ANA PAULA ANDREAZZA



CURITIBA 2004 ANA PAULA ANDREAZZA

SEQÜENCIAMENTO PARCIAL DO REPLICON pAZ6 de Azospirillum brasilense FP2

Dissertação apresentada ao Curso de Pós-Graduação em Bioquímica e Biologia Molecular da Universidade Federal do Paraná, como requisito parcial à obtenção do título de Mestre em Ciências - Bioquímica.

Orientadores: Dr. Emanuel M. de Souza Dr.^a Cyntia M. T. Fadel Picheth

Ana Paula Andreazza

SEQÜENCIAMENTO PARCIAL DO REPLICON pAZ6 de Azospirillum brasilense FP2

Dissertação aprovada pela comissão examinadora formada pelos professores abaixo assinados, como requisito parcial para a obtenção do Título de Mestre no Curso de Pós-Graduação em Ciências - Bioquímica da Universidade Federal do Paraná.

Prof. Dr. Emanuel Maltempi de Souza (orientador)

Prof. Dr. Emanuel Maltempi de Souza (orientador) Depto. de Bioquímica - UFPR.

foreluna

Prof. Dr. Fábio de Oliveira Pedrosa. Depto. de Bioquímica - UFPR.

- Leda Satie Chats

Prof^a Dr^a Leda Satie Chubatsu Depto. de Bioquímica - UFPR.

SMAand

Prof. Dr. Sílvio Marques Zanata Depto. de Patologia Básica - UFPR.

Curitiba 2004

Dedico esta tese ao meu pai Fernando, a minha mãe Marila. aos meus irmãos Fernanda e Marcus e ao meu namorado Ricardo pelo incentivo, apoio e colaboração em todos os momentos.

AGRADECIMENTOS

Agradeço:

Aos professores Emanuel M. de Souza e Cyntia M. T. Fadel-Picheth pela orientação, oportunidade, incentivo, apoio e sugestões.

Ao professor Dr. Fábio O. Pedrosa pela oportunidade, incentivo e colaboração.

À professora Dra. Leda S. Chubatsu pela colaboração, interesse e sugestões.

À professora Dra. Roselli Wassem, à Dra. Rose Adele Monteiro e ao funcionário Valter Baura pela colaboração e sugestões na construção e sequênciamento das bibliotecas genômicas.

Ao professor Dr. Leonardo M. Cruz e ao Augusto F. J. Favetti pela colaboração e apoio com a bioinformática.

Aos professores Dr. Marshal Geofrey Yates, Dra. Maria Berenice R. Steffens, Dra. Liu Un Rigo, Dra. Elaine Benelli e Dra. Glaci Zancan pelo incentivo e colaboração.

À cordenação do curso de pós-graduação em Bioquímica, às professoras Dra. Fani Reicher e Dra. Leda S. Chubatsu pelo apoio.

As amigas Daniela Seichas e Andrea Tarzia pela paciência, apoio e atenção em todas as horas. Aos amigos André Luíz F. Souza, Carolina Galvão, Ana Claudia Bonato, Marcelo Assumpção, Fabiane G. M. Rego, Lilian Noindorf, Luciano Fernandes, Luíza M^a de Araújo, Juliana Ramos e a Juliana Inaba pelo carinho e atenção.

Aos companheios de laboratório Helison, Lysangela, Stefan, Patrícia, Marcelo, Karen, Michelle, Marco Aurélio, Humberto, Rose Adele e Geovana pela descontração e apoio no dia a dia.

Aos funcionários e amigos Valter Baura, Roseli Prado e D. Julieta Pie, pelo carinho, atenção, apoio e amizade.

Ao meu pai Fernando, a minha mãe Marila, aos meus irmãos Fernanda e Marcus e ao meu namorado Ricardo pelo incentivo, apoio e colaboração em todos os momentos.

A todos, neste e em outros planos, que de uma forma ou de outra me auxiliaram nesta conquista.

SUMÁRIO

L	STA DE ABREVIATURAS	viii					
R	SUMO	ix					
1	INTRODUÇÃO	RODUÇÃO1					
	1.1 ESTRUTURA GENÔMICA EM BACTÉRIAS						
	1.2 GENOMA DE PROTEOBACTÉRIAS						
	1.2.1 GENOMA DE Sinorhizobium meliloti	8					
	1.2.2 GENOMA DO Agrobacterium tumefaciens	11					
	1.2.3 GENOMA DE OUTRAS PROTEOBACTÉRIAS	15					
	1.3 Azospirillum	17					
	1.3.1 Azospirillum brasilense	19					
	1.3.2 GENÉTICA DO Azospirillum brasilense	20					
	1.4 ESTRUTURA GENÔMICA DO GÊNERO Azospirillum	23					
2	OBJETIVO	25					
3 MATERIAIS E MÉTODOS							
	3.1 BACTÉRIAS E PLASMÍDEOS UTILIZADOS	26					
	3.2 REAGENTES	26					
3.3 MEIOS DE CULTURA							
	3.3.1 MEIOS EMPREGADOS PARA O CULTIVO DE Azo	spirillum					
	brasilense	27					
	3.3.2 MEIOS EMPREGADOS PARA O CULTIVO DE E. coli						
	3.4 CONDIÇÕES DE CULTIVO	29					
	3.5 MANIPULAÇÃO DO DNA						
	3.5.1 OBTENÇÃO DO DNA CROMOSSÔMICO INTACTO	30					
	3.5.2 ISOLAMENTO DOS REPLICONS DE Azospirillum b	rasilense					
	FP2	31					
	3.5.2.1ISOLAMENTO DO REPLICON UTILIZANDO A ENZ	ZIMA β-					
	AGARASE						
	3.6 AMPLIFICAÇÃO NÃO-PCR DO DNA GENÔMICO	32					

	3.7 QUANTIFICAÇÃO DO DNA	33
	3.8 FRAGMENTAÇÃO MECÂNICA DO DNA	33
	3.9 CONSTRUÇÃO DA BIBLIOTECA GENÔMICA	34
	3.10 TRANSFORMAÇÃO BACTERIANA POR ELETROPORAÇÃO	34
	3.11 MINI-PREPARAÇÃO DE PLASMÍDEOS EM MICROPLACAS	36
	3.12 REAÇÃO DE SEQÜENCIAMENTO DO DNA	37
	3.13 ANÁLISE E ALINHAMENTO DAS SEQÜÊNCIAS	38
	3.14 ANOTAÇÃO DA SEQÜÊNCIA	
4	RESULTADOS E DISCUSSÃO	40
	4.1 IDENTIFICAÇÃO DO REPLICON pAZ6 (630 kpb) DE Azosp	irillum
	brasilense FP2 POR ELETROFORESE EM CAMPO PULSADO	40
	4.2 PURIFICAÇÃO E AMPLIFICAÇÃO DO REPLICON pAZ6 DE Azosp	irillum
	brasilense FP2	47
	4.3 FRAGMENTAÇÃO MECÀNICA DO DNA DO REPLICON	pAZ6
	AMPLIFICADO COM TEMPLIPHI	52
	4.4 CLONAGEM DO DNA DO REPLICON pAZ6 AMPLIFICADO	COM
	TEMPLIPHI	52
	4.5 SEQÜENCIAMENTO DOS CLONES ALEATÓRIOS DO REPI	LICON
	pAZ6	53
	4.6 AGRUPAMENTO DAS SEQÜÊNCIAS EM CONTÍGUOS	55
	4.7 ANÁLISE DOS CONTÍGUOS	57
	4.8 COMPLEMENTO GÊNICO PARCIAL DO REPLICON pAZ6	63
	4.8.1 BISSÍNTESE DE PAREDE CELULAR	63
	4.8.2 BIOSSÍNTESE DE VITAMINA B ₁₂	65
	4.8.3 BIOSSÍNTESE DE NUCLEOTÍDEOS	65
	4.8.4 METABOLISMO DE CARBOIDRATOS	66
	4.8.5 METABOLISMO DE ÁCIDOS GRAXOS	67
	4.8.6 METABOLISMO DE ENERGIA	67
	4.8.7 METABOLISMO DE NITROGÊNIO	69
	4.8.8 QUIMIOTAXIA	70

	4.8.9 SISTEMAS DE TRANSPORTE	72
5	CONCLUSÃO	74
6	REFERÊNCIAS BIBLIOGRÁFICAS	76

LISTA DE TABELAS

TABELA 1 COMPOSIÇÃO DE ALGUNS GENOMAS BACTERIANOS
COMPLETAMENTE SEQÜENCIADOS E COM MÚLTIPLOS
ELEMENTOS HEREDITÁRIOS6
TABELA 2 BACTÉRIAS E PLASMÍDEOS
TABELA 3 NÚMERO DE REPLICONS DE A. brasilense ESTIRPE FP246
TABELA 4 CONCENTRAÇÃO DO DNA DO REPLICON pAZ6
AMPLIFICADO POR TEMPLIPHI
TABELA 5 BIBLIOTECAS ALEATÓRIAS DO REPLICON pAZ6 de A.
<i>brasilense</i> FP252
TABELA 6 RESUMO DO SEQÜENCIAMENTO DE 1858 CLONES DAS
BIBLIOTECAS ALEATÓRIAS DE REPLICON pAZ6 de
A. brasilense FP254
TABELA 7 ANÁLISE ESTATÍSTICA DA QUALIDADE DAS BASES DAS
SEQÜÊNCIAS OBTIDAS54
TABELA 8 ESTATÍSTICA DA MONTAGEM DOS CONTÍGUOS DO
REPLICON pAZ6 DE Azospirillum brasilense FP256
TABELA 9 PROTEÍNAS CODIFICADAS PELO REPLICON pAZ6 DE A.
brasilense ESTIRPE FP259

LISTA DE FIGURAS

FIGURA 1 MAPA FÍSICO DO VETOR pCR4BLUNT-TOPO
FIGURA 2 PERFIL ELETROFORÉTICO EM CAMPO PULSADO DO DNA
INTACTO DE A. brasilense ESTIRPE FP2 UTILIZANDO PULSOS
ESCALONADOS41
FIGURA 3 PERFIL ELETROFORÉTICO EM CAMPO PULSADO DO DNA
INTACTO DE A. brasilense ESTIRPE FP2 UTILIZANDO UM
GRADIENTE DE PULSOS44
FIGURA 4 PERFIL ELETROFORÉTICO EM CAMPO PULSADO DO DNA
INTACTO DE A. brasilense ESTIRPE FP245
FIGURA 5 ISOLAMENTO DO REPLICON pAZ6 DE A. brasilense FP2 POR
ELETROFORESE EM CAMPO PULSADO50
FIGURA 6 AMPLIFICAÇÃO ISOTÉRMICA DO REPLICON pAZ651
FIGURA 7 PREDIÇÃO DE REGIÕES CODIFICADORAS DE PROTEÍNAS DE
UM CONTÍGUO PELO PROGRAMA FRAMEPLOT57
FIGURA 8 ALINHAMENTO DAS SEQÜÊNCIAS DE AMINOÁCIDOS
DEDUZIDAS DA SEQÜÊNCIA DE NUCLEOTÍDEOS DO GENE
murA DE Xanthomonas campestris E DA ORF2C DO CONTÍGUO
14464
FIGURA 9 ESQUEMA DAS REAÇÕES ENVOLVIDAS NO METABOLISMO
DE ÁCIDOS URÔNICOS EM BACTÉRIAS66
FIGURA 10 ALINHAMENTO DAS SEQÜÊNCIAS DE NUCLEOTÍDEOS DO
GENE sbpA de Azospirillum brasilense Sp7 E DA ORF2 DO
CONTÍGUO 37

LISTA DE ABREVIATURAS

16S	rDNA gene que codifica para DNA ribossomal 16S
BLAST	do ingês "basic local alignment search tool"
D.O	densidade ótica
DNA	ácido desoxiribonucleico
EDTA	ácido etilenodiamino tetra acético
Kda	quilo daltons
Kpb	kilo pares de bases
MDa	mega daltons
Mpb	mega pares de bases
ORF	do inglês "open reading frame"
pb	pares de bases
PCR	do inglês "polymerase chain reaction"
PFGE	do inglês "pulse-field gel electrophoresis"
RNA	ácido ribonucleico
SDS	dodecilsulfato de sódio
TRIS	tris(hidroximetil)-aminometano
UV	ultravioleta

RESUMO

A organização genômica das alfa-proteobactérias é altamente variável e geralmente muito complexa. O genoma da rizobactéria A. brasilense é composto de vários replicons variando de 100 a 3000 kb. Quatro destes replicons hibridizam com a sonda do gene 16S rDNA, sugerindo que A. brasilense possui vários cromossomos. O perfil eletroforético em campo pulsado do DNA intacto de Azospirillum brasilense estirpe FP2 obtido neste trabalho revelou dois novos replicons (de 745 e 785 kb), que não haviam sido descritos anteriormente, sugerindo que o tamanho genoma do Azospirillum brasilense estirpe FP2 é de aproximadamente 8,2 Mpb. O replicon de 630 Kb, denominado pAZ6, de A. brasilense é o menor da estirpe FP2 que hibridiza com a sonda do gene 16S rDNA. Neste estudo foi realizado o següenciamento parcial deste replicon. O pAZ6 foi isolado e purificado por eletroforese em campo pulsado utilizando o sistema CHEF-DRIII (Bio-Rad) e, posteriormente, o DNA extraído do gel de agarose foi amplificado usando o sistema TempliPhi (Amersham Biosciences), uma técnica não-PCR. O DNA amplificado foi utilizado para construir uma biblioteca genômica no vetor pCR4Blunt-TOPO. Os plasmídeos (1858) foram seqüenciados e as leituras de sequências montadas em contíguos pelos programas PHRED/PHRAP e, em seguida, os contíguos formados foram analisados utilizando os programas FramePlot e BlastP. Os resultados revelaram que o replicon pAZ6 apresenta genes envolvidos em processos biossintéticos (parede celular, de vitamina B12, lipídeos e nucleotídeos), metabolismo de piruvato, produção de energia, degradação de aminoácidos, metabolismo de carboidratos e de lipídeos, regulação transcricional, metabolismo nitrogenado, transporte de solutos e quimiotaxia. A categoria funcional mais representada entre os prováveis produtos gênicos do replicon foi a de transportadores. Os resultados sugerem fortemente que o replicon pAZ6 é parte essencial do genoma de A. brasilense estirpe FP2.

1. INTRODUÇÃO

Os procariotos provavelmente estiveram entre os primeiros organismos vivos da Terra e seu aparecimento precedeu o de animais e plantas em mais de 3 bilhões de anos (FRASER *et al.*, 2000). Estima-se que as espécies microbianas compreendem aproximadamente 50% da biomassa global (FRAZIER *et al.*, 2003). Embora invisíveis a olho nu, os procariotos são um componente essencial do biota da Terra: eles catalisam transformações únicas e indispensávies nos ciclos biogeoquímicos, produzem componentes importantes da atmosfera terrestre, afetam a produtividade do solo, a qualidade da água e o clima global (COPLEY, 2002), e representam uma grande porção da biodiversidade (WHITMAN *et al.*, 1998). Certamente, a capacidade deste planeta de sustentar a vida é dependente em grande parte da atividade destes organismos (FRAZIER *et al.*, 2003).

Menos de 1% das espécies microbianas já foram identificadas (WHITMAN *et al.*,1998), mas a diversidade das espécies descritas é extraordinária, mostrando adaptação a condições extremas de temperatura, pH, radiação, pressão e salinidade, e apresentando a capacidade de utilizar uma grande variedade de substâncias como fonte de energia, mesmo algumas consideradas resíduos tóxicos. Estes organismos podem produzir fontes de energia renováveis tais como biomassa, metano e hidrogênio (FRASER *et al.*, 2000; FRAZIER *et al.*, 2003).

A diversidade genética, fisiológica e metabólica dos procariotos é muito maior que aquela encontrada em plantas e animais (FRASER *et al.*, 2000). Avanços tecnológicos nas últimas décadas, particularmente na pesquisa genética, permitiu aprofundar os conhecimentos sobre os microrganismos ao nível mais básico, e questionar como as partes mais elementares dos microrganismos agem em conjunto para formar um organismo funcional (www.ornl.gov/microbialgenomes/benefits.html).

Um pré-requisito para o entendimento completo da biologia de um organismo é a determinação da seqüência completa do genoma (FLEISCHMANN *et al.*, 1995). Em 1995, foi publicada a primeira seqüência completa do DNA de um procarioto, a bactéria *Haemophilus influenzae* (FLEISCHMANN *et al.*, 1995). Desde então, mais de

60 seqüências completas do genoma foram publicadas, e o genoma de mais 350 espécies está sendo seqüenciado (www.tigr.org/tdb/mdb/).

Estes projetos têm produzido informações sobre o genoma de um número crescente de patógenos humanos como *Mycobacterium tuberculosis*, *Neisseria meningitidis*, *Pseudomonas aeruginosa* e várias espécies de *Chlamydia*. Também foram determinadas as seqüências genômicas de microrganismos considerados "modelo" como *Escherichia coli* e *Bacillus subtilis*; de espécies ambientais como *Thermotoga maritima*, *Deinococcus radiodurans*, e *Halobacterium* sp.; e, mais recentemente, os genomas dos microrganismos de importância agrícola como *Xylella fastidiosa*, *Pasteurella multocida*, *Buchnera* sp., *Mesorhizobium loti*, *Sinorhizobium meliloti*; e patógenos de animais (NELSON *et al.*, 2001).

O conhecimento da sequência completa do genoma não somente proporciona uma riqueza de dados, mas também informações específicas que não podem ser obtidas por outras vias. Isto inclui a identificação de genes que não compartilham similaridade com genes conhecidos nos bancos de dados, famílias de genes parálogos, genes que podem ser alvos de novas drogas (ARIGONI *et al.*, 1998) e genes inativos resultantes da evolução redutiva (ANDERSSON *et al.*, 1998). Informações importantes sobre a organização do genoma também são reveladas como o conteúdo de G+C e sua variação dentro do genoma - levando a identificação de regiões que podem ter sido adquiridas por transferência horizontal de genes -, a presença de elementos de repetição ou elementos de inserção, a descoberta de novas ilhas de patogenicidade em bactérias patogênicas, a identificação de novos operons, polaridade do genoma e a identificação da origem(ns) de replicação (FRANGEUL *et al.*,1999). Os "Projetos Genoma" também podem ser pontos de partida para projetos de reconstrução metabólica (SELKOV *et al.*, 1997) e análise funcional de todos os genes de um organismo (SMITH, 1997).

Logo após o surgimento das técnicas de seqüenciamento do DNA (SANGER, & COULSON, 1975; SANGER *et al.*,1982) e o desenvolvimento de novas tecnologias de computação (STADEN *et al.*,1982; STADEN *et al.*,1984), a estratégia de seqüenciamento "shot gun" foi introduzida, e tem permanecido o método de escolha

para a determinação do genoma total (LANDER *et al.*, 2001). O seqüenciamento e montagem de fragmentos aleatórios do DNA, elimina a necessidade de informações preliminares de mapeamento do genoma e, portanto, é aplicável a um grande conjunto de organismos cujos mapas genômicos não são conhecidos (FLEISCHMANN *et al.*, 1995).

A comparação entre os genomas completos de procariotos tem revelado que a densidade de genes nos microrganismos procariotos é de aproximadamente um gene por kilobase (FRASER et al., 2000). A maior surpresa tem sido o grande número de prováveis genes com função desconhecida. Por exemplo, o genoma do Archea Aeropyrum pernix contém mais de 1.500 ORF's (57% do total de genes), cujas seqüências similaridade já determinadas não possuem com outras (KAWARABAYASHI et al., 1999). Mais de 40% das 4.000 ORFs encontradas em Mycobacterium tuberculosis, uma das bactérias mais bem estudadas no século passado, também estão nesta categoria (COLE et al., 1998). De uma forma geral, nas sequências genômicas de procariotos pelo menos 25% dos genes são considerados "hipotéticos", e nenhuma função pode ser atribuída (DOOLITTLE, 2002). As ORF's "hipotéticas" podem ser colocadas dentro de duas categorias: aquelas que são encontradas em uma variedade de organismos, e que provavelmente codificam proteínas funcionais; e aquelas que são únicas para uma espécie ou estirpe particular (DOOLITTLE, 2002). Estas últimas, se não forem um artefato, indicam uma grande diversidade biológica entre os microrganismos e sugerem que a idéia de um organismo "modelo" no mundo microbiano pode não ser apropriada (FRASER et al., 2000).

1.1 ESTRUTURA GENÔMICA EM BACTÉRIAS

Os dois microrganismos mais utilizados na genética bacteriana. Escherichia coli e Bacillus subtilis, possuem um cromossomo circular de aproximadamente 4,5 Mpb. Como esses organismos são bastante distintos entre si, era razoável assumir que a forma e o número de cromossomos - ou seja, cromossomo único - seria constante entre as espécies bacterianas (OCHMAN, 2002). A ocorrência de plasmídeos em muitas espécies de bactérias ajudou a reforçar esta visão. Estes elementos "extracromossomais", pequenos e dispensáveis, são aleatoriamente distribuídos e têm um número de cópias variável, sugerindo que todos os genes necessários para a manutenção das funções vitais seriam codificados por um único cromossomo circular, presente obrigatoriamente em todos os membros de uma espécie (OCHMAN, 2002).

A eletroforese em campo pulsado (PFGE), surgida em 1984, permitiu o isolamento e a análise de grandes moléculas de DNA (SCHWARTZ & CANTOR, 1984; LEE et al., 1996) e tem sido usada para analisar os genomas de uma variedade de bactérias que antes não podiam ser manipulados pelas metodologias da genética clássica (HOLLOWAY, 1992). Utilizando-se desta nova metodologia, foi mostrado que o genoma bacteriano é altamente variável tanto em tamanho como em forma. O menor genoma conhecido é de Mycoplasma genitalium (classe Firmiscutes) com 0,58 Mpb (FRASER et al., 1995) e o maior é o de Stigmatella erecta (delta-Proteobactéria) com 9,9 Mpb (LEE et al., 1996). É interessante notar que o tamanho do genoma parece refletir o estilo de vida da bactéria: bactérias com genomas pequenos são especialistas, como os parasitas intracelulares obrigatórios ou microrganismos que requerem condições muito especializadas; aquelas com genomas grandes apresentam uma maior diversidade metabolica ou apresentam formas de desenvolvimento e/ou diferenciação celular, como esporulação (MORENO, 1998). Esta observação sugere que a necessidade de sobreviver em ambientes diferentes requer uma extensa maquinaria molecular de grande flexibilidade metabólica para competir e se adaptar rapidamente às variações ambientais (CASJENS, 1998).

Além da diversidade de tamanhos, o genoma bacteriano pode ser constituído por múltiplos replicons, que podem ser circulares e/ou lineares (CHENG & LESSIE, 1994; RODLEY et al., 1995; CASJENS, 1998) (TABELA 1). A presença de dois replicons foi relatada em Rhodobacter sphaeroides, Brucella melitensis 16M e Leptospira interrogans enquanto três replicons compõem os genomas de Sinorhizobium meliloti e Burkholderia cepacia. O primeiro cromossomo linear foi descrito em Borrelia burgdorferi (BARIL et al., 1989; FERDOWS & BARBOUR, 1989), e nos organismos gram-positivos Streptomyces lividans (LIN et al., 1993) e Rhodococcus fascians (CRESPI et al., 1992). Os telômeros dos replicons lineares de Borrelia e Streptomyces possuem duas formas estruturais bastante distintas. As extremidades do replicon de Borrelia são grampos fechados covalentemente (CASJENS et al., 1997: HINNERBUSCH & BARBOUR, 1991) e apresentam següências invertidas de 20 a 30 pb nas duas extremidades. O processo de replicação dos telômeros ainda não foi esclarecido (MARCONI et al., 1996; CASJENS et al., 1997). Por outro lado, os telômeros de Streptomyces são extremidades abertas e apresentam proteínas específicas ligadas covalentemente às extremidades 5' (CHEN, 1996; SAKAGUCHI, 1990).

A presença destes replicons adicionais gerou dúvidas quanto a sua classificação como cromossomos ou megaplasmídeos (OCHMAN, 2002). Os cromossomos geralmente são as maiores moléculas, depositárias do material genético essencial da célula, portanto, possuem genes com funções essenciais, como operons codificando para RNA ribossomal (rRNA). Os plasmídeos, ainda que possam ter tamanhos que lhes confiram a classificação de "megaplasmídeos" (ou seja, com tamanho na ordem de milhões de pares de bases), não contêm genes essenciais e, portanto, podem ser perdidos (OCHMAN, 2002).

Espécies	Classificação*	Tamanho(kb)	Forma	Cópias de rDNA
Agrobacterium tumefacines	Cromossomo	2842	Circular	2
	Cromossomo	2057	Linear	2
	Plasmídeo	543	Circular	0
	Plasmídeo	214	Circular	0
Borrelia burgdorfei	Cromossomo	911	Linear	1
	Plasmídeo (n=11)	9 - 54	Circular/Linear	0
Brucella melitensis	Cromossomo	2117	Circular	2
	Cromossomo	1178	Circular	1
Clostridium acetobutylıcum	Cromossomo	3941	Circular	11
	Megaplasmideo	192	Circular	0
Demococcus radiodurans	Cromossomo	2649	Circular	3
	Cromossomo	412	Circular	0
	Megaplasmídeo	177	Circular	0
	Plasmideo	46	Circular	0
Ralstonia solanacearum	Cromossomo	3716	Circular	3
	Megaplasmídeo	2095	Circular	1
Salmonella typhi	Cromossomo	4809	Circular	7
	Plasmídeo	218	Circular	0
	Plasmídeo	107	Circular	0
Sinorhizobium meliloti	Cromossomo	3654	Circular	3
	Megaplasmídeo	1683	Circular	0
	Megaplasmídeo	1354	Circular	0
Vibrio cholerae	Cromossomo	2941	Circular	8
	Cromossomo	1072	Circular	0
Yersina pestis	Cromossomo	4654	Circular	6
1	Plasmídeo (n=3)	10 - 96	Circular	0

TABELA	1.	COMPOS	IÇÃO D	E ALGUNS	GEN	IOMAS	BACTERIANOS	COMPLETA-
		MENTE	SEQÜI	ENCIADOS	E	COM	MÚLTIPLOS	ELEMENTOS
		HEREDI	TÁRIOS.					

