	0,899	0,904
AB7	14	19		0,996	0,997	0,996	0,997	0,997	0,997	0,997	0,997	0,997	0,995	0,906	0,890	0,895	0,898	0,897	0,902
BA10	14	19	3		0,998	0,997	0,997	0,997	0,998	0,998	0,997	0,997	0,996	0,906	0,890	0,896	0,899	0,898	0,902
BA11	12	17	1	1		0,998	0,999	0,999	1,000	1,000	0,999	0,999	0,997	0,908	0,892	0,897	0,900	0,900	0,904
BA12	14	19	3	3	2		0,997	0,997	0,998	0,998	0,999	0,997	0,997	0,908	0,892	0,897	0,900	0,900	0,903
BA14	13	18	2	2	1	2		1,000	0,999	0,999	0,998	1,000	0,997	0,909	0,892	0,898	0,899	0,900	0,903
BA15	13	18	2	2	1	2	0		0,999	0,999	0,998	1,000	0,997	0,909	0,892	0,898	0,899	0,900	0,903
BA16	12	17	1	1	0	1	1	1		1,000	0,999	0,999	0,997	0,908	0,892	0,897	0,900	0,900	0,904
BA17	12	17	1	1	0	1	1	0	0		0,999	0,999	0,997	0,908	0,892	0,897	0,900	0,900	0,904
BA134	13	18	2	2	1	0	2	2	1	1		0,998	0,998	0,909	0,892	0,898	0,901	0,900	0,904
BA149	13	18	2	2	1	2	0	0	1	1	2		0,997	0,909	0,892	0,898	0,899	0,900	0,903
BA161	15	20	4	4	3	2	4	4	3	3	2	4		0,907	0,892	0,897	0,899	0,899	0,902
AB98	134	134	132	132	131	130	130	130	131	131	130	130	132		0,890	0,963	0,965	0,960	0,972
BA128	153	158	157	157	156	155	156	156	156	156	155	156	156	158		0,883	0,886	0,882	0,893
AB120	143	145	145	145	144	143	143	143	144	144	143	143	145	50	163		0,968	0,988	0,967
BA126	142	145	143	143	142	141	143	143	142	142	141	143	143	50	159	44		0,969	0,965
AB48	141	143	143	143	142	141	141	141	142	142	141	141	143	56	166	15	43		0,960
AB119	135	136	137	137	136	136	137	137	136	136	136	137	138	38	149	43	47	53	

Embora não se tenham valores de homologia entre seqüências de 16S rDNA pré-definidos na literatura para determinar o limite entre espécies ou qualquer outro grupo taxonômico, é possível ter uma idéia clara da proximidade evolutiva e, conseqüentemente, também da proximidade taxonômica entre os organismos comparados.

A partir de uma pesquisa de homologia feita usando o algoritmo BLAST 2.1 (APÊNDICE 3), foi possível estabelecer os grupos taxonômicos que estão mais intimamente relacionados com os Grupos formados pelas seqüências dos isolados diazotróficos aqui estudados. A pesquisa foi feita via Web (NCBI, 2001), utilizando-se o programa blastn, que permite a comparação entre seqüências de nucleotídeos. Foram utilizados os parâmetros padrões de análise, pré-definidos no programa, e o banco de dados de seqüências não redundantes (nr), contendo 852.500 seqüências no momento da pesquisa (Tabela 11). Como as seqüências dos isolados foram depositadas no banco de dados GenBank, a maior homologia ocorreu com a própria seqüência submetida ou, muitas vezes, também com outras seqüências oriundas do presente trabalho, o que não foi levado em consideração. Os isolados dos grupos ARDRA 1 e 2 mostraram alta homologia com espécies de *Herbaspirillum* (99% em mais de 1.450 bases). Por outro lado, os isolados X8 e BA153, do grupo ARDRA 1, obtiveram os mais altos *scores* com *H. seropedicae*, enquanto que para os isolados do grupo ARDRA 2 os mais altos *scores* alcançados foram com *Pseudomonas huttiensis*. Para os grupos ARDRA 4, 6, 7, 8 e 9 foi verificada uma alta homologia com espécies do gênero *Burkholderia* (entre 95 e 97% em mais de 1.430 bases). Entretanto, alguns dos isolados destes grupos apresentaram os *scores* mais altos com bactérias não classificadas. Finalmente, o isolado BA128, pertencente ao grupo ARDRA 5, apresentou alta homologias com bactérias dos gêneros *Xylophilus* e *Acidovorax* (96% em mais de 1.450 bases). A pesquisa de homologia para os grupos ARDRA 10 e 11 foi feita somente com a região Y1-Y2 do 16S rDNA, mostrando uma alta identidade para o isolado Ala com *Azospirillum lipoferum* (98%) e para o isolado BA131 com *Ochrobactrum* sp. (99% -

Tabela 11).

TABELA 11 - ORGANISMOS QUE APRESENTARAM MAIOR HOMOLOGIA PARA A REGIÃO Y1-Y2 OU Y1-Y3 DA SEQUÊNCIA DO 16S rDNA DE ISOLADOS DIAZOTRÓFICOS DE ABACAXIZEIRO E BANANAEIRA, DE ACORDO COM PESQUISA REALIZADA COM O PROGRAMA BLASTN⁽¹⁾.

continua

No. DE ACESSO NO GenBank	ESPÉCIE	ESTIRPE	COMP. (bases)	SCORE	VALOR E	IDENTIDADES ⁽²⁾		GAPS ⁽³⁾	
						Proporção	%	Proporção	%
X8 (1.478 bases) - grupo ARDRA 1									
	<i>Herbaspirillum</i>								
AJ23835	<i>frisingense</i>	75B	1466	2866	0.0	1452/1454	99	---	---
Y10146	<i>H. seropedicae</i>	---	1526	2821	0.0	1466/1479	99	1/1479	0
	<i>Herbaspirillum</i>								
AJ238361	<i>seropedicae</i>	Z67	1479	2805	0.0	1445/1455	99	---	---
	<i>Herbaspirillum</i>								
AJ238358	<i>frisingense</i>	GSF30	1466	2795	0.0	1443/1454	99	---	---
	<i>Herbaspirillum</i>								
AJ238359	<i>frisingense</i>	Mb11	1485	2793	0.0	1448/1461	99	---	---
	<i>Herbaspirillum</i>								
AB049133	sp.	B501	1487	2787	0.0	1461/1478	98	1/1478	0
	<i>Herbaspirillum</i>								
AF137508	<i>rubrisubalbicans</i>	---	1485	2785	0.0	1447/1461	99	---	---
	<i>Herbaspirillum</i>								
AB049104	sp.	B512	1480	2785	0.0	1447/1461	99	---	---
	Nitrogen-fixing								
AF214642	bacterium	COL	1520	2781	0.0	1461/1479	98	1/1479	0
	<i>Herbaspirillum</i>								
AJ238356	<i>rubrisubalbicans</i>	M4	1465	2771	0.0	1440/1454	99	---	---
BA153 (1.478 bases) - grupo ARDRA 1									
Y10146	<i>H. seropedicae</i>	---	1526	2884	0.0	1474/1479	99	1/1479	0
	<i>Herbaspirillum</i>								
AJ238361	<i>seropedicae</i>	Z67	1479	2868	0.0	1453/1455	99	---	---
	<i>Herbaspirillum</i>								
AB049133	sp.	B501	1487	2835	0.0	1467/1478	99	1/1478	0
	Nitrogen-fixing								
AF214642	bacterium	COL	1520	2829	0.0	1467/1479	99	1/1479	0
	<i>Herbaspirillum</i>								
AB049103	sp.	B65	1477	2789	0.0	1443/1455	99	---	---
	<i>Herbaspirillum</i>								
AF137508	<i>rubrisubalbicans</i>	---	1485	2785	0.0	1447/1461	99	---	---
	<i>Herbaspirillum</i>								
AJ238357	<i>frisingense</i>	75B	1466	2771	0.0	1440/1454	99	---	---
	<i>Herbaspirillum</i>								
AJ238356	<i>rubrisubalbicans</i>	M4	1465	2771	0.0	1440/1454	99	---	---
	<i>Herbaspirillum</i>								
AJ238359	<i>frisingense</i>	Mb11	1485	2746	0.0	1442/1461	98	---	---
	<i>Pseudomonas</i>	ATCC							
AB021366	<i>huttiensis</i>	14670T	1490	2738	0.0	1441/1461	98	---	---
AB7 (1.477 bases) - grupo ARDRA 2									

No. DE ACESSO NO GenBank	ESPÉCIE	ESTIRPE	COMP. (bases)	SCORE	VALOR E	IDENTIDADES ⁽²⁾		GAPS ⁽³⁾	
						Proporção	%	Proporção	%
	<i>Pseudomonas</i>	ATCC					99		
AB021366	<i>huttiensis</i>	14670T	1490	2779	0.0	1450/1462		3/1462	0
Y10146	<i>H. seropedicae</i>	---	1526	2767	0.0	1463/1480	98	4/1480	0
	<i>Herbaspirillum</i>								
AF137508	<i>rubrisubalbicans</i>	---	1485	2763	0.0	1448/1462	99	3/1462	0
	<i>Herbaspirillum</i>								
AJ238361	<i>seropedicae</i>	Z67	1479	2752	0.0	1442/1456	99	3/1456	0
	<i>Herbaspirillum</i>								
AB049103	sp.	B65	1477	2752	0.0	1445/1460	98	3/1460	0
	<i>Herbaspirillum</i>								
AJ238356	<i>rubrisubalbicans</i>	M4	1465	2750	0.0	1441/1455	99	3/1455	0
	<i>Herbaspirillum</i>								
AJ238359	<i>frisingense</i>	Mb11	1485	2740	0.0	1445/1462	98	3/1462	0
	<i>Herbaspirillum</i>								
AB049133	sp.	B501	1487	2726	0.0	1457/1479	98	4/1479	0
	<i>Herbaspirillum</i>								
AB049104	sp.	B512	1480	2724	0.0	1443/1462	98	3/1462	0
	Nitrogen-fixing								
AF214642	bacterium	COL	1520	2720	0.0	1457/1480	98	4/1480	0

BA10 (1.477 bases) - grupo ARDRA 2

	<i>Pseudomonas</i>	ATCC							
AB021366	<i>huttiensis</i>	14670T	1490	2801	0.0	1450/1461	99	1/1461	0
Y10146	<i>H. seropedicae</i>	---	1526	2789	0.0	1463/1479	98	2/1479	0
	<i>Herbaspirillum</i>								
AF137508	<i>rubrisubalbicans</i>	---	1485	2785	0.0	1448/1461	99	1/1461	0
	<i>Herbaspirillum</i>								
AJ238361	<i>seropedicae</i>	Z67	1479	2773	0.0	1442/1455	99	1/1455	0
	<i>Herbaspirillum</i>								
AB049103	sp.	B65	1477	2773	0.0	1445/1459	99	1/1459	0
	<i>Herbaspirillum</i>								
AJ238356	<i>rubrisubalbicans</i>	M4	1465	2771	0.0	1441/1454	99	1/1454	0
	<i>Herbaspirillum</i>								
AJ238359	<i>frisingense</i>	Mb11	1485	2761	0.0	1445/1461	98	1/1461	0
	<i>Herbaspirillum</i>								
AB049133	sp.	B501	1487	2748	0.0	1457/1478	98	2/1478	0
	<i>Herbaspirillum</i>								
AB049104	sp.	B512	1480	2746	0.0	1443/1461	98	1/1461	0
	Nitrogen-fixing								
AF214642	bacterium	COL	1520	2742	0.0	1457/1479	98	2/1479	0

BA11, BA16 e BA17 (1.478 bases) - grupo ARDRA 2

	<i>Pseudomonas</i>	ATCC							
AB021366	<i>huttiensis</i>	14670T	1490	2825	0.0	1452/1461	99	---	---
	<i>Herbaspirillum</i>								
AF137508	<i>rubrisubalbicans</i>	---	1485	2809	0.0	1450/1461	99	---	---
Y10146	<i>H. seropedicae</i>	---	1526	2809	0.0	1450/1461	99	---	---
	<i>Herbaspirillum</i>								
AJ238361	<i>seropedicae</i>	Z67	1479	2797	0.0	1444/1455	99	---	---
	<i>Herbaspirillum</i>								
AB049103	sp.	B65	1477	2797	0.0	1447/1459	99	---	---

No. DE ACESSO NO GenBank	ESPÉCIE	ESTIRPE	COMP. (bases)	SCORE	VALOR E	IDENTIDADES ⁽²⁾		GAPS ⁽³⁾	
						Proporção	%	Proporção	%
	<i>Herbaspirillum</i>						99		
AJ238356	<i>rubrisubalbicans</i>	M4	1465	2795	0.0	1443/1454		---	---
	<i>Herbaspirillum</i>								
AJ238359	<i>frisingense</i>	Mb11	1485	2785	0.0	1447/1461	99	---	---
	<i>Herbaspirillum</i>								
AB049133	sp.	B501	1487	2769	0.0	1445/1461	98	---	---
	<i>Herbaspirillum</i>								
AB049104	sp.	B512	1480	2769	0.0	1445/1461	98	---	---
	Nitrogen-fixing								
AF214642	bacterium	COL	1520	2761	0.0	1444/1461	98	---	---
BA12 (1.478 bases) - grupo ARDRA 2									
	<i>Pseudomonas</i>	ATCC							
AB021366	<i>huttiensis</i>	14670T	1490	2779	0.0	1449/1462	99	2/1462	0
Y10146	<i>H. eropedicae</i>	---	1526	2767	0.0	1462/1480	98	3/1480	0
	<i>Herbaspirillum</i>								
AF137508	<i>rubrisubalbicans</i>	---	1485	2763	0.0	1447/1462	98	2/1462	0
	<i>Herbaspirillum</i>								
AJ238361	<i>seropedicae</i>	Z67	1479	2752	0.0	1441/1456	98	2/1456	0
	<i>Herbaspirillum</i>								
AB049103	sp.	B65	1477	2752	0.0	1444/1460	98	2/1460	0
	<i>Herbaspirillum</i>								
AJ238356	<i>rubrisubalbicans</i>	M4	1465	2750	0.0	1440/1455	98	2/1455	0
	<i>Herbaspirillum</i>								
AB049133	sp.	B501	1487	2742	0.0	1458/1479	98	3/1479	0
	<i>Herbaspirillum</i>								
AJ238359	<i>frisingense</i>	Mb11	1485	2740	0.0	1444/1462	98	2/1462	0
	<i>Herbaspirillum</i>								
AB049104	sp.	B512	1480	2740	0.0	1444/1462	98	2/1462	0
	Nitrogen-fixing								
AF214642	bacterium	COL	1520	2720	0.0	1456/1480	98	3/1480	0
BA14, BA15 e BA149 (1.478 bases) - grupo ARDRA 2									
	<i>Pseudomonas</i>	ATCC							
AB021366	<i>huttiensis</i>	14670T	1490	2833	0.0	1453/1461	99	---	---
Y10146	<i>H. seropedicae</i>	---	1526	2805	0.0	1464/1479	98	1/1479	0
	<i>Herbaspirillum</i>								
AF137508	<i>rubrisubalbicans</i>	---	1485	2801	0.0	1449/1461	99	---	---
	<i>Herbaspirillum</i>								
AJ238359	<i>frisingense</i>	Mb11	1485	2793	0.0	1448/1461	99	---	---
	<i>Herbaspirillum</i>								
AJ238361	<i>seropedicae</i>	Z67	1479	2789	0.0	1443/1455	99	---	---
	<i>Herbaspirillum</i>								
AB049103	sp.	B65	1477	2789	0.0	1446/1459	99	---	---
	<i>Herbaspirillum</i>								
AJ238356	<i>rubrisubalbicans</i>	M4	1465	2787	0.0	1442/1454	99	---	---
	<i>Herbaspirillum</i>								
AB049133	sp.	B501	1487	2763	0.0	1458/1478	98	1/1478	0
	<i>Herbaspirillum</i>								
AB049104	sp.	B512	1480	2761	0.0	1444/1461	98	---	---
	Nitrogen-fixing								
AF214642	bacterium	COL	1520	2757	0.0	1458/1479	98	1/1479	0

No. DE ACESSO NO GenBank	ESPÉCIE	ESTIRPE	COMP. (bases)	SCORE	VALOR E	IDENTIDADES ⁽²⁾		GAPS ⁽³⁾	
						Proporção	%	Proporção	%
BA134 (1.478 bases) - grupo ARDRA 2									
AB021366	<i>Pseudomonas huttiensis</i>	ATCC 14670T	1490	2817	0.0	1451/1461	99	---	---
Y10146	<i>H. seropedicae</i>	---	1526	2805	0.0	1464/1479	98	1/1479	0
AF137508	<i>Herbaspirillum rubrisubalbicans</i>	---	1485	2801	0.0	1449/1461	99	---	---
AJ238361	<i>Herbaspirillum seropedicae</i>	Z67	1479	2789	0.0	1443/1455	99	---	---
AB049103	<i>Herbaspirillum</i> sp.	B65	1477	2789	0.0	1446/1459	99	---	---
AJ238356	<i>Herbaspirillum rubrisubalbicans</i>	M4	1465	2787	0.0	1442/1454	99	---	---
AB049133	<i>Herbaspirillum</i> sp.	B501	1487	2779	0.0	1460/1478	98	1/1478	0
AJ238359	<i>Herbaspirillum frisingense</i>	Mb11	1485	2777	0.0	1446/1461	98	---	---
AB049104	<i>Herbaspirillum</i> sp.	B512	1480	2777	0.0	1446/1461	98	---	---
AF214642	Nitrogen-fixing bacterium	COL	1520	2757	0.0	1458/1479	98	1/1479	0
BA161 (1.478 bases) - grupo ARDRA 2									
AB021366	<i>Pseudomonas huttiensis</i>	ATCC 14670T	1490	2801	0.0	1449/1461	99	---	---
Y10146	<i>H. seropedicae</i>	---	1526	2789	0.0	1462/1479	98	1/1479	0
AF137508	<i>Herbaspirillum rubrisubalbicans</i>	---	1485	2785	0.0	1447/1461	99	---	---
AJ238361	<i>Herbaspirillum seropedicae</i>	Z67	1479	2773	0.0	1441/1455	99	---	---
AB049103	<i>Herbaspirillum</i> sp.	B65	1477	2773	0.0	1444/1459	98	---	---
AJ238356	<i>Herbaspirillum rubrisubalbicans</i>	M4	1465	2771	0.0	1440/1454	99	---	---
AB049133	<i>Herbaspirillum</i> sp.	B501	1487	2763	0.0	1458/1478	98	1/1478	0
AJ238359	<i>Herbaspirillum frisingense</i>	Mb11	1485	2761	0.0	1444/1461	98	---	---
AB049104	<i>Herbaspirillum</i> sp.	B512	1480	2761	0.0	1444/1461	98	---	---
AF214642	Nitrogen-fixing bacterium	COL	1520	2742	0.0	1456/1479	98	1/1479	0
AB98 (1.478 bases) - grupo ARDRA 4									
AF184929	2,4-D-degrading bacterium	TFD3	1392	2684	0.0	1375/1382	99	---	---
AF262932	<i>Burkholderia</i> sp. Unidentified beta	JS150	1521	2543	0.0	1431/1479	96	1/1479	0
AB015530	proteobacterium	BD1-33	1492	2516	0.0	1431/1480	96	5/1480	0
AF227856	Bacterium str.	96444	1434	2480	0.0	1386/1430	96	2/1430	0
U37344	<i>Burkholderia</i> sp.	N3P2	1516	2474	0.0	1423/1479	96	3/1479	0

No. DE ACESSO NO GenBank	ESPÉCIE	ESTIRPE	COMP. (bases)	SCORE	VALOR E	IDENTIDADES ⁽²⁾		GAPS ⁽³⁾	
						Proporção	%	Proporção	%
	<i>Pseudomonas</i>						96		
U86373	sp. uncultured	LB400	1466	2448	0.0	1404/1458		3/1458	0
AJ292637	eubacterium	WD202	1494	2442	0.0	1419/1479	95	3/1479	0
	<i>Burkholderia</i>	LMG							
AF311970	<i>cepacia</i>	12614	1494	2440	0.0	1418/1479	95	1/1479	0
	<i>Burkholderia</i>	LMG							
AF097532	<i>cepacia</i>	12614t2	1494	2440	0.0	1418/1479	95	1/1479	0
AB008002	<i>Burkholderia</i> sp.	MBIC3837	1480	2440	0.0	1402/1459	96	---	---
BA128 (1.474 bases) - grupo ARDRA 5									
	<i>Xylophilus</i>								
AF078758	<i>ampelinus</i>	---	1518	2557	0.0	1433/1479	96	5/1479	0
AJ130765	<i>Acidovorax</i> sp. Unidentified	LW1	1483	2516	0.0	1411/1457	96	1/1457	0
AB021320	bacterium Unidentified	rJ1	1464	2478	0.0	1404/1455	96	4/1455	0
AB021361	bacterium uncultured	rA10	1442	2440	0.0	1389/1442	96	4/1442	0
AJ292624	eubacterium Unidentified	WD291	1491	2438	0.0	1418/1479	95	10/1479	0
AB021340	bacterium Unidentified	rM5	1474	2411	0.0	1400/1461	95	4/1461	0
AB021334	bacterium Unidentified	rJ16	1461	2395	0.0	1399/1457	96	8/1457	0
Z93960	bacterium Unidentified	---	1529	2387	0.0	1416/1481	95	8/1481	0
Z93973	bacterium Bacteria	---	1519	2385	0.0	1408/1475	95	6/1475	0
X95838	species	T41	1519	2385	0.0	1408/1475	95	6/1475	0
AB120 (1.477 bases) - grupo ARDRA 6									
AF262932	<i>Burkholderia</i> sp. Unidentified beta	JS150	1521	2541	0.0	1436/1481	96	6/1481	0
AB015530	proteobacterium	BD1-33	1492	2498	0.0	1431/1480	96	6/1480	0
AF227856	Bacterium str.	96444	1434	2462	0.0	1386/1430	96	3/1430	0
AF074711	<i>Burkholderia</i> sp.	EN-B3	1513	2397	0.0	1419/1481	95	6/1481	0
AF074712	<i>Burkholderia</i> sp.	EN-B9	1497	2361	0.0	1416/1482	95	8/1482	0
	<i>Burkholderia</i>								
U91839	<i>pseudomallei</i>	1026b	1488	2357	0.0	1396/1461	95	3/1461	0
	<i>Burkholderia</i>	ATCC							
AF110188	<i>mallei</i>	23344	1488	2349	0.0	1395/1461	95	3/1461	0
	<i>Burkholderia</i>	NCTC							
AF110187	<i>mallei</i>	10260	1488	2349	0.0	1395/1461	95	3/1461	0
AF247494	<i>Burkholderia</i> sp. uncultured	S2.1	1449	2335	0.0	1379/1442	95	3/1442	0
AJ292643	eubacterium	WD268	1494	2329	0.0	1407/1479	95	4/1479	0
BA126 (1.476 bases) - grupo ARDRA 7									

No. DE ACESSO NO GenBank	ESPÉCIE	ESTIRPE	COMP. (bases)	SCORE	VALOR E	IDENTIDADES ⁽²⁾		GAPS ⁽³⁾	
						Proporção	%	Proporção	%
	Unidentified beta						99		
AB015530	proteobacterium	BD1-33	1492	2823	0.0	1453/1460		2/1460	0
AF227856	Bacterium str.	96444	1434	2815	0.0	1426/1428	99	---	---
AF262932	<i>Burkholderia</i> sp.	JS150	1521	2565	0.0	1420/1461	97	2/1461	0
AF247494	<i>Burkholderia</i> sp.	S2.1	1449	2434	0.0	1387/1440	96	---	---
U37344	<i>Burkholderia</i> sp.	N3P2	1516	2424	0.0	1400/1459	95	---	---
AF074711	<i>Burkholderia</i> sp. uncultured	EN-B3	1513	2409	0.0	1403/1462	95	4/1462	0
AJ292637	eubacterium	WD202	1494	2409	0.0	1398/1459	95	---	---
AB041772	<i>Burkholderia</i> sp. <i>Pseudomonas</i>	TH2	1455	2409	0.0	1389/1445	96	2/1445	0
U86373	sp. <i>Burkholderia</i>	LB400	1466	2403	0.0	1396/1456	95	1/1456	0
U96941	<i>graminis</i>		1584	2399	0.0	1392/1449	96	4/1449	0

