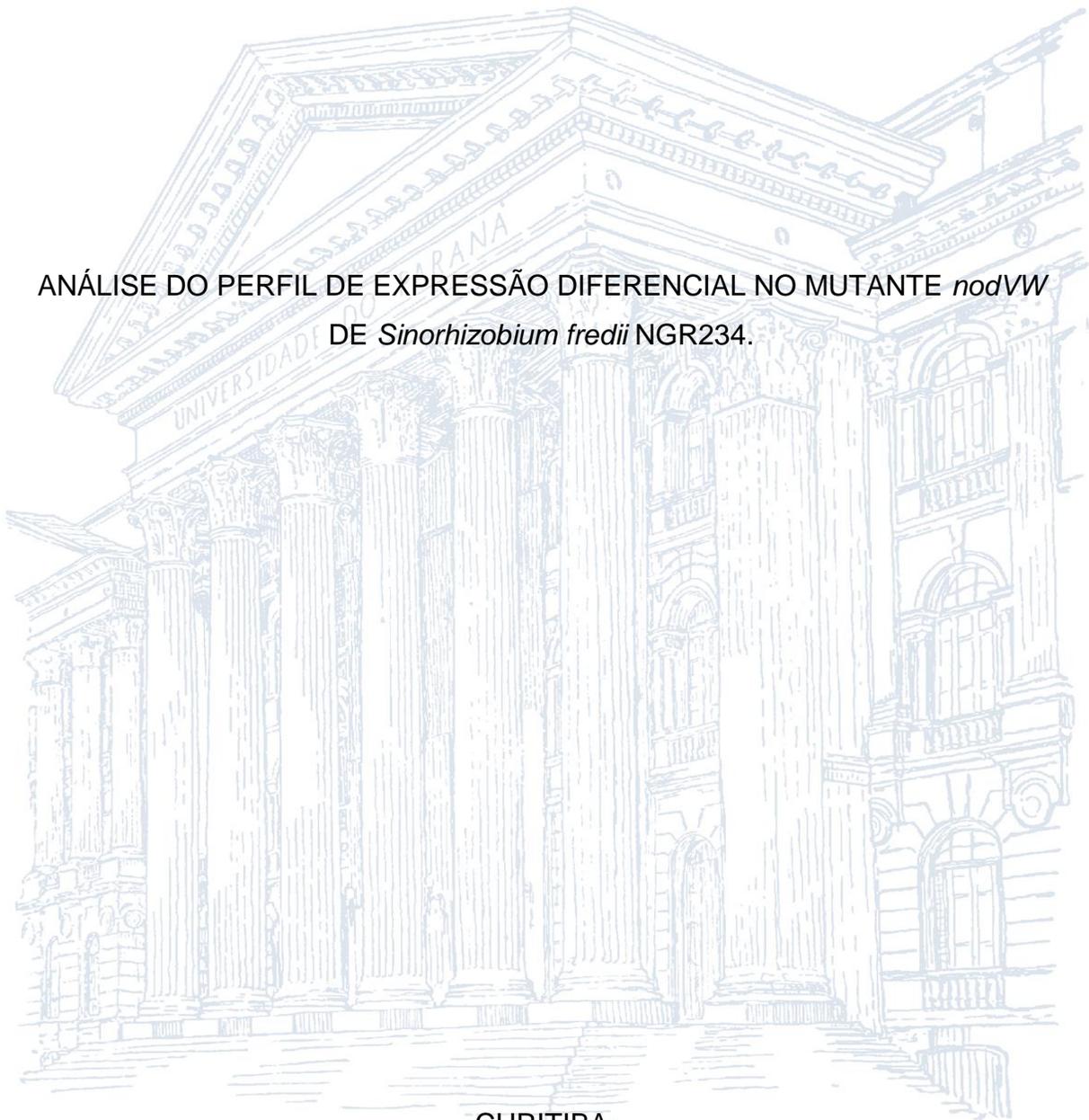


UNIVERSIDADE FEDERAL DO PARANÁ

IRIEL ARACELI JOERIN LUQUE

ANÁLISE DO PERFIL DE EXPRESSÃO DIFERENCIAL NO MUTANTE *nodVW*  
DE *Sinorhizobium fredii* NGR234.



CURITIBA

2017

UNIVERSIDADE FEDERAL DO PARANÁ

IRIEL ARACELI JOERIN LUQUE

ANÁLISE DO PERFIL DE EXPRESSÃO DIFERENCIAL NO MUTANTE *nodVW*  
DE *Sinorhizobium fredii* NGR234.

Dissertação apresentada como requisito parcial  
à obtenção do grau de Mestre em Genética ao  
Programa de Pós-Graduação em Genética,  
Setor de Ciências Biológicas, Universidade  
Federal do Paraná.

Orientadora: Prof<sup>ª</sup>. Dr<sup>ª</sup>. Roseli Wassem  
Co-orientadora: Prof<sup>ª</sup>. Dr<sup>ª</sup>. Ana Claudia Bonatto

CURITIBA

2017

Universidade Federal do Paraná  
Sistema de Bibliotecas

Joerin Luque, Iriel Araceli  
Análise do perfil de expressão diferencial no mutante *nodVW* de  
*Sinorhizobium fredii* NGR234./ Iriel Araceli Joerin Luque. – Curitiba, 2017.  
118 f.: il. ; 30cm.

Orientadora: Roseli Wassem  
Co-orientadora: Ana Cláudia Bonatto  
Dissertação (Mestrado) - Universidade Federal do Paraná, Setor de  
Ciências Biológicas. Programa de Pós-Graduação em Genética.

1. Rizóbio. 2. Nitrogênio – Fixação I. Título II. Wassem, Roseli. III.  
Bonatto, Ana Cláudia. IV. Universidade Federal do Paraná. Setor de  
Ciências Biológicas. Programa de Pós-Graduação em Genética.

CDD (20. ed.) 589.95



MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE FEDERAL DO PARANÁ  
SETOR DE CIÊNCIAS BIOLÓGICAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA  
CENTRO POLITÉCNICO - JARDIM DAS AMÉRICAS - CAIXA  
POSTAL 19071 - CEP 81531-990 - CTBA, BRASIL 41 3361 1587  
PPG-GEN@UFPR.BR



## PARECER

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Genética da Universidade Federal do Paraná foram convocados para realizar a arguição da dissertação de mestrado de IRIEL ARACELI JOERIN LUQUE, intitulada: "Análise do perfil de expressão diferencial no mutante nodVW de *Sinorhizobium fredii* NGR234", após terem inquirido a aluna e realizado a avaliação do trabalho, são de parecer pela sua aprovação.

Curitiba, 31 de março de 2017

**Doutora Leda Satie Chubatsu**  
Membro Titular

**Doutora Anelis Maria Marin**  
Membro Titular

**Doutora Roseli Wassem**  
Presidente

## **AGRADECIMENTOS**

Agradeço à minha orientadora Roseli Wassem pela possibilidade de trabalhar neste projeto e por todos os ensinamentos, a grande dedicação e a paciência brindados durante estes dois anos de trabalho.

À minha coorientadora Ana Claudia Bonatto pela disponibilidade e a ajuda que sempre esteve disposta a oferecer.

À Ana Luiza e aos meus colegas do laboratório pela parceria, por acompanhar o meu dia a dia com risadas e conversas, e pela ajuda em todos os momentos em que precisei.

À todos os bons amigos que tive o privilégio de encontrar durante este caminho e que me acompanharam em tantos momentos.

Aos meus pais e irmão por todo o seu apoio, paciência e pela força que me deram para chegar até aqui. Ao meu namorado pelo seu amor e pela companhia sem a qual teria sido muito mais difícil alcançar este objetivo.

Ao pessoal dos departamentos de Genética, Bioquímica e Biologia celular pela colaboração em muitas etapas deste trabalho.

À Universidade Federal do Paraná e a CAPES que tornaram possível a realização deste mestrado.

## RESUMO

O processo de fixação biológica de nitrogênio atmosférico por rizóbios é de grande importância uma vez que reduz a necessidade de fertilizantes nitrogenados em culturas de interesse agrônomico. Entre os rizóbios, *S. fredii* NGR234 tem sido amplamente estudado devido a sua grande faixa hospedeira. O intercâmbio de sinais moleculares requerido para uma nodulação ótima de NGR234 em leguminosas tem sido muito estudado, mas alguns fatores nele ainda têm função desconhecida. Entre eles o par regulatório NodV e NodW. Através do estudo do transcriptoma do mutante *nodVW*, em comparação com o transcriptoma da estirpe selvagem, foram determinados alguns alvos de regulação por este sistema. Dentre eles os operons NGR\_b03250-NGR\_b03190, *rmlB-wbgA*, NGR\_c06030-NGR\_c06080, *nopB-rhcU*, *nopC*-NGR\_a00520, e os genes *nopT*, *nopJ*, *fixF*, *nodU* e *rgpF* que são modulados por NodV e NodW em presença de flavonoides, enquanto que os operons NGR\_c26130-NGR\_c26200, NGR\_c26090-NGR\_26120, e os genes NGR\_c30010 e *hmgA* são modulados na ausência destas substâncias. Os operons *pqqEDCBA*, *cysHDN*, *cysG*-NGR\_c13080, *thuEFGKAB* e o gene *subB* são modulados nas duas condições. Através do uso de fusões transcricionais das regiões regulatórias dos genes *nopJ*, *nopB-rhcU*, *nopC*-NGR\_a00520, *nopT*, *fixF* e *pqqEDCBA* com o gene repórter de proteína fluorescente verde (GFP), a expressão diferencial destes foi confirmada. Ademais, foi detectada a ligação de NodW à região regulatória de *nopB*, *in vitro*, sugerindo que o sistema NodV-NodW pode participar diretamente da ativação deste gene. Além disso, foi feita uma nova descrição dos genes modulados por flavonoides em NGR234, confirmando o descrito por outros autores, através de RNA-seq. Desta maneira, este trabalho agrega informação para um maior entendimento do diálogo molecular rizóbio-leguminosa e abre portas para novas pesquisas no futuro.

**Palavras-chave:** *Sinorhizobium fredii* NGR234, *nodV*, *nodW*, RNA-seq.

## ABSTRACT

The process of biological nitrogen fixation by rhizobia is of great importance since it reduces the need for nitrogen fertilizers in crops of agronomic interest. Among rhizobia, *S. fredii* NGR234 has been extensively studied due to its large host range. The exchange of molecular signals required for optimal nodulation of NGR234 in legumes has been well studied, but some factors in it still have unknown function. Among them is the NodV and NodW regulatory pair. Through the transcriptome study of the *nodVW* mutant in comparison with the wild-type transcriptome some regulatory targets have been described for this system. Among them, the operons, NGR\_b03250-NGR\_b03190, *rmlB-wbgA*, NGR\_c06030-NGR\_c06080, *nopB-rhcU*, *nopC*-NGR\_a00520, and the *nopT*, *nopJ*, *fixF*, *nodU* and *rgpF* genes are modulated by NodV and NodW in the presence of flavonoids, while the operons NGR\_c26130-NGR\_c26200, NGR\_c26090-NGR\_26120, and the genes NGR\_c30010 and *hmgA* are in the absence of these substances. The *pqqEDCBA*, *cysHDN*, *cysG*-NGR\_c13080, *thuEFGKAB* and *suhB* gene operons are modulated under both conditions. Through the use of transcriptional fusions of the regulatory regions of the *nopJ*, *nopB-rhcU*, *nopC*-NGR\_a00520, *nopT*, *fixF* and *pqqEDCBA* genes with the green fluorescent protein (GFP) reporter gene, the differential expression of these genes has been confirmed. In addition, NodW binding to the regulatory region of *nopB-rhcU* has been detected *in vitro*, suggesting that the NodV-NodW system can participate directly in the activation of this gene. Additionally, a new description of the flavonoid-modulated genes in NGR234 has been made, confirming the one described by other authors, through RNA-seq. Therefore, this work aggregates information for a better understanding of rhizobial-legume molecular conversation, and opens the door to further research in the future.

**Key-words:** *Sinorhizobium fredii* NGR234, *nodV*, *nodW*, RNA-seq.

## LISTA DE FIGURAS

Figura 1- SISTEMA DE SECREÇÃO DE TIPO TRÊS DE <i>S. fredii</i> NGR234.....	23
Figura 2- RESUMO DA CASCATA REGULATÓRIA INDUZIDA POR FLAVONOIDES EM <i>S. fredii</i> NGR234.....	24
Figura 3- REGIÃO <i>nodVW</i> DE pNGR234b.....	26
Figura 4- ANÁLISE DA REGIÃO GENÔMICA QUE CONTÉM OS GENES <i>nodV</i> E <i>nodW</i> .....	44
Figura 5- MODELO DA ESTRUTURA DA PROTEÍNA NodV MOSTRANDO OS DOMÍNIOS CONSERVADOS. ....	46
Figura 6- COMPARAÇÃO DAS ESTRUTURAS DAS PROTEÍNAS NodW E OS POLIPEPTIDEOS CODIFICADOS PELAS ORFs NGR_b16850 E NGR_b16860.....	47
Figura 7- DISPOSIÇÃO GÊNICA DO OPERON <i>thuEFGKA</i> E DO GENE <i>thuR</i> NO PLASMÍDEO pNGR234b. ....	59
Figura 8- DISPOSIÇÃO GÊNICA DOS OPERÕES <i>pqqEDCBA</i> E NGR_b03250-NGR_b03190 NO PLASMÍDEO pNGR234b DE <i>S. fredii</i> NGR234.. ....	60
Figura 9- DISPOSIÇÃO GÊNICA DOS OPERONS NGR_c26130-NGR_c26200 E NGR_c26090-NGR_c26120 NO CROMOSSOMO DE <i>S. fredii</i> NGR234. ....	68
Figura 10- NÍVEIS DE EXPRESSÃO DE GFP CONTROLADOS PELA REGIÃO REGULATÓRIA DE <i>nopJ</i> (TB4).....	70
Figura 11- NÍVEIS DE EXPRESSÃO DE GFP CONTROLADOS PELAS REGIÕES REGULATÓRIAS DOS GENES <i>nopB-rhcU</i> (TB8), <i>nopC</i> -NGR_A00520 (TB10) E <i>nopT</i> (TB11).....	73
Figura 12- NÍVEIS DE EXPRESSÃO DE GFP CONTROLADOS PELA REGIÃO REGULATÓRIA DO GENE <i>fixF</i> (NB6). ....	74
Figura 13- NÍVEIS DE EXPRESSÃO DE GFP CONTROLADOS PELAS REGIÕES REGULATÓRIAS DOS GENES <i>thuR</i> , <i>thuEFGKA</i> . ....	76
Figura 14- NÍVEIS DE EXPRESSÃO DE GFP CONTROLADOS PELAS REGIÕES REGULATÓRIAS DOS GENES <i>cysHDN</i> E <i>pqqEDCBA</i> . ...	78

Figura 15- RESTRIÇÃO ENZIMÁTICA DE pET28-nodWCterm-His E	
pET29-nodW-C-His.....	79
Figura 16- PURIFICAÇÃO DAS PROTEÍNAS DE FUSÃO NodWCterm-His E	
NodW-C-His. ....	81
Figura 17- ENSAIO DE EMSA UTILIZANDO A REGIÃO TB8 COMO ALVO.....	82
Figura 18- RESUMO DOS GENES MODULADOS POR NodV-NodW EM	
PRESENÇA DE FLAVONOIDES. ....	85

## LISTA DE TABELAS

Tabela 1- ESTIRPES BACTERIANAS UTILIZADAS NESTE TRABALHO.....	29
Tabela 2- PLASMÍDEOS UTILIZADOS NESTE TRABALHO.....	30
Tabela 3- PRIMERS UTILIZADOS NESTE TRABALHO.....	31
Tabela 4- RESUMO DOS DADOS DE FILTRO DE QUALIDADE E MAPEAMENTOS DE TODAS AS BIBLIOTECAS UTILIZADAS USANDO CLC GENOMICS WORKBENCH.....	50
Tabela 5- GENES DIFERENCIALMENTE EXPRESSOS NA COMPARAÇÃO <i>nodVW</i> VS NGR234 COM FLAVONOIDES .....	56
Tabela 6- GENES DIFERENCIALMENTE EXPRESSOS NA COMPARAÇÃO <i>nodVW</i> VS NGR234 SEM FLAVONOIDES .....	57
Tabela 7- <i>tts-boxes</i> E <i>nod-boxes</i> MODULADOS POR NodV-NodW.....	64

## LISTA DE ABREVIATURAS

AIA	ácido-3-indolacético
Amp	ampicilina
ATP	adenosina 5' trifosfato
cDNA	<i>complementary deoxyribonucleic acid</i>
DE	diferencialmente expressos
DNase	desoxiribonuclease
EMSA	<i>electrophoretic mobility shift assay</i>
FBN	fixação biológica de nitrogênio
Fix+	fenótipo simbiótico capaz de fixar nitrogênio
GFP	<i>green fluorescent protein</i>
Gm	gentamicina
HK	<i>histidine kinase</i>
hpi	horas pós-indução
HTH	hélice-volta-hélice
IPTG	isopropil- $\beta$ -D-1-tiogalactopiranósido
Km	canamicina
LPS	lipo-polissacarídeos
LTTR	reguladores transcricionais de tipo LysR
NB	<i>nod-box</i>
ncRNA	<i>non coding deoxyribonucleic acid</i>
Nod+	fenótipo simbiótico capaz de formar nódulos
Nop	<i>nodulation outter protein</i>
ORF	<i>open reading frame</i>
PCR	<i>polimerase chain reaction</i>
pH	potencial de hidrogênio
pI	ponto isoelétrico
PMSF	<i>phenylmethane sulfonyl fluoride</i>
PQQ	pirroloquinolina-quinona

qPCR	<i>quantitative polimerase chain reaction</i>
Rif	rifampicina
R-LPS	<i>rough- lipopolysaccharides</i>
RNase	ribonuclease
rpm	rotações por minuto
RR	regulador de resposta
rRNA	RNA ribossomal
SB	<i>syr-boxes</i>
SDS	<i>sodium dodecyl sulfate</i>
SDS-PAGE	<i>sodium dodecyl sulfate polyacrylamide gel electrophoresis</i>
S-LPS	<i>smooth- lipopolysaccharides</i>
Sm	spectinomicina
T2SS	sistema de secreção de tipo 2
T3SS	sistema de secreção de tipo 3
T6SS	sistema de secreção de tipo 6
TB	<i>tts-boxes</i>
TCS	sistema de dois componentes
X-gal	5-bromo-4-cloro-3-indolil- $\beta$ -D-galactopiranosídeo

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>16</b>
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA .....</b>	<b>18</b>
2.1	FIXAÇÃO BIOLÓGICA DE NITROGÊNIO E NODULAÇÃO .....	18
2.2	<i>Sinorhizobium fredii</i> NGR234.....	20
2.2.1	Cascata regulatória da nodulação em <i>S. fredii</i> NGR234 .....	21
2.3	OBJETIVOS.....	27
2.3.1	Objetivo geral.....	27
2.3.2	Objetivos Específicos .....	27
<b>3</b>	<b>MATERIAIS E METODOS.....</b>	<b>28</b>
3.1	ESTRATÉGIA GERAL .....	28
3.2	ESTIRPES BACTERIANAS, PLASMÍDEOS E PRIMERS UTILIZADOS NESTE TRABALHO.....	28
3.3	MEIOS E CONDIÇÕES DE CULTIVO UTILIZADAS .....	31
3.4	ELETROFORESE EM GEL DE AGAROSE.....	32
3.5	PREPARAÇÃO DE CÉLULAS TERMOCOMPETENTES.....	32
3.6	TRANSFORMAÇÃO POR CHOQUE TÉRMICO .....	33
3.7	EXTRAÇÃO DE DNA PLASMIDIAL.....	33
3.8	PURIFICAÇÃO DE AMOSTRAS DE DNA PARA SEQUENCIAMENTO E REAÇÃO EM CADEIA DA POLIMERASE (PCR) .....	34
3.9	TRANSCRIPTOMA POR RNA-seq.....	34
3.9.1	Condições de cultivo e extração de RNA .....	34

3.9.2	Construção de bibliotecas de cDNA e sequenciamento .....	35
3.9.3	Mapeamento.....	36
3.9.4	Análise de genes diferencialmente expressos.....	36
3.10	VALIDAÇÃO DA EXPRESSÃO DE GENES CANDIDATOS UTILIZANDO FUSÕES TRANSCRICIONAIS COM GFP .....	36
3.10.1	Conjugações com <i>Sinorhizobium fredii</i> .....	37
3.10.2	Ensaio de GFP .....	37
3.11	SUPEREXPRESSÃO DA PROTEÍNA NODW FUSIONADA A UMA CAUDA DE HISTIDINA .....	38
3.11.1	Clonagem no vetor pTZ57R.....	38
3.11.2	Subclonagem nos vetores de expressão .....	38
3.11.3	Superexpressão em <i>E. coli</i> BL21 ( $\lambda$ DE3) e indução em larga escala .....	39
3.11.4	Purificação de proteínas de fusão .....	40
3.11.5	Eletroforese em gel de poliacrilamida com dodecil sulfato de sódio (SDS- PAGE).....	40
3.12	ENSAIO DE RETARDO DA MOBILIDADE ELETROFORÉTICA (EMSA) ....	41
3.13	ANÁLISE BIOINFORMÁTICA DA REGIÃO <i>nodVW</i> E DAS PROTEÍNAS CODIFICADAS POR ESSES GENES .....	41
<b>4</b>	<b>RESULTADOS E DISCUSSÕES .....</b>	<b>43</b>
4.1	ANÁLISES BIOINFORMÁTICAS .....	43
4.1.1	Análise da região genômica contendo os genes <i>nodV</i> e <i>nodW</i> em <i>S. fredii</i> <i>NGR234</i> .....	43
4.1.2	Análise estrutural das proteínas codificadas pelo operon <i>nodV</i> -NGR_b16860 .....	44
4.1.3	Busca de proteínas homólogas a NodV em outras espécies de proteobacteria.....	48
4.2	CONSTRUÇÃO DE BIBLIOTECAS DE cDNA E ANÁLISE DE TRANSCRIPTOMA .....	49

4.2.1	Genes diferencialmente expressos entre <i>nodVW</i> <sup>-</sup> e NGR234.....	51
4.2.2	Genes diferencialmente expressos em <i>nodVW</i> <sup>-</sup> em presença e ausência de flavonoides.....	55
4.2.3	Genes diferencialmente expressos em <i>S. fredii</i> NGR234 em presença e ausência de flavonoides .....	55
4.3	VALIDAÇÃO DE GENES DE POR ENSAIO DE GFP .....	69
4.4	PURIFICAÇÃO DE NodW.....	78
<b>5</b>	<b>CONSIDERAÇÕES FINAIS .....</b>	<b>83</b>
<b>6</b>	<b>CONCLUSÃO.....</b>	<b>86</b>
	<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>87</b>
	<b>APÊNDICE 1 - GENES DIFERENCIALMENTE EXPRESSOS EM <i>nodVW</i><sup>-</sup> À COMPARAÇÃO DA ESTIRPE NGR234 SELVAGEM EM PRESENÇA DE FLAVONOIDES .....</b>	<b>97</b>
	<b>APÊNDICE 2 - GENES DIFERENCIALMENTE EXPRESSOS EM <i>nodVW</i><sup>-</sup> À COMPARAÇÃO DA ESTIRPE NGR234 SELVAGEM EM AUSÊNCIA DE FLAVONOIDES .....</b>	<b>106</b>
	<b>APÊNDICE 3 - GENES DIFERENCIALMENTE EXPRESSOS EM NGR234 APÓS INDUÇÃO COM FLAVONOIDES .....</b>	<b>109</b>
	<b>APÊNDICE 4 - ANÁLISE DE COMPONENTE PRINCIPAL (PCA) .....</b>	<b>117</b>

## 1 INTRODUÇÃO

A fixação biológica de nitrogênio atmosférico constitui um processo essencial para a vida na biosfera ao tornar o  $N_2$  gasoso acessível para sua assimilação pelos vegetais. Os microrganismos capazes de realizar esta transformação são bactérias que possuem o complexo enzimático da nitrogenase e que são chamadas de forma geral de diazotrofos. Os diazotrofos podem estar livres no solo ou em interação simbiótica com plantas hospedeiras, alojando-se dentro de estruturas especializadas nas raízes chamadas nódulos. A interação planta-bactéria gera dentro do nódulo um microambiente apropriado para o funcionamento da nitrogenase, que requer uma concentração baixa de oxigênio. Além disso, a planta fornece fotossintatos nutritivos à bactéria e esta provê nitrogênio assimilável em forma de amônia e, as vezes, também fitohormônios.

Para o estabelecimento de uma simbiose efetiva existem diversos mecanismos que permitem o reconhecimento específico de ambos interlocutores numa conversa molecular entre o rizóbio e a planta hospedeira. Ainda na rizosfera os flavonoides liberados pela planta são detectados pelo rizóbio por meio de proteínas sensoras, as proteínas NodD, as quais assim induzidas ativam uma cascata intracelular complexa que desencadeia o processo de nodulação. Na via de transdução de sinal estão envolvidos muitos fatores numa rede que embora tenha alguns elementos conservados, é específica de cada interação planta-bactéria. O estudo destas interações está diretamente ligado à possibilidade de, através do maior entendimento deste processo, alcançar uma aplicação mais eficiente destes microrganismos na agricultura.

*Sinorhizobium fredii* NGR234 é um rizóbio de ampla faixa hospedeira cuja cascata regulatória do processo de nodulação vem sendo amplamente estudada. Foi descoberto um conjunto de *orfs* no plasmídeo simbiótico dessa estirpe cujos produtos proteicos são similares às proteínas NodV e NodW de *Bradyrhizobium japonicum*. As proteínas NodV e NodW pertencem à família de reguladores de dois

componentes e estão envolvidas na regulação da transcrição dos genes de nodulação em *B. japonicum*.

Esse trabalho teve como principal objetivo o estudo da função dos genes *nodV* e *nodW* de *S. fredii* NGR234 (também chamado *Ensifer fredii* NGR234 ou *Rhizobium* sp. NGR234) e avaliar se eles participam na cascata regulatória da nodulação nessa estirpe. Uma estirpe mutante nos genes *nodV* e *nodW* foi utilizada para avaliar quais os genes são regulados por esse par de reguladores. Inicialmente foi analisado o transcriptoma global através da técnica de RNAseq. Posteriormente foi feita a confirmação dos verdadeiros alvos de regulação desse sistema em *S. fredii* NGR234.

## 2 REVISÃO BIBLIOGRÁFICA

### 2.1 FIXAÇÃO BIOLÓGICA DE NITROGÊNIO E NODULAÇÃO

O processo de fixação biológica de nitrogênio (FBN) é de fundamental importância, uma vez que apesar de o nitrogênio gasoso ser encontrado na atmosfera numa concentração de quase 80%, ele não é assimilável pelos organismos em sua forma diatômica, sendo assim um dos principais limitantes da vida na biosfera (FRANCHE et al., 2009). A FBN consiste na redução do N<sub>2</sub> gasoso a amônia por meio do complexo enzimático da nitrogenase, presente unicamente em certos grupos de procariotos, que representam assim um elemento essencial nos ecossistemas. Eles constituem um grupo taxonômico heterogêneo com representantes na maioria dos filos de bactéria, e também algumas arqueias metanogênicas. Esses microrganismos podem ser de vida livre ou estabelecer relações associativas com plantas (DIXON; KAHN, 2004; FRANCHE et al., 2009).

Um grande grupo de bactérias gram negativas do solo, pertencentes às Proteobacterias, se associa-se simbioticamente com as raízes (ou talos, em alguns legumes aquáticos) de plantas vasculares, elicitando a formação de estruturas diferenciadas chamadas nódulos, onde ocorre a fixação de nitrogênio (DENARIE et al., 1992). Nos nódulos, as bactérias suplementam nitrogênio assimilável às células vegetais enquanto essas aportam um ambiente protegido ao microrganismo com acessibilidade a fotossintatos e nutrientes. Adicionalmente, a presença da bactéria (a qual estabelece uma associação específica com a planta hospedeira) pode evitar a invasão da planta por microrganismos patogênicos e aumentar a produção de fitohormônios favorecendo o crescimento vegetal (PERRET et al., 2000). Elas são chamadas coletivamente de rizóbios, ainda que não pertençam ao mesmo grupo taxonômico. Entre os rizóbios se incluem os gêneros de alpha-proteobacteria *Agrobacterium*, *Allorhizobium*, *Azorhizobium*, *Bradyrhizobium*, *Mesorhizobium*, *Rhizobium*, *Sinorhizobium*, *Devosia*, *Methylobacterium*, *Ochrobactrum* e *Phyllobacterium*, e as beta-proteobacterias *Burkholderia* e *Cupriavidus*

(LINDSTROM; MARTINEZ-ROMERO, 2007). A formação de nódulos se estabelece principalmente com leguminosas, porém foi descrita a associação entre bactérias do gênero *Rhizobium* com a não-leguminosa *Parasponia* sp. (DENARIE et al., 1992).

A interação entre rizóbio e leguminosa é um processo complexo que começa na rizosfera da planta e desencadeia um intrincado diálogo molecular entre ambos os organismos levando à formação do nódulo. A liberação de abundantes fotossintatos nos exsudatos dos pêlos radiculares vegetais suporta o crescimento de muitos microrganismos do solo. Entre as substâncias liberadas pela planta encontram-se carboidratos, ácidos orgânicos, vitaminas, aminoácidos e compostos fenólicos, dentre os quais os flavonoides são de importância fundamental no início da interação (BROUGHTON et al., 2000; PERRET et al., 2000). Esses compostos são detectados por uma proteína sensora da família dos reguladores transcricionais do tipo LysR (LTTR), a proteína NodD, a qual pode existir como mais de um homólogo em algumas espécies bacterianas. NodD ativa uma via de transdução de sinal intramolecular que desencadeia o processo de nodulação e que começa com a sua interação com seus genes alvos, os genes *nod*, *noe* e *nol* (PERRET et al., 2000; KOBAYASHI et al., 2008). Esses genes possuem sequências conservadas de 49 pb a montante do promotor, chamados *nod-boxes*, onde NodD liga-se para ativar a sua transcrição e produzir as enzimas que sintetizam e secretam uma série de fatores lipo-quito-oligossacarídeos (LCOs) chamados fatores Nod (ROSTAS et al., 1986; OLDROYD; DOWNIE, 2004).

A estrutura fundamental dos fatores Nod é constituída por uma cadeia de 3 a 5 resíduos de N-acetilglucosamina ligados a um ácido graxo não saturado na sua extremidade não redutora (BROUGHTON et al., 2000; PERRET et al., 2000). Cada espécie de rizóbio produz fatores Nod com modificações específicas (metilações, acetilações, fucosilações, adição de grupos fosfato, etc.) e vários estudos sugerem que os tipos de modificações e a quantidade de fatores Nod produzidos é o que determina a faixa de plantas hospedeiras que a bactéria é capaz de colonizar (RELIC et al., 1993; BROUGHTON et al., 2000; PERRET et al., 2000).

Os fatores Nod secretados pelo rizóbio induzem mudanças na anatomia vegetal permitindo a invasão da planta pela bactéria. A primeira dessas mudanças é uma deformação dos pelos radiculares da planta produzindo uma curvatura característica, com uma cavidade onde a bactéria poderá crescer e se dividir, gerando um foco de infecção. A partir desse momento é formada uma estrutura tubular, o cordão de infecção, que cresce até as células do córtex da raiz onde se ramifica. Simultaneamente, as células do córtex da raiz se dividem intensamente, iniciando a formação do primórdio do nódulo. Dentro do cordão, as bactérias se multiplicam e no momento em que alcançam o córtex elas são liberadas no interior das células, cobertas por uma porção da parede celular vegetal. Uma vez no interior das células do córtex estas se diferenciam em bacteróides capazes de fixar o nitrogênio atmosférico (DENARIE et al., 1992; OLDROYD et al., 2005). O interior do nódulo constitui um microambiente ideal para a fixação de nitrogênio, onde as fontes de carbono e outros compostos fornecidos pela planta provêm a energia necessária para o processo. A nitrogenase é muito sensível ao oxigênio e sua concentração é mantida no nódulo a um nível baixo por meio de uma hemo proteína vegetal, a leghemoglobina (WHITE et al., 2007).

## 2.2 *Sinorhizobium fredii* NGR234

Nesse trabalho, a atenção está voltada à cascata de transdução de sinal ativada por flavonoides na linhagem de *S. fredii* NGR234 (New Guinea *Rhizobium* 234). NGR234 é uma alfa-proteobactéria da ordem das *Rhizobiales* de ampla faixa hospedeira, uma das mais promíscuas conhecidas, capaz de nodular mais de 120 gêneros diferentes de leguminosas e a não-leguminosa *Parasponia andersonii* (PUEPPKE; BROUGHTON, 1999).

Segundo SCHMEISSER et al., (2009), o genoma de *S. fredii* NGR234 possui 6,9 Mpb distribuídos em três replicons:

- O cromossomo (cNGR234) de 3,93 Mpb: possui um total de 3633 *orfs* (Open Reading Frames) que codificam a maioria das proteínas envolvidas no metabolismo celular basal.
- Um megaplasmídeo (pNGR234b) de 2,43 Mpb: possui 2342 *orfs* que codificam para poucas funções essenciais, mas um grande número de sistemas transportadores (50% do total da bactéria) e a maioria dos genes de síntese de polissacarídeos de superfície.
- Um plasmídeo simbiótico (pNGR234a) de 0,54 Mpb: possui uma densidade gênica muito menor e elementos transponíveis abundantes. Não apresenta genes envolvidos em nenhuma função essencial, mas possui a maioria dos genes de nodulação, genes da replicação plasmidial e genes de transferência conjugacional.

Várias características de *S. fredii* NGR234 têm sido associadas à ampla faixa hospedeira que esse microrganismo é capaz de nodular. Entre estas está a imensa quantidade de genes que codificam para sistemas de secreção diferentes, mais do que qualquer rizóbio conhecido. Além disso, o grande conjunto de genes relacionados ao metabolismo de compostos aromáticos e não aromáticos, que poderia facilitar a sobrevivência dos rizóbios na rizosfera vegetal e a sua indução por diversas classes de flavonoides. Finalmente, a grande diversidade de fatores Nod produzidos parece ser também um fator contribuinte para essa promiscuidade (PRICE et al., 1992; HANIN et al., 1997; BERCK et al., 1999; BROUGHTON et al., 2000; SCHMEISSER et al., 2009).

### 2.2.1 Cascata regulatória da nodulação em *S. fredii* NGR234

O reconhecimento de flavonoides por NGR234 é feito por NodD1, que é capaz de responder a uma ampla faixa de indutores (LE STRANGE, 1990). Esta proteína é responsável pela ativação do processo de nodulação por meio da sua ligação aos 19 *nod* boxes contidos no plasmídeo simbiótico dessa linhagem, cinco dos quais controlam a expressão de catorze genes envolvidos na síntese dos

fatores Nod (PERRET et al., 1999; KOBAYASHI et al., 2004). Existem também outros catorze *nod* boxes os quais controlam outros genes, como os que codificam para proteínas do Sistema de Secreção de Tipo 3 (T3SS), enzimas de modificação de polissacarídeos extracelulares e de síntese de ácido-3-indolacético (AIA) (KOBAYASHI et al., 2004). O gene *nodD1* encontra-se no plasmídeo simbiótico e tem uma expressão constitutiva; seu produto pode se ligar aos *nod* boxes na ausência de flavonoides. Porém, foi mostrado que esses compostos são requeridos para a ativação da transcrição (FENG et al., 2003). NodD1 também induz a transcrição dos genes *ttsI* e *syrM2*, localizados no plasmídeo pNGR234a, os quais codificam para os fatores de transcrição TtsI e SyrM2 (representantes da família de sensores de dois componentes e da família dos LTTR, respectivamente) (KOBAYASHI et al., 2008).

TtsI ativa a transcrição de genes que possuem sequências conservadas ou *tts*-boxes (TB), presentes no plasmídeo simbiótico. Entre seus produtos encontramos os componentes do T3SS e proteínas secretadas por ele, que são codificados pelos genes *rhc* e *nop*, respectivamente. O T3SS é um transportador de membrana altamente conservado que está presente em muitas bactérias que interagem com células eucariotas (HUECK, 1998). Ele atravessa a bicamada lipídica da bactéria, a parede e membrana plasmática vegetal, injetando proteínas efetoras (Nops, do inglês *nodulation outer proteins*) no citoplasma da planta (Figura 1). Em NGR234 foram descritos cinco destes efetores: NopL, NopT, NopJ, NopM e NopP (MARIE et al., 2003; BARTSEV et al., 2004; KAMBARA et al., 2009; STAEHELIN; KRISHNAN, 2015). Existem evidências da importância da presença de um T3SS funcional para a nodulação eficiente de algumas espécies, assim como fenótipos simbióticos desfavoráveis quando plantas hospedeiras são inoculadas com estirpes mutagenizadas em cada um destes cinco efetores (MARIE et al., 2003; AUSMEES et al., 2004; KAMBARA et al., 2009; STAEHELIN; KRISHNAN, 2015).

