

UNIVERSIDADE FEDERAL DO PARANÁ

PREDIÇÃO DE RNAs CURTOS EM
***Herbaspirillum seropedicae* SmR1**

CURITIBA

2010

GUILHERME MARTINS WILLEMANN

PREDIÇÃO DE RNAs CURTOS EM
***Herbaspirillum seropedicae* SmR1**

Dissertação apresentada como requisito parcial para a obtenção do grau de mestre, pelo Programa de Pós-Graduação em Bioinformática da Universidade Federal do Paraná.

Orientadora:

Prof^a. Dr^a. Maria Berenice R. Steffens

Co-orientadora:

Prof^a. Dr^a. Jeroniza N. Marchaukoski

CURITIBA

2010

**À minha filha, Julia.
À minha esposa, Juliana.**

“If I have seen farther than others, it is because I was standing on the shoulders of giants” – **Sir Isaac Newton.**

AGRADECIMENTOS

Agradeço às minhas orientadoras, Professora Dra. Maria Berenice Reynaud Steffens e Professora Dra. Jeroniza Nunes Marchaukoski, pela paciência, aprendizado e grande ajuda no desenvolvimento desta dissertação. Muito obrigado.

À minha esposa, Juliana, por ser tão especial em minha vida e estar ao meu lado em todos os momentos, me apoiando, me incentivando e sempre me mostrando o lado positivo das coisas. Obrigado pela paciência e compreensão da minha ausência por quase dois anos.

Aos meus pais, Rita e Gonzaga, pela educação que me proporcionaram, pelo apoio e incentivo.

Ao meu avô, Lúcio, pelo apoio espiritual que me proporciona todos os dias.

A minhas avós, Fernandina e Pedra, pelo amor e preocupação.

Ao meu sogro e a minha sogra, Ivana e Mauro, pelo carinho.

Ao Prof. Nilson Ribeiro Modro, da Universidade do Estado de Santa Catarina, pelo apoio e incentivo.

Ao Prof. Claudio Valverde, pela aprendizagem e hospitalidade durante minha visita à Universidade Nacional de Quilmes.

Aos colegas da Bioinformática, em especial Ademir, Waldir, Valdemar, Rafaela, Rosa, Eduardo e Dieval.

Aos colegas da Bioquímica, Leandro, Vinícius e Marco Antonio.

Ao Programa de Pós-Graduação em Bioinformática pela oportunidade.

Ao Departamento de Bioquímica pelos genomas disponibilizados.

Ao CNPq pelo apoio financeiro.

A todos que contribuíram para a realização deste trabalho, e que de alguma forma eu não tenha citado.

Muito obrigado!!

LISTA DE FIGURAS

FIGURA 1 - DIFERENTES MECANISMOS DE AÇÃO DOS SRNAS. _____	15
FIGURA 2 – ALÇA DE RNA FORMADA DURANTE A TERMINAÇÃO INTRÍNSECA. ____	22
FIGURA 3 – CAPTURA DE TELA DO NEURAL NETWORK PROMOTER PREDICTION. 29	29
FIGURA 4 - CAPTURA DE TELA DO ARQUIVO DE SAÍDA DO NNPP. _____	29
FIGURA 5 – CAPTURA DE TELA DO SOFTWARE BPROM. _____	30
FIGURA 6 - CAPTURA DE TELA DO ARQUIVO DE SAÍDA DO SOFTWARE BPROM. __	31
FIGURA 7 - CAPTURA DE TELA DO ARQUIVO DE SAÍDA DO SOFTWARE TRANSTERMHP. _____	33
FIGURA 8 – CAPTURA DE TELA DO ARQUIVO DE SAÍDA DO SOFTWARE RNAMOTIF.35	35
FIGURA 9 - FLUXO DE TRABALHO PARA A IDENTIFICAÇÃO DE SRNAS EM GENOMAS DE PROCARIOTOS. _____	37
FIGURA 10 – CAPTURA DE TELA DA FERRAMENTA PARA A EXTRAÇÃO DAS REGIÕES INTERGÊNICAS. _____	38
FIGURA 11 – REPRESENTAÇÃO DE UMA REGIÃO INTERGÊNICA. _____	38
FIGURA 12 – FORMATAÇÃO DOS DADOS DE SAÍDA DO SOFTWARE BPROM. ____	39
FIGURA 13 - FORMATAÇÃO DOS DADOS DE SAÍDA DO SOFTWARE NNPP. _____	40
FIGURA 14 – FORMATAÇÃO DOS DADOS DE SAÍDA DO SOFTWARE RNAMOTIF. __	41
FIGURA 15 – CRUZAMENTO DOS DADOS DE PROMOTORES E TERMINADORES. __	42
FIGURA 16 – CAPTURA DE TELA DO SRNATOOL – FERRAMENTA INTEGRADORA. _	46
FIGURA 17 – FLUXO DE TRABALHO COM OS RESULTADOS OBTIDOS EM CADA ETAPA. _____	47
FIGURA 18 – 4.5S SRNA _____	57
FIGURA 19 – 6S SRNA _____	57
FIGURA 20 – CANDIDATO DA IGR 1227 (TABELA 2) _____	58
FIGURA 21 – CANDIDATO DA IGR 817 (TABELA 3) _____	58

LISTA DE QUADROS

QUADRO 1 - FATORES SIGMA INDUZIDOS POR CONDIÇÕES AMBIENTAIS ESPECÍFICAS. _____	18
QUADRO 2 - DIFERENTES SEQUÊNCIAS CONSENSO SÃO RECONHECIDAS POR DIFERENTES PROMOTORES. _____	19

LISTA DE TABELAS

TABELA 1 - CANDIDATOS COM SIMILARIDADE ENTRE ESPÉCIES E COM CONFIRMAÇÃO NO RFAM. _____	49
TABELA 2 - CANDIDATOS COM SIMILARIDADE ENTRE ESPÉCIES E SEM CONFIRMAÇÃO NO RFAM. _____	50
TABELA 3 - CANDIDATOS SEM SIMILARIDADE ENTRE ESPÉCIES E SEM CONFIRMAÇÃO NO RFAM. _____	52

LISTA DE ABREVIATURAS

ORF (Open Reading Frame) – Porção do genoma de um organismo que contém uma sequência de bases codificadora de um peptídeo ou uma proteína. Inicia no códon de início de tradução e termina no códon de término de tradução

Start Códon – Sequência de três bases que sinaliza o início da tradução

Stop Códon – Sequência de três bases que sinaliza o final da tradução

UTR (Untranslated Region) – Sequência de bases não codificada e situada entre o Sítio de Início de Transcrição e o Códon de Início de Tradução

IgR (Intergenic Region) – Região do genoma que não contém ORFs

DNA – Ácido Desoxirribonucléico

RNA – Ácido Ribonucléico

sRNA – RNA curto

bp – Pares de base

Fasta – Formato usado para armazenar sequências de aminoácidos e bases em modo texto

TCP/IP (Transmission Control Protocol Internet Protocol) – Pacote de protocolos que organizam a conexão entre os computadores e a Internet

HTTP (Hypertext Transfer Protocol) – protocolo de transferência de hipertexto usado na Internet

FTP (File Transfer Protocol) – Protocolo de transferência de arquivos

RESUMO

Herbaspirillum seropedicae é uma bactéria diazotrófica endofítica associada a diversos grupos de plantas, principalmente gramíneas de importância agrícola. Seu genoma já foi sequenciado e anotado pelo consórcio GENOPAR. Como outros seres vivos, esta bactéria tem um conjunto de RNAs regulatórios conhecidos como *small* RNAs (sRNAs). Estas moléculas têm um papel importante em vários processos celulares e possuem diversos mecanismos de ação, como o pareamento antisense entre o sRNA e o mRNA alvo. RNAs curtos podem ser identificados por meio de predições computacionais utilizando características comuns compartilhadas: tamanho entre 50 e 400 *bp*, presença de promotores e terminadores, similaridade entre gêneros ou espécies filogeneticamente relacionadas. O propósito deste projeto foi a predição *in silico* de sRNAs em regiões intergênicas do genoma de *H. seropedicae*, através do desenvolvimento e utilização do sRNATool, uma ferramenta integradora de dados. Para a predição dos promotores foram utilizados o Bprom e o NNPP. Ambos utilizam uma rede neural de sRNAs de *E. coli* confirmados experimentalmente, com uma eficiência de 80% nas predições. Para as sequências de terminadores intrínsecos foram utilizados o RNAMotif e o TranstermHP. Foi realizada a formatação e o cruzamento dos dados gerados, subtraindo a posição final do promotor e a posição final do terminador. Obtiveram-se 200 sequências candidatas que foram submetidas a comparações com espécies próximas, principalmente *Herbaspirillum rubrisubalbicans* M1. Cinquenta e sete candidatos foram selecionados por possuírem elevada similaridade, e, desta relação, confirmou-se dois sRNAs no banco de dados Rfam: 6S e 4.5S. O sRNATool mostrou-se eficaz na predição de sequências candidatas a sRNAs e pode ser aplicado em outros genomas de procariotos.

Palavras chave: *Herbaspirillum seropedicae*, Genoma, Região Intergênica, Promotor, Terminador Intrínseco, RNA curto, sRNA, ncRNA.

ABSTRACT

Herbaspirillum seropedicae SmR1 is an endophytic diazotrophic organism associated with several groups of plants, mainly crops with agricultural importance. Its genome was sequenced and annotated by the GENOPAR consortium. As other living being, this bacterium has a set of regulatory RNAs known as small RNAs (sRNAs). These molecules have important roles on several cellular process and make use of some action mechanisms like antisense pairing between the sRNA and the target mRNA. sRNAs can be identified by means of computational predictions using common features: length between 50 and 400 bp, promoters, terminators and similarity among related species. The purpose of this project was the in silico prediction of sRNAs in intergenic regions of *Herbaspirillum seropedicae* SmR1 through the development and application of the sRNATool, a data integrative software. In order to predict promoters were used BProm and NNPP. Both of them utilize a neural network of sRNAs experimentally confirmed with 80% efficiency on the predictions. For the intrinsic terminators were utilized RNAMotif and TranstermHP. After predicting promoters and terminators was realized the reformatting and the data crossing, subtracting the promoter's final position and the terminator's final position. Were obtained 200 candidates sequences that were compared with related species, mainly *Herbaspirillum rubrisubalbicans* M1. 57 candidates were selected by possessing high similarity and 2 candidates could be confirmed: 6S and 4.5S. The sRNATool proved be effective and can be applied in others prokaryotic genomes.

Keywords: *Herbaspirillum seropedicae*, Genome, Intergenic Region, Promoter, Intrinsic Terminator, Small RNA, sRNA, ncRNA.

SUMÁRIO

LISTA DE FIGURAS	4
LISTA DE QUADROS	5
LISTA DE TABELAS	6
LISTA DE ABREVIATURAS	7
RESUMO	8
ABSTRACT	9
1 INTRODUÇÃO	12
1.1 Ácido Desoxirribonucléico	12
1.2 RNAs Curtos Bacterianos	13
1.3 Promotores Bacterianos	17
1.4 Terminadores Bacterianos	19
1.5 <i>Herbaspirillum seropedicae</i> SmR1	22
2 OBJETIVO	24
3 MATERIAIS E MÉTODOS	25
3.1 Genomas	25
3.2 Sistema Operacional	25
3.3 Linguagens de Programação	25
3.3.1 Java	25
3.3.2 Bash	26
3.4 Ambiente de Programação	26
3.5 Softwares de Bioinformática	27
3.5.1 BLAST	27
3.5.2 Rfam	27
3.5.3 Preditores de Promotores	27
3.5.3.1 NNPP (Neural Network Promoter Prediction)	27
3.5.3.2 BProm	30
3.5.4 Preditores de Terminadores	31

3.5.4.1	TranstermHP	31
3.5.4.2	RNAMotif	33
3.6	Metodologia para Seleção de Candidatos a sRNAs no Genoma de <i>H. seropedicae</i>	
SmR1		36
3.6.1	Extração das Regiões Intergênicas	36
3.6.2	Predição de Promotores e Terminadores	38
3.6.3	Formatação dos Dados	39
3.6.4	Cruzamento dos Dados	41
3.6.5	Busca por Similaridade de Sequências entre Gêneros ou Espécies Bacterianas	43
3.6.6	Confirmação dos Candidatos	43
4	RESULTADOS E DISCUSSÃO	44
5	CONCLUSÃO	61
	REFERÊNCIAS	62
	APÊNDICE I – CANDIDATOS COM SIMILARIDADE DE SEQUÊNCIA ENTRE ESPÉCIES PRÓXIMAS	75
	APÊNDICE II – CANDIDATOS SEM SIMILARIDADE DE SEQUÊNCIAS ENTRE ESPÉCIES PRÓXIMAS	105

1. INTRODUÇÃO

1.1 Ácido Desoxirribonucléico

O ácido desoxirribonucléico (DNA) é a macromolécula que contém as instruções genéticas que coordenam o desenvolvimento e funcionamento de todos os seres vivos e alguns vírus. Suas funções principais são armazenar informações necessárias para a síntese de diferentes tipos de moléculas de ácido ribonucléico (RNA) e transmissão da carga genética ao longo das gerações (PIERCE, 2009). O DNA é um polímero constituído por quatro unidades nucleotídicas diferentes que se repetem aos milhares e estão ligadas por ligação fosfodiéster para formar a cadeia polinucleotídica. Cada um dos quatro nucleotídeos que compõe o DNA é composto por uma desoxirribose/trifosfato ligada a uma base nitrogenada: adenina (A), citosina (C), guanina (G) ou timina (T) (DALE, 2004). Em organismos vivos, o DNA existe como uma molécula composta por duas longas cadeias polinucleotídicas enroladas, formando uma dupla hélice para a direita e mantida estável por pontes de hidrogênio formadas entre os pares de bases AT e CG (DALE, 2004). O número de moléculas de DNA presentes nas células é bastante variável em procariotos, eucariotos e até mesmo em vírus. Cada molécula de DNA, por sua vez, está associada a proteínas para formar o cromossomo. O conjunto haplóide de cromossomos de uma célula constitui o genoma da célula (WATSON *et al.*, 2006). O tamanho e a composição do genoma variam largamente de uma espécie a outra e estão relacionados com a complexidade do organismo. Genomas bacterianos possuem tamanho variando de ~0,16 (*Carsonella ruddii*) a ~10 milhões de pares de base (*Solibacter usitatus*) e eucariotos de ~8 milhões (*Pratylenchus coffeae*) a ~130 bilhões de pares de base (*Protopterus aethiopicus*). O genoma humano possui cerca de três bilhões de pares de base (GREGORY *et al.*, 2006).

No acesso em outubro de 2010 foram encontrados 831 genomas completos depositados na base de dados do Genbank, sendo 795 de organismos procariotos e 36 de eucariotos. Outros 273 projetos de organismos eucarióticos e 564 procarióticos permanecem em desenvolvimento¹.

¹ [HTTP://www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)

O DNA contém todos os genes que serão traduzidos em proteínas e, para isso, eles são primeiramente transcritos em um tipo específico de RNA, o mensageiro. Estas moléculas então serão decodificadas em uma sequência específica de aminoácidos. Na transcrição, os genes que possuem alguma relação com as proteínas necessárias ao metabolismo celular são expressos constitutivamente ou de forma induzida, na forma de RNA mensageiro (LEWIN, 2007).

Em genomas bacterianos, grande parte do código é referente a genes que codificam proteínas, e o restante tem importância estrutural ou está envolvido na regulação do uso da informação genética, como RNAs transportadores, ribossomais e regulatórios, sendo este grupo mais conhecido como RNAs não codificantes (MOULTON, 2005).

1.2 RNAs Curtos Bacterianos

O primeiro RNA não-codificante (ncRNA) regulatório de bactéria foi descoberto no início da década de 70, e desde então várias destas moléculas têm sido identificadas. Entretanto, o avanço das pesquisas sobre o entendimento de seus diversos mecanismos de ação e o importante papel desempenhado na regulação da expressão gênica ocorreu somente na última década do século XXI. Em geral, esta classe de moléculas é também citada como micro RNAs (miRNAs), embora o termo RNA curto ou *small RNA* (sRNA) seja mais utilizado para bactérias e *archeas*. Controlando vias metabólicas ou respostas de *stress*, os ncRNAs desempenham funções importantes nos diversos processos biológicos, incluindo a regulação de transportadores de membranas, o metabolismo do ferro e a patogênese (MASSÉ, MAJDALANI e GOTTESMAN, 2003; GOTTESMAN, 2005). O grupo que melhor caracteriza os ncRNAs é o chamado RNA curto com função regulatória (*small regulatory RNA* [sRNA]), cuja função é regular a expressão gênica através do pareamento de bases com a extremidade 5' do mRNA alvo, bloqueando o acesso ao sítio de ligação ao ribossomo e/ou ao códon de início de tradução (ALTUVIA *et al.*, 1998; ZHANG *et al.*, 2003). Estas moléculas são codificadas *in trans* em relação ao RNA alvo, embora já existam evidências de sRNAs codificados

in cis (LIU e CAMILLI, 2010). Outros ncRNAs podem ligar-se a proteínas a fim de modular suas atividades (PICHON & FELDEN, 2007).

Genes codificadores de RNAs curtos regulatórios estão dispersos no genoma e algumas de suas características incluem tamanho reduzido, que pode variar de 50 a 400 pares de base, presença de sinais de promotores e sequências terminadoras Rho-independente, sequências com alta similaridade entre espécies filogeneticamente próximas e com pouca similaridade entre espécies distantes, além da conservação de sua estrutura secundária. A maioria deles estão localizados entre regiões codificadoras de proteínas, ORFs (do inglês *Open Reading Frames*), ou seja, nas regiões intergênicas do genoma (WASSARMAN *et al.*, 2001).

O número de sRNAs prováveis e confirmados tem crescido substancialmente devido à utilização de métodos computacionais que auxiliam principalmente na predição de sinais transcricionais. Outras abordagens incluem a detecção direta por meio de *microarrays* ou *northern blotting*, isolamento direto (*RNomics*), copurificação com proteínas de ligação de RNA e pirosequenciamento (WASSERMAN *et al.*, 2001; LIVNY *et al.*, 2007; ALTUVIA 2007). Estudos pioneiros foram iniciados alguns anos atrás quando vários grupos descobriram alguns sRNAs em regiões intergênicas de *E. coli* (GOTTESMAN, 2004). Nesses estudos, os sRNAs foram identificados com base em suas associações com características genéticas comuns a outros sRNAs anteriormente conhecidos. Atualmente, aproximadamente 100 sRNAs estão identificados em *E. coli* e, nos últimos anos, a busca por sRNAs foi estendida para muitas espécies bacterianas, como *Bacillus subtilis*, *Vibrio cholerae*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Streptomyces coelicolor*, *Salmonella enterica*, *Mycobacterium tuberculosis*, *Sinorhizobium meliloti* e *Listeria monocytogenes* (LIVNY *et al.*, 2006; PICHON e FELDEN, 2005; SWIERCZ *et al.*, 2008; MANDIN *et al.*, 2008; VALVERDE *et al.*, 2008).

Em relação ao mecanismo de ação dos sRNAs, o mais comum é o pareamento antisenso entre o RNA curto regulatório e o RNA mensageiro alvo. Por exemplo, a ligação de RyhB ao mRNA *sodB* promove a sua degradação (Figura 1a), enquanto a ligação do mesmo RNA curto à região UTR do mRNA *shiA* ativa a sua tradução (Figura 1b) (MASSÉ e GOTTSMAN, 2002). A interação entre sRNA e mRNA ocorre em regiões curtas e de complementaridade de sequências imperfeitas, e pode ser estabilizada pela proteína Hfq, uma chaperona de RNA. O pareamento

de bases geralmente ocorre próximo ao sítio de início de tradução dos mRNAs alvos (MASSÉ *et al.*, 2003; MAJDALANI *et al.*, 2005). É o caso do sRNA MicF, que reprime a produção da proteína de membrana externa OmpF. Vinte bases de MicF pareiam por complementaridade imperfeita com o OmpF mRNA, inibindo a sua tradução (DELIHAS e FORST, 2001).

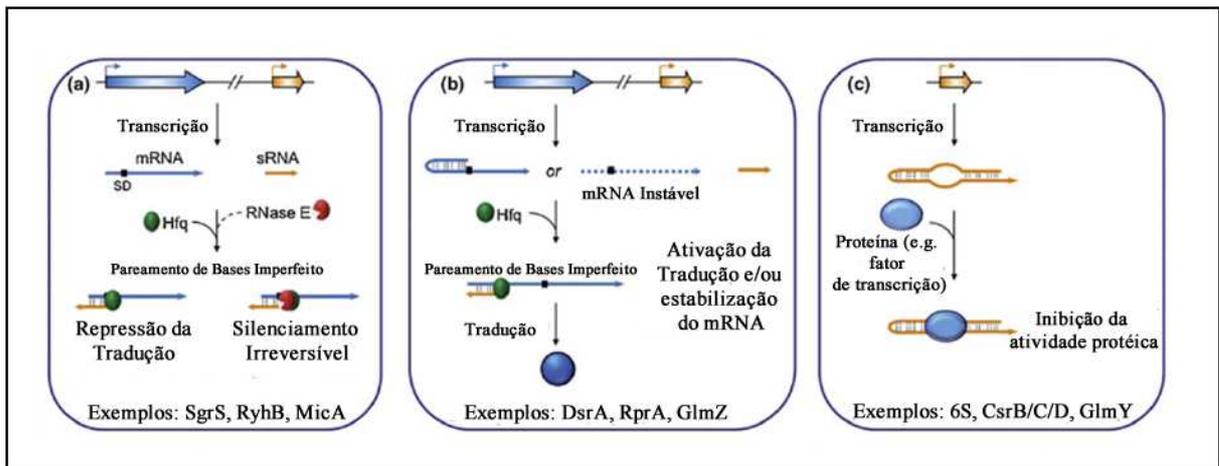


Figura 1 - Diferentes mecanismos de ação dos sRNAs.

a) Repressão da tradução. b) Ativação da tradução ou estabilização do mRNA.

c) Inibição da atividade protéica.

Fonte: modificado de LIU e CAMILLI, 2010.

A identificação de sRNAs pode se dar através de métodos experimentais ou computacionais, cada um deles dependendo de parâmetros específicos (WILLKOMM *et al.*, 2005). Na abordagem computacional utilizam-se softwares para a predição de promotores e terminadores, além da análise da conservação de sequências entre espécies próximas e a verificação de suas estruturas secundárias. Os três estudos computacionais pioneiros aplicados à busca de sRNAs foram baseados principalmente na busca de sequências conservadas em regiões intergênicas de espécies próximas juntamente com um provável promotor σ^{70} e um terminador de transcrição intrínseco (ARGAMAN *et al.*, 2001; WASSARMAN *et al.*, 2000; CHEN *et al.*, 2002). Estes estudos levaram à descoberta de 42 novos sRNAs em *E. coli*. Além da conservação da sequência, o método utilizado por ARGAMAN e colaboradores (2001) e WASSARMAN e colaboradores (2002) utilizou sinais de transcrição e limitou de 50 a 400 *bp* o tamanho do transcrito. Um quarto estudo de

predição de genes codificadores de sRNA foi desenvolvido por RIVAS e colaboradores (2001), e considerou somente a conservação da estrutura secundária.

Na abordagem experimental diversas técnicas já foram utilizadas, como o sequenciamento direto após a marcação fluorescente (TIAN-PING *et al.*, 1996), análise transcriptômica usando *microarray* (WASSARMAN *et al.*, 2001), obtenção de bibliotecas de cDNA enriquecidas (TANG *et al.*, 2002, 2005; VOGEL *et al.*, 2003; KAWANO *et al.*, 2005), RNomics (VOGEL *et al.*, 2003) e co-purificação de sRNAs com proteínas (ZHANG *et al.*, 2003; TANG *et al.*, 2005; CHRISTIANSEN *et al.*, 2006).

Alternativamente, alguns sRNAs interagem com proteínas, incluindo fatores de transcrição (Figura 1c). Estas proteínas podem facilitar o reconhecimento entre o sRNA e o mRNA. Elas também podem induzir a hidrólise específica do RNA alvo e a degradação do sRNA, assegurando que o regulador seja desativado uma vez que tenha exercido sua função (PICHON e FELDEN, 2007).

O efeito regulatório dos sRNAs pode demandar a participação de proteínas como a Hfq e a RNase E. Neste caso, ocorre o pareamento sRNA/RNA alvo, e em seguida, a formação de um complexo com as proteínas RNase E e Hfq, resultando na inibição da tradução e na degradação RNase E-dependente mRNA alvo (AIBA, 2007). A Hfq é uma proteína que possui homologia estrutural e de sequência com proteínas Sm, e é conhecida por estabilizar sRNAs *in vivo* e facilitar o pareamento dos sRNAs com os mRNAs alvos *in vitro*. Ela também se liga preferencialmente a um RNA de fita simples, interagindo com RNAs ricos em resíduos de uracila (U) e adenina (A). Um dos mRNAs alvos mais estudados é o rpoS mRNA, que codifica a subunidade sigmaS para a RNA polimerase. Hfq nesse caso é necessária para a expressão de rpoS *in vivo*. A tradução do rpoS mRNA é reprimida por um *stem loop* que bloqueia o acesso do ribossomo.

Além do envolvimento de sRNAs no metabolismo de RNA via interação com proteínas, estas moléculas podem participar de eventos regulatórios que envolvem o metabolismo do DNA (regulação da transferência e controle da replicação de plasmídeos) e o metabolismo de proteínas (controle da tradução, do trânsito e da qualidade) (PICHON e FELDEN, 2007).

Apesar dos esforços e das técnicas citadas, ainda não se pode definir com certeza a quantidade de sRNAs presente nos genomas bacterianos, porém estima-

se que deva existir entre 200 a 300 em um genoma médio de 5 milhões de bases, como é o caso da *Escherichia coli* (VALVERDE, *et al.*, 2008). Esta observação se estende para outras classes de bactérias e também para *Archaeae*, indicando um vasto campo de pesquisa.

1.3 Promotores Bacterianos

A RNA Polimerase é a enzima responsável pela síntese do RNA durante a transcrição. Nas bactérias, ao contrário do que ocorre em eucariotos, uma única RNA polimerase é responsável pela síntese de mRNAs, rRNAs e tRNAs. Aproximadamente 7000 moléculas delas estão presentes em uma célula de *E. coli* (LEWIN, 2007). Uma unidade completa de RNA Polimerase, a holoenzima, de *E. coli* possui peso molecular de aproximadamente 465 kD, sendo formada por duas subunidades alfa, uma beta, uma beta' e uma subunidade sigma, também conhecida como fator sigma. As subunidades alfa, beta e beta' são produtos dos genes *rpoA*, *rpoB* e *rpoC* respectivamente. A subunidade sigma é a única variável, sendo, portanto, codificada por diferentes genes (Quadro 1). A holoenzima $2\alpha\beta\beta'\sigma$ apresenta dois componentes estruturais: a porção catalítica, formada pelas quatro primeiras subunidades e o fator sigma. Todas as subunidades da porção catalítica participam da alongação e são necessárias para a iniciação e terminação. A subunidade sigma é o fator que ativa a interação específica entre a RNA polimerase e o DNA durante o reconhecimento de promotores. O número de fatores sigma varia entre espécies bacterianas e estes podem ser distinguidos pelo peso molecular explícito em sua designação, como por exemplo, o sigma70 e o sigmaN (ou sigma54) que possuem peso molecular de 70 kD e 54 kD respectivamente (MAKINO *et al.*, 1993).

Em condições fisiológicas normais, o fator sigma70 é o responsável pela transcrição da maioria dos genes, sendo por isso considerado um *housekeeping factor*, que pode ser interpretado como fator essencial para a manutenção do metabolismo celular. Outros fatores ditos sigma alternativos são ativados em resposta a mudanças ambientais como aumento de temperatura e falta de nitrogênio. Estes fatores são σ^S , σ^{32} , σ^E e σ^{54} (Quadro 1).

Fatores sigma de <i>E. coli</i>		
Gene	Fator	Uso
rpoD	σ^{70}	Geral
rpoS	σ^S	Estresse
rpoH	σ^{32}	heat shock
rpoE	σ^E	heat shock
rpoN	σ^{54}	falta de N
fliA	σ^{28}	síntese de flagelo

Quadro 1 - Fatores sigma induzidos por condições ambientais específicas.

Fonte: Modificado de LEWIN, 2007.

A sequência de cada tipo de promotor assegura que este será reconhecido somente pela RNA Polimerase $2\alpha\beta\beta'\sigma$ por meio do fator sigma apropriado. Como promotores são tipicamente adjacentes ao gene, as sequências reconhecidas pelo fator sigma são relativas ao sítio de início de transcrição. Uma característica dos promotores σ^{70} , σ^{28} , σ^{32} e σ^E é que eles apresentam sequências conservadas posicionadas nas regiões -10 e -35 enquanto que o fator σ^{54} está posicionado na região -12 e -24 (Quadro 2).

Existem quatro características conservadas em promotores bacterianos: o sítio de início de transcrição, a sequência -10 e -35 e a região intermediária. O sítio de início de transcrição é geralmente uma purina. A sequência ou hexâmero -10 é uma região com seis pares de base e o seu centro é posicionado geralmente 10 *bp* a montante do sítio de início de transcrição. A sequência ou hexâmero -35 posiciona-se 35 *bp* a montante do sítio de início de transcrição. A distância separando a região -10 e -35 varia entre 16 a 20 *bp* em 90% dos promotores. Nas exceções, o promotor pode ser tão pequeno quanto 15 *bp* ou tão grande quanto 21 *bp*. O promotor ideal é uma sequência consistindo de um hexâmero -10 separado por 17 *bp* do hexâmero -35.

As sequências de promotores são representadas por um consenso derivado do alinhamento de todos os exemplos conhecidos, maximizando sua homologia (Quadro 2). Para uma sequência ser aceita como consenso, cada base particular deve ser predominante em sua posição. Como exemplo pode-se citar o consenso do

σ 70 de *E. coli*. Neste fator, a região -10 possui como consenso TATAAT com as respectivas porcentagens de base: 80%, 95%, 45%, 60%, 50% e 96%. A região -35 é um pouco menos conservada, tendo como sequência TTGACA e as respectivas porcentagens de ocorrência: 82%, 84%, 78%, 65%, 54% e 45%. A sequência consenso do sítio -10 consiste na maioria das vezes de bases A-T, uma configuração que favorece a separação do DNA duplex em fitas simples.

Sequências consenso de fatores sigma				
Gene	Fator	Seq. -35	Separação	Seq. -10
rpoD	σ 70	TTGACA	16-18 bp	TATAAT
rpoH	σ 32	CCCTTGAA	13-15 bp	CCCGATNT
rpoN	σ 54	CTGGNA	6 bp	TTGCA
fliA	σ 28	CTAAA	15 bp	GCCGAATAA

Quadro 2 - Diferentes sequências consenso são reconhecidas por diferentes promotores.

Fonte: Modificado de LEWIN, 2007.

A premissa de que diferentes sequências consenso são reconhecidas por diferentes fatores σ implica no fato de que uma determinada subunidade σ deve reconhecer uma sequência de bases específica no DNA. Isso confirma a existência de uma interação estreita entre a porção catalítica e o fator σ , devendo este estar posicionado de forma a proporcionar o melhor contato da Polimerase $2\alpha\beta\beta'\sigma$ com as sequências promotoras alvo.

1.4 Terminadores Bacterianos

Além dos promotores, a presença de um terminador é uma importante característica dos sRNAs. Em bactérias existem dois tipos de terminação: Rho-dependente e Rho-independente ou Intrínseca (YACHIE *et al.*, 2006). Funcionalmente, se as características do transcrito podem causar a terminação *in vitro* na ausência de um fator protéico específico, pode-se considerar a terminação

como intrínseca. Por outro lado, a terminação dependente de fatores protéicos emprega proteínas do tipo Rho (MITRA, 2009).

Terminadores Rho-dependente diferem entre si e não podem ser definidos por uma sequência consenso. A proteína ou fator Rho atua na liberação do transcrito quando o complexo de transcrição é muito estável para uma liberação espontânea. Um terminador Rho-dependente pode ocorrer em sítios no qual a sequência de um terminador intrínseco seria incompatível (CIAMPI, 2006).

Para o fator Rho terminar a síntese do transcrito ele deve ligar-se com alta afinidade a um segmento de RNA recém-transcrito e válido. Em bactérias, a tradução ocorre de forma acoplada à transcrição dos ribossomos, o que implica na ligação do mRNA ao ribossomo logo que o sítio de ligação ao ribossomo e o códon de início de tradução estejam disponíveis. Isto ocorre normalmente antes do fator Rho poder se juntar ao mRNA que emerge do complexo de transcrição. Assim, a presença de um ribossomo traduzindo o mRNA na mesma taxa no qual o RNA está sendo sintetizado previne o acesso do fator Rho ao final do gene ou do *operon* (RICHARDSON, 2002; EPSHTEIN, 2010). Uma vez que o complexo de tradução é desfeito o mRNA nascente fica disponível e, caso a sequência seja apropriada, o fator Rho se ligará e iniciará o processo de terminação. Uma sequência apropriada consiste em um segmento de 40 ou mais nucleotídeos não pareados, contendo uma proporção alta de citosina. Tal sequência é chamada de *rut site* (RICHARDSON, 2002; EPSHTEIN, 2010).

As posições dos pontos de parada em um terminador Rho-dependente são determinadas pelas características da sequência que controlam a taxa de adição dos nucleotídeos pela RNA polimerase. Essa taxa pode variar consideravelmente de uma posição a outra, aumentando para pontos de pausa característicos que são distribuídos de forma parecida aos pontos de parada em um terminador Rho-dependente (RICHARDSON, 2002).

Alguns terminadores rho-dependente encontram-se em regiões regulatórias separando genes de seus promotores, o que pode prover à célula um mecanismo conveniente para o controle da expressão gênica. A maioria deles ocorre dentro de sequências codificantes, contribuindo para prevenir expressões desnecessárias de genes (CIAMPI, 2006).

A terminação intrínseca, também conhecida como terminação rho-independente, está presente em cerca de metade dos sítios de terminação das unidades de transcrição bacterianas. Terminadores rho-independentes têm duas características estruturais: um grampo rico em GC e uma região no final da unidade rica em uracilas (Figura 2). Ambas as características são necessárias para a terminação (CIAMPI, 2006; YACHIE *et al.*, 2006; SANTANGELO *et al.*, 2009; MITRA, 2009). A distância típica entre o grampo e a cauda poli-U varia de 7 a 9 bases. No genoma de *E. coli* existem aproximadamente 1100 sequências desse tipo (UNNIRAMAN *et al.*, 2002; ARTSIMOVITCH e HENKIN, 2009; ROBERTS *et al.*, 2008; VON HIPPEL, 1991).

Tanto as sequências do grampo quanto o comprimento da cauda influenciam a eficiência da terminação. A eficiência *in vitro*, porém, varia de 2% a 90%, e não tem correlação com o grampo ou com o número de uracilas. Elas são necessárias mas não suficientes, e parâmetros adicionais influenciam a interação com a RNA Polimerase. Em particular, as sequências a montante e a jusante do terminador intrínseco influenciam sua eficiência (ROBERTS, 2008).

O sequenciamento de genomas em larga escala tem resultado em esforços para o desenvolvimento de algoritmos para a detecção de terminadores intrínsecos (LESNIK, 2001; ERMOLAEVA *et al.*, 2000; KINGSFORD, 2007).

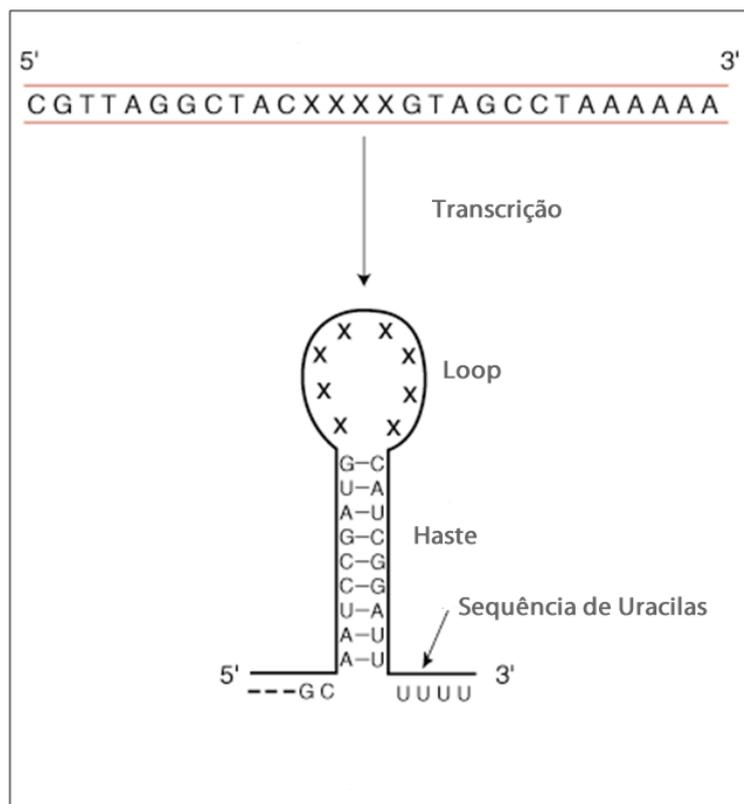


Figura 2 – Alça de RNA formada durante a terminação intrínseca.

Fonte: Autor.