AB48 (1.476 bases) - grupo ARDRA 8

	Unidentified beta								
AB015530	proteobacterium	BD1-33	1492	2517	0.0	1431/1479	96	5/1479	0
AF227856	Bacterium	96444	1434	2489	0.0	1387/1429	97	2/1429	0
AF262932	<i>Burkholderia</i> sp.	JS150	1521	2473	0.0	1425/1480	96	5/1480	0
AF074711	<i>Burkholderia</i> sp. uncultured	EN-B3	1513	2398	0.0	1416/1480	95	5/1480	0
AJ292643	eubacterium	WD268	1494	2348	0.0	1407/1478	95	3/1478	0
AF074712	<i>Burkholderia</i> sp. <i>Burkholderia</i>	EN-B9	1497	2338	0.0	1413/1483	95	11/1483	0
U96941	<i>graminis</i>	---	1584	2328	0.0	1357/1415	95	4/1415	0
AF247494	<i>Burkholderia</i> sp. uncultured	S2.1	1449	2322	0.0	1375/1441	95	2/1441	0
AJ292637	eubacterium uncultured	WD202	1494	2308	0.0	1402/1478	94	3/1478	0
AJ292644	eubacterium	WD285	1494	2301	0.0	1401/1478	94	3/1478	0

AB119 (1.478 bases) - grupo ARDRA 9

AF262932	<i>Burkholderia</i> sp.	JS150	1521	2583	0.0	1437/1479	97	2/1479	0
AF247494	<i>Burkholderia</i> sp.	S2.1	1449	2506	0.0	1400/1443	97	3/1443	0
	Unidentified beta								
AB015530	proteobacterium	BD1-33	1492	2502	0.0	1433/1481	96	8/1481	0
AF227856	Bacterium	96444	1434	2460	0.0	1388/1432	96	5/1432	0
AJ011508	<i>Burkholderia</i> sp. <i>Burkholderia</i>	DhA-54	1486	2452	0.0	1407/1462	96	3/1462	0
AF148554	<i>stabilis</i> uncultured	---	1485	2446	0.0	1408/1464	96	5/1464	0
AJ292643	eubacterium	WD268	1494	2442	0.0	1420/1479	96	4/1479	0
AB008002	<i>Burkholderia</i> sp. 2,4-D-degrading	MBIC3837	1480	2442	0.0	1407/1462	96	5/1462	0
AF184929	bacterium <i>Burkholderia</i>	TFD3	1392	2440	0.0	1349/1385	97	5/1385	0
AJ238360	<i>brasiliensis</i>	M130	1455	2434	0.0	1400/1455	96	3/1455	0

No. DE ACESSO NO GenBank	ESPÉCIE	ESTIRPE	COMP. (bases)	SCORE	VALOR E	conclusão			
						IDENTIDADES ⁽²⁾		GAPS ⁽³⁾	
						Proporção	%	Proporção	%

Ala (260 bases) - grupo ARDRA 10⁽⁴⁾

X79730	<i>Azospirillum lipoferum</i>	ATCC 29731	1333	460	10-127	251/256	98	1/256	0
X79729	<i>Azospirillum lipoferum</i>	ATCC 29708	1331	452	10-125	249/256	97	1/256	0
Z29619	<i>Azospirillum lipoferum</i>	ncimb 11861	1432	434	10-120	255/268	95	8/268	2
M59061	<i>Azospirillum lipoferum</i>	---	1439	402	10-110	248/266	93	6/266	2
AJ238567	<i>Azospirillum doebereineriae</i>	---	1303	385	10-105	242/258	93	---	---
X79736	<i>Azospirillum lipoferum</i>	F	1387	383	10-104	244/260	93	1/260	0
X79733	<i>Azospirillum brasilense</i>	DSM 1859	1387	381	10-103	244/260	93	1/260	0
X90759	<i>Azospirillum largomobile</i>	---	1422	373	10-101	243/261	93	2/261	0
X79731	<i>Azospirillum halopraeferens</i>	Au4	1323	373	10-101	239/256	93	---	---
Z29618	<i>Azospirillum halopraeferens</i>	DSM 3675	1417	373	10-101	239/256	93	---	---

BA131 (259 bases) - grupo ARDRA 11⁽⁴⁾

AF229875	<i>Ochrobactrum</i> sp.	4FB9	1409	500	10-139	259/260	99	1/260	0
AJ276036	<i>Ochrobactrum anthropi</i>	GH 1568	1426	500	10-139	259/260	99	1/260	0
AF047640	Uncultured eubacterium	TRB41	416	500	10-139	259/260	99	1/260	0
U88442	<i>Ochrobactrum anthropi</i>		1434	500	10-139	259/260	99	1/260	0
U88441	<i>Ochrobactrum anthropi</i>		1424	500	10-139	259/260	99	1/260	0
AJ242578	<i>Ochrobactrum anthropi</i>	1a	1321	500	10-139	259/260	99	1/260	0
AJ242576	<i>Ochrobactrum anthropi</i>	CLM18	697	500	10-139	259/260	99	1/260	0
AJ002812	<i>Ochrobactrum anthropi</i>	S1	1438	500	10-139	259/260	99	1/260	0
D12794	<i>Ochrobactrum anthropi</i>		1436	500	10-139	259/260	99	1/260	0
AF309080	<i>Ochrobactrum</i> sp.	MB2	680	494	10-137	256/257	99	1/257	0

(1) Os dados da pesquisa estão resumidos. Parâmetros relevantes da pesquisa (para todos os isolados):
 banco de dados: nucleotídeos, não redundante
 data: maio de 2001

no. de bases no banco de dados: 3.186.017.525

no. de seqüências no banco de dados: 852.500

valores: blastn matrix: 1 -3

penalizações para falhas: abertura = 5 e extensão = 2.

(2) As identidades mostram: número de bases idênticas entre o par de seqüências/ número total de bases do alinhamento.

- (3) Os *gaps* mostram: números de *gaps* ocorridos em ambas as seqüências/ número total de bases do alinhamento.
- (4) Seqüências da região Y1-Y2 do 16S rDNA.

3.6 RECONSTRUÇÃO FILOGENÉTICA

A partir das seqüências da região Y1-Y3 do 16S rDNA, tentou-se estabelecer o posicionamento filogenético dos isolados diazotróficos seqüenciados.

Inicialmente, as seqüências dos isolados foram alinhadas com as 50 seqüências de mais alto *score* obtido para cada uma das pesquisas de homologia (ver item "Análise comparativa de seqüências e pesquisa de homologia") após terem sido eliminadas as redundâncias e as seqüências dos próprios isolados, depositadas no GenBank. A árvore filogenética composta por 219 seqüências assim obtidas é mostrada na Figura 16. A árvore foi construída no programa Mega2 (Tabela 6) pelo método Neighbor-Joining, sendo usado um total de 1.097 sítios do alinhamento. Embora alguns ramos desta árvore não tenham se mostrado muito consistentes na análise de *bootstrap*, fica claro o posicionamento dos isolados em três grupos distintos das beta Proteobactérias: *Burkholderia*, *Herbaspirillum* e Comamonadaceae. Uma aproximação da árvore apresentada na Figura 16, feita na região onde os isolados estão posicionados, mostra com mais detalhes os organismos a eles relacionados. Cinco dos isolados foram incluídos no grupo formado pelo gênero *Burkholderia* (Figura 17). Entretanto, este grupo foi dividido em dois subgrupos, com os cinco isolados fazendo parte de um deles, próximos a bactérias ainda não classificadas (AB48, AB120 e BA126) ou a bactérias capazes de degradar xenobióticos (AB98 e AB119). Analogamente, a Figura 18 mostra a proximidade evolutiva do isolado BA153 com *Herbaspirillum seropedicae* (a seqüência destes dois organismos apresenta 100% de identidade), do isolado X8 com *Herbaspirillum frisingense* e dos isolados AB7, BA10, BA11, BA12, BA14, BA15, BA16, BA17, BA134, BA149 e BA161 com *Pseudomonas huttiensis*. Em uma análise filogenética molecular de todo o gênero *Pseudomonas*, baseada na seqüência do 16S rDNA, ANZAI *et al.* (2000) sugerem a reclassificação da espécie *P. huttiensis*, devido ao seu consistente posicionamento

filogenético entre o gênero *Herbaspirillum*. Finalmente, um único isolado, BA128, está relacionado com a família Comamonadacea, mais especificamente, com os gêneros *Xylophilus* e *Variovorax* e com espécies não classificadas (Figura 19).

Em outra análise, as seqüências dos isolados dos Grupos ARDRA 1 e 2 foram alinhadas com seqüências do 16S rDNA na região Y1-Y3 de *Herbaspirillum* sp. e espécies relacionadas. A filogenia foi inferida pelos métodos de Neighbor-Joining, Parcimônia e Máxima Verossimilhança (Figura 20A, B e C, respectivamente). Os isolados BA10, BA11, BA12, BA14, BA16, BA17, BA134, BA149 e BA161 formaram um grupo único e próximo a *Herbaspirillum rubrisubalbicans*, confirmando os resultados obtidos para a análise de ARDRA destes isolados. O isolado BA153 também confirmou o resultado obtido para a análise de ARDRA, ficando agrupado com *Herbaspirillum seropedicae*. Estes dois organismos compartilham 100% de identidade na seqüência da região Y1-Y3. Finalmente o isolado X8, que apesar de também ter sido colocado no Grupo ARDRA 1, mostrou-se mais próximo a *Herbaspirillum frisingense* do que de *H. seropedicae*. Este isolado já havia mostrado certa divergência na seqüência da região Y1-Y2. As posições dos isolados foram mantidas nas três árvores inferidas. *Herbaspirillum frisingense* é uma nova espécie proposta para o gênero *Herbaspirillum* e foi originalmente isolada de *Spartina pectinata*, *Miscanthus sinensis*, *Miscanthus sacchariflorus* e *Pennisetum purpureum*, sendo fisiologicamente diferenciada de *H. seropedicae* pela ausência de crescimento em meio contendo L-ramnose e *m*-inositol como fonte de carbono, e de *H. rubrisubalbicans* pelo crescimento em meio contendo N-acetilglucosamina e ausência de crescimento em presença de *m*-eritritol (KIRCHHOF *et al.*, 1997a). A capacidade de fixar nitrogênio em estirpes pertencentes a esta nova espécie foi confirmada por amplificação do gene estrutural da nitrogenase *nifD* e experimentos de redução de acetileno. A criação de uma nova espécie tem sido sugerida com base em análise do 16S rDNA e experimentos de hibridização DNA:DNA (KIRCHHOF *et al.*, 2001). Entretanto, pouco ainda é sabido sobre esta espécie.

De maneira análoga, os isolados AB48, AB98, AB119, AB120 e BA126 tiveram suas seqüências alinhadas e as árvores filogenéticas inferidas a partir dos diferentes métodos com seqüências de espécies de *Burkholderia* (Figura 21A, B e C). As posições do alinhamento correspondente ao intervalo entre os nucleotídeos 74 e 89 da seqüência de *Burkholderia cepacia* (número de acesso M22518) foram retiradas da análise devido a pouca confiabilidade observada no alinhamento desta região. Pequenas variações foram observadas no posicionamento de algumas espécies dentro das árvores obtidas por três diferentes métodos (ex., *Burkholderia multivorans* nas árvores B e C). Por outro lado, os isolados diazotróficos apareceram formando um grupo monofilético na "periferia" do grupo formado pelas espécies de *Burkholderia*, estando interiorizado apenas pela espécie *Burkholderia norimbergensis* e uma estirpe ainda não classificada. O grupo formado pelos isolados diazotróficos manteve-se consistente em todos os casos e com valores de *bootstrap* altos nas árvores inferidas pelos métodos de *Neighbor-Joining* e Parcimônia.

Uma análise filogenética foi conduzida ainda para seqüências do isolado BA128 e de espécies que compõem a família das Comamonadaceae (Figura 22A, B e C). As árvores obtidas para este grupo de organismos mostraram-se pouco consistentes. As árvores inferidas pelos métodos *Neighbor-Joining* e Parcimônia tiveram poucos ramos suportados por valores de *bootstrap* altos e houve uma maior variação do posicionamento dos organismos entre as árvores inferidas a partir dos três métodos. Entretanto, foi possível estabelecer uma proximidade evolutiva entre o isolado BA128 e espécies de *Acidovorax* e *Xylophilus ampelinus*.

A região Y1-Y2 do 16S rDNA contém pouca informação filogenética em relação a região Y1-Y3 e é uma região altamente variável, podendo sobreestimar os valores obtidos para os ramos de uma árvore filogenética. Entretanto a árvore inferida a partir desta região é em geral semelhante àquela obtida com a região Y1-Y3 e tem-se mostrado útil para estabelecer as relações evolutivas em vários organismos (CRUZ *et al.*, 2001; YOUNG *et al.*, 1991). Para os isolados Ala e BA131, somente a seqüência

da região Y1-Y2 foi obtida. Uma árvore inferida com o método de *Neighbor-Joining* com representantes da subdivisão alfa das Proteobactérias mostrou a posição do isolado Ala no interior do ramo formado por espécies de *Azospirillum*, próximo a *Azospirillum lipoferum* (Figura 23). Nesta mesma árvore encontra-se também o isolado BA131, próximo a *Ochrobactrum tritici*.

FIGURA 16 - ÁRVORE FILOGENÉTICA RECONSTRUÍDA A PARTIR DE 219 SEQÜÊNCIAS DO 16S rDNA DE ISOLADOS DE BANANEIRA E ABACAXIZEIRO E SEQÜÊNCIAS HOMÓLOGAS, OBTIDAS POR PESQUISA COM O PROGRAMA BLASTN. A ÁRVORE FOI CONSTRUÍDA PELO MÉTODO *NEIGHBOR-JOINING* A PARTIR DE DISTÂNCIAS CALCULADAS PELO MÉTODO DE JUKES & CANTOR; FORAM ANALISADOS 1.097 SÍTIOS NA REGIÃO Y1-Y3 COM 2.000 AMOSTRAS *BOOTSTRAP* (ALGUNS DOS VALORES ESTÃO INDICADOS EM PORCENTAGEM NA BIFURCAÇÃO DOS RAMOS); A ÁRVORE NÃO POSSUI RAIZ E A POSIÇÃO DOS ISOLADOS DIAZOTRÓFICOS É INDICADA POR UM LOSANGO.

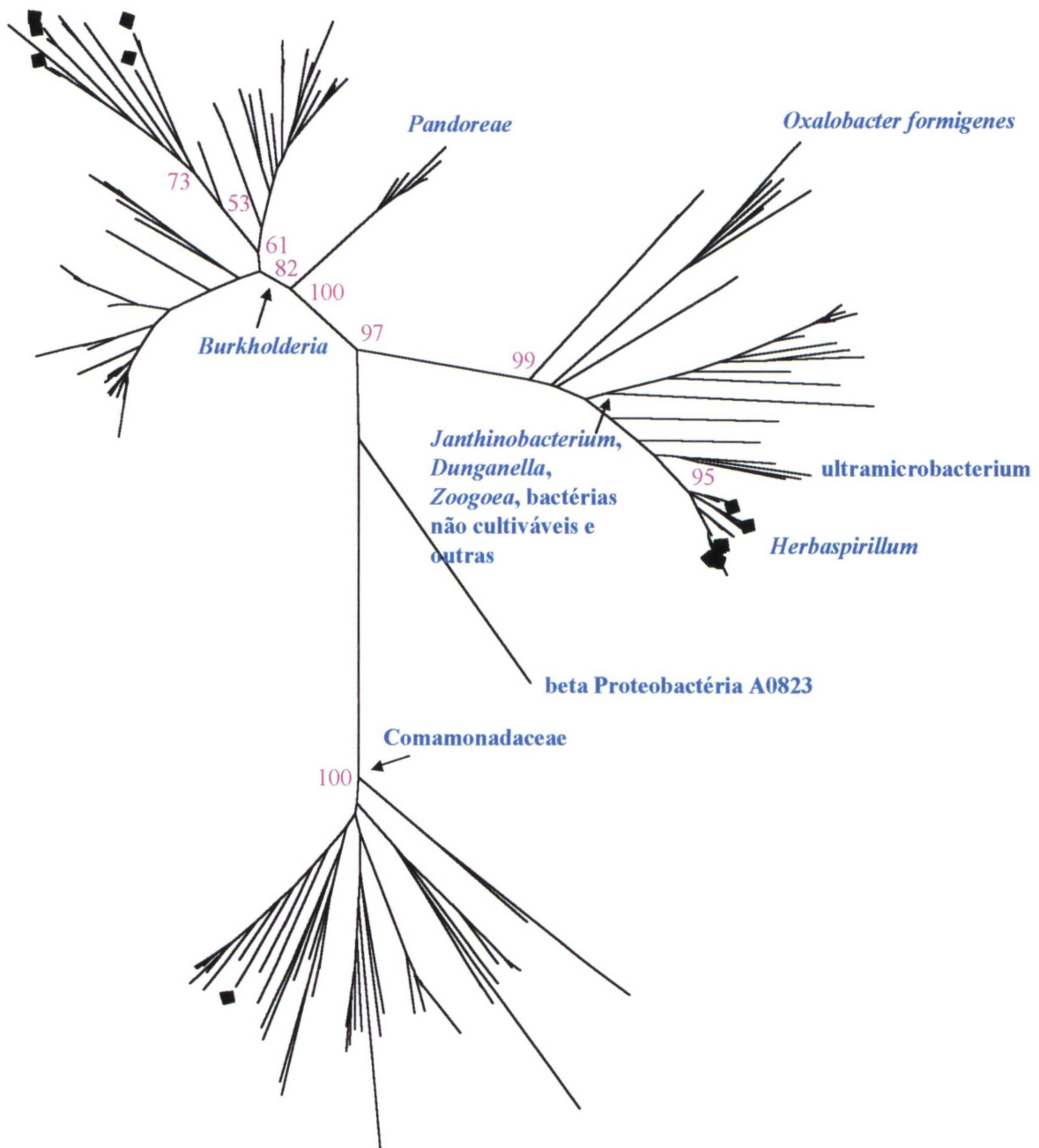


FIGURA 17 - DETALHAMENTO DA FILOGENIA DO GÊNERO *Burkholderia* E ESPÉCIES RELACIONADAS APRESENTADA NA ÁRVORE FILOGENÉTICA DA FIGURA 21. OS VALORES DE *BOOTSTRAP* ACIMA DE 50% SÃO MOSTRADOS (NÚMEROS EM LILÁS); OS ISOLADOS DIAZOTRÓFICOS SÃO MOSTRADOS EM AZUL; ESCALA EM NÚMERO DE SUBSTITUIÇÕES/ SÍTIO.

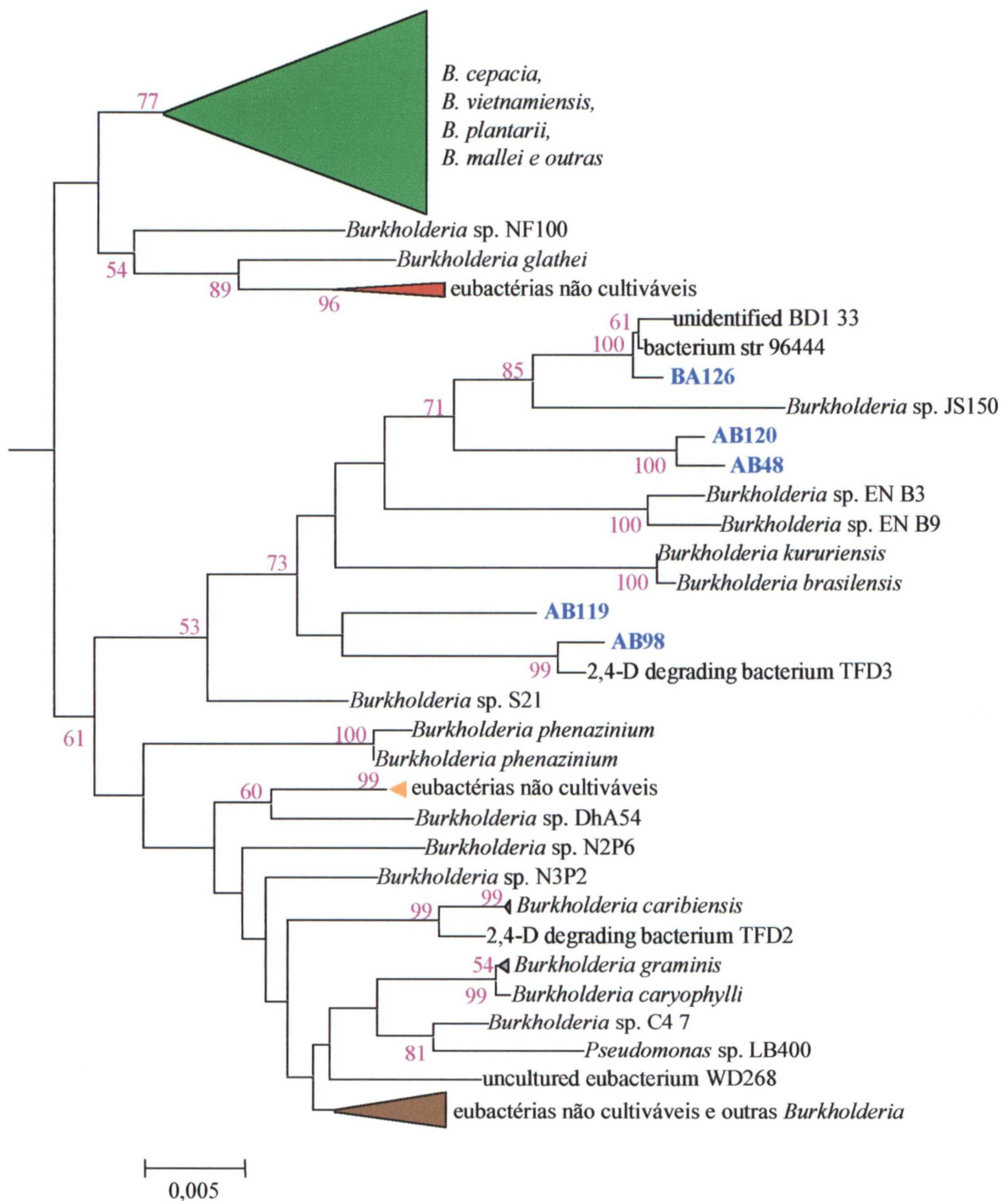


FIGURA 18 - DETALHAMENTO DA FILOGENIA DO GÊNERO *Herbaspirillum* E DOS ORGANISMOS RELACIONADOS APRESENTADA NA ÁRVORE FILOGENÉTICA DA FIGURA 21. OS VALORES DE *BOOTSTRAP* ACIMA DE 50% SÃO MOSTRADOS; ISOLADOS DIAZOTRÓFICOS EM AZUL; ESCALA EM NÚMERO DE SUBSTITUIÇÕES/ SÍTIO.