* Classificação segundo os autores. Agrobacterium tumefacines (WOOD et al,2001), Borrelia burgdorfei (FRASER et al, 1997), Brucella melitensis (DELVECCHIO et al, 2002), Clostridium acetobutylicum (NOLLING et al,2001), Deinococcus radiodurans (WHITE et al, 1999), Ralstonia solanacearum (SALANOUBAT et al, 2002), Salmonella typhi (PARKHILL et al, 2001), Sinorhizobium meliloti (GALIBERT et al, 2001), Vibrio cholerae (HEIDELBERG et al, 2000), Yersina pestis (PARKHILL et al, 2001). n = número de plasmídeos encontrados.

Fonte: OCHMAN, H., 2002

Várias hipóteses surgiram para explicar os eventos que deram origem a bactérias com mais de um cromossomo. Algumas destas hipóteses são: a quebra de um cromossomo em diferentes replicons, divisão desigual do cromossomo, mutação independente de cromossomos multicópias, transformação de plasmídeo em cromossomo pela aquisição de genes essenciais a partir do cromossomo natural ou por transferência horizontal de genes (KOLSTO, 1997). O seqüenciamento completo do genoma de bactérias, poderá auxiliar no esclarecimento de quais eventos poderiam contribuir para a estrutura e plasticidade dos genomas bacterianos (MARTIN-DIDONET, 2001).

1.2 GENOMA DE PROTEOBACTÉRIAS

A classe das Proteobactérias têm recebido atenção especial não só pela grande quantidade de organismos estudados, como também pela grande diversidade na estrutura e tamanho do genoma. As Proteobactérias são atualmente divididas em 5 subclasses (α , β , γ , δ , ϵ) (MORENO, 1998).

Dentro da subclasse α estão bactérias fotoautotróficas e quimioautotróficas, bem como espécies associadas a plantas e animais. Uma das características mais interessantes das alfa-proteobactérias é a presença de um ou mais cromossomos com diferentes topologias (linear ou circular) (MORENO, 1998). O tamanho do genoma varia de 1,1 Mpb em *Rickettsia prowazekii*, parasita intracelular obrigatório (ANDERSSON *et al.*, 1998) até 9,1 Mpb em *Bradyrhizobium japonicum*, bactéria simbiótica fixadora de nitrogênio (KANEKO, 2002), com a presença de elememtos extracromossomais como plasmídeos e megaplasmídeos (MORENO, 1998). A versatilidade metabólica das bactérias deste grupo é refletida no tamanho do genoma (MORENO, 1998).

1.2.1 GENOMA DE Sinorhizobium meliloti

O diazotrofo *Sinorhizobium meliloti* (α-Proteobactéria) estabelece simbiose com raízes de leguminosas, além de ser intimamente relacionado com bactérias patogênicas de plantas e animais, incluindo *Agrobacterium* e *Brucella* (GALIBERT *et al.*, 2001).

O genoma de *S. meliloti* consiste de três replicons: um replicon grande de 3,65 Mb e dois replicons menores, pSymA e pSymB, de 1,35 e 1,68 Mb, respectivamente (GALIBERT *et al.*, 2001). Em um genoma total de 6,7 Mb, foram encontrados 6.204 genes, sendo 59,7% dos genes com função preditas e 8,2% não apresentaram similaridade com seqüências de nucleotídeos depositados nos bancos de dados. A proporção de genes com função desconhecida foi maior nos megaplasmídeos que no cromossomo: 11,5% no pSymA e 12,3% no pSymB comparados com a porcentagem total no cromossomo de 5% (GALIBERT *et al.*, 2001).

Uma importante característica do genoma de S. meliloti é o número de genes que codificam para sistemas de transporte do tipo ABC. Estes sistemas contêm uma proteína de ligação ao ATP, uma ou duas proteínas integrais de membrana e, nos sistemas de importação, pode fazer parte, uma proteína periplasmática de ligação ao soluto. Estes genes geralmente estão dispostos como um operon (FATH & KOLTER, 1993). Dos 430 genes de sistemas de transporte ABC encontrados em todo o genoma, 235 genes estão localizados no megaplasmídeo pSymB, onde quase metade dos 64 diferentes sistemas de transporte ABC são para o transporte de açúcares (HAAS et al., 1998). Outros sistemas de transporte encontrados incluem ions Fe^{+2/+3} (HONEYCUTT et al., 1993), aminoácidos (CAPELA et al., 2001), peptídeos e oligopeptídeos (CAPELA et al., 2001), espermidina/putrescina (ALLARDET-SERVENT et al., 1993), sulfato, fosfato, colina, glicerol-3-fosfato, rizopina e taurina (JUMAS-BILAK et al., 1998). Além dos transportadores ABC, foram encontrados no pSymB prováveis genes codificando para permeases de nitrato, sulfato e xantina/uracila e genes que codificam para proteínas de efluxo transmembrana, envolvidas na exportação de compostos tóxicos da bactéria (FINAN et al., 2001).

As bactérias gram-negativas apresentam conjuntos complexos de polissacarídeos na superficie, incluindo lipopolissacarídeos (LPS) (SCHNAITMAN & KLENA, 1993), polissacarídeos capsulares (CPS) (WHITFIELD & ROBERTS, 1999), exopolissacarídeos (EPS) (STEVENSON *et al.*, 1996) e glucanas periplasmáticas. Os polissacarídeos de superficie do *S. meliloti*, cuja maquinaria de síntese são codificados principalmente pelo cromossomo e pelo pSymB, são importantes para o sucesso da infecção da planta, possivelmente por suprimir a resposta de defesa da planta (VIPREY *et al.*, 2000). Os genes envolvidos na biossíntese e exportação destes carboidratos estão geralmente agrupados. Os grupos *exo/exs* e *exp* que codificam para proteínas envolvidas na síntese do EPS succinilglucana (EPS I) (GLUCKSMANN *et al.*, 1993) e galactoglucana (EPS II) (BECKER *et al.*, 1997) foram mapeados no pSymB. As análises da seqüência do pSymB revelaram que quase 12% dos genes estão envolvidos na síntese de polissacarídeos (GALIBERT *et al.*, 2001). A maioria destes genes estão organizados em 11 grupos; a existência de 9 destes grupos, variando de 5 a 42 kb, não era conhecida anteriormente (FINAN *et al.*, 2001).

Os genes de nodulação (*nod*), necessários para a síntese e exportação dos fatores Nod, estão localizados principalmente no replicon pSymA e somente poucos genes envolvidos nesse processo foram encontrados no pSymB (*nodP2, nodQ2*) (FINAN *et al.*, 2001). Cada megaplasmídeo apresenta uma cópia do operon *nodPQ* (99% de identidade entre si) e que estão envolvidos na ativação do sulfato para a sulfatação dos fatores Nod (GALIBERT *et al.*, 2001).

O metabolismo do nitrogênio é uma característica importante do genoma do *S. meliloti*, e os genes envolvidos neste metabolismo foram encontrados principalmente no megaplasmídeo pSymA. Enquanto a síntese e atividade da nitrogenase necessita cerca de 20 genes em *Klebsiella pneumoniae*, somente nove genes *nif* foram encontrados em *S. meliloti* (*nifA*, *nifB*, *nifHDKE*, *nifX*, *nifN* e *nifS*). Exceto para um provável ortólogo de *nifS* no cromossomo e um provável gene *nifV* (CAPELA et al., 2001), todos os outros genes estão localizados no pSymA (GALIBERT et al., 2001).

Além dos genes de fixação do nitrogênio, pSymA apresenta um gene para a assimilação de amônio (glutamato desidrogenase - gdhA), um conjunto completo de

genes necessários para a desnitrificação (*nor*, *nos*, *nap* e *nir*) e genes para o transporte de nitrato (*nrtAB*). No cromossomo encontram-se os genes *ntrBC*, *glnB*, *glnA* e *glnT*, genes para uma alanina desidrogenase, um transportador de amônio *amt*, e genes que codificam para as proteínas regulatórias *ntrXY*, *glnE*, *glnK* e *glnD*; para enzima GOGAT glutamato sintase e para três enzimas glutaminas sintetases homólogas, anteriormente desconhecidas. O pSymB codifica uma nitrato redutase (*narB*), dois transportadores de nitrato (*nrtA*, SMb20436) e uma única glutamina sintetase *glnII* (GALIBERT *et al.*, 2001).

S. meliloti é uma bactéria aeróbia que deve gerar altos níveis de energia para suportar a fixação do nitrogênio em um ambiente com baixo teor de oxigênio como no nódulo. Uma enzima citocromo c oxidase do tipo cbb₃, com alta afinidade pelo oxigênio, é codificado por dois conjuntos de genes fixNOQP duplicados no pSymA. Tanto o pSymA quanto o cromossomo, apresentam um grande grupo do gene que codifica para a NADH-ubiquinona desidrogenase. (PREISIG et al.,1993; PREISIG et al.,1996).

Cada replicom apresenta um perfil de genes regulatórios distintos. Genes da família LysR (86 membros) predominam no pSymA. Genes codificando proteínas reguladoras do tipo GntR são mais freqüentemente encontrados no megaplasmídeo pSymB, e da família AsnC são mais comum no cromossomo (GALIBERT *et al.*, 2001). No megaplasmídeo pSymB foi encontrada a única cópia do gene para o tRNA da arginina e os genes *minCDE* envolvidos na divisão celular (FINAN *et al.*, 2001).

A determinação da seqüência genômica do *S. meliloti* 1021 mostrou sua arquitetura composta de três replicons com características estruturais e funcionais distintas. O cromossomo de *S. meliloti* possui as características do de uma bactéria heterotrófica aeróbica típica. A aquisição do pSymB estendeu consideravelmente a sua capacidade metabólica, por permitir metabolizar uma grande variedade de pequenos compostos encontrados no solo ou na rizosfera da planta. A maior diversidade de polissacarídeos dado pelo pSymB também pode ter melhorado significativamente o potencial de colonização deste organismo. Finalmente, a aquisição do pSymA levou a emergência da nodulação, bem como a capacidade da bactéria para colonizar o nódulo

em um ambiente com baixo teor de oxigênio. O pSymA também expandiu a capacidade de metabolizar compostos nitrogenados sob uma variedade de formas químicas, incluindo dinitrogênio molecular (GALIBERT *et al.*, 2001).

1.2.2 GENOMA do Agrobacterium tumefaciens

Agrobacterium tumefaciens é uma α-Proteobactéria da família Rhizobiacea, geralmente encontrada no solo e é o agente etiológico de um tumor em plantas. *Agrobacterium tumefaciens* infecta mais de 90 famílias de plantas dicotiledôneas, resultando em grandes perdas agronômicas (WOOD *et al.*, 2001). O tumor resulta da transferência, integração e expressão de um conjunto de genes contidos no DNA-T e localizados no plasmídeo indutor de tumor (Ti). A expressão destes genes conduzem a biossíntese de hormônios de crescimento bem como de opinas, que são uma fonte de nutrientes para a bactéria (ZUPAN *et al.*, 2000). O processamento e transferência do DNA-T é mediado pelos genes de virulência (*vir*) do plasmídeo Ti, e vários determinantes de virulência inicialmente caracterizados em *A. tumefaciens*, têm sido encontrados em simbiontes de plantas e em patógenos animais (CHRISTIE, 2001; INON *et al.*, 1998). Os genes do DNA-T podem ser trocados por qualquer seqüência de DNA, fazendo do *A. tumefaciens* um veículo ideal para a transferência de genes e uma ferramenta essencial para as pesquisas de biotecnologia em plantas (GOODNER *et al.*, 2001).

O genoma de 5,67 Mb de *A. tumefaciens* estirpe C58 é composto por quatro replicons: um cromossomo circular, um cromossomo linear e os plasmídeos pAtC58 e pTiC58 (ALLENDERT-SERVENT *et al.*, 1993; GOODNER *et al.*, 1999). O genoma contém 5.419 genes: a 3.475 (64%) destes foi atribuída uma provável função, 1.236 (22%) genes são proteínas conservadas hipotéticas e 708 genes são proteínas hipotéticas, ou seja, com nenhuma similaridade com seqüências nos bancos de dados (WOOD *et al.*, 2001). O conteúdo G+C do genoma é 58%, mas o plasmídeo TiC58 tem duas regiões com conteúdo G+C diferente: o DNA-T (46%) e a região *vir* (54%). Um conteúdo reduzido de G+C (53%) foi também observado em um segmento de 24

kb do pAtC58 (ilha AT) (WOOD *et al.*, 2001). O genoma contém 53 tRNAs que representam todos os aminoácidos. Estes tRNAs estão distribuídos desigualmente entre os cromossomos circular e linear, sendo os tRNA correspondentes aos codons da alanina, glutamina e valina encontrados apenas no cromossomo linear (WOOD *et al.*, 2001). Cada cromossomo apresenta dois grupos dos genes para os RNA's ribossomais, e os genes envolvidos na maioria dos processos essenciais estão localizados no cromossomo circular (TATUSOV, 2001).

O cromossomo linear apresenta as extremidades covalentemente ligadas, possuindo aparentemente estruturas em forma de grampo nos telômeros (GOODNER *et al.*, 2001). Proteínas associadas com a manutenção da extremidade linear, como as telomerases, estão ausentes em *A. tumefaciens*, mas uma característica das terminações do replicon linear é a presença de elementos IS próximos a cada extremidade (WOOD *et al.*, 2001).

Foram encontrados 1.882 genes que codificam para proteínas no replicon linear, incluíndo aquelas que codificam proteínas ribossomais, de replicação e enzimas de vinte e uma vias metabólicas completas. A presença destes genes confirma a designação de cromossomo para este replicon. Entretanto, características tipicamente plasmidiais foram encontradas neste replicon, como genes provavelmente envolvidos em conjugação (*traA*, *mobC* e *traG*) e um operon *repABC*, que codifica para proteínas da família de replicadores tipo-RepABC de plasmideos circulares, localizado próximo ao centro do cromossomo linear. Se verificado experimentalmente, este mecanismo de replicação provará ser único entre os replicons lineares conhecidos (WOOD et al., 2001). O cromossomo circular apresenta uma origem de replicação (*oriC*) similar a *Cori* de *Caulobacter crescentus* (BRASSINGA *et al.*, 2001).

A replicação dos plasmídeos pTiC58 e pAtC58 é mediada por um sistema tipo-RepABC comumente encontrado nos plasmídeos de Rhizobiacea. Este sistema, apresenta duas proteínas de segregação (RepA e RepB) e uma proteína de início de replicação que se liga à origem (RepC) (LI & FARRAD, 2000). Os plasmídeos codificam para toda a maquinaria necessária para conjugação e não apresentam genes essenciais (WOOD *et al.*, 2001). Genes de proteínas transportadoras constituem 15% do genoma de *A. tumefaciens*. Destes, 87% estão localizados nos cromossomos. Estes sistemas conferem uma ampla capacidade para o transporte de solutos encontrados na rizosfera, incluindo açúcares, aminoácidos e peptídeos. Como em *S. meliloti* e *M. loti, A. tumefaciens* apresenta um grande número de transportadores ABC, constituindo 60% de todos os transportadores. A preferência por transportadores ABC em *A. tumefaciens* pode refletir uma necessidade por sistemas de captação de alta afinidade, para a aquisição de nutrientes em um ambiente altamente competitivo como o solo e a rizosfera (WOOD *et al.*, 2001).

Os genes envolvidos na transformação da planta e formação do tumor estão localizados nos quatro elementos genéticos. O cromossomo circular abriga os genes *chvAB* necessários para a ligação da bactéria à planta; os genes *chvGI, chvE* e *ros* envolvidos na regulação da expressão dos genes *vir* do plasmídeo pTi (ZUPAN *et al.*, 2000). O cromossomo linear abriga o gene *exoC* (*pgm*) necessário para a síntese dos polissacarídeos β -1,2 glucana extracelular e o succinilglucana; e os genes para a síntese de celulose (*cel*) (ZUPAN *et al.*, 2000). O plasmídeo pAtC58 contém os genes envolvidos no ancoramento (*att*) necessários para a ligação inicial e específica da bactéria à planta (ZUPAN *et al.*, 2000; GOODNER *et al.*, 2001).

Os genes que codificam para as vias metabólicas estão dispersos entre os quatro replicons. Ao contrário do que ocorre em *E. coli*, a maioria dos genes destas vias não formam grupos, o que sugere que não formam operons. Foram identificadas vias para a síntese de todos os aminoácidos e de coenzimas. As vias para o metabolismo de energia incluem glicólise, ciclo do ácido tricarboxílico e Entner-Doudoroff. Foram encontradas as vias para a utilização de metabólitos da planta, tipicamente encontrados na rizosfera, que incluem D-glucose, D-frutose, sacarose, D-ribose, D-xilose, D-xilose, D-xilose e lactose bem como mio-inositol, hidantoína, uréia e glicerol (WOOD *et al.*, 2001).

Muitos componentes do metabolismo do nitrogênio são conservados entre A. tumefaciens e os simbiontes fixadores de nitrogênio S. meliloti e M. loti, como os componentes do sistema de regulação do metabolismo de fontes alternativas de nitrogênio (Ntr): *ntrBC*, *ntrA*, *glnE*, *glnD*, *glnB e glnK*. *Agrobacterium tumefaciens* apresenta sete genes para a glutamina sintetase (GS). A presença de múltiplos genes GS pode estar relacionada com a necessidade do *Agrobacterium* de altas concentrações de glutamato para um crescimento ótimo (WOOD *et al.*, 2001). Foram encontrados genes para o transporte de nitrato e vários genes de desnitrificação no cromossomo linear (GOODNER *et al.*, 2001).

As análises do genoma revelaram uma extensa similaridade entre os cromossomos circulares de A. tumefaciens e do simbionte S. meliloti, que apoia a hipótese que estas bactérias originaram-se de um ancestral comum. Modelos de evolução bacteriana sugerem que a aquisição diferencial e/ou a perda de genes em organismos que habitam um mesmo ambiente, permite a divergência entre os estilos de vida simbiótico ou patogênico (OCHMAN, 2001). A aquisição de tais elementos é aparente em A. tumefaciens e S. meliloti. Os genes nod de S. meliloti e os genes vir e DNA-T de A. tumefaciens, possuem um conteúdo de G+C e o uso de códon distinto do resto do genoma, sugerindo uma aquisição evolucionária. No caso do DNA-T, o reduzido conteúdo de G+C pode facilitar a sua expressão na planta hospedeira, onde o baixo conteúdo G+C é comum (WOOD et al., 2001). Além disso, nenhum DNA-T e poucos genes vir de A. tumefaciens tem ortólogos em S. meliloti, e os genes nod não foram encontrados em Agrobacterium. A seleção diferencial e a manutenção de genes adquiridos horizontalmente provavelmente levam a divergência entre os estados patogênicos e simbióticos. Assim, estes organismos proporcionam um sistema modelo para futuras investigações da divergência evolutiva entre patógenos e simbiontes (WOOD et al., 2001).

A disponibilidade das seqüências genômicas das α -Proteobactérias permitem estudos comparativos que tem contribuído enormemente para o entendimento da história evolutiva deste grupo de microrganismos e a manutenção de seus genomas com múltiplos cromossomos. De forma especial, estes estudos permitiram inferir o aparecimento dos mecanismos de simbiose e de patogenicidade de células animais e vegetais (GALIBERT *et al.*, 2001).

1.2.3 GENOMA DE OUTRAS PROTEOBACTÉRIAS

A bactéria *Brucella melitensis* é um patógeno intracelular facultativo e o causador da brucelose em animais e da febre de Malta em humanos. *Brucella* pertence ao grupo das α -Proteobactérias e apresenta uma relação filogenética muito próxima com *Agrobacterium, Rickettsia, Rhizobium* e *Rhodobacter* (KE *et al.*, 2000). O tamanho total do genoma de *B. melitensis* 16M é de 3,29 Mb dividido em dois cromossomos circulares de 2,12 e 1,18 Mb. O conteúdo de G+C é de 57% e nenhum plasmídeo foi encontrado. No cromossomo I (2,12 Mb) foram encontradas 2060 ORFs, e no cromossomo II (1,18 Mb) 1138 ORFs.

O genoma deste microrganismo apresenta três operons rRNA, dois no cromossomo I, separados por aproximadamente 200 kb, e o terceiro no cromossomo II. Todos os três operons estão organizados da mesma maneira: 16S, *ile* tRNA, *ala* tRNA, 23S, 5S e *met* tRNA (DELVECCHIO *et al.*, 2002).

Todas as 33 proteinas da subunidade maior do ribossoma e a maioria das da subunidade menor são codificadas no cromossomo I, enquanto a proteína S21 (subunidade menor) está codificada no cromossomo II (DELVECCHIO *et al.*, 2002). A mesma origem de replicação (*ori*) foi encontrada nos dois cromossomos, e são similares aquelas das α -Proteobactérias. Os genes essenciais, incluíndo aqueles envolvidos na replicação do DNA, transcrição, tradução, metabolismo central e biossíntese da parede celular estão distribuídos em ambos os cromossomos. Os genes envolvidos na virulência, como as adesinas, invasinas e hemolisinas, estão localizados no cromossomo I (DELVECCHIO *et al.*, 2002).

A seqüência completa do genoma da γ -Proteobactéria *Vibrio cholerae* EI Tor N16961, o agente etiológico do cólera, apresentou 4 Mpb distribuídos em dois cromossomos circulares de 3 Mpb (cromossomo 1) e 1 Mpb (cromossomo 2), com uma conteúdo médio de G+C de 46,9% e 47,7%, respectivamente. Foi observada uma grande assimetria na distribuição dos genes reconhecidos como essenciais para o crescimento e virulência entre os dois cromossomos. Oito operons que codificam para o rRNA ribossomal foram localizados no cromossomo 1 e nenhum *no cro*mossomo 2.

Além disso, um número maior de genes que codificam para proteínas envolvidas na replicação, reparo e transcrição do DNA, biossíntese da parede celular e uma variedade de vias biossintéticas e catabólicas são codificadas pelo cromossomo 1. Similarmente, a maioria dos genes considerados essenciais para a patogenicidade bacteriana como aqueles que codificam para a toxina do cólera, enzimas para síntese de lipopolissacarídeos e para a maquinaria de secreção de proteínas extracelular, estão também localizados no cromossomo 1. O cromossomo 2 apresenta uma quantidade maior de genes hipotéticos (59%) quando comparado com o cromossomo 1 (42%). A localização destes genes hipotéticos no cromossomo 2 é concentrada na chamada ilha de integração, que contém 216 ORFs (125,3 kb). Entre os genes que puderam ser identificados estão três que codificam para resistência a drogas, várias enzimas do metabolismo de DNA (MutT, transferases e uma integrase), genes de virulência (hemaglutinina e lipoproteínas) e três genes cujos produtos estão relacionados com proteínas plasmidiais na seleção de células hospedeiras (*higA*, *higB* e *doc*) (DELVECCHIO *et al.*, 2002).

A bactéria gram-negativa *Ralstonia solanacearum* é um patógeno de planta encontrado no solo e que infecta naturalmente as raízes das plantas. Esta bactéria exibe um forte tropismo tecido específico dentro do hospedeiro, invadindo e multiplicandose rapidamente nos vasos do xilema. Esta bactéria, pertencente ao grupo das β -Proteobactérias, apresenta uma ampla faixa de hospedeiros (mais de 200 espécies vegetais), o que a torna um sistema modelo para o estudo dos determinantes moleculares que governam a patogenicidade (SALANOUBAT *et al.*, 2002).

O genoma da *Ralstonia solanacearum* GMI1000 está organizado em duas moléculas circulares: um cromossomo de 3,7 Mpb e um megaplasmídeo de 2,09 Mpb, produzindo um genoma total de 5,8 Mpb. As duas moléculas apresentam conteúdo de G+C quase idêntico, 67,04% para o cromossomo e 66,86% para o megaplasmídeo (SALANOUBAT *et al.*, 2002).

O replicon maior apresenta a origem de replicação característica de cromossomos bacterianos, enquanto o menor apresenta uma origem de replicação característica de plasmídeos. No cromossomo são encontrados todos os genes que codificam para as funções essenciais. O megaplasmídeo apresenta vários genes metabolicamente essenciais que também estão presentes no cromossomo. Isto inclui uma cópia do operon rDNA com dois genes tRNA, um gene que codifica para a subunidade α da DNA polimerase III e um gene que codifica para um fator de alongamento. Várias enzimas que controlam o metabolismo primário, incluindo a biossíntese de aminoácidos e cofatores, são codificadas unicamento pelo megaplasmídeo. No megaplasmídeo também estão localizados todos os genes *hrp*, necessários para causar doença nas plantas, os constituintes do flagelo e a maioria dos genes responsáveis pela síntese de polissacarídeos (SALANOUBAT *et al.*, 2002). A análise dos genes presentes no megaplasmídeo sugerem que este replicom deve ser necessário para a adaptação a várias condições ambientais, e também é importante para a manutenção de funções celulares básicas da bactéria (SALANOUBAT *et al.*, 2002).

1.3 Azospirillum

As primeiras espécies de Azospirillum foram isoladas por BEIJERINCK (1925) e denominados Spirillum lipoferum. Em 1963, Spirullum lipoferum foi redescoberto por BECKING em solos da África, mas sua importância agronômica surgiu quando DÖBEREINER e DAY (1976) descreveram sua associação com gramíneas de várias origens geográficas. TARRAND et al. (1978) fizeram um estudo taxonômico do grupo S. lipoferum e propuseram, baseado em características morfológicas, fisiológicas e homologia de DNA, a denominação de Azospirillum para um novo gênero. Atualmente existem sete espécies de Azospirillum descritas: A. brasilense (TARRAND et al.,1978), A. lipoferum (TARRAND et al.,1978), A. irakense (KHAMMAS et al., 1989), A. halopraeferens (REINHOLD et al.,1987), A. largimobile (DEKHIL et al., 1997) e A. doebereinerae (ECKERT et al., 2001).

As bactérias do gênero *Azospirillum* são organismos fixadores de nitrogênio encontrados no solo e também em associação endofítica com raízes de gramíneas, inclusive com cereais de interesse agrícola (MAGALHÃES *et al.*, 1979; DÖBEREINER, 1992; BALDANI *et al.*,1997). Devido a sua capacidade de promover

o crescimento das plantas, o gênero *Azospirillum* spp. tem sido classificado como rizobactérias promotoras de crescimento de plantas (BASHAN *et al*, 1990).