1.5 *Herbaspirillum seropedicae* SmR1

A bactéria *Herbaspirillum seropedicae* é uma bactéria diazotrófica endofítica, não patogênica e membro da classe β das Proteobactérias (BALDANI *et al.*, 1986). Recebeu essa denominação por apresentar forma de espiral encontradas em associação com plantas herbáceas na cidade de Seropedicae (RJ) onde foi isolada. A estirpe SmR1 é resistente a estreptomicina e foi selecionada, *in vitro*, pelo prof Dr. Emanuel Matempi de Souza no Núcleo de Fixação de Nitrogenio da UFPR. A importância atribuída à *H. seropedicae* provém da sua capacidade de fixar nitrogênio atmosférico à amônia sob condição de microaerobiose, de sua capacidade de associação com gramíneas de interesse econômico, como milho, sorgo e cana-de-açúcar, e de seu potencial como bactéria promotora do crescimento vegetal (BALDANI *et al.*, 1993; BALDANI *et al.*, 1995).

H. seropedicae é uma bactéria gram-negativa, geralmente vibrióide, algumas vezes helicoidal e muito móvel, possuindo de um a três flagelos em um ou ambos os pólos (BALDANI *et al.*, 1993). Assim como outras bactérias fixadoras de nitrogênio, é capaz de metabolizar o nitrato, crescendo aerobicamente em meio contendo nitrato como única fonte de nitrogênio (BALDANI *et al.*, 1993). A infecção por *H. seropedicae* ocorre através da ligação da bactéria à superfície da planta, seguida pela proliferação, preferencialmente nas raízes secundárias e ferimentos, penetração e espalhamento da bactéria através dos espaços intercelulares e feixes vasculares das partes aéreas com subsequente colonização e estabelecimento nos vasos do xilema (BALDANI *et al.*, 1995).

O genoma da *Herbaspirillum seropedicae* foi obtido pelo consórcio GENOPAR², sendo o primeiro genoma sequenciado no Estado do Paraná. Possui 5.513.887 pares de base e um total de 4.804 genes anotados. Este genoma mostrou-se muito similar a *Janthinobacterium sp.*, *Ralstonia spp* e *Burkholderia spp* (WEISS, 2010). O GENOPAR também sequenciou parcialmente o genoma de outra bactéria do gênero *Herbaspirillum*, *H. rubrisubalbicans* M1 (BALDANI *et al.*, 1996). Esta bactéria é um diazotrofo endofítico reconhecido por causar a doença da estria mosqueada na variedade B4362 da cana de açúcar e em sorgo (BALDANI e BALDANI, 2005).

² [HTTP://www.genopar.ufpr.br](http://www.genopar.ufpr.br)

2 OBJETIVO

O presente trabalho teve por objetivo a predição *in silico* de RNAs curtos na bactéria diazotrófica *Herbaspirillum seropedicae* SmR1.

Objetivos Específicos:

- 1) Desenvolver uma ferramenta integradora dos dados gerados por softwares de bioinformática para identificar RNAs curtos (*small RNAs*) em regiões intergênicas de genomas de procariotos.
- 2) Aplicar a ferramenta para a busca de RNAs curtos no genoma de *Herbaspirillum seropedicae* SmR1.
- 3) Realizar uma análise comparativa com o genoma parcial da bactéria *Herbaspirillum rubrisubalbicans* M1.

3 MATERIAIS E MÉTODOS

3.1 Genomas

As sequências dos genomas das bactérias *Herbaspirillum seropedicae* SmR1 e *Herbaspirillum rubrisubalbicans* M1 foram obtidas pelo consórcio Genopar. Genomas das espécies dos gêneros *Burkholderia*, *Ralstonia*, *Janthinobacterium*, *Pseudomonas* e *Hermiimonas* foram obtidos no GenBank³.

3.2 Sistema Operacional

O sistema operacional utilizado no desenvolvimento das ferramentas foi o Ubuntu⁴ na sua versão 10.04. É um sistema baseado em Debian GNU/LINUX e distribuído como software de código fonte aberto. A licença usada pelo Ubuntu é a GNU, na qual declara explicitamente que usuários são livres para executar, copiar, distribuir, estudar, alterar, desenvolver e melhorar o sistema (FERREIRA, 2009).

3.3 Linguagens de Programação

Para a geração, formatação e análise dos dados foram desenvolvidos *scripts* auxiliares e uma ferramenta integradora, escritos nas linguagens Bash e Java, respectivamente.

3.3.1 Java

No desenvolvimento do software integrador foi utilizada a linguagem de programação Java, por ser amplamente difundida e com suporte multiplataforma. Diferentemente das linguagens convencionais, o código escrito em Java é compilado em um "bytecode" que é executado por uma máquina virtual (DEITEL e DEITEL, 2005).

Esta linguagem apresenta algumas características de destaque:

³ [HTTP://www.ncbi.nlm.nih.gov/genbank](http://www.ncbi.nlm.nih.gov/genbank)

⁴ [HTTP://www.ubuntu.org](http://www.ubuntu.org)

- Orientação a objeto;
- Portabilidade (Independência de plataforma);
- Recursos de Rede - Possui extensa biblioteca de rotinas que facilitam a cooperação com protocolos TCP/IP, como HTTP e FTP;
- Segurança - Pode executar aplicativos via rede, porém, com restrições de execução.

3.3.2 Bash

Bash⁵ é um interpretador de comandos de código aberto escrito para o projeto GNU⁶ e, pode ser executado na maioria dos sistemas operacionais *Unix-like*. Com ele é possível realizar sequências de comandos utilizando ferramentas externas, automatizando os processos. Também é muito útil na manipulação de dados genômicos, especialmente quando utilizado em conjunto com as ferramentas auxiliares *awk* e *sed*, possibilitando ao programador controle total da formatação dos dados (SAADE, 2001).

3.4 Ambiente de Programação

Para o desenvolvimento do software integrador foi utilizado o NetBeans, um ambiente de desenvolvimento integrado (IDE) gratuito e de código aberto para a linguagem Java. Pode ser executado em múltiplas plataformas, como Windows, Linux⁷ e Solaris⁸ (GONÇALVES, 2006).

Esta linguagem apresenta algumas características de destaque:

- Editor de código fonte integrado com o pré-visualizador;
- Visualizador de classes integrado ao de interfaces, que gera automaticamente o código dos componentes de forma bem organizada;
- Depurador de erros apurado de aplicações e componentes;
- Suporte a diversos bancos de dados.

⁵ [HTTP://www.gnu.org/software/bash/](http://www.gnu.org/software/bash/)

⁶ [HTTP://www.gnu.org](http://www.gnu.org)

⁷ [HTTP://www.linux.org](http://www.linux.org)

⁸ [HTTP://www.oracle.com/solaris](http://www.oracle.com/solaris)

3.5 Softwares de Bioinformática

3.5.1 BLAST

BLAST⁹ (Basic Local Alignment Search Tool) é um algoritmo de comparação e alinhamento de sequências primárias com base na similaridade. O programa compara sequências de nucleotídeos ou aminoácidos com uma base de dados e calcula a significância estatística dos alinhamentos (ALTSCHUL *et al.*, 1990). Neste trabalho o BLAST foi utilizado para comparar sequências candidatas a sRNA de *Herbaspirillum seropedicae* SmR1 com regiões intergênicas de outras espécies filogeneticamente próximas.

3.5.2 Rfam

Rfam¹⁰ é um banco de dados de alinhamento de sequências de RNAs não codificadores (ncRNAs) mantido pelo Sanger Institut¹¹, que faz uso de uma grande quantidade de dados de RNAs publicados. Também possui uma ferramenta de pesquisa de sequências que retorna as coordenadas iniciais e finais caso a sequência submetida tenha similaridade com a sua biblioteca de RNAs.

3.5.3 Preditores de Promotores

3.5.3.1 NNPP (Neural Network Promoter Prediction)

NNPP¹² é um algoritmo que prediz sequências de início de transcrição de genes bacterianos regulados pelo fator $\sigma 70$. A base de dados é uma rede neural baseada em promotores experimentalmente comprovados de *Escherichia coli*. A rede apresenta duas camadas, uma para reconhecer o hexâmero -10 (TATA-BOX) e outra para reconhecer o sítio de início de transcrição (REESE *et al.*, 2001).

⁹ [HTTP://www.ncbi.nlm.nih.gov/BLAST/](http://www.ncbi.nlm.nih.gov/BLAST/)

¹⁰ [HTTP://rfam.sanger.ac.uk/](http://rfam.sanger.ac.uk/)

¹¹ [HTTP://www.sanger.ac.uk](http://www.sanger.ac.uk)

¹² [HTTP://www.fruitfly.org](http://www.fruitfly.org)

O software exige sequências com comprimento mínimo de 51 pares de base. Cada predição inclui o escore da predição e uma sequência associada de 46 pares de base, que corresponde às posições -41 à +5, no qual a posição +1 é o TSS (Sítio de início de transcrição). Todos os escores variam de 0 a 1. Pelo fato dos elementos dos promotores poderem aparecer em posições relativas diferentes um dos outros, a precisão posicional do algoritmo NNPP é de +-3 pares de base (BURDEN *et al.*, 2005). A precisão do algoritmo é de aproximadamente 80% com um corte de 0,8. A fita complementar também é analisada, não sendo necessário efetuar a função complementar reversa e realizar novo processamento (KALATE *et al.*, 2003).

Alguns parâmetros do software: *Type of organism* (Figura 3a) é o tipo de organismo da sequência, *Include reverse strand* (Figura 3b) indica a possibilidade de se incluir a fita reversa na busca, *Minimum promoter score* (Figura 3c) o escore mínimo de cada predição e por fim a região no qual a sequência deve ser informada (Figura 3d).

O algoritmo gera um arquivo de saída no formato mostrado na Figura 4. Para cada predição tem-se a posição inicial e final da sequência no arquivo de intergênicas (Figura 4a) e b), o escore (Figura 4c) e a sequência promotora (Figura 4d).

Type of organism: prokaryote eukaryote a)

Include reverse strand? yes no b)

Minimum promoter score (between 0 and 1): 0.8 c)

Cut and paste your sequence(s) here: Use single-letter nucleotides: (A, C, G, T). You can include multiple sequences if each has a FASTA title line starting with >

```
AAGTTACGGTACCAATTTGCCGAGTTCCTTCTCCCGAGTTCCTCTCAAGCG
CCTTAGAATACTCATCTCGCCACCTGTGTGGTTTGCAGTACGGTCTCG
TTAGACTGAAGCTTAGAGGCTTTTCTTGAACCACTTCCGATTGCTTCGC
GAACAAGTTCGCTCGTCTCACACCCTTGAATTACGCTGCCGGATTTGCCT
AACAGCCTTCTTCGATGCAAAAACCGGGACTTCCAACACCCGGACAACCT
TCCGCGATCCGTCACCCCATCGCATCTAACGACGGTGCAGGAATATTAAC
CTGCTTCCCATCAGCTACGCATCTCTGCCTCGCCTTAGGGGCGGACTCAC
CCTGCTCCGATGAACGTTGAACAGGAAACCTTGGGCTTACGGCGTGCAGG
CTTTTCACCCGCATTATCGTACTCATGTGAGCATTGCACTTCTGATAC
CTCCAGCATCCTTTACAAGACACCTTACAGGCTTACAGAACGCTCTCCT
ACCATATCCTTACGGATATCCGCAGCTTCGGTGACTGGCTTAGCCCCGTT
ACATCTTCCGCGCAGGACGACTCGATCAGTGAGCTATTACGCTTTCTTTA
AAGGATGGCTGCTTCTAAGCCAACCTCCTGACTGTTTTAGCCTTCCCCT
TCGTTTTTCCACTTAGCCAATCTTTGGGACCTTAGCTGGCGGTCTGGGTTG
TTTCCCTCTTGACGTCGGACGTTAGCACCCGGCGTCTGTCTCCAAGCTC
GCACTC
```

d)

Figura 3 – Captura de tela do Neural Network Promoter Prediction.

a) Indicar o tipo de organismo, b) Opção de incluir a fita reversa, c) Escore mínimo da predição e d) Informar a sequência de DNA.

Promoter predictions for IG1 :

a)	b)	c)	d)
Start	End	Score	Promoter Sequence
253	298	0.94	GACGTTTTAAACCAGCTCACGTACCACTTTAAATGGCGAACAGCCATAC
671	716	0.94	TGGTATTTCAAGGTCGGCTCCACAGAACTAGCGTTCCTGCTTCAAAGCC
1049	1094	0.96	TGTTTTTATTAACAGTCGCAGCCACCATTTTTATTGCAACCCTTTTGTCC
1921	1966	0.87	GGTTTGGTAAGTCGCGATGACCCCC TAGCCATAACAGTGCTC TACCCCCG
1969	2014	0.89	CGATGGTGATAC TTGAGGCAC TACC TAAATAGTTTTTCGGAGAGAACCAGC
2285	2330	0.95	CTACTGTTTGTATGCACACGGTTTCAGGATCTATTTCACTCCCCTTCCGG
2866	2911	0.95	TAGCGTTTGCCGATTCCAAAGTAAGTCTTCTAATTGCTAAGATCACTTC
2921	2966	0.99	ACTTTGATTGATACAATCACACCCATTTTTACTTTCCGAAGAACC GAAGT

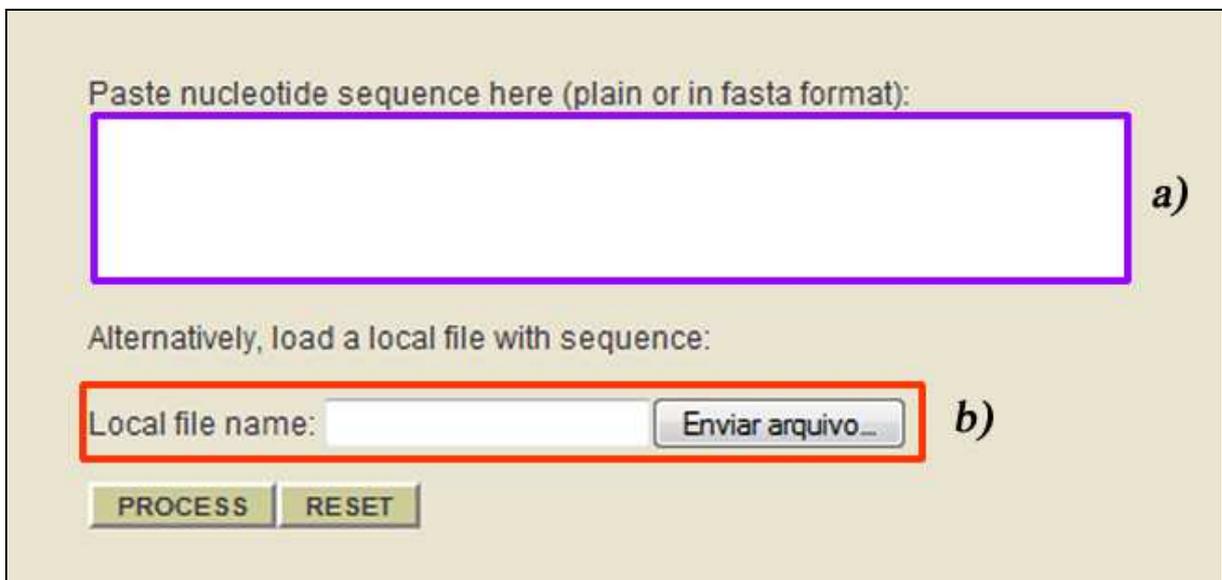
Figura 4 - Captura de tela do arquivo de saída do NNPP.

a) Posição inicial da predição, b) Posição final da predição, c) Escore e d) Sequência promotora.

3.5.3.2 BProm

O software BProm¹³ utiliza a função discriminante linear (LDF) que combina características funcionais e composição oligonucleotídica das possíveis regiões promotoras.

Como entrada de dados é exigida apenas uma sequência no padrão Fasta. É possível colar os dados no campo indicado (Figura 5a) ou carregar o arquivo com as regiões intergênicas (Figura 5b). Cada predição do arquivo de saída do BProm é composta pelo escore da LDF (Figura 6a) e pela posição, escore e sequência dos hexâmeros -10 e -35 (Figura 6b e c).



The screenshot shows the BProm software interface. At the top, it says "Paste nucleotide sequence here (plain or in fasta format):". Below this is a large empty text box labeled "a)". Below that, it says "Alternatively, load a local file with sequence:". Below this is a "Local file name:" label followed by an empty text input field and a button labeled "Enviar arquivo...". This section is labeled "b)". At the bottom, there are two buttons: "PROCESS" and "RESET".

Figura 5 – Captura de tela do software BProm.
a) Colar sequências; b) Carregar arquivo de sequências.

¹³ [HTTP://www.softberry.com](http://www.softberry.com)

```

> H.seropedicae
Length of sequence-      35842
Threshold for promoters - 0.20
Number of predicted promoters -      74
Promoter Pos: 34411 LDF- 8.79 a)
-10 box at pos. 34396 TTATATCAT Score 57 b)
-35 box at pos. 34379 TTA ACT Score 38 c)
Promoter Pos: 11824 LDF- 6.06
-10 box at pos. 11812 TGCTATAAT Score 88
-35 box at pos. 11785 TTGTAG Score 41
Promoter Pos: 16820 LDF- 6.04
-10 box at pos. 16805 GGTTATAAT Score 90
-35 box at pos. 16786 TTCTTT Score 20
Promoter Pos: 10873 LDF- 5.82
-10 box at pos. 10858 GCGTATAAT Score 77
-35 box at pos. 10835 TTGACT Score 61
Promoter Pos: 35822 LDF- 5.49
-10 box at pos. 35807 TGTTATTAA Score 41
-35 box at pos. 35789 TTTACA Score 47
Promoter Pos: 10128 LDF- 5.32
-10 box at pos. 10113 GTCTAGACT Score 51
-35 box at pos. 10093 TTGCCT Score 56
Promoter Pos: 31737 LDF- 4.91
-10 box at pos. 31722 TAATACACT Score 42
-35 box at pos. 31702 TTGCTT Score 48

```

Figura 6 - Captura de tela do arquivo de saída do software BPROM.
a) Escore da LDF, b) Posição, sequência e escore do hexâmero -10 e c) Posição, sequência e escore do hexâmero -35.

3.5.4 Preditores de Terminadores

3.5.4.1 TranstermHP

TranstermHP¹⁴ é um software para a identificação de terminadores intrínsecos, que faz buscas por grampos curtos seguidos por uma cauda de timinas (que são transcritas em uracilas) (KINGSFORD *et al.*, 2007).

Este software apresenta as seguintes características:

¹⁴ [HTTP://transterm.cbcb.umd.edu](http://transterm.cbcb.umd.edu)

- Esquema apurado de escores, no qual para cada característica da terminação é atribuído um escore.
- Leitura de anotações de genes de arquivos simples. Para utilizar o software em regiões intergênicas, é necessário criar um arquivo com dois genes, um no início da IgR e o outro no final. Desse modo, a ferramenta irá procurar por terminadores entre esses genes.
- Suporte a arquivos de coordenadas
- Tratamento de sobreposição em genes e terminadores.
- Tratamento geral de *gaps* em grampos.

O comando de execução do TranstermHP é:

Transterm coords IGRs.fas > saída

O primeiro parâmetro “coords” é o arquivo de coordenadas dos genes e o segundo é a sequência no formato Fasta. Por fim, com o uso do sinal “>” o processamento é redirecionado ao arquivo “saída”. Cada registro de terminador possui as seguintes informações: número da terminação (Figura 7a), posição inicial e final do terminador (Figura 7b), orientação da sequência e tipo de região no qual o terminador está presente (Figura 7c), escore geral do terminador (Figura 7d), escore do grampo (Figura 7e), escore da cauda (Figura 7f) e sequência predita (Figura 7g).

```

SEQUENCE ID (length 820555)

a) b) c) d) e) f)
TERM 1 5636 - 5605 - f 95 -7.2 -5.44721 | gap 1,
opp_overlap 5632
AGCACCGCAAAAAAA CCGCGACAG-GCCCG CCC CAGGCGCTCTGGCGG g)
CGCTTCGACGACTT
TERM 2 6555 - 6515 - f 100 -18.4 -3.12351 | gap 1,
opp_overlap 6526, overlap 6526
GTGGCTCATAGGGAA GCGCCCGCAAAGCAGGCC CGCC GGGCCTGCTTTGCACGC-C
TGGAAAGTCTGGCTC
TERM 3 6526 - 6545 + F 88 -15 -3.54424 | opp_overlap
6515 6526
GGAAGCGCCCGCAAA GCAGGCC CGCC GGGCCTGC
TTTGCACGCCTGGAA
TERM 4 6545 - 6526 - f 88 -15 -3.54459 | opp_overlap
6526, overlap 6515
GGAAGCGCCCGCAAA GCAGGCC CGCC GGGCCTGC
TTTGCACGCCTGGAA
TERM 5 7125 - 7098 - f 82 -8.4 -4.41348 | opp_overlap
7105, overlap 7102
TCCTGTCGAAAGAAA GGGACCATCGGA GCGT TCCGATGGTGTC
ACGAGCAATGCCCG
TERM 6 7285 - 7266 - f 100 -13.6 -4.91873 | opp_overlap
7270
GACATCAATGAAAAA GGGACGGA AAAC TCCGTCCC
TTCGCATGCAGGCTG
TERM 7 7461 - 7487 + F 84 -12.9 -3.69395 | opp_overlap
7461
CCGGCCCATGAAAAA GCGGCCCTGGC AGAAC GCTGGGCCGC
TTCTTGCGTCATCCG
TERM 8 7487 - 7461 - f 95 -11.4 -4.72272 | opp_overlap
7461
CCGGCCCATGAAAAA GCGGCCCTGGCA GAA CGCTGGGCCGC
TTCTTGCGTCATCCG

```

Figura 7 - Captura de tela do arquivo de saída do software TranstermHP.
a) Número da terminação, b) Posição inicial e final do terminador, c) Orientação da sequência e tipo de região no qual o terminador está presente, d) Escore geral do terminador, e) Escore do grampo, f) Escore da cauda e g) Sequência predita.

3.5.4.2 RNAMotif

O segundo software utilizado para a identificação de terminadores intrínsecos foi o RNAMotif¹⁵. Seu algoritmo confere ao usuário a liberdade de buscar qualquer estrutura secundária, seja ela simples ou complexa, por meio de motivos, ou seja, variações características da estrutura secundária assumida pelo RNA (MACKÉ *et al.*, 2001). Estes motivos estruturais incluem hélices e nucleotídeos não-pareados em grampos e *loops* internos. Também é possível identificar elementos estruturais de

¹⁵ [HTTP://casegroup.rutgers.edu](http://casegroup.rutgers.edu)

RNA que podem não ser reconhecidos com o uso de motivos, como sítios de ligação de proteína e conformações *backbone* (esqueleto de ribose e grupo fosfato). Os padrões estruturais são definidos em um arquivo descritor com uma linguagem de *script* que distingue, em seu nível mais baixo, posições pareadas e não-pareadas. Os descritores podem ser parametrizados em comprimento e pareamento de bases, provendo um alto grau de controle sobre as estruturas que são identificadas (VEKSLER-LUBLINSKY *et al.*, 2007).

O algoritmo permite todos os 16 tipos de pareamento de bases possíveis, incluindo Watson-Crick (GC e AU), oscilante ou *wobble* (GU), entre outros pareamentos como AC e UU que são definidos como parte da hélice. Níveis similares de controle também podem ser providos para definir *mismatches* de sequências (FOGEL *et al.*, 2002).

O descritor é composto por quatro sessões: parâmetros, descritor, sítios e escore. A sessão “descritor” define o critério requerido para gerar a comparação. A sessão “sítios” permite especificar relações entre os elementos do descritor, enquanto a sessão “escore” combina os resultados primários (YACHIE *et al.*, 2006). A sessão “escore” consiste de um conjunto de testes que são aplicados em cada candidato predito pela busca. Ela serve a dois propósitos relacionados: 1) prover um modo para adicionar restrições ao motivo que são difíceis de implementar em um padrão puro de linguagem; e 2) avaliar e ranquear os candidatos (LESNIK *et al.*, 2002).

Muitos descritores geram um grande número de candidatos, e isso é especialmente verdadeiro para aqueles que contêm somente restrições de hélice sem restrições de sequência. Em muitos casos, não há sistema de escore no qual estes candidatos possam ser ranqueados. Para auxiliar nesses casos existe a função de estabilidade termodinâmica que avalia a estabilidade termodinâmica (ΔG) do candidato, ou qualquer parte dele, e permite a sua classificação com base no cálculo de energia livre de Gibbs (MACKE *et al.*, 2001; VEKSLER-LUBLINSKY *et al.*, 2007).

O software é executado em linha de comando e necessita que dois parâmetros sejam informados: um descritor, que especifica a estrutura secundária pretendida, e um arquivo de dados no formato Fasta.

O comando de execução é:

rnamotif -descr test.descr test.fastn > resultado

O primeiro parâmetro “-descr” é o arquivo descritor do motivo e o segundo é a sequência no formato Fasta. Por fim, com o uso do sinal “>” o processamento é redirecionado ao arquivo “resultado”.

O formato do arquivo gerado é mostrado na Figura 8.

```
#RM scored
#RM descr ss h5(tag='1') ss h3(tag='1') ss ss ss ss
#RM dfile descriptor.desc
>IGR1 0-3098, + -
IGR1      | -5.70 | 0.00 | -2.80 | -1.70 | -5.70 | -
1.76 | 4.0 | 4.0 | 2.0 | 0.4 | 0 2314 35 tctatttcact
cccc ttcc gggg . ttctt ttc gcct
>IGR1 0-3098, + -
IGR1      | -7.30 | 0.00 | -0.80 | -3.60 | -4.50 | -
4.48 | 5.0 | 10.0 | 2.0 | -3.6 | 1 1396 43
cggaaggttgt cgggg tgttggaagt cccgg . ttttt gca tcga
>IGR8 12988-13101, - -
IGR8      | -9.00 | 0.00 | -4.00 | -1.10 | -1.50 | -
4.38 | 7.0 | 4.0 | 2.0 | 0.4 | 1 94 41 gcttgatcgat
ggttgct ttca ggcaacc . ttgat ctt gttt
>IGR17 24962-25447, - -
IGR17     | -13.50 | 0.00 | -0.80 | -1.80 | -3.80 | |
-11.61 | 8.0 | 4.0 | 3.0 | 0.6 | 1 404 43
ctgcatgcgaa gggacgga gttt tccgtccc . ttttt cat tgat
>IGR17 24962-25447, - -
IGR17     | -7.90 | 0.00 | -2.80 | -2.40 | -4.50 | -
3.68 | 6.0 | 6.0 | 1.0 | 0.2 | 1 134 41 ggcatttcttt
ctgcag ttcttt ctgcag . ttctt tct gcat
>IGR19 28236-28670, - -
IGR19     | -7.90 | 0.00 | -5.10 | -0.40 | -2.60 | -
2.47 | 6.0 | 4.0 | 2.0 | 0.4 | 0 174 39 ttgactactgc
ggttgc gaag gcagcc . tgtct ttt ttct
>IGR20 32181-32387, + +
IGR20     | -16.20 | -0.90 | -2.10 | -1.10 | -3.90 | |
-12.46 | 8.0 | 10.0 | 3.0 | -3.4 | 0 15 50
actccgatcag gccgggag gcagaaatgt cgcccggc c ttttc att tctc
```

Figura 8 – Captura de tela do arquivo de saída do software RNAMotif.

Exemplo de sequência predita pelo RNAMotif:

AGTGGGTAAGCCCAACGTACGGGGGGTTGAGGGAACCCGCCCCCTTTTTTG
AA

A alça ou *loop* é representada na cor azul e a haste ou *stem* em vermelho. A cauda de timinas (cauda poli-T) é representada em verde no final da sequência.

O descritor usado nas predições foi gentilmente disponibilizado pelo Dr. Jonathan Livny¹⁶ (Departamento de Biologia Molecular e Microbiologia - Faculdade de Medicina da Universidade Tufts – Boston – USA).

3.6 Metodologia para Seleção de Candidatos a sRNAs no Genoma de *H. seropedicae* SmR1

Para a seleção dos candidatos foi utilizado o fluxo de trabalho mostrado na Figura 9. Cada uma das etapas é descrita a seguir.

3.6.1 Extração das Regiões Intergênicas

Regiões intergênicas são sequências de DNA que não codificam peptídeos e proteínas, ou seja, que não contém ORFs (do inglês *Open Reading Frame* = Fase de Leitura Aberta) (Figura 11). Nestas regiões se localizam os genes codificadores de RNAs curtos. Para a extração das regiões intergênicas foi desenvolvida uma ferramenta em Java (Figura 10) que faz uso do arquivo de ORFs anotadas (ffn), do genoma (fna) e das sequências de tRNAs e rRNAs anotadas (Figura 10b). Durante a extração, um *log* é exibido (Figura 10c). Após o processamento dos dados, as IgR são carregadas em uma tabela com suas respectivas coordenadas (Figura 10a).

As ORFs podem estar localizadas tanto na fita principal quanto na fita secundária do genoma, muitas vezes uma sobrepondo a outra. Quando ocorre sobreposição, é necessário remover a sequência iniciada na posição inicial da primeira ORF e terminada na posição final da segunda ORF. Para as ORFs da fita secundária é necessário utilizar a função complementar reversa.

¹⁶ jonathan.livny@tufts.edu

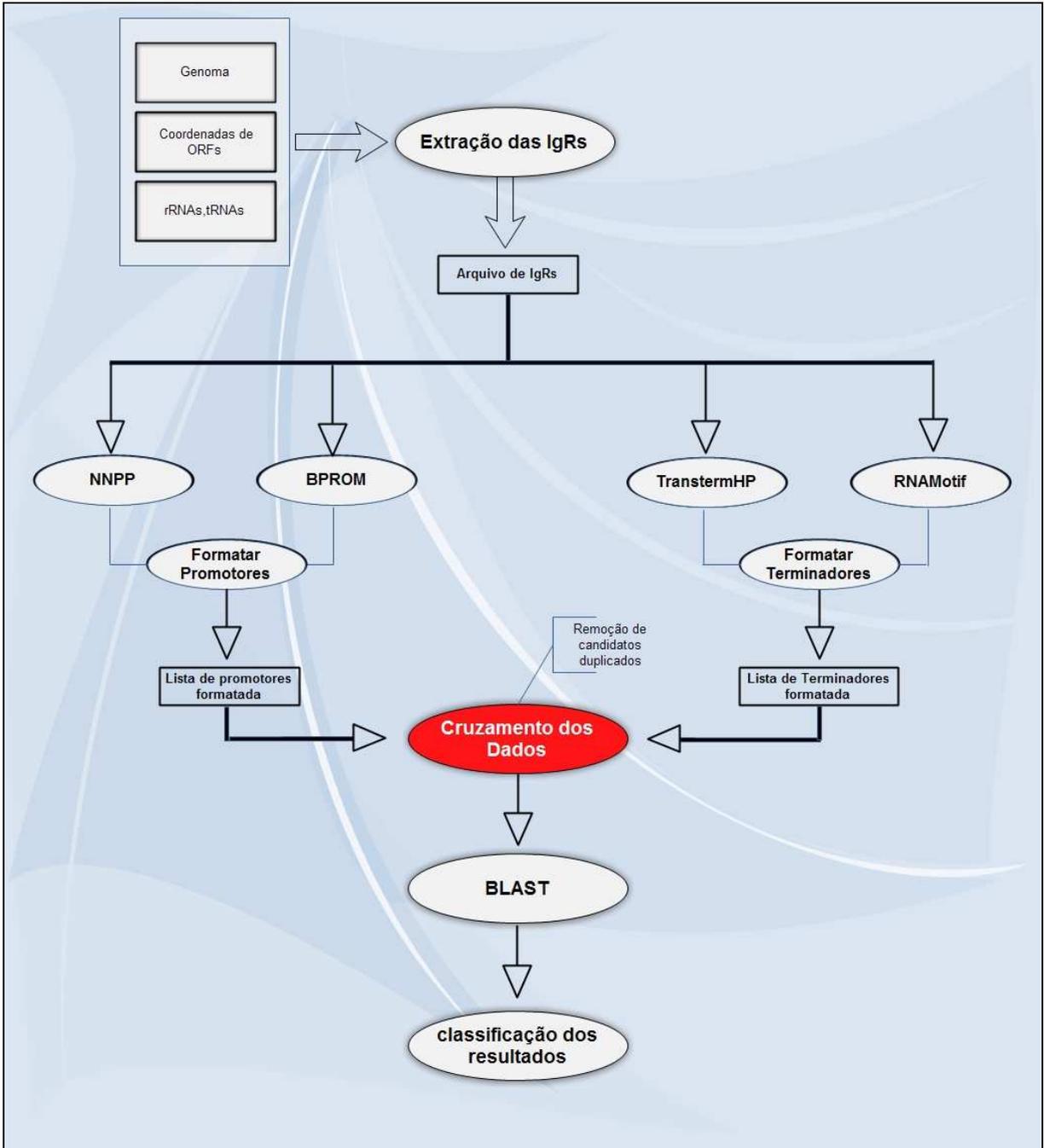


Figura 9 - Fluxo de trabalho para a identificação de sRNAs em genomas de procaríotos.

The screenshot shows the FurBox software interface. At the top, a genomic track displays coordinates from 26935 to 33741. Below the track are three main panels:

- Panel a) Remover ORFs e Gerar IGRs:** Contains buttons for 'Abrir Orfs (.ffn)', 'Abrir Fastas(Genoma)', 'Gerar IGRs', and 'Salvar IGRs Modo Fas'. It also shows 'Genoma: herbas_12ctps' and 'Tamanho(bp): 5728509 bp'.
- Panel b) IGRs:** Includes a 'File: <none>' field, 'Tamanho Minimo' (set to 150), 'Refresh', 'Abrir IGRs', 'Núm. da IGR' (set to 1), and 'Ver IGR'.
- Panel c) Log Reg.:** A list of log entries showing 'Regex Match' results for various ORF types (Principal, Inteira) with their respective genomic coordinates.

Figura 10 – Captura de tela da ferramenta para a extração das regiões intergênicas. a) IGRs carregadas com seus respectivos índices no genoma, b) Extração das IGRs e c) Log do processo de extração.

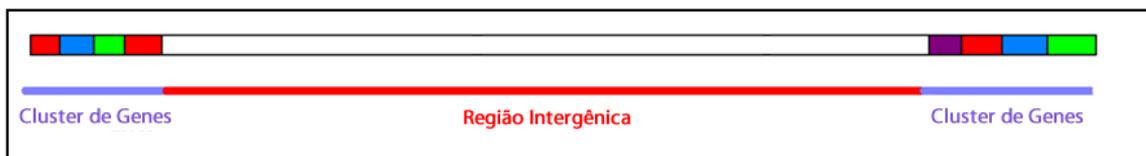


Figura 11 – Representação de uma região intergênica.
Fonte: Autor.

3.6.2 Predição de Promotores e Terminadores

Após a extração das regiões intergênicas, as sequências foram submetidas aos preditores de promotor e terminador. Para os preditores de promotor BProm e NNPP, foram utilizados cortes de 0,70 e 70, respectivamente. Ambos utilizam um $\sigma 70$ consenso de *E. coli* com os hexâmeros -35 e -10 representados por TTGACA e TATAAT, separados por uma região de 15 a 21 bp. Nos softwares terminadores RNAMotif e TranstermHP foi definido para a terminação intrínseca uma cauda mínima de 3 Timinas.

3.6.3 Formatação dos Dados

Com ambos os promotores e terminadores preditos foi necessário o desenvolvimento de um *script* para formatar todas as saídas dos dados em um arquivo padrão, uma vez que cada um dos quatro softwares (BProm, NNPP, RNAMotif e TranstermHP) gera um arquivo de saída individual (Figura 4, Figura 6, Figura 7, Figura 8). Após a formatação de dados, foi definido para cada registro de promotor o número da região intergênica (Figura 12 e Figura 13a), a posição inicial da predição (Figura 12 e Figura 13b), o número de bases entre os hexâmeros (Figura 12 e Figura 13c), a orientação da predição (Figura 12 e Figura 13d) e o escore da predição (Figura 12 e Figura 13e). Para cada registro de terminador foi definido o número da região intergênica (Figura 14a), a posição final do terminador (Figura 14b), a região da predição (Figura 14c) e a orientação da sequência (Figura 14d).