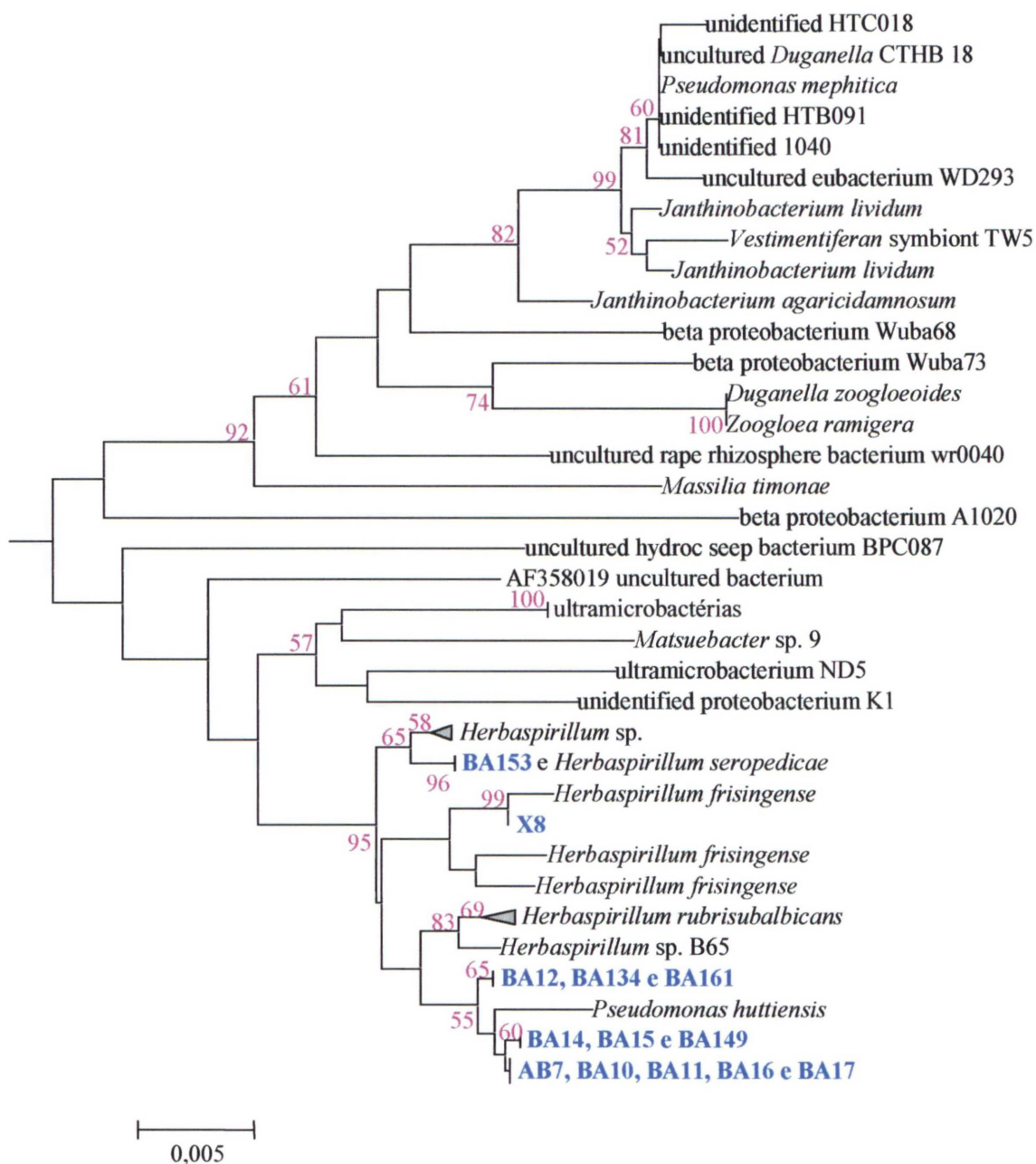


FIGURA 19 - DETALHAMENTO DA FILOGENIA DA FAMÍLIA COMAMONADACEAE E DOS ORGANISMOS RELACIONADOS APRESENTADA NA ÁRVORE FILOGENÉTICA DA FIGURA 21. OS VALORES DE *BOOTSTRAP* ACIMA DE 50% SÃO MOSTRADOS; O ISOLADO DIAZOTRÓFICO BA128 É MOSTRADO EM AZUL; ESCALA EM NÚMERO DE SUBSTITUIÇÕES/ SÍTIO.

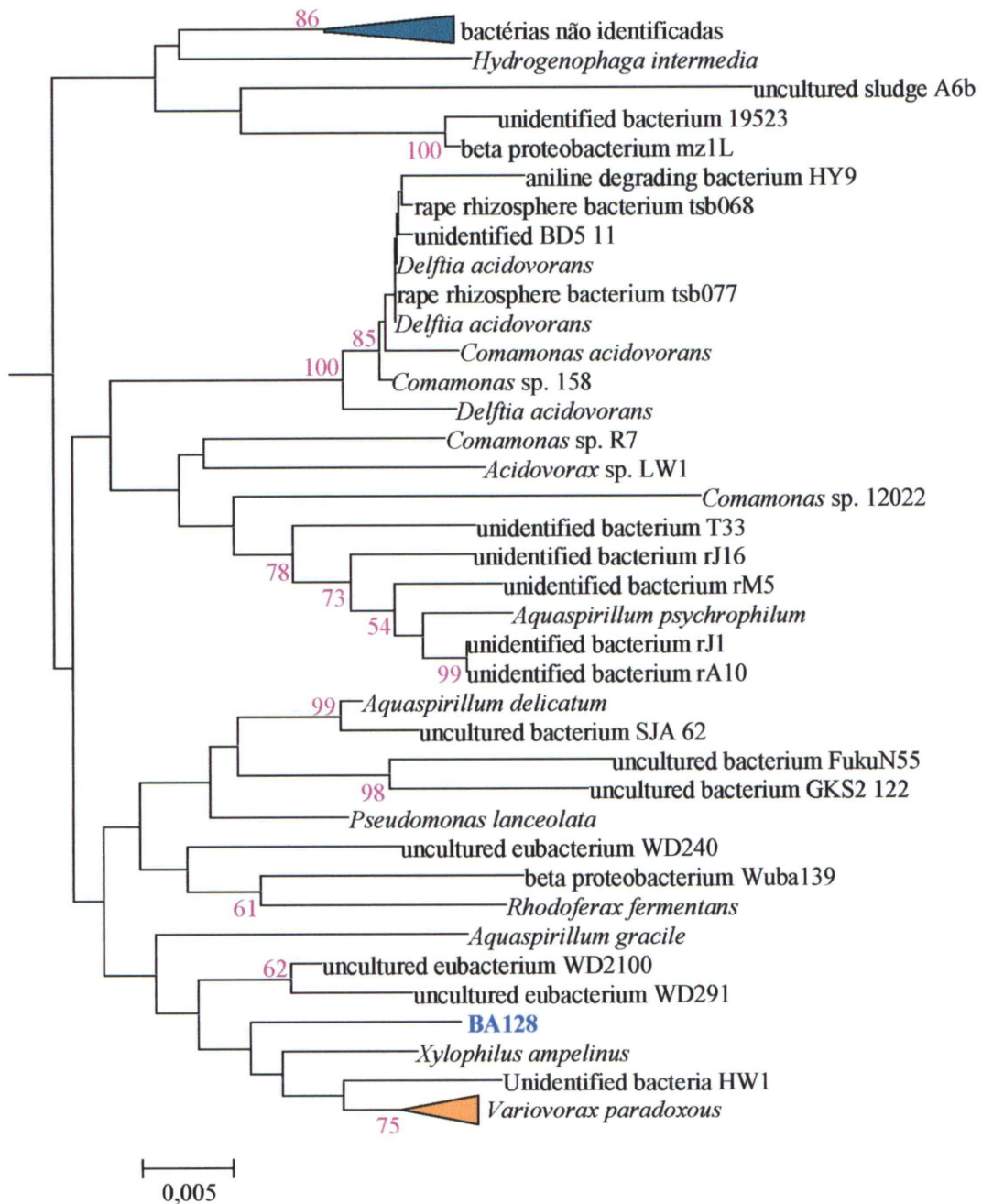


FIGURA 20 - ÁRVORE FILOGENÉTICA DA REGIÃO Y1-Y3 DO 16S rDNA DE ISOLADOS DE BANANEIRA E ABACAXIZEIRO E ESPÉCIES RELACIONADAS. AS POSIÇÕES DOS ISOLADOS DIAZOTRÓFICOS SÃO MOSTRADAS EM VERMELHO; A SEQÜÊNCIA DE *Alcaligenes faecalis* FOI USADA COMO RAIZ DA ÁRVORE; A ESCALA REPRESENTA 0,01 SUBSTITUIÇÕES DE NUCLEOTÍDEOS/ SÍTIO. **A** - OS VALORES DE *BOOTSTRAP* ACIMA DE 50% SÃO MOSTRADOS PARA 5.000 REPETIÇÕES; A TOPOLOGIA FOI INFERIDA PELO MÉTODO *NEIGHBOR-JOINING* COM DISTÂNCIAS DE JUKES & CANTOR; **B** - OS VALORES DE *BOOTSTRAP* ACIMA DE 50% (ABAIXO DESTES VALORES OS RAMOS SÃO MOSTRADOS COMO NÃO RESOLVIDOS) PARA 500 REPETIÇÕES SÃO MOSTRADOS; A TOPOLOGIA FOI INFERIDA PELO MÉTODO DA PARCIMÔNIA; A ÁRVORE APRESENTADA CORRESPONDE A ÁRVORE CONSENSO DE 117 ÁRVORES; **C** - TOPOLOGIA DE UMA ÁRVORE INFERIDA PELO MÉTODO DA MÁXIMA VEROSSIMILHANÇA.

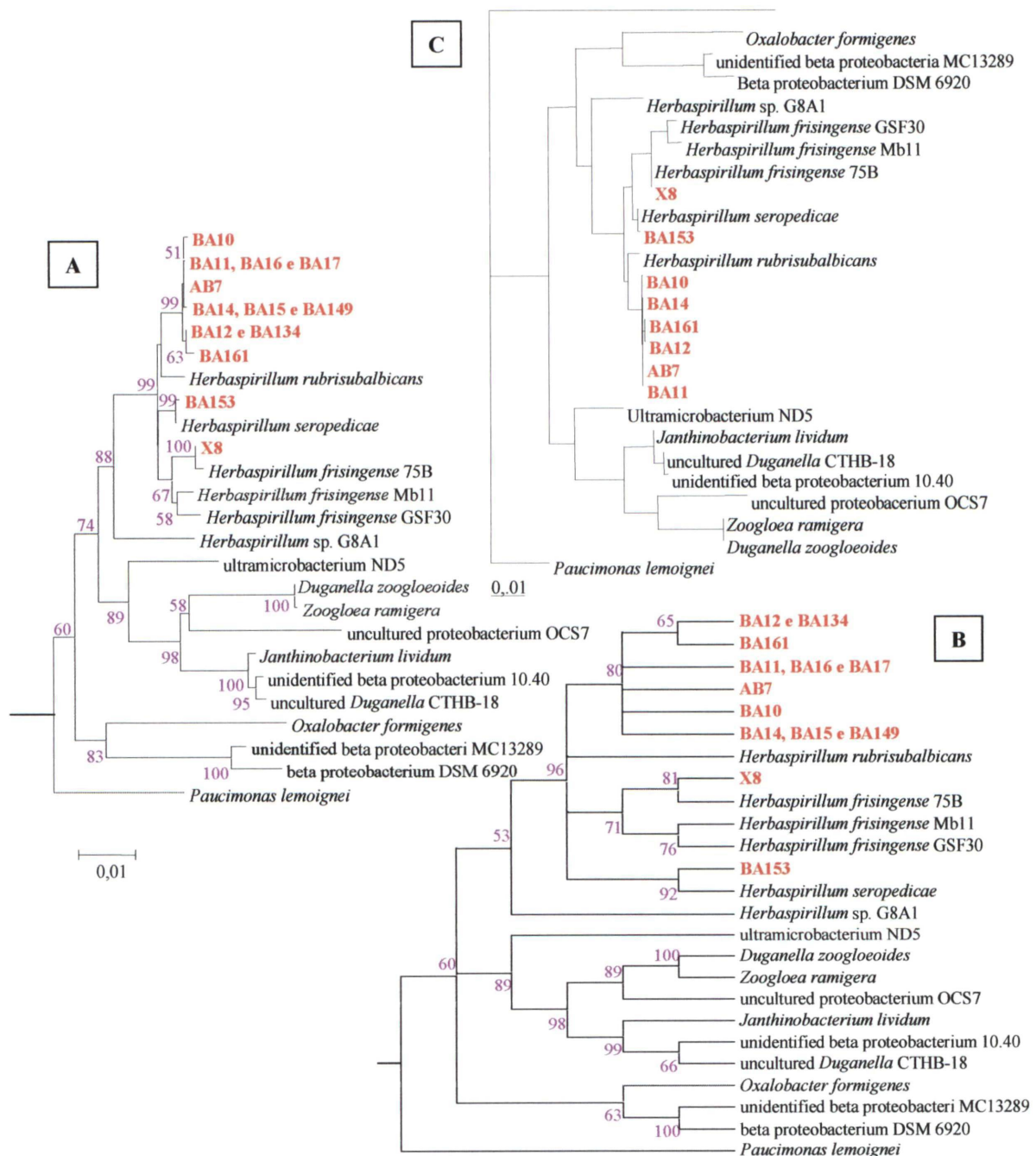


FIGURA 21 - ÁRVORE FILOGENÉTICA DA REGIÃO Y1-Y3 DO 16S rDNA DE *Burkholderia* sp. AS POSIÇÕES DOS ISOLADOS DIAZOTRÓFICOS SÃO MOSTRADAS EM VERMELHO; A SEQUÊNCIA DE *Alcaligenes faecalis* FOI USADA COMO RAIZ DA ÁRVORE; A ESCALA REPRESENTA 0,01 SUBSTITUIÇÕES DE NUCLEOTÍDEOS/ SÍTIO. **A** - OS VALORES DE 5.000 REPETIÇÕES *BOOTSTRAP* SÃO MOSTRADAS EM PORCENTAGEM PARA VALORES ACIMA DE 50%; A TOPOLOGIA FOI INFERIDA PELO MÉTODO *NEIGHBOR-JOINING* A PARTIR DE DISTÂNCIAS CALCULADAS PELO MÉTODO DE JUKES & CANTOR; **B** - OS VALORES DE 500 REPETIÇÕES *BOOTSTRAP* SÃO MOSTRADAS EM PORCENTAGEM PARA VALORES ACIMA DE 50%; OS RAMOS COM VALORES DE *BOOTSTRAP* MENORES OU IGUAIS A 50% SÃO MOSTRADOS COMO RAMOS NÃO RESOLVIDOS; A TOPOLOGIA FOI INFERIDA PELO MÉTODO DE PARCIMÔNIA; A ÁRVORE APRESENTADA CORRESPONDE A ÁRVORE CONSENSO DE 9 ÁRVORES; **C** - TOPOLOGIA INFERIDA PELO MÉTODO DA MÁXIMA VEROSSIMILHANÇA.

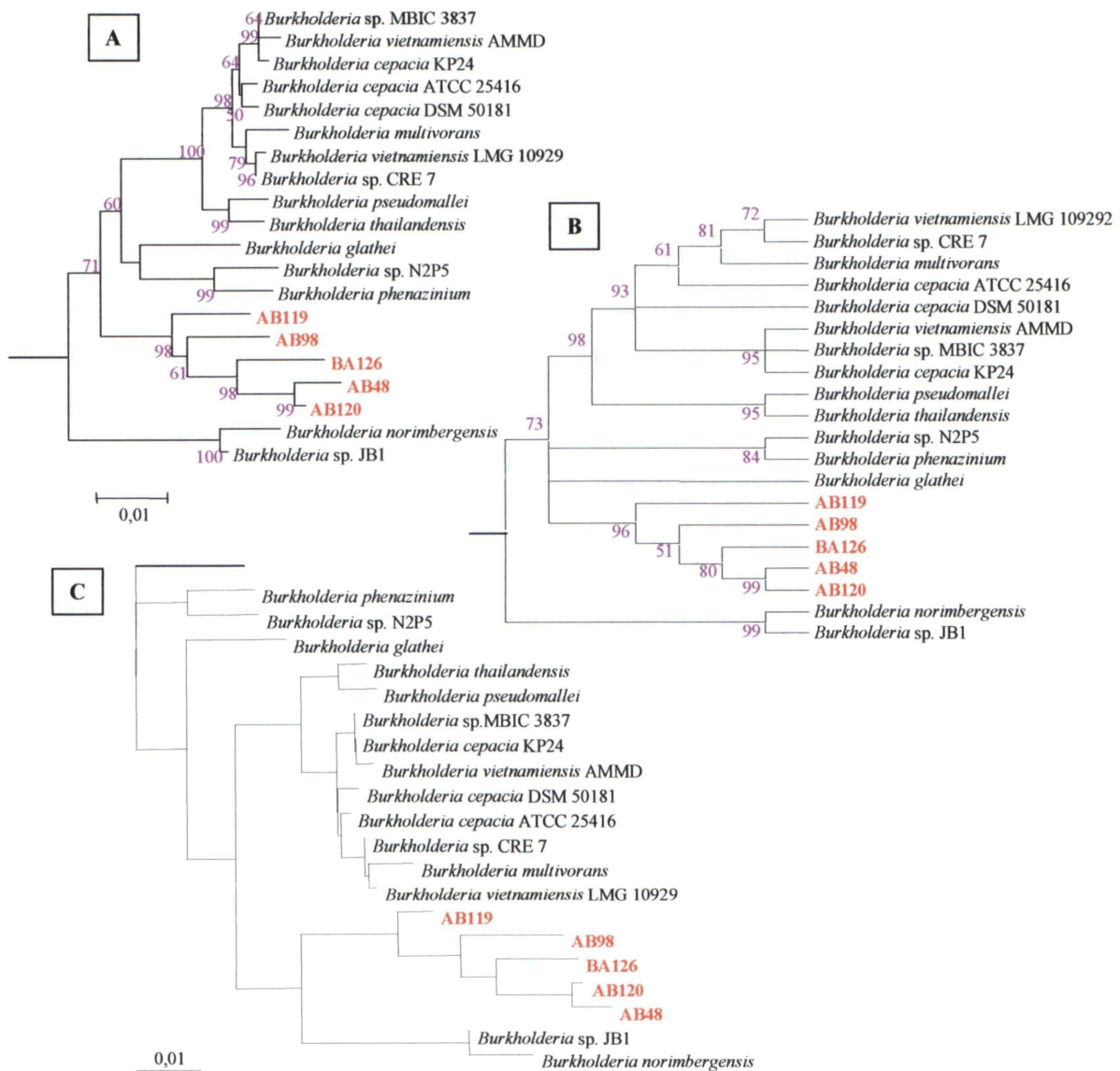


FIGURA 22 - ÁRVORE FILOGENÉTICA DA REGIÃO Y1-Y3 DO 16S rDNA DO ISOLADO BA128 E ESPÉCIES DE COMAMONADACEAE RELACIONADAS. A POSIÇÃO DO ISOLADO DIAZOTRÓFICO É MOSTRADA EM VERMELHO; A SEQÜÊNCIA DE *Escherichia coli* FOI USADA COMO RAIZ DA ÁRVORE; A ESCALA REPRESENTA SUBSTITUIÇÕES DE NUCLEOTÍDEOS/ SÍTIO. **A** - OS VALORES DE *BOOTSTRAP* ACIMA DE 50% SÃO MOSTRADOS PARA 5.000 REPETIÇÕES; A TOPOLOGIA FOI INFERIDA PELO MÉTODO *NEIGHBOR-JOINING* COM DISTÂNCIAS DE JUKES & CANTOR; **B** - OS VALORES DE *BOOTSTRAP* ACIMA DE 50% SÃO MOSTRADOS PARA 500 REPETIÇÕES; A TOPOLOGIA FOI INFERIDA PELO MÉTODO DE PARCIMÔNIA; A ÁRVORE APRESENTADA CORRESPONDE A ÁRVORE CONSENSO DE 4 ÁRVORES; **C** - TOPOLOGIA FOI INFERIDA PELO MÉTODO DA MÁXIMA VEROSSIMILHANÇA.

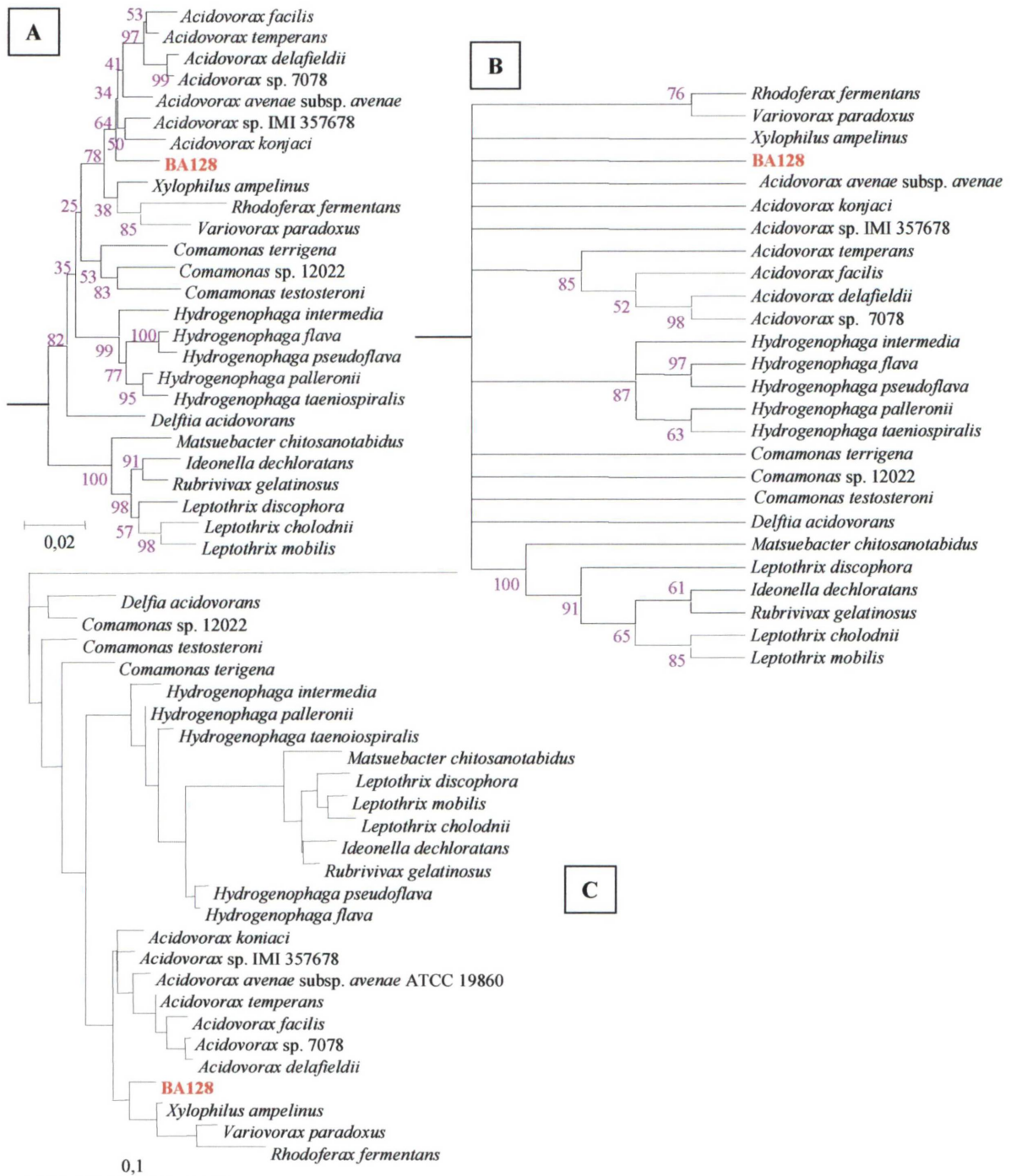
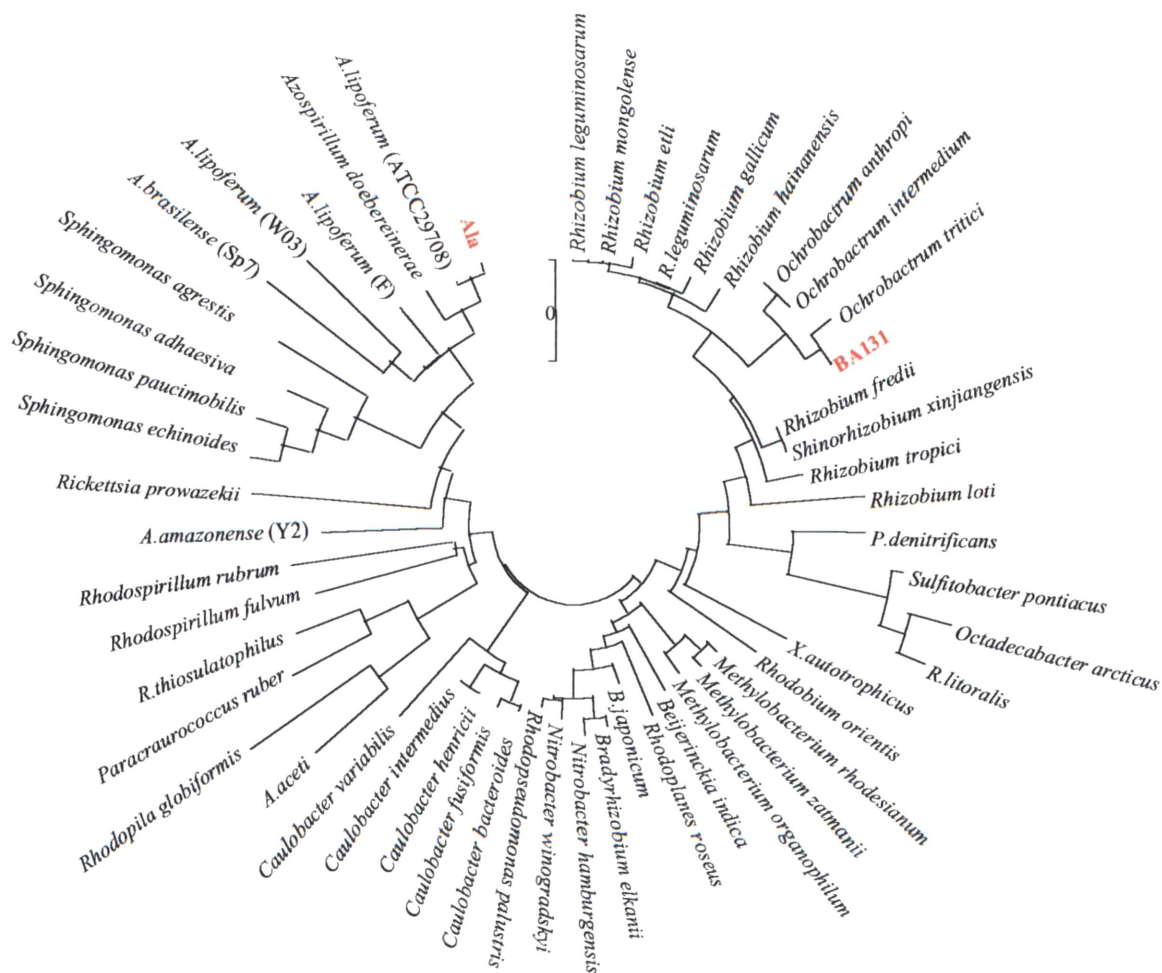


FIGURA 23 - ÁRVORE FILOGENÉTICA DA REGIÃO Y1-Y2 DO 16S rDNA MOSTRANDO O POSICIONAMENTO DOS ISOLADOS BA131 E Ala ENTRE ESPÉCIES DE ALFA PROTEOBACTÉRIAS. AS POSIÇÕES DOS ISOLADOS DIAZOTRÓFICOS SÃO MOSTRADAS EM VERMELHO; OS NÚMEROS DE ACESSO NO GENBANK PARA AS SEQUÊNCIAS USADAS ENCONTRAM-SE LISTADOS NA TABELA 12; A TOPOLOGIA FOI INFERIDA PELO MÉTODO NEIGHBOR-JOINING A PARTIR DE DISTÂNCIAS CALCULADAS PELO MÉTODO DE JUKES & CANTOR.