Azospirillum são bacilos curvos gram-negativos, apresentam uma grande versatilidade quanto a utilização de fontes de carbono e nitrogênio, o que facilita a sua adaptação no ambiente competitivo da rizosfera. Estes microrganismos utilizam preferencialmente ácidos orgânicos (malato, lactato) como fonte de carbono. Amônio, nitrato, nitrito, aminoácidos e nitrogênio molecular servem como fonte de nitrogênio (DÖBEREINER & PEDROSA, 1987). Em condições desfavoráveis, tais como dessecamento ou limitação de nutrientes, *A. lipoferum* e *A. brasilense* podem converter-se em uma forma de "cisto" denominada forma C (DÖBEREINER & PEDROSA, 1987). Esta mudança morfológica é acompanhada pelo desenvolvimento de uma camada externa de polissacarídeos e pelo acúmulo de grânulos de poli-β-hidroxibutirato, que podem servir como fonte de carbono e energia (TAL & OKON, 1985). O conteúdo percentual de G+C característico deste gênero é 67 - 70% (OKON, 1985).

As bactérias do gênero Azospirillum são móveis, apresentam um flagelo polar durante o crescimento em meio líquido e vários flagelos laterais são induzidos durante o crescimento em meio sólido (DÖBEREINER & PEDROSA, 1987). As espécies de *Azospirillum* exibem quimiotaxia positiva para ácidos orgânicos, açúcares, aminoácidos e compostos aromáticos (LOPES-DE-VITORIA *et al.*, 1994) e também para os exsudatos das raízes (HEINRICH & HESS, 1985). Uma outra característica de *Azospirillum* spp. é o movimento em direção a concentrações ótimas de oxigênio, fenômeno denominado aerotaxia (BARAK *et al.*, 1982). Esta capacidade permite o deslocamento das bactérias para regiões onde a concentração de oxigênio dissolvido é adequada para a fixação do nitrogênio (4 µmol/litro de O₂ dissolvido) (ZHULIN *et al.*, 1996).

Em certas condições ambientais e de solo, *Azospirillum* spp. pode influenciar positivamente o crescimento, produtividade e o conteúdo de nitrogênio da planta associada (PATRIQUIN *et al.*, 1983). Este efeito estimulatório tem sido atribuído a vários mecanismos, incluindo fixação biológica do nitrogênio, produção de auxinas e

um aumento da captação de sais minerais pelas raízes da planta (BASHAN, 1990). VLASSAK e REYNDERS (1978) descobriram que muitas estirpes de *Azospirillum* spp. produziam quantidades substanciais de auxinas quando o aminoácido triptofano era adicionado à cultura bacteriana. Sob condições similares, TIEN *et al.* (1979) relataram que ácido indol-acético (AIA), giberilina e substâncias similares a citocininas eram produzidas por *Azospirillum brasilense*. Quando esta cultura foi inoculada nas raízes de milheto (*Pennisetum americanum*) foi observado modificações na morfologia das raízes. Quase todas as raízes laterais foram densamente cobertas por pêlos radiculares e o número de raízes laterais foi aumentado.

Azospirillum spp. coloniza predominantemente a superficie das raízes e somente poucas linhagens são capazes de colonizar internamente a planta (PATRIQUIN, 1983; RAMOS *et al.*, 2002). RAMOS *et al* (2002) monitoraram o padrão de colonização da superficie das raízes de trigo por *A. brasilense* estirpe FP2, usando uma construção que expressa constitutivamente os genes *gusA* e *gfp*, e mostraramu que os primeiros sítios de colonização são os pontos de emergência das raízes laterais e a região dos pêlos radiculares.

1.3.1 Azospirillum brasilense

Azospirillum brasilense é a espécie melhor estudada do gênero (DÖBEREINER & PEDROSA, 1987). Apresenta características fisiológicas próprias: não é capaz de usar glucose como única fonte de carbono para o crescimento em meio semi-sólido sem nitrogênio e nem apresenta a capacidade de fermentar açúcares como ribose, manitol e sorbitol. Não requer vitaminas para o crescimento. Em meio semi-sólido livre de nitrogênio e contendo malato como fonte de carbono acrescido de 0,005% de extrato de levedura, as células de *Azospirillum brasilense* tendem a ser curtas, vibrióides, com aproximadamente 1µm de diâmetro e móveis. Células em forma de S podem ocorrer em culturas velhas. A estirpe de referência é a Sp7 (ATCC29145) (TARRAND *et al.*, 1978). As colônias podem apresentar uma coloração rósea em meio NFbHP, e há alguns mutantes espontâneos com coloração avermelhada que ocorrem principalmente na estirpe Sp7 (DÖBEREINER & PEDROSA, 1987).

1.3.2 GENÉTICA DO Azospirillum brasilense

Apesar de muitas características ecológicas e fisiológicas do Azospirillum spp. serem já conhecidas, informações sobre os mecanismos envolvidos na interação planta-bactéria são relativamente escassos (VANDE BROEK & VANDERLEYDEN, 1995). Esta interação planta-bactéria somente pode ser bem sucedida se a população bacteriana alcançar um número substancial no sistema radicular da planta (STEENHOUDT & VANDERLEYDEN, 2000). Na rizosfera, forma-se um gradiente de nutrientes, gerado pelo exsudato das raízes, que diminui a partir da epiderme em direção ao solo adjacente. O isolamento de genes de Azospirillum brasilense induzidos pela planta, descobertos a partir das análises dos padrões de proteínas da bactéria crescidas na presença e na ausência de exsudatos, resultou na identificação de uma proteína ácida de 40 kDa, denominada de SbpA (VAN BASTELAERE et al., 1999). Esta proteína é muito similar a proteína periplasmática ChvE identificada em Agrobacterium tumefaciens e envolvida no aumento da virulência. Aparentemente, a proteína SbpA está envolvida na captação de D-galactose e na quimiotaxia do A. brasilense para vários açúcares (D-frutose, L-arabinose e D-galactose) (VAN BASTELAERE et al., 1999). Estes resultados sugerem que os passos iniciais para a interação com a planta são conservados entre diferentes bactérias (STEENHOUDT & VANDERLEYDEN, 2000). A quimiotaxia não requer somente substâncias quimioatrativas, mas também depende da motilidade da bactéria. Genes relacionados com a motilidade de Azospirillum brasilense foram encontrados no plasmídeo de 90MDa (CROES et al., 1991). A presença de um plasmídeo de 90 MDa foi notada frequentemente em estirpes de A. brasilense e A. lipoferum (ONYEOCHA et al., 1990). Em outras bactérias do solo, megaplasmídeos podem conter informações essenciais para a interação com a planta (STEENHOUDT & VANDERLEYDEN, 2000).

Os genes de virulência (vir) de Agrobacterium, bem como os de nodulação (nod e hsn) de Sinorhizobium estão localizados em megaplasmídeos (STEENHOUDT & VANDERLEYDEN, 2000). A construção de mutantes contendo deleção no plasmídeo de 90 MDa de A. brasilense Sp7 levou a identificação de três loci envolvidos na motilidade: Mot1, Mot2 e Mot3. As deleções em Mot1 e Mot2 afetaram a motilidade em meio semi-sólido, mas não no meio líquido, sugerindo que estes loci provavelmente estão envolvidos na síntese ou funcionamento dos flagelos laterais. O mutante Mot3 é completamente imóvel e aparentemente perdeu tanto o flagelo polar quanto os laterais (CROES et al., 1991).

Sequências homólogas ao gene chv de A. tumefaciens (WAELKENS et al., 1987), essenciais para a ligação e formação de tumor na planta, e aos genes nod e hsn de Sinorhizobium meliloti (FOGHER et al., 1985), necessários para a formação do nódulo, foram identificadas em Azospirillum brasilense. Os genes nodPQ foram encontrados no plasmídeo de 90 MDa (VIEILLE & ELMERICH, 1990), enquanto o gene homólogo nodG parece estar localizado no cromossomo (VIEILLE & ELMERICH, 1992). Entretanto, a função destes genes na interação Azospirillumplanta permanece desconhecida (STEENHOUDT & VANDERLEYDEN, 2000). Por outro lado, o gene chvB de A. brasilense pode complementar um mutante defectivo chvB de Agrobacterium com respeito a formação de tumor nas folhas de Nicotiana tabacum. Assim, é possível que o produto do gene chvB possa estar envolvido na (STEENHOUDT & adsorção do Azospirillum raízes das plantas às VANDERLEYDEN, 2000).

Os genes de *A. brasilense* que codificam para enzimas envolvidas na síntese de exopolissacarídeos (EPS), foram isolados por complementação de mutantes *exoB* e *exoC* de *Sinorhizobium meliloti* pelo seu fenótipo calcoflúor-negativo. Calcoflúor é um corante que se liga especificamente a polissacarídeos com ligações β -1,4 e β -1,3. Estes mutantes de *S. meliloti* são deficientes na síntese de succinilglucana, o maior EPS ácido de rizóbios, e formam nódulos não fixadores de nitrogênio em alfafa. Em *A. brasilense* foram isolados dois *loci exoB* (*exoB1* e *exoB2*) e um *locus exoC*. Ambos os *loci exoB* de *A. brasilense* complementaram o mutante *exoB* de *S. meliloti* para a

produção de succinilglucana e para o fenótipo simbiótico, enquanto o *locus exoC* de *A. brasilense* não complementou o fenótipo Fix do mutante *exoC* de *S. meliloti* (MICHIELS *et al.*, 1988). Os genes *exoB* de *A. brasilense* são funcionalmente homólogos ao *exoB* de *S. meliloti*, que codifica a enzima UDP-glucose-4-epimerase envolvida na síntese de UDP-galactose e necessária para incorporação de galactose em exopolissacarídeos, polissacarídeos capsulares e lipopolissacarídeos (DE TROCH *et al.*, 1994). O gene *exoC* codifica uma proteína que apresenta homologia com o gene *algC* de *Pseudomonas aeruginosa*, que codifica uma fosfomanomutase essencial para a conversão de GDP-manose-6-fosfato em GDP-manose-1-fosfato (STEENHOUDT & VANDERLEYDEN, 2000). O *locus exoC* de *A. brasilense* codifica uma enzima diferente daquela que foi perdida no mutante *exoC* de *S. meliloti* (fosfoglicomutase). Isto explica a incapacidade deste gene de restaurar a produção de EPS e o fenótipo simbiótico no mutante *exoC* de *S. meliloti*. Os genes *exoB1* e *exoC* de *Azospirillum brasilense* foram localizados no plasmídeo de 90 MDa, e o gene *exoB2* foi localizado no cromossomo (CROES *et al.*, 1991; MICHIELS *et al.*, 1989).

A maioria dos estudos genéticos sobre a fixação do nitrogênio em Azospirillum foi realizado com A. brasilense. O operon nifHDK de A. brasilense, codificando as proteínas estruturais da nitrogenase, foram identificados e posteriormente isolados baseados na similaridade de seqüência com os genes nif de Klebsiella pneumoniae (QUIVIGER et al., 1982). Um número adicional de genes nif e fix envolvidos no processamento e transporte de elétrons para o complexo da nitrogenase, biossíntese do cofator ferro-molibdênio, bem como genes da regulação da expressão da fixação do nitrogênio foram isolados (FRAZZON & SCHRANK, 1998; MILCAMPS & VANDERLEYDEN, 1993). Exceto os genes nifA e nifB que são transcritos separadamente, todos os outros genes nif estão localizados em um grupo que compreende 30 kb do genoma (LIANG et al., 1991).

O amônio fixado pela nitrogenase é assimilado pelo A. brasilense principalmente pela via glutamina sintetase (GS)/glutamato sintase (GltS) (WESTBY et al., 1987). Dois genes, gltDB, que codificam as subunidades da glutamato sintase foram encontrados em Azospirillum brasilense (PELANDA et al., 1993). A glutamina sintetase é codificada pelo *glnA*, que neste organismo faz parte de um operon junto com o gene *glnB* (codifica a proteína P_{II}) (ZAMAROCZY *et al.*, 1993). A atividade da glutamina sintetase é regulada em resposta aos níveis de amônio por adenililação reversível, reação catalizada pela proteína GlnE ou adenililtransferase. Na presença de altas concentrações de amônio, a GS está altamente adenililada e inativada. Em condições limitantes de nitrogênio, os grupos adenilil da GS são removidos pela proteína GlnE, ativando a enzima (GAUTHIER & ELMERICH, 1977). Ao contrário do que ocorre em bactérias entéricas, o produto do gene *glnB* (proteína PII), não participa na via de regulação para a adenililação da GS de *A. brasilense* (GAUTHIER & ELMERICH, 1977).

1.4 ESTRUTURA GENÔMICA DO GÊNERO Azospirillum

Plasmídeos de *Azospirillum* foram descritos pela primeria vez por FRANCHE & ELMERICH (1981), através de eletroforese em gel de agarose. Em geral as estirpes estudadas apresentavam de um a seis plasmídeos com massa molecular variando de 4 MDa a 370 MDa (6 a 550 kb). A presença de um plasmídeo de 90 MDa (135 kb) foi notada freqüentemente em estirpes de *A. brasilense* (PLAZINSKI *et al.*, 1981; ONYEOCHA *et al.*, 1990)

A presença de plasmídeos e "minicromossomos" em espécies de *Azospirillum* foi descrita por WOOD *et al.* (1982). Neste trabalho, seis estirpes de *A. brasilense* e duas de *A. lipoferum* foram analisadas utilizando a técnica de Eckhardt (ECKHARDT, 1978), e a presença de sete a oito replicons foram observadas nas duas espécies. No entanto, a denominação de minicromossomos foi baseada somente no tamanho dos replicons.

Mais recentemente, um estudo com 26 isolados de *A. brasilense* revelou a presença de cinco a oito replicons, separados pela técnica de Eckhardt. Cada estirpe de *A. brasilense* mostrou um único perfil e o tamanho molecular dos replicons variou de 65 a mais de 1800 kb. O megareplicon de 1700 kb e outros com tamanho aproximado de 600 kb hibridizaram fortemente com uma sonda do gene 16S rDNA. Assim, a

presença de genes codificando para 16S rRNA em múltiplos replicons de *A. brasilense* sugere que o genoma desta espécie consiste de múltiplos minicromossomos (CABALLERO-MELLADO *et al.*, 1999), uma vez que replicons contendo genes com fúnções indispensáveis são classificados como cromossomos (MORENO, 1998).

MARTIN-DIDONET *et al.* (2000) descreveram a presença de vários megareplicons em cinco espécies de *Azospirillum* estudadas por eletroforese em campo pulsado (PFGE ou *Pulsed-Field Gel Electrophoresis*). As estirpes de *A. brasilense* analizadas (Sp7, FP2, Cd) mostraram cinco megareplicons, variando de 0,63 a 2,5 Mpb. Os perfis eletroforéticos de *Azospirillum lipoferum* estirpes Sp59b e JA25 mostraram de 8 e 10 replicons, respectivamente, com tamanho molecular variando de 0,15 a 2,6 Mpb. As estirpes de *A. amazonense* Y2 e Y6 possuem quatro replicons, mas de perfis distintos na eletroforese em campo pulsado, com tamanhos variando de 0,71 Mpb (Y6) a 2,8 Mpb (Y2). Vários replicons foram também observados em *A. halopraeferens* e *A. irakense*, cada espécie com um padrão de DNA específico. O tamanho total do genoma do gênero *Azospirillum* variou de 4,8 Mpb (*A. irakense*) a 9,7 Mpb (*A. lipoferum* estirpe Sp59b). O comportamento eletroforético dos replicons indicou que algumas moléculas de DNA são provavelmente lineares (MARTIN-DIDONET *et al.*, 2000).

Os estudos de hibridização com a sonda 16S rDNA de *A. brasilense* mostraram que todas as espécies testadas possuem pelo menos dois replicons com cópias deste gene. Enquanto o gene que codifica para 16S rRNA tenha sido encontrado em vários replicons, o operon *nifHDK* parece estar localizado em apenas um replicon em *A. brasilense* estirpe FP2 e Sp7. Estes genes foram localizados no maior replicon destas estirpes, como revelado pela hibridização com uma sonda *nifHDK* de *A. brasilense* (MARTIN-DIDONET *et al.*, 2000).

Azospirillum spp. apresenta um dos mais complexos padrões de DNA de alta massa molecular entre as α -Proteobactérias já estudadas, com a informação genética distribuída em vários replicons (MARTIN-DIDONET *et al.*, 2000).
2. OBJETIVO

Construir bibliotecas genômicas aleatórias do replicon pAZ6 (630 kpb) de *A*. *brasilense* estirpe FP2 a partir do DNA isolado por eletroforese em campo pulsado e seqüenciar os clones desta biblioteca para obter um panorama das informações genéticas contidas neste replicon.

OBJETIVOS ESPECÍFICOS:

- Isolar o replicon pAZ6 (630 kpb) por PFGE.

- Construir bibliotecas deste replicon pAZ6 de *A. brasilense* FP2 no vetor pCR4Blunt-TOPO.

- Seqüenciar as duas extremidades dos clones aleatórios.

- Ordenar as seqüências obtidas em contíguos.

- Identificar genes codificados pelo replicon pAZ6 de *A. brasilense* FP2 por comparação das seqüências obtidas com as do banco de dados Genbank.

3. MATERIAIS E MÉTODOS

3.1 BACTÉRIAS E PLASMÍDEOS UTILIZADOS

As bactérias e o plasmídeo que foram utilizados neste trabalho estão listados na tabela 2.

TABELA 2 - BACTERIAS E PLASMIDEO				
Estirpes/Plasmídeo	Genótipo/Fenótipo	Referência		
Azospirillum brasilense				
FP2	Estirpe selvagem, Nif [*] Sp7 (ATCC29145) Nal ^R Sm ^R	Pedrosa e Yates, 1984		
Escherichia coli				
TOP10	F mcrA D(mrr-hsdRMS-mcrBC) \$801acZDM15 DlacX74 deoR reaA1 araD139 D(ara-leu)7697 galU galk rpsL (Str ^R) endA1 nupG	Invitrogen		
Plasmídeo PCR®4Blunt-TOPO®	Amp ^R , Km ^R , P _{lac} , lacZccdB	Invitrogen		

TABELA 2 - BACTÉRIAS E PLASMÍDEO

3.2 REAGENTES

Lisozima, agarose para eletroforese em campo pulsado, agarose de baixo ponto de fusão, Tris base, glicerol e antibióticos foram adquiridos da Sigma Chemical Company. Proteinase K, dodecilsulfato de sódio (SDS), agarase, RNAse, biotina, Triton-X, desoxicolato de sódio, lauril sarcosina de sódio, acetato de sódio, fenol e padrão de tamanho molecular de DNA (1 kb Ladder) da Invitrogen. EDTA, NaCl, NaOH, etanol, isopropanol, ácido acético, clorofórmio e álcool isoamílico foram adquiridos da Merck. As enzimas modificadoras de DNA e RNAse foram adquiridas da Invitrogen. Extrato de levedura, triptona e ágar bacteriológico foram adquiridos da Biobrás ou Invitrogen. O gás nitrogênio foi adquirido da White Martins. O marcador de DNA de alto tamanho molecular (cromossomos de *Sacharomyces cereviseae*) foi adquirido da Bio-Rad ou Amersham Biosciences.

3.3 MEIOS DE CULTURA

3.3.1 MEIOS EMPREGADOS PARA O CULTIVO DE Azospirillum brasilense.

As estirpes de *A. brasilense* FP2 foram crescidas em meio NFbHP (MACHADO *et al.*, 1991) contendo NH₄Cl 20 mmol/L. A composição e preparo do meio NFbHP estão descritos abaixo.

MEIO NFb	
MgSO ₄ .7H ₂ O	0,2 g/L
NaCl	0,1 g/L
CaCl ₂	20 mg/L
Ácido nitrilo triacético	56 mg/L
FeSO ₄ .7H ₂ O	20 mg /L
Lactato de sódio	5,0 g/L
Biotina	0,1 mg/L
$Na_2MoO_4.2 H_2O$	2 mg/L
MnSO ₄ .H ₂ O	2,35 mg/L
H ₃ BO ₄	2,8 mg /L
$CuSO_{4.}5H_{2}O$	0,08 mg/L
ZnSO ₄ .7H ₂ O	0,24 mg/L

SOLUÇAU DE FUSFATUS	
K ₂ HPO ₄	1 20 g /L
KH ₂ PO ₄	80 g/L

Na hora do uso o meio NFb foi acrescido da solução de fosfatos na proporção de 50 mL/L passando a ser denominado NFbHP. O meio de cultura e a solução de fosfatos foram esterilizados separadamente por autoclavação a 120°C a 1 atmosfera de pressão durante 20 minutos. O meio deve apresentar pH final de 6,8. A fonte de nitrogênio utilizada foi uma solução 1 mol/L de cloreto de amônio, esterilizada nas

condições descritas acima, e acrescida ao meio NFbHP na proporção de 20 mL/L, passando então o meio a ser denominado NFbHPN. O meio NFbHP sólido foi obtido adicionando-se ágar bacteriológico na concentração de 15 g/L ao meio líquido.

3.3.2 MEIOS EMPREGADOS PARA O CULTIVO DE E. COLI.

O meio Luria Bertani Broth (LB) (SAMBROOK et al., 1989) apresenta a seguinte composição:

Bacto triptona	10 g/L
Extrato de levedura	5 g/L
NaCl	10 g/L

O pH do meio deve ser ajustado para 7,0 com NaOH. O meio foi esterilizado por autoclavação a 120°C e 1 atmosfera de pressão durante 20 minutos. O meio sólido Luria Bertani Ágar (LA) foi obtido pela adição de 15 g/L ágar ao meio líquido.

O meio Terrific Broth (SAMBROOK et al., 1989) foi preparado como descrito a seguir.

Bacto triptona	12 g/L
Extrato de levedura	24 g/L
Fosfato de potássio, dibásico	9,4 g/L
Fosfato de potássio, monobásico	2 g/L
Glicerol	4 mL/L

O meio foi esterilizado por autoclavação a 120°C e 1 atmosfera de pressão durante 20 minutos.

3.4 CONDIÇÕES DE CULTIVO

As culturas de *A. brasilense* foram incubadas a 30°C em meio NFbHP (MACHADO *et al.*, 1991) suplementado com 20mmol/L de cloreto de amônio. Os antibióticos foram utilizados nas seguintes concentrações: estreptomicina (Sm) 80µg/mL e ácido nalidíxico (Nal) 10µg/mL.

A estirpe de *E. coli* TOP10 contendo ou não plasmídeos foi cultivada a 37°C em meio Luria Bertani Broth (LB), Luria Bertani Ágar (LA) (SAMBROOK *et al.*, 1989) ou Terrific Broth (TARTOF e HOBBS, 1987). Os meios de cultivo foram suplementados com as seguintes concentrações de antibióticos: ampicilina (Amp) 250µg/mL, canamicina (Km) 50µg/mL ou estreptomicina (Sm) 20µg/mL.

As culturas das bactérias foram estocadas e mantidas em glicerol 50% a -20°C, e alternativamente em frascos com tampa de rosca contendo meio NFbHPN ou LA sólido inclinado, à temperatura ambiente.

3.5 MANIPULAÇÃO DE DNA

As reações de defosforilação, ligação e restrição de DNA foram realizadas em condições especificadas pelo fabricante.

A extração e purificação de plasmídeos foram realizadas pelo método de lise alcalina ("Mini-Prep") (SAMBROOK *et al.*, 1989). Plasmídeos íntegros e fragmentos de restrição foram separados por eletroforese em gel de agarose, corados com solução brometo de etídeo (0,5µg/mL) e visualizados sob luz ultravioleta (302nm) conforme descrito (SAMBROOK *et al.*,1989).

Fragmentos de DNA de interesse foram isolados dos géis de agarose de baixo ponto de fusão após separação eletroforética. O tampão de corrida utilizado foi TAE (Tris base 40mmol/L, ácido acético 40mmol/L, EDTA 2mmol/L, pH8,0). Após a eletroforese, o gel foi tratado com solução de brometo de etídeo (0,5µg/mL) e o fragmento de interesse foi identificado por comparação com o padrão de tamanho molecular (1 Kb Ladder), evitando-se excesso de gel e exposição prolongada do DNA

à luz ultravioleta. A fatia do gel contendo o fragmento de interesse foi colocada em tubo tipo eppendorf e incubada a 65°C até completa fusão da agarose. Em seguida foram adicionados um volume de água milliQ, pré-aquecida a 65°C, 1/5 do volume de uma solução de Tris-acetato 1mol/L (pH 8,0) e 1 volume de fenol-clorofórmio. Logo em seguida, os tubos foram agitados vigorosamente em vórtex. Decorridos 10 min de incubação à temperatura ambiente, os tubos foram novamente agitados em vórtex e centrifugados a 13.000 rpm por 5 min. A fase aquosa foi transferida para um novo tubo e novamente tratada com fenol-clorofórmio para a remoção da agarose remanescente. A fase aquosa foi precipitada com 0,6 volumes de isopropanol e o sedimento lavado duas vezes com etanol 70% e seco a vácuo. O precipitado foi dissolvido em 10 μ L de água milliQ estéril.

3.5.1 OBTENÇÃO DO DNA CROMOSSÔMICO INTACTO

Para a obtenção do DNA cromossômico intacto, as células de Azospirillum foram cultivadas em meio líquido NFbHPN até densidade ótica (D.O₆₀₀) de 0,6. Aproximadamente 30mL da cultura foram concentrados em tubos Falcon por centrifugação a 5000 rpm por 10 min. O sedimento de células foi ressuspenso em 250µL de tampão SET (Tris pH 7,5 10mmol/L, EDTA 100mmol/L e NaCl 200mmol/L) e foram adicionados à suspensão de células 500µL de agarose de baixo ponto de fusão a 1%, dissolvida em água milliQ e mantida a 40° para evitar geleificação. Alíquotas de 100µL desta suspensão foram transferidas para os moldes (10 x 5 x 1,5 mm) e mantidos a 4°C durante 15 minutos para a geleificação da agarose. Os blocos de agarose foram transferidos para tubos tipo eppendorf de 1,5mL com 1mL de solução de lise (Tris-HCl pH8,0 6mmol/L, NaCl 1mol/L, EDTA 100mmol/L, Triton X-100 0,5%, desoxicolato de sódio 0,2%, lauril sarcosina de sódio 0,5%) contendo 1mg/mL de lisozima, e incubados à 37°C durante 24 horas. A solução foi trocada pela solução de digestão ESP (EDTA 0,5M pH 8,0, lauril sarcosina de sódio 1%, desoxicolato de sódio 0,2%) acrescida de proteinase K 0,1mg/mL. O sistema foi então incubado a 50°C por 72 horas. Os blocos de agarose com o DNA cromossomal

foram estocados na solução de digestão ESP isenta de proteinase K a 4°C (SMITH & CANTOR, 1988).