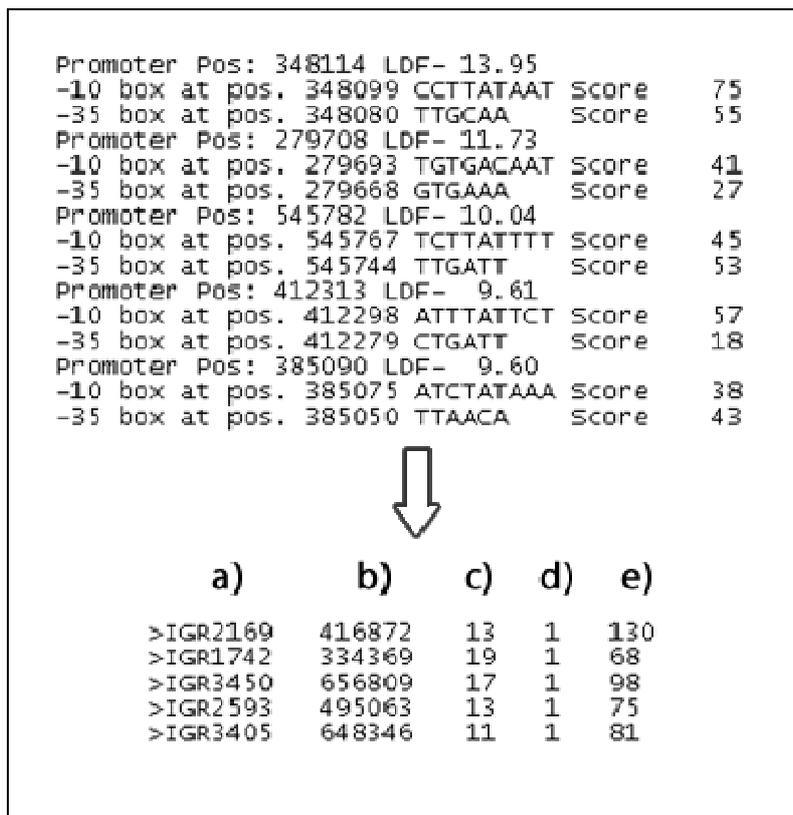


Figura 12 – Formatação dos dados de saída do software BProm.

a) Região Intergênica no qual o possível promotor se encontra, b) Posição inicial da predição no arquivo de IGRs, c) Tamanho da região intermediária dos hexâmeros, d) Orientação da sequência e e) Escore do promotor.



Figura 13 - Formação dos dados de saída do software NNPP.

a) Região intergênica no qual o possível promotor se encontra, b) Posição inicial da predição no arquivo de IGRs, c) Tamanho da região intermediária dos hexâmeros, d) Orientação da sequência e e) Escore do promotor.

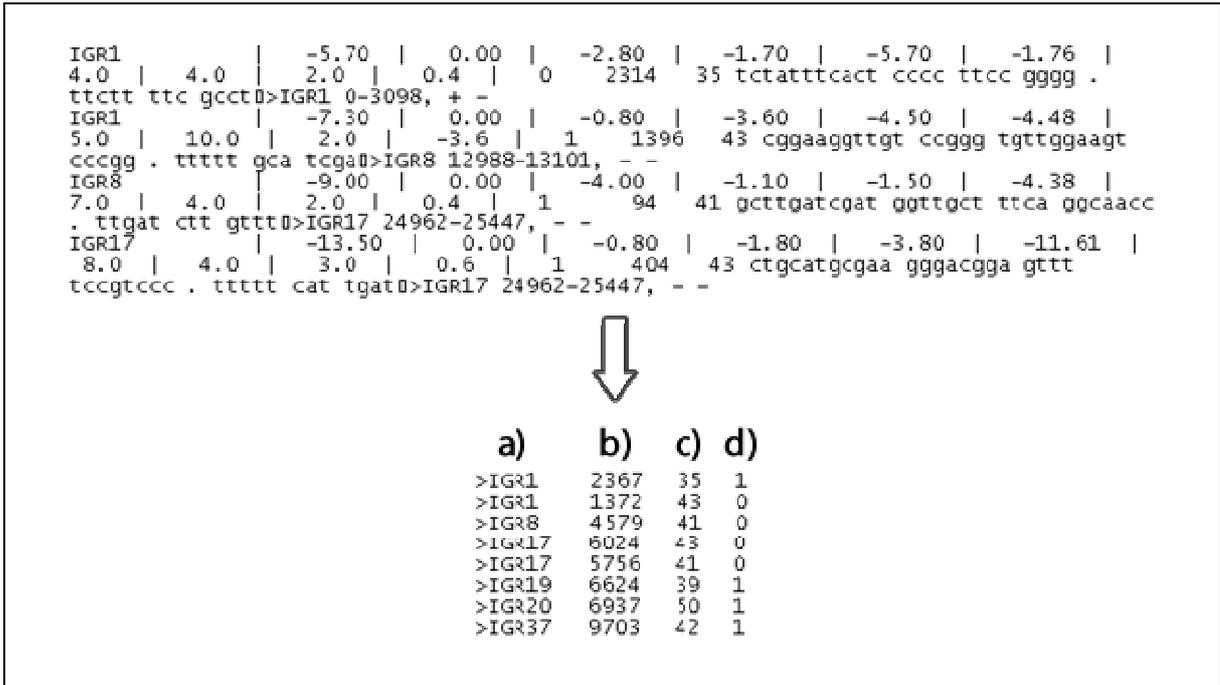


Figura 14 – Formatação dos dados de saída do software RNAMotif.

a) Região intergênica no qual o possível terminador se encontra, b) Posição final da predição no arquivo de IGRs, c) Tamanho da predição e d) Orientação da sequência.

3.6.4 Cruzamento dos Dados

Com os dados gerados pelos preditores e devidamente reformatados, foi efetuado o cruzamento dos dados (Figura 15). Nessa abordagem, a região intergênica de cada registro da lista de promotores foi comparada com todas as regiões intergênicas dos registros de terminadores, sempre verificando a orientação da sequência, promotor na fita principal com terminador na fita principal e promotor na fita reversa com terminador na fita reversa, além de um tamanho entre 50 e 400 *bp*. O tamanho do RNA candidato foi obtido subtraindo as sequências correspondentes à posição final do terminador, à posição inicial do promotor e ainda 30 pares de bases. Este último valor corresponde ao tamanho médio da região promotora, considerando os dois hexâmeros e a região intermediária (Figura 15). Os dados foram agrupados em tabelas, e cada registro foi documentado em nove colunas: número da região intergênica (Figura 15a), tamanho da região intergênica (Figura 15b), posição inicial do promotor (Figura 15c), posição final do terminador (Figura 15d), orientação da sequência (Figura 15e), tamanho da sequência (Figura

15f), orientação dos genes adjacentes (Figura 15g), escore do promotor (Figura 15h) e conteúdo GC da sequência (Figura 15i).

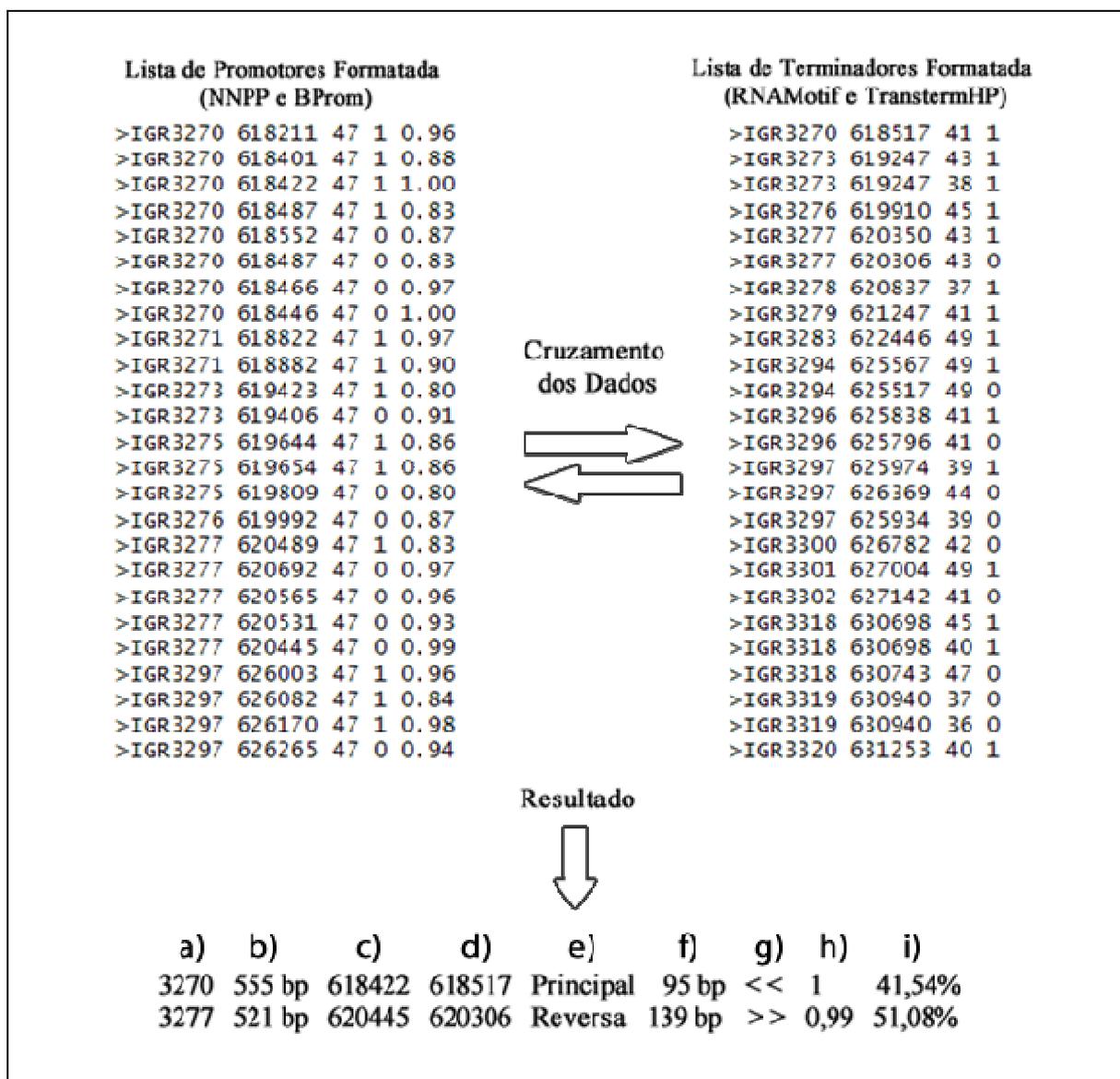


Figura 15 – Cruzamento dos dados de promotores e terminadores.

a) Região intergênica do candidato, b) Tamanho da IgR, c) Posição inicial no arquivo de IgRs, d) Posição final no arquivo de IgRs, e) Orientação da sequência, f) Tamanho da sequência, g) Orientação dos genes adjacentes, h) Escore do promotor e i) Composição GC do candidato.

3.6.5 Busca por Similaridade de Sequências entre Gêneros ou Espécies Bacterianas

A similaridade de sequências entre gêneros ou espécies filogeneticamente próximas é uma importante característica dos sRNAs (ARGAMAN *et al.*, 2001; RIVAS *et al.*, 2001; WASSARMAN *et al.*, 2001). O software mais amplamente difundido para esta tarefa é o BLAST, que exige alguns parâmetros básicos como: banco de dados alvo, *query* de pesquisa, tipo de informação que se deseja buscar (nucleotídeos ou aminoácidos) e o arquivo de saída. Na criação do banco de dados foi utilizado o genoma parcial de *Herbaspirillum rubrisubalbicans* M1 e genomas completos de espécies dos gêneros *Burkholderia*, *Ralstonia*, *Janthinobacterium*, *Pseudomonas* e *Herminiimona*, obtidos no banco de dados GenBank. A *query* é a sequência obtida na etapa do cruzamento dos dados e como este procedimento tende a retornar um elevado número de sequências, foi desenvolvido um *script* para realizar a interação. Para evitar *hits* com baixa similaridade, foi definida identidade mínima de 70%.

3.6.6 Confirmação dos Candidatos

Parâmetros como sequência com promotor, terminador, tamanho entre 50 e 400 *bp* e similaridade com sRNAs já identificados/confirmados são fatores importantes para a predição de um sRNA, porém, o refinamento da exploração *in silico* ou a confirmação *in vitro/vivo* é essencial para a obtenção da real identidade do candidato. Neste trabalho o refinamento *in silico* foi realizado utilizando o banco de dados Rfam. Não apenas as sequências com similaridade, mas todas as preditas com terminadores e promotores foram analisadas.

4 RESULTADOS E DISCUSSÃO

O conhecimento a respeito da existência de RNAs curtos ou *small* RNAs (sRNAs) em diferentes grupos de organismos e sobre as funções que estas moléculas desempenham têm aumentado muito nos últimos anos, entre 2000 e 2010. Abordagens experimentais visando à identificação e caracterização de sRNAs bacterianos, em regiões intergênicas, têm sido aplicadas e, simultaneamente, uma gama de abordagens e ferramentas computacionais têm sido desenvolvidas com objetivo de possibilitar a identificação comparativa de sRNAs (GOTTESMAN, 2004; HUTTENHOFER e VOGEL, 2006; ALTUVIA, 2007; PICHON E FELDEN, 2008; SONNENLEITENER *et al.*, 2008; SRIDHAR *et al.*, 2010). Estas moléculas atuam normalmente através do pareamento de bases com mRNAs específicos, afetando sua estabilidade e/ou o processo de tradução (GOTTESMAN, 2004; LIU & CAMILLI, 2010). Além disso, podem ligar-se a proteínas regulatórias e modificar sua atividade (BABITZKE & ROMEO, 2007).

O genoma da bactéria diazotrófica *H. seropedicae* SmR1 que foi utilizado neste trabalho para a predição de sRNAs foi sequenciado pelo consórcio GENOPAR e encontra-se disponível no GenBank sob o número de acesso CP002039. Cerca de 89% do genoma de *H. seropedicae* SmR1 corresponde a ORFs, tRNAs e rRNAs, e 11% corresponde a regiões intergênicas. Para a seleção de sequências candidatas a RNAs curtos foi necessário o desenvolvimento de uma ferramenta integradora que segue o fluxo de trabalho apresentando na Figura 9 e permite automatizar os passos necessários à identificação de prováveis sRNAs em procariotos. Esta ferramenta foi denominada sRNATool – Ferramenta Integradora para Busca de sRNAs (Figura 16), e se caracteriza por formatar e cruzar os dados gerados por dois softwares preditores de promotores, BProm e NNPP, e dois softwares preditores de terminadores, RNAMotif e TranstermHP, além de realizar a busca por similaridade entre as espécies filogeneticamente próximas. As características de cada um dos softwares utilizados estão descritas na sessão Materiais e Métodos. Os passos seguidos para seleção das sequências candidatas a sRNA e o resumo dos resultados obtidos pela aplicação desta ferramenta no genoma de *H. seropedicae* SmR1 estão apresentados no item 3.6 e no fluxograma da Figura 17.

Para a aplicação do sRNATool em diferentes genomas de procariotos é necessário que o usuário defina os seguintes parâmetros: tamanho mínimo e máximo dos candidatos, identidade mínima no BLAST, tamanho mínimo das IgRs e o carregamento dos dados gerados pelos softwares preditores (Figura 16c). Também é necessário selecionar as espécies bacterianas relacionadas (Figura 16d). Neste trabalho os parâmetros iniciais foram: tamanho mínimo de 50 e máximo de 400 *bp* e identidade mínima de 70% no BLAST. O banco de dados comparativo foi composto por genomas de bactérias dos gêneros *Burkholderia*, *Ralstonia*, *Janthinobacterium*, *Pseudomonas* e *Herminiimonas*, além do genoma de *H. rubrisubalbicans* M1, espécie pertencente ao mesmo gênero da bactéria estudada.

A caixa de saída do sRNATool está padronizada e cada registro contém as seguintes informações: sequência de DNA, número da região intergênica, tamanho em *bp* da IgR, posição inicial e final do candidato no arquivo de IgRs, orientação da sequência, tamanho do candidato, orientação dos genes adjacentes, predição de promotores e terminadores, conteúdo GC e os dois principais *hits* obtidos pelo BLAST para candidatos com similaridade (Figura 16a). Os dados da caixa de saída podem ser agrupados de acordo com a definição dos parâmetros já comentados acima. Neste trabalho, os dados da caixa de saída foram subdivididos em 3 Tabelas: Tabela 1- sRNA candidatos que apresentam promotor, terminador, tamanho entre 50 e 400 *bp*, conservação da sequência entre organismos relacionados e comprovação na base de dados do Rfam; Tabela 2 - possui as mesmas características da Tabela 1, exceto pelo fato dos candidatos relacionados não constarem no Rfam; e Tabela 3 – sRNA candidatos que apresentam promotor, terminador e tamanho adequado, porém, não possuem confirmação no Rfam e não possuem similaridade entre espécies próximas. A existência deste último grupo de sequências sugere que o banco de dados Rfam, embora abrangente, não foi suficientemente completo para a indicação de similaridade (Tabela 2), e que é possível supor a existência de sRNAs exclusivos do gênero *Herbaspirillum*. O conteúdo GC médio dos candidatos nas Tabelas 1, 2 e 3 foi de 53%, 51,5% e 49,63% respectivamente, valores abaixo dos 63,39% registrado para o genoma de *H. seropedicae*. O tamanho médio dos candidatos foi de 278, 186, e 151 *bp* (Tabelas 1, 2 e 3 respectivamente).

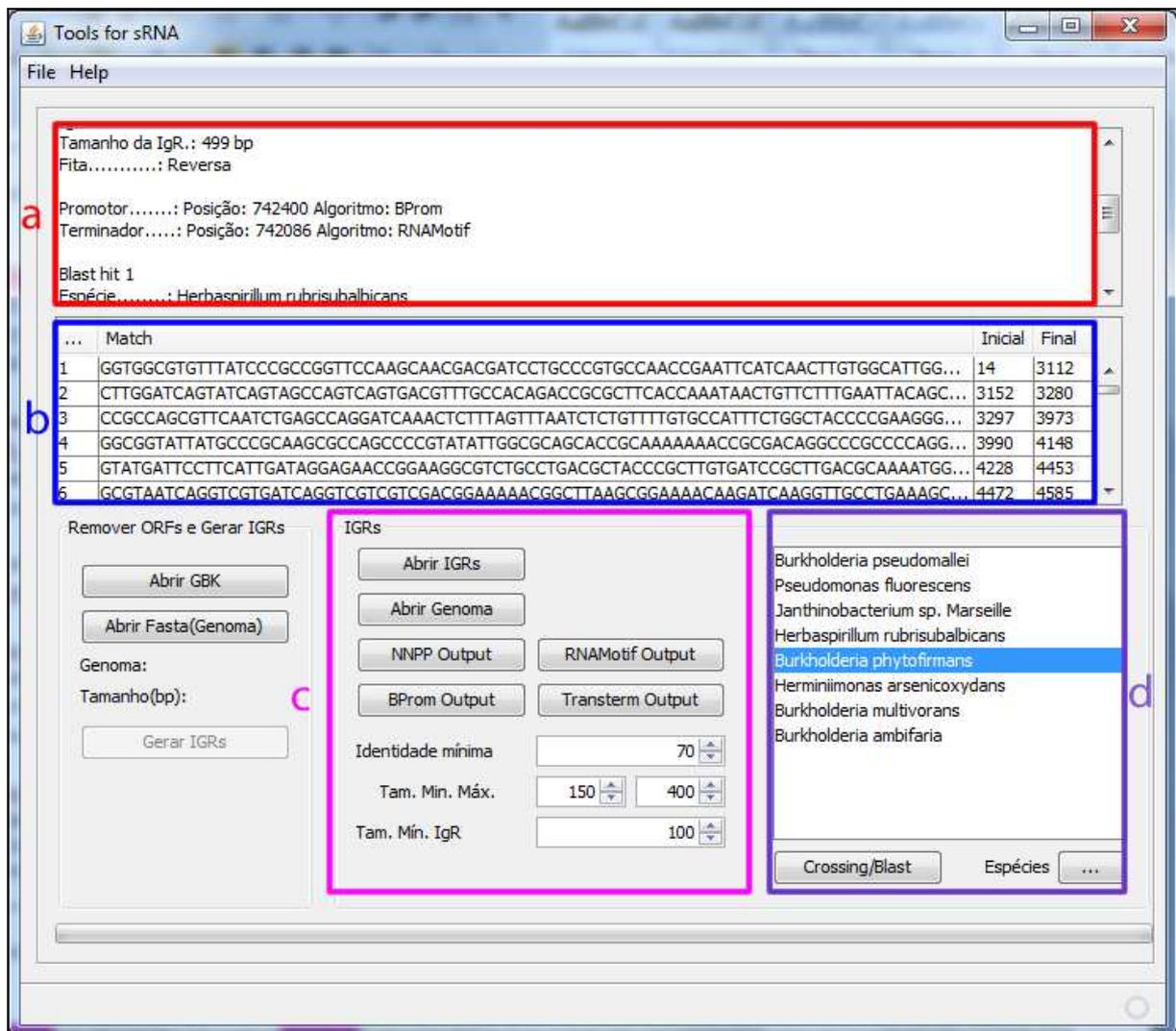


Figura 16 – Captura de tela do sRNATool – Ferramenta Integradora.

a) Caixa de saída dos dados, b) Carregamento das IGRs, c) Carregar dados de promotores e terminadores, além da definição de parâmetros e d) Seleção das espécies que serão utilizadas para a criação do banco de dados do BLAST.

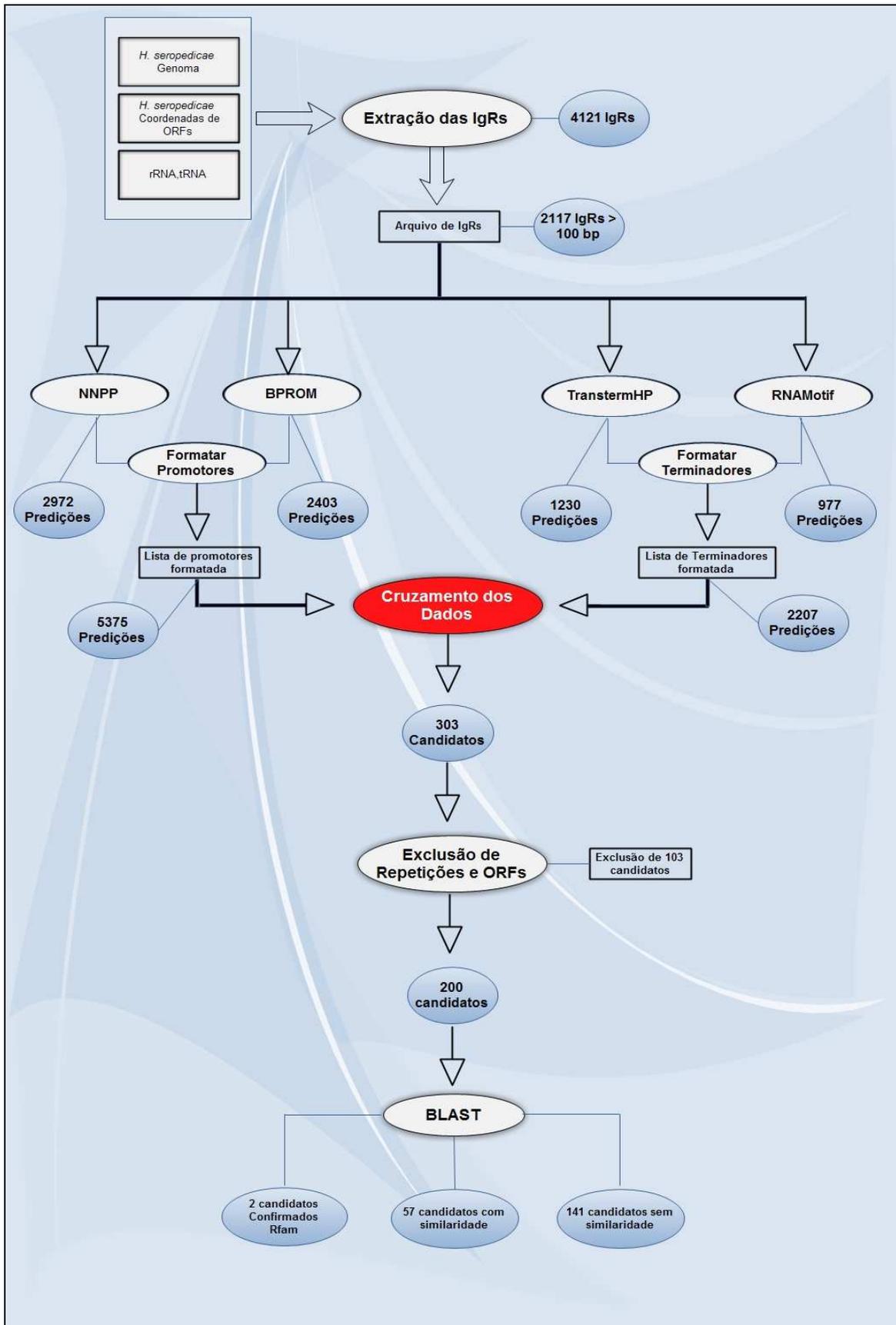


Figura 17 – Fluxo de trabalho com os resultados obtidos em cada etapa.

A carga dos dados do genoma de *H. seropedicae* SmR1 no sRNATool (item 3.6.1) permitiu extrair inicialmente 4121 IgRs e, após a definição de um tamanho mínimo de 100 *bp*, o número de IgRs foi reduzido para 2117 (Figura 17). A definição de um tamanho mínimo foi necessária pois uma IgR inferior a 100 *bp* não é capaz de comportar um RNA curto com uma região promotora. O conteúdo médio de GC nas IgRs foi de 57,32%, enquanto que no genoma completo foi de 63,39%. Examinando a distribuição dos candidatos, constatou-se não mais que um sRNA por IgR. Estas IgRs possuíam um comprimento variando de 300 a 1000 *bp* na maioria dos casos.

Quanto ao número de predições, 1048 possíveis regiões promotoras na fita principal e 1355 na fita reversa foram identificadas pelo BProm, e 2972 predições em ambas as fitas foram obtidas pelo NNPP, totalizando 5375 sequências. O número de predições obtidas para terminadores pelo RNAMotif e TranstermHP foi 977 e 1230 respectivamente, totalizando 2207 sequências.

A fase seguinte foi o cruzamento dos dados seguido pela busca por similaridade. Após o cruzamento, o número de sequências candidatas a sRNA reduziu-se para 303 (Figura 17). Uma vez que foram utilizados dois softwares para a predição tanto de promotores quanto de terminadores, verificaram-se muitas predições repetidas que foram removidas manualmente da lista. Além disso, algumas regiões codificantes de *H. seropedicae* SmR1 foram identificadas e também removidas. As 200 sequências candidatas restantes foram automaticamente comparadas com o BLAST e, como resultado da busca por similaridade, as sequências puderam ser divididas em três grupos: 1. sequências com identidade confirmada (2 candidatos), 2. sequências candidatas que possuíam certo grau de similaridade com pelo menos uma das espécies constantes no banco comparativo (57 candidatos) e 3. sequências candidatas que não apresentaram nenhum registro de similaridade (141 candidatos). As sequências distribuídas nas categorias 1, 2 e 3 estão listadas nas Tabelas 1, 2 e 3 respectivamente. Exemplos de sequências candidatas a sRNA e suas características para cada categoria, estão apresentados nas Figura 19, Figura 18, Figura 20 e Figura 21.

Tabela 1 - Candidatos com similaridade entre espécies e com confirmação no Rfam.

Núm. Seq.	Núm. IGR	Tam. IGR	Posição Inicial	Posição Final	Orient. Seq.	Tam. sRNA.	Orient. Genes Adj.	Software Prom.	Escore Prom.	Software Terminador	Escore Blast	Identidade Blast	Conteúdo GC	RNA Conf.
1	3474	466 bp	662940	663194	>	217 bp	< >	NNPP	117	TransTerm	422 bits (228)	238/243 (97%)	52,11%	6S
2	3796	499 bp	742400	742086	<	314 bp	> <	BProm	0,98	RNAMotif	340 bits (184)	278/321 (86%)	54,14%	4.5S

Tabela 2 - Candidatos com similaridade entre espécies e sem confirmação no Rfam.

Núm. Seq.	Núm. IGR	Tam. IGR	Posição Inicial	Posição Final	Orient. Seq.	Tam. Seq.	Orient. Genes Adj.	Software Promotor	Score Prom.	Software Terminador	Score BLAST	Identidade BLAST	Conteúdo GC
1	1227	631 bp	232165	232299	>	134 bp	<>	BProm	123	RNAMotif	180 bits (97)	131/145 (90%)	47,12%
2	3796	499 bp	742400	742086	<	314 bp	><	BProm	117	RNAMotif	340 bits (184)	278/321 (86%)	54,14%
3	1672	418 bp	319508	319352	<	156 bp	<<	BProm	113	RNAMotif	270 bits (146)	153/156 (98%)	51,28%
4	2823	775 bp	534223	534003	<	220 bp	<<	BProm	113	RNAMotif	235 bits (127)	127/127 (100%)	54,09%
5	1727	422 bp	331340	331173	<	167 bp	<<	BProm	104	RNAMotif	100 bits (54)	72/80 (90%)	56,29%
6	2316	478 bp	442630	442492	<	138 bp	><	BProm	97	RNAMotif	100 bits (54)	58/60 (96%)	48,55%
7	1818	775 bp	348567	348820	>	253 bp	<<	BProm	95	TranstermHP	113 bits (61)	72/77 (93%)	53,36%
8	552	575 bp	100359	100035	<	324 bp	<<	BProm	93	RNAMotif	335 bits (181)	287/334 (85%)	57,10%
9	276	290 bp	49547	49307	<	240 bp	>>	BProm	91	RNAMotif	211 bits (114)	197/237 (83%)	46,25%
10	867	497 bp	159165	158974	<	191 bp	>>	BProm	89	RNAMotif	285 bits (154)	178/189 (94%)	48,69%
11	255	316 bp	45647	45402	<	245 bp	<>	BProm	89	RNAMotif	178 bits (96)	108/114 (94%)	44,90%
12	1243	278 bp	235035	234940	<	95 bp	<>	BProm	86	RNAMotif	99,0 bits (53)	73/83 (87%)	50,53%
13	542	357 bp	98073	98196	>	123 bp	<>	BProm	81	RNAMotif	145 bits (78)	108/122 (88%)	61,29%
14	42	511 bp	10798	11157	>	359 bp	<>	BProm	74	RNAMotif	305 bits (165)	206/225 (91%)	53,19%
15	2077	473 bp	401580	401860	>	280 bp	<>	BProm	74	RNAMotif	490 bits (265)	275/280 (98%)	48,40%
16	884	1217 bp	162503	162702	>	199 bp	><	BProm	74	RNAMotif	111 bits (60)	92/106 (86%)	55,03%
17	415	783 bp	74058	74281	>	223 bp	><	BProm	73	TranstermHP	147 bits (79)	107/119 (89%)	43,52%
18	413	1006 bp	73186	73010	<	176 bp	<>	BProm	73	RNAMotif	111 bits (60)	70/74 (94%)	59,66%
19	4050	612 bp	790245	790679	>	434 bp	<>	BProm	73	RNAMotif	484 bits (262)	268/271 (98%)	59,16%
20	1353	972 bp	256989	257285	>	296 bp	<>	BProm	72	TranstermHP	398 bits (215)	271/297 (91%)	47,37%
21	2572	453 bp	489905	490297	>	392 bp	>>	BProm	72	RNAMotif	226 bits (122)	215/256 (83%)	58,56%
22	1902	502 bp	363738	363947	>	209 bp	>>	BProm	71	TranstermHP	196 bits (106)	117/122 (95%)	63,69%
23	3935	219 bp	769947	770088	>	141 bp	>>	BProm	70	TranstermHP	176 bits (95)	119/130 (91%)	54,95%
24	2564	465 bp	487844	488170	>	326 bp	<<	BProm	70	TranstermHP	169 bits (91)	181/222 (81%)	57,09%
25	3174	278 bp	597164	597280	>	116 bp	<>	NNPP	0,99	TranstermHP	99,0 bits (53)	60/63 (95%)	51,16%
26	1672	418 bp	319519	319352	<	167 bp	<<	NNPP	0,99	TranstermHP	291 bits (157)	164/167 (98%)	49,10%
27	3047	580 bp	575588	575337	<	251 bp	<<	NNPP	0,98	TranstermHP	350 bits (189)	231/250 (92%)	53,78%
28	817	289 bp	150084	150232	>	148 bp	>>	NNPP	0,98	RNAMotif	211 bits (114)	133/142 (93%)	55,08%
29	276	290 bp	49526	49307	<	219 bp	>>	NNPP	0,97	RNAMotif	195 bits (105)	180/216 (83%)	48,40%

Núm. Seq.	Núm. IGR	Tam. IGR	Posição Inicial	Posição Final	Orient. Seq.	Tam. Seq.	Orient. Genes Adj.	Software Promotor	Score Prom.	Software Terminador	Score BLAST	Identidade BLAST	Conteúdo GC
30	2291	253 bp	437136	437018	<	118 bp	><	NNPP	0,97	RNAMotif	171 bits (92)	109/117 (93%)	45,76%
31	3270	555 bp	618211	618517	>	306 bp	<<	NNPP	0,96	RNAMotif	361 bits (195)	271/307 (88%)	50,36%
32	2922	520 bp	553127	553319	>	192 bp	>>	NNPP	0,95	TranstermHP	265 bits (143)	178/194 (91%)	57,41%
33	255	316 bp	45660	45402	<	258 bp	<>	NNPP	0,95	RNAMotif	202 bits (109)	121/127 (95%)	43,80%
34	1065	365 bp	196212	196530	>	318 bp	<>	NNPP	0,95	RNAMotif	195 bits (105)	258/327 (78%)	40,63%
35	775	479 bp	141905	141690	<	215 bp	>>	NNPP	0,95	RNAMotif	104 bits (56)	73/81 (90%)	44,19%
36	3047	580 bp	575604	575337	<	267 bp	<<	NNPP	0,94	RNAMotif	372 bits (201)	249/271 (91%)	52,06%
37	83	405 bp	17770	17429	<	341 bp	<<	NNPP	0,93	TranstermHP	274 bits (148)	278/336 (82%)	50,44%
38	2029	276 bp	389650	389453	<	197 bp	<<	NNPP	0,93	RNAMotif	267 bits (144)	180/197 (91%)	58,38%
39	276	290 bp	49518	49307	<	211 bp	>>	NNPP	0,93	RNAMotif	185 bits (100)	173/208 (83%)	48,82%
40	255	316 bp	45627	45402	<	225 bp	<>	NNPP	0,93	TranstermHP	158 bits (85)	91/94 (96%)	47,11%
41	3027	545 bp	572506	572410	<	96 bp	<>	NNPP	0,92	RNAMotif	139 bits (75)	75/75 (100%)	48,96%
42	2077	473 bp	401617	401860	>	243 bp	<>	NNPP	0,91	RNAMotif	422 bits (228)	238/243 (97%)	52,11%
43	1111	315 bp	206295	206489	>	194 bp	<>	NNPP	0,9	TranstermHP	182 bits (98)	149/172 (86%)	54,27%
44	1111	315 bp	206295	206449	>	154 bp	<>	NNPP	0,9	TranstermHP	148 bits (80)	131/154 (85%)	56,45%
45	2900	800 bp	548230	548599	>	369 bp	<>	NNPP	0,89	TranstermHP	433 bits (234)	334/379 (88%)	66,08%
46	3297	545 bp	626144	625934	<	210 bp	<<	NNPP	0,89	TranstermHP	305 bits (165)	191/203 (94%)	44,29%
47	1430	394 bp	270404	270595	>	191 bp	>>	NNPP	0,87	RNAMotif	322 bits (174)	185/190 (97%)	45,34%
48	2291	253 bp	437177	437018	<	159 bp	><	NNPP	0,87	RNAMotif	191 bits (103)	134/149 (89%)	42,77%
49	867	497 bp	159137	158974	<	163 bp	>>	NNPP	0,86	TranstermHP	250 bits (135)	153/161 (95%)	50,31%
50	4097	551 bp	801628	801396	<	232 bp	><	NNPP	0,86	RNAMotif	128 bits (69)	188/242 (77%)	56,47%
51	1227	631 bp	232116	232299	>	183 bp	<>	NNPP	0,85	RNAMotif	198 bits (107)	166/192 (86%)	45,10%
52	740	213 bp	136437	136300	<	137 bp	><	NNPP	0,84	RNAMotif	84,2 bits (45)	49/51 (96%)	47,45%
53	3474	466 bp	662940	663194	>	254 bp	><	NNPP	0,84	TranstermHP	396 bits (214)	224/229 (97%)	50,00%
54	2029	276 bp	389467	389573	>	106 bp	<<	NNPP	0,83	TranstermHP	150 bits (81)	98/106 (92%)	65,79%
55	3935	219 bp	769980	770088	>	108 bp	>>	NNPP	0,83	TranstermHP	139 bits (75)	99/110 (90%)	55,13%
56	3164	303 bp	595464	595589	>	125 bp	>>	NNPP	0,83	RNAMotif	115 bits (62)	105/124 (84%)	57,89%
57	81	466 bp	17090	16853	<	237 bp	<<	NNPP	0,81	RNAMotif	405 bits (219)	232/238 (97%)	50,21%

Tabela 3 - Candidatos sem similaridade entre espécies e sem confirmação no Rfam.