4 DISCUSSÃO

4.1 ANÁLISE DE RESTRIÇÃO DO 16S rDNA E REGIÃO INTERGÊNICA 16S-23S rDNA

A simples comparação eletroforética dos fragmentos amplificados da região intergênica 16S-23S rDNA, muitas vezes, possibilita uma caracterização dos microrganismos, por ser uma região altamente variável, tanto no comprimento quanto na sua seqüência. Em uma técnica denominada RISA (*rRNA intergenic spacer analysis*) a amplificação por PCR e posterior separação destes fragmentos em gel de poliacrilamida foi usada para estimar a diversidade em comunidades microbianas do solo, rizosfera e ambientes aquáticos (BORNEMAN & TRIPLETT, 1997; ROBLETO *et al.*, 1998; ACINAS *et al.*, 1999). A variação no comprimento das regiões intergênicas 16S-23S rDNA entre as Eubactérias e Archeobactérias é da ordem de 10 vezes, indo de 143 pb a 1.529 pb. Esta ampla variação se deve, em parte, à presença ou não de seqüências codificadoras para tRNAs nestas regiões (FISHER & TRIPLETT, 1999). A amplificação da região intergênica 16S-23S rDNA dos novos isolados diazotróficos sugerem um alto grau de proximidade entre a maior parte dos isolados analisados e permitem especular sobre a presença de seqüências codificadoras de tRNAs nestes organismos. É importante lembrar que o tamanho do amplificado não corresponde exatamente ao tamanho da região intergênica, já que os iniciadores usados anelam em regiões conservadas do 16S e 23S rDNAs. A estrutura do operon *rrnB* de *E. coli* (BROSIUS *et al.*, 1981) mostra que a região de anelamento do iniciador pHR no 16S rDNA ocorre um único nucleotídeo antes da ponta 3' da molécula. Entretanto, a região de anelamento do iniciador p23 começa exatamente a 464 pb para o interior do 23S rDNA. Portanto, neste organismo, o comprimento da região intergênica é de 440 pb. Se considerarmos o tamanho do fragmento de um iniciador ao outro, incluindo as seqüências dos próprios iniciadores, teríamos um amplificado teórico de 946 pb

para *E. coli*. Admitindo-se que a porção do 23S rDNA é conservada nos isolados e estirpes referência, e igual, ao menos em comprimento, à de *E. coli*, chegamos a um valor de aproximadamente 236-336 pb, 536 pb e 636 pb para as regiões intergênicas dos amplificadores de 700-800 pb, 1.000 pb e 1.100 pb, respectivamente. O primeiro fragmento possui um tamanho próximo ao tamanho médio para a região intergênica 16S-23S rDNA de bactérias gram positivas, que é de 327 ± 111 pb e que não possuem seqüência codificadora para tRNAs. Por outro lado, os dois outros fragmentos com regiões intergênicas de cerca de 536 e 636 pb, são semelhantes ao tamanho médio da região intergênica 16S-23S rDNA de bactérias gram negativas, que é de 533 ± 233 pb e possuem tRNAs codificados no interior desta região (FISHER & TRIPLET, 1999).

Normalmente, os organismos procariotos possuem os RNAs ribossomais de 16S, 23S e 5S agrupados em um único operon co-transcrito nesta ordem (BROSIUS *et al.*, 1981). É comum a ocorrência de mais de um operon de RNA ribossomal em um mesmo organismo. *E. coli* apresenta sete operons (KENERLEY *et al.*, 1977; KISS *et al.*, 1977) com um ou dois genes que codificam para tRNAs na região intergênica entre os genes do 16S rRNA e 23S rRNA (WU & DAVIDSON, 1975). Por ser uma região altamente variável é possível ainda obter mais de um fragmento amplificado para a região intergênica 16S-23S rDNA de um mesmo organismo. O número de fragmentos obtidos indica a ocorrência de, no mínimo, igual número de operons para os DNAs ribossomais no organismo em questão. Por outro lado, uma banda única como a obtida para os isolados em estudo, sugere a presença de um ou mais operons altamente conservados nestes organismos, distribuídos pelo genoma. Em *Herbaspirillum seropedicae*, estirpe SMR1, três bandas do perfil eletroforético obtido pela técnica de Eletroforese em Gel de Campo Pulsado (PFGE) com DNA genômico clivado com *SwaI*, *PacI* e *I-CeuI* hibridizaram com sonda de 16S rDNA (J. L. R. SOARES-RAMOS, comunicação pessoal).

As diferenças nas seqüências da região intergênica 16S rDNA e 16S-23S rDNA entre os organismos podem ser, em parte, reveladas através dos sítios de

restrição contidos nos fragmentos amplificados. A diferença entre sítios conservados e mutados produz um padrão de fragmentos de DNA característico para cada organismo, após tratamento dos amplificados com endonucleases de restrição. A medida que os organismos evoluem, mutações podem alterar os sítios de restrição para uma determinada endonuclease, alterando a seqüência de bases reconhecida pela enzima com conseqüente perda do sítio de restrição ou, de modo inverso, criando um sítio de restrição onde este não existia. Desta forma o perfil de restrição para as endonucleases difere entre os organismos, mantendo uma correlação entre o número de fragmentos compartilhados, ou o número de sítios de restrição conservados, e o grau de divergência evolutiva (NEI & LI, 1979).

Na análise de restrição, nossos dados mostraram que duas enzimas atuando em conjunto (ex., *AluI* e *HaeIII*) foram suficientes para resolver as estirpes testadas nos 12 grupos a partir do 16S rDNA. Foi sugerido que uma combinação de três ou mais endonucleases é suficiente para identificação de espécies pela análise de ARDRA (LAGUERRE *et al.*, 1994; MOYER *et al.*, 1996; VINUESA *et al.*, 1998).

A análise de restrição da região intergênica 16S-23S rDNA comumente apresenta maior poder de resolução (KHBAYA *et al.*, 1998; MEHTA & ROSATO, 2001). Para os isolados diazotróficos, 15 Grupos 16S-23S rDNA foram obtidos (Tabela 6), comparados com os 12 Grupos ARDRA (Tabela 7). Entretanto poucos dos grupos obtidos por ARDRA mostraram-se heterogêneos na análise de restrição do 16S-23S rDNA (Grupos 16S-23S rDNA A, B, E, I e J). Por outro lado, os isolados BA153 e X8, incluídos no Grupo 16S-23S rDNA A formaram um grupo distinto na análise de ARDRA.

Alguns genótipos foram encontrados somente em plantas de bananeira (ex., Grupos ARDRA 5 e 7) e outros em abacaxizeiro (ex., Grupos ARDRA 4, 6 e 8). Entretanto, nenhuma correlação específica foi observada entre as variedades e locais de coleta das plantas e os genótipos detectados, confirmando o largo espectro hospedeiro para estes organismos. Apesar de nenhuma esterilização ter sido feita na

superfície das plantas antes do processo de isolamento (WEBER *et al.*, 1998), a presença de *Herbaspirillum* e *Burkholderia* sugere uma natureza endofítica desta associação. Para todos os isolados foi estimado uma ocorrência em números substanciais nos tecidos analisados, variando de 10^2 a 10^6 células por grama de material fresco, uma população similar àquela encontrada para diversos endofíticos associados a gramíneas (WEBER *et al.*, 1998).

Embora os métodos empregados não sejam suficientes para definir uma nova espécie (LAGUERRE *et al.*, 1994), a combinação de dados de ARDRA, a partir dos perfis de restrição do 16S rDNA obtidos com 3 ou 4 endonucleases, é suficiente para produzir um relacionamento evolucionário confiável (MOYER *et al.*, 1996). Além do mais, dados publicados para rizóbios e *Azospirillum* (GRIFONI *et al.*, 1995) indicam que este método é uma poderosa ferramenta para caracterização de organismos com base em relações evolutivas e detecção de novas taxa. Aplicando a técnica de ARDRA em uma simulação por computador, MOYER *et al.* (1996) foram capazes de detectar mais de 99% das espécies, a partir de um modelo com mais de 100 seqüências de 16S rDNA distribuídas entre os domínios das Bactérias e usando um mínimo de três endonucleases de restrição. Baseado nestes dados, nós concluímos que os 12 Grupos ARDRA determinados pelos dados de 4 endonucleases combinados são verdadeiros e representam 12 diferentes grupos distintos de organismos.

4.2 ANÁLISE DAS SEQÜÊNCIAS E ESTRUTURA SECUNDÁRIA

A análise da região Y1-Y2 do 16S rDNA foi empregada para diferenciar espécies de “*Pseudomonas*” (SEAL *et al.*, 1993), caracterizar isolados de *Bradyrhizobium*, *Rhizobium* e *Sinorhizobium* (VINUESA *et al.*, 1998; HAUKKA *et al.*, 1996) e sugerir a inclusão de um simbiote fototrófico (BTai1) em um grupo taxonômico diferente (YOUNG *et al.*, 1991). Esta região do 16S rDNA é altamente variável e possibilita a diferenciação de organismos próximos. Entretanto, o

seqüenciamento desta região não foi suficiente para diferenciar os isolados diazotróficos dos grupos ARDRA 1 e 2.

As seqüências do 16S rDNA, dos Grupos I ao IX e XII, variaram de 286 a 290 pb, na região Y1-Y2, excluindo-se os iniciadores (Tabela 10). Esta variação foi similar a encontrada por SEAL *et al.* (1993) para as beta Proteobactérias "*Pseudomonas*" *solanacearum*, "*Pseudomonas*" *syzygii* e "*Pseudomonas*" *pickettii*, usando os mesmos iniciadores. Por outro lado, as seqüências da região Y1-Y2 dos Grupos X e XI apresentaram somente 260 e 259 pb, respectivamente, semelhante ao tamanho (260 pb) encontrado para espécies de *Rhizobium*, *Bradyrhizobium* e outras espécies de alfa Proteobactérias (YOUNG *et al.*, 1991). Existe uma relação aparente entre o tamanho e a estrutura secundária assumida pela molécula de 16S rRNA e a subdivisão das Proteobactérias a que pertence um organismo. Duas hélices no 16S rRNA, posições 184-193 e 198-219 (numeração em *E. coli*), ajudam a definir e distinguir entre quatro subdivisões que compõem a divisão das Proteobactérias (WOESE, 1987). Esta diferença pode ser observada na formação das hélices 10 e 11, respectivamente, de organismos representantes das subdivisões das Proteobactérias na Figura 24. Observa-se que a hélice 10 apresenta um "braço" curto nas subdivisões alfa, beta, gama e epsilon, sendo composta por três pares de nucleotídeos ligados mais quatro nucleotídeos formando uma alça (*loop*). Entre as delta Proteobactérias nota-se a presença de um "braço" bem mais longo, mas com seqüência pouco conservada. Os isolados dos Grupos I ao IX seguem claramente o padrão das beta Proteobactérias, como observado na região correspondente à hélice 10 no alinhamento da Figura 25, com as variações ocorrendo de forma compensatória no sentido de manter intacta a estrutura secundária. A base ambígua observada no seqüenciamento do Grupo IX ocorre na região de alça da estrutura, onde se observa a maior variação na seqüência. Por outro lado os Grupos X e XI seguem fielmente o padrão apresentado pela subdivisão alfa. A hélice 11 (Figura 24), por sua vez, apresenta um "braço" curto nas alfa e epsilon Proteobactérias e um "braço" mais longo nas beta, gama e delta

Proteobactérias. Neste caso também são claros os padrões apresentados pelos Grupos de seqüências dos isolados, com os Grupos de I ao IX semelhantes ao padrão da subdivisão beta e os Grupos X e XI ao padrão da subdivisão alfa.

Uma terceira região, entre as posições 21 a 60 da Figura 25, onde há uma extensão relativamente grande de falhas (*gaps*) nas seqüências dos Grupos X, XI e consenso para alfa Proteobactérias em relação às demais, também contribui para que ocorra variação no comprimento das seqüências. Os Grupos X e XI mostraram ainda certa divergência em relação à seqüência consenso das alfa Proteobactérias no alinhamento, mas o caráter variável das bases nesta região não permite chegar a uma conclusão.

FIGURA 24 - ESTRUTURA SECUNDÁRIA DO 16S rRNA DE REPRESENTANTES DAS CINCO SUBDIVISÕES DAS PROTEOBACTÉRIAS (DE RIJK ET AL., 2000). A ESTRUTURA DE TODA A REGIÃO Y1-Y2 É MOSTRADA PARA *Burkholderia cepacia* (NUCLEOTÍDEOS DE 44 A 337 EM *Escherichia coli*); AS DEMAIS SÃO MOSTRADAS ENTRE OS NUCLEOTÍDEOS DE 139 A 224. O NOME DAS ESPÉCIES USADAS SÃO INDICADOS, SEGUIDO DO NÚMERO DE ACESSO DA SEQUÊNCIA NO GENBANK EM PARÊNTESES E DA SUBDIVISÃO A QUAL A ESPÉCIE PERTENCE.

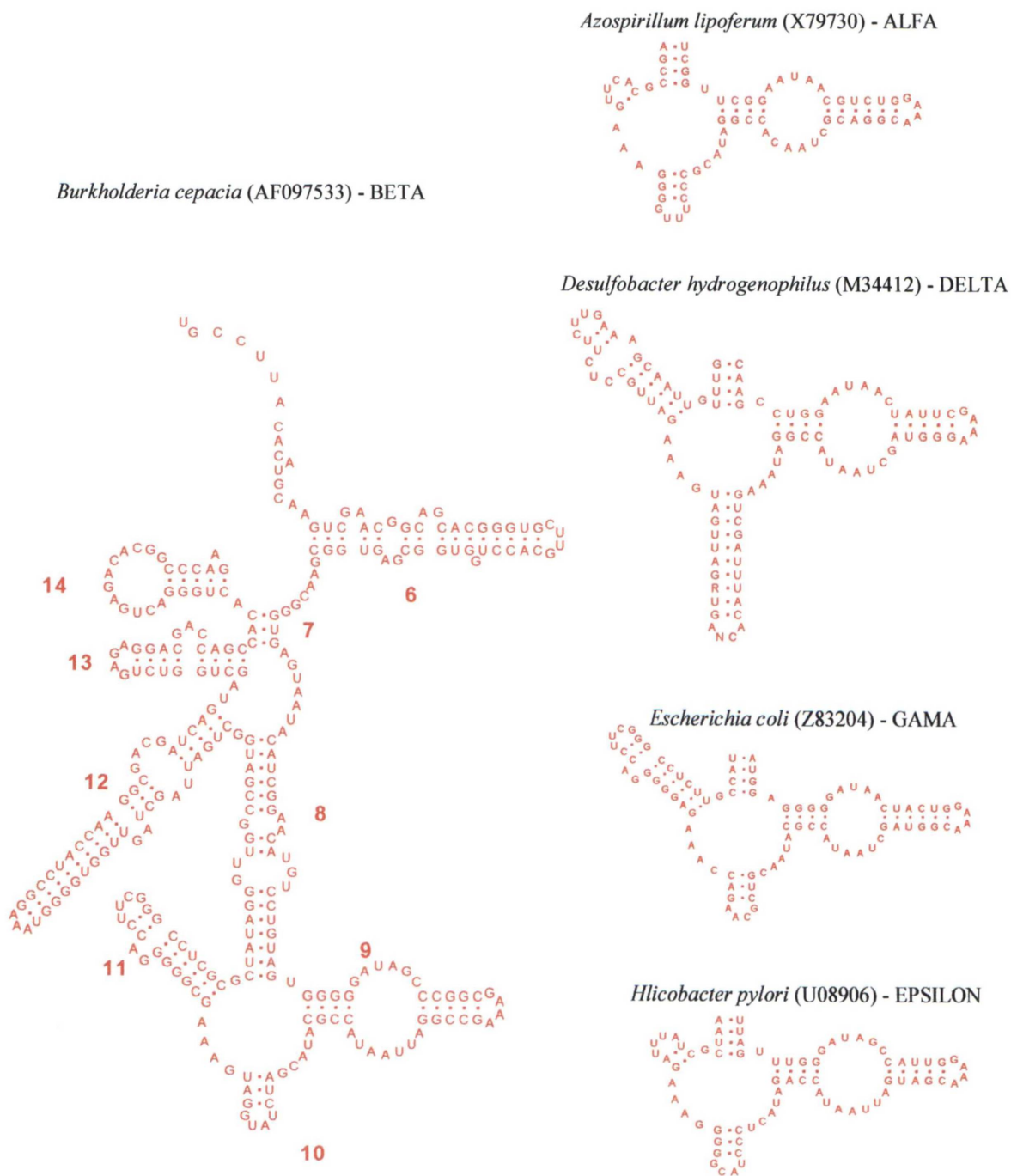


FIGURA 25 - ALINHAMENTO ENTRE OS GRUPOS DE SEQÜÊNCIAS Y1-Y2 E SEQÜÊNCIAS CONSENSO PARA AS SUBDIVISÖES DAS PROTEOBACTÉRIAS (INDICADO NO TÍTULO À ESQUERDA). O ALINHAMENTO FOI FEITO NO PROGRAMA CLUSTALW; AS ESTIRPES PERTENCENTES A CADA UM DOS GRUPOS ESTÃO LISTADAS NA TABELA 10; PONTOS DE INTERROGAÇÃO NAS SEQÜÊNCIAS CONSENSO INDICAM OS SÍTIOS ONDE HÁ UMA VARIAÇÃO DE BASES MAIOR DE 40% NAS SEQÜÊNCIAS ANALISADAS; OS RETÂNGULOS MARCAM AS POSIÇÕES DAS HÉLICES 10 E 11 DA FIGURA 24.