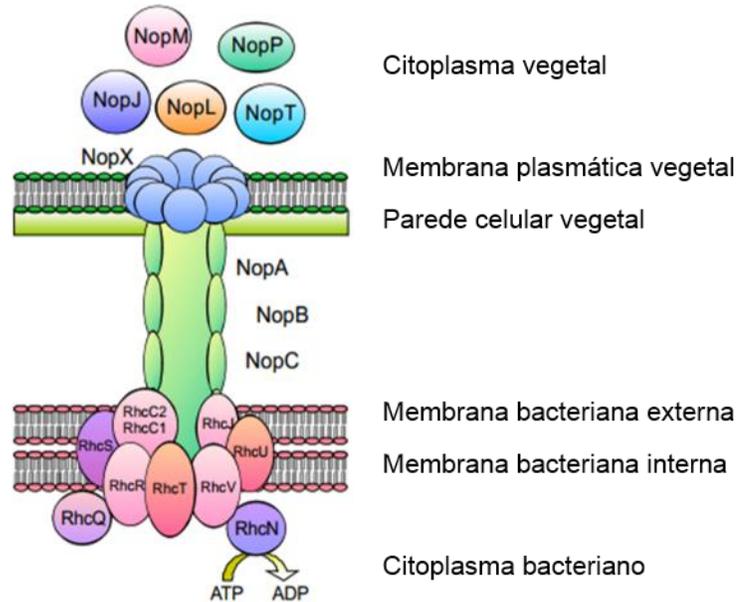


Figura 1- SISTEMA DE SECREÇÃO DE TIPO TRÊS DE *S. fredii* NGR234. FONTE: adaptado de Kambara, (2008)

TtsI participa também na ativação da síntese de uma classe especial de lipo-polissacarídeos de superfície (LPS) que é induzida durante o processo de invasão da planta hospedeira. Os LPSs parecem estar envolvidos nas primeiras etapas da simbiose, tais como reconhecimento do hospedeiro, colonização das raízes, proteção contra compostos tóxicos produzidos pela planta, formação do cordão de infecção e invasão do nódulo (KANNENBERG; CARLSON, 2001; BROUGHTON et al., 2006). Os LPSs são os principais componentes da membrana externa de bactérias Gram-negativas. Quando em vida livre, a membrana externa do rizóbio está coberta de R-LPS (*rough*-LPS). Após a indução, esses compostos são substituídos por S-LPS (*smooth*-LPS), predominantemente compostos de resíduos de ramnose e que possuem um domínio adicional de antígeno O na extremidade extracelular (KANNENBERG; CARLSON, 2001). Outros polissacarídeos envolvidos no processo de infecção são os polissacarídeos capsulares ou antígenos-K (KPS) cuja importância ainda não tem sido completamente elucidada (QUÉRÉ, LE et al., 2006).

Por outro lado, o regulador SyrM2 dirige a expressão de *syr*-boxes (SB) que se encontram no plasmídeo simbiótico (KOBAYASHI et al., 2004). Entre os alvos de regulação de SyrM2 encontra-se o homólogo de NodD1, NodD2, que age na ativação de certos *nod* boxes de ação tardia no processo de nodulação. Os genes ativados por NodD2 estão envolvidos na síntese do fitormônio ácido-3-indolacético (AIA) e na síntese de polissacarídeos ricos em ramnose. NodD2 também reprime a expressão de NodD1 e dos genes envolvidos na síntese de fatores Nod, participando assim do ajuste da cascata regulatória (KOBAYASHI et al., 2008). Um resumo dessa cascata é mostrado na Figura 2.

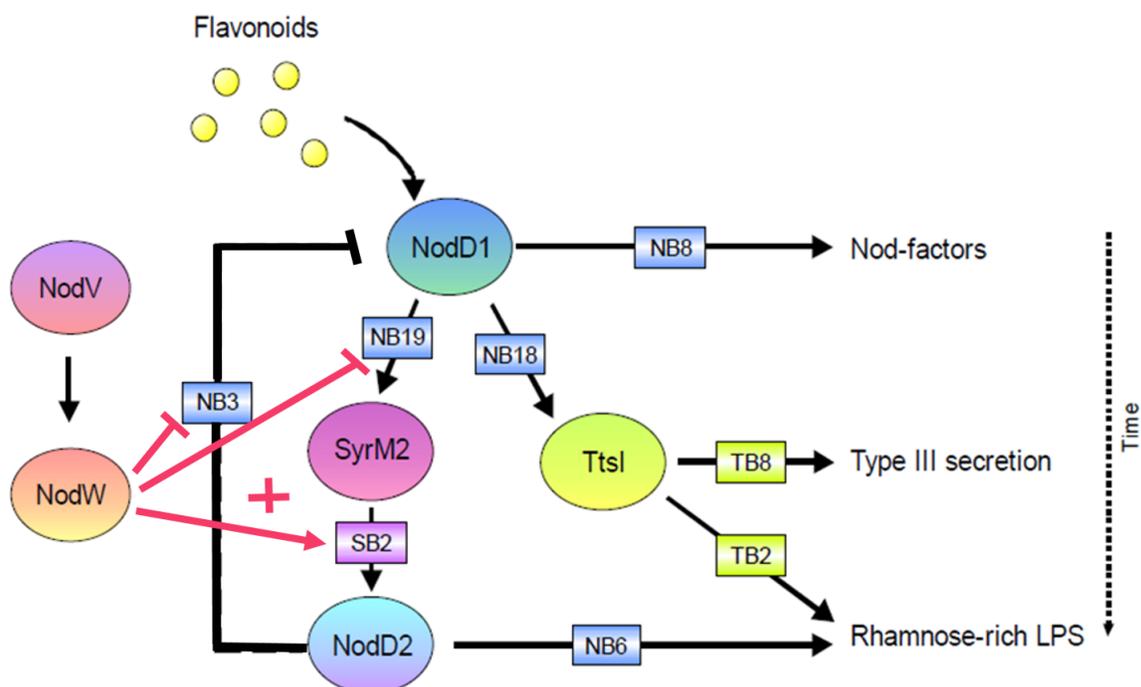


Figura 2- RESUMO DA CASCATA REGULATÓRIA INDUZIDA POR FLAVONOIDES EM *S. fredii* NGR234. À esquerda a função de NodV e NodW sugerida por Kambara, 2008. Caixas representam sequências consenso na região promotora dos genes. A linha pontilhada indica o decorrer do tempo durante a ativação da cascata de nodulação. Linhas terminadas em seta orientadas à sequências consenso indicam ativação transcricional enquanto linhas terminadas em barra indicam repressão transcricional. O sinal + indica presença de flavonoides. FONTE: adaptado de Kambara, 2008.

Outro par regulatório potencialmente envolvido com o processo de nodulação foi identificado em NGR234: NodV e NodW. Este trabalho está focado no esclarecimento da função dos genes *nodV* e *nodW* em NGR234 cujos homólogos já foram descritos em *Bradyrhizobium japonicum* (GÖTTFERT et al., 1990).

Em *Bradyrhizobium japonicum*, NodV e NodW constituem um sistema de regulação de genes *nod* independente de NodD, que é essencial para a nodulação das leguminosas *Vigna radiata*, *Vigna unguiculata* e *Macroptilium atropurpureum*, porém contribui só marginalmente à nodulação de *Glycine max*, já que na ausência destes genes a nodulação desta última espécie sofreu apenas um atraso de dois dias. Contudo, na ausência de proteínas NodD este sistema é necessário para a nodulação dessa espécie (GÖTTFERT et al., 1990). As proteínas NodV e NodW compartilham homologia com sistemas de dois componentes sensores de sinais extracelulares. Vários estudos apontam NodV como o componente quinase sensor que se autofosforila ao detectar um sinal específico extracelular e transfere o grupo fosfato a NodW, que pode assim agir como um ativador transcricional por se ligar ao DNA através de um domínio hélice-volta-hélice (STOCK et al., 1989; GÖTTFERT et al., 1990; STACEY; SANJUAN, 1995; LOH; GARCIA, 1997).

O papel hospedeiro-específico de NodV e NodW em *B. japonicum* sugere que NodV reconhece especificamente os isoflavonoides produzidos por *Vigna radiata*, *Vigna unguiculata* e *Macroptilium atropurpureum*, (GÖTTFERT et al., 1990). Por outro lado, a ausência de expressão de *ttsI* numa estirpe de *B. japonicum* USDA110 mutagenizada no gene *nodW* mostrou que a ativação deste gene é dependente do par NodV e NodW (KRAUSE et al., 2002).

Em NGR234, *nodV* e *nodW* encontram-se no megaplasmídeo pNGR234b e, a jusante deles, existem duas *orfs* cuja busca de homologia por BLAST-P na base de dados do National Center for Biotechnology Information (NCBI) indicou similaridade a *nodW*. Essas ORFs são *ngr157* (*locus tag* NGR\_b16860) e *ngr158* (*locus tag* NGR\_b16850) (STREIT et al., 2004; KAMBARA, 2008) (Figura 3). Conforme descreveu KAMBARA, (2008), enquanto a proteína codificada por *ngr157* apresenta 53% similaridade a NodW em toda a sua extensão, a codificada por *ngr158* não possui o domínio de ligação ao DNA. A presença destes genes no plasmídeo pNGR234b, a jusante do gene *nodW*, pode indicar que as proteínas codificadas por eles também respondem a NodV. Em outras situações, um membro quinase sensor pode fosforilar diferentes alvos (BIJLSMA; GROISMAN, 2003; LAUB; GOULIAN, 2007). Em *B. japonicum*, um segundo par regulatório semelhante

a NodV e NodW foi identificado. O componente regulador de resposta possui 65% de identidade com NodW e é chamado NwsB. Foi demonstrado que nessa estirpe NwsB não é essencial para nodulação de *Vigna radiata*, *Vigna unguiculata* e *Macroptilium atropurpureum*. Porém, quando NwsB foi expressa a partir de um promotor mais forte que o nativo, ele foi capaz de complementar a mutação em *nodW*, sugerindo que tem função equivalente (GROB et al., 1993). É interessante notar que a jusante, tanto de *nodW* quanto de *nwsB*, também foram identificadas ORFs semelhantes a NodW, mas que não possuem o domínio de ligação ao DNA. A função destas ainda é desconhecida.

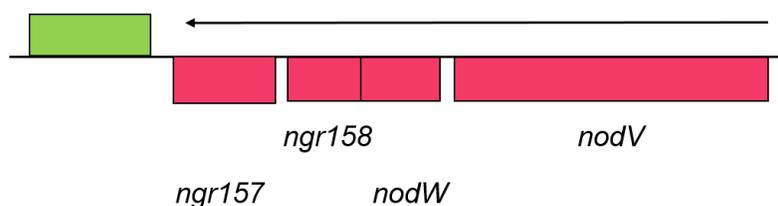


Figura 3- REGIÃO *nodVW* DE pNGR234b. NOTA: a seta preta acima do *cluster* indica o sentido da transcrição. FONTE: Adaptado de Streit et al., 2004.

Outro antecedente importante na busca de esclarecimento da função das proteínas NodV e NodW em NGR234 foi a descoberta de que a linhagem mutante *nodD1<sup>-</sup>* é capaz de ativar loci cuja expressão é dependente de flavonoides (FELLAY et al., 1995). Esse dado sugeriu a presença de outros reguladores sensíveis a flavonoides em NGR234, sendo o sistema NodV-NodW um possível candidato (KAMBARA, 2008).

Um estudo foi feito com um mutante de inserção no gene *nodV* em *S. fredii* NGR234 (KAMBARA, 2008). Os resultados desse estudo mostraram que a ausência de NodV não teve nenhum efeito na secreção de efetores Nop, na produção de LPSs ricos em ramnose ou no fenótipo simbiótico nas leguminosas *Vigna radiata*, *V. unguiculata*, *Macroptilium atropurpureum*, *Lablab purpureum*, *Pachyrhizus tuberosus*, *Crotalaria juncea* e *Leucaena leucocephala*. Além disso, foi demonstrada uma fraca ligação de NodW à região de DNA que contém NB18, o

*nod*-box presente na região promotora do gene *ttsI*. O mutante *nodVW*<sup>-</sup> mostrou níveis de expressão de *nod* boxes (NB8 e NB18) e *tts*-boxes (TB2 e TB8) reduzidos após a indução com flavonoides. Entretanto, os perfis observados de secreção de Nops, produção de LPS e nodulação são sempre indistinguíveis da linhagem selvagem. Além disso, foi observado um aumento significativo na expressão do gene *nodD1* no mutante *nodVW*<sup>-</sup> na ausência de flavonóides e por isso foi sugerido que NodV e NodW poderiam ser repressores desse gene para manter baixo o nível de expressão de NodD1 e da cascata regulatória da simbiose na ausência de flavonoides. O estudo também sugeriu que NodV e NodW aumentariam a ativação transcricional de *nodD2* na presença de flavonóides contribuindo dessa forma à repressão de *nodD1*. A função de NodV-NodW na cascata regulatória de *S. fredii* NGR234 sugerida por KAMBARA (2008) é mostrada na Figura 2.

## 2.3 OBJETIVOS

### 2.3.1 Objetivo geral

Elucidar a função dos genes *nodV* e *nodW* em *Sinorhizobium fredii* NGR234.

### 2.3.2 Objetivos Específicos

- Analisar *in silico* a região genômica que contém os genes *nodV* e *nodW*, bem como as proteínas NodV e NodW.
- Analisar o transcriptoma do mutante NGR $\Omega$ *nodVW* (daqui em diante *nodVW*<sup>-</sup>) em comparação ao do NGR234 selvagem, buscando prováveis alvos de regulação.
- Analisar as regiões regulatórias dos potenciais alvos de regulação de NodV-NodW buscando sequências conservadas às quais NodW liga-se.
- Confirmar se NodW liga-se aos alvos encontrados, usando a técnica de Ensaio de retardo de mobilidade em gel (EMSA).

### 3 MATERIAIS E METODOS

#### 3.1 ESTRATÉGIA GERAL

Com o intuito de compreender a organização estrutural e as funções dos genes *nodV* e *nodW* de *S. fredii* NGR234 foi feita uma análise da região genômica que os contém assim como das proteínas NodV e NodW utilizando softwares para avaliar domínios conservados, similaridade com outras proteínas e sequências reguladoras de transcrição. Além disso, foi realizada uma análise dos transcriptomas previamente obtidos por meio de sequenciamento de RNA (RNA-seq) do mutante *nodVW*<sup>-</sup>, comparando-o com o transcriptoma da estirpe *S. fredii* NGR234 selvagem (daqui em diante NGR234), ambos em presença e ausência de flavonóides. A partir dessa análise foram sugeridos genes candidatos a alvos de modulação por NodW. As regiões regulatórias dos genes candidatos foram analisadas em busca de sequências conservadas capazes de agir como sítios de ligação para NodW. Esses genes alvo foram validados por fusão transcricional da região regulatória ao gene de proteína fluorescente verde (GFP). As regiões regulatórias desses genes foram utilizadas em ensaio de retardo da mobilidade eletroforética (EMSA) para detectar se existe ligação *in vitro* de NodW a elas.

#### 3.2 ESTIRPES BACTERIANAS, PLASMÍDEOS E PRIMERS UTILIZADOS NESTE TRABALHO

Tabela 1- ESTIRPES BACTERIANAS UTILIZADAS NESTE TRABALHO.

ESTIRPE	GENÓTIPO/FENÓTIPO	FONTE/REFERÊNCIA
<i>Escherichia coli</i> DH10B	F- <i>mcrA</i> $\Delta$ ( <i>mrr-hsdRMSmcrBC</i> ) $\phi$ 80 <i>lacZ</i> $\Delta$ M15 $\Delta$ <i>lacX74 deoR recA1 araD139</i> ( <i>ara leu</i> ) 7697 <i>galU galK</i> rpsL <i>endA1 nupG</i>	INVITROGEN
<i>Escherichia coli</i> S17.1	Sm <sup>R</sup> Spm <sup>R</sup> Tra+ <i>pro his recA hsdR</i> (RP4-2 <i>kan::Tn7 tet</i> ; Mu, integrado no cromossomo)	SIMON et al., (1983)
<i>Escherichia coli</i> BL21 ( $\lambda$ DE3)	Expressa a RNA polimerase do fago T7. Não expressa as proteases lon e OmpT e é indicada para a expressão de proteínas recombinantes.	NOVAGEN
<i>Sinorhizobium fredii</i> NGR234	Rif <sup>R</sup> derivado do tipo selvagem de NGR234	LEWIN et al., (1990)
<i>Sinorhizobium fredii</i> NGR $\Omega$ <i>nodD1</i>	NGR234 com um cassete $\Omega$ Spm no sítio <i>Bam</i> HI de <i>nodD1</i> , Spm <sup>R</sup>	RELIC et al., (1993)
<i>Sinorhizobium fredii</i> NGR $\Delta$ <i>syrM2</i>	Mutante de NGR234 onde um fragmento de 880pb, incluindo parte das regiões regulatória e codificante de <i>syrM2</i> foi deletado.	SENA, (2013–não publicado)
<i>Sinorhizobium fredii</i> NGR $\Omega$ <i>nodVW</i>	Mutante de NGR234 Spm <sup>R</sup> onde o gene <i>nodV</i> foi interferido por inserção de um cassete de Spm.	KAMBARA, (2008 –não publicado)

FONTE: A autora (2016)

Tabela 2- PLASMÍDEOS UTILIZADOS NESTE TRABALHO

PLASMÍDEO	FENÓTIPO/GENÓTIPO	FONTE/REFERÊNCIA
pTZ57R	Vetor de clonagem. Amp <sup>R</sup>	FERMENTAS
pPROBE-GT	Vetor de clonagem, possui o gene <i>gfp</i> sem promotor. Gm <sup>R</sup>	MILLER et al., (2000)
pET28a	Vetor de expressão, possui o promotor do fago T7. Km <sup>R</sup>	NOVAGEN
pET29a	Vetor de expressão, possui o promotor do fagoT7. Km <sup>R</sup>	NOVAGEN
pET28a- <i>nodW-His</i>	Vetor de expressão, possui o gene <i>nodW</i> clonado com cauda de Histidina N-terminal	Este trabalho
pET28a- <i>nodWCterm-His</i>	Vetor de expressão, possui o fragmento C-terminal do gene <i>nodW</i> clonado com cauda de Histidina N-terminal	Este trabalho
pET29a- <i>nodW-C-His</i>	Vetor de expressão, possui o gene <i>nodW</i> clonado com cauda de Histidina C-terminal	Este trabalho

FONTE: A autora (2016)

Tabela 3- PRIMERS UTILIZADOS NESTE TRABALHO

IDENTIFICAÇÃO	SEQUÊNCIA	TEMPERATURA DE ANELAMENTO	SÍTIOS DE RESTRIÇÃO
<i>petnodWfor</i>	5'-GAAGGCAGCCATATGACGAGCGA-3'	61,3°C	<i>NdeI</i>
<i>petnodWrev</i>	5'-CATCATCAAGGATCCCGACGACAT-3'	58,7°C	<i>BamHI</i>
<i>nodWCterm</i>	5'-CAACGTTATCGCCATATGACACCG-3'	58,1°C	<i>NdeI</i>
<i>nodWrev</i>	5'-TTCATGGTCATTCCCTCCCAC-3'	59°C	-
<i>nopBfor</i>	5'-CTCGTCTTGATAAACCAAATCTGAA-3'	55°C	-
<i>nopBrev</i>	5'-GGACTCGATTACTTAACTCTTTGAC-3'	55°C	-
<i>pqq2</i> (reverse)	5'-CGGTGGTCGGAATTCAGATCA 3'	57,0 °C	-
<i>pqqA</i> (reverse)	5' CACGGAATTCAGCATCAATCTT 3'	53,6 °C	-
<i>Universal- CY3</i>	5'GACGTTGTAAAACGACGGCCAGT3'	69,8°C	-

FONTE: A autora (2016)

### 3.3 MEIOS E CONDIÇÕES DE CULTIVO UTILIZADAS

As estirpes de *E. coli* foram cultivadas a 37°C nos meios Luria Bertani (LB) e Luria Bertani Agar (LA) (SAMBROOK et al., 1989). As estirpes de *Sinorhizobium fredii* foram cultivadas a 28°C em meio triptona-levedura (TY) (BERINGER, 1974) ou meio mínimo para rizóbio contendo succinato como fonte de carbono (RMS) (BROUGHTON et al., 1986). Os cultivos em meio líquido foram incubados sob agitação a 120 rpm.

Quando necessário, o meio foi adicionado com uma mistura dos flavonoides apigenina, crisina, genisteína, kampferol e naringenina (0,2  $\mu$ M).

Os antibióticos utilizados foram: gentamicina (Gm) 15  $\mu$ g/mL, canamicina (Km) 50  $\mu$ g/mL, rifampicina (Rif) 50 ou 100  $\mu$ g/mL, e ampicilina (Amp) 250  $\mu$ g/mL.

### 3.4 ELETROFORESE EM GEL DE AGAROSE

As amostras de DNA foram analisadas por eletroforese em gel de agarose utilizando TBE 1X como tampão de corrida (Tris 90 mM, Ácido Bórico 89 mM, EDTA 2mM). Os géis foram corados com solução de brometo de etídeo 0,5  $\mu$ g/mL e visualizados em transiluminador de luz ultravioleta (Cell biosciences).

### 3.5 PREPARAÇÃO DE CÉLULAS TERMOCOMPETENTES

O pre-inóculo de *E. coli* de cada estirpe foi cultivado em 5 mL de meio LB por aproximadamente 12 horas. No dia seguinte, 500  $\mu$ L dessa cultura foram inoculados em 80 mL de LB em erlenmeyer de 500 mL, e incubado até atingir uma D.O.<sub>600</sub> entre 0,2-0,4. A cultura foi então distribuída em tubos de centrifuga estéreis e mantida em banho de gelo por 15 minutos e centrifugada a 2500 g e 4°C por 8 minutos. O sedimento foi ressuspensão em 15 mL de solução de MgCl<sub>2</sub> 100 mM estéril e gelada antes de ser centrifugado novamente nas mesmas condições. Dessa vez o pellet foi ressuspensão em 25 mL de CaCl<sub>2</sub> 100 mM e incubado em banho de gelo por 30 minutos. Em seguida, as células foram centrifugadas novamente e o pellet ressuspensão em 800  $\mu$ L de solução de CaCl<sub>2</sub> 100 mM com 20% de glicerol. Foram feitas alíquotas de 100  $\mu$ L de células em tubos eppendorf de 1,5 mL estéreis e gelados e armazenadas a -80°C até seu uso.

### 3.6 TRANSFORMAÇÃO POR CHOQUE TÉRMICO

As células termocompetentes foram descongeladas em banho de gelo e adicionado 1-2  $\mu\text{L}$  do DNA a ser transformado, mantendo o tubo no gelo durante 30 minutos. Imediatamente, a reação foi submetida a um choque térmico de  $42^{\circ}\text{C}$  durante 45 segundos e novamente incubada em gelo por 5 minutos. Em seguida foi adicionado 1 mL de meio LB no mesmo tubo e incubado por 40 minutos a  $37^{\circ}\text{C}$  e 120 rpm para a recuperação das células. A seleção de colônias transformantes foi realizada plaqueando 300  $\mu\text{L}$  de cada transformação em meio LA contendo o antibiótico adequado para cada vetor. Quando utilizado o vetor pTZ57R as placas foram adicionadas com 250  $\mu\text{g}/\text{mL}$  de ampicilina e 30  $\mu\text{g}/\text{mL}$  de 5-bromo-4-cloro-3-indolil- $\beta$ -D-galactopiranosídeo (X-gal) e 100  $\mu\text{M}$  de isopropil- $\beta$ -D-1-tiogalactopiranosídeo (IPTG).

### 3.7 EXTRAÇÃO DE DNA PLASMIDIAL

As extrações foram feitas a partir de culturas de 12 h em meio líquido com os antibióticos adequados. As culturas foram aliqüotadas em tubos de 1,5 mL e centrifugadas a 13000 rpm por 30 segundos (mini spin Eppendorf). O sobrenadante foi descartado e o sedimento foi ressuspenso adicionando 150  $\mu\text{L}$  de solução GET (glucose 50 mM, EDTA 10 mM, Tris HCl pH 8,0 25 mM), gelada. As células foram lisadas adicionando 150  $\mu\text{L}$  de solução de lise (SDS 1%, NaOH 0,18M) e misturando por inversão. A reação foi neutralizada após 1 minuto acrescentando 150  $\mu\text{L}$  de KacF (acetato de Potássio 3M, ácido fórmico 1,8M, pH 4,8). Depois foram adicionados 40  $\mu\text{L}$  de fenol:clorofórmio:álcool isoamílico (25:24:1), agitado no vortex e centrifugado por 5 minutos. Foram coletados 400  $\mu\text{L}$  da fase superior e transferidos para um novo tubo. O DNA foi precipitado com 2,5 vezes o volume de etanol absoluto e recuperado por centrifugação durante 8 minutos como se detalhou previamente. Foi feita uma lavagem do pellet de DNA com 1 mL de etanol 70%

seguida de uma nova centrifugação por 5 minutos. O pellet foi seco a temperatura ambiente e dissolvido em 20  $\mu\text{L}$  de  $\text{H}_2\text{O}$  ultrapura.

### 3.8 PURIFICAÇÃO DE AMOSTRAS DE DNA PARA SEQUENCIAMENTO E REAÇÃO EM CADEIA DA POLIMERASE (PCR)

Ao total de DNA plasmidial purificado foi adicionado água até um volume final de 100  $\mu\text{L}$  e 1  $\mu\text{L}$  de RNase 10 mg/mL e incubado durante 3 horas a 37°C. Posteriormente foram adicionados 30  $\mu\text{L}$  de clorofórmio:álcool isoamílico 24:1, seguido de agitação no vortex e centrifugação por 5 minutos. A fase aquosa foi transferida a um tubo novo e acrescentados 2/3 volumes de acetato de amônio 7,5 M e 2,5 volumes de etanol absoluto. As amostras foram homogeneizadas e o DNA precipitado foi recuperado por centrifugação durante 8 minutos. O precipitado ser lavado com etanol 70% e centrifugado 5 minutos mais. O pellet foi seco e ressuspenso em 10  $\mu\text{L}$  de  $\text{H}_2\text{O}$  ultrapura. As reações de sequenciamento, pelo método de terminação de cadeia (SANGER et al., 1977), foram realizadas com BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems), conforme orientações do fabricante. Em seguida estas foram purificadas, secas durante a noite e levadas para sequenciamento no sequenciador automático ABI 3500 (Applied Biosystem).

### 3.9 TRANSCRIPTOMA POR RNA-seq

#### 3.9.1 Condições de cultivo e extração de RNA

Foi extraído o RNA de bactérias da estirpe selvagem *Sinorhizobium fredii* NGR234 e do mutante *nodVW*, as quais foram cultivadas em meio RMS em ausência e presença de  $10^{-6}$  M de uma mistura dos flavonoides apigenina, crisina,

genisteína, kampferol e naringenina. As células foram coletadas a uma  $DO_{600} = 0.9$  (fase logarítmica de crescimento) a 4°C, adicionando 1 mL RNAlater™ a cada amostra, em pellet proveniente de 3mL de cultura. A extração de RNA foi feita com RiboPure™ RNA Purification Kit bactéria conforme orientações do fabricante. Só uma das bibliotecas, correspondente a uma das replicatas biológicas *nodVW* sem flavonoides, foi construída durante a execução deste projeto, estando as demais prontas. Depois da extração, as amostras foram tratadas com kit TURBO DNA-free (Ambion) e incubadas em estufa 37°C por 30 minutos. A DNase foi inativada com a adição de 10% do volume do sistema com solução inativadora e repouso a temperatura ambiente por 2 minutos. As amostras foram então centrifugadas a 14000 rpm (Mini Spin, Eppendorf) por 1 minuto e o sobrenadante foi transferido para um novo tubo. Foi feita uma quantificação usando NanoDrop (Thermo Scientific). Posteriormente, as amostras foram tratadas com o kit MICROBExpress™ Bacteria, para depleção do rRNA, conforme orientações do fabricante. Depois disso as amostras foram novamente quantificadas com Qubit – Quant-iT Assays (Invitrogen), conforme orientações do fabricante.

### 3.9.2 Construção de bibliotecas de cDNA e sequenciamento

A partir do RNA extraído foram construídas bibliotecas de cDNA usando o kit ION TOTAL RNA-SEQ V2 for Whole Transcriptome Libraries (Ion Torrent- Life Technologies), seguindo as instruções do fabricante. O cDNA foi purificado usando o kit PureLink™ PCR Micro Kit (Invitrogen™). Foram então selecionados os fragmentos de 200 nucleotídeos utilizando o sistema eletroforético E-Gel™ SizeSelect Agarose Gels, 2% (Invitrogen). Em seguida foram feitos 16 ciclos de amplificação desses fragmentos para inserção de barcodes. Os produtos foram purificados usando o kit PureLink™ PCR Micro Kit (Invitrogen™) e quantificados utilizando Bioanalyzer 2100 (Agilent Technologies Genomics). Posteriormente foi realizada PCR em emulsão utilizando o kit Ion PI Template v3 (Ion Technologies), e enriquecimento das bibliotecas com o mesmo kit. As bibliotecas assim obtidas foram sequenciadas na plataforma IonProton (ThermoFisher Scientific) usando o kit Ion PI Sequencing OT2 200.

### 3.9.3 Mapeamento

Tanto o mapeamento quanto a filtragem (*trimming*) de qualidade das sequências foram feitos usando o software CLC Genomics Workbench 7.5. O filtro de qualidade consistiu na remoção de sequências de baixa qualidade (limite de 0,05) e nucleotídeos ambíguos (permitindo um máximo de dois). Além disso, para aumentar o número de leituras mapeadas cada biblioteca foi subdividida em dois grupos, um deles contendo as leituras maiores a 100 pares de bases e, no outro, aquelas entre 20 e 99 pares de bases, eliminando assim as leituras menores do que 20 pares de bases. O mapeamento RNA-seq foi realizado com uma tolerância de dois *mismatches*, fração de comprimento mínima e fração de similaridade mínima de 0,80, e mapeamento de cada leitura em um único local. Esses valores significam que pelo menos 80% das bases das leituras devem alinhar com 80% de similaridade contra o genoma de NGR234.

### 3.9.4 Análise de genes diferencialmente expressos

A análise estatística para determinar genes diferencialmente expressos (DE) foi feita utilizando o software DESeq (ANDERS; HUBER, 2010). Este programa utiliza o método estatístico baseado na distribuição binomial negativa e no teste de Benjamini-Hochberg (1995) que gera um valor  $p$  levando em consideração testes de hipóteses múltiplas.

## 3.10 VALIDAÇÃO DA EXPRESSÃO DE GENES CANDIDATOS UTILIZANDO FUSÕES TRANSCRICIONAIS COM GFP

A validação da expressão diferencial de genes foi realizada utilizando fusões transcricionais das suas regiões regulatórias com o gene *gfp* (codifica Green Fluorescent Protein). As fusões foram realizadas em trabalhos anteriores utilizando o vetor pPROBE-GT ou pPROBE-GT' (MILLER et al., 2000) e transformadas na estirpe de *E. coli* S17.1 (WASSEM et al., 2008; HERBST FIORESE, 2015- não publicado; GREMSKI STABACH, 2016- não publicado).

### 3.10.1 Conjugações com *Sinorhizobium fredii*

As estirpes de *E. coli* S17.1 contendo fusões transcricionais de interesse foram conjugadas com as estirpes *S. fredii* NGR234 selvagem e as estirpes mutantes *syrM2<sup>-</sup>*, *nodD1<sup>-</sup>* e *nodVW<sup>-</sup>*. Para isso, as culturas foram crescidas em meio sólido e posteriormente foi preparada uma placa para conjugação bi-parental contendo os meios TY e LB em proporção 1:1 sem antibióticos. Cada estirpe de *Sinorhizobium fredii* foi transferida junto com uma estirpe de *E. coli* S17.1, misturadas nessa placa e incubadas a 28°C por 24 h. No dia seguinte foram feitas placas de seleção contendo meio TY com os antibióticos rifampicina (Rif) 100 µg/mL e gentamicina (Gm) 15 µg/mL. A partir dessa placa foram selecionadas cinco colônias e transferidas a uma placa nova com Rif 50 µg/mL e Gm 15 µg/mL.

### 3.10.2 Ensaio de GFP

Com o intuito de acompanhar o perfil de expressão das regiões estudadas, através da emissão de fluorescência das linhagens transconjugantes, foram montados blocos de 96 poços de 2mL, contendo as mesmas em presença e ausência de flavonoides. Para isso, as bactérias foram previamente crescidas em meio TY líquido com os antibióticos Rif 50 µg/mL e Gm 15 µg/mL durante dois dias a 28°C. Posteriormente, as mesmas culturas foram transferidas ao meio RMS líquido com Gm 15 µg/mL e incubadas dois dias a 28°C, após os quais foi feita uma última subcultura *overnight* nas mesmas condições. Para a montagem do bloco foi lida a DO<sub>600</sub> das culturas e inoculadas de forma a obter uma DO<sub>600</sub>=0,05 em todos os poços. Para acompanhar a emissão em presença de flavonoides o meio RMS foi adicionado com uma mistura de cinco flavonóides (apigenina, crisina, genisteína, kampferol e naringenina 0,2 µM cada). A emissão de fluorescência foi acompanhada no Infinite 200 PRO NanoQuant (TECAN) a cada 24 h durante cinco dias.

### 3.11 SUPEREXPRESSION DA PROTEÍNA NODW FUSIONADA A UMA CAUDA DE HISTIDINA

Foram utilizadas três estratégias de clonagem do gene *nodW* com o objetivo de facilitar a solubilização e purificação de NodW. Na primeira delas foi clonado o gene completo adicionando uma cauda de histidina na extremidade C-terminal da proteína; na segunda estratégia foi clonada somente a sequência que codifica o domínio de ligação ao DNA, presente na região C-terminal de NodW com a cauda de histidina no N-terminal da mesma. Na terceira estratégia foi clonado o gene completo mas a cauda de histidina foi adicionada na extremidade N-terminal da proteína. Para isso, foi analisada a região gênica de *nodW* e foram desenhados três *primers* específicos utilizando o programa NetPrimer (<http://www.premierbiosoft.com>). Um quarto primer foi utilizado proveniente de outro trabalho (KAMBARA, 2008 - não publicado).

#### 3.11.1 Clonagem no vetor pTZ57R

Para a primeira estratégia de clonagem foram utilizados os primers *petnodWfor* e *nodWrev* para amplificar por PCR a sequência completa do gene *nodW*. Na segunda estratégia foram utilizados os primers *nodWCtermfor* e *petnodWrev* para amplificar apenas a sequência que codifica o fragmento C-terminal de *nodW*, e na terceira estratégia empregaram-se os primers *petnodWfor* e *petnodWrev* amplificando também a sequência completa do gene *nodW*. Os produtos foram analisados através de electroforese em gel de agarose e, em seguida, foram clonados no vetor pTZ57R conforme orientações do fabricante (Fermentas) e utilizados para transformar células competentes de *E. coli* DH10B. Os clones transformantes foram avaliados pela presença de inserto através de PCR de colônia e sequenciamento de DNA plasmidial.

#### 3.11.2 Subclonagem nos vetores de expressão

Foi utilizado o vetor de expressão pET29a para a primeira estratégia de clonagem de NodW já que este vetor permite a expressão de proteínas com cauda

de histidina na extremidade C-terminal. O vetor pET28a foi usado na segunda e terceira estratégia de clonagem já que este permite a expressão de proteínas com cauda de histidina na extremidade N-terminal.

A primeira estratégia foi realizada através da digestão enzimática do plasmídeo pTZ57R-*nodW* com as enzimas *Xba*I e *Sma*I e preenchimento com DNA polimerase *Phusion*, conforme recomendações do fabricante. Ao mesmo tempo, o vetor pET29a foi cortado com a enzima *EcoRV* gerando pontas cegas que puderam ser ligadas ao inserto. Para evitar a religação do vetor este foi tratado com fosfatase alcalina SAP (Shrimp Alkaline Phosphatase), segundo orientações do fabricante.

A segunda e terceira estratégias de clonagem foram alcançadas por meio de digestão enzimática tanto dos plasmídeos pTZ57R-*nodW* e pTZ57R-*nodWCterm* para a liberação dos insertos, quanto do vetor de expressão pET28a, com as enzimas *Nde*I e *Bam*HI, conforme orientações do fabricante. A digestão gerou cortes coesivos complementares em ambas moléculas que permitiram a ligação. As ligações foram realizadas utilizando T4 DNA ligase conforme orientações do fabricante.

A seleção de clones transformantes foi feita por PCR de colônia com os mesmos *primers* utilizados na clonagem, digestão enzimática do DNA plasmidial e em alguns casos foi realizado o sequenciamento do inserto com *primers* específicos do vetor.

### 3.11.3 Superexpressão em *E. coli* BL21 ( $\lambda$ DE3) e indução em larga escala

As construções pET28a-*nodWCterm-His*, pET28a-*nodW-His* e pET29a-*nodW-C-His* foram usadas para transformar células competentes de *E. coli* BL21 ( $\lambda$ DE3).