Núm. Sequência	Número IGR	Tamanho IGR	Posição Inicial Seq.	Posição Final Seq.	Orientação da Seq.	Tamanho da Seq.	Orientação Genes Adj.	Software Promotor	Escore Promotor	Software Terminador	Conteúdo GC
1	317	375 bp	54041	53918	<	123 bp	<<	BProm	117	TranstermHP	47,15%
2	3302	179 bp	627219	627142	<	77 bp	<<	BProm	114	RNAMotif	55,84%
3	3829	546 bp	751407	751091	<	316 bp	<<	BProm	111	RNAMotif	57,59%
4	3348	442 bp	637432	637498	>	66 bp	<	BProm	108	RNAMotif	36,11%
5	83	405 bp	17759	17429	<	330 bp	<<	BProm	103	RNAMotif	50,30%
6	2230	166 bp	425878	425795	<	83 bp	>>	BProm	103	RNAMotif	34,94%
7	1075	460 bp	198296	198541	>	245 bp	<<	BProm	102	RNAMotif	53,02%
8	3270	555 bp	618425	618517	>	92 bp	<<	BProm	100	RNAMotif	43,55%
9	3811	642 bp	746741	746558	<	183 bp	>>	BProm	99	TranstermHP	50,82%
10	1049	272 bp	193553	193630	>	77 bp	>>	BProm	98	RNAMotif	46,81%
11	425	274 bp	76908	76846	<	62 bp	>>	BProm	96	RNAMotif	46,77%
12	230	275 bp	41075	41165	>	90 bp	><	BProm	95	RNAMotif	40,00%
13	1474	353 bp	279563	279679	>	116 bp	>>	BProm	94	RNAMotif	40,70%
14	2971	442 bp	561883	561663	<	220 bp	<>	BProm	94	RNAMotif	39,09%
15	817	289 bp	150117	150232	>	115 bp	>>	BProm	92	RNAMotif	54,12%
16	3989	640 bp	779741	779837	>	96 bp	<>	BProm	91	TranstermHP	65,15%
17	2484	431 bp	473127	473484	>	357 bp	<>	BProm	88	TranstermHP	48,62%
18	4108	283 bp	804934	805022	>	88 bp	<>	BProm	88	RNAMotif	46,55%
19	4108	283 bp	804934	805127	>	193 bp	<>	BProm	88	RNAMotif	52,76%
20	2930	361 bp	555421	555136	<	285 bp	<<	BProm	87	TranstermHP	41,75%
21	3595	6867 bp	690553	690452	<	101 bp	>>	BProm	87	RNAMotif	61,39%
22	3595	6867 bp	690553	690395	<	158 bp	>>	BProm	87	RNAMotif	60,13%
23	2307	445 bp	439851	440088	>	237 bp	<<	BProm	86	RNAMotif	50,24%
24	603	496 bp	109959	109798	<	161 bp	<<	BProm	84	TranstermHP	61,49%
25	4049	277 bp	789908	789995	>	87 bp	><	BProm	84	TranstermHP	49,12%
26	1970	198 bp	378093	378199	>	106 bp	<>	BProm	83	TranstermHP	55,26%
27	944	148 bp	174864	174959	>	95 bp	><	BProm	79	TranstermHP	50,77%
28	3677	922 bp	712859	713000	>	141 bp	<>	BProm	77	RNAMotif	48,65%
29	243	518 bp	43229	43444	>	215 bp	<<	BProm	76	RNAMotif	48,65%

Núm. Sequência	Número IGR	Tamanho IGR	Posição Inicial Seq.	Posição Final Seq.	Orientação da Seq.	Tamanho da Seq.	Orientação Genes Adj.	Software Promotor	Escore Promotor	Software Terminador	Conteúdo GC
30	3742	666 bp	728078	727746	<	332 bp	><	BProm	75	TranstermHP	50,60%
31	1645	452 bp	314637	314725	>	88 bp	>>	BProm	74	RNAMotif	55,17%
32	431	282 bp	77790	77699	<	91 bp	><	BProm	73	TranstermHP	37,36%
33	737	988 bp	135286	135634	>	348 bp	>>	BProm	73	TranstermHP	55,97%
34	89	355 bp	19157	18842	<	315 bp	>>	BProm	73	RNAMotif	54,60%
35	1548	722 bp	295474	295574	>	100 bp	>>	BProm	73	TranstermHP	35,71%
36	3977	267 bp	777878	778011	>	133 bp	>>	BProm	73	RNAMotif	63,11%
37	2429	487 bp	463442	463079	<	363 bp	><	BProm	73	RNAMotif	57,85%
38	3782	357 bp	738059	737909	<	150 bp	<<	BProm	72	RNAMotif	45,33%
39	2074	1112 bp	400871	400595	<	276 bp	>>	BProm	72	RNAMotif	43,12%
40	3595	6867 bp	694213	693889	<	324 bp	>>	BProm	72	RNAMotif	61,42%
41	3786	1001 bp	739225	739089	<	136 bp	>>	BProm	72	TranstermHP	49,26%
42	2161	296 bp	414517	414361	<	156 bp	>>	BProm	65	TranstermHP	54,49%
43	2572	453 bp	489905	490297	>	392 bp	>>	BProm	64	TranstermHP	58,56%
44	3560	627 bp	681552	681770	>	218 bp	<>	BProm	63	RNAMotif	57,45%
45	2301	178 bp	438469	438345	<	124 bp	<<	BProm	62	RNAMotif	60,48%
46	3177	164 bp	597886	597798	<	88 bp	><	BProm	71	TranstermHP	44,32%
47	590	435 bp	107919	108021	>	102 bp	<>	BProm	71	TranstermHP	59,72%
48	1535	1804 bp	291401	291532	>	131 bp	<>	BProm	71	RNAMotif	49,50%
49	3639	296 bp	705456	705599	>	143 bp	><	BProm	71	RNAMotif	41,59%
50	2824	330 bp	534792	535031	>	239 bp	<<	BProm	71	RNAMotif	51,67%
51	3484	996 bp	666093	666244	>	151 bp	>>	BProm	71	RNAMotif	52,07%
52	371	488 bp	63958	63796	<	162 bp	<<	BProm	71	TranstermHP	47,53%
53	3869	144 bp	757932	758032	>	100 bp	><	BProm	71	TranstermHP	41,43%
54	1	3098 bp	2015	2367	>	352 bp	><	BProm	70	RNAMotif	50,00%
55	3755	482 bp	731018	730786	<	232 bp	>>	BProm	70	RNAMotif	48,71%
56	3884	507 bp	760265	759869	<	396 bp	<>	BProm	70	RNAMotif	44,70%
57	3260	658 bp	615168	615071	<	97 bp	>>	BProm	70	TranstermHP	53,61%
58	3093	666 bp	585503	585845	>	342 bp	><	BProm	70	RNAMotif	63,78%
59	2481	590 bp	472372	472498	>	126 bp	<<	BProm	70	TranstermHP	51,04%
60	3270	555 bp	618422	618517	>	95 bp	<<	NNPP	1	TranstermHP	41,54%

Núm. Sequência	Número IGR	Tamanho IGR	Posição Inicial Seq.	Posição Final Seq.	Orientação da Seq.	Tamanho da Seq.	Orientação Genes Adj.	Software Promotor	Escore Promotor	Software Terminador	Conteúdo GC
61	3484	996 bp	665939	666244	>	305 bp	>>	NNPP	1	RNAMotif	45,82%
62	99	142 bp	20718	20628	<	90 bp	><	NNPP	0,99	TranstermHP	46,67%
63	230	275 bp	41068	41165	>	97 bp	><	NNPP	0,99	RNAMotif	37,31%
64	1548	722 bp	295228	295045	<	183 bp	>>	NNPP	0,99	RNAMotif	41,53%
65	2021	322 bp	388368	388465	>	97 bp	>>	NNPP	0,99	RNAMotif	61,19%
66	3277	521 bp	620445	620306	<	139 bp	>>	NNPP	0,99	TranstermHP	51,08%
67	3466	666 bp	661184	661345	>	161 bp	>>	NNPP	0,99	RNAMotif	51,15%
68	3068	708 bp	580980	580795	<	185 bp	<<	NNPP	0,98	TranstermHP	49,19%
69	3068	708 bp	580980	580692	<	288 bp	<<	NNPP	0,98	TranstermHP	53,47%
70	3193	169 bp	602310	602190	<	120 bp	><	NNPP	0,98	RNAMotif	52,50%
71	3639	296 bp	705457	705599	>	142 bp	><	NNPP	0,98	RNAMotif	41,07%
72	3979	153 bp	778328	778441	>	113 bp	><	NNPP	0,98	TranstermHP	36,14%
73	317	375 bp	54052	53918	<	134 bp	<<	NNPP	0,97	RNAMotif	49,25%
74	3277	521 bp	620692	620306	<	386 bp	>>	NNPP	0,97	RNAMotif	48,96%
75	3639	296 bp	705404	705599	>	195 bp	><	NNPP	0,97	TranstermHP	42,42%
76	1501	661 bp	285125	285418	>	293 bp	<<	NNPP	0,96	RNAMotif	44,87%
77	2426	241 bp	462532	462691	>	159 bp	<>	NNPP	0,96	RNAMotif	55,81%
78	3270	555 bp	618211	618517	>	306 bp	<<	NNPP	0,96	RNAMotif	50,36%
79	3277	521 bp	620565	620306	<	259 bp	>>	NNPP	0,96	RNAMotif	47,49%
80	3829	546 bp	751415	751091	<	324 bp	<<	NNPP	0,96	TranstermHP	56,79%
81	276	290 bp	49557	49307	<	250 bp	>>	NNPP	0,95	RNAMotif	46,59%
82	1065	365 bp	196212	196530	>	318 bp	<>	NNPP	0,95	RNAMotif	40,63%
83	1534	184 bp	291043	291188	>	145 bp	<<	NNPP	0,95	TranstermHP	57,39%
84	2307	445 bp	439886	440088	>	202 bp	<<	NNPP	0,95	RNAMotif	52,91%
85	2726	292 bp	518884	518684	<	200 bp	>>	NNPP	0,95	TranstermHP	59,00%
86	2922	520 bp	553127	553319	>	192 bp	>>	NNPP	0,95	RNAMotif	57,41%
87	3251	118 bp	612717	612800	>	83 bp	>>	NNPP	0,95	RNAMotif	67,92%
88	184	526 bp	35761	35654	<	107 bp	<<	NNPP	0,94	TranstermHP	50,47%
89	243	518 bp	43228	43444	>	216 bp	<<	NNPP	0,94	RNAMotif	48,39%
90	1535	1804 bp	291442	291532	>	90 bp	<>	NNPP	0,94	RNAMotif	60,00%
91	2230	166 bp	425930	425795	<	135 bp	>>	NNPP	0,94	TranstermHP	33,33%

Núm. Sequência	Número IGR	Tamanho IGR	Posição Inicial Seq.	Posição Final Seq.	Orientação da Seq.	Tamanho da Seq.	Orientação Genes Adj.	Software Promotor	Escore Promotor	Software Terminador	Conteúdo GC
92	2400	358 bp	458496	458672	>	176 bp	>>	NNPP	0,94	RNAMotif	58,22%
93	3677	922 bp	712708	713000	>	292 bp	<>	NNPP	0,94	TranstermHP	41,98%
94	3677	922 bp	712765	713000	>	235 bp	<>	NNPP	0,94	RNAMotif	42,44%
95	3677	922 bp	712790	713000	>	210 bp	<>	NNPP	0,94	RNAMotif	44,44%
96	3931	213 bp	769276	769162	<	114 bp	<>	NNPP	0,94	TranstermHP	62,28%
97	1548	722 bp	295336	295574	>	238 bp	>>	NNPP	0,93	RNAMotif	41,83%
98	3277	521 bp	620531	620306	<	225 bp	>>	NNPP	0,93	TranstermHP	50,22%
99	1065	365 bp	196359	196530	>	171 bp	<>	NNPP	0,92	RNAMotif	46,10%
100	2930	361 bp	555267	555136	<	131 bp	<<	NNPP	0,92	RNAMotif	38,93%
101	3027	545 bp	572506	572410	<	96 bp	<>	NNPP	0,92	TranstermHP	48,96%
102	3302	179 bp	627230	627142	<	88 bp	<<	NNPP	0,92	TranstermHP	59,09%
103	1535	1804 bp	291261	291532	>	271 bp	<>	NNPP	0,91	TranstermHP	43,57%
104	1548	722 bp	295140	295045	<	95 bp	>>	NNPP	0,91	TranstermHP	45,26%
105	2738	196 bp	520927	520794	<	133 bp	>>	NNPP	0,9	RNAMotif	48,87%
106	3639	296 bp	705436	705599	>	163 bp	><	NNPP	0,9	RNAMotif	42,11%
107	4098	415 bp	802099	802223	>	124 bp	<>	NNPP	0,9	RNAMotif	54,26%
108	4098	415 bp	802129	802223	>	94 bp	<>	NNPP	0,9	RNAMotif	64,06%
109	1	3098 bp	1987	2367	>	380 bp	><	NNPP	0,89	RNAMotif	49,43%
110	1065	365 bp	196327	196530	>	203 bp	<>	NNPP	0,89	RNAMotif	44,51%
111	1548	722 bp	295473	295574	>	101 bp	>>	NNPP	0,89	RNAMotif	35,21%
112	17	485 bp	5875	5756	<	119 bp	<<	NNPP	0,88	TranstermHP	48,74%
113	230	275 bp	40959	41165	>	206 bp	><	NNPP	0,88	TranstermHP	42,05%
114	310	150 bp	52628	52762	>	134 bp	<<	NNPP	0,88	TranstermHP	48,08%
115	3270	555 bp	618401	618517	>	116 bp	<<	NNPP	0,88	TranstermHP	38,37%
116	3989	640 bp	779741	779837	>	96 bp	<>	NNPP	0,88	TranstermHP	65,15%
117	2307	445 bp	439993	440088	>	95 bp	<<	NNPP	0,87	RNAMotif	61,54%
118	2314	604 bp	441893	441745	<	148 bp	>>	NNPP	0,87	RNAMotif	54,73%
119	3466	666 bp	661206	661345	>	139 bp	>>	NNPP	0,87	RNAMotif	56,88%
120	3484	996 bp	665953	666244	>	291 bp	>>	NNPP	0,87	RNAMotif	46,74%
121	3782	357 bp	738069	737909	<	160 bp	<<	NNPP	0,87	RNAMotif	44,38%
122	1535	1804 bp	291272	291532	>	260 bp	<>	NNPP	0,86	RNAMotif	43,91%

Núm. Sequência	Número IGR	Tamanho IGR	Posição Inicial Seq.	Posição Final Seq.	Orientação da Seq.	Tamanho da Seq.	Orientação Genes Adj.	Software Promotor	Escore Promotor	Software Terminador	Conteúdo GC
123	1537	309 bp	293468	293251	<	217 bp	>>	NNPP	0,86	RNAMotif	47,93%
124	3742	666 bp	728090	727746	<	344 bp	><	NNPP	0,86	TranstermHP	50,87%
125	1030	246 bp	190229	190321	>	92 bp	><	NNPP	0,85	RNAMotif	33,87%
126	1537	309 bp	293372	293251	<	121 bp	>>	NNPP	0,85	RNAMotif	51,24%
127	3466	666 bp	661149	661345	>	196 bp	>>	NNPP	0,85	TranstermHP	43,37%
128	1494	422 bp	283580	283326	<	254 bp	<>	NNPP	0,84	RNAMotif	52,76%
129	1548	722 bp	295314	295574	>	260 bp	>>	NNPP	0,84	RNAMotif	42,61%
130	1805	361 bp	346251	346148	<	103 bp	>>	NNPP	0,84	TranstermHP	58,25%
131	3884	507 bp	760185	759869	<	316 bp	<>	NNPP	0,84	RNAMotif	43,99%
132	173	137 bp	34072	33969	<	103 bp	>>	NNPP	0,83	TranstermHP	61,17%
133	2230	166 bp	425895	425795	<	100 bp	>>	NNPP	0,83	TranstermHP	31,00%
134	2261	345 bp	432131	431977	<	154 bp	><	NNPP	0,83	TranstermHP	63,64%
135	2426	241 bp	462563	462691	>	128 bp	<>	NNPP	0,83	TranstermHP	59,18%
136	1535	1804 bp	291401	291532	>	131 bp	<>	NNPP	0,82	RNAMotif	49,50%
137	1548	722 bp	295377	295045	<	332 bp	>>	NNPP	0,82	RNAMotif	43,98%
138	3783	445 bp	738167	738290	>	123 bp	<>	NNPP	0,82	RNAMotif	50,54%
139	3884	507 bp	760258	759869	<	389 bp	<>	NNPP	0,82	TranstermHP	44,73%
140	1302	109 bp	246678	246594	<	84 bp	><	NNPP	0,81	RNAMotif	47,62%
141	2849	151 bp	540034	540128	>	94 bp	<<	NNPP	0,81	TranstermHP	37,50%

```

TGGCCC GGTCGGTGTGATTTGAGATACAAT TCACCCCGGCGGGCCCTGCGCATTGTGGCATGGTCAACCT
GGTCAGGT CGGGAACGAA GCAGCCACAGCCATTTCCCGCAAGT GCCGCAGACAAGGCTCGCCTCCTTCCG
TTTT

Tamanho sRNA...: 116 bp
IgR.....: 1452
Tamanho da IgR.: 356 bp
Fita.....: Principal

Promotor.....: Posição: 275590 Algoritmo: BProm
Terminador....: Posição: 275706 Algoritmo: RNAMotif

Blast hit 1
Espécie.....: Herbaspirillum rubrisubalbicans
Escore.....: 226 bits (122)
Identidade....: 178/204 (87%)

```

Figura 18 – IgR 1452 (Tabela 1) localizada no genoma de *H. seropedicae* SmR1 e identificada no banco de dados como 4.5S RNA. Em rosa a provável região promotora e em verde a região terminadora.

```

17 bp
TTGGAT CGTGGCAGCGAATGCGT TAGACTTCG TACTCCTG CCGTGTT CGCGATTGCC ATATATTCCTTGAA
CCATTTACGAGCATAGGTTGTGGGATCGTTTGATGGGCGTGAGCGT CGCTAGTCCGATGAACCCGAAATTG
AACTACTGCAACCAACTTGAACCGTCAGGTT CAGGATGCCGGCACAGCGGCACAGGCGGGGCTTGATTCAA
GAAAAACGGCTGCATATGCAGCCGTTTTTTTTT

Tamanho sRNA...: 217 bp
IgR.....: 3474
Tamanho da IgR.: 466 bp
Fita.....: Principal

Promotor.....: Posição: 662940 Algoritmo: NNPP
Terminador....: Posição: 663194 Algoritmo: TranstermHP

Blast hit 1
Espécie.....: Herbaspirillum rubrisubalbicans
Escore.....: 396 bits (214)
Identidade....: 224/229 (97%)

Blast hit 2
Espécie.....: Janthinobacterium sp. Marseille
Escore.....: 214 bits (236)
Identidade....: 180/217 (83%)

```

Figura 19 – IgR 3474 (Tabela 1) localizada no genoma de *H. seropedicae* SmR1 e identificada no banco de dados como 6S sRNA. Em rosa a provável região promotora e em verde a região terminadora.

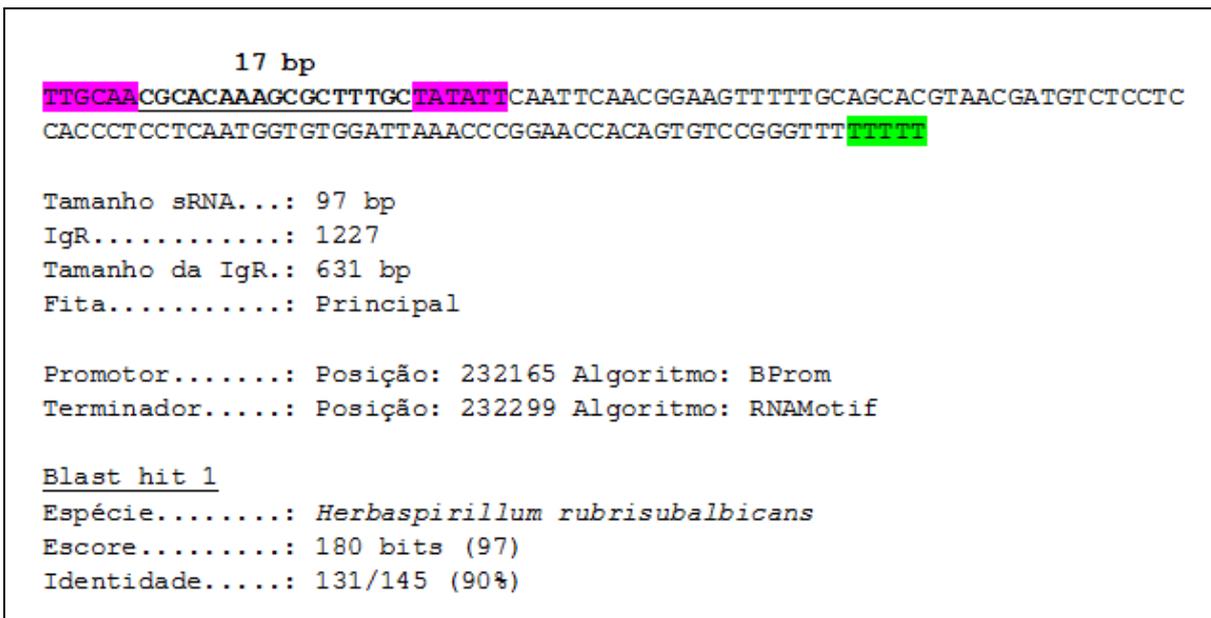


Figura 20 – IgR 1227 (Tabela 2) localizada no genoma de *H. seropedicae* SmR1 com similaridade e sem confirmação no Rfam. Em rosa a provável região promotora e em verde a região terminadora.

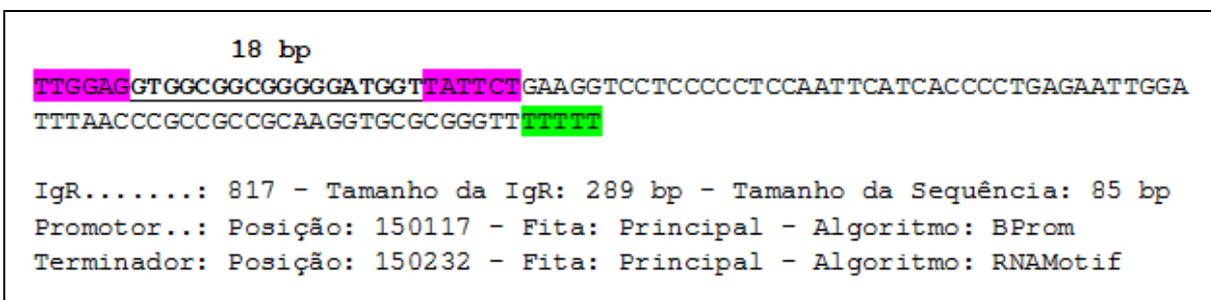


Figura 21 – IgR 817 (Tabela 3) localizada no genoma de *H. seropedicae* SmR1 sem similaridade e sem confirmação no Rfam. Em rosa a provável região promotora e em verde a região terminadora.

Dentre a totalidade de sequências obtidas como candidatas a sRNA, duas delas tiveram identidade confirmada pelo Rfam: 6S RNA e 4.5S RNA (Tabela 1, Figura 18 e Figura 19).

O 6S RNA é codificado pelo gene *sseS* e é continuamente expresso, embora seja mais abundante na fase final estacionária, quando alcança 10.000 cópias por célula (BROWNLEE, 1971). Durante a fase exponencial e estacionária, pelo menos 90% dos 6S RNAs presentes na célula são ligados ao um fator $\sigma 70$ (WASSARMAN e STORZ, 2000). A função do 6S RNA é inibir a transcrição através de sua interação

direta com a holoenzima RNA polimerase. A estrutura secundária deste RNA mimetiza um promotor de DNA aberto e tem capacidade de ligar-se ao complexo σ^{70} - RNA polimerase que é então mantido no estado inativo. Além disso, o sRNA 6S serve como molde para a síntese de um fragmento de 6S RNA de 12-14 nucleotídeos, chamado de pRNA, que desestabiliza o complexo 6SsRNA- σ^{70} - RNA polimerase, liberando o complexo σ^{70} - RNA polimerase. O nível celular do pRNA 6S eleva-se e cai rapidamente, provavelmente devido a digestão por RNase (WASSARMAN e SAECKER, 2006). A regulação da transcrição do 6S conduz à alteração da sobrevivência celular por meio do redirecionamento da utilização de recursos metabólicos sob condições de limitação de nutrientes (ARGAMAN *et al.*, 2001).

O 4.5S RNA é uma sequência de 116 nucleotídeos (Figura 18) que desempenha um papel fundamental no direcionamento de proteínas contendo sinais de peptídeos ao aparelho secretor, formando uma partícula de reconhecimento de sinal com a proteína Ffh. Também se liga independentemente ao fator G de alongação de proteína (EF-F), e funciona no processo de tradução. Este RNA contém uma estrutura secundária com um grampo GGAA e cinco alças internas conservadas filogeneticamente (NAKAMURA *et al.*, 2001).

Como esperado, todos os 57 candidatos da Tabela 2 obtiveram como primeiro *hit* a bactéria *H. rubrisubalbicans M1*, que pertence ao mesmo gênero da *H. seropedicae* SmR1. Para o segundo *hit* constatou-se a seguinte distribuição: 14 candidatos apresentaram similaridade com o gênero *Janthinobacterium*, 6 com *Burkholderia*, 5 com *Hermiimonas* e 1 com *Pseudomonas*. Entretanto, nenhuma destas sequências pôde ser identificada como um sRNA. As sequências dos candidatos listados na Tabela 2 estão detalhadas no APÊNDICE I, com a provável região promotora destacada em rosa e a região terminadora em verde. Algumas sequências promotoras não puderam ser identificadas visualmente por não possuírem uma sequência consenso típica. As sequências dos candidatos listados na Tabela 3 estão detalhadas no APÊNDICE II.

A aplicação da ferramenta integradora sRNATool mostrou-se eficiente, uma vez que foi possível identificar 200 sequências candidatas no genoma de *H. seropedicae* SmR1. Foi possível verificar que a abordagem utilizada no presente projeto pode ser aperfeiçoada pela ampliação de dois parâmetros importantes:

predição de regiões promotoras reconhecidas por fatores σ alternativos e extensão das predições além das regiões intergênicas, uma vez que estudos de *RNomics* identificaram sRNAs parcialmente codificados em fitas não codificantes de ORFs (LIVNY *et al.*, 2005; VOGEL *et al.*, 2003).

5 CONCLUSÕES

Uma ferramenta computacional que integrou os softwares NNPP, Bprom, RNAMotif, TranstermHP e BLAST foi aplicada de forma eficaz, possibilitando a identificação de 200 candidatos a sRNA na bactéria *Herbaspirillum seropedicae* SmR1.

Dentre os candidatos selecionados foram identificados o 4.5S e 6S RNA.

O uso do genoma de *Herbaspirillum rubrisubalbicans* M1 na criação do banco de dados para o uso do BLAST foi decisivo em grande parte das predições, haja vista que a maioria dos sRNAs possivelmente compartilham homologia de sequência entre espécies próximas.

A metodologia aplicada no sRNATool mostrou-se adequada ao objetivo proposto, possibilitando o cruzamento, a formatação e a busca das sequências por similaridade de forma automatizada, gerando uma saída de dados organizada.

O sRNATool é passível de aplicação em outros genomas de procaríotos, exigindo-se apenas o referido genoma juntamente com suas anotações de ORFs, tRNAs e rRNAs.

6 REFERÊNCIAS

AIBA, H. Mechanism of RNA silencing by Hfq-binding small RNAs. **Curr Opin Microbiol.** 10:134-139, 2007.

ALTSCHUL, S. F.; GISH, W.; MILLER, W.; MYERS, E. W.; LIPMAN, D. J. ALTUVIA, S.; ZHANG, A.; ARGAMAN, L.; TIWARI, A.; STORZ, G. The *Escherichia coli* OxyS regulatory RNA represses fhfA translation by blocking ribosome binding. **EMBO**, 20: 6069-6075, 1998.

ALTUVIA, S. Identification of bacterial small non-coding RNAs; experimental approaches. **Curr Opin Microbiol.**, 10: 257-261, 2007.

ARGAMAN, L.; HERSHBERG, R.; VOGEL, J.; BEJERANO, G.; WAGNER, G. H.; MARGALIT, H.; ALTUVIA, S. Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. **Curr Biol**, 11:941-950, 2001.

ARTSIMOVITCH, I.; HENKIN, T. M. In vitro approaches to analysis of transcription termination. **Methods**, 47:37-43, 2009.

BACKOFEN, R; HESS, W.R. Computational prediction of sRNAs and their targets in bacteria. **RNA Biol.**, 7:33-42, 2010.

BALDANI, J. I.; BALDANI, V. L. D.; DOEBEREINER, J.; SELDIN, L. Characterization of *Herbaspirillum seropedicae*. **Intl. J. Sys. Bact.**, 36:86-93, 1986.

BALDANI, J. I.; BALDANI, V. L. D.; KIRCHHOF, G.; FALSEN, E.; OLIVARES, F. L.; HOSTE, B.; KERSTERS, K.; HARTMANN, A.; GILLIS, M.; DÖBEREINER, J. Emended description of *Herbaspirillum*; a mild plant pathogen, as *Herbaspirillum rubrisubalbicans* comb. nov.; and classification of a group of clinical isolates (EF group 1) as *Herbaspirillum* species 3. **Int J Syst Bacteriol**, 46:802-810, 1996.

BALDANI, J.I.; BALDANI, V.L.D. History on the biological nitrogen fixation research in graminaceous plants: special emphasis on the Brazilian experience **Anais da Academia Brasileira de Ciências**, 77:549-579, 2005.

BALDANI, V. L. D.; JAMES, E. K.; BALDANI, J. I.; DÖBEREINER, J. Colonization of rice by the nitrogenfixing *Herbaspirillum* spp. and *Azospirillum brasilense*. **New Horizons in Nitrogen Fixation. Kluwer Academic Publishers**, 1993.

BALDANI, V. L. D.; OLIVARES, F. L.; DÖBEREINER, J. Selection of *Herbaspirillum* spp. strains associated with rice seedlings amended with ¹⁵N-labeled fertilizer. **International Symposium on Sustainable Agriculture for the Tropics: The Role of Biological Nitrogen Fixation**, 1995.

ALTSCHUL, S. F.; GISH, W.; MILLER, W.; MYERS, E. M.; LIPMAN, D. J. BLAST - Basic local alignment search tool. **J. Mol. Biol**, 1990.

BEISEL, C. L.; STORZ, G. Base pairing small RNAs and their roles in global regulatory networks. **FEMS Microbiol Rev.**, 34:866-882, 2010.

BRENNAN, R. G.; LINK, T. M. Hfq structure, function and ligand binding. **Current Opinion in Microbiology**, 10:125-133, 2007.

BROWNLEE, G. G. Sequence of 6S RNA of *Escherichia. coli*. **Nat New Biol.**, 229:147-149, 1971.

CARTER, R. J.; DUBCHAK, I.; HOLBROOKA, S. R. A computational approach to identify genes for functional RNAs in genomic sequences. **Nucleic Acids Res.**, 29:928–3938, 2001.

CHEN, S.; LESNIK, E. A.; HALL, T. A.; SAMPATH, R.; GRIFFEY, R. H.; ECKER, D. J.; BLYN, L. B. A bioinformatics based approach to discover small RNA genes in the *Escherichia coli* genome. **Biosystems**, 65:157-177, 2002.

CIAMPI, M. S. Rho-dependent terminators and transcription termination.

Microbiology, 152:2515-2528, 2006.

COENYE, T.; DREVINEK, P.; MAHENTHIRALINGAM, E.; SHAH, S.; GILL, R. T.; VANDAMME, P.; USSERY, D. W. Identification of putative noncoding RNA genes in the *Burkholderia cenocepacia* J2315 genome. **FEMS Microbiol Lett.**, 276:83-92, 2007.

DALE, J.; PARK, S. Molecular Genetics of Bacteria. **Wiley-Blackwell**, 2004.

DAVIS, B. M.; QUINONES, M.; PRATT, J.; DING, Y.; WALDOR, M. K. Characterization of the Small Untranslated RNA RyhB and Its Regulon in *Vibrio cholerae*. **J. Bacteriol**, 187:4005–4014, 2005.

DEITEL, H. M.; DEITEL, P. J. Java: Como Programar. **Prentice-Hall**, 2005.

DEL VAL, C.; RIVAS, E.; TORRES-QUESADA, O.; TORO, N.; JIMÉNEZ-ZURDO, J. Identification of differentially expressed small non-coding RNAs in the legume endosymbiont *Sinorhizobium meliloti* by comparative genomics. **Mol. Microbiol.**, 66:1080–1091, 2007.

DELIHAS, N.; FORST, S. MicF: an antisense RNA gene involved in response of *Escherichia coli* to global stress factors. **J. Mol. Biol.**, 313:1-12, 2001.

DICHIARA, J. M.; CONTRERAS-MARTINEZ, L. M.; LIVNY, J.; SMITH, D.; MCDONOUGH, K. A.; BELFORT, M. Multiple small RNAs identified in *Mycobacterium bovis* BCG are also expressed in *Mycobacterium tuberculosis* and *Mycobacterium smegmatis*. **Nucleic Acids Res.**, 38:4067–4078, 2010.

EDDY, S. R. Computational Genomics of Noncoding RNA Genes. **Cell**, 19:137-140, 2002.

EPSHTEIN, V.; DUTTA, D.; WADE, J.; NUDLER, E. An allosteric mechanism of Rho-dependent transcription termination. **Nature**, 463:245-249, 2010.

ERMOLAEVA, M.D.; KHALAK, H.G.; WHITE, O.; SMITH, H.O. Prediction of transcription terminators in bacterial genomes. **J. Mol. Biol.**, 301:27–33, 2000.

ERMOLAEVA, M. D.; KHALAK, H. G.; WHITE, O.; SMITH, H. O.; SALZBERG, S. L. Prediction of transcription terminators in bacterial genomes. **J Mol Biol.**, 301:27-33, 2000.

FERREIRA, R. A. Guia prático Ubuntu 9.04. **Universo dos Livros**, 2009.

GONÇALVES, E. Dominando NetBeans. **Ciência Moderna**, 2006.

GONZÁLEZ, N.; HEEB, S.; VALVERDE, C.; KAY, E.; REIMMANN, C.; JUNIER, T.; HAAS, D. Genome-wide search reveals a novel GacA-regulated small RNA in *Pseudomonas* species. **BMC Genomics**, 9:167, 2008

GREGORY, R.; NICOL, J.; TAMM, H.; KULLMAN, B.; MURRAY, B.; KAPRAUN, D.; GREILHUBER, J.; BENNETT, M. Eukaryotic genome size databases. **Nucleic Acids Research**, 25:332-338, 2006.

GUSAROV, I.; NUDLER, E. The Mechanism of Intrinsic Transcription Termination. **Molecular Cell**, 3:495-504, 1999.

HENKIN, T. M. Riboswitch RNAs: using RNA to sense cellular metabolism. **Genes Dev.**, 22:3383-3390, 2008.

HENKIN, T. M.; Transcription termination control in bacteria. **Curr. Opin. Microbiol.**, 3:149-153, 2000.

HERSHBERG, R.; ALTUVIA, S.; MARGALITB, H. A survey of small RNA-encoding genes in *Escherichia coli*. **Nucleic Acids Res.**, 31:1813–1820, 2003.

HOON, M.J.; MAKITA, Y.; NAKAI, K.; MIYANO, S. Prediction of transcriptional terminators in *Bacillus subtilis* and related species. **PLoS Comput.**, 3:25, 2005.

HUANG, H.; CHANG, H.; CHOU, C.; TSENG, C.; HO, S.; YANG, C.; JU, Y.; HUANG, H. sRNAMap: genomic maps for small non-coding RNAs, their regulators and their targets in microbial genomes. **Nucleic Acids Res.**, 37:150-154, 2009

HÜTTENHOFER, A.; VOGEL, J. Experimental approaches to identify non-coding RNAs. **Nucleic Acids Res.**, 34:635-646, 2006.

JIANG, R.P.; TANG, D.J.; CHEN, X.L.; HE, Y.Q.; FENG, J.X.; JIANG, B.L.; LU, G.T.; LIN, M.; TANG, J.L. Identification of four novel small non-coding RNAs from *Xanthomonas campestris pathovar campestris*. **BMC Genomics**, 11:316, 2010.

KINGSFORD, C. L.; AYANBULE, K.; SALZBERG, S. L. Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. **Genome Biol.**, 8:R22, 2007.

KULKARNI, R. V.; KULKARNIA, P. R. Computational approaches for the discovery of bacterial small RNAs. **Methods**, 43:131-139, 2007.

KUMAR, R.; SHAH, P.; SWIATLO, E.; BURGESS, S. C.; LAWRENCE, M. L.; NANDURI, B. Identification of novel non-coding small RNAs from *Streptococcus pneumoniae* TIGR4 using high-resolution genome tiling arrays. **BMC Genomics**, 11:350, 2010.

LANDT, S. G.; ABELIUK, E.; MCGRATH, P. T.; LESLEY, J. A.; MCADAMS, H. H.; SHAPIRO, L. Small non-coding RNAs in *Caulobacter crescentus*. **Mol. Microbiol.**, 68:600-614, 2008.

LESNIK, E. A.; SAMPATH, R.; LEVENE, H. B.; HENDERSON, T. J.; MCNEIL, J. A.; ECKERA, D. J. Prediction of rho-independent transcriptional terminators in *Escherichia coli*. **Nucleic Acids Res.**, 29:3583–3594, 2001.

LEWIN, B. Genes IX. **Jones & Bartlett Publishers**, 2007.

LISSER, S.; MARGALIT, H. Compilation of *E. coli* mRNA promoter sequences. *Nucleic Acids Res.*, 21:1507-1516, 1993.

LIU, J. M.; CAMILLI, A. A broadening world of bacterial small RNAs. **Curr. Opin. Microbiol.**, 13:18-23, 2010.