	10	20	30	40	50	60	70	80	90
GRUPO_I	ATGCCTTACA	CATGCAAGTC	GAACGGCAGC	ATAGGAG--C	TTG---CTC	CTGATGGCGA	GTGGCGAACG	GGTGAGTAAT	ATATCGG-AA
GRUPO_II	ATGCCTTACA	CATGCAAGTC	GAACGGCAGC	ATAGGAG--C	TTG---CTC	CTGATGGCGA	GTGGCGAACG	GGTGAGTAAT	ATATCGG-AA
GRUPO_III	ATGCCTTACA	CATGCAAGTC	GAACGGCAGC	ATAGGAG--C	TTG---CTC	CTGATGGCGA	GTGGCGAACG	GGTGAGTAAT	ATATCGG-AA
GRUPO_IV	ATGCCTTACA	CATGCAAGTC	GGACGGCAGC	GCGGG---C	TTC---GGC	CTGGCGGCGA	GTGGCGAACG	GGTGAGTAAT	ACATCGG-AA
GRUPO_V	ATGCCTTACA	CATGCAAGTC	GAACGGCAGC	ACGGGTG--C	TTG---CAC	CTGGTGGCGA	GTGGCGAACG	GGTGAGTAAT	ACATCGG-AA
GRUPO_VII	ATGCCTTACA	CATGCAAGTC	GGACGGCAGC	GCGGG---C	TTC---GGC	CTGGCGGCGA	GTGGCGAACG	GGTGAGTAAT	ACATCGG-AA
GRUPO_VIII	ATGCCTTACA	CATGCAAGTC	GGACGGCAGC	GCGGG---A	GCA---ATC	CTGGCGGCGA	GTGGCGAACG	GGTGAGTAAT	ACATCGG-AA
GRUPO_IX	ATGCCTTACA	CATGCAAGTC	GGACGGCAGC	GCGGG---C	TTC---GGC	CTGGCGGCGA	GTGGCGAACG	GGTGAGTAAT	ACATCGG-AA
GRUPO_XII	ATGCCTTACA	CATGCAAGTC	GAACGGCAGC	ACGGGTG--C	TTG---CAC	CTGGTGGCGA	GTGGCGAACG	GGTGAGTAAT	ACATCGG-AA
GRUPO_VI	ATGCCTTACA	CATGCAAGTC	GAACGGTAAAC	A----GG--T	CTT---CGG	ATGCTGACGA	GTGGCGAACG	GGTGAGTAAT	ACATCGG-AA
GRUPO_X	ATGCCTAACA	CATGCAAGTC	GAACGAAGG-	-----	-----C	TTGGCCCTTA	GTGGCGCACG	GGTGAGTAAC	ACGTGGG-AA
GRUPO_XI	AGGCTTAACA	CATGCAAGTC	GAGCGCC--	-----	-----	CGCAAGGGGA	GCGGCAGACG	GGTGAGTAAC	GCGTGGG-AA
Consenso_alfa	AGGC?TAACA	CATGCAAGTC	GA?CG?????	-----	-----	-----	G?GGC?GACG	GGTGAGTAAC	GCGTGGG-AA
Consenso_beta	ATGC?TTACA	CATGCAAGTC	GAACGGCAGC	??GG??--?	T??-----	??T??G??G?	GTGGCGAACG	GGTGAGTAAT	?CATCGG-AA
Consenso_gama	AGGC?TAACA	CATGCAAGTC	GA?CGG?A?C	?????????C	TTG--?????	??G??G?CGA	G?GGCGGACG	GGTGAGTAAT	GC?T?GG-AA
Consenso_epsilon	GTGCCTAATA	CATGCAAGTC	GAACGATGAA	G???-?TAGC	TTGCTAG--	??G??GATTA	GTGGCGCACG	GGTGAGTAAT	G?ATAG?T?A
Cosenso_delta	GTGC?TAACA	CATGCAAGTC	GAACG?GAAA	G?????--T	TCG----???	??GAGTA?A	GTGGCGCACG	GGTGAGTAAC	?CGTGGATAA

	100	110	120	130	140	150	160	170	180		
GRUPO_I	CGTGCCCTAG	AGTGGGGGAT	AACTAGTCGA	AAGATTAGCT	AATACCGCAT	ACG	ATC----	----TACG--	----GAT	GA AA	GTGGGGGA
GRUPO_II	CGTGCCCTAG	AGTGGGGGAT	AACTAGTCGA	AAGATTAGCT	AATACCGCAT	ACG	ATC----	----TACG--	----GAT	GA AA	GTGGGGGA
GRUPO_III	CGTGCCCTAG	AGTGGGGGAT	AACTAGTCGA	AAGATTAGCT	AATACCGCAT	ACG	ATC----	----TAAG--	----GAT	GA AA	GTGGGGGA
GRUPO_IV	CGTATCCTGG	AGTGGGGGAT	AGCCCGGCGA	AAGCCGGATT	AATACCGCAT	ACG	CTC----	----TGAG--	----GAG	GA AA	GCGGGGGA
GRUPO_V	CGTGTCCGT	AGTGGGGGAT	AGCCCGGCGA	AAGCCGGATT	AATACCGCAT	ACG	ATC----	----TACG--	----GAT	GA AA	GCGGGGGA
GRUPO_VII	CATGTCCAGT	AGTGGGGGAT	AGCCCGGCGA	AAGCCGGATT	AATACCGCAT	ACG	ATC----	----TGTG--	----GAT	GA AA	GCGGGGGA
GRUPO_VIII	CATGTCCAGT	AGTGGGGGAT	AGCCCGGCGA	AAGCCGGATT	AATACCGCAT	ACG	ATC----	----TGTG--	----GAT	GA AA	GCGGGGGA
GRUPO_IX	CATGTCCAGT	AGTGGGGGAT	AGCCCGGCGA	AAGCCGGATT	AATACCGCAT	ACG	CTC----	----TGYG--	----GAG	GA AA	GCGGGGGA
GRUPO_XII	CATGTCCAGT	AGTGGGGGAT	AGCCCGGCGA	AAGCCGGATT	AATACCGCAT	ACG	ATC----	----TTTG--	----GAT	GA AA	GCGGGGGA
GRUPO_VI	CGTGCCCGAG	AGTGGGGGAT	AACGGAGCGA	AAGCTTTGCT	AATACCGCAT	ACG	ATC----	----TACG--	----GAT	GA AA	GCAAGGGGA
GRUPO_X	CCTGCCTTC	GGTTCGGAAT	AACGTTTGG	AACGAACGCT	AATACCGCAT	ACG	CCC----	----TATG--	----GGG	GA AA	G-----
GRUPO_XI	CGTACCTTTT	GCTACGGAAT	AACTCAGGGA	AACTTGTGCT	AATACCGTAT	GTG	CCC----	----TTCG--	----GGG	GA AA	GA-----
Consenso_alfa	??T?CCCTT?	?GT?CGGAAT	AAC????GGA	AAC????GCT	AATACCG?AT	A?G	CCC----	----TT?G--	----GGG	GA AA	GA-----
Consenso_beta	CGTGCCCTGT	AGTGGGGGAT	AAC????CGA	AAG????GCT	AATACCGCAT	ACG	??C----	----T?AG--	----G??	GA AA	GCGGGGGA
Consenso_gama	TCTGCCT??T	?G?GGGGGAT	AAC????GGA	AACG????GCT	AATACCGCAT	A??	??C----	----T?AG--	----G??	?A AA	G?G?GGGGA
Consenso_epsilon	T?TGCCCTTT	AGT??GGGAT	A?C?A?TGGA	AACG?T????	AATAC??AT	ACT	CC?????	?---?????	---????G?	GA AA	G?-----
Cosenso_delta	TCCT?CC????	?G???GGGAT	AAC????CGA	AAGG????GCT	AATACCGGAT	A??	?TC?????	?????????	?????G??	?A AA	GG?GGCCT

	190	200	210	220	230	240	250	260	270
GRUPO_I	TC-GCAA-GA CCTCAT	GCTC CTGGAGCGGC	CGATATCTGA	TTAGCTAGTT	GGTGGGGTAA	AAGCCTACCA	AGGCGACGAT	CAGTAGCTGG	
GRUPO_II	TC-GCAA-GA CCTCAT	GCTC CTGGAGCGGC	CGATATCTGA	TTAGCTAGTT	GGTGGGGTAA	AAGCTACCA	AGGCGACGAT	CAGTAGCTGG	
GRUPO_III	TC-GCAA-GA CCTCAT	GCTC CTGGAGCGGC	CGATATCTGA	TTAGCTAGTT	GGTGGGGTAA	AAGCCTACCA	AGGCGACGAT	CAGTAGCTGG	
GRUPO_IV	CC-TTCG-GG CCTCGC	GCTC AAGGGGCGGC	CGATGGCGGA	TTAGGTAGTT	GGTGGGGTAA	AGGCCACCA	AGCCGACGAT	CCGTAGCTGG	
GRUPO_V	TC-TTCG-GA CCTCGC	GCTA TAGGGGCGGC	CGATGGCGGA	TTAGCTAGTT	GGTGAGGTAA	AGGCTACCA	AGGCGACGAT	CCGTAGCTGG	
GRUPO_VII	CC-TTCG-GG CCTCGC	GCTA TTGGGGTGGC	CGATGGCGGA	TTAGCTAGAT	GGTGGGGTAA	AGGCCACCA	TGGCGACGAT	CCGTAGCTGG	
GRUPO_VIII	TC-GCAA-GA CCTCGC	GCTA TTGGGGTGGC	CGATGGCGGA	TTAGCTAGTT	GGTGGGGTAA	AGGCCACCA	AGGCGACGAT	CCGTAGCTGG	
GRUPO_IX	TC-GAAA-GA CCTCGC	GCTA TTGGGGTGGC	CGATGGCGGA	TTAGCTAGAT	GGTGGGGTAA	AGGCCACCA	TGGCGACGAT	CCGTAGCTGG	
GRUPO_XII	CC-TTCG-GG CCTCGC	GCTA TTGGGGTGGC	CGATGGCGGA	TTAGCTAGTT	GGTGGGGTAA	AGGCCACCA	AGGCGACGAT	CCGTAGCTGG	
GRUPO_VI	CC-GCAA-GG CCTTGC	GCTC ATGGAGCGGC	CGATGGCAGA	TTAGGTAGTT	GGTGGGATAA	AAGCTTACCA	AGCCGACGAT	CTGTAGCTGG	
GRUPO_X	---TTTA--- ----C	GCCG AGAGAGGGGC	CCGCGTCGGA	TTAGGTAGTT	GGTGAGGTAA	TGGCTACCA	AGCCTTCGAT	CCGTAGCTGG	
GRUPO_XI	---TTTA--- ----TC	GGCA AAGGATCGGC	CCGCGTTGGA	TTAGCTAGTT	GGTGAGGTAA	AGGCTACCA	AGGCGACGAT	CCATAGCTGG	
Consenso_alfa	---TTTA--- ----TC	GC?? A?GGAT??GC	CCGCGT??GA	TTAGCTAGTT	GGTGGGGTAA	?GGC??ACCA	AGGCGACGAT	C??TAGCTGG	
Consenso_beta	?C-G?AA-G? CCTCGC	GCTA TTGGAGCGGC	CGATGTC?GA	TTAGCTAGTT	GGTGGGGTAA	AGGC?TACCA	AGGCGACGAT	C?GTAGCTGG	
Consenso_gama	CC-TTCG-GG CCT??C	GC?A T??GATGAC	C?A?GT?GGA	TTAGCTAGTT	GGTGAGGTAA	?GGCTACCA	AGGCGACGAT	CCGTAGCTGG	
Consenso_epsilon	---TTT?--- ----TC	GCTA A?GGAT?AG?	CTAT?TC?TA	TCAGCT?T	GGTGAGGTAA	TGGCT?ACCA	AGGCTATGAC	G?GTA?C?GG	
Cosenso_delta	CT??T??AA GCT??C	G?? ?GGATGAGT	CCGCGT?CCA	TTAGCT?GTT	GGTGGGGTAA	?GGCTACCA	AGGC?ACGAT	GG?TAGCTGG	

	280	290	300	310	
GRUPO_I	TCTGAGAGGA	CGACCAGCCA	CACTGGGACT	GAGACACGGC	CCAG
GRUPO_II	TCTGAGAGGA	CGACCAGCCA	CACTGGGACT	GAGACACGGC	CCAG
GRUPO_III	TCTGAGAGGA	CGACCAGCCA	CACTGGGACT	GAGACACGGC	CCAG
GRUPO_IV	TCTGAGAGGA	CGACCAGCCA	CACTGGGACT	GAGACACGGC	CCAG
GRUPO_V	TCTGAGAGGA	CGACCAGCCA	CACTGGGACT	GAGACACGGC	CCAG
GRUPO_VII	TCTGAGAGGA	CGACCAGCCA	CACTGGGACT	GAGACACGGC	CCAG
GRUPO_VIII	TCTGAGAGGA	CGACCAGCCA	CACTGGGACT	GAGACACGGC	CCAG
GRUPO_IX	TCTGAGAGGA	CGACCAGCCA	CACTGGGACT	GAGACACGGC	CCAG
GRUPO_XII	TCTGAGAGGA	CGACCAGCCA	CACTGGGACT	GAGACACGGC	CCAG
GRUPO_VI	TCTGAGAGGA	CGACCAGCCA	CACTGGGACT	GAGACACGGC	CCAG
GRUPO_X	TCTGAGAGGA	TGATCAGCCA	CACTGGGACT	GAGACACGGC	CCAG
GRUPO_XI	TCTGAGAGGA	TGATCAGCCA	CACTGGGACT	GAGAC-CGGC	CCAG
Consenso_alfa	TCTGAGAGGA	TGATCAGCCA	CACTGGGACT	GAGACACGGC	CCAG
Consenso_beta	TCTGAGAGGA	CGACCAGCCA	CACTGGGACT	GAGACACGGC	CCAG
Consenso_gama	TCTGAGAGGA	TGATCAGCCA	CACTGGAACT	GAGACACGGT	CCAG
Consenso_epsilon	?CTGAGAGG?	TGATC?G?CA	CACTGGAACT	GAGACACGGT	CCAG
Cosenso_delta	TCTGAGAGGA	TGATCAGCCA	CACTGGAACT	GA?ACACGGT	CCAG

4.3 ANÁLISE COMPARATIVA ENTRE OS MÉTODOS

No presente trabalho foram mostradas evidências da proximidade evolutiva entre isolados diazotróficos de vários cultivares de *Musa* spp. e *Ananas comosus* com espécies da subdivisão alfa e, principalmente, beta das Proteobactérias. As análises empregadas permitiram identificar vários dos isolados diazotóxicos como espécies de *Herbaspirillum*, *Burkholderia*, *Azospirillum* e outros grupos taxonômicos, não comumente encontrados em associação a plantas.

Foram também identificados vários diazotróficos com potencial para constituir novas taxa. Isto foi também observado na caracterização fisiológica e pela análise de hibridização (WEBER *et al.*, 1998) e foi confirmado aqui pelas análises de ARDRA e seqüenciamento.

A grande proximidade genética entre os grupos obtidos, principalmente entre as estirpes de *Herbaspirillum seropedicae* e *Herbaspirillum rubrisubalbicans*, observada nos experimentos de RFLP da região intergênia 16S-23S rDNA, ARDRA e seqüenciamento do 16S rDNA, mostraram que há um baixo nível de diversidade genética entre estes organismos. Análises feitas com estirpes de *Herbaspirillum frisingense* (KIRCHHOF *et al.*, 2001) e *Gluconacetobacter diazotrophicus* (CABALLERO-MELLADO & MARTINEZ-ROMERO, 1994) chegaram a esta mesma conclusão. MC ARTHUR *et al.* (1988) sugeriram que esta pequena variabilidade genética possa estar correlacionada a um habitat estável ao longo do tempo para estes organismos.

Isolados agrupados com a estirpe M4 de *H. rubrisubalbicans* por análise de ARDRA mostraram-se ambíguos, apresentando 100% de identidade na região Y1-Y2 com a estirpe Z67 de *H. seropedicae*. Os genótipos de *Herbaspirillum* foram particularmente difíceis de resolverem devido a sua grande proximidade. Eles apresentaram uma única banda, no padrão de restrição obtido com *AluI* no método de

ARDRA (Figura 9) e somente duas bases diferentes na seqüência Y1-Y2 (Figura 15). WEBER *et al.* (1998) também encontraram ambigüidade nas análises fisiológicas e de hibridização com os mesmos isolados.

Os Grupos definidos para as seqüências Y1-Y2 concordaram bem com os Grupos definidos para ARDRA, não mostrando nenhum polimorfismo aparente dentro dos Grupos. Entretanto, dois desacordos foram observados com os Grupos de *Herbaspirillum* definidos por ARDRA. O Grupo ARDRA 1 continha os Grupos de seqüência I e II e o Grupo ARDRA 2 os Grupos de seqüência I e III. É interessante notar que embora as estirpes referência de *H. seropedicae* e *H. rubrisubalbicans* mostraram diferenças nos padrões de ARDRA (1 e 2) e de seqüenciamento (I e III), onze isolados compartilharam o mesmo padrão ARDRA de *H. rubrisubalbicans*, enquanto mostraram 100% de identidade de seqüência com *H. seropedicae* na região Y1-Y2 do 16S rDNA (Tabelas 7 e 8). Segundo WEBER (1998), estes onze isolados hibridizaram com uma sonda BET42a para o 23S rRNA da subdivisão beta das Proteobactérias, mas não com uma sonda para o 23S rDNA de *H. seropedicae*. Além do mais, de acordo com o mesmo autor, enquanto alguns isolados foram capazes de crescer em meio JNFb contendo N-acetilglucosamina como única fonte de carbono (BA12, BA14, BA15, BA17, BA134, BA149 e BA161), outros cresceram usando *m*-eritritol (BA10, BA11 e BA16). A capacidade de utilização destas duas fontes de carbono diferencia as espécies de *H. seropedicae*, que utiliza N-acetilglucosamina e *H. rubrisubalbicans* que utiliza *m*-eritritol (BALDANI *et al.*, 1996). Aparentemente estes isolados formam um grupo diferente dentro do gênero *Herbaspirillum*.

O seqüenciamento da região Y1-Y3 do 16S rDNA dos isolados que compõem os Grupos ARDRA 1 e 2 permitiu confirmar a diferença existente entre estes grupos. Entre os organismos que não podem ser diferenciados pela técnica de ARDRA ocorre uma similaridade média de 96,1% na seqüência do 16S rDNA.

Através da análise de ARDRA, estes grupos diferiram somente na banda de mais alto peso molecular observada no padrão de restrição com *AluI*. As seqüências

dos isolados BA153 e X8 (Grupo ARDRA 1 e seqüência Y1-Y2 I) e de dois representantes do Grupo ARDRA 2 e seqüência Y1-Y2 I, BA15 e BA17, mostram que a banda de maior peso molecular, com 427 pb, em *H. seropedicae* e nos isolados BA153 e X8, foi dividida em duas outras bandas nos isolados BA15 e BA17 (Figura 26A): uma delas contendo 398 pb, correspondendo à banda diferencial do grupo ARDRA 2 e uma segunda banda contendo 29 pb, não resolvida nas condições eletroforéticas usadas. Esta diferença ocorre devido à presença de um sítio de restrição nas seqüências dos isolados BA15 e BA17, não existente nos isolados BA153 e X8 (figura 26B). Dois outros fragmentos de restrição diferenciam estes isolados. Entretanto, devido ao baixo peso molecular, não foram resolvidos pela eletroforese em gel de agarose. Pela figura 26A é possível observar ainda que o perfil de restrição do isolado BA153 é idêntico ao de *H. seropedicae* (as duas seqüências diferem em um único nucleotídeo na região Y1-Y3, contendo 1.429 pb) e que, estes diferem do perfil do isolado X8 nas bandas abaixo de 50 pb. A diferença ocorre devido a uma mutação pontual observada na posição 245 do isolado X8, que contém uma base T, em relação a *H. seropedicae*, contendo uma base C. Observa-se desta forma mais um sítio de restrição para a endonuclease *AluI*, que reconhece o palíndromo AGCT (corresponde ao sítio marcado somente na seqüência do isolado X8 no esquema na figura 26B).

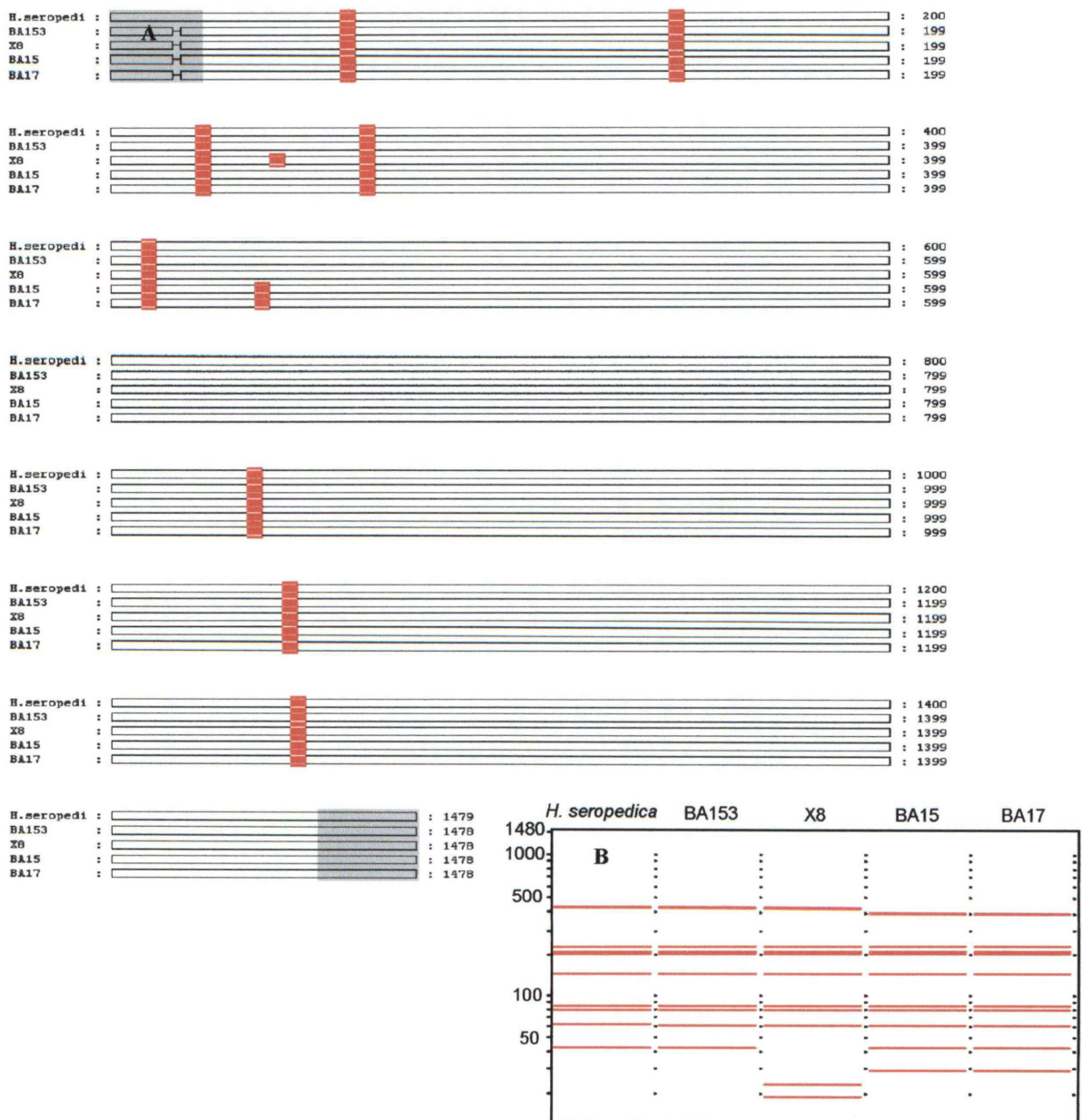
As seqüências da região Y1-Y3 de todos os isolados colocados no Grupo ARDRA 2, mostraram que este grupo está, na verdade, mais próximo de *Herbaspirillum rubrisubalbicans* do que de *Herbaspirillum seropedicae* (Figura 20), como determinado pelo método de ARDRA.

Para os demais grupos não houve contradições para os resultados obtidos por ARDRA ou seqüenciamento. Entretanto, o seqüenciamento permitiu estabelecer as relações filogenéticas entre os isolados que não puderam ser identificados pelo método de ARDRA, mostrando que grande parte dos genótipos isolados de abacaxizeiro (AB48, AB98, AB119 e AB120) aparentemente pertencem ao gênero *Burkholderia*. Um único isolado de bananeira, BA126, parece também pertencer a este gênero.

Todos os genótipos analisados no presente trabalho hibridizaram com a sonda BET42a, específica para o 23S rRNA da subdivisão beta das Proteobactérias e, com a exceção do Grupo ARDRA 5, também foram capazes de hibridizar com a sonda BC16, específica para o 16S rDNA de *Burkholderia cepacia* (WEBER, 1998).

O isolado diazotróficos incluídos no Grupo ARDRA 5 e Grupo de seqüência Y1-Y2 VI mostraram-se geneticamente mais distantes dos demais Grupos incluídos na subdivisão beta das Proteobactérias. A análise filogenética da seqüência Y1-Y3 do isolado BA128, representante deste Grupo, ilustra este fato (Figura 22), mostrando a proximidade deste organismo com espécies de Comamonadaceae. Esta-família contém bactérias saprofíticas e fitopatogênicas com perfis fisiológicos diversos e tem sido isolada de uma grande variedade de nichos ecológicos tais como solo, água, ambiente natural e industrial, amostras clínica e material de plantas infectadas. Devido a sua flexibilidade fisiológica, possuem um importante papel na degradação de xenobióticos e componentes recalcitrantes, como resíduos derivados de óleo e poluentes como nitrato (VANDAMME *et al.*, 1996). Entretanto, assim como o gênero *Burkholderia* e *Herbaspirillum*, as Comamonadaceae possuem seu núcleo originário das *Pseudomonas*, neste caso, do grupo de homologia de rRNA III (PALLERONI *et al.*, 1973).

FIGURA 26 - MAPEAMENTO DOS SÍTIOS DE RESTRIÇÃO PARA *AluI* NA SEQUÊNCIA DA REGIÃO Y1-Y3 DE *H. seropedicae* E DOS ISOLADOS BA15, BA17, BA153 E X8. **A**, POSICIONAMENTO DOS SÍTIOS DE RESTRIÇÃO PARA *AluI* (EM VERMELHO) NAS SEQÜÊNCIAS Y1-Y3 DO 16S rDNA (INCLUINDO A REGIÃO DOS INICIADORES) DE *Herbaspirillum seropedicae* (H. SEROPEDI, NÚMERO DE ACESSO NO GENBANK Y10146) E ISOLADOS DOS GRUPOS ARDRA 1 (BA153 E X8) E 2 (BA15 E BA17) (AS REGIÕES DOS INICIADORES SÃO MOSTRADAS EM CINZA); **B**, SIMULAÇÃO EM COMPUTADOR DO PERFIL ELETROFORÉTICO DA REGIÃO Y1-Y3 DO 16S rDNA DOS MESMOS ORGANISMOS, APÓS DIGESTÃO COM A ENDONUCLEASE DE RESTRIÇÃO *AluI*.