Para a indução, as células transformadas foram inoculadas em 300 mL de meio LB com o antibiótico adequado e incubadas por aproximadamente 3 h ou até atingir uma D.O.<sub>600</sub> entre 0,3-0,5. A cultura foi então induzida com IPTG 200  $\mu$ M e incubada sob agitação a 20°C, *overnight*. No dia seguinte, o volume total de cultura

foi centrifugado a 5000 g, a 4°C, por 5 minutos e o pellet ressuspensão em 30 mL de tampão A (imidazol 20 mM, Tris-HCl 50 mM pH 8, glicerol 10%, glicina 50 mM, triton 0,1%, NaCl 500 mM). A seguir foram adicionados 30 µL de PMSF 0,1 M e 30 µL de Lisozima (10 mg/mL). A reação foi incubada a 28°C por 30 minutos e após as células foram lisadas por sonicação (20 pulsos de 1 minuto e 2 minutos de descanso entre eles) no SONICATOR Ultrasonic Processor XL (HEAT SYSTEMS). O lisado foi centrifugado a 10000 g, a 4°C por 30 minutos. A fração insolúvel foi descartada e a fração solúvel levada para a purificação. Amostras de cada fração e antes da indução foram guardadas para confirmar a expressão das proteínas recombinantes através de eletroforese desnaturante em gel de poliacrilamida (SDS-PAGE), conforme descrito por LAEMMLI, (1970).

#### 3.11.4 Purificação de proteínas de fusão

A purificação das proteínas foi realizada em coluna cromatográfica HiTrap Chelating, carregada com níquel, através do sistema de cromatografia AKTA (GE Healthcare) ou manualmente. A cromatografia foi feita em tampão A e as proteínas eluídas com gradiente descontínuo de imidazol utilizando tampão B (imidazol 1 M, Tris-HCl 50 mM pH 8, glicerol 10%, glicina 50 mM, triton 0,1%, NaCl 500 mM). A purificação foi analisada por eletroforese em SDS-PAGE conforme Laemmli (1970) e as frações contendo as proteínas Cterm-His ou NodW-His foram dialisadas contra tampão de armazenamento (Tris 50 mM pH 8, NaCl 50 mM, glicerol 50%) e armazenadas a -20°C.

#### 3.11.5 Eletroforese em gel de poliacrilamida com dodecil sulfato de sodio (SDS-PAGE)

A análise das amostras por SDS-PAGE foi conforme LAEMMLI (1970), utilizando gel de poliacrilamida 12 ou 15%. Utilizou-se o sistema vertical Miniprotean (BioRad) a 120 V. A concentração de acrilamida e o tempo de corrida dependeram da amostra aplicada. Antes da aplicação cada amostra foi misturada com tampão de amostra e incubada a 95°C de 3 a 5 minutos. Após a corrida, os géis foram corados com Azul de Coomassie (0,1 % Azul de Coomassie R-250 em solução

fixadora), durante 1-12 horas em agitação contínua e posteriormente descorados com solução fixadora (40% metanol, 10% ácido acético, 50% água destilada).

### 3.12 ENSAIO DE RETARDO DA MOBILIDADE ELETROFORÉTICA (EMSA)

Regiões regulatórias de genes modulados por NodV-NodW foram amplificadas por PCR e incubados, durante 20 minutos a temperatura ambiente ou 30 minutos no gelo, em presença de diferentes concentrações de proteína, tampão de armazenamento (Tris-HCl 50 mM pH 8, NaCl 150 mM, EDTA 0,1 mM, glicerol 50%) e TAP (50 mM Tris-acetato; 100 mM acetato de potássio; 8 mM acetato de Magnésio pH 7,9). Imediatamente, as amostras foram misturadas com 2  $\mu$ L de corante (Glicerol 10%, Azul de Bromofenol) e aplicadas em gel nativo de poliacrilamida 4,5% previamente submetido a pré-corrída em tampão nativo 1X (Tris 50 mM; glicina 0,4 M; EDTA 2 mM; MgSO<sub>4</sub> 8 mM) durante até 30 minutos a voltagem constante de 70 V. A reação foi corrida por duas ou três horas, corada com SYBR gold e visualizada no BioChem System UVP (BioImaging Systems). Quando utilizados fragmentos de PCR fluorescentemente marcados com CY3 o DNA foi visualizado no Typhoon FLA 9500 (GE Healthcare).

### 3.13 ANÁLISE BIOINFORMÁTICA DA REGIÃO *nodVW* E DAS PROTEÍNAS CODIFICADAS POR ESSES GENES

Para a determinação dos limites da unidade transcricional da qual fazem parte os genes *nodV* e *nodW* foi realizada uma busca manual de sequências consenso promotoras de transcrição, sítios de ligação ao ribossomo e sequências terminadoras transcricionais. A possibilidade de genes contíguos a *nodV* e *nodW* formarem parte do mesmo operon foi também avaliada por meio do sítio microbesonline (ALM et al., 2005; DEHAL et al., 2010) (<http://www.microbesonline.org>). A busca de sequências terminadoras da

transcrição foi realizada através da ferramenta ARNold (NAVILLE et al., 2011), disponível no sítio <http://rna.igmors.u-psud.fr/toolbox/arnold/>.

A análise de domínios funcionais das proteínas codificadas pelos genes do operon foi realizada utilizando as ferramentas BLASTP, InterPro, SMART e PFam (ALTSCHUL et al., 1997; SCHULTZ et al., 1998; JONES et al., 2014; LETUNIC et al., 2015; FINN et al., 2016, 2017), disponíveis nos endereços [blast.ncbi.nlm.nih.gov](http://blast.ncbi.nlm.nih.gov), [www.ebi.ac.uk/interpro](http://www.ebi.ac.uk/interpro), [smart.embl-heidelberg.de](http://smart.embl-heidelberg.de), [pfam.xfam.org](http://pfam.xfam.org) respectivamente. Para determinação do pI e da massa molecular teórica foi utilizado o programa ProtParam disponível em [www.expasy.org](http://www.expasy.org) (GASTEIGER et al., 2005). A presença de peptídeo sinal e de regiões de hélice transmembrana na proteína NodV foi analisada usando o software SignalP 4.0 (PETERSEN et al., 2011) e os softwares TMHMM Server v. 2.0 e CCTOP respectivamente (SONNHAMMER et al., 1998; PETERSEN et al., 2011; DOBSON et al., 2015).

Análises de homologia e vizinhança génica da proteína NodV em outras espécies de rizóbios foram realizadas utilizando as ferramentas disponíveis na base de dados Integrated Microbial Genomes and Microbial Samples (IMG/M) (CHEN et al., 2017), disponível no endereço <https://img.jgi.doe.gov/>.

## 4 RESULTADOS E DISCUSSÕES

### 4.1 ANÁLISES BIOINFORMÁTICAS

#### 4.1.1 Análise da região genômica contendo os genes *nodV* e *nodW* em *S. fredii* NGR234

A região genômica do megaplasmídeo b de *S. fredii* NGR234 que contém os genes *nodV* e *nodW* foi analisada em busca de sequências consenso sinalizadoras de começo e fim da transcrição com o intuito de determinar os limites da unidade transcricional. Através da busca manual foi encontrada uma possível sequência promotora de transcrição dependente do fator  $\sigma^{70}$  na região a montante do gene *nodV* (TTGCGGCTGTGGTGCGGCGGGTTTACGGT), assim como também foram identificadas sequências consenso de Shine-Dalgarno (AGGAGG) antes do começo desse mesmo gene. Foi observada uma sobreposição entre os genes *nodV* e *nodW*, assim como também entre os genes NGR\_b16850 e NGR\_b16860. Através do uso da ferramenta online ARNold foi identificada uma provável sequência terminadora independente da proteína Rho na região a jusante do gene NGR\_b16860. Além disso, foi utilizada a base de dados Microbesonline para prever a probabilidade de os genes contíguos a *nodV* e *nodW* formarem um operon. O resultado confirmou que os genes *nodV*, *nodW*, NGR\_b16850 e NGR\_b16860 formam um operon, mas não inclui outros genes na mesma região. Um esquema da região mostrando a organização destes genes é mostrada na Figura 4.

```

TAAAGGCCGGGGAGTGAGGCAGCGGCCGGTTTACGAAGTGTAATCGATCGCAAACACCGTCGCCGAGACG
TCATGAAGCAGCTCAAGAAAAGAATGTCGTTGGCTTGCTGCCTTGTCTCACGCCAGGGCGGCAAGCAGC
GGCTCGTCGCCGACCTTGGTCGCCGAAGTCTCGGCGTTGCCGACCGTCAGCGGAAAAGGCAGAGTAATTT
CCGGCTCACTCGCTTTGATCCAGCTTTCCACGAATGCGGCGTAGCGGCGACGCCGCCGCGCAATACATGG
ATTTCCATGAATTGCGGCTGTGGTGCGGCGGGTTTACGGTGTGAGACGCGTTTCGCCAGCCCAATTTCGA
GAATAACTGCTGTACGCGCTGACGATCCCTGAGTTGTCCGGCAGGAGCAGCTGATG! ..... nodV
.....AGGCAGCGCGATGA! ..... nodW .....GGGAGGGAATGA
CGATGAAATA ..... NGR_b16850 .....TAGAGGAGGTTGTGATG! ..... NGR_b16860
TAAAGCTGAAACCGCTCTGCCGAGGATCATCCCGGCAGAACGGTTTCAAATGAAGGACTTGTGCTCCACC
TTGGCTAAGGTCGAATGCCGCAATTTGGCGGCTAAGCAACCAGTCCTTCGCGTGGCCGGTTGGACTGCCG
TCATG

```

Figura 4- ANÁLISE DA REGIÃO GENÔMICA QUE CONTÉM OS GENES *nodV* E *nodW*. As sequências consenso -10 e -35 da possível região promotora  $\sigma^{70}$  estão sublinhadas, os possíveis sítios de união ao ribossomo diante de cada gene estão em negrito, o códon de início de tradução de cada gene encontra-se em quadro de traços e o códon de terminação se mostra em quadro e itálico, o corpo dos genes se mostra como pontilhado. Uma possível sequência terminadora independente de Rho se mostra sublinhada com pontilhado. O códon de fim da ORF a montante deste operon e o códon de início da ORF a jusante se mostram ressaltados em cinza escuro. FONTE: A autora (2017)

#### 4.1.2 Análise estrutural das proteínas codificadas pelo operon *nodV*-NGR\_b16860

As sequências de aminoácidos das proteínas do operon foram analisadas utilizando os sítios InterPro, SMART, MyHits e Pfam.

A proteína NodV possui 1333 aminoácidos com uma massa molecular calculada de 148.820 Da e um ponto isoelétrico (pI) teórico de 5,52, como determinado pelo programa ProtParam. A análise de arquitetura de domínios desta proteína mostrou a presença de um domínio N-terminal sensor integral de membrana da superfamília MASE1, seis domínios PAS, um domínio HisKA e um domínio HATPase no extremo C-terminal. O domínio MASE1 (do inglês *Membrane Associated Sensor 1*) constitui um domínio sensor de sinal tipicamente encontrado em Histidina quinases e outras proteínas de sinalização bacteriana. MASE1 é geralmente constituído por oito segmentos transmembrana e a natureza do sinal

percebido por este domínio é desconhecida (NIKOLSKAYA et al., 2003). NodV possui ainda seis domínios PAS, os quais também funcionam como sensores em proteínas de sinalização. Embora este tipo de domínio possa existir em proteínas tanto de *Archaea*, *Bacteria* e *Eukaria*, em procariotos encontram-se geralmente associados a histidina quinases (HK). Os domínios PAS são citosólicos e podem monitorar diversos parâmetros celulares como luz, nível de energia celular, oxigênio e até fatores ambientais que atravessam a membrana bacteriana. Diversos domínios PAS em uma única proteína podem estar funcionalmente diferenciados para perceber diferentes estímulos (TAYLOR; ZHULIN, 1999). Na direção C-terminal aos domínios PAS, NodV possui um domínio HisKA conservado entre HK. Este domínio está implicado na dimerização da HK e inclui o resíduo de Histidina fosfoceptor que é *trans*-autofosforilado por um domínio kinase da outra subunidade da HK homodimérica ao perceber um estímulo específico (WEST; STOCK, 2001). O grupo fosforil é depois transferido ao componente regulador de resposta (RR) numa reação catalisada pela HK, ativando a resposta intracelular ao estímulo (STOCK et al., 1989). Na extremidade C-terminal de NodV, foi encontrado um domínio HATPase, que liga ao ATP doador do grupo fosforil e que é responsável pela atividade quinase (WILLIAMS; WHITWORTH, 2010). A análise de domínios transmembrana realizada com varios *softwares* não teve um número de segmentos transmembrana consenso, estando este número entre 7 e 11, presentes todos entre os aminoácidos 1 a 290 da extremidade N-terminal, aproximadamente. Através do *software* CCTOP foi predita uma estrutura com 10 segmentos transmembrana, todos eles na região predita como domínio MASE1, entre os aminoácidos 12 e 290. O mesmo programa predisse uma extensa porção intracitoplasmática a partir do aminoácido 291 e até o C-terminal de NodV, onde estariam inclusos os domínios PAS, HisKA e HATPase. O esquema da arquitetura de NodV é mostrado na Figura 5. A análise através do programa SignalP4.1 determinou ausência de peptídeo sinal N-terminal (Não mostrado).

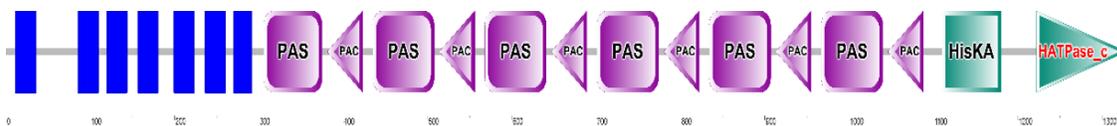


Figura 5- MODELO DA ESTRUTURA DA PROTEÍNA NodV MOSTRANDO OS DOMÍNIOS CONSERVADOS. Em azul são mostrados os segmentos transmembrana que correspondem à região predita como domínio MASE1, os domínios PAS em roxo onde são indicados também os motivos estruturais PAC, o domínio HisKA pelo quadrado verde e o domínio HATPase\_c pelo triângulo verde. FONTE: SMART

A proteína NodW possui 216 aminoácidos e uma massa molecular calculada de 23,6 kDa, com um  $pI$  teórico de 5,13. Está formada por um domínio receptor regulador de resposta N-terminal pertencente à super-família CheY (REC) e um domínio efetor C-terminal com atividade de ligação ao DNA. A fosforilação do domínio N-terminal, que acontece num resíduo de Asp conservado deste tipo de regulador, induz uma mudança conformacional que ativa o domínio efetor, desencadeando a resposta celular. Enquanto os domínios receptores são amplamente conservados em reguladores de resposta (RR), o domínio efetor não faz parte de uma única família de parálogos (LAUB; GOULIAN, 2007). O domínio C-terminal de NodW possui um motivo hélice-volta-hélice (HTH), presente em reguladores transcricionais da família LuxR/FixJ. O domínio forma uma estrutura de quatro hélices, das quais a primeira e a segunda formam o motivo que liga ao DNA através do sulco maior. A quarta hélice está envolvida com a dimerização em alguns reguladores desta família que se ligam ao DNA como multímeros (DUCROS et al., 2001). A ausência de um core AAA+, característico de fatores de transcrição ativadores associados ao fator  $\sigma^{54}$ , indica que NodW se associa ao fator  $\sigma^{70}$ , podendo atuar tanto como ativador quanto como repressor da transcrição (JOLY et al., 2012).

Por outro lado a proteína codificada pela ORF NGR\_b16850 contém 120 aminoácidos e uma massa molecular calculada de 12.892 Da, assim como um  $pI$  teórico de 5,27. A análise de domínios funcionais deste polipeptídeo mostrou homologia com o domínio REC, mas sem domínio de ligação ao DNA (Figura 6)

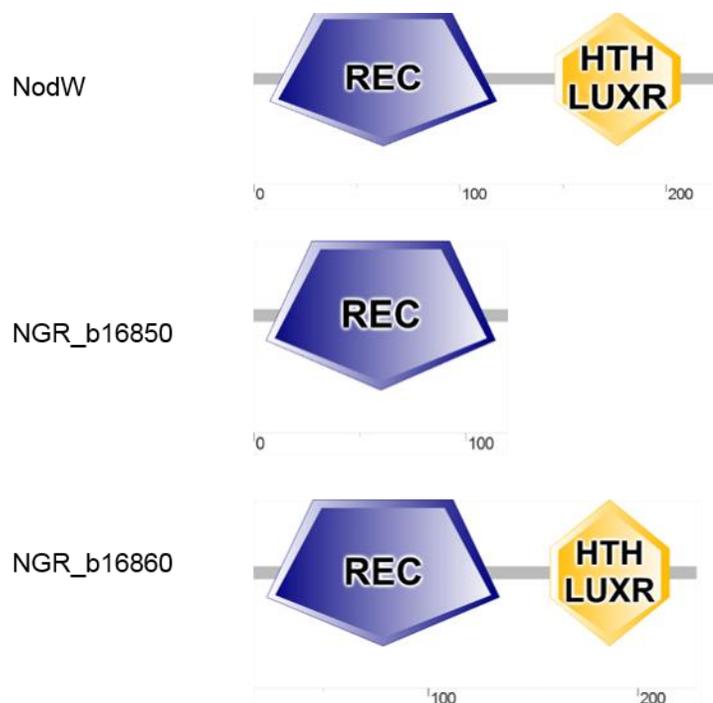


Figura 6- COMPARAÇÃO DAS ESTRUTURAS DAS PROTEÍNAS *NodW* E OS POLIPEPTÍDEOS CODIFICADOS PELAS ORFs *NGR\_b16850* E *NGR\_b16860*. FONTE: Adaptado de SMART.

O polipeptídeo codificado pela ORF *NGR\_b16860* possui 228 aminoácidos com um peso molecular calculado de 27,4 kDa e um *pI* teórico de 5,44. Com arquitetura semelhante à *NodW*, foi determinada para esta proteína a presença de um domínio REC e um domínio HTH N-terminal. Na Figura 6 observa-se a comparação das estruturas das proteínas codificadas pelos genes *nodW*, *NGR\_b16850* e *NGR\_b16860*. Como foi descrito por KAMBARA em 2008, *NGR\_b16850* apresenta 37% de identidade com *NodW* e 67% no caso de *NGR\_b16860*. A presença destas duas últimas ORFs no operon poderia indicar que elas estão diferenciadas para executar diversas funções efetoras em resposta ao estímulo percebido por *NodV*, formando parte de um sistema mais complexo de regulação como já tem sido observado em outros organismos (BIJLSMA; GROISMAN, 2003). A evidência disto é suportada pela análise de WILLIAMS e WHITWORTH (2010) que através da análise de genes de sistemas de dois componentes (TCSs) de genomas procarióticos completos da base de dados P2CS determinaram que os genes de TCSs contíguos e com distâncias menores a 200 pb entre eles muito provavelmente funcionam juntos. É necessário avaliar o efeito da

ausência de cada uma destas ORFs separadamente para uma melhor compreensão do funcionamento deste sistema.

#### 4.1.3 Busca de proteínas homólogas a NodV em outras espécies de proteobacteria

Com o intuito de avaliar se existem outros *clusters* de genes com a mesma organização que o *nodV*-NGR\_b16860 em outros membros de Proteobactéria, foi realizada uma busca de homologia na base de dados *Integrated Microbial Genomes* (IMG) utilizando o gene *nodV*. Dentre os genes semelhantes a *nodV* foram selecionados aqueles cuja identidade fosse maior ou igual a 70% e estes foram analisados no seu contexto genômico a fim de determinar se eles se encontram em uma disposição genica similar à de *nodV*-NGR\_b16860. Foi observado que todos os genes com este grau de identidade se encontram em bactérias da ordem das *Rhizobiales*, possuindo todas elas o mesmo contexto genômico, formado por uma histidina quinase e três proteínas possuindo um domínio REC. Estes genes estão presentes em 45 estirpes de *Rhizobium* sendo 37 da espécie *leguminosarum*, 4 da espécie *etli*, 1 *tropici*, 1 *gallicum*, 1 *giardinii*, 1 *bolegui* e 4 de espécie não determinada. Um gene com estas características foi encontrado no gênero *Ensifer*, pertencendo à estirpe BR816. Outras espécies de *Ensifer* possuem genes codificadores de HK com identidade mais baixa do que 70% com *nodV*. Em concordância com esses resultados, a análise de WILLIAMS e WHITWORTH, (2010) determinou que os *clusters* de genes de TCSs formados por quatro genes (ou tétrades) são o tipo menos abundante em procariotos, seguidos por tríadas, genes órfãos e pares, em ordem de abundância crescente. As tétrades estão particularmente enriquecidas em alfa-, beta- e delta-proteobacteria, estando na maioria das vezes organizadas na sequência HK-RR-HK-RR. Alfa- e beta-proteobacteria, segundo os mesmos pesquisadores, são também os grupos de organismos com mais genes de TCSs presentes em plasmídeos, característica incomum em outros táxons. Cabe destacar também que estes grupos estão compostos por muitos simbioses de parasitas e metazoários, sugerindo que possa existir alguma relação entre a presença de TCSs codificados por plasmídeos e este

estilo de vida. Por outro lado, a organização HK-RR-RR-RR vista no operon *nodV-NGR\_b16860* e outras estirpes de rizóbios não é comum em procariotos. Desta forma, se torna necessário realizar mais estudos para determinar se existe uma relação evolutiva entre todos estes *clusters*, e se eles estão cumprindo, ou não, uma função relacionada em cada um destes rizóbios.

#### 4.2 CONSTRUÇÃO DE BIBLIOTECAS DE cDNA E ANÁLISE DE TRANSCRIPTOMA

Com o objetivo de comparar os transcriptomas do mutante *nodVW<sup>-</sup>* e NGR234 selvagem em busca de genes potencialmente regulados por NodV-NodW em presença e ausência de flavonoides, foram analisadas oito bibliotecas de cDNA (uma para cada condição e uma replicata biológica de cada uma delas). No momento de realizar este trabalho as bibliotecas já se encontravam prontas, porém, foi necessário refazer uma delas (*nodVW1<sup>-</sup>*), pois esta possuía um número baixo de leituras e um comprimento médio de leituras muito diferente das demais, o que diminuiria a confiabilidade da análise. Para avaliar diferenças entre as bibliotecas foi realizada uma Análise de Componente Principal (PCA) com todas elas utilizando o programa CLC Genomics Workbench versão 7.5, o que confirmou o melhor agrupamento da nova biblioteca com as bibliotecas restantes (Apêndice 4). Em função da baixa qualidade da primeira biblioteca, os dados relativos a ela foram descartados e apenas a segunda biblioteca foi analisada.

Para a construção das bibliotecas, o RNA total das estirpes *S. fredii* NGR234 selvagem e do mutante *nodVW<sup>-</sup>* foi extraído em presença e ausência de flavonoides. As bibliotecas de cDNA geradas foram sequenciadas na plataforma Ion Proton, realizando-se duas corridas para todas as bibliotecas, exceto da *nodVW1* sem flavonoides. As bibliotecas resultantes contaram com um total de 88.624.882 leituras, que foram submetidas a um filtro de qualidade para remover sequências de baixa qualidade. Em seguida foi feito o mapeamento contra o rRNA de *S. fredii* NGR234, sendo mapeadas entre 40-72% das leituras, as quais foram excluídas de

análises posteriores. As 34.297.783 leituras restantes foram utilizadas no mapeamento RNA-seq, gerando um total de 13.361.581 leituras mapeadas. Um resumo dos dados de cada biblioteca se apresenta na Tabela 4.

Tabela 4- RESUMO DOS DADOS DE FILTRO DE QUALIDADE E MAPEAMENTOS DE TODAS AS BIBLIOTECAS UTILIZADAS USANDO CLC GENOMICS WORKBENCH. NOTA: Os valores representam o número de leituras obtidas para cada replicata biológica considerando a somatória de corridas de cada biblioteca. Os símbolos + e – representam a presença e ausência de flavonoides respectivamente.

BIBLIOTECA	NÚMERO DE LEITURAS TOTAIS	NÚMERO DE LEITURAS APÓS TRIMMING	NÚMERO DE LEITURAS MAPEADAS CONTRA rRNA	NÚMERO DE LEITURAS MAPEADAS RNA-SEQ CONTRA NGR234
NGR1 -	17.221.243	16.299.358	11.719.509	1.824.822
NGR1 +	17.307.086	16.274.314	9.739.701	1.908.386
NGR2 -	13.700.698	13.084.361	8.269.489	1.130.479
NGR2 +	3.423.993	3.236.118	2.205.200	457.221
nodVW1 -	5.042.164	4.819.092	2.860.254	489.127
nodVW1 +	2.804.241	2.611.020	1.438.230	378.731
nodVW2 -	14.893.743	14.097.423	7.574.699	4.777.414
nodVW2 +	14.231.714	13.307.126	5.623.947	2.395.401

FONTE: A autora (2016)

Com o objetivo de identificar genes candidatos a alvos de regulação do sistema de dois componentes NodV-NodW foram analisados os genes diferencialmente expressos no transcriptoma de *nodVW*<sup>-</sup> em comparação com a estirpe selvagem, em presença e ausência de flavonoides. Para obter genes controlados por flavonóides foram feitas comparações do transcriptoma de *S. fredii* NGR234 com e sem flavonoides e do transcriptoma do mutante *nodVW*<sup>-</sup> nas mesmas condições. Portanto, as seguintes comparações foram realizadas: *nodVW*<sup>-</sup>

vs NGR sem flavonoides, *nodVW*<sup>-</sup> vs NGR com flavonoides, NGR sem flavonoides vs NGR com flavonoides e *nodVW*<sup>-</sup> sem flavonoides vs *nodVW*<sup>-</sup> com flavonoides.

A análise estatística para determinar os genes DE nas quatro comparações foi realizada utilizando o programa DESeq, que gera o valor de *fold change* e o valor *p* associado a cada gene a partir das leituras mapeadas contra ele.

Foram geradas quatro tabelas de genes DE, nas quais foram mantidos para análise somente aqueles genes cuja cobertura (número de leituras do gene x comprimento médio da biblioteca/comprimento do gene) fosse igual ou maior a 3, valor *p* menor que 0,05 (equivalente a um valor de confiança de 95%) e *fold change* maior que 2 e menor que -2. As tabelas finais utilizadas na análise são apresentadas nos Apêndices 1, 2 e 3. Abaixo serão descritos os resultados obtidos.

#### 4.2.1 Genes diferencialmente expressos em *S. fredii* NGR234 em presença e ausência de flavonoides

O intercâmbio de sinais moleculares entre NGR234 e leguminosas, requerido para uma nodulação ótima, tem sido amplamente descrito na literatura. Até o momento e através de duas abordagens diferentes da nossa, dois trabalhos têm estudado a ativação por flavonóides dos loci simbióticos em NGR234 e são referência nesse campo. O primeiro trabalho foi realizado através do uso de *macroarrays* de DNA do plasmídeo simbiótico pNGR234a, e o segundo através da fusão transcricional de *nod-boxes* com o gene repórter *lacZ* (PERRET et al., 1999; KOBAYASHI et al., 2004). No estudo de Kobayashi e colaboradores foi determinada a existência de uma hierarquia temporal no padrão de ativação dos promotores que participam da cascata regulatória dos genes de nodulação.

Em concordância com os resultados dos dois trabalhos anteriormente citados, nosso estudo do transcriptoma de NGR234 após 26 h de indução com flavonoides revelou a ativação transcricional dos genes sob controle dos *nod-boxes* 2, 4, 6, 7, 9, 10, 11, 12, 14, 16, 17. Como foi descrito por Kobayashi e colaboradores, os *nod-boxes* 2, 4, 7, 9, 10, 11, 12, 14 são ativados por NodD1 logo após a indução por flavonoides (entre 1 e 6 hpi), enquanto os *nod-boxes* 6, 15, 16 e 17 tem uma

ativação tardia (após 24 hpi) a qual é dependente de NodD2. Os genes sob controle do *nod box 8* (operon *nodABCIIjnlOnoel*) não se encontraram como diferenciais na nossa comparação, como era de esperar uma vez que a sua transcrição é reprimida por NodD2 após 24 hpi (FELLAY et al., 1998; PERRET et al., 1999). Por outro lado, os genes codificadores dos reguladores transcricionais TtsI e SyrM2 (NB18 e 19), são induzidos por NodD1 24 hpi, ativando a transcrição de *nodD2* no caso do primeiro e de ORFs controlados por *tts-boxes* no segundo caso (KOBAYASHI et al., 2004). Em concordância com isto, os genes de ambos os reguladores *ttsI* e *syrM2* se mostraram induzidos por flavonóides no nosso experimento.

Após indução de *syrM2*, *nodD2* é ativado e este reprime a expressão de *nodD1*. Essa repressão não foi observada em nossa comparação, assim como também não foi vista a ativação transcricional de *nodD2*, o que leva a supor que no momento da coleta das células essas variações de expressão não seriam ainda detectáveis. Como foi relatado por Kobayashi e colaboradores (2004), a repressão de NodD1 é tardia e provavelmente dependente da síntese de quantidades suficientes de NodD2. Porém, como já foi mencionado, houve um aumento na expressão de genes controlados por *nod-boxes* ativados por NodD2, o que sugere que os níveis nesse momento existentes deste regulador seriam suficientes para ativar esses loci mas não para reprimir a transcrição de *nodD1*.

Por outro lado, uma vez ativado, TtsI liga-se às sequências conservadas chamadas *tts-boxes*, ativando a transcrição de genes envolvidos na formação da maquinaria do T3SS, fatores Nop e síntese e modificação de LPS ricos em ramnose (MARIE et al., 2004; WASSEM et al., 2008). O estudo do transcriptoma de NGR234 após a indução por flavonoides mostra expressão de todos os *tts-boxes* a exceção do TB3 e TB5, confirmando parcialmente as observações de WASSEM e col. (2008), de que todos os *tts-boxes* mostraram indução 24 hpi. No caso do TB5 essa falta de indução pode ser parcialmente explicada já que este TB tem mostrado ter uma atividade transcricional basal alta ainda na ausência de TtsI, sendo só levemente induzido por flavonoides (WASSEM et al., 2008). Além disso, as ORFs à

jusante deste TB não parecem ter funções relacionadas à nodulação, mas com a estabilização do pNGR234a (DOMBRECHT et al., 2001).

Uma vez que o complexo da nitrogenase, encarregado da fixação de nitrogênio, é marcadamente sensível ao oxigênio e que este processo tem uma grande demanda de energia, a ativação dos genes envolvidos na sua formação é estritamente regulada (FISCHER, 1994). Foi proposto que a expressão da maioria destes genes está limitada ao interior dos nódulos (PERRET et al., 1999). Porém, os mesmos autores mostraram expressão de alguns destes genes 24 hpi de NGR234 em vida livre, o que está de acordo com nossos resultados onde *fixX*, *fixA*, *fixB* e *fixC* já se expressam de 3 a 5 vezes mais após indução por flavonoides do que na ausência destas substâncias.

Vários genes relacionados ao metabolismo de energia encontraram-se induzidos após a adição de flavonóides. Entre eles os genes *atpH*, *atpA*, *atpG*, *atpF2* e *atpF1* que codificam para subunidades da ATP sintase (*fold changes* 4,25 a 2), os genes *nuoE2* e *nuoF2* (*fold changes* 2,59 e 2), subunidades da NADH-desidrogenase, e os genes NGR\_a00450, NGR\_b03130 e *cycF*, que codificam para outros componentes da cadeia respiratória (*fold changes* de 4 a 2,15). O aumento de expressão de genes envolvidos com a geração de energia já foi relatado em estudos anteriores de transcriptoma onde foi comparada a expressão gênica em vida livre com a de bacteróides de NGR234 em diferentes plantas hospedeiras (LI et al., 2013; GREMSKI STABACH, 2016- não publicado). Nestes trabalhos, a expressão dos distintos operons relacionados ao metabolismo de energia variou em função da planta hospedeira testada. Devido à grande demanda energética que implica a fixação de nitrogênio é possível que estes genes sejam ativados após a indução para prover à célula o ATP necessário para as seguintes etapas no processo de nodulação e diferenciação de bacteróides.

Um grande conjunto de genes codificadores de proteínas ribossomais foram encontrados dobrando ou triplicando a expressão após a indução com flavonoides, o que pode estar relacionado com a necessidade de síntese de novas proteínas ausentes no estado não induzido. Conjuntamente, a maior expressão de genes

relacionados à síntese de aminoácidos, metabolismo de nucleotídeos e processamento de mRNA e DNA apóiam essa hipótese.

Foi também observado um aumento na expressão de genes codificadores de sistemas de transporte, entre eles transportadores de ferro, fosfato, níquel, açúcares, as aminas espermidina e putrescina, transportadores TRAP e transportadores ABC. Embora a necessidade de captação de ferro e fosfato possa ser explicada pela maior demanda energética celular e também para a síntese de leghemoglobina, como já foi descrito em estudos de transcriptoma de bacteróides de NGR234 interagindo com *Leucaena leucocephala* e *Vigna unguiculata* (LI, Y. et al., 2013), o aumento do transporte de níquel não foi relatado anteriormente. Por outro lado, as aminas espermidina e putrescina se encontram entre os nutrientes captados por rizóbios no interior dos nódulos, já que não são comumente encontradas em rizóbios em vida livre (FUJIHARA, 2009). Não está claro por que a síntese desses transportadores começa enquanto o rizóbio ainda está em vida livre. O aumento na síntese de transportadores ABC e TRAP tem sido observado em estudos de transcriptoma de bacteróides (LI et al., 2013; GREMSKI STABACH, 2016- não publicado), mas o desconhecimento do substrato destes sistemas neste estudo dificulta a interpretação da sua importância após a indução.

Por outro lado, foi encontrado que o conjunto de genes relacionados com o metabolismo de fenilacetato (operons NGR\_c26130-NGR\_c26200, NGR\_c26090-NGR\_c26120 e o gene NGR\_c30010) foi reprimido após a indução com flavonoides (*fold changes* entre -4,10 e -2,3). Este resultado contrasta com o aumento na expressão destes genes observado em nódulos na interação *S. meliloti* com *M. truncatula* e bacteróides de NGR234 com várias plantas hospedeiras, e são necessários mais estudos para determinar a relevância do controle do metabolismo destas substâncias após a indução por flavonoides (BARNETT et al., 2004; LI et al., 2013; GREMSKI STABACH, 2016- não publicado).

Outros genes cuja expressão diminuiu após a indução incluíram reguladores transcricionais, porém, o desconhecimento dos seus alvos impediu relacioná-los com algum papel na nodulação.

O gene NGR\_a02420, potencial codificador de um regulador transcricional da família LysR, localizado no plasmídeo a, foi encontrado 3,1 vezes *down*-regulado após a indução. Este regulador diverge de um operon composto por genes potencialmente envolvidos na formação de transportadores ABC, transportadores TRAP de dicarboxilatos e enzima transcetolase. Apesar de não haver referência do envolvimento destes genes na nodulação, o fato deles estarem no plasmídeo simbiótico sugere que eles possam ter algum papel neste processo.

#### 4.2.2 Genes diferencialmente expressos em *nodVW* em presença e ausência de flavonoides

A comparação dos perfis transcricionais de *nodVW* na ausência e presença de flavonoides após a análise no DESeq gerou uma tabela com 1314 genes diferenciais, porém, a falta de genes com valor *p* abaixo de 0,05 impediu tirar conclusões estatisticamente confiáveis dessa análise.

#### 4.2.3 Genes diferencialmente expressos entre *nodVW* e NGR234

A comparação dos perfis transcricionais entre estas duas estirpes gerou duas listas de genes diferencialmente expressos. Em presença de flavonoides, 360 genes foram diferencialmente expressos na estirpe mutante enquanto na ausência destes, 137 genes. Entre eles, foram selecionados para análise os genes com maiores valores de *fold change* ou maior potencial de envolvimento no processo de nodulação. Um resumo dos genes analisados e suas funções se mostra nas Tabela 5 e Tabela 6. Números negativos representam genes controlados positivamente por NodV-NodW e positivos, reprimidos.