LIU, J. M.; LIVNY, J.; LAWRENCE, M. S.; KIMBALL, M. D.; WALDOR, M. K.; CAMILLI, A. Experimental discovery of sRNAs in *Vibrio cholerae* by direct cloning, 5S/tRNA depletion and parallel sequencing. **Nucleic Acids Res.**, 37:e46, 2009.

LIVNY, J.; FOGEL, M. A.; DAVIS, B. M.; WALDOR, M. K. sRNAPredict: an integrative computational approach to identify sRNAs in bacterial genomes. **Nucleic Acids Res.**, 33:4096-4105, 2005.

LIVNY, J.; BRENCIC, A.; LORY, S.; WALDOR, M. K. Identification of 17 *Pseudomonas aeruginosa* sRNAs and prediction of sRNA-encoding genes in 10 diverse pathogens using the bioinformatic tool sRNAPredict2. **Nucleic Acids Research**, 34:3484-3493, 2006.

LIVNY, J.; TEONADI, H.; LIVNY, M.; WALDOR, M. K. High-Throughput, Kingdom-Wide Prediction and Annotation of Bacterial Non-Coding RNAs. **PLoS ONE**, 3:3197, 2008.

LIVNY, J.; WALDOR, M. K. Identification of small RNAs in diverse bacterial species. **Curr. Opin. Microbiol.**, 10:1096-1101, 2007.

LODISH, H.; BERK, A. Molecular Cell Biology. **W. H. Freeman**, 2007.

LUBAN, S.; KIHARA, D. Comparative Genomics of Small RNAs in Bacterial Genomes. **OMICS**, 11:58-73, 2007.

MACKE, T.J.; ECKER, D.J.; GUTELL, R.R.; GAUTHERET, D.A.; SAMPATH, R. RNAMotif, an RNA secondary structure definition and search algorithm. **Nucleic Acids Res.**, 29:4724-4735, 2001.

MAJDALANI, N; VANDERPOOL, C.K.; GOTTESMAN, S. Bacterial small RNA regulators. **Crit Rev Biochem Mol Biol.** 40:93-113, 2005.

MAKINO, K.; AMEMURA, M.; NAKATA, S.; SHINAGAWA, H. Role of the sigma 70 subunit of RNA polymerase in transcriptional activation by activator protein PhoB in *Escherichia coli*. **Genes Dev.**, 7:149-160, 1993.

MANDIN, P.; REPOILA, F.; VERGASSOLA, M.; GEISSMANN, T.; COSSART, P. Identification of new noncoding RNAs in *Listeria monocytogenes* and prediction of mRNA targets. **Nucleic Acids Res.**, 35:962–974, 2007.

MASSÉ, E.; GOTTESMAN, S. A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. **Proc Natl Acad Sci U S A.**, 7:4620-4625, 2002.

MASSÉ, E.; MAJDALANI, N.; GOTTESMAN, S. Regulatory roles for small RNAs in bacteria. **Curr Opin Microbiol.**, 6:120-124, 2003.

MCCUTCHEON, J. P.; EDDY, S. R. Computational identification of non-coding RNAs in *Saccharomyces cerevisiae* by comparative genomics. **Nucleic Acids Res.**, 31:4119–4128, 2003.

MITARAI, N.; BENJAMIN, J. M.; KRISHNA, S.; SEMSEY, S.; CSISZOVSKI, Z.; MASSÉ, E.; SNEPPENB, K. Dynamic features of gene expression control by small regulatory RNAs. **Proc Natl Acad Sci U S A**, 106:10655–10659, 2009.

MITRA, A.; ANGAMUTHU, K.; JAYASHREE, H.V.; NAGARAJA, V. Occurrence, divergence and evolution of intrinsic terminators across eubacteria. **Genomics**, 94:110-116, 2009.

MOULTON, V. Tracking down noncoding RNAs. **Proc. Natl Acad. Sci USA**, 102:2269-2270, 2005.

NAKAMURA, K.; MIYAMOTO, H.; SUZUMA, S.; SAKAMOTO, T.; KAWAI, G.; YAMANE, K. Minimal functional structure of *Escherichia coli* 4.5 S RNA required for binding to elongation factor G. **J Biol Chem.**, 276:22844-22849, 2001.

PAGET, M. S.B.; HELMANN, J. D. The $\sigma 70$ family of sigma factors. **Gen. Biol.**, 4:203, 2003.

PÁNEK, J.; BOBEK, J.; MIKULÍK, K.; BASLER, M.; VOHRADSKÝ, J. Biocomputational prediction of small non-coding RNAs in *Streptomyces*. **BMC Genomics**, 9:217, 2008.

PICHON, C.; FELDEN, B. Proteins that interact with bacterial small RNA regulators. **FEMS Microbiol. Rev.**, 31:614-625, 2007.

PICHON, C.; FELDEN, B. Small RNA gene identification and mRNA target predictions(REVIEW). **Bioinformatics**, 24:2807-2813, 2008.

PIERCE, B. Genetics - A Conceptual Approach. **The Guilford Press**, 2009.

RICHARDSON, J. P. Loading Rho to Terminate Transcription. **Cell**, 114:157-159, 2003.

RICHARDSON, J. P. Rho-dependent termination and ATPases in transcript termination. **Biochem. Biophys. Acta**, 1577:251-260, 2002.

RIVAS, E.; EDDY, S. R. Noncoding RNA gene detection using comparative sequence analysis. **BMC Bioinformatics**, 2:8, 2001.

RIVAS, E.; KLEIN, R. J.; JONES, T. A.; EDDY, S. R. Computational identification of noncoding RNAs in *E. coli* by comparative genomics. **Curr Biol.**, 4:1369-1373, 2001.

ROBERTS, J. W.; SHANKAR, S.; FILTER, J. J. RNA Polymerase Elongation Factors. **Annu. Rev. Microbiol.**, 62:211-233, 2008.

SAADE, J. Bash: Guia de Consulta Rápida. **Novatec**, 2001.

SÆTROM, P.; SNEVE, R.; KRISTIANSEN, K. I.; SNØVE, O.; GRÜNFELD, T.; ROGNES, T.; SEEBERG, E. Predicting non-coding RNA genes in *Escherichia coli* with boosted genetic programming. **Nucleic Acids Res.**, 33:3263–3270, 2005.

SAITO, S.; KAKESHITA, H.; NAKAMURA, K. Novel small RNA-encoding genes in the intergenic regions of *Bacillus subtilis*. **Gene**, 428:2-8, 2009.

SANTANGELO, T.J.; CUBONOVÁ, L.; SKINNER, K.M.; REEVE, J.N. Archaeal intrinsic transcription termination in vivo. **J. Bacteriol**, 191:7102-7108, 2009.

SCHLÜTER, J.; REINKENSMEIER, J.; DASCHKEY, S.; EVGUENIEVA-HACKENBERG, E.; JANSSEN, S.; JÄNICKE, S.; BECKER, J. D.; GIEGERICH, R.; BECKER, A. A genome-wide survey of sRNAs in the symbiotic nitrogen-fixing alpha-proteobacterium *Sinorhizobium meliloti*. **BMC Genomics**, 11:245, 2010.

SHARMA, C. M.; VOGEL, J. Experimental approaches for the discovery and characterization of regulatory small RNA. **Curr. Opin..Microbiol.**, 12:536-546, 2009.

SONNLEITNER, E.; SORGER-DOMENIGG, T.; MADEJ, M. J.; FINDEISS, S.; HACKERMÜLLER, J.; HÜTTENHOFER, A.; STADLER, P. F.; BLÄSI, U.; MOLL, I. Detection of small RNAs in *Pseudomonas aeruginosa* by RNomics and structure-based bioinformatic tools. **Microbiology**, 154:3175-3187, 2008.

SOPPA, J.; STRAUB, J.; BRENNEIS, M. Small RNAs of the halophilic archaeon *Haloferax volcanii*. **Biochem. Soc. Trans.**, 37:133-136, 2009.

SOREK, R.; KUNIN, V.; HUGENHOLTZ, P. CRISPR—a widespread system that provides acquired resistance against phages in bacteria and archaea. **Rev. Microbiol.**, 6:181-186, 2008.

SRIDHAR, J.; SOWMIYA, G.; SEKAR, K.; RAFI, Z.A. PsRNA: a computing engine for the comparative identification of putative small RNA locations within intergenic regions. **Genom. Proteom. Bioinformat.**, 8:127-134, 2010.

STRAUB, J.; BRENNEIS, M. M; JELLEN-RITTER, A.; HEYER, R.; SOPPA, J.; MARCHFELDER, ANITA. Small RNAs in Haloarchaea Identification, differential expression and biological function. **RNA Biol.**, 6:281-292, 2009.

SWIERCZ, J. P.; BOBEK, J.; HAISER, H. J.; BERARDO, C.; TJADEN, B.; ELLIOT, M. A. Small non-coding RNAs in *Streptomyces coelicolor*. **Nucleic Acids Res.**, 36:7240–7251, 2008.

TEZUKA, T.; HARA, H.; OHNISHI, Y.; HORINOUCI, S. Identification and Gene Disruption of Small Noncoding RNAs in *Streptomyces griseus*. **J. Bacteriol**, 191:4896–4904, 2009.

THÉBAULT, P.; GIVRY, S.; Searching RNA motifs and their intermolecular contacts with constraint networks, **Bioinformatics**, 22:2074-2080, 2006.

TJADEN, B. Prediction of small, noncoding RNAs in bacteria using heterogeneous data. **J. Math. Biol.**, 56:183-200, 2007.

TJADEN, B.; GOODWIN, S. S.; OPDYKE, J. A.; GUILLIER, M.; FU, D. X.; GOTTESMAN, S.; STORZ, G. Target prediction for small, noncoding RNAs in bacteria. **Nucleic Acids Res.**, 34:2791–2802, 2006.

TRAN, T. T.; ZHOU, F.; MARSHBURN, S.; STEAD, M.; KUSHNER, S. R.; XU, Y. De novo computational prediction of non-coding RNA genes in prokaryotic genomes. **Bioinformatics**, 15:2897–2905, 2009.

TSUI, T.; MUKHERJEE, D.; RAY, V. A.; SHAM, L.; FEIG, A. L.; WINKLER, M. E. Identification and Characterization of Noncoding Small RNAs in *Streptococcus pneumoniae* Serotype 2 Strain D39. **J. Bacteriol.**, 192:264–279, 2010.

ULVÉ, V. M.; SEVIN, E. W.; CHÉRON, A.; BARLOY-HUBLER, F. Identification of chromosomal alpha-proteobacterial small RNAs by comparative genome analysis and detection in *Sinorhizobium meliloti* strain 1021. **BMC Genomics**, 8:467, 2007.

UNNIRAMAN, S.; PRAKASH, R.; NAGARAJA, V. Conserved economics of transcription termination in eubacteria. **Nucleic Acids Res.**, 30:675-684, 2002.

VALVERDE, C.; LIVNY, J.; SCHLÜTER, J.; REINKENSMEIER, J.; BECKER, A.; PARISI, G. Prediction of *Sinorhizobium meliloti* sRNA genes and experimental detection in strain 2011. **BMC Genomics.**, 8:416, 2008.

VERCRUYSSSE, M.; FAUVART, M.; CLOOTS, L.; ENGELEN, K.; THIJS, I. M.; MARCHAL, K.; MICHIELS, J. Genome-wide detection of predicted non-coding RNAs in *Rhizobium etli* expressed during free-living and host-associated growth using a high-resolution tiling array. **BMC Genomics**, 11:53, 2010.

VOGEL, J.; BARTELS, V.; TANG, T.; CHURAKOV, G.; SLAGTER-JÄGER, J. G.; HÜTTENHOFER, A.; WAGNER, G. RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. **Nucleic Acids Res.**, 31:6435–6443, 2003.

VOGEL, J.; SHARMA C.M. How to find small non-coding RNAs in bacteria. **Biol. Chem.**, 386:1219-1238, 2005.

VOGEL, J.; WAGNER, G. Target identification of small noncoding RNAs in bacteria. **Curr. Opin. Microbiol.**, 10:262-270, 2007.

VON HIPPEL, P.H.; YAGER, T.D. Transcript elongation and termination are competitive kinetic processes. **Proc. Nat.l Acad. Sci. U S A.**, 88:2307-2311, 1991.

VOß, B.; GEORG, J.; SCHÖN, V.; UDE, S.; HESS, W. R. Biocomputational prediction of non-coding RNAs in model cyanobacteria. **BMC Genomics**, 10:123, 2009.

WANG, C.; DING, C.; MERAZ, R.F.; HOLBROOK, S.R. PSoL: a positive sample only learning algorithm for finding non-coding RNA genes. **Bioinformatics**, 22:2590-2596, 2006.

WASSARMAN, K. M.; REPOILA, F.; ROSENOW, C.; STORZ, G.; GOTTESMAN, S. Identification of novel small RNAs using comparative genomics and microarrays. **Genes Dev.**, 15:1637–1651, 2001.

WASSARMANN, K. M. 6S RNA: a small RNA regulator of transcription. **Curr. Opin. Microbiol.**, 10:164-168, 2007.

WEINBERG, Z; BARRICK, J. E.; YAO, Z.; ROTH, A.; KIM, J. N. Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline. **Nucleic Acids Res.**, 35:4809–4819, 2007.

WEISS, V. A. Estratégias de Finalização da Montagem do Genoma da Bactéria Diazotrófica Endofítica *Herbaspirillum seropedicae*. **Dissertação de Mestrado**. Universidade Federal do Paraná, 2010.

WILLKOMM, D. K.; MINNERUP, J.; HÜTTENHOFER, A.; HARTMANN, R. K. Experimental RNomics in *Aquifex aeolicus*: identification of small non-coding RNAs and the putative 6S RNA homolog. **Nucleic Acids Res.**, 33:1949–1960, 2005.

XIAO, B.; LI, W.; GUO, G.; LI, B.; LIU, Z.; JIA, K. Identification of Small Noncoding RNAs in *Helicobacter pylori* by a Bioinformatics-Based Approach. **Curr. Opin. Microbiol.**, 58:258-263, 2009.

YACHIE, N.; ARAKAWA, K.; TOMITA, M. On the interplay of gene positioning and the role of rho-independent terminators in *Escherichia coli*. **FEBS Lett**, 580:6909-6914, 2006.

YACHIE, N.; NUMATA, K.; SAITO, R.; KANAI, A.; TOMITA, M. Prediction of non-coding and antisense RNA genes in *Escherichia coli* with Gapped Markov Model. **Gene**, 372:171-181, 2006.

ZHANG, A.; WASSARMAN, K.M.; ROSENOW, C.; TJADEN, B.C.; STORZ, G.; GOTTESMAN, S. Global analysis of small RNA and mRNA targets of Hfq. **Mol. Microbiol**, 50:1111-1124, 2003.

ZHANG, Y.; ZHANG, Z.; LING, L.; SHI, B.; CHEN, R. Conservation analysis of small RNA genes in *Escherichia coli*. **Bioinformatics**. 20:599-603, 2004.

APÊNDICE I – CANDIDATOS COM SIMILARIDADE DE SEQUÊNCIA ENTRE ESPÉCIES PRÓXIMAS

Candidato 1

17 bp

TTGCAA**CGCACAAAGCGCTTTGCTATATT**CAATTCAACGGAAGTTTTTGCAGCACGTAA
CGATGTCTCCTCCACCCTCCTCAATGGTGTGGATTAACCCGGAACCACAGTGTCCGG
GTTT**TTTT**

Tamanho sRNA...: 97 bp

IgR.....: 1227

Tamanho da IgR.: 631 bp

Fita.....: Principal

Promotor.....: Posição: 232165 Algoritmo: BProm

Terminador.....: Posição: 232299 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Score.....: 180 bits (97)

Identidade.....: 131/145 (90%)

Blast hit 2

Espécie.....: *Burkholderia pseudomallei*

Score.....: 59.0 bits (64)

Identidade.....: 56/69 (82%)

Candidato 2

16 bp

TTGAAA**CGCGCAGAGTGCTGC**TAAAGTGCGGGCTTCGCTGAAACGGCAAACGGTCAT
TGCCGACCCAGCAATCGAATAGAGCAGTCTTGTTTTTGGTTCATTGTTGTTGATGTACGC
GCCTGGCCGGCGACCCTTCCGAATCATCGGAACGGCGCACCCGGGCGGCACCGGATA
CAAAGCGGGTGGTGCATAACAAACGGTATCGATACCGGGAAACTGAGGAGACGTGCAT
TCGATTGAAGCCGTTCAAACGGCCGGGAATGCATCAGGAGTATTTTGAAGGCGCATCTG
CGGATGCGCCTT**TTTT**

Tamanho sRNA...: 280 bp

IgR.....: 3796

Tamanho da IgR.: 499 bp

Fita.....: Reversa

Promotor.....: Posição: 742400 Algoritmo: BProm

Terminador.....: Posição: 742086 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 340 bits (184)

Identidade.....: 278/321 (86%)

Blast hit 2

Espécie.....: *Pseudomonas fluorescens*

Escore.....: 57.2 bits (62)

Identidade.....: 33/34 (98%)

Candidato 3

17 bp

TTGACGCTCTGCGCAAACGTCC**TAAATAT**TGACCGCATGATCGCCAACGCCACATTTTT
ACGACGCCCGCAGCCAACCTGAGAAAGGTTTGCTGCTGCGCGCGTGAAGTGAAACGAAT
CGCGAAGCCACAGGCAAACCCTGTGGCTT**TTTTT**

Tamanho sRNA...: 123 bp

IgR.....: 1672

Tamanho da IgR.: 418 bp

Fita.....: Reversa

Promotor.....: Posição: 319508 Algoritmo: BProm

Terminador.....: Posição: 319352 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 270 bits (146)

Identidade.....: 153/156 (98%)

Blast hit 2

Espécie.....: *Janthinobacterium sp. Marseille*

Escore.....: 50.0 bits (54)

Identidade.....: 29/30 (97%)

Candidato 4

20 bp

GTGTCAATGCATGGGGCGAAATTTGC**TATAAT**CGCCGTCCAGGCTGTGCGACCGCCGT
TGGGCGGCCCAATCAGGCAATCCCTACACCGAATCTGTAAGAGCGAGAAAGGCAAAC
CGGTTATGACACCGCGAACTACCACCGAGGTTTTACGCGACGCTAGTCGGGTTCTTTT
TCGTCCTTTGAAAAGCGCGGCTGGTCCGCGTTTTT**TTT**

Tamanho sRNA...: 182 bp

IgR.....: 2823

Tamanho da IgR.: 775 bp

Fita.....: Reversa

Promotor.....: Posição: 534223 Algoritmo: BProm
Terminador.....: Posição: 534003 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 235 bits (127)
Identidade.....: 127/127 (100%)

Blast hit 2

Espécie.....: *Janthinobacterium sp. Marseille*
Escore.....: 150 bits (166)
Identidade.....: 117/138 (85%)

Candidato 5

21 bp

TTGTCA**GCCTGTGCTGGCCGGCACTGG****CAA**ACTGCATCGGCGAAAAAGATCAGTAGAT
CCTTCAATCGAATTCGACCCGCAGCCAGTGCCAGGCTTGGCTGCACCCGCCCAATGCG
CGTACAATCCTCGCTCGAGCATGCGGACCACGGTCCGCA**TTTT**

Tamanho sRNA...: 127 bp
IgR.....: 1727
Tamanho da IgR.: 422 bp
Fita.....: Reversa

Promotor.....: Posição: 331340 Algoritmo: BProm
Terminador.....: Posição: 331173 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 100 bits (54)
Identidade.....: 72/80 (90%)

Candidato 6

17 bp

TTGATG**AAGTACACCGTTTTTTT****AGTCT**TGATCCGGGGAGGTGGTTGTACACGCTAG
ATGGCCGCCGCCCGAAAGGTTTTTACTGCGCAAAAACAAAACGCCTCATTCTTGCGA
ATGAGGCG**TTTT**

Tamanho sRNA...: 101
IgR.....: 2316
Tamanho da IgR.: 478 bp
Fita.....: Reversa

Promotor.....: Posição: 442630 Algoritmo: BProm
Terminador.....: Posição: 442492 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 100 bits (54)
Identidade.....: 58/60 (96%)

Candidato 7

18 bp

TTGCCACCAAACATGTGGATAGTGATCATCAGGCCATTGTCTGAGGCGGCATTTGTCC
CTTTCCGGCAGCGGCGGCACGTGGTTGGCAGGGTGTTGCCCATGGTCTGAAGCGGAAA
TCAAAGTTGTTTCAAATCAATGTCTTGATGAGCTTGTCAAGACGTTCTCCACAGGCTTG
CCACCAGAAAATGTGGATAGTTCGTCTCAGCGCATGTCAGGAGGGGCCATGGTTCCTTCC
TGCCTG**TTTTT**

Tamanho sRNA...: 215 bp
IgR.....: 1818
Tamanho da IgR.: 775 bp
Fita.....: Principal

Promotor.....: Posição: 348567 Algoritmo: BProm
Terminador.....: Posição: 348820 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 113 bits (61)
Identidade.....: 72/77 (93%)

Blast hit 2

Espécie.....: *Burkholderia phytofirmans*
Escore.....: 48.2 bits (52)
Identidade.....: 32/36 (89%)

Candidato 8

19 bp

ATGCCGTTTTCTTTGGCGGGTGCGTATGATGAAGAACCTGCACGACCGATCGGCATCG
GTCTTGTTGGTGCAGAAAACGTCAGGCGCGGTGCAGGAAACATCAGGCTTCCCAGCTC
GCAGGAAGTGACCGGAGGGCCTGCACCAAAGCAAGAAAAATTAGAAAACCAGACCGT
CGTGGCAGACCGGTGGAACACTCCGGCAGCACGGGAGAACGGCAAACAACACTCG
GAGACACGCTCGCAAGAGCGGCAGGAGGCCATGATGGCTCACGCCATGCGGGCGC
AAAGGCGCGCTGCTGCAGCG**TTTTT**

Tamanho sRNA...: 284

IgR.....: 552
Tamanho da IgR.: 575 bp
Fita.....: Reversa

Promotor.....: Posição: 100359 Algoritmo: BProm
Terminador.....: Posição: 100035 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 335 bits (181)
Identidade.....: 287/334 (85%)

Candidato 9

18 bp

TTCAAATAGATGATG**TTTGACGT**TATCTTGGAAATATTTTCGAGCTATTTATCGTTATCTT
CGGAAAATTGCCTCCTTTGAATTAAGGATGAACAACCTTCCCTCAAGATACTTTACCTT
AGCTGCGGGTGATGCATTCACCCCGCTTCCCTTCCCTCCCTGAGCGGGGCCGATC
GTGCAAGCGAAAGGCATTCGGCCCCGGAGAATATCCCTGTTCTCCGGGG**TTTTTTT**

Tamanho sRNA....: 206 bp
IgR.....: 276
Tamanho da IgR.: 290 bp
Fita.....: Reversa

Promotor.....: Posição: 49547 Algoritmo: BProm
Terminador.....: Posição: 49307 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 211 bits (114)
Identidade.....: 197/237 (83%)

Blast hit 2

Espécie.....: *Herminiimonas arsenicoxydans*
Escore.....: 59.0 bits (64)
Identidade.....: 72/92 (79%)

Candidato 10

19 bp

TTCAAGCAAGACGTGAATGTAACAT**CCAGT**AAAAGCCATTGGACCAGCTTCGGCACT
GACGTTTAATAGACCTCACTTGAACAGCGATTGATTGGTATCCGTTACTGTGCAAGGTTT
CTCCCCTCTAGTGTTGAAAGCACCACTAGCTCGCGGATCTGGCTCCCCCTGTCAGATCC
GC**TTTTTTT**

Tamanho sRNA...: 156 bp
IgR.....: 867
Tamanho da IgR.: 497 bp
Fita.....: Reversa

Promotor.....: Posição: 159165 Algoritmo: BProm
Terminador.....: Posição: 158974 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 285 bits (154)
Identidade.....: 178/189 (94%)

Candidato 11

18 bp

GTGAAA**AAAATATTTAGAGAATTTATTCT**GATTATAAAGGCAATAAACGCGCCTAATAT
CAAGGAAGAGTCAATAGAATCCTCCCTCATAAAAAACATAACGTAATGCGGGCAGCGGG
CTCAATCGAGCGCCTCTGCACGGATGACAGCAGTAACACCGGCAATTCTGTTGGCCCA
TGTTGTTTTGCCATGCCCTGCCTCTCACAAGGAATGCAGGGTGCGCAAATCCTGC**TTT**
T

Tamanho sRNA...: 207 bp
IgR.....: 255
Tamanho da IgR.: 316 bp
Fita.....: Reversa

Promotor.....: Posição: 45647 Algoritmo: BProm
Terminador.....: Posição: 45402 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 178 bits (96)
Identidade.....: 108/114 (94%)

Candidato 12

18 bp

ACGAAA**CATTTCGGGCCGGGATCGTAAATT**CGGCCAAAATGCTGGCATTTTGGCCACAA
AACCCGGCCCCTTATAGGGG**TTT**

Tamanho sRNA...: 53 bp
IgR.....: 1243
Tamanho da IgR.: 278 bp
Fita.....: Reversa

Promotor.....: Posição: 235035 Algoritmo: BProm
Terminador.....: Posição: 234940 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 99.0 bits (53)
Identidade.....: 73/83 (87%)

Candidato 13

19 bp

TAGGCA**TTAATCCGACATTCTACCTTATAAT**TCCCGAGAAACCAGCAAGCCGGTTTGCC
AAAAAATCTGGCATTTCGGGGACCTCCCGGCCGCCATCCGGCGCGGGTTTGCCG**TTTTT**

Tamanho sRNA...: 86 bp
IgR.....: 542
Tamanho da IgR.: 357 bp
Fita.....: Principal

Promotor.....: Posição: 98073 Algoritmo: BProm
Terminador.....: Posição: 98196 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 145 bits (78)
Identidade.....: 108/122 (88%)

Candidato 14

18 bp

CTGCCG**GTTGCTTCCCTTCGCTGG****CATTAT**CCAGATCAGGTTCAAGGGTATTTCTCACC
CGGCTACCCAAGCATGAGCTTGAGCGCCAGGACCCCTAGCGGTGCGACGCTGGCTGC
TTGCGCAGTCCCTTGTTCGTGGGCCGATTCTACTCCAATTGGCGTGAGGAACAAGGTTTC
ATTTCCGATGGGTTGAGTGTGCAGCGCAACATTCCGGACCCGATTCTAGAAAGTTGCAG
GAACGCAGAGGACCGTTTTACGCTCGAAAGTGATGTCCTCGCAACCGAACAGTGGAG
AACGCCGCATGAATCAACAAGGGAAAGCTCCTTCGTCTCTGGCTGAAGGATTTCA**TTT**
T

Tamanho sRNA...: 322 bp
IgR.....: 42
Tamanho da IgR.: 511 bp
Fita.....: Principal

Promotor.....: Posição: 10798 Algoritmo: BProm
Terminador.....: Posição: 11157 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 305 bits (165)
Identidade.....: 206/225 (91%)

Blast hit 2

Espécie.....: *Janthinobacterium sp. Marseille*
Escore.....: 62.6 bits (68)
Identidade.....: 36/37 (98%)

Candidato 15

16 bp

ATGAATAGCCAAAACCATGCCATACAATATTTTTCTTTTAAATCAATGGGATAGCGACGA
TTTCTGATTGCTGCGCCAGTGGCGATGATTGTCGTAATCCGACAAAGACGACACCGTT
GTCGCCAACCCGCCAAGGCGACATCGAGGAAAGCCGGCATTGTGTGGGGTTGCTTGG
CGAGATCGATTCATGCATCTGCTTGGGCATGAAGTTTGCTTGACCAGCAAGGTA ACTAT
TCGGTAATTACACGGCGCTTTTTCTGCCG**TTTT**

Tamanho sRNA...: 241 bp
IgR.....: 2077
Tamanho da IgR.: 473 bp
Fita.....: Principal

Promotor.....: Posição: 401580 Algoritmo: BProm
Terminador.....: Posição: 401860 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 490 bits (265)
Identidade.....: 275/280 (98%)

Candidato 16

18 bp

CTGTTAAATAATTATTTATCTAATTATTTCTGGTTTGTTGTTTTGCTTGGGGTGGGCTAT
TTGGGCTTGGGTTTTTTTTGTTGCGCTCCTGGGTAGTCTGCACGGGATTTCTTCCTGCGC
TCTAGCTTAATTCTTATCTTTCTCCGGTCGGAAGGACGCGCCGGGGGCGGCCCGGCAG
CCGCTCAC**TTTT**

Tamanho sRNA...: 160 bp
IgR.....: 884
Tamanho da IgR.: 1217 bp
Fita.....: Principal

Promotor.....: Posição: 162503 Algoritmo: BProm
Terminador.....: Posição: 162702 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 111 bits (60)
Identidade.....: 92/106 (86%)

Blast hit 2

Espécie.....: *Janthinobacterium sp. Marseille*
Escore.....: 71.6 bits (78)
Identidade.....: 39/39 (100%)

Candidato 17

18 bp

TTGACC**GAATTCATTGGCTTTAAC****AAAAAT**TTCCCATTCAGCCGCCAAATCATCAAAA
CGCCTCGCCCTGTGGATAAAGTTCTGAATAACTTTTTTCATCGATTGGGGAAAACTCGAAA
AGTTATCCAAAGGCGAATTTTTATCCACGATTTGGAATCACGTCATACAGGCCTTGCCCA
AGGATTTATGCGTTCGCGCAAGCGACTGA**TTTTT**

Tamanho sRNA...: 184 bp
IgR.....: 415
Tamanho da IgR.: 783 bp
Fita.....: Principal

Promotor.....: Posição: 74058 Algoritmo: BProm
Terminador.....: Posição: 74281 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 147 bits (79)
Identidade.....: 107/119 (89%)

Candidato 18

17 bp

TTGTCT**CCTCCTCCCTCTTTAGG****AAGGAT**TAACCCCGCAGGCCATGCCTGTGGGGTTTT
TTTTTCGTCCCTGCGCCAGGCGCCCTGTGCCCCCTGTGCCCCATGCTTCCGGGGGCATC
CGGTCCAGATATTCACCGGCGTTGCCGCCAGCGTCAAATCGCGTTG**TTTTT**

Tamanho sRNA...: 140 bp
IgR.....: 413
Tamanho da IgR.: 1006 bp
Fita.....: Reversa

Promotor.....: Posição: 73186 Algoritmo: BProm
Terminador.....: Posição: 73010 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 111 bits (60)
Identidade.....: 70/74 (94%)

Candidato 19

TTGCAACGACGGTTCGCGCATAGAACTCATTTCATAAATAGGCGTGAGATGCGTTCTGCC
CAGAAAGGGCGCTGGCAAGGCGCGCGACGCGTCGTGTGGCGGTGCCACACGCAAGG
AGCGCAACGCGGCCAGCGCCCTTTCTGGGCAGAACCCGGAGGGAGCGGCCCGTTTGG
GCGAATTGCTGCGTTACGAATTTGGGTCAAGACGCCAGTCTTGACCCAAATTCGCGCC
TTGCACTTCATCCCAAACGGGACTCGCTCGCACTCACGCCTATTTATGAAATGAGTTCTA
GGCTCTTCTTCAGCGACTCTGGCAAGACTGACAGTAGTCGCCACCCCTCAGTCGAGCG
CTGCACCGATATGGCTGGCGCAAGGCTATCCCTAAGATGAGGACTCGTAGGGCAGTGA
CGGCCCTGCGGTTTTT

Tamanho sRNA...: 386 bp
IgR.....: 4050
Tamanho da IgR.: 612 bp
Fita.....: Principal

Promotor.....: Posição: 790245 Algoritmo: BProm
Terminador.....: Posição: 790679 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 484 bits (262)
Identidade.....: 268/271 (98%)

Blast hit 2

Espécie.....: *Janthinobacterium sp. Marseille*
Escore.....: 185 bits (204)
Identidade.....: 203/266 (77%)

Candidato 20

17 bp

TTACGACATCAAAAAATGAATAATATCATGTGTTTCAGGTCGTACAGGCCCTTGTTCCA
TGGGGCTTGCAATGAAGTTAGCCGCCATTCTGCCACACGGATTGCCCTTGAGCGATGCT
TGCCAATTGGTCCAGACAACGTTTTTGGGACTGTGTTGTTTCTTCGCAACGGGATATTTT
CGGGCTGTTTTTGTTCATCAGGGGCCGAAAACTTCGCTTTTTTTAAGAATATCCTTAT
TTTTCATACAGTTCGGCCATTTTCGTTGCGGCAAGTAGTTGCCTTTTT

Tamanho sRNA...: 257 bp
IgR.....: 1353
Tamanho da IgR.: 972 bp
Fita.....: Principal

Promotor.....: Posição: 256989 Algoritmo: BProm
Terminador.....: Posição: 257285 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 398 bits (215)
Identidade.....: 271/297 (91%)

Candidato 21

CCGCTTATATCGCCGCACCGGCGCGGGCAGTAATCTGGTCGTTACAGGCATACGCGTA
CAGCGCGTCGTTTCCATTGAAAAGACATACCCCATATCCAAAGGAGGTCAGCATGAGCG
CAAGCAAACCCGACAAGGACCATCGTCCCAACAAGTACGACACGCCCTCGGAAGTGCC
GCTGGATGAAGCCAACGAGATGAGCTTCCCCGGCAGCGATCCGGTGGCCACCTCCAAT
ATCACGCGCATCGAAAAAGCGCCTGAGATGGATCGCGCCAGCAATGACCACCAGAACC
GCAACCAGGTCCGCGAGAACACCCGACAGAACACCCCTGACTGAGCCTAGTCCGCCCC
ATCCAGGGCGGCCGGTGCATGCCGGCCGCCCTTTT

Tamanho sRNA...: 349 bp
IgR.....: 2572
Tamanho da IgR.: 453 bp
Fita.....: Principal

Promotor.....: Posição: 489905 Algoritmo: BProm
Terminador.....: Posição: 490297 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 226 bits (122)
Identidade.....: 215/256 (83%)

Candidato 22

TTTTTTGTTTGGTGGCCCCTAGCGATGTGCTTGAGCACATCGCGGGGAGGATTCGAAAG
GTTTCGCTTGCAAGCGAAACCGGGCCCGCGGCACGTGGGCGAATCCCTCGCTCGCGC
AGCGAGCGTTAGAGGGTGGTGGAGCGCGCGCATGACGCAATCCAAGCCAAGCCAGCC
GATCGGAAGCCCGCGCAAGCGGGCTTTTT

Tamanho sRNA...: 173 bp
IgR.....: 1902

Tamanho da IgR.: 502 bp

Fita.....: Principal

Promotor.....: Posição: 363738 Algoritmo: BProm

Terminador.....: Posição: 363947 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 196 bits (106)

Identidade.....: 117/122 (95%)

Blast hit 2

Espécie.....: *Janthinobacterium sp. Marseille*

Escore.....: 66.2 bits (72)

Identidade.....: 56/67 (84%)

Candidato 23

18 bp

CGGGCA**GCAAGGCAGAGCCGATGCGATAAT**GCATCGGCTTTTTTATTGCCCGGCTTGC
TTGCGATCATGGAGCCATCGGGCCGCAGGCCGCACAATGGCTCGCATTGCCGTTTTG
CTGCCCTGCAAAAC**TTTT**

Tamanho sRNA...: 105 bp

IgR.....: 3935

Tamanho da IgR.: 219 bp

Fita.....: Principal

Promotor.....: Posição: 769947 Algoritmo: BProm

Terminador.....: Posição: 770088 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 176 bits (95)

Identidade.....: 119/130 (91%)

Candidato 24

GACGGTGACAGGATGTGGTCTGAAGCGAATCGATTGGAAAACGCCGCGTTTTTTATCGG
AACACATTAGCAAAGGCAGGCGCCTGGCATTCCGAACAGAAAGGGCGCAGGCAAGGC
CTATTTGATGAATATGTGTCTCGTGGCAGGCTGCGGCATATGCCGTGGCATTTCGGAA
GAAACATCAGGTGGCTGGGTTTCATCGGCCTTGC GCGCGGCTCTGGAAGCAGGTCTGCG
CAAACGCGAGCCAGCCGTCGACGCGGACTCACCTTCCTGACTGACGTCCAATATCCA
GCGCGCGGACGGTATGGCGTCCGGCTTA**TTTTT**

Tamanho sRNA...: 293 bp

IgR.....: 2564
Tamanho da IgR.: 465 bp
Fita.....: Principal

Promotor.....: Posição: 487844 Algoritmo: BProm
Terminador.....: Posição: 488170 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 169 bits (91)
Identidade.....: 181/222 (81%)

Candidato 25

19 bp

TAGGTTTGTCTGTATTCTATTTTCAGTATAATAATTTTGTATAAGCTAATGCGCCGCTAGC
GCCAAGGCCGGGAAGCCTTGCCGCGCGGCC**TTT**

Tamanho sRNA...: 63 bp
IgR.....: 3174
Tamanho da IgR.: 278 bp
Fita.....: Principal

Promotor.....: Posição: 597164 Algoritmo: NNPP
Terminador.....: Posição: 597280 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 99.0 bits (53)
Identidade.....: 60/63 (95%)

Candidato 26

17 bp

TTGACGCTCTGCGCAAACGTCCTAATATTGACCGCATGATCGCCAACGCCACATTTTT
ACGACGCCCGCAGCCAACCTGAGAAAGGTTTGCTGCTGCGCGCGTGAAGTGAAACGAAT
CGCGAAGCCACAGGCAAACCCTGTGGC**TTTTTTT**