Os replicons do DNA cromossomal foram separados por eletroforese em campo pulsado (PFGE - *Pulsed-Field Gel Electrophoresis*) em tampão de corrida TBE (Trisborato 44,5mmol/L, EDTA 1mmol/L pH8,0). Durante a corrida a temperatura foi mantida a 14°C, o ângulo de reorientação foi de 120°, e a 6 V/cm com tempo de pulso de 60 seg por 15 horas e 90 seg por 8 horas ou tempo de pulso escalonado de 40 a 120 seg durante 24 horas. O sistema de eletroforese CHEF-DR III (Bio-Rad) foi utilizado em todos os ensaios.

3.5.2 ISOLAMENTO DOS REPLICONS DE Azospirillum brasilense FP2

Os replicons de interesse foram isolados dos géis de agarose de baixo ponto de fusão 1% após separação por eletroforese em campo pulsado nas condições acima. Após a eletroforese, o gel foi tratado com solução de brometo de etídio, o replicon de interesse foi identificado por comparação com o padrão de tamanho molecular (Saccharomyces cerevisiae) e a fatia de gel contendo a região de interesse foi retirada. A fatia do gel foi cortada em pedaços menores, colocada em tubo tipo eppendorf e incubada a 65°C até completa fusão da agarose. Em seguida foram adicionados um volume de água milliQ (pré-aquecida a 65°C), 1/5 do volume de uma solução de Trisacetato 1 mol/L (pH 8,0) e 1 volume de fenol-clorofórmio. Logo em seguida, os tubos foram invertidos delicadamente por 5minutos. Decorridos 10 minutos de incubação à temperatura ambiente, os tubos foram novamente homogeneizados por inversão e centrifugados a 13.000 rpm por 5 minutos. A fase aquosa foi transferida para um novo tubo e novamente tratada com fenol-clorofórmio para a remoção da agarose remanescente. A fase aquosa foi precipitada com 0,6 volume de isopropanol e o sedimento lavado duas vezes com etanol 70% e seco a vácuo. O precipitado foi dissolvido em 10 µL de T₁₀E₁ (Tris-HCl pH7,5 10mmol/L, EDTA 1mmol/L) ou água milliQ estéril.

3.5.2.1 ISOLAMENTO DO REPLICON UTILIZANDO A ENZIMA β -AGARASE

A β -agarase ou agarose 3-glicanohidrolase é uma enzima que rompe as ligações $(1\rightarrow 3)$ - β -D da agarose. A agarose é formada por ligações repetidas de $(1\rightarrow 4)$ -3,6anidro- α -L-galactopiranosil- $(1\rightarrow 3)$ - β -D-galactopirananose (DUCKWORTH & TURVEY, 1969). Como resultado desta degradação enzimática a viscosidade e a capacidade natural de gelificação da agarose de baixo ponto de fusão são eliminadas liberando assim os ácidos nucleicos. A incubação prolongada do DNA com a enzima não reduz o tamanho do DNA nem interfère em procedimentos subseqüentes .

A fatia do gel obtida na eletroforese foi colocada em um tubo Falcon de 15ml, e lavada duas vezes em $T_{10}E_1$ (Tris-HCl 10mmol/L, pH7,5 e EDTA 1mmol/L) por uma hora à temperatura ambiente. Após este período, a solução de $T_{10}E_1$ foi retirada e o tubo foi incubado em banho-maria a 65°C por aproximadamente 10 minutos ou até a completa fusão da agarose. O tubo Falcon foi transferido para um banho maria a 40°C e incubado por aproximadamente 3 minutos para equilibrar o sistema, e então foi adicionado 0,5U de β -agarase para cada 100µl de agarose. A digestão da agarose foi realizado durante uma hora a 40°C. A enzima foi inativada pela incubação do sistema a 60°C durante uma hora. O DNA foi precipitado com 2 volumes de etanol absoluto e o sedimento lavado duas vezes com etanol 70% e seco a vácuo. O precipitado foi dissolvido em 10µL de água milliQ estéril.

3.6 AMPLIFICAÇÃO NÃO-PCR DO DNA GENÔMICO

Para a amplificação do DNA do replicon de interesse foi usado o kit TempliPhi (Amersham Biosciences). A tecnologia TempliPhi utiliza a DNA polimerase do fago Phi29 que amplifica o DNA por um sistema de círculos rolantes (LEWIN, 2001) com primers aleatórios sob condições isotérmicas. A enzima apresenta alta processividade e acurácia apresentando amplificados com mais de 10kb em comprimento com uma frequência de erro menor que 1 em 10^6 .

Alíquotas de 1µL do DNA isolado da agarose de baixo ponto de fusão foram adicionados ao tampão (Tris-HCl pH8,2 10mM, EDTA 0,1mM). A mistura foi aquecida a 95°C durante 5 minutos e depois resfriada rapidamente a 4°C por mais 5 minutos. Ao DNA desnaturado foram adicionados 5µL da mistura de reação RCA (*"Rolling Circles Amplification"*) (Tampão de reação + 0,2µL da enzima, Amersham Biosciences). A reação de amplificação foi realizada à temperatura de 30°C durante 18 horas. A enzima foi inativada pelo calor a 65°C por 10 min.

3.7 QUANTIFICAÇÃO DO DNA

A concentração do DNA foi estimada espectrofotometricamente pela absorbância a 260nm (A_{260}). Os cálculos foram efetuados considerando-se que uma solução de 50µg/mL de DNA fita dupla ou 35µg/mL de DNA fita simples apresentam uma A_{260} = 1,0 (SAMBROOK *et al.*, 1989).

Alternativamente, as concentrações de DNA foram determinadas por densitometria de eletroforetogramas em gel de agarose. A concentração aproximada foi feita por densitometria das bandas de DNA, tomando-se por base a intensidade das bandas do padrão de DNA de concentração conhecida.

3.8 FRAGMENTAÇÃO MECÂNICA DO DNA

Para a fragmentação do DNA de interesse foi utilizado um nebulizador. Neste nebulizador foi adicionado tampão de fragmentação (acetato de sódio 3M, glicerol 50%), e aproximadamente 25µg de DNA e água milliQ estéril para um volume final de 1mL. O DNA foi nebulizado por 1 min e 30 seg a 2,0 Kgf/cm² com gás nitrogênio para a obtenção de fragmentos de 1 a 5Kpb. O tamanho dos fragmentos foi confirmado por eletroforese em gel de agarose (1%) com tampão de corrida TBE (Tris-borato 89mmol/L, EDTA 2mmol/L pH 8,3).

O DNA fragmentado foi transferido para um tubo eppendorf de 2,0mL e precipitado com 700µL de isopropanol 100% e 80µL de acetato de sódio 3M pH 5,2.

Após duas horas de incubação à temperatura ambiente, o tubo foi centrifugado a 13000 rpm por 15 minutos. O sobrenadante foi removido e o precipitado de DNA foi lavado uma vez com 1 volume de etanol 80% e seco a vácuo. O precipitado foi dissolvido em 10µL de água milliQ estéril.

3.9 CONSTRUÇÃO DA BIBLIOTECA GENÔMICA

As bibliotecas foram construídas utilizando o sistema TOPO Shotgun Subcloning (Invitrogen). As reações de reparo das extremidades dos fragmentos de DNA, desfosforilação e clonagem do DNA foram realizadas seguindo as instruções do fabricante.

O plasmídeo utilizado como vetor foi o pCR®4Blunt-TOPO® (FIGURA 1). Este plasmídeo é fornecido linearizado, e apresenta a enzima topoisomerase I do vírus *Vaccinia* covalentemente ligada á extremidade 3'. A topoisomerase I do vírus *Vaccinia* liga-se específicamente ao DNA dupla-fita e quebra a ligação fosfodiéster após a seqüência 5'-CCCTT em uma das extremidade. A energia proveniente da quebra da ligação fosfodiéster é conservada para a formação de uma ligação covalente entre o fosfato 3' da extremidade quebrada e um resíduo tirosil (Tyr-274) da topoisomerase I. A ligação fosfo-tirosil entre o DNA e a enzima é subseqüentemente atacada pela extremidade 5'hidroxil do DNA inserto, revertendo a reação e liberando a topoisomerase (TOPO Shotgun Subcloning Kit, Invitrogen).

3.10 TRANSFORMAÇÃO BACTERIANA POR ELETROPORAÇÃO

O preparo de células competentes e transformação por eletroporação de *E. coli* com DNA plasmidial foi realizado conforme recomendado pelo manual do eletroporador CELL-PORATOR ELECTROPORATION SYSTEM (Life Technologies, INC). O aparelho foi ajustado para a aplicação do choque elétrico numa diferença de potencial de 12 a 16 kV/cm à câmara de eletroporação, com tempo de pulso entre 6 a 10 ms.

Após a aplicação do choque elétrico, as células foram ressuspensas em 1mL meio SOC e incubadas para recuperação a 37°C por 30 minutos. Em seguida, 100µL foram plaqueados em LA contendo os antibióticos adequados.

FIGURA 1 - MAPA FÍSICO DO VETOR pCR®4BLUNT-TOPO®



O vetor pCR4Blunt-TOPO foi utilizado na construção das bibliotecas genômicas do replicon pAZ6 de *A. brasilense* estirpe FP2. Este plasmídeo contem 3954 nucleotídeos e apresenta as seguintes características:

Região do promotor lac: bases 2 - 216 Sítio de ligação CAP: bases 95 - 132 Sítio de ligação da RNA polimerase: bases 133-178 Sítio de ligação do repressor Lac: bases 179-199 Início da transcrição: base 179 Sítio do primer reverso M13: bases 205-221 Fusão do gene LacZα-ccdB *: bases 217-810 Porção do LacZ α da fusão: bases 217-497 Porção do *ccdB* da fusão: bases 508-810 Sítio do primer T3: bases 243-262 Sítio de clonagem TOPO: bases 294-295 Sítio do primer T7: bases 328-347 Sítio do primer universal M13 (-20): bases 355-370 Promotor do gene de resistência a canamicina: bases 1021-1070 Gene de resistência a canamicina: bases 1159-1953 Gene de resistência a ampicilina (bla): bases 2309-3061 (fita complementar) Promotor do gene de resistência a ampicilina (bla): bases 3062-3158 (c) Origem pUC: bases 3159-3832

* O gene letal ccdB (control of cell death) é incorporado no sítio de policionagem do vetor pCR4Blunt como uma fusão com o fragmento $lacZ\alpha$. A proteína CcdB age sobre a DNA girase bacteriana (topoisomerase II), inibindo a sua habilidade de religar as fendas que se formam no DNA dupla-fita levando a degradação do cromossomo da célula hospedeira. Devido a proteína CcdB ser tóxica, o vetor pCR4Blunt previne o crescimento de bactérias, a menos que a fusão *lacZ-ccdB* esteja interrompida pela ligação de um inserto.

3.11 MINI-PREPARAÇÃO DE PLASMÍDEOS EM MICROPLACAS

Em um bloco estéril de 96 poços foi adicionado 1 mL de meio Terrific Broth contendo 250µg/mL de ampicilina (250µL de ampicilina 100mg/mL para 100mL de meio) em cada poço. As colônias foram inoculadas com o auxílio de um repicador a partir de uma placa de LA inoculada no dia anterior e incubada a 37° C por cerca de 16 horas. Os blocos foram selados com adesivo selador, um pequeno orificio foi feito em um dos cantos de cada poço para facilitar a aeração e incubados a 37°C, 150-200 rpm, durante 16 horas. As culturas foram centrifugadas a 4000 rpm por 7 min. O sobrenadante foi descartado e ao pellet foram adicionados 180µL de GET (glucose 25mmol/L, EDTA 10mmol/L, Tris-HCl 50mmol/L pH8,0) para lavar as células que foram sedimentadas por nova centrifugação a 4000 rpm por 7 min. O sobrenadante foi descartado e os blocos mantidos invertidos sobre papel absorvente por cerca de 1 minuto.

A cada poço foi adicionado 80μ L de GET contendo $0,5\mu$ L de RNAse (20mg/mL) e os blocos foram homogeneizados em vórtex até a completa suspensão das células. As células foram transferidas para uma microplaca de 250 μ L e lisadas com 80μ L de solução de lise (NaOH 0,18mol/L, SDS 1%). As placas foram seladas, homogeneizadas cuidadosamente por inversão e incubadas por 5 minutos à temperatura ambiente. Em seguida foi adicionado 80μ L de acetato de potássio 3M pH 5,8 na mesma ordem que foi adicionado a solução de lise para evitar diferentes tempos de lise. As placas foram seladas, homogeneizadas por inversão e incubadas por 10 minutos à temperatura ambiente. Para evitar que fique solução no adesivo, as placas foram centrifugadas por alguns segundos (4.000 rpm) e posteriormente incubadas, sem o adesivo, em estufa a 90°C por 30 minutos. Após este período, as placas foram resfriadas em gelo picado por 15 minutos e centrifugadas a 4000 rpm por 10 minutos a 4° C. O sobrenadante (120 μ L) de cada poço foi transferido para os poros respectivos da placa Millipore (MAGV N22) fixada, com fita adesiva, no topo de uma microplaca de polipropileno nova de fundo em "V" de 250 μ L. O conjunto foi centrifugado a 4.000

rpm por 5 minutos a 20°C. A solução de DNA (aproximadamente 120 μ L) filtrada foi coletada na placa de fundo em "V".

A precipitação do DNA foi realizada adicionando-se 80µL de isopropanol ao filtrado. A placa foi homogeneizada por inversão e centrifugada por 45 minutos a 4.000 rpm, 20°C. O sobrenadante foi descartado e o precipitado foi lavado uma vez com 150µL de etanol 70% e seco a vácuo por 10 minutos ou em estufa à 37°C por uma hora. O DNA foi dissolvido em 30µL de água milliQ estéril. As placas com DNA purificado foram armazenadas à -20°C.

3.12 REAÇÃO DE SEQÜENCIAMENTO DO DNA

O seqüenciamento do DNA foi relizado pelo método de terminação de cadeia (SANGER *et al.*, 1977) em sequenciador automático MegaBACE 1000 (Amersham Biosciences) utilizando dideoxinucleotídeos fluorescentes. As reações de seqüenciamento e purificação foram realizadas segundo protocolo do kit *DYEnamic ET Dye Terminator Cycle Sequencing* (Amersham Biosciences).

Os seguintes oligonucleotídeos iniciadores foram utilizados nas reações de sequenciamento:

Primer	Seqüência
Universal	5'P GTAAAACGACGGCCAG 3'OH
Reverso	5'P CAGGAAACAGCTATGAC 3'OH

Os parâmetros para a reação de seqüenciamento foram:

Programa:	1 ciclo:	95°C	20 seg
	35 ciclos:	95°C	20 seg
		55°C	15 seg
		60°C	1 min e 30 seg

3.13 ANÁLISE E ALINHAMENTO DAS SEQÜÊNCIAS

Para o processamento e alinhamento das seqüências foram utilizados os programas PHRED/PHRAP/CONSED (EWING et al., 1998; EWING & GREEN, 1998; GORDON et al., 1998).

O programa PHRED reconhece as seqüências de nucleotídeos a partir do arquivo de dados brutos do seqüenciador (eletroforetograma), atribui valores de qualidade às bases da seqüência nucleotídica. Este processo é denominado identificação de bases ou "*base call*" (EWING *et al.*, 1998). Os valores de qualidade das seqüências analisadas, que são encontrados nos arquivos FASTA ou PHD, determinado pela fórmula abaixo, que calcula a probabilidade de erro na identificação de bases, onde *Pe* é a probabilidade de uma base estar errada (PROSDOCINI *et al.*, 2003)

PHRED Qualidade = $-10 \log (Pe)$

Regiões contaminantes são partes da seqüência obtida que não representam o DNA que se deseja analisar. Tais regiões representam, normalmente, partes de vetores de clonagem. Essas regiões são retiradas ou mascaradas pelo programa Cross_match, que compara a seqüência desejada com o arquivo de seqüências de vetores (EWING *et al.*, 1998; PROSDOCINI *et al.*, 2003).

O PHRAP (*Phragment Assembly Program*) é o programa responsável pela montagem dos pequenos fragmentos de DNA seqüenciados em seqüências maiores, ou contíguos (*contigs*) (EWING &GREEN, 1998). Este programa possui alguns pontos chaves para a obtenção de um resultado final satisfatório tais como: construção da seqüência do contíguo através de um mosaico de partes das seqüências de alta qualidade; utilização de informações da qualidade dos dados computados internamente e de implementações feitas pelos usuários para aumentar a qualidade da montagem (EWING &GREEN, 1998). A visualização e edição das seqüências geradas após a montagem são realizadas através do programa CONSED (GORDON *et al.*, 1998).

3.14 ANOTAÇÃO DA SEQÜÊNCIA

O principal objetivo desta etapa é determinar os produtos codificados pelo DNA seqüenciado, posicionando cada um dos possíveis genes e caracterizando as regiões não-gênicas. Nesta etapa foi utilizado o programa FRAMEPLOT (ISHIKAWA & HOTTA, 1999) para localizar as prováveis regiões codificadoras de proteínas ou ORFs (*open reading frames*) na seqüência de DNA.

Mapeados os possíveis genes, a etapa seguinte consiste em identificar quais proteínas são codificadas. Todas as prováveis regiões que codificam para proteínas, com comprimento igual ou maior que 90 pb, foram traduzidas em seqüências de aminoácidos, e foram comparadas com bancos de dados de proteínas (GeneBank) utilizando o programa BLASTP (*Basic Local Alignment Search Tool*) (ALTSCHUL *et al.*, 1997).

4. RESULTADOS E DISCUSSÃO

4.1 IDENTIFICAÇÃO DO REPLICON pAZ6 (630 kbp) DE *A. brasilense* FP2 POR ELETROFORESE EM CAMPO PULSADO

A primeira etapa deste trabalho foi a obtenção do perfil genômico do Azospirillum brasilense FP2 por eletroforese em campo pulsado (PFGE) e identificação do replicon na região de 630 kb descrito por MARTIN-DIDONET et al. (2000). Estes autores mostraram que este replicon, denominado de pAZ6, contém pelo menos uma cópia do gene que codifica para 16S rRNA podendo, portanto, ser considerado um cromossomo (MARTIN-DIDONET et al., 2000). O DNA intacto de A. brasilense foi preparado em blocos de agarose (Material e Métodos, item 3.5.1) e os replicons foram separados pelo sistema de eletroforese CHEF-DR III (Bio-Rad) utilizando agarose para eletroforese em campo pulsado a 1%, tampão TBE 0,5X, temperatura de 14°C, ângulo de reorientação de 120° e pulso de 60 segundos por 15 horas e 90 segundos por 8 horas. Estes parâmetros foram escolhidos por permitirem a separação dos replicons na faixa de tamanho molecular de 240 a 2000 Kbp. A FIGURA 2 mostra o eletroforetograma obtido de duas amostras com concentrações de células diferentes. Na amostra 1 foi utilizado uma cultura com densidade ótica $(D.O_{600})$ de 0,6 e na amostra 2, uma cultura com densidade ótica $(D.O_{600})$ de 0,85. Na região entre 500 a 1000 kb, puderam ser observadas 5 bandas na amostra 1, com tamanhos de aproximadamente 610, 680, 745, 785 e 815 Kb enquanto a amostra 2 apresentou duas bandas difusas nas regiões de 785 e 680 Kb.

Estes resultados são substancialmente diferentes daqueles obtidos para *Azospirillum brasilense* estirpe Sp7 e sua derivada estirpe FP2 por outros autores. CABALERO-MELLADO *et al.* (1999), utilizando eletroforese tipo Eckhartd, reportaram a presença de apenas 2 replicons na faixa de 600 a 800 Kbp. Posteriormente, MARTIN-DIDONET *et al.* (2000) utilizando PFGE (Amersham Biosciences), que possui maior capacidade de resolução do que eletroforese tipo Eckhartd, observaram a presença de 3 replicons: 630, 700 e 810 Kbp (para

comparação dos resultados veja a TABELA 3). O perfil mostrado na FIGURA 2 sugere que a composição de replicons da estirpe FP2 de *A. brasilense* apresenta maior complexidade do que anteriormente descrito.

FIGURA 2 - PERFIL ELETROFORÉTICO EM CAMPO PULSADO DO DNA INTACTO DE *A. brasilense* ESTIRPE FP2 UTILIZANDO PULSOS ESCALONADOS



Perfil eletroforético em campo pulsado do DNA intacto de *A. brasilense* estirpe FP2. Os parâmetros da eletroforese foram: pulsos de 60 segundos por 15 h e 90 segundos por 8 h, 6v/cm, ângulo de reorientação de 120°, 14°C. Os cromossomos de Saccharomyces cerevisae (Amersham Biosciences) foram utilizados como marcador de tamanho molecular. Nas linhas 1 e 2 foram aplicadas duas preparações diferentes de DNA intacto de *A. brasilense* denominadas amostras 1 e 2, respectivamente. Na amostra 1 foi utilizada uma cultura de células com densidade ótica (D.O₆₀₀) de 0,6 e na amostra 2 uma cultura com densidade ótica (D.O₆₀₀) de 0,85.

Dois dos parâmetros que mais influenciam a capacidade de resolução de PFGE são o tempo de pulso e a concentração de DNA. O tempo de pulso (*switch time*) é a duração do campo elétrico em cada orientação e seu valor determina a faixa de tamanho das moléculas que serão separadas. Este período pode ser programado de forma escalonada, produzindo zonas de resolução distintas no gel, ou em gradiente contínuo, produzindo uma migração mais homogênea e linear das moléculas dos DNA de diferentes tamanhos. Tempos de pulso longos, permitem que moléculas de tamanhos maiores sejam separadas, mas também diminuem a resolução devido ao espalhamento das bandas. Por outro lado, tempos de pulso curtos permitem a separação de moléculas de DNA menores e com maior resolução (LEE *et al.*, 1996).

A concentração final de DNA nas amostras de PFGE depende da quantidade inicial de células, mas também varia em função da eficiência da lise, da digestão da parede e proteínas celulares e da quebra mecânica do DNA de alta massa molecular. Estes fatores implicam numa grande variação do rendimento de DNA de alta massa molecular entre diferentes preparações (LEE *et al.*, 1996). Amostras contendo pouco DNA produzem bandas muito fracas, de difícil visualização. Por outro lado, preparações muitas concentradas em DNA produzem bandas com baixa definição, uma vez que bandas de massas moleculares próximas podem co-migrar.

A FIGURA 3 mostra o eletroforetograma do PFGE desenvolvido durante 24 horas utilizando um gradiente de pulsos de 40 a 90 segundos. A temperatura, voltagem, concentração de agarose e tampão foram os mesmos da FIGURA 2, bem como as duas amostras utilizadas. Em primeiro lugar, a corrida eletroforética com gradiente linear de pulsos não propiciou uma maior resolução das moléculas de DNA com tamanho na faixa de 600 a 800 Kbp. Isto pode ser observado pela separação dos cromossomas de Saccharomyces cerevisae utilizados com padrão de tamanho molecular (compare as FIGURAS 2 e 3). Por exemplo, os cromossomas de 815, 785 e 745 Kbp da amostra-padrão estão bem resolvidas na FIGURA 2, mas formam uma banda alongada de difícil distinção na FIGURA 3. Além disso, na amostra 1 (linha 1, FIGURA 3) duas bandas são claramente discerníveis na região de 600 a 800 Kbp, enquanto que na amostra 2 (linha 2, FIGURA 3) três bandas (650, 700 e 800 Kbp)

podem ser observadas. O padrão na linha 2 da FIGURA 3 é idêntico àquele descrito por MARTIN-DIDONET *et al.* (2000). Estes resultados indicam que tanto a concentração inicial de células no bloco de agarose quanto o método de separação podem influenciar substancialmente o resultado da eletroforese em campo pulsado. Para separação de moléculas da DNA na faixa de 600 a 800 Kbp o melhor resultado obtido foi com o aparelho CHEF-DR III (Bio-Rad), agarose para eletroforese em campo pulsado (1%), tampão TBE 0,5X, temperatura de 14°C, ângulo de reorientação de 120° e pulso de 60 segundos por 15 horas e 90 segundos por 8 horas.

Para confirmar a presença dos dois novos replicons na estirpe FP2 de A. brasilense, novas preparações de DNA foram feitas, utilizando três concentrações de células, e o DNA separado nas mesmas condições da FIGURA 2 (FIGURA 4). Nas linhas 1 e 2 da FIGURA 4, onde foram utilizados culturas com densidades óticas $(D.O_{600})$ de 1,2 e 0,86, respectivamente, os replicons não estão bem resolvidos na faixa de 600 a 800 kb e apenas 3 bandas podem ser claramente assinaladas (tamanhos aproximados de 620, 700 e 830 Kbp); uma guarta banda de aproximadamente 680 Kbp pode ser inferida pelo formato da banda de 700 Kbp. Por outro lado, com a diminuição da quantidade de células aplicada na linha 3 (D.O₆₀₀ de 0,6) 5 bandas podem ser visualizadas (tamanhos aproximados de 610, 680, 720, 780 e 810 Kbp), embora as bandas 780 e 810 Kbp não estejam claramente definidas, um padrão praticamente idêntico àquele da FIGURA 2. Na região abaixo de 295 Kbp foram identificadas 2 bandas nas linhas 1, 2 e 3. Estas bandas correspondem aos plasmídeos de 166 e 216 Kbp descrito por PLAZINSCKI et al. (1983). Na região de 1900 Kb, apenas 1 banda pode ser visualizada em todos os experimentos, que provavelmente corresponde aos replicons de 1740 e 2500 Kbp descritos por MARTIN-DIDONET et al. (2000). Nas condições utilizadas nos eletroforetogramas das FIGURAS 2 a 4, esta é uma região de compressão, o que provavelmente impediu a resolução das bandas.

FIGURA 3 - PERFIL ELETROFORÉTICO EM CAMPO PULSADO DO DNA INTACTO DE *A. brasilense* ESTIRPE FP2 UTILIZANDO UM GRADIENTE DE PULSOS



A.b FP2

Perfil eletroforético em campo pulsado do DNA intacto de *A. brasilense* estirpe FP2. Os parâmetros da eletroforese foram: gradiente de pulsos de 40 a 90s durante 24 h, 6v/cm, campos elétricos orientados a 120° e temperatura de 14°C. Os cromossomos de Saccharomyces cerevisae (Amersham Biosciences) foram utilizados como marcador de tamanho molecular. Nas linhas 1 e 2 foram aplicadas as mesmas amostras da figura 2.Na amostra 1 foi utilizada uma cultura de células com densidade ótica (D.O₆₀₀) de 0,6 e na amostra 2 uma cultura com densidade ótica (D.O₆₀₀) de 0,85.