Tabela 5- GENES DIFERENCIALMENTE EXPRESSOS NA COMPARAÇÃO *nodVW* VS NGR234 COM FLAVONOIDES

GENE/OPERON	FOLD CHANGE	FUNÇÃO
<i>thuEFGKAB</i> (NGR_b23100-NGR_b23140)	-10 a -54	Transporte e assimilação do dissacarídeo trealose
<i>thuR</i>	-11,2	Regulador transcricional da família LacI
<i>pqqEDCBA</i> (NGR_b03260-NGR_b03300)	-7 a -12	Síntese do cofator redox pirroloquinolina-quinona (PQQ)
NGR_b03250-NGR_b03190	-4 a -11	Álcool desidrogenase, regulador transcricional da família LuxR, histidina quinase transdutora de sinal, proteína precursora de pseudoazurina e 3 proteínas hipotéticas
NGR_c15620	-16	ncRNA SuhB
<i>rmlB-wbgA</i>	-2,9 a -3,9	Síntese de LPS rico em ramnose
<i>nopB</i> ; <i>nopC</i> -NGR_a00520; <i>rhcT</i>	-8,46; -6,66; -7,7	Componentes do sistema de secreção de tipo 3 (T3SS)
<i>nopT</i> ; <i>nopJ</i>	-8,71; -9,75	Proteínas efetoras de nodulação externas (Nops)
<i>fixF</i>	-6,71	Síntese de LPS rico em ramnose
<i>rgpF</i>	-8,33	Ramnosil-transferase
<i>nodU</i>	-4,54	6-O-carbamoiltransferase, modificação de fatores Nod.
NGR_c06030-NGR_c06080	2,9 a 18,8	Metabolismo de ácidos graxos, ciclo do glioxilato, potencial componente de transportador ABC, potencial tioesterase, álcool-desidrogenase que contém ferro
<i>cysHDN</i> ; <i>cysK1</i> ; <i>cysG</i> - NGR_c13080	2,1 a 5,5	Assimilação de enxofre inorgânico

FONTE: A autora (2017)

Tabela 6- GENES DIFERENCIALMENTE EXPRESSOS NA COMPARAÇÃO *nodVW* VS NGR234 SEM FLAVONOIDES

GENE/OPERON	FOLD CHANGE	FUNÇÃO
<i>thuEFGKAB</i> (NGR_b23100-NGR_b23140)	-7 a -26	Transporte e assimilação do dissacarídeo trealose
<i>thuR</i>	-6,67	Regulador transcricional da família LacI
NGR_c26130-NGR_c26200	-9,2 a -11,4	Degradação de fenilacetato
NGR_c26090- NGR_26120	-10 a -14,3	Degradação de fenilacetato
NGR_c30010	-57,1	4-hidroxifenilpiruvato dioxigenase, metabolismo de fenilalanina e tirosina
<i>hmgA</i>	-14,3	homogentisato 1,2-dioxigenase, biossíntese de fenilalanina, tirosina e triptofano
NGR_c15620	-26	ncRNA <i>SuhB</i>
<i>cysHDN; cysK1; cysG</i> -NGR_c13080	10,8 a 38,7	Assimilação de enxofre inorgânico

FONTE: A autora (2017)

Entre os genes menos expressos em *nodVW* (*down*-regulados), destacaram-se os genes do operon *thuEFGKAB* (NGR\_b23100-NGR\_b23140), os quais tiveram expressão diferencial tanto na ausência quanto em presença de flavonoides. Em presença de flavonoides, os operons *pqqEDCBA* (NGR\_b03260-NGR\_b03300), NGR\_b03250-NGR\_b03190, *rmlB-wbgA*, *nopC*-NGR\_a00520 e os genes *nopB*, *rhcT*, *nopT*, *nopJ*, *fixF*, *rgpF* e *nodU* foram menos expressos no mutante. Na ausência de flavonoides, foi observada uma forte diminuição na expressão dos operons NGR\_c26130-NGR\_c26200, NGR\_c26090-NGR\_26120 e os genes NGR\_c30010, *hmgA* e *suhB*. Por outro lado, o conjunto de genes mais expressos (*up*-regulados) em *nodVW* incluiu, na condição com flavonoides, o operon NGR\_c06030-NGR\_c06080 e, em ambas condições, os operons *cysHDN* e *cysG*-NGR\_c13080.

O operon *thuEFGKAB* (NGR\_b23100-NGR\_b23140) se destacou por seu baixo nível de expressão na estirpe mutante em relação à selvagem, com valores

de *fold change* entre -26 a -7 em ausência de flavonoides e entre -54 e -10 na presença desses compostos. Este operon codifica genes envolvidos com o transporte e assimilação do dissacarídeo trealose, uma fonte de carbono que pode sustentar o crescimento de *S. fredii* NGR234 (BOBOYE, 2004). Em *S. meliloti*, o operon é composto pelos genes *thuE*, *-F*, *-G* e *-K*, que compõem um transportador do tipo ABC específico para os açúcares trealose, sacarose e maltose e os genes *thuA* e *thuB* implicados na utilização desses açúcares, sendo o operon todo induzido unicamente por trealose. Foi demonstrado que a mutagênese deste operon em *S. meliloti* afetou a competitividade dessa estirpe para colonizar raízes de alfafa, quando comparada com a estirpe selvagem, sugerindo que os dissacarídeos podem ser uma importante fonte de energia para sustentar o crescimento microbiano na raiz dessa planta. Testes de nodulação com o mesmo mutante não mostraram diferença com a estirpe selvagem quanto ao número e a cor dos nódulos, indicando que estes genes poderiam não ser críticos para o desenvolvimento do nódulo e a fixação de nitrogênio nessa interação. Porém, foi sugerido que a evidente falta de competitividade na colonização da raiz poderia afetar a ocupação do nódulo (JENSEN et al., 2002). Entretanto, os mesmos autores mostraram que para *S. meliloti* e *S. medicae*, a capacidade de metabolizar trealose pode afetar positiva ou negativamente a competitividade de nodulação em função do genótipo da planta hospedeira (AMPOMAH et al., 2008), uma vez que, como foi sugerido por JENSEN e colaboradores (2005), a acumulação de trealose em mutantes dos genes de utilização deste dissacarídeo (*thuA* e *thuB*), pode contribuir na luta contra diversos tipos de estresses dentro do cordão de infecção. Desta forma a diferença de distribuição deste dissacarídeo entre cultivares poderia ser a causa do efeito hospedeiro-específico da atividade dos genes de utilização de trealose na competitividade na infecção (AMPOMAH et al., 2008). Em direção oposta a este operon encontra-se o gene que codifica o regulador transcricional *ThuR*, um regulador do tipo *Lacl*, que potencialmente reprime a transcrição de *thuEFGKAB* (JENSEN et al., 2002). *thuR* também apareceu nas duas comparações com uma expressão reduzida na estirpe *nodVW*. Em outros estudos, os genes *thu* foram vistos diferencialmente expressos no transcriptoma de bacteróides de NGR234,

tendo sua expressão aumentada na interação com *Tephrosia vogelii* quando comparada com *Phaseolus vulgaris* e em *Leucaena leucocephala* e *Vigna unguiculata* quando comparada com a expressão no rizóbio em vida livre (LI et al., 2013; GREMSKI STABACH, 2016- não publicado). Foi sugerido por BOBOYE e col. (2004) que a expressão destes genes poderia contribuir com a degradação de trealose que se acumula nos nódulos, podendo ser tóxica para algumas plantas e contribuindo assim com o sucesso da simbiose. Esta informação sugere que o sistema NodV-NodW poderia ser um ativador direto ou indireto tanto do operon *thuEFGKAB* quanto do regulador *thuR*, cujos produtos proteicos podem contribuir na simbiose específica com algumas plantas hospedeiras ou na competitividade durante a colonização da raiz da planta. Além disso, o fato da expressão destes genes ser afetada pela falta de NodV-NodW na ausência de flavonoides implica que o papel destes reguladores no controle do metabolismo de trealose acontece também, embora em menor medida, fora da condição de nodulação. Isso pode ser decorrente da capacidade de metabolizar trealose para a sobrevivência do rizóbio na rizosfera, aumentando a chance de interação com a planta hospedeira. A disposição génica do operon *thuEFGKA* e do gene *thuR* no plasmídeo pNGR234b de *S. fredii* NGR234 se mostra na Figura 7.

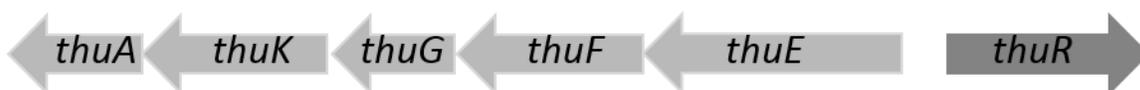


Figura 7 - DISPOSIÇÃO GÉNICA DO OPERON *thuEFGKA* E DO GENE *thuR* NO PLASMÍDEO pNGR234b. FONTE: A autora, 2017.

O operon *pqqEDCBA* (NGR\_b03260-NGR\_b03300) teve uma forte diminuição no nível de expressão no mutante, unicamente na condição com flavonoides, com valores de *fold change* de -12 a -7 para os diferentes genes que o compõem. Este operon é formado por 5 genes relacionados à síntese do cofator redox pirroloquinolina-quinona (PQQ). Esta molécula é essencial para o funcionamento de enzimas desidrogenases, envolvidas na oxidação primária de substratos do crescimento tais como álcoois, aminas e aldoses. Estas enzimas são principalmente secretadas no periplasma bacteriano (GOODWIN; ANTHONY, 1998;

MISRA et al., 2012). Estudos com uma linhagem de *S. meliloti* mutada no gene de uma glucose-desidrogenase dependente de PQQ, mostraram a necessidade dessa enzima para a competitividade e eficiência na nodulação de raízes de alfafa (BERNARDELLI et al., 2008). Além disso, tem sido mostrado que PQQ pode agir como antioxidante, protegendo as bactérias do dano oxidativo e como nutriente para suportar o crescimento bacteriano (MISRA et al., 2012). Contudo, o papel deste cofator na simbiose rizóbio-leguminosa é desconhecido.

Na direção contrária ao operon *pqqEDCBA* se encontra um segundo operon (NGR\_b03250-NGR\_b03190), que inclui genes codificadores de uma álcool-desidrogenase, um regulador transcricional da família LuxR, uma histidina quinase transdutora de sinal, uma proteína precursora de pseudoazurina e várias proteínas hipotéticas. A expressão destes genes também foi fortemente *down* regulada em *nodVW* em presença de flavonoides, com valores de *fold change* de -11 a -4. Devido a sua localização, é possível que a álcool desidrogenase seja dependente de PQQ e que o operon *pqqEDCBA* seja responsável pela síntese do cofator dessa enzima. Na Figura 8 se mostra a disposição génica dos operões *pqqEDCBA* e NGR\_b03250-NGR\_b03190 no plasmídeo pNGR234b de *S. fredii* NGR234.



Figura 8 –DISPOSIÇÃO GÉNICA DOS OPERÕES *pqqEDCBA* E NGR\_b03250-NGR\_b03190 NO PLASMÍDEO pNGR234b DE *S. fredii* NGR234. FONTE: A autora, 2017.

Também em presença de flavonoides, destacaram-se pela sua baixa expressão, vários dos genes que codificam para os componentes do T3SS NopB, NopC e RhcT e também para os efetores NopT e NopJ. Em NGR234 este complexo de secreção envolve aproximadamente 20 proteínas, algumas das quais se localizam no citoplasma e outras constituem um *pilus* que atravessa as membranas bacterianas interna e externa, e injeta proteínas no citoplasma da célula vegetal. Nove destas proteínas constituem as Nops, quatro das quais são componentes

externos da maquinaria e cinco, efetores injetados no citoplasma vegetal (VIPREY et al., 1998; MARIE et al., 2003; AUSMEES et al., 2004; SAAD et al., 2005). NopB e NopC formam parte do pilus enquanto RhcT é uma das proteínas de membrana interna (SAAD et al., 2005; STAEHELIN; KRISHNAN, 2015). Os *fold changes* destes genes foram -8,46, -6,66 e -7,7 respectivamente. Por sua vez, NopJ e NopT se encontraram *down* regulados 9,75 e 8,71 vezes em relação à estirpe selvagem. Todos estes genes se localizam no plasmídeo simbiótico e são regulados pelo ativador transcricional TtsI através da sua ligação a *tts-boxes* (TB) (WASSEM et al., 2008). Os genes *nopB* e *rhcT* são parte de um operon cuja expressão é controlada pelo TB8, enquanto *nopC* fica sob controle do TB10. A expressão de *nopJ* e de *nopT* é controlada pelos TB4 e TB11, respectivamente. Tem sido descritos fenótipos simbióticos característicos destas proteínas para várias plantas hospedeiras de *S. fredii* NGR234 (KAMBARA et al., 2009). Além disso, e também ativado por TtsI através da sua ligação ao TB2, encontrou-se diferencial e *down regulado* em nossa comparação o operon *rmlB-wbgA*, envolvido na síntese do componente rico em ramnose dos lipopolissacarídeos de superfície (MARIE et al., 2004). Estes genes mostraram *fold changes* de -2,9 à -3,9. Como já foi mencionado, a ramnose é o componente principal no LPS após a indução e é necessário na simbiose com algumas plantas hospedeiras (BROUGHTON et al., 2006). Foi demonstrado que em *Bradyrhizobium japonicum* a ativação transcricional do gene *ttsI* e, conseqüentemente, de todos os TBs, é dependente de NodW (KRAUSE et al., 2002). Porém, foi comprovado num trabalho anterior (KAMBARA, 2008), que o mutante *nodVW*<sup>-</sup> de NGR234 não apresenta um fenótipo simbiótico distinguível do selvagem quando inoculado nas leguminosas *V. unguiculata*, *V. radiata*, *M. atropurpureum*, *Lablab purpureus*, *Pachyrhizus tuberosus*, *Crotalaria juncea* e *Leucaena leucocephala*, apesar de que *L. purpureus*, *C. juncea* e *P. tuberosus* já tinham mostrado fenótipos simbióticos diferenciados quando inoculadas com estirpes T3SS<sup>-</sup>, NopT<sup>-</sup> e NopJ<sup>-</sup> (KAMBARA et al., 2009). Além disso, um mutante do operon *rmlB-wbgA* não elicitou nódulos funcionais em *V. unguiculata* (ARDISSONE et al., 2011). Apesar da falta de fenótipo simbiótico do mutante *nodVW*<sup>-</sup>, KAMBARA, (2008) mostrou a existência de ligação, embora fraca, de

NodW à região promotora de *ttsI* e também uma variação na expressão desse gene no mutante *nodVW*<sup>-</sup>. Na nossa comparação em presença de flavonoides, o gene *ttsI* foi eliminado da tabela final devido ao valor p, apesar de mostrar um valor de *fold change* de -2,9 no mutante. Isto leva a pensar que se NodV-NodW não modula a expressão de TtsI, pode ativar TBs seja em forma direta ou mediada por outro regulador em forma cooperativa com este regulador. A diminuição da expressão de *nopB*, *nopC*, *rhcT*, *nopT*, *nopJ* e *rmlB-wbgA* também sugere que a ausência de fenótipo vista nas plantas testadas por KAMBARA pode ser devido à que ainda existe um nível de expressão suficiente desses genes para não alterar a simbiose com estas espécies. Os genes controlados por *tts*-boxes analisados neste trabalho são mostrados na Tabela 7.

Tanto na ausência de flavonoides quanto na presença destas substâncias, foi detectada uma grande diminuição (*fold change* de -26 e -16 respectivamente) na transcrição do RNA não codificador (ncRNA) SuhB no mutante *nodVW*<sup>-</sup>. O gene *suhB* (NGR\_c15620) está localizado no cromossomo de NGR234 e, a montante dele e separado por 121 pares de bases existe uma segunda cópia (NGR\_c15630). Esta cópia foi encontrada 22 vezes mais expressa no mutante *nodVW*<sup>-</sup> em ausência de flavonoides, mas o valor p associado à este *fold change* foi superior a nosso limite de corte. Os ncRNAs são elementos regulatórios geralmente presentes nas regiões 5' não traduzidas dos mRNAs bacterianos e permitem perceber a presença de metabólitos específicos através de domínios aptaméricos na própria sequência, onde estas substâncias se ligam, e regulam a expressão gênica (CORBINO et al., 2005). *suhB* foi reconhecido originalmente em *Agrobacterium tumefaciens* e está presente em muitos genomas de alfa-proteobacterias, porém, não tem sido associado à regulação de genes vizinhos específicos (CORBINO et al., 2005). A forte diminuição na expressão deste gene pode indicar que ele é ativado pelo sistema NodV-NodW. Em *Pseudomonas aeruginosa*, este elemento parece estar envolvido na modulação da expressão dos componentes do T3SS, um importante fator da virulência dessa espécie. Foi sugerido que este controle se realiza de forma pós-transcricional, possivelmente através da interação de SuhB com a região regulatória do gene que codifica o regulador transcricional ExsA, que por sua vez

ativa a expressão dos genes do T3SS nessa espécie (LI, K. et al., 2013). Em *Burkholderia cenocepacia*, uma deleção no gene *suhB* também afetou a atividade dos T2SS e T6SS (ROSALES-REYES et al., 2012). O papel de SuhB na modulação da atividade de sistemas de secreção é interessante, uma vez que sugere uma potencial relação com a expressão do T3SS e Nops de NGR234. Neste sentido, a diminuição da expressão tanto de *suhB* quanto dos genes que codificam para o T3SS e Nops na estirpe *nodVW*<sup>-</sup> induzida, pode sugerir um envolvimento de SuhB na regulação desses genes. É possível que o sistema NodV-NodW ative a transcrição de *suhB* tanto na ausência quanto na presença de flavonoides, e este último participe da modulação dos genes antes citados após a percepção de algum estímulo específico. Dado o caráter leguminosa-específico dos efeitos da presença do T3SS na simbiose e do sistema NodV-NodW de *B. japonicum*, esses estímulos poderiam ser exclusivos do hospedeiro. Porém, são necessários mais estudos para esclarecer a relação entre estes elementos.

Outro gene também envolvido com a produção de LPS rico em ramnose de NGR234 é *fixF*, que também foi encontrado *down regulado* no mutante *nodVW*<sup>-</sup>. Segundo BROUGHTON e colaboradores (2006), a proteína codificada por este gene pode estar envolvida com a exportação ou ancoragem do antígeno O rico em ramnose através da membrana plasmática. *fixF* é regulado por uma via diferente à do operon *rmIB-wbgA* sendo dependente da via de ativação NodD1-SyrM2-NodD2, aonde este último regulador ativa-o através da sua ligação ao NB6 (KOBAYASHI et al., 2004). Porém, estudos tem mostrado que mutantes em *nodD2* e *syrM2* ainda produzem LPS rico em ramnose, indicando que existe outra via de regulação deste gene (BROUGHTON et al., 2006). Nossa análise de transcriptoma sugere que esta segunda via pode ser intermediada por NodV-NodW.

Ademais, o gene *rgpF*, com função semelhante à de *wbgA* foi também *down-regulado* em *nodVW*<sup>-</sup> (*fold change* -8,33). Apesar da similaridade com *wbgA*, *rgpF* mostrou ser necessário para a nodulação eficiente em *V. unguiculata* enquanto um mutante *wbgA*<sup>-</sup> induziu nódulos eficientemente (ARDISSONE et al., 2011). A

forma com que este gene é regulado ainda não foi descrita na literatura, sendo o sistema NodV-NodW o primeiro candidato para esta função.

Ainda entre os genes ativados na cascata regulatória da nodulação, o gene *nodU* também se mostrou com baixa expressão em *nodVW*<sup>-</sup> após a indução com flavonoides, com um *fold change* de -4,54. Este gene encontra-se entre os genes de modificação de fatores Nod encarregando-se especificamente de 6-O-carbamoilações. *nodU* é controlado pelo NB12 e está a jusante do gene *nodS*, que codifica uma N-metiltransferase (JABBOURI et al., 1995). O gene *nodS* apresentou um valor de *fold change* de -2, mas um valor de *p* acima do nosso ponto de corte. Apesar de estarem num mesmo operon, mais estudos são necessários para confirmar se estes genes são realmente regulados por NodV-NodW. A região *nodSU* tem mostrado ser essencial para a nodulação de *Leucaena* sp e a conjugação de *nodS* de NGR234 em *R. fredii* USDA257 confere a essa última espécie um fenótipo Nod+ Fix+, características atribuídas à presença do gene *nodS* (LEWIN; CERVANTES; CHEE-HOONG; et al., 1990; KRISHNAN et al., 1992). Embora não conhecido ainda, o gene *nodU* pode potencialmente ter efeitos simbióticos em outras leguminosas. Os genes controlados por *nod-boxes* analisados neste trabalho são mostrados na Tabela 7.

Tabela 7– *tts-boxes* E *nod-boxes* MODULADOS POR NodV-NodW.

SEQUÊNCIA CONSENSO	GENES CONTROLADOS
TB2	<i>rmlB-wbgA</i>
TB4	<i>nopJ</i>
TB8	<i>nopB-rhcU</i>
TB10	<i>nopC-NGR_a00520</i>
TB11	<i>nopT</i>
NB6	<i>fixF</i>
NB12	<i>nodSU</i>

FONTE: A autora, 2017

Diferente dos resultados obtidos por KAMBARA e col. (2008), onde foram observadas reduções à metade da expressão do regulador NodD2 a partir das 24 hpi na estirpe *nodVW*<sup>-</sup>, o estudo do transcriptoma do mutante *nodVW*<sup>-</sup> à comparação da estirpe selvagem não revelou diferenças na expressão do gene *nodD2*, nem na presença nem na ausência de flavonoides, excluindo-o da nossa lista de candidatos a alvos de regulação por NodV-NodW. De forma similar, o gene do regulador principal da cascata, *nodD1*, também não foi encontrado diferencial em nossa comparação, contrastando com as observações feitas por KAMBARA e col. de que este gene manteve uma expressão mais alta no mutante do que na estirpe selvagem tanto na presença de flavonoides quanto na ausência destas substâncias 24 hpi.

Além dos genes que compõem a cascata regulatória da nodulação, a expressão dos operons *cysHDN* e *cysG-NGR\_c13080* encontrou-se aumentada em *nodVW*<sup>-</sup> tanto na presença quanto na ausência de flavonoides. Em presença de flavonoides eles apresentaram *fold changes* que vão de 3,3 a 5,5 e na ausência deles, apresentaram valores de 10 a 38 vezes a expressão na estirpe selvagem. Os dois operons estão envolvidos com a assimilação de enxofre inorgânico em microrganismos. Os genes do operon *cysHDN* são induzidos em condições de deficiência de enxofre e reprimidos por cisteína. Eles codificam enzimas envolvidas na redução de sulfato, que será logo incorporado em compostos orgânicos, principalmente o aminoácido cisteína. Em *Bacillus subtilis*, o operon parece ser reprimido pela enzima CysK, a qual catalisa o último passo na síntese deste aminoácido, ou por seu próprio produto (ALBANESI et al., 2005). Em NGR234, esse último passo é catalisado pelos produtos dos genes *cysK1* e *cysK2*. *cysK1* também se encontrou com mais expresso no transcriptoma de *nodVW*<sup>-</sup>, 4,3 vezes mais expresso em presença de flavonoides e 10,8 vezes na ausência desses compostos. Quanto ao operon *cysG-NGR\_c13080*, foi estudado na estirpe *S. fredii* H0N1, onde foi também associado à redução de sulfato e sulfito na via de assimilação de enxofre (LU et al., 2008). O enxofre, além de ser um elemento essencial para a síntese de

proteínas e cofatores enzimáticos, é necessário para a nodulação devido à sua utilização na modificação de fatores Nod, afetando a interação com algumas plantas hospedeiras (PERRET et al., 2000). Embora a sulfatação dos fatores Nod em rizóbios ter sido fortemente associada aos genes *nodPQ*, os quais produzem o doador de sulfato PAPs (3'-fosfoadenosina-5'-fosfosulfato), estes genes não foram induzidos em NGR234 nem por flavonoides nem pelos reguladores NodD1, SyrM2 e SyrM1 (SENA, 2013- não publicado), sugerindo que os produtos de outros genes podem ser os produtores de doador de sulfato para a modificação de fatores Nod nessa estirpe. Somado a isso, outros autores têm mostrado que o *pool* de doador de sulfato APS (adenosina 5'-fosfosulfato, a forma desfosforilada de PAPs) produzido tanto pelas enzimas NodPQ quanto por CysDN, pode ser utilizado para a sulfatação de fatores Nod assim como também para a síntese de componentes celulares, sendo estas fontes intercambiáveis em *Sinorhizobium* sp BR816 (SNOECK et al., 2003). Esses dados suportam a hipótese de que em NGR234 os genes *cysHDN* podem ter o mesmo papel dual na síntese de doadores de sulfato visto em *Sinorhizobium* sp BR816. O fato da expressão destes genes se encontrar aumentada tanto na ausência quanto na presença de flavonoides, e com um *fold change* maior na primeira condição, sugere que a ação regulatória de NodV-NodW poderia ser a repressão destes genes na ausência de indução ou, na ausência de algum estímulo vegetal específico durante a interação, como também sugere um envolvimento destes reguladores em funções alheias à nodulação.

Na estirpe mutante em presença de flavonoides, o operon NGR\_c06030-NGR\_c06080 alcançou valores de expressão até 18 vezes maiores do que na estirpe selvagem. Este operon é composto por genes envolvidos com o metabolismo de ácidos graxos e com o ciclo do glioxilato, além de genes que codificam para um potencial componente de transportador ABC, uma provável tioesterase e uma álcool-desidrogenase que contém ferro. Apesar destes genes não terem sido previamente associados à interação de NGR234 com leguminosas, o envolvimento deste operon no metabolismo de ácidos graxos pode indicar uma relação com este processo. Tem sido observado um acúmulo de ácidos graxos em nódulos de soja e em rizóbios extraídos desses nódulos (JOHNSON et al., 1966).

Porém, foi sugerido que eles cumpririam uma função estrutural e não de provisão de energia, já que foi vista uma baixa atividade de enzimas chave do ciclo do glioxilato nesses nódulos. Através deste ciclo são produzidos compostos de 4 carbonos capazes de ingressar no ciclo dos ácidos tricarboxílicos para provisão de energia celular, a partir de compostos como ácidos graxos, entre outros (JOHNSON et al., 1966; LI, Y. et al., 2013). A diminuição de atividade do ciclo do glioxilato nos nódulos sugeriu que os ácidos graxos poderiam estar destinados a outros fins, como por exemplo estruturais, já que muitos compostos carbonados utilizados como fonte de energia são disponibilizados pela planta (JOHNSON et al., 1966). Por outro lado, a observação de maiores quantidades de lipídeos em bacteróides de *Rhizobium leguminosarum* do que em bactérias de vida livre, sugere que a síntese destes compostos é uma alternativa do metabolismo do acetil-CoA que não pode ser totalmente oxidado no ciclo dos ácidos tricarboxílicos devido ao ambiente microaeróbico do nódulo. Desta forma, os autores desse estudo sugerem que os bacteróides podem ser lipogênicos como uma resposta fisiológica ao processo de fixação de nitrogênio (TERPOLILLIA et al., 2016). Entretanto, o aumento de expressão destes genes no mutante *nodVW* sugere que eles são reprimidos por NodV-NodW na estirpe selvagem após a indução por flavonoides, o que pode indicar uma diminuição da síntese destes compostos e uma maior utilização deles como fonte de energia. A repressão de genes relacionados à síntese de ácidos graxos tem sido observada também num estudo de proteoma de bacteróides de *B. japonicum*, quando comparado com o proteoma de vida livre (SARMA; EMERICH, 2006).

Na ausência de flavonoides, vários genes envolvidos com o metabolismo de fenilacetato e fenilalanina foram entre 57 e 9 vezes menos expressos em *nodVW* em relação à estirpe selvagem, indicando uma forte atividade ativadora de NodV-NodW sobre eles. Estes genes pertencem aos operons NGR\_c26130-NGR\_c26200, NGR\_c26090-NGR\_26120 e os genes NGR\_c30010 e *hmgA*. Os operons NGR\_c26130-NGR\_c26200 e NGR\_c26090-NGR\_26120 se encontram na mesma região do cromossoma de *S. fredii* NGR234, em posições convergentes um do outro e codificam para proteínas da degradação de fenilacetato, enquanto o gene

NGR\_c30010 se encontra isolado em outra posição do cromossomo e codifica para uma 4-hidroxifenilpiruvato dioxigenase, envolvida no metabolismo dos aminoácidos fenilalanina e tirosina. O gene *hmgA* codifica para uma homogentisato 1,2-dioxigenase, enzima envolvida na biossíntese dos aminoácidos fenilalanina, tirosina e triptofano. O gene *hmgA* pertence a um operon com vários genes envolvidos no metabolismo de aminoácidos aromáticos, porém, os demais componentes do operon, apesar de mostrar valores de *fold changes* entre -2 e -3,77, apresentaram valores de p inaceitáveis conforme nosso critério de corte e por isso não foram incluídos nesta análise. O composto fenilacetato além de ser uma fonte de carbono e energia para muitos microrganismos é intermediário no metabolismo dos aminoácidos fenilalanina e tirosina (ISMAIL et al., 2003), e tem sido descrito como um dos produtos das vias de degradação de flavonoides em NGR234 e outros rizóbios (RAO; COOPER, 1994). As vias que envolvem essas enzimas produzem compostos que podem ingressar no ciclo dos ácidos tricarbóxicos e ser utilizados como fonte de energia (ARIAS-BARRAU et al., 2004). Um aumento da expressão dos genes de metabolismo de fenilacetato foi observado em bacteróides de NGR234 em *Leucaena leucocephala* e de *Sinorhizobium meliloti* em *Medicago truncatula* em comparação à expressão destes genes em rizóbios de vida livre (BARNETT et al., 2004; LI, Y. et al., 2013). Entretanto, os resultados da análise do transcriptoma sugerem que esses genes seriam ativados por NodV-NodW na ausência de flavonoides, uma vez que em *nodVW*<sup>-</sup> eles são reprimidos, sugerindo que este par de reguladores poderia ter outros papéis fora do processo de nodulação. A disposição gênica dos operons NGR\_c26130-NGR\_c26200 e NGR\_c26090-NGR\_c26120 no cromossomo de *S. fredii* NGR234 se mostra na Figura 9.



Figura 9 – DISPOSIÇÃO GÊNICA DOS OPERONS NGR\_c26130-NGR\_c26200 E NGR\_c26090-NGR\_c26120 NO CROMOSSOMO DE *S. fredii* NGR234. FONTE: A autora, 2017.

### 4.3 VALIDAÇÃO DE GENES DE POR ENSAIO DE GFP

Para validar os resultados de RNA-seq e assim confirmar os possíveis alvos de regulação de NodV-NodW foram utilizadas fusões transcricionais das regiões regulatórias dos genes *nopJ* (TB4), *nopT* (TB11), *fixF* (NB6) e *thuR*; e os operons *nopC*-NGR\_a00520 (TB10), *nopB-rhcU* (TB8), *pqqEDCBA*, *thuEFGKA* e *cysHDN* com o gene da proteína fluorescente verde (GFP). Enquanto os genes da cascata regulatória da nodulação são induzidos por flavonoides e dependem do regulador transcricional NodD1 para a sua ativação, a expressão dos genes *thuR*, *thuEFGKA*, *pqqEDCBA* e *cysHDN* em resposta a flavonoides e NodD1 é desconhecida. Por este motivo todas as fusões transcricionais foram introduzidas na estirpe selvagem *S. fredii* NGR234 como controle positivo e nas estirpes mutantes *nodVW*<sup>-</sup>, *syrM2*<sup>-</sup>, e *nodD1*<sup>-</sup>. Esta última estirpe foi utilizada como controle negativo no caso dos genes da cascata de nodulação. O nível de atividade foi acompanhado em presença e ausência de flavonoides durante cinco dias. Os experimentos referentes à *nopB-rhcU*, *nopJ*, *fixF* e *cysHDN* foram realizados três vezes e os referentes à *nopC*-NGR\_a00520, *nopT*, *thuR* e *pqqEDCBA*, quatro vezes. Os experimentos referentes ao operon *thuEFGKA* estão sendo repetidos devido a problemas no crescimento das estirpes transconjugantes e por este motivo são apresentados os resultados parciais correspondentes a dois experimentos. Foram utilizadas três replicatas biológicas para o primeiro experimento de cada região gênica e quatro para os experimentos seguintes. Nas figuras 4 à 8 estão representados os gráficos obtidos de um experimento representativo de cada uma das regiões regulatórias testadas, sendo obtidos resultados semelhantes todas as vezes que estes foram repetidos. Para o caso do operon *pqqEDCBA* é mostrado o gráfico correspondente à média de três experimentos realizados.

Todas as fusões contendo regiões regulatórias de genes envolvidos na cascata de nodulação (TB4, TB8, TB10, TB11 e NB6) mostraram os perfis esperados de indução por flavonoides na estirpe selvagem e a ausência de indução em *nodD1*<sup>-</sup>. Na ausência de flavonoides não houve indução de GFP em nenhuma

destas linhagens, tendo todas elas o mesmo perfil, mostrado apenas para TB4 na Figura 10.

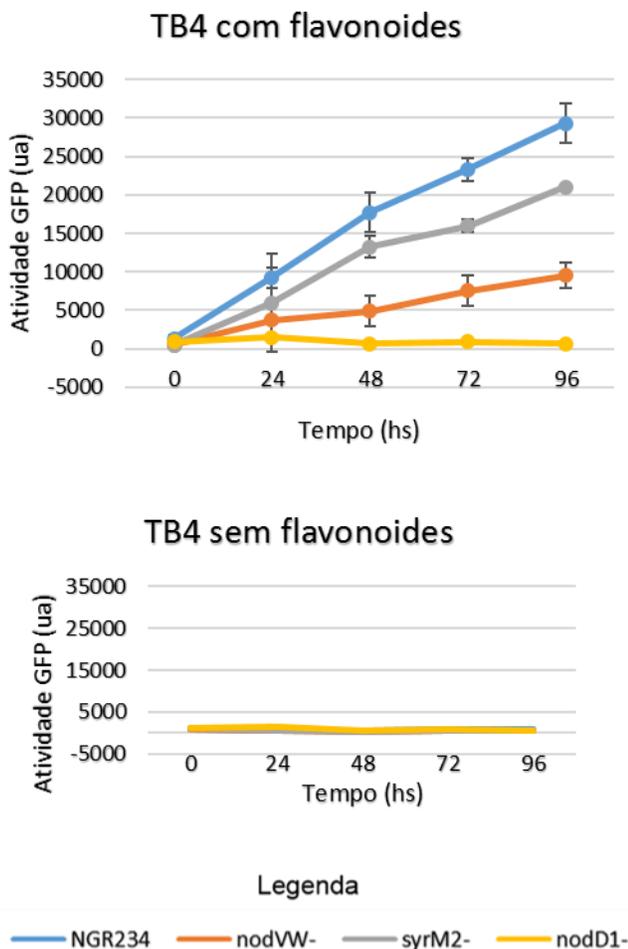


Figura 10- NÍVEIS DE EXPRESSÃO DE GFP CONTROLADOS PELA REGIÃO REGULATÓRIA DE *nopJ* (TB4). Os níveis de GFP foram medidos em presença e ausência dos flavonoides apigenina, crisina, genisteína, kaempferol e naringenina (0,2  $\mu$ M) durante 5 dias a cada 24 horas. Abreviaturas: ua unidades arbitrárias, hs horas. FONTE: A autora (2017)

No caso dos genes *nopJ* (TB4) e *nopT* (TB11), foi observada uma diminuição da expressão tanto na estirpe *nodVW*<sup>-</sup> quanto em *syrM2*<sup>-</sup> já às 24 hpi, confirmando a necessidade de NodV-NodW e de SyrM2 para obter o nível de expressão observado na estirpe selvagem, confirmando os seus *fold changes* negativos vistos no transcriptoma de *nodVW*<sup>-</sup> (Figura 10 e Figura 11). Como já foi mencionado a via conhecida de ativação transcricional destes genes responde à

ligação de TtsI aos TB4 e TB11. Os resultados aqui mostrados revelam que em presença de flavonóides estes genes precisam de NodV-NodW e SyrM2, além de TtsI para serem completamente ativados, sendo a dependência de NodV-NodW mais evidente. Como mostrado anteriormente, estes efetores são importantes para a nodulação de diversas plantas hospedeiras. Em *Tephrosia vogelii*, NopT parece ter um efeito positivo sobre o número de nódulos enquanto NopJ não mostrou afetar a simbiose (KAMBARA et al., 2009). Os resultados mostrados neste estudo motivam a testar os efeitos do mutante *nodVW*<sup>-</sup> na nodulação desta espécie, já que estes testes não foram realizados previamente.

O operon *nopB-rhcU* (TB8) mostrou na estirpe *nodVW*<sup>-</sup> um perfil de atividade intermediário entre os dos controles NGR234 e *nodD1*<sup>-</sup> (Figura 11), confirmando a baixa expressão vista no transcriptoma dos genes *nopB* e *rhcT*, e dos outros genes presentes no operon, os quais apesar de mostrarem valores de *p* acima do nosso ponto de corte, mostraram valores de *fold change* menores que -2. O papel regulador do sistema NodV-NodW na ativação deste operon foi evidente a partir das 24 hpi com flavonoides, momento em que já mostra atividade na estirpe selvagem e uma diminuição da expressão na estirpe *nodVW*<sup>-</sup>. A ativação do TB8 já foi mostrada a partir das 4 hpi por Wassem e colaboradores (2008) e por Perret e colaboradores, (1999). O fato de não ter sido vista expressão diferencial do regulador TtsI no transcriptoma de *nodVW*<sup>-</sup> sugere que a ativação deste operon por NodV-NodW poderia ser através da união direta de NodW à esta região regulatória ou, indiretamente, através de outros intermediários. Resultados similares aos mostrados para *nopB* foram obtidos por KAMBARA e col. (2008), embora os dados de qPCR não tenham sido similares ao de GFP realizados no mesmo trabalho. Por outro lado, o regulador SyrM2 não parece ter nenhum papel na regulação deste operon sendo que a curva de atividade é idêntica à selvagem. Este resultado contrasta com o observado por Sena, (2003 –não publicado), que mostrou um efeito ativador em TB8 por este regulador a partir das 60 hpi.