Tamanho sRNA...: 123 bp
IgR.....: 1672
Tamanho da IgR.: 418 bp
Fita.....: Reversa

Promotor.....: Posição: 319519 Algoritmo: NNPP
Terminador.....: Posição: 319352 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 291 bits (157)

Identidade.....: 164/167 (98%)

Blast hit 2

Espécie.....: *Janthinobacterium sp. Marseille*

Escore.....: 50.0 bits (54)

Identidade.....: 29/30 (97%)

Candidato 27

17 bp

TGGACA**GTATTAAGCGATTGCATTAAACT**ATCGTTTGTCTTCGCGACATTACTTGTTCAA
CAACGACTCTCAGCAGTGTCCGATCCGCGCTTGTGCGGGTTGTCCAAATTTCCGGCTG
GCAGCGTGAACGCCCTCCCGTTCTCCACGGTGACCCGTTCCCATCCCTGTCCGGTATG
CACTATGCCGAAGCTGGCTTCATCGCTTCGGTTTCATCGCTGCCCTTGCAGC**TTT**

Tamanho sRNA...: 202 bp

IgR.....: 3047

Tamanho da IgR.: 580 bp

Fita.....: Reversa

Promotor.....: Posição: 575588 Algoritmo: NNPP

Terminador.....: Posição: 575337 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 350 bits (189)

Identidade.....: 231/250 (92%)

Candidato 28

17 bp

TTGGCA**ACGGATGTTACAGCGTG****TAAAAA**GCCTTGAGGTGGCGGCGGGGGATGGTTAT
TCTGAAGGTCTCTCCCCTCCAATTCATCACCCCTGAGAATTGGATTTAACCCGCCGCCG
CAAGGTGCGCGGG**TTTTTTT**

Tamanho sRNA...: 108 bp

IgR.....: 817

Tamanho da IgR.: 289 bp

Fita.....: Principal

Promotor.....: Posição: 150084 Algoritmo: NNPP

Terminador.....: Posição: 150232 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 211 bits (114)

Identidade.....: 133/142 (93%)

Blast hit 2

Espécie.....: *Janthinobacterium sp. Marseille*

Escore.....: 53.6 bits (58)

Identidade.....: 55/71 (78%)

Candidato 29

CGTTATCTTGGAATATTTTCGAGCTATTTATCGTTATCTTCGGAAAATTGCCTCCTTTGAA
TTAAAGGATGAACAACCTTCCCTCAAGATACTTTACCTTAGCTGCGGGTGATGCATTAC
CCCCGCTTCCCCTTCCCTCCCTGAGCGGGGCCGATCGTGCAAGCGAAAGGCATTCGG
CCCCGGAGAATATCCCTGTTCTCCGGGG **TTTTTTT**

Tamanho sRNA...: 186 bp

IgR.....: 276

Tamanho da IgR.: 290 bp

Fita.....: Reversa

Promotor.....: Posição: 49526 Algoritmo: NNPP

Terminador.....: Posição: 49307 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 195 bits (105)

Identidade.....: 180/216 (83%)

Blast hit 2

Espécie.....: *Herminiimonas arsenicoxydans*

Escore.....: 59.0 bits (64)

Identidade.....: 72/92 (79%)

Candidato 30

18 bp

TTGGCC **TGATTGATGCGGAGCAGTAAAGT** AACACCTGTCTCCTCCAATTTCTCCTAA
AGAATTGGATTCAAGCCACACCGCAAGGTTGTGGGC **TTTTTTT**

Tamanho sRNA...: 74 bp

IgR.....: 2291

Tamanho da IgR.: 253 bp

Fita.....: Reversa

Promotor.....: Posição: 437136 Algoritmo: NNPP
Terminador.....: Posição: 437018 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 171 bits (92)
Identidade.....: 109/117 (93%)

Blast hit 2

Espécie.....: *Janthinobacterium sp. Marseille*
Escore.....: 59.0 bits (64)
Identidade.....: 69/89 (78%)

Candidato 31

19 bp

TTGTTGAGAGATCCGTCTGGTCGGATAAACCAGGCCGGGTCGGTCCGGTCATTCGGTAC
GCGATGCCCGCCGGCAGGTGATGCAGGAAAACCCATCATATTGCCTGCGCTGCAACCCC
TTCCCGGCCCTTGGGGCAGGATGTTGGAGGCGGAACAGGCTTAAATGTACGGTCTTCC
AGCCGTGTGGACGACGAATTTTATCAGGGAATTTTCAATTTTCAAATTACGATAATCGAT
TTCATAATACGAAAATATGCTGCAAAGCGGCATAAACAGGGGCTTCCAGGCCTCGTA**T**
TTT

Tamanho sRNA...: 266 bp
IgR.....: 3270
Tamanho da IgR.: 555 bp
Fita.....: Principal

Promotor.....: Posição: 618211 Algoritmo: NNPP
Terminador.....: Posição: 618517 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 361 bits (195)
Identidade.....: 271/307 (88%)

Blast hit 2

Espécie.....: *Herminiimonas arsenicoxydans*
Escore.....: 46.4 bits (50)
Identidade.....: 47/58 (82%)

Candidato 32

19 bp

TTGACGAAAGTCAGTCGGCCTTTGTATTTTGGGGTCCACGAGCAAGGCGCCTGCGCG
CCTTGCGAGAGGGACGAGAAAGTCCGGCCTTACAAGGCCGGACGGGGTTCGCGGGACG

TGACCGAGTGCCTCAGTGCGCAGCAAAACAAAAAAGGTCGATTGATGAAAATCAATCGG
CCTTTTGCGTTT

Tamanho sRNA...: 156 bp
IgR.....: 2922
Tamanho da IgR.: 520 bp
Fita.....: Principal

Promotor.....: Posição: 553127 Algoritmo: NNPP
Terminador.....: Posição: 553319 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 265 bits (143)
Identidade.....: 178/194 (91%)

Blast hit 2

Espécie.....: *Burkholderia multivorans*
Escore.....: 60.8 bits (66)
Identidade.....: 69/89 (78%)

Candidato 33

18 bp

GTGAAA AAAATATTTAGAGAATTTATTCTGATTATAAAGGCAATAAACGCGCCTAATAT
CAAGGAAGAGTCAATAGAATCCTCCCTCATAAAAAACATAACGTAATGCGGGCAGCGGG
CTCAATCGAGCGCCTCTGCACGGATGACAGCAGTAACACCGGCAATTCTGTTGGCCCA
TGTTGTTTTGCCATGCCCTGCCTCTACAAGGAATGCAGGGTGGCGCAAATCCTGC
TTT

Tamanho sRNA...: 207 bp
IgR.....: 255
Tamanho da IgR.: 316 bp
Fita.....: Reversa

Promotor.....: Posição: 45660 Algoritmo: NNPP
Terminador.....: Posição: 45402 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 202 bits (109)
Identidade.....: 121/127 (95%)

Candidato 34

20 bp

TTGATGCGCGACTTCTCGCGTGCCGAAAAAATTTTAGGAAGTCTTTACAGATTGTGGCG
CTGCATTTTCTGTTGACGGTTCATGCGATACATTTGATCGGGCCTTAAGTTCAAATTACT
TAAAACTATTAAGTCGAGAAAAAATGATAGTTGTAGTTGGAACGTGGCTAGTAATAGCTT
CCGGGTTTTCAGAAGCATCAGTTTCACTCCGAAAAGTACCGTCTCGGACTATTCCGAGTC
TAAAAATCAGGAATCTCCGGAATTTTTCGGTCTCGACAAATTCTCTTGCGCGAGCTCGC
GCAAGA**TTTT**

Tamanho sRNA...: 275 bp

IgR.....: 1065

Tamanho da IgR.: 365 bp

Fita.....: Principal

Promotor.....: Posição: 196212 Algoritmo: NNPP

Terminador.....: Posição: 196530 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 195 bits (105)

Identidade.....: 258/327 (78%)

Candidato 35

17 bp

TTCACTGGCGCTCTAAAAAGTAATGAAAT****GTACGGTAACAACGATTACTCAACTTTTCTAT
TACATTTTACAGCGCAGCGGAAGTTACGGCCGGGCATCCAGAGAACAGGCAAGACCCATG
GTGAGGAATTGATATCGCGCAAGGTATCTATCGTCCAACACCTGCATACTGTTTCCCAG
CAGAATACGTTGCTG**TTTT**

Tamanho sRNA...: 168 bp

IgR.....: 775

Tamanho da IgR.: 479 bp

Fita.....: Reversa

Promotor.....: Posição: 141905 Algoritmo: NNPP

Terminador.....: Posição: 141690 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 104 bits (56)

Identidade.....: 73/81 (90%)

Candidato 36

19 bp

TTTAGATTGCTACATTGCTGGACAGTATTAAGCGATTGCATTAAACTATCGTTTGTCTTC
GCGACATTACTTGTTCAACAACGACTCTCAGCAGTGTCCGATCCGCGCTTGTGCGGGTT
GTCCAAATTTCCGGCTGGCAGCGTGAACGCCCTCCCGTTCTCCACGGTGACCCGTTCC
CATCCCTGTCCGGTATGCACTATGCCGAAGCTGGCTTCATCGCTTCGGTTTCATCGCTG
CCCTTGCAGC**TT**

Tamanho sRNA...: 218 bp

IgR.....: 3047

Tamanho da IgR.: 580 bp

Fita.....: Reversa

Promotor.....: Posição: 575604 Algoritmo: NNPP

Terminador.....: Posição: 575337 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Score.....: 372 bits (201)

Identidade.....: 249/271 (91%)

Candidato 37

17 bp

TTGCATCCTCCCTAACCCGCCACTAGAATTAGTCCCAATCGATGTTTTGATACTAGATGT
AGTGTCCAGGTGAACTTATTTTTGTACACGGGTCTTCATCCACAGATGGCCCGGCTGGC
AGTGCTTCAGAAGGCGCAAGATCCCGCGAGGCAGTCACCAAGTACCAAGCAACAAGTA
CGAAAGCGCCCCGGTTCTCCCTTCTTATTCGGTCGTTTCGCTCATCAGTTTCAAGCAGTT
TTTGCCTCAGCATGAATATCCATGCAAGGGATAGTCATCGTGGGACGCAGTCGTCGTTT
GCCAATCCCGTGTAGTTCGGGATTCTGA**TT**

Tamanho sRNA...: 298 bp

IgR.....: 83

Tamanho da IgR.: 405 bp

Fita.....: Reversa

Promotor.....: Posição: 17770 Algoritmo: NNPP

Terminador.....: Posição: 17429 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Score.....: 274 bits (148)

Identidade.....: 278/336 (82%)

Candidato 38

19 bp

TTGACGCGAGTCAGTCGGCCTTTGTATTTTGGGGACCACGAGCAAGGCGCCTGCGCG
CCTTGCGGGAGGGACGAGAAAGTCCGGCCTTGCAAGGCCGGACGGGGTTCGAGGGAC
GTGACCGAGTCCCTCCGTGGGCACCAAATAGAAAAGGTTCGATTGACGCGAGTCAGTC
GGCCTTTTGTATTTTT

Tamanho sRNA...: 157 bp

IgR.....: 2029

Tamanho da IgR.: 276 bp

Fita.....: Reversa

Promotor.....: Posição: 389650 Algoritmo: NNPP

Terminador.....: Posição: 389453 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 267 bits (144)

Identidade.....: 180/197 (91%)

Blast hit 2

Espécie.....: *Burkholderia ambifaria*

Escore.....: 75.2 bits (82)

Identidade.....: 90/116 (78%)

Candidato 39

20 bp

TGGAATATTTTCGAGCTATTTATCGTATCTTCGGAAAATTGCCTCCTTTGAATTAAGG
ATGAACAACCTTTCCCTCAAGATACTTTACCTTAGCTGCGGGTGATGCATTCACCCCCCG
CTTCCCCTTCCCTCCCTGAGCGGGGCCGATCGTGCAAGCGAAAGGCATTCGGCCCCG
GAGAATATCCCTGTTCTCCGGGGTTTTTTT

Tamanho sRNA...: 175 bp

IgR.....: 276

Tamanho da IgR.: 290 bp

Fita.....: Reversa

Promotor.....: Posição: 49518 Algoritmo: NNPP

Terminador.....: Posição: 49307 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 185 bits (100)

Identidade.....: 173/208 (83%)

Blast hit 2

Espécie.....: *Herminiimonas arsenicoxydans*
Escore.....: 59.0 bits (64)
Identidade.....: 72/92 (79%)

Candidato 40

20 bp

CTGATTATAAAGGCAATAAACGCGCCTAATATCAAGGAAGAGTCAATAGAATCCTCCCT
CATAAAAAACATAACGTAATGCGGGCAGCGGGCTCAATCGAGCGCCTCTGCACGGATG
ACAGCAGTAACACCGGCAATTCTGTTGGCCCATGTTGTTTTGCCATGCCCTGCCTCTCA
CAAGGAATGCAGGGTGGCGCAAATCCTGC**TTTT**

Tamanho sRNA...: 177 bp
IgR.....: 255
Tamanho da IgR.: 316 bp
Fita.....: Reversa

Promotor.....: Posição: 45627 Algoritmo: NNPP
Terminador.....: Posição: 45402 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 158 bits (85)
Identidade.....: 91/94 (96%)

Candidato 41

ATGTACGGTGTGGCCTGATTTGCTTGACCAAATGATGCAACGCATTACTGTCCTGT
TTCCCGCCTCCACATCTGGGGCG**TTTTT**

Tamanho sRNA...: 59 bp
IgR.....: 3027
Tamanho da IgR.: 545 bp
Fita.....: Reversa

Promotor.....: Posição: 572506 Algoritmo: NNPP
Terminador.....: Posição: 572410 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 139 bits (75)
Identidade.....: 75/75 (100%)

Candidato 42

TCTTTTAAATCAATGGGATAGCGACGATTTCTGATTGCTGCGCCAGTGGCGATGATTGT
CGTAAATCCGACAAAGACGACACCGTTGTCGCCAACCCGCCAAGGCGACATCGAGGAA
AGCCGGCATTGTGTGGGGTTGCTTGGCGAGATCGATTCATGCATCTGCTTGGGCATGA
AGTTTGCTTGACCAGCAAGGTAAGTATTTCGGTAATTACACGGCGC**TTTT**

Tamanho sRNA...: 195 bp
IgR.....: 2077
Tamanho da IgR.: 473 bp
Fita.....: Principal

Promotor.....: Posição: 401617 Algoritmo: NNPP
Terminador.....: Posição: 401860 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 422 bits (228)
Identidade.....: 238/243 (97%)

Candidato 43

19 bp

GTGACTCATGCGGGAGCATCCGGCGTATCGTAGTCCGGTAGTCCCATAGTCCCATCAT
GCAAGCAAGCCCTCAACCAAGACCAGACCGAAAAGGCGTCGGGCGCTAGGAGACGGA
TGGCGCGGCCTGTGTTTTACCGGCTTTTTGTCTCTTGCATAGCAGCGATTGCAGTGGT
TGCAG**TTTT**

Tamanho sRNA...: 153 bp
IgR.....: 1111
Tamanho da IgR.: 315 bp
Fita.....: Principal

Promotor.....: Posição: 206295 Algoritmo: NNPP
Terminador.....: Posição: 206489 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 182 bits (98)
Identidade.....: 149/172 (86%)

Candidato 44

19 bp

GTGACTCATGCGGGAGCATCCGGCGTATCGTAGTCCGGTAGTCCCATAGTCCCATCAT
GCAAGCAAGCCCTCAACCAAGACCAGACCGAAAAGGCGTCCGGCGCTAGGAGACGGA
TGGCGCGGCCTGTGTTTTACCGGC**TTTT**

Tamanho sRNA...: 113 bp

IgR.....: 1111

Tamanho da IgR.: 315 bp

Fita.....: Principal

Promotor.....: Posição: 206295 Algoritmo: NNPP

Terminador.....: Posição: 206449 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 148 bits (80)

Identidade.....: 131/154 (85%)

Candidato 45

17 bp

TAGAATTGTTTGGATCATCAGGATAAAAGTCTGGCAAGACTCGCATCCGGCGTTCCGGC
CCAGGACAGCAGCGCACAGCCCACCGGCGACCAGCAGCGCAACAGCGGCAGGCTCAT
CCGCCTCCCCTCTTGACAGCGCATCGCACGGGCCATGCCGCAATTCAGGCCGAGGCGT
CGCGAAGCTTGCCGGTTGATGCAGTGAAACCCGATCGCCGTCGCCAACCCCTATTGATA
TGTTACGTTTGGCTTACAGACCAGCTTACCGTCCGCAGTTGCCTGCGCGACCGCCTAT
GGCGGAGCGCGCGCGGATACGGCGCTGCCCGGCGTTCCGGTCCACCGCATCCG
CGCCCC**TT**

Tamanho sRNA...: 326 bp

IgR.....: 2900

Tamanho da IgR.: 800 bp

Fita.....: Principal

Promotor.....: Posição: 548230 Algoritmo: NNPP

Terminador.....: Posição: 548599 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 433 bits (234)

Identidade.....: 334/379 (88%)

Candidato 46

20 bp

TTGCCA**AAGGTCTGGATTGTAGCAC****TATGAA**TGCCTGCTTCACTTAAACGATTGACGC
CAGACATGTCCTTCACCCGCGCTTACTTTTTATTTTTACTTTAGCTATTCCAGCCCCACGG
CGGTGAAAAGCGGTAAGCGTTCGTAAAAAGACAAAACCGATTTCTAAAAAACCGCCAG
TGATGGCGGTTTTTT**TTT**

Tamanho sRNA...: 164 bp

IgR.....: 3297

Tamanho da IgR.: 545 bp

Fita.....: Reversa

Promotor.....: Posição: 626144 Algoritmo: NNPP

Terminador.....: Posição: 625934 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 305 bits (165)

Identidade.....: 191/203 (94%)

Blast hit 2

Espécie.....: *Janthinobacterium sp. Marseille*

Escore.....: 105 bits (116)

Identidade.....: 104/129 (81%)

Candidato 47

17 bp

TTGCTG**CCGATACCTGAATCGGT****TATAAT**GTGTTAGTTAAACGCAACCTGCCGCGAATC
GCATATGTGATAAATCATTGATAGATCCGCCTGCGCCATGAACGATGGAAGTACGGTGG
TTCGCGTGATGTGTGATCTCGCGTTGCGTGAGATAACAAAAGATGGGCAGATGCCCA**TT**
TTTTTTTT

Tamanho sRNA...: 157 bp

IgR.....: 1430

Tamanho da IgR.: 394 bp

Fita.....: Principal

Promotor.....: Posição: 270404 Algoritmo: NNPP

Terminador.....: Posição: 270595 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 322 bits (174)

Identidade.....: 185/190 (97%)

Candidato 48

18 bp

GTGAAATAACCAAACCGTATATCTGATATTTGGAAGATAAATTGGCCTGATTTGATGCG
GAGCAGTAAAGTAACACCTGTCTCCTCCAATTTCTCCTAAAGAATTGGATTCAAGCCCA
CACCGCAAGGTTGTGGGC**TTTTTTT**

Tamanho sRNA...: 115 bp

IgR.....: 2291

Tamanho da IgR.: 253 bp

Fita.....: Reversa

Promotor.....: Posição: 437177 Algoritmo: NNPP

Terminador.....: Posição: 437018 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 191 bits (103)

Identidade.....: 134/149 (89%)

Blast hit 2

Espécie.....: *Janthinobacterium sp. Marseille*

Escore.....: 68.0 bits

Identidade.....: 78/101 (78%)

Candidato 49

17 bp

TTGGACGAGCTTCGGCACTGACGTTTAATAGACCTCACTTGAACAGCGATTGATTGGTA
TCCGTTACTGTGCAAGGTTCCCTCCCTCTAGTGTTGAAAGCACCCTAGCTCGCGGATC
TGGCTCCCCCTGTCAGATCCGC**TTTTTTT**

Tamanho sRNA...: 119 bp

IgR.....: 867

Tamanho da IgR.: 497 bp

Fita.....: Reversa

Promotor.....: Posição: 159137 Algoritmo: NNPP

Terminador.....: Posição: 158974 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 250 bits (135)

Identidade.....: 153/161 (95%)

Candidato 50

19 bp

TTGTCCCGTGTAGCAAATAGTTAATTATGGTTCGTACCAGCGTCGCTCCCCGTCAAC
ACAACGGCGGCAGGACAATAAAACGGACAACACTGCATGACGACGGCACCGTCCAGATAG
GCCGGGTACGTCAGGCCAACGGCACTGCGCAGATAGCGTGGGCCTGGGCCAATCTAG
TGTAATCAATGAACGGCCCGCCTGATTCAAGGCGGGC**TTTT**

Tamanho sRNA...: 185 bp

IgR.....: 4097

Tamanho da IgR.: 551 bp

Fita.....: Reversa

Promotor.....: Posição: 801628 Algoritmo: NNPP

Terminador.....: Posição: 801396 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 128 bits (69)

Identidade.....: 188/242 (77%)

Candidato 51

17 bp

ATGTGACGCCGCTTTCAGACAGTTACAATTTAGAGGCCACAATCTTATTGCAACGCACA
AAGCGCTTTGCTATATTCAATTCAACGGAAGTTTTTGCAGCACGTAACGATGTCTCCTCC
ACCTCCTCAATGGTGTGGATTAACCCGGAACACAGTGTCCGGG**TTTTTTT**

Tamanho sRNA...: 144 bp

IgR.....: 1227

Tamanho da IgR.: 631 bp

Fita.....: Principal

Promotor.....: Posição: 232116 Algoritmo: NNPP

Terminador.....: Posição: 232299 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 198 bits (107)

Identidade.....: 166/192 (86%)

Blast hit 2

Espécie.....: *Burkholderia sp.*

Escore.....: 68.0 bits (74)

Identidade.....: 60/70 (86%)

Candidato 52

17 bp

GAGACAAGTCTTTAGTACGGCTCTAGGATTGAAAATCCTTATGTTGGTGGTTCGATTCC
GCCCCGGGCCACCAAGAATGTTCTAAGCAAACCGCCAGCCCTCGCCGGTTGGCGTTT
TTCA**TTT**

Tamanho sRNA...: 95 bp
IgR.....: 740
Tamanho da IgR.: 213 bp
Fita.....: Reversa

Promotor.....: Posição: 136437 Algoritmo: NNPP
Terminador.....: Posição: 136300 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 84.2 bits (45)
Identidade.....: 49/51 (96%)

Blast hit 2

Espécie.....: *Janthinobacterium sp. Marseille*
Escore.....: 87.8 bits (96)
Identidade.....: 51/53 (97%)

Candidato 53

17 bp

TTGGATCGTGGCAGCGAATGCGTTAGACTTCGTA

ACTCCTGCCGTGTTTCGCGATTGCCAT
ATATTCCTTGAACCATTTACGAGCATAGGTTGTGGGATCGTTT
GATGGGCGTGAGCGTC
GCTAGTCCGATGAACCCGAAATTGAACTACTGCAACCAACTT
GAACCGTCAGGTT
CAGGATGCCGGCACAGCGGCACAGGGCGGGCTTGATTCAAG
AAAAACGGCTGCATATGCAGC
CG**TTTTTTTT**

Tamanho sRNA...: 217 bp
IgR.....: 3474
Tamanho da IgR.: 466 bp
Fita.....: Principal

Promotor.....: Posição: 662940 Algoritmo: NNPP
Terminador.....: Posição: 663194 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 396 bits (214)
Identidade.....: 224/229 (97%)

Blast hit 2

Espécie.....: *Janthinobacterium sp. Marseille*
Escore.....: 214 bits (236)
Identidade.....: 180/217 (83%)

Candidato 54

19 bp

CTGACTCGCGTCAATCGACCTTTTC TATTTTGGTGCCACGGAGGGACTCGGTACGTC
CCTCGACCCCGTCCGGCCTTGCAAGGCCGGAC TTT

Tamanho sRNA...: 63 bp
IgR.....: 2029
Tamanho da IgR.: 276 bp
Fita.....: Principal

Promotor.....: Posição: 389467 Algoritmo: NNPP
Terminador.....: Posição: 389573 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 150 bits (81)
Identidade.....: 98/106 (92%)

Blast hit 2

Espécie.....: *Burkholderia multivorans*
Escore.....: 44.6 bits (48)
Identidade.....: 56/73 (77%)

Candidato 55

TCGGCTTTTTTATTGCCCGGCTTGCTTGCGATCATGGAGCCATCGGGCCGCAGGCCGC
ACAATGGCTCGCATTGCGCGTTTTGCTGCCCTGCAAAAC TTTT

Tamanho sRNA...: 72 bp
IgR.....: 3935
Tamanho da IgR.: 219 bp
Fita.....: Principal

Promotor.....: Posição: 769980 Algoritmo: NNPP
Terminador.....: Posição: 770088 Algoritmo: TranstermHP

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*
Escore.....: 139 bits (75)
Identidade.....: 99/110 (90%)

Candidato 56

18 bp

TTGTTAGGTTTTGGGTAGTCTGCACAGAATGCTTTTCCTGCGCTCTAGCTTAATTACTTTC
TTTCTCCGGTTCGGGAGAACGCGCCGGGGGCGGCCCGGCAGCCGCTCAC**TTTT**

Tamanho sRNA...: 82 bp

IgR.....: 3164

Tamanho da IgR.: 303 bp

Fita.....: Principal

Promotor.....: Posição: 595464 Algoritmo: NNPP

Terminador.....: Posição: 595589 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 115 bits (62)

Identidade.....: 105/124 (84%)

Blast hit 2

Espécie.....: *Janthinobacterium sp. Marseille*

Escore.....: 71.6 bits (78)

Identidade.....: 39/39 (100%)

Candidato 57

CTTCATCGAGACGTGATGCACAACAGTGGTTCGGGTAGCTTGCAAGATGCCCGACTGCT
GAAACGGAAAAGGCCTACGCCACATTAATGAACAGTCAGAGACCGAATCAATCTTCTCT
GTTATCGCCGCTTAACCCCATCTTGTCGGGTTCAGGCGGCTTTTCCCTTTGAAAAACACCG
AGGTTTTTCAGGAACGCGATTCCGACGCTAGCGTTTGAACACGCTGGCG**TTTTTTT**

Tamanho sRNA...: 203 bp

IgR.....: 81

Tamanho da IgR.: 466 bp

Fita.....: Reversa

Promotor.....: Posição: 17090 Algoritmo: NNPP

Terminador.....: Posição: 16853 Algoritmo: RNAMotif

Blast hit 1

Espécie.....: *Herbaspirillum rubrisubalbicans M1*

Escore.....: 405 bits (219)

Identidade.....: 232/238 (97%)

Blast hit 2

Espécie.....: *Herminiimonas arsenicoxydans*

Escore.....: 51.8 bits (56)

Identidade.....: 125/185 (68%)

APÊNDICE II – CANDIDATOS SEM SIMILARIDADE DE SEQUÊNCIAS ENTRE ESPÉCIES PRÓXIMAS

Candidato 1

TTGCCAACGAAGAGCGAATCTCCTATCATCTTCAAATAATAACGATTCTCATCTGTGCGC
GGCATGGGTGAGTGAAAGACAGCATTTCATGGCGGCCTGCGTATGCAGGTTCGTTTCTTTT

IgR.....: 317 - Tamanho da IgR: 375 bp - Tamanho da Sequência: 93 bp
Promotor...: Posição: 54041 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 53918 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 2

TTGCAGTCCTGTTTGCCAGCCGGTAACATCTGATGCCCGCTTGCGATGCTTCGCTTGC
GGGGCATTITTT

IgR.....: 3302 - Tamanho da IgR: 179 bp - Tamanho da Sequência: 47 bp
Promotor...: Posição: 627219 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 627142 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 3

TTGAGAAATTTTTCCGGCTTTTCGTTAATATCTGCGACAAGCCTGACAGACAGGACGCA
GCAACAGTCCTGCAGGGAAGTAGCTCATCGCGAGGCGTACCGCGCCGATGTAGAGCTC
GTCTGTGCGCAATGACAGGCGATGCGGTGAGAAAGATGCAACGGTAGCAATCATCAATC
GGCACAACCAGCACCCAGCCAGTACCAATCAGCAGCAACCGAACGTACCAGAAGGTCCG
CGACATCGATGGCGCTGGCGGCCGTCGTCCGCCGGTGCAGCAGGACTCCGCAGGGA
GCCGGGGACGGCGTTTGCCTTTTTT

IgR.....: 3829 - Tamanho da IgR: 546 bp - Tamanho da Sequência: 286 bp
Promotor...: Posição: 751407 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 751091 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 4

AATTTAGACAAAAATCTAATTTTTTAAAATTTGTCTTATTCCGTTTGGACAAGTTTTTTT

IgR.....: 3348 - Tamanho da IgR: 442 bp - Tamanho da Sequência: 36 bp
Promotor...: Posição: 637432 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 637498 - Fita: Principal - Algoritmo: RNAMotif

Candidato 5

TTGCATCCTCCCTAACCCGCCACTAGAATTAGTCCCAATCGATGTTTTGATACTAGATGT
AGTGTCCAGGTGAACTTATTTTTGTACACGGGTCTTCATCCACAGATGGCCCGGCTGGC
AGTGCTTCAGAAGGCGCAAGATCCCGCGAGGCAGTCACCAAGTACCAAGCAACAAGTA
CGAAAGCGCCCCGGTTCTCCCTTCTTATTCCGTCGTTTCGCTCATCAGTTTTCAAGCAGTT
TTTGCCTCAGCATGAATATCCATGCAAGGGATAGTCATCGTGGGACGCAGTCGTCGTTT
GCCAATCCCGTGTAGTTCGGGATTCTGATTT

IgR.....: 83 - Tamanho da IgR: 405 bp - Tamanho da Sequência: 300 bp
Promotor...: Posição: 17759 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 17429 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 6

TTGAATTAAGACCAATTAATTATGCTTTTTATGGAATAGGGACGAAAAAATCGCAGTCC
GAAGACTGCGATTTT

IgR.....: 2230 - Tamanho da IgR: 166 bp - Tamanho da Sequência: 53 bp
Promotor...: Posição: 425878 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 425795 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 7

CGGATTGCCAGCAAATGGAAATCAATAAGATCTAAAAATTGCACATAGCCTGATGGTGC
AGGCATACCAGTTATCGCCCCGAAACGAGCGACACCCGTATTTTGGAAAGTGGAAAGCG
AATGCTCGCATCCGACGCTTCAGCACTTATTGAAGTTCGTCCACTTAATTCGATTTCTGA
TGGGGGAGAACAAGGTCTGCGGAGTGGGACGGGCAGGGTGCGGCGCGTGTGGCCG
TTATGCTTT

IgR.....: 1075 - Tamanho da IgR: 460 bp - Tamanho da Sequência: 215 bp
Promotor...: Posição: 198296 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 198541 - Fita: Principal - Algoritmo: RNAMotif

Candidato 8

AATTTTCAAATTACGATAATCGATTTTCATAATACGAAAATATGCTGCAAAGCGGCATAAAC
AGGGGCTTTCCAGGCCTCGTATTTTT

IgR.....: 3270 - Tamanho da IgR: 555 bp - Tamanho da Sequência: 62 bp
Promotor...: Posição: 618425 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 618517 - Fita: Principal - Algoritmo: TranstermHP

Candidato 9

TTCCCTCCGCATTTTGGGAACAGCTATATTAGACCAAACCCCAAGTTCTATCTGGGAAA
TTCCACGTTTCGCGTTTTGATTTTCAGTAAACATTTCCGCATTCCGACAGCCGGTGCGAG
CAATGCGCGACGGGCGTTTTGAGAAGCAGCCGCATCGGCATGCTGCGCGTGTCATATTCT
TT

IgR.....: 3811 - Tamanho da IgR: 642 bp - Tamanho da Sequência: 153 bp

Promotor...: Posição: 746741 - Fita: Reversa - Algoritmo: BProm

Terminador: Posição: 746558 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 10

ACGGTTGCAGAATTGCCGCAAATCCGCTAACCTGTCTGACAGGTAGATAAGTATGTCTG
CCTGCCTTTTTT

IgR.....: 1049 - Tamanho da IgR: 272 bp - Tamanho da Sequência: 47 bp

Promotor...: Posição: 193553 - Fita: Principal - Algoritmo: BProm

Terminador: Posição: 193630 - Fita: Principal - Algoritmo: TranstermHP

Candidato 11

TTCACGTGAAACAATCGTATCGATTTTAGGAAGCCATATAGCATCGATAAGGAGGCATG
CATACGGGCGTCAAATAAAAGCGGAGCCATTAGCTCCGCTTTACCTTCCACTTTT

IgR.....: 425 - Tamanho da IgR: 274 bp - Tamanho da Sequência: 32 bp

Promotor...: Posição: 76908 - Fita: Reversa - Algoritmo: BProm

Terminador: Posição: 76846 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 12

TTAATTA ACTCTTTATACGGTTTATATCATTTAAAAAATAGCCAGAAAATAAAAAAGCGCG
TGCATTCTGCACGCGCTTTT

IgR.....: 230 - Tamanho da IgR: 275 bp - Tamanho da Sequência: 60 bp

Promotor...: Posição: 41075 - Fita: Principal - Algoritmo: BProm

Terminador: Posição: 41165 - Fita: Principal - Algoritmo: RNAMotif

Candidato 13

ACCTTTGCCAGCGAAGCGTATTTTTCTAGTCTTTTATCGAATTAGTTCGGCATCGATCT
CGCAAGATATCGGGAATTCATGAAGACTATGCGTCATTAGCGCATTTTT

IgR.....: 1474 - Tamanho da IgR: 353 bp - Tamanho da Sequência: 86 bp
Promotor...: Posição: 279563 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 279679 - Fita: Principal - Algoritmo: RNAMotif

Candidato 14

TTTTCAACCGCATTNTTTTTATGCTTTTTGTGCGTTTATGTGTGTTTTCGTGGAATGCG
GCATTTCCCGATGGCGGGAGTTCAACTTGCTGACCATAATCAGAGTCTTAATTTTGATCA
AGTCAATTGAAAACAAACGAGTTTTTGGAACTTTTGTAAATTCTGCTGCCATGTGCCGATG
GTAGAATGACCTTGCCAAGATGGTAAGTTTT

IgR.....: 2971 - Tamanho da IgR: 442 bp - Tamanho da Sequência: 190 bp
Promotor...: Posição: 561883 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 561663 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 15

AAGCTTGGAGGTGGCGGCGGGGATGGTTATTCTGAAGGTCCTCCCCCTCCAATTCAT
CACCCCTGAGAATTGGATTAAACCCGCCGCCAAGGTGCGCGGGTTTTTTT

IgR.....: 817 - Tamanho da IgR: 289 bp - Tamanho da Sequência: 85 bp
Promotor...: Posição: 150117 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 150232 - Fita: Principal - Algoritmo: RNAMotif

Candidato 16

CGATTTACAGCCACTATCCCAAGGTCGTCATCATCAGTCATGGCCGGCGGGCGGATT
GGCCTGCCCGTTGCCCGCCGTCGGCGTTTTT

IgR.....: 3989 - Tamanho da IgR: 640 bp - Tamanho da Sequência: 66 bp
Promotor...: Posição: 779741 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 779837 - Fita: Principal - Algoritmo: TranstermHP

Candidato 17

AGTGTTGAAGTTGAACGGGCAATGTGAGGGAGTAGGCTGATTGGACTTGGGGGCGAGG
CTGCCGGTTCAAGCTTTGAATAGCGAAGTGATTACGTAGATGGATCAGGTTGCGCAGAG
GAATTCGTGCCCTGGATGAAGAGGACGGTGCCCGCTCCAGTCGCTCCTGGATCAGGC
TGCTCAGTTAGTGACAATGGTGTGGGGTTTTAATAGCTAGCAGTGGAGTTGCTTGAAG
CCATCTGGGATCGACAACTGCAAACGATATTTACATAATACAAATTATGCGAAGTTTTG
ATGGCAGCACAGCAACATTCAATATCGGGAAATGCCGCCTTTTTAGGCGGCTTTTT

IgR.....: 2484 - Tamanho da IgR: 431 bp - Tamanho da Sequência: 327 bp

Promotor...: Posição: 473127 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 473484 - Fita: Principal - Algoritmo: TranstermHP

Candidato 18

GGGGTTGAAGGAAATAAAGGAGGATAAAAAAGGAAAATAGCGAGATGACTGCTGGCA
AGCCGGTCAAACCGGCATTTT

IgR.....: 4108 - Tamanho da IgR: 283 bp - Tamanho da Sequência: 58 bp
Promotor...: Posição: 804934 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 805022 - Fita: Principal - Algoritmo: RNAMotif

Candidato 19

GGGGTTGAAGGAAATAAAGGAGGATAAAAAAGGAAAATAGCGAGATGACTGCTGGCA
AGCCGGTCAAACCGGCATTTTACTATCTCCGGCGTCAAAGGTGCAGTGCAATAGGG
CGAAATTCTACCGCAAGCCCTCTGACAAGCCTGTTTTGCGGGGGTAAATCCGCGCTACC
ATGCGGTTTTT

IgR.....: 4108 - Tamanho da IgR: 283 bp - Tamanho da Sequência: 163 bp
Promotor...: Posição: 804934 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 805127 - Fita: Principal - Algoritmo: RNAMotif