FIGURA 4 - PERFIL ELETROFORÉTICO EM CAMPO PULSADO DO DNA INTACTO DE *A. brasilense* ESTIRPE FP2.

Perfil eletroforético em campo pulsado do DNA intacto de *A. brasilense* estirpe FP2. Os parâmetros da eletroforese foram pulsos de 60 s por 15 h e 90s por 8 h, 6v/cm, 120°, 14°C. Os cromossomos de Saccharomyces cerevisae (Amersham Biosciences) foram utilizados como marcador de tamanho molecular. Nas linhas 1, 2 e 3 foram aplicadas três preparações diferentes de DNA intacto de *A. brasilense* denominadas amostras 1, 2 e 3 respectivamente. Na amostra 1 foi utilizada uma cultura de células com densidade ótica $(D.O_{600})$ de 1,2; na amostra 2 uma cultura com densidade ótica $(D.O_{600})$ de 0,86 e na amostra 3 uma cultura com densidade ótica $(D.O_{600})$ de 0,86 e na

^a Aparentemente, existe dois replicons.

Juntos, os resultados mostram que *A. brasilense* estirpe FP2 possui pelo menos 7 megareplicons: pAZ1 (2500 Kbp), pAZ2 (1740 Kbp), pAZ3 apresentando provavelmente dois replicons (815 e 785 Kbp), pAZ4 (745 Kbp), pAZ5 (680 Kbp) e pAZ6 (630 Kbp), além dos megaplasmídios de 216 e 166 Kbp (TABELA 4). MARTIN-DIDONET (2000) mostrou que os replicons pAZ1 (2500 Kbp), pAZ2 (1740 Kbp), pAZ3 (810 Kbp) e/ou pAZ4 (780 Kbp), e pAZ5 (630 Kbp) hibridizam com sondas de 16S rDNA e 23S rDNA. Estes resultados também destacam a importância do uso de diversas condições para otimizar a identificação de megareplicons de alta massa molecular de bactérias por meio de eletroforese em campo pulsado.

THEELING THEMENO DE THE ENGLING DE THE BROWN ENGLINE LIGTING DE THE				
Trabalho	Estirpe	Número de	Tamanho dos replicons (kb)	Técnica
		repricons		
PLAZINSCKI et al (1983)	Sp7	4	166, 216, 583, 616	Lise alcalina e eletroforese comum
CABALLERO-MELLADO et al (1999)	Sp7	6	130, 570, 630 ^a , 1700, >1800	Eletroforese Eckhardt
MARTIN-DIDONET et al (2000)	Sp7 e FP2	7	200, 210, 630, 700, 810, 1740, 2500	PFGE (Amersham Bioscience)
Este trabalho ^b	FP2*	8	<225, <295, >610, 680, 745, 785 ^c ,~1900	PFGE sistema CHEF-DR III (Bio-Rad)

TABELA 3 - NÚMERO DE REPLICONS DE A. brasilense ESTIRPE Sp7.

^a Aparentemente, existe dois replicons (nota do autor).

^b O tamanho dos replicons é estimado apenas por comparação visual com os marcadores de tamanho molecular de Saccharomyces cerevisae (Amersham Biosciences), tendo como referência a amostra 3 da figura 4.

^c Aparentemente, existe dois replicons.

*A estirpe FP2 é um mutante espontâneo resistente a estreptomicina e ácido nalidíxico da estirpe Sp7. O padrão de plasmídios e megareplicons das duas estirpes é idêntico (MARTIN-DIDONET *et al.*, 1999; PEDROSA, FUNAYMA & RIGO, resultados não publicados).

4.2 PURIFICAÇÃO E AMPLIFICAÇÃO DO REPLICON pAZ6 DE A. brasilense FP2

Para a construção de bibliotecas aleatórias do replicon pAZ6 foram realizadas eletroforeses preparativas em campo pulsado em gel de agarose de baixo ponto de fusão. A banda correspondente ao replicon foi visualizada sob luz ultravioleta e uma fatia de gel contendo o DNA de interesse foi retirada e purificada por uma segunda eletroforese (eletroforese preparativa II) conforme mostra a FIGURA 5. Inicialmente, o DNA foi extraído da fatia do gel pela hidrólise da agarose com agarase ou pelo método da extração com fenol (Material e Métodos, item 3.5.2 e 3.5.2.1) ou ainda por uma combinação dos dois métodos. Entretanto, as quantidades de DNA recuperadas foram muito pequenas e insuficientes para a construção das bibliotecas aleatórias representativas.

A necessidade de grandes quantidades de DNA para a construção das bibliotecas requeria a utilização de métodos de amplificação de DNA. A reação em cadeia da polimerase (PCR) permite a produção de DNA em grandes quantidades a partir de, em princípio, uma única molécula molde. Dois métodos baseados em PCR foram descritos em 1992 para a amplificação do DNA genômico. A primeira estratégia é conhecida como pré-amplificação por extensão de oligonucleotídio iniciador (Primer Extension Preamplification ou PEP) (ZHANG et al., 1992). Neste método, o DNA genômico é digerido com enzimas de restrição e os fragmentos resultantes são ligados a adaptadores. Em seguida, uma reação de PCR é realizada utilizando iniciadores que anelam às sequências dos adaptadores. A segunda estratégia é conhecida como PCR com oligonucleotídios degenerados (Degenerate oligonucleotide primed ou DOP-PCR) (TELENIUS et al., 1992, CHEUNG & NELSON, 1996), que emprega uma coleção de oligonucleotídios iniciadores (primers) aleatórios para conseguir uma amplificação não-específica do DNA genômico. Apesar das duas técnicas serem importantes, nenhuma pode realmente amplificar o genoma inteiro. Uma amostra de fragmentos genômicos pequenos pode ser útil para aplicações como a genotipagem, mas não para a construção de uma biblioteca genômica representativa.

Um terceiro método para amplificação de DNA total descrito recentemente é baseado no conceito da amplificação por deslocamento de fita (*Strand Displacement Amplification* ou *SDA*) e utiliza a DNA polimerase do bacteriófago ¢ 29. Esta polimerase sintetiza DNA pelo mecanismo de círculos rolantes (DEAN *et al.*, 2001, HAWKINS *et al.*, 2002) e o produto, que é DNA simples fita, serve de molde para um novo ciclo de síntese utilizando oligonucleotideos degenerados como iniciadores. Assim, como a PCR, a SDA é um processo cíclico que resulta em um aumento exponencial do número de cópias do DNA alvo. Entretanto, diferente da PCR, a SDA é isotérmica, ou seja, todas as fases são realizadas numa mesma temperatura.

O uso de iniciadores aleatórios para SDA permite a rápida amplificação do DNA genômico (HAWKINS *et al.*, 2002). Estes autores puderam obter mais de 30 microgramas de DNA como produto de amplificação a partir de quantidades muito pequenas de material de partida, equivalente a menos que cinco cópias do genoma humano. Outra vantagem desta técnica inclui a maior acurácia e processividade da polimerase do fago ϕ 29 comparada com as enzimas de PCR como *Taq* DNApolimerase ou *Pfu* DNApolimerase, resultando em fragmentos de DNA com tamanho acima de 10 kb e com uma freqüência de erro menor que 1 em 10⁶.

Comparações entre bibliotecas genômicas obtidas de DNA amplificado por SDA e de DNA purificado pelo método tradicional (extração do DNA, fragmentação e clonagem em plasmídeo), confirmaram que SDA gera um conjunto de fragmentos representativos da amostra inicial. DETTER *et al.* (2002) demonstraram que o genoma bacteriano total pode ser efetivamente amplificado de células ou pequenas quantidades de DNA genômico purificado. Bibliotecas aleatórias (*shotgun*) foram preparadas do DNA genômico isolado de culturas de *Xylella fastidiosa* e do DNA produzido por SDA utilizando cerca de 1000 células de *X. fastidiosa* (sem o isolamento do DNA). Em ambos os casos, aproximadamente 3000 seqüências foram obtidas e alinhadas com o genoma da *X. fastidiosa*. As duas bibliotecas apresentaram coberturas semelhantes: 39% para a biblioteca construída de maneira convencional e 34% para a biblioteca SDA. Estes são também semelhantes à cobertura esperada (40% levando em conta o tamanho do genoma da *Xylella* e a quantidade de seqüências geradas). Neste trabalho o DNA do replicon pAZ6 foi amplificado pelo método SDA utilizando o Kit TempliPhiTM (Amersham Biosciences). Este sistema contém a polimerase do fago ¢29 e foi otimizado para produção de DNA molde para sequenciamento, mas era o único sistema de amplificação pelo método SDA disponível comercialmente à época dos experimentos.

Uma vez que algumas substâncias (por exemplo, agarose e componentes de meios de cultura) são capazes de inibir a reação com o sistema TempliPhi, o DNA do replicon pAZ6 foi isolado por eletroforese PFGE em agarose de baixo ponto de fusão, extraído do gel pelo método do fenol (Material e Métodos item 3.5.2) e utilizado nas reações de amplificação isotérmicas. As reações foram preparadas contendo diferentes concentrações de DNA de diferentes preparações. A FIGURA 6 mostra um exemplo de amplificação isotérmica do replicon pAZ6. Os produtos finais da reação são moléculas de DNA dupla fita de grande tamanho molecular, o que é evidenciado pela baixa taxa de migração do amplificado e pela presença de DNA no ponto de aplicação. A TABELA 4 resume os resultados de amplificação de 5 diferentes preparações do replicon pAZ6, que produziu no total cerca de 300 microgramas de DNA.

TABELA 4- CONCENTRAÇÃO DO DNA DO REPLICON pAZ6 AMPLIFICADO POR TEMPLIPHI.

Amostra	Volume total (10µL)
2609	60µg
0811	7µg
1211	65µg
2911	170µg

FIGURA 5 - ISOLAMENTO DO REPLICON pAZ6 DE A. brasilense FP2 POR ELETROFORESE EM CAMPO PULSADO



Isolamento do replicon pAZ6 de *A. brasilense* estirpe FP2 utilizando eletroforese em campo pulsado. Os parâmetros da eletroforese foram: pulsos de 60 segundos por 15 h e 90 segundos por 8 h, 6v/cm, ângulo de reorientação de 120°, 14°C. Os cromossomos de Saccharomyces cerevisae (Amersham Biosciences) foram utilizados como marcador de tamanho molecular. A linha 1 mostra o perfil eletroforético do DNA intacto de *A. brasilense* FP2. As linhas 2, 3 e 4 representam diferentes pedaços da mesma fatia do gel preparativo II que continha o replicon pAZ6 .purificado.

FIGURA 6 - AMPLIFICAÇÃO ISOTÉRMICA DO REPLICON pAZ6.



Amplificação do replicon pAZ6 de *Azospirillum brasilense* estirpe FP2 utilizando hexâmeros aleatórios e o método de amplificação por deslocamento de fita do fago ϕ 29. Um μ L do DNA do replicon pAZ6 foi amplificado com TempliPhi durante 18 horas à 30°C. A eletroforese foi realizada em tampão TBE 1X, agarose 0,8% a 40V. A linha 1 mostra o marcador de tamanho molecular 1 Kb Ladder e a linha 2 a amostra 2911 (1 μ L do amplificado).

4.3 FRAGMENTAÇÃO MECÂNICA DO DNA DO REPLICON pAZ6 **AMPLIFICADO COM TEMPLIPHI**

Para a obtenção de fragmentos de 1 a 5 Kbp, os produtos de amplificação foram quebrados por nebulização durante 1 minuto e 30 segundos a 2,0 Kgf/cm² com gás nitrogênio (Material e Métodos item 3.8). Para verificar o tamanho dos fragmentos obtidos, uma amostra foi analisada por eletroforese em gel de agarose (1%). As condições de nebulização foram otimizadas para produzir fragmentos com tamanho entre l e 2 Kbp (resultado não apresentado).

4.4 CLONAGEM DO DNA DO REPLICON pAZ6 AMPLIFICADO COM **TEMPLIPHI**

As extremidades dos fragmentos (1 - 2 kb) foram tornadas cegas com as enzimas T4 DNApolimerase e Klenow e então defosforiladas com CIP (Calf Intestinal Phosphatase) (Material e Métodos item 3.9). O DNA foi clonado no vetor pCR® 4 Blunt-TOPO® e transformado em E. coli TOP10 por eletroporação.

No total foram construídas cinco bibliotecas a partir de diferentes amostras amplificadas com o sistema TempliPhi. O número de clones em cada biblioteca e a quantidade de clones seqüenciados está na TABELA 5.

TABELA 5 - BIBLIOTECAS ALEATORIAS DO REPLICON PAZO DE A. brasilense FP2.				
Biblioteca	Número de clones	Clones seqüenciados		
AP1	50	34		
AP2	1536	1344		
AP3	672	-		
AP5	1248	88		
AP6	2496	192		
Total	6002	1858		

DIDLIGHT ALL TOTAL OF DEDLIGOUL ARC DE AL TD0

4.5 SEQÜENCIAMENTO DE CLONES ALEATÓRIOS DO REPLICON pAZ6.

Dos 6002 clones obtidos, 1858 foram seqüenciados em um seqüenciador automático MegaBace1000 pelo método do terminador de cadeia marcado e utilizando o oligonucletídio iniciador Reverso (Material e Métodos item 3.12). As seqüências produzidas foram processadas pelos programas PHRED e PHRAP. O programa PHRED identifica cada base nitrogenada e atribui um valor de qualidade, que mede o grau de confiança na determinação daquela base (ver Material e Métodos item 3.13). As leituras foram consideradas de boa qualidade quando tinham pelo menos 400 pares de bases de comprimento com qualidade igual ou superior a 20 (TABELA 6).

Dos 1858 clones seqüenciados apenas 546 passaram pelo critério de qualidade arbitrado. Esta eficiência (29%) foi bastante baixa. A proporção total de bases com qualidade igual ou superior a 20 foi ligeiramente maior (33%) (TABELA 7). Uma das prováveis causas desta baixa eficiência é o alto conteúdo de (G+C) do genoma de *A. brasilense*, o que aumenta o número e a estabilidade de estruturas secundárias intramolecular do DNA molde dificultando, portanto, tanto a reação de seqüenciamento quanto a corrida eletroforética. Cerca de 88,5% das leituras de alta qualidade continham seqüências novas, ou seja, não-vetor, um valor aceitável para bibliotecas aleatórias (TABELA 7), indicando que as bibliotecas foram de boa qualidade quanto à freqüência de insertos.

TABELA 6 - RESUMO DO SEQÜENCIAMENTO DE 1858 CLONES DAS
BIBLIOTECAS ALEATÓRIAS DO REPLICON pAZ6 DE A.
brasilense FP2

	Número de Leituras
Leituras utilizando primer universal (FWD ou ponta b)	0
Leituras utilizando o primer reverso (REV ou ponta g)	1.858
Leituras não identificadas	0
Total de Leituras	1.858
Total de Leituras Aceitas ^A (excluindo vetores)	483
Total de Leituras Aceitas ^A (incluindo vetores)	546
Comprimento Médio de Leitura	673,80
Comprimento Médio de Leitura (qualidade ^B \geq 20)	208,26
Total de Placas Submetidas	22
Número Médio de Leitura por Placa	84,45
Total de Leituras com mais de 10% de vetor	416
Total de Leituras com vetores entre 10% e 80%	402
Total de Leituras com mais de 80% de vetor	14

^A Leituras aceitas são sequências de boa qualidade, ou seja, que atingem 400 pares de base com qualidade PHRED igual ou superior a 20.

^B Critério de qualidade: valor de PHRED igual ou superior a 20

TABELA 7 - ANÁLISE ESTATÍSTICA DA QUALIDADE DAS BASES DAS SEQÜÊNCIAS OBTIDAS

	Número de bases
Total de Bases (incluindo vetores)	1.251.959
Total de Bases com Qualidade ^A \geq 20 (incluindo vetores)	416.644
Total de Bases com Qualidade ^A ≥30 (incluindo vetores)	254.705
Total de Bases (excluindo vetores)	1.159.622
Total de Bases com Qualidade ^A \geq 20 (excluindo vetores)	386.952
Total de Bases com Qualidade ^A \geq 30 (excluindo vetores)	240.122
Qualidade Média por Base ^A	16,80

A análise foi feita pelo Programa PHRED e CROSS-MATCH.

^A Critério de qualidade: valor de PHRED igual ou superior a 20.

4.6 AGRUPAMENTO DAS SEQÜÊNCIAS EM CONTÍGUOS

Todas as seqüências obtidas foram verificadas quanto à presença de seqüências de vetor pelo programa Cross-Match e então foram alinhadas e sobrepostas pelo programa PHRAP (*Fragment Assembly Program*). Este programa utiliza um mosaico de partes das seqüências individuais com alta qualidade para a construção do alinhamento, e também fornece todas as informações sobre a montagem realizada (PROSDOCIMI *et al.*, 2003). Os resultados deste alinhamento são seqüências formadas pela sobreposição parcial das leituras individuais e que são denominadas contíguos (*contigs*) (TABELA 8). A visualização e edição das seqüências geradas pela montagem são realizadas através do programa CONSED (PROSDOCIMI *et al.*, 2003).

As 1858 seqüências obtidas produziram 144 contíguos com um tamanho total de 84.888 bases não redundantes ou únicas (TABELA 8). A qualidade média dos contíguos foi de 38 e representam uma cobertura do replicon pAZ6 de 13,5%. Considerando as leituras de clones únicos ou singlets a cobertura foi de aproximadamente 34% (TABELA 8).

O conteúdo G+C dos contíguos variou de 63,8 a 73,6 %. Estes valores coincidem com a composição percentual G+C característica de *A. brasilense* que varia de 67 a 70% (TARRAND *et al.*, 1978).

Total de Contíguos	144
Total de Contíguos com 1 leitura	6
Maior Contíguo	1.613
Menor Contíguo	31
Tamanho Médio	589
Quantidade de Leituras para Montagem de Contiguos	1.170
Maior Quantidade de Leituras/Contíguos	117
Menor Quantidade de Leituras /Contíguos	1
Tamanho Total dos Contíguos	84.888
Qualidade Média das Bases dos Contíguos	37,99
Total de Singlets	688
Total de Singlets Não-Vetor	192
Maior Singlet	3.313
Tamanho Médio dos Singlets	761
Soma dos Singlets Não-Vetor (trimmed)	129.322
Cobertura do genoma	Pares de bases
Tamanho Estimado do Genoma	630.000
Cobertura com Contíguos	84.888
Cobertura com Singlets	129.322

TABELA 8: ESTATÍSTICA DA MONTAGEM DOS CONTÍGUOS DO REPLICONpAZ6 DE Azospirillum brasilense FP2

O tamanho máximo de contíguo obtido foi de 1.613 bp. Este pequeno tamanho dos contíguos formados deve-se ao baixo número e qualidade das seqüências obtidas.

Total Coberto

A análise dos resultados de montagem revelou um alto grau de redundância das seqüências. Esta redundância provavelmente foi em decorrência do método de construção das bibliotecas e sugere que nem todo o DNA extraído do gel de agarose serviu como molde para a amplificação isotérmica. Esta conclusão é apoiada pelo fato de que as seqüências redundantes ocorreram entre clones de uma mesma biblioteca, mas não entre clones de diferentes bibliotecas.

214.210 (34%)

4.7 ANÁLISE DOS CONTÍGUOS

Para determinar quais genes estavam presentes nas seqüências dos contíguos obtidos, as prováveis regiões codificadoras foram identificadas pelo programa FramePlot (ISHIKAWA & HOTTA, 1999). Este programa identifica provável gene baseado na freqüência de G+C na terceira posição. As seqüências de aminoácidos deduzidas das prováveis regiões codificadoras foram comparadas com o banco de dados GenBank utilizando o programa BlastP (ALTSCHUL *et al.*, 1997). Finalmente, a seqüência completa de aminoácidos da proteína com maior similaridade foi recuperada do banco de dados alinhada com a seqüência de aminoácidos deduzida do contíguo analisado utilizando o programa ClustalW (THOMPSON *et al.*, 1994). Este processo permitiu identificar os prováveis genes presentes nas seqüências que são semelhantes a genes de outros organismos já seqüenciados. A FIGURA 7 mostra um exemplo dos resultados obtidos pelo uso do programa FramePlot.

FIGURA 7 - PREDIÇÃO DE REGIÕES CODIFICADORAS DE PROTEÍNAS DE UM CONTÍGUO PELO PROGRAMA FRAMEPLOT



As curvas dão a probabilidade calculada (ordenada) da presença de um gene em relação ao número de pares de bases da seqüência (abcissa). Quando a linha ultrapassa o valor limiar dado pela linha pontilhada a região é considerada com alta probabilidade de codificar para uma proteína. O método não discrimina orientação, assim o gene pode estar no sentido 5'-3' ou no sentido 3'-5'. As setas acima do gráfico representam as ORF hipotéticas nas seis fases de leitura.

Nos 144 contíguos analisados foram identificados 37 prováveis genes que codificam para proteínas conhecidas. Destes, 26% possuem alta similaridade com proteínas hipotéticas ou possuem baixa similaridade com seqüências do banco de dados; 48% não apresentaram nenhuma similaridade com seqüências de proteínas ou de nucleotídeos depositadas nos bancos de dados públicos.

A TABELA 9 resume os resultados das análises das 37 prováveis ORFs identificadas nos contíguos do replicon pAZ6.

TABELA 9 - PROTEÍNAS CODIFICADAS PELO REPLICON pAZ6 DE <i>A. brasilense</i> ESTIRPE FP2.	G+C ^A Alinhamento ^B E Função (%)	88,5 57/107 7e-22 Metbolismo e transporte de carboidratos	79,1 60/91 9e-30 Metabolismo do piruvato	100 42/73 8e-14 Metabolismo e transporte de aminoácidos	75 71/109 3e-35 Função não determinada	90,8 83/189 3e-28 Transdução de sinal	73,9 45/88 2e-15 Transporte de açúcar	97,4 111/159 8e-61 Metabolismo de ácidos graxos	70,3 36/60 2e-14 Metabolismo de aminoácidos	63 53/54 5e-23 Captação de galactose e quimiotaxia para açúcares	97,7 34/77 4e-10 Fosforilação oxidativa
	Microrganismo	Agrabacterium tumefaciens	Zymomonas mobilis	Bradyrhizobium japonicum	Bordetella pertussis	Magnetospirillum magnetotacticum	Bordetella pertussis	Burkholderia fungorum	Pseudomonas syringae	Azospirillum brasılense	Burkholderia fungorum
	Proteína codificada Gene	Transportador ABC, proteína ATPase de ligação ao nucleotídeo	Enzima ativadora piruvato <i>pflA</i> formato liase	Transportador ABC, proteína de ligação ao ATP	Provável proteína exportada	Partícula GTPase reconhecedora de sinal	Transportador de açúcar ABC	Acetil-CoA desidrogenase fadE	Aldeído desidrogenase NAD- dependente	Receptor periplasmático sbpA ChvE de ligação a açúcar	Citocromo tino-bd aninol
	N° PHRED Contig (valor médio)	7 31,50	9 40,70	16 33,30	18 72,40	27 25,40	31 28,80	33 32,50	34 51,60	37 48,60	43 33 50

ļ					e	•	o	-		6	
DE FP2.	Função	Fosforilação oxidativa	Metabolismo de ácido urônico	Metabolismo de carbiodrato	Metabolismo e transport de aminoácidos	Metabolismo de metano	Metabolismo e transport de carboidratos	Regulador transcricional	Doador de grupos metil	Quimiotaxia	
	ш	4e-80	8e-18	4e-19 3e-24		le-28	8e-23	5e-80	3e-13	le-17	
	Alinhamento ^B	141/172	46/78	47/63	57/70	60/78	54/78	143/187	32/49	72/182	
ESTIRI	G+C ^A (%)	85,1	85,2	83,7	72,2	78	94,1	76,8	79,8	83,3	
A 9 - PROTEÍNAS CODIFICADAS PELO REPLICON pAZ6 DE A. brasilense	Microrganismo	Magnetospirillum magnetotacticum	Ralstonia solanacearum	Magnetospirillum magnetotacticum	Ralstonia metallidurans	Methylobacterium extorquens	Bordetella parapertussis	Agrabacterium tumefaciens	Magnetospirillum magnetotacticum	Magnetospirıllum magnetotacticum	
	Gene	onu	uxaA	ŀ		fdh1B	•	·	ı	mcpA	
	Proteína codificada	NADH:ubiquinona oxidoredutase subunidade 20Kd e Fe-S oxidoredutase relacionada	Provável proteína altronato hidrolase (ÁCIDO AL TRONICO HIDRATASE)	Fosfoinanomutase	Sistema de transporte de níquel/oligopeptídeo/dipeptídeo tipo ABC, componente ATPase	Subunidade beta da formiato desidrogenase contendo tungstênio	Provável componente de ligação ao ATP do transportador ABC	Regulador transcricional, família TetR	Metiltransferase SAM- dependente	Proteína quimiotáxica que aceita metilação	
	PHRED (valor médio)	43,90	25,50	22,20	28,30	28,50	26,20	30,10	11,50	29,50	
TABEL	N° contig	45	54	55	56	57	59	71	72	73	
	^B E Função	5e-15 Função não determinada	2e-30 Transporte ligado a próton	2e-25 Conversão e produção de energia	1e-10 Metabolismo de carboidratos	6e-12 Metabolismo e transporte de aminoácidos	4e-12 Função não determinada	3e-12 Biossíntese de lipídeos	8e-16 Metabolismo e transporte de aminoácidos	3e-18 -	2e-40 Metabolismo e transporte
---	-------------------------	---	----------------------------------	--	--------------------------------------	---	-------------------------------------	--	---	---------------------	--------------------------------
E FP2.	Alinhamento	40/53	64/91	58/96	31/42	36/45	39/109	34/42	46/107	65/200	73/109
se ESTIR	G+C ^A (%)	70,5	88,5	87,3	67,6	77,7	86,3	67,4	94,8	80,0	74,8
TEÍNAS CODIFICADAS PELO REPLICON pAZ6 DE A. brasilens	Microrganismo	Rhodospirillum rubrum	Azotobacter vinelandii	Bradyrhizobium japonicum	Bradyrhizobium japonicum	Sinorhizobium meliloti	Magnetospirillum magnetotacticum	Bradyrhizobium japonicum	Ralstonia solanacearum	Micrococcus luteus	Mesorhizobium loti
	Gene		ı	hycE	uxuB	·	ı	fabl	•	·	•
	Proteína codificada	Sistema de transporte não caracterizado tipo-ABC, componente permease	Importador H+/gluconato	Componente [NiFe]da hidrogenase III	Manitol desidrogenase	Provável sistema de transporte spermidina/putrescina, proteína transportadora ABC	Proteína externa de membrana	Proteína NADH-enoil acil-ACP redutase	Provável substrato de ligação da proteína transportadora ABC periplasmática	Proteína hipotética	Transportador de aminoácido
A 9 - PROT	PHRED (valor méio)	67,40	59,10	34.90	37,80	50,60	64,40	34,50	31,20	32,70	52,50
TABEL/	N° Contig (76	77	84	89	92	94	76	105	106	107

61

TABEL	TOA9 - 9 A.	TEÍNAS CODIFICADAS PEL	O REPLIC	ON pAZ6 DE A. brasilens	e ESTIRI	PE FP2.		
N° Contig	PHRED (valor médio)	Proteína codificada	Gene	Microrganismo	G+C ^A (%)	Alinhamento ^B	ш	Função
112	68,60	D-manonato oxidoredutase	uxuB	Escherichia coli CFT073	69,6	39/81	8e-12	Interconversão de pentose e glucoronato
115	64,30	Proteína permease transportador ABC	ı	Bradyrhizobium japonicum	71,3	42/55	8e-18	Metabolismo e transporte de aminoácidos
121	66,80	Proteína da biosíntese de cobalamina CobT (nicotinato- mononucleotídeo: 5,6- dimetilbenzamidazol fosforibosiltransferase)	cobT	Magnetospirillum magnetotacticum	72,8	36/42	2e-13	Biosintese de cofator
122	59,80	Domínio PAS/PAC	ı	Ralstonia metallidurans	88,3	56/96	2e-27	Metabolismo de sinalização
127	71,00	Fosforibosilpirofosfato sintetase	prsA	Magnetospirillum magnetotacticum	68,8	31/35	9e-11	Metabolismo de nucleotídeos
136	75,10	Sistema de transporte de íon metal tipo ABC, componente periplasmático/adesina de superfície	ı	Rhodospirıllum rubrum	85,1	47/62	2e-22	Transporte de íons
137	37,80	Provável proteína glutamato- amonia-ligase adenililtransferase (glutamina sintetase adenililtransferase)	glnE	Ralstonia solanacearum	82,6	105/200	3e-51	Metabolismo de nitrogênio
144	57,50	UDP-N-acetilglucosamina 1-carboxiviniltransferase	murA	Xanthomonas campestris	80,8	32/37	3e-11	Biossíntese de parede celular
^A Valor p	ercentual de C +	G na terceira posição.						