Já para o operon *nopC-NGR\_a00520* (TB10), a expressão no mutante *nodVW*<sup>-</sup> mostrou uma curva de expressão mais baixa que a da estirpe selvagem,

aproximando-se mais à curva de *nodD1*<sup>-</sup>, confirmando que este operon depende de NodV-NodW para a sua expressão, como mostrou a análise do transcriptoma de *nodVW*<sup>-</sup> (Figura 11). Como já tem sido relatado para o operon *nopB-rhcU*, o resultado deste ensaio demonstra que todos os genes do operon, cujos valores de *fold change* são menores a -2, dependem de NodV-NodW para a sua expressão e não apenas aqueles com valores de *p* aceitáveis. Por outro lado, na estirpe *syrM2*<sup>-</sup> o perfil de expressão praticamente não se afastou daquele observado na estirpe selvagem, indicando que SyrM2 não é necessário para a expressão de TB10.

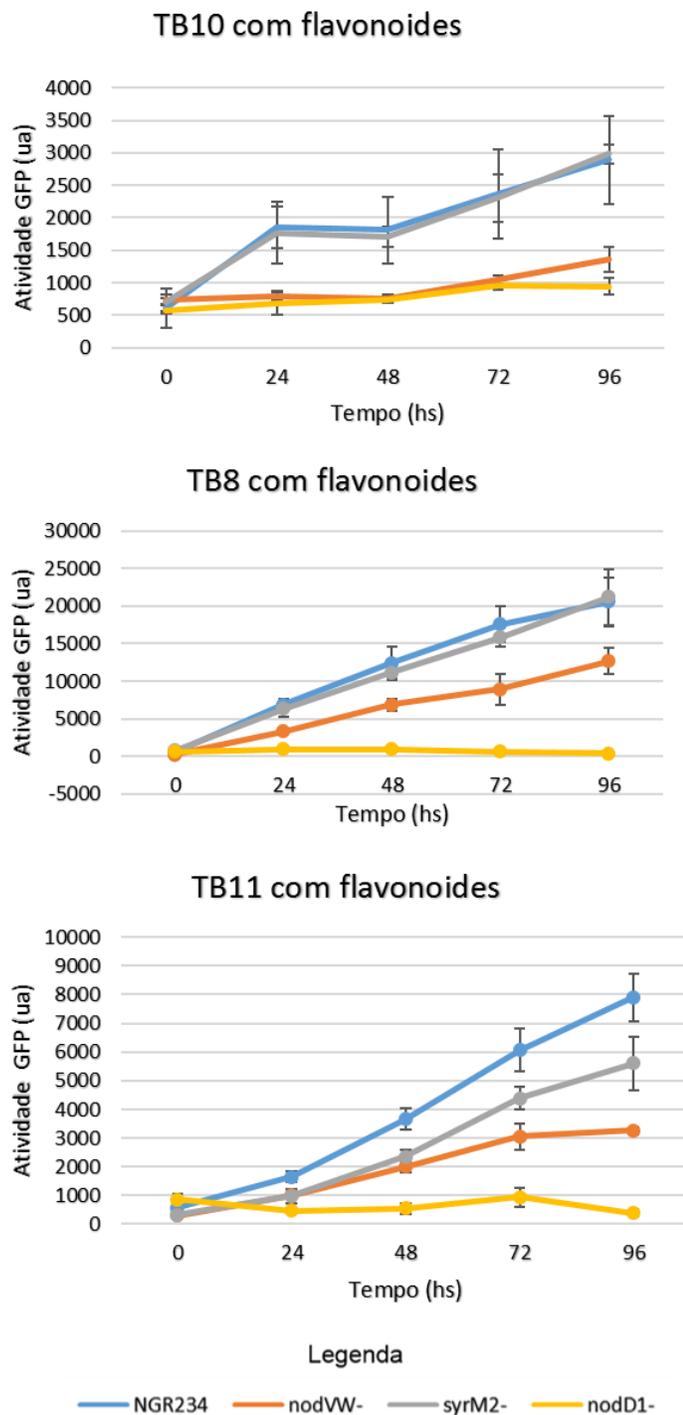


Figura 11 - NÍVEIS DE EXPRESSÃO DE GFP CONTROLADOS PELAS REGIÕES REGULATÓRIAS DOS GENES *nopB-rhcU* (TB8), *nopC-NGR\_A00520* (TB10) E *nopT* (TB11). Os níveis de GFP foram medidos em presença e ausência dos flavonoides apigenina, crisina, genisteína, kaempferol e naringenina (0,2  $\mu$ M) durante 5 dias a cada 24 horas. Abreviaturas: ua unidades arbitrárias, hs horas. FONTE: A autora (2017)

No caso do gene *fixF* (NB6), os perfis de expressão nas estirpes *nodVW*<sup>-</sup> e *syrM2*<sup>-</sup> foram intermediários entre os dos controles, sendo o perfil na linhagem *nodVW*<sup>-</sup> mais próxima ao da linhagem selvagem (Figura 12). Isto confirma a necessidade de NodV-NodW para a ativação completa de *fixF*, como foi observado no transcriptoma deste mutante. Ademais, isto corrobora as observações feitas em 2006 por Broughton e colaboradores, de que a ativação do NB6 não é unicamente dependente da via NodD1-SyrM2-NodD2 e mostrando que o sistema NodV-NodW pode ser a via alternativa que ativa a expressão deste gene. O fato da expressão de *fixF* não ser completamente abolida no mutante *syrM2*<sup>-</sup>, como mostra o ensaio de GFP também reforça esta hipótese.

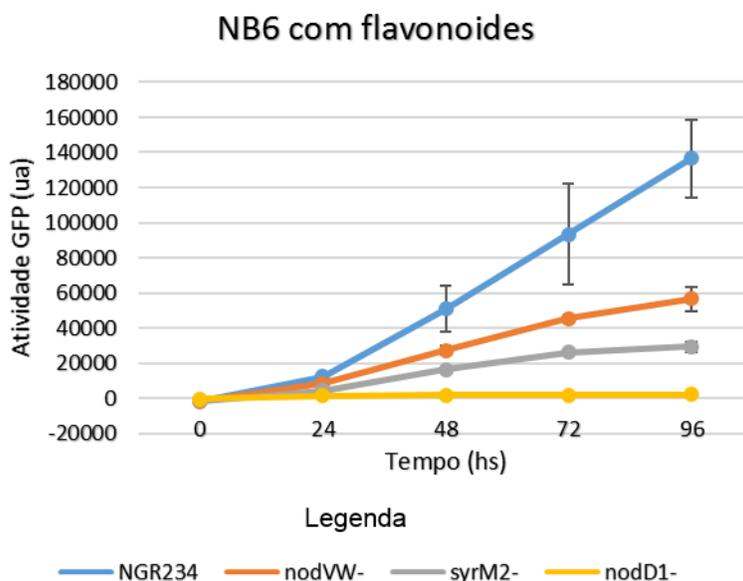


Figura 12 - NÍVEIS DE EXPRESSÃO DE GFP CONTROLADOS PELA REGIÃO REGULATÓRIA DO GENE *fixF* (NB6). Os níveis de GFP foram medidos em presença e ausência dos flavonoides apigenina, crisina, genisteína, kaempferol e naringenina (0,2  $\mu$ M) durante 5 dias a cada 24 horas. Abreviaturas: ua unidades arbitrárias, hs horas. FONTE: A autora (2017)

A atividade transcricional do operon *thuEFGKA* não pôde ser confirmada pelo ensaio de GFP devido aos problemas no crescimento das linhagens transconjugantes. Os resultados parciais parecem confirmar a total dependência deste operon da presença de um sistema NodV-NodW funcional para sua

transcrição tanto em presença quanto na ausência de flavonoides, uma vez que a sua expressão foi completamente abolida na ausência destes indutores. Da mesma forma, os testes realizados apontam para uma dependência pelo regulador SyrM2, embora novos ensaios precisem ser feitos para, além de confirmar os *fold changes* negativos observados no transcriptoma de *nodVW*<sup>-</sup> em presença e ausência de flavonoides, determinar a importância de NodD1 na expressão destes genes. Estes resultados preliminares não mostram nenhuma influência da presença de flavonoides na expressão destes genes, sendo esta visivelmente constante durante o período avaliado em presença destas substâncias e sem diferença significativa com a expressão nas culturas induzidas. Contrariando os valores de *fold change* vistos para o gene *thuR*, a expressão de GFP a partir da sua região regulatória não mostrou nenhuma diferença entre as linhagens testadas, assim como também não com a adição de flavonoides (Figura 13). Isto sugere que a expressão deste gene é independente de NodD1, SyrM2 e NodV-NodW e que os valores negativos de *fold change* observados na análise de transcriptoma podem ser devidos às flutuações na concentração de mRNA de *thuR* ocasionadas por fatores desconhecidos. Desta forma, a regulação do operon *thuEFGKA* por NodV-NodW poderia ser através da união direta de NodW à sua região regulatória ou através de outro regulador intermediário. Entretanto, em função da sua localização, é bastante provável que ThuR seja o regulador do operon *thu*. Potencialmente, sua expressão pode não ser regulada, mas sim sua atividade, através da interação com uma molécula sinalizadora.

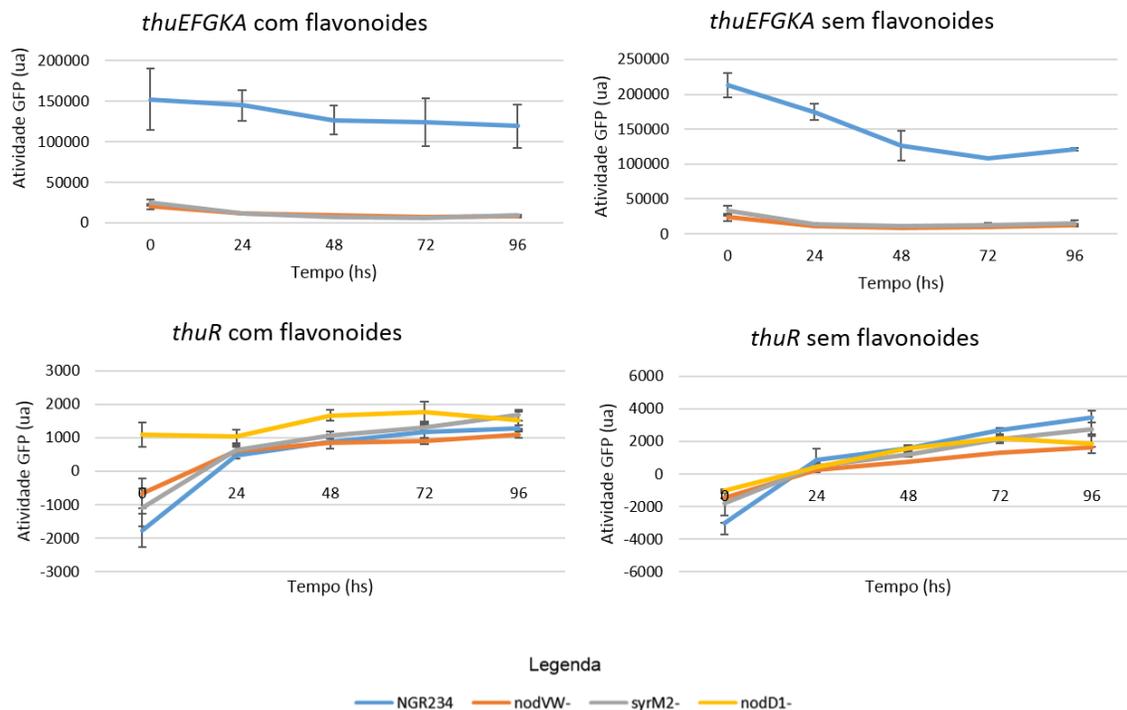


Figura 13- NÍVEIS DE EXPRESSÃO DE GFP CONTROLADOS PELAS REGIÕES REGULATÓRIAS DOS GENES *thuR*, *thuEFGKA*. Os níveis de GFP foram medidos em presença e ausência dos flavonoides apigenina, crisina, genisteína, kaempferol e naringenina (0,2  $\mu$ M) durante 5 dias a cada 24 horas. Abreviaturas: ua unidades arbitrárias, hs horas. FONTE: A autora (2017)

Da mesma forma, a atividade do operon *cysHDN* também parece não responder à presença dos reguladores NodV-NodW, NodD1, SyrM2 nem à adição de flavonoides, estando a expressão a partir da sua região regulatória nessas estirpes no mesmo nível que da estirpe selvagem e nas duas condições testadas (Figura 14). A maior quantidade de leituras de mRNA mapeadas contra este operon no transcriptoma de *nodVW*<sup>-</sup> pode ser devido a flutuações na concentração deste mRNA, não detectadas pelo ensaio de GFP. Desta forma não é possível por este meio confirmar o resultado visto no transcriptoma de *nodVW*<sup>-</sup> e devem ser feitos outros experimentos diferentes com este fim.

No caso do operon *pqqEDCBA*, o perfil de expressão mais baixa observado no transcriptoma de *nodVW*<sup>-</sup> pôde ser confirmado pelo ensaio de GFP tanto na condição com flavonoides quanto na ausência deles, já que apesar de estarem com altos valores de *p*, os *fold changes* de todos os genes do operon foram menores

que -2 nas duas condições (Figura 14). O ensaio também revelou uma independência da atividade deste operon da presença de flavonoides, sendo seu perfil de expressão praticamente o mesmo nas duas condições. O ensaio revelou também uma dependência do regulador NodD1 para a expressão destes genes, que assim como na ausência dos reguladores NodV-NodW, a ausência do regulador NodD1 parece diminuir a transcrição destes genes nas primeiras 48 hs do ensaio. Isto pode indicar que a síntese de PQQ seja importante durante o processo de nodulação antes e depois da indução por flavonoides. Num estudo realizado através da análise RNAseq do transcriptoma de mutantes de genes essenciais para as duas vias de regulação do quórum sensing em NGR234, foi determinado que este operon é reprimido pela via dependente do gene *ngrI* (KRYSCIAK et al., 2014). No mesmo estudo foi mostrado que a ausência destes genes não afetou a expressão dos genes da cascata regulatória da nodulação nem o fenótipo de nodulação nas leguminosas *V. unguiculata*, *V. radiata* e *T. vogelii*, sugerindo que o sistema de quórum sensing não tem importância crucial na nodulação de NGR234. Em conjunto, estes resultados mostram que a síntese de PQQ pode ser necessária para a nodulação e apontam que a ativação dos genes do operon *pqqEDCBA* no contexto da cascata regulatória da nodulação independe da regulação por quórum sensing.

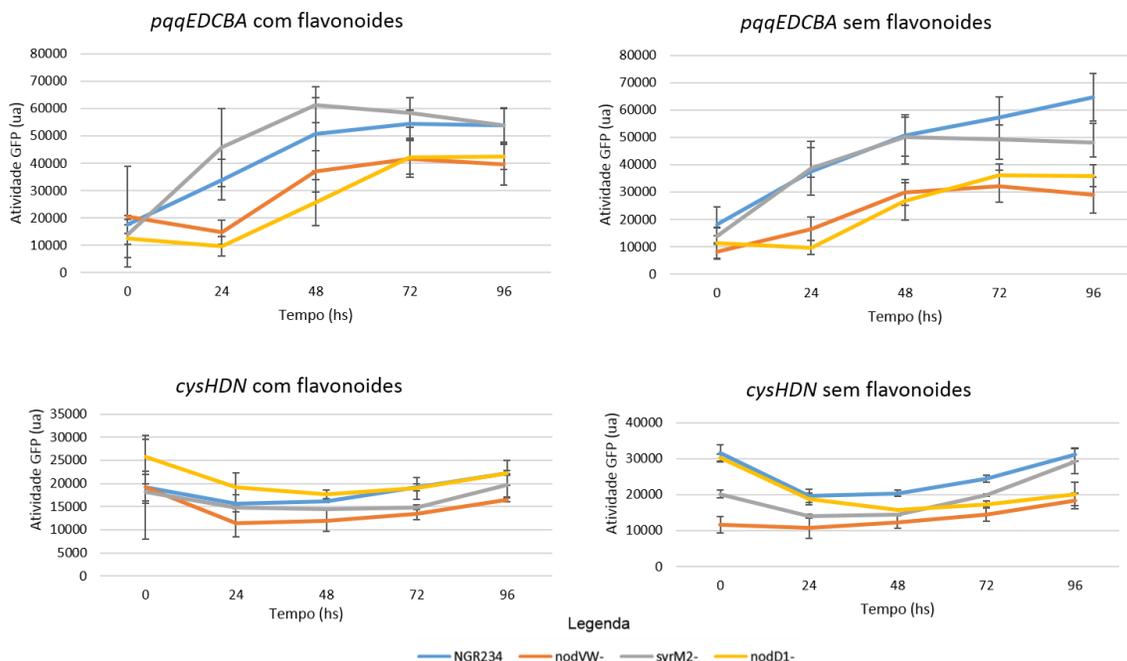


Figura 14 - NÍVEIS DE EXPRESSÃO DE GFP CONTROLADOS PELAS REGIÕES REGULATÓRIAS DOS GENES *cysHDN* E *pqqEDCBA*. Níveis de GFP foram medidos em presença e ausência dos flavonoides apigenina, crisina, genisteína, kaempferol e naringenina ( $0,2 \mu\text{M}$ ) durante 5 dias a cada 24 horas. Abreviaturas: ua unidades arbitrárias, hs horas. FONTE: A autora (2017)

#### 4.4 PURIFICAÇÃO DE NodW

Com o objetivo de avaliar a capacidade de interação direta com regiões alvo, a proteína NodW foi purificada. Para tanto, foram planejadas construções para que esta fosse obtida com cauda de histidinas na região N- e C-terminal, bem como uma contendo apenas o domínio de ligação ao DNA, presente na região C-terminal. A proteína NodW foi clonada em sua forma completa adicionando uma cauda de histidina na extremidade N-terminal e faremos referência a esta como NodW-His. Na segunda estratégia foi clonado o gene completo, mas a cauda de histidina foi adicionada na região C-terminal da proteína; faremos referência a este produto como NodW-C-His. Na terceira estratégia foi clonada só a sequência que codifica o fragmento C-terminal de NodW com a cauda de histidina no N-terminal do mesmo e esta construção foi identificada como NodWCterm-His. Uma clonagem inicial das três construções foi realizada no vetor pTZ57R e a presença de bandas de tamanho

esperado foi confirmada por PCR de colônia, assim como também por restrição enzimática dos vetores purificados. A ausência de mutação foi avaliada por sequenciamento e os insertos previamente liberados foram subclonados nos vetores de expressão pET29a para clonagem *in frame* da proteína com cauda de histidina no C-terminal e pET28a para as clonagens com cauda His no N-terminal. A subclonagem foi confirmada por restrição enzimática (Figura 15).

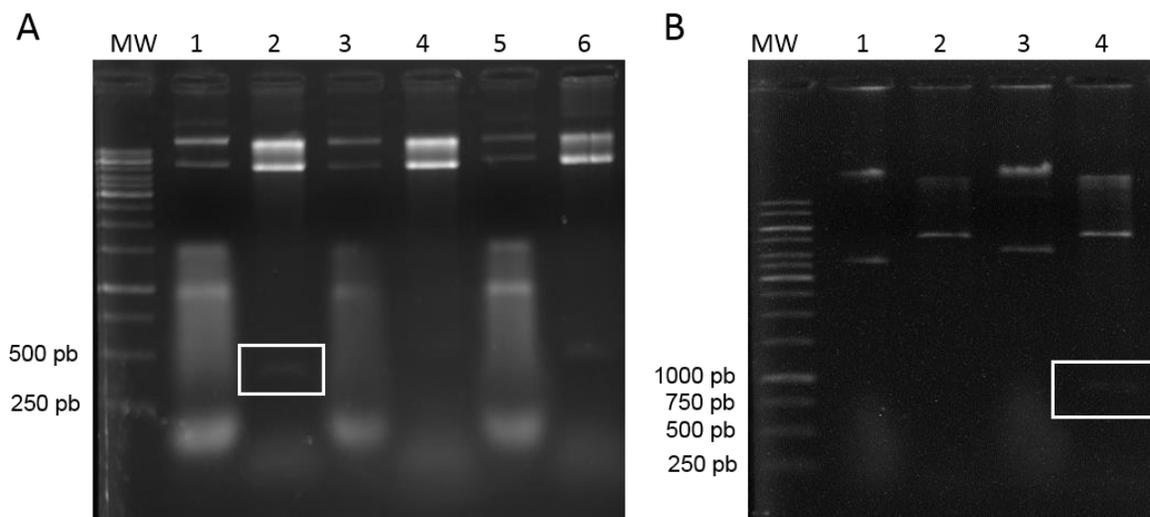


Figura 15 - RESTRIÇÃO ENZIMÁTICA DE pET28-nodWCterm-His E pET29-nodW-C-His. NOTA: Painel A - Restrição enzimática de pET28-NodWCterm-His com XbaI e EcoRI. As linhas 1, 3, 5 correspondem aos vetores sem cortar, e 2, 4 e 6 aos vetores cortados. O quadro branco indica a banda de tamanho esperado (aproximadamente 270 pb). A colônia correspondente à linha 2 foi utilizada para indução e purificação de NodWCterm-His. Painel B - Restrição enzimática de pET29-nodW-C-His com XbaI e EcoRI. As linhas 1 e 3 correspondem aos vetores sem cortar e as linhas 2 e 4 aos vetores cortados. O quadro branco indica a banda de tamanho esperado (aproximadamente 800 pb), a mesma colônia foi utilizada para indução e purificação de NodW-C-His. FONTE: A autora (2016)

As construções obtidas foram transformadas na estirpe de *E. coli* BL21 (DE3) e a presença das proteínas na fase solúvel foi avaliada por indução em pequena escala e SDS-PAGE.

Em seguida, os produtos NodW-C-His e NodWCterm-His foram induzidos em larga escala e purificados. NodW-C-His foi purificada utilizando sistema de cromatografia AKTA (GE Healthcare), sendo eluída nas frações de Imidazol 50 mM (Figura 16, Painel B). As três frações mais concentradas da proteína de tamanho esperado (25,63 kDa) foram reunidas e dialisadas separadamente das três frações mais diluídas. O produto NodWCterm-His foi purificado manualmente utilizando as mesmas condições e a proteína eluiu nas frações correspondentes a concentração de imidazol de 120-500 mM (Figura 16, Painel A). Todas as frações que apresentaram a banda correspondente a proteína de tamanho aproximado 10 kDa também copurificaram um polipeptídeo maior. Foram realizadas várias tentativas de separação destes polipeptídeos incluindo purificação através de coluna cromatográfica HiTrap Heparine e purificação com coluna HiTrap Chelating carregada com Níquel em condições desnaturantes, com tampão de Uréia, conforme recomendações do fabricante da coluna (Novagen). Estas últimas tentativas não produziram melhores resultados na purificação e os ensaios seguintes foram feitos com as frações purificadas com o primeiro método.

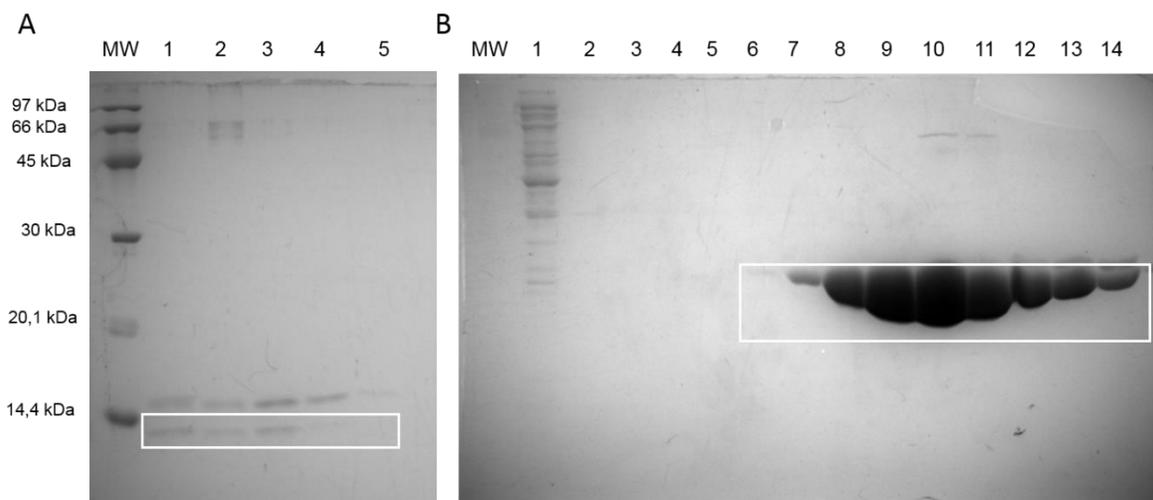


Figura 16 - PURIFICAÇÃO DAS PROTEÍNAS DE FUSÃO NodWCterm-His E NodW-C-His. Painel A- Purificação de NodWCterm-His. Gel de acrilamida 16%. A linha 1 corresponde à fração eluída com tampão A 120 mM de imidazol, 2 e 3 200 mM, 4 e 5 500 mM. O quadro branco indica a localização das proteínas no gel. As frações correspondentes às linhas 1, 3, 4 e 5 foram reunidas e dialisadas e posteriormente utilizadas no EMSA. Painel B- Purificação de NodW-C-His. Gel de acrilamida 12%. As linhas 1 a 14 correspondem a: 1-Lavagem da coluna com tampão A, 2 e 3- 50 mM imidazol, 4 e 5- 70 mM, 6- 90 mM, 7 e 8- 100 mM, 9 e 10- 120 mM, 11 e 12- 200 mM, 13 e 14- 1 M (Tampão B). As frações correspondentes às linhas 9, 10 e 11 e 8, 12 e 13 foram reunidas separadamente e dialisadas e armazenadas a  $-20^{\circ}\text{C}$ . FONTE: A autora (2016)

#### 4.4 EMSA

O ensaio de EMSA foi realizado utilizando o produto de amplificação por PCR das regiões regulatórias de *syrM2* (NB19), *nopB-rhcU* (TB8), *thuEFGKAB* e *pqqEDCBA*. As sequências dos *primers* utilizados são mostradas na Tabela 3. Todas as reações se realizaram incubando os produtos de PCR em presença de concentrações crescentes da proteína NodWCterm-His. Na Figura 17, são mostrados os resultados para o ensaio correspondente a *nopB*, sendo este o único que mostrou ligação de NodWCterm-His.

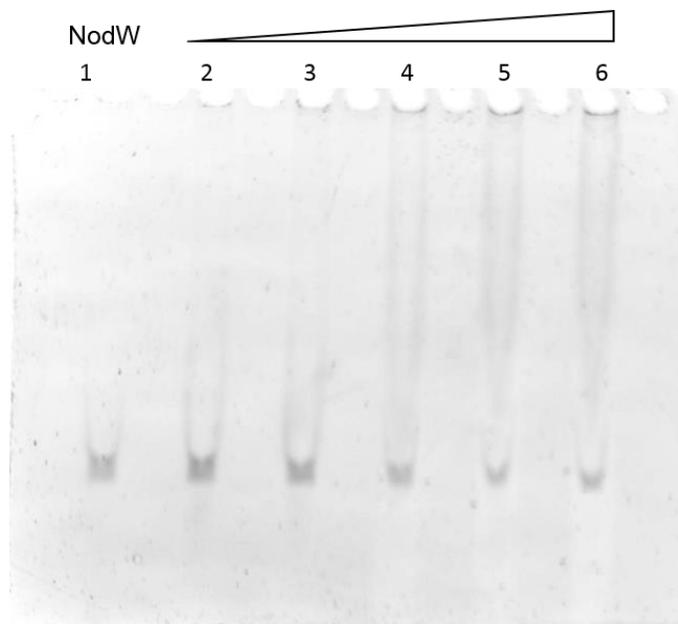


Figura 17- ENSAIO DE EMSA UTILIZANDO A REGIÃO TB8 COMO ALVO. As linhas 1 a 6 correspondem a volumes de NodWCterm-His de 0, 0,8, 2, 4, 6, e 8  $\mu$ L respectivamente. Foi utilizado um volume de reação de 10  $\mu$ L. FONTE: A autora (2016)

Na Figura 17 observa-se a presença de retardo na migração da banda correspondente à TB8, sendo isto mais visível na medida em que se aumenta a concentração de proteína na reação. Este resultado sugere que o papel ativador de NodV-NodW sobre *nopB* poderia ser através da ligação direta à TB8. Foram vistos resultados similares com a região regulatória de *thuEFGKAB* (não mostrado) porém, estes resultados não puderam ser confirmados por ensaios de competitividade já que não foi possível reproduzir as condições de ligação. Desta forma, se existe ligação de NodW às regiões regulatórias destes genes deverá ser confirmado em novos estudos.

## 5 CONSIDERAÇÕES FINAIS

Procurou-se nesse estudo esclarecer a função do sistema de dois componentes NodV-NodW de NGR234, cujos homólogos já foram estudados em *B. japonicum*. Nesse organismo, o sistema NodV-NodW constitui uma via de regulação dos genes *nod* e de *TtsI*, independente de NodD, e os seus efeitos na nodulação são hospedeiro-específicos (GÖTTFERT et al., 1990; SANJUAN et al., 1994; KRAUSE et al., 2002). Através da análise do transcriptoma do mutante *nodVW*<sup>-</sup> foi mostrado que NodV-NodW é necessário para ativação de vários genes da cascata regulatória da nodulação. Entre eles, os genes que codificam para componentes do T3SS, contidos nos operons *nopB-rhcU* e *nopC-NGR\_a00520*, os que codificam os efetores NopT e NopJ, secretados pelo T3SS, (MARIE et al., 2003) e os responsáveis pela síntese e exportação do polissacarídeo rico em ramnose, *rmlB-wbgA*, *fixF* e *rgpF* (BROUGHTON et al., 2006; ARDISSONE et al., 2011; STAEHELIN; KRISHNAN, 2015). É sabido que *nopB-rhcU*, *nopC-NGR\_a00520*, *nopT*, *nopJ* e *rmlB-wbgA* são regulados por *TtsI* através da sua ligação aos *tts-boxes* (WASSEM et al., 2008), enquanto *fixF* é dependente da ativação por NodD2 (KOBAYASHI et al., 2004; BROUGHTON et al., 2006). A via de ativação de *rgpF* não foi ainda descrita na literatura. Os dados observados do transcriptoma indicam uma via potencial para a ativação dos TBs, já que NodV-NodW também parece ativar a transcrição do gene que expressa o ncRNA *SuhB*, que já foi associado à ativação de sistemas de secreção bacterianos (ROSALES-REYES et al., 2012; LI, K. et al., 2013). A baixa expressão dos TB4, TB8, TB10, TB11 e NB6 em *nodVW*<sup>-</sup> foi confirmada *in vivo* neste trabalho através do ensaio com o gene repórter GFP. Além disso, a expressão destes genes, testada no mutante *syrM2*<sup>-</sup>, revelou que este regulador também é necessário para a ativação completa dos TB11, TB4 e NB6, e não é necessário para ativação dos TB10 e TB8. Por outro lado, NodV-NodW também parece ativar a transcrição do operon *nodSU*, normalmente ativado após a indução de flavonoides por NodD1 (KOBAYASHI et al., 2004). Amplamente estudado como um loci de especificidade simbiótica em NGR234 (LEWIN et al.,

1990; KRISHNAN et al., 1992; JABBOURI et al., 1995), o fato de ser modulado por NodV-NodW aponta mais fortemente à um papel destes reguladores no controle de sinais simbióticos específicas do hospedeiro.

Além dos genes da cascata de nodulação, genes envolvidos com o metabolismo geral bacteriano também mostraram ser modulados por NodV-NodW. Entre eles, os genes de transporte e assimilação de trealose *thuEFGKAB* e seu potencial regulador transcricional *thuR* (JENSEN et al., 2002), e o operon *pqqEDCBA*, envolvido na síntese do cofator redox PQQ, são ativados por NodV-NodW independentemente da adição de flavonoides, como foi confirmado pelo ensaio de GFP. Através deste último ensaio pôde ser confirmado que o operon *pqqEDCBA* é também dependente de NodD1, o que torna relevante o estudo da participação deste cofator na nodulação. Um conjunto de genes envolvidos com o metabolismo de fenilacetato e fenilalanina mostraram ser ativados por NodV-NodW na ausência de flavonoides, enquanto este sistema mostrou ser um repressor de genes do metabolismo lipídico após a indução com flavonoides, e dos genes de assimilação de enxofre inorgânico independentemente da adição destas substâncias.

Em termos gerais, os resultados vistos neste estudo apontam para um papel complexo do sistema NodV-NodW no metabolismo celular de NGR234, tendo uma forte influência na modulação de fatores de nodulação de conhecido efeito hospedeiro-específico assim como também de vias metabólicas gerais. Um resumo dos genes modulados por este sistema em presença de flavonoides se mostra na Figura 18.

Por outro lado, o contexto genômico em que se encontram os genes *nodV* e *nodW*, também sugere que os três reguladores de resposta codificados por *nodW*, NGR\_b16850 e NGR\_b16860 podem funcionar juntos em resposta à ativação por NodV, podendo cada um destes RR ter funções efetoras diferenciadas ou reconhecer sequências específicas nos genes alvo, como já foi observado em outros sistemas de dois componentes bacterianos (LAUB; GOULIAN, 2007).

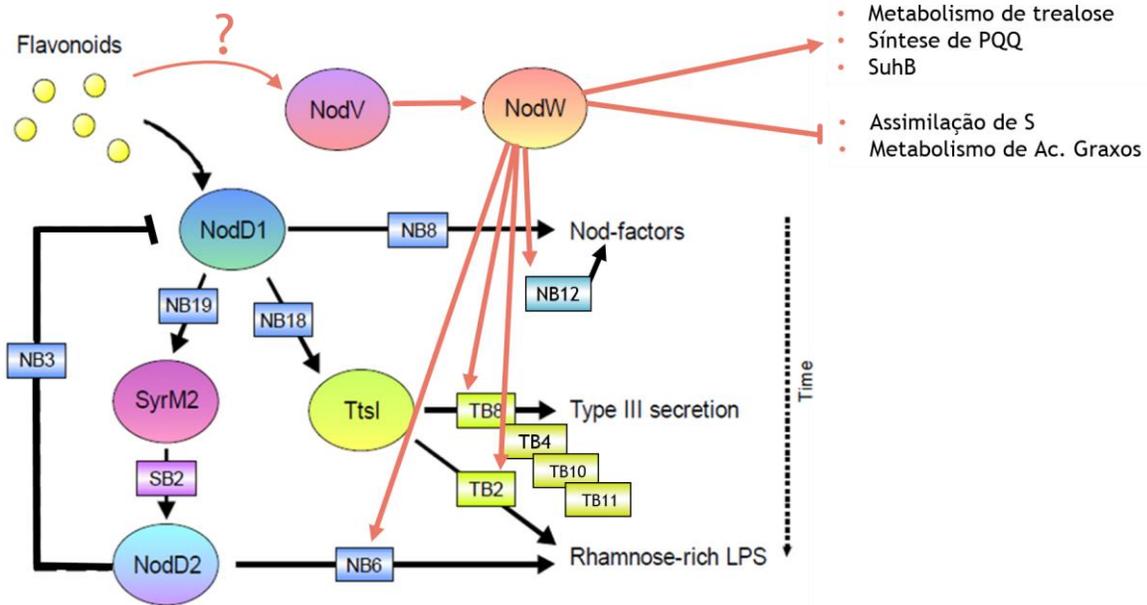


Figura 18- RESUMO DOS GENES MODULADOS POR NodV-NodW EM PRESENÇA DE FLAVONOIDES. A linha pontilhada indica o decorrer do tempo durante a ativação da cascata de nodulação. Linhas terminadas em seta orientadas às sequências consenso ou funções celulares indicam ativação transcricional enquanto linhas terminadas em barra indicam repressão transcricional. FONTE: adaptado de Kambara, 2008.