Candidato 20

TTGAAGTCAAATATATAGTGCGCTATTTAATAGTATGTAATTACGCATTCAGGAAGTCGTA
CGCAGCTTGATGTGCAGGTTCTCCTTCCCTCCGACAGGTCCGCAGGAGGGCAGTTAT
ACCGAGAGCCAATAGCGTCATATAAGTAGTTTGCTATTTAATAGTTCTTTAATAGTTATTT
GGTAGTTATTTAATGGTTGTTTGATTGAAACGCCGTTTTTTTTGCTGTAACCCGACGCAGT
CCCCTCGTTGCAGTTCTCCGGTGTCCGCACCGATTTGATATTT

IgR.....: 2930 - Tamanho da IgR: 361 bp - Tamanho da Sequência: 255 bp
Promotor...: Posição: 555421 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 555136 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 21

TTGCGGTTATCCGTGGAGCGGATGGGGGAATAATCCGGTGGAGTTGGTGTGGCCGCCGG
TGCCGTTGGTCTCGGTGGCGGCGACGGGGTCTTTTT

IgR.....: 3595 - Tamanho da IgR: 6867 bp - Tamanho da Sequência: 71 bp
Promotor...: Posição: 690553 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 690452 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 22

TTGCGGTTATCCGTGGAGCGGATGGGGGAATAATCGGTGGAGTTGGTGTGGCCGCCGG
TGCCGTTGGTCTCGGTGGCGGCGACGGGGTCTTTTTTCGTCATCGGGCGAGATGATGGG
CTTGTCGCCTTCTTCGTCTTGCTGGCGGGCTTTTTT

IgR.....: 3595 - Tamanho da IgR: 6867 bp - Tamanho da Sequência: 128 bp
Promotor...: Posição: 690553 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 690395 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 23

TCTGTTGCTCCCTGTGACCTGCAAGTTTTATGTAAATTTTTGTCCCGAAACCATTCAAG
CGGATTGAATTACTTCGGTAAATCTCCCGAAAGTGCCGTTGATCCGTTGACAAGCAACG
AATGGCTGGAGAACCCCAACGTGAAGCAACTGAAGGTCGTTCAAACCAGCAATATCTTC
GGCTGCCCTTTGGTGAGGGCGGAGAGGGGAAGGGGGCATTCCCCCTTTT

IgR.....: 2307 - Tamanho da IgR: 445 bp - Tamanho da Sequência: 207 bp
Promotor...: Posição: 439851 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 440088 - Fita: Principal - Algoritmo: RNAMotif

Candidato 24

TTGAATGCACGCATGGTAGCGGCGAGCTGTGATGCGGTGCATCAGGAAATGACTGAGG
CGATAAACATCACACCCGGCGCGGCGCGCAAGCTCTACAGCAGGAACCGCAGCGCTG
GAGGGGGGTCTGCTGGCTATCGCCGCATCGGCGCTTTTT

IgR.....: 603 - Tamanho da IgR: 496 bp - Tamanho da Sequência: 131 bp
Promotor...: Posição: 109959 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 109798 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 25

GTGCTTGCTTCTTCTTCATTGTTATTCCCAGTACAGTAGGCTCAACATCGAACAGGCAAT
GGAGTGCTTTGCCTGTCTTTT

IgR.....: 4049 - Tamanho da IgR: 277 bp - Tamanho da Sequência: 57 bp
Promotor...: Posição: 789908 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 789995 - Fita: Principal - Algoritmo: TranstermHP

Candidato 26

CACATTTCCAGTCGGCGCGAGGACTGGCAGACTGGTAAGCCTTCGAAGACGTGCCAGC
CTTCCAAGCCTTCTCAACCTTTCGCCGCGAGGGCGCTTTTCCATTTT

IgR.....: 1970 - Tamanho da IgR: 198 bp - Tamanho da Sequência: 76 bp
Promotor...: Posição: 378093 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 378199 - Fita: Principal - Algoritmo: TranstermHP

Candidato 27

CGAATGGCCGCCAGAAGGAAACGATTTATATTATTTCCCTCTGCCTGCACTTCATGAAAA
ACCGCCGCTTGCAAGACCGGCGGTTTTTT

IgR.....: 944 - Tamanho da IgR: 148 bp - Tamanho da Sequência: 65 bp
Promotor...: Posição: 174864 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 174959 - Fita: Principal - Algoritmo: TranstermHP

Candidato 28

TGCCTTACCTCACATTCCACATATTCCATATATTCAGAAATAAATTCATGGTGCAATGGC
GGCGCAGCGGTGTCATGCGCTGCCGGCGGGCTGGAATTTGCACATTCTCATAGGGATA
AGTGCTGTCTCTTTTT

IgR.....: 3677 - Tamanho da IgR: 922 bp - Tamanho da Sequência: 111 bp
Promotor...: Posição: 712859 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 713000 - Fita: Principal - Algoritmo: RNAMotif

Candidato 29

GTTTTTGCAAGTGCAAAAGCGAGCTTTGGAAAAGTTACAAATGCAGCCGCCAGCGAGGG
CTATTGGATCGCATGGGAAAGCGATCGCAAAGGGAATGCGAGGCCATCGTCTGGCTCG
CATTGATGCAAATAATTTATAAGACGAGCCAGGGAACTTCAACACCAAAAATAATAAAA
AAGCGCCGCATGAAACTCATGCGGCGCTTTT

IgR.....: 243 - Tamanho da IgR: 518 bp - Tamanho da Sequência: 185 bp
Promotor...: Posição: 43229 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 43444 - Fita: Principal - Algoritmo: RNAMotif

Candidato 30

TTGATATTTATAAATTCCTCGCAGATTCTTAATCAACAGGGCAGGCGGCTAAGCTAGTT
TCACACGCTGTTTCCGAATAACAAATCCTGCCAGGATTTTCGAGAGATAATTCCCAGCTT
GCTTGCTGGGGAATATTTTGCCCAAACGAGCGCAAACCCGCGAATTGGCGCACTGTAT
GCGTAACCAGTGGCCCGCGTCGGATGGGTCTTTGCGGCGTGCCGTCAACATGCGCTTT
TGA CT CATAATCCGTTGGTGCCGTATTGACTCACGGGGAGGGCTGCCACTATCTAAAG
GGGGATTCCGAGAGATCGGAGTCCCTTTTTTTT

IgR.....: 3742 - Tamanho da IgR: 666 bp - Tamanho da Sequência: 302 bp
Promotor...: Posição: 728078 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 727746 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 31

TAGGTTGATGAGTTGCAGGCTCGGCCTGCTTTCCTTCAGACCGTGAAAGCGGCGCATG
TCGAAAGGCATGCGCCGCTTTTT

IgR.....: 1645 - Tamanho da IgR: 452 bp - Tamanho da Sequência: 58 bp
Promotor...: Posição: 314637 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 314725 - Fita: Principal - Algoritmo: RNAMotif

Candidato 32

ATGAGTATTCGGAGAAGATCGTTATTCTATAAATTATTTATATCAATAAAAAATGGGCTGC
TCCGAATCGGGCAGCCATTTGCTTT

IgR.....: 431 - Tamanho da IgR: 282 bp - Tamanho da Sequência: 61 bp
Promotor...: Posição: 77790 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 77699 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 33

GGTCTTTCTAGGACCTACACATAGCGACTTGAAAATGACCAAGAGTAAAAAGAAGCCG
AAGCAGACTATGAGTGCCTCTCGATGGGCGACCATTTCCGAGCGAGCATGAGCGCCG
TGAGCGTGAGGAAAGAAGGAGGCGCCGAACAGCGCATCGCCGGTAACGGTCTATAAG
CCTGACGCCGGTTCGCAATATGGTTCGCGGCGGAGGAGTTGCGGTGCCGGTACTGCCA
ACCGCTTCGCGCAATCGCAATTGAAGCGCATCTGCTCGACCATTTAAGGACGACCCAG
CACCTCAAGGCGTATCACAAGTAGTCCTTTCAAGCCCGCACCTGTGGTTTTTTT

IgR.....: 737 - Tamanho da IgR: 988 bp - Tamanho da Sequência: 318 bp
Promotor...: Posição: 135286 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 135634 - Fita: Principal - Algoritmo: TranstermHP

Candidato 34

TTGGCAATGAGCGCGCAAGCTCATCATTGGATGGCATGGTAATAGACCTGCTGATGCTT
GTTTCAGTTCTGTTTACTTTCAGTTCTACGGGTGGGGAGGTGAGTGTCGGTACATCGGTAC
ACGGGCAGCGCCGACACATCATGCAGAATGATCAGAATAAGCAGGATAACCAGGCGTC
GCCTTGCGCGGCATCAGAGAAGCTCGCCAATGGGCGTCAGACAAGAGGAAACCTGCC
GATCCGGCAGGGCCGCTGCAGACGTGCGGACACATCTGGAAACAAAACGCCCGCAT
GACATCTGCGGCGTTTTT

IgR.....: 89 - Tamanho da IgR: 355 bp - Tamanho da Sequência: 285 bp
Promotor...: Posição: 19157 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 18842 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 35

TTATTTCAAAAAGAAACCAATTCCCGCTACAAAAGGGAAATCATCTGTTTCTTACGCC
AAACATTGGACGCAGATATGTTACTGTGTTTT

IgR.....: 1548 - Tamanho da IgR: 722 bp - Tamanho da Sequência: 70 bp
Promotor...: Posição: 295474 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 295574 - Fita: Principal - Algoritmo: TranstermHP

Candidato 36

TCCTCTGTCAAAAATTTCAAGCCTCTGTTAGTCTTTCCTCATGCACGCTGGCTCTGGCGT
GCAGCGGCAGTCCCTTAGCGCCATCGCATCACGCTGTAGTCTGGCGGCAGTCACGCTG
CCGTTTT

IgR.....: 3977 - Tamanho da IgR: 267 bp - Tamanho da Sequência: 103 bp
Promotor...: Posição: 777878 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 778011 - Fita: Principal - Algoritmo: RNAMotif

Candidato 37

TTGTCATATCGTTGTCAGGCCGGTACATCAGAAAAACCCTGATATCTGGCCTTTCACCTG
GGGAATAGGATGAGGCCATCGCATCTGGTCTTATCCGTGAGAGGGAAAAATGGACGTC
ATCCAAAGCTTGCTGCACTACGGTAGCTACGCCGTGGCTTTCTATGCACTTTATCTGGT
GGCGGGCGCGCTGGCGTTCGGCCTCTACATCTGGTTGCGCGTCTATCGCCAGCCGGA
GCTCAAGCCCGTGCCGGTGCGCGTGCGCCGCAACCTGCGCGACTCGGATACACCGCT
GCGCTGATACCGGCTGTTACAGCTATGCACGCAAGGGGCTGGATTGATGAATCCGGCC
CCTTTGCTTT

IgR.....: 2429 - Tamanho da IgR: 487 bp - Tamanho da Sequência: 333 bp
Promotor...: Posição: 463442 - Fita: Reversa - Algoritmo: BProm

Terminador: Posição: 463079 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 38

TTCAAAAACATCAGATCAGGAAATTA AAAAGATTCCGTTGTAATTTAAAGATGACGCATC
GCGAAGAGTCATCGACGCGTCGGAAGAAAGAACGAGCCAGTGACACCATCGGACAG
CAAGGAATTCGGCATCGCCGGTTTT

IgR.....: 3782 - Tamanho da IgR: 357 bp - Tamanho da Sequência: 120 bp

Promotor...: Posição: 738059 - Fita: Reversa - Algoritmo: BProm

Terminador: Posição: 737909 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 39

TTAAAAGCTGCAAACTTTTTAACTCATATTA ACTGGCACCGACCGATGCCGAGCGATG
CGAACTGATATTA ACTGATACCAGCTGATGCAATGATGGGAATATCCCATCCGGAAAAA
CAGGAAATAAGCGCCGATCTGGCCTGCAGATCGGAACGGGAAGAAACAAGAATGACCT
CTCTCCCTTCCCGTTCTTTGTTTCAGCTTATTTCTGCATATTTCTGCTTGTTTCTCTTATTG
CTGTTCTGATCCTGGCTGATCAGGTATTTCTTTTT

IgR.....: 2074 - Tamanho da IgR: 1112 bp - Tamanho da Sequência: 246 bp

Promotor...: Posição: 400871 - Fita: Reversa - Algoritmo: BProm

Terminador: Posição: 400595 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 40

TTGACAGCGTTGTTGCGCGTAGCTGGGCTGGCGTTTTTCGGCGGCGCTCTTGTCCAGGT
CCTGGACGCCGATGTTGCGGTAATCCTCTGCCGTGGGCAGTTCTCTGGCATTGCCTGC
CACACCGTCGGCCAGTGCCAGGATGTGGTTATAGGAGCCGACCAGGGTAGCCAGCCT
GGGCCAGGAGTTGATCTTTGGACCCTTCACTTCACGCAGGGCGCTGTTGATGGCGGCG
AGGTTGCTGGCGTCTACGGCGGTGTCGCCATTGAGAAGCGCGTTCCTGTAGTTGCTGA
CGAACGGCTCCGTTTGGGCCGGATTTT

IgR.....: 3595 - Tamanho da IgR: 6867 bp - Tamanho da Sequência: 294 bp

Promotor...: Posição: 694213 - Fita: Reversa - Algoritmo: BProm

Terminador: Posição: 693889 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 41

TTTCCGGTGTTATTTTTTTGCAGGCCCTATTGTGCAGGAATTCACCGTAAATAGTTGTTT
CTGAGTGCGACAGCAGGTGCGCCACGACAAGGTGGGCCAAACAAAAAAGGCAGCCGG
AGCTGCCTTTTTT

IgR.....: 3786 - Tamanho da IgR: 1001 bp - Tamanho da Sequência: 106 bp
Promotor...: Posição: 739225 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 739089 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 42

TTGACAGCATAGCCGGCCTGGCGATCTGGCCGGCATGCCAGCGCGCAGTTTACCGTCT
AACCGCTCCTTGCAACACTTAACAGACGTAAAGAATTACATTCAGGCATGCCCGCAGCC
CAGCAAAAAGGACGCATGATCGCGTCCTTTTT

IgR.....: 2161 - Tamanho da IgR: 296 bp - Tamanho da Sequência: 126 bp
Promotor...: Posição: 414517 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 414361 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 43

CCGCTTATATCGCCGCACCGGCGCGGGCAGTAATCTGGTCGTTACAGGCATACGCGTA
CAGCGCGTCGTTTCCATTGAAAAGACATACCCCATATCCAAAGGAGGTCAGCATGAGCG
CAAGCAAACCCGACAAGGACCATCGTCCCAACAAGTACGACACGCCCTCGGAAGTGGC
GCTGGATGAAGCCAACGAGATGAGCTTCCCCGGCAGCGATCCGGTGGCCACCTCCAAT
ATCACGCGCATCGAAAAAGCGCCTGAGATGGATCGCGCCAGCAATGACCACCAGAACC
GCAACCAGGTCCGCGAGAACACCCGACAGAACACCCCTGACTGAGCCTAGTCCGCCCC
ATCCAGGGCGGCGCGGTGCATGCCGGCCGCCCTTTTTCTTTT

IgR.....: 2572 - Tamanho da IgR: 453 bp - Tamanho da Sequência: 362 bp
Promotor...: Posição: 489905 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 490297 - Fita: Principal - Algoritmo: TranstermHP

Candidato 44

TTTTGTGCCGAGAAGAGGGGGCAGATCGCCATAAGATTGCGAATTCACGCATGGGAAG
ACTGCCGTTCTGCGTCGCACCGCGTGCATCATCGCTGGGGTAGCGCTGCGGTCCGGCT
GGCGTACCTGGCCAGCGCAGCGCAGCGCTCGCGAAGGCGCAATGATCTAAACGTTTT
ATATGATTTACCGGCAGCAGGTCTTGGCTGCCGTTTGATTT

IgR.....: 3560 - Tamanho da IgR: 627 bp - Tamanho da Sequência: 188 bp
Promotor...: Posição: 681552 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 681770 - Fita: Principal - Algoritmo: RNAMotif

Candidato 45

TTCACCCCATCTGTGGCGGTGATGCACTACGATACTGGGCAATCTGCAATGCTCAGCAG
CCCCGATTCCGGTTCGTGCCTGCCCGACCATCGCCGGCGGCATCCAGCCGCGTTT

IgR.....: 2301 - Tamanho da IgR: 178 bp - Tamanho da Sequência: 94 bp
Promotor...: Posição: 438469 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 438345 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 46

TAGACGTGATGTAACGGTTGCGTGATTATCACGTAATCGAGAATGGATGAAAAAAGCC
ACGGATCGCTCCGTGGCTTTT

IgR.....: 3177 - Tamanho da IgR: 164 bp - Tamanho da Sequência: 58 bp
Promotor...: Posição: 597886 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 597798 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 47

GAGATTGTCGGGGTATTGCTGTGGTGGTGCGCATAAACCACAGGACGCGCGCCAGAAT
CAGCCTTTCTGCAGGGGCGCAAGGCGGTCCCCTTTT

IgR.....: 590 - Tamanho da IgR: 435 bp - Tamanho da Sequência: 72 bp
Promotor...: Posição: 107919 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 108021 - Fita: Principal - Algoritmo: TranstermHP

Candidato 48

GTAATTAACAAAAGAGCGAAGGTCGGTGTACAGTTGGTAGCTTTTTTTAGAAAATTTCTT
CAGCAAACGGAATCGTTCGCAAGCTTACGCATGGCAATGGCTGCCAGCCGGGCCGGCT
ATTTTT

IgR.....: 1535 - Tamanho da IgR: 1804 bp - Tamanho da Sequência: 101 bp
Promotor...: Posição: 291401 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 291532 - Fita: Principal - Algoritmo: RNAMotif

Candidato 49

CCATTTGTGTAGCGTGATTTTTCTGCACTCGAAAATAAAAAAAGCAACTGCAAAAAACA
ACAGCCCGCCGCCCGATATCTCAGTTTTTCTCTCTCCACCAATGAAAAAAGCGTCGCC
AGACAGCGACGCTTTTTT

IgR.....: 3639 - Tamanho da IgR: 296 bp - Tamanho da Sequência: 113 bp
Promotor...: Posição: 705456 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 705599 - Fita: Principal - Algoritmo: TranstermHP

Candidato 50

AAACTCGCAAAAAATTCGTTAGCAAGTTTAATTCATTTTTTCGATGCTTCCGTGAAACATCA
GGCCGGACCAGCGCAAAAACGAGGTTTTACCCGGCTGTCATTGCGCAAGCATTGATCG
AGGACAACCCGGCCCATGACAGATCACGGCAGGTCACGCACAGCAGGAAAGTGCGGG
GAGAGAAAAAGAGAGGCACAAAACAAAAAGGAAGCCGCCAGTGCGCTTCCTTTT

IgR.....: 2824 - Tamanho da IgR: 330 bp - Tamanho da Sequência: 209 bp
Promotor...: Posição: 534792 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 535031 - Fita: Principal - Algoritmo: RNAMotif

Candidato 51

AACCTAGCCGGATCATTACCAGACTGGAAAGATGTTTCATATATCGGCATGACCAATTCA
CGTCAGGGATTGATAGGGCGCTGGCAGCAGTTCTTTAACTCAATTCGAGGCACACCTG
GGCATAGTGGCGGCAACGCCGTTTT

IgR.....: 3484 - Tamanho da IgR: 996 bp - Tamanho da Sequência: 121 bp
Promotor...: Posição: 666093 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 666244 - Fita: Principal - Algoritmo: TranstermHP

Candidato 52

TTTCTACACCCCTTTTTCTACACCCCTTTTTCTACACCCCTTTTTCTACACCCCTTTTTCTA
CCCCCCTTTTTTCTATCCCCCTTGGCAAGCCCAGGGTGGTAGTTCCATCCTGTTACCA
ATGCTTTCAAGCGCTCACGTTGCTTGCATTTTT

IgR.....: 371 - Tamanho da IgR: 488 bp - Tamanho da Sequência: 132 bp
Promotor...: Posição: 63958 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 63796 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 53

TGAGTTGGTAGTTAGTAATTAGTAGTGATTAATTAATGTTTCAGTAGTCAGCAAATGCGGG
CATAAAAAAACGGAACCATCGGTTCCGTTTTTTT

IgR.....: 3869 - Tamanho da IgR: 144 bp - Tamanho da Sequência: 70 bp
Promotor...: Posição: 757932 - Fita: Principal - Algoritmo: BProm

Terminador: Posição: 758032 - Fita: Principal - Algoritmo: TranstermHP

Candidato 54

ATAGTTTTCGGAGAGAACCAGCTATTTCCAGATTTGTTTAGCCTTTCACCCCTATCCACA
GCTCATCCCCTAATTTTTCAACATTAGTGGGTTTCGGTCTCCAGTGCGTGTTACCGCAC
CTTCAACCTGGCCATGGATAGATCATCTGGTTTCGGGTCTACACCCAGCGACTGAACGC
CCTATTCGGACTCGATTTCTCTGCGCCTTCCCTATTCGGTTAAGCTTGCCACTGAATGTA
AGTCGCTGACCCATTATACAAAAGGTACGCAGTCACCCCTTTCGAGGCTCCTACTGTTT
GTATGCACACGGTTTCAGGATCTATTTCACTCCCCTTCGGGGTTCTTTT

IgR.....: 1 - Tamanho da IgR: 3098 bp - Tamanho da Sequência: 322 bp
Promotor...: Posição: 2015 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 2367 - Fita: Principal - Algoritmo: TranstermHP

Candidato 55

TTTCGGTCTACGAACTAGTTGTTCCACTATGTGTGCGGCATAGCGCGAATAGAGCTAT
TCATGCGCTCGACTTATCATTAGCAGCGGCTCCCTCTTCCCGTATTGACAAGTAATAT
CTCAAGTTGAAGACACATCGGCTTCATTTGGCCAATTAATGCCGAAGCGACGCCATTA
CCTCACGGCACTCTTGTCGCTCAGCAGGTAGCTGAAGGCTGCTTTT

IgR.....: 3755 - Tamanho da IgR: 482 bp - Tamanho da Sequência: 202 bp
Promotor...: Posição: 731018 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 730786 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 56

CTGTTAGTAATCGATGGTCTTTTCAACTCTGCCCGTGCGTTTTATTTGTGTAAGCGGGA
GGTGCTGATGCACTCGATTTAGAAAGTTTTTACATATTGCGGAACTGGACCGCTACTG
TTCTGTCTATGCGATACGTTTGAGCAACATTAAGTTCAACTTACGAAAAACAATTACTTA
AAACTTAAACTTAATTAATTAGTTGGTTGCCAGCCTCATCAACCGGCTGGTCGCGATT
GGAGCGCGGGATGGATAAAAAGTGCATTTTTTTCGCACAAATTTTCGCGCAACGGTCGCT
CAGGGAATTCATTGATGTCTCAACAGTAAGCGTTTCGGAGATCTCCGGGTGCTGCATGGA
TCGTCTGCGCGCCCTGCGCGCAGATGTTTTTTTTT

IgR.....: 3884 - Tamanho da IgR: 507 bp - Tamanho da Sequência: 366 bp
Promotor...: Posição: 760265 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 759869 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 57

TTGTGGTGGCGCATGAAGGGCGGCTTACTCAGCATTTTTCAAAAACATTTGAACGTGCA
GGTCTATGCGATGGCCGGGCGCCGGCTTTGTTT

IgR.....: 3260 - Tamanho da IgR: 658 bp - Tamanho da Sequência: 67 bp
Promotor...: Posição: 615168 - Fita: Reversa - Algoritmo: BProm
Terminador: Posição: 615071 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 58

GCGGCGGCAGCGGCAAGGCTCCCACCTATACTCCCCTGGCAGTCACAAGTCTGGGCC
ACCGGCCCGGGGGGCCACACTTTACCGAGGAGATGGTCATGTCCATCAACAACGTCA
CTTCCACCTATCACAGCAACGCCTACACCAACCTGCCACGCGCGCCAAGAAGCCTGC
GGTCAGCTTCACCGACGTGCTCAACTCCGCTGGCCAGCCTGCCCTGGGCAGCAATGGC
GTCGTCAACTATGGTGTGCCGGCCAGCCAGAACCACCCGCGCCTGAGCGAGCGGC
GCAGAGGCAAGAAAGCCCGGTCACGTTGTTGCGTGGCCGGGCTTTTT

IgR.....: 3093 - Tamanho da IgR: 666 bp - Tamanho da Sequência: 312 bp
Promotor...: Posição: 585503 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 585845 - Fita: Principal - Algoritmo: RNAMotif

Candidato 59

GCGCCTCCAGCGCGCTGCTCCGATCTGTGTGATCTCCGGCAAGACAAATAGTAGAGAG
ACTGACGGCGTTCGATTGCGCCAAAACCATGATTGCCATTGCCGCCATCAATGGCGCTTT
TT

IgR.....: 2481 - Tamanho da IgR: 590 bp - Tamanho da Sequência: 96 bp
Promotor...: Posição: 472372 - Fita: Principal - Algoritmo: BProm
Terminador: Posição: 472498 - Fita: Principal - Algoritmo: TranstermHP

Candidato 60

TTCAATTTTCAAATTACGATAATCGATTTTCATAATACGAAAATATGCTGCAAAGCGGCATA
AACAGGGGCTTTCCAGGCCTCGTATTTTT

IgR.....: 3270 - Tamanho da IgR: 555 bp - Tamanho da Sequência: 65 bp
Promotor...: Posição: 618422 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 618517 - Fita: Principal - Algoritmo: TranstermHP

Candidato 61

TTTTTTATTGGCGCAAATATTAGGCAGAGGAAATTCTGTGAGATCTATTAAGTCCCTAG
AATGAAGTTCTGCCCATGGTTTTCTTGGGAGGATCGAAAGGATATTCCTCTGTCTGAAAT
GCCGGAATTTACCTTCTCTCAATCACAAAAAAGAACCTAGCCGGATCATTACCAGACT
GGAAAGATGTTTCATATATCGGCATGACCAATTCACGTCAGGGATTGATAGGGCGCTGG
CAGCAGTTCTTTAACTCAATTCGAGGCACACCTGGGCATAGTGGCGGCAACGCCGTTTT

IgR.....: 3484 - Tamanho da IgR: 996 bp - Tamanho da Sequência: 275 bp

Promotor...: Posição: 665939 - Fita: Principal - Algoritmo: NNPP

Terminador: Posição: 666244 - Fita: Principal - Algoritmo: RNAMotif

Candidato 62

CTTGTTTGATTGTTTGAAGATCGATGGCGCAGGGGAAAGTGAAGCGGAAATGAAAAAG
CCCAGCATTGCTGCTGGGCTTTTTT

IgR.....: 99 - Tamanho da IgR: 142 bp - Tamanho da Sequência: 60 bp

Promotor...: Posição: 20718 - Fita: Reversa - Algoritmo: NNPP

Terminador: Posição: 20628 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 63

GCTAGTTTTAATTAACCTTTATACGGTTTATATCATTTAAAAAATAGCCAGAAAATAAAAA
AGCGCGTGCATTCTGCACGCGCTTTT

IgR.....: 230 - Tamanho da IgR: 275 bp - Tamanho da Sequência: 67 bp

Promotor...: Posição: 41068 - Fita: Principal - Algoritmo: NNPP

Terminador: Posição: 41165 - Fita: Principal - Algoritmo: RNAMotif

Candidato 64

TGACTTTTTTGAATTTGTGGAGCAAGAATCCACATAAACTATTGATCTAATGAAATTACTT
CTAGATTTTCCGTCCCTTGGGCAGGCGCGCTATGTTAGCTTCACTCGCTGTTTCTAAAC
AACAAATTCCGCCCGTGATTTGAATATACAAACGCCCGCACAAATGCGGGCGTTTTTT

IgR.....: 1548 - Tamanho da IgR: 722 bp - Tamanho da Sequência: 153 bp

Promotor...: Posição: 295228 - Fita: Reversa - Algoritmo: NNPP

Terminador: Posição: 295045 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 65

CTTTTTTTGTGACGCAGAATGTTCCGGTGTGCAATTTTCCCGGAATGTGAGACGAGGGG
CGGCAGGTTTTCCCGGACCGCCCGGTTTTT

IgR.....: 2021 - Tamanho da IgR: 322 bp - Tamanho da Sequência: 67 bp
Promotor...: Posição: 388368 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 388465 - Fita: Principal - Algoritmo: RNAMotif

Candidato 66

GCCTGTATTGGCTTTGACGCAGCTTGCCTGCTATCAGAATGGCGCCCTCCATTCTCAGG
GGCTGCGCAATGGACATGGATGGCCCCTTGCGAATACTTATAAAAAAATCGCCGGCTTG
CACCGGCGATTTGTTTT

IgR.....: 3277 - Tamanho da IgR: 521 bp - Tamanho da Sequência: 109 bp
Promotor...: Posição: 620445 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 620306 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 67

AAGTGATTTAAACGAATAAAACAAATTAATTAATTTATTATCATTCTCTGGTAAGATTGTCA
CTGTAACCAGAACGGTCAGCGGGAATCCGTAGCAAACGTCGTGACGGCGGGACGATG
CGCATCACCAGCATGCGGTCATACGCGTGGTTTT

IgR.....: 3466 - Tamanho da IgR: 666 bp - Tamanho da Sequência: 131 bp
Promotor...: Posição: 661184 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 661345 - Fita: Principal - Algoritmo: TranstermHP

Candidato 68

GAGTTTTCTTATTTTCACTCCTCGCTGCTTAAAGATACTTCGATCGGTCGTCTGCTCGT
GATACTGGTACGCGGGGAAAGCAGGCGACCAATGTCGTAAAGGTAGCAAGTGAGTAAG
TAGTAACGTAGCAGAGGTAGTCGTGGGTAAAGCCCGCATCCGAAAGGATGCGGGCTTT
TTT

IgR.....: 3068 - Tamanho da IgR: 708 bp - Tamanho da Sequência: 155 bp
Promotor...: Posição: 580980 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 580795 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 69

GAGTTTTCTTATTTTCACTCCTCGCTGCTTAAAAGATACTTCGATCGGTCGTCTGCTCGT
GATACTGGTACGCGGGGAAAGCAGGCGACCAATGTCGTAAAGGTAGCAAGTGAGTAAG
TAGTAACGTAGCAGAGGTAGTCGTGGGTAAAGCCCGCATCCGAAAGGATGCGGGCTTT
TTTGTGCGCCTTCACTTCAGGCCATGGCTGCCTGCGGGCAACCTTGCAGCCGTGCTG
CCTTGCTGCCATGGTGCCTTGGCGCCGACCATGGGGCGTTGCTTTT

IgR.....: 3068 - Tamanho da IgR: 708 bp - Tamanho da Sequência: 258 bp

Promotor...: Posição: 580980 - Fita: Reversa - Algoritmo: NNPP

Terminador: Posição: 580692 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 70

TCTGCCAATTGGATTTCTTCCTCGCCCCCTGTTTAAAGTAAAGGCAGCAGAGCAGGAG
GCCCCAGGAAGACGCGCAAAAACAAAAAGCCCGGTCGCAAGGACCGGGCTTTTCATTT

IgR.....: 3193 - Tamanho da IgR: 169 bp - Tamanho da Sequência: 90 bp

Promotor...: Posição: 602310 - Fita: Reversa - Algoritmo: NNPP

Terminador: Posição: 602190 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 71

CATTTGTGTAGCGTGATTTTTCTGCACTCGAAAATAAAAAAGCAACTGCAAAAAACAA
CAGCCCGCCGCCGATATCTCAGTTTTTCTCTCTCCACCAATGAAAAAAGCGTCGCCA
GACAGCGACGCTTTTTT

IgR.....: 3639 - Tamanho da IgR: 296 bp - Tamanho da Sequência: 112 bp

Promotor...: Posição: 705457 - Fita: Principal - Algoritmo: NNPP

Terminador: Posição: 705599 - Fita: Principal - Algoritmo: RNAMotif

Candidato 72

TCCTTTTGGAACTTAGTCTTTGTTTTAACAAAATAAGTTTTTCATTGGACGGGAAATTTT
CTACCGAAATAAAAAAGGCTGCATTCCAGGGAATGCAGCCTTTTT

IgR.....: 3979 - Tamanho da IgR: 153 bp - Tamanho da Sequência: 83 bp

Promotor...: Posição: 778328 - Fita: Principal - Algoritmo: NNPP

Terminador: Posição: 778441 - Fita: Principal - Algoritmo: TranstermHP

Candidato 73

CAGCGCAGGCATTGCCAACGAAGAGCGAATCTCCTATCATCTTCAAATAATAACGATTCT
CATCTGTGCGCGGCATGGGTGAGTGAAAGACAGCATTTCATGGCGGCCTGCGTATGCAG
GTCGTTTCTTTT

IgR.....: 317 - Tamanho da IgR: 375 bp - Tamanho da Sequência: 104 bp
Promotor...: Posição: 54052 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 53918 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 74

TACTTCGCCATTGAGACAGTGTCCGGAACCGATGCAAGAATCACGCCATGGTCCGGGCG
GCGAAGCACAAGGATTTTACACGCTGGTATGAGCTAGGAAGAAGCAGTTTCTGCAAACA
GGTGCCGAAAAAATCACTTTGGTTTCCATGTGGAAATATTTAGTTAACGTGATTTTCGAT
TACGTGCGTATTTGCTGATGTCATGACAGTCCGTTTGCATACGCAATTTCCGCATGCCCT
GCATCGCCGCGCCTGTATTGGCTTTGACGCAGCTTGCCTGCTATCAGAATGGCGCCCT
CCATTCTCAGGGGCTGCGCAATGGACATGGATGGCCCCTTGCGAATACTTATAAAAAA
TCGCCGGCTTGCACCGGCGATTTGTTTT

IgR.....: 3277 - Tamanho da IgR: 521 bp - Tamanho da Sequência: 356 bp
Promotor...: Posição: 620692 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 620306 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 75

GTTTTATTAAGAAATGTAAGAAAGCATTAGAGAATCTGGAAACCGGATTCACCCATTTGT
GTAGCGTGATTTTTCTGCACTCGAAAATAAAAAAGCAACTGCAAAAAACAACAGCCCG
CCGCCCGATATCTCAGTTTTTCTCTCTCCACCAATGAAAAAAGCGTCGCCAGACAGCG
ACGCTTTTTT

IgR.....: 3639 - Tamanho da IgR: 296 bp - Tamanho da Sequência: 165 bp
Promotor...: Posição: 705404 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 705599 - Fita: Principal - Algoritmo: TranstermHP

Candidato 76

GTTTTGGGAACAGTGGACGAAGAAGACTGTAGAAGGGACGCCGTTGAAGCTGTTAGCC
ACTTTTAGCAATTTTCAGAAACCAACAAAGACGGCGCTGGACGAAAATGGGGCAGAAGC
TGAAAAAGCGAATGCAGGAGCGGACTTTAGGAAGGTGCGATCGACTAGGGGAAAGCTTT
AAATCGATCTGGTATAACCAATGGACGATGTCGGCCGCAGAATCTCACCGCTTTTCA
GGGCACTGATTGTTGATCAAATAAAATAAATCGACCAATTGGTCGATTTT

IgR.....: 1501 - Tamanho da IgR: 661 bp - Tamanho da Sequência: 263 bp

Promotor...: Posição: 285125 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 285418 - Fita: Principal - Algoritmo: RNAMotif

Candidato 77

ATGGATTTGTGTTTACTGCAAATCCATACTAGCATTACGTATGGGAGCAACAATCCAGA
TAACATCCATACAAGCGTCAGCGCAACGCACTCACTCTCCGGTATCCACGTGGACCCAT
CCGCACTTCCCTGGCCTGGGTAGGCCGTTTT

IgR.....: 2426 - Tamanho da IgR: 241 bp - Tamanho da Sequência: 129 bp
Promotor...: Posição: 462532 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 462691 - Fita: Principal - Algoritmo: RNAMotif

Candidato 78

CTGGTTGTTGAGAGATCCGTCTGGTCCGATAAACCAGGCCGGTCCGGTCCGGTCATTCCG
GTACGCGATGCCGCCGGCAGGTGATGCAGGAAAACCCATCATATTGCCTGCGCTGCAA
CCCCTTCCCGGCCCTTGGGGCAGGATGTTGGAGGCCGGAACAGGCTTAAATGTACGGTC
TTCCAGCCGTGTGGACGACGAATTTTATCAGGGAATTTTCAATTTTCAAATTACGATAAT
CGATTTATAATACGAAAATATGCTGCAAAGCGGCATAAACAGGGGCTTCCAGGCCTC
GTATTTTT

IgR.....: 3270 - Tamanho da IgR: 555 bp - Tamanho da Sequência: 276 bp
Promotor...: Posição: 618211 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 618517 - Fita: Principal - Algoritmo: TranstermHP

Candidato 79

AAATCACTTTGGTTTCCATGTGGAAATATTTAGTTAACGTGATTTTCGATTACGTGCGTAT
TTGCTGATGTCATGACAGTCCGTTTGCATACGCAATTTGCCATGCCCTGCATCGCCGC
GCCTGTATTGGCTTTGACGCAGCTTGCCTGCTATCAGAATGGCGCCCTCCATTCTCAGG
GGCTGCGCAATGGACATGGATGGCCCCTTGCGAATACTTATAAAAAAATCGCCGGCTTG
CACCGGCGATTTGTTTT

IgR.....: 3277 - Tamanho da IgR: 521 bp - Tamanho da Sequência: 229 bp
Promotor...: Posição: 620565 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 620306 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 80