5 CONCLUSÕES

A análise dos padrões eletroforéticos obtidos para os fragmentos de restrição da região Y1-Y3 do 16S rDNA amplificado (ARDRA), permitiu separar os 38 isolados diazotróficos analisados em 12 grupos, cada um deles formado por isolados compartilhando padrões eletroforéticos idênticos. Dezesesseis destes isolados foram identificados nos grupos formados por ARDRA como *Herbaspirillum seropedicae*, *Herbaspirillum rubrisubalbicans*, “*Burkholderia brasilensis*” e “*Burkholderia tropicalis*”. Os vinte e dois isolados diazotróficos restantes foram distribuídos em oito grupos não identificados.

A análise feita a partir do seqüenciamento da região Y1-Y2 do 16S rDNA dos isolados diazotróficos confirmou a maioria dos grupos obtidos, anteriormente, por ARDRA. Uma única exceção ocorreu com os isolados diazotróficos identificados como *Herbaspirillum rubrisubalbicans* (Grupo ARDRA 2), que, na análise de seqüenciamento, apresentaram 100% de identidade com a seqüência da região Y1-Y2 de *Herbaspirillum seropedicae*.

A pesquisa de homologia e análise filogenética das seqüências da região Y1-Y2 e Y1-Y3 do 16S rDNA mostrou que os isolados estão distribuídos entre a subdivisões alfa e beta das Proteobactérias, mais especificamente, relacionados a *Azospirillum lipoferum* e *Ochrobactrum* sp., da subdivisão alfa e a *Burkholderia* sp., *Herbaspirillum* sp. e à família Comamonadaceae, da subdivisão beta.

6 REFERÊNCIAS

- ACINAS, S. G.; ANTON, J.; RODRIGUEZ-VALERA, F. Diversity of free-living and attached bacteria in offshore western Mediterranean waters as depicted by analysis of genes encoding 16S rRNA. **Appl. Environ. Microbiol.**, Washington, v. 65, p. 514-522, 1999.
- ALTSCHUL, S. F.; MADDEN, T. L.; SCHÄFFER, A. A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Res.**, Oxford, v. 25, p. 3389-3402, 1997.
- ALTSCHUL, S. F. Phylogenetic estimation. In: BISHOP, M. J.; RAWLINGS, C. J. (eds.). **DNA and protein sequence analysis - a practical approach**. New York, IRL press, 1997. p. 279-312.
- AMORIM, D. S. **Elementos básicos de sistemática filogenética**. Segunda ed. Ribeirão Preto: Holos Editora, 1997.
- ANZAI, Y.; KIM, H.; PARK, J.-U.; WAKABAYASHI, H.; OYAIZU, H. Phylogenetic affiliation of the pseudomonads based on 16S rRNA sequence. **Int. J. Syst. Evol. Microbiol.**, Reading, v. 50, p. 1563-1589, 2000.
- APP, A. A.; WATANABE, I.; ALEXANDER, M.; VENTURA, W.; DAEZ, C.; SANTIAGO, T.; DEDATTA, S. K. Non-symbiotic nitrogen-fixation associated with the rice plant in flooded soils. **Soil Sci.**, Philadelphia, v. 130, n. 5, p. 283-289, 1980.
- ATCHLEY, W. R.; FITCH, W. M.; BRONNER-FRASER, M. Molecular evolution of the MyoD family of transcription factors. **Proc. Natl. Acad. Sci. USA**, Washington, v. 91, p. 11522-11526, 1994.
- BALDANI, J. I.; OLIVEIRA, A. L. M.; GUIMARÃES, S. L.; BVALDANI, V. L. D.; REIS Jr., F. B.; SILVA, L. G.; REIS, V. M.; TEIXEIRA, K. R. S.; DÖBEREINER, J. Biological nitrogen fixation (BNF) in non-leguminous plants: the role of endophytic diazotrophs. In: NITROGEN FIXATION: FROM MOLECULES TO CROP PRODUCTIVITY, 12th, 1999, Foz do Iguaçu. **Proceedings of the International Congress on Nitrogen Fixation**. Dordrecht: Klumer Academic Publishers, 2000. p. 397-400.
- BALDANI, J. I.; CARUSO, L.; BALDANI, V. L. D.; GOI, S. R.; DÖBEREINER, J. Recent advances in BNF with non-legume plants. **Soil Biol. Biochem.**, Bethesda, v. 29, n. 5/6, p. 911-922, 1997.
- BALDANI, J. I.; POT, B.; KIRCHHOF, G.; FALSEN, E.; BALDANI, V. L. D.; OLIVARES, F. L.; HOSTE, B.; KERSTERS, K.; HARTMANN, A.; GILLIS, M.; DÖBEREINER, J. Emended description of *Herbaspirillum*; inclusion of [*Pseudomonas*] *rubrisubalbicans*, a mild plant pathogen, as *Herbaspirillum rubrisubalbicans* comb. nov.; and classification of a group of clinical isolates (EF Group 1) as *Herbaspirillum* species 3. **Int. J. Syst. Bacteriol.**, Washington, v. 46, n. 3, p. 802-810, 1996.
- BALDANI, J. I.; BALDANI, V. L. D.; SELDIN, L.; DÖBEREINER, J. Characterization of *Herbaspirillum seropedicae* gen., nov. sp. nov., a root-associated nitrogen-fixing bacterium. **Int. J. Syst. Bacteriol.**, Washington, v. 36, p. 86-93, 1986.
- BALDANI, V. L. D. **Efeito da inoculação de *Herbaspirillum* spp. no processo de colonização e infecção de plantas de arroz e, ocorrência e caracterização parcial de uma**

nova bactéria diazotrófica. Itaguaí, 1996. 238 p. Tese (Doutorado em Ciência do Solo) - Instituto de Agronomia, Universidade Federal Rural do Rio de Janeiro.

BASHFORD, D.; CHOTHIA, C.; LESK, A. M. Determinants of a protein fold – unique features of the globin amino-acid-sequences. **J. Mol. Biol.**, London, v. 196, n. 1, p. 199-216, 1987.

BAZZICALUPO, M.; OKON, Y. Associative and endophytic symbiosis. In: NITROGEN FIXATION: FROM MOLECULES TO CROP PRODUCTIVITY, 12th, 1999, Foz do Iguaçu. **Proceedings of the International Congress on Nitrogen Fixation**. Dordrecht: Klumer Academic Publishers, 2000. p. 409-410.

BEN DEKHIL, S.; CAHILL, M.; STACKEBRANDT, E.; SLY, L. I. Transfer of *Conglomeromonas largomobilis* subsp. *largomobilis* to the genus *Azospirillum* as *Azospirillum largomobile* comb. nov., and elevation of *Conglomeromonas largomobilis* subsp. *parooensis* to the new type species of *Conglomeromonas*, *Conglomeromonas parooensis* sp. Nov. **Syst. Appl. Microbiol.**, v. 20, p. 72-77, 1997.

BORNEMAN, J.; TRIPLETT, E. W. Molecular microbial diversity in soils from Eastern Amazonia: evidence for unusual microorganisms and population shifts associated with deforestation. **Appl. Environ. Microbiol.**, Washington, v. 63, p. 2647-2653, 1997.

BÖTTGER, E. C. Rapid determination of bacterial ribosomal RNA sequences by direct sequencing of enzymatically amplified DNA. **FEMS Microbiol. Lett.**, Amsterdam, v. 65, p. 171-176, 1989.

BROSIUS, J.; DULL, T. J.; SLEETER, D. D.; NOLLER, H., F. Gene organization and primary structure of a ribosomal RNA operon from *Escherichia coli*. **J. Mol. Biol.**, London, v. 148, p. 107-127, 1981.

BROWN, J. W. Phylogenetic comparative analysis on Macintosh computers. **Comput. Appl. Biosci.**, Oxford, v. 7, p. 391-393, 1991.

BUSSE, H. J.; LUBITZ, W. Classification and identification of bacteria: current approaches to an old problem. Overview of methods used in bacterial systematic. **J. Biotechnol.**, Amsterdam, v. 47, p.3-38, 1996.

CABALLERO-MELLADO, J.; MARTINEZ-ROMERO, E. Limited genetic diversity in the endophytic sugarcane bacterium *Acetobacter diazotrophicus*. **Appl. Environ. Microbiol.**, Washington, v. 60, n. 5, p. 1532-1537, 1994.

CAVALLI-SFORZA, L. L.; EDWARDS, A. W. F. Analysis of human evolution. In: GENETICS TODAY, 11th, 1964. **Proc. Int. Congr. Genet.** The Hague, The Netherlands: Pergamon, 1964. p. 923-933.

CHISNELL, J. R.; PREMAKUMAR, R.; BISHOP, P. E. Purification of a 2nd alternative nitrogenase from a nifHDK deletion strain of *Azotobacter vinelandii*. **J. Bacteriol.**, Washington, v. 170, n.1, p. 27-33, 1988.

COSTACURTA, A.; VANDERLEYDEN, J. Synthesis of phytohormones by plant-associated bacteria. **Crit. Rev. Microbiol.**, Boca Raton, v. 21, n. 1, p. 1-18, 1995.

CRUZ, L. M.; SOUZA, E. M.; WEBER, O. B.; BALDANI, J. I.; DÖBEREINER, J.; PEDROSA, F. O. 16S ribosomal DNA characterization of nitrogen-fixing bacteria isolated from banana (*Musa* spp.) and pineapple (*Ananas comosus* (L.) Merril). **Appl. Environ. Microbiol.**, Washington, v. 67, n. 5, p. 2375-2379, 2001.

DE RIJK, P.; WUYTS, J.; VAN DE PEER, Y.; WINKELMANS, T.; DE WACHTER, R. The European small subunit ribosomal RNA database. **Nucleic Acids Res.**, Oxford, v. 28, n. 1, p. 175-176, 2000.

DE RIJK, P.; DE WACHTER, R. **RnaViz 2.0**. Antwerp, 1997 - 2001. 1 arquivo (1,1 Mb); Disponível em: <http://rrna.uia.ac.be/rnaviz/> A program to create RNA secondary structure drawings.

DE SMEDT, J.; BAUWENS, M.; TYTGAT, R.; DE LEY, J. Intra- and intergeneric similarities of ribosomal ribonucleic acid cistrons of free-living, nitrogen-fixing bacteria. **Int. J. Syst. Bacteriol.**, Washington, v. 30, p. 106-122, 1980.

DÖBEREINER, J. Biological nitrogen fixation in the tropics: social and economic contributions. In: INTERNATIONAL SYMPOSIUM ON SUSTAINABLE AGRICULTURE FOR THE TROPICS - THE ROLE OF BIOLOGICAL NITROGEN FIXATION, 1995, Angra dos Reis. **Programme and Abstracts**. Viçosa: JARD Produções Gráficas Ltda. 1995. p. 3.

DÖBEREINER, J.; BALDANI, V. L. D.; BALDANI, J. I. **Como isolar e identificar bactérias diazotróficas de plantas não-leguminosas**. Embrapa: Embrapa - SPI, 1995.

DÖBEREINER, J. Recent changes in concepts of plant bacteria interactions: endophytic N₂ fixing bacteria. **Journal of the Brazilian Association for the Advancement of Science**, v. 44, n. 5, p. 310-313, 1992.

DÖBEREINER, J.; REIS, V. M.; LAZARINI, A. C. New N₂-fixing bacteria in association with cereals and sugar-cane. In: NITROGEN FIXATION: HUNDRED YEARS AFTER, Stuttgart. **Proceeding**. Gustav Fischer, 1988 p. 717-722.

DÖBEREINER, J.; RUSCHEL, A. P. Uma nova espécie de Beijerinckia. **Revista de Biologia**, Rio de Janeiro, v. 1, p. 261-272, 1958.

DEUTSCHE SAMMLUNG VON MIKROORGANISMEN UND ZELLKULTUREN GMBH (DSMZ). **Bacterial nomenclature up-to-date**. Disponível em: <<http://www.dsmz.de/bactnom/bactname.htm>> Acesso em: 02 jun. 2001.

EMBLEY, T. M. The linear PCR reaction: a simple and robust method for sequencing amplified rRNA genes. **Lett. Appl. Microbiol.**, Oxford, v. 13, p. 171-174, 1991.

FARRIS, J. S. A successive approximations approach to character weighting. **Syst. Zool.**, Washington, v. 18, p. 374-385, 1969.

FENG, D.-F.; DOOLITTLE, R. F. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. **J. Mol. Evol.**, New York, v.25, n. 4, p. 351-360, 1987.

FELSENSTEIN, J. **Phylip 3.6 (alfa)**. Seattle, 1989 - 2001. 2 arquivos (2 Mb); Disponível em: < <http://evolution.genetics.washington.edu/phylip.html>>. Phylogeny Inference Package.

FELSENSTEIN, J. Phylogenies from molecular sequences: inference and reliability. **Annu. Rev. Genet.**, Palo Alto, v. 22, p. 521-565, 1988.

FERREIRA, A. C.; COZZOLINO, K.; CARVALHO, A. R. V.; DÖBEREINER, J. Isolation and characterization of diazotrophic bacteria in oil palm trees. In: INTERNATIONAL SYMPOSIUM ON SUSTAINABLE AGRICULTURE FOR THE TROPICS - THE ROLE OF BIOLOGICAL NITROGEN FIXATION, 1995, Angra dos Reis. **Programme and abstracts**. Viçosa: JARD Produções Gráficas, 1995. p. 210.

- FERREIRA, M. C. B.; FERNANDES, M. S.; DÖBEREINER, J. Role of *Azospirillum brasilense* nitrate reductase in nitrate assimilation by wheat plants. **Biology and Fertility of Soils**, New York, v. 4, p. 47-53, 1987.
- FISHER, M.; TRIPLETT, E. W. Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. **Appl. Environ. Microbiol.**, Washington, v. 65, n. 10, p. 4630-4636, 1999.
- FOX, G.E.; STACKEBRANDT, E.; HESPELL, R.B.; GIBSON, J.; MANILOFF, J.; DYER, T.A.; WOLFE, R.S.; BALCH, W.E.; TANNER, R.S.; MAGRUN, L.J.; ZABLEN, L.B.; BLAKEMORE, R.; GUPTA, R.; BONEN, L.; LEWIS, B.J.; STAHL, D.A.; LUEHRSEN, K.R.; CHEN, K.N.; WOESE, C.R. The phylogeny of prokaryotes. **Science**, Washington, v. 209, p. 457-463, 1980.
- FOX, G. E.; PECKMAN, K. J.; WOESE, C. R. Comparative cataloging of 16S ribosomal ribonucleic acid: molecular approach to prokaryotic systematics. **Int. J. Syst. Bacteriol.**, Washington, v. 27, p. 44-57, 1977.
- GILLIS, M.; VAN, T. V.; BARDIN, R.; GOOR, M.; HEBBAR, P.; WILLEMS, A.; SEGERS, P.; KERSTERS, K.; HEULIN, T.; FERNANDEZ, M. P. Polyphasic taxonomy in the genus *Burkholderia* leading to a new description of the genus and proposition of *Burkholderia vietnamiensis* sp. nov. for N₂-fixing isolates from rice in Vietnam. **Int. J. Syst. Bacteriol.**, Washington, v. 45, p. 274-289, 1995.
- GILLIS, M.; KERSTERS, K.; HOSTE, B.; JANSSENS, D.; KROPPESTEDT, R. M.; STEPHAN, M. P.; TEIXEIRA, K. R. S.; DÖBEREINER, J.; DE LEY, J. *Acetobacter diazotrophicus* sp. nov., a nitrogen-fixing acetic acid bacterium associated with sugarcane. **Int. J. Syst. Bacteriol.**, Washington, v. 39, p. 361-364, 1989.
- GOLDMAN, N. Phylogenetic estimation. In: BISHOP, M. J.; RAWLINGS, C. J. (eds.). **DNA and protein sequence analysis - a practical approach**. New York: IRL press, 1997. p. 279-312.
- GRIFONI, A.; BAZZICALUPO, M.; DI SERIO, C.; FANCELLI, S.; FANI, R. Identification of *Azospirillum* strains by restriction fragment length polymorphism of the 16S rDNA and of the histidine operon. **FEMS Microbiol. Lett.**, Amsterdam, v. 127, p. 85-91, 1995.
- GÜRTLER, V.; MAYALL, B. C. Genomic approaches to typing, taxonomy and evolution of bacterial isolates. **Int. J. Syst. Evol. Microbiol.**, Reading, v. 51, p. 3-16, 2001.
- GUTELL, R. R. Comparative anatomy of 16S-like ribosomal RNA. **Prog. Nucl. Acids Res. Mol. Biol.**, v. 32, p. 155-216, 1985.
- HALL, T.A. **BioEdit 4.8**. Raleigh, 1997 – 2001. 1 arquivo (11,5 Mb); Disponível em: <<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>> BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.
- HALLMANN, J.; QUADT-HALLMANN, A.; MAHAFFEE, W. F.; KLOEPPER, J. W. Bacterial endophytes in agricultural crops. **Can. J. Microbiol.**, Ottawa, v. 43, p. 895-914, 1997.
- HART, M. H. Was the pre-biotic atmosphere of the earth heavily reducing? **Orig. Life Evol. Biosph.**, Dordrecht, v. 9, n. 4, p. 261-275, 1979.
- HARTMANN, A.; BALDANI, J. I.; KIRCHHOF, G.; ASSMUS, B.; HUTZLER, P.; SPRINGER, N.; LUDWING, W.; BALDANI, V. L. D.; DÖBEREINER, J. Taxonomic and

- ecologic studies of diazotrophic rhizosphere bacteria using phylogenetic probes. In: FENDRIK, I.; DEL GALLO, M.; VANDERLEYDEN, J.; ZAMAROCZY, M. DE (eds.). **Azospirillum VI and related microorganisms**. Berlin: Springer Verlag, 1995.
- HASEGAWA, M.; KISHINO, H.; TANO, T. Dating the human-ape splitting by a molecular clock of mitochondrial DNA. **J. Mol. Evol.**, New York, v. 22, p. 160-174, 1985.
- HAUKKA, K.; LINDSTRÖM, K.; YOUNG, J. P. W. Diversity of partial 16S rRNA sequences among and within strains of African rhizobia isolated from *Acacia* and *Prosopis*. **System. Appl. Microbiol.**, Jena, v. 19, p. 352-359, 1996.
- HENNECKE, H.; KALUZA, K.; THÖNY, B.; FUHRMANN, M.; LUDWIG, W.; STACKEBRANDT, E. Concurrent evolution of nitrogenase genes and 16S rRNA in *Rhizobium* species and other nitrogen fixing bacteria. **Arch. Microbiol.**, v. 142, p. 342-348, 1985.
- HERSHKOVITZ, M. A.; LEIPE, D. D. Phylogenetic analysis. In: BAXEVANIS, A. D.; OUELLETTE, B. F. F. (eds.). **Bioinformatics - a practical guide to the analysis of genes and proteins**. New York: John Wiley & Sons, 1998. p. 189-230.
- HIGGINS, D. G.; BLEASBY, A. J.; FUCHS, R. CLUSTAL V: improved software for multiple sequence alignment. **Comput. Appl. Biosci.**, Oxford, v. 8, p. 189-191, 1991.
- HIGGINS, D. G.; SHARP, P. M. Fast and sensitive multiple sequence alignments on a microcomputer. **Comput. Appl. Biosci.**, Oxford, v. 5, n. 2, p. 151-153, 1989.
- HIGGINS, D. G.; SHARP, P. M. CLUSTAL – a package for performing multiple sequence alignment on a microcomputer. **Gene**, Amsterdam, v. 73, n. 1, p. 237-244, 1988.
- HUANG, X. A contig assembly program based on sensitive detection of fragment overlaps. **Genomics**, San Diego, v. 14, n. 1, p. 18-25, 1992.
- JAMES, E. K. Nitrogen fixation in endophytic and associative symbiosis. **Field Crops Res.**, Amsterdam, v. 65, p. 197-209, 2000.
- JERMANN, R. M.; OPITZ, J. G.; STACKHOUSE, J.; BENNER, S. A. Reconstructing the evolutionary history of the artiodactyl ribonuclease superfamily. **Nature**, London, v. 374, p. 57-59, 1995.
- JIMENEZ-SALGADO, T.; FUENTES-RAMIREZ, L. E.; TAPIA-HERNANDEZ, A.; MASCARUA-ESPARZA, M. A.; MARINEZ-ROMERO, E.; CABALLERO-MELLADO, J. *Coffea arabica* L., a new host plant for *Acetobacter diazotrophicus*, and isolation of other nitrogen-fixing Acetobacteria. **Appl. Environ. Microbiol.**, Washington, v. 63, n. 9, p. 3676-3683, 1997.
- JUKES, T. H.; CANTOR, C. R. Evolution of protein molecules. In: MUNRO, H. H. (ed.). **Mammalian protein metabolism**. New York: Academic Press, 1969. p. 21-132.
- KARLIN, S.; WWINSTOCK, G.M.; BRENDDEL, V. Bacterial classifications derived from recA protein sequence comparisons. **J. Bacteriol.**, Washington, v. 177, n. 23, p. 6881-6893, 1995.
- KENERLEY, M. E.; MORGAN, E. A.; POST, L.; LINDAHL, L.; NOMURA, M. Characterization of hybrid plasmids carrying individual ribosomal ribonucleic acid transcription units of *Escherichia coli*. **J. Bacteriol.**, Washington, v. 132, p. 931-949, 1977.

KENNEDY, C. Recent progress in characterization of associative and endophytic diazotrophs and their influence on host plant growth. In: NITROGEN FIXATION: FROM MOLECULES TO CROP PRODUCTIVITY, 12th, 1999, Foz do Iguacu. **Proceedings**. Dordrecht: Klumer Academic Publishers, 2000. p. 395-396.

KHBAYA, B.; NEYRA, M.; NORMAND, P.; ZERHARI, K.; FILALI-MALTOUF, A. Genetic diversity and phylogeny of rhizobia that nodulate *Acacia* spp. in Morocco assessed by analysis of rRNA genes. **Appl. Environ. Microbiol.**, Washington, v. 64, n. 12, p. 4912-4917, 1998.

KIMURA, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. **J. Mol. Evol.**, New York, v. 16, p. 111-120, 1980.

KIRCHHOF, G.; ECKERT, B.; STOFFELS, M.; BALDANI, J. I.; REIS, V. M.; HARTMANN, A. *Herbaspirillum frisingense* sp. nov., a new nitrogen-fixing bacterial species that occurs in C4-fibre plants. **Int. J. Syst. Evol. Microbiol.**, Reading, v. Pt1, p. 157-168, 2001.

KIRCHHOF, G.; REIS, V. M.; BALDANI, J. I.; ECKERT, B.; DÖBEREINER, J.; HARTMANN, A. Occurrence, physiological and molecular analysis of endophytic diazotrophic bacteria in gramineous energy plants. **Plant and Soil**, Dordrecht, v. 194, p. 45-55, 1997a.

KIRCHHOF, G.; SCHLOTTER, M.; ASSMUS, B.; HARTMANN, A. Molecular microbial ecology approaches applied to diazotrophs associated with non-legumes. **Soil Biol. Biochem.**, Bethesda, v. 29, n. 5/6, p. 853-862, 1997b.

KISS, A.; SAIN, B.; VENETIANER, P. The number of rRNA genes in *Escherichia coli*. **FEBS Lett.**, Amsterdam, v. 79, n. 77-79, 1977.

KLAPPENBACH, J. A.; SAXMAN, P. R.; COLE, J. R.; SCHMIDT, T. M. rrndb: the ribosomal RNA operon copy number database. **Nucleic Acids Res.**, Oxford, v. 29, n. 1, p. 181-184, 2001.

KLOEPPER, J. W.; BEAUCHAMP, C. J. A review of issues related to measuring colonization of plant roots by bacteria. **Can. J. Microbiol.**, Ottawa, v. 38, p. 1219-1232, 1992.

KUMAR, S.; TAMURA, K.; JAKBSEN, I. B.; NEI, M. **MEGA2 2.0**. Pennsylvania, 1993 – 2000. 1 arquivo (3,1 Mb); Disponível em: <<http://www.megasoftware.net>> MEGA2: Molecular Evolutionary Genetics Analysis Software.

KUMAR, S. A stepwise algorithm for finding minimum evolution trees. **Mol. Biol. Evol.**, Lawrence, v. 13, p. 584-593, 1996.