## 6 CONCLUSÃO

- Através da análise do transcriptoma de *nodVW*<sup>-</sup> foi determinado que o sistema NodV-NodW ativa a transcrição de genes da cascata regulatória da nodulação em presença de flavonoides. Entre eles se encontraram vários genes envolvidos com a especificidade simbiótica deste rizóbio na sua interação com leguminosas, como componentes do T3SS, Nops, genes de síntese de LPS e de modificação de fatores Nod.
- Na presença de flavonoides NodV-NodW também modulam a transcrição de genes do metabolismo geral bacteriano, como metabolismo de trealose, síntese de PQQ, assimilação de enxofre inorgânico, metabolismo de ácidos graxos e o ncRNA SuhB.
- Na ausência de flavonoides, foi observada modulação de genes do metabolismo de trealose, síntese de PQQ, assimilação de enxofre inorgânico e degradação de fenilacetato.
- A expressão diferencial dos genes *nopB-rhcU*, *nopC-NGR\_a00520*, *nopJ*, *nopT*, *fixF*, *pqqEDCBA* no mutante *nodVW*<sup>-</sup> foi confirmada através do ensaio com o gene repórter GFP, assim como também a necessidade de SyrM2 para a ativação de TB11, TB4 e NB6 após a indução por flavonoides. Também por meio deste ensaio foi detectada uma dependência de NodD1 para a ativação do operon *pqqEDCBA*, independentemente da adição de flavonoides.
- Foi vista uma fraca ligação da proteína de fusão NodWCterm-His à região regulatória dos genes *nopB* e *thuEFGKA*, porém, a especificidade desta interação precisa ser confirmada por novos ensaios.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ALBANESI, D.; MANSILLA, M. C.; SCHUJMAN, G. E.; MENDOZA, D. DE. Bacillus subtilis Cysteine Synthetase Is a Global Regulator of the Expression of Genes Involved in Sulfur Assimilation. **Journal of Bacteriology**, v. 187, n. 22, p. 7631–7638, 2005.
- ALM, E. J.; HUANG, K. H.; PRICE, M. N.; et al. The MicrobesOnline Web site for comparative genomics. **Genome Research**, v. 15, p. 1015–1022, 2005.
- ALTSCHUL, S. F.; MADDEN, T. L.; SCHÄFFER, A. A.; et al. Gapped BLAST and PSI-BLAST : a new generation of protein database search programs. **Nucleic Acids Research**, v. 25, n. 17, p. 3389–3402, 1997.
- AMPOMAH, O. Y.; JENSEN, J. B.; BHUVANESWARI, T. V. Lack of trehalose catabolism in Sinorhizobium species increases their nodulation competitiveness on certain host genotypes. **New Phytologist**, v. 179, p. 495–504, 2008.
- ANDERS, S.; HUBER, W. Differential expression analysis for sequence count data. **Genome Biology**, v. 11, p. R106, 2010.
- ARDISSONE, S.; NOEL, K. D.; KLEMENT, M.; BROUGHTON, W. J.; DEAKIN, W. J. Synthesis of the Flavonoid-Induced Lipopolysaccharide of Rhizobium Sp . Strain NGR234 Requires Rhamnosyl Transferases Encoded by Genes rgpF and wbgA. **Molecular Plant-Microbe Interactions**, v. 24, n. 12, p. 1513–1521, 2011.
- ARIAS-BARRAU, E.; OLIVERA, R.; LUENGO, M.; et al. The Homogentisate Pathway : a Central Catabolic Pathway Involved in the Degradation of L - Phenylalanine , L -Tyrosine , and 3-Hydroxyphenylacetate in Pseudomonas putida. **Journal of Bacteriology**, v. 186, n. 15, p. 5062–5077, 2004.
- AUSMEES, N.; KOBAYASHI, H.; DEAKIN, W. J.; et al. Characterization of NopP , a Type III Secreted Effector of Rhizobium sp . Strain Characterization of NopP , a Type III Secreted Effector of Rhizobium. **Journal of Bacteriology**, v. 186, n. 14, p. 4774–4780, 2004.
- BARNETT, M. J.; TOMAN, C. J.; FISHER, R. F.; LONG, S. R. A dual-genome Symbiosis Chip for coordinate study of signal exchange and development in a prokaryote – host interaction. **PNAS**, v. 2675, n. 6, 2004.
- BARTSEV, A. V; DEAKIN, W. J.; BOUKLI, N. M.; et al. NopL , an Effector Protein of Rhizobium sp . NGR234 , Thwarts Activation of Plant Defense Reactions 1. **Plant Physiology**, v. 134, n. February, p. 871–879, 2004.

BERCK, S.; PERRET, X.; BROUGHTON, W. J.; JABBOURI, S. NodL of *Rhizobium* sp. Strain NGR234 Is Required for O<sup>6</sup>-Acetyltransferase Activity. **Journal of Bacteriology**, v. 181, n. 3, p. 957–964, 1999.

BERINGER, J. E. R Factor Transfer in *Rhizobium leguminosarum*. **Journal of General Microbiology**, v. 84, n. 1, p. 188–198, 1974.

BERNARDELLI, C. E.; LUNA, M. F.; GALAR, M. L.; BOIARDI, J. L. Symbiotic phenotype of a membrane-bound glucose dehydrogenase mutant of *Sinorhizobium meliloti*. **Plant Soil**, v. 313, p. 217–225, 2008.

BIJLSMA, J. J. E.; GROISMAN, E. A. Making informed decisions : regulatory interactions between two-component systems. **Trends in Microbiology**, v. 11, n. 8, p. 359–366, 2003.

BOBOYE, B. Degradation of trehalose by rhizobia and characteristics of a trehalose-degrading enzyme isolated from *Rhizobium* species NGR234. **Journal of Applied Microbiology**, v. 97, n. 2, p. 256–261, 2004.

BROUGHTON, W. J.; HANIN, M.; RELIC, B.; et al. Flavonoid-Inducible Modifications to Rhamnan O Antigen Are Necessary for *Rhizobium* sp. Strain NGR234-Legume Symbioses. **Journal of Bacteriology**, v. 188, n. 10, p. 3654–3663, 2006.

BROUGHTON, W. J.; JABBOURI, S.; PERRET, X. Keys to symbiotic harmony. **Journal of Bacteriology**, v. 182, n. 20, p. 5641–5652, 2000.

BROUGHTON, W. J.; WONG, C. H.; LEWIN, A.; et al. Identification of *Rhizobium* Plasmid Sequences Involved in Recognition of *Psophocarpus*, *Vigna*, and Other Legumes. **The Journal of Cell Biology**, v. 102, n. April, p. 1173–1182, 1986.

CHEN, I. A.; MARKOWITZ, V. M.; CHU, K.; et al. IMG / M : integrated genome and metagenome comparative data analysis system. **Nucleic Acids Research**, v. 45, n. October 2016, p. 507–516, 2017.

CORBINO, K. A.; BARRICK, J. E.; LIM, J.; et al. Evidence for a second class of S-adenosylmethionine riboswitches and other regulatory RNA motifs in alpha-proteobacteria. **Genome Biology**, v. 6, n. 8, 2005.

DEHAL, P. S.; JOACHIMIAK, M. P.; PRICE, M. N.; et al. MicrobesOnline : an integrated portal for comparative and functional genomics. **Nucleic acids Research**, v. 38, n. November 2009, p. 396–400, 2010.

DENARIE, J.; DEBELLE, F.; ROSENBERG, C. Signaling and Host Range Variation in Nodulation. **Annual Review of Microbiology**, v. 46, n. 1, p. 497–531,

1992. Disponível em:

<<http://www.annualreviews.org/doi/abs/10.1146/annurev.mi.46.100192.002433>>. .

DIXON, R.; KAHN, D. GENETIC REGULATION OF BIOLOGICAL NITROGEN FIXATION. **Nature Reviews Microbiology**, v. 2, n. August, p. 621–631, 2004.

DOBSON, L.; REMÉNYI, I.; TUSNÁDY, G. E. CCTOP : a Consensus Constrained TOPology prediction web server. **Nucleic Acids Research**, v. 43, n. May, p. 408–412, 2015.

DOMBRECHT, B.; VANDERLEYDEN, J.; MICHIELS, J. Stable RK2-Derived Cloning Vectors for the Analysis of Gene Expression and Gene Function in Gram-Negative Bacteria. **Molecular Plant-Microbe Interactions**, v. 14, n. 3, p. 426–430, 2001.

DUCROS, A. M.; LEWIS, R. J.; VERMA, C. S.; et al. Crystal Structure of GerE , the Ultimate Transcriptional Regulator of Spore Formation in *Bacillus subtilis*. **Journal of Molecular Biology**, v. 306, p. 759–771, 2001.

FELLAY, R.; HANIN, M.; MONTORZI, G.; et al. nodD 2 of *Rhizobium* sp . NGR234 is involved in the repression of the nodABC operon. **Molecular Microbiology**, v. 27, n. 2, p. 1039–1050, 1998.

FELLAY, R.; PERRET, X.; VIPREY, V.; BROUGHTON, W. J.; BRENNER, S. Organization of host-inducible transcripts on the symbiotic plasmid of *Rhizobium* sp. NGR234. **Molecular Microbiology**, v. 16, n. 4, p. 657–667, 1995.

FENG, J.; LI, Q.; HU, H. L.; CHEN, X. C.; HONG, G. F. Inactivation of the nod box distal half-site allows tetrameric NodD to activate nodA transcription in an inducer-independent manner. **Nucleic Acids Research**, v. 31, n. 12, p. 3143–3156, 2003.

FINN, R. D.; ATTWOOD, T. K.; BABBITT, P. C.; et al. InterPro in 2017 — beyond protein family and domain annotations. **Nucleic Acids Research**, v. 45, n. November 2016, p. 190–199, 2017.

FINN, R. D.; COGGILL, P.; EBERHARDT, R. Y.; et al. The Pfam protein families database : towards a more sustainable future. **Nucleic Acids Research**, v. 44, n. December 2015, p. 279–285, 2016.

FISCHER, H. Genetic Regulation of Nitrogen Fixation in Rhizobia. **Microbiological Reviews**, v. 58, n. 3, p. 352–386, 1994.

FRANCHE, C.; LINDSTRÖM, K.; ELMERICH, C. Nitrogen-fixing bacteria associated with leguminous and non-leguminous plants. **Plant and Soil**, v. 321, n. 1-2, p. 35–59, 2009.

FUJIHARA, S. Minireview Biogenic Amines in Rhizobia and Legume Root Nodules. **Microbes Environ.**, v. 24, n. 1, p. 1–13, 2009.

GASTEIGER, E.; HOOGLAND, C.; GATTIKER, A.; et al. Protein Identification and Analysis Tools on the ExPASy Server. **The Proteomics Protocols Handbook**. p.571–608, 2005.

GOODWIN, P. M.; ANTHONY, C. The Biochemistry , Physiology and Genetics of PQQ and PQQ-containing Enzymes. **Advances in Microbial Physiology**. v. 40, 1998.

GÖTTFERT, M.; GROB, P.; HENNECKE, H. Proposed regulatory pathway encoded by the nodV and nodW genes, determinants of host specificity in *Bradyrhizobium japonicum*. **Proceedings of the National Academy of Sciences of the United States of America**, v. 87, n. 7, p. 2680–2684, 1990.

GREMSKI STABACH, C. ANÁLISE DE EXPRESSÃO DIFERENCIAL DE *Sinorhizobium fredii* NGR234. , 2016.

GROB, P.; MICHEL, P.; HENNCCKE, H.; et al. of a *Bradyrhizobium japonicum* nodW mutant. **Molecular Genetics and Genomics**, v. 241, p. 531–541, 1993.

HANIN, M.; JABBOURI, S.; QUESADA-VINCENS, D.; et al. Sulphation of *Rhizobium* sp. NGR234 Nod factors is dependent on noeE, a new host-specificity gene. **Molecular microbiology**, v. 24, n. 6, p. 1119–1129, 1997.

HERBST FIORESE, W. WHFIORESE. , 2015.

HUECK, C. J. Type III Protein Secretion Systems in Bacterial Pathogens of Animals and Plants Type III Protein Secretion Systems in Bacterial Pathogens of Animals and Plants. **Microbiology and Molecular Biology Reviews**, v. 62, n. 2, p. 379–433, 1998.

ISMAIL, W.; MOHAMED, M. E.; WANNER, B. L.; et al. Functional genomics by NMR spectroscopy Phenylacetate catabolism in *Escherichia coli*. **European Journal of Biochemistry**, v. 3054, p. 3047–3054, 2003.

JABBOURI, S.; FELLAY, R.; TALMONT, F.; et al. Involvement of nodS in N - Methylation and nodU in 6- O - Carbamoylation of *Rhizobium* sp . NGR234 Nod Factors. **The Journal of Biological Chemistry**, v. 270, n. 39, p. 22968–22973, 1995.

JENSEN, J. B.; AMPOMAH, O. Y.; DARRAH, R.; PETERS, N. K.; BHUVANESWARI, T. V. Role of Trehalose Transport and Utilization in *Sinorhizobium meliloti* – Alfalfa Interactions. **Molecular Plant-Microbe Interactions**, v. 18, n. 7, p. 694–702, 2005.

JENSEN, J. B.; PETERS, N. K.; BHUVANESWARI, T. V. Redundancy in periplasmic binding protein-dependent transport systems for trehalose, sucrose, and maltose in *Sinorhizobium meliloti*. **Journal of Bacteriology**, v. 184, n. 11, p. 2978–2986, 2002.

JOHNSON, G. V; EVANS, H. J.; CHING, T. Enzymes of the Glyoxylate Cycle in Rhizobia and Nodules of Legumes<sup>1</sup>. **Plant Physiology**, v. 1, n. 90, p. 1330–1336, 1966.

JOLY, N.; ZHANG, N.; BUCK, M.; ZHANG, X. Coupling AAA protein function to regulated gene expression. **Biochimica et Biophysica Acta**, v. 1823, n. 1, p. 108–116, 2012. Elsevier B.V. Disponível em: <<http://dx.doi.org/10.1016/j.bbamcr.2011.08.012>>. .

JONES, P.; BINNS, D.; CHANG, H.; et al. Sequence analysis InterProScan 5 : genome-scale protein function classification. **Bioinformatics**, v. 30, n. 9, p. 1236–1240, 2014.

KAMBARA, K. Thesis Regulation and effects of the type-three secretion system of “ Rhizobium ” species NGR234 Regulation and Effects of the Type-three Secretion System of Rhizobium species NGR234. **Infection**, 2008.

KAMBARA, K.; ARDISSONE, S.; KOBAYASHI, H.; et al. Rhizobia utilize pathogen-like effector proteins during symbiosis. **Molecular Microbiology**, v. 71, n. October 2008, p. 92–106, 2009.

KANNENBERG, E. L.; CARLSON, R. W. Lipid A and O-chain modifications cause Rhizobium lipopolysaccharides to become hydrophobic during bacteroid development. **Molecular Microbiology**, v. 39, n. 2, p. 379–391, 2001.

KOBAYASHI, H.; BROUGHTON, W. J.; BIOLOGIE, L. DE; SUPÉRIEURES, P. Chapter 5 FINE-TUNING OF SYMBIOTIC GENES IN RHIZOBIA : FLAVONOID SIGNAL TRANSDUCTION CASCADE. **Media**, p. 117–152, 2008.

KOBAYASHI, H.; NACIRI-GRAVEN, Y.; BROUGHTON, W. J.; PERRET, X. Flavonoids induce temporal shifts in gene-expression of nod-box controlled loci in Rhizobium sp. NGR234. **Molecular Microbiology**, v. 51, n. 2, p. 335–347, 2004.

KRAUSE, A.; DOERFEL, A.; GÖTTTFERT, M. Mutational and transcriptional analysis of the type III secretion system of *Bradyrhizobium japonicum*. **Molecular plant-microbe interactions : MPMI**, v. 15, n. 12, p. 1228–1235, 2002.

KRISHNAN, H. B.; LEWIN, A.; FELLAY, R.; BROUGHTON, W. J.; PUEPPKE, S. G. Differential expression of nodS accounts for the varied abilities of *Rhizobium fredii* USDA257 and *Rhizobium* sp . strain NGR234 to nodulate *Leucaena* spp . **Molecular Microbiology**, v. 6, n. 22, p. 3321–3330, 1992.

KRYSCIAK, D.; GROTE, J.; ORBEGOSO, R.; et al. RNA Sequencing Analysis of the Broad-Host-Range Strain *Sinorhizobium fredii* NGR234 Identifies a Large Set of Genes Linked to Quorum Sensing-Dependent Regulation in the Background of a *tral* and *nglI* Deletion Mutant. **Applied and Environmental Microbiology**, v. 80, n. 18, p. 5655–5671, 2014.

LAEMMLI, U. K. Cleavage of Structural proteins during the Assembly of the Head of bacteriophage T4. **Nature**, v. 227, p. 680–685, 1970.

LAUB, M. T.; GOULIAN, M. Specificity in Two-Component Signal Transduction Pathways. **Annual Review of Genetics**, v. 41, p. 121–145, 2007.

LETUNIC, I.; DOERKS, T.; BORK, P. SMART : recent updates , new developments and status in 2015. , v. 43, n. October 2014, p. 257–260, 2015.

LEWIN, A.; CERVANTES, E.; CHEE-HONG, W.; BROUGHTON, W. J. *nodSU*, two new *nod* genes of the broad host *Rhizobium* strain NGR234 encode host-specific nodulation of the tropical tree *Leucaena*. **Molecular Plant-Microbe Interactions**, v. 3, n. 5, p. 317–326, 1990.

LEWIN, A.; CERVANTES, E.; CHEE-HOONG, W.; BROUGHTON, W. J. *nodSU*, two new *nod* genes of the broad host range *Rhizobium* strain NGR234 encode host-specific nodulation of the tropical tree *Leucaena leucocephala*. **Mol Plant Microbe Interact.**, v. 15, n. 12, p. 317–326., 1990.

LI, K.; XU, C.; JIN, Y.; et al. *SuhB* Is a Regulator of Multiple Virulence Genes and Essential for Pathogenesis of *Pseudomonas aeruginosa*. **mBio**, v. 4, n. 6, p. 1–10, 2013.

LI, Y.; TIAN, C. F.; CHEN, W. F.; et al. High-Resolution Transcriptomic Analyses of *Sinorhizobium* sp. NGR234 Bacteroids in Determinate Nodules of *Vigna unguiculata* and Indeterminate Nodules of *Leucaena leucocephala*. **PLOS ONE**, v. 8, n. 8, p. 1–12, 2013.

LINDSTROM, K.; MARTINEZ-ROMERO, E. International Committee on Systematics of Prokaryotes; Subcommittee on the taxonomy of *Agrobacterium* and *Rhizobium*: Minutes of the meeting, 23-24 July 2006, Arhus, Denmark. **International Journal of Systematic and Evolutionary Microbiology**, v. 57, n. 6, p. 1365–1366, 2007. Disponível em: <<http://ijs.sgmjournals.org/content/journal/ijsem/10.1099/ijs.0.65255-0>>. .

LOH, J.; GARCIA, M. *NodV* and *NodW* , a Second Flavonoid Recognition System Regulating *nod* Gene Expression in *Bradyrhizobium japonicum*. **Journal of Bacteriology**, v. 179, n. 9, p. 3013–3020, 1997.

- LU, Z.; CAO, Y.; LONG, W.; et al. Isolation and characterization of an operon involved in sulfate and sul  $\phi$  te metabolism in *Sinorhizobium fredii*. **FEMS Microbiology Letters**, v. 282, p. 89–99, 2008.
- MARIE, C.; DEAKIN, W. J.; OJANEN-REUHS, T.; et al. TtsI , a Key Regulator of Rhizobium Species NGR234 Is Required for Type III-Dependent Protein Secretion and Synthesis of Rhamnose-Rich Polysaccharides. **Molecular Plant-Microbe Interactions**, v. 17, n. 9, p. 958–966, 2004.
- MARIE, C.; DEAKIN, W. J.; VIPREY, V.; et al. Characterization of Nops, nodulation outer proteins, secreted via the type III secretion system of NGR234. **Molecular plant-microbe interactions : MPMI**, v. 16, n. 9, p. 743–751, 2003.
- MILLER, W. G.; LEVEAU, J. H.; LINDOW, S. E. Improved gfp and inaZ broad-host-range promoter-probe vectors. **Molecular plant-microbe interactions : MPMI**, v. 13, n. 11, p. 1243–1250, 2000.
- MISRA, H. S.; RAJPUROHIT, Y. S.; KHAIRNAR, N. P. Review Pyrroloquinoline-quinone and its versatile roles in biological processes. **Journal of Biosciences**, v. 37, n. June, p. 313–325, 2012.
- NAVILLE, M.; GHUILLOT-GAUDEFFROY, A.; MARCHAIS, A.; GAUTHERET, D. A web tool for the prediction of Rho-independent transcription terminators. **RNA biology**, v. 8, n. February, p. 11–13, 2011.
- NIKOLSKAYA, A. N.; MULKIDJANIAN, A. Y.; BEECH, I. B.; GALPERIN, M. Y. MASE1 and MASE2: Two Novel Integral Membrane Sensory Domains. **Journal of Molecular Microbiology and Biotechnology**, v. 5, p. 11–16, 2003.
- OLDROYD, G. E. D.; DOWNIE, J. A. CALCIUM , KINASES AND NODULATION SIGNALLING IN LEGUMES. **Nature Reviews Molecular Cell Biology**, v. 5, n. July, p. 566–576, 2004.
- OLDROYD, G. E. D.; HARRISON, M. J.; UDVARDI, M. Peace talks and trade deals. Keys to long-term harmony in legume-microbe symbioses. **Plant physiology**, v. 137, n. 4, p. 1205–1210, 2005.
- PERRET, X.; FREIBERG, C.; ROSENTHAL, A.; BROUGHTON, W. J.; FELLAY, R. High-resolution transcriptional analysis of the symbiotic plasmid of *Rhizobium* sp . NGR234. **Molecular Microbiology**, v. 32, n. 2, p. 415–425, 1999.
- PERRET, X.; STAEHELIN, C.; BROUGHTON, W. J. Molecular basis of symbiotic promiscuity. **Microbiology and molecular biology reviews : MMBR**, v. 64, n. 1, p. 180–201, 2000.

PETERSEN, T. N.; BRUNAK, S.; HEIJNE, G. VON; NIELSEN, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. **Nature Methods**, v. 8, n. 10, p. 785–786, 2011. Nature Publishing Group. Disponível em: <<http://dx.doi.org/10.1038/nmeth.1701>>. .

PRICE, N. P.; RELIĆ, B.; TALMONT, F.; et al. Broad-host-range *Rhizobium* species strain NGR234 secretes a family of carbamoylated, and fucosylated, nodulation signals that are O-acetylated or sulphated. **Molecular microbiology**, v. 6, n. 23, p. 3575–3584, 1992.

PUEPPKE, S. G.; BROUGHTON, W. J. *Rhizobium* sp. Strain NGR234 and *R. fredii* USDA257 Share Exceptionally Broad, Nested Host Ranges. **Molecular Plant-Microbe Interactions**, v. 12, n. 4, p. 293–318, 1999.

QUÉRÉ, A. J. L. LE; DEAKIN, W. J.; SCHMEISSER, C.; et al. Structural characterization of a K-antigen capsular polysaccharide essential for normal symbiotic infection in *Rhizobium* sp. NGR234: Deletion of the rkpMNO locus prevents synthesis of 5,7-diacetamido-3,5,7,9-tetradecyloxy-non-2-ulosonic acid. **Journal of Biological Chemistry**, v. 281, n. 39, p. 28981–28992, 2006.

RAO, J. R.; COOPER, J. E. *Rhizobia* Catabolize nod Gene-Inducing Flavonoids via C-Ring Fission Mechanisms. **Journal of Bacteriology**, v. 176, n. 17, p. 5409–5413, 1994.

RELIC, B.; FELLAY, R.; LEWIN, A.; et al. nod genes and Nod factors of *Rhizobium* species NGR234. In: R. et al. PALACIOS (Ed.); **New Horizons in Nitrogen Fixation**. p.183–189, 1993. Kluwer Academic Publishers.

ROSALES-REYES, R.; AUBERT, D. F.; EL-HALFAWY, O. M.; VALVANO, M. A. The *suH* gene of *Burkholderia cenocepacia* is required for protein secretion, biofilm formation, motility and polymyxin B resistance. **Microbiology**, v. 158, p. 2315–2324, 2012.

ROSTAS, K.; KONDOROSI, E.; HORVATH, B.; SIMONCSITS, A.; KONDOROSI, A. Conservation of extended promoter regions of nodulation genes in *Rhizobium*. **Proceedings of the National Academy of Sciences of the United States of America**, v. 83, n. 6, p. 1757–1761, 1986.

SAAD, M. M.; KOBAYASHI, H.; MARIE, C.; et al. NopB, a Type III Secreted Protein of *Rhizobium* sp. Strain NGR234, Is Associated with Pilus-Like Surface Appendages. **Journal of Bacteriology**, v. 187, n. 3, p. 1173–1181, 2005.

SAMBROOK, J.; FRITSCH, E. F.; MANIATIS, T. **Molecular cloning a laboratory manual**. 2 ed. ed. New York: Cold Spring Harbor Laboratory Press, 1989.

SANGER, F.; NICKLEN, S.; COULSON, A R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the National Academy of Sciences of the United States of America**, v. 74, n. 12, p. 5463–7, 1977. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=431765&tool=pmcentrez&rendertype=abstract>>. .

SANJUAN, J.; GROB, P.; GOTTFERT, M.; HENNECKE, H.; STACEY, G. NodW is essential for full expression of the common nodulation genes in *Bradyrhizobium japonicum*. **Molecular Plant-Microbe Interactions**, v. 7, n. 3, 1994.

SARMA, A. D.; EMERICH, D. W. A comparative proteomic evaluation of culture grown vs nodule isolated *Bradyrhizobium japonicum*. **Proteomics**, v. 6, n. 10, p. 3008–3028, 2006.

SCHMEISSER, C.; LIESEGANG, H.; KRYSCIAK, D.; et al. *Rhizobium* sp. strain NGR234 possesses a remarkable number of secretion systems. **Applied and Environmental Microbiology**, v. 75, n. 12, p. 4035–4045, 2009.

SCHULTZ, J.; MILPETZ, F.; BORK, P.; PONTING, C. SMART , a simple modular architecture research tool : Identification of signaling domains. **Proc. Natl. Acad. Sci.**, v. 95, n. May, p. 5857–5864, 1998.

SENA, J. **Análise da função dos ativadores transcricionais SyrM1 e SyrM2 de *Rhizobium* sp. NGR234 no processo de nodulação.**, 2013. Universidade Federal do Paraná. Conselho Nacional de Desenvolvimento Científico e Tecnológico.

SIMON, R.; PRIEFER, U.; PUHLER, A. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. **Biotechnology**, v. 1, n. 9, p. 784–791, 1983.

SNOECK, C.; VERRETH, C.; HERNA, I.; MARTÍNEZ-ROMERO, E. Identification of a Third Sulfate Activation System in *Sinorhizobium* sp . Strain BR816 : the CysDN Sulfate Activation Complex. **Applied and Environmental Microbiology**, v. 69, n. 4, p. 2006–2014, 2003.

SONNHAMMER, E. L. L.; HEIJNE, G. VON; KROGH, A. A hidden Markov model for predicting transmembrane helices in protein sequences. **Proc. of Sixth Int. Conf. on Intelligent Systems for Molecular Biology**. p.175–182, 1998.

STACEY, G.; SANJUAN, J. SIGNAL EXCHANGE IN THE BRAD YRHIZOBIUM-SOYBEAN SYMBIOSIS. **Soil Biology and Biochemistry**, v. 21, n. 415, p. 473–483, 1995.

STAEHELIN, C.; KRISHNAN, H. B. Nodulation outer proteins : double-edged swords of symbiotic rhizobia. **Biochemical Journal**, v. 470, n. 3, p. 263–274, 2015.

STOCK, J. B.; NINFA, A. J.; STOCK, A. M. **Protein phosphorylation and regulation of adaptive responses in bacteria**. 1989.

STRANGE, K. K. LE. The Rhizobium Strain NGR234 nodD 1 Gene Product Responds to Activation by the Simple Phenolic Compounds Vanillin and Isovanillin Present in Wheat Seedling Extracts. **Molecular Plant-Microbe Interactions**, v. 3, n. 4, p. 214, 1990. Disponível em: <<http://www.apsnet.org/publications/mpmi/backissues/Documents/1990Abstracts/Microbe03-214.htm>> . .

STREIT, W. R.; SCHMITZ, R. A; PERRET, X.; et al. An Evolutionary Hot Spot : the pNGR234 b Replicon of Rhizobium sp . Strain NGR234. **Journal of Bacteriology**, v. 186, n. 2, p. 535–542, 2004.

TAYLOR, B. L.; ZHULIN, I. B. PAS Domains : Internal Sensors of Oxygen , Redox Potential , and Light. **Microbiology and Molecular Biology Reviews**, v. 63, n. 2, p. 479–506, 1999.

TERPOLILLIA, J. J.; SHYAM; MASAKAPALLIC, K.; et al. Lipogenesis and redox balance in nitrogen-fixing pea bacteroids. **Journal of Bacteriology**, v. 198, n. 20, p. 2864–2875, 2016.

VIPREY, V.; GRECO, A. DEL; GOLINOWSKI, W.; BROUGHTON, W. J.; PERRET, X. Symbiotic implications of type III protein secretion machinery in Rhizobium. **Molecular Microbiology**, v. 28, n. 6, p. 1381–1389, 1998.

WASSEM, R.; KOBAYASHI, H.; KAMBARA, K.; et al. TtsI regulates symbiotic genes in Rhizobium species NGR234 by binding to tts boxes. **Molecular Microbiology**, v. 68, n. 3, p. 736–748, 2008.

WEST, A. H.; STOCK, A. M. Histidine kinases and response regulator proteins in two-component signaling systems. **Trends in Biochemical Sciences**, v. 26, n. 6, p. 369–376, 2001.

WHITE, J.; PRELL, J.; JAMES, E. K.; POOLE, P. Nutrient sharing between symbionts. **Plant physiology**, v. 144, n. 2, p. 604–614, 2007.

WILLIAMS, R. H. N.; WHITWORTH, D. E. The genetic organisation of prokaryotic two-component system signalling pathways. **BMC Genomics**, v. 11, n. 1, p. 720, 2010. BioMed Central Ltd. Disponível em: <<http://www.biomedcentral.com/1471-2164/11/720>> . . d

## APÊNDICE 1 - GENES DIFERENCIALMENTE EXPRESSOS EM *nodVW* À COMPARAÇÃO DA ESTIRPE NGR234 SELVAGEM EM PRESENÇA DE FLAVONOIDES

Identificação do gene	Fold Change	Valor p	Cobertura NGR1+	Cobertura NGR2+	Cobertura VW1+	Cobertura VW2+
thuA	-54,9618787	2,69E-09	217,276336	112,271692	0,80534351	18,3599237
thuG	-38,7028238	6,21E-12	116,085612	51,9164269	0,9294964	7,62601918
thuF	-31,0320622	3,86E-09	140,465046	58,8151469	1,16170213	14,6633232
thuK	-30,6674705	1,22E-07	180,990058	74,2918129	1,56432749	17,6712476
NGR_c36670	-29,3307901	0,00010683	3,60251444	1,47901235	0	0,71975309
thuE	-28,5597517	0,00484667	631,532542	269,403642	5,8	72,5131433
NGR_b01430	-25,6100337	0,01601501	3,02622951	0,63879781	0	0,3136612
NGR_c27520	-19,09871	0,01105615	15,7290605	1,37701149	0	0,67011494
NGR_b23090	-18,9858085	1,97E-06	171,873913	72,8178442	2,83804348	25,4265399
NGR_c15620	-16,7556886	0,00041934	13,7691358	44,735443	6,39113924	5,90379747
NGR_c27640	-12,7324734	0,02336256	40,2197848	1,29895833	0	0,83463542
NGR_b03270	-12,4900427	1,55E-06	97,4815789	43,4246711	1,62828947	29,1330044
NGR_b03250	-11,9120279	0,00781519	311,066893	126,061375	6,11103565	79,5160158
NGR_b15200	-11,2080557	1,74E-05	4,97435897	1,24893162	0,23333333	0
NGR_c14940	-11,1811729	1,10E-07	20,0099564	8,81431624	0,11987179	11,0192308
thuR	-10,9395702	2,60E-08	13,2930233	4,71070736	0,15872093	3,78013566
NGR_b03210	-10,7552035	5,58E-06	94,4910345	52,1047126	2,63448276	32,6303448
NGR_c03050	-10,5857259	2,29E-05	41,3884458	24,9525745	0	23,4536585
NGR_b03280	-10,4733122	1,90E-08	37,0125	15,7042969	0,78203125	11,7914063
rhcT_a	-9,90021946	4,35E-06	9,52136752	2,6515873	0	3,62283272
nopJ	-9,7511473	0,00404098	136,111494	76,5814815	0,28659004	85,4403576
NGR_b01440	-9,73762767	4,98E-11	172,323288	40,3011416	6,98082192	12,9141553
NGR_c03510	-9,61142094	0,00089794	4,76788202	1,73623188	0	3,33594203
NGR_c03060	-9,23823099	0,00102715	11,0243771	15,2677249	0	19,7082892
NGR_b02810	-9,12053223	0,01031165	5,55238095	2,00343915	0	1,82222222
NGR_b13480	-9,04884068	0,01496031	5,16106195	0,51725664	0,16106195	0,31425762
NGR_b03160	-8,91982388	3,34E-05	12,0328947	3,7004386	0	6,02039474
nopT	-8,71780989	0,00021353	29,1801527	13,3767176	0	19,0712468
nopB	-8,4608744	0,01091781	476,983232	242,696364	0	295,874949
NGR_a00970	-8,3009758	0,00019734	117,420548	62,8019406	2,23082192	64,9819635
NGR_b03300	-8,23555153	4,67E-08	23,8186667	8,41217778	0,63093333	6,83386667
NGR_a03500	-8,00290266	1,11E-07	6,62109375	2,60078125	0,19479167	2,44704861
NGR_c01440	-7,81047916	6,72E-05	3,51143903	3,33671498	0,27101449	2,64628019

NGR_c19860	-7,76607179	9,11E-08	4,5848623	5,63162393	1,43846154	0,4982906
rhcR_a	-7,70087262	3,65E-05	20,8702541	9,26136024	0	14,1136024
NGR_b18310	-7,52551317	0,0006305	19,1285714	9,44900794	0	9,82261905
NGR_b03260	-7,44198317	0,00544086	687,793388	313,369972	28,214876	303,843251
NGR_b03150	-7,39940886	0,00459889	4,93582555	2,93182762	0,05669782	3,44683281
NGR_a00730	-7,16038714	0,01382057	210,96548	97,7394164	0,6735905	162,273195
NGR_c19020	-6,92007095	0,00400507	7,38570395	2,88678679	0,43288288	3,67522523
NGR_b03240	-6,9060435	8,70E-05	81,2356846	35,5849239	2,77759336	46,3041494
fixF	-6,71446738	0,03404154	96,8823821	59,0543011	0,09925558	95,3795699
nopC	-6,66635201	0,01094303	585,405387	239,959764	0,18888889	460,326599
NGR_c03800	-6,60339125	0,00234754	33,8943283	36,9121005	0,90547945	57,9746575
NGR_b03170	-6,46171356	0,0001099	3,83613445	1,26708683	0,15294118	1,79047619
NGR_b03290	-6,07256115	1,78E-06	26,8632653	11,1159864	2,15306122	10,8996599
NGR_a00710	-5,93113349	0,00282722	175,928387	81,2558065	1,24	148,677204
NGR_b06210	-5,87354381	0,03575099	4,89115646	3,43157596	0,12380952	4,32097506
NGR_c11250	-5,69651598	0,00298423	3,362255	1,57013201	0,23787129	1,31023102
NGR_a00740	-5,65183691	0,02323601	149,513421	69,7459649	0,50578947	142,020965
NGR_c20910	-5,57106957	6,62E-05	11,59	12,5033333	1,435	16,1470833
NGR_c08090	-5,53133724	0,0010096	9,87576459	7,00224719	0,42022472	10,9674157
NGR_c10360	-5,46792792	0,00281867	99,7268817	2,59333333	0,28769231	2,98974359
NGR_c25590	-5,45833182	0,00468383	3,57996357	1,96393443	0,30655738	2,62786885
NGR_c22810	-5,29701285	0,0007484	159,243958	222,960185	27,0333333	174,038889
NGR_a03370	-5,28740698	2,60E-07	13,159736	5,68539604	1,09034653	4,09051155
NGR_b10460	-5,1622563	0,02574305	5,4625	2,4265625	0	6,65833333
NGR_c29050	-5,12688454	7,75E-05	5,21131131	6,92222222	1,38518519	3,28364198
NGR_b03540	-5,11715358	0,01025453	102,397033	46,0418892	1,14480712	94,9019782
NGR_c20890	-5,05462632	0,04478354	5,25356245	3,73003663	0,20549451	7,31172161
NGR_b03130	-5,01114014	0,00017185	9,63755459	3,01477438	0,31790393	5,20189229
NGR_a00630	-5,00787307	0,01350485	160,092179	75,5470205	2,31005587	159,252886
NGR_b03220	-4,95712089	0,00036712	7,05632653	3,69954649	0,33170068	6,08825397
NGR_c18830	-4,93472359	0,01482961	6,4877192	1,40610329	0,26338028	3,07746479
NGR_c01730	-4,91388674	0,00159126	5,84768451	3,90295359	0,74303797	7,43839662
NGR_c30640	-4,84277792	0,00107272	10,4574987	2,72272727	0,425	5,90530303
NGR_c16270	-4,83879155	0,01217484	18,8273668	6,45151515	0,46363636	8,83002755
NGR_a02560	-4,71910931	0,00036788	15,3967787	5,16022409	0,82941176	11,8445378
NGR_a00980	-4,66043955	0,00595533	66,6498168	28,7512821	0,61648352	72,5844322
NGR_c24710	-4,63957084	0,00011781	1,57819235	1,9992016	0,52724551	3,76566866
NGR_b09430	-4,59993933	0,00090777	10,6753488	5,54744186	0,82883721	9,43984496
NGR_c16280	-4,58441629	3,05E-05	58,3349303	9,16869919	1,85609756	8,52926829
NGR_b04560	-4,57508656	0,0051966	3,08992248	1,28049096	0,14108527	1,81098191
NGR_b10150	-4,56689975	0,00081803	11,8634146	4,02195122	0,44390244	5,44254743