^B O alinhamento mostra o número de aminoácidos idênticos/ número de aminoácidos alinhados. O valor de E indica a probabilidade do alinhamento das seqüencias de aminoácidos serem ao acaso.

62

4.8 COMPLEMENTO GÊNICO PARCIAL DO REPLICON pAZ6

A análise das seqüências dos contíguos obtidos permitiu a identificação de genes envolvidos na biossíntese de parede celular, vitamina B_{12} , lipídeos e nucleotídeos, no catabolismo de carboidratos e ácidos graxos, no metabolismo de energia e no transporte de solutos.

4.8.1 BIOSSINTESE DE PAREDE CELULAR

A seqüência parcial de aminoácidos deduzida da provável região codificadora identificada pelo programa FramePlot no contig 144 possui alta similaridade com a proteína deduzida do gene *murA* de *Xanthomonas campestris* (86% de resíduos idênticos em 37 alinhados) (FIGURA 8). O gene *murA* codifica para a proteína UDP-N-acetilglucosamina 1-carboxivinil transferase (MurA transferase). Esta enzima catalisa a transferência do grupamento enolpiruvato do fosfoenolpiruvato para o difosfouridina N-acetilglucosamina liberando o fosfato:

Fosfoenolpiruvato + N-acetil-D-glucosamina-UDP - N-acetilcarboxivinil-D-glucosamina-UDP + Pi

Esta é a primeira etapa da via de biossíntese do peptideoglicano da parede celular bacteriana (ZEMEL & ANWAR, 1975). A molécula de N-acetilcarboxivinil-D-glucosamina-UDP formada é então reduzida a ácido N-acetilmurâmico-UDP, um dos componentes da parede celular bacteriana.

FIGURA 8 - ALINHAMENTO DAS SEQÜÊNCIAS DE AMINOÁCIDOS DEDUZIDAS DA SEQÜÊNCIA DE NUCLEOTÍDEOS DO GENE murA DE Xanthomonas campestris E DA ORF2C DO CONTÍGUO 144. ETGTFLVAAAMTGGSVTVRRARPDTLDAVLDKLTEAGATITTTADSVTLDMHGKRPRAVN 30 A-Xanthomonas **B**-frame2c/144 H----- 25 ** A-Xanthomonas LTTAPYPAFPTDMQAQFMALNCVADGVGVINETIFENRFMHVNELLRLGADIQVEGHTAI 360 **B**-frame2c/144 -----OTDMQAOFMALNCVAEGVGVINETIFENRFMHVRPTCR-----63 **** * VRGAERLSGAPVMATDLRASASLILAGLVADGDTTIDRIYHLDRGYENIEEKLGALGATI 420 A-Xanthomonas B-frame2c/144 -RSSWRSTAWP-----RAWA----- 77 ** * *.: * :. * RRIA 424 A-Xanthomonas B-frame2c/144 ____

Alinhamento das sequências de aminoácidos deduzidas da sequência de nucleotídeos do gene *murA* de Xanthomonas campestris e da ORF2c do contíguo 144 utilizando o programa ClustalW. A linha A mostra a sequência de aminoácidos do gene *murA* de Xanthomonas campestris. A linha B mostra a sequência de aminoácidos deduzida da provável ORF encontrada na Frame 2c do contig 144 da biblioteca genômica do replicon pAZ6 de A. brasilense FP2..

* indica aminoácidos idênticos; : substituições conservadas; . substituições semi-conservadas

A identidade da seqüência de nucleotídeos do gene *murA* de *Xanthomonas campestris* e a ORF2c do contig 144 é surpreendente e levanta a hipótese de contaminação da biblioteca com DNA exógeno. Cerca de 93% dos nucleotídios da região comparada do gene *murA* são idênticos aos da seqüência obtida do replicon pAZ6. Entretanto, quando os 33 resíduos mais conservados de aminoácidos de MurA (TDMQAQFMALNCVAEGVGVINETIFENRFMHV, FIGURA 8) foram comparados com o banco de dados GenBank, a percentagem de identidade foi de 96, 81, 84, 84 e 84% com proteínas ortólogas de *Xylella fastidiosa, Desulfovibrio desulfuricans, Enterobacter cloacae, Yersinia pestis* e *Chromobacterium violaceum*, respectivamente. Esta alta identidade sugere um alto grau de conservação desta região do gene *murA*.

4.8.2 BIOSSINTESE DE VITAMINA B₁₂

O gene *cobT* foi identificado no contíguo 121 e codifica para a enzima CobT ou nicotinato-mononucleotídeo:5,6-dimetilbenzamidazol fosforibosiltransferase. Esta enzima cataliza condensação dimetilbenzimidazol de e ácido nicotínico mononucleotídeo formando o α -ribasol fosfato. O α -ribasol, produzido pela hidrólise do fosfato, é incorporado a GDP-cobinamida na última etapa da via biossintética da vitamina B₁₂. A enzima CobT é parte da chamada via anaeróbica de biossíntese de cobalamina e está bem caracterizada em Salmonella tiphymurium (SCOTT & **ROESSNER**, 2002).

A vitamina B_{12} (cobalamina) atua como co-enzima em três tipos de reações: rearranjos intramoleculares, metilações (por exemplo na síntese de metionina) e redução de ribonucleotídeos a desoxirribonucleotídeos (STRYER *et al.*, 2002).

4.8.3 BIOSSÍNTESE NUCLEOTÍDIOS

O gene *prsA* que codifica para a proteína fosforibosil pirofosfato quinase também foi encontrado no megaplasmídeo pAZ6 (contíguo 127). Esta enzima é a responsável pela formação do 5'-fosforibosil-1-pirofosfato (PRPP), um doador de unidade ribose fosfato na biossíntese de ribonucleotídeos.

Nucleotídeos desempenham importantes funções em quase todos os processos bioquímicos: são os precursores do DNA e RNA; derivados de nucleotídeos são intermediários ativos em muitas reações biossintéticas (por exemplo, UDP-glicose); nucleotídeos da adenina são componentes de importantes coenzimas (NAD+, FAD e CoA) e do intermediário universal de energia em sistemas biológicos, o ATP; o GTP supre a energia em processos como translocação de cadeias peptídicas nascentes em

ribossomos; e podem atuar como reguladores metabólicos como o AMP cíclico (STRYER et al., 2002).

4.8.4 METABOLISMO DE CARBOIDRATOS

Dois genes que codificam para proteínas envolvidas no metabolismo de ácidos urônicos foram encontrados no replicon pAZ6. O gene *uxaA* (contíguo 54), que faz parte do operon *uxaC-uxaA* em *E. coli*, codifica para a altronato hidrolase, e o gene *uxuB* (contiguo 112), codifica para a D-manonato oxidoredutase.

O metabolismo dos ácidos urônicos possui duas vias (via do galacturonato e do glucoronato) (FIGURA 9) que convergem por meio de uma reação de desidratação para a formação de ácido pirúvico e triose fosfato (ASHWELL *et al.*, 1960). A manonato oxidoredutase, produto do gene *uxaA*, é a segunda enzima da via de degradação do ácido D-glucurônico, enquanto a altronato hidrolase que converte o D-altronato em 2-ceto,3-desoxi,6-fosfogluconato catalisa a última etapa da via específica do galacturonato. Estes resultados sugerem que *A. brasilense* é capaz de metabolizar tanto os ácidos urônicos derivados de glucose e galactose.

FIGURA 9 - ESQUEMA DAS REAÇÕES ENVOLVIDAS NO METABOLISMO DE ÁCIDOS URÔNICOS EM BACTÉRIAS



Esquema de reações envolvidas no metabolismo dos ácidos urônicos (ácido glucurônico e galacturônico) em bactérias. De acordo com esta sequência de reações, ao ácidos glucurônico e galacturônico são isomerizados para formar os seus correspondentes ceto-análogos, ácido fruturônico e tagaturônico respectivamente. Na presença de difosfopiridina nucleotídeo reduzido (DPNH), o ácido fruturônico é reduzido a ácido D-manonico enquanto o ácido tagaturônico forma ácido D-altrônico. Neste ponto, as duas vias convergem por meio de uma reação de desidratação onde o ácido 2-ceto-3-deoxi-D-gluconico é formado. Este último composto é então fosforilado pela adenosina trifosfato (ATP) produzindo ácido 2-ceto-3-deoxi-6-fosfoglucônico que é clivado para formar ácido pirúvico e triose fosfato. FONTE: ASHWELL et al., 1960

4.8.5 METABOLISMO DE ÁCIDOS GRAXOS

O gene *fadE*, que codifica para acil-CoA desidrogenase (ACAD), foi identificado no contíguo 33. Esta enzima cataliza a α , β -desidrogenação dos ácidos graxos conjugados a CoA que são originados da β -oxidação (GHISA &THORPE, 2004). O ciclo de alongamento na síntese de ácidos graxos consiste de quatro etapas: condensação, redução, desidratação e redução. A última etapa, que é redução do transenoil-acil-ACP, é catalizada pela NADH-enoil acil-ACP redutase (STRYER, 2002). O gene *fabI* que codifica para esta proteína foi identificado no contíguo 97 encontrado no pAZ6.

4.8.6 METABOLISMO DE ENERGIA

No megaplasmídeo pAZ6 foram encontrados dois genes que codificam para proteínas da cadeia respiratória. O gene *cydA* que codifica para a subunidade 1 do citocromo tipo-*bd* quinol oxidase, e gene *nuo* que codifica para a subunidade 20 KDa da NADH^{*}ubiquinona oxidoredutase.

Existem poucos estudos sobre a cadeia respiratória em *Azospirillum* spp. Alguns dos componentes da cadeia respiratória de *A. brasilense* foram determinados enzimaticamente e espectroscopicamente em células crescidas diazotroficamente (N₂ e baixas tensões de oxigênio) ou na presença de NH₄⁺ e altas tensões de oxigênio (BURRIS *et al.*, 1978 em PEDROSA, 1988). O oxigênio apresenta um grande efeito sobre a composição das membranas de *A. brasilense*. Quando *A. brasilense* cresce sob baixa tensão de O₂, apresenta citocromos tipo *b* (λ 559 a 560 nm) e *c* (λ 551 a 553 nm) (OKON *et al.*, 1983). Uma banda de absorção adicional (λ 603 nm) correspondente a um citocromo *aa3* foi encontrada em preparações de membranas isoladas de células cultivadas sob alta tensão de oxigênio (OKON *et al.*, 1983). Posteriormente, estudos espectroscópicos de preparações de membranas de *Azospirillum brasilense* revelaram a presença de duas oxidases terminais adicionais, citocromo *o* e *d* em células crescidas em baixa e alta tensão de O₂. As enzimas succinato e NADH oxidase também foram encontradas na membrana de *A. brasilense* e suas atividades aumentam com a tensão de O_2 no meio de cultura. Estudos utilizando inibidores da cadeia respiratória sugeriram a presença de uma cadeia transportadora de elétrons ramificada em *A. brasilense* (OKON *et al.*, 1983).

Em *Azotobacter vinelandii*, que também apresenta uma cadeia respiratória ramificada, o complexo citocromo *bd* está envolvido no mecanismo de tolerância ao oxigênio e na proteção respiratória da nitrogenase. A proteção respiratória consiste no rápido consumo do O_2 por uma oxidase terminal que mantenha uma concentração de O_2 intracelular compatível com a atividade da nitrogenase (KELLY *et al.*, 1990).

A identificação de subunidades da NADH desidrogenase e do citocromo *bd* oxidase suportam a indicação de uma cadeia respiratória ramificada em *A. brasilense* estirpe FP2 provavelmente semelhante a de *Azotobacter venelandii*.

Uma região codificadora de proteína foi identificada no contíguo 57 e a comparação com o bando de dados revelou alta identidade (76% de resíduos idênticos em 78 alinhados) com a subunidade β da formato desidrogenase da alfa-proteobacteria *Methylobacterium extorquens*. A enzima formato desidrogenase NADP-dependente é um heterodímero com duas subunidades $\alpha_1\beta_1$, e responsável pela oxidação do formato a CO₂ com a formação de NADH e H⁺. A formato desidrogenase de *M. extorquens* contém tungstênio (LAUKEL *et al.*, 2003), sugerindo portanto que a enzima codificada pelo replicon pAZ6 de *A. brasilense* também contém este elemento. A presença de genes codificando para formato desidrogenase em *A. brasilense* é suportada por ensaios fisiológicos que mostraram que *A. lipoferum* e *A. brasilense* apresentam a capacidade de utilizar metano, metanol ou formato como única fonte de carbono (SAMPAIO *et al.*, 1982).

Muitos microrganismos usam hidrogênio molecular em suas vias metabólicas como fonte de energia ou para eliminar o excesso de elétrons. A enzima que cataliza reversivelmente a conversão do hidrogênio molecular para dois elétrons e dois prótons são conhecidas como hidrogenases. Estas proteínas são metaloenzimas onde todas contem ferro (Fe) e em muitos casos também níquel (Ni) (DEMENTIN *et al.*, 2004).

No contig 84 foi identificado um gene que codifica para um componente da enzima [NiFe] hidrogenase III. A hidrogenase é uma enzima importante para reciclar a energia durante o processo de fixação de nitrogênio. Isto porque a nitrogenase produz compulsoriamente 1 mol de hidrogênio para cada 2 mols de nitrogênio reduzidos. A oxidação do hidrogênio pela hidrogenase permite a recuperação de uma parte da energia dispendida na redução de prótons pela nitrogenase (POSTGATE, 1992). Muitos organismos possuem mais do que um tipo de hidrogenase. Os resultados obtidos não permitem definir se *A. brasilense* possui apenas uma hidrogenase, que seria codificada pelo replicon pAZ6 e, no caso da ocorrência de mais do que uma hidrogenase, se o gene identificado é essencial para reclicagem de energia durante crescimento diazotrófico.

4.8.7 METABOLISMO DE NITROGÊNIO

O metabolismo do nitrogênio envolve a expressão coordenada de um grande número de enzimas envolvidas com a utilização de fontes de nitrogênio extracelular e biossíntese intracelular de compostos nitrogenados. O controle desta expressão é determinada pela disponibilidade do nitrogênio fixado na célula e é efetuado por complexos regulatórios que envolvem a regulação tanto ao nível transcricional quanto pós-transcricional (MERRICK & EDWARDS, 1995). A glutamina sintetase é a enzima chave no metabolismo do nitrogênio e é responsável pela incorporação do amônio no glutamato para formar a glutamina em baixas concentrações de nitrogênio. A enzima glutamina sintetase adenililtransferase (GlnE) é uma proteína regulatória que está envolvida na controle da atividade da glutamina sintetase. A atividade da glutamina sintetase é regulada em resposta aos níveis de amônio pela adenililação reversível, reação catalisada pela proteína GlnE ou adenililtransferase. Na presença de altas concentrações de amônio, a GS está altamente adenililada e inativada. Em condições limitantes de nitrogênio, os grupos adenilil da GS são removidos, ativando a enzima (MERRICK & EDWARDS, 1995). O gene *glnE* foi identificado no replicon pAZ6 baseado na alta identidade da ORF do contíguo 137 com a proteína GlnE de *Ralstonia solanacearum* (Valor de E= 3e-51). A presença do gene que codifica para a proteína controladora da atividade da glutamina sintetase, a adenililtransferase (GlnE), no megaplasmídeo pAZ6, mostra que genes que codificam para as proteínas do sistema de regulação do metabolismo nitrogenado não estão restritos ao cromossomo principal ou maior de *A. brasilense* estirpe FP2. Este resultado contrasta, por exemplo, com o que ocorre no genoma de *S. meliloti*, onde os genes que codificam para as proteínas do sistema Ntr (por exemplo, *ntrBC, glnA, ntrXY, glnE, glnK* e *glnD*) estão localizados no cromossomo e não nos megaplasmídeos (GALIBERT et al., 2001)

4.8.8 QUIMIOTAXIA

O contíguo 37 contém um seqüência idêntica à do gene *sbpA* de *A. brasilense* já descrito (VAN BASTELAERE *et al*, 1999; FIGURA10). Este gene codifica para um receptor periplasmático de ligação a açúcares similar à proteína ChvE encontrada em *Agrobacterium tumefaciens* (VAN BASTELAERE *et al*, 1999).

A proteína SbpA (SbpA - <u>sugar binding protein A</u>) de A. brasilense foi inicialmente isolada de células cultivadas na presença de exsudatos de raízes de trigo, utilizando eletroforese bidimensional. Análises da seqüencia desta proteína mostrou alta similaridade com o gene ChvE de A. tumefaciens. O gene correspondente foi posteriormente identificado de uma biblioteca genômica de A. brasilense por hibridização com a sonda do gene chvE de A. tumefaciens (VAN BASTELAERE et al., 1999). O produto deste gene está envolvido na quimiotaxia do Azospirillum brasilense para açúcares específicos como D-galactose, L-arabinose e D-fucose, e na captação de D-galactose. A expressão do gene sbpA é induzida pela presença destes mesmos açúcares no meio de crescimento, e um aumento na expressão do gene foi observado em condições de privação da fonte carbono (malato) e na presença de D-galactose (VAN BASTELAERE et al., 1999).

Em Agrobacterium ChvE é necessária para a indução dos genes vir em resposta a açúcares (HUANG et al., 1990). Além disto, a proteína ChvE também é importante na captação de açúcares e na quimiotaxia do *A. tumefaciens* em direção a estes açúcares (CANGELOSI et al., 1990). A identificação de *sbpA* no contíguo 37 confirma a origem do DNA utilizado para a construção das bibliotecas genômica e localiza este gene no replicon pAZ6.

FIGURA 10 - ALINHAMENTO DAS SEQÜÊNCIAS DE NUCLEOTÍDEOS DO GENE sbpA DE Azospirillum brasilense Sp7 E DA ORF2 DO CONTÍGUO 37.

A-A zospirillum B- frame2/37	GAGACGATGGTCGCCAAGAACAGCAAGGTCCTGGTGATTGCCGCCATCGACGGCACGACG CAGCTGATTGCCGCCATCGACGGCACGACG * * *********************************	300 243
A- Azospirillum B- frame2/37	CTGACCGACGTGCTCCAGCAGGCCAAGGACCGCGGCGTGAAGGTCATCGCCTACGACCGG CTGACCGACGTGCTCCAGCAGGCCAAGGACCGCGGGCGTGAAGGTCATCGCCTACGACCGG *******************************	360 303
A- Azospirillum B- frame2/37	CTGATCCGCGGGTCGGAGAATGTGGACTATTACGCGACCTTCGACAATTTCCAGGTCGGC CTGATCCGCGGGTCGGAGAATGTGGACTATTACGCGACCTTCGACAATTTCCAGGTCGGC ********************************	420 363
A- Azospirillum B- frame2/37	GTGCTGCAGGGCAGCTACATCGTCGACGCGCTGGGCCTGAAGGACGGCAAAGGTCCCTTC GTGCTGCAGGGCAGCTGA	480 381

Alinhamento das seqüências de nucleotídeos do gene *sbpA* de *Azospirillum* brasilense Sp7 e da ORF2 do contíguo 37 utilizando o programa ClustalW. A linha A mostra a seqüência de nucleotídeos do gene *sbpA* de *Azospirillum* brasilense Sp7. A linha B mostra a seqüência de nucleotídeos da provável ORF encontrada na Frame 2 do contig 37 da biblioteca genômica do replicon pAZ6 de *A. brasilense* FP2. * indica nucleotídeos idênticos.

Um gene que codifica para uma proteína quimiotática aceptora de grupamento metil também foi identificado no replicon pAZ6. Esta proteína, pertencente a família das MCP's (<u>methyl-accepting chemotaxis proteins</u>), serve como um receptor específico e/ou transdutor de sinal e tem o seu nível de metilação mudado em resposta a ligação de quimioefetores (ZHULIN & ARMITAGE, 1993). As MCP's são tipicamente proteínas ligadas à membrana que contêm dois segmentos transmembrana, um domínio periplasmático amino-terminal e um domínio de sinalização

citoplasmático que gera o sinal intracelular para o flagelo. O domínio amino-terminal contém sítios para o reconhecimento do ligante (BLAIE *et al.*, 1995). ZHULIN e ARMITAGE (1993) não conseguiram encontrar proteínas quimiotáticas aceptoras de grupos metil em *A. brasilense* utilizando ensaios de metilação *in vivo* e anticorpos policionais heterólogos. É possível que estes autores tenham falhado em demonstrar a presença de MCP em *A. brasilense* pela falta de sensibilidade da técnica ou pelo baixo número de cópias destas proteínas na bactéria.

A quimiotaxia e a motilidade são considerados fatores importantes para uma eficiente colonização da planta pela bactéria (STEENHOUDT & VANDERLEYDEN, 2000). A sinalização molecular envolvida na interação entre o *Azospirillum* e a planta ainda é pouco conhecida. Algum tipo de sinalização específica pode ser sugerido uma vez que diferentes estirpes de *Azospirillum* exibem comportamento quimiotático diferencial e dependente da composição do exsudato das raízes (REINHOLD *et al.*,1985). Quimioatrativos como açúcares, ácidos orgânicos e compostos aromáticos provavelmente são importantes nas etapas iniciais da associação entre o *Azospirillum* e a planta (VANDE BROEK & VANDERLEYDEN, 1995). Assim, o replicon pAZ6 contém informações importantes para a interação *Azospirillum*-planta.

4.8.9 SISTEMA DE TRANSPORTE

Dos 37 genes identificados no megaplasmídeo pAZ6, 11 genes aparentemente codificam para um dos componetes do sistema de transporte ABC e estão envolvidos com o transporte de açúcares, aminoácidos e íons. O sistema de transporte ABC (<u>ATP</u> <u>binding cassette</u>) é um dos sistemas de transporte ativo envolvido tanto na importação quanto na exportação de moléculas (HIGGINS, 1992). Estes transportadores possuem uma organização global comum. Tipicamente, ele consiste de uma ou duas proteínas integrais de membrana (permeases), duas proteínas periféricas de membrana que ligam e hidrolizam ATP e uma proteína periplasmática de ligação ao substrato (FINAN *et al.*, 2001). O componente proteico de ligação ao ATP é o mais conservado e as

permeases são as menos conservadas (FATH & KOLTER, 1993). Os genes que codificam para os três polipetídeos geralmente formam um operon.

Em *S. meliloti* mais da metade dos genes que codificam para o transportadores ABC estão localizados no megaplasmídeo pSymB, o que consideralmente estendeu a flexibilidade metabólica do microrganismo por permitir metabolizar uma grande variedade de pequenas moléculas encontradas no solo ou na rizosfera da planta (GALIBERT *et al.*, 2001). A alta proporção de transportadores identificados nas sequencias derivadas do replicon pAZ6 de *A. brasilense* FP2 sugerem que este organismo é capaz de se adaptar metabolicamente a uma grande variedade de ambientes.

Um gene que codifica para uma permease simporte $H^+/gluconato$ (contíguo 77) também foi identificado no pAZ6 de *A. brasilense*.

5. CONCLUSÃO

A presença de minicromossomos no genoma de *A. brasilense* foi sugerida primeiramente por WOOD *et al.* (1981). Posteriormente, CABALLERO-MELLADO *et al.*(1999) e MARTIN-DIDONET *et al.*(2000) demonstraram a presença do gene 16S rDNA em vários destes megareplicons. Neste trabalho foram encontrados nove replicons no genoma de *Azospirillum brasilense* estirpe FP2 utilizando o sistema CHEF-DRIII de eletroforese em campo pulsado, resultando em um tamanho total do genoma de 8,2 Mb. Este tamanho de genoma sugere que este microrganismo possua uma grande capacidade metabólica e é capaz de se adaptar a muitos ambientes.

O menor replicon de *A. brasilense* da estirpe FP2 que é capaz de hibridizar com o 16S rDNA possui 630 Kpb (MARTIN-DIDONET *et al.*, 2000) e foi denominado pAZ6. O sequenciamento de clones aleatórios de uma biblioteca do replicon pAZ6 permitiu a identificação de um gene essencial que codifica para a síntese de parede celular (*murA*), reforçando a sugestão da existência de múltiplos cromossomos no genoma de *A. brasilense* como tem sido sugerido para genomas complexos de *Agrobacterium tumefaciens* C58, *Brucella melitensis, Rhodobacter sphaeroides* (KOLSTO, 1997) e *Sinorhizobium meliloti* (GALIBERT *et al.*, 2001).