LAGUERRE, G.; ALLARD, M.; REVOY, F.; AMARGER, N. Rapid identification of rhizobia by Restriction Fragment Length Polymorphism analysis of PCR-amplified 16S rRNA genes. **Appl. Environ. Microbiol.**, Washington, v. 60, p. 56-63, 1994.

LANE, D. J.; PACE, B.; OLSEN, G. J.; STAHL, D. A.; SOGIN, M. L.; PACE, N. R. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analysis. **Proc. Natl. Acad. Sci. USA**, Washington, v. 82, p. 6955-6959, 1985.

LAWRENCE, J.; OCHMAN, H. Molecular archaeology of the *Escherichia coli* genome. **Proc. Natl. Acad. Sci. USA**, Washington, v. 95, p. 9413-9414, 1998.

- LEVIN, B. R.; BERGSTROM, C. T. Bacteria are different: observations, interpretations, speculations, and opinions about the mechanisms of adaptive evolution in prokaryotes. **Proc. Natl. Acad. Sci. USA**, Washington, v. 97, n. 13, p. 6981-6985, 2000.
- LEVINE, J. S.; AUGUSTSSON, T. R. **Orig. Life Evol. Biosph.**, Dordrecht, v. 15, p. 299-318, 1983.
- LUDWIG, W.; SCHLEIFER, K.-H. Phylogeny of bacteria beyond the 16S rRNA standard. **ASM News**, Washington, v. 65, n. 11, p. 752-757, 1999.
- MACHADO, H. B.; FUNAYAMA, S.; RIGO, L. U.; PEDROSA, F. O. Excretion of ammonium by *Azospirillum brasilense* mutants resistant to ethylenediamine. **Can. J. Microbiol.**, Ottawa, v. 37, p. 549-553, 1991.
- MADIGAN, M. T.; MARTINKO, J. M.; PARKER, J. **Biology of Microorganisms**. Eighth edition. Prentice Hall: 1997.
- MARTÍNEZ-ROMERO, E.; CABALLERO-MELLADO, J. *Rhizobium* phylogenies and bacterial genetic diversity. **Crit. Rev. Plant Sci.**, Boca Raton, v. 15, n. 2, p. 113-140, 1996.
- MC ARTHUR, J. V.; KOVACIC, D. A.; SMITH, M. H. Genetic diversity in natural populations of a soil bacterium across a landscape gradient. **Proc. Natl. Acad. Sci. USA**, Washington, v. 85, p. 9621-9624, 1988.
- MEDLIN, L.; ELWOOD, H. J.; STICKEL, S.; SOGIN, M. L. The characterisation of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. **Gene**, Amsterdam, v. 71, p. 491-499, 1988.
- MEHTA, A.; ROSATO, Y. B. Phylogenetic relationships of *Xylella fastidiosa* strains from different hosts, based on 16S rDNA and 16S-23S intergenic spacer sequences. **Int. J. Syst. Evol. Microbiol.**, Reading, v. 51, p. 311-318, 2001.
- MOYER, G. L.; TIEDJE, J. M.; DOBBS, F. C.; KARL, D. M. A computer simulated Restriction Fragment Length Polymorphism analysis of bacteria small-subunit rRNA genes: efficacy of selected tetrameric restriction enzymes for studies of microbial diversity in Nature. **Appl. Environ. Microbiol.**, Washington, v. 62, p. 2501-2507, 1996.
- MYERS, E. W.; MILLER, W. Optimal alignments in linear-space. **Comput. Appl. Biosci.**, Oxford, v. 4, n. 1, p. 11-17, 1988.
- NCBI (NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION). **BLAST**. Disponível em: <<http://www.ncbi.nlm.nih.gov/BLAST/>> Acesso em: maio de 2001.
- NEI, M. Phylogenetic analysis in molecular evolutionary genetics. **Annu. Ver. Genet.**, Palo Alto, v. 30, p. 371-403, 1996.
- NEI, M.; TAKEZAKI, N.; SITNIKOVA, T. Assessing molecular phylogenies. **Science**, Washington, v. 267, p. 253-255, 1995.
- NEI, M.; LI, W.-H. Mathematical model for studying genetic variation in terms of restriction endonucleases. **Proc. Natl. Acad. Sci. USA**, Washington, v. 76, p. 5269-5273, 1979.
- NEWTON, W. E. Nitrogen fixation in perspective. In: NITROGEN FIXATION: FROM MOLECULES TO CROP PRODUCTIVITY, 12th, 1999, Foz do Iguaçu. **Proceedings**. Dordrecht: Klumer Academic Publishers, 2000. p. 3-8.

- NICHOLAS, K. B.; NICHOLAS, H. B. Jr. **GeneDoc 2.6**. Pittsburgh, 1997 - 2000. 1 arquivo (1,3 Mb); Disponível em: < <http://www.psc.edu/biomed/genedoc/> > GeneDoc: a tool for editing and annotating multiple sequence alignments.
- ODUM, E. P. **Ecologia**. Tradução de: Ricardo Iglesias Rios e Christopher J. Tribe. Rio de Janeiro: Editora Guanabara S. A., 1988.
- OLIVE, D. M.; BEAN, P. Principles and applications of methods for DNA-based typing of microbial organisms. **J. Clin. Microbiol.**, Washington, v. 37, n. 6, p. 1661-1669, 1999.
- OLSEN, G. J.; LANE, D. J.; GIOVANNONI, S. J.; PACE, N. R. Microbial ecology and evolution: a ribosomal RNA approach. **Ann. Rev. Microbiol.**, Palo Alto, v. 40, p. 337-365, 1986.
- PALLERONI, N. J.; KUNISAWA, R.; CONTOPOUOU, R.; DOUDOROFF, M. Nucleic acid homologies in the genus *Pseudomonas*. **Int. J. Syst. Bacteriol.**, Washington, v. 23, p. 333-339, 1973.
- PATTEN, C. L.; GLINCK, B. R. Bacterial biosynthesis of indole-3-acetic acid. **Can. J. Microbiol.**, Ottawa, v. 42, n. 3, p. 207-220, 1996.
- PENNY, D. From macromolecules to trees. In: STACKEBRANDT, E.; GOODFELLOW, M. (eds.). **Nucleic acid techniques in bacterial systematics**. New York: John Wiley & Sons. Ltd., 1991. p. 281-324.
- POSTGATE, J. R.; EADY, R. R. In: NITROGEN FIXATION: HUNDRED YEARS AFTER, 1988, Stuttgart. Gustav Fescher .p 31-40.
- POSTGATE, J. R. **Symp. Soc. Gen. Microbiol.**, v. 24, p. 265-292, 1974.
- ROBLETO, E. A.; BORNEMAN, J.; TRIPLETT, E. W. Effects of bacterial antibiotic production on rhizosphere microbial communities from a culture-independent perspective. **Appl. Environ. Microbiol.**, Washington, v. 64, p. 5020-5022, 1998.
- RUSSO, C. A. M.; TAKEZAKI, N.; NEI, M. Efficiencies of different genes and different tree-building methods in recovering a known vertebrate phylogeny. **Mol. Biol. Evol.**, Lawrence, v. 13, p. 525-536, 1996.
- RZHETSKY, A.; NEI, M. Theoretical foundation of the minimum-evolution method of phylogenetic inference. **Mol. Biol. Evol.**, Lawrence, v. 10, p. 1073-1095, 1993.
- SAIKI, R. K.; GELFAND, D. H.; STOFFEL, S.; SHARF, S. J.; HIGUCHI, R.; HORN, G. T.; MULLIS, K. B.; ERLICH, H. A. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. **Science**, Washington, v. 239, p. 487-491, 1988.
- SAITOU, N.; NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Mol. Biol. Evol.**, Lawrence, v. 4, n. 4, p. 406-425, 1987.
- SANKOFF, D. Minimal mutation trees of sequences. **SIAM J. Appl. Math.**, v. 28, n. 1, p. 35-42, 1975.
- SEAL, S. E.; JACKSON, L. A.; YOUNG, J. P. W.; DANIELS, M. J. Differentiation of *Pseudomonas solanacearum*, *Pseudomonas syzygii*, rRNA sequencing: construction of oligonucleotide primers for sensitive detection by polymerase chain reaction. **J. Gen. Microbiol.**, Reading, v. 139, p. 1587-1594, 1993.

SEVILLA, M.; KENNEDY, C. Genetic analysis of nitrogen fixation and plant-growth stimulating properties of *Acetobacter diazotrophicus*, an endophyte of sugarcane. In: TRIPLETT, E. (ed.). **Prokaryotic Nitrogen Fixation**. Wyndham: Horizon Scientific Press, 2000. p. 737-760.

SMITH, J. M.; DOWSON, C. G.; SPRATT, B. G. Localized sex in bacteria. **Nature**, London, v. 349, p. 29-31, 1991.

SNEATH, P. H. A.; SOKAL, R. R. **Numerical taxonomy**. San Francisco: W. H. Freeman, 1973.

SOKAL, R. R.; SNEATH, P. H. A. **Principles of numerical taxonomy**. San Francisco: W. H. Freeman, 1963.

STALEY, J. T. Bacterial biodiversity: a time for place. **ASM News**, Washington, v. 65, n. 10, p. 681-687, 1999.

SUBBA RAO, N. S. Nitrogen-fixing bacteria associated with plantation and orchard plants. **Can. J. Microbiol.**, Ottawa, v. 29, p. 863-866, 1983.

SWOFFORD, D. L.; OLSEN, G. J.; WADDELL, P. J.; HILLIS, D. M. Phylogenetic inference. In: HILLIS, D. M.; MORITZ, C.; MABLE, B. K. (eds.). **Molecular Systematics**. Sunderland: Sinauer Associates, 1996. p. 407-514.

TAJIMA, F.; TAKEZAKI, N. Estimation of evolutionary distance for reconstructing molecular phylogenetic trees. **Mol. Biol. Evol.**, Lawrence, v. 11, p. 278-286, 1994.

TAKAHATA, N. Allelic genealogy and human evolution. **Mol. Biol. Evol.**, Lawrence, v. 10, p. 2-22, 1993.

TAPIA-HERNANDEZ, A.; BUSTILLOS-CRISTALES, M. R.; JIMENEZ-SALGADO, T.; CABALLERO-MELLADO, J.; FUENTES-RAMIREZ, L. E. Natural endophytic occurrence of *Acetobacter diazotrophicus* in pineapple plants. **Microbial Ecol.**, New York, v. 39, n. 1, p. 49-55, 2000.

THOMPSON, J. D.; HIGGINS, D. G.; GIBSON, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Res.**, Oxford, v. 22, n. 22, p. 4673-4680, 1994a.

THOMPSON, J. D.; HIGGINS, D. G.; GIBSON, T. J. Improved sensitivity of profile searches through the use of sequence weights and gap excision. **Compt. Appl. Biosci.**, Oxford, v. 10, p. 19-29, 1994b.

TRAN VAN, V.; GILLIS, M.; HEBBER, K. P.; FERNANDEZ, M.; SEGERS, P.; MARTEL, M. H.; BERGE, O.; MEYER, J. M.; HEULIN, T. Isolation from the rice rhizosphere of a new species of nitrogen-fixing Proteobacteria, belonging to the genus *Burkholderia*. In: NITROGEN FIXATION WITH NON-LEGUMES, sixth, 1994, Cairo. **Proceedings**. Cairo: American University in Cairo Press, 1994. p. 299-309.

URQUIAGA, S.; CRUZ, K. H. S.; BODDEY, R. M. Contribution of nitrogen-fixation to sugar-cane-N-15 and nitrogen-balance estimates. **Soil Sci. Soc. Am. J.**, v. 56, n. 1, p. 105-114, 1992.

- VAN DE PEER, Y.; CHAPELLE, S.; DE WACHER, R. A quantitative map of nucleotide substitution rates in bacterial rRNA. **Nucleic Acid Res.**, Oxford, v. 24, n. 17, p. 3381-3391, 1996.
- VAN DE PEER, Y. **TREECON for Windows 1.3**. Antwerpen, 1994 - 1998. 2 disquetes (1,9 Mb); 3 ½ pol. A software package for the construction and drawing of evolutionary distance trees.
- VANDAMME, P.; POT, B.; GILLIS, M.; DE VOS, P.; KERSTERS, K.; SWINGS, J. Polyphasic taxonomy, a consensus approach to bacterial systematics. **Microbiol. Rev.**, Washington, v. 60, n. 2, p. 407-438, 1996.
- VINUESA, P.; RADEMAKER, J. L. W.; DE BRUIJN, F. J.; WERNER, D. Genotypic characterization of *Bradyrhizobium* strains nodulating endemic woody legumes of the Canary Islands by PCR-Restriction Fragment Length Polymorphism analysis of genes encoding 16S rRNA (16S rDNA) and 16S-23S rDNA intergenic spacers, Repetitive Extragenic Palindromic PCR genomic fingerprinting, and partial 16S rDNA sequencing. **Appl. Environ. Microbiol.**, Washington, v. 64, p. 2096-2104, 1998.
- WÄCHTERSCHÄUSER, G. Before enzymes and templates – theory of surface metabolism. **Microbiol. Rev.**, Washington, v. 52, n. 4, p. 452-484, 1988.
- WANG, Y.; ZHANG, Z.; RAMANAN, N. The actinomycete *Thermobispora bispora* contains two distinct types of transcriptionally active 16S rRNA genes. **J. Bacteriol.**, Washington, v. 179, p. 3270-3276, 1997.
- WAYNE, L. G.; BRENNER, D. J.; COLWELL R. R.; GRIMONT, P. A. D.; KANDLER, O.; KRICHEVSKY, M. I.; MOORE, L. H.; MOORE, W. E. C.; MURRAY, R. G. E.; STACKEBRANDT, E.; STARR, M. P.; TRUPER, H. G. Report of the Ad Hoc Committee on reconciliation of approaches to bacterial systematics. **Int. J. Syst. Bacteriol.**, Washington, v. 37, n. 4, p. 463-464, 1987.
- WEBER, O. B. **Ocorrência e caracterização de bactérias diazotróficas em bananeiras (*Musa spp.*) e abacaxizeiros (*Ananas comosus* (L.) MERRIL) e seus efeitos no crescimento de mudas micropropagadas**. Rio de Janeiro, 1998. 192 p. Tese (Doutorado em Ciência do Solo) - Instituto de Agronomia, Universidade Federal Rural do Rio de Janeiro.
- WILSON, A. C.; CARLSON, S. S.; WHITE, T. J. Biochemical evolution. **Annu. Rev. Biochem.**, Palo Alto, v. 46, p. 573-639, 1977.
- WISTOW, G. Lens crystallins: gene recruitment and evolutionary dynamism. **Trends Biochem. Sci.**, London, v. 18, p. 301-306, 1993.
- WOESE, C. R. Bacterial evolution. **Microbiol. Rev.**, Whashington v. 51, n. 2, p. 221-271, 1987.
- WU, M.; DAVIDSON, N. Use of gene 32 protein staining of single-strand polynucleotides for gene mapping by electron microscopy: application to the phi80d3ilvsu⁺7 system. **Proc. Nat. Acad. Sci. USA.**, Washington, v. 72, p. 4506-4510, 1975.
- XIA, Y.; EMBLEY, T. M.; O'DONNELL, A. G. Phylogenetic analysis of *Azospirillum* by direct sequencing of PCR amplified 16S rDNA. **System. Appl. Microbiol.**, Jena, v. 17, p. 197-201, 1994.
- YABUUCHI, E.; KOSAKO, Y.; OYAIZU, H.; YANO, Y.; HOTTA, H.; HASHIMOTO, Y.; EZAKI, T.; ARAKAWA, M. Proposal of *Burkholderia* gen. nov. and transfer of seven

species of the genus *Pseudomonas* homology group II to the new genus, with the type species *Burkholderia cepacia* (Palleroni and Holmes 1981) comb. nov. **Microbiol. Immunol.**, Tokyo, v. 36, p. 1251-1275, 1992.

YAMADA, Y.; HOSHINO, K.; ISHIKAWA, T. The phylogeny of acetic acid bacteria based on the partial sequences of 16S ribosomal RNA: the elevation of the subgenus *Gluconoacetobacter* to generic level. **Biosci. Biotechnol. Biochem.**, Tokyo, v. 61, p. 1244-1251, 1997.

YANG, W. M.; INOUE, C. J.; ZENG, Y.; BEARSS, D.; SETO, E. Transcriptional repression by YY1 is mediated by interaction with mammalian homolog of the yeast global regulator RPD3. **Proc. Natl. Acad. Sci. USA**, Washington, v. 93, p. 12845-12850, 1996.

YOUNG, J. P. W. Taxonomy and evolution of diazotrophs. In: NITROGEN FIXATION: FROM MOLECULES TO CROP PRODUCTIVITY, 12th, 1999, Foz do Iguaçu **Proceedings**. Dordrecht: Klumer Academic Publishers, 2000a. p. 153-154.

YOUNG, J. P. W. Molecular evolution in diazotrophs: do the genes agree? In: NITROGEN FIXATION: FROM MOLECULES TO CROP PRODUCTIVITY, 12th, 1999, Foz do Iguaçu **Proceedings**. Dordrecht: Klumer Academic Publishers, 2000b. p. 161-164.

YOUNG, J. P. W. Taxonomy of rhizobia. In: DIVERSITY AND TAXONOMY OF RHIZOBIA, 1996, Beijing. **Proceeding**. Beijing: China Agricultural Sciencetech Press., 1996. p. 3-10.

YOUNG, J. P. W. Phylogenetic classification of nitrogen-fixing organisms. In: STACEY, G.; BURRIS, R. H.; EVANS, H. J. (eds.). **Biological nitrogen fixation**. New York: Chapman & Hall, 1992. p. 43-86.

YOUNG, J. P. W.; DOWNER, H. L.; EARDLY, B. D. Phylogeny of the phototrophic *Rhizobium* strain BTAi1 by polymerase chain reaction-based sequencing of a 16S rRNA gene segment. **J. Bacteriol.**, Washington, v. 173, p. 2271-2277, 1991.

ZHANG, H.; HANADA, S.; SHIGEMATSU, T.; SHIBUYA, K.; KAMAGATA, Y.; KANAGAWA, T.; KURANE, R. *Burkholderia kururiensis* sp. nov., a trichloroethylene (TCE)-degrading bacterium isolated from an aquifer polluted with TCE. **Int. J. Syst. Evol. Microbiol.**, Reading, v. 50, p. 743-749, 2000.

APÊNDICES

APÊNDICE 1.....	149
APÊNDICE 2.....	153
APÊNDICE 3.....	157

APÊNDICE 1 – COMPOSIÇÃO DOS MEIOS DE CULTURA

SUMÁRIO

1 MEIO JMV.....	151
2 MEIO NF _b HPN.....	151
3 SOLUÇÃO DE MICRONUTRIENTES PARA MEIO DE CULTURA.....	152
4 SOLUÇÃO DE VITAMINAS PARA MEIO DE CULTURA.....	152

APÊNDICE 1 – COMPOSIÇÃO DOS MEIOS DE CULTURA

1 MEIO JMV (BALDANI, 1996)

As quantidades são calculadas para 1 litro de meio de cultura:--

Manitol	5,0 g
K ₂ HPO ₄	0,6 g
KH ₂ PO ₄	1,8 g
MgSO ₄ .7H ₂ O	0,2 g
NaCl	0,1 g
CaCl ₂ .2H ₂ O	0,2 g
Azul de bromotimol sol. 0,5% em 0,2 N de KOH	2,0 mL
Solução de micronutrientes (ver abaixo)	2,0 mL
Fe EDTA solução 1,64%	4,0 mL
Solução de vitaminas (ver abaixo)	1,0 mL

O pH foi ajustado para 4,2 - 4,5. Adicionados 100 mg de extrato de levedura e autoclavado. Para o meio líquido foram adicionado 10 mmol/ L de glutamato de sódio.

2 MEIO NFb HPN (MACHADO *et al.*, 1991)

As quantidades apresentadas são suficientes para o preparo de 1,0 litro de meio:

Ácido málico ou lactato	5,0 g
K ₂ HPO ₄	0,5 g
MgSO ₄ .7H ₂ O	0,2 g
NaCl	0,1 g
CaCl ₂ .2H ₂ O	0,02 g

Solução de micronutrientes (ver abaixo)	2,0 mL
Azul de bromotimol (solução 0,5% em KOH 0,2 N)	2,0 mL
FeEDTA (solução 1,64%)	4,0 mL
Solução de vitaminas (ver abaixo)	1,0 mL
KOH	4,5 g

O pH foi ajustado para 6,5 – 6,8 com NaOH. Para o meio semi-sólido, foram adicionados 1,75 a 1,8 g de agar/ L e para o meio sólido 15 g/ L.

3 SOLUÇÃO DE MICRONUTRIENTES PARA MEIO DE CULTURA (DÖBEREINER *et al.*, 1995)

Quantidades suficientes para 200 mL de solução:

Na ₂ MoO ₄ .2H ₂ O	0,2 g
MnSO ₄ .H ₂ O	0,235 g
H ₃ BO ₃	0,28 g
CuSO ₄ .5H ₂ O	0,008 g
ZnSO ₄ .7H ₂ O	0,024 g

4 SOLUÇÃO DE VITAMINAS PARA MEIO DE CULTURA (DÖBEREINER *et al.*, 1995)

Quantidades suficiente para 100 mL de solução:

Biotina	10 mg
Pyridoxol - HCl	20 mg
Dissolvidos em banho-maria.	

APÊNDICE 2 - TERMINOLOGIA BÁSICA DE ÁRVORES

APÊNDICE 2 – TERMINOLOGIA BÁSICA DE ÁRVORES (BASEADA EM PENNY, 1991)

As árvores usadas para representar os relacionamentos filogenéticos entre um grupo de organismos são casos particulares de gráficos, que por sua vez são estudados por uma área da matemática. Os componentes básicos destas estruturas são apresentados a seguir e mostrados na Figura 27.

Os gráficos são formados a partir de dois elementos básicos, os *pontos* e as *retas* (as quais sempre conectam somente dois *pontos*); uma *reta* pode ser considerada como dois *pontos* conectados. Os *pontos* no final de uma *reta* são *incidentes* com a *reta* e um par de *pontos* são *adjacentes* se eles são *incidentes* à mesma *reta*. Para se chegar a um determinado *ponto* partindo-se de outro qualquer é preciso percorrer um conjunto de *retas* conectadas e a este percurso dá-se o nome de *caminho*. Finalmente, uma *árvore* representa um caso particular de gráfico, sendo um gráfico completamente conectado (não há *pontos* soltos) e acíclico.

Uma *árvore* pode ser *enraizada* ou *sem raiz* o que, em termos biológicos significa ter ou não um ponto de ancestralidade comum a todos os organismos analisados. Quase todos os métodos para análise de dados de seqüências produzem árvores *sem raiz*. As *retas* da *árvore* podem ser *pendentes* ou *internas*. O *grau* de um *ponto* é o número de *retas* incidentes com aquele *ponto*. Uma *árvore binária* tem *pontos* de *grau* 1 (*pontos terminais*) ou de *grau* 3 (*pontos internos*). Sinônimos para os termos empregados são *linha*, *internó* e *ramo* para as *retas*; *nó* para *pontos*; *ponto do ramo* para *pontos internos*; *folha*, *extremidade* e *nó terminal* para um *ponto pendente*; e *rede* para uma *árvore sem raiz*. O termo *ramo* é particularmente ambíguo, já que pode significar uma *reta* simples ou uma sub-*árvore* com várias *retas* e *pontos*. *Rede* também é ambíguo porque pode ser usado para um gráfico conectado com ciclos.