NGR_b03230	-4,54289773	7,73E-05	10,2598985	4,24192047	0,65583756	6,44137056
nodU	-4,5305655	0,03498734	8,41132976	3,66472868	0,06690519	9,86189624
pcaG	-4,437986	0,00499471	3,50439024	1,08674797	0,08878049	2,18585366
NGR_c12540	-4,36580016	2,70E-05	31,4719604	15,9741259	3,06223776	23,6701632
NGR_c35900	-4,35550222	2,54E-06	4,75004939	4,07878788	1,04615385	2,27575758
NGR_c06980	-4,31779648	0,0034929	8,33124843	8,26748971	1,20246914	10,0423868
NGR_c29310	-4,30715686	0,00424733	35,6076231	7,70142857	0,83857143	17,9028571
NGR_c10080	-4,26770204	0,00979639	90,0870504	50,887963	15,0138889	109,479167
NGR_c18090	-4,26323542	0,01716042	10,1245614	1,05087719	0,16403509	2,51374269
rhcC2_a	-4,25893093	0,03846677	8,7581761	4,4317217	0,22051887	8,96226415
NGR_c09250	-4,22489314	8,80E-05	7,65641049	4,86981982	0,88445946	3,31463964
NGR_c19140	-4,13687227	0,00027344	13,839944	20,4469512	3,22439024	26,1477642
rpIF	-4,12473961	0,00411388	16,6848295	12,0567416	0,85505618	19,4775281
NGR_c19060	-4,04625646	0,01893072	3,83722697	1,15192308	0,35961538	2,61602564
NGR_b02950	-3,98341557	0,0145642	12,6444444	1,85555556	0,57777778	4,33809524
noI	-3,97959679	0,00336751	18,9354223	7,70476839	0,68365123	21,1010899
NGR_c09510	-3,96370941	0,01085864	10,4761561	4,15602241	0,47142857	6,93753501
NGR_b20130	-3,9571872	0,00450576	3,48073394	1,95840979	0,33394495	1,57910296
NGR_c31740	-3,95228004	0,00607787	353,267081	180,016555	39,9080537	240,357271
atpF2	-3,9497119	0,00016472	4,45530879	1,65261438	0,45833333	2,52320261
NGR_b03140	-3,83635657	0,00376115	4,19393939	0,98186027	0,13787879	3,07020202
NGR_b20140	-3,81359085	0,0107611	2,99915433	1,45778717	0,1154334	3,71353066
NGR_c13890	-3,77030168	0,01110118	7,77701441	4,56766169	0,55820896	11,1616915
rmlB	-3,76575154	0,04643359	10,8892688	5,13518519	0,21310541	15,0747388
NGR_c03810	-3,7216804	0,02850433	3,45721125	3,72861842	0,50065789	7,81217105
noI	-3,6590402	0,00033256	8,28698413	4,00486772	0,90698413	4,60582011
NGR_a02850	-3,64306548	0,02499793	4,12066028	1,99957401	0,18753994	4,54696486
NGR_c27960	-3,61280858	0,00176478	4,29961263	3,8727381	0,68642857	6,01845238
NGR_c25500	-3,60529797	0,00344926	7,41769881	1,23389513	0,21011236	4,01067416
NGR_c19430	-3,55125622	0,01685852	18,930423	1,37234679	0,2632287	3,52780269
NGR_c28620	-3,55125622	0,01685852	2,46860217	1,22986799	0,29059406	4,49521452
NGR_c06390	-3,54855624	0,03297542	21,4836357	0,33956916	0,13605442	1,96553288
NGR_b09830	-3,52397813	0,0200988	9,15769231	3,73044872	0,35	9,01025641
NGR_b15970	-3,49758848	0,00704621	8,65	2,3295	0,24266667	6,73055556
attA1	-3,49665014	0,04000001	3,86684642	0,33446878	0,04549878	0,70900243
rmlD	-3,48875521	0,00487972	6,38417508	2,04006734	0,18888889	5,60157127
cycG	-3,48552407	0,02708332	3,61031414	2,47256831	0,32360656	7,88131148
rkpT2	-3,39573543	0,00026531	8,8957529	1,98854569	0,56216216	4,75714286
NGR_a01720	-3,39371198	0,00070974	56,5963382	19,9831449	7,08610662	20,9854066
NGR_b05600	-3,39168066	0,03653799	3,17454545	2,12545455	0,66181818	1,04363636
NGR_b15190	-3,31638855	0,00148368	29,05	13,410779	2,9076087	23,9630435

NGR_c11310	-3,30520453	0,04271123	2,20256932	2,27468354	0,74303797	6,45443038
NGR_b07440	-3,27845815	0,04210284	6,86066066	3,67837838	0,72132132	6,79279279
NGR_b00600	-3,25860603	0,00102118	12,7154639	7,20463918	1,68865979	5,43780069
NGR_c35250	-3,23169107	0,01806647	5,94162522	6,47636004	1,77418398	11,5277943
NGR_c12170	-3,20918277	0,01087219	9,34771079	7,88690096	0,95591054	23,7537806
NGR_b03520	-3,20089858	0,02671182	111,561111	49,2486111	4,01805556	135,066204
NGR_c08110	-3,18985914	0,01673146	5,38940288	4,69641026	0,49282051	14,542906
NGR_c12460	-3,1819299	0,00395083	61,7604909	26,5567204	8,09758065	64,5397849
NGR_a03310	-3,12244642	0,01266296	13,155144	5,6090535	0,80802469	14,4353909
NGR_a03490	-3,05745282	0,016286	8,28759058	3,10253623	0,36277174	10,1701087
exsl	-3,05469754	0,01632441	17,9309524	5,21369048	0,86666667	12,6571429
NGR_b09740	-3,05243172	0,00177139	6,11521739	3,90398551	1,18695652	3,53297101
NGR_c05580	-3,04832161	0,00724074	4,27552964	3,02614622	0,4866171	11,4095415
NGR_c23310	-3,03427328	0,00215943	14,1309293	10,516129	3,1	24,5166667
NGR_a02940	-3,0161157	0,04006401	4,63095975	1,5869969	0,11578947	6,54602683
NGR_c17950	-3,01284703	0,00530579	5,38965517	2,40977011	0,80603448	6,23850575
wbgA	-2,96887565	0,01520719	8,3552381	2,88071429	0,3110989	10,2361905
NGR_c10790	-2,96602574	0,0456055	32,4713095	6,41511111	1,28133333	23,5026667
NGR_b14420	-2,93026077	0,00151808	19,2807229	10,1599398	2,45542169	11,487751
NGR_b14010	-2,91002974	0,00312761	5,79224806	3,46059432	1,3255814	2,60568475
NGR_c19290	-2,88382263	0,0164537	8,29678116	2,74234234	0,77567568	6,00472973
NGR_c31090	-2,85704504	4,77E-05	4,31976182	7,22572464	3,43858696	9,23731884
NGR_c19130	-2,82584399	0,00749396	29,5255794	5,16368159	1,11641791	10,5116915
NGR_a03270	-2,82125569	0,00784919	4,95294118	2,28970588	0,68921569	5,4453159
NGR_c30730	-2,82123361	0,01727696	8,37053311	2,17668712	0,85092025	4,70777096
NGR_b07450	-2,81668482	0,02582192	11,6173913	5,91123188	1,18695652	12,7094203
rpmJ	-2,8108728	0,02654645	19,3945175	7,34126984	1,78095238	23,5880952
NGR_c25070	-2,79273214	0,04908366	2,36893204	2,55177994	0,55728155	6,41245955
NGR_c32390	-2,78531986	0,03501629	35,1200011	13,9348837	1,82209302	47,8677003
NGR_c07670	-2,77323518	0,00802991	26,4753734	35,92625	7,0125	100,70625
NGR_b14410	-2,7538898	0,02832204	8,64375	3,72591146	0,853125	10,0963542
NGR_c12160	-2,73929925	0,02427641	22,9686593	8,7811463	1,81221719	24,6105581
NGR_c01610	-2,73453072	0,01645027	3,07932504	1,08467153	0,40948905	2,37542579
actP	-2,69270198	0,00055095	5,81260404	1,61460563	0,76575505	3,65727309
noeJ	-2,69164124	0,01402123	11,7412606	4,07472385	0,85360624	15,9692008
speB	-2,68824448	0,00495977	7,71818583	0,67414773	0,38664773	1,31070076
NGR_b10020	-2,67863815	0,03145588	3,43361345	0,49117647	0,67226891	1,01036415
NGR_c31470	-2,65470915	0,00120736	7,07660526	6,32262443	2,58552036	17,6665158
rkpT1	-2,63086952	0,03004693	3,28715953	1,70077821	0,42490272	3,47613489
NGR_b20480	-2,61191866	0,01222215	15,8433898	6,95480226	1,99864407	17,8723164
arcA1	-2,60124008	0,00551913	9,41808883	2,95963203	0,98831169	9,46006494

NGR_b15640	-2,59578329	0,02483569	11,5389423	4,57347756	0,7875	14,9071314
trnT2	-2,59225506	0,00766656	599,357895	143,76	53,108	282,172
NGR_c11280	-2,57061776	0,01303797	13,0005861	3,58518519	1,47275986	8,26750299
NGR_c19990	-2,52844378	0,0136349	8,14172535	7,02066958	2,10960699	16,7740902
NGR_b08050	-2,50830806	0,04030799	8,69403974	4,12328918	1,7589404	8,10485651
rpsO	-2,50430005	0,03715223	8,84337778	13,2088889	2,93777778	40,1455556
NGR_b20950	-2,49975401	0,02651926	7,99796954	2,21878173	0,66497462	8,99898477
pilA1	-2,49733367	0,0382706	78,2502347	122,067213	38,2606557	288,471585
NGR_b10850	-2,47161708	0,00212396	33,6990291	8,85744337	4,45048544	19,3122977
NGR_c00160	-2,46885203	0,02250451	59,4435777	55,3115183	18,8554974	89,9598604
NGR_c01060	-2,46371221	0,04609617	2,00547945	0,95730594	0,38424658	4,32420091
NGR_b22330	-2,44957233	0,03045038	5,87887324	2,18333333	0,76901408	3,80938967
NGR_c12450	-2,44129125	0,04378371	70,702574	3,19018265	0,64041096	13,4719178
NGR_c16670	-2,43600622	0,01286316	31,0650884	9,32168022	3,14634146	32,1802168
nuoE1	-2,41222926	0,04581826	4,4728123	2,46335283	0,68654971	6,27641326
NGR_c27610	-2,40128164	0,00312959	11,884456	8,4010101	3,43939394	20,2602694
NGR_c25510	-2,39085527	0,00725773	6,72651293	5,65555556	1,90465116	8,09431525
NGR_a02280	-2,38777536	0,02843564	2,92856341	1,36717172	0,5526936	3,13641975
NGR_c30010	-2,36915131	0,016173	23,9966559	7,27810545	5,33378016	4,10160858
NGR_c07840	-2,35713713	0,00296255	5,72591248	4,33048433	1,35854701	5,35470085
NGR_b23070	-2,32465364	0,00165884	91,6033613	34,2680672	17,4	47,0663866
NGR_b08360	-2,30723004	0,03262676	8,44666667	2,09652778	0,97	4,39361111
NGR_c18850	-2,29009326	0,00380507	3,04109925	1,99644128	0,73202847	2,8688019
NGR_c28870	-2,27277391	0,03151128	14,0503054	21,3410761	7,26614173	44,7413386
NGR_c26550	-2,27170099	0,04150762	47,3140453	7,14413146	3,26056338	8,16353678
stbC	-2,25725489	0,03973877	15,7976744	8,24418605	2,39186047	18,0100775
NGR_b00480	-2,24158812	0,01191959	14,3794393	2,27211838	2,20934579	7,81277259
NGR_c10810	-2,21703538	0,01182223	6,61557565	7,96071429	3,33928571	19,4297619
NGR_b10110	-2,21519184	0,02194285	12,0728682	3,30297158	1,3255814	9,91007752
NGR_c20650	-2,20060579	0,01774373	4,48658673	2,77664399	1,4170068	3,53582766
NGR_c31840	-2,18369683	0,03404111	63,7163558	10,2769608	4,63308824	35,0693627
NGR_b05480	-2,17086296	0,03652899	10,9965753	3,60308219	1,53287671	11,2687215
NGR_b00580	-2,16883676	0,02356473	7,25233645	2,18504673	1,19065421	6,12834891
NGR_c01960	-2,15478457	0,04863191	4,92656638	2,1744186	0,82751938	6,04186047
cyaF	-2,15374857	0,0218032	2,35187211	1,52746331	0,6591195	3,59654088
NGR_b03200	-2,14068599	0,03107926	48,0619048	15,7013889	4,85238095	51,6440476
NGR_b12750	-2,13351315	0,03144006	6,26125	1,57770833	0,81875	5,48020833
NGR_b10740	-2,12394715	0,01955955	34,6891429	12,609619	5,61485714	36,0464762
paaB	-2,10801611	0,0331699	5,36944843	2,70807018	2,07789474	9,71298246
NGR_b01920	-2,10133938	0,04855838	13,8188235	6,85647059	3,38117647	15,4290196
NGR_c13990	-2,09892057	0,01398796	3,55422136	3,1540107	1,21390374	6,90784314

NGR_c22780	-2,09243104	0,02909796	16,3883249	0,92153846	0,86307692	0,89692308
NGR_b16600	-2,05469881	0,01268885	5,33762376	2,2099835	1,53069307	4,19224422
NGR_c08360	-2,02521017	0,0373909	4,9838318	5,17460317	2,13714286	7,26126984
NGR_b09400	-2,01898231	0,00672364	5,1858209	1,52350746	1,12686567	2,84427861
NGR_c22650	-2,00458483	0,0150458	19,8282886	6,75807018	4,44	18,1482456
NGR_c28270	-2,00282077	0,01217979	354,008738	4,02330383	2,50530973	6,23215339
NGR_b05060	2,00883048	0,0291338	9,28409091	4,59160354	6,05151515	47,7325758
NGR_c17800	2,01626758	0,04036164	5,15941221	2,59134438	3,99779006	27,4071823
NGR_b21980	2,04194196	0,02842914	7,18043478	3,17047101	3,77101449	50,1666667
NGR_c01600	2,04465772	0,03618535	2,99927007	0,5935446	0,59260563	6,05457746
NGR_c05080	2,04927403	0,02779135	1,55208501	1,21701977	1,88961864	10,9273305
NGR_c09660	2,05089544	0,04506584	5,67671185	0,34172779	0,66552567	7,30268949
sufE	2,05310724	0,03559823	6,89329729	1,66685934	2,87052023	34,6132948
NGR_c04780	2,05958827	0,04792691	11,2844797	0,40098401	0,27601476	3,69138991
NGR_c24850	2,06409669	0,0080315	1,95059264	1,54438306	2,46243094	23,5567219
NGR_c36030	2,08071129	0,02429404	5,60648227	1,79326241	1,81808511	26,8641844
NGR_c03120	2,08199075	0,00882446	7,74730044	0,70078329	0,68355091	10,1308964
ilvH	2,08728118	0,04344907	0,94841475	0,82425829	0,69895288	8,51710297
NGR_b21970	2,09063725	0,01607396	2,87785548	1,47105672	1,40839161	22,2428127
NGR_b13030	2,09494781	0,03978919	0,41403509	0,29298246	0,36491228	3,88897243
NGR_c13020	2,11038484	0,02388273	3,68049887	1,21201814	1,99183673	17,3104308
rpoD	2,11658862	0,03653278	10,537698	6,13192214	8,56919708	102,730122
hisI	2,11846966	0,02356187	1,52380089	1,73921569	2,42418301	21,7413943
NGR_c13060	2,12646334	0,01464785	3,89681702	2,42704403	2,29339623	34,7726415
groEL_1_b	2,13810545	0,01036619	11,5524862	6,68468386	8,54917127	100,450031
NGR_b01030	2,14287343	0,01064686	4,51524164	2,41530359	2,84089219	35,2672862
dme	2,14944312	0,02564173	1,93922815	0,67034155	1,31245136	11,981323
amaB	2,17243572	0,0109471	6,05460899	17,7469225	18,1098321	181,360911
NGR_c17830	2,20917259	0,0322361	6,1592233	11,2016181	16,9378641	171,412298
NGR_c27850	2,21996947	0,02063921	0,93537362	0,90408006	1,83429561	13,1268283
uvrC	2,24700049	0,04367457	1,62876682	0,42443126	0,75682493	8,17794263
hemH	2,26271816	0,0433691	1,61644472	1,62803468	3,70375723	21,6561657
NGR_c24820	2,26286927	0,00779725	5,60628561	9,25498442	14,3359813	121,38271
NGR_c30590	2,2628736	0,01250886	13,3698378	2,47908745	3,57490494	31,4631179
NGR_c23600	2,26759126	0,04538371	4,92967894	0,4691906	1,24099217	11,373107
dht	2,28691903	0,00197553	0,88927699	4,36783505	6,85360825	51,4766323
groEL_2_b	2,31577844	0,00825565	11,0088398	6,55988336	9,50128913	93,6349294
fusA_c	2,33168483	0,00740925	4,53373637	3,599	4,96528571	53,263
NGR_c24860	2,34959404	0,00658496	2,14225989	1,22874317	1,42721311	19,1065574
trnR2	2,361654	0,01192986	1869,25195	27,2272727	40,0714286	383,871429
NGR_b13240	2,37383385	0,01967756	2,23821429	0,79717262	1,75428571	11,688869

NGR_c34500	2,38531105	0,01631114	7,49019511	7,42011719	11,8277344	116,482552
NGR_c33710	2,39382084	0,00684211	0,56105137	0,61143695	0,71290323	12,3608016
NGR_b12880	2,4190829	0,00438206	1,95292621	0,95163274	1,50025445	11,3506361
NGR_b02560	2,43206015	0,00658276	6,78830409	2,49917154	4,16081871	44,1522417
NGR_c01840	2,45088837	0,00554876	2,14238507	1,79722222	2,64537037	36,5564815
NGR_c25760	2,45631233	0,00934067	3,54844909	1,83853779	3,14758364	38,8480793
NGR_c24890	2,45754873	0,00619978	7,05293596	3,47089627	7,82175227	48,6636455
NGR_b21920	2,46973917	0,01006072	1,32277778	0,56708333	1,04111111	9,60740741
NGR_b19300	2,47547254	0,01828957	4,63261538	1,79323077	2,76553846	40,5222564
NGR_b12780	2,48529388	0,01151464	0,8460199	0,31502902	0,70621891	4,97815091
NGR_c24470	2,50184564	0,00747923	2,57878788	0,81022639	1,63138173	16,491413
NGR_b11530	2,52697582	0,00598434	4,04954683	1,70443102	4,07854985	29,5205438
repB2	2,53156713	0,03841599	0,57627119	0,38286252	0,41129944	4,4220339
NGR_b18500	2,53567277	0,03445441	0,48615385	0,52128205	0,49	5,25307692
NGR_b13100	2,54031142	0,00877223	3,04839506	0,88893004	2,42567901	20,7193416
cpaE1	2,54140291	0,04133969	0,8397711	0,3211838	0,30584112	5,09501558
NGR_c25450	2,55349547	0,02128077	2,64916865	0,62175732	1,17364017	16,1539749
NGR_c36610	2,56440912	0,03953876	14,9685085	0,27860465	0,9627907	8,92217054
NGR_a03190	2,56682456	0,0039693	1,87129187	1,8238437	2,66937799	17,5805423
deoA	2,57113375	0,00204957	1,86885409	1,06097191	1,53348519	15,7197418
NGR_c33110	2,5716637	0,00758288	0,81963595	0,52485207	0,66390533	7,21282051
NGR_c28570	2,57634292	0,00878372	5,37913271	0,70256898	1,15899144	8,28024104
NGR_c28430	2,60484686	0,01342939	5,93089058	0,36208981	0,43601036	7,67443869
NGR_c24210	2,61024505	0,04638526	1,58400127	0,92246094	1,0328125	18,2583333
NGR_b06770	2,61899	0,03650799	1,05022026	0,38436123	0,6814978	3,92408223
recN	2,61983961	0,0234514	1,32193238	0,33787336	0,77078853	5,54396655
NGR_c19370	2,64720528	0,01755531	2,14278812	0,73928889	2,3064	17,6499556
NGR_c13480	2,66678634	0,0388634	4,06207665	0,15763889	0,39914216	2,52332516
NGR_b22310	2,68599555	0,0296049	0,70446927	0,32653631	0,81340782	4,78500931
NGR_b00350	2,68973092	0,03177403	15,766998	5,52548045	15,0270378	73,174619
NGR_c35590	2,75910057	0,01016633	2,89947739	1,84876543	6,18055556	44,4709877
guaA	2,76099984	0,04218169	1,13237102	0,39577735	1,01996161	5,08055022
NGR_c33970	2,76751659	0,04999461	1,80447651	0,53267974	2,15931373	17,1387255
NGR_c35390	2,77366513	0,0176697	32,462682	2,66711712	6,3527027	54,0112613
NGR_c24870	2,78765622	0,0169274	0,90760764	0,4659887	0,64271186	8,95536723
acnA	2,80426698	0,00461804	0,50783201	0,15480515	0,31396141	3,34025728
NGR_b22410	2,8055442	0,04190091	0,66212425	0,21389446	0,4741483	3,83767535
NGR_c09330	2,84212733	0,03072019	0,89119339	0,22371615	0,37394958	4,36778711
hslU	2,84886231	0,00301855	3,21879496	1,24904215	2,70114943	22,3395402
NGR_c30600	2,84951043	0,00668917	6,82640764	8,78283582	18,5840796	137,764262
NGR_c03790	2,86093535	0,00541954	44,6974149	0,41163435	0,73961219	9,69953832

nuoL	2,87188369	0,00554525	1,60924534	0,3435589	0,94360902	6,03298246
NGR_b21720	2,89029964	0,00124101	7,89651475	4,07873101	6,93619303	69,4638963
NGR_c06080	2,91063204	0,00799288	0,32972973	0,53963964	0,75810811	7,58040541
NGR_b00210	2,93181667	0,00311576	0,2591716	0,28821499	0,37692308	5,51193294
NGR_c04050	2,94916389	0,03542858	4,05516534	0,43694581	1,15788177	6,35172414
NGR_c30580	2,97091809	0,00364652	4,80773781	0,85327635	2,76153846	26,2954416
NGR_c30290	2,97216886	0,01819096	0,82090039	0,34277539	0,7223176	3,96108727
NGR_c30520	2,98563985	0,02513798	1,45685689	0,20690846	0,74015544	4,12728843
NGR_c26060	3,00195748	0,0169086	1,90894083	0,35167206	1,01035599	6,63225458
NGR_c33060	3,00497251	0,00499895	9,74840632	8,43831325	17,4036145	174,147871
exsH	3,00674114	0,0227418	1,25021459	0,22904149	0,74206009	5,0786123
NGR_b14800	3,01849539	0,04905042	1,39189189	0,15797297	0,89513514	5,07351351
NGR_c35620	3,02977353	0,04691867	1,39140741	0,21017544	1,11	9,96842105
NGR_c32980	3,04316306	0,02947374	0,34851569	0,13430493	0,55089686	3,36569507
NGR_c14280	3,05403635	0,00587506	4,675	1,23946078	3,63235294	33,8855392
atpD	3,060864	0,01711717	4,17594544	2,09064559	4,22450593	47,9514493
NGR_b20410	3,0753247	0,03882219	0,07132353	0,28651961	0,85588235	3,27401961
IrnK1	3,1043377	0,00468522	2991,38289	9,45789474	38,3842105	275,390789
cysN	3,12310942	0,00162824	2,38300725	1,48436874	2,98657315	36,1285237
cobC	3,12359071	0,04047846	0,53156317	0,11956088	0,40748503	3,21337325
rho	3,13028981	0,00211648	2,1693863	0,94083728	4,20379147	17,0311216
rpsL	3,14982655	0,00016911	23,6446368	7,79731183	16,8072581	125,683602
NGR_c13300	3,16913747	0,04820611	1,83044077	0,29935721	1,18484848	5,70027548
NGR_b14790	3,22833103	0,01182184	0,35374732	0,1251606	0,27280514	3,68536759
NGR_b23030	3,24813983	0,00170888	2,50091896	0,87539683	2,67502089	17,6068505
NGR_b16040	3,2877327	0,01896746	0,32119205	0,31975717	0,2410596	3,45640177
ilvC	3,30392423	0,00825147	2,04241564	1,05029412	2,33147059	25,1413725
cysG	3,31060847	0,00113077	2,22182159	1,19295003	1,79301848	24,9993155
NGR_c13070	3,36744305	0,00368808	5,90122857	2,26123059	4,02903226	59,5091995
NGR_c13560	3,39659516	0,02517395	0,71128785	0,30412979	0,95619469	8,87330383
trnM1	3,43836451	0,00141027	0	5,51710526	14,7631579	75,1763158
NGR_b20560	3,4815714	0,01906226	0,69032258	0,18854839	0,78645161	4,91709677
NGR_c01630	3,48959554	0,00464521	3,46679773	0,49847095	1,21284404	12,0298165
NGR_b00500	3,569833	0,04599728	1,04513889	0,06765046	0,70763889	6,63946759
NGR_b22390	3,57844839	0,00391937	0,98038147	0,29082652	0,75367847	6,43251589
NGR_b02610	3,59575862	0,00010634	0,89108434	0,70012048	1,20096386	12,157992
NGR_b11810	3,6794226	0,00959905	0,24871795	0,12489316	0,30320513	3,64316239
NGR_b22340	3,72870846	0,00483451	1,59508197	0,43743169	1,7147541	10,1998634
NGR_c27160	3,72898391	0,03555351	5,4198748	0,20517413	0,34268657	6,0159204
NGR_c28990	3,75924924	0,04108552	12,3197531	0,03048346	0,08564885	1,47547074
aspB	3,79328025	0,00184417	0,49244966	0,28151986	0,53290155	4,15310881

cysD	3,85360814	0,02726514	20,8068749	0,71844864	2,04842767	30,7421384
adhC4	3,87287797	0,00120295	0,62918919	0,21063063	0,64918919	6,6836036
NGR_c25410	3,99932087	6,19E-05	1,61643794	2,20425384	6,26485356	38,153417
NGR_c34860	4,03937796	0,00787602	2,15166094	0,56244131	4,47746479	27,7779343
NGR_c03980	4,07088364	0,03911688	13,5704301	0,08834808	0,59070796	5,5219764
NGR_b19290	4,11139876	0,02763541	1,10331754	0,09233807	1,15545024	7,53538705
NGR_c06040	4,25403101	0,00067624	0,42285527	0,31197917	0,80351563	9,3328125
rhcJ	4,33990565	0,03635892	0,14477612	0,072699	0,1358209	3,09900498
cysK1	4,34775347	9,58E-06	5,2288653	4,18819255	9,74359673	129,714623
lysC	4,55800618	7,25E-05	0,96904696	0,64690196	1,78447059	8,42415686
NGR_c22830	4,63148923	0,00145685	6,36823928	0,27990654	1,60934579	11,2331776
trnH	4,76145263	0,00316951	1,9012987	149,361039	598,157143	4246,05714
cyoB	4,83171518	2,57E-06	1,55598802	0,76973553	2,46197605	18,9799401
NGR_c02600	4,86403967	0,0382456	1,01460894	0,15478036	0,43488372	3,0878553
NGR_c13080	5,07509408	0,00118045	8,23099631	2,45883534	6,09879518	98,4164659
hisF	5,0813651	0,00010364	2,04387676	0,38545689	1,37181467	6,0958816
NGR_c36920	5,15335276	0,00790403	2,23619257	0,03781566	1,01534091	3,75246212
NGR_b20570	5,23834863	0,00589197	0,43294574	0,0377584	0,46550388	3,01815245
NGR_c11730	5,51044505	0,00029858	8,78602041	6,81666667	19,5619898	223,881803
cysH	5,54404033	0,00023451	5,64589334	0,31693122	3,03015873	15,9763228
NGR_c06070	5,56762053	7,47E-05	0,76269816	0,15478036	1,11136951	12,6116279
trnF	5,64287319	3,10E-05	35989,3613	63,8407895	283,452632	1880,94211
NGR_c31650	5,88163757	0,00282086	32,4915095	37,3168067	178,357143	1208,34538
trnA4	6,03195739	1,19E-07	162643,797	7,88157895	28,05	219,392105
xyfF	6,09608945	2,24E-07	1,37865597	1,05456292	3,61066282	50,8023055
NGR_c15990	6,17828373	0,02053222	0	0,32511111	0,374	4,49955556
NGR_c11890	6,27868366	1,87E-06	10,4817478	2,31369048	10,3005102	96,1278912
trnG2	6,36584344	0,00026938	7828,232	72,0418919	493,528378	3084,38514
NGR_b06430	6,37737866	1,45E-06	1,49201681	0,34702381	1,87352941	11,1338235
NGR_c02130	6,72913994	4,14E-07	9,15784229	5,87323944	45,2661972	415,362441
NGR_b09690	6,86086531	0,01003857	0,16440678	0,16511299	0,9559322	5,69519774
NGR_c02140	8,3685269	0,01425135	0,5183223	0	2,28979592	15,0666667
NGR_c26450	8,64333726	0,00029058	1,91282884	0,48848114	5,34862385	24,8328236
fabG3	8,97783103	0,00060892	0,47434343	0,27275132	0,92142857	13,3927249
NGR_b19750	9,23190422	0,04224055	0,56833333	0	1,06083333	3,67611111
groEL_1_c	9,66515452	4,90E-07	47,5977617	4,86098901	26,2093407	247,078144
NGR_c06060	9,72523321	7,98E-10	0,8995866	0,50091683	2,33398821	21,4755075
NGR_c36860	16,1638456	8,13E-07	0,63631436	0,09462875	0,97488152	8,31153239
NGR_c06030	18,8476621	2,64E-14	0,50325327	0,39636063	6,81612903	72,6238213

## APÊNDICE 2 - GENES DIFERENCIALMENTE EXPRESSOS EM *nodVW* À COMPARAÇÃO DA ESTIRPE NGR234 SELVAGEM EM AUSÊNCIA DE FLAVONOIDES

Identificação do gene	Fold Change	Valor p	Cobertura NGR1-	Cobertura NGR2-	Cobertura VW1-	Cobertura VW2-
NGR_c30010	-57,1299576	0,00047809	77,7912869	50,9168007	0,58873995	4,08699732
NGR_b07280	-48,2643872	0,00055183	4,75397727	6,11458333	0	0,79715909
trnL4	-28,8406742	0,00370088	12342,5006	13644,3558	95,2627907	2802,47093
NGR_c15620	-26,6599163	0,00010286	170,458861	116,182278	5,53797468	1,55063291
thuA	-26,0803871	0,0075345	133,216603	113,889949	0,15903308	30,031743
thuE	-26,0119482	0,0027417	500,78266	478,59323	0,14845606	116,739549
thuG	-23,5092993	0,00774072	76,2868705	66,8645084	0,49016787	16,2193046
thuK	-23,4943325	0,00980297	115,151754	144,032407	0,26013645	29,8076998
trnE2	-21,355351	0,01210679	225,036	202,048	1,66666667	62,0666667
thuF	-20,2019725	0,01404347	90,9930091	97,6846505	0,2070922	25,8316109
NGR_c18420	-17,3571387	0,00325046	9,60461538	5,17512821	0	3,41487179
NGR_c31090	-16,8762849	0,01727166	27,0032609	16,2827899	0,11322464	9,91675725
NGR_b08470	-16,6301056	0,00477647	4,23181818	6,42340067	0	2,62811448
NGR_c35900	-16,4823614	0,00205253	7,24615385	6,5037296	0,14568765	1,57051282
NGR_c35570	-16,1468644	0,01127593	520,444581	567,959606	10,6428571	199,081363
NGR_b23090	-16,1024769	0,02047756	119,066576	121,228261	0,4115942	45,2051178
NGR_b12440	-15,9258021	0,00357951	30,7227437	13,18572	0,73112716	3,39366225
NGR_c36670	-14,9997308	0,00241855	8,73611111	8,89104938	0,38580247	2,05524691
paaB1	-14,3751141	0,00744867	3,54715909	2,87089646	0	1,86994949
hmgA	-14,3281651	0,0196936	16,2905286	8,99357562	0,15007342	6,28784875
NGR_c03520	-14,1912113	0,02386262	3,15	3,51754386	0	1,43274854
NGR_b23070	-12,8667254	0,0122809	202,721849	209,642857	8,31848739	45,179972
trnP2	-11,9194916	0,040812	1062,53182	745,6	12,987013	448,636364
NGR_b03210	-11,4866299	0,02696178	74,9162069	64,0977011	1,0445977	27,9913793
NGR_c26140	-11,4715974	0,02325394	13,5994318	9,70612374	0,15782828	5,58604798
NGR_c10080	-11,2202195	0,02161279	251,802083	172,183333	12,5912037	69,6118056
NGR_b23360	-10,6902369	0,01745447	11,0917808	16,0821918	0	7,44383562
NGR_c26200	-10,3504219	0,04460996	39,4003597	26,8982014	0,49016787	27,0227818
NGR_b14060	-10,237945	0,02386885	14,0093617	20,6074468	0,35460993	7,67439716
NGR_c26120	-10,0231403	0,01755245	13,0409548	8,54589615	0,580067	6,69773869
NGR_c30640	-9,65164188	0,01223836	12,075	9,90284091	0,5375	4,97329545
paaJ	-9,47132826	0,03099071	19,9086207	8,44099617	0,1197318	9,34789272
NGR_c03620	-9,4116599	0,02689174	32,3116279	13,5624031	0,73333333	15,0599483
paaB	-9,24468618	0,02872942	40,8236842	22,764386	0,99578947	20,8631579
trnG1	-9,24462216	0,01012536	23372,34	17999,834	384,166667	13448,406
sucB1	-8,67289508	0,0085266	4,68402439	6,56199187	0,51227642	1,3598374

NGR_c17210	-8,33422071	0,03881796	46,2151685	24,8964419	1,17041199	20,7872659
NGR_b08480	-8,14578405	0,0006421	27,0689189	13,2207207	1,48378378	2,84391892
NGR_c30620	-8,12401877	0,03481951	8,84444444	8,37654321	0,19290123	7,87685185
NGR_b02380	-8,05096777	0,04708817	13,0762082	17,0935564	0,60495663	7,67744734
NGR_c22320	-7,9755348	0,01913036	7,27358491	7,5091195	0,39308176	5,51163522
thuR	-7,9470694	0,01689919	14,2011628	8,10537791	0,51724806	5,2502907
NGR_b08170	-7,88499646	0,03098139	6,70824742	7,06013746	0	4,48487973
NGR_c18880	-7,87184289	0,00224325	3,04580153	2,72455471	0,31806616	0,6913486
actP	-7,85657648	0,01495611	25,4548751	15,4718589	1,63166865	5,23196591
NGR_c20580	-7,49253217	0,00960849	4,36714286	2,01880952	0,14880952	2,08654762
NGR_b14420	-7,05899444	0,01821663	31,5885542	34,7720884	3,60240964	6,71726908
NGR_c00500	-6,98329043	0,04939157	437,701744	209,510271	28,3244186	200,905039
NGR_c30390	-6,9065756	0,04366204	6,13846154	2,52717949	0	4,08333333
NGR_b07290	-6,80308829	0,03237346	4,25	7,93243243	0,18768769	2,57507508
NGR_b02370	-6,73409514	0,04458626	5,69387755	5,49064626	0	6,83231293
trnS1	-6,54090359	0,00646289	105763,21	88969,175	2969,18778	89308,4644
NGR_b03130	-6,48321222	0,03533616	4,72991266	3,63136827	0	5,28537118
NGR_b10680	-6,43706151	0,01079753	5,0546875	3,05729167	0,21701389	2,09947917
NGR_c19060	-6,3931178	0,0328731	7,40192308	5,28429487	0,90961538	3,14102564
NGR_b05580	-6,33774651	0,01227614	7,00833333	4,07638889	0,65694444	1,70138889
NGR_b14490	-6,15636716	0,03062675	4,284375	5,09548611	0,49270833	1,48871528
NGR_b04820	-6,03915108	0,03874688	8,02641509	12,9213836	0,89245283	5,00786164
trnL2	-6,01384325	0,02182835	55,523494	28,2024096	3,76506024	22,1385542
trnV2	-6,00769884	0,03260461	6411,132	6697,97333	745	3281,36667
NGR_c03560	-5,82228365	0,04522445	33,9180723	18,0012048	2,32690763	16,815261
NGR_c20830	-5,7720836	0,01870131	8,26832061	4,47175573	0,67913486	4,61564885
NGR_b07420	-5,6639137	0,04009235	8,58095238	15,2702822	1,05167549	5,96631393
NGR_b21190	-5,60233122	0,03023245	14,0939597	18,0564877	1,29619687	5,99395973
NGR_b11430	-5,55376708	0,04561072	11,562931	16,8678161	0,94655172	9,59913793
NGR_c12680	-5,4968183	0,00975752	24,0395833	6,90729167	1,66967593	5,50046296
NGR_b08960	-5,45469382	0,02329871	4,71771255	4,65401484	0,55168691	1,9159919
NGR_c12520	-5,45103327	0,03595246	3,92179487	1,75498575	0,4042735	3,02022792
NGR_c03570	-5,41322343	0,01518649	13,8923077	7,31205128	1,77589744	8,00564103
NGR_c22840	-5,32503561	0,0449348	24,3411058	10,2027244	2,29294872	9,12612179
NGR_b10990	-5,11005929	0,04840734	6,59902597	6,98809524	0,27056277	5,17727273
NGR_b18680	-5,05397624	0,03808396	26,586	20,3493333	2,55866667	11,5004
pilA3	-5,02412101	0,01976646	9,59134615	10,3477564	0,80128205	3,92628205
NGR_c11240	-4,76098044	0,01843015	10,0777778	7,66080247	1,40246914	3,46095679
NGR_c31740	-4,62248506	0,03478732	711,498322	508,211409	123,365324	186,857047
NGR_b14050	-4,07511632	0,03215419	8,98835821	7,59303483	1,26567164	2,17840796
NGR_c27520	-4,01134651	0,03936922	5,72586207	5,31034483	1,07758621	2,96954023