CATATTTGTTGAGAAATTTTTCCGGCTTTTCGTTAATATCTGCGACAAGCCTGACAGAC
AGGACGCAGCAACAGTCCTGCAGGGAAGTAGCTCATCGCGAGGCGTACCGCGCCGAT
GTAGAGCTCGTCTGTCGCAATGACAGGCGATGCGGTGAGAAAGATGCAACGGTAGCAA
TCATCAATCGGCACAACCAGCACCAGCCAGTACCAATCAGCAGCAACCGAACGTACCA
GAAGGTCGGCGACATCGATGGCGCTGGCGGCCGTGTCGCGCCGGTGCAGCAGGACT
CCGCAGGGAGCCGGGGACGGCGTTTGCCTTTTTT

IgR.....: 3829 - Tamanho da IgR: 546 bp - Tamanho da Sequência: 294 bp

Promotor...: Posição: 751415 - Fita: Reversa - Algoritmo: NNPP

Terminador: Posição: 751091 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 81

GGGCATTCTTTCAAATAGATGATGTTTGACGTTATCTTGAATATTTTCGAGCTATTTAT
CGTTATCTTCGGAAAATTGCCTCCTTTGAATTAAGGATGAACAACCTTCCCTCAAGATA
CTTTACCTTAGCTGCGGGTATGATTACCCCCGCTTCCCTTCCCTCCCTGAGCGG
GGCCGATCGTCAAGCGAAAGGCATTGCGCCCCGGAGAATATCCCTGTTCTCCGGGGT
TTTTTTT

IgR.....: 276 - Tamanho da IgR: 290 bp - Tamanho da Sequência: 220 bp

Promotor...: Posição: 49557 - Fita: Reversa - Algoritmo: NNPP

Terminador: Posição: 49307 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 82

TTTTTGATGCGCGACTTCTCGCGTGCCGAAAAATTTTAGGAAGTCTTTACAGATTGTGG
CGCTGCATTTTCTGTTGACGTTTCATGCGATACATTTGATCGGGCCTTAAGTTCAAATTA
CTTAAACTATTAAGTCGAGAAAAATGATAGTTGTAGTTGGAACGTGGCTAGTAATAGC
TTCCGGGTTTTCAGAAGCATCAGTTTCACTCCGAAAAGTACCGTCTCGGACTATTTCGGAG
TCTAAAATCAGGAATCTCCGGAATTTTTCGGTCTCGACAAATTCTCTTGCGCGAGCTCG
CGCAAGATTTT

IgR.....: 1065 - Tamanho da IgR: 365 bp - Tamanho da Sequência: 288 bp

Promotor...: Posição: 196212 - Fita: Principal - Algoritmo: NNPP

Terminador: Posição: 196530 - Fita: Principal - Algoritmo: RNAMotif

Candidato 83

GCTCTTTATACAGGTTGAACCGATGAGCTATATCGGGAAGAACCTGATAGCCGAGCGC
TAGGGATGTCCCTACAGGGGGAGGCCGAGATAATTGCCTCGCTTCTTGGCAGAACGC
CCGCTTGGTGCGGGCGTTCGTTT

IgR.....: 1534 - Tamanho da IgR: 184 bp - Tamanho da Sequência: 115 bp
Promotor...: Posição: 291043 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 291188 - Fita: Principal - Algoritmo: TranstermHP

Candidato 84

ATTTTTTGTCCCGAAACCATTCAAGCGGATTGAATTACTTCGGTAAATCTCCCGAAAGTG
CCGTTGATCCGTTGACAAGCAACGAATGGCTGGAGAACCCCAACGTGAAGCAACTGAA
GGTCGTTCAAACCAGCAATATCTTCGGCTGCCCTTTGGTGAGGGCGGAGAGGGGGAAG
GGGGCATTCCCCCTTTT

IgR.....: 2307 - Tamanho da IgR: 445 bp - Tamanho da Sequência: 172 bp
Promotor...: Posição: 439886 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 440088 - Fita: Principal - Algoritmo: RNAMotif

Candidato 85

CCGATAAGTTGCGAGACTAGCATCCTGCGTCCGGATAGAACAATATGCAGCAACGCAGC
ATACAGCGTCAAACAGACGCCAGCGGCCATCTGTTGCTTGACGAAACGACAATGCTG
CGAGCGCAGCAGCGGGCGGCGGTGACTGCAACAGCGGGCGGCGGAAAAACAAAACC
GCGCCGAAAGCGCGTTTTT

IgR.....: 2726 - Tamanho da IgR: 292 bp - Tamanho da Sequência: 170 bp
Promotor...: Posição: 518884 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 518684 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 86

TCGATTGACGAAAGTCAGTCCGGCCTTTTGTATTTTGGGGTCCACGAGCAAGGCGCCTGC
GCGCCTTGCGAGAGGGACGAGAAAGTCCGGCCTTACAAGGCCGGACGGGGTTCGCGG
GACGTGACCGAGTGCCTCAGTGCAGCAAACAAAAAGGTTCGATTGATGAAAATCAA
TCGGCCTTTTGCCTTTT

IgR.....: 2922 - Tamanho da IgR: 520 bp - Tamanho da Sequência: 162 bp
Promotor...: Posição: 553127 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 553319 - Fita: Principal - Algoritmo: RNAMotif

Candidato 87

CCGCTCTTGAAGGCGCGCGGCGAAGCGGCTACACTGCCGCCCTGACCCGCAACCGC
CCGGCTGCGGGTCGTTTCTTTT

IgR.....: 3251 - Tamanho da IgR: 118 bp - Tamanho da Sequência: 53 bp
Promotor...: Posição: 612717 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 612800 - Fita: Principal - Algoritmo: TranstermHP

Candidato 88

CAAGAGTAGCTGACTGGCGAAAAGTCCGCCCGCCGAAAATGTTTATTGCTGTCGGATTA
AAGCAGCTTGTGCTCTTTATTGCCGGCTTTCTGGCATT

IgR.....: 184 - Tamanho da IgR: 526 bp - Tamanho da Sequência: 77 bp
Promotor...: Posição: 35761 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 35654 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 89

GGTTTTTGCAGTGCAAAAGCGAGCTTTGGAAAAGTTACAAATGCAGCCGCCAGCGAGG
GCTATTGGATCGCATGGGAAAGCGATCGCAAAGGGAATGCGAGGCCATCGTCTGGCTC
GCATTGATGCAAATAATTTATAAGACGAGCCAGGGAACTTCAACACCAAAAATAATAAA
AAAGCGCCGCATGAACTCATGCGGCGCTTTT

IgR.....: 243 - Tamanho da IgR: 518 bp - Tamanho da Sequência: 186 bp
Promotor...: Posição: 43228 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 43444 - Fita: Principal - Algoritmo: RNAMotif

Candidato 90

CTTTTTTAGAAAATTTCTTCAGCAAACGGAATCGTTCGCAAGCTTACGCATGGCAATGGC
TGCCAGCCGGGCCGGCTATTTTT

IgR.....: 1535 - Tamanho da IgR: 1804 bp - Tamanho da Sequência: 60 bp
Promotor...: Posição: 291442 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 291532 - Fita: Principal - Algoritmo: RNAMotif

Candidato 91

ATACCATTCTATTGGTCATTTCGCTTAACGCAAGTACATTTAATTATTCAAATTGAATTAA
GACCAATTAATTATGCTTTTTATGGAATAGGGACGAAAAAATCGCAGTCCGAAGACTGC
GATTTT

IgR.....: 2230 - Tamanho da IgR: 166 bp - Tamanho da Sequência: 105 bp
Promotor...: Posição: 425930 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 425795 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 92

AATTTGCACCTAGAATCCGAGGCGAAACGTATCGTATCGGTCTGGGAAGGCGTGCTTTT
GTCATAAGAACAAGCAGTTCATCGCCTTGCGCGGCGCTGCCTTGGTTGCCGCAAATG
CGAGCCGTACATCCACGCCCTTCTTCTCGCCTGCGCTCGCGCAGGCGTCTTGCTTT

IgR.....: 2400 - Tamanho da IgR: 358 bp - Tamanho da Sequência: 146 bp
Promotor...: Posição: 458496 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 458672 - Fita: Principal - Algoritmo: RNAMotif

Candidato 93

TTTTGTTTTGATTGATAAGACTACAGCAGTCACTTTGCACCAACATCTGCGGGCATTCTG
TCATTTGAGTTGGAGTTAATCAATATTAATTTGAGTGGTAGTTAATCAGTATCAATTTGGT
ACCAATTGCTGGTGTGAATTTGTTTCGTTTTGCCTTACCTCACATTCCACATATTCCATAT
ATTCAGAAATAAATTCATGGTGCAATGGCGGCGCAGCGGTGTCATGCGCTGCCGGCGG
GCTGGAATTTGCACATTCTCATAGGGATAAGTGCTGTCTCTTTTT

IgR.....: 3677 - Tamanho da IgR: 922 bp - Tamanho da Sequência: 262 bp
Promotor...: Posição: 712708 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 713000 - Fita: Principal - Algoritmo: TranstermHP

Candidato 94

CTGTCATTTGAGTTGGAGTTAATCAATATTAATTTGAGTGGTAGTTAATCAGTATCAATTT
GGTACCAATTGCTGGTGTGAATTTGTTTCGTTTTGCCTTACCTCACATTCCACATATTCC
ATATATTCAGAAATAAATTCATGGTGCAATGGCGGCGCAGCGGTGTCATGCGCTGCCGG
CGGGCTGGAATTTGCACATTCTCATAGGGATAAGTGCTGTCTCTTTTT

IgR.....: 3677 - Tamanho da IgR: 922 bp - Tamanho da Sequência: 205 bp
Promotor...: Posição: 712765 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 713000 - Fita: Principal - Algoritmo: TranstermHP

Candidato 95

ATATTAATTTGAGTGGTAGTTAATCAGTATCAATTTGGTACCAATTGCTGGTGTGAATTT
GTTTCGTTTTGCCTTACCTCACATTCCACATATTCCATATATTCAGAAATAAATTCATGGTG
CAATGGCGGCGCAGCGGTGTCATGCGCTGCCGGCGGGCTGGAATTTGCACATTCTCAT
AGGGATAAGTGCTGTCTCTTTTT

IgR.....: 3677 - Tamanho da IgR: 922 bp - Tamanho da Sequência: 180 bp
Promotor...: Posição: 712790 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 713000 - Fita: Principal - Algoritmo: TranstermHP

Candidato 96

GAGGTGCGTTGACTGAAGATTGGCGGGTGTGGCAGCATTGGCTGCAGAACCGCAAT
CGCCCGGGCCGCAATTCTACCATTGGGCGGGCCACGGCCAGTTTT

IgR.....: 3931 - Tamanho da IgR: 213 bp - Tamanho da Sequência: 84 bp
Promotor...: Posição: 769276 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 769162 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 97

AATTTTGTATATGGTATCGCCCCTGATGCTCAATTAGCAATTCTCCGTTGCCAGCACAGC
AGGAACAAGCTTCTATGCCCTTCCACCTAACGCGCCTATGTCCGATAGCGCGCAAGT
GCAAAATTTCTTTGAGAATTATTTCAAAAAAGAAACCAATTCCCGCTACAAAAAGGGAA
ATCATCTGTTTCTTACGCCAACATTGGACGCAGATATGTTACTGTGTTTTT

IgR.....: 1548 - Tamanho da IgR: 722 bp - Tamanho da Sequência: 208 bp
Promotor...: Posição: 295336 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 295574 - Fita: Principal - Algoritmo: TranstermHP

Candidato 98

TAACGTGATTTTCGATTACGTGCGTATTTGCTGATGTCATGACAGTCCGTTTGCATACGC
AATTTGCCATGCCCTGCATCGCCGCGCCTGTATTGGCTTTGACGCAGCTTGCCTGCTA
TCAGAATGGCGCCCTCCATTCTCAGGGGCTGCGCAATGGACATGGATGGCCCCTTGCG
AATACTTATAAAAAATCGCCGGCTTGCACCGGCGATTTGTTTT

IgR.....: 3277 - Tamanho da IgR: 521 bp - Tamanho da Sequência: 195 bp
Promotor...: Posição: 620531 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 620306 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 99

GATAGTTGTAGTTGGAACGTGGCTAGTAATAGCTTCCGGGTTTCAGAAGCATCAGTTTC
ACTCCGAAAAGTACCGTCTCGGACTATTCGGAGTCTAAAAATCAGGAATCTCCGGAATT
TTTCGGTCTCGACAAATTCTTTCGCGGAGCTCGCGCAAGATTTT

IgR.....: 1065 - Tamanho da IgR: 365 bp - Tamanho da Sequência: 141 bp
Promotor...: Posição: 196359 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 196530 - Fita: Principal - Algoritmo: RNAMotif

Candidato 100

ATTTAATAGTTCTTTAATAGTTATTTGGTAGTTATTTAATGGTTGTTTGATTGAAACGCCGT
TTTTTTGCTGTAACCCGACGCAGTCCCCTCGTTGCAGTTCTCCGGTGTCCGCACCGATT
TGATATTT

IgR.....: 2930 - Tamanho da IgR: 361 bp - Tamanho da Sequência: 101 bp
Promotor...: Posição: 555267 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 555136 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 101

ATGTACGGTGTGTTGGCCTGATTTGCTTGACCAAATGATGCAACGCATTACTGTCCTGT
TTCCGCCTCCACATCTGGGGCGTTTTT

IgR.....: 3027 - Tamanho da IgR: 545 bp - Tamanho da Sequência: 66 bp
Promotor...: Posição: 572506 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 572410 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 102

GCGCAGGACGGTTGCAGTCCTGTTTGCCAGCCGGTAACATCTGATGCCCCGCTTGCGA
TGCTTCGCTTGCGGGGCATTTTT

IgR.....: 3302 - Tamanho da IgR: 179 bp - Tamanho da Sequência: 58 bp
Promotor...: Posição: 627230 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 627142 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 103

AATTGAATGCGAGCTATTTAATCCATATCAAAAATCCGAGTGTTATCCTTCCTGGCTAAA
CCCTTCAAGATCAGACCATATGGAGTTCTCGAATTTTCCCTAAGAGCAAAGTGTGTGTG
TTATTCCACAATAAAAACGAGTAATTAACAAAAGAGCGAAGGTGCGGTGTACAGTTGGTA
GCTTTTTTAGAAAATTTCTTCAGCAAACGGAATCGTTCGCAAGCTTACGCATGGCAATGG
CTGCCAGCCGGGCCGGCTATTTTT

IgR.....: 1535 - Tamanho da IgR: 1804 bp - Tamanho da Sequência: 241 bp
Promotor...: Posição: 291261 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 291532 - Fita: Principal - Algoritmo: RNAMotif

Candidato 104

CGCTATGTTAGCTTCACTCGCTGTTTCTAAACAACAAATTCCGCCCGTGATTTGAATATA
CAAACGCCCGCACAATGCGGGCGTTTTTT

IgR.....: 1548 - Tamanho da IgR: 722 bp - Tamanho da Sequência: 65 bp
Promotor...: Posição: 295140 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 295045 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 105

GGTCCTTGTGGTTGGGTGCTGTAGTTGCACCGGATATAGCTTTCCAGCAGGGATGATAT
GCGCGCCGAAGTCCAGGCGAAGAGGAAAACAGAAAGCAAAAACCTTCGTTTTAGACG
AAGTTTTT

IgR.....: 2738 - Tamanho da IgR: 196 bp - Tamanho da Sequência: 103 bp
Promotor...: Posição: 520927 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 520794 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 106

AATCTGGAAACCGGATTCACCCATTTGTGTAGCGTGATTTTTCTGCACTCGAAAATAAAA
AAAGCAACTGCAAAAAACAACAGCCCGCCCGCCGATATCTCAGTTTTTCTCTCTCCAC
CAATGAAAAAAGCGTCGCCAGACAGCGACGCTTTTTT

IgR.....: 3639 - Tamanho da IgR: 296 bp - Tamanho da Sequência: 133 bp
Promotor...: Posição: 705436 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 705599 - Fita: Principal - Algoritmo: TranstermHP

Candidato 107

GTATTGTACACCACCACAAACAAGATGACGAAAATTAGTGACTTTTTGTGTGGCATTTCGT
CCAATCTCGCCGCTGGCAGGGGCTAGAGCGCGGGCATGGTGGGCTTAACCGCTTTTT

IgR.....: 4098 - Tamanho da IgR: 415 bp - Tamanho da Sequência: 94 bp
Promotor...: Posição: 802099 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 802223 - Fita: Principal - Algoritmo: RNAMotif

Candidato 108

AAAATTAGTGACTTTTTGTGTGGCATTTCGTCCAATCTCGCCGCTGGCAGGGGCTAGAGC
GCGGGCATGGTGGGCTTAACCGCTTTTT

IgR.....: 4098 - Tamanho da IgR: 415 bp - Tamanho da Sequência: 64 bp
Promotor...: Posição: 802129 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 802223 - Fita: Principal - Algoritmo: RNAMotif

Candidato 109

CGATGGTGATACTTGAGGCACTACCTAAATAGTTTTCGGAGAGAACCAGCTATTTCCAG
ATTTGTTTAGCCTTTCACCCCTATCCACAGCTCATCCCCTAATTTTTCAACATTAGTGGGT
TCGGTCCTCCAGTGCCTGTTACCGCACCTTCAACCTGGCCATGGATAGATCATCTGGTT
TCGGGTCTACACCAGCGACTGAACGCCCTATTTCGACTCGATTTCTCTGCGCCTTCCC
TATTCGGTTAAGCTTGCCACTGAATGTAAGTCGCTGACCCATTATACAAAAGGTACGCA
GTCACCCCTTTCGAGGCTCCTACTGTTTGTATGCACACGGTTTCAGGATCTATTTCACTC
CCCTCCGGGGTTCTTTT

IgR.....: 1 - Tamanho da IgR: 3098 bp - Tamanho da Sequência: 350 bp
Promotor...: Posição: 1987 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 2367 - Fita: Principal - Algoritmo: RNAMotif

Candidato 110

AATTAATAAACTATTAAGTCGAGAAAAAATGATAGTTGTAGTTGGAACGTGGCTAGTA
ATAGCTTCCGGGTTTCAGAAGCATCAGTTTCACTCCGAAAAGTACCGTCTCGGACTATT
CGGAGTCTAAAAATCAGGAATCTCCGGAATTTTTCGGTCTCGACAAATTCTCTTGCGCG
AGCTCGCGCAAGATTTT

IgR.....: 1065 - Tamanho da IgR: 365 bp - Tamanho da Sequência: 173 bp
Promotor...: Posição: 196327 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 196530 - Fita: Principal - Algoritmo: RNAMotif

Candidato 111

ATTATTTCAAAAAAGAAACCAATTCCCGCTACAAAAGGGAAATCATCTGTTTCTTACGC
CAAACATTGGACGCAGATATGTTACTGTGTTTTT

IgR.....: 1548 - Tamanho da IgR: 722 bp - Tamanho da Sequência: 71 bp
Promotor...: Posição: 295473 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 295574 - Fita: Principal - Algoritmo: TranstermHP

Candidato 112

TGGTCCCTTTCTTTTCGACAGGACACGAGACCATGCATATTGCCCGCGCGCCCGCTCTTT
TTTCTCGCCGTTTTTATTGGCATTCTTTCTGCAGTTCTTTCTGCAGTTCTTT

IgR.....: 17 - Tamanho da IgR: 485 bp - Tamanho da Sequência: 89 bp
Promotor...: Posição: 5875 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 5756 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 113

AGGTTGTAGCGATTGCAAGTGCTGGAAGGGAGAATTTTCGACGAGGCTGGAGGTGGTGT
TTGACGGTCCCGGACAGCTTGACCTATGCGTATGCATACGCAAATTATCCGGCTAGTTT
TAATTAActCTTTATACGGTTTATATCATTTAAAAAATAGCCAGAAAATAAAAAAGCGCGT
GCATTCCTGCACGCGCTTTT

IgR.....: 230 - Tamanho da IgR: 275 bp - Tamanho da Sequência: 176 bp
Promotor...: Posição: 40959 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 41165 - Fita: Principal - Algoritmo: RNAMotif

Candidato 114

CTGCGATTGGGCGGGGCTGATGTTTCGTACTACCATAAATTATGGGCCTTGCAGACGCAA
AACCCACATCGAAGCCGAGCATTATAGCCCGTAAACA AAAAGCGCACAACCGAGGTTGT
GCGCTTTT

IgR.....: 310 - Tamanho da IgR: 150 bp - Tamanho da Sequência: 104 bp
Promotor...: Posição: 52628 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 52762 - Fita: Principal - Algoritmo: TranstermHP

Candidato 115

GACGAATTTTATCAGGGAATTTTCAATTTTCAAATTACGATAATCGATTTTCATAATACGAA
AATATGCTGCAAAGCGGCATAAACAGGGGCTTTCCAGGCCTCGTATTTT

IgR.....: 3270 - Tamanho da IgR: 555 bp - Tamanho da Sequência: 86 bp
Promotor...: Posição: 618401 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 618517 - Fita: Principal - Algoritmo: TranstermHP

Candidato 116

CGATTTACAGCCACTATCCCAAGGTCGTCATCATCAGTCATGGCCGGCGGGCGGATT
GGCCTGCCCGTTGCCCGCGTCGGCGTTTTT

IgR.....: 3989 - Tamanho da IgR: 640 bp - Tamanho da Sequência: 66 bp
Promotor...: Posição: 779741 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 779837 - Fita: Principal - Algoritmo: TranstermHP

Candidato 117

AAGCAACTGAAGGTCGTTCAAACCAGCAATATCTTCGGCTGCCCTTTGGTGAGGGCGG
AGAGGGGGAAGGGGGCATTCCCCCTTTT

IgR.....: 2307 - Tamanho da IgR: 445 bp - Tamanho da Sequência: 65 bp
Promotor...: Posição: 439993 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 440088 - Fita: Principal - Algoritmo: RNAMotif

Candidato 118

ATTTCAATTTCAAGTGACAGCGTGTGACCCAACGACAGAGTCCCCTCACCGTCGCTTGCC
TCGCGGAAAAATTCCGCAGCGGTATTAACAAGCCGCTAACCCAGGTGTCAAGCGAG
TCTGCGGCCTATGCCGCGATTTT

IgR.....: 2314 - Tamanho da IgR: 604 bp - Tamanho da Sequência: 118 bp
Promotor...: Posição: 441893 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 441745 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 119

AAATTAATTAATTTATTATCATTCTCTGGTAAGATTGTCACTGTAACCAGAACGGTCAGCG
GGAATCCGTAGCAAACGTCGTGACGGCGGGACGATGCGCATCACCAGCATGCGGTCAT
ACGCGTGTTTTT

IgR.....: 3466 - Tamanho da IgR: 666 bp - Tamanho da Sequência: 109 bp
Promotor...: Posição: 661206 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 661345 - Fita: Principal - Algoritmo: TranstermHP

Candidato 120

AAATATTAGGCAGAGGAAATTCTGTGAGATCTATTAAGTCCCTAGAATGAAGTTCTGCC
CATGGTTTTCTTGGGAGGATCGAAAGGATATTCCTCTGTCTGAAATGCCGGGAATTTAC
CTTCTCTCAATCACAAAAAAGAACCTAGCCGGATCATTACCAGACTGGAAAGATGTTTCA
TATATCGGCATGACCAATTCACGTCAGGGATTGATAGGGCGCTGGCAGCAGTTCTTTAA
CTCAATTGAGGCACACCTGGGCATAGTGGCGGCAACGCCGTTTT

IgR.....: 3484 - Tamanho da IgR: 996 bp - Tamanho da Sequência: 261 bp
Promotor...: Posição: 665953 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 666244 - Fita: Principal - Algoritmo: RNAMotif

Candidato 121

ATTCCAGAAATTCAAAACATCAGATCAGGAAATTA AAAAGATTCCGTTGTAATTTAAAGA
TGACGCATCGCGAAGAGTCATCGACGCGTCGGGAAGAAAGAACGAGCCAGTGACACCA
TCGGACAGCAAGGAATTCGGCATCGCCGGTTTT

IgR.....: 3782 - Tamanho da IgR: 357 bp - Tamanho da Sequência: 130 bp
Promotor...: Posição: 738069 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 737909 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 122

AGCTATTTAATCCATATCAAAAATCCGAGTGTTATCCTTCCTGGCTAAACCCTTCAAGAT
CAGACCATATGGAGTTCTCGAATTTCCCTAAGAGCAAAAGTGTGTGTGTTATTCCACAA
TAAAACGAGTAATTAACAAAAGAGCGAAGGTCGGTGTACAGTTGGTAGCTTTTTTTAGA
AAATTTCTTCAGCAAACGGAATCGTTTCGCAAGCTTACGCATGGCAATGGCTGCCAGCCG
GGCCGGCTATTTTT

IgR.....: 1535 - Tamanho da IgR: 1804 bp - Tamanho da Sequência: 230 bp
Promotor...: Posição: 291272 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 291532 - Fita: Principal - Algoritmo: TranstermHP

Candidato 123

TTGGAGACTTTAGCAAACAATCGCATTCTGGAACGAAAATTCCGCTTCGCAATTTTTTAA
CGACCGAGGCCCAATAAGGGCTTACGAGGCTCTAAATTTCTCAATTGCGAAGCAAAAAT
CGGCCTAGCTATCTGATTTCTATAGTTTTTCGCACCGGACTTGTAATCAGGCGCTCCGCA
CGAGAAAACGCCCGCGACCTGCGGGCGTTTTTT

IgR.....: 1537 - Tamanho da IgR: 309 bp - Tamanho da Sequência: 187 bp
Promotor...: Posição: 293468 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 293251 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 124

CAGCATGCCTTGTTGATATTTATAAATTCCCTCGCAGATTCTTAATCAACAGGGCAGGGC
GCTAAGCTAGTTTTACACGCTGTTTCCGAATAACAAATCCTGCCAGGATTTGAGAGA
TAATTCCCAGCTTGCTTGCTGGGGAATATTTTGCCCAAACGAGCGCAAACCGCGAATT
GGCGCACTGTATGCGTAACCAGTGGCCCGCGTCGGATGGGTCTTTGCGGCGTGCCGT
CAACATGCGCTTTTGA CTCATAATCCGTTGGTGCCGTATTGACTCACGGGGAGGGCTG
CCACTATCTAAAGGGGGATTCCGAGAGATCGGAGTCCCTTTTTTTTT

IgR.....: 3742 - Tamanho da IgR: 666 bp - Tamanho da Sequência: 314 bp

Promotor...: Posição: 728090 - Fita: Reversa - Algoritmo: NNPP

Terminador: Posição: 727746 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 125

GAAATGCTGGAAAACGGCCTAAACTTGGTAATAAAGAAATATTTTATGTCTATGGCCCG
CAGAGAAATCTTGCGGGCCATTTCTTTTTTTTT

IgR.....: 1030 - Tamanho da IgR: 246 bp - Tamanho da Sequência: 62 bp

Promotor...: Posição: 190229 - Fita: Principal - Algoritmo: NNPP

Terminador: Posição: 190321 - Fita: Principal - Algoritmo: RNAMotif

Candidato 126

TTTCTCAATTGCGAAGCAAAAATCGGCCTAGCTATCTGATTTCTATAGTTTTCGCACCGG
ACTTGTAATCAGGCGCTCCGCACGAGAAAACGCCCGCGACCTGCGGGCGTTTTTTTG

IgR.....: 1537 - Tamanho da IgR: 309 bp - Tamanho da Sequência: 91 bp

Promotor...: Posição: 293372 - Fita: Reversa - Algoritmo: NNPP

Terminador: Posição: 293251 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 127

GGCCGTTGCCTTGTTTTACATCCTGAAGCCATATTAAGTGATTTAAACGAATAAAACAAA
TTAATTAATTTATTATCATTCTCTGGTAAGATTGTCACTGTAACCAGAACGGTCAGCGGG
AATCCGTAGCAAACGTCGTGACGGCGGGACGATGCGCATCACCAGCATGCGGTCATAC
GCGTGGTTTT

IgR.....: 3466 - Tamanho da IgR: 666 bp - Tamanho da Sequência: 166 bp

Promotor...: Posição: 661149 - Fita: Principal - Algoritmo: NNPP

Terminador: Posição: 661345 - Fita: Principal - Algoritmo: TranstermHP

Candidato 128

AGAAATCTTGACATCAAGCAAACACTGACTCGCAATTTTATGGGGGATTTTCGCTGCGTC
GCGAACGCTGACAGGCAGGCAGAATGACTGACAAAACACTGACACCTCCACACGTCTT
CCACACTCTCCGCACAAAATCCTTGCATTCTGCTCACACGACAAAACATCCGACCGCC
CGAATTTCCGCTTCCGGCCTGGCTCAAATGCCTTCCCTTGCCTTGCCTTGCCTGCGGCGGGCCTG
GGCAGGAATATTTTT

IgR.....: 1494 - Tamanho da IgR: 422 bp - Tamanho da Sequência: 224 bp

Promotor...: Posição: 283580 - Fita: Reversa - Algoritmo: NNPP

Terminador: Posição: 283326 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 129

GTTCCGGTGAAAAACGCGAACGCAATTTTGTATATGGTATCGCCCCTGATGCTCAATTAG
CAATTCTCCGTTGCCAGCACAGCAGGAACAAGCTTCCTATGCCCTTCCCACCTAACGCG
CCTATGTTCGGATAGCGCGCAAGTGCAAAATTTCTTTTGAGAATTATTTCAAAAAGAAAC
CAATTTCCGCTACAAAAGGGAAATCATCTGTTTCTTACGCCAACATTGGACGCGAGATA
TGTTACTGTGTTTTT

IgR.....: 1548 - Tamanho da IgR: 722 bp - Tamanho da Sequência: 230 bp

Promotor...: Posição: 295314 - Fita: Principal - Algoritmo: NNPP

Terminador: Posição: 295574 - Fita: Principal - Algoritmo: TranstermHP

Candidato 130

AGGTCGATCGCGTGAAATTAGCCAGGAGACCGGGAATGATTGCCCGCCTCCGCGCCTG
ATTTCCGCTGCCCTAGGAGCCGTTGATACCGGCTTTTTT

IgR.....: 1805 - Tamanho da IgR: 361 bp - Tamanho da Sequência: 73 bp

Promotor...: Posição: 346251 - Fita: Reversa - Algoritmo: NNPP

Terminador: Posição: 346148 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 131

TAGAAAGTTTTTACATATTGCGGAACTGGACCGCTACTGTTCTGTCTATGCGATACGTTT
GAGCAACATTAAGTTCAACTTACGAAAAACAATTAACCTAAAACTTAAACTTAATTAAT
TAGTTGGTTGCCAGCCTCATCAACCGGCTGGTCGCGATTGGAGCGCGGGATGGATAAA
AAGTGCATTTTTTCGCACAAATTTTCGCGCAACGGTCGCTCAGGGAATTCATTGATGTCT
CAACAGTAAGCGTTCCGAGATCTCCGGGTGCTGCATGGATCGTCTGCGCGCCCTGCGC
GCAGATGTTTTTTTTT

IgR.....: 3884 - Tamanho da IgR: 507 bp - Tamanho da Sequência: 286 bp

Promotor...: Posição: 760185 - Fita: Reversa - Algoritmo: NNPP

Terminador: Posição: 759869 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 132

TCCGAATGTGGGTTCAAACAGCTTGGTCTTTGTTAGCGTGCGCAGGTTCTGCCGCCTG
CTTGGGCCACGGCGGACAGCGCCGCGTGGCGTCTTT

IgR.....: 173 - Tamanho da IgR: 137 bp - Tamanho da Sequência: 73 bp

Promotor...: Posição: 34072 - Fita: Reversa - Algoritmo: NNPP

Terminador: Posição: 33969 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 133

ACATTTAATTATTCAAATTGAATTAAGACCAATTAATTATGCTTTTTATGGAATAGGGACG
AAAAAATCGCAGTCCGAAGACTGCGATTTT

IgR.....: 2230 - Tamanho da IgR: 166 bp - Tamanho da Sequência: 70 bp

Promotor...: Posição: 425895 - Fita: Reversa - Algoritmo: NNPP

Terminador: Posição: 425795 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 134

CGAAGAACAGGTGGAGTGGTTCACCGACCTGTTCTACGATGGCCTGACCCCGGAAGAA
GCCGCCAAGGAAGCCCTGGCCTGATCCTCATCAGTGGCCGCCCTACCCCCACCCTCAC
CCGTATCCCGGCCCGCAAGGCCGGGATTTT

IgR.....: 2261 - Tamanho da IgR: 345 bp - Tamanho da Sequência: 124 bp

Promotor...: Posição: 432131 - Fita: Reversa - Algoritmo: NNPP

Terminador: Posição: 431977 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 135

GCATTCACGTATGGGAGCAACAATCCAGATAACATCCATACAAGCGTCAGCGCAACGCA
CTCACTCTCCGGTATCCACGTGGACCCATCCGCACTTTCCCTGGCCTGGGTAGGCCGT
TTT

IgR.....: 2426 - Tamanho da IgR: 241 bp - Tamanho da Sequência: 98 bp

Promotor...: Posição: 462563 - Fita: Principal - Algoritmo: NNPP

Terminador: Posição: 462691 - Fita: Principal - Algoritmo: TranstermHP

Candidato 136

GTAATTAACAAAAGAGCGAAGGTCCGGTGTACAGTTGGTAGCTTTTTTAGAAAATTTCTT
CAGCAAACGGAATCGTTCGCAAGCTTACGCATGGCAATGGCTGCCAGCCGGGCGGCT
ATTTTT

IgR.....: 1535 - Tamanho da IgR: 1804 bp - Tamanho da Sequência: 101 bp
Promotor...: Posição: 291401 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 291532 - Fita: Principal - Algoritmo: TranstermHP

Candidato 137

ATTGCTAATTGAGCATCAGGGGCGATACCATATACAAAATTGCGTTCGCGTTTTTCACCG
AACGGTAGAGATTACTATCGCATTTCGCAACGGAAGTTCTTTTCCACAACCTCCAGATCGA
GCGGGACTATAACCGGCTTTGTGAGCCGCTGACTTTTTTGAATTTGTGGAGCAAGAATC
CACATAAACTATTGATCTAATGAAATTAATTCTAGATTTTCCGTCCCTTGGGCAGGCGCG
CTATGTTAGCTTCACTCGCTGTTTCTAAACAACAAATTCGCCCGTGATTTGAATATACA
AACGCCCGCACAAATGCGGGCGTTTTTT

IgR.....: 1548 - Tamanho da IgR: 722 bp - Tamanho da Sequência: 302 bp
Promotor...: Posição: 295377 - Fita: Reversa - Algoritmo: NNPP
Terminador: Posição: 295045 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 138

GGGCTGATTGTATTTTTCTCTGGCATCAAGAGATTTTTTCAATCAGCCACAAGGGAAAAC
AGAGCTAAAACTGCTGCGCATTGTATGCCAGGCGCGTGCGGCGCTGCCGCCTTTTTT

IgR.....: 3783 - Tamanho da IgR: 445 bp - Tamanho da Sequência: 93 bp
Promotor...: Posição: 738167 - Fita: Principal - Algoritmo: NNPP
Terminador: Posição: 738290 - Fita: Principal - Algoritmo: RNAMotif

Candidato 139

TAATCGATGGTCTTTTTCAACTCTGCCCGTGCGTTTTATTTGTGTAAGCGGGAGGTGCC
TGATGCACTCGATTTAGAAAGTTTTTACATATTGCGGAACTGGACCGCTACTGTTCTGTC
TATGCGATACGTTTGAGCAACATTAAGTTCAACTTACGAAAAACAATTAACTTAAAACTT
AAAACCTAATTAATTAGTTGGTTGCCAGCCTCATCAACCGGCTGGTCGCGATTGGAGCG
CGGGATGGATAAAAAGTGCATTTTTTCGCACAAATTTTCGCGCAACGGTCGCTCAGGGA
ATTCATTGATGTCTCAACAGTAAGCGTTTCGGAGATCTCCGGGTGCTGCATGGATCGTCT
GCGCGCCCTGCGCGCAGATGTTTTTTTTT

IgR.....: 3884 - Tamanho da IgR: 507 bp - Tamanho da Sequência: 359 bp
Promotor...: Posição: 760258 - Fita: Reversa - Algoritmo: NNPP

Terminador: Posição: 759869 - Fita: Reversa - Algoritmo: TranstermHP

Candidato 140

TTTCCTGCTGATTCAATGCGGGCAATTTGGGGAGGATAAGAAAACCGCTGTTCCCCTTG
CAAGGAACAGCGGTTTTT

IgR.....: 1302 - Tamanho da IgR: 109 bp - Tamanho da Sequência: 54 bp

Promotor...: Posição: 246678 - Fita: Reversa - Algoritmo: NNPP

Terminador: Posição: 246594 - Fita: Reversa - Algoritmo: RNAMotif

Candidato 141

CTTTTTTCGACATCCGCCGAAGCCAAACTTTCAATGCTCGCTCACCTATCACAATAAAAA
AAGGAAGCTTTGTTAGGCTTCCTTTTTT

IgR.....: 2849 - Tamanho da IgR: 151 bp - Tamanho da Sequência: 64 bp

Promotor...: Posição: 540034 - Fita: Principal - Algoritmo: NNPP

Terminador: Posição: 540128 - Fita: Principal - Algoritmo: TranstermHP