O megaplasmídeo pAZ6 de *A. brasilense* estirpe FP2 apresenta também: a)genes envolvidos na biossíntese de parede celular, vitamina B12, lipídeos e nucleotídeos; b)genes que codificam para enzimas envolvidas nas reações de degradação de aminoácidos, carboidratos, lipídeos e ácidos urônicos; c)genes envolvidos no metabolismo do piruvato e do formato; d)genes envolvidos na produção de energia; e) genes envolvidos na quimiotaxia de *A. brasilense* para açúcares; f)genes que codificam para proteínas do sistema de transporte ABC; e g)gene envolvido na regulação do metabolismo nitrogenado.

Os produtos de vários dos genes identificados no replicon pAZ6 estão envolvidos na biossíntese de precursores chaves, transformação de nutrientes importantes e na transdução de energia, processos que são essenciais para a manutenção da viabilidade celular ou da competividade do organismo em muitos ambientes. Considerando que se não existir duplicação gênica ou de funções, é necessário que estes genes estejam localizados em unidades replicativas estáveis e devem fazer parte de um núcleo principal de informações genéticas deste organismo. A esta conclusão, segue que a estrutura do genoma de A. *brasilense* estirpe FP2 é mais semelhante àquela de eucariotos, onde a informação genética essencial está distribuída entre vários cromossomas, do que do paradigma de procariotos, onde um único cromossoma contém toda a informação genética essencial.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ALLARDET-SERVENT, A.; MICHAUX-CHARACHON, S.; JUMAS-BILAK, E.; KARAYAN, L.; RAMUZ, M. Presence of one linear and one circular chromosome in the Agrobacterium tumefaciens C58 genome. J. Bacteriol., v.175, p.7869-7874, 1993.

ALTSCHUL, S.F.; MEDDEN, T.L.; SCHÄFFER, A.A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D.J. Gapped BLAST and PSI-BLAST: a new generation database search programs. Nucl. Acids Res., v.25, p.3389-3402, 1997.

ANDERSSON, S.G.E.; ZOMORODIPOUR, A.; ANDERSSEN, J.O.; SICHERITZ-PONTÉN, T.; ALSMARK, U.C.M.; PODOWSKI, R.M.; NÄSLUND, A.K.; ERIKSSON, A.S.; WINKLER, H.H.; KURLAND, C.G. The genome sequence of *Rickettsia prowazekii* and the origen of mitochondria. Nature, v.396, p.133-140, 1998.

ARIGONI, F.; TALABOT, F.; PEITSCH, M. A genome based approach for the identification of essencial bacterial genes. **Nat. Biotechnol**, v.16, p.851-856, 1998.

ASHWELL, G.; WAHBA, A.J.; HICKMAN, J. Uronic acid metabolism in bacteria. J. Biol. Chem., v.235, p.1559-1565, 1960.

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.**, v.29, p.911-922, 1997.

BARAK, R.; NUR, I.; OKON, Y. Detection of chemotaxis in Azospirillum brasilense. J. Appl. Bacteriol., v.53, p.399-403, 1983.

BARAK, R.; NUR, I.; OKON, Y.; HENIS, Y. Aerotactic response of *Azospirillum brasilense*. J. Bacteriol., v.152, p.643-649, 1982.

BARIL, C.; RICHAUD, C.; BARANTON, G.; SAINT GIRONS, I. Linear chromosome of *Borrelia burgdorferi*. Res. Microbiol., v.140, p.507-516, 1989.

BARNETT, M.J.; FISHER, R.F.; JNES, T.; KOMP, C.; ABOLA, A.P.; BARLOY-HUBLER, F.; BOWSER, L.; CAPELA, D.; GALIBERT, F.; GOUZY, J.; GURJAL, M.; HONG, A.; HUIZAR, L.; HYMAN, R.W.; KAHN, D.; KAHN, M.L.; KALMAN, S.; KEATING, D.H.; PAL, C.; PECK, M.C.; SURZYCKI, R.; WELLS, D.H.; YEH, K.; DAVIS, R.W.; FEDERSPIEL, N.A.; LONG, S.R. Nucleotide sequence and predited functions of the entire Sinorhizobium meliloti pSymA megaplasmid. Proc. Natl. Acad. Sci., v.98, p.9883-9888, 2001.

BASHAN, Y. Current status of *Azospirillum* inoculation technology: *Azospirillum* as a challenge for agriculture. **Can. J. Microbiol.**, v.36, p.591-608, 1990.

BECKER, A.; RUBERG, S.; KUSTER, H.; ROXLAU, A.A.; KELLER, M.; IVASHINA, T.; CHENG, H.P.; WALKER, G.C.; PUHLER, A. The 32-kilobase exp gene cluster of *Rhizobium meliloti* directing the biosynthesis of galactoglucan: genetic organization and properties of the encoded gene products. J. Bacteriol., v.179, p.1375-1384, 1997.

BECKER, A.; KLEICKMANN, A.; KELLER, M.; ARNOLD, W.; PUHLER, A. Identification and analysis of the *Rhizobium meliloti exoAMONP* genes involved in exopolysaccharide biosynthesis and mapping of promoters located on the *exoHKLAMONP* fragment. **Mol. Gen. Genet.**, v.241, p.367-379, 1993.

BECKING, J.H. Fixation of molecular nitrogen by na aerobic *Vibrio* or *Spirillum*. Antonie van Leeuwenhoek. J. Microbiol. Serol., v.29, p.326, 1963.

BEIJERINCK, M.W. Über ein *Spirillum*, welches freien Stickstoff binden kann? Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Abt.2, v. 63, p.353-359, 1925.

BERNS, C. & HILLEN, W. Gene regulation by tetracycline, Contraints of resistance regulation in bateria shape TetR for application in eukaryotes. **Eur. J. Biochem.**, v.270, p.3109-3121, 2003.

BLAIR, D.F. How bacteria sense and swim. Ann. Rev. Microbiol., v.49, p.489-522, 1995.

BRASSINGA, A.K.; SIAM, R.; MARCZYNSKI, G.T. Conserved gene cluster at replication origins of the alpha-proteobacteria *Caulobacter crescentus* and *Rickettsia prowazekii*. J Bacteriol., v.183, p.1824-1829, 2001.

CABALLERO-MELLADO, J.; LÓPEZ-REYES, L.; BUSTILLOS-CRISTALES, R. Presence of 16S rRNA genes in multiple replicons in *Azospirillum brasilense*. **FEMS Microbiol. Lett.**, v.178, p.283-288, 1999.

CANGELOSI, G.A.; ANKENBAUER, R.G.; NESTER, E.W. Sugar induce the *Agrobacterium* virulence genes through a periplasmic binding protein and a trnamenbrane signal protein. **Proc. Natl. Acad. Sci. USA**, v. 87, p.6708-6712, 1990.

CAPELA, D.; BARLOY-HUBLER, F.; GOUZY, J.; BOTHE, G.; AMPE, F.; BATUT, J.; BOISTARD, P.; BECKER, A.; BOUTRY, M.; CADIEU, E.; DRÉANO, S.; GLOUX, S.; GODRIE, T.; GOFFEAU, A.; KAHN, D.; KISS, E.; LELAURE, V.; MASUY, D.; POHL, T.; PORTETELLE, D.; PÜHLER, A.; PURNELLE, B.; RAMSPERGER, U.; RENARD, C.; THÉBAULT, P.; VANDENBOL, M.; WEIDNER, S.; GALIBERT, F. Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. **Proc. Natl. Acad. Sci USA**, v.98, p.9877-9882, 2001.

CASJENS, S.; MURPHY, M.; DELANGE, M.; SAMPSON, L.; VAN VUGT, R.; HUANG, W.M. Telomeres of linear chromosomes of Lyme disease spirochaetes: nucleotide sequence and possible exchange with linear plasmids telomeres. **Mol. Microbiol.**, v.26, p.581-596, 1997.

CASJENS, S. The diverse and dynamic structure of bacterial genomes. Annu. Rev. Genet., v.32, p.339-377, 1998.

CHEN, C.W. Complications and implications of linear bacterial chromosomes. **Trends Genet.**, v.12, p.192-196, 1996.

CHENG, H.P. & LESSIE, T.G. Multiple replicons constituting the genome of *Pseudomonas cepacia* 17616. J. Bacteriol., v.176, p.4034-4042, 1994.

CHEUNG, V.; NELSON, S. Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA. **Proc. Natl. Acad. Sci. USA.**, v. 93, p.14676-14679.

CHRISTIE, P.J. Type IV secretion: intercellular transfer of macromolecules by systems ancestrally related to conjugation machines. **Mol. Microbiol.**, v.40, p.294-305, 2001.

COLE, S.T. *et al.* Deciphering the biology of *Mycobacterium tuberculosis* from the complete geneome sequence. **Nature**, v.393, p.537-544, 1998.

COPLEY, J. All at sea. Nature, v.415, p.572-574, 2002.

COSTA, D.M.; SUZUKI, K.; SATOU, M.; YOSHIDA, K. Genome analysis of *Agrobacterium tumefaciens*:Linkage map and genetic feature of the left region of the linear chromosome. **Genes Genet. Syst.**, v.76, p.363-371, 2001.

CRESPI, M.; MESSENS, E.; CAPLAN, A.B.; VAN MONTAGU, M.; DESOMER, J. Fasciation induction by the phytopatogen *Rhodococcus fascians* depends upon a linear plasmid enconding a cytokinin synthase gene. **EMBO J.**, v.11, p.795-804, 1992.

CROES, C.; VAN BASTELAERE, E.; DECLERCQ, E.; EYERS, M.; VANDERLEYDEN, J.; MICHIELS, K. Identification and mapping of loci involved in motility, adsorption to wheat roots, colony morphology, and growth in minimal medium on the *Azospirillum brasilense* Sp7 90-MDa plasmid. **Plasmid**, v.26, p.83-93, 1991.

DE TROCH, P.; KEIJERS, V.; VANDERLAYDEN, J. Sequence analysis of the *Azospirillumbrasilense exoB* gene, enconding UDP-glucose 4'-epimerase. **Gene**, v. 144, p.143-144, 1994.

DE ZAMAROCZY, M.; PAQUELIN, A.; ELMERICH, C. Functional organization of the *glnA-glnB* cluster of *Azospirillum bresilense*. J. Bacteriol., v.175, p.2507-2515, 1993.

DEAN F.; NELSON, J.; GIESLER, T.; LASKEN, R. Rapid amplification of plasmid and phage DNA using ϕ 29 DNA polymerase and multiply-primed rolling circle amplification. **Genome Res.**, v.11, p.1095-1099, 2001.

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

DELCHER, A.L.; HARMON, D.; KASIF, S.; WHITE, O.; SALZBERG, S.L. Improved microbial gene identification with Glimmer. Nucl. Acids Res., v.27, p.4636-4641, 1999.

DELVECCHIO, V.G.; KAPATRAL, V.; REDKAR, R. J.; PRATA, G.; MUJER, C.; LOS, T.; IVANOVA, N.; ANDERSON, I.; BHATTACHARYYA, A.; LYKIDIS, A.; REZNIK, G.; JABLOSNSKI, L.; LARSEN, N.; D'SOUZA, M.; BERNAL, A.; MAZUR, M.; GOLSTMAN, E.; SELKOV, E.; ELZER, P.H.; HAGIUS, S.; OÇALLAGHAN, D.; LETESSON, J.J.; HASELKORN, R.; KYRPIDES, N.; OVERBEEK, R. The genome sequence of the facultative intracellular pathogen *Brucella melitensis*. **Proc. Natl. Acad. Sci. USA**, v.99, p.443-448, 2002.

DEMENTIN, S.; BURLAT, B.; DE LACEY, A.L.; PARDO, A.; ADRYANCZYK-PERRIER GUIGLIARELLI, B.; FERNANDEZ, V.M.; ROUSSET, M. A glutamate is the essential proton transfer gate during the catalytic cycle of the [NiFe] hydrogenase. J. Biol. Chem., v.279, p.10508-10513, 2004.

DETTER, J.C.; JETT, J.M.; LUCAS, S.M.; DALIN, E.; ARELLANO, A.R.; WANG, M.; NELSON, J.R.; CHAPMAN, J.; LOU, Y.; ROKHSAR, D.; HAWKINS, T.L.; RICHARDSON, P.M. Isothermal strand-displacement amplification applications for high-throughput genomics. Genomics, v.80, p.691-698, 2002.

DÖBEREINER, J. & PEDROSA, F.O. Nitrogen-fixing in non leguminous crop plants. Sringer-Verlag, Berlin, p.1-155, 1987.

DÖBEREINER, J. History and new perspectives of diazotrophs in association with non-leguminous plants. **Symbiosis**, v.13, p.1-13, 1992.

DÖBEREINER, J.; MARRIEL, I.E.; NERY, M. Ecological distribution of *Spirillum lipoferum* Beijerinck. Can. J. Microbiol., v. 22, p.1464-1473, 1976.

DOOLITLE, R. F. Microbial genomes multiply. Nature, v.416, p.697-700, 2002.

DUCKWORTH, M.; TURVEY, JR. The action of a bacterial agarase on agarose, porphyran and alkali-treated porphyran. **Biochem. J.**, v.113, p.687-692, 1969.

ECKERT, B.; WEBER, O.B.; KIRCHHOF, G.; HALBRITTER, A.; STOFFELS, M.; HARTMANN, A. *Azospirillum doebereinerae* sp. nov., a nitrogen-fixing bacterium associated with the C4-grass *Miscanthus*. Int. J. Syst. Bacteriol., v.51, p.17-26, 2001.

ECKHARDT, T A rapid method for the identification of plasmid DNA in bacteria. **Plasmid**, v.1, p.584-588, 1978.

EWING, B. & GREEN, P. Base calling of automated sequencer using *phred*. II. Error probabilities. Genome Res., v.8, p.186-194, 1998.

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

FATH, M.J. & KOLTER, R. ABC transporters: Bacterial exporters. Microbiol. Rev., v.57, p.995-1017, 1993.

FERDOWS, M.S. & BARBOUS, A.G. Megabase-sized linear DNA in the bacterium *Borrelia burgdorferi*, the Lyme disease agent. **Proc. Natl. Acad. Sci.** USA, v.86, p.5969-5973, 1989.

FINAN, T.M.; WEIDNER, S.; WONG, K.; BUHRMESTER, J.; CHAIN, P.; VORHÖLTER, F.J.; HERNANDEZ-LUCAS, I.; BECKER, A.; COWIE, A.; GOUZY, J.; GOLDING, B.; PÜHLER. The complete sequence of the 1.683-kb pSymB megaplasmid from the N₂-fixing endosymbiont *Sinorhizobium meliloti*. **Proc. Natl. Acad. Sci USA**, v.98, p. 9889-9894, 2001.

FLEISCHMANN, R.D. et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science, v.269, p.496-512, 1995

FOGHER, C.; DUSHA, I.; BARBOT, P.; ELMERICH, C. Heterologous hybridization of *Azospirillum* DNA to *Rhizobium nod* and *fix* genes. FEMS Microbiol. Lett., v.30, p.245-249, 1985.

FRANCHE, C. & ELMERICH, C. Physiological properties and plasmids content of several strains of *Azospirillum brasilense* and *A. lipoferum*. Ann. Microbiol. Inst. Pasteur 132A, p.29-37, 1981.

FRANGEUL, L.; NELSON, K.E.; BUCHRIESER, C.; DANCHIN, A.; GLASER, P.; KUNST, F. Cloning and assembly in microbial genome projects. **Microbiology**, v.145, p.2625-2634, 1999.

FRASER, C.M.; CASJENS, S.; HUANG, W.M.; SUTTON, G.G.; CLAYTON, R.; LATHIGRA, R.; WHITE, O.; KETCHUM, K.A.; DODSON, R.; HICKEY, E.K. *et al.* Genomic sequence of the Lyme disease spirochaete, *Borrelia burgdorferi*. Nature, v.390, p.580-587, 1997.

FRASER, C.M.; EISEN, J.A.; SALZBERG S.L. Microbial genome sequencing. Nature, v.406, p.799-803, 2000.

FRASER, C.M.; GOCAYNE, J.D.; WHITE, O.; ADAMS, M.D.; CLAYTON, R.A.; FLEISCHMANN, R.D.; BULT, C.J.; KERLAVAGE, A.R.; SUTTON, G.; KELLEY, J.M.; FRITCHMAN, J.L.; WEIDMAN, J.F.; SMALL, K.V.; SANDUSKY, M.; FUHRMANN, J.; NGUYEN, D.; UTTERBACK, R.; SAUDEK, D.M.; PHILLIPS, C.A.; MERRICK, J.M.; TOMB, J. DOUGHERTY, B.A.; BOTT, K.F.; HU, P.; LUCIER, T.S.; PETERSON, S.N.; SMITH, H.O.; HUTCHISON III, C.A.; VENTER, J.C. The minimal gene complete of *Mycoplasma genitalium*. Science, v.270, p.397-403, 1995.

FRAZIER, M.E.; JOHNSON, G.M.; THOMASSEN, D.G.; OLIVIER, C.E.; PATRINOS, A. Realizing the potencial of the genome revolution: The genomes of life program. **Science**, v.300, p.290-293, 2003.

FRAZZON, J & SCHRANK, I.S. Sequencing and complementation analysis of the *nifUSV* genes from *Azospirillum brasilense*. **FEMS Microbiol. Lett.**, v.159, p.151-158, 1998.

GALIBERT, F.; FINAN, T.M.; LONG, S.R.; PÜHLER, A.; ABOLA, A.; AMPE, F.; BARLOY-HUBLER, F.; BARNETT, M.J.; BECKER, A.; BOISTARD, P.; BOTHE, G.; BOUTRY, M.; BOWSER, L.; BUHRMESTER, J.; CADIEU, E.; CAPELA, D.; CHAIN, P.; COWIE, A.; DAVIS, R.W.; DRÉANO, S.; FEBERSPIEL, N.A.; FISHER, R.F.; GLOUX, S.; GODRIE, T.; GOFFEAU, A.; GOLDING, B.; GOUZY, J.; GURJAL, M.; HERNANDEZ-LUCAS, I.; HONG, A.; HUIZAR, L.; HYMAN, R.W.; JONES, T.; KAHN, D.; KAHN, M.L.;KALMAN, S.; KEATING, D.H.; KISS, E.; KOMP, C.; LELAURE, V.; MASUY, D.; PALM, C.; PECK, M.C.; POHL, T.M.; PORTETELLE, D.; PURNELLE, B.; RAMSPERGER, U.; SURZYCKI, R.; THÉBAULT, P.; VANDENBOL, M.; VORHÖLTER, F.J.; WEIDNER, S.; WELLS, D.H.; WONG, K.; YEH, K.C.; BATUT, J. The composite genome of the legume symbiont *Sinorhizobium meliloti*. **Science**, v.293, p.668-672, 2001.

GAUTHIER, D. & ELMERICH, C. Relationship between glutamine synthetase and nitrogenase in *Spirillum lipoferum*. **FEMS Microbiol. Lett.**, v.2, p.101-104, 1977.

GHISLA, S. & THORPE, C. Acyl-CoA dehydrogenases. Eur. J. Biochem., v. 271, p.494-508, 2004.

GLUCKSMANN, M.A.; REUBER, T.L.; WALKER, G.C. Genes needed for the modification, polymerization, export, and processing of succinoglycan by *Rhizobium meliloti*: a model for succinoglycan biosynthesis. J. Bacteriol., v.175, p.7045-7055, 1993.

GOODNER, B.; HINKLE, G.; GATTUNG, S.; MILLER, N.; BLANCHARD, M.; QUROLLO, B.; GOLDMAN, B.S.; CAO, Y.; ASKENAZI, M.; HALLING, C.; MULLIN, L.; HOUMIEL, K.; GORDON, J.; VAUDIN, M.; IARTCHOUK, O.; EPP, A.; LIU, F.; WOLLAM, C.; ALLINGER, M.; DOUGHTY, D; SCOTT, C.; LAPPAS, C.; MARKELZ, B.; FLANAGAN, C.; CROWELL, C.; GURSON, J.; LOMO, C.; SEAR, C.; STRUB, G.; CIELO, C.; SLATER, S. Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58.**Science**, v. 294, p.2323-2328, 2001.

GOODNER, B.W.; MARKELZ, B.P.; FLANAGAN, M.C.; CROWELL, C.B.; RACETTE, J.L.; SCHILLING, B.A.; HALFON, L.M.; MELLORS, J.S.; GRABOWSKI, G. Combined genetic and physical map of the complex genome of *Agrobacterium tumefaciens*. J. Bacteriol., v.181, p.5160-5166, 1999. GORDON, D.; ABAJIAN, C.; GREEN, P. Consed: A graphical tool for sequence finishing. Genome Res., v.8, p. 195-202, 1998.

HAAS, S.; VINGRON, M.; POUSTKA, A.; WIEMANN, S. Primer design for large scale sequencing. Nucleic Acids Res., v.26, p.3006-3012, 1998.

HAWKINS, T.L.; DETTER, J.C.; RICHARDSON, P.M. Whole genome amplification – implications and advances. Curr. Opin. Biotechnol., v.13, p.65-67, 2002.

HEIDELBERG, J.F.; EISEN, J.A.; NELSON, W.C.; CLAYTON, R.A.; GWINN, M.L.; DODSON, R.J.; HAFT, D.H.; HICKEY, E.K.; PETERSON, J.D.; UMAYAM, L.; GILL, S.R.; NELSON, K.E.; READ, T.D.; TETTELIN, H.; RICHARDSON, D.; ERMOLAEVA, M.D.; VAMATHEVAN, J.; BASS, S.; QIN, H.; DRAGOI, I.; SELLERS, P.; MCDONALD, L.; UTTERBACK, T.; FLEISHMANN, R.D.; NIERMAN, W.C.; WHITE, O.; SALZBERG, S.; SMITH, H.O.; COLWELL, R.R.; MEKALANOS, J.J.; VENTER, J.C.; FRASER, C.M. DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. **Nature**, v.406, p.477-483, 2000.

HEINRICH, D.& HESS, D. Chemotatic attraction of *Azospirillum lipoferum* by wheat roots and characterization of some attractants. **Can. J. Microbiol.**, v.31, p.26-31, 1985.

HIGGINS, C.F. ABC transporters: from microorganisms to man. Annu. Rev.Cell Biol., v.8, p.67-113, 1992.

HINNEBUSCH, J.& BARBOUR, A. Linear plasmids of *Borrelia burgdorferi* have a telomeric structure and sequence similar to those of a eukaryotic virus. J. **Bacteriol**, v.173, p.7233-7239, 1991.

HINNEBUSCH, J.; BERGSTROM, S.; BARBOUR A. Cloning and sequence analysis of linear plasmid telomeres of the bacterium *Borrelia burgdorferi*. Mol. Microbiol., v.4, p.811-820, 1990.

HOLLOWAY, B.W.; ESCUADRA, M.D.; MORGAN, A.F.; SAFFERY, R.; KRISHNAPILLAI, V. The new approaches to whole genome analysis of bacteria. FEMS Microbiol. Lett., v.79, p.101-105, 1992.

HONEYCUTT, R.J.; MCCLELLAND, M.; SOBRAL, B.W. Physical map of the genome of *Rhizobium meliloti* 1021. J Bacteriol., v.175, p.6945-6952, 1993.

HUANG, M.W.; CANGELOSI, G.A.; HALPERIN, W.; NESTER, E.W. A chromosomal *Agrabacterium tumefaciens* gene required for effective plant signal transduction. J. Bacteriol., v. 172, p.1814-1822, 1990.

INON DE IANNINO, N.; BRIONES, G.; TOLMASKY, M.; UGALDE, R.A. Molecular cloning and characterization of *cgs*, the *Brucella abortus* cyclic beta(1-2) glucan synthetase gene: genetic complementation of *Rhizobium meliloti ndvB* and *Agrobacterium tumefaciens chvB* mutants. J. Bacteriol., v. 180, p.4392-4400,1998.

ISHIKAWA, J.; HOTTA, K. Frame-Plot: a new implementation of Frame analysis for predicting protein-coding regions in bacterial DNA with a high G+C content. **FEMS Microbiol. Letters.**, v. 174, p.251-253, 1999.

JUMAS-BILAK, E.; MICHAUX-CHARACHON, S.; BOURG, G.; RAMUZ, M.; ALLARDET-SERVENT, A. Unconventional genomic organization in the alpha subgroup of the Proteobacteria. J. Bacteriol., v. 180, p.2749-2755, 1998.

KANEKO, T.; NAKAMURA, Y.; SATO, S.; MINAMISAWA, K.; UCHIUM, T.; SASAMOTO, S.; WATANABE, A.; IDESAWA, K.; IRIGUCHI, M.; KAWASHIMA, K.; KOHARA, M.; MATSUMOTO, M.; SHIMPO, S.; TSURUOKA, H.; WADA, T.; YAMADA, M.; TABATA, S. Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. **DNA Res.**, v.9, p.189-197, 2002.

KAWARABAYASI, Y.; HINO, Y.; HORIKAWA, H.; YAMAZAKI, S.; HAIKAWA, Y.; JIN-NO, K.; TAKAHASHI, M.; SEKINE, M.; BABA, S.; ANKAI, A.; KOSUGI, H.; HOSOYAMA, A.; FUKUI, S.; NAGAI, Y.; NISHIJIMA, K.; NAKAZAWA, H.; TAKAMIYA, M.; MASUDA, S.; FUNAHASHI, T.; TANAKA, T.; KUDOH, Y.; YAMAZAKI, J.; KUSHIDA, N.; OGUCHI, A.; KIKUCHI, H. *et al.* Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, *Aeropyrum pernix* K1. **DNA Res.**, v.6, p.83-101 e 145-52, 1999.

KE, D.; BOISSINOT, M.; HULETSKY, A.; PICARD, F.J.; FRENETTE, J.; OUELLETTE, M.; ROY, P.H.; BERGERON, M.G. Evidence for horizontal gene transfer in evolution of elongation factor Tu in enterococci. J. Bacteriol., v.182, p.6913-6920, 2000.

KELLY, M.J.S.; POOLE, R.K.; YATES, M.G.; KENNEDY, C. Cloning and mutagenesis of genes encoding the cytochrome *bd* terminal oxidase complex in *Azotobacter vinelendii*: Mutants deficient in the cytochrome *d* complex are unable to fix nitrogen in air. J. Bacteriol., v.172, p.6010-6019, 1990.