NGR_b02950	-3,8736587	0,01705035	16,8571429	19,4113757	2,31481481	4,96084656
trnA5	-3,86423109	0,03903744	142405,176	109898,262	6608,55263	189734,315
NGR_c19050	-3,53106409	0,04096412	7,27763158	3,01959064	0,73099415	4,22938596
NGR_c27840	-3,48493225	0,04729936	3,84368421	1,23538012	0,29239766	2,1619883
NGR_c00990	-3,41201202	0,03882625	15,73	5,50428571	2,34349206	6,51142857
NGR_c08790	-3,37737295	0,04801066	5,11149635	1,97043796	0,7459854	3,22128954
NGR_b07460	-3,1877785	0,04853341	3,27537313	7,30099502	1,72288557	1,65646766
NGR_c21170	-3,14918816	0,03638173	7824,81143	5858,14761	1810,14271	5419,80289
corA	4,05794516	0,04138718	0,52272727	0,39712121	1,28484848	7,34338384
metF	4,11293944	0,04892352	1,30369775	0,9903537	4,14437299	24,8623794
NGR_c12650	4,15853671	0,04260737	1,31940299	0,65199005	3,10679934	14,7337479
NGR_c10550	4,16475011	0,04882537	1,56650943	0,77484277	3,80345912	19,1400943
NGR_c03120	4,19855755	0,0321782	0,96971279	0,32167102	0,98355091	14,0239774
NGR_c36530	4,2420067	0,04577917	1,49577465	0,28920188	1,67073552	16,7915493
soxA2	4,39367617	0,03287552	0,23517206	0,09107625	0,2701417	3,68309717
NGR_c25890	4,43708953	0,04150706	1,30386905	0,62668651	1,9417328	19,1661376
cpaE1	4,45651637	0,0375086	0,71425234	0,19190031	0,76962617	12,6448988
cpaD	4,58147393	0,02841837	0,08489362	0,26212766	0,75716312	4,67368794
NGR_c01870	4,62500556	0,02577977	0,27300771	0,33688946	1,55175664	5,43290488
ribH1	4,67025848	0,02477272	0,63132911	0,51983122	2,41983122	13,6840717
NGR_c32680	4,67639371	0,03434484	3,5880531	3,25471976	9,36312684	74,5325959
NGR_c20050	4,72216128	0,03817513	1,39348404	0,47863475	2,22987589	19,3809397
argD	5,04617042	0,03940515	1,02951389	0,5410108	3,62746914	12,1024691
NGR_c28700	5,17274969	0,01957805	0,62707736	0,29417383	1,43753582	9,27163324
NGR_c34100	5,18432358	0,04760231	0,58004292	0,11750119	1,47577492	5,52041011
queA	5,18853952	0,04413841	0,52438692	0,11189827	1,01108084	4,53896458
NGR_c30520	5,34856807	0,02629299	0,36178756	0,28631261	1,49084629	4,933981
NGR_c31230	5,43890849	0,0214044	0,15896414	0,16361222	0,83678619	5,80863214
NGR_b20630	5,58002381	0,03914045	9,84237805	8,35162602	14,8770325	215,844309
NGR_c08010	5,67399115	0,02426623	0,62470588	0,18117647	1,78696078	6,73784314
NGR_c12790	5,72586409	0,01718654	0,52178571	0,29333333	2,07535714	8,28636905
NGR_c35820	6,10168538	0,02339258	2,45076923	0,47384615	7,22051282	29,2560256
dctP	6,15225842	0,04794254	7,05528455	3,54331527	31,2139115	104,028997
NGR_c12420	6,20475219	0,02251856	0,17654867	0,0605703	0,24582104	4,54169125
celC	6,33301954	0,03012423	1,14482759	0,70282567	4,49176245	13,952682
glxX	6,358528	0,03390638	0,57345679	0,18515089	1,71358025	6,82283951
NGR_c13310	6,49409503	0,01262132	0,6303719	0,37183196	1,6822314	8,06460055
NGR_b20620	6,55599731	0,02093697	28,5533333	34,3193122	59,350582	879,535979
rpIV	6,90549306	0,02330187	4,18846154	1,48192308	12,1623077	61,8464103
glyS	7,09382403	0,03716569	0,47867036	0,2843952	3,06966759	8,86119575
NGR_c04460	7,3374775	0,02228649	4,21056338	0,86760563	17,4319249	63,053169

rho	7,38020681	0,01820007	1,275	0,63254344	7,18933649	21,9078989
NGR_c17470	7,73103018	0,01924566	1,2872093	0	3,16899225	14,7921189
NGR_c12690	7,85521982	0,01534197	0,12954545	0,31764069	0,7495671	8,2741342
NGR_c27950	7,87449568	0,01933575	1,7060241	0	3,86921017	13,5964525
NGR_c23290	8,58914647	0,03134234	0,80350318	0	2,30254777	9,10636943
NGR_c08760	8,93470485	0,01541316	1,02975	0,10266667	3,79516667	13,7499167
cysK1	10,8664865	0,01832033	6,93392371	3,91970936	56,788465	160,77089
NGR_b22900	15,1866066	0,03643075	0,32177419	0	1,00806452	7,20241935
NGR_c13070	15,8449842	0,00949701	2,35645161	1,76845878	31,3651732	86,5456989
aatA3	19,3867091	0,02918871	0,44775561	0,17319202	5,93275145	15,9245636
cysN	19,4827599	0,00519154	2,07655311	0,86412826	25,5495658	66,9773213
NGR_c13060	19,7599956	0,00067423	1,94292453	0,58113208	20,3399371	82,220283
NGR_c13080	19,8088364	0,00132535	3,28373494	1,20783133	29,7381526	154,431225
cysG	24,2448087	0,03260013	1,7325462	0,45386721	24,4345654	47,3220055
cysD	38,7997845	0,01372555	0,50188679	0,56593291	20,2212788	45,9385744
NGR_c02520	Inf	0,04891767	0	0	0,6952381	4,17738095
NGR_c26680	Inf	0,00372363	0	0	0,83647343	6,18345411

### APÊNDICE 3 - GENES DIFERENCIALMENTE EXPRESSOS EM NGR234 APÓS INDUÇÃO COM FLAVONOIDES

Identificação do gene	Fold Change	Valor p	Cobertura NGR1-	Cobertura NGR1+	Cobertura NGR2-	Cobertura NGR2+
NGR_c07750	-6,226515528	0,00689363	8,94285714	2,77142857	8,77718254	0
NGR_a02330	-6,157156304	0,00569152	3,73512397	0,96198347	2,96253444	0
NGR_c09610	-5,951137924	9,82E-07	97,4782051	22,4102564	60,8451567	1,74487179
NGR_b15000	-4,626626282	8,92E-08	7,57299703	1,41127596	3,40825915	0,31671612
NGR_c20560	-4,625145022	7,27E-05	10,8328767	0,79726027	5,63424658	1,98424658
NGR_b07280	-4,513828391	0,02063894	4,75909091	1,76363636	2,54375	0
NGR_c07770	-4,145248706	0,00457992	4,96363636	1,91385281	3,87867965	0,16868687
paal	-4,103171743	8,80E-09	19,9396552	4,68965517	8,3756705	1,06130268
NGR_c03820	-4,072869448	0,00075527	6,62697161	1,5022082	2,73690852	0,48901157
NGR_c16900	-4,02135509	1,08E-07	13,2627907	3,29767442	6,51627907	0,67767442
NGR_c30620	-3,87815811	0,03373407	8,86111111	3,24444444	8,33410494	0,72160494
NGR_c26200	-3,802333558	2,74E-05	39,4521583	13,4158273	26,7491607	2,0293765
NGR_b04820	-3,759940337	0,01180952	8,02924528	3,29433962	5,91446541	0,36761006
NGR_b08960	-3,665968115	2,51E-07	4,72317814	1,10080972	2,39875169	0,31551957
NGR_c15740	-3,659469058	0,00383141	10,8645161	5,54408602	5,95430108	0,51917563

paaN	-3,512216812	2,82E-08	26,230102	7,60408163	13,8612488	1,84113217
NGR_b03800	-3,502798138	0,00013115	5,47859425	1,76932907	2,58935037	0,27875399
NGR_b22100	-3,467530884	0,01291135	4,51764706	1,59084967	1,43747277	0,12734205
paaK	-3,456687665	1,31E-08	24,0108635	7,06629526	12,8521356	1,43700093
actP	-3,443986304	0,00054122	25,5012485	5,81260404	9,32374158	1,61460563
NGR_b21190	-3,435933556	3,32E-06	14,1010067	4,88456376	8,47035794	0,91532438
NGR_b05580	-3,410096141	0,02183836	7,01666667	1,61666667	2,09259259	0,27060185
NGR_b07460	-3,315189524	0,02366287	3,27537313	0,86865672	4,31641791	0,58159204
paaB	-3,230943393	1,82E-06	40,9184211	11,4610526	22,6852632	2,65719298
paal	-3,191872031	2,34E-08	27,4242366	7,8221374	17,5655216	3,22022901
NGR_b21490	-3,144788302	0,00264157	7,69444444	2,87407407	5,05123457	0,80787037
NGR_c27790	-3,135522796	2,71E-05	252,211017	77,2011299	124,316008	19,0332392
paaF	-3,113247538	2,89E-05	5,79542334	2,33409611	3,46655225	0,31209001
NGR_c20830	-3,113234329	0,00110488	8,27290076	3,18625954	4,44516539	0,44618321
NGR_c18580	-3,108860665	4,68E-05	21,7852174	5,91478261	9,27985507	1,5173913
NGR_c30010	-3,101232634	1,07E-05	77,8938338	27,9008043	50,6518767	7,13556747
NGR_b03790	-3,048794876	0,00292644	3,65070423	0,68450704	1,08951487	0,54483568
NGR_a02420	-3,047664447	0,00715683	4,40167224	0,9083612	1,08896321	0,45328874
NGR_c26120	-3,024460576	0,02408571	13,059799	5,80502513	8,49154104	0,87688442
NGR_c07640	-2,993280491	0,00010856	16,6945714	6,93714286	11,5311429	1,19561905
NGR_b08480	-2,956757444	0,00941774	27,1216216	4,86216216	7,81621622	3,39977477
NGR_c17560	-2,917550195	0,0315841	7,04819277	3,62650602	6,37931727	0,46947791
NGR_c15010	-2,900162836	0,00097843	10,1256881	4,09357798	7,20932722	0,89373089
NGR_b18700	-2,893228262	1,72E-05	27,9393204	8,01941748	12,2987055	2,26165049
NGR_b11430	-2,890405779	0,00099534	11,5758621	4,34827586	8,49482759	1,75258621
ocd1	-2,840846967	0,0010966	7,38325301	3,30361446	3,8964257	0,46742972
NGR_c26140	-2,815986152	0,00025131	13,6204545	5,66969697	9,65751263	1,32518939
NGR_c20580	-2,811787684	0,01551161	4,36928571	1,52428571	2,00702381	0,27833333
NGR_c21340	-2,803589917	0,00182646	4,01918605	1,84806202	2,57254522	0,60083979
paaA	-2,801808586	1,55E-06	30,5534328	10,4137313	15,0625373	2,72333333
NGR_b14380	-2,801746624	0,00023122	5,88670213	2,48138298	3,45545213	0,51591312
NGR_c36050	-2,782307646	0,0169423	25,7365385	13,0769231	16,6521368	1,49871795
NGR_c12680	-2,775547542	0,02266908	24,0895833	6,275	6,87291667	1,68240741
NGR_b14990	-2,770187554	0,00015085	7,35206287	2,956778	5,0651277	0,53588736
NGR_c22610	-2,716728476	1,09E-05	113,50084	43,5882353	60,6382353	9,23263305
NGR_b21380	-2,650925077	0,00031727	7,20108696	2,10869565	3,0839372	0,49414251
NGR_b14000	-2,645012889	0,00125116	16,123125	7,04	9,56833333	2,18125
NGR_c34060	-2,642636974	0,0056349	5,65517241	1,67241379	2,24683908	0,94444444
NGR_c03570	-2,64246457	0,00249272	13,9084615	4,03384615	7,26692308	1,94179487
NGR_c21140	-2,641360839	0,00103406	20,1447761	9,34477612	9,75149254	1,30858209
NGR_b14370	-2,619330794	0,00041387	6,54826303	2,7235732	2,82874276	0,38676592

NGR_c17310	-2,550918572	7,38E-05	30,3102539	12,6984375	17,7241862	2,74973958
NGR_b00550	-2,539486954	5,06E-05	94,2675573	35,2015267	45,0402036	7,56564885
NGR_c13780	-2,521216836	0,01146908	142,265039	63,6800781	88,3392904	16,4570638
NGR_c22010	-2,51808985	0,00508448	149,569565	82,7043478	113,858696	18,3415459
sucB1	-2,471004038	0,0004276	4,68658537	2,10634146	3,21760163	0,66321138
NGR_c13660	-2,454996994	0,01805598	103,740493	42,602261	54,294964	10,2939363
NGR_b02380	-2,435311175	0,00426454	13,0957249	7,12267658	9,6598513	1,65954151
Int	-2,435063041	0,00024309	3,4612782	1,36917293	2,22669173	0,53023183
NGR_c18660	-2,43329047	0,03710636	40,4033742	21,8392638	27,943456	3,33118609
NGR_c10270	-2,417778874	0,01556305	3,76678322	1,49230769	3,47284382	0,81748252
NGR_c35390	-2,408427235	0,00206235	26,3148649	8,52432432	12,6603604	2,6213964
NGR_c22840	-2,402180101	1,05E-05	24,3807692	7,75576923	10,1410256	2,66346154
NGR_c03620	-2,388574514	4,90E-05	32,3581395	11,1426357	13,478553	2,94186047
NGR_b20290	-2,380179531	0,009208	4,51737452	2,66718147	2,37374517	0,37612613
NGR_b14050	-2,374425648	0,00095667	9,00626866	3,56776119	4,26736318	0,92547264
NGR_c18650	-2,368580966	0,00041925	241,692857	69,8485714	87,7604762	26,3928571
paaB1	-2,36586821	0,00560022	3,55056818	1,61893939	2,85353535	0,62569444
NGR_c07710	-2,362225425	1,78E-05	40,0903846	15,6551282	18,6438034	5,03365385
NGR_b12670	-2,346910451	0,00015077	25,0450413	8,84297521	11,4782369	2,4922865
NGR_c17620	-2,315637893	0,0001544	81,0727273	40,2939394	46,5222222	10,5116162
NGR_c04320	-2,313082202	0,00129408	135,669886	54,2265152	84,3253157	21,5912247
NGR_b02650	-2,311913558	0,03246533	11,2120482	5,14939759	3,92289157	1,05120482
NGR_c26670	-2,307312991	0,00082823	285,635047	118,229907	166,096106	34,097352
cheB3	-2,306181811	0,04051679	6,56814159	3,92979351	5,04587021	0,97202557
NGR_b19220	-2,292586859	0,00994265	63,8963415	37,6536585	47,954065	9,25609756
ctaG	-2,2885061	0,01548075	5,38425	2,725	1,77083333	0,67766667
NGR_b15010	-2,255528291	1,91E-05	31,7896226	12,081761	16,099109	4,69895178
NGR_c27440	-2,246446851	0,00024482	11,1015707	4,07225131	5,71239092	1,68089005
pleD	-2,22600164	0,00786984	17,2128289	9,29692982	11,9796053	1,89480994
NGR_c33740	-2,224037718	0,02461992	6,64941176	2,96705882	4,54882353	1,14607843
ddl	-2,211851464	0,00057479	13,8361564	6,51856678	6,60564604	1,23615635
NGR_c20200	-2,208172262	0,00104469	16,0554745	6,91970803	7,8053528	1,27682482
dctA1_A	-2,19837805	8,28E-05	11,5930693	4,41227723	5,56554455	1,47947195
rnd	-2,196924188	0,00077771	8,41289063	3,49375	3,39366319	0,83606771
NGR_b15270	-2,195508798	0,00106506	37,727027	16,3891892	26,8015766	5,25427928
NGR_c08580	-2,189920304	0,0023042	243,806667	129,357778	158,45	34,4201852
NGR_c19900	-2,186273543	0,00475882	7,5391791	1,59253731	3,58631841	1,01778607
NGR_a01340	-2,166078982	0,00259079	5,47871622	2,47207207	3,51325075	0,95964715
nuoM2	-2,162103071	0,00028373	6,26166951	2,66609881	4,00406019	0,77927314
NGR_c22650	-2,157856103	0,00077348	54,9544737	27,0126316	29,2785088	6,63850877
NGR_c35880	-2,157149699	0,00227149	6,09778846	3,04346154	3,10064103	0,69070513

NGR_c12810	-2,146648335	0,0111364	233,346729	119,076636	116,951869	21,4859813
secD1	-2,120886453	0,00020739	9,22393617	4,17021277	5,1929275	1,37878251
NGR_b07890	-2,109979896	0,00934693	12,9327273	6,44	10,0325758	1,77121212
NGR_b16940	-2,080732809	0,00136869	11,2931373	3,5620915	5,43017429	1,90457516
NGR_c21480	-2,080465047	0,02913077	14,5904437	8,34880546	7,83754266	1,36171786
NGR_c27590	-2,072929183	0,00468215	155,597107	89,1123967	95,7429752	20,2392562
NGR_b22320	-2,070516827	0,03504046	5,6625	1,70925926	4,37361111	1,26280864
NGR_c35550	-2,068499183	0,01990774	5,80956522	3,0826087	3,02695652	0,71818841
NGR_c17640	-2,067608163	0,00682258	97,0806402	50,7914634	51,2453252	13,2409045
NGR_c20190	-2,064718659	0,00335741	27,8997791	14,1917526	17,1279823	4,0757732
trnG2	-2,062483638	0,03118971	469,903378	305,156757	347,325	70,297973
NGR_c26560	-2,061557306	0,01385938	118,050502	66,509699	72,453233	17,4036232
NGR_b02960	-2,061466366	0,00569503	12,0147783	5,17241379	5,11633826	1,24351396
NGR_c06110	-2,055664162	0,0417021	9,59509346	4,40654206	3,95911215	0,68084112
NGR_c12580	-2,054890053	0,00107683	76,9449664	34,3194631	30,3092841	8,14686801
NGR_b14820	-2,054628685	0,03101559	4,50833333	2,41695906	1,8331384	0,51023392
NGR_c20720	-2,052250021	0,00147277	14,2603448	6,77777778	7,98876117	2,34003831
NGR_c24650	-2,044850598	0,00558796	34,8830769	23,9138462	23,9261538	4,93923077
NGR_c14990	-2,041646077	0,01150597	54,3138158	26,9342105	28,2798246	5,38355263
NGR_b20740	-2,039039066	0,00575	5,06357143	2,03666667	2,14305556	0,60103175
NGR_c12470	-2,036491976	0,00477426	52,8396794	25,5507014	29,3059118	7,56780227
dctA2	-2,026810512	0,00980065	15,7374718	8,4469526	9,04676448	1,75346125
NGR_c12460	-2,01886593	0,03192066	177,525	110,801613	121,631048	26,1162634
NGR_c00500	-2,017169174	0,006027	438,310465	196,181395	208,243992	63,5614341
NGR_c33040	-2,005971302	0,00120357	4,2026839	1,81749503	2,07263088	0,92455268
fabG1	2,005439595	0,04136422	2,0804878	3,43902439	1,30765583	1,77682927
NGR_b01380	2,008090489	0,00407207	7,55786802	10,2568528	2,58248731	4,88045685
rplX	2,016835478	0,00150011	4,15528846	12,5153846	4,54182692	4,29246795
NGR_c12170	2,018618815	0,00043275	10,7041534	19,0645367	5,69760383	7,72939297
glpD	2,01904454	0,00026891	1,97075099	5,24466403	1,90721344	2,03402503
msrA2	2,03087531	0,03009063	25,4910138	41,5917051	13,86298	21,6848694
pyrG	2,032028456	0,01737153	2,29219601	3,61597096	1,27843315	2,30450696
NGR_b11200	2,039726123	0,00048687	3,86904762	7,71428571	1,87389771	2,64660494
atpG	2,039809757	0,02724434	1,92193878	3,20748299	0,65090703	1,0574263
NGR_c13070	2,040434732	0,00028543	2,35887097	5,41863799	1,7589307	2,22398447
rmlC	2,041782126	0,02126983	3,73086735	6,05306122	1,43358844	2,18256803
thuG	2,042172676	0,01094531	76,3920863	116,085612	32,1323741	51,9164269
NGR_c32010	2,046847938	0,00912957	7,19733728	9,43076923	2,72317554	3,96982249
NGR_c10940	2,054171677	0,02167832	1,9512766	3,14212766	0,90092199	1,27964539
NGR_c28290	2,064347556	0,01368423	5,475	12,4323529	5,02990196	5,29436275
nuof2	2,067474834	0,00290505	1,78793103	3,66528736	1,03908046	1,36118774

rhcS_A	2,073451083	0,02940469	3,90326087	8,98478261	2,43315217	4,01449275
atpF2	2,083131321	0,0059258	3,64852941	6,19607843	1,37736928	1,61944444
rpmJ	2,085937188	0,00351414	6,33214286	21,0380952	7,26785714	7,18015873
NGR_c35730	2,096205504	0,00021283	2,33314763	5,24373259	1,2933844	1,71541318
NGR_b18920	2,116792217	0,00022336	20,2396552	39,5802956	12,4536125	18,9949097
sucC	2,127920542	0,00102691	7,86708543	12,3030151	3,42671692	6,34254606
NGR_b18310	2,143649799	0,0068191	9,81071429	19,1285714	6,35456349	9,44900794
sucA	2,147382421	0,00010321	4,85720721	10,2598599	3,00382049	4,55427094
cycF	2,15613633	0,00034166	23,5520134	38,7946309	7,84519016	16,4470917
NGR_c01050	2,161546186	0,01204854	1,72595745	3,63744681	1,02248227	1,07780142
NGR_a00510	2,170190737	0,00187753	2,20084459	5,05878378	0,92150901	1,27922297
NGR_b12570	2,176764365	0,00060465	15,6677748	36,4862958	10,5763003	14,3384668
thuA	2,210901636	0,02289898	133,41584	217,276336	60,7908397	112,271692
rpsB	2,217525816	0,00055776	5,71230469	9,22265625	2,67213542	3,95091146
NGR_c30950	2,220194579	0,00094823	4,24621849	6,52773109	1,3910831	3,0751634
rpsT	2,227918288	0,00348814	2,91404494	8,5011236	3,52209738	4,47340824
cysN	2,232487945	0,00011162	2,07835671	4,65250501	0,85641283	1,45821643
atpF1	2,234420448	0,0177372	1,10833333	3,71975309	1,23199588	0,96213992
rplF	2,236310935	0,00375226	12,6926966	22,2786517	5,58426966	11,9019663
metK	2,243772547	0,04406037	9,60106383	14,2931442	4,6498424	9,23987392
noI	2,251395057	0,00066782	5,40238095	8,23619048	1,88714286	3,94502646
NGR_b03130	2,251805761	0,00078307	4,73253275	9,63755459	2,00967977	3,01477438
sucD	2,275149706	4,57E-05	6,64734219	15,9262458	2,94872647	3,71483942
NGR_c12130	2,276034089	0,02376215	6,74884615	11,8215385	3,85320513	7,01128205
rpsO	2,279216513	0,00020284	13,5933333	28,5	7,73555556	12,9511111
NGR_c32550	2,279340677	0,00099064	1,01404358	3,17869249	0,87744149	1,01117837
NGR_c36180	2,296877708	9,08E-05	23,0898551	41,8009662	9,08784219	17,5392915
galE	2,306744441	0,04369748	1,09826087	3,49681159	0,63748792	0,70222222
NGR_b03150	2,312361929	0,01474937	2,48551402	4,93582555	1,3569055	2,93182762
rpoA	2,325934141	0,01444377	15,1090504	25,0409496	7,26246291	14,3653314
NGR_c16980	2,330494786	0,01451634	9,39586466	12,124812	1,74486216	5,18446115
NGR_c27960	2,338903369	0,00172597	4,98535714	7,04428571	1,36690476	3,80886905
NGR_a03490	2,341286365	4,16E-05	3,21603261	8,24184783	1,60366848	3,05919384
NGR_b05480	2,366686105	5,46E-06	4,2359589	10,9965753	3,05302511	3,60308219
spsC	2,399407022	2,23E-06	12,9506494	46,0779221	13,3768398	17,2718615
NGR_b12600	2,401609818	0,00020074	1,85294118	4,57544757	0,99987212	1,2912191
sucB2	2,418438447	3,51E-06	2,52101449	7,58164251	1,97657005	2,41851852
NGR_c14310	2,430960024	8,37E-06	6,20790441	13,5845588	2,42503064	4,81626838
NGR_c01440	2,45150187	0,0075926	3,14086957	6,62202899	2,06560386	3,26806763
rplM	2,472679501	0,00771691	1,4583871	4,45290323	0,78774194	1,0055914
NGR_c25590	2,474629496	0,02755287	3,48688525	6,84262295	1,33442623	1,91639344

NGR_a03310	2,476708773	0,00019128	5,86944444	13,0790123	2,3867284	5,51656379
NGR_b20130	2,509784766	0,02257004	2,01192661	3,48073394	0,98374108	1,95840979
NGR_c21730	2,51306995	0,00142591	1,92731092	4,14621849	1,18060224	2,00028011
NGR_c06980	2,557362531	0,01474989	7,38611111	14,5382716	2,73940329	8,13621399
NGR_c32180	2,562279161	0,00322462	1,57555556	3,31185185	0,70876543	1,22358025
rpsD	2,570357745	0,00034179	3,12961165	8,31650485	1,70024272	3,24441748
NGR_c31670	2,57180603	0,00364715	8,49666667	19,2044444	5,20055556	11,3322222
rmID	2,57369442	4,16E-05	1,81313131	6,35218855	1,2201459	2,02216611
rplN	2,58761501	0,00016608	11,4560976	26,6520325	9,10894309	13,5602981
nuoE2	2,590629553	0,00152451	1,46371841	3,7566787	0,97057762	1,60854392
NGR_b03550	2,646108537	3,02E-06	11,1607895	26,7563158	5,05758772	12,5212719
NGR_c12160	2,659247604	0,00055198	10,0167421	20,3375566	3,6300905	8,6127451
NGR_c22860	2,667040941	0,0037037	1,30705128	3,48974359	0,73084046	1,20548433
NGR_b03160	2,683515615	0,00017446	5,72763158	12,0328947	2,09057018	3,7004386
exsI	2,686887481	0,00018782	7,36071429	17,9309524	2,66488095	5,21369048
atpA	2,707146363	2,10E-06	2,05970588	6,34941176	1,65284314	2,49643791
metF	2,709223129	0,0001049	1,30418006	3,99614148	0,98151125	1,62609861
NGR_c34820	2,712767993	0,00010014	19,0521809	47,60937	10,6162089	23,4639742
nodU	2,735490326	1,84E-06	3,6509839	8,36422182	1,66759094	3,5990161
NGR_c05580	2,763076591	8,87E-07	3,21301115	9,96802974	1,73996283	2,96325898
NGR_a03370	2,764554608	2,32E-07	3,90037129	13,0970297	3,40775578	5,61720297
NGR_c17010	2,781419113	0,00125852	1,46126374	3,89230769	1,16327839	1,76401099
NGR_a03630	2,786793059	0,0022757	4,17207447	7,1393617	1,06161348	4,22189716
NGR_b03170	2,799487208	0,00051238	1,2012605	3,83613445	0,88956583	1,26708683
NGR_a02560	2,801009114	1,17E-06	5,08361345	15,3302521	3,19978992	5,06477591
NGR_c19960	2,81071194	0,01397823	1,24066667	3,33422222	0,36177778	0,77933333
wbgA	2,816124507	6,79E-08	2,80516484	8,31604396	1,38260073	2,84565934
NGR_c18100	2,828157806	5,59E-07	4,73684211	12,3506073	2,15951417	3,30951417
NGR_c03800	2,876029366	0,0001384	30,4828767	72,7938356	15,4071347	36,2334475
NGR_b20640	2,917284297	0,00020782	6,30592885	13,5438735	2,95362319	7,03955863
NGR_c29330	2,919950886	0,000213	1,85472103	4,4583691	1,04806867	2,1304721
rhcC2_A	2,923500135	0,00126733	4,20106132	8,71886792	1,59398585	4,37130503
nodZ	2,931713899	2,88E-06	2,0995356	6,58885449	1,07105263	2,70123839
NGR_c29300	2,933300795	1,60E-07	3,18094556	10,5083095	1,66852913	3,30582617
NGR_b15970	2,943534131	5,17E-05	2,3275	8,65	1,18055556	2,3295
NGR_b20620	2,986644819	5,29E-06	28,5766667	74,872381	16,0910582	34,7012698
rpmH	3,005235394	0,01276884	6,79666667	14,0311111	2,71333333	8,84740741
noI	3,04659887	1,38E-08	5,83283379	18,8386921	3,86117166	7,59150772
rhcU_A	3,060118922	4,95E-06	1,26806358	5,51213873	1,27129094	1,96348748
nodS	3,060979488	4,04E-07	3,55368664	11,9179724	2,19477727	5,05130568
NGR_b14270	3,062319824	7,16E-05	1,83033708	4,94606742	0,59470974	2,23192884

nopC	3,072661393	0,00019576	188,930303	581,523232	111,430471	235,697643
NGR_c03810	3,095043381	0,00175859	2,69013158	7,51315789	1,72757675	3,64753289
NGR_a00970	3,098911017	0,00015672	42,5691781	116,824658	25,9358447	61,9941781
NGR_c31300	3,100707543	1,03E-08	5,03757764	16,5801242	2,69321946	5,25093168
NGR_b20630	3,171860286	8,97E-08	9,85060976	26,3231707	3,99695122	10,8933943
NGR_a00980	3,192841571	1,97E-08	22,392033	66,2945055	12,0154762	28,3635531
NGR_b03560	3,221122174	1,57E-07	10,5507143	29,8019048	4,4787037	13,9379365
NGR_c03510	3,224896376	0,00052149	0,8673913	6,76869565	1,48710145	1,6942029
ndk	3,247706517	2,06E-08	11,4095745	34,0822695	4,27423168	11,1503546
NGR_c13060	3,277316783	0,01420395	1,94433962	3,57169811	0,5759434	2,38144654
NGR_c29310	3,289593871	7,44E-07	4,74857143	23,3228571	5,3502381	7,515
hpnD	3,300405787	4,40E-06	1,97089286	6,35357143	1,2802381	2,25690476
spsD	3,321444653	1,09E-09	7,92755344	29,3786223	4,76769596	9,92367379
NGR_c03050	3,377490122	6,00E-07	13,945122	48,304065	8,80365854	24,5242547
NGR_c32750	3,425518499	0,00115063	1,87777778	6,52407407	1,31898148	3,55308642
NGR_c03060	3,509651902	3,78E-05	9,88492063	28,9904762	5,2117284	14,9608466
rhcT_A	3,510100145	2,59E-07	2,53296703	9,47985348	0,9991453	2,6002442
fixC	3,514197285	0,01215059	2,75951835	7,22477064	1,55011468	4,82801988
NGR_c23240	3,522319583	2,39E-09	4,05871429	13,0190476	1,80688889	5,15112698
rmIB	3,607695856	2,88E-09	3,74957265	10,82849	1,89093067	5,06082621
NGR_a00750	3,711096163	3,75E-06	22,7136725	70,3046105	9,79292528	31,3329624
NGR_b03520	3,799242205	2,71E-10	30,6072917	111,561111	13,4306713	49,2486111
NGR_c29320	3,804472731	2,52E-10	13,2122318	42,0060086	4,66423462	17,341774
NGR_c03040	3,833877799	5,59E-11	18,5909341	70,889011	9,73195971	30,0995421
NGR_b03540	3,854621635	4,76E-08	25,8017804	102,397033	15,1032641	46,0418892
NGR_a00990	3,892268657	3,68E-07	8,66457726	32,7586006	5,14105928	15,4723518
nopM	3,909656913	3,02E-05	9,06608775	26,3956124	3,79609994	13,4868678
NGR_a00920	3,956212528	9,69E-05	32,992997	120,670181	18,522239	59,2619729
NGR_a00630	3,985888558	2,58E-08	46,2787709	159,290503	19,490689	74,5344507
NGR_a00450	4,028113722	0,00027623	1,09736842	3,23349282	0,45781499	1,58070175
rhcN_A	4,029539597	1,47E-05	40,0964602	143,477434	20,7178097	67,5199115
NGR_c36530	4,041786118	0,00030042	1,49788732	3,3286385	0,28661972	1,45954617
NGR_a00740	4,101205455	1,79E-06	36,4788158	148,759474	19,9688596	68,8327193
nopT	4,104321646	9,47E-09	7,13187023	29,0045802	4,14796438	13,1332061
nopL	4,105622041	0,00496423	36,1818584	121,972271	22,2419371	76,8153884
NGR_a00960	4,141917061	2,78E-05	3,31108949	12,5451362	1,77399481	7,18540856
NGR_a00730	4,145654207	1,15E-05	52,101632	209,893769	30,6385757	96,299456
NGR_b10020	4,147032101	0,01195964	0,67058824	3,43361345	0,1710084	0,49117647
NGR_c09120	4,238028033	4,62E-06	28,443913	109,294783	15,897029	56,696087
atpH	4,252654775	0,00111602	1,26587302	3,34708995	0,32301587	1,84656085
nopJ	4,255286822	0,00029405	38,2609195	135,327969	20,8222222	75,3814815

NGR_a00540	4,279940672	0,00052915	100,598485	395,852525	59,3670034	209,043266
NGR_c10670	4,287714247	0,00789585	2,06833333	5,07555556	0	1,08240741
fixF	4,287791838	0,00034795	28,4575682	96,3315136	14,466129	58,2831266
NGR_a00710	4,321545965	1,36E-09	46,6683871	175,04129	18,9915054	80,2766667
NGR_b03530	4,356919599	1,64E-10	10,0742358	40,2253275	5,7522198	18,6748544
NGR_a01000	4,394468361	8,64E-09	21,5300813	83,7121951	10,6352304	41,4936314
nopP	4,409396804	3,58E-08	22,8534926	99,3330882	13,8912377	47,0976103
fixB	4,434647409	0,00010409	1,06013514	5,54702703	0,97941441	3,09301802
rmlA	4,437177883	0,03132407	1,80679443	4,94285714	1,07857143	4,12328688
rhcJ_A	4,46202983	3,27E-06	57,9998276	224,542759	27,6153448	109,607586
NGR_a00760	4,480682462	0,00012724	6,64814815	28,1688889	4,56469136	16,0925926
NGR_a00720	4,565320254	3,18E-10	14,1101528	57,5	7,76961426	25,7105895
rhcQ_A	4,605536586	1,33E-12	11,2408078	55,1097493	6,46425255	23,9850975
rhcR_A	4,695205991	3,04E-12	5,66569507	20,7372197	1,7338565	9,10530643
ttsI	4,704670667	6,40E-12	4,83171806	19,1603524	1,93773862	7,85888399
rhcL	4,737140711	5,19E-07	71,9454545	328,010526	38,5369219	168,898246
NGR_a00430	4,878061564	1,86E-05	6,21552901	19,6457338	1,9447099	10,9254266
NGR_a00780	4,887304398	8,18E-09	2,28650794	9,98201058	0,96904762	5,3335097
nopB	4,893334993	3,33E-06	109,268182	474,711515	57,2238384	239,541515
noeE	4,91146824	9,94E-15	5,16107143	31,0195238	3,45126984	13,821627
NGR_a00520	4,924861442	0,00044898	22,8491803	78,6765027	9,81138434	47,0343352
noIU	5,144526308	7,79E-11	27,4330986	113,761502	12,6938185	60,4812207
rhcV_A	5,210434222	2,99E-07	17,0020057	85,4624642	9,15787966	43,5230898
syrM2	5,323447858	3,88E-05	3,73367647	16,4235294	2,21568627	7,97014706
fixX	5,451945487	0,01128696	2,98469388	10,7204082	1,3297619	6,93231293
fixA	5,522281425	5,19E-09	1,13916084	5,74545455	0,69784382	2,6479021
NGR_b14280	6,662778171	0,02438078	0,88909884	3,08313953	0,11831395	2,53386628
NGR_c08250	8,718181435	0,02957394	0,78235294	3,80392157	0,39901961	5,31503268
NGR_a01210	10,05220295	1,56E-06	5,77850467	39,1981308	1,1411215	22,7288162

## APÊNDICE 4 - ANÁLISE DE COMPONENTE PRINCIPAL (PCA)

NOTA: Os números representam as bibliotecas 1) *nodVW1-* descartada, 2) *nodVW1-* utilizada, 3) NGR2+, 4) *nodVW2-*, 5) *nodVW2+*, 6) NGR1-, 7) NGR2-, 8) NGR1+, 9) *nodVW1+*