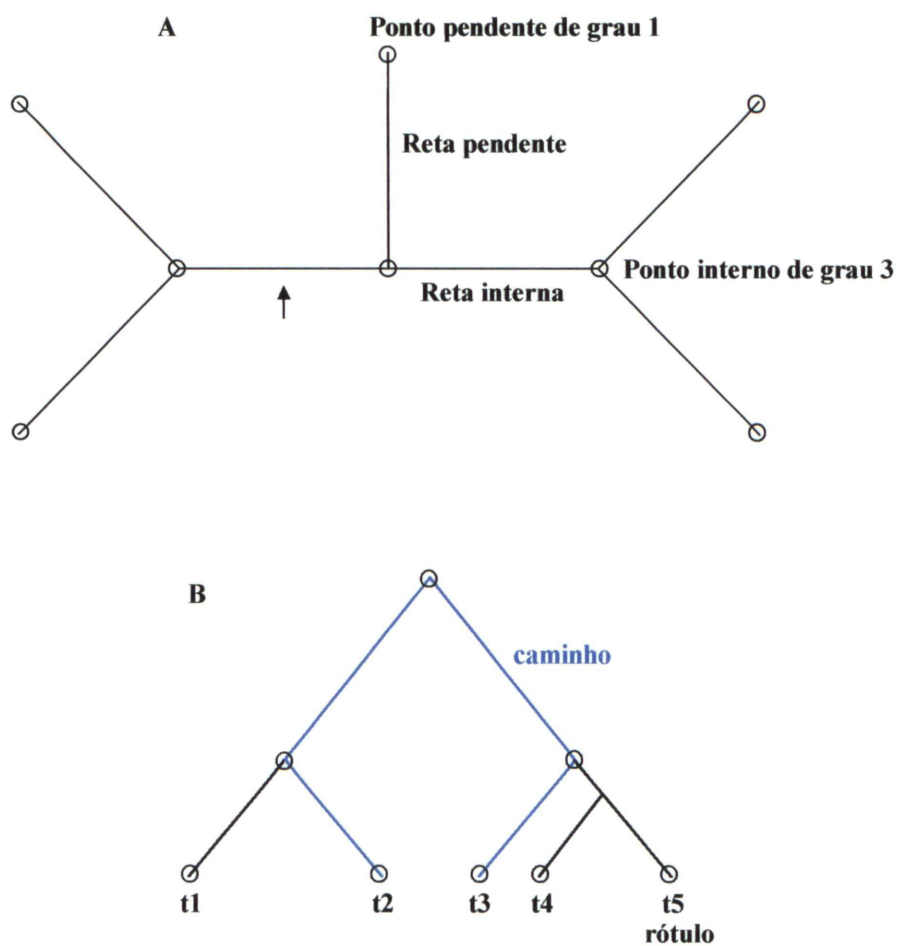
As *árvores* podem ainda ser *não rotuladas* ou *rotuladas*. Um rótulo (geralmente o nome ou número de um taxon) é normalmente designado a um *ponto pendente*, originando uma *árvore filogenética*. Quando as *árvores* são *rotuladas* desta

maneira, há $(2n - 5)!!$ *árvores binárias sem raiz* para n taxa (a notação de duplo fatorial representa $1 \times 3 \times 5 \times \dots \times [2n - 5]$). A *raiz* pode ser colocada em qualquer uma das $2n - 3$ retas existentes em uma *árvore*, originando $(2n - 3)!!$ *árvores binárias enraizadas*.

Uma *árvore enraizada* significa que a direção da evolução através da *árvore* é mostrada. A posição da *raiz* na *árvore* pode ser inferida de uma das seguintes maneiras:

- a) assumir um grau de constância no "relógio molecular" e colocar a *raiz* na metade do mais longo ramo ou
- b) usar um "grupo externo", uma seqüência de um organismo que se sabe previamente não pertencer ao grupo de seqüências analisadas.

FIGURA 27 - REPRESENTAÇÃO DOS ELEMENTOS BÁSICOS DE UMA ÁRVORE. **A.** UMA ÁRVORE BINÁRIA SEM RAIZ; **B.** UMA ÁRVORE FILOGENÉTICA BINÁRIA ENRAIZADA, DERIVADA DE **A** POR INSERÇÃO DE UM PONTO ADICIONAL NA RETA INDICADA POR UMA SETA.



**APÊNDICE 3 – ALGORITMOS DOS PROGRAMAS DE
COMPUTADOR UTILIZADOS**

SUMÁRIO

1 CLUSTAL.....	159
1.1 HISTÓRICO.....	159
1.2 CARACTERÍSTICAS.....	159
1.3 ALGORITMO.....	160
1.4 O ALINHAMENTO DOS PARES.....	160
1.5 A ÁRVORE GUIA.....	160
1.6 O ALINHAMENTO PROGRESSIVO.....	161
2 CONTIG ASSEMBLY PROGRAM (CAP).....	162
2.1 ALGORITMO.....	162
3 BASIC LOCAL ALIGNMENT SEARCH TOOL.....	164
3.1 CARACTERÍSTICAS.....	164
3.2 ALGORITMO.....	164
4 UPGMA (<i>UNWEIGHTED PAIR-GROUP METHOD USING ARITHMETIC AVERAGES</i>).....	167
4.1 CARACTERÍSTICAS.....	167
4.2 ALGORITMO.....	167
5 <i>NEIGHBOR-JOINING</i>	168
5.1 CARACTERÍSTICAS.....	168
5.2 ALGORITMO.....	169
6 TESTE DE <i>BOOTSTRAP</i>	171

APÊNDICE 3 - ALGORITMOS DOS PROGRAMAS DE COMPUTADOR UTILIZADOS

1 CLUSTAL

1.1 HISTÓRICO

O programa foi originalmente escrito por HIGGINS & SHARP (1988 e 1989), em linguagem Fortran para microcomputadores com sistema operacional MSDOS. Um alinhamento completo de seqüências era feito rodando-se três programas sucessivamente. Mais tarde, estes foram fundidos em um único programa para VAX/VMS. Em seguida, foi lançada a série "V" do programa (ClustalV), escrita em linguagem C, contendo todas as características do programa anterior, além de outras novas e capaz de rodar em diferentes computadores, incluindo PCs e Macs (HIGGINS *et al.*, 1991). Posteriormente foi lançada a série "W" (THOMPSON *et al.*, 1994a), sendo aumentada a sensibilidade no alinhamento, principalmente de seqüências distantemente relacionadas. Importantes mudanças no cálculo dos pesos dados às seqüências e às falhas e o uso das matrizes para comparação entre resíduos de aminoácidos permitiram esse aumento de precisão. Os programas da série "X" foram adaptados para ambiente Windows e apresentam algumas facilidades no tratamento e edição das seqüências, embora o programa permaneça com as mesmas características dos programas da série "W" (THOMPSON *et al.*, 1997).

1.2 CARACTERÍSTICAS

A principal proposta do programa é o alinhamento múltiplo (global) de seqüências de DNA e proteínas. Entretanto, é possível ainda fazer edição e converter formatos de arquivos de seqüências e calcular uma árvore filogenética pelo método *Neighbor-Joining* (SAITOU & NEI, 1987), com opção para *bootstrap* e correção para múltiplas substituições pelo método de KIMURA (1980). O alinhamento múltiplo é calculado usando uma abordagem freqüentemente referida como "alinhamento progressivo" (FENG & DOOLITTLE, 1987). Esta abordagem envolve uma previsão

inicial da filogenia das seqüências para que se possa usar a ordem dos ramos da árvore inferida no alinhamento de grupos de seqüências progressivamente maiores e cada vez mais distantemente relacionados.

1.3 ALGORITMO

O algoritmo de alinhamento múltiplo básico consiste de três estágios principais:

- a) todos os pares de seqüências são alinhadas separadamente para calcular uma matriz de distância dando a divergência entre cada par de seqüências;
- b) uma árvore guia é calculada a partir da matriz de distâncias calculada e;
- c) as seqüências são progressivamente alinhadas de acordo com a ordem dos ramos na árvore guia.

1.4 O ALINHAMENTO DOS PARES

Na série "V" do programa, as distâncias entre os pares são calculadas usando um método aproximado e rápido (BASHFORD *et al.*, 1987). Ele permite o alinhamento de um grande número de seqüências, mesmo em um microcomputador. As pontuações (*scores*) são calculados como o número de pareamento entre *k-tuples* (fragmentos curtos de resíduos, normalmente 1 ou 2 para proteína e 2 a 4 para nucleotídeos) no melhor alinhamento entre duas seqüências menos uma penalidade dada a cada *gap*. Nas séries "W" e "X" do programa, um método de pontuação mais lento e preciso é também oferecido a partir de alinhamentos de programação dinâmica completo (*dynamic programming alignments*) usando duas penalidades para falhas (abertura e extensão) e uma matriz de pesos para aminoácidos. Estas pontuações são calculadas como o número de identidades no melhor alinhamento dividido pelo número de resíduos comparados (posições contendo falhas são excluídas).

1.5 A ÁRVORE GUIA

A árvore usada para guiar o processo de alinhamento múltiplo final é calculada a partir da matriz de distâncias do passo 1, usando o método de *Neighbor-*

Joining (SAITOU & NEI, 1987). Esta é outra mudança apresentada nos programas da série "W"; anteriormente a árvore guia era calculada a partir do método UPGMA (SNEATH & SOKAL, 1973), menos robusto e confiável. O método *Neighbor-Joining* processa uma árvore sem raiz com comprimentos de ramos proporcionais à divergência estimada ao longo do ramo. A raiz é colocada pelo método do ponto médio (*mid-point method*), numa posição onde a média do comprimento dos ramos de cada lado da raiz é igual. A árvore é também usada para calcular um peso para cada seqüência, que é dependente da distância da seqüência à raiz da árvore, ou seja, a soma dos comprimentos dos ramos da raiz à seqüência em questão. Entretanto se o ramo possui mais de uma seqüência, seu comprimento é dividido pelo número de seqüências nele contido e o valor obtido usado na soma (THOMPSON *et al.*, 1994).

1.6 O ALINHAMENTO PROGRESSIVO

O procedimento básico neste estágio é usar uma série de alinhamentos aos pares (*pairwise alignments*) para alinhar grupos cada vez maiores de seqüências, seguindo a ordem dos ramos na árvore guia. Procede-se o alinhamento a partir das extremidades da árvore enraizada em direção a raiz. A cada estágio, um algoritmo de programação dinâmica completo (*full dynamic programming*; MYERS & MILLER, 1988; THOMPSON, 1994) é usado com uma matriz de peso para os resíduos (*residue weight matrix*) e penalidades para abertura e extensão de falhas. Cada passo consiste em alinhar dois alinhamentos pré-existentes ou seqüências. As posições das falhas que foram introduzidos durante uma fase inicial do alinhamento de seqüências próximas, não são alterados a medida que novas seqüências são adicionadas. Isto é justificado porque a colocação de falhas no alinhamento entre seqüências próximas é muito mais preciso do que entre seqüências distantes. Quando todas as seqüências são muito divergentes (ex., menos de 25 - 30% de identidade entre qualquer par de seqüências), esta abordagem progressiva torna-se muito menos precisa.

2 CONTIG ASSEMBLY PROGRAM (CAP; HUANG, 1992)

2.1 ALGORITMO

O programa usa um algoritmo de programação dinâmica (*dynamic programming algorithm*) para calcular a pontuação máxima do alinhamento de sobreposição entre dois fragmentos (HUANG, 1992). Fragmentos em orientação ao acaso são montados em uma seqüência contígua (*contig*). O programa é eficiente no uso de memória: um grande número de fragmentos de comprimento e orientação arbitrários podem ser montados.

Dois grupos de pontuações são feitas pelo algoritmo de programação dinâmica para conduzir o alinhamento nas regiões de sobreposição: o grupo de pontuação "forte" (*heavy set*) usado para regiões de baixa taxa de erro de seqüenciamento e o grupo de pontuação "leve" (*light set*), para pontas de fragmentos de alta taxa de erro de seqüenciamento (mais do que 5%). Neste processo é atribuído um valor para a identidade entre bases (*MATCH*), para bases diferentes (*MISMATCH*) e para a inclusão de falhas (*EXTEND*). Dessa forma a sobreposição entre as seqüências é aquela em que foi obtido o mais alto valor para a soma das pontuações dadas.

	<i>Heavy set</i>	<i>light set</i>
<i>MATCH</i>	2	2
<i>MISMATCH</i>	-6	-3
<i>EXTEND</i>	4	2

Na montagem inicial das seqüências contíguas, qualquer sobreposição deve ter um comprimento e uma porcentagem de identidade mínimos, que são definidos nos parâmetros *OVERLEN* e *PERCENT*, definidos pelo usuário (20 bases e 85% foram os valores usados para montagem das seqüências do isolados diazotróficos). Após a montagem inicial, o programa tenta juntar as seqüências contíguas em seqüências

contíguas maiores usando parâmetros de sobreposição menos estridentes. Duas seqüências contíguas são montadas se a pontuação da região alinhada atinge o valor mínimo dado ao parâmetro *CUTOFF*. Por sua vez, o valor para o *CUTOFF* é escolhido de acordo com o valor para o parâmetro *MATCH*.

As regiões de sobreposição que serão alinhadas são determinadas pelo parâmetro *DELTA*. As condições deste valor são usadas para rapidamente rejeitar pares de fragmentos que não poderiam ter uma relação de sobreposição. O algoritmo de programação dinâmica é então aplicado somente a pares de fragmentos que passam por este crivo. Um valor alto para *DELTA* significa condições estridentes, onde o valor para *DELTA* é um número real de no mínimo 8,0.

3 BASIC LOCAL ALIGNMENT SEARCH TOOL (BLAST; ALTSCHUL *et al.*, 1997)

3.1 CARACTERÍSTICAS

É um método conveniente para detectar homologia de uma seqüência biológica com seqüências caracterizadas já existentes que compõem um banco de dados. O programa usa um algoritmo heurístico e procura por homologia através de regiões alinhadas localmente de identidade e/ ou similaridade entre uma seqüência desconhecida e seqüências em um banco de dados. É composto por cinco programas que são, na verdade variações para o tipo de comparação que será feita entre a seqüência para a qual se deseja fazer a pesquisa de homologia e aquelas compondo o banco de dados. Os programas e suas funções são:

- a) *blastp*: compara uma seqüência de aminoácido com um banco de dados de seqüências de proteínas;
- b) *blastn*: compara uma seqüência de nucleotídeos com um banco de dados de seqüências de nucleotídeos (foi o programa usada para a pesquisa de homologia feita com as seqüências dos isolados diazotróficos);
- c) *blastx*: compara o produto da tradução das seis fases de leitura (*frames*) conceituais de uma seqüência de nucleotídeos com um banco de dados de seqüências de proteínas;
- d) *tblastn*: compara uma seqüência de proteína com um banco de dados de seqüências de nucleotídeos automaticamente traduzido em todas nas seis fases de leitura conceituais;
- e) *tblastx*: compara a tradução das seis fases de leitura conceituais de uma seqüência de nucleotídeos com um banco de dados de nucleotídeos traduzido nas seis fases de leitura conceituais.

3.2 ALGORITMO

O algoritmo segue os seguintes passos básicos:

- a) Uma seqüência para a qual se deseja fazer a pesquisa de homologia é dividida em seqüências curtas (também denominadas de "palavras" e contendo entre 3 a 8 resíduos, dependendo se é uma seqüência de aminoácidos ou nucleotídeos);
- b) uma tabela (denominada de *lookup table*) de todos os fragmentos de seqüência com o mesmo tamanho e que podem parear com cada uma destas "palavras", com uma pontuação (*score*) acima de um limite definido é construída;
- c) o banco de dados de seqüências é varrido em busca da ocorrência de seqüências da *lookup table*;
- d) quando uma "palavra" é encontrada no banco de dados, a qual pode alinhar a uma "palavra" na seqüência desconhecida, acima de um limite crítico, o alinhamento é estendido em ambas as direções. Esta extensão continua enquanto seu comprimento não exceder um limite definido, sem aumentar ainda mais a pontuação do alinhamento;
- e) quando uma extensão é terminada, e a pontuação total do sub-alinhamento fica acima de outro limite definido, o alinhamento é informado;
- f) quando um destes sub-alinhamentos é produzido, aquela seqüência particular do banco de dados é novamente varrida para a ocorrência de outros pares de segmentos de alta pontuação não redundantes (*high-scoring segment pairs*, HSPs) as quais pontuam acima de outro valor previamente definido (a soma de vários sub-alinhamentos não significantes dentro das mesmas duas seqüências pode, quando tomados em conjunto, indicar uma similaridade significativa, indicativo de homologia);
- g) uma medida estatística é mostrada, indicando a probabilidade de uma HSP com pontuação similar ou um grupo de HSPs encontrado para uma dada seqüência resultaria da pesquisa do mesmo banco de dados com uma

seqüência gerada ao acaso com o mesmo comprimento da seqüência que foi submetida.

4 UPGMA (*UNWEIGHTED PAIR-GROUP METHOD USING ARITHMETIC AVERAGES*; SNEATH & SOKAL, 1973)

4.1 CARACTERÍSTICAS

Este método consiste em um dos muitos métodos de análise de grupo (*Cluster Analysis*). Em geral estes métodos baseiam-se em uma série de passos que compactam gradativamente uma matriz de distâncias entre dados, unindo os dados ou grupos de dados dois a dois a cada passo, até que restem apenas dois grupos de dados. O princípio deste método é bastante simples e implica em que a taxa evolutiva permanece constante em todos os ramos da árvore produzida (neste caso, também denominada de dendrograma).

4.2 ALGORITMO

O método parte de uma matriz de distâncias calculadas para cada par de dados (aqui representados por cada um dos isolados). Nesta matriz, é feita uma pesquisa pelo valor de menor dissimilaridade (ou maior similaridade). Quando o valor é encontrado o par de dados referente a este valor é então “fundido” para que uma nova matriz de distâncias seja calculada. A nova matriz é semelhante a primeira com a diferença de que a distância entre os dados e os dados fundidos no passo anterior é agora calculado como sendo a média aritmética das distâncias entre o dado e cada um dos dados fundidos. Em cada uma das etapas dois dados são fundidos e uma nova matriz, com menos um dado é recalculada. O processo continua até que somente dois grupos de dados restem e sejam também unidos pelo princípio das médias usado pelo método. Em quaisquer dos passos, quando ocorrem dois ou mais valores de dissimilaridade que sejam iguais e os menores dentro da matriz, a escolha do par de dados a ser fundido é feita aleatoriamente entre estes valores.

5 NEIGHBOR-JOINING (SAITOU & NEI, 1987)

5.1 CARACTERÍSTICAS

O método *Neighbor-Joining* (NJ) é um método baseado em distâncias, o que significa dizer que é preciso converter os dados de seqüências (após alinhamento) em um valor numérico que é uma estimativa da distância evolutiva, obtido para cada par de seqüências. Como mencionado na INTRODUÇÃO, os métodos de distância utilizam-se de um determinado critério para a escolha da "melhor árvore", sendo os mais empregados:

- a) método do quadrado mínimo (*least square* – LS) – o princípio do método baseia-se em calcular a soma mínima dos quadrados das diferenças entre os pares de distâncias observados e estimados para uma dada topologia e escolher a topologia que mostra a menor soma mínima dos quadrados das diferenças. Um problema inerente ao método é a possibilidade de fornecer comprimento de ramos negativos (NEI, 1996). Uma maneira de contornar este problema é conduzir a estimativa dos quadrados mínimos dos comprimentos dos ramos com a restrição de comprimentos de ramos negativos (FELSENSTEIN, 1988). Entretanto este artifício não é satisfatório quando se trabalha com mais do que quatro seqüências (NEI, 1996).
- b) método de evolução mínima (*minimum evolution* – ME) – neste método o comprimento dos ramos de uma árvore são estimados por um certo algoritmo a partir de dados de distâncias dos pares de seqüências e a soma total (S) do comprimento dos ramos é calculada para cada uma das topologias possíveis. A topologia que mostra o menor valor de "S" irá então ser escolhida como a árvore mais provável. Embora o método ME seja estatisticamente interessante ele requer uma grande quantidade de tempo de processamento computacional para examinar todas as diferentes

topologias se o número de seqüências é maior que 10 (NEI, 1996).

O método NJ é, na verdade, uma simplificação do método ME para inferir uma árvore bifurcada. Neste método, o valor "S" não é computado para todos ou várias topologias diferentes, mas o exame de diferentes topologias está embebida no algoritmo, então somente uma árvore final é produzida. Na prática, qualquer medida de distância está sujeita a erros estocásticos, então a árvore NJ obtida pode não necessariamente ser a verdadeira árvore. De acordo com simulações em computador, este método quase sempre produz a mesma topologia que a árvore ME quando a extensão das diferenças entre as seqüências é suficientemente grande e o número de nucleotídeos examinados também é grande (>500). Quando a última condição não é satisfeita, entretanto, a árvore NJ pode ser consideravelmente diferente da árvore ME (RZHETSKY & NEI, 1993), apesar disto a diferença no valor "S" entre as árvores NJ e ME é normalmente estatisticamente não significativa. Além do mais, usando simulação de computador, Kumar (KUMAR, 1996) mostrou que quando as árvores NJ e ME são diferentes, a última árvore não é necessariamente próxima da verdadeira árvore (topologia). Em uma dada taxa, a alta eficiência do método NJ em obter a árvore ME ou a verdadeira árvore parece originar do fato de que em cada passo do agrupamento das seqüências o princípio da evolução mínima é aplicado e a aplicação repetida deste princípio poderia reduzir os efeitos dos erros de amostragem na construção da topologia (NEI, 1996).

O método é conceitualmente relacionado a análise de grupos, mas não assume que a taxa evolutiva é a mesma em todas as linhagens. No método NJ, uma matriz de distâncias modificada é construída, na qual a separação entre cada par de nós é ajustada com base na sua média de divergência a partir de todos os outros nós, que tem o efeito de normalizar a divergência de cada seqüência a taxa média (SWOFFORD *et al.*, 1996).

5.2 ALGORITMO

Neste método a inferência da árvore começa com uma topologia onde não há

nenhum grupo formado (chamada pelos autores de “topologia estrela”). Entretanto, na prática algumas seqüências estão mais relacionadas entre si do que com outras. O algoritmo irá juntar duas seqüências (denominadas de *neighbors*) e calcular a soma (S) do comprimento total dos ramos resultante desta topologia. Para isso os valores de distância calculados para cada par de seqüência é usado. Como não é sabido *a priori* quais as duas seqüências que se deve unir para reconstruir a verdadeira árvore filogenética, o algoritmo une todas as seqüências, duas a duas e calcula o valor S, compondo uma matriz S_{ij} com “i” colunas (número de seqüências sendo analisada) e “j” linhas (número de seqüências menos um). A partir da matriz, o algoritmo escolhe o menor valor de S_{ij} e une o par de seqüências a partir do qual o valor foi obtido. Assim é finalizado o primeiro ciclo do algoritmo. No próximo ciclo o processo se repete, entretanto as seqüências *neighbors* da etapa seguinte são fundidas, derivando-se um valor de distância médio para esta “nova” seqüência. A partir daí, o processo se repete da mesma forma, sendo que agora a matriz S_{ij} possui uma coluna e uma linha a menos que a anterior. Nas etapas seguintes novas matrizes S_{ij} serão calculadas com uma coluna e uma linha sendo subtraídas a cada ciclo, até que somente três colunas restem, gerando uma topologia de árvore sem raiz.

6 TESTE DE *BOOTSTRAP*

Este teste consiste em fazer uma re-amostragem utilizando os próprios dados e foi introduzido na análise filogenética, para testar a reprodutibilidade dos ramos de uma árvore filogenética, por FELSEINSTAIN (1988). A utilização deste teste na análise filogenética parte do alinhamento de seqüências, onde são feitas as re-amostragens. As colunas do alinhamento, ou seja, os sítios, são tomados ao acaso para "construir" um novo alinhamento, com o mesmo número de colunas do alinhamento original. A partir deste "alinhamento amostrado" (amostra *bootstrap*) é calculada uma árvore filogenética usando o mesmo método e os mesmos parâmetros usados para o cálculo da árvore a partir do alinhamento original. Diversas amostras *bootstrap* podem ser obtidas e uma árvore calculada para cada uma delas individualmente. Na verdade, quanto maior o número de amostras *bootstrap* produzidos, mais confiável será a análise. É comum encontrar na literatura análises feitas com 100 amostras *bootstrap*, valor limitado apenas pelo grande consumo da capacidade de processamento dos computadores. Entretanto, com os computadores atuais, mesmo os de uso doméstico, aliado aos algoritmos em uso, é possível analisar de 1.000 a 5.000 "amostras *bootstrap*" em um tempo razoável, quando se utiliza determinados métodos (como por ex. NJ).