KHAMMAS, K.M.; AGERON, E.; GRIMONT, P.A.D.; KAISER, P. *Azospirillum irakense* sp. nov. a nitrogen-fixing bacterium associated with rice roots and rhizosphere soil. **Res. Microbiol.**, v.140, p.679-693, 1989.

KOLSTO, A.B. Dynamics bacterial genome organization. **Mol. Microbiol.**, v.24, p.241-248, 1997.

KONEMAN, E.W.; ALLEN, S.D.; JANDA, W.M.; SCHRECKENBERGER, P.C.; WINN, W.C. Diagnostic Microbiology, (5 ed.) Lippincott, Philadelphia, New York, 1395p,1997.

LANDER, E. S. *et al.* Inicial sequencing and analysis of the human genome. Nature, v.409, p.860-921,2001.

LARIMER, F.W.; CHAIN, P.; HAUSER, L.; LAMERDIN, J.; MALFATTI, S.; DO, L.; LAND, M.L.; PELLETIER, D.A.; BEATTY, J.T.; LANG, A.S.; TABITA, F.R.; GIBSON, J.L.; HANSON, T.E.; BOBST, C.; TORRES Y TORRES, J.L.; PERES, C.; HARRISON, F.H. Complete genome sequence of the metabolically versatile photosynthetic *Rhodopseudomonas palustris*. **Nature Biotechnology**. Published online 14 december 2003; doi: 10.1038/nbt923

LAUKEL, M.; CHISTOSERDOVA, L.; LIDSTROM, M.E.; VORHOLT, J.A. The tungsten-containing formate dehydrogenase from *Methylobacterium extorquens*AM1: Purification and properties. **Eur. J. Biochem.**, v. 270, p.325-333, 2003.

LEE, J.S.; BIRREN, B.; LAI, E. Introduction to Pulsed-field gels and preparation and analysis of large DNA. *In* Nonmammalian genomic analysis. A practical guide. BIRREN, B.; LAI, E. (ed.) Academic Press, California, USA, p.1-24, 1996.

LESSIE, T.G.; HENDRICKSON, W.; MANNING, B.D.; DEVEREUX, R. Genomic complexity and plasticity of *Burkholderia cepacia*. **FEMS Microbiol.** Let., v.144, p.117-128, 1996.

LEWIN, B. Genes VII. (ed.) Artmed, São Paulo, 955p, 2001.

LI, P.L; FARRAND, S.K. The replicator of the nopaline-type Ti plasmid pTiC58 is a member of the repABC family and is influenced by the TraR-dependent quorum-sensing regulatory system. J. Bacteriol., v.182, p.179-188, 2000.

LIANG, Y.Y.; KAMINSKI, P.A.; ELMERICH, C. Identification of a *nifA*-like regulatory gene of *Azospirillum brasilense*Sp7 expressesd under conditions of nitrogen fixation and in the presence of air and ammonia. **Mol. Microbiol.**, v.5, p.2735-2744, 1991.

LIN, Y.S.; KIESER, H.M.; HOPWOOD, D.A.; CHEN, C.W. The chromosomal DNA of *Streptomyces lividans* 66 is linear. Mol. Microbiol., v.10, p.923-933, 1993.

LOPEZ-DE-VITORIA, G.; FIELDER, D.R.; ZIMMER-FAUST, R.K.; LOVELL, C.R. Motility behavior of *Azospirillum* species in response to aromatic compounds. **Can. J. Microbiol.**, v.40, p.705-711, 1994.

MACHADO, H.B.; FUNAYAMA, S.; RIGO, L.U.; PEDROSA, F.O. Excretion of ammonium by *Azospirillum brasilense* mutants resistant to ethylenediamine. **Can. J. Microbiol.**, v.37, p.549-553, 1991.

MAGALHÃES, F.M.M.; PATRIQUIN, D.; DÖBEREINER, J. Infection of field grown maize with *Azospirillim* spp. **Rev. Brasil. Biol.**, v.39, p.587-596, 1979.

MARCONI, R.T.; CASJENS, S.; MUNDERLOH, U.G.; SAMUELS, D.S. Analysis of linear plasmid dimers in *Borrelia burgdorferi* sensu lato isolates: implications concernig the potential mechanism of linear plasmid replication. J. Bacteriol., v.178, p.3357-3361, 1996.

MARTIN-DIDONET, C.C.G.; CHUBATSU, L.S.; SOUZA, E.M.; KLEINA, M.; REGO, F.G.M.; RIGO, L.U.; YATES, M.G.; PEDROSA, F.O. Genome structure of the genus *Azospirillum*. J. Bacteriol., v.182, p.4113-4116, 2000.

MERRICK, M.J.; EDWARDS R.A. Nitrogen control in bacteria. Microbiol. Rev., v.59, p.604-622, 1995.

MCLACHLAN, A.D.; STADEN, R.; BOSWELL, D.R. A method for measuring the non-random bias of a codon usage table. Nucl. Acids Res., v.12, p.9567-9575, 1984.

MICHAUX, S.; PAILLISON, J.; CARLES-NURIT, M.J.; BOURG, G.; ALLANDERT-SERVENT, A.; RAMUZ, M. Presence of two independent chromosomes in the *Brucella melitensis* 16M genome. J. Bacteriol., v.175, p.701-705, 1993.

MICHIELS, K.W.; VANDERLAYDEN, J.; VAN GOOL, A.P.; SINGER, E.R. Isolation and characterization of *Azospirillum brasilense loci* that correct *Rhizobium melilotiexoB* and *exoC* mutations. J. Bacteriol., v.170, p.5401-5404, 1988.

MICHIELS, K.; DE TROCH, P.; ONYEOCHA, I.; VAN GOOL, A.; ELMERICH, C.; VANDERLEYDEN, J. Plasmid localization and mapping of two *Azospirillum brasilense loci* that affect exopolysaccharide synthesis. **Plasmid**, v.21, p.142-146, 1989.

MILCAMPS, A. & VANDERLAYDEN, J. *Azospirillum* genetics. Curr. Topics Mol. Genet., v.1, p.215-232, 1993.

MINERDI, D.; FANI, R.; BONFANTE, P. Identification and evolutionary analysis of putative cytoplasmic McpA-like protein in bacterial strain living in symbiosis with a mycorrhizal fungus. J. Mol.Evol., v.54, p.815-824, 2002.

MORENO, E. Genome evolution within the alpha *Proteobacteria*: why do some bacteria not posses plasmids and others exhibit more than one different chromosome? **FEMS Microbiol Rev.**, v.22, p.255-275, 1998.

NELSON K.E.; PAULSEN, I.T.; FRASER, C.M. Microbial genome sequencing: a window into evolution and physiology. American Society for Microbiology, v.67, p.310-317, 2001.

NOLLING, J.; BRETON, G.; OMELCHENKO, M.V.; MAKAROVA, K.S.; ZENG, Q.; GIBSON, R.; LEE, H.M.; DUBOIS, J.; QIU, D.; HITTI, J. *et al.* Genome sequencing and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. J. Bacteriol., v.183, p.4823-4838, 2001.

OCHMAN, H.; MORAN, N.A. Genes lost and genes found: evolution of bacterial pathogenesis and symbiosis. **Science**, v.292, p.1096-1099, 2001

OCHMAN, H. Bacterial evolution: Cromosome arithmetic and geometry. Curr. Biol., v.12, p.427-428, 2002.

OKON, Y.; NUR, I.; HINIS, Y. Efect of oxygen concentration on electron transport components and microaerobic properties of *Azospirillum brasilense*. *In* **Azospirillum II : Genetics, Physiology, Ecology** KLINGMÜLLER, W. (ed.), Birkhauser Verlag, Basel: p 115, 1983.

OKON, Y. *Azospirillum* as a potential inoculant for agriculture. **Trends Biotech.**, v.3, p.223-228, 1985.

OKON, Y. Azospirillum-Plant Associations. CRC Press, Boca Raton, Fla, 1994.

ONEYOCHA, I.; VIEILLE, C.; ZIMMER, W.; BACA, B.E.; FLORES, M.; PALACIOS, R.; ELMERICH, C. Physical map and properties of a 90-MDa plasmid of *Azospirillum brasilense* Sp7. **Plasmid**, v.23, p.169-182, 1990.

PARKHILL, J.; DOUGAN, G.; JAMES, K.D.; THOMSON, N.R.; PICKARD, D.; WAIN, J.; CHURCHER, C.; MUNGALL, K.L.; BENTLEY, S.D.; HOLDEN, M.T. *et al.* Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. **Nature**, v.413, p.848-852, 2001.

PARKHILL, J.; WREN, B.W.; THOMSON, N.R.; TITBALL, R.W.; HOLDEN, M.T.; PRENTICE, M.B.; SEBAIHIA, M.; JAMES, K.D.; CHURCHER, C.; MUNGALL, K.L. *et al.* Genome sequence of *Yersinia pestis*, the causative agent of plague. **Nature**, v.413, p.523-527, 2001.

PATRIQUIN, D.G.; DÖBEREINER, J.; JAIN, D.K. Sites and processes of association between diazotrophs and grasses. **Can. J. Microbiol.**, v.29, p.900-915, 1983.

PEDROSA, F.O.; YATES, M.G. Regulation of nitrogen fixation (*nif*) genes of *Azospirillum brasilense* by *nifA* (*gln*) type gene products. **FEMS Microbiol.** Lett., v.23, p.95-101, 1984.

PEDROSA, F.O. Physiology, biochemistry, and genetics of *Azospirillum* and other root-associated nitrogen-fixing bacteria. CRC Critical Rev. in Plant Sciences, v.6, p.345-384, 1988.

PELANDA, R.; VANONI, M.A.; PEREGO, M.; PIUBELLI, L.; GALIZZI, A.; CURTI, B.; ZANETTI, G. Glutamate synthase genes of the diazotroph *Azospirillum brasilense*. Cloning, sequencing, and analysis of functional domains. J. Biol. Chem., v.268, p.3099-3106, 1993.

PLAZINSKI, J.; DART, P.J.; ROLFE, B.G. Plasmid visualization and *nif* gene location in nitrogen-fixing *Azospirillum* strains. J. Bacteriol., v. 155, p.1429-1433, 1983.

PORTALIER, R., ROBERT-BAUDOUY, J., STOEBER, F. Regulation of *Escherichia coli* K-12 hexuronate system genes: *exu* regulon. J. Bacteriol., v.143, p.1095-1107, 1980.

POSTGATE, J. Trends and perspectives in nitrogen fixation research. Adv. Microb. Physiol., v.30, p.1-22, 1989.

PREISIG, O.; ANTHAMATTEN, D.; HENNECKE, H. Genes for a microaerobically induced oxidase complex in *Bradyrhizobium japonicum* are essential for a nitrogen-fixing endosymbiosis. **Proc. Natl. Acad. Sci**. USA, v.90, p.3309-3313, 1993.

PREISIG, O.; ZUFFEREY, R.; THONY-MEYER, L.; APPLEBY, C.A.; HENNECKE, H. A high-affinity cbb3-type cytochrome oxidase terminates the symbiosis-specific respiratory chain of *Bradyrhizobium japonicum*. J Bacteriol., v.178. p.1532-1538, 1996.

PROSDOCINI, F.; CERQUEIRA, G.C.; BINNECK. E.; SILVA, A.F.; REIS, A.N.; JUNQUEIRA, A.C.M.; SANTOS, A.C.F.; NBANI Jr., A.; WUST. C.I.; CAMARGO FILHO, F.; KESSEDJIAN, J.L.; PETRETSKI, J.H.; CAMARGO, L.P.; FERREIRA, R.G.M.; LIMA, R.P.; PEREIRA, R.M.; JARDIM, S.; SAMPAIO, V.S.; FOLGUERAS-FLATSCHART, A.V. Bioinformática: Manual do usuário. Biotecnologia Ciência & Desenvolvimento, v.29, p.12-22, 2003.

QUIVINGER, B.; FRANCHE, C.; LUTFALLA, G.; RICE, D.; HASELKORN, R.; ELMERICH, C. Cloning of a nitrogen fixation (*nif*) gene cluster of *Azospirillum brasilense*. **Biochimie**, v.64, p.495-502, 1982.

RAMOS, H.J.O.; RONCATO-MACCARI, L.D.B.; SOUZA, E.M.; SOARES-RAMOS, J.R.L.; HUNGRIA, M.; PEDROSA, F.O. Monitoring *Azospirillum*wheat interactions using the *gfp* and *gusA* constitutively expressed from a new broad-host range vector. J. Biotech., v. 97, p.243-252, 2002.

REINHOLD, B.; HUREK, T.; FENDRIK, I. Strain-specific chemotaxis of *Azospirillum* spp..**J. Bacteriol.**, v.162, p.190-195, 1985.

REINHOLD, B.; HUREK, T.; FENDRIK, I.; POT, B.; GILLIS, M.; KESTERS, K.; THIELEMANS, S.; DE LEY, J. *Azospirillum halopraeferens* sp. nov. a nitrogen-fixing organism associated with the roots of Kallar grass (*Leptochloa fusca* (L.) Kunth). Int. Syst. Bacteriol., v.37, p.43-51, 1987.

RODLEY, P.D.; RÖMLING, U.; TÜMMIER, B. A physical genome map of the *Burkholderia cepacia* type strain. **Mol. Microbiol.**, v.17, p.57-67, 1995.

SAKAGUCHI, K. Invertrons, a class of structurally and functionally related genetic elements that includes linear DNA plasmids, transposable elements, and genomes of adeno-type viruses. **Microbiol. Rev.**, v.54, p.66-74, 1990.

SALANOUBAT, M.; GENIN, S.; ARTIGUENAVE, F.; GOUZY, J.; MANGENOT, S.; ARLAT, M.; BILLAULT, A.; BROTTIER, P.; CAMUS, J.C.; CATTOLICO, L.; CHANDLER, M.; CHOISNE, N,; CLAUDEL-RENARD, C.; CUNNAC, S.; DEMANGE, N,; GASPIN,C.; LAVIE, M.; MOISAN, A.; ROBERT, C.; SAURIN, W.; SCHIEX, T.; SIGUIER, P.; THÉBAULT, P.; WHALEN, M.; WINCKER, P.; LEVY, M.; WEISSENBACH, J.; BOUCHER, C.A. Genome sequence of the plant pathogen *Ralstonia solanacearum*. **Nature**, v.415, p.497-502, 2002.

SAMBROOK, J.; FRITSCH, E.F.; MANIATIS, T. Molecular cloning: a laboratory manual. N.Y., Cold Spring Harbor Laboratory Press, 1989.

SAMPAIO, M.J.A.M.; PEDROSA, F.O.; DÖBEREINER, J. Growth of *Derxia* gummosa and azospirillum spp. on Cl-compounds. Ann. Acad. Bras. Cienc., v.54, p.457, 1982

SANGER, F.; COULSON R. A rapid method for determining sequence in DNA by primed synthesis with DNA polymerase. J. Mol. Biol., v.94, p.441-448, 1975.

SANGER, F.; COULSON, A.R.; HONG, G.F.; HILL, D.F.; PETERSEN, G.B. Nucleotide sequence of bacteriophage lambda DNA. J Mol Biol., v.162, p.729-773, 1982.

SCHNAITMAN, C.A.; KLENA, J.D. Genetics of lipopolysaccharide biosynthesis in enteric bacteria. **Microbiol. Rev.**, v.57, p.655-685, 1993.

SCOTT, A.I. & ROESSNER, C.A. Biosynthesis of cobalamin (vitamin B_{12}). Biochem. Society, v.30, p.613-620, 2002.

SELKOV, E.; MALTSEV, N.; OLSEN, G.J.; OVERBEEK, R.; WHITMAN, W.B. A reconstruction of the metabolism of *Methanococcus jannaschii* from sequence data. **Gene**, v.197, p.11-26, 1997.

SMITH, C.L.; KLCO, S.R.; CANTOR, C.R. Pulsed-field gel electrophoresis and the technology of large DNA molecules. *In* Genome analysis. A practical approach, DAVIES, K.E. (ed.), IRL Press, Oxforg, p.41-72, 1988.

SMITH, D.R.; DOUCETTE-STAMM, L.A.; DELOUGHERY, C. Complete genome sequence of *Methano-bacterium thermoautotrophicum* Δ H: functional analysis and comparative genomics. J. bacteriol., v.179, p.7135-7155, 1997.

SOBRAL, B.W.S.; HONEYCUTT, R.J.; ATHERLY, A.G.; MCCLELLAND, M. Eletrophoretic separation of the three *Rhizobium meliloti* replicons. J. Bacteriol., v.173, p.5173-5180, 1991.

STADEN, R. Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. Nucl. Acids Res., v.10, p.4731-4751, 1982.

STADEN, R. Graphic methods to determine the function of nucleic acid sequences. Nucl. Acids Res., v.12, p.521-538, 1984.

STEENHOUDT, O. & VANDERLEYDEN, J. Azospirillum, a free-living nitrogen-fixing bacterium closely associated with grasses: genetic, biochemical and ecological aspects. FEMS Microbiol. Lett., v.24, p.487-506, 2000.

STEVENSON, G.; ANDRIANOPOULOS, K.; HOBBS, M.; REEVES, P.R. Organization of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. **J. Bacteriol.**, v.178, p.4885-4893, 1996.

STRYER, L.; BERG, J.M.; TYMOCZKO, J.L. Biochemistry. W. H.Freeman and Co., New York, 2002.

SUWANTO, A. & KAPLAN, S. Physical and genetic mapping of the *Rhodobacter sphaeroides* 2.4.1. genome: presence of two unique circular chromosomes. J. Bacteriol., v. 171, p.5850-5859, 1989.

SWARTZ, D.C.; CANTOR, C.R. Separation of yeast chromosome sized DNAs by pulsed-field gradient gel eletrophoresis. Cell, v.37, p.67-75, 1984.

TAL, S.; OKON, Y. Production of the reserve material poly- β -hydroxybutyrate and its function in *Azospirillum brasilense* Cd. Can. J. Microbiol., v.31, p.608-613, 1985.

TARRAND, J.J.; KRIEG, N.R.; DÖBEREINER, J. a taxonomic study of the *Spirillum lipoferum* group, with descriptions of new genus, *Azospirillum* gen. nov. and two species, *Azospirillum* lipoferum (Beijerinck) comb. nov. and *Azospirillum brasilense* sp. nov. **Can. J. Microbiol.**, v.24, p.967-980, 1978.

TARTOF, K.D.; HOBBS, C.A. Improved media for growing plasmid and cosmid clones. Focus, v.9, p.12, 1987.

TATUSOV, R.L.; NATALE, D.A.; GARKAVTSEV, I.V.; TATUSOVA, T.A.; SHANKAVARAM, U.T.; RAO, B.S.; KIRYUTIN, B.; GALPERIN, M.Y.; FEDOROVA, N.D.; KOONIM, E. The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucl. Acids. Res., v.29, p.22-28, 2001.

TAYLOR, B.L. & ZHULIN, I.B. PAS domains: Internal sensora of oxygen, redox potential, and light. Microbiol. Mol. Biol. Rev., v.63, p.479-506, 1999.

TELENIUS, H.; CARTER, N.; BEBB, C.; NORDENSKJOLD, M. *et al.* Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. **Genomics**, v.13, p.718-725, 1992.

TIEN, T.M.; GASKINS, M.H.; HUBBELL, D.H. Plant growth produced by *Azospirillum brasilense* and their effect on the growth of Pearl Millet (*Pennisetum americanum* L.). Appl. Envir. Microbiol., v.37, p.1016-1024, 1979.

VAN BASTELAERE, E.V.; LAMBRECHT, M.; VERMEIREN, H.; DOMMELEN, A.V.; KEIJERS, V.; PROOST, P.; VANDERLEYDEN, J. Characterization of a sugar-binding from *Azospirillum brasilense* mediating chemotaxis to and uptake of sugars. **Mol. Microbiol.**, v.32, p.703-714, 1999.

VANDE BROEK, A.; VANDERLEYDEN, J. The genetics of the *Azospirillum*plant root association. **Crit. Rev. Plant Sci.**, v.14, p.445-466, 1995a.

VAN PEE, K.H. Biosynthesis of halogenated metabolites by bacteria. Annu. Rev. Microbiol., v.50, p.375-399, 1996.

VIEILLE, C.; ELMERICH, C. Characterization of an *Azospirillum brasilense* Sp7 plasmid genes homologous to *Alcaligenes eutrophus phbB* and to *Rhizobium meliloti nodG*. **Mol. Gen. Genet.**, v.231, p.375-384, 1992.

VIEILLE, C. & ELMERICH, C. Characterization of two Azospirillum brasilense Sp7 plasmid genes homologous to *Rhizobium meliloti nodPQ*. **Mol. Plant-Microbe Interact.**, v.3, p.389-400, 1990.

VIPREY, V.; PERRET, X.; BROUGHTON, W.J. Host-plant invasion by rhizobia. Subcell Biochem., v.33, p.437-456, 2000.

VLASSAK, K; REYNDERS, L. Paper present at Steenbock-Kettering International Symp. Nitrogen Fixation (Madison, Wisconsis), June 12-16.

WAELKENS, F.; MARIS, M.; VERRETH, C.; VAMDERLEYDEN, J.; VAN GOOL, A. *Azospirillum* DNA shows homology with *Agrobacterium* chromosomal virulence genes. **FEMS Microbiol. Lett.**, v.43, p.241-246, 1987. WETSBY, C.A.; ENDERLIN, C.S.; STEINBERG, N.A.; JOSEPH, C.M.; MEEKS, J.C. Assimilation of ¹⁵NH₄⁺ by *Azospirillum brasilense* under nitrogen limitation and excess. J. Bacteriol., v.169, p.4211-4214, 1987.

WHITE, O.; EISEN, J.A.; HEIDELBERG, J.F.; HICKEY, E.K.; PETERSON, J.D.; DODSON, R.J.; HAFT, D.H.; GWINN, M.L.; NELSON, W.C.; RICHARDSON, D.L.; MOFFAT, K.S.; QIN, H.; JIANG, L.; PAMPHILE, W.; CROSBY, M.; SHEN, M.; VAMATHEVAN, J.J.; LAM, P.; MCDONALD, L.; UTTERBACK, T.; ZALEWSKI, C.; MAKAROVA, K.S.; ARAVIND, L.; DALY, M.J.; MINTON, W.; FLEISCHMANN, R.D.; KETCHUM, K.A.; NELSON, K.E.; SALZBERG, S.; SMITH, H.O.; VENTER, J.C.; FRASER, C.M. Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1. Science, v.286, p.1571-1577, 1999.

WHITFIELD, C.; ROBERTS, I.S. Structure, assembly and regulation of expression of capsules in *Escherichia coli*. Mol. Microbiol., v.31, p.1307-1319, 1999.

WHITMAN, W.B.; COLEMAN, D.C.; WIEBES, W.J. Prokaryotes: The unseen majority. **Proc. Natl. Acad. Sci.**, v.95, p.6578-6583, 1998.

WOOD, A.G.; MENEZES, E.M.; DYKSTRA, C.; DUGGAN, D.E. Methods to demonstrate the megaplasmids (or minichromosomes) in *Azospirillum*, p.18. *In* W. Klingmüller (ed), *Azospirillum* I: genetics, physiology, ecology. Burkhäuser Verlag, Basel, Switzerland. 1982

WOOD, D.W.; SETUBAL, J.C.; KAUL, R.; MONKS, D.E.; KITAJIMA, J.P.; OKURA, V.K.; ZHOU, Y.; CHEN, L.; WOOD, G.E.; ALMEIDA, N.F.; WOO, L.; CHEN, Y.; PAULSEN, I.T.; EISEN, J.A.; KARP, P.D.; BOVEE, D.; CHAPMAN, P.; CLENDENNING, J.; DEATHERAGE, G.; GILLET, W.; GRANT, C.; KUTYAVIN, T.; LEVY, R.; LI, M.; McCLELLAND, E.; PALMIERI, A.; RAYMOND, C.; ROUSE, G.; SAENPHIMMACHAK, C.; WU, Z.; ROMERO, P.; GORDON, D.; ZHANG, S.; YOO, H.; TAO, YUMIN, T.; BIDDLE, P.; JUNG, M.; KRESPAN, W.; PERRY, M.; GORDON-KAM, B.; LIAO, L.; KIM, S.; HENDRICK, C.; ZHAO, Z.; DOLAN, M.; CHUMLEY, F.; TINGEY, S.V.; TOMB, J.; GORDON, M.P.; OLSON, M.V.; NESTER, E.W. The genome of natural genetic engineer *Agrobacterium tumafaciens* C58. Science, v.294, p.2317-2323, 2001.

YAMAMOTO, I.; SAIKI, T.; LIU, S.; LJUNGDAHL, L.G. Purification and properties of NADP-dependent formate dehydrogenase from *Clostridium thermaceticum*, a tungsten-selenium-iron protein. J. Biol. Chem., v. 258, p. 1826-1832, 1983.

ZAMAROCZY, M.; PAQUELIN, A.; ELMERICH, C. Functional organization of the *glnB-glnA* cluster of *Azospirillum brasilense*. J. Bacteriol., v.175, p.2507-2515, 1993.

ZEMEL, R.I.; ANWAR, R.A. Pyruvate-uridine diphospho-N-acetylglucosamine transferase. J. Biol. Chem., v.250, p,3185-3192, 1975.

ZHANG, L.; CUI, X.; SCHIMITT, K.; HUBERT, R. *et al.* Whole genome amplification from a single cell: implications for genetic analysis. **Proc. Nat.** Acad. Sci. USA., v.89, p.5847-5851, 1992.

ZHULIN, I.B. & ARMITAGE, J.P. Motility, chemokinesis, and methylationindependent chemotaxis in *Azospirillum brasilense*. J. Bacteriol., v. 175, p. 952-958, 1993. ZHULIN, I.B.; BESPALOV, V.A.; JOHNSON, M.S.; TAYLOR, B.L. Oxygen taxis and proton motive force in *Azospirillum brasilense*. J. Bacteriol., v.178, p.5199-5204, 1996.

ZUERNER, R.I.; HERMANN, J.L.; SAINT GIRONS, I. Comparison of genetic maps for two *Leptospira interrogans* serovars provides evidence for two chromosomes and intraspecies heterogeneity. J. Bacteriol., v.175, p.5445-5451, 1993.

ZUPAN, J.; MUTH, T.R.; DRAPER, O.; ZAMBRYSKI, P. The transfer of DNA from *Agrobacterium tumefaciens* into plants: a feast of fundamental insights. **Plant J.**, v.23, p.11-28, 2000.